HILBERT'S PROGRAM THEN AND NOW

Richard Zach

1 INTRODUCTION

Hilbert's program is, in the first instance, a proposal and a research program in the philosophy and foundations of mathematics. It was formulated in the early 1920s by German mathematician David Hilbert (1862–1943), and was pursued by him and his collaborators at the University of Göttingen and elsewhere in the 1920s and 1930s. Briefly, Hilbert's proposal called for a new foundation of mathematics based on two pillars: the axiomatic method and finitary proof theory. Hilbert thought that by formalizing mathematics in axiomatic systems, and subsequently proving by finitary methods that these systems are consistent (i.e., do not prove contradictions), he could provide a philosophically satisfactory grounding of classical, infinitary mathematics (analysis and set theory). Had it been successful, Hilbert's program would perhaps not only have eclipsed in subsequent influence other foundational projects of the time, such as the logicist projects pursued by Frege and Russell and Brouwer's intuitionism, but it would also have achieved Hilbert's stated goal, viz., to "dispose of the foundational questions in mathematics as such once and for all" [Hilbert, 1929, 228]. Unfortunately, Gödel's theorems show that the program as originally envisaged by Hilbert cannot be carried out.

Although Hilbert's own project for the foundation of mathematics was ultimately unsuccessful, the project itself and technical advances made in its pursuit have had an enormous impact on logic and the foundations of mathematics more generally. In order to carry out the first part of the program, the axiomatization of mathematics in a formal system, Hilbert and his collaborators had pushed forward the development of logical formalisms in which such an axiomatization could be carried out. This led, in particular, to the first axiomatizations of propositional and first-order logic as independent systems (i.e., other than as fragments of more comprehensive higher-order systems, such as Frege's *Begriffsschrift* and the Whitehead-Russell system of *Principia mathematica*) and their metalogical investigation, such as the proof (by Paul Bernays in 1918) of the completeness of the propositional calculus and the work in Hilbert's school from 1921 onward on the decision problem. The investigation of the metalogical properties of logical systems led directly to some of the most important metalogical results in logic, viz., Gödel's completeness theorem and the negative solution by Church and Turing of the decision problem. The development of proof theory itself is an outgrowth of Hilbert's program. Gentzen's development of natural deduction and the sequent

calculus was carried out in the tradition of Hilbert's program and with the aim of constructing a logical system which facilitates consistency proofs. Gödel obtained his incompleteness theorems while trying to prove the consistency of analysis. And the tradition of reductive proof theory of the Gentzen-Schütte school and others is itself a direct continuation of Hilbert's program.

The present chapter is divided into three parts: The first part provides a sketch of the historical development of logic, proof theory, and philosophy of mathematics in the work of Hilbert and his followers through the 1930s. The second part deals with the philosophical interpretation and assessment of Hilbert's program. The third part presents recent work in proof theory which bears on the aims of Hilbert's program.

2   HILBERT'S PROGRAM THEN

2.1   Hilbert's early work on foundations

Hilbert's work on the foundations of mathematics can be traced to his work on geometry of the 1890s which resulted in his influential textbook *Foundations of Geometry* [1899]. One philosophical advance of this work was the development of Hilbert's conception of the axiomatic method. Hilbert believed that the proper way to develop any scientific subject rigorously required an axiomatic approach. In providing an axiomatic treatment, the theory would be developed independently of any need for intuition, and it would facilitate an analysis of the logical relationships between the basic concepts and the axioms. Of basic importance for an axiomatic treatment are, so Hilbert, investigation of the independence and, above all, of the consistency of the axioms. In his 1902 lectures on the foundations of geometry, he puts it thus:

> Every science takes its starting point from a sufficiently coherent body of facts as given. It takes form, however, only by *organizing* this body of facts. This organization takes place through the *axiomatic method*, i.e., one constructs a *logical structure of concepts* so that the relationships between the concepts correspond to relationships between the facts to be organized.
>
> There is arbitrariness in the construction of such a structure of concepts; we, however, *demand* of it:
>
> 1) completeness, 2) independence, 3) consistency. [Hilbert, 2004, 540]

From the time of his work on geometry forward, the last consideration, consistency, was of special importance in Hilbert's conception of the axiomatic method in general and the foundations of mathematics in particular. Hilbert was heavily influenced by the foundational views of late-19th century mathematicians, in particular, Cantor, Dedekind, and Kronecker. He shared with Dedekind and Cantor the view that mathematical activity should be free of constraints, which led to

his view, highlighted in his correspondence with Frege, that consistency of an axiomatic theory guarantees the existence of the structure described, and is in this sense sufficient to justify the use of the theory. And he shared with Kronecker a recognition that elementary arithmetic has a privileged role in mathematics, although he was of course opposed to the converse espoused by Kronecker, viz., that the natural numbers, and constructions based on elementary arithmetic, exhaust legitimate mathematics. These two influences in Hilbert's thought are at the root of his investigations of consistency.

Proofs of consistency for the axioms of geometry can be given by providing an interpretation of the system in the real plane, and thus the consistency of geometry is reduced to the consistency of analysis. Analysis, of course, itself requires justification. In [1900b], Hilbert approached the problem from the axiomatic standpoint by proposing an axiomatization of the real numbers. In order to show the consistency of this system, Hilbert expressly rejected the construction of a model, e.g., a construction based on Dedekind cuts of rationals, as an option. He considered the construction of the reals from the rationals and ultimately the natural numbers using the "genetic method" as insufficient: "Despite the high pedagogical and heuristic value of the genetic method, for the final presentation and the complete logical grounding of our knowledge the axiomatic method deserves the first rank" [Hilbert, 1900b, 1093]. Hilbert thus was after a *direct* consistency proof of analysis, i.e., one not based on reduction to another theory. He proposed the problem of finding such a proof as the second of his 23 mathematical problems in his address to the International Congress of Mathematicians in 1900 [1900a].

The discovery of Russell's paradox in 1902 made it clear that an axiomatic foundation of arithmetic and set theory requires a more precise development of the underlying logical systems. Hilbert knew of the paradoxes of set theory from Cantor and Zermelo, but it was apparently not until Russell's [1902] publication of the contradiction in Frege's system that Hilbert and Zermelo realized their importance. Hilbert's exchange with Frege on the axiomatic approach to geometry led him to realize that his conceptions of "axiom," "definition," "proof" were in need of clarification. In response, Hilbert formulated an early version of his proof-theoretical program in his 1904 Heidelberg talk [1905]. After criticizing the foundational views of Kronecker as dogmatic, and those of Frege and Dedekind as suffering from "unavoidable contradictions," he writes:

> Arithmetic is often considered to be a part of logic, and the traditional fundamental logical notions are usually presupposed when it is a question of establishing a foundation for arithmetic. If we observe attentively, however, we realize that in the traditional exposition of the laws of logic certain fundamental arithmetic notions are already used, for example, the notion of set and, to some extent, also that of number. Thus we find ourselves turning in a circle, and that is why a partly simultaneous development of the laws of logic and of arithmetic is required if paradoxes are to be avoided. [Hilbert, 1905, 131]

Hilbert's sketch of this "simultaneous development" of logic and arithmetic in the case of a very basic axiom system for the natural numbers is very close to the approach Hilbert's proof theoretic program would take 20 years later: Hilbert gives a direct argument that no contradiction can arise from the five axioms of his system.

This was a promising start, but several factors delayed the further development of Hilbert's proof theoretic program. One was Poincaré's [1906] criticism of what he saw as a viciously circular use of induction in Hilbert's sketched consistency proof (see [Steiner, 1975], Appendix). Moreover, Hilbert realized that axiomatic investigations required a well worked-out logical formalism in which axiomatic systems could be developed. At the time he used a logical formalism based on Schröder's algebraic logic, which was not particularly suited as a formalism for the axiomatization of mathematics. Following the 1905 lectures on foundations, Hilbert turned his immediate attention to work in other areas of mathematics and theoretical physics. He did, however, actively support others who worked on foundational questions in Göttingen, in particular Ernst Zermelo and Leonard Nelson.[1]

The publication of Whitehead and Russell's *Principia Mathematica* [1910; 1912; 1913] provided the required logical basis for a renewed attack on foundational issues. Beginning in 1914, Hilbert's student Heinrich Behmann and others studied the system of *Principia*.[2] Hilbert himself returned to work on the foundations of mathematics in 1917. In September 1917, he delivered an address to the Swiss Mathematical Society entitled "Axiomatic Thought" [1918a]. It is his first published contribution to mathematical foundations since 1905. In it, he again emphasized the requirement of consistency proofs for axiomatic systems: "The chief requirement of the theory of axioms must go farther [than merely avoiding known paradoxes], namely, to show that within every field of knowledge contradictions based on the underlying axiom-system are *absolutely impossible*." He posed the proof of the consistency of the integers (and of set theory) again as the main problems. In both these cases, there seems to be nothing more fundamental available to which the consistency could be reduced other than logic itself. Hilbert at the time considered the problem as essentially solved by Whitehead and Russell's work in *Principia*. Nevertheless, other fundamental problems of axiomatics remained unsolved, including the problem of the "decidability of every mathematical question," which also traces back to Hilbert's 1900 address.

---

[1]See [Sieg, 1999; Sieg, 2002], [Stein, 1988], [Hallett, 1990; Hallett, 1994], [Mancosu, 1998b], and [Avigad and Reck, 2001] for further discussion of the conceptual framework and historical background of Hilbert's thought, and [Resnik, 1974a] on the Frege-Hilbert correspondence. On Hilbert's foundational interests before 1917, and his engagement for Husserl, Zermelo, and Nelson in Göttingen, see [Peckhaus, 1990]. On general discussions of formalism and the place of Hilbert's thought in the mathematical context of the late 19th century, see [Webb, 1997] and [Detlefsen, 2005].

[2]See [Mancosu, 1999] and [2003] on Behmann's role in Hilbert's school and the influence of Russell.

These unresolved problems of axiomatics led Hilbert to devote significant effort to work on logic in the following years. In 1917, Paul Bernays joined him as his assistant in Göttingen. In a series of courses from 1917–1921, Hilbert, with the assistance of Bernays and Behmann, made significant new contributions to formal logic. The course from 1917 [Hilbert, 1918b], in particular, contains a sophisticated development of first-order logic, and forms the basis of Hilbert and Ackermann's textbook *Principles of Theoretical Logic* [1928]. In 1918, Bernays submitted a treatise on the propositional calculus of *Principia mathematica* as a *Habilitationsschrift*; it contains the first completeness proof of the propositional calculus for truth-functional semantics.[3]

The 1917–18 lectures were only the beginning of a strand of work on logic and metalogic in Hilbert's school, including work on the decision problem. The decision problem for first-order logic was tightly bound up with the aim of finding a completeness proof for the first-order predicate calculus (the "restricted calculus of functions" in Hilbert's terminology). This aim was stated in the 1917–18 lectures, but since completeness does not hold for first-order logic in any purely syntactic sense (an early result due to Ackermann), a development of the semantics of first-order logic was needed first. The decision problem, one of Hilbert's main aims for metamathematics in the 1920s, was already an issue in the lectures from 1905, and has its roots in Hilbert's belief, first explicitly stated in the Paris address, that "in mathematics, there is no ignorabimus," i.e., that every mathematical question can be solved either affirmatively or negatively. The questions of completeness and decidability thus became closely linked in the 1920s, with Behmann, Bernays, Schönfinkel, and later Ackermann working on special cases of the decision problem for first-order logic throughout the 1920s.

2.2 The consistency program, finitism, and proof theory

In about 1920, Hilbert came to reject Russell's logicist solution to the consistency problem for arithmetic, mainly for the reason that the axiom of reducibility cannot be accepted as a purely logical axiom. In lectures from the Summer term 1920, he concluded that "the aim of reducing set theory, and with it the usual methods of analysis, to logic, has not been achieved today and maybe cannot be achieved at all" [Hilbert, 1920]. At the same time, Brouwer's intuitionist mathematics gained currency. In particular, Hilbert's former student Hermann Weyl converted to intuitionism. Weyl's 1920 address to the Hamburg Mathematical Seminar, "The new foundational crisis in mathematics" [1921] was answered by Hilbert in three talks in Hamburg in the Summer of 1921 [1922b]. Here, Hilbert presented his own mature proposal for a solution to the problem of the foundation of mathematics. This proposal incorporated Hilbert's ideas from 1904 regarding direct consistency proofs, his conception of axiomatic systems, and also the technical developments in the axiomatization of mathematics in the work of Russell as well as the further

---

[3]See [Moore, 1997], [Sieg, 1999] and [Zach, 1999] for more detail on the development of logic in Hilbert's school around 1918.

developments carried out by him and his collaborators. What was new was the way in which Hilbert wanted to imbue his consistency project with the philosophical significance necessary to answer Brouwer and Weyl's criticisms: the finitary point of view.

According to Hilbert, there is a privileged part of mathematics, contentual elementary number theory, which relies only on a "purely intuitive basis of concrete signs." Whereas the operating with abstract concepts was considered "inadequate and uncertain," there is a realm of

> extra-logical discrete objects, which exist intuitively as immediate experience before all thought. If logical inference is to be certain, then these objects must be capable of being completely surveyed in all their parts, and their presentation, their difference, their succession (like the objects themselves) must exist for us immediately, intuitively, as something which cannot be reduced to something else.[4]

The objects in questions are *signs*, both numerals and the signs that make up formulas and formal proofs. The domain of contentual number theory consists in the finitary numerals, i.e., sequences of strokes. These have no meaning, i.e., they do not stand for abstract objects, but they can be operated on (e.g., concatenated) and compared. Knowledge of their properties and relations is intuitive and unmediated by logical inference. Contentual number theory developed this way is secure, according to Hilbert: no contradictions can arise simply because there is no logical structure in the propositions of contentual number theory.

