Reprinted in P. Benacerraf & H. Putnam (eds.), 1983. *Philosophy of Mathematics*, 2nd edition (pp. 41-65). Cambridge: Cambridge University Press.

Symposium on the foundations of mathematics

1. The logicist foundations of mathematics

RUDOLF CARNAP

The problem of the logical and epistemological foundations of mathematics has not yet been completely solved. This problem vitally concerns both mathematicians and philosophers, for any uncertainty in the foundations of the "most certain of all the sciences" is extremely disconcerting. Of the various attempts already made to solve the problem none can be said to have resolved every difficulty. These efforts, the leading ideas of which will be presented in these three papers, have taken essentially three directions: *Logicism,* the chief proponent of which is Russell; *Intuitionism,* advocated by Brouwer; and Hilbert's *Formalism*.

Since I wish to draw you a rough sketch of the salient features of the logicist construction of mathematics, I think I should not only point out those areas in which the logicist program has been completely or at least partly successful but also call attention to the difficulties peculiar to this approach. One of the most important questions for the foundations of mathematics is that of the relation between mathematics and logic. *Logicism* is the thesis that mathematics is reducible to logic, hence nothing but a part of logic. Frege was the first to espouse this view (1884). In their great work, *Principia Mathematica,* the English mathematicians A. N. Whitehead and B. Russell produced a systematization of logic from which they constructed mathematics.

We will split the logicist thesis into two parts for separate discussion:

1. The *concepts* of mathematics can be derived from logical concepts through explicit definitions.
2. The *theorems* of mathematics can be derived from logical axioms through purely logical deduction.

I. The derivation of mathematical concepts

To make precise the thesis that the concepts of mathematics are derivable from logical concepts, we must specify the logical concepts to be employed

The first three essays in this chapter form part of a symposium on the foundations of mathematics which appeared in Erkenntnis (1931), pp. 91–121. They were translated by Erna Putnam and Gerald J. Massey and appear here with the kind permission of Rudolf Carnap, Arend Heyting, and Klara von-Neumann Eckart. The last of these appears in A. H. Taub, ed., *John von Neumann Collected Works,* Vol. 2 (New York: Pergamon Press, 1961).

in the derivation. They are the following: In propositional calculus, which deals with the relations between unanalyzed sentences, the most important concepts are: the negation of a sentence p, 'not-p' (symbolized '$\sim p$'); the disjunction of two sentences, 'p or q' ('$p \vee q$'); the conjunction, 'p and q' ('$p \cdot q$'); and the implication, 'if p, then q' ('$p \supset q$'). The concepts of functional calculus are given in the form of functions, e.g., '$f(a)$' (read 'f of a') signifies that the property f belongs to the object a. The most important concepts of functional calculus are universality and existence: '$(x)f(x)$' (read 'for every x, f of x') means that the property f belongs to every object; '$(\exists x)f(x)$' (read 'there is an x such that f of x') means that f belongs to at least one object. Finally there is the concept of identity: '$a=b$' means that 'a' and 'b' are names of the same object.

Not all these concepts need be taken as undefined or primitive, for some of them are reducible to others. For example, '$p \vee q$' can be defined as '$\sim(\sim p \cdot \sim q)$' and '$(\exists x)f(x)$' as '$\sim (x) \sim f(x)$'. It is the logicist thesis, then, that the logical concepts just given suffice to define all mathematical concepts, that over and above them no specifically mathematical concepts are required for the construction of mathematics.

Already before Frege, mathematicians in their investigations of the interdependence of mathematical concepts had shown, though often without being able to provide precise definitions, that all the concepts of arithmetic are reducible to the natural numbers (i.e., the numbers 1, 2, 3, ... which are used in ordinary counting). Accordingly, the *main problem* which remained for logicism was to derive the natural numbers from logical concepts. Although Frege had already found a solution to this problem, Russell and Whitehead reached the same results independently of him and were subsequently the first to recognize the agreement of their work with Frege's. The crux of this solution is the correct recognition of the logical status of the natural numbers; they are logical attributes which belong, not to things, but to concepts. That a certain number, say 3, is the number of a concept means that three objects fall under it. We can express the very same thing with the help of the logical concepts previously given. For example, let '$2_m(f)$' mean that at least two objects fall under the concept f. Then we can define this concept as follows (where '$=_{Df}$' is the symbol for definition, read as "means by definition"):

$$2_m(f) =_{Df} (\exists x)(\exists y)[\sim(x=y) \cdot f(x) \cdot f(y)]$$

or in words: there is an x and there is a y such that x is not identical with y and f belongs to x and f belongs to y. In like manner, we define 3_m, 4_m, and so on. Then we define the number two itself thus:

$$2(f) =_{Df} 2_m(f) \cdot \sim 3_m(f)$$

or in words: at least two, but not at least three, objects fall under f. We can also define arithmetical operations quite easily. For example, we can define addition with the help of the disjunction of two mutually exclusive concepts. Furthermore, we can define the concept of natural number itself.

The derivation of the other kinds of number – i.e., the positive and negative numbers, the fractions, the real and the complex numbers – is accomplished, not in the usual way by adding to the domain of the natural numbers, but by the construction of a completely new domain. The natural numbers do not constitute a subset of the fractions but are merely correlated in obvious fashion with certain fractions. Thus the natural number 3 and the fraction 3/1 are not identical but merely correlated with one another. Similarly we must distinguish the fraction 1/2 from the real number correlated with it. In this paper, we will treat only the definition of the real numbers. Unlike the derivations of the other kinds of numbers which encounter no great difficulties, the derivation of the real numbers presents problems which, it must be admitted, neither logicism, intuitionism, nor formalism has altogether overcome.

