

Mathematical Logic, Axiomatics

To the Greeks we owe the insight that the structure of space, which manifests itself in the relations between spatial configurations and their mutual lawful dependences, is something entirely rational. Whereas in examining a real object we have to rely continually on our sense perception in order to bring to light ever new features, capable of description in concepts of vague extent only, the structure of space can be exhaustively characterized with the help of a few exact concepts and in a few statements, the axioms, in such a manner that all geometrical concepts can be defined in terms of those basic concepts and every true geometrical statement follows as a logical consequence from the axioms. Thereby geometry has become the prototype of a *deductive science*. And in view of this its character, mathematics is eminently interested in the methods by which concepts are *defined* in terms of others and statements are *inferred* from others. (Aristotelian logic, too, was essentially a product of abstraction from mathematics.) What is more, it does not seem possible to lay the foundations of mathematics itself without first giving a complete account of these methods.

1. RELATIONS AND THEIR COMBINATION, STRUCTURE OF PROPOSITIONS

In Euclidean geometry we are concerned with three categories of objects, *points*, *lines*, and *planes*, which are not defined but assumed to be intuitively given, and with the basic relations of *incidence* (a point lies on a line, a line lies in a plane, a point lies in a plane), *betweenness* (a point z lies between the points x and y), and *congruence* (congruence of line segments and of angles). Analogously, in the domain of natural numbers 1, 2, 3, . . . we have a single basic relation in terms of which all others are definable, namely that between a number n and the number n' immediately following upon n . Again, the kinship relations among people furnish an excellent illustration of the general theory of relations. In this case there are two basic categories, *males* and *females*, and two basic relations, *child* (x is child of y) and *spouse* (x is married to y).

The propositional scheme of a *relation*, e.g. ' x follows upon y ,' contains one or more blanks x , y , . . . , each of which refers to a certain category of objects. From the propositional scheme a definite

proposition is obtained, e.g. '5 follows upon 4,' when each blank is filled by (the name of) a certain object of the corresponding category. Language does not reflect the structure of such a relational proposition correctly; we have no subject, copula, and predicate, but a relation with two blanks, neither subordinate to the other, which are filled by objects. One might, in order to get rid of the grammatical accidents of language, represent the propositional schemata of relations by wooden boards provided with so many holes corresponding to the blanks, and the objects by little pegs which fit into the holes. In principle these would be symbols as suitable as words. Two propositions such as '5 follows upon 4' and '4 precedes 5' are expressions of one and the same relation between 4 and 5. It is unwarranted to speak here of two relations inverse to each other. The blanks in a relational proposition, though, do each have a specific position; and it is a particular property (commutativity) if the relation $R(xy)$ (e.g. x is a cousin of y) is equivalent (or coextensive) with $R(yx)$.

Properties will have to be counted among the relations, just as 1 is taken to be a natural number. Their propositional scheme possesses exactly one blank.

[In §47 of his fifth letter to Clarke, Leibniz speaks of a "relation between L and M , without consideration as to which member is preceding or succeeding, which is the subject or object." "One cannot say that both together, L and M , form the subject for such an *accidens*; for we would then have one *accidens* in two subjects, namely one which would stand, so to speak, with one foot in one subject and with the other in the other subject, and this is incompatible with the concept of an *accidens*. It must be said, therefore, that the relation . . . is something outside of the subjects; but since it is neither substance nor *accidens* it must be something purely ideal, which is nevertheless well worthy of examination." The (explicit or implicit) assumption that every relation must be based on properties has given rise to much confusion in philosophy. A statement asserting, say, that one rose is differently colored from a second is indeed founded on the fact that one is red, the other yellow. But the relation 'the point A lies on the left of B ' is not based on a qualitatively describable position of A alone and of B alone. The same holds for kinship relations among people. The view here opposed evidently originates within the domain of sense data, which — it is true — can yield but quality and not relation. It is for this reason that Leibniz, in the above quotation, refers to the relation as something purely ideal. More than two-place relations are hardly ever mentioned in the logico-philosophical literature.

The introduction of propositional schemata with blanks represents

an important progress of mathematical beyond traditional logic. In analogy to mathematical functions, which yield a number when their arguments, or blanks, are filled by numbers, propositional schemata are often also referred to as "propositional functions."

Aside from relations, *operations* play a part in the axioms of arithmetic; e.g. the operation of addition which, when applied to two numbers, a and b , produces a third, $a + b$. This operation can be replaced, however, by the relation $a + b = c$ between the three numbers a , b , c ; it is 'single-valued' with respect to the argument c , in the sense that for any two numbers a and b there exists one and only one number c which stands in the relation $a + b = c$ to them. Thus we are able to subordinate genetic construction to the static existence of relations. Later, however, we shall proceed conversely, inasmuch as we shall replace all relations by constructive processes. }

The principles of the *combination of relations* are as follows:

1. In a relation scheme with several blanks it is possible to *identify* several of these blanks. For instance, from the scheme

$N(xy): x$ is a nephew of y

we may obtain

$N(xx): x$ is a nephew of himself.

2. *Negation*. Symbol: \sim . $N(xy)$ becomes

$\sim N(xy): x$ is not a nephew of y .

3. *and*. Symbol: $\&$. Thus $N(xy)$ and, say, $F(xy) — x$ is father of $y —$ yield the relation with three blanks

$F(xy) \& N(yz): x$ is father of y and y is nephew of z .

It must be stated which blanks of the combined schemata are to be identified. Symbolically this is indicated by choosing the same letter for the blanks.

4. *or*. Symbol: \vee . For instance,

$F(xy) \vee N(yx): x$ is father of y or y is nephew of x .

The combination by means of 'or' can also be expressed in terms of negation and the 'and' combination, and vice versa.¹

¹ Leibniz employs the signs \cdot and $+$ for 'and' and 'or' respectively. We deviate from his notation in order to avoid confusion with the arithmetical operations of multiplication and addition. The formal analogy becomes apparent in J. H. Lambert's distributive law:

$$a \cdot (b + c) = (a \cdot b) + (a \cdot c)$$

(*Acta erudit.* 1765, p. 441). Our use of the product sign Π and summation sign Σ in 6 and 7 are in agreement with Leibniz's usage.

5. Filling a blank by an immediately given object of the corresponding category (substitution). $F(I, x)$ means: I am father of x . This is the scheme of that property with one blank x which appertains exclusively to my children.
6. *all*. Symbol: Π_x . For instance, $\Pi_x R(xy)$ means: all x (of the corresponding category) are in the relation $R(xy)$ to y .
7. *some*. Symbol: Σ_x . Thus $\Sigma_y R(xy)$ means: there exists a y to which x is in the relation $R(xy)$. Σ_x and Π_x are reducible to each other in the same way with the help of negation as \vee and $\&$. The presence of a prefixed symbol Π_x or Σ_x (with index x) deprives the blank x of its capability of substitution just as much as if it had been filled in according to 5. For the sake of these last two principles of construction, it will always be necessary to add the two-place relation of logical identity, $x = y$, to the immediately given relations of our domain of investigation.

[*Examples*. 1. Let (xl) mean: the point x lies on the line l . In plane geometry, according to Euclid, parallelism of two lines, $l \parallel l'$, consists in their having no point (x) in common:

$$\sim \Sigma_x \{ (xl) \& (x'l') \}$$

is therefore the definition of the relation $l \parallel l'$.

2. The statement that through two distinct points (x, y) there always exists a line (l) would have to be written thus:

$$\Pi_y \Pi_x ((x = y) \vee \Sigma_l \{ (xl) \& (yl) \}).$$

3. In the domain of natural numbers, p is called a prime number if no numbers x and y , both different from 1, exist which stand to p in the relation $x \cdot y = p$. This property of p , of being a prime number, is to be defined as follows:

$$\Pi_y \Pi_x ((x = 1) \vee (y = 1) \vee \sim (x \cdot y = p)). \rfloor$$

Starting with the immediately given basic relations of a field of objects we may by applying the above principles in arbitrary combination obtain an unlimited array of 'derived' relations (among which the basic relations will of course be counted too). In particular we shall thus arrive at relations with only one blank, the 'derived properties.' How such a property $E(x)$ may serve as 'differentia specifica' in the sense of Aristotelian logic to demarcate a new concept within the 'genus proximum' of the object category to which its blank x refers, will be sufficiently clear from the definition of 'prime number' in Example 3. Among the derived propositional schemata we find,

furthermore, those which no longer possess any blank at all, such as in Example 2; they are the *pertinent propositions* of our discipline. If we knew of each of these propositions whether or not it is true, then we should have complete knowledge of the objects of the basic categories as far as they are connected by the basic relations. The logical structure of a proposition of this kind can be described adequately only by stating the manner, order, and combination, in which our seven principles have contributed to its construction. This is a far cry from the old doctrine, according to which a proposition must always consist of subject, predicate, and copula. The syntax of relations, as indicated here, offers a firm starting point for a logical critique of language.

[Compare, for instance, Russell's remarks (*Introduction to Mathematical Philosophy*, Chap. 16) on the definite article in non-deictic application (such as in the proposition: *the* line through the distinct points *A*, *B* also passes through *C*).

A proposition is called *general* if it is constructed without recourse to the fifth principle, of substitution of an immediately given object ('this here'). A non-general proposition is called *particular*. (Here one might still distinguish between the singular case, in which Principle 5 only, and neither Π_x nor Σ_x , is used for elimination of a blank x , and the mixed general-singular case.) An object a shows itself to be an *individual* if it can be completely characterized by a pertinent general property; that is, if without recourse to Principle 5 a property can be constructed that applies to a but to no other object of the same category. Existence can be asserted only of something described by a property in this manner, not of something merely named, it being essential that Σ_x carries a blank x as an index. (This remark is of use in a critique of the ontological proof of the existence of God.) Within the domain of natural numbers, 1 is an individual, for it is the only such number which does not follow upon any other. Indeed, all natural numbers are individuals. The mystery that clings to numbers, the magic of numbers, may spring from this very fact, that the intellect, in the form of the number series, creates an infinite manifold of well distinguishable individuals. Even we enlightened scientists can still feel it e.g. in the impenetrable law of the distribution of prime numbers. On the other hand, it is the free constructibility and the individual character of the numbers that qualify them for the exact theoretical representation of reality. The very opposite holds for the points in space. Any property derived from the basic geometric relations without reference to individual points, lines, or planes that applies to any one point applies to every point. This conceptual

homogeneity reflects the intuitive homogeneity of space. Leibniz has this in mind when he gives the following 'philosophical' definition of similar configurations in geometry, "Things are similar if they are indistinguishable when each is observed by itself." (*Math. Schriften*, V, p. 180.) }

2. THE CONSTRUCTIVE MATHEMATICAL

DEFINITION

Aside from the combinatorial definition of derived relations, as discussed in Section 1, mathematics has a *creative definition* at its disposal, through which new ideal objects can be generated. Thus, in plane geometry, the concept of a circle is introduced with the help of the ternary point relation of congruence, $OA = OB$, which appears in the axioms, as follows, "A point O and a different point A determine a circle, the 'circle about O through A '; that a point P lies on this circle means that $OA = OP$." For the mathematician it is irrelevant what circles are. It is of importance only to know in what manner a circle may be given (namely by O and A) and what is meant by saying that a point P lies on the circle thus given. Only in statements of this latter form or in statements explicitly defined on their basis does the concept of a circle appear. Therefore the circle about O through A is identical with the circle about O' through A' if and only if all points lying on the first circle also lie on the second, and vice versa. The axioms of geometry show that this criterion, which refers to the infinite manifold of all points, may be replaced by a finite one: O' must coincide with O , and we must have $OA' = OA$.

[*Further examples.* 1. Nobody can explain what a *function* is, but this is what really matters in mathematics: "A function f is given whenever with every real number a there is associated a number b (as for example, by the formula $b = 2a + 1$). b is then said to be the *value* of the function f for the argument value a ." Consequently, two functions, though defined differently, are considered the same if, for every possible argument value a , the two corresponding function values coincide.

2. In Euclidean geometry the "*points at infinity*," in which parallel lines allegedly intersect, are such ideal elements added to the real points by a creative mathematical definition. By a suitable introduction of ideal points one can, more generally, extend a given limited portion S of space, the 'accessible' space, so as to comprise the whole space of projective geometry. The task is to decide through geometric constructions within S whether two real lines, i.e. lines passing through S , issue from the same ideal point. Such a point is defined most

simply as the vertex of a trilateral corner (formed by three real lines). Thus we arrive at the following definition: "Three non-coplanar lines a, b, c , any two of which are coplanar, determine an ideal point $[a, b, c]$. To say that a line l passes through this point means that l is coplanar with each of the lines a, b, c ." Again this definition implies a criterion for the coincidence of two such ideal points. To every real point p there corresponds exactly one ideal point π such that every line through p passes, in the sense of our definition, through π . Thus a part of the ideal points may be identified with the real points. (Compare Pasch, *Vorlesungen über neuere Geometrie*, 2d ed., p. 40.) According to the same scheme, mathematics always accomplishes the extension of a given domain of operation through the introduction of ideal elements. Such an extension is made in order to enforce the validity of simple laws. For example, as a consequence of the addition of the points at infinity it is true not only that two distinct points can always be joined by a line, but also that two distinct coplanar lines always intersect at a point. The introduction of *imaginary* elements in geometry (in order to enforce simple and universally valid theorems on intersection of algebraic curves and surfaces) and the introduction by Kummer of *ideal numbers* in number theory (in order to restore the laws of divisibility, which at first were lost in the transition from rational to algebraic numbers) are among the most fruitful examples of this *method of ideal elements*.]

A special case is the process of *definition by abstraction*. A binary relation $a \approx b$ in a domain of objects is called an *equivalence* (a relation of the character of equality), if the following is universally true:

- (i) $a \approx a$;
- (ii) if $a \approx b$, then $b \approx a$ (commutativity);
- (iii) if $a \approx b$ and $b \approx c$, then $a \approx c$ (transitivity).

By agreeing to consider two objects a and b as distinct if and only if they do not satisfy the equivalence relation $a \approx b$, a new object domain is derived *by abstraction* from the original one.

Examples and comments. 1. Similarity of geometrical figures is an equivalence. Every figure is attributed a certain *shape*, and two figures have the same shape if and only if they are similar. In a more philosophical mode of expression one is used to say that the concept of shape results from that of figure by abstracting from position and magnitude. In scientific practice the introduction of a concept thus abstracted expresses the intention of exclusively considering *invariant* properties and relations among the originally given objects. $R(xy)$ is invariant with respect to the equivalence \approx , if $R(ab)$ always entails $R(a'b')$, provided $a' \approx a$ and $b' \approx b$.

2. Two sets A and B of objects (say, the persons and chairs in a room) are said to be numerically equivalent, $A \approx B$, if it is possible to pair off the elements of A with those of B (if it is possible to assign one person to each chair, so that no chair remains vacant and no person remains unseated). Numerical equivalence obviously is an equivalence in the above sense. 'Every set determines a (*cardinal*) number; two sets determine the same number if and only if they are numerically equivalent.' (This explanation can already be found in Hume, *Treatise of Human Nature*, Book I, Part III, Section 1.)² In more careless formulation one would say that the concept of (*cardinal*) number results from that of set by abstracting from the nature of the elements of the set and merely considering their discernibility. The objection occasionally put forward that all elements, if degraded into mere Ones, collapse into one, is met by the above precise formulation.

{The example of number may serve to illustrate that the definition by abstraction is a special case of the creative definition. It is subordinated to the latter as follows: 'Every set A determines a number (A). To say that an arbitrary set M consists of (A) elements means that M and A are numerically equivalent.' Consequently the number (A) is the same as the number (B), if every set M that is $\approx A$ also is $\approx B$, and vice versa. But according to the rules (ii) and (iii) for equivalences, such is the case if and only if $A \approx B$. Finally, rule (i) guarantees that, in particular, A itself consists of (A) elements.

3. Two integers, according to Gauss, are congruent modulo 5 if their difference is divisible by 5. Congruence is a relation of the character of equality. Through the corresponding abstraction we obtain, from the integers, the congruence-integers modulo 5. Since the operations of addition and multiplication are invariant with respect to congruence, the result is a finite domain of only 5 elements, within which the usual algebra can be carried on just as well as in the infinite domain of the ordinary rational integers. We have here, for instance, $2 + 4 = 1$, $3 \cdot 4 = 2$ (modulo 5). Not only subtraction but even division can be carried out, by virtue of the fact that 5 is a prime number. This example is of fundamental importance for number theory.

4. The most significant physical concepts are likewise obtained in accordance with the scheme of mathematical abstraction. We shall return to this in Part II when the process of measurement is discussed.

² The passage is worth quoting. "When two numbers are so combined, as that the one has always an unit answering to every unit of the other, we pronounce them equal; and it is for want of such a standard of equality in extension, that geometry can scarce be esteemed a perfect and infallible science."

The principle of definition by abstraction I find alluded to by Leibniz in his fifth letter to Clarke, §47. He says there, "Incidentally, I have proceeded here by and large as Euclid did. The latter, since he found himself unable to define the concept of geometrical ratio absolutely, stipulated what was to be understood by equal ratios." And shortly before that, "The mind, however, is dissatisfied with this equality. It seeks an identity, a thing which would truly be the same, and it imagines it to be in a manner outside of the subjects." The principle has shown its full importance for mathematics only in the 19th century. It was consciously formulated in all generality by Pasch in his book quoted on p. 9 (1882), still more clearly by Frege (*Die Grundlagen der Arithmetik*, Breslau 1884, Sections 63–68). Compare also Helmholtz (*Zählen und Messen*, 1887, *Wissenschaftliche Abhandlungen*, III, p. 377).

Beside the above-mentioned mathematical form of abstraction one might be inclined to place another, the originary abstraction. In looking at a flower I can mentally isolate the abstract feature of color as such. This act of abstraction would here be primary while the statement that two flowers have the same color 'red' would be based on it; whereas in mathematical abstraction it is the equality which is primary, while the feature with regard to which there is equality comes second and is derived from the equality relation. But the integers of the same congruence class modulo 5 can also be characterized by the fact that upon division by 5 they all leave the same remainder; the similarity of two triangles by the fact that the angles in both have the same numerical values and corresponding sides have the same ratios. The general procedure of constructing these remainders and these numerical values of angles and ratios, respectively, takes the place of the feature 'color,' its identical result for two integers or triangles that of the identical 'red' of two flowers. Originary abstraction thus is subordinate to mathematical abstraction. But that which is common to all *congruent* triangles or to all bodies occupying the same spatial position, I find myself unable to represent by an objective feature (it is the latter example that Leibniz *loc. cit.* had in mind), but merely by the indication: congruent to *this* triangle, occupying *this* spatial position. Our question is connected with the problem of relativity (Section 13), with the difference between conceptual definition and intuitive exhibition. But in both cases alike, those of originary and of mathematical abstraction, the transformation of a common feature into an ideal object, e.g. of the property 'red' into an objectified 'red color,' of which the red things 'partake,' is an essential step (Plato's $\mu\epsilon\theta\epsilon\epsilon\iota\varsigma$).]

With every property $E(x)$ which is meaningful for the objects x of a given category we correlate a set, namely 'the set of objects x having the property E .' Thus we speak of the set of all even numbers, or of the set of all points on a given line. The conception that such a set be obtained by assembling its individual elements should by all means be rejected. To say that we know a set means only that we are given a property characteristic of its elements. Only in the case of a finite set do we have, in addition to such *general* description, the possibility of an *individual* description which would exhibit each one of its elements. [Formally, by the way, the latter mode of description is a special case of the former; e.g. the set consisting of three given objects a, b, c corresponds to the property of being either a or b or c : $(x = a) \vee (x = b) \vee (x = c)$.] It is possible that the same set is correlated with two properties E and E' . This happens when every object (of our category) having the property E also has the property E' , and conversely. Hence, what is decisive for the identity of the two sets is not the manner of their definition (in terms of the principles enumerated in Section 1), but solely the question whether each element of one is an element of the other and vice versa, a question referring to a domain of *existing* objects and unanswerable by recourse to the meaning alone. If the concept of set is understood in this way, then the creative definition is seen to be nothing but the transition from a property to a set, so that the mathematical construction of new classes of ideal objects can quite generally be characterized as *set formation*. Now there is no longer anything objectionable in describing the circle about O through A as the set of all points P whose distance from O equals OA , or the color of an object as the set of objects having the same color, or the cardinal number 5 as the set of all those aggregates which are numerically equivalent to the exhibited aggregate of the fingers of my right hand. But it is an illusion — in which Dedekind, Frege, and Russell indulged for a time, because they apparently conceived of a 'set' after all as a collective — to think that thereby a concrete representation of the ideal objects has been achieved. On the contrary, it is through the principle of creative definition that the meaning of the general set concept is elucidated as well as safeguarded against false interpretations.

The properties employed in the creation of new abstracts Φ generally depend on one or more arguments u, v, \dots , which are allowed to vary freely within certain domains: Φ is a *function* of u, v, \dots . In the definition of a circle, for instance, the ternary point relation $OP = OA$ is interpreted as a property of P (relation with one blank P) depending on O and A ; the 'circle about O through A ' is a function of O and A . The criterion for the coincidence of two values of an

abstract, $\Phi(u, v, \dots)$ and $\Phi(u', v', \dots)$, refers to a totality of existing objects. But of special importance are those cases where this transfinite criterion can be converted by virtue of certain universally valid facts into a finite criterion requiring recourse to the meaning of the defining relation only. Instances are our definition of the circle and the definitions by abstraction. Not only properties, but more generally *relations* may serve to define ideal elements. If we want to adhere to the set-theoretical terminology throughout, it will be necessary to have a 'binary set' (R) correspond to every binary relation R ; such that (R) and (R') are identical if, for arbitrary elements a, b , it never happens that one of the statements $R(a, b)$, $R'(a, b)$ is true, the other false. The same for ternary, quaternary, \dots relations. We thus arrive at the following final version of the *principle of creative definition*: A relation $R(xy \dots / uv \dots)$, whose blanks are separated into two groups, $xy \dots$ and $uv \dots$, determines an abstract $\Phi(uv \dots)$ depending on the arguments u, v, \dots ; equality $\Phi(uv \dots) = \Phi(u'v' \dots)$ for any two sets of values of the arguments, u, v, \dots and u', v', \dots , holds if and only if any objects x, y, \dots of the appropriate categories which stand in the relation R to u, v, \dots also stand in the relation R to u', v', \dots , and conversely.

3. LOGICAL INFERENCE

Having dealt with definitions we now come to *proofs*. If one turns a geometrical proposition into a hypothetical statement whose premiss consists of all geometrical axioms, replacing mentally at the same time any abbreviatory expressions by what they mean according to their definitions, one will arrive at a '*formally valid*,' '*analytic*' *proposition*, the truth of which is in no way tied to the meanings of the concepts entering into it (point, line, plane, incidence, betweenness, congruence). The logic of inference has the task of characterizing those propositional structures which assure the formal validity of the proposition. *Barbara*, *Baralipton*, and so on, are of little help in this connection. Leibniz considered the doctrine of the *argumens en forme* "une espèce de Mathématique universelle, dont l'importance n'est pas assez connue" (*Nouveaux Essais*, Livre IV, Chap. 17, §4).

That part of logic which operates exclusively with the logical connectives 'not,' 'and,' 'or' will be referred to as *finite logic*, as opposed to *transfinite logic*, which in addition uses the propositional operators 'some' (or 'there is') and 'all'. The reason for this subdivision is as follows. Suppose several pieces of chalk are lying in front of me; then the statement 'all these pieces of chalk are white' is merely an abbreviation of the statement 'this piece is white & that piece is white & \dots '

(where each piece is being pointed at in turn). Similarly 'there is a red one among them' is an abbreviation of 'this is red \vee that is red $\vee \dots$ '. But only for a finite set, whose elements can be exhibited individually, is such an interpretation feasible. In the case of infinite sets, the meaning of 'all' and 'some' involves a profound problem which touches upon the core of mathematics, the very secret of the infinite; it will unfold itself to us in the next chapter. The situation here may be compared to the transition from finite to infinite sums; the meaning of the latter is tied to special conditions of convergence, and one may not deal with them in every respect as with finite sums.

