"Über die Geometrieen in denen die Geraden die Kürzesten sind," Math. Ann. 57 (1903), 231-264.

# On the geometries in which the lines are shortest (\*)

By

GEORG HAMEL in Karlsruhe.

Translated by D. H. Delphenich

### **Table of Contents**

Introduction:	The problem and the axioms that will be used	2
	The historical basis for the problem	3
	Chapter I: Geometries in the plane.	
§1.	The "ideal elements and their "naming" conventions	5
§ 2.	The "length"	6
§ 3.	The monodromy axioms	13
§ 4.	The normal line: elliptic geometries	15
§ 5.	Examples. The Minkowskian and Hilbertian geometries	19
§ 6.	On the effect of discontinuities in the function <i>w</i>	21
§ 7.	Generalizing the concept of an "H-curve"	22
	Chapter II: Geometries in space.	
§ 8.	The structure of the postulated geometries	23
§ 9.	The monodromy axioms	27
§ 10.	Examples	28
§ 11.	On the differential equation $\partial^2 h = \partial^2 h$	
	On the differential equation $\frac{\partial u}{\partial q} = \frac{\partial v}{\partial v} \frac{\partial p}{\partial p}$	29
§ 12.	On the geometric interpretation of the results. Extending the concept of "H-surface"	30

<sup>(\*)</sup> The present article is not just an excerpt from the dissertation (Göttingen, 1901) that appeared under the same title. The line of reasoning and the main results indeed remain the same, but they have been completely reworked at some essential places, namely, the Introduction and Sections 1, 2, 8. § 4, which treats elliptic geometries, is completely new, At the same time, I hope that many of the details have been made more concise and unambiguous than they were in my dissertation.

#### Introduction.

The goal of the following investigations is this: Exhibit all geometries that satisfy the following requirements:

A) The axioms of "linking" and "grouping" (Verknüpfung and Anordnung), in Hilbert's (\*) terminology, or the "projective" axioms in Schur's (\*\*) terminology.

If we restrict ourselves to just the plane then we can add **Desargues** theorem as an axiom (\*\*\*).

B) The following *continuity* axiom, which includes both the Archimedean (<sup>†</sup>) axiom and the completeness axiom (<sup>†</sup>) in this conception of things:

"If an infinite sequence of points P are generated inside of a line segment by geometric constructions [in accord with Axioms (A)] in such a way that each point P lies between the previous one and the following one then the aforementioned sequence shall possess a well-defined limit point inside of or at the ends of the line segment, i.e., a point that separates all of the points P in the sequence that have been overtaken once from all of the points that were never attained by the point-sequence P." (<sup>††</sup>)

C) The "linear congruence axioms," which deal with the removal (*Abtragen*) of line segments (IV, 1, 2, 3, with Hilbert's numbering; that notation shall always be used in what follows). However, the first congruence axiom will be modified in the following way:

There might be points from which it might not be possible to remove line segments along all, or perhaps individual, lines that go through them. We would like to briefly call such points "singular" or "relatively singular," respectively. All other points will be called "regular", but the ones that are not simply singular will be called "attainable."

However, the appearance of singular points might be restricted in the following way:

 $\alpha$ ) If one can remove a line segment a along a line from *one side* of a point on it then that shall also be possible on the other side.

 $\beta$ ) If one can remove a line segment a from a point on a line then it shall also be possible to remove all line segments at that location that are "smaller" than a, i.e., line segments whose points all lie inside of *a*.

Hilbert: "Grundlagen der Geometrie," Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in (\*) Göttingenm Teil I (Teubner, 1899).

 <sup>(\*\*)</sup> Schur: "Ueber die Grundlagen der Geometrie," Math
 (\*\*\*) See Hilbert: "Grundlagen," Chap. V, among others. Schur: "Ueber die Grundlagen der Geometrie," Math. Ann. 55. One will also find further bibliography there.

 $<sup>(^{\</sup>dagger})$ Hilbert: "Grundlagen," Axiom V, § 8; In addition: "Ueber den Zahlbegriff" Ber. deutschen Math.-Ver. (1900); finally, the translation of the "Grundlagen" by Langel in the Annales de l'École normale (3) 17, pps. 122, 123.

Cf., Klein, "Ueber die sogenannten Nicht-Euklidische Geometrie," Math. Ann. 6, pp. 136. (††)

 $\gamma$ ) If two lines *a* and *b* are given (which might also lie inside of each other) with attainable points *A* (*B*, resp.) on each of them then the neighborhoods of *A* and *B* must be "comparable," i.e., there must be at least one line segment *AA*' on *a* that is congruence to a line segment *BB*' on *b*.

 $\delta$ ) There is a region of finite extent inside of which only regular points lie (\*).

D) The axiom: The straight line segments shall be the shortest link between two points.

E) An axiom of differentiability (see § 2, pp. 10).

What *will not* be required are:

a) The parallel axiom.

b) The first law of congruence or the congruence axiom VI.6, with Hilbert's numbering: "If the congruences AB = A'B', AC = A'C',  $\ll BAC = \ll B'A'C'$  are true for two triangles ABC and A'B'C' then the congruences:

$$\checkmark ABC \equiv \checkmark A'B'C', \quad \measuredangle ACB \equiv \checkmark A'C'B'$$

shall also be always fulfilled."

(We will never speak of the congruence of angles again in what follows.)

The gist of the entire problem statement lies in the requirement D). Since itarywas probably first posed by Archimedes (\*\*), we would like to briefly call it the "Archimedean requirement" in what follows. (In order to distinguish it from the aforementioned Archimedean axiom of continuity.)

So much for the statement of the problem in the introduction itself. As far as its historical context and its relationship to the calculus of variations are concerned, allow me to make the following remarks:

The geometries that we speak of here include not only elementary Euclidian geometry, but also the two so-called non-Euclidian geometries that can be constructed on the basis of a projective geometry [which will satisfy our requirements A) and B)] by introducing a suitable metric, as **Klein** first showed (\*\*\*). **Hilbert** (<sup>†</sup>) then exhibited a more general geometry in a letter to **Klein** 

<sup>(\*)</sup> The presentation here deviates essentially from the one that I gave in my dissertation. There, the concept of "length" was introduced directly as a number, by definition (§ 2). Here, it will be shown later in § 2 how all of the requirements of an arithmetic character can be derived from our Axiom C, which I introduced axiomatically in my dissertation (§ 2). On the other hand, I have restricted myself here to some simpler cases insofar as I shall likewise demand from the outset what I call the "strong monodromy axiom," i.e., that one must always have  $AB \equiv BA$ . For that, one might confer §§ 3 of my dissertation and this treatise.

<sup>(\*\*) &</sup>quot;Περι σφαιρας και κυλινδρου." Λαμβανομενον α'.

<sup>(\*\*\*)</sup> Klein: "Ueber die sogenannte Nicht-Euclidische Geometrie," Math. Ann. 4.

<sup>(&</sup>lt;sup>†</sup>) Hilbert: "Ueber die gerade Linie als kürzeste Verbindung zweier Punkte," Math. Ann. 34.

whose metric is arranged in such a way that the straight lines will remain shortest. Finally, **Minkowski** (\*) drew attention to another type of geometry that will satisfy our demands.

That is why the question of the most-general geometry of that type would probably seem to be justified in its own right (\*\*).

However, it also deserves further attention due to the fact that it defines a special case of the question that I would like to call the inverse problem in the calculus of variations:

"Given a system of differential equations, find the most general variational problem to which those differential equations will belong as the Lagrange equations."

The creation of that problem probably has mechanics to thank, and indeed Jacobi's investigations into Hamilton's principle, in particular. In the case where an ordinary differential equation is given, to my knowledge, the first solution was given by Imchenetzky (\*\*\*). Further works on that topic were published by **Königsberger** ( $^{\dagger}$ ), **Böhm** ( $^{\dagger\dagger}$ ), and in an especially-elegant form by **Hirsch** (<sup>†††</sup>), who also brought partial differential equations into his sphere of consideration. As far as systems of differential equations are concerned, one probably finds the first result in **Helmholtz** (<sup>‡</sup>); **A. Mayer** (<sup>‡‡</sup>) gave the proof that was missing from the latter. In general, those investigations went only so far as to give the conditions for when the differential equations would have the form of a Lagrange equation. Whether or not a system could be brought into that form by a suitable combination of the individual equations was not mentioned. The question was not pursued in any case beyond the purely-formal one either. No other conditions from the calculus of variations than the Lagrangian ones were ever discussed.

In the problem that is announced here, the greatest weight shall be placed upon precisely a discussion of the sufficient conditions for the calculus of variations, to the extent that the problem is one of the calculus of variations.

Minkowski: Geometrie der Zahlen, Teubner, Leipzig, 1896, Chap. I.

<sup>(\*)</sup> (\*\*) On this, confer: Hilbert: "Mathematische Probleme," Vortrag gehalten auf dem internationalen Mathemaiker-Congress Paris, 1900, pp. 15. Darboux: Leçons sur la théorie générale des surfaces, t. III, Paris, 1894, pp. 54.

