Hilbert's Programme

C. Smoryński
429 S. Warwick
Westmont, IL 60559
USA

The present paper is the introduction to the fourth chapter, that dealing with Gödel's Incompleteness Theorems, of a book in progress on the metamathematics of arithmetic. This introduction covers Hilbert's programme and the *Grundlagenstreit* of which the programme was a central feature. This introduction does not entirely stand alone, nor was it intended to. However, so much has been written about Hilbert's programme by those who have not read Hilbert's papers that I thought, imperfect as it is, the present account might serve the purposes of countering the nonsense promulgated by the ignoscenti and of correcting the historical error of identifying Hilbert's late advertisement for the programme (a proof of consistency yields conservation [A strong theory T is conservative over a weaker theory S with respect to some class of formulae Γ if every formula $\varphi\in\Gamma$ which is provable in T is provable already in S]) with the rationale for the programme. Such an identification appears to have been a commonplace by 1930, and has been uncritically accepted by modern expositors (e.g. by me in the *Handbook of Mathematical Logic* and by Girard in his recent text on proof theory). A study of Hilbert's papers reveals his early belief in such conservation, but also reveals that the recognition that his consistency programme yields conservation came relatively late. This historical fact has, however, no relevance to the simple observation - first made by Kreisel? - that the First Incompleteness Theorem, by blocking the conservation result effectively kills Hilbert's programme; the Second Incompleteness Theorem is merely a refinement.

The present account is based primarily on a careful reading of all (but the very last) of Hilbert's papers on foundations, less careful readings of Brouwer's and Weyl's papers of the period, secondary sources, hearsay, and heresy. As I consider it merely a first approximation to an historical account of Hilbert's programme and the *Grundlagenstreit*, I have not bothered to document the sources other than the papers of Hilbert et al.

Two important sources must, however, be singled out: First, I have relied heavily on Dirk van Dalen's knowledge of Brouwer and Brouwer's relation to Hilbert. Without van Dalen, the present account would not have been half as complete as it is. Second, I have generally not consulted Bernays' writings on the subject as, for the most part, these came (or, at least, were published) *after* the end of the whole affair. A more definitive study of Hilbert's programme will have to take Bernays more fully into account.

The thought, perhaps - the wish, the will -
Those I could understand; but really
To do the deed! Ah, no - that beats me!
Peer Gynt

The story I want to tell is one of a great controversy, a fight between two mathematical geniuses on the course of mathematical research. The antagonists in the dispute are David Hilbert, whose philosophical views we are already partially familiar with, and Luitzen Egbertus Jan Brouwer, whose views are much more difficult to describe. We will come to him a bit later. Those playing lesser rôles in the dispute include initially Gottlob Frege and Henri Poincaré; later Paul Bernays and Hermann Weyl; then Wilhelm Ackermann and Johann von Neumann; and ultimately Kurt Gödel.

The dispute between Hilbert and Brouwer was, to some extent, a retelling of that between Leopold Kronecker and Georg Cantor. Indeed, this is how Hilbert viewed it. It was, however, both more and less than this. It was a bit less in that it was, to some extent, a fight between Hilbert and the ghost of Kronecker come back from the grave to haunt the sexagenarian Hilbert. It was more in that both the Kronecker-Cantor and Hilbert-Brouwer controversies were just episodes - admittedly the most outstanding ones - in a general antagonism between the cautious mathematical conservatives and the daring mathematical liberals. Prior to the Kroneckerian spectre there had been, for example, an early 19th century battle in England over the acceptability of negative numbers; and even today there are those who would proscribe mathematical practice. The difference between the Hilbert-Brouwer dispute and these other episodes is the mathematics that developed out of it.

Hilbert's first brush with controversy was in connection with his first major result. This was Hilbert's solution in 1888 to Gordan's Problem in the theory of invariants, in which Hilbert proved, without construction, the existence of a finite basis for any ideal in the polynomial ring $K[X_0, \ldots, X_{n-1}]$ over a field K. Hilbert drew fire for the failure of his proof to provide such a basis even in the simplest examples: Some denied he had a proof; his formal thesis advisor, Ferdinand Lindemann, called the proof 'unheimlich'; and Paul Gordan labelled the proof 'theology'. The extent to which this controversy traumatised the young Hilbert - in his twenties and only a few years from having received his doctorate - was to be quite apparent three and a half decades later when Hilbert set out to remove Brouwer from the mathematical landscape.

History - even chronology - tends to have an annoyingly non-linear character. Another, apparently independent, prologue to the Hilbert-Brouwer dispute was Hilbert's work on the foundations of geometry. The earliest recorded comment of his on the topic was an aphoristic remark he made concerning a lecture on geometry he had just attended: 'It must be possible to replace in all geometric statements the words *point, line, plane* by *table, chair, mug*'. This remark, encapsulating, as Hermann Weyl put it, 'the axiomatic standpoint in a nutshell', was made in 1891. In the next decade, Hilbert was to think deeply on matters both geometrical and foundational, lecturing several times on geometry, and finally, in 1899, publishing his lectures on Euclidean geometry, the *Grundlagen der Geometrie.*

A first approximation to an understanding of the foundational importance of Hilbert's book on the foundations of geometry is to be had by comparing Hilbert's work to Euclid's *Elements.* Although Euclid's *Elements* had been regarded as the epitome of rigour for two millenia, the fact is that Euclid's treatment is very far from being rigorous; there are gaps in Euclid's reasoning one could drive a truck through. It seems Euclid did not replace the (Greek equivalents of) 'point', 'line', and 'plane' by 'table', 'chair', and 'mug' (nor, as Hilbert also suggested, by 'love', 'law', and 'chimney-sweep'); Euclid allowed himself to use evident properties of points, lines, and planes not explicitly assumed in his axioms. Hilbert's treatment of geometry may be viewed as a stopping of Euclid's gaps. Hilbert gave a complete axiomatisation of geometry and used only the axioms cited in his further development of geometry. His treatment was fully rigorous.

Describing Hilbert's *Grundlagen der Geometrie* as a rigorisation of Euclid's *Elements*, or as the first purely deductive treatment of geometry, is, however correct, a superficial description from the foundational point of view. As Hermann Weyl put it, 'It is one thing to build up geometry on sure foundations, another to inquire into the logical structure of the edifice thus erected. If I am not mistaken, Hilbert is the first who moves freely on this higher "metageometric" level'. The *Grundlagen* is a milestone in the history of the axiomatic method. It marks a break with the old axiomatic viewpoint and a full acceptance of the newer, not previously recognised one.

In the traditional Aristotelian approach to axiomatisation, the axioms are evident truths about the structure being axiomatised, and truth is borne from the axioms to the theorems via the proofs. In the 18th and 19th centuries, this view was slowly changing. In 1733, Gerolamo Saccheri attempted to derive Euclid's parallel postulate from the remaining ones by assuming the negation of the postulate and deriving a contradiction therefrom. Mistakenly, he thought he had succeeded. What he had really done, and what was consciously done in the following century by various mathematicians (including Ferdinand Karl Schweikert, Carl Friedrich Gauss, Janos Bolyai, Nicolai Lobachevskii, and Bernhard Riemann), was to develop axiomatically a consistent non-Euclidean geometry. By the mid-19th century, with three major geometries and only one physical space to be described by them, it was clear that the geometries couldn't all have true axioms. Whereas initially most mathematicians ignored the non-Euclidean geometries, those familiar with these new geometries recognised their coherence. While some asked which of the three gave a correct description of space, no-one denied the significance of any of the geometries. This became especially true in the latter half of the 19th century when models of the non-Euclidean geometries were found.

A minor episode of a similar nature occurred in England in the early 19th century in response to questions of the validity of the use of negative numbers. In the face of such doubts, William Frend attempted an algebra of non-negative quantities. Such an algebra is an awkward reversion to the practice of the 16th century by which, for example, the equations,

$$x^3 = ax+b \quad \text{and} \quad x^3+ax = b,$$

are of completely different character (as a,b must be non-negative) and had to be solved separately. In response to this, George Peacock proposed his 'symbolical algebra'. In it, he acknowledged that he had no idea of what 'negative quantities' were, and proposed an algebra not dealing with quantities at all. His was a purely symbolic system of algebraic manipulation based on simple rules (like commutativity, associativity, etc.). Further British developments - Hamilton's invention of quaternions with their non-commutative multiplication, Cayley's numbers with their non-commutative and non-associative multiplication, etc. - revealed the extreme axiomatic freedom mathematics could offer.

In short, there was a shift from the Aristotelian conception of axiomatics by which axioms were true about a given structure to the modern mathematical conception by which a consistent set of axioms determines the subject matter under study. Hilbert's *Grundlagen der Geometrie* is the first major work written in this completely modern spirit; it barely casts a nostalgic glance at the old practice. Central to the book is not so much an axiomatic development of geometry, but a study of the axiomatics of geometry. The axioms are divided into various groups and the relevance of the various axiom groups for certain important results is examined. Further, the axioms are proven independent and the consistency of the system as a whole is established by modelling the geometric axioms in the arithmetic of the reals.

When the definitive account of the Hilbert-Brouwer dispute is finally written, it will devote an entire chapter to the *Grundlagen der Geometrie* - and Hilbert's resulting correspondence on this book with the philosopher Gottlob Frege. For, it is in this correspondence that Hilbert, as far as his lack of patience with philosophy allowed him, explained his views on axiomatics and what he intended to accomplish. Given the limitations of space, I shall merely offer two quotes from Hilbert's responses to Frege. These quotes do not encompass the full range of Hilbert's thoughts on the matter and are not truly representative of his views; they do, however, illustrate how modern his outlook was. The first is, more-or-less, his dismissal of Frege and ends their correspondence on the subject:

> My opinion is simply this, that a concept can only be logically fixed through its relations to other concepts. These relations, formulated in precise statements, I call axioms and I add, that the axioms (possibly including the giving of names to the concepts) are the definitions of the concepts. I have not thought out this view to pass the time, but rather I saw myself pressed to it through the demand of rigour in logical inference and in the logical construction of a theory. I have come to the conviction that in mathematics and the natural sciences one can handle more subtle things with security only in such a way; otherwise one merely turns in circles.

So, rigour requires the use of the axiomatic method; but what check is there on the axiom systems? Hilbert had answered this question some months earlier:

> If the arbitrarily given axioms do not contradict one another with all their consequences, then they are true and the things defined by the axioms exist. This is for me the criterion of truth and existence.

In his own axiomatic work, Hilbert allowed himself some liberties in choosing the axioms - particularly in relaxing the Aristotelian requirement of the evident truth of the axioms - but he was far from arbitrary in such choice of axioms. He had, if not criteria, guidelines in the selection of axioms. Completeness and simplicity were two desiderata he cited in the introduction to the *Grundlagen*; consistency was, of course, another. None of these desiderata was entirely unproblematic.

Hilbert's knowledge of logic, indeed logic itself, was then too vague for us to determine exactly what Hilbert meant at the time by (axiomatic) completeness. Hilbert's own axiomatisations are complete in that they are categorical: Any two models are isomorphic. I cite here his remark on completeness from the second of his problems in his 1900 Paris lecture:

> 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 between the elementary ideas of that science. The axioms so set up are at the same time the definitions of those elementary ideas; and no statement within the realm of the science whose foundations we are testing is held to be correct unless it can be derived from those axioms by means of a finite number of logical steps.

One could take the 'exact and complete description' to be complete enough to decide the truth or falsity of every statement. *Semantically*, such completeness follows from categoricity; *syntactically*, such completeness - derivation from the axioms by finitely many logical steps - fails, as is now well-known. Ultimately, Hilbert would definitely mean completeness in this syntactic sense: Every sentence or its negation would be derivable from the axioms in finitely many steps.

Possibly the only one who understood the distinction between syntax and semantics at the turn of the century was Gottlob Frege; the equivalence - for first-order logic - would first be posed by Hilbert as a problem in the late 1920s and solved by Gödel in 1929. Hilbert's axiomatisation of geometry in the *Grundlagen* and his axiomatisation of arithmetic published in 1900 were not first-order. Indeed, each had a second-order Archimedean axiom and both (the axiomatisation of arithmetic and the axiomatisation of geometry given in later editions of the *Grundlagen*) had a 'completeness axiom' asserting that the structure under consideration was maximal with respect to the remaining axioms, i.e. no proper extension of the structure could satisfy the remaining axioms. This 'completeness axiom' is not even properly a second-order axiom, but goes beyond that logic.

What Hilbert meant by the simplicity of an axiomatisation is unclear, and not relevant to our discussion. I remark only that, both in the *Grundlagen* and his paper on axioms for real arithmetic, he paid close attention to the dependence and independence of the axioms given, but that, given the elegance of his work, 'simplicity' must have had an aesthetic component as well.

Although the meaning of consistency was clear - a theory is consistent if it doesn't derive two contrary assertions (equivalently, if it leaves some assertion unproven) - the necessity of and means for proving consistency were unclear. The attitude of most mathematicians, even today, is that, since the axioms are true and derivations preserve truth, consistency is obvious. Frege had said as much in his correspondence with Hilbert. He had not experienced, as Hilbert had, the opposition to the results of mathematical proof, nor (apparently) the strong rejection by Kronecker of the infinite in mathematics - which rejection would preclude any proof of the consistency of theories of the infinite by reference to models. The problem, as Frege asked Hilbert, was: Is there any other way of proving consistency than by exhibiting a model?

Earlier mathematicians had proven the consistency of non-Euclidean geometry by modelling it in Euclidean geometry, and Hilbert had reduced the consistency of the latter to that of his axioms for the arithmetic of the reals. In the second of his 23 problems at Paris, Hilbert re-iterated this and added that a 'direct proof' of the consistency of this arithmetic was now needed. The only comment he offered on how such a proof could be given was so vague ('I am convinced that it must be possible to find a direct proof for the compatibility of the arithmetical axioms, by means of a careful study and suitable modification of the known methods of reasoning in the theory of irrational numbers.') that it is impossible to decide whether or not his later attempts conform with this remark. His earliest consistency proofs follow a different line.

Before Hilbert published his first direct consistency proofs, Bertrand Russell discovered his famous paradox of the set of all sets which are not elements of themselves. If

$$R = \{x: x \notin x\},$$

then

$$R \in R \quad \text{iff} \quad R \notin R.$$

Russell communicated his result to Frege, in whose axiomatic system the paradox was derivable. Frege went ahead with the publication of the second volume of his book on the inconsistent system, but added an appendix including Russell's paradox and acknowledging there to be serious flaws in his work.

News of Russell's paradox created quite a stir. Philosophers uncovered newer paradoxes, and even an occasional mathematician took them seriously. Richard Dedekind, whose *Was sind und was sollen die Zahlen?* used the set of all sets, delayed publication of the third edition of his book which was due to come out in 1903 until 1911 because he was concerned about the foundations of his treatment.

Hilbert was largely unperturbed. Writing to Frege in 1903 in thanks for a copy of Frege's book, he noted that his student Ernst Zermelo had shown

Hilbert the Russell paradox some 3 or 4 years earlier and he himself had known other, more convincing ones for 4 to 5 years. Indeed, we now know that Georg Cantor had written to him as early as 1897 to explain that there is no set of all cardinal numbers - a fact Hilbert cited in his 1900 paper on axiomatising arithmetic and in his 1900 problems address. The paradoxes convinced Hilbert that, just as the paradoxes in calculus had engendered a need for rigour in that field, one needed rigour in all foundational work. As Hilbert had said in his problems paper, 'the requirement of logical deduction by means of a finite number of processes is simply the requirement of rigour in reasoning. Indeed the requirement of rigour, which has become proverbial in mathematics, corresponds to a universal philosophical necessity of our understanding'. The paradoxes also convinced Hilbert that, contrary to Frege's programme and to what would become Russell's programme of reducing mathematics to logic, logic alone was insufficient for the foundations of mathematics.

In August 1904, at another International Congress of Mathematicians, this one in Heidelberg, Hilbert offered his first examples of consistency proofs. Because logic already uses arithmetic notions, Hilbert insisted on simultaneously developing logic and arithmetic. The fragments of mathematics shown consistent are also logically fragmentary and the results are quite unimpressive. They do, however, illustrate both his approach and its weakness, and it will be instructive to consider an example.

We begin by specifying a (non-first-order) language as follows. First, there is one constant $\bar{1}$ and three function symbols: I (for an infinite set), S (for successor), and S' (an accompanying operation); there are no variables. Terms are built up in the usual way. The only relation symbol is that for equality and equations between terms constitute the atomic formulae. Equations can be negated. We also use parentheses for readability's sake, but these are not part of the formal language.

The axioms and rules of inference are as follows: For any terms t,u and any formula φ,

AXIOMS.
$$t = t$$
$$S(It) = I(S't)$$
$$\neg(S(It) = I\bar{1})$$

RULES.
$$\frac{t = u \quad \varphi(t)}{\varphi(u)}$$
$$\frac{S(It) = S(Iu)}{It = Iu}$$

THEOREM 1. *The system described is consistent.*

PROOF. Define an equation $t=u$ to be *homogeneous* if t and u contain the same number of symbols (not counting parentheses, which are only used for readability's sake). A simple induction on the number of inferences applied shows any derivable equation to be homogeneous. (Such also shows that the rule of substitution of equals need only be applied to equations in deriving equations.) It follows that

$$S(It) = I\overline{1}$$

is not derivable, and the system is consistent. QED

The range of the function I, i.e. the set of terms $I(\overline{1})$, $I(S'\overline{1})$, $I(S'S'\overline{1})$, . . ., is infinite in a weak sense: The addition of any equation of distinct such terms to the system results in an inconsistent system. Had the system included some logic, this would have meant the derivability of the inequations; as formulated, however, the system cannot prove even $\neg I(S'\overline{1}) = I(S'S'\overline{1})$. Nonetheless, Hilbert maintained that Theorem 1 proved the existence of an infinite set and that this method showed the incorrectness of the widely held view that the elements of a set must have their existence prior to the existence of the set itself. Through the axiomatic method, both the set (here represented by I) and its elements (here represented by the terms $I(t)$) come into existence simultaneously with the proof of consistency.