In the propositional calculus it is convenient to introduce, in addition to the symbols for 'not,' 'and,' 'or,' the symbol $a \rightarrow b$ (read: a implies b). It has the same meaning as $\sim a \vee b$ (a does not hold or b holds) and does not beyond that signify any deeper connection between the propositions a and b .

[Incidentally two of the four symbols \sim , $\&$, \vee , \rightarrow would suffice; for the propositional calculus it is convenient to choose \rightarrow and \sim . Nay, it is even possible to get along with one symbol, a/b , denoting the incompatibility of the propositions a and b ($\sim a \vee \sim b$). For, in place of

$$\sim a, \quad a \rightarrow b, \quad a \& b, \quad a \vee b$$

we may write

$$a/a, \quad a/(b/b), \quad (a/b)/(a/b), \quad (a/a)/(b/b).$$

However, for the sake of greater lucidity, we shall here use all four symbols.]

In a finite-logical formula the letters (propositional variables) for which arbitrary propositions (without blanks) may be substituted are combined by those four symbols \sim , $\&$, \vee , \rightarrow . For example:

$$b \rightarrow (a \rightarrow b).$$

There exists a general rule by which the formal validity of a formula of this kind can be recognized. In fact, assign to each letter occurring in the formula one of the values 'true' (T) or 'false' (F) in all possible combinations, and determine in each instance the value of the entire formula according to the following direction for evaluating compound propositions:

a	$\sim a$	a	b	$a \rightarrow b$	$a \& b$	$a \vee b$
T	F	T	T	T	T	T
F	T	T	F	F	F	T
		F	T	T	F	T
		F	F	T	F	F

(The number of combinations to be tested, for instance, when the formula contains 5 different propositional variables is 2^5 .) If the resulting value of the formula is T in every case, then the formula is formally valid. This rule, which may be said to be based on the *law of contradiction* and the *law of the excluded middle* (*tertium non datur*), I shall call briefly the *finite rule*.

Example: $b \rightarrow (a \rightarrow b)$.

a	b	$a \rightarrow b$	$b \rightarrow (a \rightarrow b)$
T	T	T	T
T	F	F	T
F	T	T	T
F	F	T	T

On this level, then, it is possible to ascertain directly by a combinatorial procedure following a fixed scheme, whether a given assertion is a logical consequence of certain other propositions, provided premiss and conclusion are built up of propositions a, b, \dots (whose meanings do not matter) with the help of the four operations $\sim, \rightarrow, \&, \vee$.

All this changes completely as soon as 'some' and 'all' (and their concomitants, the blanks) are admitted into our formulas. Σ_x and Π_x compel us to *construct* — we set up a number of formally valid basic formulas, the *logical axioms*, and state a rule by which further formally valid propositions may be obtained from formally valid propositions. The rule is none other than the one by which logic is applied to all theoretical sciences, namely the *sylogism*: if you have a proposition A , and a proposition $A \rightarrow B$ in which the first proposition A reappears on the left of \rightarrow , then set down the proposition B . All propositional structures obtained from the axioms by repeated application of this rule are of analytic character. It is impossible, however, to characterize descriptively the infinite manifold of these individual structures independently of the constructive manner in which they are generated. Hence the necessity of step by step demonstration. Using a phrase coined by J. Fries in a somewhat different sense, one

may speak of a 'primordial obscurity of reason.' We do not possess the truth, it wants to be attained by action.

[Galileo ("Dialogo," *Opere*, VII, p. 129) expresses a widely spread view when he interprets this as the difference between human and divine understanding. "We proceed in step-by-step discussion from inference to inference, whereas He conceives through mere intuition. Thus in order to gain insight into some of the properties of the circle, of which it possesses infinitely many, we begin with one of the simplest; we take it for a definition and proceed from it by means of inferences to a second property, from this to a third, hence a fourth, and so on. The divine intellect, on the other hand, grasps the essence of a circle *senza temporaneo discorso* and thus apprehends the infinite array of its properties." (But intensively, i.e. as regards the objective certainty of an individual mathematical truth, the human intellect does not fall short of the divine intellect).]

Concerning 'some' and 'all,' Σ_x and Π_x , we may lay down, first of all, the following two axioms, in which $a(x)$ stands for an arbitrary propositional scheme containing the one blank x , and c for any given object of the corresponding category:

$$\text{I. } \Pi_x a(x) \rightarrow a(c); \quad \text{II. } a(c) \rightarrow \Sigma_x a(x).$$

The first of these axioms tells us merely how to derive something from a universal proposition, but fails to show how other propositions can ever lead us to a universal proposition. The converse is true of the second axiom.

[Everybody knows the classical example of an inference: (α) all men are mortal, (β) Caius is a man; hence (γ) Caius is mortal. Our formalism decomposes it into several steps. Let H and M designate the properties of being (hu)man and mortal respectively, and let c designate Caius. Then

$$(\alpha) \quad \Pi_x (H(x) \rightarrow M(x)),$$

in connection with I,

$$\Pi_x (H(x) \rightarrow M(x)) \rightarrow (H(c) \rightarrow M(c)),$$

according to our syllogism rule of inference yields

$$H(c) \rightarrow M(c).$$

This, together with

$$(\beta) \quad H(c),$$

again by the rule of inference, yields

$$(\gamma) \qquad M(c).$$

\rightarrow does not, of itself, involve the idea of universality; but (α) illustrates how it may combine with 'all' to form a *universal hypothetical statement*. The grounds for the validity of a universal implication of the form

$$\Pi_x(a(x) \rightarrow b(x))$$

may of course be several. If they are solely to be found in the logical axioms, then the symbol \rightarrow expresses purely logical consequence. But the grounds may well be of a factual nature, such as a causal relation or some other empirically observed regularity. This remark may suffice to clarify the question as to how the relation of cause and effect is connected with that of logical reason and consequence. The symbol \rightarrow remains neutral with respect to all this.

The finite-logical axioms can be found listed in D. Hilbert and P. Bernays, *Grundlagen der Mathematik*, I, Berlin, 1934, p. 66. They are of course constructed in such a manner that their formal validity can be established by means of the 'finite rule.' Conversely it can be shown — but this already requires an essentially mathematical and not altogether trivial proof — that the list of axioms is complete, in the sense that all logical formulas containing only the symbols \sim , \rightarrow , $\&$, \vee which are formally valid according to the 'finite rule' can be obtained from these few axioms by substitution and repeated application of the syllogism. The group of transfinite axioms, of which we know as yet only Axioms I and II, remains in need of supplementation.

From the syllogism rule of inference other, derived, rules of inference may be obtained by means of the logical axioms. Indeed every formally valid proposition of the form $A \rightarrow B$ (where A and B are built up from the propositional variables a, b, \dots with the help of the logical connectives), by virtue of the syllogism, leads to the following rule of inference: If you have a proposition of the form A , then you may set down the corresponding proposition of the form B . Conversely, the syllogism also has its representative among the logical formulas:

$$a \rightarrow ((a \rightarrow b) \rightarrow b).$$

And yet, since construction means action, one does not get along with formulas alone; some practical rule of inference that tells how to handle the formulas is needed. This is probably the truth behind the opinion as to the normative character of logic.

Kant's distinction between analytic and synthetic judgments (*Critique of Pure Reason*, Introduction) is so obscurely phrased as to render a comparison with the precise concept of formal validity in mathematical logic almost impossible. The latter concept is in agreement, however, with Husserl's definition (*Logische Untersuchungen*, II, 2d ed., p. 254): "Analytic laws are unconditionally universal propositions containing no concepts other than formal. As opposed to the analytic laws we have their particular instances, which arise through the introduction of material concepts or of ideas positing individual existence. And as particular cases of laws always yield necessities, so particular cases of analytic laws yield analytic necessities." }

REFERENCES

- G. W. LEIBNIZ, *Philosophische Schriften*, VII, pp. 43-247, 292-301.
 — *Mathematische Schriften*, VII, 49-76, 203-216.
 L. COUTURAT, *La Logique de Leibniz*, Paris, 1901.
 — *Opusculs et fragments inédits de Leibniz*, Paris 1903.
 G. BOOLE, *The Mathematical Analysis of Logic*, London and Cambridge, 1847.
 — *An Investigation of the Laws of Thought*, London, 1854; reprinted in: *Collected Logical Works*, Chicago and London, 1916.
 E. SCHRÖDER, *Vorlesungen über die Algebra der Logik*, 3 Vols., Leipzig, 1890-95.
 A. N. WHITEHEAD and B. RUSSELL, *Principia Mathematica*, 3 Vols., Cambridge, 1910-13; 2d ed., 1925-27.
 B. RUSSELL, *The Principles of Mathematics*, Cambridge, 1903; 2d ed., New York, 1938.
 — *Introduction to Mathematical Philosophy*, 2d ed., London, 1920.
 L. WITTGENSTEIN, *Tractatus Logico-Philosophicus*, New York and London, 1922.
 C. I. LEWIS and C. H. LANGFORD, *Symbolic Logic*, New York, 1932.
 A. TARSKI, *Introduction to Logic* (translated by O. Helmer), New York 1941.

4. THE AXIOMATIC METHOD

The axiomatic method consists simply in making a complete collection of the basic concepts as well as the basic facts from which all concepts and theorems of a science can be derived by definition and deduction respectively. If this is possible, then the scientific theory in question is said to be *definite* according to Husserl. Such is the case for the theory of space. Of course, from the axioms of geometry I cannot possibly deduce the law of gravitation. Hence it was necessary to explain above what is to be considered a *pertinent* proposition of a given field of inquiry. Similarly the axioms of geometry fail to disclose whether Zurich is farther from Hamburg than Paris. Though this question deals with a geometrical relation, the relation is one between individually exhibited locations. Thus, precisely speaking,

what is supposed to be deducible from the axioms are the *pertinent general true propositions*.

{“Such, then, is the whole art of convincing. It is contained in two principles: to define all notations used, and to prove everything by replacing mentally the defined terms by their definitions.” Thus Pascal in a discourse *de l'esprit géométrique* (Oeuvres complètes, ed. F. Strowski, Paris (Librairie Ollendorff), I, p. 427). But this is more easily said than done. Euclid's *Elements* fail to afford a complete solution of the problem of axiomatizing geometry. He begins with *ῥοι*, definitions; but they are only in part definitions in our sense, the most important among them are descriptions, indications of what is intuitively given. Nothing else, in fact, is possible after all for the basic geometrical concepts such as ‘point,’ ‘between,’ etc.; but as far as the deductive construction of geometry is concerned, descriptions of this kind are evidently irrelevant. There follow, under the name of *αἰτήματα*, certain geometrical axioms, in particular the axiom of parallels: Given a plane P , a line l in P , and a point p in P not lying on l ; all lines in P which pass through p , except one, intersect l . Finally a few general axioms of magnitude: *κοινὰ ἔννοια*. They play their part in the development of geometry, inasmuch as certain geometrical relations such as congruence, or equality of areas, are tacitly assumed to satisfy these axioms. Behind them are concealed an indefinite number of proper axioms of geometry. In later books of the *Elements* the list of axioms is supplemented as the occasion demands. Because the geometrical postulates are intuitively self-evident and because a purely logico-deductive attitude is not natural to the human mind, it has required great pains to compile a complete list of geometrical axioms. ‘Non-Euclidean geometry,’ established by Bolyai and Lobatschewsky around 1830, becomes the driving force for axiomatic research in the second half of the 19th century. The most hidden axioms, those of order, are disclosed by Pasch around 1880. Finally, at the turn of the century, the goal is reached completely and finds its classical expression in Hilbert's *Grundlagen der Geometrie*. Hilbert arranges the axioms in five groups: the axioms of incidence, of order (‘betweenness’), of congruence, of parallelity, and of continuity.

The axiomatic procedure of the ancients, which aside from Euclid was also handled by Archimedes with admirable facility, became exemplary for the foundation of modern mechanics. It dominates Galileo's doctrine of uniform and uniformly accelerated motion (“Discorsi,” 3rd and 4th days), and even more so Huyghens' establish-

ment of the laws of the pendulum in his *Horologium oscillatorium*. In more recent times the axiomatic program was carried out completely (outside of mathematics proper) for the statics of rigid bodies, the space-time theory of special relativity, and other parts of physics.

An axiom system is by no means uniquely determined by the discipline in question; rather, the choice of the basic concepts and basic facts is arbitrary to a considerable extent. The question as to whether it is possible to differentiate between essentially originary and essentially derived notions lies beyond the competence of the mathematician.³ The definition of a geometrical relation concept that was originally chosen may with equal justification be replaced by any criterion which, in accordance with geometrical facts, is a necessary and sufficient condition under which the relation holds.]

An axiom system must under all circumstances be free from contradictions, in which case it is called *consistent*; that is to say, it must be certain that logical inference will never lead from the axioms to a proposition a while some other proof will yield the opposite proposition $\sim a$. If the axioms reflect the truth regarding some field of objects, then, indeed, there can be no doubt as to their consistency. But the facts do not always answer our questions as unmistakably as might be desirable; a scientific theory rarely provides a faithful rendition of the data but is almost invariably a bold construction. Therefore the testing for consistency is an important check; this task is laid into the mathematician's hands. Not indispensable but desirable is the *independence* of the individual axioms of an axiom system. It should contain no superfluous components, no statements which are already demonstrable on the basis of the other axioms. The question of independence is closely connected with that of consistency, for the proposition a is independent of a given set of axioms if and only if the proposition $\sim a$ is consistent with them.

The dependence of a proposition a on other propositions A (an axiom system) is established as soon as a concrete proof of a on the basis of A is given. In order to establish the independence, on the other hand, it is required to make sure that no combination of inferences, however intricate, is capable of yielding the proposition a . There are three methods at one's disposal of reaching this goal; by what has been said above, each of them qualifies also for proving the consistency of an axiom system.

(1) The first method is based on the following principle: if a con-

³ Sometimes this is certainly the case; e.g. among the kinship relations, 'child' and 'spouse' are the essentially originary ones.

tains a new original concept, not defined in terms of those occurring in A , then a cannot be a consequence of A . For example: a ship is 250 feet long and 60 feet wide; how old is its captain? Only in the most trivial cases does this simple idea accomplish our objective.

(2) *The construction of a model.* Objects and relations are exhibited which, upon suitable naming, satisfy all of the propositions A , and yet fail to satisfy a . This method has been the most successful so far invented.

The most famous example is furnished by the axiom of parallels. From the beginning, even in antiquity, it was felt that it was not as intuitively evident as the remaining axioms of geometry. Attempts were made through the centuries to secure its standing by deducing it from the others. Thus doubt of its actual validity and the desire to overcome that doubt were the driving motives. The fact that all these efforts were in vain could be looked upon as a kind of inductive argument in favor of the independence of the axiom of parallels, just as the failure to construct a perpetuum mobile is an inductive argument for the validity of the energy principle. Negating the axiom of parallels amounts to the assumption that, given a point P and a line l not passing through P , there exist in the plane determined by P and l an infinitude of lines through P not intersecting l . Therefore this is what the constructors of non-Euclidean geometry did: they drew the consequences of that assumption, and in doing so they found, even though they made free use of the remaining axioms of Euclidean geometry, that no contradiction arose, *as far as they followed the matter up*. But they could not claim security for all future. Klein was the first to offer a Euclidean model for non-Euclidean geometry; the objects of Euclidean geometry itself, upon an assignment of names differing from the customary one, satisfy the non-Euclidean axioms. Let S be a sphere in Euclidean space. The dictionary which furnishes the translation into non-Euclidean language consists of only a few words (here characterized by quotation marks): by a '*point*' we understand any point in the interior of S . Several such '*points*' are said to lie on a '*line*' or in a '*plane*,' and a '*point*' is said to lie '*between*' two others, if they do so in the customary sense. A '*motion*' is any collinear transformation which transforms the sphere S into itself; two configurations are '*congruent*' if one results from the other by a '*motion*.' For anyone who believes in the truth and thus in the consistency of Euclidean geometry, the consistency and thus conceivability of non-Euclidean geometry is thereby established.

The consistency of Euclidean geometry, on the other hand, can be demonstrated quite independently of the belief in its truth and of the intuitive content of its basic concepts. For analytic geometry, which

can best be based on the concept of vector (see Section 12), has shown that Euclidean geometry is but a different expression of the facts of *linear algebra*, of the theory of linear equations; it has thus provided us with a simple *arithmetical model* of Euclidean space. Linear algebra accounts for the affine concepts of geometry, while the adjunction of a positive definite quadratic form that serves as the 'metric ground form' leads to the metrical concepts. In algebra the number n of variables (or 'unknowns') may be left indeterminate. One has to choose $n = 3$ in order to get the geometry of the intuitive 3-dimensional space. Arithmetic and geometry, by virtue of this correspondence, are so closely interwoven that today even in pure analysis we constantly make use of geometrical terms. Any contradiction in geometry would at the same time show up as a contradiction in arithmetic. This may be looked upon as a reduction, since the numbers are to a far greater measure than the objects and relations of space a free product of the mind and therefore transparent to the mind.

These examples indicate that the method of models need not be restricted to those cases where the truth about the objects and relations employed in the construction of the model is known, but that it may serve to reduce the consistency of an axiom system A (e.g. that of geometry) to the consistency of another, B (e.g. that of arithmetic). This is achieved whenever the basic concepts of the system A are defined in terms of the basic concepts of the system B in such a manner that the axioms A become a logical consequence of the axioms B . No attention has to be paid for this purpose to the intuitive meaning of the basic concepts in A and in B ; the assignment of the names given to the basic concepts of A to certain concepts derived from B is purely arbitrary.

More than anybody else has Hilbert, through the ingenious construction of suitable arithmetical models, contributed to the clarification of the logical relations that connect the various parts of the geometrical system of axioms.

If we are dealing with a finite number of objects only which are explicitly exhibited one by one and designated by symbols, we may be able to prove consistency by stating for each single instance in terms of the symbols whether or not the basic relations obtain. As an example we give a *combinatorial model* that ensures the consistency of the incidence axioms in plane projective geometry (which deal with the single relation 'point lies on line'). The model consists of seven symbols for points, 1, 2, 3, 4, 5, 6, 7, and seven symbols for lines, I, II, III, IV, V, VI, VII, and incidence is defined by the following table, in which a *, say, at the crossing of row 3 and column VI indicates that point 3 lies on line VI:

	I	II	III	IV	V	VI	VII
1		*	*	*			
2	*		*		*		
3	*	*				*	
4	*			*			*
5		*			*		*
6			*			*	*
7				*	*	*	

For example, verify from this table the axioms stating that through any two distinct points there goes exactly one line (i.e. any two rows contain exactly one pair of *'s in the same column) and that any two distinct lines intersect at exactly one point!

The case of a finite system of objects exhibited one by one is comparatively trivial. In all other cases the method of models is merely capable of reducing the consistency of one system to that of another. Ultimately it will become necessary to prove consistency in an absolute way for one basic system of axioms. For the larger part of mathematics and for the whole of physics this basic system deals with the concept of *real number*.

(3) For the purpose of an absolute proof of consistency we have none but the *direct method* at our disposal, which endeavors to show that by following the rules of deductive inference one will never arrive at two propositions of which one is the negation of the other. Complete enumeration of the logical rules of the game is here a necessary presupposition (compare Section 3); for only then can one apply the method to propositions, blind against their meaning, as one applies the rules of chess to chessmen. Only in recent years has Hilbert attacked the problem of securing the consistency of the arithmetical axioms in this manner. (Should a new and evidently stringent method of logical inference be discovered and thus the set of rules of the game be augmented, one would have to be prepared to see a consistency proof conducted by the direct method become obsolete. The method of models, on the other hand, is independent of the 'rules of the game.')

{The following might serve as an analogue in chess: it is required to see that a game of chess, no matter what the various moves, as long as it is played in accordance with the rules, can never lead to a position in which there are 10 queens of the same color on the board. Here the 'direct method' is applicable. For it can be gathered from the

rules of the game that no move increases the sum of the numbers of queens and pawns of the same color. Hence, as this sum is 9 initially, it must remain ≤ 9 . Incidentally, method (1) is a special case of the direct method, but it seemed to deserve special mention because of its simplicity.

In addition to consistency and independence, the *completeness* of the axioms which form the basis of a science will be required. What is meant by that? That for every pertinent general proposition a the question 'does a or $\sim a$ hold?' be decidable by logical inference on the basis of the axioms? Just as consistency guarantees that not *both* a and $\sim a$ can be obtained, completeness would then guarantee that always *one* of them can be obtained. Completeness in this sense would only be ensured by the establishment of such procedural rules of proof as would lead demonstrably to a solution for every pertinent problem. Mathematics would thereby be trivialized. But such a philosopher's stone has not been discovered and never will be discovered. Mathematics does not consist in developing the logical consequences of given assumptions omnilaterally, but intuition and the life of the scientific mind pose the problems, and these cannot be solved by mechanical rules like computing exercises. The deductive procedure that may lead to their solution is not predesigned but has to be discovered in each case. Analogy, experience, and an intuition capable of integrating multifarious connections are our principal resources in this task. As was already mentioned in Section 3, there is no descriptive characteristic of all propositions deducible from given premisses; we have to rely on construction. It is not feasible in practice to proceed like Swift's scholar, whom Gulliver visits in Balnibarbi, namely, to develop in systematic order, say according to the required number of inferential steps, all consequences and discard the "uninteresting" ones; just as the great works of world literature have not come into being by taking the twenty-six letters of the alphabet, forming all 'combinations with repetition' up to the length 10^{10} , and selecting and preserving the most meaningful and beautiful among them.

Suppose we make a continuous deformation of space (as if it were filled out with plasticine), and suppose we understand now by lines, planes, and congruent figures such curves, surfaces, and figures as result from real lines, real planes, and really congruent figures by this deformation. Then evidently all the facts of geometry hold for these newly introduced concepts. It is therefore impossible to distinguish conceptually between the system of lines and the system of curves resulting from them by a spatial deformation. }

This brings us to the idea of *isomorphism*, which is of fundamental importance for epistemology. Let us assume that we have a system Σ_1 of objects (such as the points, lines, and planes of geometry) and certain appertaining basic relations R, R', \dots . Let there be a second system Σ_2 , with corresponding basic relations which (though they may have entirely different meanings) are correlated, say, by the use of the same names, to the relations R, R', \dots within the first domain of objects. Then, if it is possible to state a rule by which the elements of the system Σ_1 are paired in a mutually unique manner with the elements of the system Σ_2 , so that elements in Σ_1 between which R (or R', \dots) holds correspond to elements in Σ_2 between which the relation with the same name R (or R', \dots respectively) holds, then the two domains are said to be *isomorphic*. The correlation in question is said to be an *isomorphic mapping* of Σ_1 into Σ_2 . Isomorphic domains may be said to possess the same *structure*. For every pertinent true proposition about Σ_1 (whose sense can be understood by virtue of the meanings of R, R', \dots within Σ_1), there is a corresponding and identically phrased proposition about Σ_2 , and conversely. Nothing can be asserted of the objects in Σ_1 that would not be equally valid in Σ_2 . Thus, for example, Descartes' construction of coordinates maps the space isomorphically into the operational domain of linear algebra. These considerations induce us to conceive of an axiom system as a *logical mold* ('*Leerform*') of possible sciences. A concrete interpretation is given when designata have been exhibited for the names of the basic concepts, on the basis of which the axioms become true propositions. One might have thought of calling an axiom system complete if in order to fix the meanings of the basic concepts present in them it is sufficient to require that the axioms be valid. But this ideal of uniqueness cannot be realized, for the result of an isomorphic mapping of a concrete interpretation is surely again a concrete interpretation. Hence the final formulation has to be as follows: an axiom system is complete, or *categorical*, if any two concrete interpretations of it are necessarily isomorphic. In this sense the categoricity of Hilbert's axiom system of Euclidean geometry is guaranteed. Indeed it can easily be shown that a space satisfying these axioms is isomorphic to the algebraic model provided by Descartes' analytic geometry.