<sup>(\*\*\*)</sup> Imschenetzky, "Sur la transformation d'une équation différentielle de l'ordre pair à la forme d'une équation isopérimétrique," Bulletin de St. Pétersbourg 31 (1887).

Königsberger: "Ueber die Principien der Mechanik," Berliner Berichte (1896), II. "Ueber die allgemeinen (†) kinetischen Potentiale," Crelle's Journal, 121.

 $<sup>(^{\</sup>dagger\dagger})$ Böhm: "The Existenzbedingungen eines von den ersten und zweiten Differentialquotienten der Koordinaten abhängigen kinetischen Potentiale," Crelle's Journal, 121.

Hirsch, "Die Existenzbedingungen des verallgemeinerten kinetischen Potentials," Math. Ann. 50, and: (†††) "Uener eine characteristische der Differentialgleichungen der Variations-Rechnung," Math. Ann. 49.

<sup>(\*)</sup> Helmholtz: "Ueber die physikalsiche Bedeutung des Prinzips der of kleinsten Wirkung," Crelle's Journal 100.

<sup>(‡‡)</sup> A. Mayer: "Die Existenzbedingungen eines kinetischen Potentials," Sächsische Berichte (1896), pp. 519.

#### CHAPTER I

#### Geometries in the plane.

## § 1. – The "ideal" elements and their "naming" conventions.

We imagine a certain system of points, lines, and planes that satisfy our requirements A) and B). As **Schur** (\*) has shown, we can then add "ideal" points, lines, and planes to those points, lines, and planes that fulfill the modified parallel axiom:

### Any two lines in a plane have one (and only one) common point.

The axioms of the first group remain preserved for the extended domain, just like the continuity axiom that emerges from it, namely, that a line segment of ideal points can be projected onto a line segment of actual points from an actual point. Let the precise conception of the notion of "between" that generally becomes necessary be implemented with the help of a "normal plane" (normal line, resp.), and indeed in the same way that **Dehn** (\*\*) did using the procedure of **Pasch** (\*\*\*).

We shall now go further with this completed geometry. Therefore, we shall distinguish between the original points, lines, and planes and the "ideal" ones by the term *actual* (<sup>†</sup>). (With our previous terminology, the "actual" points again split into "attainable" ones and "singular" ones.)

Now, it is a known consequence of Axioms A) and B) that each point of the (extended) plane (to which we shall now restrict ourselves to begin with) *can be associated with a real triple of numbers x, y, t in such a way that the point will be represented by the ratios of those three numbers, but the line will be represented by a homogeneous linear equation in the three numbers.* 

The consequence of Axiom B) and the modified parallel axiom is that one also has that, conversely, every triple of numbers (besides the triple x = 0, y = 0, t = 0) will belong to a point, and every linear equation whose coefficients are not all zero will belong to a line. In particular, we can succeed in showing that t = 0 represents the "normal line." (<sup>††</sup>)

That introduction of a number system shall only serve to characterize the points and lines, so it represents a sort of *naming convention*. However, it will not, by any means, define any sort of length.

<sup>(\*)</sup> Schur, "Ueber die Einführung der sogennanten idealen Elemente in die projective Geometrie," Math. Ann. **39** (1891).

<sup>(\*\*)</sup> **Dehn**, "Die Legendreschen Sätze über die Winkelsumme im Dreieck," Math. Ann. **53** (1900).

<sup>(\*\*\*)</sup> Pasch, Vorlesungen über neuere Geometrie, Teubner, Leipzig, 1882.

<sup>(&</sup>lt;sup>†</sup>) It might very well be the case that there are no ideal elements at all. The geometry will then have the elliptic type. The normal line must be introduced from the outset in order to make the concept of "between" precise.

<sup>(&</sup>lt;sup>††</sup>) In addition to the abundant literature on this subject that was cited before, let us mention: Klein: "Ueber die sogennante Nicht-Euclidische Geometrie," Math. Ann., Bd. **4**, **6**, **7**. Fiedler: *Darstellende Geometrie*, Bd. III, as well as the *Vierteljahrsschrift der nat. Ges. Zürich*, XV, 2 (1871). Anne Lucy **Bosworth**: "Begründung einer vom Parallelenaxiom unabhängigen Steckenrechnung," Diff. Göttingen, 1900.

In place of the triple x, y, t, we will mostly consider x / t, y / t to be coordinates and denote them by x, y.  $x = \infty$ , as well as  $y = \infty$ , will now mean points of the normal line (<sup>†</sup>).

If we initially exclude that normal line from further consideration [we will address it again later (§ 4)] then one can represent all lines with the help of a parameter s by the differential equation:

$$\frac{d^2x}{ds^2} = 0$$
 and  $\frac{d^2y}{ds^2} = 0$ 

in which we have set  $ds = +\sqrt{dx^2 + dy^2}$ .

In what follows, we would recommend using:

$$y\cos \theta - x\sin \theta = a$$

as the linear equation of a line. We can then call  $\mathcal{G}$  the *direction* of the line. In the event that it deviates from  $\pm \pi/2$ , we might also write:

$$y = \tan \theta \cdot x + b$$
 and  $\frac{d^2 y}{ds^2} = 0$ .

From now on, the concept of *continuous curve* might be introduced in such a way that we will understand it to mean a manifold whose x and y can be represented as continuous functions of a continuously-varying parameter.

## § 2. – The "length."

We shall now attempt to infer the consequences of axioms C) and D) (see Introduction).

a) If one can measure out a line segment  $OA_1 \equiv a$  along a line from a point *O* then one can measure out *a* along the line successively arbitrarily often:

$$OA_1 \equiv A_1 A_2 \equiv A_2 A_3 \equiv \dots$$

(Consequence of C,  $\alpha$ , and IV.1)

In that way, it can happen that the starting point O will once more be overstepped (when the line is a closed curve in our extended plane!) or there might be a certain limit point G that separates the points in the sequence  $A_1$ ,  $A_2$  from the other ones as a result of continuity.

We shall initially assume that a limit point G exists. The following theorems will then be true:

<sup>(†)</sup> The fact that the conception of continuity that we gave in B) coincides with the usual concept of continuity in the numerical manifold of x, y (except for the normal plane) follows from the fundamental theorem of projective geometry.

b) The concept of congruence cannot be applied to GO itself. In particular, GO is not congruent to itself. That is because otherwise, according to C. $\beta$ , one must also be able to measure out *a* from *G*, which would contradict the fact that *G* is the limit of the sequence  $A_1, A_2, \ldots$  The same thing will be true for any segment that includes the point *G* (according to C. $\beta$ ).

c) The limit G is the same for all segments that are smaller than a.

That is because if b < a then the theorem will be obvious when one can measure out b a finite number of times (say, n times) in such a way that one will have nb > a (consequence of the congruence axioms IV.3 and the continuity axiom). However, there must be a number n such that besides the limit point that belongs to b, which already lies inside of a, the concept of congruence can otherwise be already no longer applicable to a (Theorem b).

d) The limit G is the same for all points that lie between O and G (a consequence of Theorem c and IV.3).

e) It is no longer possible at all to measure out a segment from a limit point G along the line in question.

That is because if that were the case then one could, in particular, also measure out the segment g from G to O (according to  $C.\alpha$ ), say,  $GB \equiv g$ . However, there are places inside of GB from which one can measure out a to G, so a < g. Therefore, from (C. $\beta$ ), one must also be able to measure out a from G, which is not the case.

Thus, such a limit point G will be a "relative singular point" relative to the line OG in the event that it is an actual point at all. (See Introduction) In any event, all points between O and G are actual points. If actual points lie beyond OG that are not all singular in regard to the line OG then one must repeat the argument.

f) If a limit *G* does not exist then *O* will once more be overstepped by the point sequence  $A_1$ ,  $A_2$ , ... (and therefore repeatedly) then that will be true for all point-sequences  $B_1, B_2, ...$  that arise by continually measuring out another segment *b* along the same line. All points of the line will then be actual, attainable points on the line. We would like to say that the geometry that relates to those lines is "of elliptic type."

All of those theorems collectively, along with Axiom C. $\gamma$ , will now prove the following theorem:

g) Let  $s_1$  and  $s_2$  be two segments that are enclosed by relatively-singular points relative to their lines but include no relatively-singular points within themselves. One can then map  $s_1$  and  $s_2$  to each other congruently, and that map will be one-to-one and well-defined when one associates an attainable point A on the segment  $s_1$  with another such point B on the segment  $s_2$  and has established a sense of direction on both segments.

The same thing will be true in a somewhat-altered way for two lines, one (both, resp.) of which is of elliptic type, except that the map will first become single-valued when one specifies how often one must imagine that the starting point A(B, resp.) is supposed to be overstepped.

With that, one has achieved all theorems that admit the concept of "length."

I) We initially define length for a line with the attainable point P (according to C. $\delta$ , such a thing must exist) as follows:

We denote any segment OA with no relatively-singular points by the number 1. We give lengths of 2, 3, ... to the segments  $OA_2$ ,  $OA_3$ , ..., resp., that arise by continually measuring out the segment OA in succession. We understand the segment of length 1 / m (*m* is a whole number) to mean the segment OB that is produced by measuring out the segment 1 *m* times in succession. (Such a segment OB will exist as a result of continuity.)

