BIBLIOTHECA MATHEMATICA

*A Series of Monographs on Pure and Applied Mathematics*

*Volume I*

*

*Edited with the cooperation of*

THE „MATHEMATISCH CENTRUM"

and

THE „WISKUNDIG GENOOTSCHAP"

at Amsterdam

*



# INTRODUCTION TO METAMATHEMATICS

BY

STEPHEN COLE KLEENE †

PROFESSOR OF MATHEMATICS AT THE UNIVERSITY OF WISCONSIN (MADISON, WIS., U.S.A.)

WOLTERS-NOORDHOFF PUBLISHING – GRONINGEN

NORTH-HOLLAND PUBLISHING COMPANY – AMSTERDAM

NEW YORK

The following riddle also turns upon the paradox. A traveller has fallen among cannibals. They offer him the opportunity to make a statement, attaching the conditions that if his statement be true, he will be boiled, and if it be false, he will be roasted. What statement should he make? (A form of this riddle occurs in Cervantes' "Don Quixote" (1605), II, 51.)

**§ 12. First inferences from the paradoxes.** The reader may try his hand at solving the paradoxes. In the half century since the problem has been open, no solution has been found which is universally agreed upon.

The simplest kind of solution would be to locate a specific fallacy, like a mistake in a student's algebra exercise or geometry proof, with nothing else needing to be changed.

Ideas for solving the paradoxes in this sense come to mind on first considering them. One may propose that the error in the paradoxes (A) — (C) consists in using too large sets, such as the set of all sets or the set of all cardinal numbers; or in permitting sets to be considered as members of themselves, which again argues against the set of all sets. These suggestions are not necessarily wrong, but they are not after all simple. They leave us the problem of refounding set theory on a drastically altered basis, the details of which are not fully implicit in the suggestions. For example, if we ban the set of all cardinal numbers, we are unable to introduce the set of the natural numbers, unless we already know that they are not all the cardinal numbers; and the same difficulty will arise at higher stages. If we ban the set of all sets, we find ourselves in conflict with Cantor's definition of set. In order to have set theory at all, we must have theorems about all sets, and all sets then constitute a set under Cantor's definition. If not so, we must say what other definition of set we shall use instead, or we must supplement Cantor's definition with some further criterion to determine when a collection of objects as described in his definition shall constitute a set (Skolem 1929-30).

AXIOMATIC SET THEORY. Reconstructions of set theory can be given, placing around the notion of set as few restrictions to exclude too large sets as appear to be required to forestall the known antinomies. Since the free use of our conceptions in constructing sets under Cantor's definition led to disaster, the notions of set theory are governed by axioms, like those governing 'point' and 'line' in Euclidean plane geometry. The first system of *axiomatic set theory* was Zermelo's (1908). Refinements in the axiomatic treatment of sets are due to Fraenkel (1922, 1925), Skolem (1922-3, 1929), von Neumann (1925, 1928), Bernays (1937-48), and others. Analysis can be founded on the basis of axiomatic set theory, which perhaps is the simplest basis set up since the paradoxes for the deduction of existing mathematics. Some very interesting discoveries have been made in connection with axiomatic set theory, notably by Skolem (1922-3; cf. § 75 below) and Gödel (1938, 1939, 1940).



THE BROADER PROBLEM OF FOUNDATIONS. Assuming that the paradoxes are avoided in the axiomatization of set theory — and of this the only assurance we have is the negative one that so far none have been encountered — does it constitute a full solution of the problem posed by the paradoxes?

In the case of geometry, mathematicians have recognized since the discovery of a non-Euclidean geometry that more than one kind of space is possible. Axiom systems serve to single out one or another kind of space, or certain common features of several spaces, for the geometer to study. A contradiction arising in a formal axiomatic theory can mean simply that an unrealizable combination of features has been postulated.

But in the case of arithmetic and analysis, theories culminating in set theory, mathematicians prior to the current epoch of criticism generally supposed that they were dealing with systems of objects, set up genetically, by definitions purporting to establish their structure completely. The theorems were thought of as expressing truths about these systems, rather than as propositions applying hypothetically to whatever systems of objects (if any) satisfy the axioms. But then how could contradictions have arisen in these subjects, unless there is some defect in the logic, some error in the methods of constructing and reasoning about mathematical objects, which we had hitherto trusted?

To say that now these subjects should instead be established on an axiomatic basis does not of itself dispose of the problem. After axiomatization, there must still be some level at which we have truth and falsity. If the axiomatics is informal, the axioms must be true. If the axiomatics is formal, at least we must believe that the theorems do follow from the axioms; and also there must be some relationship between these results and some actuality outside the axiomatic theory, if the mathematicians' activity is not to reduce to nonsense. The formally axiomatized propositions of mathematics cannot constitute the whole of mathematics; there must also be an intuitively understood mathematics. If we must give up our former belief that it comprises all of arithmetic, analysis and set theory, we shall not be wholly satisfied unless we learn wherein that belief was mistaken, and where now instead to draw a line of separation.

The immediate problem of eliminating the paradoxes thus merges with the broader problem of the foundations of mathematics and logic.

What is the nature of mathematical truth? What meaning do mathematical propositions have, and on what evidence do they rest? This broad problem, or complex of problems, exists for philosophy apart from the circumstance that paradoxes have arisen in the fringes of mathematics. Historically, this circumstance has led to a more intensive study of the problem on the part of mathematicians than would otherwise have been likely; and the paradoxes obviously impose conditions on the solution of the problem.