The system of Theorem 1 was only intended as an example and we should thus be careful in criticising the proof for what it doesn't accomplish. That the system does not accommodate simple propositional logic (i.e. logic without quantifiers) is not a critique, but a desideratum for extending the result. Similarly, that it does not include any arithmetic beyond the trivial successor function ought not to be put forth as an objection; more arithmetic can be added to the system with the extension then to be proven consistent. Hilbert viewed mathematics and foundations as an open-ended enterprise. One would start with something like the infinite set I and its proof of consistency, and add additional structure and axioms, proving the consistency of such as one went along. Indeed, in his paper of 1904, Hilbert attempted an example of the extension of the system by adjoining an abstract set and some rules of inference, but this adjunction is (at least to me) not very convincing.

One real criticism - valid of practically all of Hilbert's contemporaries as well - is that Hilbert was fundamentally confused about the nature of logic. He included the notion of set in logic, held second-order axiomatisations like Dedekind's as paradigms, and spoke of logical derivations proceeding in finitely many steps. Yet the system he treated was virtually logic-free and, when he did mention adding some logic to the system, it was purely propositional logic. At this point in time, only Frege had any experience in working within a formal logical system - and his was inconsistent. At the end of the decade, Bertrand Russell and Alfred North Whitehead would be working on *Principia Mathematica,* a monumental case study in the formalisation of mathematics in logical systems. But it really wouldn't be until the 1920s that the nature of the quantifier and the distinction between first-order logic on the one hand and second-order logic and set theory on the other would be understood. A good part of this understanding would come from Hilbert and his disciples. [For the non-logician: First-order logic deals with the elements of arbitrary, but fixed given structures. It allows one to say things like, 'there is an element such that...' or 'for any element...'. In second-order logic, one can also talk about arbitrary subsets of the given structure. Thus, in Dedekind's axiomatisation of the natural numbers, one can state that all non-empty sets of natural numbers have least elements. In first-order arithmetic, one can only schematically assert that each definable such set has a minimum.]

Having said all of this, it must yet be stated that Hilbert had achieved something - he had shown how the consistency of a 'theory' could be recognised without the construction of a model - and it must yet be stated that there was one bit of criticism that was levelled and that Hilbert could not at the time answer. This was Henri Poincaré's observation that the proof of Theorem 1 used induction: Hilbert's stated goal of providing a 'rigorous and completely satisfying foundation for the notion of number' rested on the notion of number itself. For Poincaré, induction was a special mathematical intuition which did not need any justification - a fortunate thing since any justification, like Hilbert's, required induction and was therefore circular. It would be over a decade and a half before Hilbert offered a formal response to Poincaré's critique, and by then Poincaré would be a side issue. For, by then, Hilbert would already be locked in battle with Brouwer.

The contrasts between Hilbert and Brouwer were many. Hilbert had little patience with philosophy, his own philosophy of mathematics being perhaps best described as naïve optimism - a faith in the mathematician's ability to solve any problem he might set for himself. His interest in the foundations of mathematics was largely a concern for rigour - establishing the rules of the game as it were; his proposal of consistency as the criterion of existence was ostensibly a libertarian suggestion intended to allow maximum freedom to the mathematician. As he was eventually to say, his foundational goal was to abolish foundational questions once and for all, i.e. to allow mathematicians to get back to business-as-usual. Brouwer, on the other hand, was as much a philosopher as a mathematician and was deeply interested in the foundations of mathematics for its own sake. Moreover, unlike Hilbert, who took the central problem of foundations to be the clarification and justification of (then) current practice, Brouwer took the epistemological problem to be central. According to Brouwer we have one fundamental mathematical intuition and mathematics is the working out of the consequences of this intuition. Those parts of mathematics that cannot be based on this fundamental intuition must be rejected.

I must be careful here in referring to Brouwer's rejection of parts of mathematics. Brouwer was not the cardboard replica of Kronecker that Hilbert was to caricature him as. When in 1912 Brouwer became Extraordinary Professor in Amsterdam with the recommendations of Hilbert and Felix Klein in Göttingen and Poincaré and Emile Borel in Paris, he chose as the topic of his inaugural address the two philosophies, intuitionism and formalism, represented by himself and Hilbert, respectively. A quote from the English translation of the following year will illustrate the dispassionate style of the prose and the level of objectivity Brouwer could bring to bear on such issues:

> On what grounds the conviction of the unassailable exactness of mathematical laws is based has for centuries been an object of philosophical research, and two points of view may here be distinguished, intuitionism (largely French) and formalism (largely German). In many respects these two viewpoints have become more and more definitely opposed to each other; but during recent years they have reached agreement as to this, that the exact validity of mathematical laws as laws of nature is out of the question. The question where mathematical exactness does exist, is answered by the two sides; the intuitionist says: in the human intellect, the formalist says: on paper.

Another quote:

> In the domain of finite sets in which the formalist axioms have an interpretation perfectly clear to the intuitionists, unreservedly agreed to by them, the two tendencies differ solely in their method, not in their results; this becomes quite different however in the domain of infinite or transfinite sets.

There are points of agreement: The notion of a '"denumerably infinite ordinal number" ... has a clear and well-defined meaning for both formalist and intuitionist'. However, the theory of higher cardinal numbers has no meaning for the intuitionist and 'illustrates so clearly the impassable chasm which separates the two sides'. Following some details on this separation, he concludes with 'So far my exposition of the fundamental issue, which divides the mathematical world. There are eminent scholars on both sides and the chance of reaching an agreement within a finite period is practically excluded'.

Like Kronecker, Brouwer was led to constructivity and the rejection as meaningless of certain parts of mathematics. Whereas Kronecker's rejection had been externally political - Kronecker wrote to Sofya Kovalevskaya criticising Mittag-Leffler for publishing the 'blind assertions' of Cantor, he told Lindemann that the latter's proof of the transcendence of π was a wasted effort since the number didn't exist, etc. - we find Brouwer simply stating his views and acknowledging, with typically Dutch tolerance, that there are other views. Brouwer's writings would not become polemical until it was time to respond to Hilbert's polemics.

Brouwer formed his philosophical views in his youth and never changed them. They were pessimistic to say the least and included a strong disapproval of 'extrovert science' (read: technology), the goal of which was to conquer nature. His dissertation, begun at the age of 23 in 1904, was largely an attempt to reconcile his interest in mathematics with the rôle of mathematics as the chief tool of extrovert science. The excessive philosophical content - and its excessively pessimistic nature - alarmed his advisor D.J. Korteweg (now remembered primarily for the Korteweg-de Vries equation in partial differential equations) and a small battle over the contents of the thesis ensued.

When Brouwer finally published his thesis in 1907, the offending material was gone.

Brouwer's connection with Hilbert may be said to have begun with this thesis. Not only do some of the problems tackled come directly from Hilbert's problem list, but - more pertinent to our purposes - the thesis was on the foundations of mathematics and Brouwer offered critiques of Hilbert's work in the area, both the *Grundlagen der Geometrie* and Hilbert's 1904 paper on consistency proofs. I will not repeat these criticisms here as they are not central to our story. For one thing, Brouwer criticised aspects of Hilbert's 1904 paper that I have not covered because the paper led in a direction Hilbert was never again to follow. One point perhaps worth mentioning is that the style of the criticism is that of a self-confident young man: Brouwer states simply that, where, and how Hilbert is confused or outright wrong. For example, he says, 'Making a mathematical study of linguistic symbols ... can teach us nothing about mathematics', and, 'the most uncompromising conclusion of the methods we attack, which illustrates most lucidly their inadequacy, has been drawn by Hilbert'. The matter-of-fact style applied to matters of opinion was inappropriate, perhaps mildly offensive, but hardly hostile.

The following year (1908) Brouwer published two articles that must be cited here. The first of these, entitled 'Die moeglichen Maechtigkeiten' ('The possible powers'), cited our one mathematical ur-intuition and its application to the development of the denumerable infinite both in the form of the order type of the natural numbers (iterate 'next') and that of the order type of the rational numbers (iterate 'between'). He pointed out that such iteration allows only the construction of countable sets. As there occur higher powers in mathematics, there is a problem of reconciling this difference. According to Brouwer, there are two ways in which higher powers occur - in diagonalising and in constructing the continuum. In the first case, one really has a *method* and not a set. For example, constructing newer and larger countable ordinals from any sequence of countable ordinals does not show the uncountability of the *set* of countable ordinals, but exhibits a method of construction of larger ordinals and shows, incidentally, that the set of such ordinals does not exist. As for the continuum, i.e. the real number line, we can compare it to the collection of all paths through the infinite binary tree. This tree really contains only denumerably many nodes, not uncountably many paths. While individual paths may exist, their entire collection does not - only the countable tree exists. The 'collection' of paths ought to be thought of, not as a set which is an object, but as a matrix in which to place objects. The continuum is similarly a matrix and not a set.

Coming in 1908, Brouwer's interpretation was not so revolutionary as it might now seem. Poincaré similarly rejected the uncountable, Zermelo's axioms for set theory were just appearing that year, and Russell and Whitehead's *Principia Mathematica,* in which logic was a theory of types rather than sets, was yet to appear. Sets had not yet disengaged themselves from properties and the paradoxes of set theory had not yet been resolved. In short, the situation was completely unsettled and many solutions were being proposed. Brouwer's solution to the problem of higher powers was as plausible as anyone else's. It does illustrate, however, the direction his research was taking and the extent to which his philosophy would lead him: He was willing to jettison the majority of Cantor's theory of cardinal numbers.

The other paper of 1908 that I wish to mention is more relevant to our discussion and illustrates more dramatically the depth of Brouwer's thought and of his commitment to his philosophy. Entitled 'De onbetrouwbaarheid der logische principes' ('The unreliability of the logical principles'), it is a paper worth reading. If in the former paper Brouwer was willing to jettison uncountable sets, he was now questioning the laws of logic. Galileo had already shown that not all properties of finite cardinals carry over to the infinite case. Brouwer now suggested that among the lost properties one might find some logical laws - in particular, the Law of the Excluded Middle, or *principium tertii exclusi* (or, *tertium non datur*):

$$\varphi \vee \neg \varphi.$$

According to Brouwer, to assert $\varphi \vee \neg \varphi$, i.e. that φ is either true or false, one must either have a construction accomplishing the task demanded by φ or a construction that stops the process of performing this task. It was by no means clear to Brouwer that this could always be done:

> ... the question of the validity of the principium tertii exclusi is equivalent to the question *whether unsolvable mathematical problems can exist.* There is not a shred of a proof for the conviction, which has sometimes been put forward, that there exist no unsolvable mathematical problems.

The point here is that Brouwer placed mathematics in the intellect; to be true or false, a mathematical assertion had to be *known* to be true or known to be false. With infinite sets (as well as infinite methods and infinite matrices) we cannot necessarily perform the infinitely many verificational tasks necessary in finitely many steps; the principle of mathematical induction would occasionally allow this for denumerable collections, but would be of no use elsewhere. Brouwer never was to assert that the Law of the Excluded Middle was false in the sense that,

$$\neg(\varphi \vee \neg \varphi)$$

was true, because this was contradictory. In other words, he accepted $\neg\neg(\varphi \vee \neg \varphi)$ - whence the consistency of the Law of the Excluded Middle. But he distinguished between 'provable truths' and 'provable non-contradictories'. Eventually, as he refined his intuitionistic development, he would assert things like,

$$\neg \forall F(\exists x(Fx=0) \vee \neg \exists x(Fx=0)),$$

which appears absurd at first glance. On a constructive-intuitionistic reading, however, it asserts the unassailable proposition that it is impossible to construct a uniform algorithm which, given a function F, will tell us whether or not F has a zero.

The questioning of the validity of the Law of the Excluded Middle and the introduction of the dichotomy between truth and consistency stand in clear opposition to Hilbert's views, but - despite all my advance remarks - neither they nor the remarks of Brouwer's dissertation ought to be taken as the opening shots in the battle to come. Hilbert was the leading mathematician of the day, and he had spoken on the foundations of mathematics. Ergo, Hilbert was the yardstick by which all serious foundational remarks would be measured. Anyone putting forward original views on the subject would have to discuss them in light of Hilbert's views. [Besides, Brouwer published his dissertation and his 1908 paper in Dutch; for effect he would have to have published in French or German.]

The next few years saw two major developments in our story - Brouwer's growing friendship with Hilbert and the spread of Brouwer's fame. Let us begin with the second of these. Under some coaxing from Korteweg, Brouwer shifted much of his energy from foundations to mathematics itself. He did this with such success that, by the time of his inauguration as extraordinarius in 1912, he was one of the leading mathematicians of the day. This rise is reflected in the recognition accorded him:

1907 - Ph. D.
1909 - Privaatdocent (unpaid lecturer) at the University of Amsterdam
1912 - Extraordinary Professor; member of the Royal Academy
1913 - Ordinary Professor
1915 - Associate Editor of *Mathematische Annalen*.

Two episodes of this development will give us a hint to Brouwer's personality. The first of these also introduces a minor character who was again to be a minor character in the later fight - Otto Blumenthal, the managing editor of the *Annalen*. At the time *Mathematische Annalen* was one of the leading mathematical journals and Brouwer published several of his important early papers on topology in it. One of these was his proof of the invariance of dimension. [The unit interval can be mapped one-one onto the unit square, and the former can be mapped continuously onto the latter. Brouwer now showed that, replacing the dimensions 1 and 2 by arbitrary unequal m and n, the mapping cannot be one-one, onto, and bi-continuous.] When Brouwer submitted this paper to the *Annalen*, Blumenthal was visiting Paris and the subject came up in a conversation with Henri Lebesgue, who remarked to Blumenthal that for some time he had had a number of proofs of the result. Blumenthal did not understand Brouwer's proof. Nor did he understand the proof Lebesgue sketched for him, but such was his faith in Lebesgue that, when Brouwer's paper appeared it was immediately followed by Lebesgue's simpler proof in the form of an extract from a letter from Lebesgue to Blumenthal.

Lebesgue's proof was simply wrong. Blumenthal had treated Brouwer unfairly, and Brouwer's priority on this obviously major result was needlessly in question - and all because neither Lebesgue nor Blumenthal had a sufficient grasp of the complexity of the problem. Brouwer's reaction betrayed no appreciation of the irony of the situation. He immediately submitted another paper to the *Annalen*, the tone of which Blumenthal found 'unfriendly and unpleasant'. The paper went unpublished. However, Brouwer did manage to get a few digs against Lebesgue into print that year (1911): In one paper he remarked in a footnote that Lebesgue's proof was incorrect, and in a second paper he omitted the proof of an important lemma noting that Lebesgue was going to publish such - knowing full well the result lay beyond Lebesgue's reach. This latter was not mere impishness, but a tactic aimed at forcing Lebesgue's public acknowledgement of mathematical impotence. The priority battle continued - as late as 1924 Brouwer was attacking Lebesgue in print. [The whole fight was pointless: Priority was predictably assigned along national lines, the French lining up behind Lebesgue and the Germans behind Brouwer. *Now* that both parties are dead, everyone agrees that Brouwer had priority and that Lebesgue's proof was incorrect - but that it had the germ of an important idea.]

The obvious causes of an excessive and inappropriate response like Brouwer's - youth, financial insecurity, ambition - were all there. But they do not fully explain his actions; for, he was perennially embroiled in battles of this kind. The facts are that Brouwer was a highly emotional man who (i) had a rigid conception of right and wrong, and (ii) never completely matured emotionally. In theory Brouwer hated all people; in practice he enjoyed the company of others and formed a number of friendships. However, when one of his friends or acquaintances would behave in a manner consistent with his general theory of human behaviour, Brouwer seems to have been taken completely by surprise and reacted emotionally and strongly.

A second anecdote of this period illustrates Brouwer's immaturity and childish egocentrism. In 1913, Brouwer vacillated between keeping his extraordinary professorship in Amsterdam and answering the call to the more desirable ordinary professorship - but in the more provincial Groningen. Brouwer stuck with Amsterdam. Subsequently, in one of the rare altruistic gestures of mathematics, Korteweg vacated his chair and took the lesser position of extraordinarius so that Brouwer could be elevated to ordinarius in Amsterdam. [There is an apocryphal story that Isaac Barrow did the same for Newton, but his motives were different.] Brouwer soon thoughtlessly complained to Korteweg of all the work involved!

As I remarked earlier, Brouwer's growing acquaintance with Hilbert was another major development of the years 1909-1919 preceding the battle. They met in the summer of 1909. Writing to a friend in November, Brouwer described the experience thus:

> This summer the first *mathematicus* of the world was in Scheveningen. Through my work I was already in touch with him, [but] now I have over and over again walked with him, and talked as a young apostle with a prophet. He was 46 [actually: 47], but young in spirit and body, swam powerfully and climbed with the greatest pleasure over walls and fences of barbed wire. It was a beautiful new shaft of light through my life.

What did the young apostle and his prophet talk about? The mathematics in the ensuing correspondence - at least that which has been published - is topology. Later, Brouwer would maintain - in a footnote - that they discussed the foundations of mathematics. The only other comment on these discussions I know of is in a letter from Brouwer to Hilbert of 28 October 1909: 'If I come into the dunes, I think always on our beautiful excursions, and for me so instructive and stimulating conversations'.

It is evident from the above remarks that Brouwer had a very high opinion of Hilbert. Hilbert may have been something of a father-figure to Brouwer: When, in 1913, Brouwer was a young extraordinarius trying to decide whether or not to let Groningen spirit him out of Amsterdam with its offer of an ordinarius, he wrote to Hilbert for advice.

Hilbert could not possibly have had as high an opinion of Brouwer as the latter had of him. He was, after all, the 'first *mathematicus* of the world' and was intelligent enough to realise it. But he could and did appreciate talent. As I've already mentioned, in 1912 Hilbert was one of those recommending Brouwer's promotion to the position of extraordinary professor. Moreover, between 1909 and 1919 Brouwer visited Göttingen frequently and made numerous acquaintances there; it was, so to speak, his second scientific home. And, in 1919, on the eve of their battle, Hilbert tried to make Göttingen Brouwer's first home through the offer of a professorship - not the kind of offer lightly or easily made.

