6

FORMALISM: DO MATHEMATICAL STATEMENTS MEAN ANYTHING?

CASUAL observation reveals, or seems to reveal, that much mathematical activity consists of the manipulation of linguistic symbols according to certain rules. If someone doing arithmetic establishes a sentence in the form $a \times b = c$, then he can write the corresponding $b \times a = c$. If he also gets to a sentence like $a \neq 0$, then he is entitled to write $c/a = b$. The elementary and advanced parts of mathematics alike have this feature of at least appearing as rule-governed manipulation.

What is the significance of this observation about the practice of mathematics? The various philosophies that go by the name of 'formalism' pursue a claim that the *essence* of mathematics is the manipulation of characters. A list of the characters and allowed rules all but exhausts what there is to say about a given branch of mathematics. According to the formalist, then, mathematics is not, or need not be, about anything, or anything beyond typographical characters and rules for manipulating them.

Formalism seizes on one aspect of mathematics, perhaps neglecting or downplaying all else. For better or worse, much elementary arithmetic is taught as a series of blind techniques, with little or no indication of what the techniques do, or why they work. How many schoolteachers could explain the rules for long division, let alone the algorithm for taking square roots, in terms other than the execution of a routine? But perhaps this is more of a critique of some pedagogy than an attempt to justify a philosophy.[1]

Formalism has a better pedigree among mathematicians than among philosophers of mathematics. Throughout history, mathematicians have had occasion to introduce symbols which, at the time, seemed to have no clear interpretation. The very names 'negative numbers', 'irrational numbers', 'transcendental numbers', 'imaginary numbers', and 'ideal points at infinity' indicate ambivalence. Fortunately, the profession of mathematics has had its share of bold, imaginative souls, but it seems that more sceptical folk provide the names. Although the newly introduced 'entities' proved useful for applications within mathematics and science, in their philosophical moments some mathematicians did not know what to make of them. What are imaginary numbers, really? A common response to such dilemmas is to retreat to formalism. The mathematician asserts that symbols for complex numbers, for example, are to be manipulated according to (most of) the same rules as real numbers, and that is all there is to it.

Mathematicians themselves, however, do not always develop their philosophical positions in depth. One of the most detailed articulations of the basic versions of formalism is found in Gottlob Frege's (1893: §86–137) vigorous critique of the view.

1. Basic Views; Frege's Onslaught

There are at least two different general positions that have some historical claim to the title 'formalism'. Although the philosophies stand in opposition to each other in crucial ways, both opponents and defenders of formalism sometimes run them together.

[1] The advent of calculators may increase the tendency toward formalism. If there is a question of justifying, or making sense of, the workings of the calculator, it is for an engineer (or a physicist), not a teacher or student of elementary mathematics. Is there a real need to assign 'meaning' to the button-pushing? We hear (or used to hear) complaints that calculators ruin the younger generation's ability to think, or at least their ability to do mathematics. It seems to me that if the basic algorithms and routines are taught by rote, with no attempt to explain what they do or why they work, then the children might as well use calculators. Formalism cuts deeply.

1.1. Terms

Term formalism is the view that mathematics is about characters or symbols—the systems of numerals and other linguistic forms. That is, the term formalist *identifies* the entities of mathematics with their names. The complex number $8 + 2i$ is just the symbol '$8 + 2i$'. A thorough term formalist would also identify the natural number 2 with the numeral '2', but perhaps one can be a formalist about some branches of mathematics and not others. One might adopt formalism only for those branches that one is queasy about.

According to term formalism, then, mathematics has a subject-matter, and mathematical propositions are true or false. The view proposes simple answers to (seemingly) difficult metaphysical and epistemological problems with mathematics. What is mathematics about? Numbers, sets, and so on. What *are* these numbers, sets, and so on? They are linguistic characters. How is mathematics known? What is mathematical knowledge? It is knowledge of how the characters are related to each other, and how they are to be manipulated in mathematical practice.

Consider the simplest possible equation:

$$0 = 0.$$

Presumably it comes out true. How does the term formalist interpret it? She cannot say that the equation says that the leftmost hunk of ink (or burnt toner) shaped like an oval is identical to the rightmost hunk of ink also shaped line an oval. Clearly, those are two different hunks of ink.

The term formalist might take the equation to assert that those two hunks of ink have the same shape. But this seems to presuppose the existence of entities called 'shapes'. When discussing linguistic items like letters and sentences, contemporary philosophers distinguish *types* from *tokens*. Tokens are physical objects made up of ink, pencil, chalk marks, burned toner, and so on. As physical objects, they can be created and destroyed at will. Types are the abstract forms of tokens. The word 'concatenation' has two instances of the one type 'c'. The type 'c' is shared by all letter-tokens of that shape. When we say that the Roman alphabet has twenty-six letters, we are talking about the types, not the tokens. The statement would remain true if every token of the letter 'a' were destroyed. From this perspective, the term formalist might assert that mathematics is about *types*. The above equation would thus be a simple, straightforward instance of the law of identity. The equation says that the type '0' is identical with itself.

What are we to make of these shapes or types? Notice that shapes and types are abstract objects, much like numbers. What, then, is the advantage of term formalism over realism in ontology that asserts the existence of numbers outright? Perhaps the term formalist can maintain that, unlike numbers, types have straightforward instances, their tokens, and we learn things about them through their tokens.

A rudimentary term formalism was put forward (at least temporarily) by two mathematicians, E. Heine and Johannes Thomae, around the turn of the twentieth century. Heine (1872: 173) wrote, 'I give the name *numbers* to certain tangible signs, so that the existence of these numbers is thus unquestionable'. Thomae (1898: §§1–11) claimed that the 'formal standpoint rids us of all metaphysical difficulties; this is the advantage it affords us'. This remains to be seen.

Frege (1893: §§86–137) launched a sustained articulation of, and harsh attack on, their views. Consider the equation:

$$5 + 7 = 6 + 6.$$

What can this come to? Perhaps it means that the symbol '5 + 7' is identical to the symbol '6 + 6'. But this is absurd. Even the *types* are different. The former '5 + 7' has an occurrence of the type '5' and the latter '6 + 6' does not. It is not open to the formalist to claim that the two symbols denote the *same number*, since the central thesis of term formalism is that we need not consider extra-linguistic entities that the terms supposedly denote. All that matters are the *characters*. They denote themselves. So the term formalist cannot interpret the ' = ' sign as identity. On behalf of term formalism, Frege suggests that the equation be interpreted as saying that in arithmetic, the symbol '5 + 7' can be substituted anywhere for '6 + 6' without a change in truth-value. That is, a sentence of the form $A = B$ says that the symbol corresponding to A is inter-substitutable with the symbol corresponding to B in any mathematical context. So the above identity '0 = 0' asserts the truism that the type '0' can be substituted for itself without a change in truth-value.

Term formalism can perhaps be extended to the integers and rational numbers, but what are the real numbers supposed to be? We cannot identify them with their names, since most real numbers do not have names. A term formalist might attempt to identify the real number π with the Greek letter 'π', but what would he say about real numbers that do not have names? How would he understand a statement about *all* real numbers? A straightforward attempt would be to identify π with its decimal expansion: 3.14159 . . . However, the expansion is an infinitary object, and not a linguistic symbol. The term formalist might introduce a theory of 'limits' of terminating decimals, and identify π with the 'limit' of the symbols '3', '3.1', '3.14', . . . If this route is followed, however, it is hard to see any advantage of term formalism. The 'limit' of the symbols looks too much like the ordinary understanding of π as the limit of the rational *numbers* 3, 3.1, 3.14, . . . We seem to have lost the sense of formalism.