Impredicative definition. When a set $M$ and a particular object $m$ are so defined that on the one hand $m$ is a member of $M$, and on the other hand the definition of $m$ depends on $M$, we say that the procedure (or the definition of $m$, or the definition of $M$) is *impredicative*. Similarly, when a property $P$ is possessed by an object $m$ whose definition depends on $P$ (here $M$ is the set of the objects which possess the property $P$). An impredicative definition is circular, at least on its face, as what is defined participates in its own definition.

Each of the antinomies of § 11 involves an impredicative definition. In (B), the set M of all sets includes as members the sets $\mathfrak{U}$M and $\mathfrak{S}$M defined from M. The impredicative procedure in the Russell paradox (C) stands out when the definition of $T$ is elaborated thus. We divide the set M of all sets into two parts, the first comprising those members which contain themselves, and the second (which is $T$) those which do not. Then we put $T$ (defined by this division of M into two parts) back into M, to ask into which part of M it falls. In the Richard paradox (D), the totality of expressions in the English language which constitute definitions of a function (real number, natural number) is taken as including the quoted expression, which refers to that totality. In the Epimenides paradox (E), the totality of statements is divided into two parts, the true and the false statements. A statement which refers to this division is reckoned as of the original totality, when we ask whether it is true or false.

Poincaré (1905-6, 1908) judged the cause of the paradoxes to lie in these impredicative definitions; and Russell (1906, 1910) enunciated the same explanation in his vicious circle principle: No totality can contain members definable only in terms of this totality, or members involving or presupposing this totality. Thus it might appear that we have a sufficient solution and adequate insight into the paradoxes, except for one circumstance: parts of mathematics we want to retain, particularly analysis, also contain impredicative definitions.

An example is the definition of u = l.u.b. M (§ 9 (A)). Under the



Dedekind cut definition of the real numbers, the set C of real numbers is the set of all sets x of rationals having three properties (a), (b), (c). Now this totality has been divided into two parts, M and C − M. We define u as $\mathfrak{S}$M, and then reckon this set $\mathfrak{S}$M as a member of C. This definition u = $\mathfrak{S}$M depends on C in the general case, since in the general case M will have been defined from C as the set of those members of C which have a certain property $P$.

One can attempt to defend this impredicative definition by interpreting it, not as defining or creating the real number u for the first time (in which interpretation the definition of the totality C of real numbers is circular), but as only a description which singles out the particular number u from an already existing totality C of real numbers. But the same argument can be used to uphold the impredicative definitions in the paradoxes.

Weyl's constructive continuum. The impredicative character of some of the definitions in analysis has been especially emphasized by Weyl, who in his book "Das Kontinuum (The continuum)" (1918) undertook to find out how much of analysis could be reconstructed without impredicative definitions. A fund of operations can be provided for constructing many particular categories of irrationals. Weyl was thus able to obtain a fair part of analysis, but not the theorem that an arbitrary non-empty set M of real numbers having an upper bound has a least upper bound. (Cf. also Weyl 1919.)

There have arisen three main schools of thought on the foundations of mathematics: (i) the logicistic school (Russell and Whitehead, English), (ii) the intuitionistic school (Brouwer, Dutch), and (iii) the formalistic or axiomatic school (Hilbert, German). (Sometimes "logistic" is used instead of "logicistic"; but "logistic" also has another meaning § 15.) This broad classification does not include various other points of view, which have not been as widely cultivated or do not comprise to a similar degree both a reconstruction of mathematics and a philosophy to support it.

Logicism. The logicistic thesis is that mathematics is a branch of logic. The mathematical notions are to be defined in terms of the logical notions. The theorems of mathematics are to be proved as theorems of logic.

Leibniz (1666) first conceived of logic as a science containing the ideas and principles underlying all other sciences. Dedekind (1888) and Frege (1884, 1893, 1903) were engaged in defining mathematical notions in terms of logical ones, and Peano (1889, 1894-1908) in expressing mathematical theorems in a logical symbolism.

To illustrate how mathematical notions can be defined from logical ones, let us presuppose the Frege-Russell definition of cardinal number (§ 3), and the definitions of the cardinal number 0 and of the cardinal number $n + 1$ for any cardinal number $n$ (§ 4). Then a *finite cardinal* (or *natural number*) can be defined as a cardinal number which possesses every property $P$ such that (1) 0 has the property $P$ and (2) $n + 1$ has the property $P$ whenever $n$ has the property $P$. In brief, a natural number is defined as a cardinal number for which mathematical induction holds. The viewpoint here is very different from that of §§ 6 and 7, where we presupposed an intuitive conception of the natural number sequence, and elicited from it the principle that, whenever a particular property $P$ of natural numbers is given such that (1) and (2), then any given natural number must have the property $P$. Here instead we presuppose the totality of all properties of cardinal numbers as existing in logic, prior to the definition of the natural number sequence. Note that this definition is impredicative, because the property of being a natural number, which it defines, belongs to the totality of properties of cardinal numbers, which is presupposed in the definition.