The following theorem will then be true:

The segment of length 1 / m can be made arbitrarily small by increasing m, i.e., there is no segment OC that would be smaller than any segment 1 / m (m = 1, 2, 3, ..., ad inf.)

It is now entirely clear what we should understand a segment c to mean when c is a rational number. However, continuity also allows us to associate a segment with any irrational number. Conversely, any segment from O will also be associated a unique positive number as its "length."

II) For all other segments, we then define the length such that we establish that *the length will be invariant under congruent maps* (Theorem g). That definition is independent of the choice of the point that was called *A* in Theorem g (from IV.2).

We can now derive the following theorems (compare § 2 of my Dissertation):

1) Let the points 1:  $x_1$ ,  $y_1$  and 2 :  $x_2$ ,  $y_2$  be connected by a segment with no relatively-singular points. That segment will then acquire a length that is a positive function of the four variables  $x_1$ ,  $y_1$ ;  $x_2$ ,  $y_2$ . We denote that function by  $F(x_1, y_1; x_2, y_2)$ . Since one should always have  $AB \equiv BA$ , one will also have (\*):

$$F(x_1, y_1; x_2, y_2) = F(x_2, y_2; x_1, y_1).$$

This is a consequence of the definitions I) and II), as well as the congruence axioms IV.3.

The relation that was written down is also true when two points coincide, i.e., let:

$$F(x, y; x, y) = 0.$$

In addition, one can also infer the following from the general equation (2):

If we set  $x_2 = x_1 + s \cos \theta$ ,  $y_2 = y_1 + s \sin \theta$  (s > 0, see § 1) then  $F(x_1, y_1; x_2, y_2)$  will be a function that *always increases with s*.

<sup>(\*)</sup> This differs here from my Dissertation; see the remark on page 3.

3) *F is a continuous function of its four arguments* in a finitely-extended region in which only regular points lie.

 $\alpha$ ) If we initially alter the points 1, 2 only in such a way that they remain on the lines 1, 2, then the theorem will be a consequence of Theorem 2), as well as the definitions I) and II), that is easy to prove.

 $\beta$ ) If we would like to prove the theorem in general then would have to add Axiom D) (the Archimedean requirement).

However, we must first prove the following lemma:

If we measure out a segment  $\varepsilon$  along all lines from a point A inside of the aforementioned region then the endpoints of those segments will not come arbitrarily close to A in the sense of the naming convention.

In fact, if that were not the case then there would have to be a ray S in whose arbitrary neighborhood such endpoints would have to approach A arbitrarily closely (in the sense of the



naming convention). However, one now determines a point *B* on *S* such that  $AB \equiv \frac{1}{2}\varepsilon$ . One erects a perpendicular (in the sense of the naming convention) to *S* at *B* and measures out a segment *BD* (*BD'*, resp.) on both sides that is likewise of length  $\varepsilon / 2$ . (All of those operations are possible, from the assumption about the region in question.) As a result of Axiom D), all points of the triangle *ADD'* will have a distance from *A* that is less than  $\varepsilon$ . Therefore, none of the

aforementioned endpoints can be found inside of ADD'. Thus, none of them can simultaneously lie arbitrarily close to S and arbitrarily close to A.

Therefore, one can bound a finite region around A in the sense of the naming convention inside of which each point will possess a distance from A that is less than  $\varepsilon$ .

Now let *AB* be a segment of length *l* and let a region around *AB* be bounded inside of which no singular point lies. One can then give a region about *A* such that for each point *A'* inside of it one will have  $AA' < \varepsilon/4$ , in which  $\varepsilon$  is a given arbitrarily-small quantity. (Consequence of the Lemma)

If one now lets the points A' and B' move towards to the point A, B in some way inside of that region then from the Archimedean requirement, one will always have:



$$AB < A'B' + BB' + AA' < A'B' + \frac{\varepsilon}{2},$$
  
$$A'B' < AB + BB' + AA' < AB + \frac{\varepsilon}{2},$$

so:

$$|A'B'-AB|<\varepsilon\,,$$

and with that the continuity is proved.

However, in order to be able to work more conveniently, we would now like to introduce the following *axiom*:

E) The function F will possess the first four derivatives with respect to variables, in general.

Whether or not that axiom can possibly be proved or whether it is possible to get along without it shall not be discussed any further here.

In my dissertation, I then showed how as a result of theorems 1), 2), 3), and axiom E), the length dl of a line element ds that starts from the location x, y in the direction 9 can be represented as follows:

$$dl \equiv ds \cdot f(x, y, \tan \theta)$$

independently of which finite segment ds it belongs to, in which f is a positive, single-valued, and generally continuous and three-fold differentiable function of its arguments.

(In my dissertation, it is not required to be single-valued relative to the last argument tan 9.)

In the event that the segment in question does not run precisely parallel to the *y*-axis, it is convenient to apply the less-symmetric form:

$$dl = g(x, y, \tan \theta) dx$$
.

Naturally, one simply has  $g \cdot \cos \vartheta = f$ .

In general, we will now understand the length of an arbitrary curve segment that should nonetheless initially possesses a generally continuous tangent at each point to mean the expression:

$$\int f(x, y, \tan \theta) ds \quad \text{or} \quad \int g(x, y, y') dx, \text{ resp.}$$

That integral extends along the curve segment in question, such that one sets:

$$y = y(x);$$
  $y' = \tan \vartheta = \frac{dy}{dx}.$ 

That definition agrees with the one that was given for segments as a result of out theorems 1), 2), and 3).

From now on, we shall restrict the arbitrariness of the function f(g, resp.) quite appreciably by the demand that:

4) The function g(f, resp.) shall be chosen such that the length of a straight line segment shall be shorter than that of any other path of the aforementioned type that connects the endpoints of the segment and whose elements x, y, y' define g completely everywhere.

That is the content of the *Archimedean requirement* that was mentioned to begin with. The calculus of variations teaches us that it splits into the following sub-requirements:

4.a) The Lagrange equations for the variational problem that arises from  $\int g \, dx$  must assume the form:

$$\frac{d^2 y}{dx^2} = 0$$

Therefore, g must satisfy the partial equation:

$$\frac{\partial^2 g}{\partial p \,\partial x} + p \frac{\partial^2 g}{\partial p \,\partial y} - \frac{\partial g}{\partial y} = 0$$

in which one sets y' = p.

One obtains that partial differential equation when one exhibits the Lagrange equation and then considers the fact that one should have y'' = 0 (\*).

If one differentiates the differential equation for g with respect to p once more then one will get a linear partial differential equation for  $M = \frac{\partial^2 g}{\partial p^2}$  whose general integral will read (\*\*):

$$M = W(p, y - p x) .$$

*W* is an arbitrary function of the arguments *p* and y - p x. Thus, one initially gives *g* the form:

$$g = \int_{c}^{p} \int_{c}^{p} W(p, y - px) dp dp + p \cdot v(x, y) + w(x, y),$$

in which *c* is an arbitrary constant.

If one substitutes that result in the differential equation for g then one will get (\*):

<sup>(\*)</sup> Hirsch, Math. Ann. 49, § 7; Darboux, loc. cit., pp. 53, et seq.

<sup>(\*\*)</sup> **Darboux**, *loc. cit.*, pp. 58 and 59.

$$g = \int_{c}^{p} \int_{c}^{p} W(p, y - px) dp dp + \frac{\partial u(x, y)}{\partial x} + p \frac{\partial u}{\partial y}$$

after an easy calculation, in which c and u remain arbitrary.

Now that we have appealed to the formally-simple function g up to this point, we would like to employ the more generally-valid (namely, for  $\mathcal{G} = \pm \pi/2$ , as well) form:

$$\int f(x, y, \tan \theta) \, ds \, .$$

If we set:

$$\frac{1}{\cos^2 \vartheta} W = w (\tan \vartheta, y - x \tan \vartheta)$$

and

$$c = \tan \theta_0$$

then we will get:

$$f = \int_{\mathcal{G}}^{\mathcal{G}} \sin(\mathcal{G} - \tau) w(\tan \tau, y - x \tan \tau) d\tau + \frac{\partial u}{\partial x} \cos \mathcal{G} + \frac{\partial u}{\partial y} \sin \mathcal{G}$$

after an easy calculation. The  $\tau$  in that means the integration variable.

4b) The Weierstrass E function must always possess a negative sign.

Let  $\overline{x}$ ,  $\overline{y}$ ,  $\tan \overline{\vartheta} = \overline{p}$  be the values of x, y, p along any curve  $\overline{C}$  that connects the given endpoints 1 and 2. Let  $\vartheta = \arctan p$  be the directions of the extremals (i.e., lines) that osculate  $\overline{C}$ and start from 1. With the use of the function g(x, y, p), E will be represented by (\*):

$$E = \left[g\left(\overline{x}, \overline{y}, p\right) - p\frac{\partial g\left(\overline{x}, \overline{y}, p\right)}{\partial p} - g\left(\overline{x}, \overline{y}, \overline{p}\right) + \overline{p}\frac{\partial g\left(\overline{x}, \overline{y}, p\right)}{\partial p}\right]\frac{\partial \overline{x}}{\partial \overline{s}}.$$

If one performs the differentiations that occur in the representation of g in terms of W and then substitutes w in place of W then the expression for E can be reduced in the following way:

$$E = \int_{g_0}^{g} \sin(\overline{g} - \tau) w(\tan \tau, \overline{y} - \overline{x} \tan \tau) d\tau.$$

One sees immediately from this that the necessary condition for E to always be negative is that w must possess a positive sign in all of its domain of definition, and that condition will also be

<sup>(\*)</sup> Compare Kneser, Lehrbuch der Variationsrechnung, Braunschweig, 1900. pp. 77, formula 58.

sufficient since we can assume the restriction on  $\overline{\mathcal{G}}$  that  $|\overline{\mathcal{G}} - \mathcal{G}| < \pi$ . (More details on that are in § 3.)