{A science can only determine its domain of investigation up to an isomorphic mapping. In particular it remains quite indifferent as to the 'essence' of its objects. That which distinguishes the real points in space from number triads or other interpretations of geometry one

can only *know* (*kennen*) by immediate intuitive perception. But intuition is not blissful repose never to be broken, it is driven on toward the dialectic and adventure of cognition (*Erkenntnis*).⁴ It would be folly to expect cognition to reveal to intuition some secret essence of things hidden behind what is manifestly given by intuition. The idea of isomorphism demarcates the self-evident insurmountable boundary of cognition. This reflection has enlightening value, too, for the metaphysical speculations about a world of things in themselves behind the phenomena. For it is clear that under such a hypothesis the absolute world must be isomorphic to the phenomenal one (where, however, the correlation needs to be unique only in the direction thing in itself \rightarrow phenomenon); for "we are justified, when different perceptions offer themselves to us, to infer that the underlying real conditions are different" (Helmholtz, *Wissenschaftliche Abhandlungen*, II, p. 656). Thus even if we do not *know* the things in themselves, still we have just as much *cognition* about them as we do about the phenomena. The same idea of isomorphism clarifies the problem which Leibniz, stimulated by Hobbes' nominalistic theory of truth, treats in his dialogue on the connection between things and words; Leibniz evidently wrestles with giving expression to that idea (*Philosophische Schriften*, VII, pp. 190–193).

Through the disclosure of isomorphic relations it is possible to transfer any insights gained in one field to the isomorphic field. A service of this kind is rendered, for instance, by the principle of duality in plane projective geometry. Its only relational concept is the incidence of point and line (point lies on line, line passes through point). It is possible to pair off uniquely the points and the lines in the plane in such a manner that, whenever a point P lies on a line q , the line p paired with P passes through the point Q paired with q . Consequently any valid theorem of projective geometry (phrased in terms of the directionless relation of incidence) at once becomes another valid theorem if the words 'point' and 'line' are interchanged. S. Lie discovered that the lines of (complex) space may be uniquely correlated with the spheres in such a manner that intersecting lines correspond to tangent spheres. An important part of analytic function theory, the so-called theory of uniformization, may be treated most naturally in the language of Bolyai-Lobatschewskyan geometry. Let an electrical network be given which consists of individual homo-

⁴ Unfortunately English uses the same word 'know' for the two meanings that the author's German distinguishes as *kennen* and *erkennen*, and that the Latin, French and Greek languages express by the pairs *cognoscere* vs. *scire*, *connaître* vs. *savoir*, *γινῶναι* vs. *εἰδέναι*. Our translation is inconsistent in so far as it uses the terms *cognize*, *cognition* in contrast to *know* and *knowledge* only in places where the distinction is essential. [Translator's note.]

geneous wires connecting at various branch points; if by a 'Point' we understand an arbitrary current distribution, which assigns to each (oriented) wire s the intensity I_s of the electric current in s , then these Points satisfy the laws of a Euclidean space with a center O and as many dimensions as there are branches in the network. Here the central point O is represented by the absence of current where every I_s vanishes, and the square of the distance of a 'Point' from O is defined as the Joule heat developed by the current distribution per unit of time. This isomorphism is of value since it correlates the simple and important notions of geometry and the simple and important physical notions concerning electrical circuits. For instance, the basic problem of finding the current distribution when the various electromotive forces in the wires s are given is identical with the geometrical problem of finding the perpendicular projection of a Point onto a plane. The existence of a unique solution is thereby at once mathematically established, and a method for computing the solution made available. }

Pure mathematics, in the modern view, amounts to a general hypothetico-deductive theory of relations; it develops the theory of logical 'molds' without binding itself to one or the other among the possible concrete interpretations. Concerning this *formalization*, as "a point of view, without which an understanding of mathematical methods is out of the question," compare Husserl, *Logische Untersuchungen*, I, Sections 67-72. "The presupposition for the erection of a general arithmetic," Hankel declares (*Theory of Complex Numbers*, 1867, p. 10) "is thus a purely intellectual mathematics, dissociated from all intuition, a pure theory of forms, which has as its object not the combination of quanta or their images, the numbers, but intellectual objects, to which there may (but need not) correspond actual objects or relations." The axioms become *implicit definitions* of the basic concepts occurring in them. The concepts, admittedly, retain a certain range of indeterminacy; but the logical consequences of the axioms are valid, no matter what concrete interpretation may be adopted within this range. Pure mathematics acknowledges but one condition for truth, and that an irremissible one, namely *consistency*.

{Perhaps there already is an inkling of this modern view in Euclid's term for axioms: *αἰρήματα*, postulates. Leibniz takes some decisive steps towards the realization of a *mathesis universalis* in the sense here indicated and clearly understood by him. The theory of groups above all, that shining example of "purely intellectual mathematics," belongs within the framework of his *ars combinatoria*. A finite group

G is a system of a finite number of objects within which, in some way, an operation is defined which generates, from two (equal or different) elements a, b (in this order), an object ab of the system. The only postulates, or axioms, are these:

the associative law $a(bc) = (ab)c$;
if $a \neq b$ (a different from b), then also $ac \neq bc, ca \neq cb$.

From these insignificant looking assumptions springs an abundance of profound relationships; and mathematics offers an astounding variety of different interpretations of this simple axiom system. The group is perhaps the most characteristic concept of the mathematics of the 19th century.

The method of implicit definition is of importance also within the sciences themselves, and not only in the laying of their foundations. The *area* of a *piece*, where by the latter I will understand a piece of the plane that is bounded by line segments, satisfies the following requirements:

- (i) The area is a positive number.
- (ii) If a piece is dissected into two parts by a sequence of line segments in its interior, then the area of the whole is equal to the sum of the areas of the parts.
- (iii) Congruent pieces have the same area.

These are the really essential properties of the concept of area but they contain no explicit definition of it. It can be shown, however, that these requirements are consistent and that a procedure can be devised by which every piece γ is assigned a positive number $J(\gamma)$ as its area which satisfies requirements (ii) and (iii). The requirements fail to determine the concept unambiguously; they are also satisfied, apart from $J(\gamma)$, by $c \cdot J(\gamma)$, where c is any positive constant independent of γ . But beyond this there are no further possibilities. The remaining arbitrariness as expressed by the factor c , can only be eliminated by the exhibition of an individual piece, say, a square, and the stipulation that it be assigned the area 1 (relativity of size). The significance of the implicit definition within all sciences, not only mathematics, is expounded very aptly in Schlick's *Allgemeine Erkenntnistheorie* (Berlin, 1918, pp. 30-37). "From the viewpoint of exact science, which strings inference after inference, a concept is indeed nothing but that of which certain propositions may be asserted. Thereby it should consequently also be defined." A suitable field of application, aside from the exact sciences, might be jurisprudence.]

R E F E R E N C E S

- M. PASCH, *Vorlesungen über neuere Geometrie*, 2d ed., Berlin, 1926.
 D. HILBERT, *Grundlagen der Geometrie*, 1899; 7th ed. Leipzig, 1930.
 ——— *Axiomatisches Denken*, *Mathematische Annalen* 78, 1918.
 M. GEIGER, *Systematische Axiomatik der euklidischen Geometrie*, Augsburg, 1924.
 F. GONSETH, *Les mathématiques et la réalité*, Paris, 1936.

CHAPTER II

Number and Continuum, the Infinite

5. RATIONAL NUMBERS AND COMPLEX NUMBERS

THE genetic construction of the mathematical realm of numbers takes as its point of departure the sequence of natural numbers 1, 2, 3, The first step to be made is the rise from the natural numbers to the fractions. Historically *fractions* owe their creation to the transition from counting to *measuring*. All measuring is based on a domain of magnitudes, such as the segments on a line. We have here (1) a relation of equality, $a = b$ (congruence), satisfying the axioms set up for such a relation (p. 9), and (2) an operation applicable to any two segments a, b and producing a segment $a + b$. From the segment a we obtain, say, the segment $5a$ by forming the sum $a + a + a + a + a$ with 5 terms a . This brings out the connection between counting and measuring. This process of iteration which leads from a to $1a, 2a, 3a, \dots$ can be exactly explained as follows:

$\alpha)$ $1a = a$;

$\beta)$ if n is a natural number, then $(n + 1)a$ results from na in accordance with the formula

$$(n + 1)a = (na) + a.$$

Within the domain of line segments, the operation of iteration admits of a unique inversion, partition: given a segment a and a natural number n , there exists one and (in the sense of equality) only one segment x such that $nx = a$; it is denoted by a/n . The operation of partition may be combined with that of iteration. Thus e.g. we get $5a/3$, called '5/3 times' a . The fractional symbol m/n serves as the symbol of the composite operation, so that two fractions are equal if the two operations denoted by them lead to the same result, no matter to what segment a they are being applied. *Multiplication* of fractions is performed by carrying out one after another the operations denoted by them. The possibility of *adding* fractions is due to the fact that the operation (applied to an arbitrary segment x) that is expressed by

$$(mx/n) + (m^*x/n^*)$$

can be represented by a single fraction.

It is unnecessary to introduce special fractions for each domain of magnitudes. Since their laws are independent of the nature of these magnitudes, it is more expedient to define them in purely arithmetical terms.⁵ This can be achieved by simply choosing as domain of magnitudes in the above considerations the natural numbers themselves. The fact that in this domain a relation between x and y , such $5x = 3y$, cannot always for a given x be solved with respect to y , does not impede the development of the theory. We thus arrive at the following formulation: "Two natural numbers m , n determine a fraction m/n . The statement that, of any two natural numbers x and y , the second is m/n times the first is merely another form of expressing the equation $mx = ny$." This is a creative definition in the sense of Section 2. Two fractions m/n , m^*/n^* are equal provided any numbers x , y which stand in the relation $mx = ny$ also stand in the relation $m^*x = n^*y$, and conversely. The operational rules for natural numbers permit one to replace this transfinite criterion, whose phrasing seems to require a checking through of all possible numbers x , y , by the following finite one:

$$(C) \qquad m \cdot n^* = n \cdot m^*.$$

Hence we deal with a special case of the definition by abstraction: the equality of the fractions m/n , m^*/n^* may be explained directly by (C), after one has convinced oneself that this relation is an equivalence. The introduction of fractions as 'ideal elements' can also be motivated purely arithmetically without reference to applications. Indeed after the numerical operations have been suitably extended to fractions it is found that all the important arithmetical axioms remain in force. Moreover division, the inverse operation of multiplication, can now always be carried out while this was only exceptionally so in the arithmetic of natural numbers.

[If the same idea is applied once again for the purpose of ensuring the invertibility of addition, then we get from the fractions to the *rational numbers* (which include 0 and the negative). (This, though, calls for one rather serious sacrifice — the possibility of division has to be abandoned for the divisor 0.) There are nowhere in this pro-

⁵ This is in line with the oldest mathematical tradition, that of the Sumerians. Only after the discovery of the irrational did the Greeks abandon the algebraic road and find themselves compelled to couch algebraic facts in geometric terms. The post-classical Occident, partly stimulated by the algebraic achievements of the Arabs, reversed this development. There was little justification, however, for the modern viewpoint subsuming all quantities under a universal concept of number, before Dedekind gave Eudoxus' analysis of the irrational its constructive twist (cf. Section 7).

cedure any logical obscurities or philosophical difficulties. A much more serious matter is the starting point, the system of natural numbers, and then the irrational, the transition from the rational numbers to the continuum of real numbers. But once we have climbed to this level, the further advance toward the complex and hypercomplex numbers no longer leads past any abysses. In order to introduce the *complex numbers* it is only necessary to describe how any such number is given and how one is to operate with them. A complex number is given by its two components; thus we might as well say that we understand by a complex number any pair (α, β) of real numbers (Hamilton, 1837). We shall not set down the rules of operation here explicitly. According to them, $e = (1, 0)$ plays the part of unity in the complex domain, since its multiplication by any complex number (α, β) reproduces (α, β) . And $(0, 1)$ is that imaginary unit i which satisfies the equation $i \cdot i = -e$. The inner reason for the stipulations is again to be seen in their extending the formal rules of computation from real to complex numbers. Nothing remains of the mystic flavor that was so long attached to the imaginary quantities.⁶ From the complex it is possible to ascend to the *hypercomplex numbers* with 3 or more components. But it could be shown quite generally that, no matter how addition and multiplication be defined in their domain, the continued validity of all rules of operation of arithmetic is unattainable. In this respect the complex numbers denote a natural boundary for the extension of the number concept. Yet also hypercomplex number systems play their role in mathematics; thus the 4-component *quaternions*, which satisfy all rules of operation except the commutative law of multiplication, are a useful tool in dealing with the rotations of a rigid body in space.

Instead of constructing the realm of numbers genetically, arithmetic may also be based on an axiom system. From this viewpoint the genesis merely serves to reduce the consistency of that system to the consistency of the axioms governing the natural numbers. The

⁶ For instance, Huyghens declares in 1674 (see Leibniz, *Mathematische Schriften*, II, p. 15) with reference to a complex formula: "Il y a quelque chose de caché là-dedans, qui nous est incompréhensible." Even Cauchy, in 1821, still has a somewhat obscure idea as to the manipulation of complex quantities. But negative quantities had produced almost as many headaches at an earlier time. Referring to the rule "minus times minus is plus," Clavius says in 1612: "debilitas humani ingenii accusanda (videtur), quod capere non potest, quo pacto id verum esse possit." Descartes, in accordance with contemporary usage, still designates the negative roots of an algebraic equation as false roots. The explanation, surviving in some textbooks, of i as that number which, when multiplied by itself, yields -1 is of course pure nonsense as long as only the real numbers are at one's disposal; it merely contains the demand that the number concept be so expanded and the sense of multiplication be so extended to the expanded number domain as to produce the desired equation.

axioms of arithmetic fall into two groups, the algebraic axioms and the axioms of magnitude. The algebraic group deals with the operations of addition and multiplication. It contains the formal rules of operation (such as $a + b = b + a$), requires the existence of a 0 and a 1 with the properties

$$a + 0 = 0 + a = a, \quad 1 \cdot a = a \cdot 1 = a$$

and the invertibility of addition and multiplication (with the exception of division by 0). The axioms of magnitude (which do not carry over to the domain of complex numbers) deal with the relation $a > b$ (a greater than b). Compare the table in Hilbert's *Grundlagen der Geometrie*.]

REFERENCES

- W. R. HAMILTON, *Lectures on Quaternions*, Dublin, 1853.
 H. HANKEL, *Theorie der komplexen Zahlen*, Leipzig, 1867.
 D. HILBERT, *Grundlagen der Geometrie*.
 O. HÖLDER, *Die Arithmetik in strenger Begründung*, 2d ed., Leipzig, 1928.

6. THE NATURAL NUMBERS

"The integers were created by God; all else is man-made," is a frequently quoted statement of Kronecker's. In the natural numbers, the problem of cognition presents itself to us in its simplest form. Let us once more begin with the purely mathematical aspect.

The sequence of natural numbers commences with 1 and is generated by a process which yields from a number already obtained the next following number; never does an earlier number recur in this progression. A concept (a characteristic or an operation) referring to arbitrary numbers can therefore be introduced only by *complete induction* (also called mathematical induction), namely by stating (α) what the concept means for the first number 1, and (β) how it carries over from any number n to the next following n' ($= n + 1$). Examples: The definition of na in the preceding section. The concepts even and odd: (α) 1 is odd; (β) n' is even or odd according as n is odd or even. The general notion of addition $a + n$ of two natural numbers a and n :

$$(\alpha) \quad a + 1 = a'; \quad (\beta) \quad a + n' = (a + n)'$$

What is true of the concepts similarly holds for the proofs. To prove that a certain theorem holds for every number one shows (α) that it holds for 1, and (β) that it holds for n' if n is a number for which it holds. With the help of this method of definition and of proof by

complete induction, of inference from n to $n + 1$, the theory of natural numbers can be completely built up step by step. That inference introduces an entirely new and peculiar feature unknown to Aristotelian logic into the mathematical method; it is the very essence of the art of mathematical demonstration. The first explicit mention of the principle of complete induction seems to be with B. Pascal (1654) and Jacob Bernoulli (1686).

In the building up of number theory by complete induction, the *succession* of numbers appears as their constitutive characteristic. They occur primarily as *ordinal numbers* and are distinguished only by their position in the sequence. Justly Schopenhauer (*Vierfache Wurzel vom zureichenden Grunde*, Section 38) says of this conception of number, "Every number presupposes the preceding ones as reasons for its being: I can get to ten only through all the preceding numbers. . . ." The well-known method of counting, applied to a given aggregate of objects, produces a certain natural number as the *number* (*Anzahl*) of *elements* in the aggregate. By virtue of the counting process the elements of the aggregate are themselves arranged in a sequence (first, second, third, . . .); and a special consideration is required to ensure the fundamental fact that the result of counting is independent of the order. Only thus is the concept of cardinal number put on a safe basis. Compare, for instance, the treatment by Helmholtz (*Zählen und Messen*, *Wissenschaftliche Abhandlungen*, III, p. 356); further L. Kronecker (*Werke*, III, 1, p. 249).

{The question has been argued extensively whether the concept of cardinal, rather than ordinal, number is not the primary one. The former, if it is to be introduced independently of an ordinal arrangement, has to be defined by abstraction (as on p. 10). This definition is not even restricted to finite sets; a theory of infinite cardinal numbers based thereon was developed by G. Cantor within the framework of his general set theory. But the criterion of numerical equivalence makes use of the possibility of pairing, which can only be ascertained if the acts of correlation are carried out one after another in temporal succession and the elements of the sets themselves are thereby arranged in order. Even if one follows the road of abstraction and splits up the act of numerical comparison of two sets by first ascribing a number to each set and then comparing these numbers, it remains indispensable to order each individual set itself by exhibiting its elements one by one in temporal succession. (Such a one-by-one exhibition is necessary anyhow if an aggregate is to be considered as concretely given; and the numbers employed by us in everyday life concern only such aggregates.) For this reason it seems to me unquestionable that the

concept of ordinal number is the primary one. Modern research in the foundations of mathematics, which has destroyed dogmatic set theory, confirms this view.

Another point of debate is the question whether the numbers are independent ideal objects or whether arithmetic is concerned merely with the concrete *numerical symbols* "whose shape is recognizable by us with certainty independently of place and time, of the particular conditions of their manufacture, and of trifling differences in their execution" (Hilbert). Thus e.g. Helmholtz (*Zählen und Messen*, *loc. cit.*, p. 359): "I consider arithmetic, or the theory of pure numbers, as a method built upon purely psychological facts, by which the consistent application of a system of symbols of unlimited extent and unlimited possibility of refinement is taught. In particular, arithmetic investigates what different modes of combination of these symbols (numerical operations) lead to the same result." Only recently Hilbert carried this point of view consistently into effect (compare Section 10), in a manner unassailable even by the criticism directed against it by Frege (*Grundgesetze der Arithmetik*, 1893). A succession of strokes ('ones') offers itself as a suitable symbol. If I hear a sequence of tones, I put down a stroke upon hearing each one, placing one stroke after another: $////$. A second time I proceed similarly, again obtaining a symbol consisting of a succession of strokes. If I were immediately able to judge the equality or disparity of the 'shape' of the two symbols, a numerical comparison would be accomplished. Here the representation of the data by strokes has the function of putting these data into a 'normal form' of such a kind that a difference in shape at once indicates a difference in number. (For a directly given whole, number is meant to describe a relation between the whole and such parts of it as are considered as units. A difference in the shape of two wholes does not necessarily imply a difference in the number of units; e.g. $:\cdot:$ and $:\ddot{}$. An act of assembling is said to be the basis for determining the number of elements. It seems to me that the application of the symbolic method of counting to a structural whole of units does not require that a mere 'aggregate' be abstracted by dissolving the structural tie; nor need individually given elements, such as successive tones, be assembled to form an aggregate. The statement 'there were this many tones: $////$ ' is quite intelligible in itself, and it is unnecessary to search for an 'aggregate of the tones heard.') The immediate recognition of equality or disparity of two symbols consisting of successions of strokes is possible, however, for the lowest numbers only. In general one has to proceed by using the strokes recorded during the first sequence over again, say, by crossing them out one by one; for this purpose it is required that the first

sequence stays put (and does not disappear like the tones themselves). In principle, symbols can be dispensed with for the verification of a statement such as 'this time there were more tones than the first time,' provided the tones of the first sequence (which may have been falling in pitch) can be reproduced in their temporal succession while the second sequence is being listened to. Symbols become indispensable only when the comparison is torn up into two number determinations ('the first time there were 4, now there are 5 tones; 5 is greater than 4'); for then part of the mental operation ('5 is greater than 4') is shifted onto the permanent symbols, which are at the same time expedient for preservation and communication. Thus it is not the comparison of numbers but the determination of numbers which is of an essentially symbolic character. 'There were 4 tones' is unintelligible without reference to a symbol.

If one wants to speak, all the same, of numbers as concepts or ideal objects, one must at any rate refrain from giving them independent existence; their being exhausts itself in the functional role which they play and their relations of more or less. (They certainly are not concepts in the sense of Aristotle's theory of abstraction.)

The employment of several digits and the *positional system* (developed in Mesopotamia and later consistently by the Indians for written numbers) permit a quick decision about greater and smaller for much larger numbers than the simple numerical symbols composed of successive Ones; this considerable practical advantage is not one of principle however. The basis of the number system, which in our system is ten, is different with different cultures. The Indian, and particularly the Buddhistic, literature revels in the possibilities of producing and designating prodigious numbers by means of the positional system, that is, by combination of addition, multiplication, and exponentiation. In spite of their fantastic aspect there is something truly great in these efforts; the human mind for the first time senses its full power to fly, through the use of the symbol, beyond the boundaries of what is attainable by intuition. Something akin we find among the Greeks only in the latest epoch, namely, in Archimedes' paper addressed to Gelon "The Sand-reckoner"; and here is manifested the delight, not in the step by step opening-up of the infinite, but in the rational subjugation of the unbounded.

Regarding the *relation of number to space and time* we may say that time, as the form of pure consciousness, is an essential, not an accidental, presupposition for the mental operations on which the sense of a numerical statement is founded. Contrary to the opinion of some philosophers (e.g. Hobbes), this does not apply to space, although permanent symbols having a spatial configuration are the

most convenient means of putting down a result of counting, of storing and communicating it, and of safeguarding the manipulation of numbers. Kant above all has emphasized the bond between the number concept and time, but it would be going too far if one were to claim arithmetic as the science of time in the same sense that geometry is the science of space.

With reference to two concretely given numerical symbols, m and n , the sense of the proposition $m + n = n + m$ can be described without having to 'generate' any other numbers. It is also possible to see that this proposition holds in *any* concrete situation. Something new happens, however, when I imbed the actually occurring numerical symbols in the *sequence of all possible numbers*. That sequence is produced through a generating process according to the principle that any given number gives rise to a new, the next following, number by the addition of One. Here the being is projected upon the background of the *possible*, of a manifold of possibilities which is produced by a fixed process and yet is open towards infinity. This is the standpoint held by us at the beginning of the present section when arithmetic was founded on the principle of complete induction. We rely on it when we speak of a trillion ($= 10^{12}$) dollars. By repeated application of definitions by complete induction we obtain from the prime arithmetical process of changing n into $n + 1$ the operation of multiplication by 10, and by performing this operation 12 times (beginning with 1), we arrive at the desired number 10^{12} . The numbers 10 and 12 can be written out in strokes; as for 10^{12} , this has never been done, and yet we maintain the 'fiction' of such a number.]

Thus it is already in the field of numbers that we encounter the following basic features of *constructive cognition*:

1. We ascribe to that which is given certain characters which are not manifest in the phenomena but are arrived at as the result of certain mental operations. It is essential that the performance of these operations is held universally possible and that their result is held to be uniquely determined by the given. But it is not essential that the operations which define the character be actually carried out.

2. By the introduction of symbols the assertions are split so that one part of the operations is shifted to the symbols and thereby made independent of the given and its continued existence. Thereby the free manipulation of concepts is contrasted with their application, ideas become detached from reality and acquire a relative independence.

3. Characters are not individually exhibited as they actually occur, but their symbols are projected on the background of an ordered mani-

fold of possibilities which can be generated by a fixed process and is open into infinity.