To adapt the logicistic construction of mathematics to the situation arising from the discovery of the paradoxes, Russell excluded impredicative definitions by his *ramified theory of types* (1908, 1910). Roughly, this is as follows. The primary objects or individuals (i.e. the given things not being subjected to logical analysis) are assigned to one type (say *type* 0), properties of individuals to *type* 1, properties of properties of individuals to *type* 2, etc.; and no properties are admitted which do not fall into one of these logical types (e.g. this puts the properties 'predicable' and 'impredicable' of § 11 outside the pale of logic). A more detailed account would describe the admitted types for other objects, such as relations and classes. Then, to exclude impredicative definitions within a type, the types above type 0 are further separated into orders. Thus for type 1, properties defined without mentioning any totality belong to *order* 0, and properties defined using the totality of properties of a given order belong to the next higher order. (The logicistic definition of natural number now becomes predicative, when the $P$ in it is specified to range only over properties of a given order, in which case the property of being a natural number is of the next higher order.) But this separation into orders makes it impossible to construct the familiar analysis, which as we saw above contains impredicative definitions. To escape this outcome, Russell postulated his *axiom of reducibility*, which asserts that to any property belonging to an order above the lowest, there is a coextensive



property (i.e. one possessed by exactly the same objects) of order 0. If only definable properties are considered to exist, then the axiom means that to every impredicative definition within a given type there is an equivalent predicative one.

The deduction of mathematics as a province of logic was carried out on this basis, using a logical symbolism, in the monumental "Principia mathematica" of Whitehead and Russell (three volumes, 1910-13). This work has had a great influence on subsequent developments in symbolic logic.

This deduction of mathematics from logic was offered as intuitive axiomatics. The axioms were intended to be believed, or at least to be accepted as plausible hypotheses concerning the world.

The difficulty is now: on what grounds shall we believe in the axiom of reducibility? If properties are to be constructed, the matter should be settled on the basis of constructions, not by an axiom. As the authors admitted in the introduction to their second edition (1925), "This axiom has a purely pragmatic justification: it leads to the desired results, and to no others [so far as is known]. But clearly it is not the sort of axiom with which we can rest content."

Ramsey 1926 found that the desired results and no others can apparently be obtained without the hierarchy of orders (i.e. with a *simple theory of types*). He classified the known antinomies into two sorts, now called 'logical' (e.g. the Burali-Forti, Cantor and Russell) and 'epistemological' or 'semantical' (e.g. the Richard and Epimenides); and he observed that the logical antinomies are (apparently) stopped by the simple hierarchy of types, and the semantical ones are (apparently) prevented from arising within the symbolic language by the absence therein of the requisite means for referring to expressions of the same language. But Ramsey's arguments to justify impredicative definitions within a type entail a conception of the totality of predicates of the type as existing independently of their constructibility or definability. This has been called "theological". Thus neither Whitehead and Russell nor Ramsey succeeded in attaining the logicistic goal constructively. (An interesting proposal for justifying impredicative definitions within a type, by Langford 1927 and Carnap 1931-2, is also not free of difficulties.)

Weyl 1946 says that, in the system of "Principia mathematica", "mathematics is no longer founded on logic, but on a sort of logician's paradise . . ."; and he observes that one who is ready to believe in this "transcendental world" could also accept the system of axiomatic set theory (Zermelo, Fraenkel, etc.), which, for the deduction of mathematics, has the advantage of being simpler in structure.

Logicism treats the existence of the natural number series as an hypothesis about the actual world ('axiom of infinity'). A quite different handling of the problem of infinity is proposed by the intuitionists (§ 13) and the formalists (§ 14).

From both the intuitionistic and the formalistic standpoints, the (abstract) natural number sequence is more elementary than the notions of cardinal number and of all properties of cardinal numbers, which are used in the logicistic characterization of it.

The logicistic thesis can be questioned finally on the ground that logic already presupposes mathematical ideas in its formulation. In the intuitionistic view, an essential mathematical kernel is contained in the idea of iteration, which must be used e.g. in describing the hierarchy of types or the notion of a deduction from given premises.

Recent work in the logicistic school is that of Quine 1940*. A critical but sympathetic discussion of the logicistic order of ideas is given by Gödel 1944. Introductory treatments are provided by Russell 1919 and Black 1933.

**§ 13. Intuitionism.** In the 1880's, when the methods of Weierstrass, Dedekind and Cantor were flourishing, Kronecker argued vigorously that their fundamental definitions were only words, since they do not enable one in general to decide whether a given object satisfies the definition.

Poincaré, when he defends mathematical induction as an irreducible tool of intuitive mathematical reasoning (1902, 1905-6), is also a forerunner of the modern intuitionistic school.

In 1908 Brouwer, in a paper entitled "The untrustworthiness of the principles of logic", challenged the belief that the rules of the classical logic, which have come down to us essentially from Aristotle (384—322 B.C.), have an absolute validity, independent of the subject matter to which they are applied. Quoting from Weyl 1946, "According to his view and reading of history, classical logic was abstracted from the mathematics of finite sets and their subsets. . . . Forgetful of this limited origin, one afterwards mistook that logic for something above and prior to all mathematics, and finally applied it, without justification, to the mathematics of infinite sets."

Two obvious examples will illustrate that principles valid in thinking about finite sets do not necessarily carry over to infinite sets. One is the principle that the whole is greater than any proper part, when applied to 1-1 correspondences between sets (§§ 1, 3, 4). Another is that a set of natural numbers contains a greatest.



A principle of classical logic, valid in reasoning about finite sets, which Brouwer does not accept for infinite sets, is the law of the excluded middle. The law, in its general form, says *for every proposition A, either A or not A*. Now let $A$ be the proposition *there exists a member of the set (or domain) D having the property P*. Then *not A* is equivalent to *every member of D does not have the property P*, or in other words *every member of D has the property not-P*. The law, applied to this $A$, hence gives *either there exists a member of D having the property P, or every member of D has the property not-P*.

For definiteness, let us specify $P$ to be a property such that, for any given member of $D$, we can determine whether that member has the property $P$ or does not.