Whether their relationship was purely professional or a genuine friendship is hard to say. Their first - and evidently prolonged - meeting was at the seaside resort of Scheveningen in the Netherlands and it was a meeting of families, not just men. Brouwer's letters of the period seem friendly enough - the one cited above continues to say how nice the weather by the sea had been, to express concern about Hilbert's news that Mrs. Hilbert's knee had not recovered, and to introduce some comments Brouwer's wife was to add. This certainly suggests friendship, but in those pre-telephone days letter writing was more of an art than it is today and I am hesitant to read too much into such remarks. I note, for example, that scarcely a century earlier Thomas Carlyle and Johann Wolfgang von Goethe had a similar correspondence - right down to the women's appending their own comments to the letters - and yet the Carlyles and Goethes never met. Indeed, while Carlyle was expressing to Goethe his undying friendship and deep devotion, he was also writing a friend for an opinion of Goethe's mental state. On the other hand, one of Brouwer's psychological make-up is more prone to exaggeration than dissimulation, and I am inclined to believe Brouwer's feelings, however overexpressed, to be genuine enough. Hilbert's feelings remain a mystery to me.

I have as yet said nothing about Hilbert's personality. This remains something of a mystery to me. Constance Reid's biography gives one the impression of a generally nice man, as does Blumenthal's hagiography published in Hilbert's collected works. Blumenthal had been Hilbert's first doctoral student, receiving his degree in 1898, and he paints a picture at odds with the image of the later Hilbert we will encounter. Our Hilbert was in his 60s. It was the decade in which he killed the career of his student Wilhelm Ackermann simply because the latter married too early for Hilbert's taste. Indeed, it has been reported that students of this later generation today sit up straight at the mention of his name. The image of the dictatorial German professor may thus be the one we should keep before us. And the poor old Hilbert haunted by the spectre of Kronecker - as I suggested some pages back? There may be some of that too. Barring genuine historical research, I can only suggest these possibilities and caution the reader that Hilbert's assault on Brouwer will not reveal him in the best possible light.

But enough about personalities! Let us get back to our story. From the time of his thesis until that of his inauguration in 1912, Brouwer devoted most of his energy to topology. With the security of his job, Brouwer returned to his foundational interests, but not completely: He continued his topological studies into the mid-1920s and even afterwards kept in touch through his young assistants and international visitors. However, he did return to his foundational work and in 1918 and 1919 he published, in two parts, an important paper - 'Begründung der Mengenlehre unabhängig vom logischen Satz vom ausgeschlossenen Dritten' ('Founding of set theory independently of the logical principle of the excluded middle'). As Brouwer was to say in an accompanying paper - a sort of commentary on the present one - his previous writings on intuitionistic mathematics had merely been fragmentary and his non-philosophical, mathematical work had, although he had strived for results which could be established intuitionistically, been proven in the classical manner. In the present paper, he actually gave an intuitionistic development that went some distance. This success was to be his undoing.

Hilbert during this period was primarily occupied with mathematical physics, but in 1917 he did three things that concern us here. In the spring he hired Paul Bernays as his assistant, in the fall he gave a lecture in Zürich entitled 'Axiomatisches Denken' ('Axiomatic thinking'), and in the winter semester (1917/1918) he gave a course on 'Principles of mathematics and logic' in Göttingen. Bernays had impressed Hilbert with his knowledge of philosophy and would become the latter's primary logical co-worker in the years to come. Bernays was an extremely modest man and kept mostly in the background; he will be underrepresented in the sequel. Of the three occurrences of the year the one I wish to discuss is the Zürich lecture. This lecture, presented on 11 September 1917 and published the following year in *Mathematische Annalen,* is a transitional one lying somewhere between his attempt of 1904 and his future programme to be undertaken with the modest Bernays. Mathematically, it is the least of Hilbert's papers on foundations: It attempts nothing and announces no attempts. It is primarily a paean to the axiomatic method, and sings of the victories the method had had in mathematics and physics. Towards the end, however, there are some points worth mentioning.

With respect to *in*consistency, Hilbert cited Cantor's paradox of the set of all sets (which couldn't possibly be smaller than its power set) and noted that, because of it, respected mathematicians like Kronecker and Poincaré denied the legitimacy of set theory - and this despite the fact that set theory was 'one of the most fruitful and most powerful knowledge-branches of all of mathematics'. Fortunately, said Hilbert, Zermelo had saved the day by axiomatising set theory. [Tiny commentary: (i) Hilbert did not cite Russell's paradox - the one usually cited - because he had long known Cantor's earlier one; (ii) Kronecker objected to set theory long before the paradoxes were known; (iii) Hilbert did not mention Brouwer; and (iv) Zermelo's axioms seem not to have been motivated by the paradoxes, but by the desire to clarify his proof of the Well-Ordering Theorem - it would be the later explanation of the cumulative hierarchy of sets as an intuitive model of Zermelo's axioms that would rescue set theory.]

As for proving consistency, Hilbert cited a few examples of consistency proofs by interpretation - e.g. his proof of the consistency of geometry by interpretation of his axioms in a theory of the arithmetic of the reals and the consistency of the latter theory by interpretation in the (second-order) theory of the arithmetic of the integers. For the arithmetic of the integers and set theory, however, the only possibility left along these lines was to reduce these subjects to logic. Such a reduction, begun by Frege, had been - Hilbert said - accomplished by Russell, whose axiomatisation of logic could be viewed as the high point [Krönung] of work on axiomatics.

Four years later Hilbert would definitely be of the opinion that the work of Russell (and Whitehead) was insufficient and one needed a better consistency proof. At this time, however, he had a few other things to say. Related to the consistency problem were 5 other epistemological questions of a 'specifically mathematical colour':

(i) the problem of the solvability in principle of each mathematical question;
(ii) the problem of the additional [nachträglichen] controllability of the results of a mathematical investigation;
(iii) the question of a criterion of simplicity of mathematical proofs;
(iv) the question of the relation between content and formalism in mathematics and logic;

and

(v) the problem of the decidability of a mathematical question through a finite number of operations.

Of these, the only one he elaborated on was the fifth, which he remarked was the best known and most frequently discussed. The first example he cited was his own solution to Gordan's invariant problem. His first solution solved the problem, but didn't allow the effective calculation of the invariants; for this one needed Hilbert's second solution.

Finally, just before a short summary of his views, Hilbert noted that problems like this fifth one of the decidability via finitely many operations of mathematical questions seemed to require a new field of research, one in which proofs themselves were the objects of investigation. He added that the execution of this programme was yet an unaccomplished task. He did not add that he would soon be working on this task, but, given his hiring of Bernays, it can be assumed he was inclining toward, if not yet definitely preparing for, such a task. [Indeed, as I have already mentioned, Hilbert gave a course on 'Principles of mathematics and logic' in Göttingen during the ensuing winter semester.]

I should say a few words about Hilbert's questions (i)-(iv), particularly the ones the meanings of which are not obvious, i.e. (i), (ii), and (iv). Question (i) will be discussed shortly with Brouwer's reaction to Hilbert's paper, and question (iv) will be clarified in discussing Hilbert's further papers. This leaves us with question (ii) and the meaning of the 'controllability of the results of a mathematical investigation'. Hilbert was to repeat this phrase and even give something of an example of this controllability. Basically, he was referring to the practice of accountants of controlling, or partially checking, their calculations. Indeed, according to Samuel Johnson's dictionary (admittedly not the most up-to-date source, but I am so taken with this explanation that I have not subjected it to the control of a more modern etymologist), the word 'control' (in German: Kontrolle) derives from the word '*counter*ol' indicating just this practice of ac*count*ants (sitting at their *count*ing boards with their copper *count*ers) of checking their results. With pencil and paper computation, one can cite the technique of casting out nines as a control for large sums. Hilbert appears to be calling for something similar for all mathematical investigations - not just arithmetic calculations.

I remarked earlier that Brouwer published a separate commentary on his 1918/1919 paper on intuitionistic set theory. In it he began by remarking that since 1907 he had in several publications defended two theses: (i) that the chief set-existence axiom of set theory (the Comprehension Axiom) was - even in Zermelo's weaker form - inadmissible; and (ii) that Hilbert's Axiom of the Solvability of Every Problem was equivalent to the Law of the Excluded Middle, and was impermissible as a means of proof. In a footnote, he commented on Hilbert's 1917 paper: This proposition, an axiom to Hilbert in 1900, was now - as question (i) - considered by Hilbert to be an open problem. To Brouwer, who saw the question as an open problem in 1908, the principle was now simply false (albeit not in the strong sense of being contradictory). Moreover, to Brouwer Hilbert's questions (i) and (v) were the same - solvability (in principle or in practice) meant solvability by a finite number of operations. In particular, he did not accept Hilbert's assessment of what Hilbert's first solution to the invariant problem had and had not accomplished. Brouwer said simply that proving a set (here: the set of invariants) not to be infinite was not the same as proving the set to be finite. In effect, Brouwer said that Hilbert's first proof did not establish the existence of a finite number of invariants - just as Gordan, Kronecker, et al. had said years earlier.

Although it had been written in German, Brouwer's 1918/1919 paper appeared in a Dutch journal. The commentary was published in both a Dutch and a German journal, the latter being the *Jahresbericht der Deutschen Mathematiker-Vereinigung* ('Annual Report of the German Mathematicians-Union', hereafter simply *Jahresbericht*). This was the main organ of the German mathematical society and every German mathematician would see it.

In September of 1920 Brouwer lectured at a meeting of natural scientists in Bad Nauheim, in Germany, on the question of the intuitionistic existence of decimal expansions of all real numbers. He published the paper with its negative answer to this question the following year in *Mathematische Annalen.* The single direct reference Hilbert would make to one of Brouwer's papers would be to this one; everything else Hilbert would say about Brouwer would be about a Brouwer filtered through Hermann Weyl.

It is now time for Weyl to enter our story. Weyl, who had got his degree from Hilbert, was one of the leading young mathematicians of the day. Like Hilbert he worked in mathematical physics as well as in mathematics, and like Brouwer he was something of a philosopher as well. As a philosopher, he took the paradoxes more seriously than the average mathematician. On his view, paradoxes in the more remote regions of mathematics were symptomatic of a deeper problem in ordinary mathematics. His diagnosis of the problem agreed with Russell's and his solution was in line with Poincaré's: The paradoxes arose from circular reasoning, particularly in using *impredicative* definitions - definitions in which the object being defined is defined in terms of some collection in which the object is included. This naturally leads to *vicious circles.* For example, the definition of the set R of Russell's paradox,

$$R = \{x: x \notin x\},$$

is impredicative in that R itself is an element of the possible range of its values. That this is bad manifests itself when one asks if $R \in R$. To determine the truth of this statement, one must check if $R \notin R$, which of course presupposes one to have checked if $R \in R$ - thus taking one full circle. Like Poincaré, Weyl proposed to base analysis on the natural numbers and predicatively defined sets thereof. His notion of predicativity was, however, extremely narrow and his reconstruction of analysis did not go as far as he would have liked.

In 1918 Weyl published a short monograph, *Das Kontinuum,* containing his attempted reconstruction of analysis. In the following year he published a note entitled 'Der Circulus vitiosus in der heutigen Begründung der Analysis' ('The circulus vitiosus in the modern foundation of analysis'): Modern analysis was based on vicious circles and was in danger of collapsing under the weight of paradox. As late as 1930 he was to say, 'It is true that so far no actual contradictions in analysis proper have resulted; we do not completely understand this fact at present'. Of greatest immediate importance for us, however, were his lectures in 1920 in the mathematics colloquium in Zürich. Published the following year under the title 'Über die neue Grundlagenkrise der Mathematik' ('On the new crisis in the foundations of mathematics'), these lectures caused quite a stir and roused Hilbert to action.

Zürich, where Weyl was a professor (ordinarius) at the time, was not Weyl's only pulpit. On 11 May 1920 he had, for example, lectured on 'Das Kontinuum' in Hilbert's Göttingen, and on 28-30 July he gave a series of lectures on his new foundation of analysis in Hamburg. His Zürich lectures and the paper to come out of them had, however, a bigger effect. Everything about the paper was provocative. Weyl himself would later refer to the style as 'bombastic', reflecting the excitement of the postwar years. With less caution than Brouwer's 'I believe' one cannot apply the Law of the Excluded Middle, Weyl at one point says, 'But this standpoint is absurd and untenable'. Moreover, at one point - after dramatically announcing he was abandoning his own foundational programme and joining Brouwer - he sings the praises of Brouwer to an embarrassing excess: '...Brouwer - that is the revolution!' and 'Brouwer is the one to thank for the new solution to the problem of the continuum... Whether I have the right to describe Brouwer's theory... is to me, mind you, doubtful'.

It was not style alone that was provocative about Weyl's lectures. There was also the content: Dedekind's definition of the set of natural numbers is circular; there are circular definitions in analysis. In his analysis the convergence of Cauchy sequences still holds, as does the Intermediate Value Theorem, but the existence of suprema of bounded sets fails and one cannot hope to save - one of Hilbert's results - Dirichlet's Principle. In accepting Brouwer's continuum one had to give up points; Brouwer's continuum was not a collection of points, but a 'medium of free becoming' - the exact meaning of which (fortunately!) need not concern us here.

Weyl gave the first clear and accessible (though not entirely faithful) account of Brouwer's ideas. Brouwer's big 1918/1919 paper had been primarily straight mathematics, and his commentary on it was somewhat sparse. Weyl, by then an experienced expositor, offered a more expansive explanation. Moreover, he phrased things more vividly. Discussing the problem with classical logic, he described the use of quantifiers as follows:

> An existential assertion - e.g. 'there is an even number' - is not at all a judgment in the proper sense, which asserts a fact.

Rather, it was a 'judgment abstract' to be compared with a slip of paper offering a treasure, but not telling where the treasure was to be found. Continuing,

> The general 'Every number has property E' - e.g. 'for each number m, $m+1=1+m$' - is equally little a genuine judgment, rather [it is] a general instruction [Anweisung] on judgment.

Thus, Weyl more-or-less considers,

$\exists$: an I.O.U. ('I owe you')

$\forall$: a payment slip.

'In this light mathematics appears as a dreadful [ungeheuer] paper-economy'. For good measure he adds, 'It is not the existence theorem that is valuable, but rather the construction carried out in the proof. Mathematics is, as Brouwer occasionally said, more of a doing than a teaching' ['mehr ein Tun dann eine Lehre' - 'teaching' here in the sense of a doctrine].

Weyl's lectures were widely discussed. In Göttingen, Richard Courant joined Bernays on 1 and 8 February 1921 in reporting 'On the new arithmetic theories of Weyl and Brouwer' in the Göttingen mathematical colloquium. Also that year, the *Jahresbericht* reports lectures 'On the Brouwer-Weyl number concept' in Breslau (now: Wrocław) by L. Koschmieder and 'On Weyl's researches in the foundations of mathematics' in Frankfurt by E. Hellinger.

Hilbert had been thinking of returning to the problem of foundations since 1917 when he had hired Bernays as his assistant. Nonetheless, for the next three years he continued to work mainly in mathematical physics. Weyl's public defection changed that: In the winter of 1920/1921, Hilbert once again turned his attention to foundations. Indeed, it seems he first announced his views only two weeks after Bernays and Courant lectured on Weyl's lectures when, on 21 and 22 February 1921, he lectured on 'A new laying of foundations for the number concept' in the Göttingen mathematics colloquium. In the spring he lectured in Copenhagen and on 25-27 July, after an advance notice in the *Jahresbericht*, he lectured in Hamburg. These Hamburg lectures drew 'numerous listeners - including also famous mathematicians' according to Kurt Reidemeister's report in the *Jahresbericht*, the first, but inadequate, description of Hilbert's new programme to appear in print. In September (in particular, on 23 September) of that year Bernays lectured at a meeting of the DMV (the German Mathematicians-Union mentioned above) in Jena 'On Hilbert's thoughts on laying the foundations of arithmetic'.

The year 1922 was not so full of public action for Hilbert's programme. The texts of his and Bernays' lectures were published, and on 22 September Hilbert gave another lecture, the text of which was published in *Mathematische Annalen* the following year, at a meeting of the DMV in Leipzig. There were also some personal matters that, however, are best postponed until after discussing Hilbert's new programme and the papers on it.

Hilbert's programme was, of course, to prove the consistency of arithmetic (with arithmetic taken in a broad sense to include analysis and set theory). What was new were the ground rules and method, as well as perhaps a sense of urgency. What was still missing was a convincing reason for the emphasis on consistency. Obviously, an inconsistent theory has no informational content, but why should consistency be enough? Bernays, who was more philosophically inclined than Hilbert, went farther than Hilbert had previously gone in explaining this importance. By this I do not mean that Bernays directly addressed the issue, but that here and there in his paper one can find statements bearing on the problem:

> Whereas logic has to do with *contentual* universality [inhaltliche Allgemeinsten], (pure) mathematics is the general study of *formal* relations and properties;

and:

> What matters for the question of pure mathematics is only whether the usual, axiomatically characterised mathematical continuum is possible in itself, that is it is a consistent creation [Gebilde].

To Hilbert, whose fundamental views on this subject may well have come from geometry where the success of non-Euclidean geometry exemplifies the merely formal nature of mathematics and the sufficiency of consistency, this was second nature. Neither he nor Bernays realised at the time that, to Brouwer, consistency was a red herring: Brouwer was after *truth*, not mere consistency. He would hardly have been impressed by Bernays' further remark that

> And what we have gathered from the investigations of Weyl and Brouwer is the conclusion that a proof of consistency via a replacement of existence axioms by construction postulates is not possible.

For, Brouwer could easily have objected that his goal was not to establish the consistency of a false theory, but to learn the truth. It never seems to have occurred to Hilbert, and did not occur to Bernays at the time, that the very goal of the programme - once achieved - would not be convincing and was itself in need of explanation. Perhaps, had Hilbert not been so impatient with philosophers, he would have been led to this realisation through his earlier correspondence with Frege. Eventually Hilbert would - accidentally? - hit upon such an explanation and the myth of Hilbert's Programme (with a capital 'P') would be generated. But we will get to this later.

