"Räumliche Kollineation bei optischen Instrumenten," Zeit. Math. Phys. **46** (1901), 376-382.

Spatial collineation in optical instruments

By **F. KLEIN** in Göttingen

Translated by D. H. Delphenich

The result that will be derived in what follows is not new, in and of itself, but is already found in the treatise (which was discussed in the foregoing notice) by **Bruns** on the eikonal. However, whereas it appeared there only occasionally in the middle of extensive analytical developments, here it shall be derived directly by merely geometrical considerations. The problem is to decide what relationship might exist between the object and image in an *absolute* optical instrument, i.e., an instrument that recombines all rays that start from an arbitrary point in object space into precisely one point in the image space.

The closely-related remark that one makes from the geometric standpoint is that the relationship between object space and image space must be *collinear* in any event (cf., **Czapski**, *Theorie der optischen Instrumente nach Abbe*, Breslau, 1893). In fact, the two spaces are indeed related to each other from the outset in such a way that every straight line in the one space (i.e., every light ray) always corresponds to a straight line in the other space (i.e., the associated light ray). However, since the relationship should also be, by assumption, a *pointwise* one, one will arrive at a collineation in a known way on the basis of **Moebius**'s net construction. In that way, as far as the function-theoretic character of the map of the one space to the second one is concerned, one must assume nothing but the *continuity* of the relationship. The fact that the map is an *analytic* one is deduced from the line of reasoning that implies that it is a collineation as an incidental result.

One now comes to the realization that the collineation that is thus allowed is one of a very special type. To that end, I shall draw upon a tool that is really quite familiar to the geometers, but hardly ever finds employment in optics, namely, the consideration of *imaginary* straight lines or light rays. (In that way, "light ray" and "straight line" shall be synonyms, i.e., nothing further shall be said of the direction in which light traverses the straight line.) Indeed, I shall consider the course of refraction under the assumption that the incident ray is a *minimal line*, i.e., an imaginary straight line that intersects the spherical circle. In that way, I will apply the same formulas for imaginary lines that I apply to real ones. However, in order to eliminate all doubt that might arise in that regard from the outset, I would like to expressly assume (and this implies no sort of restriction *in practice*) that all refracting surfaces in the instrument are *algebraic surfaces*.

On the basis of the convention that was made in that way, we next propose the elementary law of refraction: For a minimal line, the sine of the angle that it makes with the surface normal is known to be infinitely large, and conversely, a minimal line is characterized by the demand that its sine must be infinitely large. It then follows that *when the incident ray traverses a minimal line,* *the same thing must be true for the refracted ray*. With that conclusion, we already basically have an adequate foundation for the following argument. However, for the sake of precision, we must first make a small detour:

There are *two* minimal lines that run through the plane of incidence at the point where the incident ray meets it: One of them coincides with the incident ray, while the other coincides with its mirror image. *Which of those two lines represents the refracted ray remains undetermined*. Namely, the law of refraction includes a square root when one expresses it in Cartesian coordinates, but nothing can be said about its sign here, where we work with imaginaries. There is no point in my explaining that in detail here, but rather I will say briefly only that *a minimal line will be converted into two minimal lines by any refraction* (one of which coincides with the incident ray and the other, with its mirror image). If we have *n* refracting surfaces then we will have 2^n minimal rays as the final result of the refraction. One of them always coincides with the original minimal ray, so it has passed through the instrument "as a Röntgen ray would," while one will get the other one when one lets an arbitrary number of reflections occur in the *n* successive refracting surfaces.

Now, the complication that was thus discussed will not prevent one from reaching a simple conclusion in regard to the collinear map that the assumed absolute instrument mediates. In fact: A collinear map is single-valued for all lines in space. There will then be 2^n minimal rays in it that arise from an incident minimal ray by refraction in the instrument, but *only one* can participate. The whole complication will go away as long as we restrict ourselves to the consideration of the collinear map that we speak of. We say briefly:

The collineation between object space and image space is arranged such that every minimal line in the former will produce a minimal ray in the latter,

or even more briefly:

The spherical circle in the object space goes to the spherical circle in the image space.