Now suppose $D$ is a finite set. Then we could examine every member of $D$ in turn, and thus either find a member having the property $P$, or verify that all members have the property *not-P*. There might be practical difficulties, e.g. when $D$ is a very large set having say a million members, or even for a small $D$ when the determination whether or not a given member has the property $P$ may be tedious. But the possibility of completing the search exists in principle. It is this possibility which for Brouwer makes the law of the excluded middle a valid principle for reasoning with finite sets $D$ and properties $P$ of the kind specified.

For an infinite set $D$, the situation is fundamentally different. It is no longer possible in principle to search through the entire set $D$.

Moreover in this situation the law is not saved for Brouwer by substituting, for the impossible search through all the members of the infinite set $D$, a mathematical solution of the problem posed. We may in some cases, i.e. for some sets $D$ and properties $P$, succeed in finding a member of $D$ having the property $P$; and in other cases, succeed in showing by mathematical reasoning that every member of $D$ has the property *not-P*, e.g. by deducing a contradiction from the assumption that an arbitrary (i.e. unspecified) member of $D$ has the property $P$. (An example for the second kind of solution is when $D$ is the set of all the ordered pairs $(m, n)$ of positive integers, and $P$ is the property of a pair $(m, n)$ that $m^2 = 2n^2$. The result is then Pythagoras' discovery that $\sqrt{2}$ is irrational.) But we have no ground for affirming the possibility of obtaining either one or the other of these kinds of solutions in every case.

An example from modern mathematical history is afforded by Fermat's "last theorem", which asserts that the equation $x^n + y^n = z^n$ has no solution in positive integers $x, y, z, n$ with $n > 2$. (For $n = 2$, there are triples of positive integers, called Pythagorean numbers, which satisfy, e.g. $x = 3, y = 4, z = 5$ or $x = 5, y = 12, z = 13$.) Here $D$ is the set of all ordered quadruples $(x, y, z, n)$ of positive integers with $n > 2$, and $P$ is the property of a quadruple $(x, y, z, n)$ that $x^n + y^n = z^n$. About 1637 Fermat wrote on the margin of his copy of Bachet's "Diophantus" that he had discovered a truly marvellous demonstration of this "theorem" which the margin was too narrow to contain. Despite an immense expenditure of effort, no one since then has succeeded in proving or disproving the alleged "theorem"; and moreover we lack the knowledge of any systematic method, the pursuit of which must in principle ultimately lead to a determination as to its truth or falsity. (Cf. Vandiver 1946 for details.)

Brouwer's non-acceptance of the law of the excluded middle for infinite sets $D$ does not rest on the failure of mathematicians thus far to have solved this particular problem, or any other particular problem. To meet his objection, one would have to provide a method adequate in principle for solving not only all the outstanding unsolved mathematical problems, but any others that might ever be proposed in the future. How likely it is that such a method will be found, we leave for the time being to the reader to speculate. Later in the book we shall return to the question (§ 60).

The familiar mathematics, with its methods and logic, as developed prior to Brouwer's critique or disregarding it, we call *classical*; the mathematics, methods or logic which Brouwer and his school allow, we call *intuitionistic*. The classical includes parts which are intuitionistic and parts which are non-intuitionistic.

The non-intuitionistic mathematics which culminated in the theories of Weierstrass, Dedekind and Cantor, and the intuitionistic mathematics of Brouwer, differ essentially in their view of the infinite. In the former, the infinite is treated as *actual* or *completed* or *extended* or *existential*. An infinite set is regarded as existing as a completed totality, prior to or independently of any human process of generation or construction, and as though it could be spread out completely for our inspection. In the latter, the infinite is treated only as *potential* or *becoming* or *constructive*. The recognition of this distinction, in the case of infinite magnitudes, goes back to Gauss, who in 1831 wrote, "I protest ... against the use of an infinite magnitude as something completed, which is never permissible in mathematics." (Werke VIII p. 216.)

According to Weyl 1946, "Brouwer made it clear, as I think beyond any doubt, that there is no evidence supporting the belief in the existential character of the totality of all natural numbers .... The sequence of numbers which grows beyond any stage already reached by passing to the



next number, is a manifold of possibilities open towards infinity; it remains forever in the status of creation, but is not a closed realm of things existing in themselves. That we blindly converted one into the other is the true source of our difficulties, including the antinomies — a source of more fundamental nature than Russell's vicious circle principle indicated. Brouwer opened our eyes and made us see how far classical mathematics, nourished by a belief in the 'absolute' that transcends all human possibilities of realization, goes beyond such statements as can claim real meaning and truth founded on evidence."

Brouwer's criticism of the classical logic as applied to an infinite set $D$ (say the set of the natural numbers) arises from this standpoint respecting infinity. We see this clearly by considering the meanings which the intuitionist attaches to various forms of statements.

A generality statement *all natural numbers $n$ have the property $P$*, or briefly *for all $n$, $P(n)$*, is understood by the intuitionist as an hypothetical assertion to the effect that, if any particular natural number $n$ were given to us, we could be sure that that number $n$ has the property $P$. This is a meaning which does not require us to take into view the classical completed infinity of the natural numbers.

Mathematical induction is an example of an intuitionistic method for proving generality propositions about the natural numbers. A proof by induction of the proposition *for all $n$, $P(n)$* shows that any given $n$ would have to have the property $P$, by reasoning which uses only the numbers from 0 up to $n$ (§ 7). Of course, for a particular proof by induction to be intuitionistic, also the reasonings used within its basis and induction step must be intuitionistic.

