"Anwendung der Vektorrechnung auf die Grundlagen der Geometrischen Optik," Ann. Phys. (Leipzig) **35** (1911), 277-298.

The application of vector calculus to the foundations of geometrical optics

By A. Sommerfeld and J. Runge

Translated by D. H. Delphenich

Contents

In the present article, a method shall be given by which the known laws of geometrical optics can be represented in an especially intuitive form. The method consists of assigning a unit vector $\mathfrak{S}$ to each point in the direction of the light ray in such a way that the tools of vector calculus will become fruitful for our class of problems ([1]).

([1]) The method was developed by Sommerfeld in a lecture (Munich, 1910) and further elaborated by J. Runge for the questions of curvilinear light rays, in particular, in § 7 and 11.

I. Rectilinear light rays.

Geometrically speaking, the light rays in a homogeneous medium are a family of straight lines through which one can construct a normal surface at any point – i.e., a surface that is perpendicular to every ray.

These two conditions – viz., rectilinearity and the existence of normal surfaces – will now be expressed in the language of vector calculus.

§ 1. Rectilinearity

Light rays are the streamlines of a vector field $\mathfrak{S}$; the condition requires that they must be straight, and thus, of zero curvature.

If we first start with the arbitrary curvilinear case of a given ray curve through the unit vector $\mathfrak{S}$ then since the length of $\mathfrak{S}$ is assumed to be constant (= 1) as we advance along such a curve, the infinitely small vector $d\mathfrak{S}$ will be perpendicular to $\mathfrak{S}$ and equal to the change in angle between two neighboring $\mathfrak{S}$ (Fig. 1). If one then denotes the line element of the ray curve by ds then the curvature or change of direction per line element will be equal to $d\mathfrak{S} / ds$, and one must then have:

(1) $$\frac{d\mathfrak{S}}{ds} = 0$$

for rectilinear rays.

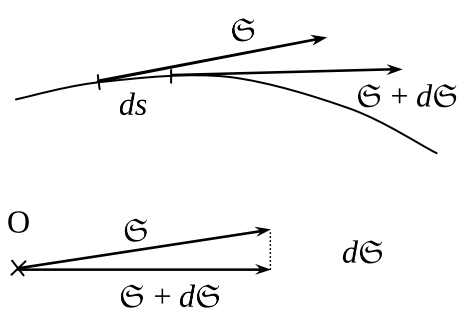

Figure 1.

This condition can be altered somewhat by a component calculation.

It is:

$$\frac{d\mathfrak{S}}{ds} = \frac{\partial\mathfrak{S}}{\partial x}\frac{dx}{ds} + \frac{\partial\mathfrak{S}}{\partial y}\frac{dy}{ds} + \frac{\partial\mathfrak{S}}{\partial z}\frac{dz}{ds}.$$

However, the quantities dx / ds, dy / ds, dz / ds are nothing but $\mathfrak{S}_x$, $\mathfrak{S}_y$, $\mathfrak{S}_z$ so we have:

(1a) $$\frac{d\mathfrak{S}}{ds} = \mathfrak{S}_x \frac{\partial\mathfrak{S}}{\partial x} + \mathfrak{S}_y \frac{\partial\mathfrak{S}}{\partial y} + \mathfrak{S}_z \frac{\partial\mathfrak{S}}{\partial z}.$$

On the other hand, since $|\mathfrak{S}|^2 = 1$ for any direction of the gradient:

(1b) $$0 = \tfrac{1}{2}\,\text{grad}\,|\mathfrak{S}|^2 = \mathfrak{S}_x\,\text{grad}\,\mathfrak{S}_x + \mathfrak{S}_y\,\text{grad}\,\mathfrak{S}_y + \mathfrak{S}_z\,\text{grad}\,\mathfrak{S}_z\,.$$

By subtracting (1a) and (1b), it then follows that:

$$\frac{d\mathfrak{S}}{ds} = \mathfrak{S}_x\left(\frac{\partial\mathfrak{S}}{\partial x} - \text{grad}\,\mathfrak{S}_x\right) + \mathfrak{S}_y\left(\frac{\partial\mathfrak{S}}{\partial y} - \text{grad}\,\mathfrak{S}_y\right) + \mathfrak{S}_z\left(\frac{\partial\mathfrak{S}}{\partial z} - \text{grad}\,\mathfrak{S}_z\right).$$

If one takes – e.g. – the x-component of this vector equation:

$$\left(\mathfrak{S} = \mathfrak{S}_x, \quad \text{grad} = \frac{\partial}{\partial x}\right)$$

then the first term on the right will drop out, and bracket of the second will be:

$$\frac{\partial\mathfrak{S}_x}{\partial y} - \frac{\partial\mathfrak{S}_y}{\partial x} = -\,\text{rot}_x\,\mathfrak{S},$$

and that of the third will correspondingly be equal to $+\,\text{rot}_y\,\mathfrak{S}$, so:

$$\frac{d\mathfrak{S}_x}{ds} = -\,\mathfrak{S}_y\,\text{rot}_z\,\mathfrak{S} + \mathfrak{S}_z\,\text{rot}_y\,\mathfrak{S} = [\text{rot}\,\mathfrak{S}, \mathfrak{S}]_x\,.$$

As a result, one will have:

(2) $$\frac{d\mathfrak{S}}{ds} = [\text{rot}\,\mathfrak{S}, \mathfrak{S}]\,.$$

From (1), it then follows that the vectorial form for the condition of rectilinearity will be:

(3) $$[\text{rot}\,\mathfrak{S}, \mathfrak{S}] = 0.$$

§ 2. The existence of normal surfaces

A vector field that is surface-normal must have the same direction as the gradient of a function φ of position that is constant on the surfaces of the family and can then be made identical to it by multiplying by a suitable position-dependent factor; we will then have:

$$\lambda\,\mathfrak{S} = \text{grad}\,\varphi\,.$$

If one computes the rotation then one will get:

$$\text{rot}\ \lambda\, \mathfrak{S} = \text{rot grad}\ \varphi = 0.$$

Now, one has:

$$\text{rot}\ \lambda\, \mathfrak{S} = \lambda\, \text{rot}\ \mathfrak{S} - [\mathfrak{S}, \text{grad}\ \lambda],$$

so it will follow that:

$$\lambda\, \text{rot}\ \mathfrak{S} = [\mathfrak{S}, \text{grad}\ \lambda],$$

i.e.:

$$\text{rot}\ \mathfrak{S} \perp \mathfrak{S},$$

or

(4) $$(\text{rot}\ \mathfrak{S}, \mathfrak{S}) = 0.$$

§ 3. The characteristic condition for optical ray bundles

The two conditions (3) and (4) can now be combined into a single one. The equations:

$$[\text{rot}\ \mathfrak{S}, \mathfrak{S}] = 0,$$
$$(\text{rot}\ \mathfrak{S}, \mathfrak{S}) = 0$$

are compatible with each other only if:

(5) $$\text{rot}\ \mathfrak{S} = 0,$$

since they require that rot $\mathfrak{S}$ is simultaneously $\parallel \mathfrak{S}$ and $\perp \mathfrak{S}$.

This vanishing of rot $\mathfrak{S}$ – or, in the language of the theory of currents, *the irrotationality of the current* $\mathfrak{S}$ – is the characteristic condition for optical rays for general Kummer ray systems.

One can make the rotation character of the general ray bundle more intuitive thus: Once one has distinguished a "principal ray," one considers the rays of the system that are infinitely close to it to be an "infinitely-thin bundle," and marks the points at which a plane E that is perpendicular to the principal ray, as well as a parallel plane E' that is at an infinitely-small distance δ from it, are met by the rays of the bundle. The associated points of E and E' are related by a general affine transformation. From the fundamental theorem of the kinematics of plane continua, it can always be decomposed into a deformation along two mutually-perpendicular directions (a transformation of an even character in the coefficients) and a rotation (a transformation of an odd character). If one draws an infinitely-small circle in the plane E around its intersection point with the principal ray then it will be converted into an ellipse by the deformation; this ellipse will be rotated by the rotation. The angular velocity – i.e., the infinitely-small rotation, divided by δ – will now be equal to $\frac{1}{2}$rot S, which is similar to the vorticial velocity in hydrodynamics; the rotation must then be calculated at the midpoint of the plane E (or E') and represents the component of that vector along the principal ray. (The components that are perpendicular to it will vanish, from (8), due to the rectilinearity of the bundle.) *A* *particular screw-sense around the principal ray is then defined by the sign of* rot $\mathfrak{S}$ *at any location of the general Kummer ray bundle.*

For the other one – viz., the deformation component – the two mutually-perpendicular directions – viz., the "principal dilatation directions" are fixed; i.e., *each of them remains parallel to itself when it advances along the principal ray.* One concludes that from a simple calculation that lies beyond the scope of this article and will be passed over here.

If rot $\mathfrak{S} = 0$ then the deformation ellipse will not rotate as one advances along the principal ray, and the principal axes of the deformation ellipses will lie in two fixed, mutually-perpendicular planes through the principal ray; they are *symmetry planes* for the structure of the ray bundle. In particular, they go through the two degenerate cases of the deformation ellipse for which it contracts to one of its two principal axes. The two points of the principal ray in whose associated cross-sectional planes one finds the degenerate case are called the *focal points* of the bundle, and the planes through the principal ray and the directions into which the ellipse degenerates will be called *principal planes*. **Sturm's** theorem states that the principal planes for an optical bundle are perpendicular to each other. In our way of looking at things, it is then an immediate consequence of the fundamental theorem of the kinematics of two-dimensional continua, together with the vanishing vorticity of optical ray bundles. The fact that the focal points must always be real in the optical cases is connected with the fact that in the absence of rotation components, the advancing deformation of the ellipse must lead to a degeneracy of the ellipse in one and the other principal direction.

On the other hand, if rot $\mathfrak{S} \neq 0$ then there cannot be any symmetry plane for the ray bundle; furthermore, a sense of rotation will be given by the sign of rot $\mathfrak{S}$ in any cross-section of the bundle. Whether the focal points are real or imaginary will depend upon the magnitude of rot $\mathfrak{S}$. For small values of $|\text{rot}\ \mathfrak{S}|$, the effect of the deformation on the form of the deformation ellipse will predominate, and one will then arrive at a degeneracy of the ellipse in two cross-sections. However, the positions of the principal planes that are defined as above are influenced by the presence of the rotation and will no longer be mutually perpendicular. By contrast, for large values of $|\text{rot}\ \mathfrak{S}|$, the influence of the rotation will predominate. The deformation that takes place along the spatially-fixed dilatation directions, but relative to the deformation ellipse, along rapidly-alternating diameters, will then lead to no real degeneracy of the ellipse. The entire ray bundle is arranged around the principal ray in the sense of a right-wound or left-wound screw, according to whether rot $\mathfrak{S}$ is greater than or less than zero, respectively.

The existence of two mutually-perpendicular symmetry planes for the optical ray bundle is then the external feature of its rotation-free structure; likewise, the presence of a screw sense and the possible absence of focal points will point to the rotational character of the general ray bundle.

II. Rays in inhomogeneous media

Before we go on to relate the further theorems of geometrical optics with the condition (5), we might mention a generalization that is especially interesting due to the fact that it exhibits the connection between that purely geometrical condition and its physical basis.

One addresses the presentation of the corresponding condition for the ray vector in an arbitrary homogeneous medium. Naturally, the path that we shall embark upon can not be the same as the above path. There, we started with the idea that the rectilinear form of the ray curve was a fact of experience, while here the formulas that we will present shall give us the means to construct the paths of the rays, which are generally curvilinear. Moreover, the knowledge of the inhomogeneous medium and its physical influence on the light ray must define the starting point here.