Suppose that the term formalist manages to solve this problem and come up with a decent linguistic surrogate for real numbers. Still, the view only captures mathematical *calculation*. How is the term formalist to make sense of mathematical *propositions*, like the prime number theorem or the fundamental theorem of calculus? In what sense can those be said to be about symbols?

1.2. Games

The other basic version of formalism likens the practice of mathematics to a game played with linguistic characters. Just as, in chess, one can use a pawn to capture one square forward on a diagonal, so in arithmetic one can write '$x = 10$' if one has previously gotten to '$x = 8 + 2$'. Call this *game formalism*.

Radical versions of this view assert outright that the symbols of mathematics are meaningless. Mathematical formulas and sentences do not express true or false propositions about any subject-matter. The view is that mathematical characters have no more meaning than the pieces on a chessboard. The 'content' of mathematics is exhausted by the rules for operating with its language. More moderate versions of game formalism concede that the languages of mathematics may have some sort of meaning, but if so, this meaning is irrelevant to the practice of mathematics. As far as the working mathematician is concerned, the symbols of mathematical language may as well be meaningless.

The difference between radical and moderate versions of game formalism has little significance for the philosophy of mathematics. The two views agree on the lack of *mathematical* interpretation for the typographical characters of a branch of mathematics. Against this, the term formalist holds that mathematics is about its terminology.

Like term formalism, game formalism either solves or sidesteps difficult metaphysical and epistemological problems with mathematics. What is mathematics about? Nothing. What *are* numbers, sets, and so on? They do not exist, or they might as well not exist. How is mathematics known? What is mathematical knowledge? It is knowledge of the rules of the game, or knowledge that certain moves that accord with these rules have been made. The equation '$2^{10} = 1024$' and the theorem that for every natural number x there is a prime number $y > x$ (in symbols, $\forall x \exists y(y > x \,\&\, y$ is prime)) each indicate the outcome of a certain play in accordance with the rules of arithmetic.[2]

In the context of game formalism, the phrases like 'language' and 'symbol' are misleading. In just about any other context, the purpose of language, first and foremost, is to communicate. We use language to talk *about* things, usually things other than language itself. In its normal usage, a symbol *symbolizes something*. The word 'Stewart' stands for the person Stewart. So one would think that the numeral '2' stands for the number 2. This is just what the game formalist denies, or demurs from. Either the numeral does not stand for anything, or else it might as well not stand for anything. For mathematics, all that matters is the numeral, and the role of the numeral in the game of mathematics.

It is ironic that Frege's own work in logic (see ch. 5, §1) gives impetus to a sophisticated version of game formalism. Frege claimed that one of the purposes of his logic was to codify correct inference. To determine the epistemic significance of a derivation, there can be no 'gaps' in the reasoning; all premisses must be made

[2] Since Wittgenstein 1953, there has been much philosophical discussion of rule-following. What is it for someone to be following one rule, rather than another? Can we distinguish the following of one rule incorrectly from the following of a different rule correctly? See, for example, Kripke 1982. If there is an issue here, it is a problem for any philosophy of mathematics, not just game formalism.

explicit. For this purpose, Frege developed a *formal* system, or to be precise, he presented a deductive system that could be understood formally: 'my concept writing . . . is designed to . . . be *operated like a calculus* by means of a small number of standard moves, so that no step is permitted which does not conform to the rules which are laid down once and for all?' (Frege 1884: §91, emphasis mine). Frege was aware that this feature could feed a version of formalism:

Now it is quite true that we could have introduced our rules and other laws of the *Begriffsschrift* [e.g. Frege 1879] as arbitrary stipulations, without speaking of the meaning and the sense of the signs. We would then have been treating the signs as figures. What we took to be the external representation of an inference would then be comparable to a move in chess, merely the transition from one configuration to another. We might give someone our [axioms] and . . . definitions . . . —as we might the initial position of the pieces in chess—tell him the rules permitting transformations, and then set him the problem of deriving our theorem . . . all this without his having the slightest inkling of the sense and meaning of these signs, or of the thoughts expressed by the formulas . . . (Frege 1903: §90)

Frege pointed out that the *meaning* that we attribute to the sentences is what makes them interesting, and that this meaning suggests strategies for the derivations. The game formalist might agree with this, but will add that the meaning of mathematical expressions is extraneous to mathematics itself. As far as mathematics goes, all that matters is that the rules are followed. Meaning is merely heuristic, no more than a psychological aid. Mathematics need have no subject-matter at all.

The game formalist, however, is left with a daunting problem. Why are the mathematical games so useful in the sciences? After all, no one even looks for useful applications of chess. Why think that the meaningless game of mathematics should have any applications? It clearly does, and we have to explain those applications. A similar problem arises for applications of mathematics within mathematics. Why is the game of complex analysis useful in the game of real analysis or arithmetic? This issue is all the more troubling for someone who is a game formalist about, say, complex analysis, but not about real analysis or arithmetic.

In this sense, game formalism is much like a philosophy of science called *instrumentalism*, which was designed to alleviate worries about unobserved theoretical entities, like electrons. According to instrumentalism, theoretical science is no more than a complicated instrument for making predictions about the observable, physical world. The scientist need not believe that theoretical entities exist. The instrumentalist is thus spared the epistemological problem of accounting for our knowledge of theoretical entities, but she is left with a gaping problem of explaining just why the instrument works so well, or why it works at all. Similarly, the game formalist is spared the problem of saying what mathematics is about, and perhaps she has a clean solution to the problem of how mathematics is known, but the issue of why mathematics is useful now looks intractable.

Frege's (1903: §91) main criticism of game formalism goes along these lines:

> an arithmetic without thought as its content will also be without possibility of application. Why can no application be made of a configuration of chess pieces? Obviously, because it expresses no thought. If it did so and every chess move conforming to the rules corresponded to a transition from one thought to another, applications of chess would also be conceivable. Why can arithmetical equations be applied? Only because they express thoughts. How could we possibly apply an equation which expressed nothing and was nothing more than a group of figures, to be transformed into another group of figures in accordance with certain rules? [I]t is applicability alone which elevates arithmetic from a game to the rank of a science.

The formalist could retort that applications are not part of mathematics itself, but are extraneous to it. Frege (1903: §88) quotes Thomae (1898: §§1–11):

> The formal conception of numbers accepts more modest limitations than does the logical conception. It does not ask what numbers are and what they do, but rather what is demanded of them in arithmetic. For the formalist, arithmetic is a game with signs which are called empty. That means that they have no other content (in the calculating game) than they are assigned by their behaviour with respect to certain rules of combination (rules of the game). The chess player makes similar use of his pieces; he assigns them certain properties determining their behavior in the game . . . To be sure, there is an important difference between arithmetic and chess. The rules of chess are arbitrary, the system of rules for arithmetic is such that by means of simple axioms the numbers can be referred to manifolds and can thus make important contributions to our knowledge of nature.

Thomae here seems to adopt the view I call 'moderate game formalism'. The idea is that the mathematician treats his 'language' *as if* it is a bunch of meaningless characters. The rules for arithmetic were perhaps chosen for the purpose of some applications, but these applications are of no concern to the mathematician as such. As Frege puts it on behalf of this game formalist, 'in formal arithmetic we absolve ourselves from accounting for one choice of the rules rather than another' (Frege 1903: §89).