The intuitive-contentual operations with signs form the basis of Hilbert's metamathematics. Just as contentual number theory operates with sequences of strokes, so metamathematics operates with sequences of symbols (formulas, proofs). Formulas and proofs can be syntactically manipulated, and the properties and relationships of formulas and proofs are similarly based in a logic-free intuitive capacity which guarantees certainty of knowledge about formulas and proofs arrived at by such syntactic operations. Mathematics itself, however, operates with abstract concepts, e.g., quantifiers, sets, functions, and uses logical inference based on principles such as mathematical induction or the principle of the excluded middle. These "concept-formations" and modes of reasoning had been criticized by Brouwer and others on grounds that they presuppose infinite totalities as given, or that they involve impredicative definitions (which were considered by the critics as viciously circular). Hilbert's aim was to justify their use. To this end, he pointed out that they can be formalized in axiomatic systems (such as that of *Principia* or those developed by Hilbert himself), and mathematical propositions and proofs thus turn into formulas and derivations from axioms according to strictly circumscribed rules of derivation. Mathematics, so Hilbert, "becomes an inventory of provable formulas." In this way the proofs of mathematics are subject to metamathematical, contentual investigation. The goal of Hilbert's program is then to

---

[4][Hilbert, 1922b, 202]. The passage is repeated almost verbatim in [Hilbert, 1926, 376], [Hilbert, 1928, 464], and [Hilbert, 1931b, 267]

give a contentual, metamathematical proof that there can be no derivation of a contradiction, i.e., no formal derivation of a formula A and of its negation $\neg A$.

This sketch of the aims of the program was fleshed out by Hilbert and his collaborators in the following 10 years. On the conceptual side, the finite standpoint and the strategy for a consistency proof were elaborated by Hilbert [1923; 1926; 1928] and Bernays [1922; 1928b; 1930], of which Hilbert's article "On the infinite" [1926] provides the most detailed discussion of the finitary standpoint. In addition to Hilbert and Bernays, a number of other people were involved in technical work on the program. The ε-operator was first introduced in the Hamburg lecture of 1921 [Hilbert, 1922b], and developed in a number of lectures given in Göttingen [Hilbert, 1922a; Hilbert and Bernays, 1923], as well as in [Hilbert, 1923]. Hilbert and Bernays developed the ε-calculus as their definitive formalism for axiom systems for arithmetic and analysis, and the so-called ε-substitution method as the preferred approach to giving consistency proofs.

Briefly, the ε-calculus is a formalism that includes ε as a term-forming operator. If $A(x)$ is a formula, then $\varepsilon_x A(x)$ is a term, which intuitively stands for a witness for $A(x)$. In a logical formalism containing the ε-operator, the quantifiers can be defined by: $\exists x\, A(x) \equiv A(\varepsilon_x A(x))$ and $\forall x\, A(x) \equiv A(\varepsilon_x \neg A(x))$. The only additional axiom necessary is the so-called "transfinite axiom," $A(t) \to A(\varepsilon_x A(x))$. Based on this idea, Hilbert and his collaborators developed axiomatizations of number theory and analysis. Consistency proofs for these systems were then given using the ε-substitution method. The idea of this method is, roughly, that the ε-terms $\varepsilon_x A(x)$ occurring in a formal proof are replaced by actual numerals, resulting in a quantifier-free proof. The simplest case, outlined in Hilbert's papers, is as follows. Suppose we had a (suitably normalized) derivation of $0 = 1$ that contains only one ε-term $\varepsilon_x A(x)$. Replace all occurrences of $\varepsilon_x A(x)$ by 0. The instances of the transfinite axiom then are all of the form $A(t) \to A(0)$. Since no other ε-terms occur in the proof, $A(t)$ and $A(0)$ are basic numerical formulas without quantifiers and, we may assume, also without free variables. So they can be evaluated by finitary calculation. If all such instances turn out to be true numerical formulas, we are done. If not, this must be because $A(t)$ is true for some t, and $A(0)$ is false. Then replace $\varepsilon_x A(x)$ instead by n, where n is the numerical value of the term t. The resulting proof is then seen to be a derivation of $0 = 1$ from true, purely numerical formulas using only modus ponens, and this is impossible. Indeed, the procedure works with only slight modifications even in the presence of the induction axiom, which in the ε-calculus takes the form of a least number principle: $A(t) \to \varepsilon_x A(x) \leq t$, which intuitively requires $\varepsilon_x A(x)$ to be the *least* witness for $A(x)$.

The ε-substitution method is simple enough for the basic cases considered by Hilbert, but becomes extremely complex when ε-operators are nested. In his 1924 dissertation, [Ackermann, 1925] presented an (erroneous) consistency proof based on the ε-substitution method for a version of analysis. John von Neumann, then visiting Göttingen, gave a corrected consistency proof for a system of the ε-formalism (which, however, did not include the induction axiom) in 1925 [1927].

Building on von Neumann's work, Ackermann devised a new ε-substitution proce-
dure which he communicated to Bernays (see [Bernays, 1928b]). Ackermann and
Bernays considered the proof to be correct for the entire first-order fragment of
arithmetic and were confident that it could be extended to a consistency proof
of analysis. In his address "Problems of the grounding of mathematics" to the
International Congress of Mathematicians in Bologna in 1928 [1929], Hilbert opti-
mistically claimed that the work of Ackermann and von Neumann had established
the consistency of number theory and that the proof for analysis had already been
carried out by Ackermann "to the extent that the only remaining task consists in
the proof of an elementary finiteness theorem that is purely arithmetical."[5]

2.3   The impact of Gödel's incompleteness theorems

Gödel's incompleteness theorems [Gödel, 1931] showed that Hilbert's optimism
was unfounded. In September 1930, Kurt Gödel announced his first incomplete-
ness theorem at a conference in Königsberg. Von Neumann, who was in the
audience, immediately recognized the significance of Gödel's result for Hilbert's
program. Shortly after the conference he wrote to Gödel, telling him that he
had found a corollary to Gödel's result. Gödel had found the same result already
independently: the second incompleteness theorem, asserting that the system of
Principia does not prove the formalization of the claim that the system of *Prin-
cipia* is consistent (provided it is). All the methods of finitary reasoning used in
the consistency proofs up till then were believed to be formalizable in *Principia*,
however. Hence, if the consistency of *Principia* were provable by the methods
used in Ackermann's proofs, it should be possible to formalize this proof in *Prin-
cipia*; but this is what the second incompleteness theorem states is impossible.
Bernays also immediately realized the importance of Gödel's results after he stud-
ied Gödel's paper in January 1931. He wrote to Gödel that (under the assumption
that finitary reasoning can be formalized in *Principia*) the incompleteness theo-
rems seem to show that a finitary consistency proof of *Principia* is impossible.
Shortly thereafter, von Neumann showed that Ackermann's consistency proof is
flawed and provided a counterexample to the proposed ε-substitution procedure.[6]

   Although the impact of Gödel's incompleteness theorems for Hilbert's pro-
gram was recognized soon after their publication, Hilbert's program was by no
means abandoned. Hilbert himself was no longer actively involved in founda-
tional research, but Bernays continued to work on the foundations of mathemat-
ics. The two-volume *Grundlagen der Mathematik* [Hilbert and Bernays, 1934;
Hilbert and Bernays, 1939] was prepared by Bernays alone, and included new re-

---

[5]See [Avigad and Zach, 2002] for a general introduction to the ε-calculus and [Zach, 2003b]
and [2004] on the history of the ε-calculus and the substitution method. [Sieg, 1999] presents a
detailed and perceptive analysis of the development of the program and its influence as a whole.
   [6]The correspondence between Gödel and Bernays is published in [Gödel, 2003a, 41–313] and
that with von Neumann in [Gödel, 2003b, 327–377]. See also the informative introductions by
Feferman and Sieg, respectively, to these sections of the correspondence, as well as [Mancosu,
2004] and the last section of [Zach, 2003b].

sults by Ackermann and Bernays on the ε-calculus. It also included Herbrand's [1930] work on proof theory, and a sketch of Gentzen's [1936] consistency proof of first-order arithmetic. Bernays's and Gentzen's work, in particular, focused on possible extensions of the finitary standpoint which could yield consistency proofs for substantial parts of mathematics in spite of Gödel's theorems.

Gentzen's first consistency proof for number theory, the so-called galley proof [1935], was the result of a combination of Gentzen's [1934] earlier work on the logical formalism of the sequent calculus, which provided a proof-theoretically more convenient axiomatization of arithmetic, and a new strategy of proving consistency. This strategy involved defining certain "reduction steps" on proofs: local transformations of parts of a derivation in the new formalism. The consistency proof then proceeds by showing that by iterating these reductions on a proof one eventually arrives at a proof of a special form (a proof free of the cut rule and the induction rule), and no proof of such a form can be the proof of a contradiction. The first version of the proof relied on the notion of a reduction rule, which itself cannot be formalized in arithmetic.[7] This notion met with some objections, and in the revised, published version [1936], Gentzen replaced the appeal to reduction rules by a proof that the iteration of reduction steps itself terminates. He did this by assigning a measure for the complexity of a given derivation, and showed that the result of the application of a reduction step to a proof reduces the complexity measure of that proof. The complexity measure Gentzen used was certain finite strings of numbers which may be interpreted as naming countable ordinals less than ε_0.[8] The consistency result then follows if one accepts that there are no infinite descending sequences of such ordinal notations, or, more precisely, by using transfinite induction up to ε_0. This principle, by Gödel's incompleteness theorem, cannot itself be formalized in first-order arithmetic [Gentzen, 1943]. Gentzen's proof allowed Ackermann [1940] to give a correct consistency proof based on the ε-substitution method for first-order arithmetic, also based on transfinite induction up to ε_0.

Gentzen's work marks the beginning of post-Gödelian proof theory. In the proof-theoretic tradition of Gentzen, axiomatic theories are analyzed according to which transfinite induction principles are required to prove the consistency of the theory. However, Gentzen's contribution and influence goes beyond this: He emphasized that proof-theoretic methods do not just allow us to prove the consistency of a theory, but that they also enable us to extract information from proofs beyond the fact that the formula proved follows from the axioms.

3 PHILOSOPHICAL INTERPRETATION OF HILBERT'S PROGRAM

The philosophical importance and influence of Hilbert's work on foundations is twofold. First, the epistemological standpoint of Hilbert's finitism is of inter-

---

[7]On the galley proof, see [Bernays, 1970], [Kreisel, 1971, Appendix II], and [Negri, 1980].

[8]If $\omega_0 = \omega$, and $\omega_{n+1} = \omega^{\omega_n}$, then the ordinal ε_0 is the limit of ω_n for $n = 1, 2, \ldots$. In other words, ε_0 is the least fixed point of $\alpha = \omega^\alpha$.

est in the philosophy of mathematics quite independently of the success of the proof-theoretic program which it underlies. Like other important proposals in the philosophy of mathematics such as intuitionism, predicativism, and logicism, Hilbert's finitism is, *inter alia*, a philosophical view about the nature of mathematical knowledge and delineates a particular set of (finitarily) meaningful propositions, as well as finitarily admissible constructions and methods of proof. Debate about the status of finitary evidence and proof are still very much alive today. Second, Hilbert's program can and has been seen as a version of reductive instrumentalism in mathematics. That is to say, one can read Hilbert as proposing that only a certain part of mathematics (propositions, proofs) is meaningful, viz., the finitary part. The rest, which includes classical infinitary mathematics (full first-order arithmetic, analysis, and set theory, in particular) are mere instruments. Hilbert's program has thus been an important inspiration for related instrumentalist proposals in the philosophy of mathematics (e.g., [Field, 1980; Detlefsen, 1986]).

This section will expand on these two themes. In the case of both the debate on the philosophy of finitism, and on the view of Hilbert's program as an instrumentalist philosophy of mathematics, questions of historical interpretation interact with conceptual analysis. The distinction between these aspects should be kept in mind.

3.1 The finitary point of view

The cornerstone of Hilbert's philosophy of mathematics, and the substantially new aspect of his foundational thought from [1922b] onward, was the so-called finitary standpoint. This methodological standpoint consists in a restriction of mathematical thought to objects which are "intuitively present as immediate experience prior to all thought," and to those operations on and methods of reasoning about such objects which do not require the introduction of abstract concepts, in particular, require no appeal to completed infinite totalities.

Hilbert characterized the domain of finitary reasoning in a well-known paragraph which appears in roughly the same formulation in all of Hilbert's more philosophical papers from the 1920s [1922b; 1926; 1928; 1931b]:

> [A]s a condition for the use of logical inferences and the performance of logical operations, something must already be given to our faculty of representation, certain extra-logical concrete objects that are intuitively present as immediate experience prior to all thought. If logical inference is to be reliable, it must be possible to survey these objects completely in all their parts, and the fact that they occur, that they differ from one another, and that they follow each other, or are concatenated, is immediately given intuitively, together with the objects, as something that can neither be reduced to anything else nor requires reduction. This is the basic philosophical position that I consider req-

uisite for mathematics and, in general, for all scientific thinking, understanding, and communication. [Hilbert, 1926, 376]

These objects are, for Hilbert, the *signs*. For the domain of contentual number theory, the signs in question are sequences of strokes ("numerals") such as

$$|, ||, |||, ||||| \ .$$

The question of how exactly Hilbert understood the numerals is difficult to answer. What is clear in any case is that they are logically primitive, i.e., they are neither concepts (as Frege's numbers are) nor sets. For Hilbert, the important issue is not primarily their metaphysical status (abstract versus concrete in the current sense of these terms), but that they do not enter into logical relations, e.g., they cannot be predicated of anything. In Bernays's most mature presentations of finitism [Bernays, 1930; Hilbert and Bernays, 1939], the objects of finitism are characterized as *formal objects* which are recursively generated by a process of repetition; the stroke symbols are then concrete representations of these formal objects (see [Sieg, 2002]).