It might be remarked that in the present case, it would make no sense to demand only the occurrence of a weak minimum, so to be satisfied with only fulfilling the Legendre condition. That is because, as a more rigorous argument would show, the requirement that a weak extremum should exist for *all* segments along which f(x, y, y') exists at all would be equivalent to the demand of a strong minimum.

As far as the Jacobi condition is concerned, in the case where we fix the limits of the integral to be minimized, it is fulfilled automatically since there will exist no envelope for all of the lines that emanate from a point. It is only in the case where all points and lines of the extended geometry are attainable points and lines that we must add yet another remark, which we would like to defer until § **4**, however:

In the absence of it, we can now state the following result:

All plane geometries that correspond to our axioms, and especially Archimedean requirement, are given by a metric in which the arc-length element dl is expressed by:

$$dl = ds \left\{ \int_{\theta_0}^{\theta} \sin(\theta - \tau) w(\tan \tau, y - x \tan \tau) d\tau + \frac{\partial u(x, y)}{\partial x} \cos \theta + \frac{\partial u(x, y)}{\partial y} \sin \theta \right\},\$$

in which w must be a function that is positive everywhere in its domain of definition.

The previous restriction that one should appeal to only curves with continuous tangents for comparison also drops away. The line is also shorter than any curve for which the length can be defined, say, in the **Hilbertian** sense ( $^*$ ).

### § 3. – The monodromy axioms.

The single demand that has still not been followed up on in our representation of the line element is the part of the first congruence axiom that says that:

$$AB \equiv BA$$
 .

I did not demand that congruence in my dissertation: There, it took the form of the following definition:

I) When  $AB \equiv BA$ , or what amounts to the same thing, when  $f(x, y, \tan \theta)$  is also a single-valued function of the last argument, we would like to say that the *strong monodromy axiom* is fulfilled.

<sup>(\*)</sup> See C. A. Noble, "Eine neue Methode in der Variationsrechnung," Dissertation, Göttingen, 1901. Chap. II.

Since we have constantly demanded the satisfaction of this axiom here, we must now ask what conditions it would impose upon the functions w and u, as well as on the constants c in the previous paragraph.

Now, in my dissertation, I showed that the necessary and sufficient conditions for the fulfillment of the strong monodromy axiom are the following ones:

- 1) w must be a single-valued function of x, y, tan  $\mathcal{G}$ .
- 2) The function *u* cannot be assumed to be arbitrary. Rather, it is determined as follows:

$$\frac{\partial u}{\partial x} = \frac{1}{2} \int_{g_0}^{g_0 + \pi} \sin \tau \cdot w \, d\tau \,,$$
$$\frac{\partial u}{\partial y} = \frac{1}{2} \int_{g_0}^{g_0 + \pi} \cos \tau \cdot w \, d\tau \,.$$

The integrability condition is fulfilled.  $dl [f(x, y, \tan \vartheta), \operatorname{resp.}]$  will then be completely independent of  $\vartheta_0$  and can be put into the simple form:

$$dl = \frac{ds}{2} \int_{\theta-\pi}^{\theta} \sin(\theta-\tau) w(\tan\tau, y-x\tan\tau) d\tau.$$

This is the element of length in the case where the strong monodromy axiom is fulfilled.

However, if one does without the strong monodromy axiom then the representation that was given in § 2 will still not guarantee that the determination of the length of a segment AB is single-valued, If I fix A and let B wander about in some way until it once more returns to its original position then the length of the segment AB after a complete cycle of  $\mathcal{G}$  though  $2\pi$  will no longer need to be the same as it started. In my dissertation, that situation gave rise to the following definition:

II) When dl or  $f(x, y, \tan \vartheta)$  again takes its old value after a complete cycle of  $\vartheta$  through  $2\pi$ , so F is a single-valued function in the full sense, we would like to say that the *weak monodromy axiom* has been fulfilled.

In my dissertation, I showed that the necessary and sufficient conditions for the weak monodromy axiom to be satisfied are the following ones:

- 1) w must be a single-valued function of x, y, sin  $\mathcal{G}$ , cos  $\mathcal{G}$ .
- 2) The relations must exist:

$$\int_{0}^{2\pi} \sin \tau \cdot w \, d\tau = 0 \qquad \text{and} \qquad \int_{0}^{2\pi} \cos \tau \cdot w \, d\tau = 0 \, .$$

The only demand upon the function u is that it must be single-valued with respect to the variables *x*, *y*.

However, in the course of the following considerations, we would always like to assume that the strong monodromy axiom has been fulfilled, with the exception of the example of Minkowskian geometry (see § 5).

# § 4. – The normal line. Elliptic geometries.

Up to now, the normal line has remained excluded from consideration, but we can belatedly establish its behavior with no further calculations when we first prove the following general theorem:

1) If all of our axioms A, B, C, D, E are true for all points and lines with the exception of a single one for which that is still uncertain, and if all of the segments that osculate the distinguished line have a certain finite value then every segment on the line in question will likewise have a certain finite value, viz., the length of a (variable) segment whose endpoints move to the endpoints of the segment in question in some way and converge to a fixed finite limiting value independently of the type of passage to the limit.

2) If the assumptions of Theorem 1) are fulfilled, so in particular the Archimedean requirement is true for all segments that lie in a certain neighborhood of the line in question, and for which at most one point lies in the line, then the segment in question will itself also be the shortest distance between its endpoints.

The proof of the first theorem takes the following form:

Let  $G_1 G_2$  be a segment on the line whose length we still do not know. Let AB be a segment whose ends A and B should go to  $G_1$  and  $G_2$ , resp. Let  $A_1 B_1$  be another segment of that type.

We then connect A,  $G_1$ ,  $A_1$  with each other by segments, and likewise  $B, G_2, B_1$ .

We next prove that we can delimit a finite region around  $G_1$ whose points all have a distance from  $G_1$  that is smaller than the arbitrarily-chosen quantity  $\varepsilon$ , except for the points of the critical line. To that end, we draw two lines through  $G_1$  on which we measure out  $\varepsilon/2$  on both sides. If we connect the four endpoints of those segments to each other then all points inside of the rectangle that arises will have the required property, which would emerge from a

double application of the Archimedean requirement.



As a result, we can choose  $A_1$  and A to be close enough to  $G_1$  and choose B and  $B_1$  to be close enough to  $G_2$  that:

$$AG_1 < \frac{\varepsilon}{2}, \qquad A_1G_1 < \frac{\varepsilon}{2}, \qquad BG_2 < \frac{\varepsilon}{2}, \qquad B_1G_2 < \frac{\varepsilon}{2},$$

in which  $\varepsilon$  means an arbitrarily-given small number. From the Archimedean requirement, we will also have:

$$AA_1 < \varepsilon$$
 and  $BB_1 < \varepsilon$ .

From the same requirement, that will imply that:

$$A B < A_1B_1 + 2 \varepsilon,$$
  

$$A_1B_1 < A B + 2 \varepsilon,$$

so

$$\lim_{\varepsilon=0} AB = \lim_{\varepsilon=0} A_1 B_1,$$

and we will call that limit the *distance*  $G_1$   $G_2$ .

With that, the first theorem is proved. Now, the second theorem says that the segment  $G_1$   $G_2$  is also smaller than any other link between the points  $G_1$  and  $G_2$ .

Namely, if there were a shorter link then one could always construct another one that is no longer than the first one and would consist of two segments  $G_2C$  and  $CG_2$ . However, one can choose a point A on  $CG_1$  and a point B on  $CG_2$  such that:

$$AG_{1} < \frac{\varepsilon}{3},$$
$$BG_{1} < \frac{\varepsilon}{3},$$
$$G_{1}G_{2} - AB \mid < \frac{\varepsilon}{2}.$$

If one continues in that way then:

AB < AC + BC,

and it will follow that:

or

$$G_1 G_2 \leq G_1 C + G_2 C ,$$

 $G_1G_2 < G_1C + G_2C + \varepsilon$ 



Figure 4.

because  $\varepsilon$  can be made arbitrarily small, but nothing enters into that inequality except for quantities that are independent of  $\varepsilon$ .

However, the fact that the inequality is also impossible follows from the fact that one can construct other paths between  $G_1$  and  $G_2$  that are shorter than  $CG_1 + CG_2$ , e.g., e.g.,  $G_1EFG_2$  (Fig. 4), but on the other hand, they can once more be only equal to  $G_1G_2$ , which is a contradiction.