Frege responds that the problem of applicability does not go away just because the formalist, or even the mathematician, refuses to deal with it. He sarcastically asks what is gained by the dodge: 'To be sure, arithmetic is relieved of some work, but does this dispose of the problem? The [formalist] shifts it to the shoulders of his colleagues, the geometers, the physicists, and the astronomers; but they decline the occupation with thanks; and so it falls into a void between the sciences. A clear cut separation of the domains of the sciences may be a good thing, provided that no domain remains for which no one is responsible' (Frege 1903: §92). Frege then points out that the applications in question are extremely wide. Mathematics applies to anything that can be counted or measured. The same number 'may arise with lengths, time intervals, masses, moments of inertia, etc.' Thus, the problem of 'the usefulness of arithmetic is to be solved—in part, at least—independently of those sciences to which it is to be applied'. And so it will not do to avoid the problem in this way.[3] Even if Frege's dismissal of formalism is premature, it is clear that the formalist does owe us an account of the applicability of mathematics.

2. Deductivism: Hilbert's *Grundlagen der Geometrie*

One of Frege's criticisms of game formalism suggests a variation on the moderate version of that view. Suppose that someone—the

[3] The wide applicability of numbers is one of Frege's considerations in favour of logicism. His own account of the natural numbers explicitly begins with one of their applications: to mark cardinality (see Chapter 5, §1). Frege's (1903) account of the real numbers turns on their application in measuring ratios of quantities (see Simons 1987 and Dummett 1991: ch. 22).

mathematician, the physicist, the astronomer—manages to interpret the basic axioms of, say, arithmetic so that they come out true. This is not enough to secure an application for arithmetic, since by itself this interpretation would not guarantee that the *theorems* are true under the same interpretation. How do we know that the rules of the arithmetic-game take us from truths (so interpreted) to truths? Frege (1903: §91) wrote:

> Whereas in an arithmetic with content equations and inequations are senses expressing thoughts, in formal arithmetic they are comparable with the positions of chess pieces, transformed in accordance with the rules without consideration for any sense. For if they were viewed as having sense, the rules could not be arbitrarily stipulated; they would have to be chosen so that from formulas expressing true propositions [one] could [derive] only formulas likewise expressing true propositions. Then the standpoint of formal arithmetic would have been abandoned, which insists that the rules for the manipulation of signs are quite arbitrarily stipulated.

In contemporary terms, for the application of a branch like arithmetic to succeed, the rules of the game cannot be arbitrary, but must constitute *logical consequences*. No matter how the language is interpreted, if the axioms come out true, then the theorems should be true under the same interpretation.

The advent of rigorous deductive systems—thanks in large part to Frege—suggests a tempting philosophy that has something in common with game formalism, but avoids this particular pitfall. A *deductivist* accepts Frege's point that rules of inference must preserve truth, but she insists that the *axioms* of various mathematical theories be treated as if they were arbitrarily stipulated. The idea is that the practice of mathematics consists of determining logical consequences of otherwise uninterpreted axioms. The mathematician is free to regard the axioms (and the theorems) of mathematics as meaningless, or to give them an interpretation at will.

To articulate this view rigorously, one would distinguish the logical terms like 'and', 'if . . . then', 'there exists', and 'for all' from the non-logical, or specifically mathematical, terminology such as 'number', 'point', 'set', and 'line'. The logical terminology is understood with its normal meaning, while the non-logical terminology is left uninterpreted, or is treated as if it were uninterpreted.[4] Let Φ be a theorem of, say, arithmetic. According to deductivism, the 'content' of Φ is that Φ follows from the axioms of arithmetic. Deductivism is sometimes called 'if-then-ism'.

The affinity between game formalism and deductivism results from the development of logical systems that can be 'operated like a calculus', as Frege put it. Deductivism is consonant with the slogan that logic is topic-neutral. From the modern, model-theoretic point of view, if an inference from a set of premisses Γ to a conclusion Φ is valid, then Φ is true under any interpretation that makes all of the premisses Γ true. The idea behind deductivism is to ignore the interpretation and stick to the inferences.

Like the game formalist, our deductivist proposes clean answers to philosophical questions. What is mathematics about? Nothing, or it can be regarded as about nothing. What is mathematical knowledge? It is knowledge of what follows from what. Mathematical knowledge is *logical* knowledge.[5] How is a branch of mathematics applied? By finding interpretations that make its axioms true.

Deductivism is a philosophy that goes well with developments in the foundations of mathematics, especially geometry, during the nineteenth and early twentieth centuries. The crucial events included the advent and success of analytic geometry, with projective geometry as a response; the attempt to accommodate ideal and imaginary elements, such as points at infinity; the development of *n*-dimensional geometry; and the assimilation of non-Euclidean geometry into mainstream mathematics alongside, not replacing, Euclidean geometry. These themes helped to undermine the Kantian thesis that mathematics is tied to intuitions of space and time (see ch. 4, §2). The mathematical community took on a growing interest in rigour, in the axiomatizations of various branches of mathematics, and ultimately in the understanding of deduction as independent of content. It is perhaps a small and natural step from these mathematical and logical developments to the philosophical thesis that the 'interpretation' of the axioms does not matter. The physicist can worry about whether real space-time is Euclidean or

[4] This approach is foreign to Frege's logicism. For Frege, every term of mathematics is logical, and so would be fully interpreted. See van Heijenoort 1967a and Goldfarb 1979.

[5] Deductivism has this much in common with logicism (see ch. 5).

4-dimensional, but the mathematician is free to explore the consequences of all kinds of geometries.

Moritz Pasch developed the idea that logical inference should be topic-neutral. Pasch wrote that geometry should be presented in a formal manner, without relying on intuition or observation when making inferences:

> If geometry is to be truly deductive, the process of inference must be independent in all its parts from the meaning of the geometrical concepts, just as it must be independent of the diagrams; only the relations specified in the propositions and definitions may legitimately be taken into account. During the deduction it is useful and legitimate, but in no way necessary, to think of the meanings of the terms; in fact, if it is necessary to do so, the inadequacy of the proof is made manifest. (Pasch 1926: 91)

Ernest Nagel (1939: §70) wrote that Pasch's work set the standard for geometry: 'No work thereafter held the attention of students of the subject which did not begin with a careful enumeration of the undefined or primitive terms and unproved or primitive statements; and which did not satisfy the condition that all further terms be defined, and all further statements proved, solely by means of this primitive base.'

David Hilbert's work in geometry around the turn of the twentieth century represents the culmination of these foundational developments. The programme executed in his *Grundlagen der Geometrie* (1899) marked an end to an essential role for intuition in geometry. Although spatial intuition or observation remains the source of the axioms of Euclidean geometry, in Hilbert's writing the role of intuition and observation is explicitly limited to motivation and is heuristic. Once the axioms have been formulated, intuition and observation are banished. They are not part of mathematics.

One result of this orientation is that *anything at all* can play the role of the undefined primitives of points, lines, planes, and so on, so long as the axioms are satisfied. Otto Blumenthal reports that, in a discussion in a Berlin train station in 1891, Hilbert said that in a proper axiomatization of geometry 'one must always be able to say, instead of "points, straight lines, and planes", "tables, chairs, and beer mugs"' (see Hilbert 1935: 388–429; the story is related on p. 403).