Let us assume that we have already constructed the series of fractions (ordered according to magnitude). Our task, then, is to supply definitions of the real numbers based on this series. Some of the real numbers, the rationals, correspond in obvious fashion to fractions; the rest, the irrationals, correspond as Dedekind showed (1872) to "gaps" in the series of fractions. Suppose, for example, that we divide the (positive) fractions into two classes, the class of all whose square is less than 2, and the class comprising all the rest of the fractions. This division forms a "cut" in the series of fractions which corresponds to the irrational real number $\sqrt{2}$. This cut is called a "gap" since there is no fraction correlated with it. As there is no fraction whose square is two, the first or "lower" class contains no greatest member, and the second or "upper" class contains no least member. Hence, to every real number there corresponds a cut in the series of fractions, each irrational real number being correlated with a gap.

Russell developed further Dedekind's line of thought. Since a cut is uniquely determined by its "lower" class, Russell defined a real number as the lower class of the corresponding cut in the series of fractions. For example, $\sqrt{2}$ is defined as the class (or property) of those fractions whose square is less than two, and the rational real number 1/3 is defined as the class of all fractions smaller than the fraction 1/3. On the basis of these definitions, the entire arithmetic of the real numbers can be developed. This development, however, runs up against certain difficulties connected with so-called "impredicative definition," which we will discuss shortly.

The essential point of this method of introducing the real numbers is that they are *not postulated but constructed*. The logicist does not establish the existence of structures which have the properties of the real numbers by laying down axioms or postulates; rather, through explicit definitions, he produces logical constructions that have, by virtue of these definitions, the usual properties of the real numbers. As there are no "creative definitions," definition is not creation but only name-giving to something whose existence has already been established.

In similarly constructivistic fashion, the logicist introduces the rest of the concepts of mathematics, those of analysis (e.g., convergence, limit, continuity, differential, quotient, integral, etc.) and also those of set theory (notably the concepts of the transfinite cardinal and ordinal numbers). This "constructivist" method forms part of the very texture of logicism.

II. The derivation of the theorems of mathematics

The second thesis of logicism is that the *theorems of mathematics* are derivable from logical axioms through logical deduction. The requisite system of logical axioms, obtained by simplifying Russell's system, contains four axioms of propositional calculus and two of functional calculus. The rules of inference are a rule of substitution and a rule of implication (the *modus ponens* of ancient logic). Hilbert and Ackermann have used these same axioms and rules of inference in their system.

Mathematical predicates are introduced by explicit definitions. Since an explicit definition is nothing but a convention to employ a new, usually much shorter, way of writing something, the *definiens* or the new way of writing it can always be eliminated. Therefore, as every sentence of mathematics can be translated into a sentence which contains only the primitive logical predicates already mentioned, this second thesis can be restated thus: Every provable mathematical sentence is translatable into a sentence which contains only primitive logical symbols and which is provable in logic.

But the derivation of the theorems of mathematics poses certain difficulties for logicism. In the first place it turns out that some theorems of arithmetic and set theory, if interpreted in the usual way, require for their proof besides the logical axioms still other special axioms known as the *axiom of infinity* and the *axiom of choice* (or multiplicative axiom). The axiom of infinity states that for every natural number there is a greater one. The axiom of choice states that for every set of disjoint non-empty sets, there is (at least) one selection-set, i.e., a set that has exactly one member in common with each of the member sets. But we are not concerned here with the content of these axioms but with their logical

character. Both are existential sentences. Hence, Russell was right in hesitating to present them as logical axioms, for logic deals only with possible entities and cannot make assertions about whether something does or does not exist. Russell found a way out of this difficulty. He reasoned that since mathematics was also a purely formal science, it too could make only conditional, not categorical, statements about existence: if certain structures exist, then there also exist certain other structures whose existence follows logically from the existence of the former. For this reason he transformed a mathematical sentence, say S, the proof of which required the axiom of infinity, I, or the axiom of choice, C, into a conditional sentence; hence S is taken to assert not S, but $I \supset S$ or $C \supset S$, respectively. This conditional sentence is then derivable from the axioms of logic.

A greater difficulty, perhaps the greatest difficulty, in the construction of mathematics has to do with another axiom posited by Russell, the so-called *axiom of reducibility*, which has justly become the main bone of contention for the critics of the system of *Principia Mathematica*. We agree with the opponents of logicism that it is inadmissible to take it as an axiom. As we will discuss more fully later, the gap created by the removal of this axiom has certainly not yet been filled in an entirely satisfactory way. This difficulty is bound up with Russell's *theory of types* which we shall now briefly discuss.

We must distinguish between a "simple theory of types" and a "ramified theory of types." The latter was developed by Russell but later recognized by Ramsey to be an unnecessary complication of the former. If, for the sake of simplicity, we restrict our attention to one-place functions (properties) and abstract from many-place functions (relations), then type theory consists in the following classification of expressions into different "types": To type 0 belong the names of the objects ("individuals") of the domain of discourse (e.g., $a, b, \ldots$). To type 1 belong the properties of these objects (e.g., $f(a), g(a), \ldots$). To type 2 belong the properties of these properties (e.g., $F(f), G(f), \ldots$); for example, the concept $2(f)$ defined above belongs to this type. To type 3 belong the properties of properties of properties, and so on. The basic rule of type theory is that every predicate belongs to a determinate type and can be meaningfully applied only to expressions of the next lower type. Accordingly, sentences of the form $f(a), F(f), 2(f)$ are always meaningful, i.e., either true or false; on the other hand combinations like $f(g)$ and $f(F)$ are neither true nor false but meaningless. In particular, expressions like $f(f)$ or $\sim f(f)$ are meaningless, i.e., we cannot meaningfully say of a property either that it belongs to itself or that it does not. As we shall see, this last result is important for the elimination of the antinomies.