Hilbert's first paper, 'Neubegründung der Mathematik, erste Mitteilung' (usually referred to simply as Hilbert's First Hamburg Lecture), may be considered the first shot fired in the battle between Hilbert and Brouwer. Indeed, Hilbert fired a whole salvo of polemics - but polemics that appear to be aimed more at Weyl than at Brouwer: His criticisms refer directly to Weyl's Zürich lectures - not Brouwer's writings; and whenever he refers to Brouwer and Weyl, he mentions Weyl's name first. Was Weyl in the audience in Hamburg? A few quotes will illustrate the tone and the extent to which the battle between Hilbert and Brouwer, before Brouwer was even engaged in it, was already becoming one between Hilbert and the dead Kronecker:

> Weyl and Brouwer are searching for the solution to the problem - in my opinion - down the wrong path.

In fact:

> What Weyl and Brouwer are doing comes in principle to this, that they are wandering down the former paths of Kronecker: They seek to found mathematics in such a way, that they throw overboard everything that has an uncomfortable [unbequem] appearance to them and erect a *Verbotsdiktatur* à la Kronecker.

Against Weyl alone we read,

> The circulus vitiosus is artificially dragged into analysis by Weyl;

and

> if Weyl notices an 'inner groundlessness of the foundations on which rests the construction of our Empire' and worries about 'the menacing dissolution of the State of Analysis', then he's imagining things [literally: seeing ghosts].

Hilbert had his own political metaphors to offer, and offer them he did - in one of the most famous quotations from the entire battle:

> I believe that, as little as Kronecker in his day succeeded to banish the irrational numbers... so little will Weyl and Brouwer succeed today; no: Brouwer is not, as Weyl believes, the Revolution, but rather only the repetition of an attempted putsch with old means, which, in its day more energetically undertaken, miscarried completely and now especially with the State so well-prepared and strengthened by Frege, Dedekind and Cantor, is condemned at the outset to failure.

In describing one of Hilbert's later lectures, Weyl said that there were 'anger and determination in Hilbert's voice'. Can we hear that in the present paper? If so, there is also some humour in what we read: It cannot be with mere vicarious viciousness that we enjoy the 'Revolution/putsch' comparison. Moreover, Hilbert's good humour is further revealed when he gets down to introducing his theory. In stating that he is going to start with concrete symbols, he biblically intones, *'in the beginning* - so it goes here - *is the sign'.* In any event, any anger directed at Weyl was short-lived. Hilbert gave the lecture in 1921; in 1922 he tried to spirit Weyl out of Zürich to join him in Göttingen. Weyl declined, but in future years Hilbert would try again and eventually succeed in luring Weyl home geographically, if not philosophically.

Although Hilbert's indulgence in polemics is certainly a sufficient reason to mention his First Hamburg Lecture, it is not the main reason for doing so. The main point of interest in this lecture is, of course, the announcement of his new programme for defeating Weyl and Brouwer. As in 1904, the key to this programme was to be a consistency proof for arithmetic in its broadest sense (i.e. including set theory). There was, however, a major difference between his 1904 attempt and his new one. His approach in 1904 was to build up mathematics and logic a bit at a time, at each stage using what was currently available to prove the consistency of what was to be added next. His development at the time never reached a stage where the mathematics proven consistent was powerful enough to do anything, and he did not make clear what was to be allowed in the initial stages. His new attempt avoided these difficulties.

To begin with, Hilbert distinguished between two kinds of mathematics - actual [eigentlich] mathematics (or: mathematics proper) and *meta*mathematics. Actual mathematics is what mathematicians do - analysis, set theory, etc. Actual mathematics is abstract, infinitary, and has no empirical meaning; the sole check one has on its validity is its non-contradictoriness.

Metamathematics, on the other hand, is intuitive and contentual mathematics; it is the direct combinatorial study of signs and their combinations. As we have already seen, Hilbert felt that the need for rigour imposed the axiomatic method on actual mathematics. He now went one step farther and called for its formalisation: Actual mathematics was to be thought of as or replaced by a formal system with precisely defined axioms and rules of inference and the consistency of this formal system was to be proven metamathematically.

The distinction between the formal mathematics - actual mathematics - and the contentual metamathematics was made by Brouwer in his dissertation in 1907. Hilbert very probably picked it up from Brouwer during their discussion in the dunes of Scheveningen in 1909; indeed, Brouwer would maintain this in years to come. In any event, the separation suited Hilbert's purposes well. In the first place, metamathematical reasoning was acceptable to Weyl and Brouwer and they would have to accept the consistency proof - though, as I said earlier, they need not accept the main consequence of consistency Hilbert envisioned - the correctness, in some sense, of the consistent theory. Moreover, by distinguishing between metamathematics and mathematics proper, Hilbert hoped to sidestep Poincaré's earlier charge of circularity - that Hilbert had been using induction to justify induction. According to Hilbert, metamathematical induction was of a simpler character than the full principle of induction. Whether one agrees with this is largely a matter of perspective, of whether one considers the properties to which induction is applied as central (in which case one will agree with Hilbert) or one considers induction as a property of natural numbers rather than of the formulae inducted on (in which case one will still agree with Poincaré).

Hilbert named his programme *Beweistheorie* [proof theory] and, as in 1904, gave a sample metamathematical consistency proof of a formal system. Again as in 1904, the result was not very impressive and it need not concern us here. He would do better the following year in his Leipzig lecture.

The Leipzig lecture of September 1922 was published in 1923 in *Mathematische Annalen* under the title 'Die logischen Grundlagen der Mathematik' ('The logical foundations of mathematics') and shows some real progress, both expository and mathematical. The tone is much milder, with only an occasional touch of the polemical - as in the opening remark where he describes the goal of his new programme to be 'to banish definitively from the world the general doubt of the security of mathematical inference'. Moreover, he came very close to determining exactly what constituted his metamathematics. In doing so, he more-or-less reversed himself: In Hamburg he had said that the metamathematics, being intuitive and contentual, was not axiomatic; it had no axioms and could not be inconsistent. In Leipzig, he now offered a formal system of metamathematical arithmetic (i.e. the arithmetic portion of his metamathematics) and sketched a proof of the consistency of a fragment thereof. Neither the formalisation of nor the proof of the consistency of his metamathematics was necessary for the working out of his programme, but they illustrated most simply how he thought the programme could be carried out. Indeed, in the ensuing years his student Wilhelm Ackermann would almost succeed in carrying through Hilbert's programme by following the lines of Hilbert's sketch.

It will be instructive for us to consider Hilbert's formalisation of metamathematical arithmetic, his formalisation of 'transfinite' arithmetic, and his sketches of consistency proofs for fragments thereof. First, however, a few words of explanation. Hilbert introduced a new word into his proof-theoretic vocabulary - 'finit', a Latin form of 'finite' to stand alongside the usual German 'endlich'. English equivalents abound - 'finitist', 'finitistic', and 'finitary' are generally used. The use of the new word is conceptual as well as emphatic: One can imagine a finite set of infinite objects, e.g. $\{\mathbb{Z}, \mathbb{Q}, \mathbb{R}\}$; but one would not call such a thing finitistic. In commenting on his metamathematical arithmetic, Hilbert said:

> The derivable formulas, which are obtained on this standpoint, all have the character of the finitary, i.e. the thoughts, whose images they are, can also be obtained without any axioms contentually and immediately through consideration of finite universes.

In such finite universes, there is no questioning of the applicability of traditional Aristotelian logic; in particular, the use of the Law of the Excluded Middle is unexceptionable. However, Hilbert went on to say, when we thoughtlessly apply procedures that are reliable in the finite case to the infinite case, we make mistakes left and right. This happens in analysis when we ignore convergence criteria in dealing with infinite sums and products. In logic it happens when one starts using quantifiers $\exists$ and $\forall$, which merely serve to abbreviate infinite logical sums and products, respectively.

A universal arithmetic assertion $\forall v \varphi v$ was to Hilbert an infinite conjunction $\varphi(\overline{0}) \wedge \varphi(\overline{1}) \wedge \ldots$ Its negation had no precise content or meaning. Similarly, an existential assertion $\exists v \varphi v$ was an infinite disjunction $\varphi(\overline{0}) \vee \varphi(\overline{1}) \vee \ldots$, incapable of being negated. Exactly what Hilbert meant here is unclear, and the clarity would not improve in future presentations of his programme. In 1925, he would maintain that universal assertions $\forall v \varphi v$, with φ quantifier-free, were finitistic because they were mere infinite conjunctions, but that existential assertions were not finitistic because they were only partial judgments, omitting crucial information. In accordance with the necessity of considering finite universes, one should perhaps be more inclined to accept existential assertions as finitistic and condemn the universal ones. Hilbert may have preferred the universal ones because, as we shall shortly see, they were subject to some control. More probably, he questioned existential assertions because he had to establish their validity (in some sense) because it was his existential theorems that had drawn fire from the critics - a fact he cited in this paper.

Ignoring the question of the exact borderline between finitistic and infinitistic assertions, one can say that, according to Hilbert, quantifiers introduce an infinitistic element into logic. Quantificationally complex formulae are meaningless and Aristotelian logic - in particular, the Law of the Excluded Middle - does not apply to them. Nonetheless, Hilbert would add such formulae to his formal system of finitary arithmetic and apply the Law of the

Excluded Middle to them. He compared this to practice in analysis and geometry:

> In my proof theory [we] adjoin to the finite axioms the transfinite axiom and formulae, just as one [introduces] imaginary elements to the reals in the theory of complex numbers and ideal objects in geometry. And the motive for doing this and the success of the procedure is in my proof theory the same as there: namely, the addition of the transfinite axiom achieves in a sense the simplification and rounding off of the theory.

['Transfinite axiom' will be explained shortly.] The use of ideal elements in algebra and geometry can, in principle, be dispensed with. Any theorem, for example, about the real numbers proven via a detour into the complex realm can be proven anew, usually in a more cumbersome manner, without appeal to the complex numbers. Does the analogy go this far? According to Hilbert,

> To be sure [freilich] one can presumably prove a finitistic statement also without application of transfinite means of proof... but this claim is of the sort of the claim that in general every mathematical assertion must allow itself either to be verified correct or refuted.

He illustrated this with the example of his theorem on invariants: After his first transfinite proof, he had given his second finitistic proof. Incidentally, in citing this example, Hilbert went on to recall how Gordan had distrusted the first proof and labelled it 'theological'; it would appear that the old bitterness was resurfacing.

If Hilbert was not yet ready to tackle the problem of the solvability of every mathematical problem or the conservation of transfinite methods over finitary ones, he was ready to commit himself as to the consistency of the transfinite. Indeed, and as I have already said, he sketched the outline of his approach. It is to this that we now turn.

In presenting Theorem 1, I took the liberty of simplifying the system Hilbert proved consistent as the point there was merely to give an example of a consistency proof. Because the present system offers an 'exact' description of finitistic arithmetic, I shall present it in its original form. By way of preliminary explanation, I note that because of the equivalences,

$$\varphi\lor\psi. \Leftrightarrow .\neg\varphi\rightarrow\psi$$

$$\varphi\land\psi. \Leftrightarrow .\neg(\varphi\rightarrow\neg\psi),$$

Hilbert's logic only included the two connectives, $\neg$, $\rightarrow$. Moreover, Hilbert used variables for formulae as well as for numbers. Variable formulae were used in this paper merely as place-holders allowing the axioms to be stated as single sentences rather than as schemata; such variables do not appear in the theorems of the system and can be dispensed with. In line with Hilbert's general ambivalence concerning the meaning of universal assertions, one might ask whether the same was to hold for the number variables in the finitistic system:

Was the axiom, $a=a$, with variable a, to be taken as an abbreviation for the schema $t=t$, for closed terms t, or was it implicitly to be taken as the universal assertion $\forall a(a=a)$? Given that consistency is itself a universal assertion ('for any proof p, the end formula of p is not a contradiction'), and that it was to be proven finitistically, we should probably assume even the finitist system to be aimed at proving general assertions with free variables. Such an assumption does, however, require care in interpreting Hilbert's remark on establishing provable formulae by considering finite universes. We will return to this question later.

The non-logical primitives of Hilbert's initial system are the constant 0 (with no overbar), the successor function $(\cdot)+1$, and the predecessor function δ. Initially, the only rule of inference is *modus ponens*: from φ and $\varphi \rightarrow \psi$ infer ψ. There are ten axioms broken into four groups: axioms of consequence (1-4), axioms of negation (5-6; axiom 6 is his formulation of the Law of the Excluded Middle), axioms of equality (7-8), and axioms of number (9-10). These axioms follow:

A1. $A \rightarrow (B \rightarrow A)$
A2. $(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$
A3. $(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$
A4. $(B \rightarrow C) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$
A5. $A \rightarrow (\neg A \rightarrow B)$
A6. $(A \rightarrow B) \rightarrow ((\neg A \rightarrow B) \rightarrow B)$
A7. $a=a$
A8. $a=b \rightarrow (A(a) \rightarrow A(b))$
A9. $\neg(a+1=0)$
A10. $\delta(a+1)=a$.

Hilbert overlooked the axiom,

$$\neg a=0 \rightarrow \delta(a)+1=a,$$

as well as some determination of a value for $\delta(0)$, e.g. $\delta(0)=0$. Even with such additional axioms, one has at best a fragment of finitist arithmetic. To obtain his full metamathematical arithmetic, Hilbert said one should add 'recursion and intuitive induction'. With no quantifiers yet allowed, induction would have to be added as a rule of inference:

$$\text{from } A(0) \text{ and } A(a) \rightarrow A(a+1) \text{ infer } A(b).$$

The question of which recursions were to be allowed is premature: Hilbert was familiar with Dedekind's recursive definitions of addition and multiplication, as well as Dedekind's general theorem allowing definition of functions by recursion; but in 1922 the class of functions generated by recursions had not yet been subjected to any general study. By coincidence, Thoralf Skolem was already developing a quantifier-free number theory based on recursion, but his paper would only first be published in 1923 - the year in which Hilbert's Leipzig lecture would be published. The next few years would see the beginning of a general study - by the Hilbert school - of recursions, and in 1925

Hilbert would report that they had examples of recursively generated non-primitive recursive functions. In short, at the time of his Leipzig lecture Hilbert's conception of recursion, and hence of finitism, was only partially formed and I can be no more specific in clarifying the situation than to say that finitist arithmetic is obtained from the initial fragment already described by adding 'recursion and intuitive induction'.

[If the historical question of what Hilbert meant by finitism is replaced by the philosophical question of what he could or should have meant, an answer is at hand: Allow only primitive recursions. This is what Ackermann did in 1924, and there is today a consensus of opinion that finitistic arithmetic is adequately captured by a quantifier-free theory of primitive recursive functions. Through the encoding introduced by Gödel, such a theory is taken to embody finitism itself.]

Before discussing the addition of quantifiers through the adjunction of a 'transfinite axiom', I would like to give Hilbert's proof-sketch of the consistency of the theory, say $\mathfrak{T}_0$, given by axioms 1-10 above and the rule modus ponens. To do this, I must first carefully define what a formal derivation is. This is easy: A formal derivation of a formula φ is a sequence of formulae $\varphi_0, \varphi_1, \ldots, \varphi_n = \varphi$ (say, containing no variable formulae) such that each φ_i is an instance of an axiom, e.g.

$$\varphi \rightarrow (\psi \rightarrow \varphi) \quad \text{or} \quad t = t,$$

or follows from two earlier formulae φ_j and φ_k (of the form $\varphi_j \rightarrow \varphi_i$) of the list (i.e. $j,k < i$) by modus ponens. With some effort it can be shown that axioms 1-6 are complete with respect to purely propositional reasoning, whence, in particular, the consistency of $\mathfrak{T}_0$ reduces to the unprovability of $\neg 0 = 0$.

THEOREM 2. *The system $\mathfrak{T}_0$ is consistent.*

PROOF SKETCH. Suppose $D = \varphi_0, \varphi_1, \ldots, \varphi_n$ is a formal derivation of $\neg 0 = 0$.

As a first step, we can modify D by throwing away any φ_i (other than φ_n) which is not used as a premise of an application of modus ponens. Moreover, by repetition of formulae, we can assume each φ_i occurs only once as such a premise. Call the resulting derivation D'.

Second, we can omit number variables from D' by substituting, say, 0 for each occurrence of a variable in D'. Call the result D''.

Third, we can simplify the terms to the point that each formula is a propositional combination of equations of *numerals*,

$$0,\ 0+1,\ 0+1+1, \ldots$$

Fourth, every formula can be brought into a logical normal form (of some sort - Hilbert doesn't say which kind).

Each formula of the derivation is now subject to a control, i.e. one can check each formula for 'correctness' or 'falsity'. But it can be shown that each formula of this final derivation is correct, whence the end formula $\neg 0 = 0$ is correct - a contradiction. Thus, there was no derivation D of $\neg 0 = 0$.

QED

This proof is as Hilbert gave it and the reader should not fault me if he finds it less than clear. I myself do not find it completely clear. The first step has the character of something one finds one has to do to carry out the real work of the proof-transformations and we can probably ignore it. The fourth step is an enigma to me - I simply don't know what 'normal form' Hilbert has in mind. The second and third steps, however, seem crucial. The second step guarantees that all the axioms of the derivation hold in some finite universe, thereby showing in principle the finitary nature of the derivation (in the sense of our earlier quote from Hilbert). More to the point, it turns every term into a computable term capable of reducing to a numeral. [This last is only true if one adds the axioms $\delta(0)=0$ and $\neg a=0 \rightarrow \delta(a)+1=a$.]

Hilbert's subsequent partial extension of Theorem 2 to the 'transfinite case' suggests the following interpretation of step 3: Calculate the value of each term t occurring in the given derivation D'' and replace t in each occurrence by the numeral for this value. Every formula is now a propositional combination of equations of numerals. The 'correctness' or falsity of each equation is simply a matter of comparison, and the calculation of truth values of propositional combinations is a simple matter. One can now show by induction on the length of D'' that the formulae occurring in the latest transformation are all true. [If this is indeed what Hilbert had in mind, then the reference to putting formulae into a 'logical normal form' must refer to the simple calculation of the truth value of a propositional combination of sentences once the individual truth values are known.]