An existence statement *there exists a natural number $n$ having the property $P$*, or briefly *there exists an $n$ such that $P(n)$*, has its intuitionistic meaning as a partial communication (or abstract) of a statement giving a particular example of a natural number $n$ which has the property $P$, or at least giving a method by which in principle one could find such an example.

Therefore an intuitionistic proof of the proposition *there exists an $n$ such that $P(n)$* must be *constructive* in the following (strict) sense. The proof actually exhibits an example of an $n$ such that $P(n)$, or at least indicates a method by which one could in principle find such an example.

In classical mathematics there occur *non-constructive* or *indirect* existence proofs, which the intuitionists do not accept. For example, to prove *there exists an $n$ such that $P(n)$*, the classical mathematician may deduce a contradiction from the assumption *for all $n$, not $P(n)$*. Under both the classical and the intuitionistic logic, by reductio ad absurdum this gives *not for all n, not* $P(n)$. The classical logic allows this result to be transformed into *there exists an n such that* $P(n)$, but not (in general) the intuitionistic. Such a classical existence proof leaves us no nearer than before the proof was given to having an example of a number $n$ such that $P(n)$ (though sometimes we may afterwards be able to discover one by another method). The intuitionist refrains from accepting such an existence proof, because its conclusion *there exists an n such that* $P(n)$ can have no meaning for him other than as a reference to an example of a number $n$ such that $P(n)$, and this example has not been produced. The classical meaning, that somewhere in the completed infinite totality of the natural numbers there occurs an $n$ such that $P(n)$, is not available to him, since he does not conceive the natural numbers as a completed totality.

As another example of a non-constructive existence proof, suppose it has been shown for a certain $P$, by intuitionistic methods, that if Fermat's "last theorem" is true, then the number 5013 has the property $P$, and also that if Fermat's "last theorem" is false, then 10 has the property $P$. Classically this suffices to demonstrate the existence of a number $n$ such that $P(n)$. But with the problem of the "last theorem" unsolved, Brouwer would disallow such an existence proof, because no example has been given. We do not know that 5013 is an example, nor do we know that 10 is an example, nor do we know any procedure which would in principle (i.e. apart from practical limitations on the length of procedures we can carry out) lead us to a particular number which we could be sure is an example. Brouwer would merely accept what has been given as proving the implication (or conditional statement) *if F or not F, then there exists an n such that* $P(n)$, where $F$ is the statement *for all* $x, y, z > 0$ *and* $n > 2$, $x^n + y^n \neq z^n$. The classical mathematician, by his law of the excluded middle, has the premise *F or not F* of this implication, and so he can infer its conclusion *there exists an n such that* $P(n)$. But in the present state of knowledge, Brouwer does not accept the premise *F or not F* as known.

As appears in this example, intuitionistic methods are to be distinguished from non-intuitionistic ones in the case of definitions as well as in the case of proofs. In the present state of our knowledge, Brouwer does not accept *the number n which is equal to* 5013 *if F, and equal to* 10 *if not F* as a valid definition of a natural number $n$.

A disjunction *A or B* constitutes for the intuitionist an incomplete communication of a statement telling us that $A$ holds or that $B$ holds, or at least giving a method by which we can choose from $A$ and $B$ one which holds. A conjunction *A and B* means that both $A$ and $B$ hold. An implication *A implies B* (or *if A, then B*) expresses that $B$ follows from $A$ by intuitionistic reasoning, or more explicitly that one possesses a method which, from any proof of $A$, would procure a proof of $B$; and a negation *not A* (or *A is absurd*) that a contradiction *B and not B* follows from $A$ by intuitionistic reasoning, or more explicitly that one possesses a method which, from any proof of $A$, would procure a proof of a contradiction *B and not B* (or of a statement already known to be absurd, such as $1 = 0$). Additional comments on these intuitionistic meanings will be given in § 82. See Note 1 on p. 65.



Quoting from Heyting 1934, "According to Brouwer, mathematics is identical with the exact part of our thinking. . . . no science, in particular not philosophy or logic, can be a presupposition for mathematics. It would be circular to apply any philosophical or logical principles as means of proof, since mathematical conceptions are already presupposed in the formulation of such principles." There remains for mathematics "no other source than an intuition, which places its concepts and inferences before our eyes as immediately clear." This intuition "is nothing other than the faculty of considering separately particular concepts and inferences which occur regularly in ordinary thinking." The idea of the natural number series can be analyzed as resting on the possibility, first of considering an object or experience as given to us separately from the rest of the world, second of distinguishing one such from another, and third of imagining an unlimited repetition of the second process. "In the intuitionistic mathematics, one does not draw inferences according to fixed norms, which can be collected in a logic, but each single inference is immediately tested on its evidence." But also "There are general rules, by which from given mathematical theorems new theorems can be formed in an intuitively clear way; the theory of these connections can be treated in a 'mathematical logic', which is then a branch of mathematics and is not sensibly applied outside of mathematics."

We turn now to the question: How large a part do the non-intuitionistic methods play in the classical mathematics?

The fact that non-intuitionistic methods occur in classical elementary number theory is significant, since it enables elementary number theory to serve as the first and simplest testing ground in research on foundations growing out of the intuitionistic and formalistic thinking. We shall be almost wholly concerned with elementary number theory in this book.

Actually, in the existing body of elementary number theory, the non-intuitionistic methods do not play a large part. Most non-constructive existence proofs can be replaced by constructive ones.