Sometimes Hilbert's view is presented as if Hilbert claimed that the numbers are signs on paper. It is important to stress that this is a misrepresentation, that the numerals are not physical objects in the sense that truths of elementary number theory are dependent only on external physical facts or even physical possibilities (e.g., on what sorts of stroke symbols it is possible to write down). Hilbert [1926] made too much of the fact that for all we know, neither the infinitely small nor the infinitely large are actualized in physical space and time, yet he certainly held that the number of strokes in a numeral is unbounded. It is also essential to the conception that the numerals are sequences of one kind of sign, and that they are somehow dependent on being grasped as such a sequence, that they do not exist independently of our intuition of them. Only our seeing or using "||||" as a sequence of 4 strokes as opposed to a sequence of 2 symbols of the form "||" makes "||||" into the numeral that it is. This raises the question of individuation of stroke symbols. An alternative account would have numerals be mental constructions. However, Bernays denied also this, writing that "the objects of intuitive number theory, the number signs, are, according to Hilbert, also not 'created by thought'. But this does not mean that they exist independently of their *intuitive construction*, to use the Kantian term that is quite appropriate here" [Bernays, 1923, 226]. [Kitcher, 1976] proposes the view that, whatever the numerals are, the strokes on paper or the stroke sequences contemplated by the mind merely *represent* the numerals. According to Hilbert and Bernays, the numerals are given in our representation, but they are not merely subjective "mental cartoons" (Kitcher's term).

> If we want [...] the ordinal numbers as definite objects free of all inessential elements, then in each case we have to take the mere schema of the relevant figure of repetition [*Wiederholungsfigur*] as an object; this requires a very high abstraction. We are free, however, to represent these purely formal objects by concrete objects ("number signs"); these

contain then inessential, arbitrarily added properties, which, however,
are also easily grasped as such. [Bernays, 1930, 244]

One version of this view would be to hold that the numerals are *types* of stroke-symbols as represented in intuition. This is the interpretation that Tait [1981] gives. At first glance, this seems to be a viable reading of Hilbert. It takes care of the difficulties that the reading of numerals-as-tokens (both physical and mental) faces, and it gives an account of how numerals can be dependent on their intuitive construction while at the same time not being created by thought. The reasoning that leads Tait to put forward his reading lies in several constraints that Hilbert and Bernays put on the numerals. For instance, [Bernays, 1923, 159] writes that "figures [i.e., numerals] *are* not shapes, they *have* a shape." Thus it is the shape of the numerals, and not the numerals themselves, which is supposed to be independent of place and time, independent of the circumstances of production, independent of inessential differences in execution, and capable of secure recognition in all circumstances [Hilbert, 1922b, 163]. Tait infers from this that identity between numerals is type identity, and hence, that numerals should be construed as types of stroke symbols.

Types are ordinarily considered to be abstract objects and not located in space or time. Taking the numerals as intuitive representations of sign types might commit us to taking these abstract objects as existing independently of their intuitive representation. That numerals are "space- and timeless" is a consequence that already [Müller, 1923, 158] thought could be drawn from Hilbert's statements, and that was in turn denied by Bernays. The reason is that a view on which numerals are space- and timeless objects existing independently of us would be committed to them existing simultaneously as a completed totality, and this is exactly what Hilbert is objecting to.

> It is by no means compatible, however, with Hilbert's basic thoughts to introduce the numbers as ideal objects "with quite different determinations from those of sensible objects," "which exist entirely independent of us." By this we would go beyond the domain of the immediately certain. In particular, this would be evident in the fact that we would consequently have to assume the numbers *as all existing simultaneously.* But this would mean to assume at the outset that which Hilbert considers to be problematic. [Bernays, 1923, 225–26]

This is not to say that it is *incoherent* to consider the numbers as being abstract objects, only that the finitary viewpoint prohibits such a view. Bernays goes on to say:

> Hilbert's theory does not exclude the possibility of a philosophical attitude which conceives of the numbers [but not the finitist's numerals] as existing, non-sensible objects (and thus the same kind of ideal existence would then have to be attributed to transfinite numbers as well, and in particular to the numbers of the so-called second number class).

> Nevertheless the aim of Hilbert's theory is to make such an attitude
> dispensable for the foundation of the exact sciences. [Bernays, 1923,
> 226]

Another open question in this regard is exactly what Hilbert meant by "concrete."
He very likely did not use the term in the same sense as it is used today, i.e., as
characteristic of spatio-temporal physical objects in contrast to "abstract" objects.
However, sign types certainly are different from full-fledged abstracta like pure sets
in that all their tokens are concrete. Parsons takes account of this difference by
using the term "quasi-concrete" for such abstracta. Tait, on the other hand, thinks
that even the tokens are not concrete physical objects, but abstract themselves.

Now what is the epistemological status of the finitary objects? In order to carry
out the task of providing a secure foundation for infinitary mathematics, access
to finitary objects must be immediate and certain. Hilbert's philosophical back-
ground was broadly Kantian, as was Bernays's, who was closely affiliated with the
neo-Kantian school of philosophy around Leonard Nelson in Göttingen. Hilbert's
characterization of finitism often refers to Kantian intuition, and the objects of
finitism as objects given intuitively. Indeed, in Kant's epistemology, immediacy is
a defining characteristic of intuitive knowledge. The question is, what kind of in-
tuition is at play? Mancosu [1998b] identifies a shift in this regard. He argues that
whereas the intuition involved in Hilbert's early papers was a kind of perceptual
intuition, in later writings (e.g., [Bernays, 1928a]) it is identified as a form of pure
intuition in the Kantian sense. Hilbert [1931b, 266–267] later sees the finite mode
of thought as a separate source of *a priori* knowledge in addition to pure intuition
(e.g., of space) and reason, claiming that he has "recognized and characterized the
third source of knowledge that accompanies experience and logic." Both Bernays
and Hilbert justify finitary knowledge in broadly Kantian terms (without how-
ever going so far as to provide a transcendental deduction), characterizing finitary
reasoning as the kind of reasoning that underlies all mathematical, and indeed,
scientific, thinking, and without which such thought would be impossible.[9]

The simplest finitary propositions are those about equality and inequality of
numerals. The finite standpoint moreover allows operations on finitary objects.
Here the most basic is that of concatenation. The concatenation of the numerals
|| and ||| is communicated as "$2 + 3$," and the statement that || concatenated
with ||| results in the same numeral as ||| concatenated with || by "$2 + 3 = 3 +
2$." In actual proof-theoretic practice, as well as explicitly in [Bernays, 1930;
Hilbert and Bernays, 1934], these basic operations are generalized to operations
defined by recursion, paradigmatically, primitive recursion, e.g., multiplication and
exponentiation. Roughly, a primitive recursive definition of a numerical operation
is one in which the function to be defined, f, is given by two equations

$$f(0, \mathfrak{m}) = g(\mathfrak{m})$$
$$f(\mathfrak{n}', \mathfrak{m}) = h(\mathfrak{n}, \mathfrak{m}, f(\mathfrak{n}, \mathfrak{m})),$$

---

[9]See [Kitcher, 1976] and [Parsons, 1998] for more discussion on the metaphysics and episte-
mology of finitism, and [Sieg, 1999] for historical remarks on the development of finitism.

where g and h are functions already defined, and $\mathfrak{n}'$ is the successor numeral to $\mathfrak{n}$. Fraktur letters are used here, as in Hilbert's writings, as meta-variables for numerals. For instance, if we accept the functions $g(\mathfrak{m}) = \mathfrak{m}$ (the constant function) and $h(\mathfrak{n}, \mathfrak{m}, \mathfrak{k}) = \mathfrak{m} + \mathfrak{k}$ as finitary, then the equations above define a finitary function, in this case, multiplication $f(\mathfrak{n}, \mathfrak{m}) = \mathfrak{n} \times \mathfrak{m}$.

Similarly, finitary judgments may involve not just equality or inequality but also basic decidable properties, such as "is a prime." This is finitarily acceptable as long as the characteristic function of such a property is itself finitary: For instance, the operation which transforms a numeral to | if it is prime and to || otherwise can be defined by primitive recursion and is hence finitary. Such finitary propositions may be combined by the usual logical operations of conjunction, disjunction, negation, but also bounded quantification. Hilbert [1926] gives the example of the proposition that "there is a prime number between $\mathfrak{p} + 1$ and $\mathfrak{p}! + 1$" where $\mathfrak{p}$ is a certain large prime. This statement is finitarily acceptable since it "serves merely to abbreviate the proposition" that either $\mathfrak{p} + 1$ or $\mathfrak{p} + 2$ or $\mathfrak{p} + 3$ or ... or $\mathfrak{p}! + 1$ is a prime.

The problematic finitary propositions are those that express general facts about numerals such as that $1 + \mathfrak{n} = \mathfrak{n} + 1$ for any given numeral $\mathfrak{n}$. It is problematic because, as Hilbert puts it, it "is from the finitist point of view *incapable of being negated*" [Hilbert, 1926, 378]. By this he means that the contradictory proposition that there is a numeral $\mathfrak{n}$ for which $1 + \mathfrak{n} \neq \mathfrak{n} + 1$ is not finitarily meaningful. "One cannot, after all, try out all numbers" [Hilbert, 1928, 470]. For the same reason, a finitary general proposition is not to be understood as an infinite conjunction but "only as a hypothetical judgment that comes to assert something when a numeral is given" [Hilbert, 1926, 378]. Even though they are problematic in this sense, general finitary statements are of particular importance to Hilbert's proof theory, since the statement of consistency of a formal system T is of such a general form: for any given sequence p of formulas, p is not a derivation of a contradiction in T.

Even though in general existential statements are not finitarily meaningful, they may be *given* finitary meaning if the witness is given by a finitary function. For instance, the finitary content of Euclid's theorem that for every prime $\mathfrak{p}$ there is a prime $> \mathfrak{p}$, is that given a specific prime $\mathfrak{p}$ one can produce, by a finitary operation, another prime $> \mathfrak{p}$ (viz., by testing all numbers between $\mathfrak{p}$ and $\mathfrak{p}! + 1$. This view is discussed by Bernays [1930] and plays an important role in the uses Gentzen [1936] and others make of proof theory.

3.2   Analyses of finitism

Hilbert's substantial philosophical claims about the finitary standpoint are difficult to flesh out. For instance, Hilbert and Bernays both appeal to the role of Kantian intuition for our apprehension of finitary objects (they are given in the faculty of representation). Supposing one accepts this line of epistemic justification in principle, it is plausible that the simplest examples of finitary objects and propositions, and perhaps even simple cases of finitary operations such as concate-

nations of numerals can be given a satisfactory account. However, it is unclear exactly which more complex objects, propositions, and operations should be admitted as finitary, and how the account can be extended to cover them. This has led to substantial debate since the 1920s about the nature of finitary reasoning and its justification.

Of crucial importance to both an understanding of finitism and of Hilbert's proof theory is the question of what operations and what principles of proof should be allowed from the finitist standpoint. That a general answer is necessary is clear from the demands of Hilbert's proof theory, i.e., it is not to be expected that given a formal system of mathematics (or even a single sequence of formulas) one can "see" that it is consistent (or that it cannot be a genuine derivation of an inconsistency) the way we can see, e.g., that $|| + ||| = ||| + ||$. What is required for a consistency proof is an operation which, given a formal derivation, transforms such a derivation into one of a special form, plus proofs that the operation in fact succeeds in every case and that proofs of the special kind cannot be proofs of an inconsistency. To count as a finitary consistency proof, the operation itself must be acceptable from the finitist standpoint, and the proofs required must use only finitarily acceptable principles.

Hilbert never gave a general account of which operations and methods of proof are acceptable from the finitist standpoint, but only examples of operations and methods of inference in contentual finitary number theory which he accepted as finitary. Contentual induction was accepted in its application to finitary statements of the hypothetical, general kind explicitly in [1922b]. Hilbert [1923, 1139] said that intuitive thought "includes recursion and intuitive induction for finite existing totalities," and used exponentiation in an example in 1928. Bernays [1930] explained how exponentiation may be understood as a finitary operation on numerals. Hilbert and Bernays [1934] give the only general account of finitary contentual number theory; according to it, operations defined by primitive recursion and proofs using induction are finitarily acceptable. All of these methods, in their application in the domain of numbers, can be formalized in a system known as primitive recursive arithmetic (PRA), which allows definitions of functions by primitive recursion and induction on quantifier-free formulas. However, neither Hilbert nor Bernays ever claimed that *only* primitive recursive operations count as finitary, and non-primitive recursive methods were used in ostensibly finitary consistency proofs already in 1923 (see [Tait, 2002] and [Zach, 2003b]). These include, in particular, the consistency proof of a formal system of number theory corresponding to primitive recursive arithmetic in [Hilbert and Bernays, 1923], and a stronger system in Ackermann's dissertation [Ackermann, 1925].[10]

Although Hilbert and his collaborators used methods which go beyond the primitive recursive and accepted them as finitary, it is still unclear whether they (a) realized that these methods were not primitive recursive and (b) whether they

---

[10]Ackermann in fact used transfinite induction up to ω^{ω^ω} to justify a non-primitive recursive reduction procedure.

would still have accepted them as finitary if they had.[11] The conceptual issue is which operations *should* be considered as finitary. Since Hilbert was less than completely clear on what the finitary standpoint consists in, there is some leeway in setting up the constraints, epistemological and otherwise, an analysis of finitist operation and proof must fulfill. Hilbert characterized the objects of finitary number theory as "intuitively given," as "surveyable in all their parts," and said that their having basic properties must "exist intuitively" for us. Bernays [1922, 216] suggests that in finitary mathematics, only "primitive intuitive cognitions come into play," and uses the term "point of view of intuitive evidence" in connection with finitism [Bernays, 1930, 250]. This characterization of finitism as primarily to do with *intuition* and intuitive knowledge has been emphasized in particular by Parsons [1998] who argues that what can count as finitary on this understanding is not more than those arithmetical operations that can be defined from addition and multiplication using bounded recursion. In particular, according to Parsons, exponentiation and general primitive recursion are not finitarily acceptable.