Hilbert (1899) sums up the idea as follows: 'We think of ...

points, straight lines, and planes as having certain mutual relations, which we indicate by means of such words as "are situated", "between", "parallel", "congruent", "continuous", etc. The complete and exact description of these relations follows as a consequence of the *axioms of geometry*.' To be sure, Hilbert also says that the axioms express 'certain related fundamental facts of our intuition', but in the subsequent development of the book all that remains of the intuitive content is the use of *words* like 'point', 'line', and so on (and the diagrams that accompany some of the theorems). Hilbert's protégée Paul Bernays (1967: 497) sums up the aims of Hilbert (1899):

> A main feature of Hilbert's axiomatization of geometry is that the axiomatic method is presented and practised in the spirit of the abstract conception of mathematics that arose at the end of the nineteenth century and which has generally been adopted in modern mathematics. It consists in abstracting from the intuitive meaning of the terms . . . and in understanding the assertions (theorems) of the axiomatized theory in a hypothetical sense, that is, as holding true for any interpretation . . . for which the axioms are satisfied. Thus, an axiom system is regarded not as a system of statements about a subject matter but as a system of conditions for what might be called a relational structure . . . [On] this conception of axiomatics . . . logical reasoning on the basis of the axioms is used not merely as a means of assisting intuition in the study of spatial figures; rather logical dependencies are considered for their own sake, and it is insisted that in reasoning we should rely only on those properties of a figure that either are explicitly assumed or follow logically from the assumptions and axioms.

The second of Hilbert's famous 'Mathematical Problems' (Hilbert 1900) extends the deductivist approach to every corner of mathematics:[6] 'When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relations subsisting

[6] In a lecture before the 1900 International Congress of Mathematicians in Paris, Hilbert presented twenty-three problems for mathematicians to tackle. The list provided much of the agenda for mathematics, and mathematical logic in particular, through much of the twentieth century. One of the most famous problems, the tenth, was to find an algorithm for determining whether a given diophantine equation has a solution over the natural numbers. This issue was only resolved when Matijacevič (1970) showed that there is no such algorithm.

between the elementary ideas of that science. The axioms set up are at the same time the definitions of those elementary ideas . . .'

One important development in this context, and with logicism, was that the formal languages and deductive systems were formulated with sufficient clarity and rigour for them to be studied as mathematical objects in their own right. That is, the mathematician can prove things *about* formal systems. Such efforts became known as *meta-mathematics*. Interest in meta-mathematical questions grew from the developments in non-Euclidean geometry, as a response to the failure to prove the parallel postulate. In effect (and with hindsight), the axioms of non-Euclidean geometry were shown to be consistent by describing a structure that makes them true.

Using techniques from analytic geometry, Hilbert (1899) constructed a model of all of the axioms using real numbers, thus showing that the axioms are 'compatible', or consistent. In contemporary terms, he showed that the axioms are satisfiable. If spatial intuition were playing a role beyond heuristics, this proof would not be necessary. Intuition alone would assure us that all of the axioms are true (of real space), and thus that they are all compatible with each another. Geometers in Kant's day would wonder about the point of proving 'compatibility' or satisfiability in this context. As we shall see in a moment, Frege also balked at it.

Hilbert then gave a series of models in which one of his axioms is false, but all the other axioms hold, thus showing that each axiom is independent of the others. The various domains of 'points', 'lines', and so on of each model are sets of numbers, sets of pairs of numbers, or sets of sets of numbers. Not quite tables, chairs, and beer mugs, but in the same spirit.

Presumably, this meta-mathematics is not itself the derivation of theorems from axioms regarded as meaningless. The goal of meta-mathematics is to shed light on a subject-matter, namely formal languages and axiomatizations. Thus, meta-mathematics seems to be an exception to the theme of deductivism (and game formalism), which holds that mathematics need have no subject-matter.

One option would be for the deductivist to hold that meta-mathematics is not mathematics, but this is close to an oxymoron. Meta-mathematics has the same appearances and methods as any other branch of mathematics. To be sure, meta-mathematics can be (and subsequently was) formalized. To be consistent, our deductivist should propose that the 'mathematics' in meta-mathematics is just the derivation of consequences from the axioms of this meta-mathematics, with these axioms regarded as meaningless. The 'application' of meta-mathematics to formal languages and deductive systems is irrelevant to its essence as a branch of mathematics. Just as arithmetic can be applied to counting, meta-mathematics can be applied to deductive systems. The role and importance of meta-mathematics varies among the formalist authors.

Frege and Hilbert carried on a spirited correspondence, which highlights the differences in their philosophical approaches to mathematics.[7] Frege asked about Hilbert's (1899) claim that his axiomatization provides *definitions* of the primitives of geometry, so that the very same sentences serve as axioms and definition. Frege tried to correct Hilbert on the nature of definitions and axioms. According to Frege, while definitions should give the *meanings* and fix the denotations of terms, axioms should express *truths*. In a letter dated 27 December 1899 Frege argued that Hilbert (1899) does not provide a definition of, say, 'between', since the axiomatization 'does not give a characteristic mark' that can be used to determine whether the relation 'between' holds:

> the meanings of the words 'point', 'line', 'between' are not given, but are assumed to be known in advance . . . [I]t is also left unclear what you call a point. One first thinks of points in the sense of Euclidean geometry, a thought reinforced by the proposition that the axioms express fundamental facts of our intuition. But afterwards you think of a pair of numbers as a point . . . Here the axioms are made to carry a burden that belongs to definitions . . . [B]eside the old meaning of the word 'axiom' . . . there emerges another meaning but one which I cannot grasp.

The idea of thinking 'of a pair of numbers as a point' refers to some of Hilbert's meta-mathematical theorems. For example, Hilbert showed that his axiomatization is consistent by constructing a Cartesian model in which 'points' are pairs of numbers. In the same letter, Frege told Hilbert that a definition should specify the meaning of a single word whose meaning has not yet been given, and the definition should employ other words whose meanings are

[7] The correspondence is published in Frege 1976 and translated in Frege 1980. See Resnik 1980, Coffa 1991: ch. 7, Demopoulos 1994, and Hallett 1994 for insightful analyses of it. See also Shapiro 1997: ch. 5.

already known. In contrast to definitions, axioms and theorems 'must not contain a word or sign whose sense and meaning . . . was not already completely laid down, so that there is no doubt about the sense of the proposition and the thought it expresses. The only question can be whether this thought is true . . . Thus axioms and theorems can never try to lay down the meaning of a sign or word that occurs in them, but it must already be laid down.' Frege's point is a simple dilemma: if the terms in the proposed axioms do not have meaning beforehand, then the statements cannot be true (or false), and thus they cannot be axioms. If they do have meaning beforehand, then the axioms cannot be definitions.

In contemporary terms, Hilbert provided *implicit*, or *functional* definitions of terms like 'point', 'line', and 'plane'. These are simultaneous characterizations of several items, in terms of their relations to each other. A successful implicit definition captures a structure (see Shapiro 1997: chs. 4, 5). Frege did not accept this notion, at least not as a *definition*.

Frege added that from the truth of axioms, 'it follows that they do not contradict one another' and so there is no further need to show that the axioms are consistent. That is, Frege did not see the point of Hilbert's meta-mathematics. The truth of the axioms is guaranteed by intuition, and there is no reason to show that they are consistent.

In reply, on 29 December, Hilbert told Frege that the purpose of the *Grundlagen* (1899) is to explore logical relations among the principles of geometry, to see why the 'parallel axiom is not a consequence of the other axioms' and how the fact that the sum of the angles of a triangle is two right angles is connected with the parallel axiom. I presume that Frege, the pioneer in mathematical logic, could appreciate *this* project. Concerning Frege's assertion that the meanings of the words 'point', 'line', and 'plane' are 'not given, but are assumed to be known in advance', Hilbert replied:

> This is apparently where the cardinal point of the misunderstanding lies. I do not want to assume anything as known in advance. I regard my explanation . . . as the definition of the concepts point, line, plane . . . If one is looking for other definitions of a 'point', e.g. through paraphrase in terms of extensionless, etc., then I must indeed oppose such attempts in the most decisive way; one is looking for something one can never find

because there is nothing there; and everything gets lost and becomes vague and tangled and degenerates into a game of hide and seek.