On the other hand, in analysis (and still more transcendental branches of mathematics) the non-intuitionistic methods of definition and proof permeate the whole methodology. The real numbers in the Dedekind cut representation are infinite sets of rationals (§ 9). Thus to treat them as objects in the usual way, we are already using the completed infinite. In particular, we do apply the law of the excluded middle to these sets, in connection with the simplest definitions of the subject. For example, to show that for any two real numbers x and y, either $x < y$ or $x = y$ or $x > y$, we use it twice, thus: Either there exists a rational $r$ in y which does not belong to x, or all rationals in y belong to x; and similarly interchanging x and y. In the impredicative definition of l.u.b. M (§ 9 (A), § 12), we use the totality of the real numbers in the same way. Another instance of non-constructive reasoning occurs in the proof of (B) § 9, where we assumed the right to choose an element $a_n$ from a set $M_n$, simultaneously for infinitely many values of $n$, without giving any property to determine which element is chosen. (This is a case of the 'axiom of choice', first noticed as an assumption by Zermelo 1904. We used it also for Theorem B § 4.)

Although the completed infinite has been banned for magnitudes (as Gauss called upon us to do), it reappears in full force for collections. As Hilbert and Bernays describe the situation in their "Grundlagen der Mathematik (Foundations of mathematics)", vol. I (1934), p. 41, "The ... arithmetization of analysis is not without a residue left over, as certain systematic fundamental conceptions are introduced which do not belong to the domain of intuitive arithmetical thinking. The insight which has given us the rigorous foundation of analysis consists in this: that these few fundamental assumptions do suffice for building up the theory of magnitudes as a theory of sets of integers."

The next question is: What kind of a mathematics can be built within the intuitionistic restrictions? If the existing classical mathematics could be rebuilt within the intuitionistic restrictions, without too great increase in the labor required and too great sacrifices in the results achieved, the problem of its foundations would appear to be solved.

The intuitionists have created a whole new mathematics, including a theory of the continuum and a set theory (cf. Heyting 1934). This mathematics employs concepts and makes distinctions not found in the classical mathematics; and it is very attractive on its own account. As a substitute for classical mathematics it has turned out to be less powerful,



and in many ways more complicated to develop. For example, in Brouwer's theory of the continuum, we cannot affirm that any two real numbers $a$ and $b$ are either equal or unequal. Our knowledge about the equality or inequality of $a$ and $b$ can be more or less specific. By $a \neq b$, it is meant that $a = b$ leads to a contradiction, while $a \# b$ is a stronger kind of inequality which means that one can give an example of a rational number which separates $a$ and $b$. Of course $a \# b$ implies $a \neq b$. But there are pairs of real numbers $a$ and $b$ for which it is not known that either $a = b$ or $a \neq b$ (or $a \# b$). It is clear that such complications replace the classical theory of the continuum by something much less perspicuous in form.

Despite this, the possibility of an intuitionistic reconstruction of classical mathematics in a different way involving reinterpretation (recently undertaken) is not to be ruled out (cf. § 81).

**§ 14. Formalism.** Brouwer has revealed what the genetic or constructive tendency involves in its ultimate refinement; Hilbert does the same for the axiomatic or existential (§ 8). The axiomatic method had already been sharpened from the material axiomatics of Euclid to the formal axiomatics of Hilbert's "Grundlagen der Geometrie" (1899). Formalism is the result of a further step, to meet the crisis caused by the paradoxes and the challenge to classical mathematics by Brouwer and Weyl. This step was forecasted by Hilbert in 1904, and seriously undertaken by him and his collaborators Bernays, Ackermann, von Neumann and others since 1920 (cf. Bernays 1935a, Weyl 1944).

Hilbert conceded that the propositions of classical mathematics which involve the completed infinite go beyond intuitive evidence. But he refused to follow Brouwer in giving up classical mathematics on this account.

To salvage classical mathematics in the face of the intuitionistic criticism, he proposed a program which we can state preliminarily as follows: Classical mathematics shall be formulated as a formal axiomatic theory, and this theory shall be proved to be consistent, i.e. free from contradiction.

Prior to this proposal of Hilbert's, the method used in consistency proofs for axiomatic theories, especially in Hilbert's earlier axiomatic thinking, was to give a 'model'. A *model* for an axiomatic theory is simply a system of objects, chosen from some other theory and satisfying the axioms (§ 8). That is, to each object or primitive notion of the axiomatic theory, an object or notion of the other theory is correlated, in such a way that the axioms become (or correspond to) theorems of the other theory.

If this other theory is consistent, then the axiomatic theory must be. For suppose that, in the axiomatic theory, a contradiction were deducible from the axioms. Then, in the other theory, by corresponding inferences about the objects constituting the model, a contradiction would be deducible from the corresponding theorems.

In a famous early example, Beltrami (1868) showed that the lines in the plane non-Euclidean geometry of Lobatchevsky and Bolyai (the plane hyperbolic geometry) can be represented by the geodesics on a surface of constant negative curvature in Euclidean space. Thus the plane hyperbolic geometry is consistent, if the Euclidean geometry is consistent. (Another model for the same was given by Klein (1871) in terms of plane projective geometry with Cayley's metric (1859); this can be construed as a model in the Euclidean plane. Cf. Young 1911 Lectures II and III.)

The analytic geometry of Descartes (1619), i.e. the use of coordinates to represent geometrical objects, constitutes a general method for establishing the consistency of geometric theories on the basis of analysis, i.e. the theory of the real numbers.

Consistency proofs by the method of a model are relative. The theory for which a model is set up is consistent, if that from which the model is taken is consistent.

Only when the latter is unimpeachable does the model give us an absolute proof of consistency. Veblen and Bussey 1906 achieve absolute proofs of consistency for certain rudimentary projective geometries by setting up models using only a finite (sic!) class of objects to represent the points (cf. Young 1911 Lectures IV and V).