The thesis that finitism coincides with primitive recursive reasoning has received a forceful and widely accepted defense by Tait [1981]. Tait, in contrast to Parsons, rejects the aspect of representability in intuition as the hallmark of the finitary; instead he takes finitary reasoning to be "a minimal kind of reasoning supposed by all nontrivial mathematical reasoning about numbers" and analyzes finitary operations and methods of proof as those that are implicit in the very notion of number as the form of a finite sequence. This analysis of finitism is supported by Hilbert's contention that finitary reasoning is a precondition for logical and mathematical, indeed, any scientific thinking [Hilbert, 1931b, 267]. The crucial difference between Tait's conception of finitism and Parsons (as well as Hilbert's own) is that according to Tait there is no ultimate epistemological foundation for finitism which yields the security of finitary reasoning for which Hilbert appealed to intuition. Tait argues that

> . . . no absolute conception of security is realized by finitism or any other kind of mathematical reasoning. Rather, the special role of finitism consists in the circumstance that it is a minimal kind of reasoning presupposed by all nontrivial mathematical reasoning about numbers. And for this reason it is *indubitable* in a Cartesian sense that there is no preferred or even equally preferable ground on which to stand and criticize it. Thus finitism is fundamental to number-theoretical mathematics even if it is not a foundation in the sense Hilbert wished. [Tait, 1981, 525]

Another interesting analysis of finitary proof, which, however, does not provide as detailed a philosophical justification, was proposed by Kreisel [1960]. It yields the result that exactly those functions are finitary which can be proved to be well-defined in first-order arithmetic *PA*.[12] Kreisel's proposal makes use of the notions

---

[11]See Tait's discussion in the Appendix to Chapters 1 and 2 in [Tait, 2005b].

[12]Kreisel [1970, Section 3.5] provides another analysis by focusing on what is "visualizable."

of formalizations of provability predicates and ordinal progressions of theories. Kreisel argues that if $Pr(\ulcorner A \urcorner)$ has been recognized to be a provability predicate for a partial formalization Σ_μ of finitary reasoning, and $Pr(\ulcorner A(0^{(x)}) \urcorner)^{13}$ is provable in Σ_μ (and hence established by finitary means), then the finitist is entitled to also accept $A(x)$ as finitarily established. If that is the case, we may add $A(x)$ as an axiom to Σ_μ to obtain a new theory Σ_ν which is also finitarily justified. On Kreisel's account, finitary provability coincides with the provability in some Σ_ν so obtained, starting from $\Sigma_0 = PRA$. If some Σ_ν proves $\exists y\, A(x, y)$, for $A(x, y)$ a primitive recursive equation, then $f(x) =$ the least y such that $A(x, y)$, is finitarily justified. Kreisel sketches a proof of the theorem that the functions so justified are exactly those which are provably total in PA, and hence there are finitary functions which are not primitive recursive.

Tait [1981, §13] also contains a discussion of Kreisel's analysis. In order to obtain a non-primitive recursive function on Kreisel's account, we must properly extend Σ_0 since the provably total functions of $\Sigma_0 = PRA$ are just the primitive recursive functions. So suppose we have that PRA proves the arithmetization of the claim that $\exists y\, A(0^{(x)}, y)$ is provable.[14] This, according to Kreisel, justifies the acceptance of $f(x)$ as defined above as finitary, because a finitary proof of the general fact that it is provable in PRA that $f(x)$ is defined, together with the acceptance of PRA as finitarily acceptable, amounts to a finitary proof that $f(x)$ is defined for all x. The weak point in this justification, according to Tait, is this:

> For the finitist to recognize the validity of primitive recursive arithmetic, he must recognize the *general* validity of definition of functions by primitive recursion. But he cannot even formulate this since it involves the notion of function.

Tait's point here is that there is a significant difference between accepting each primitive recursive definition individually as finitary, and accepting primitive recursion in general as a finitarily valid principle. The finitist is able to do the former, but not the latter. For to accept primitive recursion in general as a finitarily valid principle of definition, one would either, as Tait puts it, need to invoke the notion of a function (which is not a finitary object), or one needs a justification for why, say, all primitive recursive *terms* are calculable for every argument — and for this a finitary evaluation procedure for primitive recursive terms is necessary. Such an evaluation procedure, however, cannot be primitive recursive. And prior to the extension of PRA to include the new non-primitive recursive function $f(x)$ there is no reason to suppose that such an evaluation procedure exists. Although Tait's objection is directed at Kreisel's analysis of finitary function, it of course also raises doubts about Kreisel's account of finitary proof.

The result is the same: finitary functions turn out to be just those provably total in PA.

[13]Here, x is a free variable, and $\ulcorner A(0^{(x)}) \urcorner$ is the function of x which computes $\ulcorner A(0^{\prime \cdots \prime}) \urcorner$ with x occurrences of $\prime$.

[14]In other words, there are primitive recursive functions $h(x)$ and $g(x)$ so that PRA proves that $g(x)$ codes a derivation in PRA of the formula $\ulcorner A(0^{(x)}, t) \urcorner$, where t is the primitive recursive term (containing only the free variable x) which is coded by $h(x)$.

3.3   Hilbert's program and instrumentalism

Hilbert's program has often been interpreted as an instrumentalist account of mathematics. This reading relies on the distinction Hilbert makes between the finitary part of mathematics and the non-finitary rest which is in need of grounding (via finitary meta-mathematics). The finitary part Hilbert calls "contentual," i.e., its propositions and proofs have content. The infinitary part, on the other hand, is "not meaningful from a finitary point of view." This distinction corresponds to a distinction between formulas of the the axiomatic systems of mathematics for which consistency proofs are being sought. Some of the formulas correspond to contentual, finitary propositions: they are the "real" formulas. The rest are called "ideal." They are added to the real part of our mathematical theories in order to preserve classical inferences such as the principle of the excluded middle for infinite totalities, i.e., the principle that either all numbers have a given property or there is a number which does not have it. This disjunction is not finitarily valid, as we saw above. Hilbert first mentioned "ideal" propositions in [1926], although the distinction was hinted at in [1923]. In the latter paper, Hilbert presented a formal system of quantifier-free number theory about which he says that "the provable formulae we acquire in this way all have the character of the finite" (1139). Then the transfinite axioms (i.e., quantifiers) are added to simplify and complete the theory (1144). Here he draws the analogy with the method of ideal elements: "In my proof theory, the transfinite axioms and formulae are adjoined to the finite axioms, just as in the theory of complex variables the imaginary elements are adjoined to the real, and just as in geometry the ideal constructions are adjoined to the actual" (ibid). When Hilbert, in [1926], explicitly introduces the notion of an ideal proposition, and in [1928], when he first speaks of *real propositions* in addition to the ideal, he is quite clear that the real part of the theory consists only of decidable, variable-free formulas. They are supposed to be "directly capable of verification" — akin to propositions derived from laws of nature which can be checked by experiment [Hilbert, 1928, 475].[15] It is this extension of the real part of the theory by the ideal, infinitary part that is in need of justification by a consistency proof:

> For there is a condition, a single but absolutely necessary one, to which the use of the method of ideal elements is subject, and that is the *proof of consistency*; for, extension by the addition of ideals is legitimate only if no contradiction is thereby brought about in the old, narrower domain, that is, if the relations that result for the old objects whenever the ideal objects are eliminated are valid in the old domain. [Hilbert, 1926, 383]

Weyl [1925] described Hilbert's project as replacing meaningful mathematics by a meaningless game of formulas. He noted that Hilbert wanted to "secure not *truth*,

---

[15]This reading is not universally accepted. [Detlefsen, 1990], for instance, considers the real formulas to also include the general formulas, i.e., formulas with free variables. See [Zach, 2004] for a defense of the reading given here.

but the *consistency* of analysis" and suggested a criticism that echoes an earlier one by Frege: Why should we take consistency of a formal system of mathematics as a reason to believe in the truth of the pre-formal mathematics it codifies? Is Hilbert's meaningless inventory of formulas not just "the bloodless ghost of analysis"? Weyl suggested a solution:

> [I]f mathematics is to remain a serious cultural concern, then some *sense* must be attached to Hilbert's game of formulae, and I see only one possibility of attributing to it (including its transfinite components) an independent intellectual meaning. In theoretical physics we have before us the great example of a [kind of] knowledge of completely different character than the common or phenomenal knowledge that expresses purely what is given in intuition. While in this case every judgment has its own sense that is completely realizable within intuition, this is by no means the case for the statements of theoretical physics. In that case it is rather *the system as a whole* that is in question if confronted with experience. [Weyl, 1925, 140]

The analogy with physics is striking. Hilbert himself used a similar analogy in [1928]. He there suggested that consistency is not the only virtue ideal mathematics has: transfinite inference simplifies and abbreviates proofs, "brevity and economy of thought are the *raison d'être* of existence proofs" (476). The formal system of transfinite logic is not arbitrary: "This formula game is carried out according to certain definite rules, in which the *technique of our thinking* is expressed... The fundamental idea of my proof theory is none other than to describe the activity of our understanding, to make a protocol of the rules according to which our thinking actually proceeds" (ibid).

Although these remarks are suggestive, they do not force an interpretation of Hilbert as an instrumentalist. Most commentators have taken this reading (including [**?**; Giaquinto, 1983; Sieg, 1990], and in particular [Detlefsen, 1986]), in that they interpret Hilbert's explanation that the ideal propositions "have no meaning in themselves" [Hilbert, 1926, 381] as claiming that classical mathematics is a *mere* instrument, and that statements of transfinite mathematics have no truth value. To the extent that this is accurate, however, it must be understood as a methodological instrumentalism: A successful execution of the proof-theoretic program would show that one could pretend *as if* mathematics was meaningless.[16] The analogy with physics is therefore not: transfinite propositions have no meaning just as propositions involving theoretical terms have no meaning, but: transfinite propositions require no direct intuitive meaning just as one does not have to directly see electrons in order to theorize about them. Hallett [1990], taking into account the 19th century mathematical background from which Hilbert came as well as published and unpublished sources from Hilbert's entire career (in particular [Hilbert, 1992], the most extensive discussion of the method of ideal elements) comes to the following conclusion:

---

[16]On this point see also [Sieg, 1999], esp. B3 and the conclusion, and [Sieg, 2002].

> [Hilbert's treatment of philosophical questions] is *not* meant as a kind of instrumentalist agnosticism about existence and truth and so forth. On the contrary, it is meant to provide a non-skeptical and positive solution to such problems, a solution couched in cognitively accessible terms. And, it appears, the same solution holds for both mathematical and physical theories. Once new concepts or "ideal elements" or new theoretical terms have been accepted, then they exist in the sense in which *any* theoretical entities exist. [Hallett, 1990, 239]

This conclusion is in line with Weyl's assessment in [1928]. When Weyl eventually turned away from intuitionism,[17] he emphasized the purpose of Hilbert's proof theory, not to turn mathematics into a meaningless game of symbols, but to turn it into a theoretical science which codifies scientific (mathematical) practice.

The reading of Hilbert as an instrumentalist goes hand in hand with a reading of the proof-theoretic program as a reductionist project. The instrumentalist reading interprets ideal mathematics as a meaningless formalism, which simplifies and "rounds out" mathematical reasoning. But a consistency proof of ideal mathematics by itself does not explain what ideal mathematics is an instrument *for*. Thus, commentators have sought to elaborate on Hilbert's discussion of consistency by pointing out that consistency proofs do not just establish that ideal theories are free from formal contradictions, but that they establish more than mere consistency. They establish *conservativity* of the ideal over the real part of the theory, in the following sense: If the ideal theory proves a real statement, then the real statement is also provable by real, finitary means. Such reductivist projects were common in the philosophy of science at the time, as was pointed out by Giaquinto [1983]. A conservativity proof justifies the use of transfinite mathematics: it is not only internally consistent, but it proves only true intuitive propositions.

On this picture, classical mathematics is to be formalized in a system which includes formalizations of all the directly verifiable (by calculation) propositions of contentual finite number theory. The consistency proof should show that all real propositions which can be proved by ideal methods are true, i.e., can be directly verified by finite calculation. Actual proofs such as the ε-substitution procedure are of such a kind: they provide finitary procedures which eliminate transfinite elements from proofs of real statements. In particular, they turn putative ideal derivations of $0 = 1$ into derivations in the real part of the theory; the impossibility of such a derivation establishes consistency of the theory. Indeed, Hilbert saw that something stronger is true: not only does a consistency proof establish truth of real formulas provable by ideal methods, but it yields finitary proofs of finitary *general* propositions if the corresponding free-variable formula is derivable by ideal methods [Hilbert, 1928, 474].

It bears pointing out that Hilbert never clearly articulated conservativity of the ideal over the real for finitary general statements as an aim of his foundational project. There are contemporary commentators, e.g., von Neumann [1931] who

---

[17]For the reasons for Weyl's rejection of intuitionism, see [Mancosu and Ryckman, 2002].

attribute to Hilbert an interest in conservativity proofs, but it was only Bernays in the *Grundlagen der Mathematik* who pointed out that consistency proofs themselves established not only the truth of variable-free formulas provable by ideal methods, but also of free-variable theorems. In this context, Bernays used the term 'verifiable' (*verifizierbar*): a free-variable formula $A(x)$ is verifiable if each numerical instance $A(\mathfrak{z})$ is (finitarily) true. The proof transformation methods used in consistency proofs for systems of arithmetic in [Hilbert and Bernays, 1934, 248, 298] can be applied not only to putative proofs of $0 = 1$, but generally to proofs of formulas with free-variables. If we have a proof of $A(x)$, then the following method constitutes a finitary proof that, for any $\mathfrak{z}$, $A(\mathfrak{z})$ is true. From the derivation of $A(x)$ we obtain a derivation of $A(\mathfrak{z})$ by substitution. The procedure given in the consistency proof transforms this derivation into a variable-free derivation of $A(\mathfrak{z})$ in the real part of the theory, which codifies a finitary calculation that $A(\mathfrak{z})$ is true.