However, we would like to say that our collineation is, in fact, a very special one, namely, that it is a *similarity transformation* ([1]). That similarity transformation can be a direct or inverse one (i.e., one that switches left and right).

With that, we already have the main part of the result to be derived. We will complete it when we establish the *modulus of the similarity transformation*. For the sake of generality, I would like to assume that the speed of light *c* in the object space is different from the speed of light c' in image space. The theorem then says simply that *the dimensions in the object space relate to the dimensions in the image space as c does to* c' ([2]). In particular, if $c = c'$ then the object space and

([1]) Cf., **Bruns**, *Eikonal*, page 370.

([2]) That theorem is found between the lines in **Bruns**. Bruns wrote to me in regard to it: "The modulus μ was found to be equal to *E* in line 5 of page 370 (of his treatise on the eikonal). However, the quantity *E* is, as the theorems in the text between formulas (91) and (92) show us, identical to the quotient $h : N$ of the spatial indices that is quoted in (51.*b*)."

image space will have the same dimensions, so *they are directly or mirror-image congruent* (which is the actual result that will be derived here).

In order to prove that, we just bring more real space elements into consideration and accept the conceptual way of connecting things that we spoke of in the previous notice ("Über das Brunssche Eikonal"). In that way, we will express ourselves as if the similarity transformation that our instrument mediates is a direct one. Should it be an inverse one, then we would complete the instrument by adding a plane mirror, and in that way convert an initially inverse similarity into a direct one.

We would now like to simply consider the *time* that the light requires in order to go from an arbitrary object point (that I would like to call x, y, z) to the corresponding image point (which I shall call x', y', z'). That time must be the same for all rays that emanate from x, y, z. Otherwise, not all of those rays could recombine at x', y', z', which was, however, assumed, but rather, from the principle of **Johann Bernoulli**, the only rays that would connect the object point to the image point would be the ones for which that time was a minimax. I can then denote the time in question as a function of x, y, z alone by:

$$T = \mathrm{X}\,(x, y, z)\,.$$

Now, let x_1, y_1, z_1 and x_2, y_2, z_2 be two new object points that protrude from the point x, y, z along segments of equal length r (but shall otherwise be assumed to be arbitrary). We shall temporarily denote the still-unknown similarity ratio of the image space to the object space by λ. The image points x_1', y_1', z_1' and x_2', y_2', z_2' of our new object point will both protrude from the image point x', y', z' of the original object point by $\lambda\, r$ then. I will now express myself by saying that I assume that the light ray that runs from x, y, z to x_1, y_1, z_1 further passes through our instrument and reaches the associated image point after a finite path length ([1]). For the *direct* similarity, it will first meet x', y', z' and then x_1', y_1', z_1'. The time that the light requires in order to go from x, y, z to x_1', y_1', z_1' is r / c, and the corresponding time that the segment from x', y', z' to x_1', y_1', z_1' accounts for is $\lambda\, r / c'$. We conclude that the function X has the value:

$$\mathrm{X}\,(x_1, y_1, z_1) = \mathrm{X}\,(x, y, z) - \frac{r}{c} + \frac{\lambda\, r}{c'} \tag{1}$$

for the point x_1, y_1, z_1.

Naturally, in precisely the same way, one will get (under the corresponding assumptions):

$$\mathrm{X}\,(x_2, y_2, z_2) = \mathrm{X}\,(x, y, z) - \frac{r}{c} + \frac{\lambda\, r}{c'}\,. \tag{2}$$

Therefore:

$$\mathrm{X}\,(x_1, y_1, z_1) = \mathrm{X}\,(x_2, y_2, z_2)\,. \tag{3}$$

([1]) There is nothing fundamental in that assumption, except that it fixes the signs that will appear later.

However, the points x_1, y_1, z_1 and x_2, y_2, z_2 that are used here are essentially two entirely-arbitrary object points. That is because the condition under which they were originally introduced – namely, that they have same distance r from another object point x, y, z – imposes no restriction, in reality, and the other assumptions that we made had only the goal of simplifying the notation. *It will then follow that the time* $\mathrm{X}(x, y, z)$ *is the same for all object points.* It is constant that is characteristic of our "absolute" instrument. However, we will also have that $\mathrm{X}(x_1, y_1, z_1)$ [$\mathrm{X}(x_2, y_2, z_2)$, resp.] is equal to $\mathrm{X}(x, y, z)$ in (1) [(2), resp.], from which it will follow that $\lambda = c' / c$, which was to be proved.