[Digression: Another plausible interpretation of step 3, which may shed a little light on Hilbert's idea of transforming derivations, is this: Transform the derivation D'' by repeatedly replacing terms $\delta(t+1)$ by t. With a little luck, the result is (or is easily made into) a derivation itself. For example, consider the following derivation of $0=\delta(0+1)$:

(1)	$\delta(0+1) = 0$	A10
(2)	$\delta(0+1) = 0 \rightarrow (\delta(0+1)=\delta(0+1) \rightarrow 0=\delta(0+1))$	A8
(3)	$\delta(0+1) = \delta(0+1) \rightarrow 0=\delta(0+1)$	by (1), (2)
(4)	$\delta(0+1) = \delta(0+1)$	A7
(5)	$0 = \delta(0+1)$	by (3), (4).

The replacement of $\delta(0+1)$ by 0 results in the correct derivation:

(1)	$0 = 0$	A7
(2)	$0 = 0 \rightarrow (0=0 \rightarrow 0=0)$	A8 (or A1)
(3)	$0 = 0 \rightarrow 0=0$	by (1), (2)
(4)	$0 = 0$	A7
(5)	$0 = 0$	by (3), (4).

The point to notice is that the instance of A10 has been replaced by an instance of A7. Similarly replacing $\delta(t)+1$ by t for $t \neq 0$ will transform an instance of the auxiliary axiom,

$$\neg a = 0 \rightarrow \delta(a)+1=a,$$

into the formula,

$$\neg a = 0 \rightarrow a = a,$$

which, although not an axiom, is readily given a quick logical derivation:

(1)	$a = a$	A7
(2)	$a = a \rightarrow (\neg a = 0 \rightarrow a = a)$	A1 or A5
(3)	$\neg a = 0 \rightarrow a = a$	by (1), (2).

Replacing $\delta(0)$ by 0 changes the axiom $\delta(0)=0$ into $0=0$. It is tempting to jump to the conclusion that such replacements will eliminate the arithmetic axioms other than $\neg(t+1=0)$ from the derivation and yield a (nearly) purely logical derivation. I haven't checked this, but I do note that it is not as trivial a matter as the above examples indicate.]

Accepting the pre-digressive consistency proof for $\mathfrak{T}_0$ as the one Hilbert had in mind, let us see how Hilbert intended to extend the proof to obtain the consistency of *transfinite arithmetic*, i.e. arithmetic with quantifiers.

To begin with, Hilbert decided to go a step beyond quantifiers by dealing with *choice functions*: Given a formula $\varphi(a)$ with free variable a, Hilbert added a new term $\tau_a(\varphi)$ attempting to choose a *counter*example to φ. Together, τ and φ satisfied the transfinite axiom :

$$\varphi(\tau_a(\varphi)) \rightarrow \varphi(a).$$

[Formally, in terms of variable formulae, the axiom was

A11. $A(\tau(A)) \rightarrow A(a)$.

Note that, since A is a variable for formulae, it has no numerical variable a occurring in it, whence $\tau(A)$ has no subscript indicating which variable is being *bound*. Although there are no quantifiers the new abstraction operators τ_a bind variables and the familiar problems with substitution arise; however, Hilbert was not yet aware of this.]

Using τ, quantifiers can be introduced as abbreviations:

$$\forall a\varphi(a)\colon\ \varphi(\tau_a(\varphi))$$

$$\exists a\varphi(a)\colon\ \varphi(\tau_a(\neg\varphi)).$$

With these, the usual laws for quantifiers can be derived:

$$\forall a\varphi(a) \rightarrow \varphi(a), \qquad \varphi(a) \rightarrow \exists a\varphi(a)$$

$$\neg\forall a\varphi(a) \Leftrightarrow \exists a\neg\varphi(a), \qquad \neg\exists a\varphi(a) \Leftrightarrow \forall a\neg\varphi(a).$$

Hilbert did not claim that he had already extended his consistency proof to include the transfinite axiom. There were, as he noted, difficulties with the nestings of τ's. What he did do was to indicate how the proof of Theorem 2 could be extended to cover one special instance of the new axiom.

Let f be a function variable and define,

$$\tau(f) = \tau_a(f(a)=0).$$

The corresponding instance of axiom A11 reads

A12. $f(\tau(f))=0 \rightarrow f(a)=0$.

As Hilbert quickly pointed out, the function $\tau(f)$ would not be allowed by Brouwer and Weyl (Brouwer's name now came first). Following the sketch of his consistency proof for A12, Hilbert gave a couple of examples of what this consistency proof would yield that Brouwer and Weyl forbade. First of all, we could talk about decimal expansions of numbers for which we cannot compute the expansions - contrary to the conclusion drawn by Brouwer in his 1920 paper cited a few pages back. Indeed, using τ one could define the function,

$$F(n) = \begin{cases} 0\ , & n^{\sqrt{n}} \text{ is rational} \\ 1\ , & n^{\sqrt{n}} \text{ is irrational,} \end{cases}$$

and then consider the number with dyadic expansion $.F(2)F(3)\cdots$, i.e. the number

$$r = \sum_{n=0}^{\infty} \frac{F(n+2)}{2^{n+1}}.$$

Moreover, one could do this despite the fact that one could not (in 1922) even calculate the first bit of r - i.e. one did not even know $F(2)$.

Second, Hilbert showed in some detail - this time in response to Weyl - how using the τ's made it possible to prove the existence of the least upper bound of a bounded set of real numbers.

The question, then, is the proof of the consistency of A12 with A1-A10 (and the recursions). To simplify matters, we only allow one fixed function F (obtained by recursions) in A12. Thus, A12 becomes

A12′. $F(\tau(F)) = 0 \rightarrow F(a)=0$.

Moreover, F is assumed unary. Let the resulting system with axioms A1-A10, A12′, and whatever recursion equations are necessary to compute F be called $\mathfrak{T}_1$.

THEOREM 3. *The system $\mathfrak{T}_1$ is consistent.*

PROOF SKETCH. Using the first few steps of the proof of Theorem 2, we can transform any derivation D into a derivation D' which consists solely of propositional combinations of equations involving numerals and the symbol $\tau(F)$. [Not true, but let's overlook this.]

We now attempt a control of the proof by assigning $\tau(F)$ the value 0. That is, we replace all occurrences of $\tau(F)$ in D' by the numeral 0 to obtain a new 'derivation' D''. As in the remarks following the proof of Theorem 2, we would like to show that all sentences in D'' are correct. Those that come from axioms A1-A10 or the recursion equations defining F are correct, and modus ponens preserves correctness. Thus, any possible false sentences must have been introduced by A12′:

$$F(0) = 0 \rightarrow F(z) = 0,$$

for some numerals z. If all of these are correct, then every sentence in D'' is correct and the original D could not have derived $\neg 0=0$ as this formula is false and is left unchanged by the proof transformations.

If, on the other hand, an instance,

$$F(0) = 0 \rightarrow F(z) = 0,$$

is incorrect, we simply go back to D' and replace all occurrences of $\tau(F)$ by the numeral z. The instances of A12′ now become instances like

$$F(z) = 0 \rightarrow F(x) = 0,$$

for various numerals x. But these implications are correct formulae because $F(z)=0$ is false! Again, all sentences in D'' are correct, whence D could not have been a derivation of $\neg 0=0$.

QED

Hilbert ended his lecture with the remark that one had but to carry through the details of his proof sketch in order to complete the laying of foundations for analysis, and therewith begin the corresponding work on set theory. That there would be difficulties is something Hilbert was aware of; the extent of these difficulties would not become clear until 1930. Indeed, it would only after that become clear that Theorem 2 itself was not without hidden difficulties.

So we finally come to the end of Hilbert's Leipzig lecture of 1922. Three things of a more personal nature happened in 1922. I have already cited Hilbert's offer of a position in Göttingen to Weyl and Weyl's refusal. Bernays was appointed extraordinarius without tenure, but was able to keep his job until the Nazis came to power - Bernays was Jewish. And, as Reid reports in her biography, Hilbert's new physics assistant was unimpressed: Hilbert was showing signs of aging. Hilbert was only 60 and was acting like someone much older. At first it was just assumed Hilbert was suffering from an early, but natural, decline. By the fall of 1925, however, it was known that he was suffering from pernicious anaemia, in those days a generally fatal illness. This illness would play a major rôle in Hilbert's battle with Brouwer.

Before getting too involved in personal matters, however, let us return briefly to 1923 and Brouwer. In 1921 Bernays had spoken to the DMV in Jena, and in 1922 Hilbert lectured to the DMV in Leipzig. The 1923 meeting of the DMV was in Marburg and on 21 September Brouwer spoke on 'Die Rolle des Satzes vom ausgeschlossenen Dritten in der Mathematik' ('On the rôle of the principle of the excluded third in mathematics') and, in addition to a brief summary in the *Jahresbericht*, a version of the lecture was published in another German journal under the title 'Über die Bedeutung des Satzes vom ausgeschlossenen Dritten in der Mathematik, insbesondere in der Funktionentheorie' ('On the significance of the principle of the excluded third in mathematics, in particular in the theory of functions').

As we have already seen, Hilbert's papers on the foundations of mathematics were long on commentary and short on mathematical detail. Brouwer's papers on intuitionism during this period had, on the other hand, more mathematics than philosophy. As the mathematics does not concern us here, I need not say much about Brouwer's papers other than that they exist. The Marburg lecture is an exception, for Brouwer prefaces it with comments on the validity (in finite domains) and the invalidity (in infinite domains) of the Law of the Excluded Middle. Moreover, he finishes this preface with a few comments against formalism, ending with a pungent remark on the pointlessness of giving a consistency proof:

> An incorrect theory unrestrained by any refuting contradiction is thereby no less incorrect, just as a criminal policy unrestrained by a reprimanding court is thereby no less criminal.

Such a remark is certainly enjoyable, especially in the midst of Brouwer's generally dry prose of the period. But it is still hardly an indication of a battle with Hilbert; it is more of a playful continuation of the political metaphor earlier used by Weyl and Hilbert. A real fight would soon be breaking out - the high points being the years 1925, 1927, and 1928.

Through these next few years Brouwer continued to publish papers on topology and on his intuitionistic mathematics, the most important of the latter being a three-part reworking and extension of his 1918/1919 paper, published from 1925 to 1927 in *Mathematische Annalen,* and another paper published in the same journal: 'Über Definitionsbereiche von Funktionen' ('On the domains of definition of functions'). In this latter paper, published in 1927, Brouwer devotes a footnote to his two main objections to Hilbert's formalism. The first of these would take us too far afield to be discussed intelligibly. The second reveals the great difference in Hilbert's and Brouwer's perspective. To Hilbert the consistency of a theory implied the existence of the objects of the theory; to Brouwer the objects were constructed and a consistent false theory proved no truths. To go from consistency to truth required, Brouwer said, the 'Principle of the Reciprocity of Complements', i.e.

$$\neg\neg\varphi \rightarrow \varphi,$$

which was equivalent to the Law of the Excluded Middle. Assuming Hilbert proved the consistency of the Law of the Excluded Middle, he could not conclude its *truth* without assuming its truth, i.e. Hilbert's programme, once carried out, would rest on a circularity.

The point here is philosophically subtle and this is not the correct place to go into it. I will simply say that it seems to me Hilbert was misled by Weyl's remarks on the meaninglessness of quantified formulae into believing them to be meaningless to Brouwer. Thus, to Hilbert 'transfinite formulae' did not have to be true, merely consistent; while, to Brouwer the issue was truth. One of the many ironies of the dispute between Hilbert and Brouwer was that anyone who agreed with Hilbert did not need this programme, and anyone who agreed with Brouwer would not be convinced by its positive outcome should it occur, but yet Hilbert pushed on.

In citing Brouwer's 1927 paper I have once again deserted the chronology. Let me return to 1923 - or, even better: 1924. On 8 January, Bernays lectured in Berlin on 'Neuere Untersuchungen über Hilberts axiomatische Methode' ('New investigations on Hilbert's axiomatic methods'). Perhaps he spoke on Ackermann's dissertation, which Ackermann himself spoke on in Göttingen on 12 and 26 February, and in which Ackermann thought he had completed Hilbert's above-sketched consistency proof. Ackermann, however, found an error and eventually claimed only the consistency of a fragment of 'transfinite' arithmetic.

On 22 July 1924, Brouwer spoke at Göttingen on intuitionism. Is this a sign that he and Hilbert were still on friendly terms? In her biography of Hilbert, Reid describes the end of a talk Brouwer gave in Göttingen. She says, 'After a lively discussion Hilbert finally stood up. "With your methods", he said to Brouwer, "most of the results of modern mathematics would have to be abandoned, and to me the important thing is not to get fewer results but to get more results." He sat down to enthusiastic applause'. Was this the 1924 talk, an earlier one, or a talk Brouwer gave in 1926? Is there hostility in Hilbert's stereotyped remark - a remark that applied to Kronecker and Weyl, but not quite as fairly to Brouwer? In the next year, the mood would indeed be hostile - albeit over matters more political.

The year 1925 was a major one in the dispute, for now there was a real dispute and not just two competing mathematical philosophies. Hilbert gave a major lecture which, unfortunately, confuses rather than clarifies his views. And Weyl published an article reviewing the debate on the nature of the continuum (i.e. real number line) from the time of the ancient Greeks to Hilbert's Leipzig lecture.

I have been promising a fight for so long now that I feel compelled to begin with the real battle between Hilbert and Brouwer. This battle was political and nationalistic: The reader will recall that Europe had lain locked in war from 1914 to 1918, and that Germany had been the villain. Perhaps because this was the first mechanised war to be held in Europe and the damage so great, French mathematicians forgot that, under Napoleon, the French had been the villains only a century earlier. With some vindictiveness, in organising the first post-war International Congress of Mathematicians in Strasbourg in 1920, they ruled that Germans could not attend. The bar to German attendance at the ICM was still in effect in Toronto in 1924.

It so happened that 1926 would mark the century since Riemann's birth and the editors of *Mathematische Annalen* wanted to commemorate the occasion with a special volume. Hilbert, although proud of German achievements in mathematics, was something of an internationalist. He felt that international cooperation was good for mathematics and wanted French contributions to the volume. Brouwer felt otherwise and sought to bar the French, particularly Paul Painlevé who had issued a number of anti-German statements during the war and who was becoming involved in the project. Ludwig Bieberbach, a Berlin mathematician and by then a close friend of Brouwer's, managed a compromise - Einstein, also an editor, would contact friendly French mathematicians who could be asked to contribute to the volume. In suggesting this compromise, however, Bieberbach demanded unanimous agreement by the editors (including Brouwer and himself) and not just the agreement of the main editors (Hilbert, Blumenthal, and Constantine Carathéodory). Hilbert suspected Brouwer was behind this demand. Ultimately, the volume was published with French participation.

Such, in brief, was the nature of the first serious clash between Hilbert and Brouwer. It wouldn't appear to deserve more than a footnote were it not for the fact that the French vs. German issue would again come between Hilbert and Brouwer. The one question it does raise is why Brouwer, a Dutchman, should go to the trouble of opposing Hilbert and Blumenthal on such a matter. To me the most compelling answer is this: The two leading groups of mathematicians in those days were the French and the Germans. Brouwer moved primarily in the German mathematical circle - he had been a member of the DMV since 1909 and had many German friends, particularly in Göttingen which, as already noted, was his second scientific home. If Brouwer himself was not being discriminated against, his friends and closest colleagues were. Moreover, his priority dispute with Lebesgue was still rankling him and he probably hadn't forgotten that the French mathematicians had sided with Lebesgue in this matter.

Weyl's 1925 paper is an impartial account attempting to do justice to both sides. He repeats his earlier account of Brouwer, but acknowledges that his description is not true so much to Brouwer's views as to his own reworking thereof. Although one can detect the beginning of Weyl's transition from supporting Brouwer to supporting Hilbert, his treatment of Hilbert's programme is none too flattering. Weyl compares Hilbert's formalised mathematics to the game of chess: Once the axioms and rules of inference are set up, one merely follows the rules. The consistency proof is comparable to a proof that there can never be 10 queens of a given colour on the chessboard during a game. Weyl even went so far as to call Hilbert's 'actual mathematics' a 'Formelspiel', a 'formula-game'.

The 'reproach', as Hilbert would eventually call it, that Hilbert was trying to turn mathematics into a meaningless game had long been a criticism of formalism in mathematics and, except perhaps in direct reference to Hilbert's programme, did not originate with Weyl's remarks. Hilbert himself would, two years hence, blame Brouwer for the phrase 'formula-game', although it seems to have been used by everyone except Brouwer. (At least, I haven't been able to find such a remark in Brouwer's published papers.)

That Hilbert did indeed deny any meaning to 'actual mathematics', whether by design or accident, dismayed some. Weyl himself said,

> Without doubt: For mathematics to remain a serious cultural concern, some sense must attach itself to Hilbert's formula-game; and I see only one possibility of ascribing an independent intellectual meaning to it, including its transfinite component.

To Weyl, Brouwer had demanded every mathematical assertion to be contentual and, in so doing, had revealed how little content most of mathematics had. Hilbert, on the other hand, he saw as denying any content at all to mathematics. Weyl saw room for compromise: Mathematics would become a theoretical science, like physics, where not every assertion had intuitive content.

Was Hilbert really turning mathematics into a formula-game? Was this his intent? Was Weyl's representation of Hilbert's views as liberally re-interpreted as his representation of the views of Brouwer? I would answer these questions yes, no, and yes, respectively. It is nowadays a commonplace among mathematicians that mathematics is a meaningless game of formal symbols, and Hilbert's authority is invoked in support of this view. Hilbert's contemporaries were more subtle thinkers. G.H. Hardy, perhaps the quintessential purist in mathematics, is quoted by Reid:

> ... is it really credible that this is a fair account of Hilbert's views, the view of the man who has probably added to the structure of significant mathematics a richer and more beautiful aggregate of theorems than any other mathematician of his time? I can believe that Hilbert's philosophy is as inadequate as you please, but not that an ambitious mathematical theory which he has elaborated is trivial or ridiculous. It is impossible to suppose that Hilbert denies the significance and reality of mathematical concepts, and we have the best of reasons for refusing to believe it: 'The axioms and demonstrable theorems,' he says himself, 'which arise in our formalistic game, are the images of the ideas which form the subject-matter of ordinary mathematics'.