This completes our outline of the simple theory of types, which most proponents of modern logic consider legitimate and necessary. In his system, Russell introduced the ramified theory of types, which has not found much acceptance. In this theory the properties of each type are further subdivided into "orders." This division is based, not on the kind of objects to which the property belongs, but on the form of the definition which introduces it. Later we shall consider the reasons why Russell believed this further ramification necessary. Because of the introduction of the ramified theory of types, certain difficulties arose in the construction of mathematics, especially in the theory of real numbers. Many fundamental theorems not only could not be proved but could not even be expressed. To overcome this difficulty, Russell had to use brute force; i.e., he introduced the axiom of reducibility by means of which the different orders of a type could be reduced in certain respects to the lowest order of the type. The sole justification for this axiom was the fact that there seemed to be no other way out of this particular difficulty engendered by the ramified theory of types. Later Russell himself, influenced by Wittgenstein's sharp criticism, abandoned the axiom of reducibility in the second edition of *Principia Mathematica* (1925). But, as he still believed that one could not get along without the ramified theory of types, he despaired of the situation. Thus we see how important it would be, not only for logicism but for any attempt to solve the problems of the foundations of mathematics, to show that the simple theory of types is sufficient for the construction of mathematics out of logic. A young English mathematician and pupil of Russell, Ramsey (who unfortunately died this year, i.e., 1930), in 1926 made some efforts in this direction which we will discuss later.

III. The problem of impredicative definition

To ascertain whether the simple theory of types is sufficient or must be further ramified, we must first of all examine the reasons which induced Russell to adopt this ramification in spite of its most undesirable consequences. There were two closely connected reasons: the necessity of eliminating the logical antinomies and the so-called "vicious circle" principle. We call "logical antinomies" the contradictions which first appeared in set theory (as so-called "paradoxes") but which Russell showed to be common to all logic. It can be shown that these contradictions arise in logic if the theory of types is not presupposed. The simplest antinomy is that of the concept "impredicable." By definition a property is "impredicable" if it does not belong to itself. Now is the property "impredicable" itself impredicable? If we assume that it is, then since it belongs to

itself it would be, according to the definition of "impredicable," not impredicable. If we assume that it is not impredicable, then it does not belong to itself and hence, according to the definition of "impredicable," is impredicable. According to the law of excluded middle, it is either impredicable or not, but both alternatives lead to a contradiction. Another example is Grelling's antinomy of the concept "heterological." Except that it concerns predicates rather than properties, this antinomy is completely analogous to the one just described. By definition, a predicate is "heterological" if the property designated by the predicate does not belong to the predicate itself. (For example, the word 'monosyllabic' is heterological, for the word itself is not monosyllabic.) Obviously both the assumption that the word 'heterological' is itself heterological as well as the opposite assumption lead to a contradiction. Russell and other logicians have constructed numerous antinomies of this kind.

Ramsey has shown that there are two completely different kinds of antinomies. Those belonging to the first kind can be expressed in logical symbols and are called "logical antinomies" (in the narrower sense). The "impredicable" antinomy is of this kind. Ramsey has shown that this kind of antinomy is eliminated by the simple theory of types. The concept "impredicable," for example, cannot even be defined if the simple theory of types is presupposed, for an expression of the form, a property does not belong to itself ($\sim f(f)$), is not well-formed, and meaningless according to that theory.

Antinomies of the second kind are known as "semantical" or "epistemological" antinomies. They include our previous example, "heterological," as well as the antinomy, well-known to mathematicians, of the smallest natural number which cannot be defined in German with fewer than 100 letters. Ramsey has shown that antinomies of this second kind cannot be constructed in the symbolic language of logic and therefore need not be taken into account in the construction of mathematics from logic. The fact that they appear in word languages led Russell to impose certain restrictions on logic in order to eliminate them, viz., the ramified theory of types. But perhaps their appearance is due to some defect of our ordinary word language.

Since antinomies of the first kind are already eliminated by the simple theory of types and those of the second kind do not appear in logic, Ramsey declared that the ramified theory of types and hence also the axiom of reducibility were superfluous.

Now what about Russell's second reason for ramifying the theory of types, viz., the vicious circle principle? This principle, that "no whole may contain parts which are definable only in terms of that whole", may also be called an "injunction against impredicative definition." A definition is said to be "impredicative" if it defines a concept in terms of a totality to which the concept belongs. (The concept "impredicative" has nothing to do with the aforementioned pseudo concept "impredicable.") Russell's main reason for laying down this injunction was his belief that antinomies arise when it is violated. From a somewhat different standpoint Poincaré before, and Weyl after, Russell also rejected impredicative definition. They pointed out that an impredicatively defined concept was meaningless because of the circularity in its definition. An example will perhaps make the matter clearer:

We can define the concept "inductive number" (which corresponds to the concept of natural number including zero) as follows: A number is said to be "inductive" if it possesses all the hereditary properties of zero. A property is said to be "hereditary" if it always belongs to the number $n+1$ whenever it belongs to the number n. In symbols,

$$\mathrm{Ind}(x) =_{\mathrm{Df}} (f)[(\mathrm{Her}(f) \cdot f(0)) \supset f(x)]$$

To show that this definition is circular and useless, one usually argues as follows: In the *definiens* the expression '(f)' occurs, i.e., "for all properties (of numbers)". But since the property "inductive" belongs to the class of all properties, the very property to be defined already occurs in a hidden way in the *definiens* and thus is to be defined in terms of itself, an obviously inadmissible procedure. It is sometimes claimed that the meaninglessness of an impredicatively defined concept is seen most clearly if one tries to establish whether the concept holds in an individual case. For example, to ascertain whether the number three is inductive, we must, according to the definition, investigate whether every property which is hereditary and belongs to zero also belongs to three. But if we must do this for every property, we must also do it for the property "inductive" which is also a property of numbers. Therefore, in order to determine whether the number three is inductive, we must determine among other things whether the property "inductive" is hereditary, whether it belongs to zero, and finally – this is the crucial point – whether it belongs to three. But this means that it would be impossible to determine whether three is an inductive number.