§ 4. The general rotation condition

We have already used the idea that the light rays are perpendicular to surfaces. Now, we shall add that these surfaces are wave surfaces, by whose advance the speed of light will be measured. In that, the index of refraction n, which is given to us as a function of position, will be the ratio of the speed of light *in vacuo* to the speed at that location in the medium in question:

$$n = \frac{c}{\upsilon}. \tag{6}$$

The connection between the position of the wave surfaces – or surfaces of constant phase – and the index of refraction is now expressed by the equation:

$$|\,\text{grad}\,\varphi\,| = \frac{v}{\upsilon} = \frac{vn}{c} = \frac{n}{c}\frac{r}{2\pi}, \tag{7}$$

if φ means the phase at any moment in time as a function of position, v means the frequency (= $2\pi / \tau$), and τ means the period of oscillation.

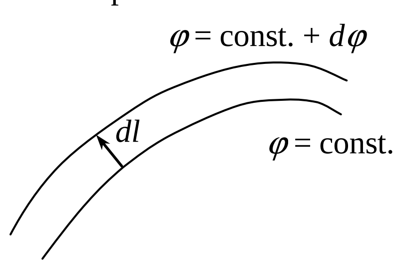

Figure 2.

If one imagines that at some moment in time one has, perhaps, fixed all of the surfaces for which the phase is a multiple of 2π then an individual wave surface that passes between them will pass each of these surfaces at equal time intervals τ, and if one chooses the successive surfaces to have a small phase difference then the time interval will be proportional to the time interval. In the limit, one will then have:

(8) $$\frac{d\varphi}{dt} = \text{const.} = \frac{2\pi}{\tau} = \upsilon.$$

If we denote the geometrical distance between neighboring surfaces by *dl* (Fig. 2) then one will have:

$$\frac{dl}{dt} = \upsilon,$$

which equals the speed of light at the location considered, and it will follow that:

$$|\operatorname{grad} \varphi| = \frac{d\varphi}{dl} = \frac{d\varphi}{dl}\cdot\frac{dl}{dt} = \frac{\nu}{\upsilon} = \frac{\nu \cdot n}{c},$$

as was asserted in (7).

Now, $\mathfrak{S}$ is a unit vector that is perpendicular to the wave surfaces, so one will have:

$$\mathfrak{S} = \frac{\operatorname{grad}\varphi}{|\operatorname{grad}\varphi|} = \operatorname{grad}\varphi \cdot \frac{1}{\nu n / c}$$

or

(9) $$\frac{\nu \cdot n}{c} \cdot \mathfrak{S} = \operatorname{grad} \varphi.$$

With that, we have found the physical meaning of the multiplier λ that was left undetermined in § 2.

However, due to the fact that rot grad $\varphi = 0$, it will follow immediately that:

$$\operatorname{rot} \frac{\nu \cdot n \mathfrak{S}}{c} = 0,$$

or if one drops the constant ν / c:

(10) $$\operatorname{rot} n\, \mathfrak{S} = 0.$$

This is the *general condition for the ray vector* $\mathfrak{S}$ *in any medium.* For a constant *n*, it will go to the special form:

$$\operatorname{rot} \mathfrak{S} = 0,$$

and can also revert to rectilinearity, since from the relation (2):

$$\frac{d\mathfrak{S}}{ds} = [\operatorname{rot} \mathfrak{S}, \mathfrak{S}],$$

the curvature will vanish with rot $\mathfrak{S}$.

§ 5. Consequences for curvilinear ray progressions

In the general case, (2) will yield an expression for the curvature of the ray path in an inhomogeneous medium, and indeed its position and magnitude.

The magnitude of the curvature is:

$$K = \left|\frac{d\mathfrak{S}}{ds}\right| = |\,[\mathrm{rot}\,\mathfrak{S}, \mathfrak{S}]\,|,$$

in which rot $\mathfrak{S}$ is calculated from (10):

(10a) $$0 = \mathrm{rot}\, n\,\mathfrak{S} = n\,\mathrm{rot}\,\mathfrak{S} - [\mathfrak{S}, \mathrm{grad}\, n], \quad \mathrm{rot}\,\mathfrak{S} = [\mathfrak{S}, \mathrm{grad}\, n].$$

Since $\mathfrak{S}$ and rot $\mathfrak{S}$ will be mutually perpendicular from now on, the magnitude of their vector product will be equal to simply the product of their magnitudes, or, since $|\,\mathfrak{S}\,| = 1$:

(11) $$K = |\,\mathrm{rot}\,\mathfrak{S}\,| = \left|\frac{1}{n}[\mathfrak{S}, \mathrm{grad}\, n]\right| = \frac{|\,\mathrm{grad}\, n\,|}{n}\sin(\mathfrak{S}, \mathrm{grad}\, n).$$

As far as its position is concerned, the vector $d\mathfrak{S} / ds$ has the direction of the principal normal of the ray curve, so it will lie in the plane of two neighboring tangents – i.e., in the osculating plane – and since $|\,\mathfrak{S}\,| = 1$, it will be perpendicular to $\mathfrak{S}$. Now, since rot $\mathfrak{S} \perp \mathfrak{S}$, from (10a), and rot $\mathfrak{S} \perp d\mathfrak{S} / ds$, from (2), and therefore $\mathfrak{S}$, rot $\mathfrak{S}$, $d\mathfrak{S} / ds$ will define a system of mutually-perpendicular directions, so rot $\mathfrak{S}$ will have the direction of the binormal of the curve. However, from (10a), the direction of the vector rot $\mathfrak{S}$ is given in such a way that it is perpendicular to $\mathfrak{S}$ and grad n, and as a result, the plane that is determined by $\mathfrak{S}$ and grad n will be the osculating plane of the curve.

Thus, if a starting position for $\mathfrak{S}$ is given at a point, and n is known as a function of position, in addition, then the osculating plane will be determined at that place, along with the curvature, and therefore the position of the center of curvature will be determined; however, the ray curve can be constructed from its successive curvature circles.

§ 6. The law of refraction

For the case of a boundary between two homogeneous media, one must give the law of refraction from condition (10). It can be derived in the same way that the continuity of the tangential components is derived from **Maxwell**'s equations in electrodynamics.

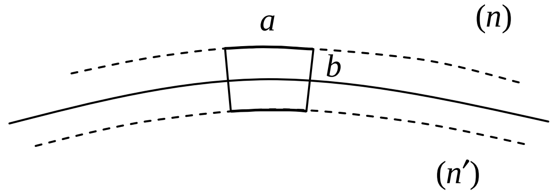

Figure 3.