['Formalistic game' is a mistranslation of 'Wechselspiel'. The passage from Hilbert's Leipzig lecture quoted by Hardy directly follows Hilbert's remark on how mathematics grows alternately by deriving theorems and adding new axioms. It was this 'interplay' Hilbert described as 'Wechselspiel'.] Hardy may not have chosen the most convincing quotation; I am inclined to point to the earlier cited remark that 'one can presumably prove a finitist statement also without application of transfinite means of proof' and the contentual nature of finitist truths. This and the comparison with the use of ideal elements in algebra and geometry suggests that *implicitly*, if not yet explicitly, Hilbert's programme was already a rendering of mathematics as the theoretical science suggested by Weyl. By the end of 1927 this would be clear.

In Hilbert's Münster lecture of 1925, published in two slightly different versions in *Mathematische Annalen* and *Jahresbericht* in the next two years, one can find remarks that seem to support the claim that Hilbert was trying to create a theoretical science out of mathematics. One can also find evidence supporting virtually any philosophy of mathematics other than the extreme sort of formalism (mathematics is a meaningless activity) generally attributed to Hilbert and the platonism implicit in his equation of consistency and existence - a platonism which would explain his blindness to the realisation that consistency alone might be unconvincing to others and which might also explain the horrible mess he made of explaining the 'finitist point of view' in this lecture. Another explanation of the latter could be his illness: Hilbert was ill at the time of this talk.

The Münster lecture divides into two parts. The second part outlined an incorrect proof of Cantor's continuum hypothesis (any uncountable set of real numbers can be put into one-one correspondence with the set of all real numbers) and is not particularly relevant to our discussion. I could propose it as an explanation of why, when Hilbert got down to details, he did not say anything about the actual workings of his consistency programme; or, I could cite it as additional evidence of the effect his illness was having on his powers of concentration. Let us leave both interpretations in the realm of superficial asides and consider the more pertinent first part. Instead of building his exposition on his earlier two lectures, Hilbert explained anew the whole situation - finite vs. infinite, the ideal nature of existential assertions, etc. At the level of high generality, it is a good discussion and well merits Jean van Heijenoort's description as a 'clear and forceful presentation of Hilbert's ideas at the time on the foundations of mathematics', and one can debatably agree with Van Heijenoort's further remark that among Hilbert's papers on foundations it 'stands out as the most comprehensive presentation of Hilbert's ideas'. When it gets down to details, however, only the adjective 'forceful' still applies.

It is detail that I wish to discuss - specifically the detail that is new. Unfortunately, it is on just this point that Hilbert is not very clear, and I have had to consult both German and English versions of the paper and both past and future papers ('future' relative to 1925) in order to come to a clear understanding - one so clear and fitting his remarks so well that I declare it the correct interpretation: No-one, though he speak with the tongues of angels, will convince me otherwise. This detail is a trichotomy replacing the old dichotomy between finitary propositions and transfinite, or ideal, formulae. There are now *real propositions, finitary general propositions,* and *ideal propositions.*

The real propositions - the term 'real' being added first in 1927 - are, Hilbert says, 'of no essential interest' in themselves. They are simple propositional combinations of equations involving primitive recursive functions and fixed numerals:

$$2+3=3+2, \quad 1+1=2, \quad \text{etc.}$$

They are directly contentual assertions verifiable by direct computation. Their importance lies primarily in affording a control on the results of formal mathematical proofs:

> The science of mathematics is by no means exhausted by numerical equations and it cannot be reduced to these alone. One can claim, however, that it is an apparatus that must always yield correct numerical equations when applied to integers.

I think we can read in these lines just the sort of theoretical view Weyl demanded: Mathematics is an abstract theoretical science subject to numerical control just as physics is an abstract theoretical science subject to experimental control.

Real propositions do not exhaust the class of finitistic propositions. There are also what I shall call here the finitary general propositions - assertions of the form 'for every numeral n, $n+1=1+n$', which Hilbert would have written in Leipzig as the free-variable formula,

$$x+1 = 1+x.$$

The general assertion 'is from the finitist point of view *incapable of being negated*' because the negation ('for some numeral n ...') 'cannot be interpreted as a combination, formed by means of "and", of infinitely many numerical equations'. [It may seem odd that infinite conjunctions are finitistic assertions. However, (i) one cannot seriously consider any science which does not propose universal laws, (ii) there are finitary schematic methods of proof, (iii) consistency, which must be proven finitistically, is such a universal assertion, and (iv) it was his *existential* theorems that Hilbert had been criticised for.]

Finally, there are the ideal propositions that are not really propositions at all, but formal symbols manipulated according to pre-determined rules - those of what Hilbert did not like being called the 'formula-game'.

Though it is unnecessary for our general discussion, I would consider myself remiss in my responsibility as a chronicler if I didn't mention the delightful irony and famous quotes to be found in Hilbert's Münster lecture. The former is given by one of his polemics:

> ... the literature of mathematics is replete with absurdities and inanities... for example, some stress the stipulation, as a kind of restrictive condition, that, if mathematics is to be rigorous, only a *finite* number of inferences is admissible in a proof - as if anyone had ever succeeded in carrying out an infinite number of them!

The irony is not so much that he is here criticising himself, but that he himself would in 5 years' time be considering proofs with infinitely many inferences.

His quotable remarks include 'mathematical analysis is but a single symphony of the infinite', 'no-one shall be able to drive us from the paradise that Cantor created for us', 'no-one, though he speak with the tongues of angels, will keep people from negating arbitrary assertions, forming partial judgments, or using the principle of the excluded middle', and, concerning what to do when confronted with a foundational problem, 'let us remember that *we are mathematicians*' and simply solve the problem. Moreover, he reached all the way back to 1900 to recall 'that within us we always hear the call: here is the problem, search for the solution; you can find it by pure thought, for in mathematics there is no *ignorabimus*'.

If the year 1925 saw Hilbert and Brouwer fighting, Hilbert giving a confused lecture attacking even himself, and Weyl inadvertently sowing the seeds of future discord with his caricature of Hilbert's programme, the year 1926, on the other hand, brought good news. Hilbert discovered that a treatment for pernicious anaemia of some promise was being experimented with in America and, after some manoeuvring, he was receiving the treatment and his health was improving. And: Hilbert and Brouwer had a reconciliation, albeit a short one.

The occasion of the reconciliation was a summer meeting of topologists in Göttingen. By 1926, Brouwer was no longer actively working in topology, but he did keep in touch with the subject throughout the 1920s. Not only did he have young assistants in topology in Amsterdam, but his home was a Mecca for young topologists - and occasional other mathematicians - from Germany and Russia. Moreover, despite his differences with Hilbert, Brouwer was still friends with many of the other Göttingen mathematicians. Thus, the mathematicians at Göttingen wanted Brouwer to come and they wanted a reconciliation.

The method of reconciling Hilbert and Brouwer was simplicity itself. Emmy Noether, whom Weyl once described as 'warm like a loaf of bread', supplied the necessary *Gemütlichkeit*: Dinner was at her home. Sitting at the table with Hilbert and Brouwer were also Richard Courant, Edmund Landau, and a host of younger mathematicians, among them P.S. Alexandroff and Heinz Hopf, later joint authors of a famous topology book. The task of getting the conversation going fell to Alexandroff. According to him, the best way of getting two warring factions together is to find a common enemy. Thus, Alexandroff brought up the subject of a 'famous Luckenwalder function theorist', P. Koebe. His success exceeded all expectations: Hilbert and Brouwer were soon 'falling all over themselves in a spirited exchange of opinion, during which they agreed ever more in their views of that function theorist'. The two became progressively friendlier and finished toasting each other. These good feelings lasted the entire period of Brouwer's stay in Göttingen.

Unfortunately, Brouwer eventually had to return to Amsterdam and the two were publicly attacking each other the following year. Hilbert's attack may have been occasioned by the heavy interest in Brouwer's lectures, in March of that year, in Berlin. Success in lectures on intuitionistic mathematics might be a bit hard to believe nowadays. Most mathematicians have little understanding of, hence little patience with, matters philosophical and if they comment on such matters at all it is usually in derision, as in the following opening remark in a paper of Oskar Perron of 1926:

> A set of numbers bounded from below (above) has a greatest lower (least upper) bound. Despite Brouwer and Weyl this fact, fundamental and indispensible for analysis, is known to have been clearly enunciated by Bolzano in 1817; it can be found just as clearly already in a posthumous manuscript of Gauss, which was written around 1800.

Berlin, however, was an exception. Under the leadership of Klein, and continuing under Hilbert, Göttingen - not Berlin - had become the capital of German mathematics. The rivalry between the two universities was intense and much of the interest in Brouwer's lectures may well have been in Brouwer himself, the thorn in Hilbert's side, and not in intuitionistic mathematics itself. Brouwer spoke to an overflowing hall, and his lectures were mentioned in the newspapers.

Perhaps Hilbert, who had been roused to action at Weyl's 1920 conversion, now saw another clear and present danger. In any event, in July he was in Hamburg again firing off a fresh salvo of polemics:

> ... it is part of the task of science to liberate us from arbitrariness, sentiment, and habit and to protect us from the subjectivism that already made itself felt in Kronecker's views and, it seems to me, finds its culmination in intuitionism.

Probably the most famous line of the paper is:

> Taking the principle of excluded middle from the mathematician would be the same, say, as proscribing the telescope to the astronomer or the boxer the use of his fists.

Apparently Hilbert was aware of his image as a polemicist; for, he prefaces his criticism of Brouwer with the disavowal, 'Not because of any inclination for polemics, but in order to express my views clearly and to prevent misleading conceptions of my own theory, I must look more closely into certain of Brouwer's assertions'. He then proceeds to compare Brouwer to Kronecker in declaring existence statements to be meaningless (fair enough, if a bit insulting), and then attributes to Brouwer Weyl's paper economy and formula-game metaphors. Hilbert was, to put it bluntly, angry; this was the lecture about which Weyl, who was in the audience, would say that there were 'anger and determination in Hilbert's voice'.

The anger, if not the determination, was irrational. For, progress on the consistency proof was so great that everyone in Hilbert's camp was certain success was forthcoming, and, moreover, Hilbert had a trump card to play - an application of consistency proofs that even Brouwer would have to admit the significance of: a method of obtaining finitistic proofs of finitary general propositions from formal proofs of their ideal representations.

I have offered an awkward statement of this last in order to remain as faithful as possible to Hilbert's views. In a lecture in December 1930 - again in Hamburg - Hilbert would point out the difference between a finitary general proposition, say,

$$1+x = x+1,$$

and the similar ideal proposition,

$$\forall x(1+x=x+1).$$

The former is the assertion that,

$$1+n = n+1,$$

for all numerals n; the latter also asserts the equation for meaningless infinitary constructs involving the τ-function - or, rather, a replacement to be cited shortly. If we agree to ignore such subtleties, we can say that Hilbert said the following: Let $\mathfrak{S}$ be a formal system of finitary arithmetic and let $\mathfrak{T}$ be some system of transfinite mathematics. Suppose $\mathfrak{S}$ proves the consistency of $\mathfrak{T}$. Then: For any universal assertion φ, if $\mathfrak{T} \vdash \varphi$ then $\mathfrak{S} \vdash \varphi$.

I quickly note that it is the finitary general propositions that Hilbert is talking about here. For the more restricted class of real propositions, i.e. simple closed instances of quantifier-free formulae, Hilbert had already shown, if not mentioned (after all, such is of no essential interest), that the corresponding conservation result follows from the controllability of the results of mathematical derivations - which controllability follows from the method of his consistency proof. What Hilbert was doing now was settling the epistemological question he had brought up but not dared to answer in his Leipzig lecture - namely, the question of the finitistic derivability of any finitistic statement obtained via transfinite means of proof. It was just this breakthrough that transformed Hilbert's programme into Hilbert's *P*rogramme and gave rise to the myth that, all along, Hilbert had known what he was doing. We shall discuss this new version of the programme when we reach the year 1930, when this new version of the programme finds its clearest (?) statement.

Hilbert proved his method by means of an example, more-or-less in the way in which one proves theorems in Euclidean geometry by drawing specific triangles. The example he chose was Fermat's theorem,

$$\text{FT: } \forall xyzw(x>1 \wedge y>1 \wedge z>1 \wedge w>2 \rightarrow x^w + y^w \neq z^w).$$

We reason finitistically: Suppose we are given numerals k,m,n,p such that,

$$k>1 \wedge m>1 \wedge n>1 \wedge p>2 \wedge k^p+m^p = n^p. \tag{*}$$

This assertion can be verified by a simple computation, which translates directly into a proof in the transfinite system. But, by the provability of FT within the transfinite system, we see that,

$$k>1 \wedge m>1 \wedge n>1 \wedge p>2 \rightarrow k^p+m^p \neq n^p, \tag{**}$$

is provable within the system. This means that

$$k^p+m^p = n^p \quad \text{and} \quad k^p+m^p \neq n^p$$

are both provable transfinitely, contrary to the consistency of the transfinite system. Hence (*) is false and we have proven that for any numerals k,m,n,p (**) holds, i.e. we have proven FT. [Actually, we have only proven $\neg\neg$(**), but the validity of $\neg\neg$(**) $\rightarrow$ (**) is intuitionistically acceptable because of the simple nature of (**). In other words, Hilbert was right that Brouwer would agree to the validity of Hilbert's method - once the finitistic consistency proof was given.]

As we shall see, Brouwer never had a chance to accept Hilbert's method. Weyl, on the other hand, was in the audience and made some remarks immediately after the lecture. Despite his occasional stylistic excesses, Weyl was perhaps the only one of the major participants who could see clearly and dispassionately all the issues involved. He began with the words, 'Permit me first to say a few words in defence of intuitionism'. He emphasised that the predominant view of mathematics had been that it was 'a system of contentual, meaningful, and evident truths'. Brouwer, he said, was the first to realise that this was no longer the case, that mathematics had transcended content, meaning, and evidence. Brouwer's solution was the obvious one and 'it does not seem strange to me that Brouwer's ideas have had a following'. Weyl did not say that Hilbert had been a bit unfair to Brouwer in his talk when he said, 'To make it a universal requirement that each individual formula then be interpretable by itself is by no means reasonable; on the contrary, a theory by its very nature is such that we do not need to fall back upon intuition or meaning in the midst of some argument'. Rather more tactfully, Weyl pointed out that Hilbert was proposing a radical re-interpretation of the meaning of mathematics, in effect - though Weyl did not say this explicitly - turning the subject into the theoretical science Weyl had suggested in his 1925 paper. Following this, Weyl noted that 'as I am very glad to confirm, there is nothing that separates me from Hilbert in the epistemological appraisal of the new situation thus created', which I think means he was now supporting Hilbert - maybe. [Lecturing before an American audience in 1930, Weyl would say,

> My opinion may be summed up as follows: if mathematics is taken by itself, one should restrict oneself with Brouwer to the intuitively cognizable truths... nothing compels us to go farther. But in the natural sciences we are in contact with a sphere which is impervious to intuitive evidence; here cognition necessarily becomes symbolical construction. Hence we need no longer demand that when mathematics is taken into the process of theoretical construction in physics it should be possible to set apart the mathematical element as a special domain in which all judgments are intuitively certain; from this higher viewpoint which makes the whole of science appear as one unit, I consider Hilbert to be right.

Query: From this higher viewpoint, is it necessary to have an intuitively certain consistency proof?]

Taken by itself, transfinite mathematics could only be rescued by the consistency proof. How was the search for this holy Grail of proof theory progressing? In 1922, Hilbert had sketched his proof of consistency for an exceedingly limited case of his transfinite axiom for the counterexample-seeking τ-function. The τ-function was soon replaced by a choice function ϵ satisfying,

$$A(a) \rightarrow A(\epsilon_a(A(a))).$$

Following this cosmetic change, Ackermann published a consistency proof in 1924. However, he discovered an error and tacked a footnote onto his paper restricting the construction of ϵ-terms. Soon he was working on an improvement and wrote to Bernays about his new proof. Johann von Neumann criticised the proof and worked out his own consistency proof - with restricted induction - and published it in 1927. However, in giving a sort of progress report on the consistency proof towards the end of his Second Hamburg Lecture, Hilbert briefly discussed Ackermann's second proof, and the published version, coming out in 1928, of this lecture was accompanied by a note by Bernays explaining the proof in more detail. The Hilbert school firmly believed that the proof was almost complete. There was only the small matter of proving the finiteness of the number of substitutions of numerals for ϵ-terms before a given derivation would be transformed into a derivation-like collection of correct real propositions. Indeed, the following year Hilbert would announce that Ackermann and Von Neumann had already proven the consistency of the arithmetic of the integers and that Ackermann had just to prove the corresponding finiteness result for analysis to complete his consistency proof for this latter theory.

If Hilbert broke the truce in 1927, Brouwer was not sitting idly himself. On 17 December in Amsterdam and again a couple of months later in Berlin, Brouwer made his own attack in a lecture entitled 'Intuitionistische Betrachtungen über den Formalismus' ('Intuitionistic reflections on formalism'). The contents of this lecture were published early in 1928, both in the proceedings of the Dutch Royal Academy and, under the sponsorship of his friend Bieberbach, in the *Sitzungsberichte der Preussischen Akademie der Wissenschaften* ('Reports of the Meetings of the Prussian Academy of Sciences'). The lecture begins with a list of those papers of Brouwer and Hilbert that Brouwer wished to discuss. This list did not include Hilbert's Second Hamburg Lecture, which had not yet been published.

The first section of this new paper has a list of 4 insights which would, upon their ultimate acceptance, render the choice between formalism and intuitionism a matter of taste. Briefly, these were:

FIRST INSIGHT. The formalist distinction between transfinite mathematics and finitary mathematics is indispensable, as is the recognition of the need of the intuitionistic mathematics of the set of natural numbers for this finitary mathematics.

SECOND INSIGHT. One must not use the Law of the Excluded Middle thoughtlessly, but ought rather to investigate where it can be used. For intuitive (contentual) mathematics it is valid only in finite systems.

THIRD INSIGHT. The Principle of the Solvability of Every Mathematical Problem is to be identified with the Law of the Excluded Middle.

FOURTH INSIGHT. The justification of the formal system of transfinite mathematics via a finitistic consistency proof contains a circulus vitiosus.