With that, the theorem is proved.

We shall now consider a geometry of completely elliptic type, i.e., a geometry for which we know that every line possesses a finite length except for the normal line. Therefore, there are otherwise no relative singular points and no ideal points at all.

In addition, we might know that the sum of the two sides of any triangle that *is not* intersected by the normal line will be greater than the third.

We now pose the questions:

- 1) When does there also exist a way of determining length along the normal line?
- 2) When will it be so arranged that the Archimedean requirement is also fulfilled by it?

At any rate, on the basis of the theorems that were proved, we can just as well state:

If we consider a well-defined segment *s* along the normal line then the lengths of all segments that lie on a well-defined side of the segment *s* in question will converge to a certain limit, and that limiting value will also be the lower limit of the lengths of all links that lie along the side of the segment *s* considered.

In order to be able to support that theorem on the arguments at the beginning of this section, we must show that even in a triangle that has one vertex on the normal line, each side will still remain smaller than the sum of the other two. However, the fact that this is indeed the case can be proved by considerations that are very similar to the ones that were presented at the beginning of this section.

However, if we would like to show that the limiting values for the length of a segment that is achieved by approaching the normal line from *different* sides coincide then we must know that the Archimedean requirement is also fulfilled for triangles that are intersected by the normal line and that the considerations of § 2 say nothing about such triangles.

[Furthermore, this is the place to make a remark. In elliptic geometry, there are two segments of finite length that connect two points (when we ignore mainly segments that are longer than the circumference of the entire line, but partially overlap with it). Naturally, only one of them will actually be the shortest link between the two points, in general. Nonetheless, the theorem that each side of a triangle will be shorter than the sum of the other two will always remain true when we select triples from all of the segments that connect three points that cut out a piece of the plane. The fact that this theorem is correct as long as the normal line does not cut the triangle follows from the considerations of § 2.]

However, should the sum of two sides of a triangle also be still greater than the third in the event that the triangle is cut by the normal line then, as we will see directly, the necessary and sufficient condition for that will be that all lines must have the same total length.

We would now like to show that the condition is necessary.

We would like to consider the two lines  $g_1$ and  $g_2$  that intersect at A. We can then choose a point B along  $g_2$  such that  $AB < \varepsilon$ , in which  $\varepsilon$  is an arbitrarily small number. We draw a parallel pto  $g_1$  through B.  $g_1$  and p then intersect along the normal line at the point C. Let  $l_1$  be the length of the one segment AC (so  $g_1 - l_1$  is that of the other), so  $l_2 < l_1 + AB$  or  $l_2 < l_1 + \varepsilon$ . Now, should the triangle theorem that we speak of also be true for triangles that are cut by the normal line, then it would also have to be fulfilled for the crosshatched triangle ABC. We must then have:

 $g_1 - AB < g_1 - l_1 + l_2$ 

$$g_2 < g_1 + 2\varepsilon$$
.

We can prove that:

in it, or:

$$g_1 < g_2 + 2\varepsilon'$$

in an entirely-analogous way. Now, since  $\varepsilon$  and  $\varepsilon'$  are small quantities, it will follow that  $g_1 = g_2$ . Q. E. D.



The case in which  $g_1$  and  $g_2$  are parallel can be resolved by a double application of the theorem that was just proved.

Now, we can also show that the condition that all lines possess equal lengths is sufficient for the sum of two sides to also be greater than the third one for triangles that are cut by the normal line (in the event that this theorem is valid for all triangles that are not cut by the normal line, as we have assumed).

We again consider the cross-hatched triangle *ABC*. Now, if the stated theorem were false, so perhaps  $(g_2 - AB) \ge g_1 - l_1 + l_2$ , then since  $g_1 = g_2$ , it would follow that  $AB + l_2 \le l_1$ , which is impossible, AB,  $l_2$ ,  $l_1$  are the sides of a triangle that *is not* cut by the normal line.

Neither could one have  $g_1 - l_1 \ge g_2 - AB + l_2$ , since one would then need to have  $AB \ge l_1 + l_2$ , which would be false on the same grounds. With that, one has succeeded in building the foundation for being able to apply the theorems that were posed in the beginning completely. There will then exist a uniquely-established metric along the normal line when:

1) The geometry has elliptic type everywhere and fulfills our axiom everywhere outside the normal line, and

2) All lines possess the same total length.

What conditions does that imply for the function *w*? Condition 1) demands that:

$$\lim_{s=\infty}\int_{0}^{s}\frac{ds}{2}\int_{g-\pi}^{g}\sin\left(g-\tau\right)w(\tan\tau,s\sin\left(g-\pi\right)+y_{0}\cos\tau-x_{0}\sin\tau\right)d\tau$$

must have a well-defined finite value for all finite  $x_0$ ,  $y_0$  and for all  $\mathcal{G}$ .

Condition 2) demands that:

$$\int_{-\infty}^{\infty} \frac{ds}{2} \int_{\vartheta-\pi}^{\vartheta} \sin(\vartheta-\tau) w(\tan\tau, s\sin(\vartheta-\pi) + y_0\cos\tau - x_0\sin\tau) d\tau$$

must be independent of  $\mathcal{G}$ ,  $x_0$ ,  $y_0$ .

Now, it can also be shown that this does, in fact, always happen when the two integrations over s and  $\tau$  commute.

If we assume that those integrations commute and that the lengths of all lines possess finite values then our geometry of elliptic type will fulfill the Archimedean requirement everywhere with no exceptions, and all lines will have the same total length.

That is understood to mean:

One of the two segments that connect two points will always be the absolute shortest link between the two points when both segments are not actually equal. However, *each* of the two segments will be shorter than any curve segment that can be created by fixing the ends of the segment and continuously deforming it, so by cutting out a piece of the elliptic plane with it. (The two segments into which an elliptic line will be divided by two points *cannot* be translated to each other. That last consideration points to an essential difference between elliptic geometry and spherical geometry, in which none of those theorems would be true in that form.)

### § 5. – Examples. The Minkowskian and Hilbertian geometries.

I) The next-simplest case is the one in which the points of one and only line are unattainable points. We will call that line the *line at infinity*.

We can assume, with no loss of generality, that it coincides with the normal line. *The parallel axiom is not fulfilled in the Euclidian sense*.

There is a certain interest in looking for those geometries in which every family of parallel extremals (lines) belongs to a family of parallel lines as *transversals* in the **Kneser** sense.

A simple calculation will show that the necessary and sufficient condition for a geometry that is characterized in the foregoing way to exist is that *w* must be independent of its second argument

and that  $\frac{\partial u}{\partial x} = \alpha$ ,  $\frac{\partial u}{\partial y} = \beta$  must be constants. (Here, we will allow the strong monodromy axiom to not be fulfilled.)

The element of length will now have the form:

$$dl = \left\{ \int_{\beta_0}^{\vartheta} \sin\left(\vartheta - \tau\right) w(\tau) \, d\tau + \alpha \cos\vartheta + \beta \sin\vartheta \right\} ds$$

then, and the length of the segment  $\overline{12}$  will be equal to:

$$\overline{12} = s_{1,2} \left\{ \int_{g_0}^{g} \sin(\theta - \tau) w(\tau) d\tau + \alpha \cos \theta + \beta \sin \theta \right\}.$$

The metric is proportional to the usual Euclidian metric on each line and is independent of parallel translation then.

However, that means:

The single geometry of our type in which parallel lines will again belong to parallel lines as transversals is the Minkowski geometry (see Introduction) because it is characterized by the aforementioned property.

The *H*-curve is represented in polar coordinates r, g by:

$$1 = r \cdot \left\{ \int_{\vartheta_0}^{\vartheta} \sin\left(\vartheta - \tau\right) w(\tau) d\tau + \alpha \cos\vartheta + \beta \sin\vartheta \right\}$$

It will follow from this that:

$$\frac{d^2}{d\vartheta^2}\left(\frac{1}{r}\right) + \frac{1}{r} = w\left(\vartheta\right).$$

However, since w(9) was subject to the single condition that it should always be positive, and on the other hand  $\frac{d^2}{d9^2} \left(\frac{1}{r}\right) + \frac{1}{r}$  is equal to the radius of curvature of the curve, up to an alwayspositive factor, we will then get the single condition for the H-curve that it should represent an everywhere-convex curve around the origin.

The fact that it actually surrounds the origin expresses the fact that 1 / r will possess a finite value for all  $\mathcal{G}$  in the event that we assume that all points are regular.

Moreover, it is clear that the demand of strong monodromy axiom comes from the fact that the *H*-curve must be symmetric about the origin.

II) Hilbert has *exhibited a geometry in which the lines are shortest in any event* in a letter that he wrote to Klein (see Introduction).

If we consider a convex, closed curve in the *xy*-plane and define the distance between two points in its interior by the logarithm of the double ratio that those two points define with the points of intersection of the line that they generate with the aforementioned curve then we will get the metric in question.

Let H(x, y) = 0 be the equation of the curve to be determined, let y = p x + b be a line. The value of the abscissa of the point of intersection of the line with the convex curve will then be obtained from the equation H(x, p x + b) = 0. That equation will have either two real roots for x, one multiple root, or no real root at all.