One then takes a continuous transition and defines the tangent component of rot $n\,\mathfrak{S}$, which is the line integral:

$$\frac{1}{a\cdot b}\int(n\,\mathfrak{S}, d\,\mathfrak{s})\,,$$

when it is taken around a small rectangle that lies in the interface perpendicular to the boundary surface and has the sides a, b (Fig. 3). If one denotes the values in the second medium by a prime then the magnitude that the sides of the tangential rectangle yield will be:

$$\frac{1}{a\cdot b}(n\,\mathfrak{S}_{\text{tan}} - n'\,\mathfrak{S}'_{\text{tan}})\cdot a\,.$$

The component that originates in the normal sides is:

$$\frac{1}{a\cdot b}\frac{\partial}{\partial a}\left(\int n\,\mathfrak{S}_{\text{norm.}}\,db\right)a\,.$$

One then has:

$$\text{rot } n\,\mathfrak{S} = \frac{n\,\mathfrak{S}_{\text{tan.}} - n'\mathfrak{S}'_{\text{tan.}}}{b} + \frac{1}{b}\frac{\partial}{\partial a}\int n\,\mathfrak{S}_{\text{norm.}}\,db\,.$$

One sees that the first term on the right-hand side increases beyond all limits with vanishing b and a non-vanishing numerator, while the second one will remain finite; now, since rot $n\,\mathfrak{S} = 0$, from the condition (10), it will follow in any case that:

$$n\,\mathfrak{S}_{\text{tan.}} - n'\,\mathfrak{S}'_{\text{tan.}} = 0.$$

The equation represents the law of refraction; agreement with the usual form will be obtained immediately when one observes that the tangential component of $\mathfrak{S}$ is equal to the sine of the incident (refracted, resp.) angle (Fig. 4):

$$\mathfrak{S}_{\text{tan.}} = \sin i, \qquad \mathfrak{S}'_{\text{tan.}} = \sin i', \qquad\qquad n \sin i = n' \sin i'.$$

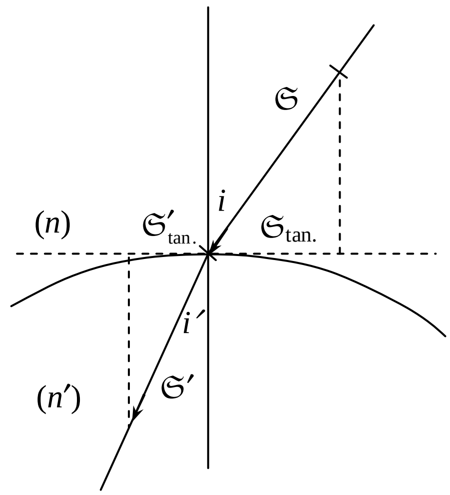

Figure 4.

In order to have a relation here that involves the entire vector $\mathfrak{S}$, as well, the normal components can be included in the equation. One gets:

$$n\,\mathfrak{S} - n'\,\mathfrak{S}' = \mathfrak{N}, \tag{12}$$

in which $\mathfrak{N}$ means a vector that is perpendicular to the boundary surface, and whose magnitude is simply the difference between the normal components of $n\,\mathfrak{S}$ and $n'\,\mathfrak{S}'$:

$$|\,\mathfrak{N}\,| = N = n\,\mathfrak{S}_{\text{norm.}} - n'\,\mathfrak{S}'_{\text{norm.}} = n \cos i - n' \cos i'. \tag{12a}$$

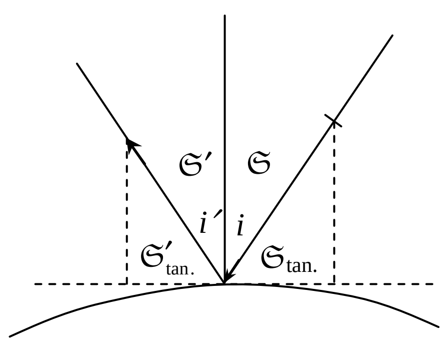

Figure 5.

The reflection law is also included in this form of the theorem. If the ray turns back into the first medium then $n' = n$, and equation (12) will say that:

$$\mathfrak{S}_{\text{tan.}} = \mathfrak{S}'_{\text{tan.}},$$

so the incident and reflected angles will be equal.

III. General theorems for geometrical optics

For the further theorems, one no longer needs to return to the basic condition rot $n\,\mathfrak{S} = 0$ now. Moreover, since, in reality, one must deal almost exclusively with homogeneous media and discontinuous boundaries between them, the condition rot $\mathfrak{S} = 0$ and the law of refraction will be satisfied. We thus turn to the purely geometrical optics that starts with the rectilinearity of rays and the law of refraction as empirical data, and knows nothing of wave surfaces.

§ 7. Malus's theorem

From this standpoint, the condition rot $\mathfrak{S} = 0$ is, at first, applicable to only those ray systems for which one knows that they are surface normal – i.e., parallel or radial rays. **Malus**'s theorem now asserts that the property is preserved for arbitrary reflections and refractions.

The proof follows from the law of refraction in the form (12); at all points of the refracting (or reflecting) surface, it is:

$$n\,\mathfrak{S} - n'\,\mathfrak{S}' = \mathfrak{N}.$$

If we construct the component of the rotation that is normal to the surface then we will have:

$$\mathrm{rot}_{\mathrm{norm.}}\,(n\,\mathfrak{S} - n'\,\mathfrak{S}') = n\,\mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{S} - n\,\mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{S}' = \mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{N}.$$

However, one now has:

$$\mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{N} = 0,$$

since $\mathfrak{N}$ has no tangential components and:

$$\mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{S} = 0,$$

since $\mathfrak{S}$ was assumed to be surface-normal. One then also has:

$$\mathrm{rot}_{\mathrm{norm.}}\,\mathfrak{S}' = 0,$$

and rot $\mathfrak{S}'$ must then be tangential to the refracting surface. Similarly, however, the condition (3) that was derived in § 1 – viz., that rot $\mathfrak{S}' \parallel \mathfrak{S}'$ – is true for $\mathfrak{S}'$. Since that is not consistent with the position of rot $\mathfrak{S}'$, rot $\mathfrak{S}'$ must vanish, and as a result, $\mathfrak{S}'$ will be surface normal.