Before we consider how Ramsey tried to refute this line of thought, we must get clear about how these considerations led Russell to the ramified theory of types. Russell reasoned in this way: Since it is inadmissible to define a property in terms of an expression which refers to "all properties," we must subdivide the properties (of type 1): To the "first order" belong those properties in whose definition the expression 'all properties' does not occur; to the "second order" those in whose definition the expression 'all properties of the first order' occurs; to the "third order"

those in whose definition the expression 'all properties of the second order' occurs, and so on. Since the expression 'all properties' without reference to a determinate order is held to be inadmissible, there never occurs in the definition of a property a totality to which it itself belongs. The property "inductive," for example, is defined in this no longer impredicative way: A number is said to be "inductive" if it possess all the hereditary properties of the first order which belong to zero.

But the ramified theory of types gives rise to formidable difficulties in the treatment of the real numbers. As we have already seen, a real number is defined as a class, or what comes to the same thing, as a property of fractions. For example, we say that $\sqrt{2}$ is defined as the class or property of those fractions whose square is less than two. But since the expression 'for all properties' without reference to a determinate order is inadmissible under the ramified theory of types, the expression 'for all real numbers' cannot refer to all real numbers without qualification but only to the real numbers of a determinate order. To the first order belong those real numbers in whose definition an expression of the form 'for all real numbers' does not occur; to the second order belong those in whose definition such an expression occurs, but this expression must be restricted to "all real numbers of the first order," and so on. Thus there can be neither an admissible definition nor an admissible sentence which refers to all real numbers without qualification.

But as a consequence of this ramification, many of the most important definitions and theorems of real number theory are lost. Once Russell had recognized that his earlier attempt to overcome it, viz., the introduction of the axiom of reducibility, was itself inadmissible, he saw no way out of this difficulty. The *most difficult problem* confronting contemporary studies in the foundations of mathematics is this: How can we develop logic if, on the one hand, we are to avoid the danger of the meaninglessness of impredicative definitions and, on the other hand, are to reconstruct satisfactorily the theory of real numbers?

IV. Attempt at a solution

Ramsey (1926a) outlined a construction of mathematics in which he courageously tried to resolve this difficulty by declaring the forbidden impredicative definitions to be perfectly admissible. They contain, he contended, a circle but the circle is harmless, not vicious. Consider, he said, the description 'the tallest man in this room'. Here we describe something in terms of a totality to which it itself belongs. Still no one thinks this description inadmissible since the person described already exists and is only singled out, not created, by the description. Ramsey believed that the same considerations applied to properties. The totality of properties already exists in itself. That we men are finite beings who cannot name individually each of infinitely many properties but can describe some of them only with reference to the totality of all properties is an empirical fact that has nothing to do with logic. For these reasons Ramsey allows impredicative definition. Consequently, he can both get along with the simple theory of types and still retain all the requisite mathematical definitions, particularly those needed for the theory of the real numbers.

Although this happy result is certainly tempting, I think we should not let ourselves be seduced by it into accepting Ramsey's basic premise; viz., that the totality of properties already exists before their characterization by definition. Such a conception, I believe, is not far removed from a belief in a platonic realm of ideas which exist in themselves, independently of *if* and *how* finite human beings are able to think them. I think we ought to hold fast to Frege's dictum that, in mathematics, only that may be taken to exist whose existence has been proved (and he meant proved in finitely many steps). I agree with the intuitionists that the finiteness of every logical-mathematical operation, proof, and definition is not required because of some accidental empirical fact about man but is required by the very nature of the subject. Because of this attitude, intuitionist mathematics has been called "anthropological mathematics." It seems to me that, by analogy, we should call Ramsey's mathematics "theological mathematics," for when he speaks of the totality of properties he elevates himself above the actually knowable and definable and in certain respects reasons from the standpoint of an infinite mind which is not bound by the wretched necessity of building every structure step by step.

We may now rephrase our crucial question thus: Can we have Ramsey's result without retaining his absolutist conceptions? His result was this: Limitation to the simple theory of types and retention of the possibility of definitions for mathematical concepts, particularly in real number theory. We can reach this result if, like Ramsey, we allow impredicative definition, but can we do this without falling into his conceptual absolutism? I will try to give an affirmative answer to this question.

Let us go back to the example of the property "inductive" for which we gave an impredicative definition:

$$\mathrm{Ind}(x) =_{\mathrm{Df}} (f)[(\mathrm{Her}(f)\cdot f(0)) \supset f(x)]$$