Let the real roots (if they exist) be  $v_1(p, b)$  and  $v_2(p, b)$ . One will then get the value:

$$W(p, y-px) = \frac{\partial^2 (v_2 - v_1)}{\partial b^2}$$

for the expression W(p, y - px) with little effort, in which  $v_2$  is taken to be the larger or smaller of the two roots according to whether  $+ \pi/2 < 9 < 3\pi/2$  or  $- \pi/2 < 9 < \pi/2$ , resp.

That will then show that the function:

$$w = \frac{1}{\cos^2 \mathcal{G}} W$$

will possess an always-positive sign, when the curve H(x, y) = 0 is everywhere convex. Moreover, *w* is defined for all points in the interior of the curve H = 0.

III) Now, we can obviously create a much-more-general geometry of our type by setting:

$$W = \sum_{\lambda=1}^{\infty} \frac{\partial^2 (v_{\lambda,2} - v_{\lambda,1})}{\partial b^2},$$

in which the  $v_{\lambda}$  correspond to curves  $H_{\lambda} = 0$ .

However, I have still not been able to ascertain whether one can obtain the general geometry of our type by such a *superposition of Hilbertian geometries*.

### § 6. – On the effect of discontinuities in the function *w*.

Allow me to highlight only one theorem in regard to the investigations that are suggested by the title of this section; otherwise, refer to my dissertation (§ 5 - 8).

The theorem in question reads:

If the function w has singularities of the type such that  $\int dl$  also possesses a well-defined finite value along a line that goes through such places (so the points in question for which the singularities occur are still referred to as attainable points) then the geometry that belongs to w will not indicate any deviation from the Archimedean requirement in the event that only w remains

everywhere-positive outside of the places in question. The straight line segments will then remain the shortest link between their endpoints.

### § 7. – Generalizing the concept of "*H*-curve."

Assume that *w* is regular along a piece of the *y*-axis, which is an assumption that implies no restriction.

If we now define the curve:

$$1 = r \left\{ \int_{\theta_0}^{\theta} \sin(\theta - \tau) w(\tau, b) d\tau + \left(\frac{\partial u}{\partial x}\right)_{x=0} \cos\theta + \left(\frac{\partial u}{\partial y}\right)_{x=0} \sin\theta \right\}$$

in polar coordinates with each point x = 0, y = b as its starting point then it will coincide with Minkowski's *H*-curve in the case where *w* is independent of *b*.

The differential equation:

(1) 
$$\frac{d^2}{dg^2}\left(\frac{1}{r}\right) + \frac{1}{r} = w\left(g, b\right)$$

is true for our curve.

As long as w is arbitrary, the curve will also be arbitrary since  $\left(\frac{\partial u}{\partial x}\right)_{x=0}$  and  $\left(\frac{\partial u}{\partial y}\right)_{x=0}$  are

naturally arbitrary functions of y (b, resp.).

If the weak monodromy axiom is fulfilled then that obviously means that the same *r* will belong to each  $\vartheta + 2k\pi$  (*k* is a whole number) that belongs to  $\vartheta$ , so the curve will be closed. If the strong monodromy axiom is fulfilled then each  $\vartheta + k\pi$  will belong to the same *r* that belongs to  $\vartheta$ , so the curve will be symmetric relative to x = 0, y = b.

However, since:

$$\frac{d^2}{d\theta^2} \left(\frac{1}{r}\right) + \frac{1}{r}$$

is equal to the curvature of the curve, up to a positive factor, as we saw before, w > 0 will say that the curve is always convex.

We therefore have the result:

*One gets the most general plane geometry in which the line is shortest in the following way:* 

We describe a convex curve around each point on a line that we shall make the *y*-axis. We then put the equation of that curve in polar coordinates r,  $\vartheta$  around the point of the line into the form it always has:

$$1 = r g \left( \mathcal{G}, b \right)$$

or

$$1 = r \left\{ \int_{\vartheta_0}^{\vartheta} \sin(\vartheta - \tau) w(\tau, b) d\tau + \left( \frac{\partial u}{\partial x} \right)_{x=0} \cos \vartheta + \left( \frac{\partial u}{\partial y} \right)_{x=0} \sin \vartheta \right\},\$$

in which *b* is the parameter that indicates the point on the line, and *u*(*x*, *b*) can still be chosen in many ways, but only in such a way that  $\left(\frac{\partial u}{\partial x}\right)_{x=0}$  and  $\left(\frac{\partial u}{\partial y}\right)_{x=0}$  are functions of *b* that are given by

g. However, w is determined completely by g, according to equation (1).

Therefore, the element of length of the desired geometry will be equal to:

$$ds\left\{\int_{\mathcal{G}}^{\mathcal{G}}\sin\left(\mathcal{G}-\tau\right)w(\tau,y-x\tan\tau)\,d\tau+\frac{\partial u}{\partial x}\cos\mathcal{G}+\frac{\partial u}{\partial y}\sin\mathcal{G}\right\}.$$

The length is then determined up to merely a function of position (which is independent of the path that connects the endpoints).

If the chosen curve is symmetric then the strong monodromy axiom will be fulfilled by a suitable choice of u that is likewise determined completely then (\*).

If the curve closes after one circuit then the weak monodromy axiom will be fulfilled, no matter how *u* might be chosen (as long as it is consistent with what was said above).

These curves that are so essential in determining the geometry might be called *H*-curves.

#### CHAPTER II.

#### Geometries in space.

#### § 8. – The structure of the postulated geometries.

There are no difficulties that might stand in the way of generalizing the arguments of § 1 to space now. We can then associate each point outside of the normal plane with a triple of numbers x, y, z. The line will then be represented by the equations:

$$y = p x + u ,$$
  
$$z = q x + v ,$$

or by the differential equations:

$$\frac{d^2x}{ds^2} = 0, \qquad \frac{d^2y}{ds^2} = 0, \qquad \frac{d^2z}{ds^2} = 0 \quad (ds = +\sqrt{dx^2 + dy^2 + dz^2}).$$

<sup>(\*)</sup> See § 3.

Now, as far as the investigations in § 2 are concerned, our axiom will now have a result that is entirely analogous to the one there: The length of a rectilinear segment can generally be represented by the integral:

$$\int_{x_1}^{x_2} y(x, y, z, y', z') dx \qquad \left( y' = \frac{dy}{dx}, z' = \frac{dz}{dx} \right),$$
$$\int_{x_1}^{x_2} f\left( x, y, z, \frac{dx}{ds}, \frac{dy}{ds}, \frac{dz}{ds} \right) ds ,$$

and always by:

in which f is homogeneous of degree one in the last three variables.

For other curves, we then define the length of a section by the integral above, which we extend along the curve in question.

4) The Archimedean requirement that the line should be the shortest link between two given points again splits into two others:

4.a) The two partial differential equations must be fulfilled that arise upon comparing the differential equations of the line with the Lagrange equations for the variational problem, namely:

(1) 
$$0 = \frac{\partial^2 g}{\partial p \partial x} + \frac{\partial^2 g}{\partial p \partial y} p + \frac{\partial^2 g}{\partial p \partial z} q - \frac{\partial g}{\partial y},$$
(2) 
$$0 = \frac{\partial^2 g}{\partial q \partial x} + \frac{\partial^2 g}{\partial q \partial y} p + \frac{\partial^2 g}{\partial q \partial z} q - \frac{\partial g}{\partial z}.$$

One initially gets the necessary conditions from them that  $\frac{\partial^2 g}{\partial p^2}$ ,  $\frac{\partial^2 g}{\partial p \partial q}$ ,  $\frac{\partial^2 g}{\partial q^2}$  must each be

functions of only p, q, u = y - px, v = z - qz.

If one correspondingly sets:

$$\frac{\partial^2 g}{\partial p^2} = U(p, q, u, x) ,$$
$$\frac{\partial^2 g}{\partial p \partial q} = V(p, q, u, x) ,$$
$$\frac{\partial^2 g}{\partial q^2} = W(p, q, u, x)$$

then the condition for the compatibility of those three equations will imply the relations:

(I.a) 
$$\frac{\partial U}{\partial q} = \frac{\partial V}{\partial p}; \qquad \frac{\partial V}{\partial q} = \frac{\partial W}{\partial p}; \qquad \frac{\partial U}{\partial v} = \frac{\partial V}{\partial u}; \qquad \frac{\partial V}{\partial v} = \frac{\partial W}{\partial u},$$

from which it will again follow that U, V, W must each satisfy the partial differential equation:

(I.b) 
$$\frac{\partial^2 h}{\partial u \,\partial q} = \frac{\partial^2 h}{\partial v \,\partial p} \qquad (h = U, V, W)$$

If one considers all of that then one will get the following necessary Ansatz for the element of length:

(II) 
$$dl = g \, dx = dx \int_{p_0,q_0}^{p,q} \left\{ dp \int_{p_0,q_0}^{p,q} (U \, dp + V \, dq) + dq \cdot \int_{p_0,q_0}^{p,q} (V \, dp + W \, dq) \right\} + dt \, (x, y, z) \, ,$$

in which t means an arbitrary function of x, y, z, while  $p_0$ ,  $q_0$  are arbitrary constants.