Kreisel was most influential in promoting the interpretation of the aim of Hilbert's program as an attempt to establish conservativity of the ideal theory for finitary general propositions all along. Kreisel [1951] cites Bernays's results; but in [Kreisel, 1960] and later, he instead points to the argument in [Hilbert, 1928, 474]. This argument, unlike Bernays's, does not rely on a particular form of the consistency proof. It assumes only that a finitary consistency proof for an ideal theory is available. Assume there is a derivation of $A(x)$. Now suppose that for a given $\mathfrak{z}$, $A(\mathfrak{z})$ is not true. Then $\neg A(\mathfrak{z})$ would be true,[18] and so there would be a derivation of $\neg A(\mathfrak{z})$ in the ideal theory (which includes all real theorems). But from the derivation of $A(x)$ we obtain, by substitution, a derivation of $A(\mathfrak{z})$, and hence a contradiction. Since we assume that we have a finitary consistency proof of T, this cannot be the case. Hence, $A(\mathfrak{z})$ must be true.

3.4   Hilbert's program and Gödel's incompleteness theorems

Gödel announced the second incompleteness theorem in an abstract published in October 1930: no consistency proof of systems such as *Principia*, Zermelo-Fraenkel set theory, or the systems investigated by Ackermann and von Neumann is possible by methods which can be formulated in these systems. In the full version of his paper, Gödel [1931] left open the possibility that there could be finitary methods which are not formalizable in these systems and which would yield the required consistency proofs. Bernays's first reaction in a letter to Gödel in January 1931 was likewise that "if, as von Neumann does, one takes it as certain that any and every finitary consideration may be formalized within the system P — like you, I regard that in no way as settled — one comes to the conclusion that a finitary demonstration of the consistency of P is impossible" [Gödel, 2003a, 87].

Through a careful ("Gödel"-) coding of sequences of symbols (formulas, proofs) by numbers, Gödel showed that in theories T which contain a sufficient amount

---

[18]This inference uses tertium non datur, but only regarding the unproblematic finitary statement $A(\mathfrak{z})$.

of arithmetic, it is possible to produce a formula $Pr(x, y)$ which expresses that x is (the code of) a proof of (the formula with code) y. Specifically, if $\ulcorner 0 = 1 \urcorner$ is the code of the formula $0 = 1$, then $Con_T \equiv \forall x \, \neg Pr(x, \ulcorner 0 = 1 \urcorner)$ expresses that T is consistent (no number is the code of a derivation in T of $0 = 1$). The second incompleteness theorem (G2) says that under certain assumptions about T and the coding apparatus, T does not prove Con_T. Now suppose there were a finitary consistency proof of T. The methods used in such a proof would presumably be formalizable in T. ("Formalizable" means that, roughly, if the proof uses a finitary operation f on derivations which transforms any derivation d into a derivation $f(d)$ of a simple form; then there is a formula $F(x, y)$ so that, for all derivations d, $T \vdash F(\ulcorner d \urcorner, \ulcorner f(d) \urcorner)$.) The consistency of T would be finitarily expressed as the general hypothetical claim that, if d is any given sequence of symbols, d is not a derivation in T of the formula $0 = 1$. The formalization of this proposition is the formula $\neg Pr(x, \ulcorner 0 = 1 \urcorner)$ in which the variable x occurs free. If there were a finitary proof of the consistency of T, its formalization would yield a derivation in T of $\neg Pr(x, \ulcorner 0 = 1 \urcorner)$, from which Con_T can be derived in T by simple universal generalization on x. Yet, a derivation of Con_T in T is ruled out by G2.

Gödel and Bernays initially thought that the difficulty for the consistency proof of Peano arithmetic PA could be overcome by employing methods which, although not formalizable in PA, are nonetheless finitary. Bernays did not seem to have any particular candidates for such a method in mind, and he thought that all methods which were up to then employed in finitary considerations were in fact formalizable in PA. Another option he considered was an extension of the notion of an axiomatic theory by a finitary version of the ω-rule proposed by Hilbert [1931a; 1931b]. Such theories might avoid the impact of Gödel's incompleteness theorem since they do not satisfy the conditions of the incompleteness theorems: the set of axioms would not be decidable. It is fair to say, however, that since about 1934 it has been almost universally accepted that the methods of proof accepted as finitary prior to Gödel's results are all formalizable in PA and that the incompleteness theorems do show that there can be no finitary consistency proofs for axiomatic theories of the kind originally considered by Hilbert.

The reaction to the incompleteness theorems in the Hilbert school then focused on extensions of the original finitary standpoint in which consistency proofs for substantial theories like PA can be carried out. Such extensions have been proposed and defended on broadly finitary grounds, e.g., Gentzen [1936] defended the use of transfinite induction up to ε_0 in his consistency proof for PA as "indisputable," and Takeuti [1987] gave another defense. In the Gentzen-Schütte tradition of proof theory by ordinal analysis, the proof methods used to give consistency proofs are all of this sort. To wit, one uses transfinite induction on ordinal notation systems which name larger and larger ordinals. The more complicated the ordering, the more difficult it is to see that the induction principle in question is finitarily justified. Another extension of the finitary standpoint is due to Gödel [1958].

Smoryński [1977], following earlier suggestions by Kreisel, has argued that already the first incompleteness theorem defeats Hilbert's program. This argument uses the fact that a finitary consistency proof of an ideal theory T implies the conservativity of T over finitary, real mathematics for general finitary statements of the form $A(x)$ (with free variable x). Now Gödel's first incompleteness theorem (G1) states that for any sufficiently strong, consistent formal theory S there is a sentence G_S which is not derivable in S if S is consistent. G_S is a general real sentence. Consider a theory T which formalizes ideal mathematics and contains the theory S, which formalizes real mathematics, as a subtheory. S satisfies the conditions of G1 and hence S does not derive G_S. Yet T, being a formalization of all of mathematics, proves (via a formalization of G1) that if S is consistent, then G_S, but it also proves the consistency (indeed, the soundness) of S. Hence T proves G_S. Thus, we have a true real statement which is provable in ideal mathematics but not in real mathematics.[19]

4 HILBERT'S PROGRAM NOW

4.1 Detlefsen's Hilbertian instrumentalism

Detlefsen [1979; 1986; 2001] has proposed a wide-ranging instrumentalist account of mathematics based on Hilbert's program which is designed to escape the difficulties that Gödel's incompleteness theorems poses for Hilbert's original consistency project. The project has several parts: Detlefsen [1986] first gives a detailed analysis and defense of a general instrumentalist view of mathematics along Hilbertian lines. This includes an analysis of the epistemic import of ideal proofs of real statements, which answers a question that was hardly addressed by Hilbert, either in his mature writings in the 1920s or in his exchange with Frege on formalism and consistency. This is the question of how manipulation of meaningless strings of symbols can ever lead to knowledge (of finitary truths). Detlefsen then analyzes the characteristics of formal systems of ideal mathematics *qua* instruments. On Detlefsen's view, even though, say, full set theory is commonly accepted as a formalization of infinitary mathematics, only parts of set theory are in fact instrumentally useful. In particular, (1) ideal proofs of real theorems which are more complex than any real proof of the same theorem do not yield an instrumental advantage, and hence are not instrumentally useful; and (2) ideal proofs which are too long or complex to be comprehended by humans, and hence never play a role in actual mathematical reasoning, are also of no instrumental value. A proof theoretic justification of instrumental mathematics, i.e., the proof of the conservativity of the ideal theory over real mathematics, is only required, so Detlefsen, for the instrumentally useful part. Detlefsen introduces the term "Hilbertian residue" for that part of ideal mathematics that is instrumentally useful and hence in need of

---

[19]The argument appeals to a number of perhaps contentious assumptions, such as that T proves the soundness of S. For a critique, see [Detlefsen, 1990].

proof-theoretic justification. On his view, then, we only need a consistency proof for the Hilbertian residue, not for all of ideal mathematics.

This move from a required justification for all of infinitary mathematics to a justification of only the Hilbertian residue is one step toward Detlefsen's defense of instrumentalism against Gödel's incompleteness theorems. For the incompleteness theorems only apply under certain conditions, e.g., only when the theory in question contains enough basic arithmetic to carry out Gödel coding, formalization of a proof predicate, and to prove the diagonal lemma. The Hilbertian residue of a theory, however, need not contain a sufficiently strong arithmetical subtheory because of (1) above. This provides part of Detlefsen's defense against the challenge posed by the first incompleteness theorem [Detlefsen, 1986, Appendix]. Detlefsen [1990] also argues that instrumentalism escapes the argument from G1 by denying that ideal mathematics must be conservative over the real part. According to him, Hilbertian instrumentalism requires only that the ideal theory not prove anything which is in conflict with the real theory; it is not required that all its real theorems are also provable by real means. If this defense is successful, Detlefsen is right to claim that not G1, but only G2 poses a real challenge to instrumentalism.

Detlefsen presents several lines of defense against the argument from G2, one of which [1979] echoes [Hilbert, 1931b]. If a version of the ω-rule is finitarily acceptable, then we would have found a finitarily acceptable method of proof which is not capable of formalization. Hence, real mathematics is not a subtheory of the ideal instrument, but this was an assumption necessary to draw the conclusion that there can be no real consistency proof of the ideal theory. Ignjatović [1994] raised serious doubts about the acceptability of Detlefsen's version of the ω-rule, however. Detlefsen's other argument against the common interpretation of Gödel's second theorem focuses on the notion of formalization. That the particular formalization of "T is consistent" by Gödel's formula Con_T is not provable does not imply that there could not be other formulas, which *are* provable in T, and which have as much right to be called "formalizations of the consistency of T." These rely on different formalizations of the provability predicate Pr_T than the standard ones. It is known that formalized consistency statements are unprovable whenever the provability predicate obeys certain general derivability conditions. Detlefsen argues that these conditions are not necessary for a predicate to count as a genuine provability predicate, and indeed there are provability predicates which violate the provability conditions and which give rise to consistency formulas which *are* provable in their corresponding theories. These, however, depend on nonstandard conceptions of provability which would likely not have been accepted by Hilbert. One quite basic example is the use of Rosser provability instead of ordinary provability. On this approach, a derivation of a formula A only counts as a proof if no derivation with smaller Gödel number is a derivation of $\neg A$. If $Prov(x, \ulcorner A \urcorner)$ is the standard formalization of "x is the code of a derivation of the formula A," then the Rosser provability predicate is given by

$$RPr(\ulcorner A \urcorner) \equiv \exists x (Prov(x, \ulcorner A \urcorner) \wedge \forall y < x \, \neg Prov(y, \ulcorner \neg A \urcorner)).$$

For this provability predicate, $\neg RPr(\ulcorner 0 = 1 \urcorner)$ *is* provable in, e.g., first-order Peano arithmetic. Provability of a formula A, however, is no longer just a matter of deriving it from the axioms; one also has to check that all shorter derivations do not end in $\neg A$. Other "consistency minded" theories which prove their own consistency are discussed, e.g., in [Jeroslow, 1971; Jeroslow, 1975] and especially [Visser, 1989]. The Rosser provability predicate is studied in, e.g., [Guaspari and Solovay, 1979] and [Arai, 1990].[20]

Another interesting aspect of Detlefsen's approach to instrumentalism and Hilbert's program related to technical work in proof theory is the emphasis on instrumental utility of ideal proofs. Hilbert, as we saw above, himself noted the theoretical and cognitive advantage of ideal methods, such as increased simplicity of proofs. In Detlefsen's instrumentalism, such considerations take center stage. Even if it is conceded that Gödel's theorems call the success of instrumentalism in its most general form into question, it would still be of substantial interest to study restricted cases of conservative extensions of real mathematics which are instrumentally useful. To flesh out the notion of "instrumental usefulness," one obvious characteristic of formal proofs is length. For instance, one might take an ideal theory to be useful if its proofs are substantially shorter than proofs in, say, PRA of the same theorems. This question is amenable to precise proof theoretical study. Caldon and Ignjatovič [2005] prove some related, but on the whole, negative results: The subsystem of first-order arithmetic $I\Sigma_1$ in which induction is limited to Σ_1 formulas has super-exponential "speed-up" over PRA. This indicates that using induction over non-finitary formulas (Σ_1 formulas have unbounded existential quantifiers) yields significantly shorter proofs than proofs without. However, more comprehensive theories (RCA, WKL, see below) which contain some second-order apparatus, do not significantly shorten proofs vis-à-vis $I\Sigma_1$.

4.2   Generalized Hilbert programs

The work of Gentzen on consistency proofs for mathematical theories using methods that go beyond the strictly finitary marks the beginning of an important line of proof-theoretic research. As outlined in 2.3 above, Gentzen's approach was to retain the aim of Hilbert's program, viz., to give consistency proofs for strong mathematical theories by restricted means. Because of Gödel's incompleteness theorems, these restricted means are necessarily not themselves formalizable in the theories whose consistency is established by them. Nevertheless, they retain a constructive character, and attempts have been made to justify them on finitary grounds.

The consistency proof of Gentzen [1936], as discussed above, uses the principle of transfinite induction up to ε_0 in order to establish the consistency of first-order Peano arithmetic. Gentzen's use of a system of notations for ordinals less

---

[20] For technical background, discussion of intensional provability predicates and examples, see [Feferman, 1960]. For discussion, see also [Resnik, 1974b], [Auerbach, 1985; Auerbach, 1992] and [Steiner, 1991].

than ε_0, and the proof of the termination of a reduction procedure for derivations in PA based on induction on these ordinal notations, provide the model for the proof theoretic analysis of axiomatic systems along these lines. In order to give an "ordinal analysis" of a theory T, one typically produces an ordinal notation system for ordinals less than some ordinal α such that for every $\beta < \alpha$, the formalization $TI(\beta)$ of the transfinite induction principle for β is provable in T. In practice, using transfinite induction up to α itself and otherwise only strictly finitary methods, one can prove the consistency of T. The fact that induction up to ε_0 establishes the consistency of PA, together with the result of [Gentzen, 1943] that shows that for all $\beta < \varepsilon_0$, PA proves $TI(\beta)$ for all $\beta < \varepsilon_0$ constitutes an ordinal analysis of PA, and we say that ε_0 is the proof theoretic ordinal of PA.