Cognition has not stopped here. The leap into the beyond occurs when the sequence of numbers that is never complete but remains open toward the infinite is made into a closed aggregate of objects existing in themselves. Giving the numbers the status of ideal objects becomes dangerous only when this is done. The belief in the absolute is deeply implanted in our breast; no wonder, then, that mathematics was bold and naive enough to perform that leap. Whoever accepts as meaningful the definition ' n is an even or odd number according as a number x does or does not *exist* such that $n = 2x$,' which refers to the infinite totality of all numbers (the definition of even and odd by complete induction, as mentioned earlier, is a different matter), already stands on the other shore; for him the system of numbers has become a realm of absolute existences which is 'not of this world' and from which only gleams here and there are caught and reflected in our consciousness. The vindication of this transcendental point of view forms the central issue of the violent dispute which has flamed up again today over the foundations of mathematics. The issue is symptomatic for all knowledge and may, in the field of mathematics sooner than elsewhere, lead to a clear decision.

REFERENCES

- R. DEDEKIND, *Was sind und was sollen die Zahlen?*, Braunschweig, 1888; 3rd ed., 1911.
 G. FREGE, *Die Grundlagen der Arithmetik*, Breslau, 1884; reprinted 1934. *Grundgesetze der Arithmetik*, 2 vols., Jena, 1893-1903.
 E. HUSSERL, *Philosophie der Arithmetik*, Halle, 1891.

7. THE IRRATIONAL AND THE INFINITELY SMALL

In a different form than in the sequence of integers we encounter the infinite in the *continuum*, which is capable of infinite division. Cases of special importance are the continua of time and of space. Here we find the second open place in the above described construction of the mathematical realm of numbers. Antiquity has bequeathed to us two important contributions to the problem of the continuum: (a) a far-reaching analysis of the mathematical question of how to fix a single position in the continuum, and (b) the discovery of the philosophical paradoxes which have their origin in the intuitively manifest nature of the continuum.

[(a) The pure geometry of the Greeks, in elevating itself above the inexactitude of the sense data, applies the idea of existence (not

only to the natural numbers but also) to the points in space. The discovery of the irrationality of the ratio $\sqrt{2}$ of the diagonal and side of a square made it clear that the fractions are not the only possible quantities measuring ratios of line segments, and thus not the only 'real numbers.' In the Platonic dialogues the deep impression can be sensed which this mathematical discovery made upon the rising scientific consciousness of his time. Independently of the particular geometrical constructions which led to individual irrationalities such as $\sqrt{2}$, Eudoxus recognized the general foundations of this phenomenon.

1. In place of the untenable commensurability he sets down the axiom: if a and b are any two segments, then a can always be added to itself so often that the sum na exceeds b . This means that all segments are of a comparable order of magnitude, or that there exists neither an actually infinitely small nor an actually infinitely large in the continuum.

2. And what is it that characterizes the individual segment ratio? Eudoxus replies: two segment ratios, $a:b$ and $a':b'$, are equal to each other if, for arbitrary natural numbers m and n , the fulfillment of the condition in the first line below invariably entails the validity of the corresponding condition in the second line:

$$(I) \begin{cases} na > mb \\ na' > mb' \end{cases} \quad (II) \begin{cases} na = mb \\ na' = mb' \end{cases} \quad (III) \begin{cases} na < mb \\ na' < mb' \end{cases}$$

Hence what is characteristic of the individual real number α is the cut which it creates in the domain of rational numbers by dividing all fractions m/n into three classes, those which are (I) less than α , (II) equal to α , and (III) greater than α . The second class is either empty or contains only a single fraction. The first axiom guarantees that no two different segments can have the same ratio to the fixed unit segment. Euclid's theory of proportions is likewise erected on this foundation, while Archimedes bases on it his general method of exhaustion.

Only in the 19th century did mathematics go beyond Eudoxus, and settled the problem in a more definite fashion. For Eudoxus the real number is given as the ratio of two given segments, and thus it is up to the axioms of geometry to tell us what segment ratios exist. But in Euclidean geometry it is not possible to construct (by means of ruler and compass) from a given segment 1, the segment $\sqrt[3]{2}$, which would solve the Delian problem of duplicating the cube, or of the segment π , which equals the circumference of a circle of diameter 1. Yet we are convinced of their existence on the basis of continuity considerations: if the edge of a cube increases from 1 to twice that size, the

volume of the cube rises continuously from 1 to 8, hence must pass the value 2 at some time. As for the segment π , we can approximate it from below and from above with any degree of accuracy by the Euclideanly constructible perimeters of regular 6-, 12-, 24-, . . . sided polygons inscribed to and circumscribed about the circle. Thus we are turning the tables: *any arbitrarily given cut* in the domain of rational numbers, that is to say, any division of all rational numbers into three classes I, II, III, no matter in what way effected, determines a real number. (The only requirements to be satisfied are the following: neither I nor III must be empty; II contains at most one fraction; I contains no largest, III no smallest fraction; any number in I is smaller than any number in II or III; any number in III is greater than any number in I or II.) According to Dedekind (*Stetigkeit und Irrationalzahlen*, 1872), we have no reason to admit only part of these cuts as real numbers. And in geometry we then postulate (Dedekind's axiom) the existence of that segment which stands to the given unit segment in the ratio determined arithmetically by the cut. Since conversely, according to Eudoxus, the ratio of any segment a to the unit segment determines a cut, the axiom of Dedekind guarantees the *completeness* of the geometrical elements: the system of points is incapable of extension, provided all axioms (including that of Eudoxus) are maintained (Hilbert). This logical completeness (absence of gaps) reflects the intuitive continuity among the points in space. With Dedekind's number concept, analysis makes itself independent of geometry. Thereby, at last, it is in a position to analyze continuity and to provide geometry with the means of proving continuity theorems of the following kind: a continuous curve joining the center of a circle to a point outside the circle meets the circumference. In Euclid, the proofs of such theorems are incomplete, as was already pointed out by Leibniz with reference to the first construction occurring in Euclid, namely that of the equilateral triangle ABC from the points A and B ; Euclid fails to show that the circle about A through B and the circle about B through A have a point in common.

Another means of characterizing a real number, equivalent to that of the cut, is the infinite sequence of 'nested' rational intervals a_nb_n ($n = 1, 2, 3, \dots$), each of which lies within the preceding one, and the length $b_n - a_n$ of which converges to 0 as the index n increases indefinitely (compare the example of π). Since the fraction is logically no more complicated than the natural number — it is determined by two natural numbers, its numerator and denominator — we may sum up the result of the historical development of Problem (a) as follows:

The individual natural numbers form the subject of number theory, the possible sets (or the infinite sequences) of natural numbers are the subject of the theory of the continuum.

(b) The essential character of the continuum is clearly described in this fragment due to Anaxagoras: "Among the small there is no smallest, but always something smaller. For what is cannot cease to be no matter how far it is being subdivided." The continuum is not composed of discrete elements which are "separated from one another as though chopped off by a hatchet." Space is infinite not only in the sense that it never comes to an end; but at every place it is, so to speak, inwardly infinite, inasmuch as a point can only be fixed step-by-step by a process of subdivision which progresses *ad infinitum*. This is in contrast with the resting and complete existence that intuition ascribes to space. The 'open' character is communicated by the continuous space and the continuously graded qualities to the things of the external world. A real thing can never be given adequately, its 'inner horizon' is unfolded by an infinitely continued process of ever new and more exact experiences; it is, as emphasized by Husserl, a limiting idea in the Kantian sense. For this reason it is impossible to posit the real thing as existing, closed and complete in itself. The continuum problem thus drives one toward epistemological idealism. Leibniz, among others, testifies that it was the search for a way out of the "labyrinth of the continuum" which first suggested to him the conception of space and time as orders of the phenomena. "From the fact that a mathematical solid cannot be resolved into primal elements it follows immediately that it is nothing real but merely an ideal construct designating only a possibility of parts" (correspondence Leibniz-De Volder, Leibniz, *Philosophische Schriften*, II, p. 268).

[In contrast to this nature of the continuum, Leibniz conceives the idea of the *monads*, since — differently from Kant — he feels compelled to give the phenomena metaphysically a foundation in a world of absolute substances. "Within the ideal or the continuum the whole precedes the parts. . . . The parts are here only potential; among the real [i.e. substantial] things, however, the simple precedes the aggregates, and the parts are given actually and prior to the whole. These considerations dispel the difficulties regarding the continuum — difficulties which arise only when the continuum is looked upon as something real, which possesses real parts before any such division as we may devise, and when matter is regarded as a substance" (letter to Remond, *Philosophische Schriften*, III, p. 622).

The impossibility of conceiving the continuum as rigid being cannot be formulated more concisely than by Zeno's well-known paradox of

the race between Achilles and the tortoise. The remark that the successive partial sums $1 - \frac{1}{2^n}$ ($n = 1, 2, 3, \dots$) of the series

$$\frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \dots$$

do not increase beyond all bounds but converge to 1, by which one nowadays thinks to dispose of the paradox, is certainly relevant and elucidating. Yet, if the segment of length 1 really consists of infinitely many subsegments of lengths $\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \dots$, as of 'chopped-off' wholes, then it is incompatible with the character of the infinite as the 'incompletable' that Achilles should have been able to traverse them all. If one admits this possibility, then there is no reason why a machine should not be capable of completing an infinite sequence of distinct acts of decision within a finite amount of time; say, by supplying the first result after $\frac{1}{2}$ minute, the second after another $\frac{1}{4}$ minute, the third $\frac{1}{8}$ minute later than the second, etc. In this way it would be possible, provided the receptive power of the brain would function similarly, to achieve a traversal of all natural numbers and thereby a sure yes-or-no decision regarding any existential question about natural numbers!

Descartes struggles with the idea that the material corpuscles of a liquid in motion have to divide *in infinitum*, "or at least *in indefinitum*, and that into so many parts that it is impossible to imagine one, however small, of which one would not know that it was actually subdivided into still smaller parts." To him this remains a mystery, confronted with which he takes recourse to the incomprehensibility of the divine omnipotence. Euler, in his "Anleitung zur Naturlehre" (*Opera postuma*, II, 1862, pp. 449-560), which in magnificent clarity summarizes the foundations of the philosophy of nature of his time, declares that although the bodies are infinitely divisible the statement that every body consists of infinitely many ('ultimate') parts is entirely false and is even obviously incompatible with the infinite divisibility (Euler, *op. cit.*, Chap. II, §12). In the Kantian system, the first two antinomies of pure reason refer to the continuum.⁷]

Three attempts have been made in the history of thought to conceive of the continuum as Being in itself. According to the first

⁷ The first of these, however, is formulated misleadingly. According to the argument presented, it is not a question of whether the world does or does not have a temporal beginning, but whether the number of temporal moments up to the present time is finite or infinite. In a continuously filled time, the latter will be the case, no matter whether (by virtue of an intrinsic or extrinsic measuring principle) it be of finite or infinite length.

and most radical the continuum consists of countable discrete elements, atoms. With regard to *matter*, this path, initiated by Democritus in antiquity, has been followed with brilliant success in modern physics. Plato, clearly conscious of the goal of 'saving' the phenomenon by means of the idea, was the first to design a consistent atomism with respect to *space*. In Islamic philosophy the atomistic theory of space was renewed by the Mutakallimûn (see Lasswitz, *Geschichte der Atomistik*, I, 1890, pp. 139-150), and in the Occident by Giordano Bruno's doctrine of the minimum. Hume, too, in his space-time theory (*Treatise of Human Nature*, Book I, Part II, Section 4) transforms the vagueness of the sense data, at which he aims, into a composition out of indivisible elements. Stimulated by quantum theory the idea again arises today in discussions about the foundations of physics. But so far it has always remained mere speculation and has never achieved sufficient contact with reality. How should one understand the metric relations in space on the basis of this idea? If a square is built up of miniature tiles, then there are as many tiles along the diagonal as there are along the side; thus the diagonal should be equal in length to the side. Hume, consequently, is forced to admit that the "just as well as obvious" principle of comparing the measures of curves and surfaces by means of the number of component elements is, in fact, *useless*. B. Riemann, in his inaugural lecture *Über die Hypothesen, welche der Geometrie zugrunde liegen* (1854), states the alternative "that for a discrete manifold the principle of measurement is already contained in the concept of this manifold, but that for a continuous one it must come from elsewhere."

The second attempt is that of the infinitely small. This is discussed ingeniously and in detail on the first day of Galileo's "Discorsi." Just as I can bend a straight line segment into an octagon or a thousand-sided polygon, so, according to Galileo, I may also transform it into a polygon with infinitely many infinitely small sides by simply winding it around a circle, and thus do not have to rely on a limiting process which never reaches the goal.⁸

{If a wheel is rolled off along a horizontal line, then every one of the smaller concentric circles appears to be stretched out in the form

⁸ Hankel says (*Zur Geschichte der Mathematik im Altertum und Mittelalter*, Leipzig, 1874): "The idea of never reaching the area of the circle, no matter how far one might go in the sequence of polygons, although one approaches it arbitrarily closely, strains the power of imagination to such a degree that it will tend, at all cost, to bridge this gap extending, as it were, between reality and the ideal. Under this psychological pressure the — infinitely small or infinitely large? — step is taken that leads to the assertion: the circle is a polygon with infinitely many infinitely small sides. The Ancients, however, have refrained from this step; as

of a line h of equal length (*rota Aristotelis*). However, if the circular wheel is replaced by a many-sided regular polygon, then the 'covered' segments along h , into which the sides of the polygon fall successively, form a disrupted line. Thus, in the case of the circular wheel, one must assume that h consists of an infinitely dense succession of covered and uncovered segments. "This method," says Galileo in the "Discorsi" (*Opere* VIII, p. 93), "perhaps better than any other, enables us to avoid many intricate labyrinths such as are encountered in the question of cohesion in solids, mentioned before, and that of rarefaction and contraction, without forcing upon us the objectionable admission of empty spaces and thereby of the penetrability of bodies. We escape all these difficulties, so it seems to me, by assuming a composition out of indivisibles." If a curve consists of infinitely many straight 'line elements,' then a tangent can simply be conceived as indicating the direction of the individual line element; it joins two 'consecutive' points on the curve. However, he who rejects Galileo's hypothesis has to define the tangent at the point P of a curve as the limiting line approached indefinitely by the secant line PP' as the second moving point P' on the curve converges toward P . The discussion between Johann Bernoulli and Leibniz on this question is very instructive. Leibniz says (*Mathematische Schriften*, III, p. 536), "For if we suppose that there actually exist the segments on the line that are to be designated by $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, . . . , and that *all* members of this sequence actually exist, you conclude from this that an infinitely small member must also exist. In my opinion, however, the assumption implies nothing but the existence of any *finite* fraction of arbitrary smallness." But Bernoulli replies (*op. cit.*, p. 563). "If 10 members are present the 10^{th} necessarily exists, if 100 then necessarily the 100^{th} , . . . , if therefore their number is ∞ then the ∞^{th} [infinitesimal] member must exist." }

The limiting process was victorious. For the *limit* is an indispensable concept, whose importance is not affected by the acceptance or rejection of the infinitely small. But once the limit concept has been grasped, it is seen to render the infinitely small superfluous. Infinitesimal analysis proposes to draw conclusions by integration from the behavior in the infinitely small, which is governed by elementary laws, to the behavior in the large; for instance, from the universal law of attraction for two material 'volume elements' to the magnitude of attraction between two arbitrarily shaped bodies with homogeneous or non-homogeneous mass distribution. If the infinitely

long as there were Greek geometers, they have always halted in front of the precipice of the infinite. . . . "

small is not interpreted 'potentially' here, in the sense of the limiting process, then the one has nothing to do with the other, the processes in infinitesimal and in finite dimensions become independent of each other, the tie which binds them together is cut. Here Eudoxus undoubtedly saw right.

[Incidentally, as far as I can see, the 18th century remained far behind the Greeks with regard to the clarity of its conception of the infinitely small. More than one writer of this enlightened era complains of the 'incomprehensibilities of mathematics,' and vague and incomprehensible indeed is their notion of the infinitesimal. As a matter of fact, it is not impossible to build up a consistent 'non-Archimedean' theory of quantities⁹ in which the axiom of Eudoxus (usually named after Archimedes) does not hold. But as was just pointed out, such a theory fails to accomplish anything for analysis. Newton and Leibniz seemed to have the correct view, which they formulated more or less clearly, that the infinitesimal calculus is concerned with the approach to zero by a limiting process. But they lack the ultimate insight that the limiting process serves not only to determine the value of the limit but also to establish its existence. For that reason Leibniz is still quite unclear as to the summation of infinite series. Only slowly does the theory of limits gain a foothold. In 1784 D'Alembert declares emphatically in the *Encyclopédie*, "La théorie de la limite est la base de la vraie métaphysique du calcul différentiel. Il ne s'agit point, comme on le dit ordinairement, des quantités infiniment petites; il s'agit uniquement des limites des quantités finies." It was left to Cauchy, at the beginning of the 19th century, to carry these ideas out consistently. In particular he discovers the correct criterion for the convergence of infinite series, the condition under which a number is generated as limiting value through an infinite process. The proof of the criterion, however, requires that fixation of the number concept which was later accomplished by the principle of the Dedekind cut.]

The third attempt to 'save' the continuum in the Platonic sense may be seen in the modern set-theoretic foundations of analysis.

REFERENCES

R. DEDEKIND, *Stetigkeit und irrationale Zahlen*, Braunschweig, 1872; third edition 1905.

Concerning the ancient history of the problem of the continuum and the irrational cf.:

⁹ Compare, for instance, Hilbert, *Grundlagen der Geometrie*, Chapter II, §12. An example of infinitely small magnitudes, discussed already by Leibniz and Wallis, are the *anguli contactus* (between, say, a circle and its tangent) as opposed to the angles formed by straight lines.

- P. TANNERY, *Pour l'histoire de la science hellène*, Paris, 1887.
 E. FRANK, *Platon und die sogenannten Pythagoreer*, Halle, 1923.
 H. HASSE and H. SCHOLZ, *Die Grundlagenkrise der griechischen Mathematik*, Berlin, 1928.
 B. L. VAN DER WAERDEN, *Mathematische Annalen*, 117 (1940), pp. 141-161.
 K. VON FRITZ, *Annals of Mathematics*, 46 (1945), pp. 242-264.

8. SET THEORY

At a first glance it might seem as though with the limiting process the rigid *Being* is definitely resolved into *Becoming*; as though, thereby alone, Aristotle's doctrine is mathematically realized which taught that the infinite is forever being on the way and therefore exists only *δυνάμει* not *ἐνέργεια* (potentially, not actually). This appearance is deceptive. For the individual convergent sequence, such as the sequence of partial sums of the Leibniz series

$$\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots,$$

which converges to $\pi/4$, does not unfold itself according to a lawless process which we have to accept blindly in order to find out what it produces step by step; but it is fixed once and for all by a definite *law*, which correlates with every natural number n the corresponding approximate value (the n^{th} partial sum). A classification of the infinitely many rational numbers into the three classes I, II, III of a Dedekind cut is not made by taking one fraction after another and assigning it to its class, but rather according to a law, namely, by stating that all rational numbers with such and such a property are to belong to class I. (It suffices to define class I, since the other two classes are defined automatically along with it.) This law, or this property, fixes the intended real number exactly.

It is said that a function $f(x)$ is continuous at the place $x = a$ if $f(x)$ converges to $f(a)$ when the variable x approaches a . But how is this notion of convergence defined? "For *every* positive ϵ there should *exist* a positive δ of such a kind that, for *all* real numbers x which satisfy the condition $a - \delta < x < a + \delta$, we also have $f(a) - \epsilon < f(x) < f(a) + \epsilon$." Our attitude thus remains static. It is characterized by the unlimited application of the terms 'there exists' and 'all' not only to natural numbers but also to the places in the continuum, i.e. to the possible sequences or sets of natural numbers. This is the essence of set theory: It considers not only the sequence of numbers but also the totality of its subsets as a closed aggregate of objects existing in themselves. In this sense it is based on the actually infinite. But once this has been accepted, the vast structure of analysis has an unshakeable firmness; it is securely founded, in all its parts based on sound argument, exact in its concepts, without gaps

in its proofs. It has thus gained a foundation which guarantees the unconditional intersubjective agreement of all workers in its field.

[To be sure, considerable mathematical acumen was required thus to establish such general facts concerning continuity as are suggested by intuition; for instance, that a continuous function assumes all intermediate values, that a closed planar curve without multiple points divides its plane into two domains, or that a two-dimensional domain cannot be mapped continuously and in a one-to-one fashion into a three-dimensional domain. We experience again and again with our students what assiduous training is necessary in order to acquire that freedom from prejudice which is indispensable for a proper understanding of these proofs and their stringency. Besides such theorems confirming our intuition, analysis also reveals numerous occurrences which appear to run counter to it, such as continuous curves being everywhere without a tangent or filling out an entire square. It was the work of the 19th century from Cauchy and Gauss to Weierstrass to test all unproved suppositions of analysis on the above foundation.]

The set-theoretical method has permeated not only analysis but also the first beginning of mathematics, the theory of the natural numbers. From the point of view of set theory, the number sequence is a completed set N , within which a mapping $n \rightarrow n'$ is defined that uniquely correlates an element n' with every element n of the set. This very fact, the existence of a one-to-one mapping of N onto a subset of N that is not identical with the entire N (the correlations $n \rightarrow 2n$ or $n \rightarrow n^2$ have the same effect), shows N to be an *infinite* set. The finiteness of a set is established only when the impossibility of such a mapping has been demonstrated.

[Thus, for set theory, there is no difference in principle between the finite and the infinite. The infinite even appears to it as the simpler of the two (in agreement with Descartes, who maintained that the infinite is prior to the finite [letter to Clerselier, *Corr.*, ed. Adam and Tannery, V, p. 356, "Or je dis que la notion que j'ai de l'infini est en moi avant celle du fini"; also *Méditations métaphysiques*, third meditation, *Oeuvres de Descartes*. I, pp. 280–281]). The fact that, in the definite sense stated, Euclid's axiom of magnitude *καὶ τὸ ὅλον μέρους μείζον* ("the whole is greater than the part") fails to hold for an infinite set was pointed out already by Galileo (*Discorsi*, *Opere*, VIII, p. 79). From this, Leibniz concludes (letter to Bernoulli, *Math. Schriften*, III, p. 536) that "the number, or set, of all numbers entails

a contradiction if one conceives of it as a completed whole." Bolzano sees in it a "paradox of the infinite" (*Paradoxien des Unendlichen*, 1851, §20). Dedekind, finally, elevates this fact to the status of a definition of the infinite (*Was sind und was sollen die Zahlen?*, 1887). }

Following Dedekind, a set C of natural numbers is said to be a *chain* if, for every number x contained in C , its 'image' $x' = x + 1$ likewise belongs to C . The fact that every natural number can be reached by starting with 1, going on to its image $1' (= 2)$, obtaining $2' (= 3)$ by repeating the mapping, *and so on*, — the idea of this 'and so on,' that seems logically irreducible, but constitutes the essence of the natural number sequence, is then expressed in the form of the following principle: *Every chain which contains 1 as an element is identical with the whole of N .* Complete induction can therefore be based on the transfinite use of the concepts 'all' and 'there is'; in this way set-theory abolishes the partition between mathematics and logic. The investigations of Dedekind, Frege, and Russell aim at logicizing mathematics completely.

The question as to when a natural number n is less than a given number m , which common sense answers by the finite specifically arithmetical criterion: 'if the enumeration of the numbers from 1 to m leads to n before m is reached,' is decided in set theory by the following transfinite purely logical criterion: 'if there exists a chain containing m but not n .' But such a thing is possible only after one has climbed to that level of application of 'there exists' where this term refers to the *sets* of natural numbers.

And it is for this purpose alone that we require that objectification of sets which everyday language, strangely enough, has carried out all along. A proposition such as 'the rose is red' is no longer subordinated to the scheme ' x is red,' having one blank, x , but to the more general one ' x has the property X ,' from which the proposition results by the substitution $x = \text{rose}$, $X = \text{red}$. The words 'has the property' denote a certain relation ϵ , which may hold between the arbitrary object x and the arbitrary property X . Only in this connection do we encounter the *copula* ϵ ; it changes the originally bipartite proposition into a tripartite one, $x \epsilon X$. (The grotesque confusion of the copula with existence and with equality is one of the saddest indications of the dependence of philosophical speculation on accidental linguistic forms.) The way is now open for a formal application of the definitional principles 6. and 7. of Section 1 to the blank X . The introduction of the general set concept thus consists of two essentially different steps; the first is the objectification just described, the second is the stipulation that two properties X and Y , or the corresponding sets,

be considered equal if all elements of X also belong to Y , and vice versa.