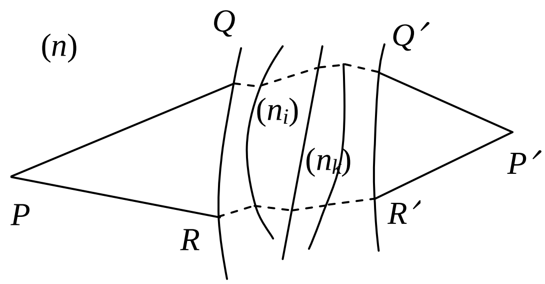

Figure 6.

§ 8. The constancy of the optical length

With the help of this theorem, one can now derived the law of the constancy of the optical length. Suppose a light bundle that emanates from a point P is refracted by some optical system in such a way that it converges to a point P'. We choose two rays, imagine that there is a surface that spans the space between them, and define the integral:

$$\int \text{rot}\, n\mathfrak{S}\, d\sigma$$

over that surface (Fig. 6). The surface is composed of a number of pieces, each of which lies entirely within a homogeneous medium, and is bounded by a piece of the two rays, and in general, by two curves on the two bounding surfaces of the medium. Inside of these surface patches, $n\,\mathfrak{S}$ will always be continuous, since n is constant, and the integral can then be converted into a boundary integral by using **Stokes**'s theorem:

$$\int \text{rot}\, n_i\, \mathfrak{S}_i\, d\sigma_i = \int (n_i\, \mathfrak{S}_i, d\mathfrak{s}).$$

If one adds all of these boundary integrals then one can remove the boundary lines between two pieces that relate to two media i, k, since, from the law of refraction, the two integrals that belong to the two boundary lines are:

$$\int (n_i\, \mathfrak{S}_i, d\mathfrak{s}) - \int (n_k\, \mathfrak{S}_k, d\mathfrak{s}) = \int (\mathfrak{N}_{ik}, d\mathfrak{s}),$$

and $(\mathfrak{N}, d\mathfrak{s})$, since $\mathfrak{N}$ is perpendicular to $d\mathfrak{s}$.

What remains is then precisely the line integral over the two rays itself:

$$\iint \text{rot}\, n\mathfrak{S}\, d\sigma = \sum n_i \iint \text{rot}\, \mathfrak{S}_i\, d\sigma_i = \int_{PQQ'P'} (n\mathfrak{S}, d\mathfrak{s}) + \int_{P'R'RP} (n\mathfrak{S}, d\mathfrak{s}).$$

Now, since rot $\mathfrak{S}$ vanishes in any homogeneous part, one will have:

$$\int_{PQQ'P'} (n\mathfrak{S}, d\mathfrak{s}) = \int_{PRR'P'} (n\mathfrak{S}, d\mathfrak{s}) . \tag{13}$$

Now, $\mathfrak{S}$ points in the direction of the path of integration on both rays, and therefore a total magnitude of 1 will enter into the calculations. One also has:

$$\int_{PQQ'P'} n\,ds = \int_{PRR'P'} n\,ds \tag{14}$$

then. However, this integral expresses nothing but the optical length along the two rays, or the time duration for the transition from P to P', up to the factor c. In fact, from (6), if $n = c / v$, and at the same time, $ds / \upsilon = dt$ then:

$$\int n\,ds = c\int \frac{ds}{\upsilon} = c\int dt .$$

In regard to this, we remark that equation (13) is true not only for actual light rays $PQQ'P'$ ($PRR'P'$, resp.), but for any two entirely arbitrary (and also curvilinear) connecting paths between any two points P and P'. In fact, the first use of the assumption that the direction of this connecting path coincides with that of our light ray $\mathfrak{S}$ was made in the transition from (13) to (14). Equation (13), by means of our rotation condition, generally expresses the idea that the line integral of the vector $n\,\mathfrak{S}$ is independent of the path.

§ 9. The eikonal and the limits of geometrical optics

We shall now exhibit the connection between our method of ray vectors and the method of the eikonal that was introduced by H. Bruns (the general **Hamiltonian** theory of characteristic functions, resp.).

The eikonal is a function of the coordinates of two points, namely, the line integral of the vector $n\,\mathfrak{S}$, that indeed for a given distribution of various media, from the remark at the end of the last paragraph, is independent of the path of integration. In fact, in connection with (13), one defines

$$E\,(P, P') = \int_P^{P'} (n\mathfrak{S}, d\mathfrak{s}) .$$

If one now fixes the point P then the eikonal will be purely a function of the position of P', and from the meaning of the line integral, one will have:

$$\text{grad}'\, E = n'\,\mathfrak{S}', \tag{14a}$$

in which the right-hand side and left-hand side both refer to the point P'. Likewise, when one fixes P' and varies:

(14b) $$\text{grad } E = - n\, \mathfrak{S}\,.$$

From its double dependency upon position, the function E will then be the *potential* of the two vectors $n\,\mathfrak{S}$ and $n'\,\mathfrak{S}'$, which is demanded directly by the condition rot $n\,\mathfrak{S} = 0$. It is then also identical to the phase function φ of § 4, up to a constant factor, except that it also includes the position of the starting point P as a variable.