Proof theory in the tradition of Gentzen and Schütte as well as Takeuti has focused on such ordinal analyses of theories of increasing strength. In recent work, Rathjen [2005b; 2005a] has pushed the boundaries of this approach in giving an ordinal analysis of a very strong subsystem of analysis called Π^1_2-comprehension.[21] The consistency proofs in this tradition are, for the most part, based on the approach of [Schütte, 1960], which uses a variant of Gentzen's formalization using infinitary derivations. A second tradition has pursued ordinal analysis using extensions of Ackermann's ε-substitution method [1940], for examples see [Mints and Tupailo, 1999] and [Arai, 2003].

Although generalized Hilbert programs in this tradition have certainly produced important mathematical work, its philosophical underpinnings are thin. Takeuti [1987] attempted to give a finitary justification for the proof theoretic ordinal ε_0, but it is unclear to what extent more complex ordinal notation systems are finitarily acceptable. Even if one concedes, as, e.g., Schütte does, that the consistency proofs in question are constructive (although no longer strictly finitary), it is still unclear what the philosophical significance of the technical results is. Feferman [1988, 366] offers this assessment:

> [A]s the systems of ordinal notation used for consistency proofs of stronger and stronger theories become more and more complicated, the significance to noncognoscenti of what is thereby accomplished decreases in inverse proportion. Thus, on the one hand, to say that one has obtained a constructive consistency proof of a theory T — without saying anything more — is too general to be informative; and, on the other hand, to say that the proof has been carried out by transfinite induction on a certain complicated recursive ordering for some very large ordinal tells us nothing about what constructive principles are involved in the proof of this well-ordering.[22]

Another important proof-theoretical approach in which the analysis of systems of classical mathematics is accomplished using a generalization of the finitary

---

[21]See [Pohlers, 1987] for a survey of the work in the Schütte school, and [Pohlers, 1998] for a more recent technical survey.

[22]For a more forceful criticism of proof theory in this tradition, see [Kreisel, 1976].

standpoint is that of functional interpretations. The model for this approach is Gödel's *Dialectica* interpretation [1958]. The *Dialectica* interpretation shows how one can reduce an infinitary theory T (in this case, intuitionistic first-order arithmetic) to a quantifier-free theory F.[23] An ordinal analysis of a theory does something similar, for instance, one can view Gentzen's consistency proof as reducing Peano arithmetic to a quantifier-free theory (PRA) extended by a certain infinitary induction principle ($TI(\varepsilon_0)$). In the case of functional interpretations, the quantifier-free theory F is also not strictly finitary: it does not just mention finite objects but also certain infinitary objects, viz., functionals of finite type. A functional interpretation can be seen as a reduction of the infinitary theory T to the theory of functionals F in question. The approach using functional interpretations has the following advantage over the Gentzen-Schütte approach. It is a consequence of the reduction of T to F that every recursive function which can be proved to be total in T is represented by a term of F. Because the functionals of F in practice characterize natural classes of functions, a functional interpretation yields an appealing analysis of the computational content of F. Moreover, the conceptual import of the reduction is more apparent than in the case of ordinal analysis: already in the case of *PA*, Gödel's functionals of finite type provide a much clearer account of the character of the constructive methods appealed to than induction up to ε_0.[24]

4.3   Relativized Hilbert programs

A philosophically more satisfactory continuation of Hilbert's program in proof theoretic terms has been suggested by Kreisel [1954; 1968; 1983] and has been elaborated especially by Feferman. This work proceeds from a wider conception of Hilbert's program as an attempt to justify ideal mathematics by restricted means. On this conception, the aim of Hilbert's proof theory was to show that, at least as far as a certain class of real propositions is concerned, ideal mathematics does not go beyond real mathematics, and in this sense finitary mathematics is a foundation for ideal mathematics. A finitary consistency proof of the kind envisaged by Hilbert would have accomplished this for all of classical mathematics.

The scope of the foundational project, however, need not necessarily be all of higher mathematics. So-called relativized Hilbert programs are projects in which one considers certain fragments of higher mathematics as the theory for which a foundation is sought (and, indeed, also theories stronger than finitism as candidates for the reducing theory, e.g., predicative theories). Examples of these are Feferman's work on explicit mathematics and predicative subsystems of analysis, and to some extent also the Friedman-Simpson program of reverse mathematics (see below). What is common to these approaches to mathematical

---

[23]Via the interpretation of classical arithmetic in intuitionistic arithmetic [Gentzen, 1933; Gödel, 1933], the *Dialectica* interpretation also yields a functional interpretation of classical arithmetic.

[24]For an excellent survey of this approach, see [Avigad and Feferman, 1998].

foundations is that they concentrate on so-called proof-theoretic reductions of systems of classical mathematics to more restricted systems. The reduction is carried out using finitist means, and therein lies its philosophical significance.

A foundational reduction in Feferman's sense [1988; 1993a] is accomplished if it can be shown that a body of mathematics which is justified by a foundational framework $\mathcal{F}_1$ (e.g, finitary, constructive, predicative, infinitary, set-theoretic) can already be justified, in a certain sense, in a weaker, or stricter foundational framework $\mathcal{F}_2$. This is in general not possible in a wholesale fashion, however, partial foundational reductions can and have been achieved. Suppose a theory T_1 is justified by a foundational framework $\mathcal{F}_1$, and a theory T_2 by a weaker framework $\mathcal{F}_2$. A proof theoretic reduction of T_1 to T_2 (conservative for Φ) is a partial recursive function f such that

1. whenever x is (the code of) a proof in T_1 of a formula (with code) y in Φ, then $f(x)$ is (the code of) a proof of y in T_2, and

2. T_2 proves the formalization of (1).

If there is such a function f, we write $T_1 \leq T_2[\Phi]$. Now if T_1 is directly justified by a foundational framework $\mathcal{F}_1$, and T_2 by $\mathcal{F}_2$, then, so Feferman, a proof-theoretic reduction that establishes $T_1 \leq T_2[\Phi]$ is a partial foundational reduction of $\mathcal{F}_1$ to $\mathcal{F}_2$. Clause (2) in the definition ensures that the reduction (the function f) itself is justified by the weaker framework $\mathcal{F}_2$. In the reductions achieved in practice, it turns out that f is actually primitive recursive and the formalization of (1) can even be proved in primitive recursive arithmetic PRA. Since PRA is directly justified by the finitary framework, such partial foundational reductions are therefore all finitarily justified. Feferman's main philosophical conclusion from the possibility of giving such foundational reductions is this: The main argument for set-theoretical realism is the Quine-Putnam indispensability argument, which proceeds from the premises that set-theory is indispensable to science. Feferman has shown, first, that much, if not all, of scientifically applicable mathematics can actually be formalized in much weaker systems (e.g., Feferman's system W, which is justified by a predicative foundational framework), and second, that predicative mathematics can be reduced to the countably infinite (in the sense that there is a partial foundational reduction of predicative mathematics to countably infinite mathematics, given by a proof-theoretic reduction of W to Peano Arithmetic PA). He concludes that,

> even if one accepts the indispensability argument, practically nothing philosophically definite can be said of the entities which are then supposed to have the same status — ontologically and epistemologically — as the entities of natural science. That being the case, what do the indispensability arguments amount to? As far as I'm concerned, they are completely vitiated. [Feferman, 1993b]

Independently of the question of mathematical realism and of the scope and force of the indispensability arguments, proof-theoretic reductions give precise answers to

questions of the relation between foundational frameworks. Since a proof-theoretic reduction of T_1 to T_2 also yields a consistency proof of T_1 in T_2 (i.e., a relative consistency result), establishing a proof-theoretic reduction also provides a solution to Hilbert's program relativized to T_1 and T_2. Feferman summarizes the importance of proof-theoretic reductions thus:

> In general, the kinds of results presented here serve to sharpen what is to be said in favor of, or in opposition to, the various philosophies of mathematics such as finitism, predicativism, constructivism, and set-theoretical realism. Whether or not one takes one or another of these philosophies seriously for ontological and/or epistemological reasons, it is important to know which parts of mathematics are in the end justifiable on the basis of the respective philosophies and which are not. The uninformed common view — that adopting one of the non-platonistic positions means pretty much giving up mathematics as we know it — needs to be drastically corrected, and that should also no longer serve as the last-ditch stand of set-theoretical realism. On the other hand, would-be nonplatonists must recognize the now clearly marked sacrifices required by such a commitment and should have well-thought out reasons for making them. [Feferman, 1993a]

Proof theorists have obtained a number of such results, including reductions of theories which on their face require a significant amount of ideal mathematics for their justification (e.g., subsystems of analysis) to finitary systems.[25]

The program of so-called reverse mathematics developed by, in particular, Friedman and Simpson, is another continuation of Hilbert's program. In the face of Gödel's results showing that not all of classical mathematics can be reduced to the finitary, they seek to answer the question: how much of classical mathematics can be so reduced? Reverse mathematics aims to give a precise answer to this question by investigating which theorems of classical mathematics are provable in weak subsystems of analysis which are reducible to finitary mathematics (in the sense discussed above). A typical result is that the Hahn-Banach theorem of functional analysis is provable in a theory known as WKL_0 (for "weak König lemma"); WKL_0 is proof-theoretically reducible to PRA for Π^0_2 sentences (i.e., sentences of the form $\forall x \exists y\, A(x, y)$).[26]

Reverse mathematics in this tradition is primarily concerned with infinitary theories, e.g., subsystems of analysis. Gödel's theorems show, however, that not even all truths of first-order number theory are provable in Peano arithmetic, and hence that not even the domain of all arithmetical truths can be given a foundation on finitary principles. This suggests the questions of whether there are "mathematically interesting" statements of number theory which are not provable

---

[25]For a discussion of the philosophical significance of such proof theoretic reductions, see [Feferman, 2000] and [Hofweber, 2000].

[26]See [Simpson, 1988] for an overview, [Simpson, 1999] for a technical introduction to reverse mathematics, and also the collection [Simpson, 2005].

in systems that can be reduced to the finitary. The most celebrated result in this regard is the proof by Paris and Harrington [1977] that a version of the finite Ramsey theorem is not provable in Peano arithmetic. However, this and other examples of independent number theoretic statements are constructed specifically to be independent of Peano arithmetic. It turns out that a great many "ordinary" number theoretic results are provable even in weak fragments of first-order number theory, and this has led Friedman to conjecture that "every theorem published in the *Annals of Mathematics* whose statement involves only finitary mathematical objects (i.e., what logicians call an arithmetical statement) can be proved in elementary arithmetic." (Here, elementary arithmetic is a very weak theory which can be proved consistent by primitive recursive methods.) Avigad [2003] gives an excellent survey of the issues and results related to this conjecture and places it in the philosophical context of Hilbert's program.[27]

The results surveyed in this section are all natural continuations of Hilbert's original ideas. A central aspect of recent proof-theoretical investigations and of Hilbert's original program alike is that they study formalized systems using metamathematical tools with the aim of understanding the structure and content of these systems. Hilbert's original consistency project, the conservativity project that Kreisel and others interpret Hilbert as having been engaged in, as well as reductive proof theory are all examples of this, and this is also part of the reason why many practicing proof theorists see themselves as still working on Hilbert's program. Ordinal analysis, functional interpretations, proof theoretic reductions and reverse mathematics are only some of the most prominent areas of proof theory, and those most explicitly situated in the tradition of Hilbert's program. Many other areas of proof theory other than those directly concerned with consistency and foundational reductions of theories are related to the aims of Hilbert's program, e.g., the no-counterexample interpretation [Kreisel, 1951; Tait, 2005a] and work on the structure and complexity of formal proofs more generally [Pudlák, 1998].

5   CONCLUSION

Although it has been a commonplace in the literature on the philosophy of mathematics and logic from 1950 onward that Hilbert's program has not only been "killed" by Gödel's incompleteness theorems but that it was over-ambitions if not ill-conceived from the start, in the current literature a more positive evaluation has emerged. This is in large part due to the attention which unpublished writings in the Hilbert school (especially lecture notes to Hilbert's courses) have received recently, as well as to the availability of more of the published writings in English translation (e.g., in [Ewald, 1996] and [Mancosu, 1998a]). But it is also due to a growing recognition that the common characterizations of Hilbert's program are

---

[27]See also [Raatikainen, 2003] on the current status of the various branches of proof-theoretic research relating to Hilbert's program.

caricatures, and to a clearer philosophical engagement with the recent results of proof theory. For instance, Ramsey's characterization that, according to Hilbert, "Mathematics proper is thus regarded as a sort of game, played with meaningless marks on paper rather like noughts and crosses" [Ramsey, 1926, 231], and the view that Hilbert held a naive formalist and instrumentalist view of mathematics, have been criticized by various writers. It remains to be seen in what ways Hilbert's philosophical views can be resuscitated (in the manner in which, e.g., Frege's logicist program has experienced a renaissance in the neo-logicist writings of, e.g., Boolos, Heck, and Hale and Wright). It should be clear in any case from the discussion in the preceding section that ideas closely related to Hilbert's own *have* been hugely successful. And it is also clear from the recent historical studies of Hilbert's unpublished writings as well as from the study of the proof theoretical practice of the Hilbert school that the ideas which underpin much of recent and current proof theoretical research are not merely "inspired by" Hilbert's program. Hilbert's fundamental aim was, all along, to make mathematical reasoning amenable to mathematical investigation, and to carry out such an investigation which yields an analysis of non-constructive reasoning in terms of restricted methods. Hilbert, of course, emphasized consistency of non-constructive systems as the most interesting property to be investigated, and emphasized finitary methods as those in terms of which such an analysis should be carried out. But even in the 1920s, in the practice of consistency proofs in the work of Ackermann, Bernays, and von Neumann, among others, more broadly constructive methods were employed in this analysis, and results about properties other than consistency were obtained. Gentzen's work of the 1930s and subsequent proof theoretical studies should thus not be seen as merely a response to Gödel's incompleteness results, but more broadly as advancing Hilbert's original aim of investigating the structure of mathematical reasoning. Seen in this light, again, proof theory as a foundational enterprise is very much alive. Although Gödel's theorems show that Hilbert's original expectations about what exactly can be analyzed in which way and with what restricted methods can not be fulfilled, proof theory and Hilbert's aims more generally have been advanced tremendously over the last half-century.