For proving absolutely the consistency of classical number theory, of analysis, and of set theory (suitably axiomatized), the method of a model offers no hope. No mathematical source is apparent for a model which would not merely take us back to one of the theories previously reduced by the method of a model to these.

The impossibility of drawing upon the perceptual or physical world for a model is argued in Hilbert and Bernays 1934 pp. 15—17. They illustrate it by considering Zeno's first paradox (fifth century B.C.), according to which a runner cannot run a course in a finite time. For before he can do so, he must run the first half, then the next quarter, then the next eight, and so on. But this would require him to complete an infinite number of acts. The usual solution of the paradox consists in observing that the infinite series of the time intervals required to run the successive segments converges. "Actually there is also a much more radical solution of the paradox. This consists in the consideration that



we are by no means obliged to believe that the mathematical space-time representation of motion is physically significant for arbitrarily small space and time intervals; but rather have every basis to suppose that that mathematical model extrapolates the facts of a certain realm of experience, namely the motions within the orders of magnitude hitherto accessible to our observation, in the sense of a simple concept construction, similarly to the way the mechanics of continua completes an extrapolation in which a continuous filling of the space with matter is assumed .... The situation is similar in all cases where one believes it possible to exhibit directly an [actual] infinity as given through experience or perception .... Closer examination then shows that an infinity is actually not given to us at all, but is first interpolated or extrapolated through an intellectual process."

Therefore, if consistency is to be proved for number theory (including its non-intuitionistic portions), for analysis, etc., it must be by another method. It is Hilbert's contribution now to have conceived a new direct approach, and to have recognized what it involves for the axiomatization. This direct method is implicit in the meaning of consistency (at least as we now think of it), namely that no logical contradiction (a proposition $A$ and its negation *not* $A$ both being theorems) can arise in the theory deduced from the axioms. Thus to prove the consistency of a theory directly, one should prove a proposition about the theory itself, i.e. specifically about all possible proofs of theorems in the theory. The mathematical theory whose consistency it is hoped to prove then becomes itself the object of a mathematical study, which Hilbert calls "metamathematics" or "proof theory". How this is possible, and what the methods of the study may be, we shall examine in the next section.

Meanwhile let us consider further the import of Hilbert's proposal. Hilbert (1926, 1928) draws a distinction between 'real' and 'ideal' statements in classical mathematics, in essence as follows. The ***real statements*** are those which are being used as having an intuitive meaning; the ***ideal statements*** are those which are not being so used. The statements which correspond to the treatment of the infinite as actual are ideal. Classical mathematics adjoins the ideal statements to the real, in order to retain the simple rules of the Aristotelian logic in reasoning about infinite sets.

The addition of 'ideal elements' to a system to complete its structure and simplify the theory of the system is a common and fruitful device in modern mathematics. For example, in Euclidean plane geometry two distinct lines intersect in a unique point, except when the lines are parallel. To remove this exception, Poncelet in his projective geometry (1822) introduced a *point at infinity* on each of the original lines, such that parallel lines have the same point at infinity and non-parallel lines have different points at infinity. The totality of these points at infinity make up a *line at infinity*. As a line through a finite point of the projective plane rotates, its point at infinity traces out the line at infinity. By this device, the relationships of incidence between points and lines is simplified. Two distinct points determine a unique line (which is 'on' both points, i.e. through both of which the line passes); and two distinct lines determine a unique point (which is on both lines). These two propositions are *duals* of each other. There is a general principle, called the *principle of duality* for plane projective geometry, which says that to each theorem of the subject the statement obtained from it by interchanging the words "point" and "line" is also a theorem.

As other examples of the addition of elements to a previously constituted system of elements to serve some theoretical purpose, we may take the successive enlargements of the number system, starting say with the natural numbers, then adjoining the negative integers, then the fractions, then the irrationals, and finally the imaginary numbers. The adjunction of the negative integers simplifies the theory of addition by making the inverse operation (subtraction) always possible; etc.

Hilbert's problem is crudely analogous to the problem which existed when imaginary numbers first came into use. As they were then not clearly understood, one might have proposed to justify their use to doubters by proving that, if imaginaries are used according to prescribed rules to derive a result expressed in terms of reals only, then that result must be correct. Of course, this kind of justification for imaginaries relative to reals is not needed now, since their interpretation by points in the plane (Wessel 1799) and by pairs of reals (Gauss 1831) have become known.

This analogy suggests asking whether, if a proof of consistency in Hilbert's sense should succeed for a portion of classical mathematics comprising both real and ideal statements, we could then infer that the real statements proved therein by an excursion through the ideal are true intuitionistically? The extent to which we could will be discussed later (end § 42, end § 82); it will depend on what reasonings are covered by the consistency proof, and what class of statements is being taken as real. To this extent, success in Hilbert's program would give to classical mathematics a role as a method of proof for the intuitionists.

A sharp controversy arose between Brouwer and Hilbert in the early years after Hilbert's program took shape. Brouwer 1923 said, "An incorrect theory which is not stopped by a contradiction is none the less incorrect, just as a criminal policy unchecked by a reprimanding court is none the less criminal." Hilbert 1928 retorted, "To take the law of the excluded middle away from the mathematician would be like denying the astronomer the telescope or the boxer the use of his fists."