From an aggregate of individually exhibited objects we may by selection produce all possible subsets and thus make a survey of them one after another. But when one deals with an infinite set like N , then the existential absolutism for the subsets becomes still more objectionable than for the elements. Since one can lay hands on such subsets only as are determined by a characteristic property of their elements, it is difficult to rid oneself of the feeling that a chaotic abundance of possibilities, of sets put together haphazardly and without rule or law, goes by the board. But the paradoxical character of the elusive 'aggregate of all possible properties of natural numbers' can be laid bare even more precisely. Suppose we had somehow succeeded in the demarcation of an 'extensionally definite' aggregate of such properties (I shall call them properties of the first level), so that we have the right to believe that the question 'is there a property of the first level of such and such a well-described kind A ?' is answered by the facts with a clear-cut yes or no. We may then speak of the property P_A which applies to a number x if and only if there *exists* any property at all of the first level which appertains to x . This property P_A , however, according to its meaning certainly lies outside the circle of properties of the first level; it belongs to a higher, the second level, since it has been defined in terms of the totality of properties of the first level. Russell formulates this insight somewhat vaguely by his "vicious circle principle": "No totality can contain members defined in terms of itself." Similarly, the third level is constructed above the second, and so forth. Correspondingly, sets of natural numbers — and hence real numbers — of the first, second, third, . . . levels should be distinguished. The mode of construction of the property P_A occurs in analysis, for instance, in determining the least upper bound of a point set on a line. The obliteration by the existential absolutism of these differences in level, which were first brought out in Russell's theory of types, constitutes an unquestionable vicious circle.

{One could escape this dilemma only if, for every property of the second level, there existed a property of the first level equal to it (not in meaning but) in extension. As long as the sequence of natural numbers is accepted as an extensionally definite aggregate, one might consider as the properties of the first level those which are generated by the definitional principles of Section 1 from the one basic relation ' n follows upon m .' In this case, our wish will hardly be fulfilled. We would have the task of extending the principles of construction for the properties of the first level in such a manner that every set of the

second level demonstrably coincides with one of the first. But there is not the slightest indication that this is possible. Russell, in order to extricate himself from the affair, causes reason to commit hara-kiri, by postulating the above assertion in spite of its lack of support by any evidence ('axiom of reducibility'). In a little book *Das Kontinuum*, published in 1918, I have tried to draw the honest consequence and constructed a field of real numbers of the first level, within which the most important operations of analysis can be carried out.]

In spite of its paradoxical character, the idea of absolute existence in the domain of natural numbers and sets of natural numbers has so far not yet led to any contradiction. G. Cantor, however, freed himself of all fetters and manipulated the set concept without any restriction, in particular permitting the formation of the set of all subsets of any given set. He developed a general theory of cardinal and ordinal numbers of infinite sets. Here, at the farthest frontiers of set theory, actual contradictions did show up. But their root can only be seen in the boldness perpetrated from beginning in mathematics, namely, of treating a field of constructive possibilities as a closed aggregate of objects existing in themselves.

REFERENCES

- B. BOLZANO, *Paradozien des Unendlichen*, posthumous edition Pfihonsky, Leipzig 1851.
 R. DEDEKIND, *Was sind und was sollen die Zahlen?*
 G. FREGE, *Die Grundlagen der Arithmetik*.
 G. CANTOR, *Gesammelte Abhandlungen*, Berlin, 1932; in particular Sections III and IV.
 B. RUSSELL, *The Principles of Mathematics*.
 H. WEYL, *Das Kontinuum*, Berlin, 1918.
 A. FRAENKEL, *Einleitung in die Mengenlehre*, 3d. ed., Berlin, 1928.

9. INTUITIVE MATHEMATICS

This situation was first clearly recognized by L. E. J. Brouwer (since 1907). He designed a system of mathematics which does not make that leap into the beyond of which we spoke at the end of Section 6. An existential statement, such as 'there exists an even number,' is not considered a proposition in the proper sense that asserts a fact. An 'infinite logical summation' such as is called for by a statement of this kind (1 is even or 2 is even or 3 is even or . . . *ad infinitum*) is evidently incapable of execution. '2 is an even number,' this is a real proposition (provided 'even' has been defined recursively as on p. 33); 'there exists an even number' is nothing but a *propositional abstract* derived from that proposition. If I consider an insight a valuable treasure, then the propositional abstract is merely a document

indicating the presence of a treasure without disclosing its location. Its only value may lie in the fact that it causes me to look for the treasure. It is a worthless piece of paper as long as it is not endorsed by a real proposition such as '2 is an even number.' Whenever nothing but the *possibility* of a construction is being asserted, we have no meaningful proposition; only by virtue of an effective construction, an executed proof, does an existential statement acquire meaning. In any of the numerous existential theorems in mathematics, what is valuable in each case is not the theorem as such but the construction carried out in its proof; without it the theorem is an empty shadow.

[The question, put in Section 3, as to how conclusions may be drawn from existential statements, must here be answered by denying that possibility in principle. It can be done only after the existential statement has been replaced by the meaningful whole from which it was isolated as a propositional abstract. All proofs that depend on the construction of auxiliary elements fall under this remark. On the other hand, how do we obtain universal theorems on natural numbers? In order to explain this by means of a very simple example, let the number-theoretical function $\varphi(n)$ be defined by complete induction as follows:

$$(\alpha) \varphi(1) = 1; \quad (\beta) \varphi(n') = (\varphi(n))'.$$

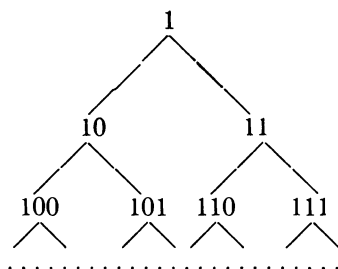
Here, (β) represents a universal proposition, from which, in connection with (α) , we may infer by complete induction that generally $\varphi(n) = n$. Thus the definition itself is seen to be the root of universality, and from there it spreads by complete induction. The principle of complete induction (as an instrument of definition or inference), not pressed into a formula but concretely applied at every step, is the true and only power of mathematics, the mathematical prime intuition. In this point Brouwer is in agreement with Poincaré ("*Science et hypothèse*"). The negation of a universal proposition about numbers would be an existential proposition; since this is void in itself, universal propositions are incapable of negation. Even a universal statement does not refer to a fact, it is not to be interpreted as the logical product of infinitely many singular propositions but as a *hypothetical* statement: if applied to a single definite given number it yields a definite proposition. There is no occasion here for the application of a principle of *tertium non datur* (either all numbers have the property A , or else there exists a number with the property $\sim A$). The belief in it, according to Brouwer (*Jahresberichte der Deutschen Mathematiker-Vereinigung*, 28, 1920) "was caused historically by the fact that, firstly, classical logic was abstracted from the mathematics of the

subsets of a definite finite set [i.e. a set given by exhibition of its elements], that, secondly, an *a priori* existence independent of mathematics was ascribed to this logic, and that, finally, on the basis of this supposititious apriority it was unjustifiably applied to the mathematics of infinite sets." }]

In Brouwer's analysis, the individual place in the *continuum*, the real number, is to be defined not by a set but by a *sequence* of natural numbers, namely, by a law which correlates with every natural number n a natural number $\varphi(n)$. (The two definitions cease to be equivalent, as soon as the natural numbers may no longer be treated as an extensionally definite aggregate.) How then do assertions arise which concern, not all natural, but all real numbers, i.e. all values of a real variable? Brouwer shows that frequently statements of this form in traditional analysis, when correctly interpreted, simply concern the totality of natural numbers. In cases where they do not, the notion of sequence changes its meaning: it no longer signifies a sequence determined by some law or other, but rather one that is created *step by step by free acts of choice*, and thus necessarily remains in *statu nascendi*. This 'becoming' *selective sequence* (*werdende Wahlfolge*) represents the continuum, or the variable, while the sequence determined *ad infinitum* by a law represents the individual real number falling into the continuum. The continuum no longer appears, to use Leibniz's language, as an aggregate of fixed elements but as a medium of free 'becoming.' Of a selective sequence *in statu nascendi*, naturally only those properties can be meaningfully asserted which already admit of a yes-or-no decision (as to whether or not the property applies to the sequence) when the sequence has been carried to a certain point; while the continuation of the sequence beyond this point, no matter how it turns out, is incapable of overthrowing that decision.

[In accordance with intuition, Brouwer sees the essential character of the continuum, not in the relation between element and set, but in that between part and whole. The continuum falls under the notion of the 'extensive whole,' which Husserl characterizes as that "which permits a dismemberment of such a kind that the pieces are by their very nature of the same lowest species as is determined by the undivided whole" (*Logische Untersuchungen*, second edition, II, p. 267). The division scheme of the one-dimensional continuum is best illustrated by the example of a finite line segment. By halving it, one decomposes it into two parts, a left (10) and a right one (11); each of the latter, by again halving them, decomposes into a left and right one, 100, 101 and 110, 111 respectively, and so on. This process may

be described purely combinatorially and thus furnishes the arithmetical blank-form of the open one-dimensional continuum. This must



be distinguished from the realization of the process for a concretely exhibited continuum, such as the segment in space. In carrying out its continued subdivision according to the arithmetical scheme it is clearly irrelevant whether the two parts are always of the same length, as long as only the fineness of the parts eventually drops below any possible threshold of exactness. (It may even be that comparison of length has no foundation in the nature of the given continuum.) The process of subdivision, which *in concreto* can never have been carried out except to a certain point, determines a coordinate system within the continuum and thus makes it possible to designate the individual points in arithmetical terms by binary fractions. Since in a concrete continuum no exact boundaries can be set, one must imagine that the division framework is at no stage of the procedure fixed with complete accuracy, but that, as the subdivision continues, the earlier points of division steadily increase in precision. Any two adjacent parts of the i^{th} division step may be joined into a 'division interval of the i^{th} level.' The division intervals of the i^{th} level overlap in such a manner that for any approximately given number, as soon as the approximation is sufficiently accurate, a division interval of the i^{th} level can be found into which that number falls. Thus the individual real number will have to be defined as an *infinite sequence of nested division intervals of increasing level*.

Two real numbers α , β coincide if, for every value of n , the n^{th} interval of the sequence α and the n^{th} interval of the sequence β partially or wholly overlap; they are distinct if a number n exists for which these two intervals are disjoint. Because of the inapplicability of the *tertium non datur* to statements of this kind, Brouwer does not recognize this as a clear-cut alternative. This view fits in well with the character of the intuitive continuum. For there the separateness of two places, upon moving them toward each other, slowly and in vague gradations passes over into indiscernibility. In a continuum,

according to Brouwer, there can be only continuous functions. *The continuum is not composed of parts.* Thus I may well distinguish within the continuum the sub-continuum of the positive numbers by using only positive binary fractions in the formation of intervals and interval sequences; but it is not true that the entire continuum is composed of the continuum of the positive numbers, that of the negative numbers, and that of the numbers coinciding with zero, in the sense that every number must belong to one of these three continua. An old truth thus finds a precise mathematical formulation; one that Aristotle (*περὶ ἀτόμων γραμμῶν*) expressed by saying, "That which moves does not move by counting," or (*Physics*, Bk. VIII, Ch. 8), "If the continuous line is divided into two halves, the one dividing point is taken for two; it is both beginning and end. But as one divides in this manner, neither the line nor the motion are any longer continuous . . . In the continuous there is indeed an unlimited number of halves, but only in possibility, not in reality." Compare in this connection the passages quoted earlier from Leibniz's letters. The principle comes into its own again that "nothing is separable which is not already separate" (Gassendi).

Mathematics with Brouwer gains its highest intuitive clarity. He succeeds in developing the beginnings of analysis in a natural manner, all the time preserving the contact with intuition much more closely than had been done before. It cannot be denied, however, that in advancing to higher and more general theories the inapplicability of the simple laws of classical logic eventually results in an almost unbearable awkwardness. And the mathematician watches with pain the larger part of his towering edifice which he believed to be built of concrete blocks dissolve into mist before his eyes.]

REFERENCES

- L. E. J. BROUWER, *Over de grondslagen der wiskunde*, Dissertation, Amsterdam and Leipzig, 1907.
 — *Intuitionisme en Formalisme*, Groningen, 1912 (English translation in *Bull. Am. Math. Soc.*, 20 (1913-14));
 — Zur Begründung der intuitionistischen Mathematik, *Mathematische Annalen*, 93, 95, 96 (1924-27).
 H. WEYL, Über die neue Grundlagenkrise der Mathematik, *Mathematische Zeitschrift*, 10 (1921).
 O. BECKER, Beiträge zur phänomenologischen Begründung der Geometrie und ihrer physikalischen Anwendungen, *Husserls Jahrbuch für Philosophie*, 6; especially pp. 398-436, 459-477.

10. SYMBOLIC MATHEMATICS

Is there no way to escape such radical consequences? The resolution to make this sacrifice is doubly hard in view of the historical fact that in set-theoretical analysis we find, in spite of the boldest and most

elaborate combinations, complete certainty of deduction and an obvious accord among all the results. Hilbert set for himself the goal of saving mathematics in its entirety through the axiomatic method. He, too, admittedly is convinced that the power of intuitive thought does not reach farther than was asserted by Brouwer, that it is incapable of supporting the transfinite modes of deduction in mathematics, and that none of the transfinite statements of mathematics can be justified as being evident material truths (*einsichtige inhaltliche Wahrheiten*). What Hilbert proposes to secure is not the truth but the *consistency* of traditional analysis.

For this purpose he has to formalize mathematics, including logic, so that it becomes a game with symbols played according to fixed rules. (The symbols are not meant to be symbols *for* something.) The mathematical formulas which are made up of these symbols do not throughout admit of a material interpretation. Along with the meaningful propositions, 'ideal propositions' had to be introduced in order to reestablish artificially the validity of the simple logical rules that, as Brouwer had shown us, were lost in the transition to the infinite — just as in algebraic number theory ideal numbers were introduced in order to enforce the validity of the simple divisibility theorems. There are four different kinds of symbols,¹⁰ which are distinguished, like the pawns and knights in chess, by the different rules of the game that apply to them: *constants* (such as 1), *variables* (symbols for blanks, x, y, \dots), *one-place and many-place operations*, and *integrations*. The most important one-place operations are \sim (negation), σ (transition from a natural number to the next following one), and N (Na , to be read: a is a natural number). The most important two-place operations are \rightarrow , $=$, and ϵ . We construe all these as operations; in particular, N is the operation which, when applied to a , produces the proposition: a is a number; $=$ is the operation which, when applied to a and b , produces the proposition: a equals b . In order to arrive at a convenient general formulation of the rules of the game, these operational symbols may consistently be written

in front of the terms (formulas) to which they apply, e.g. $\epsilon \begin{matrix} a \\ \swarrow \\ b \end{matrix}$ instead

of $a \epsilon b$. Among the integrations (which are always followed by one formula only) we have, above all, the quantifiers Σ_x , Π_x and the symbol ϵ_x to be introduced presently; they carry one (or several) arbitrary variables as subscripts. A prefixed integration symbol with the

¹⁰ Deviating somewhat from Hilbert's original version, I here follow von Neumann's simplified formalism (*Zur Hilbertschen Beweistheorie*, *Mathematische Zeitschrift*, 1926).

subscript x has the effect that the variable x becomes 'bound' at all places in the formula following the symbol, thus losing its capacity of being substituted for. In the course of the development of mathematics it is always possible to introduce new symbols. What a *formula* is, is defined recursively: "(α) every constant or variable by itself is a formula; (β) from one or two (or several) formulas already constructed a new formula is obtained by writing down respectively a one- or two- (or several-) place symbol o of operation or a (one-place) symbol of integration, and having it followed by the formula(s) in question in their proper order, each written on a separate line and its initial symbol joined to o by a dash." The complete formula then looks like a (parthenogenetic) genealogical tree of symbols, from which the "grammatical structure" of the formula, i.e. the manner of its recursive construction can be read off unambiguously. One also can decide in this way whether a given tree-like arrangement of symbols is or is not a formula.

[The linear arrangement, which is more convenient to print, has to make use of parentheses if the recursive construction is to remain uniquely recognizable. We return to the usual symbolism, which is less systematic, whenever it is a question of merely outlining the procedure in its essential features.

It is unnecessary to worry about the fact that in the formal construction the operations are applied indiscriminately to all kinds of things. Who is afraid of such generosity may prefer to discriminate between 'numerical' and 'factual' formulas, in accordance with the following recursive stipulations: "(α) A constant or variable by itself, as well as any formula beginning with σ or ϵ_x , is a *numerical* formula; formulas beginning with \sim , \rightarrow , \vee , $\&$, N , $=$, ϵ , Σ_x , Π_x , on the other hand, are *factual*. (β) The symbols σ and N must be followed by one, ϵ by two *numerical* formulas, while \sim , ϵ_x , Σ_x , Π_x must be followed by one, \rightarrow , $\&$, \vee by two *factual* formulas." Similar restrictions will then have to accompany the axiomatic rules and the syllogistic rule of inference below.

If $A(x)$ (as always in what follows) is an arbitrary formula containing only the one 'free' variable x (free in the sense that it is not bound at every place where it occurs), and if b is a 'closed formula' (i.e. one containing no free variables), then b may be substituted for x in A wherever x occurs free (i.e. is not bound). The result of this process of *substitution*, which thus has been described intuitively, is again a formula; it is denoted by the abbreviating sign $A(b)$.¹¹

¹¹ Here the letters A , b are clearly not symbols of the game, but are used as signs of communication that enable us to speak of formulas etc. in general. Hil-

Formulas serving as *axioms* form the starting point of any proof. Instead of individual axioms, however, we formulate general rules for the formation of axioms. First come the axiomatic rules of finite logic, such as

$$c \rightarrow (b \rightarrow c).$$

It says: take any two formulas b and c without free variables and construct out of them the formula $c \rightarrow (b \rightarrow c)$; the result you may use as an axiom. Secondly, there are the two axiom rules of equality; they establish the connection between logic and arithmetic:

$$\begin{aligned} b &= b. \\ (b = c) &\rightarrow (A(b) \rightarrow A(c)). \end{aligned}$$

Thirdly, we have specifically arithmetical rules of a finite character. In them the constant 1 appears, which is the material starting point of all construction:

$$\begin{aligned} N1. \\ Nb &\rightarrow N(\sigma b). \\ (\sigma b = \sigma c) &\rightarrow (b = c). \\ \sim (\sigma b = 1). &\} \end{aligned}$$

Next we come to the *transfinite* part. Taking for granted the alternative, denied by Brouwer, that either an honest man exists or all men are dishonest, one is sure to find an Aristides of whom it can be said: if any man be honest then Aristides is. For, in the first case, we may choose as Aristides one of the honest, and, in the second case, any man at all. In order to be able to construct such an Aristides, not just for the property of honesty, but for every property, i.e. for every formula A containing one free variable x , we invent a fictitious divine automaton which produces, whenever an arbitrary property A is fed into it, that individual $\epsilon_x A$ which certainly possesses the property A provided such an individual exists at all. ϵ_x is an integration symbol. (Indulgently following the fatal custom of employing the word 'is' to denote both the copula and existence we too use the same letter ϵ for both; but the confusion is avoided by the variable attached as subscript to the existential ϵ .) If such an automaton were at our disposal, then we would be rid of all the trouble caused by 'some' and 'all.' But the belief in its existence is, of course, sheer nonsense. Mathematics, however, proceeds as if it existed. This can be expressed in the form of an axiom rule, and if the application of this rule does not lead to

bert employs the Gothic alphabet to distinguish them from the symbols proper. Because of the aversion of the English-speaking reader to Gothic type, this practice has not been followed in our translation, although it is undoubtedly a valuable help in keeping the issue clear.

contradictions, then its use is legitimate in formalized mathematics. Thus we have the following transfinite logical axiom rules:

$$\begin{aligned} A(b) \rightarrow \Sigma_x A(x); \quad \Pi_x A(x) \rightarrow A(b); \\ \Sigma_x A(x) \rightarrow A(\epsilon_x A); \quad A(\epsilon_x (\sim A)) \rightarrow \Pi_x A(x). \end{aligned}$$

Those stated in the second line were still omitted in Section 3; they permit us to infer something from Σ_x and to infer Π_x from other formulas. Of course, they do not offer the same service as the fictitious automaton; for, given a formula A , they fail to reveal the identity of $\epsilon_x A$. Only in special circumstances may a formula such as $\epsilon_x A = 1$ appear as the terminal formula of a proof starting with the axioms.

[Among the arithmetical axioms, the principle of complete induction is still absent. It may be interpreted as a transfinite arithmetical axiom rule, expressing the fact that a property appertaining to 1 and 'handed on' from x to σx is a property of every arbitrary number. But, as we know, this rule becomes superfluous if it is admissible to introduce for every property A a new object y , namely, the corresponding set, such that the proposition ' x is an element of y ' is equivalent with the subsistence of $A(x)$. If this hypothesis is formulated as an axiom rule, it turns out that its application leads inescapably to a contradiction — a fact tantamount to a *forfeiture of the unlimited right of objectification*. For the purposes of analysis, however, it is sufficient to restrict the argument x to the range of natural numbers, so that we may lay down the following narrower transfinite set-theoretical rule:

$$(I) \quad \Sigma_y \Pi_x \{Nx \rightarrow ((x \in y) \Leftrightarrow A(x))\},$$

where $B \Leftrightarrow C$ serves as an abbreviation of $(B \rightarrow C) \& (C \rightarrow B)$. It seems to be desirable, though not indispensable, for the construction of analysis to add the *axiom of definiteness*, according to which two sets of numbers are equal if they contain the same elements:

$$\Pi_x \{Nx \rightarrow ((x \in b) \Leftrightarrow (x \in c))\} \rightarrow (b = c).]$$

A mathematical *proof* consists in manufacturing axioms by means of the given rules — these axioms never contain free variables — and in progressing to ever new formulas by applying the syllogistic rule of inference to such axioms or to formulas already obtained. We repeat the rule (cf. Sect. 3): Given two formulas b and $b \rightarrow c$ in the second of which the first reappears at the left of \rightarrow one may pass on to the formula c . To survey in advance what *demonstrable formulas* will be obtained as the result of this game is impossible, mainly

because the syllogism leads from two formulas b and $b \rightarrow c$ to the new formula c , which is shorter than the second of the premisses, so that in the proof game shrinkage interchanges with growth.

Up to this point all is game, not knowledge. But now the game is made the subject of investigation in what Hilbert calls *metamathematics*, the aim being to make certain that the game will never lead to a contradiction. Such a contradiction would arise if the actual play of two proof games would terminate, the one with a formula b , the other with the opposite formula $\sim b$. Only in order to arrive at this one insight does Hilbert require the finite, material, meaningful mode of thought, which cannot be pressed into any 'axioms.' In particular, this material thinking makes use of an intuitive inference by complete induction, such as we drew when we came to the conclusion (Section 4) that a correctly played game of chess can never produce 10 queens of the same color.

[One of the rules of the elementary propositional calculus that either figures among the axiomatic rules or is readily deduced from them is

$$(1) \qquad \qquad \qquad \sim b \rightarrow (b \rightarrow c),$$

where b and c are any closed formulas. Let c be an arbitrary formula of this kind, and suppose that a certain formula b and its negation $\sim b$ have been demonstrated. Under these circumstances, two syllogistic steps lead from (1) first to $b \rightarrow c$ and then to c . Hence in case the formalism is known to be inconsistent, *any* closed formula c may be demonstrated, and thus the proof game loses all interest. Consistency may also be defined by saying that the formula $\sim (1 = 1)$ is not demonstrable.]