Let us examine once again whether the use of this definition, i.e., establishing whether the concept holds in an individual case or not, really

leads to circularity and is therefore impossible. According to this definition, that the number two is inductive means:

$$(f)[(\mathrm{Her}(f)\cdot f(0)) \supset f(2)]$$

in words: Every property f which is hereditary and belongs to zero belongs also to two. How can we verify a universal statement of this kind? If we had to examine every single property, an unbreakable circle would indeed result, for then we would run headlong against the property "inductive." Establishing whether something had it would then be impossible in principle, and the concept would therefore be meaningless. But the verification of a universal logical or mathematical sentence does not consist in running through a series of individual cases, for impredicative definitions usually refer to infinite totalities. The belief that we must run through all the individual cases rests on a confusion of "numerical" generality, which refers to objects already given, with "specific" generality (cf. Kaufmann 1930). We do not establish specific generality by running through individual cases but by logically deriving certain properties from certain others. In our example, that the number two is inductive means that the property "belonging to two" follows logically from the property "being hereditary and belonging to zero." In symbols, '$f(2)$' can be derived for an arbitrary f from '$\mathrm{Her}(f)\cdot f(0)$' by logical operations. This is indeed the case. First, the derivation of '$f(0)$' from '$\mathrm{Her}(f)\cdot f(0)$' is trivial and proves the inductiveness of the number zero. The remaining steps are based on the definition of the concept "hereditary":

$$\mathrm{Her}(f) =_{\mathrm{Df}} (n)[f(n) \supset f(n+1)]$$

Using this definition, we can easily show that '$f(0+1)$' and hence '$f(1)$' are derivable from '$\mathrm{Her}(f)\cdot f(0)$' and thereby prove that the number one is inductive. Using this result and our definition, we can derive '$f(1+1)$' and hence '$f(2)$' from '$\mathrm{Her}(f)\cdot f(0)$', thereby showing that the number two is inductive. We see then that the definition of inductiveness, although impredicative, does not hinder its utility. That proofs that the defined property obtains (or does not obtain) in individual cases can be given shows that the definition is meaningful. If we reject the belief that it is necessary to run through individual cases and rather make it clear to ourselves that the complete verification of a statement about an arbitrary property means nothing more than its logical (more exactly, tautological) validity for an arbitrary property, we will come to the conclusion that impredicative definitions are logically admissible. If a property is defined impredicatively, then establishing whether or not it obtains in an individual case may, under certain circumstances, be difficult, or it may even be impossible if there is no solution to the decision problem for that logical system. But in no way does impredicativeness make such decisions impossible in principle for all cases. If the theory just sketched proves feasible, logicism will have been helped over its greatest difficulty, which consists in steering a safe course between the Scylla of the axiom of reducibility and the Charybdis of the allocation of the real numbers to different orders.

Logicism as here described has several features in common both with intuitionism and with formalism. It shares with intuitionism a constructivistic tendency with respect to definition, a tendency which Frege also emphatically endorsed. A concept may not be introduced axiomatically but must be constructed from undefined, primitive concepts step by step through explicit definitions. The admission of impredicative definitions seems at first glance to run counter to this tendency, but this is only true for constructions of the form proposed by Ramsey. Like the intuitionists, we recognize as properties only those expressions (more precisely, expressions of the form of a sentence containing one free variable) which are constructed in finitely many steps from undefined primitive properties of the appropriate domain according to determinate rules of construction. The difference between us lies in the fact that we recognize as valid not only the rules of construction which the intuitionists use (the rules of the so-called "strict functional calculus"), but in addition, permit the use of the expression 'for all properties' (the operations of the so-called "extended functional calculus").

Further, logicism has a methodological affinity with formalism. Logicism proposes to construct the logical-mathematical system in such a way that, although the axioms and rules of inference are chosen with an interpretation of the primitive symbols in mind, nevertheless, *inside the system* the chains of deductions and of definitions are carried through formally as in a pure calculus, i.e., without reference to the meaning of the primitive symbols.

2. The intuitionist foundations of mathematics

[Die intuitionistische Grundlegung der Mathematik]

AREND HEYTING

The intuitionist mathematician proposes to do mathematics as a natural function of his intellect, as a free, vital activity of thought. For him, mathematics is a production of the human mind. He uses language, both

natural and formalized, only for communicating thoughts, i.e., to get others or himself to follow his own mathematical ideas. Such a linguistic accompaniment is not a representation of mathematics; still less is it mathematics itself.

It would be most in keeping with the active attitude of the intuitionist to deal at once with the construction of mathematics. The most important building block of this construction is the concept of unity which is the architectonic principle on which the series of integers depends. The integers must be treated as units which differ from one another only by their place in this series. Since in his *Logischen Grundlagen der exakten Wissenschaften* Natorp has already carried out such an analysis, which in the main conforms tolerably well to the intuitionist way of thinking, I will forego any further analysis of these concepts. But I must still make one remark which is essential for a correct understanding of our intuitionist position: we do not attribute an existence independent of our thought, i.e., a transcendental existence, to the integers or to any other mathematical objects. Even though it might be true that every thought refers to an object conceived to exist independently of it, we can nevertheless let this remain an open question. In any event, such an object need not be completely independent of human thought. Even if they should be independent of individual acts of thought, mathematical objects are by their very nature dependent on human thought. Their existence is guaranteed only insofar as they can be determined by thought. They have properties only insofar as these can be discerned in them by thought. But this possibility of knowledge is revealed to us only by the act of knowing itself. Faith in transcendental existence, unsupported by concepts, must be rejected as a means of mathematical proof. As I will shortly illustrate more fully by an example, this is the reason for doubting the law of excluded middle.

Oskar Becker has dealt thoroughly with the problems of mathematical existence in his book on that subject. He has also uncovered many connections between these questions and the most profound philosophical problems.