This is an allusion to 'definitions' like Euclid's 'a point is that which has no parts'. Hilbert claimed that such definitions do not help. These 'definitions' do not get used in the mathematical development. All we can do is specify the *relations* of points, lines, and planes to each other—via the axiomatization. All we can provide is an implicit definition of the terminology. To try to do better is to lapse into 'hide and seek'. Hilbert also responded to Frege's complaint that Hilbert's notion of 'point' is not 'unequivocally fixed':

it is surely obvious that every theory is only a scaffolding or schema of concepts together with their necessary relations to one another, and that the basic elements can be thought of in any way one likes. If in speaking of my points, I think of some system of things, e.g., the system love, law, chimney-sweep . . . and then assume all my axioms as relations between these things, then my propositions, e.g., Pythagoras' theorem, are also valid for these things . . . This circumstance is in fact frequently made use of, e.g. in the principle of duality . . . [This] . . . can never be a defect in a theory, and it is in any case unavoidable.

Note the similarity with Hilbert's quip in the Berlin train station.

Hilbert vehemently rejected Frege's claim that there is no need to worry about the consistency of the axioms, because they are all true: 'As long as I have been thinking, writing and lecturing on these things, I have been saying the exact reverse: if the arbitrarily given axioms do not contradict each other with all their consequences, then they are true and the things defined by them exist. This is for me the criterion of truth and existence.' Literally, Hilbert claimed that if a collection of axioms is consistent, then they are true and the things the axioms speak of exist. This makes for a sharp contrast to the way we think in other areas. A more cautious statement for Hilbert would be that the consistency of a collection of axioms is sufficient for them to constitute a legitimate branch *of mathematics*. Consistency is all the 'truth' and 'existence' that the mathematician needs.

In his response, dated 6 January 1900, Frege noted that Hilbert wanted 'to detach geometry from spatial intuition and to turn it into a purely logical science like arithmetic', and Frege was able to recapture much of Hilbert's perspective, in his own framework.

However, the two great minds remained far apart. Frege said that the only way to establish consistency is to give a model: 'to point to an object that has all those properties, to give a case where all those requirements are satisfied.' As we will see in the next section, the later Hilbert programme attempted to provide another way to establish consistency.

Frege complained that Hilbert's 'system of definitions is like a system of equations with several unknowns'. I think that Hilbert would accept this analogy. In the example at hand, three 'unknowns' are 'point', 'line', and 'plane'. We only get the relations among those. Frege wrote: 'Given your definitions, I do not know how to decide the question whether my pocket watch is a point.' Hilbert would surely agree, but he would add that the attempt to resolve this issue of the pocket watch is to play the game of hide and seek. Frege's issue here is reminiscent of the so-called 'Caesar problem' raised in his own logicism (see ch. 5, §1). For Frege, the sentence 'my pocket watch is a point' must have a truth value, and our theory must determine this truth value, just as the theory of arithmetic must determine a truth value to the equation '2 = Julius Caesar'.

Hilbert took the rejection of Frege's perspective on concepts—indicated by the pocket watch issue—to be a major *innovation*, and strength to his approach. In a letter to Frege dated 7 November 1903 he wrote that 'the most important gap in the traditional structure of logic is the assumption . . . that a concept is already there if one can state of any object whether or not it falls under it . . . [Instead, what] is decisive is that the axioms that define the concept are free from contradiction.' Showing some exasperation, Hilbert summed it up:

a concept can be fixed logically only by its relations to other concepts. These relations, formulated in certain statements I call axioms, thus arriving at the view that axioms . . . are the definitions of the concepts. I did not think up this view because I had nothing better to do, but I found myself forced into it by the requirements of strictness in logical inference and in the logical construction of a theory. I have become convinced that the more subtle parts of mathematics . . . can be treated with certainty only in this way; otherwise one is only going around in a circle.

3. Finitism: The Hilbert Programme

To paraphrase Dickens, mathematics at the turn of the twentieth century was 'the best of times, the worst of times'. Powerful and fruitful developments in real analysis, due to mathematicians like Augustin Louis Cauchy, Bernard Bolzano, and Karl Weierstrass, overcame the problems with infinitesimals and put the calculus on a solid foundation. Hilbert (1925: 187) wrote that real and complex analysis is 'the most aesthetic and delicately erected structure of mathematics'. Although infinitely small and infinitely large quantities were not needed, the new theories still relied on infinite collections. According to Hilbert, 'mathematical analysis is a symphony of the infinite'. At the same time, there was an exhilarating account of the infinite in Georg Cantor's set theory.

Despite these breathtaking developments, or because of them, there was a feeling of foundational crisis. Mathematics seems to be, and should be, the most exact and certain of all disciplines, and yet challenges and doubts were arising. In light of antinomies like Russell's paradox (see ch. 5, §§1–2), there was no certainty that the set theory was even consistent. The sense of crisis was not helped by Cantor's use of what he called 'inconsistent multitudes', collections of sets that are too big to be collected together into one set. The antinomies led to attacks on the legitimacy of some mathematical methods, leading some mathematicians to impose severe restrictions on mathematical methods, restrictions that would cripple real and complex analysis (see ch. 1, §2, ch. 5, §2, and ch. 7).

Hilbert's response to these developments incorporated aspects of deductivism, term formalism, and game formalism. Whatever its philosophical merits, the *Hilbert programme* led to a fruitful era of meta-mathematics that thrives today. For Hilbert, the programme had an explicit epistemic purpose: 'The goal of my theory is to establish once and for all the certitude of mathematical methods' (Hilbert 1925: 184). It would build on the early work in axiomatizing branches of mathematics, as well as the monumental efforts of logicists like Frege in developing rigorous logical systems:

There is . . . a completely satisfactory way of avoiding the paradoxes without betraying our science. The desires and attitudes which help us

> find this way . . . are these: (1) . . . [W]e will carefully investigate fruitful definitions and deductive methods. We will nurse them, strengthen them, and make them useful. No one shall drive us out of the paradise which Cantor has created for us. (2)We must establish throughout mathematics the same certitude for our deductions as exists in ordinary elementary number theory, which no one doubts and where contradictions and paradoxes arise only through our own carelessness. (Hilbert 1925: 191)

The idea behind the programme is to carefully and rigorously formalize each branch of mathematics, together with its logic, and then to study the formal systems to make sure they are coherent.

To describe the programme, we begin with its core, which is sometimes called 'finitary arithmetic'. Most emphatically, finitary arithmetic is not understood as a meaningless game (like chess), or as the deduction of consequences from meaningless axioms. On the contrary, the assertions of finitary arithmetic are meaningful, and they have a subject-matter.

The formulas of finitary arithmetic include equations like '$2 + 3 = 5$' and '$12{,}553 + 2{,}477 = 15{,}030$', as well as simple combinations of these, like '$7 + 5 = 12$ or $7 + 7 \neq 10$', or even '$2^{10,000} + 1$ is prime'. Notice that, so far, the only statements to be considered are those that refer to specific natural numbers, and that all of the properties and relations mentioned are *effectively decidable* in the sense that there is an algorithm (or computer program) that computes whether the properties and relations hold.

Consider the following two sentences:

(1) there is a number p greater than 100 and less than $101! + 2$ such that p is prime.[8]
(2) there is a number p greater than 100 such that both p and $p + 2$ are prime.

Both of these contain a *quantifier*, 'there is a number p', but there is a difference between them. The quantifier in sentence (1) is 'limited' to the (finitely many) natural numbers less than $100! + 2$. Call this a *bounded quantifier*. In contrast, the quantifier in sentence (2) has no limits, and so it 'ranges' over *all* natural numbers, an infinite collection. This is called an *unbounded* quantifier. Hilbert regards sentences with only bounded quantifiers to be finitary, while sentences, like (2), with unbounded quantifiers are not finitary.