One gets two differential equations for E from the relations:

$$|\,\mathfrak{S}\,| = 1 \qquad \text{and} \qquad |\,\mathfrak{S}'\,| = 1,$$

and it will follow from that, by (14a) and (14b), when one goes to the absolute values in those equations, that:

$$|\text{ grad } E\,| = - n,$$
$$|\text{ grad}'\, E\,| = n',$$

or

(15) $$\begin{cases} \left(\dfrac{\partial E}{\partial x}\right)^2 + \left(\dfrac{\partial E}{\partial y}\right)^2 + \left(\dfrac{\partial E}{\partial z}\right)^2 = D(E) = n^2, \\ \left(\dfrac{\partial E}{\partial x'}\right)^2 + \left(\dfrac{\partial E}{\partial y'}\right)^2 + \left(\dfrac{\partial E}{\partial z'}\right)^2 = D'(E) = n'^2. \end{cases}$$

From a casual verbal remark of **Debye**, these differential equations can give way to *the differential equation of wave optics by passing to the limit.*

It reads:

$$\Delta u + k^2\, u = 0,$$

in which u is the light excitation (free of the factor $e^{i\nu t}$), as measured by any of its electric or magnetic components, and k is the "wave number" $2\pi / \lambda$, and is thus *a very large number* when measured in cm^{-1}.

The introduction of the eikonal into this equation comes about by way of the following argument: The light rays that geometrical optics works with are to be regarded physically, not as infinitely-thin tubes, but as pieces of a plane wave whose temporal extension must be large in comparison to the wave length of light. Indeed, by the abstraction that is implied by the words "light ray," the peculiarity exists that we might drift far from the ideal case of a rectilinearly-bounded light ray by seeking to realize it in isolation by an infinitely-thin light bundle. Now, the wave-optical Ansatz for a plane wave is:

$$u = u_0\, e^{ik\,(\alpha x + \beta y + \gamma z)},$$

in which $n\,(\alpha x + \beta y + \gamma z)$ is the light path along the propagation of the wave. While recalling the meaning of the eikonal, we then correspondingly set the light path equal to:

$$u = u_0\, e^{i\frac{k}{n}E}.$$

In that way, we then see to determine E as a function of position in such a way that this Ansatz satisfies the wave equation approximately, if one keeps in mind the magnitude of k, and one no longer regards u_0 as strictly constant, but as a slowly-varying function of position that is appreciably constant only along segments whose order of magnitude is that of the wave length.

One now calculates:

$$\text{grad}\, u = e^{\frac{ik}{n}E}\left(\frac{ik}{n}u_0 \text{grad}\, E + \text{grad}\, u_0\right)$$

and

$$\text{div grad}\, u = \Delta u = e^{\frac{ik}{n}E}\left(-\frac{k^2}{n^2}u_0\, D(E) + \frac{ik}{n}u_0 \Delta E + \frac{ik}{n}(\text{grad}\, u_0, \text{grad}\, E) + \Delta u_0\right).$$

If one substitutes this into the differential equation for wave optics then one will get the differential equation for E:

$$0 = \frac{k^2 u_0}{n^2}\left\{n^2 - D(E) + \frac{in}{k}\Delta E + \frac{in}{u_0 k}(\text{grad}\, u_0, \text{grad}\, E) + \frac{n^2}{u_0 k^2}\Delta u_0\right\}. \tag{16}$$

The factor $1 / k$ is considered to be small in geometrical optics; in that approximation, (16) will then go to the differential equation for the eikonal that was already given above:

$$0 = n^2 - D(E).$$

If one integrates this – e.g., for the case of a variable n – then the course of the light ray will follow from the relation (14a):

$$n\, \mathfrak{S} = -\, \text{grad}\, E.$$

However, at the same time, one can (also according to Debye) extract from (16) the cases in which neglecting terms that are endowed with $1 / k$ is not permissible, namely, when the numerator of one of those terms has the order of magnitude k. That will yield the following two exceptional cases:

1. *ΔE becomes large.* One has $\Delta E = n$ div $\mathfrak{S}$; however, from its vectorial-geometric definition, $d\mathfrak{S}$ is the difference between two neighboring cross-sections of an infinitely-thin ray bundle, divided by the space that is enclosed between them. If one calculates the cross-section in question as the product of its lengths in the directions of principal curvature then its difference will be equal to its so-called mean curvature (viz., $1 / r_1 + 1 / r_2$, where r_1, r_2 are the principal radii of curvature of one of the cross-sections), multiplied by the volume of the enclosed space. One then has simply:

$$\operatorname{div} \mathfrak{S} = \frac{1}{r_1} + \frac{1}{r_2}.$$

div $\mathfrak{S}$ then becomes very large at all locations where the bundle, in the sense of geometrical optics, has a *focal surface*, on in particular, a *focal line,* or a *focal point,* at which one or both values r_1, r_2 vanish.

2. grad u_0 *becomes large*, i.e., u_0 is no longer slowly-varying, but varies considerably along small distances; that happens, e.g., for geometric shadow boundaries.

In both cases, the laws of geometrical optics are no longer a sufficient approximation, and diffraction phenomena appear.

It seems to us that this simple argument can, to certain degree, replace the circumstantial considerations by which, e.g., **Kirchhoff** restricted the validity of the rectilinear ray path in his first lectures on optics.

IV. Two examples of special applications

§ 10. The sine law

If a point P of an axially-symmetric optical system were mapped to a point P' by a light bundle with a finite aperture, and at the same time, a small circular disc of radius l that is perpendicular to the axis at P were mapped to a likewise small circular disc of radius l' that is perpendicular to the axis at P' then for all rays that start from P and end at P', one will have the **Abbe** relation:

$$n\, l \sin u = n'\, l' \sin u', \tag{17}$$

if u and u' mean the angles between the ray and the axes.

Due to the rotational symmetry, one needs only to consider a meridian plane.

Let (1) be the axial ray from P to P', let (2) be any other ray that arrives at P' after refracting from P, and let (3) and (4) be two rays that start from the neighboring point P_1 and arrive at P_1', of which, (3) lies close to (1), and (4) lies close to (2) (Fig. 7). From the constancy of the optical lengths, the optical lengths will satisfy:

$$(1) = (2), \qquad (3) = (4). \tag{18}$$

Now, if x and x are the abscissas of P and P, resp., as measured along the axis, then the optical lengths will satisfy:

$$(1) = (2) = E\,(x, 0, x', 0)$$

and

$$(3) = (4) = E\,(x, l, x', l'),$$

and therefore:

$$(3) - (1) = E\,(x,\,l,\,x',\,l') - E\,(x,\,0,\,x',\,0) = \left(\frac{\partial E}{\partial y}\right)_1 l + \left(\frac{\partial E}{\partial y'}\right)_1 l',$$

in which the differential quotients on the right-hand side are taken for the ray path (1).