We return now to the construction of mathematics. Although the introduction of the fractions as pairs of integers does not lead to any basic difficulties, the definition of the irrational numbers is another story. A real number is defined according to Dedekind by assigning to every rational number either the predicate 'Left' or the predicate 'Right' in such a way that the natural order of the rational numbers is preserved. But if we were to transfer this definition into intuitionist mathematics in exactly this form, we would have no guarantee that Euler's constant C is a real number. We do not need the definition of C. It suffices to know that this definition amounts to an algorithm which permits us to enclose C within an arbitrarily small rational interval. (A rational interval is an interval whose end points are rational numbers. But, as absolutely no ordering relations have been defined between C and the rational numbers, the word 'enclose' is obviously vague for practical purposes. The practical question is that of computing a series of rational intervals each of which is contained in the preceding one in such a way that the computation can always be continued far enough so that the last interval is smaller than an arbitrarily given limit.) But this algorithm still provides us with no way of deciding for an arbitrary rational number A whether it lies left or right of C or is perhaps equal to C. But such a method is just what Dedekind's definition, interpreted intuitionistically, would require.

The usual objection against this argument is that it does not matter whether or not this question can be decided, for, if it is not the case that $A=C$, then either $A<C$ or $A>C$, and this last alternative is decided after a finite, though perhaps unknown, number of steps N in the computation of C. I need only reformulate this objection to refute it. It can mean only this: either there exists a natural number N such that after N steps in the computation of C it turns out that $A<C$ or $A>C$; or there is no such N and hence, of course, $A=C$. But, as we have seen, the existence of N signifies nothing but the possibility of actually producing a number with the requisite property, and the non-existence of N signifies the possibility of deriving a contradiction from this property. Since we do not know whether or not one of these possibilities exists, we may not assert that N either exists or does not exist. In this sense, we can say that the law of excluded middle may not be used here.

In its original form, then, Dedekind's definition cannot be used in intuitionist mathematics. Brouwer, however, has improved it in the following way: Think of the rational numbers enumerated in some way. For the sake of simplicity, we restrict ourselves to the numbers in the closed unit interval and take always as our basis the following enumeration:

$$\text{(A)} \quad 0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots$$

A real number is determined by a cut in the series (A); i.e., by a rule which assigns to each rational number in the series either the predicate 'Left' or the predicate 'Right' in such a way that the natural order of the rational numbers is preserved. At each step, however, we permit one individual number to be left out of this mapping. For example, let the rule be so formed that the series of predicates begins this way:

$$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \ldots$$

L, R, L, L, ?, L,

Here 2/3 is temporarily left out of the mapping. We need not know whether or not the predicate for 2/3 is ever determined. But it is also a possibility that 3/4 should become a new excluded number and hence that 2/3 would receive the predicate 'Left'.

It is easy to give a cut for Euler's constant. Let d_n be the smallest difference between two successive numbers in the first n numbers of (A). Now if we compute C far enough to get a rational interval i which is smaller than d_n, then at most one of these n numbers can fall within i. If there is such a number, it becomes the excluded number for the cut. Thus, we can see how closely Brouwer's definition is related to the actual computation of a real number.

We can now take an important step forward. We can drop the requirement that the series of predicates be determined to infinity by a rule. It suffices if the series is determined step by step in some way, e.g., by free choices. I call such sequences "infinitely proceeding." Thus the definition of real numbers is extended to allow infinitely proceeding sequences in addition to rule-determined sequences. Before discussing this new definition in detail, we will give a simple example. We begin with this "Left-Right" choice-sequence:

$$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5},$$

L, R, L, L, R, L, R, L, L,

Here the question about which predicate 3/5 receives cannot be answered yet, for it must still be decided which predicate to give it. The question about the predicate which 4/5 receives, on the other hand, can be answered now by 'Right,' since that choice would hold for every possible continuation of the sequence. In general, only those questions about an infinitely proceeding sequence which refer to every possible continuation of the sequence are susceptible of a determinate answer. Other questions, like the foregoing about the predicate for 3/5, must therefore be regarded as meaningless. Thus choice-sequences supplant, not so much the individual rule-determined sequences, but rather the totality of all possible rules. A "Left-Right" choice-sequence, the freedom of choice for which is limited only by the conditions which result from the natural order of the rational numbers, determines not just one real number but the spread of all real numbers or the continuum. Whereas we ordinarily think of each real number as individually defined and only afterwards think of them all together, we here define the continuum as a totality. If we restrict this freedom of choice by rules given in advance, we obtain spreads of real numbers. For example, if we prescribe that the sequence begin in the way we have just written it, we define the spread of real numbers between 1/2 and 2/3. An infinitely proceeding sequence gradually becomes a rule-determined sequence when more and more restrictions are placed on the freedom of choice.

We have used the word 'spread' exactly in Brouwer's sense. His definition of a spread is a generalization of this notion. In addition to choice-sequences, Brouwer treats sequences which are formed from choice-sequences by mapping rules. A spread involves two rules. The first rule states which choices of natural numbers are allowed after a determinate finite series of permitted choices has been made. The rule must be so drawn that at least one new permissible choice is known after each finite series of permitted choices has been made. The natural order of the rational numbers is an example of such a rule for our "Left-Right" sequence previously given. The second rule involved in a spread assigns a mathematical object to each permissible choice. The mathematical object may, of course, depend also on choices previously made. Thus it is permissible to terminate the mapping at some particular number and to assign nothing to subsequent choices. A sequence which results from a permissible choice-sequence by a mapping-rule is called an "element" of the spread.

To bring our previous example of the spread of real numbers between 1/2 and 2/3 under this general definition, we will replace the predicates 'Left', 'Right', and 'temporarily undetermined', by 1, 2, and 3; and we will derive the rule for permissible choices from the natural order of the rational numbers and from the requirement that the sequence begin in a particular way; and we will take identity for the mapping-rule.

A spread is not the sum of its elements (this statement is meaningless unless spreads are regarded as existing in themselves). Rather, a spread is identified with its defining rules. Two elements of a spread are said to be equal if equal objects exist at the nth place in both for every n. Equality of elements of a spread, therefore, does not mean that they are the same element. To be the same, they would have to be assigned to the same spread by the same choice-sequence. It would be impractical to call two mathematical objects equal only if they are the same object. Rather, every kind of object must receive its own definition of equality.