According to Brouwer, Hilbert had already accepted the First and Third Insights and it was just a matter of time before he would accept the other two. Actually, Hilbert had accepted the First, but only half of the Third Insight - this latter because he did not yet accept Brouwer's equation of the Principles of the Solvability of a Problem *in Principle* and *in Practice*. In any event, there was no hope of Hilbert's accepting the Second Insight until he had accepted the Fourth and here, unbeknownst to Brouwer, Hilbert had shown that in a special case the justification was not circular.

It is in his discussion of these insights that Brouwer, for the first time since Hilbert began attacking him in Hamburg in 1921, responded in kind in print. For example, with respect to the First Insight, he notes that it was hinted at by Poincaré, appeared first in print by Brouwer in his dissertation in 1907, and had been discussed by Brouwer with Hilbert in the fall of 1909. (This last was noted in a footnote; it is the only reference in print by either party to their conversations in the dunes of Scheveningen when their acquaintance/friendship began.) Brouwer added that Hilbert subsequently published this distinction under a new nomenclature. In case the meaning of this should not be clear, he followed the individual discussions with a few closing remarks. These begin with the observation that formalism had got nothing but good from intuitionism and could expect more. There follows:

> Accordingly, the formalist school should afford some recognition to intuitionism instead of polemicising against it in a jeering tone and in the process not once keeping proper mention of authorship. Moreover, the formalist school should reflect on [the fact] that up till now nothing of mathematics proper has been secured in the frame of formalism... If thus the formalist school, according to its remarks..., has noticed modesty in intuitionism, then it should find cause therein not to take second place with respect to this virtue.

These are strong words, but apt.

Hilbert's and Brouwer's mutual attacks were published in 1928. The real fight that year, however, was far less public. It began with the preparations for the coming International Congress of Mathematicians in Bologna. As already noted, German mathematicians had been barred from attendance in 1920 in Strasbourg and in 1924 in Toronto. The Germans were allowed to attend in 1928, but not quite on a level of equality with the others. Moreover, the programme of the meeting included an outing to 'liberated' areas, a direct slap in the German face as it were. Bieberbach, who had partially sided with Brouwer in the affair of the Riemann volume, announced that it would be a shame if a large number of Germans attended such a meeting and Brouwer called for an all-out boycott. Hilbert, who felt international contact more important than nationalistic feeling, used the full force of his prestige to break the boycott and personally led the German delegation to Bologna, his health much improved by the American treatment. His lecture was met at beginning and end with thunderous applause.

The content of this lecture will be discussed later. First, I shall dispose of Brouwer - something Hilbert probably would have liked to have done as he discussed the Brouwer problem with colleagues that August in Bologna. He soon came up with a solution: In October, he wrote to Brouwer telling the latter that he was fired from the editorship of *Mathematische Annalen.* The actual dismissal was a complicated affair that went on for several months while Hilbert's forces tried to muster the proper authority to effect Hilbert's pogrom.

The *Mathematische Annalen* was, as I said earlier, one of - if not *the* - leading mathematical journals of the day. To be an editor thereof, as Brouwer had been since 1915, was a mark of great distinction. To be fired therefrom was a great insult, particularly if the editor being fired was the most conscientious and hard working of the lot, as Brouwer apparently was. Of Hilbert's forces, the only one who appears to have had any qualms about firing Brouwer was Constantine Carathéodory, who did his duty to Hilbert and then resigned as editor. The others - particularly Blumenthal and Courant - did not share Carathéodory's moral reservation; indeed, Blumenthal executed his task with unbecoming zeal. Once he had set the wheels in motion, however, Hilbert remained aloof from the whole process.

But for Hilbert's aloofness, it would be tempting to compare the situation with the English Royal Society's investigation into the charge that Leibniz had plagiarised Newton: Leibniz, having faith in the integrity of the society expected an impartial investigation, not realising that Newton himself was ghost-writing the report. Brouwer was now trying to prepare his defence for the ultimate tribunal of all the editors, unaware that there was to be no tribunal - the delay in his receiving official notification not being caused by any question of whether or not to fire him, but of how: The decision had been made by Hilbert and few questioned it.

The details of the struggle are fascinating, but hardly relevant to our purpose and I refer the reader to Van Dalen's report cited in the Reading List. Only a few things need be said here.

The effect on Brouwer was devastating. Brouwer had not merely been publicly humiliated, but he had been betrayed by his friends - not all of them: Most of his friends in Göttingen had had nothing to do with the affair, Carathéodory had tried to spare Brouwer's feelings, and Bieberbach stood steadfastly by Brouwer during the struggle. Nonetheless, Brouwer was left a broken man: Throughout the next decade he hardly published anything, and when he did resume publication in the 1940s he never reached the heights he had achieved in the two decades from 1908 to 1928; his creative life, as Walter van Stigt put it in his moving account of Brouwer's philosophical activity of the period, was over.

Alas, matters are never as black and white as we like to paint them. Bieberbach - the true and faithful - later became a Nazi and he and Brouwer disagreed on the latter's acceptance of Jewish editors for *Compositio Mathematica*, a journal founded by Brouwer in 1930. The photographs I've seen of Blumenthal, twice the villain of our story, show such a pathetic character that it is difficult to condemn him. A Jew, he was ultimately deprived of his position by the Nazis and, after a temporary escape to the Netherlands, died in a German concentration camp in 1944.

That leaves us with Hilbert. It is doubtful that he ever realised the effect his firing of Brouwer had. He was, it seems, temporarily unbalanced. His illness had taken a drastic downward turn - apparently through some bad medicine taken sometime after the Bologna meeting - and he, and everyone around him, thought he might die. Should he die, he feared Brouwer would take over *Mathematische Annalen*. Was he afraid that, with such a tool in his hands Brouwer would alter the course of mathematics?

At this point it might be worthwhile to consider how successful Brouwer had been at conversion. In 1919, in a reprint of one of his 1908 publications, Brouwer added a note that 'the opinions which it defends have as yet not found many supporters'. His major convert was Weyl in 1920, but Weyl was only a philosophical convert - not for long either - and did not do intuitionistic mathematics. To a somewhat lesser extent, one could add Bieberbach to the list of converts. Brouwer did have a few students - he cited one in his Marburg lecture of 1923 - but the only one whose name is remembered today was Arend Heyting. [A small aside: In 1927, under the suggestion of Gerrit Mannoury, a self-taught Dutch mathematician, the Dutch Mathematical Association put up a prize question on formalising Brouwer's mathematics insofar as such was possible. Heyting submitted a paper and early in 1928 he was declared the winner. The paper was to have appeared in *Mathematische Annalen,* but appeared instead in the *Sitzungsberichte* of the Prussian Academy, of which Bieberbach was a member.] This was approximately the limit of his success.

Hilbert's fear of Brouwer was not only objectively irrational in the sense that he manifestly overestimated Brouwer's influence - Perron's derisive remark cited earlier is a typical mathematician's reaction - but it was also irrational in that it was not Brouwer, but Kronecker, with whom he had been fighting all along. Bernays would eventually say as much, but we do not need such authority to state this as it is clear already from Hilbert's often recalling his first solution of the invariant problem, his constant comparison of Brouwer with Kronecker, and his attribution to Brouwer of attitudes foreign to the latter's character, such as the reproach that Hilbert's formalism reduced mathematics to a meaningless game. [In his youth he might have said such a thing, but in the 1920s his opinion of Hilbert remained high. Despite the polemics of his last paper cited here, when Hilbert wrote to fire him, Brouwer wrote to Hilbert's wife asking her to intercede, saying that Hilbert was too good a man to go forward with such action. I don't believe this was mere rhetoric - Brouwer was no dissembler.]

Upon reflexion it appears that Brouwer was not the only tragic casualty of the affair. Hilbert himself was destroyed spiritually. Like Alexander the Great who forgot his contempt for the Persian despots and died a more despotic ruler than any of them, Hilbert, in doing battle with the restrictive policies of Kronecker, was more effective than Kronecker in his own restrictive policy - this time not restricting mathematical freedom, but restricting philosophical freedom.

Following such dramatic events, the rest of the story will appear something of a tame anticlimax. Hilbert stopped polemicising against Brouwer. Whether or not he realised that Brouwer was safely out of the way, his programme was so far advanced that it continued to roll along and, like a snowball, gain in volume. The extension of his list of proof theoretic desiderata had begun already in Bologna.

Hilbert's Bologna address consists of a core of four problems embedded in discussion. For the first problem, Hilbert incorrectly noted that Ackermann and Von Neumann had proven the consistency of the arithmetic of the integers. Hilbert now wanted to extend the proof to allow the choice function ϵ_x to apply not merely to arithmetic formulae, but to function variables. He remarked that Ackermann had almost proved this, there remaining only an arithmetically elementary finiteness theorem to be established.

The second problem was to give a consistency proof for a stronger theory in which more advanced parts of analysis and some set theory could be carried out.

The third and fourth problems were variants of the completeness of arithmetic, conjectured on the basis of Dedekind's categoricity result. Problem III is stated in terms of consistency:

III. If one can prove the consistency of φ with the axioms of number theory, then one cannot prove such consistency for $\neg\varphi$.

The point to such an odd statement is simply this: Consistency implies existence. If φ and $\neg\varphi$ were both consistent, there would be two non-isomorphic systems of arithmetic, contrary to Dedekind's result. Problem IV is a more familiar statement of completeness:

IV. If φ is not provable from the axioms of arithmetic, then adding φ as an axiom yields a contradiction [i.e. $\neg\varphi$ is derivable].

Kurt Gödel's refutation was two years away.

Actually, Gödel has another connection with this paper. Immediately following the statement of Problem IV, Hilbert also raised the question of the completeness of predicate logic. Gödel would read a statement of this problem in Hilbert's joint book with Ackermann (a revision of Hilbert's lecture notes from a decade earlier) published in 1928, and prove the completeness theorem in his dissertation in 1929, publishing the result in 1930.

The completeness theorem finally gave rigorous expression to Hilbert's contention that consistency implied existence. This contention was such a commonplace that in the (unpublished) introduction to his dissertation Gödel felt it necessary to defend his having bothered to prove the result at all. He also explained that, unlike the consistency problem, the question of completeness could be 'meaningfully' posed within the transfinite system. Thus, he saw no reason to restrict his methods. This would not be the case a year later with his proof of incompleteness.

The impromptu announcement of the First Incompleteness Theorem was the big non-event of 1930. This took place during the first of no fewer than 3 conferences held in September in the east Prussian town of Königsberg (now: Kaliningrad). This first conference was a meeting organised by philosophers to take advantage of the second - a meeting of the Society of German Scientists and Physicians. The third was a meeting of the DMV and doesn't enter our story.

The first conference was a three-day affair devoted to the foundations of mathematics. The three major competing philosophies of mathematics were presented by three major proponents thereof: Rudolf Carnap spoke on logicism, Heyting on intuitionism, and Von Neumann on formalism à la Hilbert.

Von Neumann described not so much Hilbert's actual programme as a variant of the mythic Programme of showing the use of the transfinite to be an unnecessary but generally efficient detour:

> The problems which Hilbert's proof theory has to solve are the following:
>
> 1. To enumerate all symbols which find application in mathematics and logic...
>
> 2. To characterise unambiguously all combinations of these symbols which stand for expressions classified as 'meaningful' in classical mathematics. These are called 'formulae'...
>
> 3. To give a constructive procedure that allows the successive generation of all formulae which correspond to 'provable' assertions of classical mathematics. This procedure is consequently called 'proving'.
>
> 4. To show (in a finite-combinatorial way) that those formulae corresponding to finitistically controllable (arithmetically checkable) assertions of classical mathematics can be proven (i.e. constructed) as in 3 if and only if the actual 'check' just mentioned yields the correctness of the corresponding mathematical assertions.
>
> Were 1-4 secured, the absolute reliability of classical mathematics for the following purpose would be established: as a shorter method for calculation of arithmetical expressions, for which the elementary working out would be too involved.

Von Neumann added that problems 1-3 had been solved by the work of Russell and his school, and the real problem was thus 4. For this he said that a consistency proof sufficed and he even sketched a proof of the correctness of any controllable assertion represented by a formula derived in a consistent formalism.

It is not clear from these remarks, or indeed from the rest of his paper, exactly what constituted a 'controllable assertion'. There is an implicit definition of controllability in his introductory remarks: An assertion is controllable if, when we've made a mistake in its proof, we can detect this mistake through a finite procedure other than rereading the proof. The naturally controllable assertions would be the finitary general propositions as, when we've proven such an assertion, say, 'for all numerals x, E holds of x', we can make simple calculations to see if E holds of 0, if E holds of 1, etc.; and, if the assertion is false, we will see this when we check if E holds of the wrong n. However, when proving that problem 4 reduces to proving consistency, it is real propositions, i.e. numerical formulae, he considers. Moreover, the formulation of problem 4 and its ensuing comment suggests a restriction to decidable assertions - the obvious ones being the real propositions.

Von Neumann was describing a general conservation programme of clear, if not definite, meaning, not just the intuitionistically unconvincing consistency programme. For this reason, during an organised discussion two days later, Brouwer's disciple Heyting announced his pleasure with the conference and his full acceptance of Hilbert's Programme. It was then that Gödel spoke up on the issue of the exact extent of conservation to be expected to follow from a consistency proof:

> According to the formalist conception one adjoins to the meaningful statements of mathematics transfinite (pseudo-)statements which in themselves have no meaning but only serve to make the system a well-rounded one just as in geometry one achieves a well-rounded system by the introduction of points at infinity. This conception presupposes that when one adds to the system $\mathfrak{S}$ of meaningful statements the system $\mathfrak{T}$ of transfinite statements and axioms and then proves a statement from $\mathfrak{S}$ via a detour over statements from $\mathfrak{T}$ then this statement is also correct in its content so that through the addition of the transfinite axioms no conceptually false statements become provable. One commonly replaces this requirement with that of consistency. I would like to indicate that these two requirements cannot by any means be regarded as equivalent. For, if a meaningful sentence p is provable in a consistent formal system $\mathfrak{A}$ (say that of classical mathematics), then all that follows from the consistency of $\mathfrak{A}$ is that not-p is not provable within the system $\mathfrak{A}$. Nevertheless it remains conceivable that one could recognise not-p through some conceptual (intuitionistic) considerations which can*not* be formally represented in $\mathfrak{A}$. In this case, despite the consistency of $\mathfrak{A}$, a sentence would be provable in $\mathfrak{A}$ the falsehood of which one could recognise through finite considerations. However, as soon as one construes the concept 'meaningful statement' sufficiently narrowly (for example restricted to finite numerical equations) such a thing cannot occur. In contrast it would be, e.g., entirely possible that one could prove with the transfinite methods of classical mathematics a sentence of the form $\exists x F(x)$ where F is a finite property of natural numbers (e.g. the negation of the Goldbach conjecture has this form) and on the other hand recognise through conceptual considerations that all numbers have the property not-F; and what I want to indicate is that this remains possible even if one had verified the consistency of the formal system of classical mathematics. For, one cannot claim with certainty of any formal system that all conceptual considerations are representable in it.

This much is clear from Gödel's remarks: He knew of Hilbert's Programme prior to Von Neumann's talk as his terminology differs; and his knowledge is not based on a careful reading of Hilbert's papers. The phrase 'one commonly replaces ...' suggests that, like the description of Hilbert's proof theory as turning mathematics into a formula-game, the new conservation Programme was being discussed and the myth that it preceded the consistency programme consciously was being promulgated.

For our purposes, it doesn't really matter whether Hilbert wanted to prove consistency in order to establish some conservation result or whether Hilbert merely wanted to prove consistency because of some platonistic belief that consistency implied existence, and merely realised by accident that the conservation result he had clearly believed in since the advent of his programme followed from the successful completion of the consistency programme. It doesn't matter because Gödel was only a few breaths away from destroying the conservation programme for finitary general propositions, and therewith the consistency programme.

Mistaking the last sentence of Gödel's remark already cited for Brouwer's doubts about the formal representability of our mathematical thoughts, Von Neumann noted that 'It is not settled that all modes of inference that are intuitionistically permitted can be represented formally'. To avoid misunderstanding, Gödel now stated explicitly his incompleteness result:

> One can (under the assumption of the consistency of classical mathematics) even give examples of statements (and even of the sort of Goldbach's or Fermat's), which are conceptually correct but unprovable in the formal system of classical mathematics. Therefore, if one adjoins the negation of such a statement to the axioms of classical mathematics, then one obtains a consistent system in which a conceptually false sentence is provable.

Gödel's announcement of his First Incompleteness Theorem, being given as a critique - at first hypothetical - of Hilbert's Programme and the inadequacy of consistency for Hilbert's newly perceived purpose, did not make a big impression on everybody: Kurt Reidemeister finished off the discussion with a quick recapitulation - *sans* Gödel's remarks - of the discussion. Von Neumann understood what Gödel said and the two discussed Gödel's work that day... but that story is for another occasion.

What Gödel had done, though it was not immediately clear because of his doubts about what constituted finitary mathematics and the general unfamiliarity of the day with the formal systems involved, was this: He had given a finitistic construction of a sentence asserting its own unprovability. On assumption of consistency, it was finitistically verifiable that the sentence was unprovable. Transfinitely, it was clear that the sentence was true. Applying the construction to the finitary system $\mathfrak{S}$ would produce a finitary general sentence asserting its unprovability in $\mathfrak{S}$. This unprovability is recognised in the transfinite system $\mathfrak{T}$, whence $\mathfrak{T}$ proves the sentence. However, if $\mathfrak{S}$ could prove the consistency of $\mathfrak{T}$, then by Hilbert's 1927 argument, $\mathfrak{S}$ could also prove Gödel's sentence. Then $\mathfrak{T}$ would prove both the provability and unprovability of Gödel's sentence, and $\mathfrak{T}$ would be inconsistent! It follows that Hilbert's conservation Programme and consistency programme must both fail. All that remained were the fine-tuning of the result, the working out of further consequences of Gödel's technique, and getting the message across. Only the third of these will be touched on briefly in the remainder of the present paper.