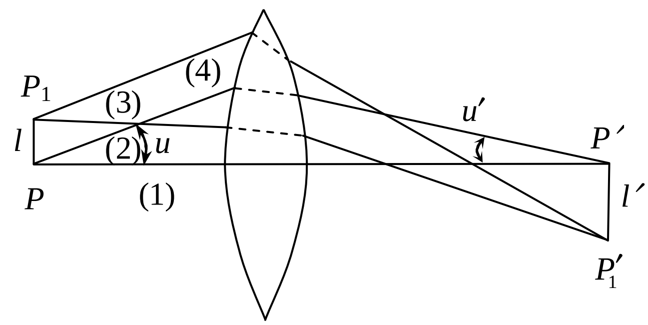

Figure 7.

However, one has:

$$-\frac{\partial E}{\partial y} = n\,\mathfrak{S}_y\,, \qquad \frac{\partial E}{\partial y'} = n'\,\mathfrak{S}'_y\,,$$

and both y-components are equal to zero for the ray path (1). One will therefore also have:

$$(3) - (1) = 0,$$

and thus, due to (18), also:

$$(4) - (2) = 0.$$

However, if we calculate the infinitely-small difference between the light paths (4) and (2) and the eikonal then, just as before, that will give:

$$(4) - (2) = 0 = \left(\frac{\partial E}{\partial y}\right)_2 l + \left(\frac{\partial E}{\partial y'}\right)_2 l',$$

in which the differential quotients of E are now taken for the ray path (2); that is:

$$0 = -\,n\,\mathfrak{S}_y \cdot l + n'\,\mathfrak{S}'_y\;l',$$

but, as is clear from the figure, $\mathfrak{S}_y$ ($\mathfrak{S}'_y$, resp.) are equal to sin u (sin u', resp.), so one will have:

$$0 = n\,l\sin u - n'\,l'\sin u',$$

which was to be proved.

An entirely analogous theorem can be derived in the same way when one displaces P_1 to P through the small distance l in the axial direction, instead of the perpendicular one, and poses the requirement that under the map by a light bundle of finite aperture, it should correspond to a point P_1' that is displaced by l' and then also lies on the axis. The single difference consists of the fact that the differential quotients of the eikonal with respect to x and x' enter in place of the ones with respect to y and y', resp. That will then yield:

$$(3) - (1) = - n\,l + n'l',$$
$$(4) - (2) = - n\,l \cos u + u'l' \cos u',$$

so one will have the relation:

$$n\,l\,(1 - \cos u) = n'l'(1 - \cos u'). \tag{19}$$

It emerges from the incompatibility of this condition with (17) that no optical system can sharply map laterally and axially-neighboring points at the same time by a wide-aperture bundle.

If one compares this proof with the proof of **Hockin** that was presented by **von Drude** ([1]) then one will recognize immediately that both of them rest upon entirely the same notions, but that **Hockin**'s proof seems to be somewhat more complicated by the introduction of some unnecessary geometric specializations. Our proof is also identical to the one that was given by **Schwarzschild** ([2]), in principle. **Straubel** ([3]) emphasized the connection between the sine law and a general reciprocity law in geometrical optics.

§ 11. Refraction by a spherical surface for oblique incidence

A ray bundle that is refracted by a spherical surface with oblique incidence will no longer have any rotational symmetry, and, from **Sturm**'s theorem, will then possess two focal points. Their distance from the refracting surface shall be determined.

It is clear that when the incident bundle starts from a point (i.e., is "stigmatic"), the two distinguished planes of the refracted bundle can only be the incidence plane of the principal ray and the plane that is perpendicular to it. However, the same thing will also be true when the incidence plane itself is already "astigmatic," in such a way that its distinguished planes simultaneously coincide with the incidence plane and the plane that is perpendicular to it. When we consider this for the case of an astigmatic bundle, which is important in the applications, we shall denote the starting point of the ray in the plane of incidence by another symbol (T) from the starting point (S) of the ray in the plane that is perpendicular to it, in such a way that T and S will refer to the two distinct focal points of the astigmatic incident bundle. The neighboring rays in the *plane perpendicular to the plane of incidence*, which will be referred to as *sagittal rays*, lie on the surface of a cone whose vertex lies at the focal point S and whose axis goes through the center of the sphere C. On the grounds of symmetry, it will then be obvious that its meeting point S'

(1) **P. Drude**, *Lehrbuch der Optik*, 2nd ed., pp. 55.
(2) **K. Schwarzschild**, Abhandl. d. Göttinger Ges. d. Wiss. 4, no. 1, 1905; *Untersuchungen zur geometrischen Optik I*, § 3.
(3) **R. Straubel**, Physik. Zeitschr. **4** (1920), 114.

after the refraction will likewise lie along that axis. One then needs only to look for the point of intersection of the principal ray with the axis.

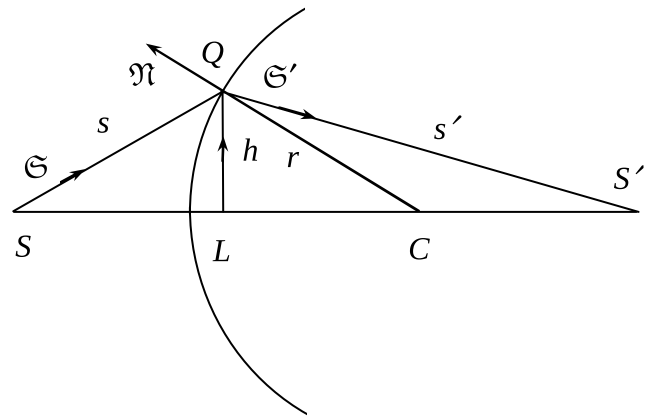

Figure 8.