The axiom system may be continually expanded, but it must be shown that the consistency is not overthrown by the expansion. In particular, definitions may be introduced in the form of new axiom rules; e.g.

$$\sigma 1 = 2, \quad \sigma(\sigma b) = \sigma_2 b.$$

This applies especially to the recursive definitions of $b + c$, $b \cdot c$ and other arithmetical operations. It can be shown once and for all that consistency, if it prevailed before, is preserved by the addition of axioms of this kind, that stand for simple or recursive definitions.¹²

¹² It can also be shown that once the definitional axioms for $b + c$, $b \cdot c$ and the corresponding operational symbols $+$, \cdot have been introduced, all other recursively definable arithmetical operations are expressible in the formalism. Compare e.g. Hilbert and Bernays, *Grundlagen der Mathematik*, vol. I, pp. 412-422.

Regarding the natural numbers, Hilbert's construction, in contrast to Brouwer's, gets along without that 'possibility *ad infinitum*' which was described in Section 6 as the third step of constructive cognition. For Hilbert, 10^{12} is a transfinite symbol, which does not denote a number of the form $\sigma\sigma \dots \sigma 1$. Geometry and physics may be adjoined, as soon and insofar as they have been strictly axiomatized. Hilbert even believes (*Axiomatisches Denken*, 1917), "Every potential subject of scientific thought, as soon as it is ripe for the formation of a theory, is bound to fall under the axiomatic method and, therefore, indirectly to the lot of mathematics."¹³

[As long as the transfinite components are left out of consideration, the consistency proof can easily be carried out by means of a 'valuation' of formulas. By a precisely described recursive procedure, every formula, according to its origin, is ascribed one of the values *T* or *F* (true or false) in such a manner that the finite axioms obviously get the value *T* and that the rules of evaluation given in Section 3 hold for the logical combinations. Hence, as long as the transfinite is excluded, the syllogism and thus the deductive method, remains impotent; for a decision as to the truth or falsehood of the premiss $b \rightarrow c$ is made only after the conclusion *c* has been evaluated.

The consistency proof can no longer be carried out along those lines if the transfinite axiom rules are taken into consideration. This brings out the fact that, with them, the insight into true and false ceases. After Hilbert and P. Bernays had developed more indirect methods, W. Ackermann and J. von Neumann in 1926 seemed to have succeeded in establishing the consistency of 'arithmetic,' i.e. of an axiomatic system including the transfinite logical axioms and the principle of complete induction, excluding however the dangerous axiom (I) about the conversion of predicates into sets. This result would vindicate the standpoint taken by the author in *Das Kontinuum*, that one may safely treat the sequence of natural numbers as a closed aggregate of existing objects. Justification of the same standpoint with respect to the 'aggregate of all possible sets of natural numbers' would depend on extending the consistency proof to the set-theoretical axiom rule (I); at the moment we do not see how that could be done.

¹³ Out of an entirely different conception of mathematics, Kant (*Metaphysische Anfangsgründe der Naturwissenschaft*, Preface) comes to the conclusion "that in every specific natural science there can be found only so much science proper as there is mathematics present in it." In the same sense as Hilbert, on the other hand, Husserl (*Logische Untersuchungen*, I, §71) declares with particular reference to mathematical logic that "the mathematical form of treatment . . . is for all strictly developed theories (this word taken in its true sense) the only scientific one, the only one that affords systematic completeness and perfection and gives insight into all possible questions and their possible forms of solution."

Even in the consistency proof for arithmetic just referred to a serious gap was later discovered. Concerning this development after 1926 and the catastrophe precipitated by an important discovery by K. Gödel in 1931, see Appendix A. But whatever the ultimate value of Hilbert's program, his bold enterprise can claim one merit: it has disclosed to us the highly complicated and ticklish logical structure of mathematics, its maze of back-connections, which result in circles of which it cannot be gathered at a first glance whether they might not lead to blatant contradictions.

The described symbolism evidently attacks again, in a refined form, the task which Leibniz had set himself with his "general characteristic" and *ars combinatoria*. But is it really more than a bloodless ghost of the old analysis that confronts us here? Hilbert's mathematics may be a pretty game with formulas, more amusing even than chess; but what bearing does it have on cognition, since its formulas admittedly have no material meaning by virtue of which they could express intuitive truths? The subject of mathematical investigation, according to Hilbert, is the concrete symbols themselves. It is without irony, therefore, when Brouwer says (*Intuitionisme en formalisme*, p. 7), "Op de vraag, waar de wiskundige exactheid dan wel bestaat, antwoorden beide partijen verschillend; de intuitionist zegt: in het menselijk intellect, de formalist: op het papier." The question why he sets up just these rules must remain unanswered by the consistent formalist. He will have to refer us to philosophy, psychology, or anthropology, so Brouwer thinks, in order to justify his "lustgevoel van echtheitsovertuiging" and his belief that the chosen axiom system is more suitable than any other to be projected onto the world of experience. }

This last remark reminds us that it is the function of mathematics to be at the service of the natural sciences. The propositions of theoretical physics, however, certainly lack that feature which Brouwer demands of the propositions of mathematics, namely, that each should carry within itself its own intuitively comprehensible meaning. Rather, what is tested by confronting theoretical physics with experience is the system as a whole. It seems that we have to differentiate carefully between phenomenal knowledge or insight — such as is expressed in the statement: 'This leaf (given to me in a present act of perception) has this green color (given to me in that same perception)' — and theoretical construction. Knowledge furnishes truth, its organ is 'seeing' in the widest sense. Though subject to error, it is essentially definitive and unalterable. Theoretical construction seems to be bound only to one strictly formulable rational principle, that of

concordance (compare Section 17, p. 121), which in mathematics, where the domain of sense data remains untouched, reduces to consistency; its organ is creative imagination. In connection with physics we shall have to discuss in greater detail the question what its determining factors, besides concordance, are. Intuitive truth, though not the ultimate criterion, will certainly not be irrelevant here. Hilbert himself expresses the following opinion (*Über das Unendliche, Mathematische Annalen*, 95, p. 190), "The function left to the infinite . . . is merely that of an idea — if, with Kant, one understands by an idea a concept of reason (*Vernunftbegriff*) transcending all experience and supplementing the concrete in the sense of totality." But perhaps this question can be answered only by pointing toward the essentially historical nature of that life of the mind of which my own existence is an integral but not autonomous part. It is light and darkness, contingency and necessity, bondage and freedom, and it cannot be expected that a symbolic construction of the world in some final form can ever be detached from it.

R E F E R E N C E S

Concerning LEIBNIZ, see references at the end of Section 3.
D. HILBERT and P. BERNAYS, *Grundlagen der Mathematik*, 2 Vols., Berlin, 1934-39.

11. ON THE CHARACTER
OF MATHEMATICAL COGNITION

[From time immemorial mathematics has been looked upon as the science of quantity, or of space and number. (Though we also find this definition with Leibniz, the *mathesis* thus delineated is to him but a part of the more comprehensive *ars combinatoria*.) Today this view appears much too narrow in consideration of such fields as projective geometry or group theory. Consequently we need not worry particularly over an exact determination of what is meant by quantitative. In fact, the development of mathematics itself raises doubts as to whether quantity is a well-determined and philosophically important category. Geometry, inasmuch as it is concerned with real space, is no longer considered a part of pure mathematics; like mechanics and physics, it belongs among the applications of mathematics. Under the influence of the general arithmetic of hypercomplex numbers and later of the axiomatic investigations, of set theory and symbolic logic, the distinction between mathematics and logic is gradually obliterated. "Mathematics is the science which draws necessary conclusions," B. Peirce declares in 1870. The definition of 'mathematics or logic' is discussed in detail in Chapter XI of Husserl's

Logische Untersuchungen (Vol. I, *Die Idee der reinen Logik*) and in the last chapter of Russell's *Introduction to Mathematical Philosophy*.

The crisis brought on by the set-theoretical antinomies — no matter if one follows Brouwer's radical intuitionism or Hilbert's symbolism — again throws into sharper relief the peculiar character of mathematics. Like Plato, Brouwer looks upon the two-oneness as the root of mathematical thinking. "Dit neo-intuitionisme zieht het uiteenvallen van levensmomenten in kwalitatief verschillende deelen, die alleen gescheiden door den tijd zich weer kunnen vereenigen, als oergebeuren in het menschelijk intellect, en het abstraheeren van dit uiteenvallen van elken gevoelsinhoud tot de intuïtie van twee-eenigheid zonder meer, als oergebeuren van het wiskundig denken." We have seen how the division scheme of the one-dimensional continuum results from "one becoming two"¹⁴ again and again (compare the diagram on p. 53). The integers when written in the binary system are obtained in the same manner. Stenzel (*Zahl und Gestalt bei Plato und Aristoteles*, 1924) makes it appear probable that Plato thought of his numbers as being arranged according to this scheme; but since the splitting of one into two here leads to larger and larger numbers, while in the continuum we descend to smaller and smaller parts, he refers to that two-ness as the "great-and-small." (See, however, for a different interpretation: H. Cherniss, *The Riddle of the Early Academy*, Univ. of Calif. Press, 1945.) More appropriate for the integers is their natural order, which Aristotle (*Metaphysics* A6 and M6) sets up in opposition to Plato's number concept. But it, too, can be generated out of the two-oneness; starting with an undivided whole, we separate it into an element (the 1), to be preserved as a unit, and an undivided remainder, the latter we then separate again into an element (2) and an undivided remainder, and so forth. (This can be visualized as the continued chopping-off of a segment from a half-line; time is open toward the future, but whenever we stop we find that another segment of time has been lived through.) In this scheme, not every part but only the last remaining part is subject to further bipartition.}]

Independently of the value attached to this last reduction of the mathematical thought process to the two-oneness, *complete induction* appears, from the intuitionist point of view, as that which prevents mathematics from becoming one huge tautology and which confers upon its assertions a synthetic non-analytic character. The pro-

¹⁴ Allusion to the phrase "Da wurde Eins zu Zwei" by which Nietzsche described his Zarathustra experience in several of his poems; e.g. in "Sils Maria":

"Da, plötzlich, Freundin, wurde Eins zu Zwei —
— und Zarathustra ging an mir vorbei . . ."

cedure of complete induction is, indeed, a decisive feature throughout. If at first it does not appear to play any part in elementary geometry (especially in elementary projective geometry), the reason is to be seen in the naive application of 'some' and 'all' to the points. In the intuitionist view, this is inadmissible; the field of construction of geometry is a continuum and hence capable of exact mathematical treatment only after it has been spun over with a division net as described above (compare also Section 15).

From the formalist standpoint, the transfinite component of the axioms takes the place of complete induction and imprints its stamp upon mathematics. The latter does not consist here of evident truths but is *bold theoretical construction*, and as such the very opposite of analytical self-evidence. The material reasoning of metamathematics, on the other hand, in running over the steps of a proof, operates by means of an intuitive inference from n to $n + 1$ and concerns itself with "extra-logical, concrete objects, which can be overlooked completely in all their parts and whose exhibition, differentiation, and succession or coordination are intuitively given along with the objects as something neither capable nor in need of reduction to anything else" (Hilbert). Thus Hilbert agrees with Kant — who, incidentally, likewise emphasized the symbolic construction with concrete tokens in algebra (*Critique of Pure Reason*, ed. Max Müller, p. 576, = p. 717 of the first edition, 1781) — that "mathematics possesses a content that is secure independently of all logic and therefore can never be based upon logic alone" (*Über das Unendliche*, p. 171).

[It should be recognized, however, that according to the Kantian usage of the words 'analytic' and 'synthetic' at least an individual equation such as $3 + 2 = 5$ ought to be called analytic; for, as Leibniz explained, it follows logically from the definitions

$$3 + 1 = 4, \quad 4 + 1 = 5, \quad (a + 1) + 1 = a + 2,$$

and thus "lies in the concepts" of the numbers 3, 5 and of the operation $+2$. Or else what meaning did Kant connect with these symbols?

Mathematics undoubtedly is *a priori*. It is not, as J. S. Mill wants to make us believe, founded on experience, in the sense that only repeated observations of numerical examples confer an increasing measure of verisimilitude upon such arithmetical theorems as

$$m + n = n + m$$

that are pretended to hold for arbitrary numbers.]

A conspicuous feature of all mathematics, which makes it so inaccessible to the layman, is the abundant use of *symbols*. The

intuitionist does not consider this an essential characteristic, he sees in them, as he does in all spoken or written language, merely a tool of communication and of support for the memory by fixation. Not so the formalist. He thinks of mathematics as consisting wholly of symbols, which have no meaning verifiable in sensual or mental intuition and which are manipulated according to fixed rules. Language, on the other hand — for instance in the description of substitution or of the practical rule of inference, as well as in metamathematical reasoning — serves him as a means for communicating modes of procedure and acts of meaningful thought. (Communication remains forever exposed to the risk of misunderstanding.) “In the geometrical figure and, later, in the mathematical formula,” A. Speiser says (*Klassische Stücke der Mathematik*, 1925, p. 148), “mathematics has liberated itself from language; and one who knows the tremendous labor put into this process and its ever-recurring surprising success, cannot help feeling that mathematics nowadays is more efficient in its particular sphere of the intellectual world than, say, the modern languages in their deplorable condition of decay or even music are on their fronts.” In his transcendental methodology (*Critique of Pure Reason*, Part II), Kant sees the essence of mathematics in the construction, “Philosophical knowledge is that which reason gains from concepts, mathematical that which it gains from the construction of concepts” (ed. Müller, p. 572, = p. 713 of the first edition, 1781). Using the theorem of the sum of the angles in a triangle as an example, he illustrates how geometrical theorems are found, not by conceptual analysis, but by construction of suitable auxiliary points and lines. The details of his description of the constructive procedure can no longer be considered satisfactory today. This much is true, however, that in the proof of a mathematical theorem it is almost always necessary to go far beyond its immediate content. The reason is to be seen in the fact emphasized before that a proof proceeding according to the syllogistic rule of inference is not a monotonically progressing construction — in contrast to a formula, whose manufacture always advances in the same direction and whose constructive parts are therefore preserved in the final form — but a constant change of adding on and removing. This circumstance, together with the points 1, 2, and 3 enumerated in Section 6 (p. 37), seem to me to give a fairly adequate characterization of construction as opposed to pure reflection.

The stages through which research in the foundations of mathematics has passed in recent times correspond to the three basic possibilities of epistemological attitude. The set-theoretical approach is the stage of *naïve realism* which is unaware of the transition from the given to the transcendent. Brouwer represents *idealism*, by demanding the reduction of all truth to the intuitively given. In axiomatic

formalism, finally, consciousness makes the attempt to 'jump over its own shadow,' to leave behind the stuff of the given, to represent the *transcendent* — but, how could it be otherwise?, only through the *symbol*. Basically, the idealist viewpoint in epistemology has been adhered to by occidental philosophy since Descartes; nevertheless, it has searched again and again in metaphysics for an access to the realm of the absolute, and Kant, who meant to shoot the bolt once and for all, was yet followed by Fichte, Schelling, and Hegel. It cannot be denied that a theoretical desire, incomprehensible from the merely phenomenal point of view, is alive in us which urges toward totality. Mathematics shows that with particular clarity; but it also teaches us that that desire can be fulfilled on one condition only, namely, that we are satisfied with the symbol and renounce the mystical error of expecting the transcendent ever to fall within the lighted circle of our intuition. So far, only in mathematics and physics has symbolical-theoretical construction gained that solidity which makes it compelling for everyone whose mind is open to these sciences. Their philosophical interest is primarily based on this fact.

[If in summing up a brief phrase is called for that characterizes the life center of mathematics, one might well say: mathematics is the *science of the infinite*. It was the great achievement of the Greeks to have made the tension between the finite and the infinite fruitful for the analysis of reality. It has been attempted here to bring out the past and present importance of this tension — and of the attempts to overcome it — for the history of theoretical knowledge. "The infinite, like no other problem, has always deeply moved the soul of men. The infinite, like no other idea, has had a stimulating and fertile influence upon the mind. But the infinite is also more than any other concept, in need of clarification" (Hilbert, *Über das Unendliche*).

For a survey of the various issues and problems in which mathematical research is interested today, the reader may be referred to Courant and Robbins, *What is Mathematics?* }

R E F E R E N C E S

I. KANT, *Kritik der reinen Vernunft*.

A thorough discussion of Kant's philosophy of mathematics is given by L. COUTURAT in *Revue de métaph. et mor.*, May 1904.

H. POINCARÉ, *La science et l'hypothèse*, Paris; many editions.

R. COURANT and H. ROBBINS, *What is Mathematics?*, New York, 1941.

L. E. J. BROUWER, *Mathematik, Wissenschaft und Sprache*, *Monatshefte der Mathematik und Physik* 36, 1929, pp. 153-164.

H. WEYL, *The Mathematical Way of Thinking*, Univ. Penn. Bicenten. Conf. 1940; also *Science* 92, 1940, pp. 437-446.

Geometry

NOWHERE do mathematics, natural sciences, and philosophy permeate one another so intimately as in the problem of space. The presuppositions for the discussion of this problem, inasmuch as they have emerged from mathematical investigation, are to be briefly outlined in this chapter.

12. NON-EUCLIDEAN, ANALYTIC, MULTI-DIMENSIONAL, AFFINE, PROJECTIVE GEOMETRY; THE COLOR SPACE

[Little has to be added concerning the topic of non-Euclidean geometry to what has been said in Section 4 in connection with axiomatics. If all remaining axioms are maintained, then there are these three possibilities: given a point P and a line l in a plane, with P not on l , there are either infinitely many lines in that plane which pass through P but do not intersect l , or just one such line, or none (known since Klein as the 'hyperbolic,' 'parabolic,' and 'elliptic' cases respectively). The sum of the angles in a triangle in these cases is respectively less than, equal to, and greater than, 180° . The last-named possibility, pointed out only toward the middle of the 19th century by Riemann, exists only if the axioms of order are modified to the effect that the line appears no longer as an open but as a closed curve. Plane elliptic geometry is none other than that which holds on a sphere in Euclidean space, except that diametrically opposite points have to be identified. Or, in other words, while all other terms referring to the geometry of the plane p retain their ordinary 'Euclidean' meaning, the meaning of the notion of congruence is to be modified to the effect that two configurations in p are considered 'congruent' if their projections from a central point O , not in p , onto a sphere about O are congruent in the ordinary sense. The plane, in this case, has to be enriched by the inclusion of the 'points at infinity,' whose rays of projection are the lines through O parallel to p . The mappings of p which map 'congruent' configurations into each other can be characterized, without reference to the space, as collinear transformations that have an invariance property similar to that prevailing in the Klein model of Bolyai-Lobatschewskyan geometry. Thus the way is open to the development of an elliptic geometry not only in the plane

but also in space. The true relation between the three kinds of geometry is brought out best if the non-metrical projective space is taken as the starting point and a 'Cayley metric' built into it. According to the type of absolute conic on which this metric is based, one or another of the three metric geometries is obtained. Klein himself interpreted his construction in this sense, namely, as endowment of projective space with a Lobatschewskyan metric, not as construction of a model by means of metric Euclidean space.]

Analytic geometry reduces every geometrical problem to an algebraic one. This presupposes that the number concept, by the inclusion of fractions and irrational numbers, has acquired that width which makes it suitable, not only for counting, but also for measuring. The Greeks had been deterred from this step because they took the discovery of the irrational seriously.¹⁵ The post-classical Western civilization, less scrupulous than they, resumed the old algebraic traditions of the Sumerians, Indians, and Arabs. It attained to independent achievements in geometry only after the science of space, through Descartes' *Géométrie* (1637), became subjected to algebraic calculus.

[Today probably the best approach to analytic geometry is by means of the vector concept, following the procedure of Grassmann's *Ausdehnungslehre*. The vector calculus is a computational device whose objects are not numbers but simple geometrical entities. A treatment of geometry along these lines was demanded and even partially executed by Leibniz in his work *De analysi situs* and his design for a geometrical characteristic (*Mathematische Schriften*, V, p. 178, and II, p. 20), which belong within the framework of his "universal characteristic." The translations, or parallel displacements, of space are called *vectors*. A point A is mapped by a translation \vec{a} into a point $A\vec{a} = B$, the 'endpoint of the vector \vec{a} laid off from A .' Conversely, if A, B are any two points in space, there exists one and only one translation \vec{a} which carries A into B . Among the translations we have the 'identity,' under which all points remain fixed; this is the vector 0 . Translations can be combined, they form a *group*; the effect of carrying out first one translation, \vec{a} , then another, \vec{b} , is the same as that of a single translation, the resultant $\vec{a} + \vec{b}$. The number concept enters geometry through the process of iteration of a

¹⁵ Descartes speaks of the "misgivings of the Ancients regarding the use of terms of arithmetic in geometry, which can only have had their origin in a lack of understanding concerning the connection between these two disciplines."

translation \vec{a} (consisting in adding \vec{a} arbitrarily often to itself; compare the beginning of Section 5). Starting with a point A and repeating the same step \vec{a} again and again, one obtains the skeleton of a line, namely, a sequence of equidistant points beginning with A . The line itself results, so to speak, by continuous iteration of the same infinitely small translation. By partition (as in Section 5) we contrive to apply not only integral but also fractional multipliers λ to the vector \vec{a} , and the continuity requirement finally removes the restriction to rational numbers. Thus arises an axiomatic construction of geometry (strictly speaking, of affine geometry, in which only parallel line segments can be measured against one another) that presupposes the fully formed concept of real number — into which the entire analysis of continuity is thrown — and uses as the only basic geometric concepts ‘point’ and ‘vector.’ Three basic operations connect these objects: (1) two vectors \vec{a}, \vec{b} generate a third vector, $\vec{a} + \vec{b}$; (2) a number λ and a vector \vec{a} generate the vector $\lambda\vec{a}$; (3) a point A and a vector \vec{a} generate a point $A\vec{a}$. The axioms referring to these operations form a system that, also in logical respect, is of a much more transparent and homogeneous structure than the purely geometrical axioms of Euclid or Hilbert. Indeed, they determine, as has already been pointed out in Section 4, nothing other than the operational field of linear algebra. They reveal a wonderful harmony between the given on one hand and reason on the other. Moreover the simplest derived geometrical concepts, to which here belong especially the line and the plane, correspond to those which suggest themselves most naturally from the logical standpoint. All vectors \vec{x} which are obtained from two given ones, \vec{e}_1 and \vec{e}_2 , through the formula

$$(1) \quad \vec{x} = x_1\vec{e}_1 + x_2\vec{e}_2,$$

with arbitrary numerical coefficient x_1, x_2 , form a ‘linear vector manifold of dimension 2.’ For the sake of the uniqueness of the representation (1) it is assumed here that \vec{e}_1, \vec{e}_2 are linearly independent, i.e. that the expression on the right furnishes the vector 0 only if x_1 and x_2 are both equal to 0. If all these vectors \vec{x} are laid off from a fixed initial point O , then the endpoints $O\vec{x} = P$ form a ‘linear point manifold of dimension 2,’ or a plane. The coordinate system here consists of the point O and the two linearly independent vectors \vec{e}_1, \vec{e}_2 . Relative to these, the point P is characterized by its ‘coordinates’ x_1, x_2 . Similarly

linear vector manifolds and linear point manifolds of dimensions 1, 2, 3, . . . (line, plane, . . .) may be introduced.}]

Only here do we meet the concept of *dimension*. In real space we cannot go beyond the third dimension; there exist 3, but no more, linearly independent vectors. Measured against the transparent lawfulness that finds its expression in our axiom system, this dimension number 3 appears as a *contingent* feature. We might just as well replace the number 3 by any number n of dimensions, by postulating that there be n , but no more, linearly independent vectors. A coordinate system for the space then consists of an initial point O and n such vectors. For $n = 1, 2, 3$ we thus obtain respectively the geometry of the line, of the plane, and of space. Only on the basis of the notion of an n -dimensional geometry to which this formalization leads in a cogent manner does the *problem of the number of dimensions* become meaningful: What inner peculiarities distinguish the case $n = 3$ among all others? If God, in creating the world, chose to make space 3-dimensional, can a 'reasonable' explanation of this fact be given by disclosing such peculiarities?

[If all vectors are laid off from a fixed initial point, it is seen that the geometry of vectors is identical with the (affine) geometry of a point space provided with an absolute center O . If one identifies any two non-vanishing vectors resulting from one another through multiplication by a number, i.e. if one considers as elements the *rays* through O , the n -dimensional affine vector geometry becomes the $(n - 1)$ -dimensional *projective geometry* (of the family of rays through O).