ACKNOWLEDGEMENTS

BIBLIOGRAPHY

[Ackermann, 1925] W. Ackermann.   Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Mathematische Annalen*, 93:1–36, 1925.

[Ackermann, 1940] W. Ackermann. Zur Widerspruchsfreiheit der Zahlentheorie. *Mathematische Annalen*, 117:162–194, 1940.

[Arai, 1990] T. Arai. Derivability conditions on Rosser's provability predicates. *Notre Dame Journal of Formal Logic*, 31:487–497, 1990.

[Arai, 2003] T. Arai. Epsilon substitution method for $ID_1(\Pi_1^0 \vee \Sigma_1^0)$. *Annals of Pure and Applied Logic*, 121:163–208, 2003.

[Auerbach, 1985] D. Auerbach. Intentionality and the Gödel theorems. *Philosophical Studies*, 48:337–351, 1985.

[Auerbach, 1992] D. Auerbach. How to say things with formalisms. In Michael Detlefsen, editor, *Proof, Logic, and Formalization*, pages 77–93. Routledge, London, 1992.

[Avigad, 2003] J. Avigad. Number theory and elementary arithmetic. *Philosophia Mathematica*, 11:257–284, 2003.

[Avigad and Feferman, 1998] J. Avigad and S. Feferman. Gödel's functional ("Dialectica") interpretation. In Buss [1998].

[Avigad and Reck, 2001] J. Avigad and E. Reck. "Clarifying the nature of the infinite": the development of metamathematics and proof theory. Technical report cmu-phil-120, Carnegie Mellon University, 2001. Online at `http://www.andrew.cmu.edu/user/avigad/Papers/infinite.pdf`.

[Avigad and Zach, 2002] J. Avigad and R. Zach. The epsilon calculus. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Summer 2002. `http://plato.stanford.edu/entries/epsilon-calculus/`.

[Barwise, 1977] J. Barwise, editor. *Handbook of Mathematical Logic*. North-Holland, Amsterdam, 1977.

[Benacerraf and Putnam, 1983] P. Benacerraf and H. Putnam, editors. *Philosophy of Mathematics*. Cambridge University Press, Cambridge, 2nd edition, 1983.

[Bernays, 1922] P. Bernays. Über Hilberts Gedanken zur Grundlegung der Arithmetik. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 31:10–19, 1922. English translation in [Mancosu, 1998a, 215–222].

[Bernays, 1923] P. Bernays. Erwiderung auf die Note von Herrn Aloys Müller: Über Zahlen als Zeichen. *Mathematische Annalen*, 90:159–63, 1923. English translation in [Mancosu, 1998a, 223–226].

[Bernays, 1928a] P. Bernays. Über Nelsons Stellungnahme in der Philosophie der Mathematik. *Die Naturwissenschaften*, 16:142–45, 1928.

[Bernays, 1928b] P. Bernays. Zusatz zu Hilberts Vortrag über "Die Grundlagen der Mathematik". *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg*, 6:88–92, 1928. English translation in: [van Heijenoort, 1967, 485–489].

[Bernays, 1930] P. Bernays. Die Philosophie der Mathematik und die Hilbertsche Beweistheorie. *Blätter für deutsche Philosophie*, 4:326–67, 1930. Reprinted in [Bernays, 1976, 17–61]. English translation in [Mancosu, 1998a, 234–265].

[Bernays, 1970] P. Bernays. On the original Gentzen consistency proof. In A. Kino, J. Myhill, and R. E. Veseley, editors, *Intuitionism and Proof Theory*, pages 409–417. North-Holland, Amsterdam, 1970.

[Bernays, 1976] P. Bernays. *Abhandlungen zur Philosophie der Mathematik*. Wissenschaftliche Buchgesellschaft, Darmstadt, 1976.

[Buss, 1998] S. R. Buss, editor. *Handbook of Proof Theory*. Elsevier, Amsterdam, 1998.

[Caldon and Ignjatovič, 2005] P. Caldon and A. Ignjatovič. On mathematical instrumentalism. *Journal of Symbolic Logic*, 70:778–794, 2005.

[Detlefsen, 1979] M. Detlefsen. On interpreting Gödel's second theorem. *Journal of Philosophical Logic*, 8:297–313, 1979. Reprinted with a postscript in [Shanker, 1988, 131–154].

[Detlefsen, 1986] M. Detlefsen. *Hilbert's Program*. Reidel, Dordrecht, 1986.

[Detlefsen, 1990] M. Detlefsen. On an alleged refutation of Hilbert's program using Gödel's first incompleteness theorem. *Journal of Philosophial Logic*, 19:343–377, 1990.

[Detlefsen, 2001] M. Detlefsen. What does Gödel's second theorem say? *Philosophia Mathematica*, 9:37–71, 2001.

[Detlefsen, 2005] M. Detlefsen. Formalism. In Stewart Shapiro, editor, *The Oxford Handbook of Philosophy of Mathematics and Logic*, pages 236–317. Oxford University Press, New York and Oxford, 2005.

[Ewald, 1996] W. B. Ewald, editor. *From Kant to Hilbert. A Source Book in the Foundations of Mathematics*, volume 2. Oxford University Press, Oxford, 1996.

[Feferman, 1960] S. Feferman. Arithmetization of metamathematics in a general setting. *Fundamenta Mathematicae*, pages 35–92, 1960.

[Feferman, 1988] S. Feferman. Hilbert's Program relativized: Proof-theoretical and foundational reductions. *Journal of Symbolic Logic*, 53(2):364–284, 1988.

[Feferman, 1993a] S. Feferman. What rests on what? The proof-theoretic analysis of mathematics. In Johannes Czermak, editor, *Philosophy of Mathematics. Proceedings of the Fifteenth International Wittgenstein-Symposium, Part 1*, pages 147–171, Vienna, 1993. Hölder-Pichler-Tempsky. Reprinted in [Feferman, 1998, Ch. 10, 187–208].

[Feferman, 1993b] S. Feferman. Why a little bit goes a long way: Logical foundations of scientifically applicable mathematics. *Proceedings of the Philosophy of Science Association 1992*, 2:442–455, 1993. Reprinted in [Feferman, 1998, Ch. 14, 284–298].

[Feferman, 1998] S. Feferman. *In the Light of Logic*. Oxford University Press, New York and Oxford, 1998.

[Feferman, 2000] Solomon Feferman. Does reductive proof theory have a viable rationale? *Erkenntnis*, 53:63–96, 2000.

[Field, 1980] H. Field. *Science without Numbers. A Defence of Nominalism*. Princeton University Press, Princeton, 1980.

[Gentzen, 1933] G. Gentzen. Über das Verhältnis zwischen intuitionistischer und klassischer Logik. First published in *Archiv für mathematische Logik und Grundlagenforschung* 16 (1974), 119–132. English translation in [Gentzen, 1969, 53–67], 1933.

[Gentzen, 1934] G. Gentzen. Untersuchungen über das logische Schließen I–II. *Mathematische Zeitschrift*, 39:176–210, 405–431, 1934. English translation in [Gentzen, 1969, 68–131].

[Gentzen, 1935] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. Published as "Der erste Widerspruchsfreiheitsbeweis für die klassische Zahlentheorie" in *Archiv für mathematische Logik und Grundlagenforschung* 16 (1974), 97–118. English translation in [Gentzen, 1969, 132–213], 1935.

[Gentzen, 1936] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936. English translation in [Gentzen, 1969, 132–213].

[Gentzen, 1943] G. Gentzen. Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie. *Mathematische Annalen*, 119:140–161, 1943. English translation in [Gentzen, 1969, 287–308].

[Gentzen, 1969] G. Gentzen. *The Collected Papers of Gerhard Gentzen*, M. E. Szabo, editor. North-Holland, Amsterdam, 1969.

[Giaquinto, 1983] M. Giaquinto. Hilbert's philosophy of mathematics. *British Journal for Philosophy of Science*, 34:119–132, 1983.

[Gödel, 1931] K. Gödel. Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931. Reprinted and translated in [Gödel, 1986, 144–195].

[Gödel, 1933] K. Gödel. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines mathematisches Kolloquiums*, 4:34–38, 1933. Reprinted and translated in [Gödel, 1986, 286–295].

[Gödel, 1958] K. Gödel. Über eine bisher noch nicht benütze Erweiterung des finiten Standpunktes. *Dialectica*, pages 280–287, 1958. Reprinted and translated in [Gödel, 1995, 217–251].

[Gödel, 1986] K. Gödel. *Collected Works*, volume 1, Solomon Feferman et al., editors. Oxford University Press, Oxford, 1986.

[Gödel, 1995] K. Gödel. *Collected Works*, volume 2, Solomon Feferman et al., editors. Oxford University Press, Oxford, 1995.

[Gödel, 2003a] K. Gödel. *Collected Works*, volume 4, Solomon Feferman et al., editors. Oxford University Press, Oxford, 2003.

[Gödel, 2003b] K. Gödel. *Collected Works*, volume 5, Solomon Feferman et al., editors. Oxford University Press, Oxford, 2003.

[Guaspari and Solovay, 1979] D. Guaspari and R. M. Solovay. Rosser sentences. *Annals of Mathematical Logic*, 16:81–99, 1979.

[Hallett, 1990] M. Hallett. Physicalism, reductionism and Hilbert. In Andrew D. Irvine, editor, *Physicalism in Mathemtics*, pages 183–257. Reidel, Dordrecht, 1990.

[Hallett, 1994] M. Hallett. Hilbert's axiomatic method and the laws of thought. In Alexander George, editor, *Mathematics and Mind*, pages 158–200. Oxford University Press, Oxford, 1994.

[Herbrand, 1930] J. Herbrand. *Recherches sur la théorie de la démonstration*. Doctoral dissertation, University of Paris, 1930. English translation in [Herbrand, 1971, 44–202].

[Herbrand, 1971] J. Herbrand. *Logical Writings*, Warren D. Goldfarb, editor. Harvard University Press, 1971.

[Hilbert, 1899] D. Hilbert. Grundlagen der Geometrie. In *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen*, pages 1–92. Teubner, Leipzig, 1st edition, 1899. reprinted in [Hilbert, 2004].

[Hilbert, 1900a] D. Hilbert. Mathematische Probleme. *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Math.-Phys. Klasse*, pages 253–297, 1900. Lecture given at the International Congress of Mathematicians, Paris, 1900. Partial English translation in [Ewald, 1996, 1096–1105].

[Hilbert, 1900b] D. Hilbert. Über den Zahlbegriff. *Jahresbericht der Deutschen Mathematiker-Vereinigung*, 8:180–84, 1900. English translation in [Ewald, 1996, 1089–1096].

[Hilbert, 1905] D. Hilbert. Über die Grundlagen der Logik und der Arithmetik. In A. Krazer, editor, *Verhandlungen des dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*, pages 174–85, Leipzig, 1905. Teubner. English translation in [van Heijenoort, 1967, 129–38].

[Hilbert, 1918a] D. Hilbert. Axiomatisches Denken. *Mathematische Annalen*, 78:405–15, 1918. Lecture given at the Swiss Society of Mathematicians, 11 September 1917. Reprinted in [Hilbert, 1935, 146–56]. English translation in [Ewald, 1996, 1105–1115].

[Hilbert, 1918b] D. Hilbert. Prinzipien der Mathematik. Lecture notes by Paul Bernays. Winter-Semester 1917–18. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen. Published in [Hilbert, 2006], 1918.

[Hilbert, 1920] D. Hilbert. Probleme der mathematischen Logik. Vorlesung, Sommer-Semester 1920. Lecture notes by Paul Bernays and Moses Schönfinkel. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen. Published in [Hilbert, 2006], 1920.

[Hilbert, 1922a] D. Hilbert. Grundlagen der Mathematik. Vorlesung, Winter-Semester 1921–22. Lecture notes by Paul Bernays. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen. Published in [Hilbert, 2006], 1922.

[Hilbert, 1922b] D. Hilbert. Neubegründung der Mathematik: Erste Mitteilung. *Abhandlungen aus dem Seminar der Hamburgischen Universität*, 1:157–77, 1922. Series of talks given at the University of Hamburg, July 25–27, 1921. Reprinted with notes by Bernays in [Hilbert, 1935, 157–177]. English translation in [Mancosu, 1998a, 198–214] and [Ewald, 1996, 1115–1134].

[Hilbert, 1923] D. Hilbert. Die logischen Grundlagen der Mathematik. *Mathematische Annalen*, 88:151–165, 1923. Lecture given at the Deutsche Naturforscher-Gesellschaft, September 1922. Reprinted in [Hilbert, 1935, 178–191]. English translation in [Ewald, 1996, 1134–1148].

[Hilbert, 1926] D. Hilbert. Über das Unendliche. *Mathematische Annalen*, 95:161–90, 1926. Lecture given Münster, 4 June 1925. English translation in [van Heijenoort, 1967, 367–392].

[Hilbert, 1928] D. Hilbert. Die Grundlagen der Mathematik. *Abhandlungen aus dem Seminar der Hamburgischen Universität*, 6:65–85, 1928. English translation in [van Heijenoort, 1967, 464-479].