[8] The number 101! is the result of multiplying 1, 2, 3, . . . , 101. It is very large.

Like the combinations of simple equations, sentences with only bounded quantifiers are effectively decidable, in the sense that there is an algorithm for computing whether they are true. Since the bounds can be very large, there is some idealization involved, but with bounded quantifiers there are only finitely many cases to be considered, and so such propositions represent computations. Sentences with unbounded quantifiers do not have this property. There is no limit to the number of cases to be considered, even in principle.

Hilbert introduces letters to represent generality. Consider the sentence:

$$(3)\quad \mathfrak{a} + 100 = 100 + \mathfrak{a}.$$

The instances of (3), like '$0 + 100 = 100 + 0$' and '$47 + 100 = 100 + 47$', are all legitimate, finitary statements. The sentence (3) says that each such instance is true. Hilbert regards such generalizations to be finitary. The commutative law thus has a finitary formulation:

$$(4)\quad \mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}.$$

The negation of an equation, like '$3 + 5 \neq 8$', is a legitimate finitary statement. It expresses the falsehood that the sum of 3 and 5 is not 8. However, it is not clear what to make of the negations of statements, like (3) and (4), that contain letters for generality. Hilbert (1925: 194) said that sentences with generality letters do not have finitary negations. He wrote: 'the statement that if $\mathfrak{a}$ is a numerical symbol, then $\mathfrak{a} + 1 = 1 + \mathfrak{a}$ is universally true, is from our finitary perspective *incapable of negation*. We will see this better if we consider that this statement cannot be interpreted as a conjunction of infinitely many numerical equations by means of "and" but only as a hypothetical judgment which asserts something for the case when a numerical symbol is given.' Thus, the negation of a statement of generality would assert that *there is* an instance—a numerical symbol—for which it is false. Similarly, the negation of (3) would say that *there is* a number p such that $p + 100$ is not identical to $100 + p$. Thus, the negation of a statement of generality contains an *unbounded* quantifier, and so is not finitary.

There is no serious epistemological issue concerning those finitary sentences that lack letters for generality. All such sentences represent routine (if long) computations, and so determining their truth value is only a matter of executing an algorithm (but see note 2 above). Hilbert is not explicit about how we legitimately come to assert finitary sentences that do have letters for generality, and there is disagreement among scholars as to the proof techniques in finitary arithmetic. The most common interpretation is that finitary arithmetic corresponds to what is today called 'primitive recursive arithmetic', but some take the extent of finitary methods to be more open-ended.[9]

Our next item concerns the *content* of finitary arithmetic. What is it about? Apparently, the subject-matter of finitary arithmetic is the natural numbers. So, once again, we ask what those are. Hilbert explicitly rejected the logicist perspective: 'we find ourselves in agreement with the philosophers, notably with Kant. Kant taught . . . that mathematics treats a subject matter which is given independently of logic. Mathematics, therefore, can never be grounded solely on logic. Consequently, Frege's and Dedekind's attempts to do so were doomed to failure' (Hilbert 1925: 192). Hilbert holds that finitary arithmetic concerns what is, in a sense, a *precondition* to all (human) thought—even logical deduction. Using Kantian language, Hilbert wrote that to think coherently at all,

> something must be given in conception, viz., certain extralogical concrete objects which are intuited as directly experienced prior to all thinking. For logical deduction to be certain, we must be able to see every aspect of these objects, and their properties, differences, sequences, and contiguities must be given, together with the objects themselves, as something which cannot be reduced to something else . . . This is the basic philosophy which I find necessary, not just for mathematics, but for all scientific thinking, understanding, and communicating. (Hilbert 1925: 192)

Hilbert proposed that the subject-matter of finitary arithmetic is 'the concrete symbols themselves, whose structure is immediately clear and recognizable'. He proposed that in finitary arithmetic, we identify the natural numbers with the 'numerical symbols':

$$|, ||, |||, |||| \ldots$$

He emphasized that, so understood, 'each numerical symbol is

[9] See any treatment of proof theory for an account of primitive recursive arithmetic (e.g. Smorynski 1977: 840 or, for a fuller treatment, Takeuti 1987). See also Detlefsen 1986 and Tait 1981.

intuitively recognizable by the fact that it contains only |'s'. The symbol '2' is then introduced as an abbreviation of '| |', etc. So the inequality '3 > 2 serves to communicate the fact that the symbol 3, i.e., | | |, is longer than the symbol 2, i.e., | |; or, in other words, that the latter symbol is a proper part of the former'.

Hilbert thus shows an affinity with what I call 'term formalism' (see §1.1 above). As with game formalism, the use of the word 'symbol' is misleading here. Hilbert is concerned with the characters themselves. In a sense, the numerical symbols symbolize themselves.

Despite the use of the word 'concrete', Hilbert intends the characters studied in finitary arithmetic to be understood more as abstract types than as physical tokens.[10] The physical hunk of ink (or burnt toner) | | is not a proper part of the physical hunk | | |. The two tokens occur at different locations in space, and so are distinct hunks. Notice also that Hilbert said that the 'concrete symbols' are 'given in conception' and 'intuited as directly experienced prior to all thinking'. Hilbert does not say that the concrete symbols are perceived. This is another indication that the 'concrete symbols' are not physical objects. He seems to have had something like Kant's form of intuition in mind (see Chapter 4, §2).

Hilbert also held that the subject of finitary arithmetic is essential to all human thought. Here as well we have seen similar ideas in Kant. The idea is that in order to think and reason at all, we have to use symbols and manipulate them in some fashion or other. Finitary arithmetic may not be absolutely incorrigible, or immune from doubt, but it is as certain as is humanly possible. There is no more preferred, or more epistemically secure, standpoint than finitary arithmetic (see Tait 1981).

To be sure, finitary arithmetic is only a small (and potentially trivial) chunk of the wonderful tapestry of mathematics. The first foray beyond finitary arithmetic consists of statements about natural numbers (or character types) that contain unbounded quantifiers. As above, this includes the negations of finitary statements that contain letters for generality. Then there is real analysis, com-

[10] See §1.1 above. In philosophical jargon, 'concrete' usually means 'physical' or 'spatio-temporal'. Mathematicians sometimes use the word 'concrete' for something more like 'specific', as opposed to 'general'. In this sense, number theory is more 'concrete' than the branches of abstract algebra like group theory.

plex analysis, functional analysis, geometry, set theory, and so on. Hilbert dubbed all of this 'ideal mathematics', to make the analogy with ideal points at infinity in geometry. Just as ideal points simplify and unify much geometry, so ideal mathematics allows us to streamline and deal more efficiently with finitary arithmetic. Therefore ideal mathematics is treated instrumentally:

> We . . . conclude that [the symbols and formulas of ideal mathematics] mean nothing in themselves, no more than the numerical symbols meant anything. Still we can derive from [the ideal formulas] other formulas to which we do ascribe meaning, viz., by interpreting them as communications of finitary statements. Generalizing this conclusion, we conceive mathematics to be a stock of two kinds of formulas: first, those to which the meaningful communications of finitary statements correspond; and, secondly, other formulas which signify nothing and which are the *ideal structures of our theory*. (Hilbert 1925: 196)

This ideal mathematics is to be treated formally, pretty much along the lines of *game* formalism (see §1.2 above). The syntax and rules of inference for each branch of ideal mathematics are to be formulated explicitly, and the branch is to be pursued as if it were just a game with characters. As Hilbert (1925: 197) put it, 'material deduction is thus replaced by a formal procedure governed by rules'. The 'rules' are those of the deductive systems developed by logicians like Frege.