Evidently, Hilbert did not get the message until some time early in 1931 - and then he more-or-less rejected it. In Königsberg it seems no-one told him about Gödel's result and, apparently on the day after Gödel's unexciting announcement, Hilbert gave his lecture - or lectures: This was the occasion of Hilbert's retirement and Königsberg was the site of Hilbert's youth and first professional position. Following his lecture to the German scientists and physicians, Hilbert was whisked off to the local radio station to deliver a shortened form of his speech to the general populace. His subject, natural science and logic, was a broad one and he spoke in generalities, finishing with the observation that the reason the French positivist philosopher Auguste Comte had been unable to find any unsolvable problems is that there aren't any, emphasising this point with the motto,

> We must know;
> We will know.

On 24 December, Bernays wrote to Gödel to ask for the proofs of the paper Gödel was writing on incompleteness. Courant and Issai Schur had told him that Gödel had obtained 'significant and surprising' results. Hilbert, unaware of what Gödel had done, gave his final published talk on the foundations of elementary number theory that month in Hamburg. After describing his formal system of number theory, Hilbert noted incorrectly that Ackermann and Von Neumann had proven the consistency thereof and added two open problems - versions of completeness from his Bologna talk (problems III and IV cited above). He could, he said, prove this in a special case by adding a new finitistic rule of inference: if, for every numeral n, the numerical formula φn can be checked to be a correct one, then conclude $\forall x \varphi x$. Hilbert noted quickly that $\forall x \varphi x$ says more than that φn is correct for all numerals n because it asserts the 'truth' of φt for all terms t. Nonetheless, he accepted the rule and proceeded to show how the consistency proof extended to include it and how completeness for assertions $\forall x \varphi x$ followed.

By the early part of 1931 Hilbert had learned of Gödel's Theorem. He was angry at first, but was soon trying to find a way around it. His solution was to extend the rule cited above. The simplest form of the ω-rule reads as follows:

> from $\varphi 0$, $\varphi 1$, ...
> infer $\forall x \varphi x$.

Under a variety of restrictions, the ω-rule allows the derivation of all true arithmetic sentences. None of them would be considered finitistic by anyone. Nonetheless, in his last full publication on foundations, Hilbert proved the Law of the Excluded Middle by deriving all true arithmetic sentences via the ω-rule. Gödel, who voiced caution in interpreting his incompleteness results and did not want to commit himself on the issue of whether or not he had destroyed Hilbert's programme, now complained to his friend Olga Taussky-Todd, 'How can he write such a paper after what I have done?'

Hilbert would have one last thing to say. In 1934 the first volume of *Grundlagen der Mathematik* was published. Ostensibly a joint work with Bernays, the latter wrote the text and Hilbert wrote a short preface, in which he said,

> ... the occasionally held opinion, that from the results of Gödel follows the non-executability of my Proof Theory, is shown to be erroneous. This result shows indeed only that for more advanced consistency proofs one must use the finite standpoint in a deeper way than is necessary for the consideration of elementary formalisms.

If we accept Hilbert's simple dichotomy of actual mathematics and finitary metamathematics, with the latter now 'deepened' (i.e., strengthened), then Hilbert's remarks lead either to nonsense (since the same refutation of the original programme obtains) or to the conclusion that the formal codification of actual mathematics inadequately represents finitary mathematics. For, Gödel had already proven his Second Incompleteness Theorem, according to which no sufficiently strong consistent formal theory could prove its own consistency. In particular, the formal system of actual, transfinite, mathematics cannot prove its own consistency. It therefore cannot contain the deepened finitary mathematics.

A second reading of Hilbert's remarks, a more sensible and more commonly accepted one, is this: There is a never-ending hierarchy - arithmetic, analysis, set theory, ... - of formal systems of actual mathematics, and a corresponding hierarchy of deepenings of metamathematics in which to prove the desired consistencies. As with the early Hilbert, the practitioners of this modified Hilbertian programme have given no thought to what the newly constructive consistency proofs are supposed to accomplish. If this new programme is philosophically suspect, resembling more a continuation through force of habit than a reasoned course of action, it has nevertheless been mathematically a successful development - a development, however, which extends beyond the scope of the present account. On this point, it may be well to remember that our rather long discussion of Hilbert's programmes has not been made for the sake of these programmes themselves, but to give the background to Gödel's Incompleteness Theorems.

I wish to finish with a few comments on the significance of Gödel's Incompleteness Theorems and the philosophical demise of Hilbert's programme. In this I offer merely a few quick remarks and not a carefully reasoned philosophical discussion.

The chief issue of the philosophy of mathematics is generally taken to be the question of the truth of mathematical laws. In what sense are they true? The traditional platonic answer that they are truths about some non-physical reality, although never refuted philosophically, is considered hopelessly naïve, especially in these materialistic times. Logicism attempted an unconvincing reduction of mathematics to logic, but never had any strong mathematical adherents and died a natural death. Intuitionism and formalism offered the most convincing explanations of this truth during the present century.

Barring a platonic belief in the actual existence of mathematical entities, intuitionism currently offers the only acceptable philosophy of mathematics under which mathematical truths are genuinely true. The honest intuitionist must find much of mathematics to be meaningless. He has the option of rejecting the meaningless material as simply mistaken (in much the way the modern physicist ignores questions about the ether, or modern scholars no longer attempt to read the book of nature for the messages God has written in it - as, e.g., when Physiologus tells us that the fact that lion cubs are born dead and only come to life three days after being born is one of the ways God chose to remind us of the death and resurrection of Jesus Christ), or of attempting to explain how classical, non-constructive mathematics has been so successful. Intuitionism is often rejected as a philosophy of mathematics for not adequately explaining all of mathematics, i.e. for not choosing this second option. Such a criticism of intuitionism as *a* philosophy of mathematics (as opposed to intuitionism as *the* philosophy of mathematics) is unfair, for, as Weyl pointed out, nothing compels the intuitionist to go beyond constructive mathematics.

Hilbert's Programme may be viewed as an attempt to offer an intuitionistic justification of classical mathematics. Gödel proved that Hilbert's approach, that of proving the outright consistency of classical mathematics, was too naïve. Since Gödel proved his Incompleteness Theorems, there have been successful partial attempts: Gödel himself proved finitistically that Hilbert's transfinite arithmetic is conservative over intuitionistic arithmetic with respect to finitary general propositions, and this result has been much improved. Thus, for some fragments of mathematics there is nothing to fear. Currently, there is much research that can be viewed as seeing how much classical mathematics can be rescued through such a modification of Hilbert's Programme.

The success of a more realistic version of Hilbert's Programme cannot, however, go all the way. Modern set theory far transcends the mathematics that is conservative over constructive mathematics with respect to, say, finitary general propositions. Thus there arises again the problem of choosing between demonstrably safe fragments - rejecting the higher infinite - and accounting for the latter.

Weyl's suggestion of mathematics as a theoretical science, with its meaningful finitary core and meaningless transfinite theoretical component, offers the only coherent philosophy of modern, classical, infinitary mathematics that I know of. On this account, as with Hilbert's programme, the transfinite mathematics is significant only as a means of deriving, say, finitary general propositions. Hilbert believed that every true such proposition had a finitary proof. Hence, the justification of transfinite mathematics was its mere consistency. By Gödel's First Incompleteness Theorem, no attempt to formalise mathematics can prove all finitary general propositions. Thus, when one views mathematics as a theoretical science, the justification of transfinite mathematics must become, like the justification of scientific theories in general, not the fact that it yields nothing new, but the fact that it does yield something new - and the ease with which it does so. Indeed, in 1946 Gödel explicitly called for an effort to use progressively more powerful transfinite theories to derive new arithmetical theorems.

In short, (if we ignore the unjustly maligned platonism) we are presented with two reasonable but antagonistic philosophies of mathematics, and Gödel's theorems tell us this mutual antagonism is necessary. In broader terms, we see here the familiar opposition of caution and daring, of the search for TRUTH and the search for truths. Gödel's theorems rule out the naïve hope that the two approaches will, upon closer examination, happen to coïncide.

As I said earlier, the fight between Brouwer and Hilbert is just an episode in the long running debate between the cautious mathematical conservative and the daring mathematical liberal. The debate continues today, but with less self-knowledge and only polemics coming from the conservatives, while the liberals have developed catastrophe theory, chaos, and large cardinals into coherent mathematical theories. To try to hide my bias, I quickly add that the Hilbert-Brouwer dispute shows that it is not always the overcautious that are the villains.

Reading List

We have a long reading list covering a number of topics: the main characters of our story, philosophy of mathematics, history of logic, and controversies in mathematics.

The *Dictionary of Scientific Biography* has articles on Hilbert, Brouwer, Weyl, and Von Neumann and can give the reader quick overviews of their achievements. The collected works of famous mathematicians are also often good sources.

1. David Hilbert, *Gesammelte Abhandlungen I-III,* Springer-Verlag, Berlin, 1935.
2. L.E.J. Brouwer, *Collected Works I-II,* North-Holland, Amsterdam, 1975-76.
3. Hermann Weyl, *Gesammelte Abhandlungen I-IV,* Springer-Verlag, Heidelberg, 1968.

Hilbert's collected works are incomplete and contain (in the third volume) only some of Hilbert's papers on foundations (most importantly: the First Hamburg Lecture (1921) and the Leipzig Lecture (1922)). The Heidelberg Lecture (1904), a shortened form of the Münster Lecture (1925), and the Second Hamburg Lecture (1927) were reprinted as appendices in the 7th edition of Hilbert's geometry book, and also appear in English translation in Van Heijenoort's collection.

4. David Hilbert, *Grundlagen der Geometrie,* Teubner, Leipzig, 1899; 7th edition 1930; English translation: *Foundations of Geometry,* Open Court, Chicago, 1903.
5. Jean van Heijenoort, *From Frege to Gödel; A Source Book in Mathematical Logic 1879-1931,* Harvard, Cambridge (Mass.), 1967.

Hilbert's correspondence with Frege is reprinted in Frege's correspondence:

6. Gottlob Frege, *Wissenschaftliche Briefwechsel,* ed. by Gottfried Gabriel et al, Felix Meiner Verlag, Hamburg, 1976; English translation: *Gottlob Frege, Philosophical and Mathematical Correspondence,* ed. by B. McGuiness, Univ. Chicago Press, Chicago, 1980.

Finally, the following book contains a reprinting of Hilbert's Bologna Lecture as well as a small recording of part of Hilbert's radio broadcast.

7. Kurt Reidemeister, ed., *Hilbert; Gedenkband*, Springer-Verlag, Heidelberg, 1971.

As a geometric aside, let me cite two references. The first discusses the gaps in Euclid's reasoning in quite some detail, and the second discusses the genesis of Hilbert's *Grundlagen der Geometrie.*

8. L. Bunt, P. Jones, and J. Bedient, *The Historical Roots of Elementary Mathematics,* Prentice-Hall, Englewood Cliffs (N.J.), 1976.
9. M.-M. Toepell, *Über die Entstehung von David Hilbert's 'Grundlagen der Geometrie',* Vandenhoeck and Ruprecht, Göttingen, 1986.

As for biography, there is a short biography of Hilbert written by Blumenthal in the third volume of Hilbert's collected works. There is also the book by Reid which, despite its shortcomings, remains the most complete treatment of Hilbert's life in print.

10. Constance Reid, *Hilbert*, Springer-Verlag, New York, 1970.

Those wishing to read Brouwer's papers on foundations will have an easier time of it. They are all collected in the first volume of his collected works. Three appear in English translation in the Van Heijenoort collection (reference 5, above). The second volume of Brouwer's collected works contains a short biography by the editors H. Freudenthal and A. Heyting of the two volumes and numerous historical notes (on topological matters) by Freudenthal. This volume also includes some correspondence between Brouwer and Hilbert. This biographical material is in English. To date, however, the fullest biography of Brouwer is only available in Dutch as an appendix to the following book.

11. L.E.J. Brouwer and C.S. Adema van Scheltema, *Droeve snaar, vriend van mij; Brieven,* ed. by D. van Dalen, Uitgeverij de Arbeiderspers, Amsterdam, 1984.

There are, however, some additional works on specific aspects of Brouwer's life available in English. The first of these is adequately described by its title; the second concerns Brouwer's philosophical activity; and the third handles the affair of the *Mathematische Annalen* which has been dubbed a 'frog-mouse war' by Albert Einstein, alluding to an old mock-homeric parody of the *Iliad.*

12. Walter P. van Stigt, 'The rejected parts of Brouwer's dissertation on the foundations of mathematics', *Historia Mathematica* 6 (1979), 385-404.
13. Walter P. van Stigt, 'L.E.J. Brouwer, the signific interlude', in: A.S. Troelstra and D. van Dalen, eds., *The L.E.J. Brouwer Centenary Symposium,* North-Holland, Amsterdam, 1982.
14. Dirk van Dalen, 'The war of the frogs and the mice, or the crisis of the *Mathematische Annalen*', to appear in *Mathematical Intelligencer.*

In German I might add the following.

15. P. Alexandroff, 'Die Topologie in und um Holland in den Jahren 1920-1930', *Nieuw Archief v. Wiskunde* (3) 17 (1969), 109-127.

Hermann Weyl's remarks on Hilbert's Second Hamburg Lecture appear in English translation, along with Bernays' expansion of Hilbert's remarks on Ackermann's consistency proof, in the Van Heijenoort volume cited above. I should add a couple of additional references on each of Weyl and Bernays. The first two are by Weyl. The first of these is his book on the philosophy of science and the interested reader might like to leaf through it. The second is a much slimmer volume that can be read in one sitting, and gives an adequate presentation of his views of the Grundlagenkrise on the eve of Gödel's discovery.

16. HERMANN WEYL, *Philosophy of Mathematics and Natural Science,* Princeton, 1949.

17. HERMANN WEYL, *The Open World,* Yale, New Haven (Conn.), 1932.

My discussion in the text slights Bernays, who didn't publish much on the subject until the battle was over. He wrote a not particularly critical account of Hilbert's programme that appeared in the third volume of Hilbert's collected works, and a number of papers on the philosophy of mathematics. These latter were collected in the following volume.

18. PAUL BERNAYS, *Abhandlungen zur Philosophie der Mathematik,* Wissenschaftliche Buchgesellschaft, Darmstadt, 1976.

I must also mention the monumental *Grundlagen der Mathematik.* The second volume gives much more on consistency proofs, the second edition giving correct non-finitistic consistency proofs for arithmetic.

19. DAVID HILBERT and PAUL BERNAYS, *Grundlagen der Mathematik I-II,* Springer-Verlag, Berlin, 1934&1939; 2nd edition: Springer-Verlag, Heidelberg, 1968 & 1970.

The references to Weyl's books and Bernays' collected philosophical papers bring us to philosophy. The proceedings of the philosophical meeting in Königsberg in September 1930 were published in 1931 in volume 2 of *Erkenntnis,* a philosophical journal. English translations of the talks of Carnap, Heyting, and Von Neumann appear in the next reference, while Dawson gives an English translation of the panel discussion from that meeting.

20. P. BENACERRAF and H. PUTNAM, eds., *Philosophy of Mathematics,* Prentice-Hall, Englewood Cliffs (N.J.), 1964; 2nd edition: Cambridge, 1983.

21. JOHN DAWSON, 'Discussion on the foundations of mathematics', *History and Philosophy of Logic* 5 (1984), 111-129.

Hilbert's programme has often been discussed by philosophers. The classic philosophical study is Kreisel's paper; the other two papers cited directly below are competent modern views and are strongly recommended. At a much lower level of sophistication, but mildly entertaining, is the published transcript in the second volume of *Historia Mathematica* (1975, pp. 503-533) of a discussion on the foundations of mathematics.

22. GEORG KREISEL, 'Hilbert's programme', *Dialectica* 12 (1958), 346-372; reprinted in: Benacerraf and Putnam, cited above.

23. VITO M. ABRUSCI, 'Proof', 'theory', and 'foundations' in Hilbert's mathematical work from 1885 to 1900', in: M.L. Dalla Chiara, ed., *Italian Studies in the Philosophy of Science,* Reidel, Dordrecht, 1981.

24. MARCUS GIAQUINTO, 'Hilbert's philosophy of mathematics', *British J. Phil. Science* 34 (1983), 119-132.

A general account on the development of logic during the period considered here that the reader might wish to consult is the following.

25. WARREN GOLDFARB, 'Logic in the twenties: the nature of the quantifier', *J. Symbolic Logic* 44 (1979), 351-368.

Before finishing with some references on mathematical controversy in general, let me cite a specifically mathematical reference. The following book gives the details of quantifier-free consistency proofs for quantifier-free systems and clarifies the hierarchical complexity of these proofs overlooked by Hilbert and his school.

26. H.E. ROSE, *Subrecursion; Functions and Hierarchies*, Oxford, 1984.

Controversy on the proper course of mathematics is not as uncommon as one might think. The first paper below discusses the British debate on algebra and the second covers the battle between the quaternionists and the vector analysts.

27. HELENA M. PYCIOR, 'Internalism, externalism, and beyond: 19th century British algebra', *Historia Mathematica* 11 (1984), 424-441.

28. MICHAEL CROWE, *A History of Vector Analysis*, Univ. of Notre Dame Press, Notre Dame, 1967.

The Hilbert-Brouwer dispute did not end the constructive vs. nonconstructive debate. Kronecker has been resurrected in the person of Errett Bishop. Two samples of Bishop's rather dry polemical style are the following.

29. ERRETT BISHOP, 'Review of H. Jerome Keisler, Elementary Calculus', *Bull. AMS* 83 (1977), 205-208.

30. ERRETT BISHOP, 'Schizophrenia in contemporary mathematics', in: Murray Rosenblatt, ed., *Errett Bishop: Reflection on Him and His Research,* AMS, Providence, 1985.

Bishop's review of Keisler overstepped the bounds of propriety. The best response I've seen was Ian Stewart's review of Douglas Bridges' reworking of Bishop's own book.

31. IAN STEWART, 'Frog and mouse revisited: a review of Errett Bishop & Douglas Bridges, *Constructive Analysis,* and A.E. Hurd & P.A. Loeb, *An Introduction to Nonstandard Real Analysis', Mathematical Intelligencer* 8 (4) (1986), 78-82.

Stewart's review led to a pleasant, if insipid, exchange between Stewart and Fred Richman, a disciple of Bishop, in the *Mathematical Intelligencer.* This journal is a good source for, among other things, current controversies in mathematics. (It has even published my own polemics.)