With that, the question that was posed to begin with is answered completely. The result is a bit disappointing. In order to keep the assumption $c = c'$, the instrument will act like a plane mirror or a composition of several plane mirrors. It will be as useless for a telescope as it is for a microscope. There will be nothing to change here. What I must further add relates more to the abandonment of a mathematical reservation that one might have in regard to the validity of the result.

Namely, the result apparently contradicts the well-known fact that the object points and image points that lie on the axis of an optical instrument correspond to each other linearly *in the most general way* and that, consistent with that notion, for small solid angles of the field of view, one can then speak of the approximation of a collinear map of an object point close to the axis to its corresponding image point that is certainly not a similarity transformation, or even a congruent transformation. I will shortly show that this contradiction will disappear when one suitably clarifies the way that the cited fact can come about ([1]).

To that end, as we usually do, we shall be content to direct our attention to those rays in the object space that lie in an arbitrary meridian plane that goes through the axis of the instrument. The corresponding rays of the image space will fill up the same meridian plane. *We have a relation between the rays of two planar fields of rays*. As I asserted, it will now suffice to assume that this relationship is an *analytic* one and that we need to keep only *the linear terms in the Taylor development* when considering the neighborhood of an isolated ray in the first approximation in order to get all of the relations that are exhibited for the axis of the instrument (its neighborhood, resp.), as well as the ones that relate to the rays of the individual meridian plane, in their most general form. (The axis has absolutely no distinguishing features inside of the individual meridian plane in that way. It acquires its special position only because of the fact that it simultaneously belongs to all meridian planes.)

In fact, in order to focus upon more familiar relationships, for the moment, one substitutes two planar fields of points for the two planar fields of lines. Let their mutual relationship be given by the *analytical* equations:

$$x' = \varphi(x, y), \quad y' = \psi(x, y).$$

If one then treats only those points (x, y) that lie close to a fixed location (x_0, y_0) then one can write, in the first approximation:

([1]) Here, I can also refer to **Bruns**: *Eikonal*, page 410, formula (176).

$$x' = x'_0 + \left(\frac{\partial \varphi}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial \varphi}{\partial y}\right)_0 (y - y_0),$$

$$y' = y'_0 + \left(\frac{\partial \psi}{\partial x}\right)_0 (x - x_0) + \left(\frac{\partial \psi}{\partial y}\right)_0 (y - y_0).$$

One usually expresses that (e.g., in cartography) by saying: *A neighborhood of the point* x_0, y_0 *will be mapped affinely to a neighborhood of the point* x'_0, y'_0 *in the first approximation.* In particular, the pencil of directions of advance $\frac{y - y_0}{x - x_0}$ that emanate from x_0, y_0 *will be mapped projectively in the most general way* to the pencil of directions of advance $\frac{y' - y'_0}{x' - x'_0}$ that emanate from x'_0, y'_0 (which is not merely an approximate statement, but an exact one).

In the developments and statements that are thus given, one needs only to replace the point x, y [x', y', resp.] with a straight line, using the principle of duality, in order to obtain the theorem that will be true for plane fields of rays (the optical relationships that exist inside of the individual meridian planes in the neighborhood of the axis of the instrument, resp.). (One must then carry out an extended investigation for light rays that run skew to the axis of the instrument.)

The stated apparent contradiction has now been resolved by the fact that the considerations that we have now made have nothing at all to do with the previous ones that were based upon the **Moebius** net construction. Our new considerations start from the possibility of a **Taylor** development (the assumption that we can truncate it after the linear terms, resp.), so they will be correct only to the extent that we treat the general correspondence of points close to the axis, and they belong to the realm of the mathematics of approximation, moreover, while the **Moebius** net construction has the character of the modern mathematics of precision. It mainly works with only finitely-different lines and assumes nothing from the outset besides the continuity of the map in question. The two approaches could not be more different. The impression that one might be dealing with related arguments is created by only the superficial fact that both of them conclude with a linear relationship.

____