Of course, ideal mathematics must be useful for finitary arithmetic. The only strict requirement on a formalized branch of ideal mathematics is that one cannot use it to derive a formula that corresponds to a false finitary statement. Suppose that T is a proposed formalization of some ideal mathematics and let Φ be any finitary statement, such as a simple equation. Then we should not be able to derive (a formula corresponding to) Φ in T unless Φ can be determined as true within finitary mathematics. In contemporary terms, the formal system T should be a *conservative extension* of finitary arithmetic.

Let us say that the formalized theory T is *consistent* if it is not possible to derive a contradictory formula, like '$0 = 0$ and $0 \neq 0$', using the axioms and rules of T. If every true finitary statement corresponds to a theorem of T and if T uses a standard deductive system (such as Frege's), then the conservativeness of T is equivalent to its consistency.[11] So the requirement on ideal mathematics is consistency.

The emphasis on consistency thus carries over from Hilbert's earlier deductivist writing (see the previous section). Recall that he wrote to Frege that 'if the arbitrarily given axioms do not contradict each other with all their consequences, then they are true and the things defined by them exist. This is for me the criterion of truth and existence.' Here, of course, the notion of 'consistency' is more fully articulated, and the philosophical role of consistency explicit.

Whether or not one follows Hilbert (or the term formalist) in *identifying* the natural numbers with their names, there is clearly a close structural connection between numbers and symbols. This connection has been exploited by logicians and other mathematicians ever since (see, for example, Corcoran *et al.* 1974). Crucially for the Hilbert programme, the identification of natural numbers with character types allows finitary arithmetic to be applied to *meta-mathematics*. That is, formal systems themselves now come under the purview of *finitary arithmetic*. As Hilbert put it, 'a formalized proof, like a numerical symbol, is a concrete and visible object. We can describe it completely.' And using finitary arithmetic, we can prove things about such formalized proofs.

Notice also that if T is a formalized axiomatic system, then the statement that T is consistent is itself finitary, formulable using a letter for generality. The statement that T is consistent has the form:

α is not a derivation in T whose last line is '$0 \neq 0$'.

The final stage of the Hilbert programme is to provide *finitary* consistency proofs of the fully formalized mathematical theories. That is, in order to use a theory of ideal mathematics we have to formalize it and then show, within finitary arithmetic, that the theory is consistent. Once this is accomplished for a theory T, then we have achieved the epistemic goal. We have maximal confidence that

[11] With standard logical rules, if Φ is a contradiction and Ψ is any formula, then 'if Φ then Ψ' is derivable. So if a formal theory T is inconsistent, then every formula can be derived in T. A fortiori, false finitary statements can be derived in T Conversely, let Φ be a true finitary statement, such as an equation, and suppose that the negation of Φ is a theorem of T. By hypothesis, both Φ and its negation are theorems of T, and so T is inconsistent.

using *T* will not bring us to contradiction, nor will it produce any false finitary statements. This is all that we can ask of an ideal mathematical theory. If *T* is a formalization of Cantorian set theory, then once we have a finitary consistency proof, we *know* with maximal certainty that we will not be driven from the paradise.

John von Neumann (1931) provided a succinct summary of the Hilbert programme, as involving four stages:

> (1) To enumerate all the symbols used in mathematics and logic . . .
>
> (2) To characterize unambiguously all the combinations of these symbols which represent statements classified as 'meaningful' in classical mathematics. These combinations are called 'formulas' . . .
>
> (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 . . . way) that those formulas which correspond to statements of classical mathematics which can be checked by finitary arithmetical methods can be proved . . . by the process described in (3) if and only if the check of the corresponding statement shows it to be true.

Items (1)–(3) call for the formalization of various branches of mathematics. This much was accomplished, brilliantly, and the study of the resulting formal systems is now a thriving branch of mathematical logic. Item (4), the crucial culmination, proved to be problematic.

4. Incompleteness

Kurt Gödel (1931, 1934) established a result that dealt a blow—many say a death blow—to the epistemic goals of the Hilbert programme. Let *T* be a formal deductive system that contains a certain amount of arithmetic. Assume that the syntax of *T* is *effective* in the sense that there is an algorithm that determines whether a given sequence of characters is a grammatical formula, and an algorithm that determines whether a given sequence of formulas is a legitimate deduction in *T*. Arguably, these conditions are essential for *T* to play a role in the Hilbert programme. Under these assumptions, Gödel showed that there is a sentence G in the language of T such that (1) if T is consistent, then G is not a theorem of T, and (2) if T has a property a bit stronger than consistency, called 'ω-consistency',[12] then the negation of G is not a theorem of T. That is, if T is ω-consistent, then it does not 'decide' G one way or another. This result, known as *Gödel's (first) incompleteness theorem*, is one of the major intellectual achievements of the twentieth century.

The formula G has the form of a finitary statement (using letters for generality). Roughly speaking, G is a formalization of a statement that G is not provable in T. So, if T is consistent, then G is true but not provable. Gödel's result thus dashes the hope of finding a single formal system that captures all of classical mathematics, or even all of arithmetic. If someone puts forward a candidate for such a formal system, then we can find a sentence that the system does not 'decide', although we can see that the sentence is true.

The incompleteness theorem thus raises doubts about any philosophy of mathematics (formalist or otherwise) that requires a single deductive system for all of arithmetic—a single formal method for deriving every arithmetic truth.[13] However, the dream of finding a single formal system for all of ideal mathematics was not an official (or essential) part of the Hilbert programme. The trouble, if that is what it is, comes elsewhere.

Gödel showed that the reasoning behind the incompleteness theorem can be reproduced *within* the given formal system T. In particular, if the formalization of 'provable in T' meets some straightforward requirements, then we can derive, in T, a sentence that expresses the following:

If T is consistent, then G is not derivable in T.

[12] An arithmetic theory T is ω-consistent if there is no formula $\Phi(x)$ such that $\Phi(0)$, $\Phi(1)$, $\Phi(2)$, . . . , are all provable as well as a statement that there is a natural number x such that $\Phi(x)$ fails. J. Barkley Rosser (1936) proved a result similar to Gödel's from the weaker assumption that T is consistent.

[13] Although one might argue that the original Fregean logicism would not be successful without such a deductive system, contemporary neo-logicists are not committed to a claim that there is a single deductive system that yields every arithmetic truth (see ch. 5, §§1, 4).

But, as noted above, '*G* is not derivable in *T*' is equivalent to *G*. So, we can derive, in *T*, a sentence to the effect that

If *T* is consistent then *G*.

Assume that *T* is consistent, and that we can derive, in *T*, the requisite statement that *T* is consistent; then it would follow that we can derive *G* in *T*. This contradicts the incompleteness theorem. So if *T* is consistent, then one cannot derive in *T* the requisite statement that *T* is consistent. This is known as Gödel's *second incompleteness theorem*. Roughly, it asserts that no consistent theory (that contains a certain amount of arithmetic) can prove its own consistency.

This result does indicate trouble for the Hilbert programme. Let PA be a formalization of (ideal) arithmetic, say the classical theory of the natural numbers. The Hilbert programme requires a *finitary proof* of the consistency of PA. But the second incompleteness theorem is that if PA is in fact consistent, then a straightforward statement of the consistency of PA is not derivable in PA itself, let alone in the finitary portion of PA. The same goes for any other formal system, so long as it contains a certain amount of arithmetic. The Hilbert programme requires a finitary proof that the deductive system is consistent, and yet, it seems, the consistency cannot be proved in the system itself, let alone in a more secure subsystem.