According to Brouwer (1928) and Heyting (1931-2, 1934), agreement between intuitionism and formalism is possible, provided (as in von Neumann 1931-2) the formalist refrains from attributing to the non-intuitionistic classical mathematics a material meaning or content, in terms of which the consistency proof justifies it. Such a justification, says Brouwer, "contains a vicious circle, because this justification depends on the (material (inhaltlichen)) correctness of the proposition that from the consistency of a statement the correctness of that statement follows, i.e. on the (material) correctness of the law of the excluded middle", which is part of the formalistic mathematics that is to be justified.

The delicate point in the formalistic position is to explain how the non-intuitionistic classical mathematics is significant, after having initially agreed with the intuitionists that its theorems lack a real meaning in terms of which they are true.

Classical mathematics constructs theories in quite a different sense from intuitionistic mathematics. Hilbert 1928 says, "It is by no means reasonable to set up in general the requirement that each separate formula should be interpretable taken by itself ...." In theoretical physics "only certain combinations and consequences of the physical laws can be checked experimentally — likewise in my proof theory only the real statements are immediately capable of a verification".

A theory in classical mathematics can be regarded as a simple and elegant systematizing scheme, by which a variety of (presumably) true real statements, previously appearing as heterogeneous and unrelated, and often previously unknown, are comprised as consequences of the ideal theorems in the theory. (Cf. von Neumann 1947, Einstein 1944 p. 288.)

The example of analytic number theory illustrates that theorems of analysis (lacking a meaning acceptable to the intuitionist) often entail theorems of number theory, which are meaningful intuitionistically, and for which either no non-analytic proofs have been discovered or only much more complicated ones.

For a theory to be valuable in this way, the real statements comprised must be true. Formerly mathematicians supposed this to be guaranteed by the truth of the theorems which we now recognize as ideal; now we hope to guarantee it instead by a consistency proof.

By easy stages of transition, the theorizing may climb to higher levels, from which it is only very indirectly concerned with systematizing the real propositions at the original level, but rather with systematizing ideal propositions at intermediate levels. In this connection it is of interest whether successively higher theoretical constructions actually add to the body of real propositions of the original sort which are comprised, as well as whether they do actually permit substantial simplifications of the proofs of those previously comprised. (Cf. end § 42.)

It is debatable how high a theoretical structure is justified for systematizing a given sort of real truths, e.g. whether classical analysis is justified as a systematization of number-theoretic truths. Historically analytic number theory was a by-product, and the actual impetus to the development of classical analysis came from the sciences, including geometry in its physical application.

Hilbert and Bernays 1934 emphasize that in the sciences "we have to do ... predominately with theories which do not reproduce the actual state of affairs completely, but represent a *simplifying idealization* of the state of affairs and have their meaning therein" (pp. 2—3). Analysis serves as a "formation of ideas (Ideenbildung)", in terms of which those theories can be expressed, or to which they can be reduced by the method of models. A proof of the consistency of analysis would assure us of the consistency of the idealizations effected in those theories (p. 19).

Weyl (1926, 1928, 1931) observes that in theoretical physics it is not the separate statements which are confronted with experience, but the theoretical system as a whole. What is afforded here is not a true description of what is given, but theoretical, purely symbolic construction of the world. (Also he argues that our theoretical interest is not exclusively or even primarily in the 'real statements', e.g. that this pointer coincides with that scale division, but rather in the ideal suppositions, e.g. the supposition of the electron as a universal electrical quantum.) It is a deep philosophical question what the 'truth' or objectivity is which pertains to this theoretical world construction going far beyond the given. This is closely connected with the question, what motivates us to take as basis the particular axiom system chosen. For this consistency is a necessary but not sufficient argument. When mathematics is taken for itself alone, he would restrict himself with Brouwer to the intuitive truths; he does not find a sufficient motive to go further. But when mathematics is merged completely with physics in the process of theoretical world construction, he sides with Hilbert.



A verdict on the formalists' thinking will depend partly on the fruits of the program they propose. This program calls for a subject called "metamathematics", in which they aim in particular to establish the consistency of classical mathematics.

We note in advance that metamathematics will be found to provide a rigorous mathematical technique for investigating a great variety of foundation problems for mathematics and logic, among which the consistency problem is only one. For example, metamathematical methods are applied now in studies of systematizations of mathematics arising from the logicistic and intuitionistic schools, as well as from Hilbert's. (Inversely, metamathematics owes much for its inception to the logicistic and intuitionistic investigations.) Our aim in the rest of this book is not to reach a verdict supporting or rejecting the formalistic viewpoint in any preassigned version; but to see what the metamathematical method consists in, and to learn some of the things that have been discovered in pursuing it.

**§ 15. Formalization of a theory.** We are now about to undertake a program which makes a mathematical theory itself the object of exact mathematical study. In a mathematical theory, we study a system of mathematical objects. How can a mathematical theory itself be an object for mathematical study?

The result of the mathematician's activity is embodied in propositions, the asserted propositions or theorems of the given mathematical theory. We cannot hope to study in exact terms what is in the mathematician's mind, but we can contemplate the system of these propositions.

The system of these propositions must be made entirely explicit. Not all of the propositions can be written down, but rather the disciple and student of the theory should be told all the conditions which determine what propositions hold in the theory.

As the first step, the propositions of the theory should be arranged deductively, some of them, from which the others are logically deducible, being specified as the axioms (or postulates).

This step will not be finished until all the properties of the undefined or technical terms of the theory which matter for the deduction of the theorems have been expressed by axioms. Then it should be possible to perform the deductions treating the technical terms as words in themselves without meaning. For to say that they have meanings necessary to the deduction of the theorems, other than what they derive from the axioms which govern them, amounts to saying that not all of their properties