Brouwer calls "species" those spreads which are defined, in classical terminology, by a characteristic property of their members. A species,

like a spread, is not regarded as the sum of its members but is rather identified with its defining property. Impredicative definitions are made impossible by the fact, which intuitionists consider self-evident, that only previously defined objects may occur as members of a species. There results, consequently, a step-by-step introduction of species. The first level is made up of those spread-species whose defining property is identity with an element of a particular spread. Hence, to every spread M there corresponds the spread-species of those spread-elements which are identical with some element of M.[1] A species of the first order can contain spread-elements and spread-species. In addition, a species of the second order contains species of the first order as members, and so on.

The introduction of infinitely proceeding sequences is not a necessary consequence of the intuitionist approach. Intuitionist mathematics could be constructed without choice-sequences. But the following set-theoretic theorem about the continuum shows how much mathematics would thereby be impoverished. This theorem will also serve as an example of an intuitionist reasoning process.

Let there be a rule assigning to each real number a natural number as its correlate. Assume that the real numbers a and b have different correlates, e.g., 1 and 2. Then, by a simple construction, we can determine a third number c which has the following property: in every neighborhood of c, no matter how small, there is a mapped number other than c; i.e., every finite initial segment of the cut which defines c can be continued so as to get a mapped number other than c. We define the number d by a choice-sequence thus: we begin as with c but we reserve the freedom to continue at an arbitrary choice in a way different from that for c. Obviously the correlate of d is not determined after any previously known finite number of choices. Accordingly, no definite correlate is assigned to d. But this conclusion contradicts our premise that every real number has a correlate. Our assumption that the two numbers a and b have different correlates is thus shown to be contradictory. And, since two natural numbers which cannot be distinguished are the same number, we have the following theorem: if every real number is assigned a correlate, then all the real numbers have the same correlate.

As a special result, we have: if a continuum is divided into two subspecies in such a way that every member belongs to one and only one of these subspecies, then one of the subspecies is empty and other other is identical with the continuum.

The unit continuum, for example, cannot be subdivided into the species of numbers between 0 and 1/2 and the species of numbers between

[1]This definition of spread-species is taken from a communication of Professor Brouwer.

and which consists basically in constructing successively certain combinations of primitive symbols which are considered "correct" or "proved." This construction-procedure, moreover, is "finitary" and directly constructive. To see clearly the essential difference between the occasionally non-constructive handling of the "content" of mathematics (real numbers and the like) and the always constructive linking of the steps in a proof, consider this example: Assume that there exists a classical proof of the existence of a real number x with a certain very complicated and deep-seated property $E(x)$. Then it may well happen that, from this proof, we can in no way derive a procedure for constructing an x such that $E(x)$. (We shall give an example of such a proof in a moment.) On the other hand, if the proof somehow violated the conventions of mathematical inference, i.e., if it contained an error, we could, of course, find this error by a finitary process of checking. In other words, although the content of a classical mathematical sentence cannot always (i.e., generally) be finitely verified, the formal way in which we arrive at the sentence can be. Consequently, if we wish to prove the validity of classical mathematics, which is possible in principle only by reducing it to the *a priori* valid finitistic system (i.e., Brouwer's system), then we should investigate, not statements, but methods of proof. We must regard classical mathematics as a combinatorial game played with the primitive symbols, and we must determine in a finitary combinatorial way to which combinations of primitive symbols the construction methods or "proofs" lead.

As we promised, we now produce an example of a non-constructive existence proof. Let $f(x)$ be a function which is linear from 0 to 1/3, from 1/3 to 2/3, from 2/3 to 1, and so on. Let

$$f(0)=-1; \quad f\left(\frac{1}{3}\right)=-\sum_{n=1}^{n=\infty}\frac{\epsilon_{2n}}{2^n}; \quad f\left(\frac{2}{3}\right)=\sum_{n=1}^{n=\infty}\frac{\epsilon_{2n}}{2^n}; \quad \text{and} \quad f(1)=1$$

ϵ_n is defined as follows: if $2k$ is the sum of two prime numbers, then $\epsilon_k=0$; otherwise $\epsilon_k=1$. Obviously $f(x)$ is continuous and calculable with arbitrary accuracy at any point x. Since $f(0)<0$ and $f(1)>0$, there exists an x, where $0\leqslant x\leqslant 1$, such that $f(x)=0$. (In fact we readily see that $1/3\leqslant x\leqslant 2/3$.) However the task of finding a root with an accuracy greater than $\pm 1/6$ encounters formidable difficulties. Given the present state of mathematics, these difficulties are insuperable, for if we could find such a root, then we could predict with certitude the existence of a root $<2/3$ or $>1/3$, according as its approximate value were $\leqslant 1/2$ or $\geqslant 1/2$, respectively. The former case (where the approximate value of the root is $\leqslant 1/2$) excludes both that $f(1/3)<0$ and that $f(2/3)=0$; the latter case (where the approximate value of the root $\geqslant 1/2$) excludes

both that $f(1/3)=0$ and that $f(2/3)>0$. In other words, in the former case the value of ϵ_n must be 0 for all even n but not for all odd n; in the latter case the value of ϵ_n must be 0 for all odd n but not for all even n. Hence we would have proved that Goldbach's famous conjecture (that $2n$ is always the sum of two prime numbers), instead of holding universally, must already fail to hold for odd n in the former case and for even n in the latter. But no mathematician today can supply a proof for either case, since no one can find the solution of $f(x)=0$ more accurately than with an error of 1/6. (With an error of 1/6, 1/2 would be an approximate value of the root, for the root lies between 1/3 and 2/3, i.e., between $1/2-1/6$ and $1/2+1/6$.)