The fact that the curve integral that was written down has a value that depends upon only the limits is linked with the fulfillment of the relations (I).

If we always extend the inner integral over the same curve as the outer one then we will also write:

(II.a) 
$$dl = g \, dx = dx \int_{p_0, q_0}^{p, q} \int_{p_0, q_0}^{p, q} (U \, dp \, dp + 2V dp \, dq + W \, dq \, dq) + dt \, (x, y, z) ,$$

and the fact that the expression for g that is thus-determined will also satisfy the differential equations (1) and (2) when only U, V, W satisfy the conditions (I) is easy to verify.

4.b) In order for a minimum to actually occur, the Weierstrass *E*-function must always be negative.

If  $\overline{p}$ ,  $\overline{q}$  are the values of p, q along the curve  $\overline{C}$  that connects the endpoints of the segment we have in mind, while p, q are the values along the extremal (viz., line) that osculates  $\overline{C}$  and starts from one end of the segment, then the *E*-function will read (\*):

$$E = \left[g(p,q) - g(\overline{p},\overline{q}) + (\overline{p} - p)\frac{\partial g(p,q)}{\partial p} + (\overline{q} - q)\frac{\partial g(p,q)}{\partial q}\right] \cdot \frac{d\overline{x}}{d\overline{s}}$$

With the help of (II), one now easily gets:

<sup>(\*)</sup> Compare Noble, *loc. cit.*, pp. 65. The factor  $d\bar{x} / ds$  must necessarily be added when we do not assume that y and z are single-valued functions of x along  $\bar{C}$ , which would define an essential restriction here.

Hamel – On the geometries in which the lines are shortest.

$$E = -\frac{d\overline{x}}{ds} \int_{p,q}^{\overline{p},\overline{q}} \int_{p,q}^{\overline{p},\overline{q}} (U \, dp \, dp + 2V dp \, dq + W \, dq \, dq) \, .$$

However, in order for that expression to always be negative, it is necessary and sufficient that one must have:

$$\frac{dx}{ds}(U\,dp\,dp+2Vdp\,dq+W\,dq\,dq)>0$$

at each location in the five-dimensional space p, q, x, y, z and for each dp, dq, from which it again follows that:

(III) 
$$U \cdot W - V^2 > 0$$
 and  $\frac{dx}{ds}U > 0$ 

which we can also write:

(III.a) 
$$\frac{\partial^2 g}{\partial p^2} \frac{\partial^2 g}{\partial q^2} - \left(\frac{\partial^2 g}{\partial p \partial q}\right)^2 > 0, \qquad \frac{dx}{ds} \cdot \frac{\partial^2 g}{\partial p^2} > 0$$

We can then express the following result:

If the otherwise-arbitrary functions U, V, W in the four arguments p, q, u = y - xp, v = z - qx fulfill the four conditions (I) and (III) then the expression for the element of length dl that is given by (II) will produce the most-general metric that corresponds to the axioms that were mentioned in the Introduction. (Except for the strong monodromy axiom)

If we make the inessential assumption that U, V, W should be regular at the location p = 0, q = 0 then we can set  $p_0 = 0$ ,  $q_0 = 0$ .

Furthermore, if a quantity  $\varepsilon$  is determined by the equation  $p / q = \cot \varepsilon$  then we can choose  $\varepsilon =$  const. as a special integration path in the event that U, V, W are regular on the entire path.

If we correspondingly set  $p = r \cos \varepsilon$ ,  $q = r \sin \varepsilon$  then the expression for the element of length will read:

$$dl = dx \int_{0}^{r} dp \int_{0}^{r} (U\cos^{2}\varepsilon + 2V\cos\varepsilon\sin\varepsilon + W\sin^{2}\varepsilon) dr + dt (x, y, z)$$

Finally, if we introduce the new variable  $\mathcal{G}$  by way of the transformation:

$$r = \tan \vartheta$$

then we will get:

(II.b) 
$$dl = ds \int_{0}^{9} \sin(\vartheta - \tau) w d\tau + dt (x, y, z) ,$$

in which we have set:

$$w = \frac{1}{\cos^2 \tau} (U \cos^2 \varepsilon + 2V \cos \varepsilon \sin \varepsilon + W \sin^2 \varepsilon).$$

The conditions that w must fulfill can thereupon be easily given by using (I) and (III). In particular, it should be mentioned that the second condition (III) reads simply:

(III.b) 
$$w > 0$$

# § 9. – The monodromy axioms.

I) If the weak monodromy axiom is to be fulfilled, i.e., if the segment AB is to take on a uniquely-defined length at all, then the following conditions must be fulfilled:

1) U, V, W can possess only those singularities that make:

a) The integrals in the independent variables p, q that occur in g remain independent of the path as long as the line at infinity in the pq-plane is not exceeded.

b) The same thing will be true after one introduces the variables  $\mathcal{G}$ ,  $\varepsilon$  in place of p, q, and indeed for all values of  $\mathcal{G}$  and  $\varepsilon$  for which U, V, W are defined at all.

2) w, and therefore U, V, W, as well, must possess periods of  $2\pi$  as functions of both  $\tau$  and  $\varepsilon$ .

3) The equations:

$$\int_{0}^{2\pi} \sin \tau w d\tau = 0 \qquad \text{and} \qquad \int_{0}^{2\pi} \cos \tau w d\tau = 0$$

must be fulfilled. The function *t* remains completely arbitrary, but naturally apart from the fact that it must be single-valued. The conditions 1), 2), 3) are also sufficient for the fulfillment of the weak monodromy axiom.

II) If the *strong monodromy axiom* is to be satisfied, i.e., if the length of the segment *AB* is to always be equal to that of *BA*, then along with the condition I.1), the following conditions must be fulfilled:

2) w must be a single-valued function of tan  $\tau \cos \varepsilon$  and tan  $\tau \sin \varepsilon$ , i.e., of p and q.

3) The function *t* must be assumed to be such that:

$$\frac{dt}{ds} = -\frac{1}{2}\int_{0}^{\pi}\sin\left(\vartheta-\tau\right)wd\tau.$$

The fact that this determination of t is possible can be easily shown with the use of the assumptions I.1).

*The fulfillment of* 1), 2), 3) *is also sufficient for the strong monodromy axiom. The element of length itself will then take the form:* 

$$dl = \frac{ds}{2} \int_{\vartheta-\pi}^{\vartheta} \sin\left(\vartheta-\tau\right) w(\tau) \, d\tau \, .$$

### § 10. – Examples.

I) If we assume that U, V, W are independent of u, v, so the conditions (I.b) will be fulfilled automatically, then  $\frac{\partial t}{\partial x} = \alpha$ ,  $\frac{\partial t}{\partial y} = \beta$ ,  $\frac{\partial t}{\partial z} = \gamma$ , where  $\alpha$ ,  $\beta$ ,  $\gamma$  are constants, and one will have:

$$dl = ds \left\{ \int_{0}^{\theta} \sin((\theta - \tau)w(\varepsilon, \tau)) d\tau + \alpha \cos\theta + \beta \sin\theta \cos\varepsilon + \gamma \sin\theta \sin\varepsilon \right\}$$

That metric is singled out from the other ones by the fact that it will be the same for parallel lines, i.e., it will depend upon only the direction  $\mathcal{G}$ ,  $\varepsilon$  of the line element, but not upon its location (*x*, *y*, *z*). We will obviously get the *Minkowski geometry* in space (see Introduction).

The *H*-surface, which has the distance 1 from the coordinate origin everywhere, is represented in polar coordinates r,  $\vartheta$ ,  $\varepsilon$  as follows:

$$1 = r \left\{ \int_{0}^{\vartheta} \sin \left( \vartheta - \tau \right) w(\tau, \varepsilon) d\tau + \alpha \cos \vartheta + \beta \sin \vartheta \cos \varepsilon + \gamma \sin \vartheta \sin \varepsilon \right\}.$$

The single condition that this surface is subject to is that it must be an everywhere-convex surface that surrounds the origin. Other than that, it can be chosen to be entirely arbitrary.

If it is simply closed, i.e., if *r* is a single-valued function of the direction  $\mathcal{G}$ ,  $\varepsilon (0 \le \mathcal{G} \le \pi, |\varepsilon| \le \pi)$ , then the weak monodromy axiom will be fulfilled. However, if the surface is symmetric about the starting point then the strong monodromy axiom will be fulfilled.