The projective geometry holds in the space of the perceptively given *color qualities* of colored light. (The manifold of the objective physical colors has infinitely many dimensions; of these, the normal non-colorblind eye produces a 2-dimensional 'projection,' a huge manifold of physically different colors giving the same color impression.) If two colors of definite intensity are composed (mixed), the result is a new definite color of a definite intensity. The various intensities of one color may be compared with one another, so that, after a unit intensity has been chosen, every intensity can be measured by a number (iterated composition of a color of unit intensity with itself producing a scale of intensities without change of the color quality). The intensities of two different color qualities, on the other hand, are incommensurable. Thus the colors with their various qualities and intensities fulfill the axioms of vector geometry if addition is interpreted as mixing; consequently, projective geometry applies to the

color qualities. All colors resulting from the mixing of three basic colors A, B, C form the 'triangle' ABC . The color space turns out to be 2-dimensional, by virtue of the fact that three basic colors suffice to produce all colors by mixing, or at least that the entire color field can be composed of such color triangles. For the real colors fill out only a restricted section of the entire projective plane. But it can, by the procedure described in Section 2, be extended ideally into a full projective plane; *ideal* colors must be chosen as the basic colors A, B, C if the field of real colors is to fall entirely within the triangle ABC . In the projective color plane, the pure spectral colors lie on a curve whose extremities come very close together and are connected by purple. Epistemologically it is not without interest that in addition to ordinary space there exists quite another domain of intuitively given entities, namely the colors, which forms a continuum capable of geometric treatment. }

REFERENCES

- F. KLEIN, Über die sogenannte nicht-euklidische Geometrie, *Gesammelte mathematische Abhandlungen*, I, Berlin, 1921, pp. 254-305, 311-350.
 O. VEBLEN and J. W. YOUNG, *Projective Geometry*, 2 Vols., New York, 1910-1918.
 H. WEYL, *Raum Zeit Materie*, fifth edition, Berlin, 1923; Sections 1-4.
 H. VON HELMHOLTZ, *Handbuch der physiologischen Optik*, 3 Vols., third edition, Hamburg and Leipzig, 1909-11; Section 20.

13. THE PROBLEM OF RELATIVITY

Our knowledge stands under the norm of *objectivity*. He who believes in Euclidean geometry will say that all points in space are objectively alike, and that so are all possible directions. However, Newton seems to have thought that space has an absolute center. Epicurus certainly thought that the vertical is objectively distinguishable from all other directions. He gives as his reason that all bodies when left to themselves move in one and the same direction. Hence the statement that a line is vertical is elliptic or incomplete, the complete statement behind it being something like this: the line has the direction of gravity at the point P . Thus the gravitational field, which we know to depend on the material content of the world, enters into the complete proposition as a contingent factor, and also an individually exhibited point P on which we lay our finger by a demonstrative act such as is expressed in words like 'I,' 'here,' 'now,' 'this.' Only if we are sure that the truth of the complete statement is not affected by free variation of the contingent factors and of those that are individually exhibited (here the gravitational field and the point P) have we a right to omit these factors from the statement and still to claim objective significance for it. Epicurus's belief is

shattered as soon as it is realized that the direction of gravity is different in Princeton and in Calcutta, and that it can also be changed by a redistribution of matter. Without claiming to give a mechanically applicable criterion, our description bears out the essential fact that objectivity is an issue decidable on the ground of experience only. It also accounts for the two main sources of the error so often committed in the history of knowledge, that of mistaking a statement for objective that is not: (1) one overlooked certain relevant circumstantial factors on which the meaning of the statement depends although they are not mentioned explicitly in its elliptic form, (2) though these factors were recognized, one did not investigate carefully enough whether or not the truth of the statement is affected by their variation. It is no wonder then that at several phases in the course of the history of science the realm of that which is considered objective has shrunk.

Whereas the philosophical question of objectivity is not easy to answer in a clear and definite fashion, we know exactly what the adequate mathematical concepts are for the formulation of this idea. Let us start with a completely axiomatized science like Euclidean geometry. For simplicity's sake we assume only one fundamental category, the points of space. According to Hilbert the fundamental relations that enter into the axioms would then be (1) the ternary relation: three points lie on a straight line, (2) the relation: (three distinct points A, B, C lie on a straight line and) B lies between A and C ; (3) the relation: four points lie in a plane; (4) the relation of congruence $AB \equiv CD$ between two pairs of points AB and CD . What we are going to say applies to any domain of objects the axioms of which deal with a few basic relations. Without prejudicing what the objects are we may call them points and thus speak of the domain as the point-field.

In Section 4 the notion of isomorphic mapping was introduced. We now consider the special case when our domain of objects is mapped not upon another domain but upon itself, and thus arrive at the notion of *automorphism*: an automorphism is a one-to-one mapping $p \rightarrow p'$ of the point-field into itself which leaves the basic relations undisturbed; i.e. whenever points a, b, \dots satisfy the basic relation $R(ab \dots)$ then the points a', b', \dots into which a, b, \dots are carried over by the mapping satisfy the same relation, and vice versa. In other words $R(ab \dots)$ implies $R(a'b' \dots)$, and $R(a'b' \dots)$ implies $R(ab \dots)$. A mapping σ carries every point of the point-field into a point $p' = \sigma p$. The simplest mapping is the identity ι carrying every point p into p itself. Two mappings $\sigma: p \rightarrow p'$ and $\tau: p' \rightarrow p''$ may be carried out one after the other and then give rise to a new mapping $\sigma\tau: p \rightarrow p''$. A mapping $\sigma: p \rightarrow p'$ is one-to-

one if it has an inverse σ^{-1} that carries p' back into p : $\sigma\sigma^{-1} = \sigma^{-1}\sigma = \iota$. Then σ^{-1} is also one-to-one. The identity is a one-to-one mapping; and if σ and τ are, so is $\sigma\tau$, its inverse being $\tau^{-1}\sigma^{-1}$. The word transformation will be used as a synonym for one-to-one mapping. The fundamental fact about the automorphisms is that they form a *group*. This means the following three things: (1) the identity is an automorphism; (2) if σ is an automorphism, then σ^{-1} is; (3) if σ and τ are automorphisms, then $\sigma\tau$ is. These three facts are an immediate consequence of the definition.

A figure F in its widest sense, or a configuration of points, is nothing but a point-set; F is given if for every point p it is determined whether or not it belongs to F . A ternary relation $R(xyz)$ between points is *invariant* with respect to a given transformation σ : $p \rightarrow p'$ and its inverse $p' \rightarrow p$ if $R(abc)$ always implies $R(a'b'c')$ and vice versa. We can now say in precise terms what is meant by the objective equality or 'indiscernibility' of all points in Euclidean space. It means that, given any two points p_1 and p_0 , there is always an automorphism carrying p_0 into p_1 . Two figures F and F' are *similar* if one can be carried into the other by an automorphism. That is now our interpretation of Leibniz's definition of similar figures as figures that are indiscernible if each is considered by itself. The three postulates for a group simply state that each figure is similar to itself and that similarity is symmetric and transitive (see the axioms for equivalence on p. 9). A point relation is said to be objective if it is invariant with respect to every automorphism. In this sense the basic relations are objective, and so is any relation logically defined in terms of them by means of the principles enumerated in Section 2, provided no use is made of Principle 5 permitting a blank to be filled out by an individually exhibited point. (Whether every objective relation may be so defined raises a question of logical completeness which is as unlikely to be answerable as the corresponding question of completeness for the axioms in the form whether every true universal statement about points can be deduced from the axioms.)

When our task is to investigate the real space, neither the axioms nor the basic relations are given to us. On the contrary: in our attempt to axiomatize geometry we select as our basic relations some of the point relations of which we are convinced that they have an objective significance (for instance Epicurus would have included the basic relation: A, B lie on a vertical; Euclid did not). Hence in order to do justice to the real state of affairs we shall have to invert the order in the development of our ideas. We start with a group Γ of transformations. It describes, as it were, to what degree our point field is homogeneous. Once the group is given we know what like-

ness or similarity means — namely two figures are similar (or alike, or equivalent) that arise from each other by a transformation of Γ —, and also under what condition a relation is objective, namely if it is invariant with respect to all transformations of Γ . It is in this sense that Felix Klein in his famous Erlanger Program (1872) promulgated the conception that a geometry is determined by a group of transformations. The question of axiomatizing this geometry is now relegated to the background. (As a first step it would require the finding of a few objective relations R_1, R_2, \dots such that the group of all transformations leaving R_1, R_2, \dots invariant is not larger than Γ but coincides with Γ .) While we need not close our eyes to the fact that objective relations can be logically constructed from other such relations, we refrain from making a distinction between basic and derived. We are equally interested in all invariant relations.

[If Newton were right in ascribing to space an absolute center O , the true group Γ_0 of automorphisms would consist of those transformations of the Euclidean group Γ of automorphisms which leave O fixed; Newton's Γ_0 is a subgroup of Euclid's Γ . On the other hand, in studying Euclidean geometry we may be primarily interested in such properties as are invariant with respect to all affine or all projective transformations. (The affine and the projective transformations of a plane are those that result from carrying out one after the other any number of parallel projections or central projections respectively.) The groups Γ' and Γ'' of these transformations are wider than Γ ; more precisely, Γ is part of Γ' , and Γ' part of Γ'' . The importance of affine and projective geometry for the theory of perspective is obvious. One sees how helpful Klein's point of view proves in surveying, and bringing to light the mutual relationship of, various kinds of geometries such as are either suggested by the nature of things or spring from arbitrary but logically useful abstraction. (Klein had a predecessor in Möbius who stressed the group-theoretical viewpoint for a number of special types of geometries.) The widest group of automorphisms one can possibly envisage for a continuum consists of all continuous transformations; the corresponding geometry is called topology. It was a lucky chance for the development of mathematics that the relativity problem was first tackled, not for the continuous point space, but for a system consisting of a finite number of distinct objects, namely the system of the roots of an algebraic equation with rational coefficients (Galois theory). This circumstance has greatly benefited the exactness of the relevant concepts. The objective relations are here those which can be constructed by means of the four basic operations of algebra (addition, subtraction, multiplication, division), in other words the

algebraic relations with rational coefficients. This sort of problems gave rise to a general theory, not only of transformation groups, but also of abstract groups.}]

Having explained automorphism we now come to a second phase of the relativity problem. How is it possible to assign to the points of a point-field marks or labels which could serve for their identification or distinction? The labels are supposed to be self-created, distinctive and always reproducible symbols, such as names, numbers (or number triples x, y, z , etc.). Only after this has been accomplished can one think of representing the spectacle of the actually given world by construction in a field of symbols. All knowledge, while it starts with intuitive description, tends toward symbolic construction. No serious difficulty is encountered as long as one deals with a domain consisting of a finite number of points only, which can be 'called up' one after the other. The problem becomes a serious one when the point-field is infinite, in particular when it is a continuum. A conceptual fixation of points by labels of the above-described nature that would enable one to reconstruct any point when it has been lost, is here possible only in relation to a *coordinate system*, or frame of reference, that has to be exhibited by an individual demonstrative act. The objectification, by elimination of the ego and its immediate life of intuition, does not fully succeed, and the coordinate system remains as the necessary residue of the ego-extinction.¹⁶ It is good to remember here that in practice two- or three-dimensional point-sets are usually given by actually putting a body or a figure drawn with pencil on paper before our eyes, and not by a logico-arithmetical construction of set-defining properties. It took a long time for mathematics before it had acquired the constructive tools to cope with the complexity and variety of such intuitively given figures. But once it had reached that stage the superiority of its symbolic methods became evident.

{Take as an example the points on a line. The coordinate system consists here of a point O and a unit segment OE , or of two distinct points O, E . When this frame of reference is given, any point P can be characterized by its abscissa x , the number measuring the length OP with OE as the unit yardstick (x is positive for points lying on the same side of O as E , negative for points on the opposite side). Any

¹⁶ Against the establishment of an essential difference between conceptual determination and intuitive exhibition, the objection might be raised that even the objective geometrical relations upon which the conceptual determination is based require intuitive exhibition. But these are a few isolated relational concepts, while the points themselves form a continuum. I am inclined to admit that this fact alone constitutes the essential difference.

two frames of reference, OE and $O'E'$, are objectively alike, for there is exactly one automorphism (similarity) that maps O into O' and E into E' . Hence by exhibiting an individual coordinate system no more is exhibited than is absolutely necessary. The field for the symbol x consists of all real numbers. Relative to a given coordinate system the correspondence $P \leftrightarrow x$ is a one-to-one mapping of the point-field onto the variability range of the symbol. The coordinates x and x' of the same arbitrary point in two coordinate systems are connected by a relation $x = ax' + b$ where $a \neq 0$ and b are two constants characteristic of the relative position of the two coordinate systems.]

With this example in mind, one will be able to understand the following general description. A class Σ of frames of reference f is supposed to be given. The class as such should be objectively distinguished; i.e. if f belongs to it, so does any similar frame $f\sigma = f'$ arising from f by an automorphism σ . But the class is supposed to contain no more elements than this requirement makes absolutely necessary, i.e. any two frames f, f' of the class are similar. Moreover an objective rule A is supposed to be given by which each point p with respect to any frame f of the class Σ determines a definite (reproducible) symbol $x = A(p; f)$ as its coordinate. For a given f the correspondence $p \leftrightarrow x$ between points p and symbols x is one-to-one. That x is objectively determined by p and f means that

$$(1) \quad A(p; f) = A(p\sigma; f\sigma)$$

for any automorphism σ .

From these conditions there flow the following consequences. Let σ be an automorphism $p \rightarrow p'$ of the point-field and f be a fixed frame of class Σ . The coordinates x of p and x' of p' in this frame are connected by a transformation S , $x' = xS$, which represents the automorphism σ in terms of f . To the identity $\sigma = \iota$ there corresponds the identity $S = I$; σ^{-1} and $\sigma\tau$ are represented by S^{-1} and ST if S and T represent σ and τ . In this sense the transformations S corresponding to the several σ of Γ form a group \mathbf{G} that is isomorphic with Γ . \mathbf{G} is nothing but the representation of Γ in terms of f . Take, on the other hand, a fixed frame f of our class Σ and an arbitrary frame $f' = f\sigma$ that arises from f by the automorphism σ . I maintain that the coordinates x, x' of the same arbitrary point with respect to f and f' are connected by the equation $x = x'S$. Indeed, denote the arbitrary point by $p\sigma$ instead of p ; we then have $x = A(p\sigma; f)$ and, because of (1),

$$x' = A(p\sigma; f\sigma) = A(p; f),$$

and thus our assertion follows. The group G that represents Γ in terms of f must be independent of f . Indeed, representation of Γ by two different groups G, G^* in terms of the two similar frames f and f^* would constitute an objective difference between f and f^* , which is impossible. It is easy to verify this explicitly. Let $f^* = f\gamma$, where γ is an automorphism. Moreover, let x and x' be the coordinates of an arbitrary point p and its image $p' = p\sigma$ with respect to f , and y and y' with respect to f^* . The transformations representing γ and σ in terms of f may be called C and S . Write the equation $x' = xS$ in

the more suggestive form $x \xrightarrow{(S)} x'$. After what has been said, we then have the following diagram

$$\begin{array}{ccc} & (S) & \\ x & \rightarrow & x' \\ (C) \uparrow & & \downarrow (C^{-1}) \\ y & & y' \end{array}$$

Hence the transformation that leads from y to y' and thus represents σ in terms of the frame f^* is $S^* = CSC^{-1}$. With S also $CSC^{-1} = S^*$ is in the group G , and vice versa: $S = C^{-1}S^*C$.

As long as the points could not be characterized conceptually, the transformations of the point field could not be either, and it was thus perhaps not perfectly clear what was meant by saying that the group of automorphisms is known or given. A stage has now been reached where this last shadow of obscurity disappears. Every point is replaced by its coordinate x (with respect to a fixed frame), and thus the group Γ of automorphisms σ appears as a group G of transformations S . The individual transformation S carrying x into $x' = xS$, is a reproducible symbol like any individual value of x . But while the coordinate x is not only dependent on p but also on f , the group G is independent of f and hence free from anything in need of individual exhibition. To fulfill the demand of objectivity we construct an image of the world in symbols. The pure mathematician will say: Given a group G of transformations in a field of symbols, a geometry is established by agreeing to study, and consider as objective, only such relations in that field as are invariant under the transformations of G .

[A last remark of a purely logical nature concerns the frames. It is quite legitimate to regard as the frame of reference f the coordinate assignment $p \rightarrow x = f(p)$ itself established by f . This seems even preferable if one has to be prepared for a group of automorphisms so wide as to comprise all continuous transformations. The symbol f is then simply a token for the function f whose argument ranges over the

points p and whose value is an element x in the field of symbols. If $\sigma: p \rightarrow p'$ is any transformation, then the transformed function $f' = f\sigma$ will be defined by the equation $f'(p') = f(p)$ for $p' = p\sigma$, or $f'(p) = f(p\sigma^{-1})$. When we write $x = A(p; f)$ for $x = f(p)$ then A stands for the universal logical operator 'value of'; $x = A(p; f)$ means: x is the value of the function f for the argument p .]

R E F E R E N C E S

F. KLEIN, *Vergleichende Betrachtungen über neuere geometrische Forschungen*, Erlangen, 1872 (*Gesammelte mathematische Abhandlungen*, I, 460–497).

14. CONGRUENCE AND SIMILARITY.

LEFT AND RIGHT

There is no doubt that the conviction which Euclidean geometry carries for us is essentially due to our familiarity with the handling of that sort of bodies which we call rigid and of which it can be said that they remain the same under varying conditions. The portions of space which such a solid fills in two of its positions are called *congruent*. Measurement depends on rigid bodies to the same degree as counting does on the use of concrete number symbols. (About the physical foundation of geometry cf. also Sections 16 and 18.) Once geometry has been abstracted from the behavior of actual bodies that are approximately rigid it provides a standard for the physical investigation of all bodies, and we can judge how far a given body realizes the ideal of rigidity. This process is not essentially different from the one by which a scale of temperature is first based on the behavior of actual gases and then reduced to the 'ideal gas scale' by postulating the exact validity of such laws as are approximately satisfied by the existing gases. Since places on a rigid body can be tagged, congruence is a point-by-point mapping of the two congruent volumes V and V' . The notion of congruence at first is relative to a given rigid body b . Its factual independence of b is one of our most fundamental experiences. Indeed, let V, V' be two portions of space filled by the solid b in two of its positions. Let b^* be another solid that fits into V ; then it may be so moved as to fill V' . Since one may extend a rigid body so as to cover any given point, the mapping $V \rightarrow V'$ can be extended to the whole space. The congruent mappings of space form a group Δ^+ of transformations which we call the *group of Euclidean motions*. Once this group is known, congruent volumes may be defined as portions of space that can be carried into each other by a transformation S of Δ^+ . The facts suggest an interpretation according to which the

group Δ^+ of congruent mappings expresses an intrinsic structure of space itself; a structure stamped by space on all spatial objects.

If this view is correct, congruence should be made the one and only basic concept of geometry. Let us first investigate what the consequences of this conception of geometry are for the automorphisms of space (*similarities*). We know quite generally that once the basic relational concepts of a geometry are fixed the group Γ of automorphisms is also fixed. In our case the criterion for an automorphism C is this: C as well as C^{-1} must transform any pair of congruent portions of space v_1, v_2 into a congruent pair. Consider the pair v_1^*, v_2^* arising from v_1, v_2 by the transformation C . Let S be the motion that

$$\begin{array}{ccc} & (S) & \\ v_1 & \rightarrow & v_2 \\ (C^{-1}) \uparrow & & \downarrow (C) \\ v_1^* & & v_2^* \end{array}$$

carries v_1 into v_2 . As the above diagram indicates, v_1^* goes into v_2^* by the mapping $C^{-1}SC$. Hence the criterion demands that the transformations $C^{-1}SC$ and CSC^{-1} should belong to Δ^+ whenever S does. A transformation C is said to commute with a given group Δ of transformations if $C^{-1}SC$ and CSC^{-1} are in Δ whenever S is. The transformations commuting with Δ form a group called the normalizer of Δ . This group necessarily contains Δ as a subgroup, be it that Δ is identical with its normalizer or a proper part of it. Our analysis can now be summarized thus: *The group Γ of similarities is the normalizer of the group Δ^+ of motions.* Hence congruent figures are necessarily similar. The converse need not be true. Indeed, since Δ^+ happens to be a proper subgroup of its normalizer there exist similar figures in Euclidean space which are not congruent; as for instance a body and its mirror image, or a building and a small scale model of it.

Let us now invert the procedure and follow Klein by starting with a given group Γ of automorphisms. Take a subgroup Δ of Γ and declare two figures to be Δ -equivalent if one is carried into the other by a transformation of Δ . Under what circumstances has this relation of Δ -equivalence objective significance? If and only if Δ -equivalent figures are carried into Δ -equivalent figures by every transformation C of Γ , or in other words, if every element C of Γ commutes with Δ . In that case the mathematician says that Δ is an invariant subgroup of Γ . Hence Δ -equivalence is an objective relation provided Δ is an invariant subgroup of Γ . For instance, the parallel displacements form an invariant subgroup of the group of Euclidean similarities; and indeed the relation \parallel between two figures arising from each other by parallel displacement is clearly of objective geometric sig-

nificance — although our language lacks a suggestive word for it. The normalizer of the group of parallel displacements consists of all affine transformations; hence affine geometry may be based on the one relation \parallel between figures. Or, still more simple, the subgroup consisting of the identity only is an invariant subgroup, and indeed the relation of identity between two figures is of objective significance. (There is none that has a better claim to objectivity, owing to the fact that the identity is contained in every possible group Γ of transformations.) The smaller the group Δ the larger its normalizer, and thus the wider the gap between congruence and similarity; or more precisely, if Δ' is a subgroup of Δ then the normalizer Γ' of Δ' contains the normalizer Γ of Δ . The normalizer of an invariant subgroup Δ of Γ always comprises Γ . A geometry whose group of automorphisms is Γ can be based on the objective relation of Δ -equivalence *alone*, provided the normalizer of Δ is not larger than Γ but coincides with Γ .

{A last remark will conclude this analysis. Space is a continuum, and when we speak of any transformation in space it is reasonable to interpret this as meaning any continuous transformation. We indicate by Ω the group of those transformations that are taken into account at all; in the case of a continuum this would be the group of all continuous transformations. By putting this explicitly in evidence our definition of normalizer may be repeated as follows. Given a subgroup Δ of the group Ω ; those elements of Ω that commute with Δ constitute the normalizer Γ of Δ . In this form the notion of normalizer makes sense even for abstract groups Ω and Δ .

Kant speaks about the divergence between congruent and similar in *Prolegomena*, §13, and claims that “by no single concept, but only by pointing to our left and right hand, and thus depending directly on intuition [Anschauung] can we make comprehensible the difference between similar yet incongruent objects (such as oppositely wound snails)”; and in his opinion only transcendental idealism offers a solution for this riddle. No doubt the meaning of congruence is based on spatial intuition, but so is similarity. Kant seems to aim at some subtler point, but just this point is one which can be completely clarified by an analysis in terms of a group Γ and its invariant subgroups Δ , or of a group Δ and its normalizer Γ . Whenever Δ is a proper invariant subgroup of Γ , the notions of congruence = Δ -equivalence and similarity = Γ -equivalence do not coincide although the former is of objective significance (= Γ -invariant). The phenomenon about which Kant wonders can thus be most satisfactorily subsumed under general and abstract ‘concepts.’ }

Whoever raises congruence to the rank of the only basic relation of geometry is obliged to develop geometry from this one notion. Several ways are open to accomplish this. A deeper insight than by the elementary approach in the style of Euclid's axioms would be gained if one succeeded in formulating the fundamental facts of geometry as simple axioms concerning the group Δ^+ of Euclidean motions. Following Ueberweg, Helmholtz first carried out this program with surprising success in his essay "Ueber die Tatsachen, die der Geometrie zugrunde liegen." Later S. Lie, who established a general theory of transformation groups, resumed the problem with his more powerful mathematical tools and generalized it from 3 to n dimensions. The Euclidean group of motions Δ^+ turns out to be almost completely characterized by the fact that it permits the rigid body that measure of free mobility with which we are familiar by experience. In more exact terms: it is possible by suitable congruent mappings to carry any point into any other, and, if a point is kept fixed, to carry any line direction at that point into any other at the same point; furthermore, if a point and line direction are kept fixed, it is possible to carry by congruent mapping any surface direction through them into any other such direction, and so forth, up to the $(n - 1)$ -dimensional direction elements. If, on the other hand, a point and a line direction through it, and a surface direction through the latter, and so forth, up to an $(n - 1)$ -dimensional direction element, are given, then there exists no congruent mapping besides the identity under which this system of incident elements remains fixed. We just said that this axiom *almost* completely characterizes the Euclidean group of motions. In fact one thus obtains the group of congruent transformations of a slightly more general space, namely of a projective space endowed with a Cayley metric. That group contains a numerically indeterminate parameter λ , the constant space curvature, of which nothing but the sign is essential. According as λ is positive, zero, or negative, the resulting space is of the elliptic, parabolic (i.e. Euclidean), or hyperbolic type. These then are the only homogeneous spaces, in which all points are equivalent, likewise all directions at a point, and so on.