II

Accordingly, the tasks which Hilbert's theory of proof must accomplish are these:

1. To enumerate all the symbols used in mathematics and logic. These symbols, called "primitive symbols," include the symbols '$\sim$' and '$\rightarrow$' (which stand for "negation" and "implication" respectively).
2. To characterize unambiguously all the combinations of these symbols which represent statements classified as "meaningful" in classical mathematics. These combinations are called "formulas." (Note that we said only "meaningful," not necessarily "true." '$1+1=2$' is meaningful but so is '$1+1=1$', independently of the fact that one is true and the other false. On the other hand, combinations like '$1+\rightarrow=1$' and '$++1=\rightarrow$' are meaningless.)
3. To supply a construction procedure which enables us to construct successively all the formulas which correspond to the "provable" statements of classical mathematics. This procedure, accordingly, is called "proving."
4. To show (in a finitary combinatorial way) that those formulas which correspond to statements of classical mathematics which can be checked by finitary arithmetical methods can be proved (i.e., constructed) by the process described in (3) if and only if the check of the corresponding statement shows it to be true.

To accomplish tasks 1–4 would be to establish the validity of classical mathematics as a short-cut method for validating arithmetical statements whose elementary validation would be much too tedious. But since this is in fact the way we use mathematics, we would at the same time sufficiently establish the empirical validity of classical mathematics.

We should remark that Russell and his school have almost completely accomplished tasks 1–3. In fact, the formalization of logic and mathematics suggested by tasks 1–3 can be carried out in many different ways. The real problem, then, is (4).

In connection with (4) we should note the following: If the "effective check" of a numerical formula shows it to be false, then from that formula we can derive a relation $p=q$ where p and q are two different, effectively given numbers. Hence (according to task 3) this would give us a formal proof of $p=q$ from which we could obviously get a proof of $1=2$. Therefore, the sole thing we must show to establish (4) is the formal unprovability of $1=2$; i.e., we need to investigate only this one particular false numerical relation. The unprovability of the formula $1=2$ by the methods described in (3) is called "consistency." The real problem, then, is that of finding a finitary combinatorial proof of consistency.

III

To be able to indicate the direction which a proof of consistency takes, we must consider formal proof procedure – as in (3) – a little more closely. It is defined as follows:

3_1. Certain formulas, characterized in an unambiguous and finitary way, are called "axioms." Every axiom is considered proved.

3_2. If a and b are two meaningful formulas, and if a and $a \rightarrow b$ have both been proved, then b also has been proved.

Note that, although (3_1) and (3_2) do indeed enable us to write down successively all provable formulas, still this process can never be finished. Further, (3_1) and (3_2) contain no procedure for deciding whether a given formula e is provable. As we cannot tell in advance which formulas must be proved successively in order ultimately to prove e, some of them might turn out to be far more complicated and structurally quite different from e itself. (Anyone who is acquainted, for example, with analytic number theory knows just how likely this possibility is, especially in the most interesting parts of mathematics.) But the problem of deciding the provability of an arbitrarily given formula by means of a (naturally finitary) general procedure, i.e., the so-called decision problem for mathematics, is much more difficult and complex than the problem discussed here.

As it would take us too far afield to give the axioms which are used in classical mathematics, the following remarks must suffice to characterize them. Although infinitely many formulas are regarded as axioms (for example, by our definition each of the formulas $1=1$, $2=2$, $3=3, \ldots$ is

an axiom), they are nevertheless constructed from finitely many schemata by substitution in this manner: 'If a, b, and c are formulas, then $(a \rightarrow b) \rightarrow ((b \rightarrow c) \rightarrow (a \rightarrow c))$ is an axiom', and the like.

Now if we could succeed in producing a class R of formulas such that

(α) Every axiom belongs to R,
(β) If a and $a \rightarrow b$ belong to R, then b also belongs to R,
(γ) '$1=2$' does not belong to R,

then we would have proved consistency, for according to (α) and (β) every proved formula obviously must belong to R, and according to (γ), $1=2$ must therefore be unprovable. The actual production of such a class at this time is unthinkable, however, for it poses difficulties comparable to those raised by the decision problem. But the following remark leads from this problem to a much simpler one: If our system were inconsistent, then there would exist a proof of $1=2$ in which only a finite number of axioms are used. Let the set of these axioms be called M. Then the axiom system M is already inconsistent. Hence the axiom system of classical mathematics is certainly consistent if every finite subsystem thereof is consistent. And this is surely the case if, for every finite set of axioms M, we can give a class of formulas R_M which has the following properties:

(α) Every axiom of M belongs to R_M.
(β) If a and $a \rightarrow b$ belong to R_M, then b also belongs to R_M.
(γ) $1=2$ does not belong to R_M.

This problem is not connected with the (much too difficult) decision problem, for R_M depends only on M and plainly says nothing about provability (with the help of all the axioms). It goes without saying that we must have an effective, finitary procedure for constructing R_M (for every effectively given finite set of axioms M) and that the proofs of (α), (β), and (γ) must also be finitary.

Although the consistency of classical mathematics has not yet been proved, such a proof has been found for a somewhat narrower mathematical system. This system is closely related to a system which Weyl proposed before the conception of the intuitionist system. It is substantially more extensive than the intuitionist system but narrower than classical mathematics (for bibliographical material, see Weyl 1927).

Thus Hilbert's system has passed the first test of strength: the validity of a non-finitary, not purely constructive mathematical system has been established through finitary constructive means. Whether someone will succeed in extending this validation to the more difficult and more important system of classical mathematics, only the future will tell.