II) One then obtains the *Hilbert geometry* for space in the following way: Take a closed, convex surface H(x, y, z) = 0. For p, q, u, v such that the lines:

$$y = p x + u ,$$
  
$$z = q x + v$$

go through the interior of the aforementioned surface, there will exist two real roots  $\omega_1$  and  $\omega_2$  of the equation:

$$H(\omega, \omega p + u, \omega p + v) = 0$$

If one then sets:

$$U = \frac{\partial^2 (\omega_2 - \omega_1)}{\partial u^2},$$
  

$$V = \frac{\partial^2 (\omega_2 - \omega_1)}{\partial u \, \partial v},$$
  

$$W = \frac{\partial^2 (\omega_2 - \omega_1)}{\partial v^2} \qquad (\omega_1 > \omega_2, \text{ in case} - \frac{\pi}{2} < \vartheta < \frac{\pi}{2}, \text{ otherwise } \omega_1 < \omega_2)$$

then the conditions (I) and (III) will both be fulfilled. If H = 0 is a simply-closed surface then the strong monodromy axiom will be satisfied in the event that one has determined *t* correctly, and indeed one will get:

$$\int_{0}^{s} \frac{ds}{2} \int_{\vartheta-\pi}^{\vartheta} \sin(\vartheta-\tau) w d\tau = \ln \frac{(x_2 - \omega_2)(x_1 - \omega_1)}{(x_2 - \omega_1)(x_1 - \omega_2)}$$

for the length of the segment  $\overline{12}$  in the interior of H = 0. The values of p, q, u, v that belong to the line that is determined by the segment  $\overline{12}$  are then substituted in  $\omega_1$  and  $\omega_2$ .

With that, we have arrived at the form that **Hilbert** gave:

The distance is equal to the logarithm of the double ratio that is defined by the following four points: The endpoints of the segment to be measured and the points of intersection of the line that it determines with the basic **convex** surface.

That geometry exists only in the interior of the aforementioned surface.

### § 11. – On the differential equation.

$$\frac{\partial^2 h}{\partial u \,\partial q} = \frac{\partial^2 h}{\partial v \,\partial p}$$

**§ 8** teaches that the actual determination of all spatial geometries of our type will follow from the integration of that differential equation.

Apart from the known existence theorem for partial differential equations, the following theorem might be briefly mentioned that I have been able to establish in regard to that equation up to now, which will be important to us:

1) Let h be given as a function U(p, q, v) along u = 0 and as a function Q(p, u, v) of the included quantities along q = 0. Let the initial values U, Q be analytic in p, v, but not necessarily in q (u, resp.); however, they should be continuous and twice-differentiable in those variables.

There will then be a function h(p, q, u, v) that is regular in a certain region around u = 0, q = 0 that fulfills the differential equation that we speak of and assumes the prescribed initial values.

[Naturally, one must have U(p, 0, v) = Q(p, 0, v)!] Indeed, when one sets:

$$h_0 = U + Q - (U)_{q=0}$$

and understands  $\int$  to mean an *n*-fold integral, the integral will read as follows:

$$h = \sum_{n=0}^{\infty} \int_{0}^{q} dq \int_{0}^{u} du \frac{\partial^{2n} h_{0}}{\partial p^{n} \partial v^{n}}$$

One can also find the following closed expression for it:

$$h(u, q, v, p) = h_0(u, q, v, p) + \frac{1}{2\pi} \int_0^u d\alpha \int_0^q d\beta \int_0^{2\pi} \frac{\partial^2 h_0}{\partial p \, \partial v} \, d\vartheta,$$

in which  $\frac{\partial^2 h_0}{\partial p \, \partial v}$  is regarded as a function of the following arguments:

$$\alpha, \beta, \nu - \sqrt{(u-\alpha)(q-\beta)} e^{\vartheta_i}, \ p - \sqrt{(u-\alpha)(q-\beta)} e^{-\vartheta_i}.$$

I have still not been able to produce a proof of existence for the case in which  $h_0$  is not analytic in p, v.

2) There exists a function h that satisfies the partial differential equation  $\frac{\partial^2 h}{\partial u \partial q} = \frac{\partial^2 h}{\partial v \partial p}$ 

and is equal to U(p, q, v) and for u = 0 and equal to Q(p, u, v) for q = 0. In a certain region about the location u = 0, q = 0, p, v, that function h is also the only function that satisfies all of those conditions, assuming that h is a continuous function that is differentiable up to second order, but by no means needs to be analytic.

The region that we speak of can be given *a priori*, and independently of the values of the functions U and Q.

### § 12. – On the geometric interpretation of the results. Extending the concept of "*H*-surface."

Now that we have formulated the analytical conditions that the element of length dl must fulfill on the basis of our axiom in §§ 8 – 11, we shall now pursue a geometrically-intuitive explanation for it.

The peculiar form of U, V, W initially refers to the fact that we essentially know g everywhere when it is given on the yz-plane along for each point in all directions.

The existence theorem that was proved in the previous paragraph then says (but generally only under certain temporary restrictions on the analytical character of certain functions) that it is sufficient to know just g for the z-axis in all directions (u = 0). However, one must also know it at all points in the yz-plane, but only for q = 0, i.e., in the directions that are parallel to the xy-plane.

We now imagine a surface that is constructed at each point x, y, z in such a way that:

$$1 = r \cos \vartheta g$$

is its equation in polar coordinates r,  $\vartheta$ ,  $\varepsilon$ , when referred to the point x, y, z.

That says nothing more than the fact that the surface that is constructed in that way will be concave with respect to the fixed point x, y, z, just like in the special example of Minkowskian geometry with the inequality conditions (III), (§ 8).

From its very geometry, one can imagine that the surface is produced in the following way:

We construct the surface that has a well-defined constant distance a < 1 from the chosen point x, y, z in the geometry in question, but then enlarge it in the sense of the naming convention in such a way that we divide each radius vector by a. We then let a converse to zero. There will then be a limit surface that reads:

$$1 = r \cos \vartheta g$$

in polar coordinates. That surface shall generally be called an *H-surface*.

With the help of such H-surfaces, we can now produce the **most-general geometry** that satisfies our axioms as follows:

One constructs a convex surface about each point on a line that we make the z-axis, and convex curve in each plane z = const. through each point of a plane that goes through the z-axis that we make the yz-plane.

Those surfaces and curves might continually intertwine each other, in general. If one writes them in the form  $1 = r \cdot \cos \vartheta \cdot \overline{\overline{g}}$  by means of polar coordinates r,  $\vartheta$ ,  $\varepsilon$  relative to the associated point then the values of  $\frac{\partial^2 \overline{\overline{g}}}{\partial p^2}$ ,  $\frac{\partial^2 \overline{\overline{g}}}{\partial p \partial q}$ ,  $\frac{\partial^2 \overline{\overline{g}}}{\partial q^2}$  (\*) ( $p = \tan \vartheta \cos \varepsilon$ ,  $q = \tan \vartheta \sin \varepsilon$ ), to the extend that

they can be defined, as well as the differential equation:

<sup>(\*)</sup> We assume that these differential quotients exist, from our previous constructions.

Hamel – On the geometries in which the lines are shortest.

$$\frac{\partial^2 h}{\partial u \,\partial q} = \frac{\partial^2 h}{\partial v \,\partial p} \qquad (u = y, v = z)$$

and the relations I.a in § 8 will determine the three functions U, V, W completely. If we then set:

$$w = \frac{1}{\cos^2 g} (\cos^2 \varepsilon U + 2\cos \varepsilon \sin \varepsilon V + \sin^2 \varepsilon W)$$

and define the function:

$$\overline{g} = \frac{1}{\cos \vartheta} \int_{0}^{\vartheta} \sin(\vartheta - \tau) w(p, q, u, v) d\tau + \alpha(y, z) + \beta(y, z) p + \gamma(y, z) q,$$

then  $\alpha$ ,  $\beta$ ,  $\gamma$  will be determined completely when we demand that  $\overline{g}$  should coincide with  $\overline{\overline{g}}$  in the domain of definition of the latter and that one should simultaneously have  $\frac{\partial \gamma}{\partial y} = \frac{\partial \beta}{\partial z}$ .

Finally, we derive the function g from  $\overline{g}$  by setting:

$$g = \frac{1}{\cos \theta} \int_{0}^{\theta} \sin(\theta - \tau) w(p, q, y - x p, z - x q) d\tau + \frac{\partial t}{\partial x} + p \frac{\partial t}{\partial y} + q \frac{\partial t}{\partial z},$$

in which t can be chosen in a number of ways such that one will have:

$$\left(\frac{\partial t}{\partial x}\right)_{x=0} = \alpha(y, z), \quad \left(\frac{\partial t}{\partial y}\right)_{x=0} = \beta(y, z), \quad \left(\frac{\partial t}{\partial z}\right)_{x=0} = \gamma(y, z).$$

The equation:

$$dl \equiv g \, dz$$

will then represent the element of length for the geometry that we seek.

It will exist as long as the function g exists, and it will satisfy the Archimedean requirement as long as w and  $\frac{\partial^2 g}{\partial p^2} \frac{\partial^2 g}{\partial q^2} - \left(\frac{\partial^2 g}{\partial p \partial q}\right)^2$  remain positive. That last condition is certainly fulfilled in a

certain neighborhood of the initial value (u = 0, q = 0) on the grounds of continuity.

The only thing that still remains undetermined in regard to the choice of the given *H*-function is, in part, the function *t*. However, that indeterminacy will also go away *when we demand the fulfillment of the strong monodromy axiom. In that case, the geometry will be established completely by the H-surface that is assumed.* 

A necessary condition for the occurrence of the strong monodromy axiom is, in any event, that the original *H*-surfaces and *H*-curves should be symmetric. Whether that condition is also sufficient can first be established by a detailed investigation of the differential equation that was discussed in the previous section.