A much-discussed paper (Gödel 1958) opens by paraphrasing Bernays:

> since the consistency of a system cannot be proved using means of proof weaker than those of the system itself, it is necessary to go beyond the framework of what is, in Hilbert's sense, finitary mathematics if one wants to prove the consistency of classical mathematics, or even that of classical number theory . . . [I]n the proofs we make use of insights . . . that spring not from the combinatorial (spatiotemporal) properties of the sign combinations . . . but only from their *meaning*.

Gödel pointed out that since we have no 'precise notion of what it means to be evident', we cannot rigorously prove Bernays's claim, but Gödel added that 'there can be no doubt that it is correct'.

There is a near, but not universal, consensus on the Bernays–Gödel conclusion. A post-Gödel defence of a Hilbert-style programme has at least two options. One is to challenge the formalization of consistency used in the proof of the second incompleteness theorem. There are other ways to express consistency-properties that escape the second incompleteness theorem (see Feferman 1960, Gentzen 1969, and Detlefsen 1980). The issue, then, turns on just what counts as expressing consistency, and what a proof of consistency must show in order to meet the epistemic goals of the Hilbert programme.

A second option would be to show, or claim, that the methodology of finitary arithmetic cannot be captured in PA or in any other formalized theory. That is, even though the purpose of a branch of ideal mathematics is to streamline the derivation of finitary statements, the proof-methods of any given formalized theory do not include every finitary proof-method. The thesis is that finitary arithmetic is *inherently informal*. See Detlefsen 1986.

5 Curry

Any contemporary philosophy of mathematics that relies heavily on the rigorous formalization of mathematical theories thereby shows some influence of formalism, and probably owes a debt to the Hilbert programme. Although formalism still has advocates among mathematicians, after the 1940s (or so) few philosophers and logicians explicitly avowed it. A notable exception is Haskell Curry.

Curry's philosophy begins with an observation that, as a branch of mathematics develops, it becomes more and more rigorous in its methodology, the end result being the codification of the branch in a formal deductive system. Curry takes this process of formalization to be the essence of mathematics.

He argues that all other philosophies of mathematics are 'vague' and, more importantly, they 'depend on metaphysical assumptions'. Mathematics, he claims, should be free from any such assumptions, and he argues that the focus on formal systems provides this freedom. He thus echoes Thomae's claim that formalism has no extraneous metaphysical assumptions.

The main thesis of Curry's formalism is that assertions of a mature mathematical theory be construed not so much as the results of moves *in* a particular formal deductive system (as Hilbert or a game formalist might say), but rather as assertions *about* a formal system. An assertion at the end of a research paper would be interpreted as something in the form 'Φ is a theorem in formal system *T*'. For Curry, then, mathematics is an objective science, and it has a subject-matter. He wrote that 'the central concept in mathematics is that of a formal system' and 'mathematics is the science of formal systems' (Curry 1954). Curry is thus allied more with term formalism than with game formalism. An appropriate slogan is that mathematics is meta-mathematics.

Unlike Hilbert, however, Curry does not restrict meta-mathematics to finitary arithmetic: 'In the study of formal systems, we do not confine ourselves to the derivation of elementary propositions step by step. Rather, we take the system . . . as datum, and . . . study it by any means at our command' (Curry 1954). Curry concedes that some 'intuition' is involved in this meta-mathematics, but he claims that 'the metaphysical nature of this intuition is irrelevant'.

Stepping back one level, on Curry's view, meta-mathematics is itself a branch of mathematics. As such, the meta-mathematics should be formalized. That is, the non-finitary results in meta-mathematics (like most of contemporary mathematical logic) are accommodated by producing a formal system for meta-mathematics, and construing the results in question as theorems about that formal system. Presumably, this does not constitute a vicious infinite regress.

For Curry, there is no real issue concerning the *truth* of a given formal system. Instead, there is only a question of 'considerations which lead us to be interested in one formal system rather than another'. This matter of 'interest' is largely pragmatic: 'Acceptability is relative to a purpose, and a system acceptable for one purpose may not be for another.'[14] Curry mentions three 'criteria of acceptability' for formal systems: '(1) the intuitive evidence of the premisses; (2) consistency . . .; (3) the usefulness of the theory as a whole' (Curry 1954).

Of course the second criterion, consistency, is important. An inconsistent formal system has limited use (assuming a standard logic, see note 11 above). Unlike Hilbert, however, Curry does not require a *proof* of consistency:

[14] Curry's notion of acceptability is quite similar to Carnap's 'external question' concerning the acceptability of a 'linguistic framework' (e.g., Carnap 1950). See chapter 5, §3.

> The criterion of consistency has been stressed by Hilbert. Presumably, the reason for this is that he . . . seeks an *a priori* justification. But aside from the fact that for physics the question of an *a priori* justification is irrelevant, I maintain that a proof of consistency is neither a necessary nor a sufficient condition for acceptability. It is obviously not sufficient. As to necessity, so long as no inconsistency is known, a consistency proof, although it leads to our knowledge about the system, does not alter its usefulness. Even if an inconsistency is discovered this does not mean complete abandonment of the theory, but its modification and refinement . . . The peculiar position of Hilbert in regard to consistency is thus no part of the formalist conception of mathematics . . . (Curry 1954)

Since there is no need to prove consistency before accepting a formal system, Curry's philosophy is not affected by Gödel's second incompleteness theorem. Since Curry does not restrict mathematics to a single formal system, his views are also unaffected by Gödel's first incompleteness theorem.

Like most formalists, Curry seems to require that every legitimate branch of mathematics be formalized. What is the formalist (or deductivist) to make of the practice of, say, arithmetic, before it was formalized in the nineteenth century? Were Archimedes, Cauchy, Fermat, and Euler not doing mathematics? On the contemporary scene, what is the status of *informal* mathematical practice, which does not explicitly invoke a rigorous deductive system? Indeed, what is the status of informal *meta*-mathematics?

Opponents of Curry-style formalism question the philosophical significance of the observation that as a branch of mathematics develops and becomes rigorous, it gets formalized. With Frege and Gödel, some philosophers maintain that something essential is *lost* in the formalism. Mathematical language has meaning and it is a gross distortion to attempt to ignore this meaning. At best, formalism and deductivism focus on a small aspect of mathematics, deliberately leaving aside what is essential to the enterprise. In the next chapter, we turn to a philosophy that insists that mathematics is inherently informal.

6. Further Reading

Many of the primary sources noted above are available in English translation. Geach and Black 1980: 162–213 contains a translation of the sections (§§86–137) of Frege 1893 on formalism (i.e., concerning Thomae and Heine). Benacerraf and Putnam 1983 contains translations of von Neumann 1931 and Hilbert 1925 (the above quoted passages from Hilbert 1925 are from that version). Van Heijenoort 1967 contains another translation of Hilbert 1925, as well as a translation of Hilbert 1904 and 1927. Other relevant papers are Hilbert 1918, 1922, and 1923. See also Hilbert and Bernays 1934. Curry 1954 is also reprinted in the Benacerraf and Putnam 1983 anthology, with a note indicating that this paper represents his views in 1939. Curry 1958 is a fuller elaboration of his mature formalism. Resnik 1980: chs. 2,3 is an excellent secondary source on the various types of formalism (and Frege's critique of game formalism). For a sample of the large literature on the Hilbert programme, see Detlefsen 1986, Feferman 1988, Hallett 1990, Sieg 1988, 1990, Simpson 1988, and Tait 1981. Bernays 1967 is a lucid and sympathetic reconstruction of Hilbert's views. Reid 1970 is a book-length intellectual biography of Hilbert.