If Q is the point at which the principal ray is refracted (Fig. 8) then the law of refraction:

$$n\,\mathfrak{S} - n'\,\mathfrak{S}' = \mathfrak{N}$$

will be valid at Q. We drop the perpendicular $QL = h$ from Q onto the axis and find the component of this equation in the direction of the perpendicular, which is measured in the sense of LQ being positive:

$$n\,\mathfrak{S}_h - n'\mathfrak{S}'_h = \mathfrak{N}_h\,.$$

If we now denote the distances SQ and $S'Q$ by s and s', respectively, and the radius of the sphere by r then, since one has $|\,\mathfrak{S}\,| = |\,\mathfrak{S}'\,| = 1$ and $|\,\mathfrak{N}\,| = N$, one will have:

$$+\,\frac{nh}{s} + \frac{n'h}{s'} = +\,N\,\frac{h}{r}$$

or

$$\frac{n}{s} + \frac{n'}{s'} = \frac{N}{r}\,.$$

The derivation is somewhat more complicated for the rays that lie in the *plane of incidence* and start from the focal point T – viz., the so-called *tangential rays* – since their meeting point T' is not determined by finite relations, but must be found by a differential consideration. Let $TQ'T'$ (Fig. 9) be a tangential ray that is close to the principal ray $T\,Q\,T'$. Differentiation of the law of refraction yields:

(20) $$n\,d\mathrm{S} - n'\,d\mathfrak{S}' = d\mathfrak{N}\,.$$

We find the component in the direction of the infinitely-small chord QQ'.

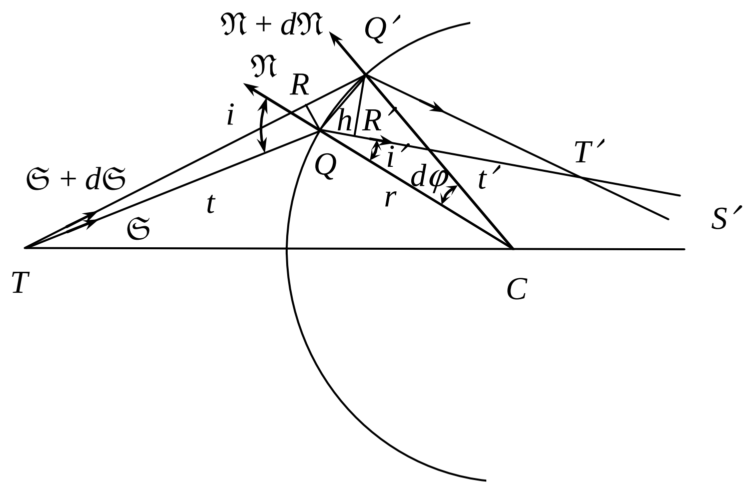

Figure 9.

Since $\mathfrak{S}$ is a unit vector, $d\mathfrak{S}$ must be perpendicular to $\mathfrak{S}$ and equal to the angle $Q\,T\,Q'$, so when one makes $R\,T = T\,Q$, to a sufficient approximation:

$$d\mathfrak{S} = \frac{QR}{TQ} = \frac{QR}{t} = \frac{h\cos(RQQ')}{t},$$

$$d\mathfrak{S}' = \frac{Q'R'}{T'Q'} = \frac{Q'R'}{t'} = -\frac{h\cos(R'Q'Q)}{t'},$$

and the components in the h direction are:

$$d\mathfrak{S}_h = \frac{h\cos^2(RQQ')}{t},$$

$$d\mathfrak{S}'_h = -\frac{h\cos^2(R'Q'Q)}{t'}.$$

Now, if QR, QQ', $Q'R'$ are infinitely-small chords in three circles with the centers T, C, and T' and radii $TQ = t$, $CQ = r$, and $T'Q' = t'$, resp., then they will be almost perpendicular to the associated radii, and therefore the angles $(R\,Q\,Q')$ and $(R'Q'Q)$ will agree, up to quantities of order h, with the angles between the radii $T\,Q$ and $C\,Q$ ($T'Q'$ and $C\,Q'$, resp.), that is, the angle of incidence i and the angle of refraction i'. Therefore, the cosines of these angles will equal $\cos i$ ($\cos i'$, resp.) to the same error. We then have, up to quantities of order two:

$$d\mathfrak{S}_h = \frac{h\cos^2 i}{t},$$

$$d\mathfrak{S}'_h = -\frac{h\cos^2 i'}{t'}.$$

On the right-hand side of (20), one finds the h-component of the differential $d\mathfrak{N}$. If one decomposes $d\mathfrak{N}$ parallel and perpendicular to $\mathfrak{N}$ then the parallel component will be equal to the change in magnitude dN, and the perpendicular one will equal $N\,d\varphi$, if $d\varphi$ is the change in direction of $\mathfrak{N}$. The latter component is parallel to h, so one will have:

$$d\mathfrak{N}_h = N\,d\varphi = \frac{N\,h}{r}.$$

We then have:

$$n\,\frac{h\cos^2 i}{t} + n'\frac{h\cos^2 i'}{t'} = \frac{N\,h}{r},$$

or, after dividing by h, and up to quantities of first order in h:

$$n\,\frac{\cos^2 i}{t} + n'\frac{\cos^2 i'}{t'} = N\frac{1}{r}.$$

Finally, one might replace N with its value in (12a):

$$N = n\cos i - n'\cos i'.$$

The well-known formulas then arise:

For sagittal rays:

$$\frac{n}{s} + \frac{n'}{s'} = \frac{n\cos i - n'\cos i'}{r}, \tag{21}$$

For tangential rays:

$$\frac{n\,\cos^2 i}{t} + \frac{n'\cos^2 i'}{t'} = \frac{n\cos i - n'\cos i'}{r}. \tag{22}$$

Whereas in the first three sections, our vector-analytic method proved to be especially suitable for giving the simplest expressions to the general concepts and theorems of geometrical optics, from the examples of the last section, one might expect that they might also prove themselves in the treatment of some special problems in geometrical optics, due to their adaptability to the special behavior of problems in which one must carry our calculations by taking the components along a certain direction and due to the elimination of all superfluous calculations.

Munich, Institut f. theoretische Physik, March 1911.

(Received on 17 April 1911)