[It is hard to talk intelligently about these problems without an exact description of the Euclidean groups Γ and Δ^+ before our eyes. A Cartesian frame of reference in three-dimensional Euclidean space consists of a point O , the origin, and three mutually perpendicular vectors $\vec{e}_1, \vec{e}_2, \vec{e}_3$ of equal length. The coordinates x_1, x_2, x_3 of a point P are defined by

$$\vec{OP} = x_1 \vec{e}_1 + x_2 \vec{e}_2 + x_3 \vec{e}_3.$$

Relative to such a frame a similarity mapping the point (x_1, x_2, x_3) into the point (x'_1, x'_2, x'_3) is represented by a linear transformation

$$(1) \quad S: \quad x'_i = a_i + a_{i1}x_1 + \cdots + a_{in}x_n \quad (i = 1, 2, \cdots, n)$$

with constant coefficients a_i, a_{ik} that satisfy the following condition: $(x'_1 - a_1)^2 + \cdots + (x'_n - a_n)^2$ is a constant positive multiple a of $x_1^2 + \cdots + x_n^2$. (Here the number n of dimensions has been left indeterminate.) The similarity is 'non-enlarging' and called an orthogonal transformation if $a = 1$. The orthogonal transformations form an invariant subgroup Δ of Γ . The condition mentioned above as one satisfied by every similarity implies the equation $d^2 = a^n$ for the determinant d of the a_{ik} . Hence an orthogonal transformation is either of signature $+$ ($d = +1$) or of signature $-$ ($d = -1$). The orthogonal transformations of signature $+$ form the group Δ^+ of Euclidean motions. Δ^+ is a subgroup of Δ of index 2, i.e., if S_1, S_2 are any two transformations of Δ of signature $-$ then $S_1^{-1}S_2$ has the signature $+$. (The fundamental fact of the distinction of left and right: two screws oppositely winding to a given screw turn in the same sense.) It makes little difference whether we claim Δ^+ or Δ as the group of congruent mappings. Assume we decide in favor of the larger group Δ . Then the continuous motion of a rigid body would be represented by an orthogonal transformation $S(t)$ depending continuously on the time parameter t and reducing to the identity I at the initial moment $t = 0$. Since the determinant of $S(t)$ is capable of the two values $+1$ and -1 only, since it equals $+1$ at the beginning $t = 0$ and varies continuously with t , it must always remain equal to $+1$. Hence even if we had admitted arbitrary orthogonal transformations, the requirement of continuity for $S(t)$ automatically eliminates those of signature $-$; a rigid body could go over into its mirror image only by a discontinuous jump.]

A far deeper aspect of the group Δ than that of describing the mobility of rigid bodies is revealed by its role as the group of *automorphisms of the physical world*. In physics we have to consider not only points but also various types of physical quantities, velocity, force, electromagnetic field strength, etc. But it is a fact that relative to a Cartesian frame, not only points but all physical quantities can be represented by numbers; e.g. a force by its components f_i ($i = 1, 2, \cdots, n$), an electromagnetic field strength by a set of skew-symmetric components $F_{ik} = -F_{ki}$, etc. And under the influence of any orthogonal mapping S , (1), of the points of space they undergo a related transformation that is uniquely determined by S ; e.g. the

force components transform according to the equations

$$f'_i = \sum_{\lambda} a_{i\lambda} f_{\lambda} \quad (i, \lambda = 1, 2, \dots, n),$$

the components of the electromagnetic field strength according to the rule

$$F'_{ik} = \sum_{\lambda, \mu} a_{i\lambda} a_{k\mu} F_{\lambda\mu} \quad (i, k, \lambda, \mu = 1, 2, \dots, n),$$

etc. All the laws of nature are invariant under the transformation thus induced by the group Δ . It is not true however that they are invariant under all similarities, although it seems so on a certain level of natural phenomena. But the facts of atomism teach us that *length is not relative but absolute*. The atomic constants of charge and mass of the electron and Planck's quantum of action h fix an absolute standard of length, that through the wave lengths of spectral lines is also made available for practical measurements. Thus we no longer depend on the preservation of the platinum-iridium meter bar that is kept in the vaults of the Comité International des Poids et Mesures in Paris. We now prescribe the absolute length 1 for the basic vectors of a Cartesian frame of reference. The orthogonal transformations of signature — must be included in Δ . For there is no indication in the laws of nature of an intrinsic difference between left and right. Now it is clear why a body all of whose places undergo a transformation $S(t)$ of the group Δ depending continuously on the time parameter t and whose physical characteristics change accordingly, has a perfectly good claim to say of himself: I have remained physically the same during my motion.

[The extensive medium of the external world is one of time as well as space. How time is included as a fourth coordinate in the above scheme will be discussed in Section 16. It was in preparation for this step that we left the number n of dimensions indeterminate. For physics the case $n = 4$ is even more important than $n = 3$. At present however we shall limit ourselves to space.]

We summarize: The group of physical automorphisms in space is the group Δ of orthogonal transformations. The group of geometric automorphisms, by virtue of the very meaning of this term, is the normalizer Γ of Δ . It is larger than Δ , inasmuch as it includes the dilatations $x'_i = ax_i$ with any constant $a > 0$. This divergence between Δ and Γ proves conclusively that *physics can never be reduced to geometry* as Descartes had hoped.

[*Left and right.* Were I to name the most fundamental mathematical facts I should probably begin with the fact (F_1) that the counting of a set of elements leads to the same number in whatever order one picks up its elements, and mention as a second the fact (F_2) that among the permutations of n (≥ 2) things one can distinguish the even and the odd ones. The even permutations form a subgroup of index 2 within the group of all permutations. The first fact lies at the bottom of the geometric notion of dimensionality, the second of that of 'sense.' Consider affine vector geometry. A basis for its vectors consists of n vectors $\vec{e}_1, \dots, \vec{e}_n$ such that every vector can be uniquely expressed as a linear combination $x_1\vec{e}_1 + \dots + x_n\vec{e}_n$, and the theorem of the invariance of dimensionality states that every basis necessarily consists of the same number n of vectors. This assertion clearly implies the fact (F_1); for by any regrouping of the basic vectors one passes to a new basis. Vice versa, the theorem of invariance is an algebraic proposition easily deduced from the fact (F_1) in conjunction with the rule for addition and multiplication of numbers. Any arrangement of n given linearly independent vectors fixes a 'sense,' and two arrangements fix the same sense provided they arise from each other by an even permutation (definition by abstraction). An odd permutation changes the sense into its opposite. That is clearly the combinatorial root of the distinction between left and right. Again, in combination with the basic operations of affine vector geometry (addition of vectors, multiplication of a vector by a number) it leads to a comparison of sense for any two bases $\vec{e}_1, \dots, \vec{e}_n$ and $\vec{e}_1^*, \dots, \vec{e}_n^*$. When one expresses the vectors \vec{e}^* in terms of the vectors e ,

$$\vec{e}_i^* = a_{i1}\vec{e}_1 + \dots + a_{in}\vec{e}_n,$$

the coefficients a_{ki} have a non-vanishing determinant. The senses of the two bases are the same or opposite according to whether the determinant is positive or negative. But the definition of a determinant is based on the distinction between even and odd permutations!

Kant finds the clue to the riddle of left and right in transcendental idealism. The mathematician sees behind it the combinatorial fact of the distinction of even and odd permutations. The clash between the philosopher's and the mathematician's quest for the roots of the phenomena which the world presents to us can hardly be illustrated more strikingly.]

15. RIEMANN'S POINT OF VIEW. TOPOLOGY

The notions of dimensionality and sense are not restricted to metric Euclidean or affine space. They apply to continuous manifolds in

general. Riemann was the first to analyze mathematically the general concept of an n -dimensional manifold. A sufficiently small neighborhood of an arbitrary point in an n -dimensional manifold may be mapped one-to-one and continuously upon a region of the n -dimensional number space, the points of the latter being the n -tuples of real numbers (x_1, x_2, \dots, x_n) . Any one-to-one transformation of the coordinates

$$\begin{aligned} y_i &= \varphi_i(x_1, \dots, x_n) & (i = 1, \dots, m); \\ x_k &= \psi_k(y_1, \dots, y_m) & (k = 1, \dots, n) \end{aligned}$$

yields a new coordinate assignment suitable for the representation of the same neighborhood. Is m necessarily equal to n ? This is the question of the topological invariance of dimensionality.

[Let $P = (x_1, \dots, x_n)$ be a given point and $P^* = (x_1 + dx_1, \dots, x_n + dx_n)$ any point infinitely near to P . If the transformation functions are differentiable then the components (dx_1, \dots, dx_n) of all infinitesimal vectors $\vec{PP^*}$ issuing from P transform according to linear formulas

$$(1) \quad dy_i = \sum_k a_{ik} \cdot dx_k, \quad dx_k = \sum_i b_{ki} \cdot dy_i$$

the coefficients a_{ik}, b_{ki} of which depend on the point P but not on P^* . (Infinitesimal quantities may be avoided by introducing an imaginary time τ and letting a point move in the manifold according to an arbitrary law $x_k = x_k(\tau)$. Suppose the point passes P at the moment $\tau = 0$; its velocity at that moment will be a vector at P with the x -components $u_k = (dx_k/d\tau)_{\tau=0}$. The y -components v_i of the same velocity are related to the x -components by the equations (1),

$$v_i = \sum_k a_{ik} u_k, \quad u_k = \sum_i b_{ki} v_i,$$

which hold for all possible velocities in P .) But these linear transformations can be inverse to each other only if $m = n$ and the determinant of the a_{ik} , the so-called Jacobian, is different from zero. Only such 'differentiable' transformations of the coordinates are now admitted at all to the totality Ω . Under these circumstances one speaks of a differentiable manifold. As the Jacobian varies continuously with P , it is either positive throughout the region covered by the two coordinate assignments, or negative throughout. We give the transformation the signature $+$ in the first case, $-$ in the second. Hence a 'sense' can be fixed over the whole region. One sees that both dimensionality and sense derive from the fact that affine geometry

holds in the infinitely small. While topology has succeeded fairly well in mastering continuity, we do not yet understand the inner meaning of the restriction to differentiable manifolds. Perhaps one day physics will be able to discard it. At present it seems indispensable since the laws of transformation of most physical quantities are intimately connected with that of the differentials dx_i , (1).]

Inspired by Gauss's theory of curved surfaces, Riemann assumed that Euclidean geometry holds in the infinitely small. Then the square of the length ds of the infinitesimal vector \vec{PP}^* with the components dx_i will be expressed by a positive quadratic form

$$(2) \quad ds^2 = \sum_{ik} g_{ik} dx_i dx_k$$

of the dx_i . Its coefficients g_{ik} are independent of the vector \vec{PP}^* with the components dx_i but will in general depend on the point P with the coordinates x_i and be continuous functions of these coordinates. It is clear from the invariant significance of ds^2 how the components g_{ik} of the 'metric field' will transform under transition to a new coordinate system y_i . The metric of a 3-dimensional Riemann space of this kind imposes itself upon any surface lying in it, which is thereby branded as a 2-dimensional Riemann space. For a 3-dimensional Euclidean space, however, it is not true that every surface in it is a 2-dimensional Euclidean space; rather, all possible 2-dimensional Riemann spaces occur as subspaces of a Euclidean 3-space. Thus in Euclid's geometry the space appears as something much more special (namely, non-curved) than the possible surfaces in it, while Riemann's space concept has just the right degree of generality to do away with this discrepancy.

As the true lawfulness of nature, according to Leibniz's continuity principle, finds its expression in laws of nearby action, connecting only the values of physical quantities at space-time points in the immediate vicinity of one another, so the basic relations of geometry should concern only infinitely closely adjacent points ('near-geometry' as opposed to 'far-geometry'). *Only in the infinitely small may we expect to encounter the elementary and uniform laws*, hence the world must be comprehended through its behavior in the infinitely small.

If one requires the space to be metrically homogeneous — and a space that can serve as 'form of phenomena' is necessarily homogeneous — then one is thrown back at once from the Riemannian to the classical space concept, to which Helmholtz's postulates for the group of motions lead. But Riemann had an entirely different conception of the nature and origin of the metrical properties of space. For him the metric field is not given rigidly once and for all, but is causally

connected with matter and thus changes with the latter. He considers the metric not as part of the static homogeneous form of phenomena, but of their ever-changing material content. Riemann asks for the inner reason of the metrical relationships in space, and having distinguished (in the words quoted on p. 43) between the cases of discrete and continuous manifolds, he continues, "Therefore, either the reality on which our space is based must form a discrete manifold, or else the reason for the metrical relationships is to be looked for externally in binding forces acting upon it." The metric field makes itself felt through the physical effects which it has upon rigid bodies, upon light rays, and all events in nature, and these effects alone permit us to ascertain the quantitative state of the metric field. But whatever acts must suffer too; it must itself be something real and cannot be enthroned in unattackable 'geometric' rigidity above the forces of matter. Thereby, in spite of the non-homogeneity of the metric field, the free mobility of bodies without changes in measure is regained, since a body in motion will 'take along' the metric field that is generated or deformed by it. Einstein, after having extended space by the inclusion of time to the full four-dimensional medium of the external world, has developed Riemann's idea into a detailed physical *theory of gravitation* and, in particular, has ascertained the laws according to which matter acts upon the metric field.

Riemann and Einstein maintain that the group of — geometric or physical — automorphisms coincides with the totality Ω of all differentiable transformations. In this respect their theories differ radically from the standpoint expounded in the previous section. Their principle of general relativity is acceptable only after inserting the metric field among the physical quantities that act upon, and are reacted upon by, matter. Nevertheless Euclidean geometry is preserved for the infinitesimal neighborhood of any given point P_0 . For it is a mathematical fact that for all line elements at a given point P_0 the metric equation (2) takes on the special form

$$ds^2 = dx_1^2 + dx_2^2 + \cdots + dx_n^2$$

if appropriate coordinates x_i are chosen for the neighborhood of P_0 . In this form there is no room for any indeterminacy, and we may therefore say that the *nature* of the metric is the *same* at every point. But the coordinate system in which the metrical law assumes this fixed standard form and which, as we shall say, is characteristic for the *orientation* of the metric is in general different from place to place. We use an analogous phrase in Euclidean geometry when we say that all cubes (of given size) are of the same *nature* and differ only by their *orientation*. The nature of the metric is *one*, and is absolutely given;

only the mutual orientation in the various points is capable of continuous changes and dependent upon matter. Euclidean space may be compared to a crystal, built up of uniform unchangeable atoms in the regular and rigid unchangeable arrangement of a lattice; Riemannian space to a liquid, consisting of the same indiscernible unchangeable atoms, whose arrangement and orientation, however, are mobile and yielding to forces acting upon them.

Perhaps this is brought out better by a different formulation of Riemann's conception, which has become indispensable in quantum physics when the quantities characterizing a spinning electron are to be fitted into general relativity theory. From the above illustration by velocities it is clear what is meant by the body of tangent vectors (velocities) at P . They form an n -dimensional vector space. The coordinate assignment $P \rightarrow x_i$ determines a vector basis $\vec{e}_1, \dots, \vec{e}_n$ in this tangent vector space $V(P)$ at P such that $u_1 \vec{e}_1 + \dots + u_n \vec{e}_n$ is the vector with the x -components u_i . Assuming that the vector space at P bears a Euclidean metric (with an absolute standard of length) we can introduce in it a local Cartesian frame of reference $\mathfrak{f} = \mathfrak{f}(P)$ consisting of n mutually perpendicular vectors of length 1. The arbitrariness in the choice of this frame is expressed by the group Δ_0 of Euclidean rotations. That group consists of all linear transformations

$$S: \quad z'_\beta = \sum_\gamma a_{\beta\gamma} z_\gamma \quad (\beta, \gamma = 1, \dots, n)$$

for which

$$z'^2_1 + \dots + z'^2_n = z^2_1 + \dots + z^2_n.$$

Here the variables z_β designate the components of an arbitrary vector of $V(P)$ with respect to the Cartesian frame \mathfrak{f} . The numerical values $e_{i\beta}$ ($\beta = 1, \dots, n$) of the components of each of the vectors \vec{e}_i ($i = 1, \dots, n$) with respect to \mathfrak{f} describe the embedment of the frame \mathfrak{f} into space. Thus the n^2 quantities $e_{i\beta}$, which depend on the choice of the coordinates x_i as well as on the Cartesian frame $\mathfrak{f}(P)$ at P and are functions of P , now serve to characterize the metric field. Riemann's g_{ik} are easily computed to have the values

$$g_{ik} = e_{i1}e_{k1} + \dots + e_{in}e_{kn}.$$

Only after coordinates x_i and a Cartesian frame $\mathfrak{f}(P)$ at each point P have been chosen can all physical quantities be represented by numbers. The laws of nature are invariant (1) with respect to arbitrary transformations of the coordinates x_i , and (2) with respect to a rota-

tion S of the frame $\{P\}$ that may depend in an arbitrary (continuous) manner on the point P . Hence there is this double invariance, the one described by the group Ω of all transformations of the coordinates x_i , the other by an element of the group Δ_0 that can vary arbitrarily with the position P .

What has happened in the transition from special to general relativity theory is obviously this. The physical automorphisms forming the group Δ as described in the previous section have been split into their translatory and rotatory parts. The group of translations has been replaced by that of all possible transformations of the coordinates, whereas the rotations have remained Euclidean rotations but are now tied to a center P and must be allowed to vary freely while the center P moves over the manifold. Space, the extensive medium of the material world, is clearly the seat of the group Ω of coordinate transformations; but the group Δ_0 seems to have its origin in the ultimate elementary particles of matter. The quantities $e_{i\beta}$ thus mediate between matter and space.

[The question arises for what inner reasons nature has picked Δ_0 among all possible groups of homogeneous linear transformations. One answer is provided by Helmholtz's theory, according to which Δ_0 is completely characterized by the axiom of free mobility: Any incident set σ of 1-, 2-, . . . $(n - 1)$ -dimensional directions can be carried into any other such set by a transformation of Δ_0 while those transformations of Δ_0 that leave a given set σ of incident directions fixed form a subgroup containing two elements only (namely the identity and the reflection in σ). However, this characterization carries less conviction now where the group can no longer be interpreted as describing the mobility of a rigid body. (Moreover it breaks down for the Lorentz group, which in the four-dimensional world takes the place of the orthogonal group in 3-space.)

The group Δ_0 could be considered as an abstract group of which various representations by linear transformations are characteristic for various physical quantities; e.g. the representation Δ_0 by orthogonal transformations itself for the vectors, a certain 'tensor' representation for the electromagnetic field strength, and a very remarkable one, the so-called spinor representation, for the electronic wave field.]

Topology. In general a coordinate assignment covers only part of a given continuous manifold. The 'coordinate' (x_1, \dots, x_n) is a symbol consisting of real numbers. The continuum of real numbers can be thought of as created by iterated bipartition. In order to account for the nature of a manifold as a whole, topology had to

develop combinatorial schemes of a more general nature. By this combinatorial approach it also got rid of the restriction to differentiable manifolds.

In order to subject a continuum to mathematical treatment it is necessary to assume that it is divided up into 'elementary pieces' and that this division is constantly refined by repeated subdivision according to a fixed scheme (which in the one-dimensional case consists in the bipartition of each elementary segment). The effect is that the continuum is spun over with a subdivision net of increasing density. Thus, properly speaking, every continuum has its own arithmetical scheme which is already completely determined by the combinatorial description of the manner in which the individual elementary pieces of the initial division border on each other; we call this the 'topological skeleton' of the manifold. The introduction of numbers as coordinates by reference to the particular division scheme of the open one-dimensional continuum is an act of violence whose only practical vindication is the special calculatory manageability of the ordinary number continuum with its four basic operations. The topological skeleton determines the connectivity of the manifold in the large. It is an important but difficult mathematical question to decide when two such skeletons are equivalent, i.e. when they represent two different ways of decomposition of the *same* continuum into elementary pieces. In the case of an n -dimensional closed manifold, the skeleton consists of a finite number of elements of rank 0, 1, 2, . . . , n (vertices, edges, . . .); these elements are to be represented by arbitrary symbols. An element of the i^{th} rank is bounded by certain elements of rank $i - 1$, and the skeleton is completely described by telling which element is bounded by which. The requirements such a skeleton has to meet, the properties which it possesses, and the question of equivalence constitute the subject of combinatorial topology.

[Topology has the peculiarity that questions belonging in its domain may under certain circumstances be decidable, even though the continua to which they are addressed may not be given exactly but only vaguely, as is always the case in reality. For instance, the topological skeleton of an undamaged brick is recognizable with certainty. Or an endless thread, which determines only approximately a curve in the exact sense of geometry, is definitely either knotted or not. Whenever the possible cases form a discrete manifold, an individual case can be fixed with absolute accuracy. Thus the rational analysis of continua proceeds in three steps: (1) morphology, which operates with vaguely circumscribed types of forms; (2) topology, which, guided by conspicuous singularities or even in

free construction, places into the manifold a vaguely localized but combinatorially exactly determined skeleton; and (3) geometry proper, whose ideal structures could only be carried with exactness into a real continuum after this has been spun over with a subdivision net of a fineness increasing *ad infinitum*. (Such geometrical properties of configurations in the continuum as are independent of the arbitrariness involved in the construction of the subdivision net may be conceived as based on a structural field spread over the continuum after the fashion of the metric field.) The significance which the idealizing geometry has for reality, in spite of the evident impossibility of fulfilling the above requirement for its application, will be discussed in Part II. The three steps described reveal the sensual-categorical ambivalence of geometry, which caused Plato to assign to geometrical configurations an intermediate position between ideas and sensory objects. For a more careful phenomenological analysis of the contrast between vagueness and exactness and of the limit concept, the reader may be referred to the work by O. Becker quoted at the end of Section 9. Carrying out the subdivision of the topological skeleton according to a fixed scheme implies the assumption that in dealing with a concretely given continuum we were not in error as to the topological character of the pieces generated by the first division. That is to say, we disregard the possibility that a more detailed scrutiny of a surface might disclose that, what we had considered an elementary piece, in reality has tiny handles attached to it which change the connectivity character of the piece, and that a microscope of ever greater magnification would reveal ever new topological complications of this type, *ad infinitum*.

The Riemann point of view allows, also for real space, topological conditions entirely different from those realized by Euclidean space. I believe that only on the basis of the freer and more general conception of geometry which had been brought out by the development of mathematics during the last century, and with an open mind for the imaginative possibilities which it has revealed, can a philosophically fruitful attack upon the space problem be undertaken. }

REFERENCES

- H. VON HELMHOLTZ, Über die Tatsachen, die der Geometrie zugrunde liegen (1868), *Wissenschaftliche Abhandlungen*, II, p. 618.
 B. RIEMANN, Über die Hypothesen, welche der Geometrie zugrunde liegen (1854), *Gesammelte mathematische Werke*, 1876 (separate edition with notes by H. Weyl, Berlin, 1923).
 H. WEYL, *Mathematische Analyse des Raumproblems*, Berlin, 1923.
 O. VERLEN, *Analysis situs*, Am. Math. Soc. Colloquium Publications, second edition, New York, 1931.
 P. ALEXANDROFF, *Einfachste Grundbegriffe der Topologie*, Berlin, 1932.