[Hilbert, 1929] D. Hilbert. Probleme der Grundlegung der Mathematik. *Mathematische Annalen*, 102:1–9, 1929. Lecture given at the International Congress of Mathematicians, 3 September 1928. English translation in [Mancosu, 1998a, 227–233].

[Hilbert, 1931a] D. Hilbert. Beweis des Tertium non datur. *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen. Math.-phys. Klasse*, pages 120–25, 1931.

[Hilbert, 1931b] D. Hilbert. Die Grundlegung der elementaren Zahlenlehre. *Mathematische Annalen*, 104:485–94, 1931. Reprinted in [Hilbert, 1935, 192–195]. English translation in [Ewald, 1996, 1148–1157].

[Hilbert, 1935] D. Hilbert. *Gesammelte Abhandlungen*, volume 3. Springer, Berlin, 1935.

[Hilbert, 1992] D. Hilbert. *Natur und mathematisches Erkennen*. Birkhäuser, Basel, 1992. Vorlesungen, 1919–20.

[Hilbert, 2004] D. Hilbert. *David Hilbert's Lectures on the Foundations of Geometry, 1891–1902*, Ulrich Majer and Michael Hallett, editors. Springer, New York, 2004.

[Hilbert, 2006] D. Hilbert. *David Hilbert's Lectures on the Foundations of Arithmetic and Logic, 1917–1933*, William Ewald and Wilfried Sieg, editors. Springer, New York, 2006. forthcoming.

[Hilbert and Ackermann, 1928] D. Hilbert and W. Ackermann. *Grundzüge der theoretischen Logik*. Springer, Berlin, 1928.

[Hilbert and Bernays, 1923] D. Hilbert and P. Bernays. Logische Grundlagen der Mathematik. Vorlesung, Winter-Semester 1922–23. Lecture notes by Paul Bernays, with handwritten notes by Hilbert. Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek, Cod. Ms. Hilbert 567. Published in [Hilbert, 2006], 1923.

[Hilbert and Bernays, 1934] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*, volume 1. Springer, Berlin, 1934.

[Hilbert and Bernays, 1939] D. Hilbert and P. Bernays. *Grundlagen der Mathematik*, volume 2. Springer, Berlin, 1939.

[Hofweber, 2000] T. Hofweber. Proof-theoretic reduction as a philosopher's tool. *Erkenntnis*, 53:127–146, 2000.

[Ignjatovič, 1994] A. Ignjatovič. Hilbert's program and the omega-rule. *Journal of Symbolic Logic*, 59:322–343, 1994.

[Jeroslow, 1971] R. G. Jeroslow. Consistency statements in formal theories. *Fundamenta Mathematicae*, 72:17–40, 1971.

[Jeroslow, 1975] R. G. Jeroslow. Experimental logics and Δ_2^0-theories. *Journal of Philosophical Logic*, 4:253–267, 1975.

[Kitcher, 1976] P. Kitcher. Hilbert's epistemology. *Philosophy of Science*, 43:99–115, 1976.

[Kreisel, 1951] G. Kreisel. On the interpretation of non-finitist proofs. Part I. *Journal of Symbolic Logic*, 16:241–267, 1951.

[Kreisel, 1954] G. Kreisel. A variant to Hilbert's theory of the foundations of arithmetic. *British Journal for the Philosophy of Science*, 4:107–129, 1954.

[Kreisel, 1960] G. Kreisel. Ordinal logics and the characterization of informal notions of proof. In J. A. Todd, editor, *Proceedings of the International Congress of Mathematicians. Edinburgh, 14–21 August 1958*, pages 289–299, Cambridge, 1960. Cambridge University Press.

[Kreisel, 1968] G. Kreisel. A survey of proof theory. *Journal of Symbolic Logic*, 33:321–388, 1968.

[Kreisel, 1970] G. Kreisel. Principles of proof and ordinals implicit in given concepts. In A. Kino, J. Myhill, and R. E. Veseley, editors, *Intuitionism and Proof Theory*, pages 489–516. North-Holland, Amsterdam, 1970.

[Kreisel, 1971] G. Kreisel. Review of *The Collected Papers of Gerhard Gentzen* [Gentzen, 1969]. *Journal of Philosophy*, 68:238–265, 1971.

[Kreisel, 1976] G. Kreisel. Wie die Beweistheorie zu ihren Ordinalzahlen kam und kommt. *Jahresberichte der Deutschen Mathematiker-Vereinigung*, 78:177–223, 1976.

[Kreisel, 1983] G. Kreisel. Hilbert's programme. In Benacerraf and Putnam [1983], pages 207–238.

[Mancosu and Ryckman, 2002] P. Mancosu and T. Ryckman. Mathematics and phenomenology: The correspondence between O. Becker and H. Weyl. *Philosophia Mathematica*, 10:130–202, 2002.

[Mancosu, 1998a] P. Mancosu, editor. *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s*. Oxford University Press, Oxford, 1998.

[Mancosu, 1998b] P. Mancosu. Hilbert and Bernays on metamathematics. In *From Brouwer to Hilbert* [1998a], pages 149–188.

[Mancosu, 1999] P. Mancosu. Between Russell and Hilbert: Behmann on the foundations of mathematics. *Bulletin of Symbolic Logic*, 5(3):303–330, 1999.

[Mancosu, 2003] P. Mancosu. The Russellian influence on Hilbert and his school. *Synthèse*, 137(1–2):59–101, 2003.

[Mancosu, 2004] P. Mancosu. Review of Kurt Gödel, *Collected Works*, vols. IV and V, Solomon Feferman, et al., eds. Oxford: Oxford University Press, 2003. *Notre Dame Journal of Formal Logic*, 45:109–125, 2004.

[Mints and Tupailo, 1999] G. Mints and S. Tupailo. Epsilon-substitution method for the ramified language and Δ_1^1-comprehension rule. In Andrea Cantini et al., editors, *Logic and Foundations of Mathematics*, pages 107–130. Kluwer, Dordrecht, 1999.

[Moore, 1997] G. H. Moore. Hilbert and the emergence of modern mathematical logic. *Theoria (Segunda Época)*, 12:65–90, 1997.

[Müller, 1923] A. Müller. Über Zahlen als Zeichen. *Mathematische Annalen*, 90:153–158; 163, 1923.

[Negri, 1980] M. Negri. Constructive sequent reduction in Gentzen's first consistency proof for
    arithmetic. In Maria Luisa Dalla Chiara, editor, *Italian Studies in the Philosophy of Science*,
    pages 153–168. Reidel, Dordrecht, 1980.
[Paris and Harrington, 1977] J. Paris and L. Harrington. A mathematical incompleteness in
    Peano arithmetic. In Barwise [1977], pages 1133–1142.
[Parsons, 1998] C. Parsons. Finitism and intuitive knowledge. In Matthias Schirn, editor, *The
    Philosophy of Mathematics Today*, pages 249–270. Oxford University Press, Oxford, 1998.
[Peckhaus, 1990] V. Peckhaus. *Hilbertprogramm und Kritische Philosophie*. Vandenhoeck und
    Ruprecht, Göttingen, 1990.
[Pohlers, 1987] W. Pohlers. Contributions of the Schütte school school in Munich to proof
    theory, 1987. Appendix to [Takeuti, 1987].
[Pohlers, 1998] W. Pohlers. Subsystems of set theory and second-order number theory. In Buss
    [1998].
[Poincaré, 1906] H. Poincaré. Les mathématiques et la logique. *Revue de métaphysique et de
    morale*, 14:294–317, 1906. English translation in [Ewald, 1996, 1038–1052].
[Pudlák, 1998] P. Pudlák. The length of proofs. In Buss [1998].
[Raatikainen, 2003] Panu Raatikainen. Hilbert's program revisited. *Synthèse*, 137:157–177,
    2003.
[Ramsey, 1926] F. P. Ramsey. Mathematical logic. *The Mathematical Gazette*, 13:185–94, 1926.
    Reprinted in [Ramsey, 1990, 225–244].
[Ramsey, 1990] F. P. Ramsey. *Philosophical Papers*, D. H. Mellor, editor. Cambridge University
    Press, Cambridge, 1990.
[Rathjen, 2005a] M. Rathjen. An ordinal analysis of parameter free Π_2^1-comprehension. *Archive
    for Mathematical Logic*, 44:263–362, 2005.
[Rathjen, 2005b] M. Rathjen. An ordinal analysis of stability. *Archive for Mathematical Logic*,
    44:1–62, 2005.
[Resnik, 1974a] M. D. Resnik. The Frege-Hilbert controversy. *Philosophy and Phenomenological
    Research*, 34:386–403, 1974.
[Resnik, 1974b] M. D. Resnik. On the philosophical significance of consistency proofs. *Journal
    of Philosophical Logic*, 3:133–47, 1974.
[Resnik, 1980] M. D. Resnik. *Frege and the Philosophy of Mathematics*. Cornell University
    Press, Ithaca, 1980.
[Russell, 1902] B. Russell. Letter to Frege, June 16, 1902. In van Heijenoort [1967], pages
    124–125.
[Schütte, 1960] K. Schütte. *Beweistheorie*. Springer, Berlin, 1960.
[Shanker, 1988] S. G. Shanker. *Gödel's Theorem in Focus*. Routledge, London, 1988.
[Sieg, 1990] W. Sieg. Reflections on Hilbert's program. In Wilfried Sieg, editor, *Acting and
    Reflecting*, pages 171–82. Kluwer, Dordrecht, 1990.
[Sieg, 1999] W. Sieg. Hilbert's programs: 1917–1922. *Bulletin of Symbolic Logic*, 5(1):1–44,
    1999.
[Sieg, 2002] W. Sieg. Beyond Hilbert's reach? In David B. Malament, editor, *Reading Natural
    Philosophy. Essays in the History and Philosophy of Science and Mathematics*, pages 363–
    405. Open Court, Chicago and La Salle, Ill., 2002.
[Simpson, 1988] S. G. Simpson. Partial realizations of Hilbert's program. *Journal of Symbolic
    Logic*, 53(2):349–363, 1988.
[Simpson, 1999] S. G. Simpson. *Subsystems of Seond Order Arithmetic*. Springer, Berlin, 1999.
[Simpson, 2005] S. G. Simpson. *Reverse Mathematics 2001*. Lecture Notes in Logic 21. Associ-
    ation for Symbolic Logic and A K Peters, Wellesley, Mass., 2005.
[Smoryński, 1977] C. Smoryński. The incompleteness theorems. In Barwise [1977], pages 821–
    865.
[Stein, 1988] Howard Stein. *Logos*, logic, and *logisitiké*: Some philosophical remarks on
    nineteenth-century transformation of mathematics. In William Aspray and Philip Kitcher,
    editors, *History and Philosophy of Modern Mathematics*, volume 11 of *Minesota Studies in
    the Philosophy of Science*, pages 238–259. University of Minnesota Press, 1988.
[Steiner, 1975] M. Steiner. *Mathematical Knowledge*. Cornell University Press, Ithaca, 1975.
[Steiner, 1991] M. Steiner. Review of *Hilbert's Program: An Essay on Mathematical Instru-
    mentalism* [Detlefsen, 1986]. *Journal of Philosophy*, 88(6):331–336, 1991.
[Tait, 1981] W. W. Tait. Finitism. *Journal of Philosophy*, 78:524–546, 1981. Reprinted in [Tait,
    2005b, 21–42].

[Tait, 2002] W. W. Tait. Remarks on finitism. In Wilfried Sieg, Richard Sommer, and Carolyn
Talcott, editors, *Reflections on the Foundations of Mathematics. Essays in Honor of Solomon
Feferman*, Lecture Notes in Logic 15. Association for Symbolic Logic and A K Peters, 2002.
Reprinted in [Tait, 2005b, 43–53].

[Tait, 2005a] W. W. Tait. Gödel's reformulation of Gentzen's first consistency proof for arith-
metic: the no-counterexample interpretation. *Bulletin of Symbolic Logic*, 11:225–238, 2005.

[Tait, 2005b] W. W. Tait. *The Provenance of Pure Reason. Essays in the Philosophy of Math-
ematics and Its History*. Oxford University Press, New York, 2005.

[Takeuti, 1987] G. Takeuti. *Proof Theory*. Studies in Logic 81. North-Holland, Amsterdam, 2nd
edition, 1987.

[van Heijenoort, 1967] J. van Heijenoort, editor. *From Frege to Gödel. A Source Book in Math-
ematical Logic, 1897–1931*. Harvard University Press, Cambridge, Mass., 1967.

[Visser, 1989] A. Visser. Peano's smart children: A provability logical study of systems with
built-in consistency. *Notre Dame Journal of Formal Logic*, 30:161–196, 1989.

[von Neumann, 1927] J. von Neumann. Zur Hilbertschen Beweistheorie. *Mathematische Zeit-
schrift*, 26:1–46, 1927.

[von Neumann, 1931] J. von Neumann. Die formalistische Grundlegung der Mathematik. *Erken-
ntnis*, 2:116–34, 1931. English translation in: [Benacerraf and Putnam, 1983, 61–65].

[Webb, 1997] J. C. Webb. Hilbert's formalism and arithmetization of mathematics. *Synthèse*,
110:1–14, 1997.

[Weyl, 1921] Hermann Weyl. Über die neue Grundlagenkrise der Mathematik. *Mathematis-
che Zeitschrift*, 10:37–79, 1921. Reprinted in [Weyl, 1968, 143-180]. English translation in
[Mancosu, 1998a, 86–118].

[Weyl, 1925] H. Weyl. Die heutige Erkenntnislage in der Mathematik. *Symposion*, 1:1–23, 1925.
Reprinted in: [Weyl, 1968, 511–42]. English translation in: [Mancosu, 1998a, 123–42].

[Weyl, 1928] H. Weyl. Diskussionsbemerkungen zu dem zweiten Hilbertschen Vortrag über die
Grundlagen der Mathematik. *Abhandlungen aus dem Mathematischen Seminar der Univer-
sität Hamburg*, 6:86–88, 1928. English translation in [van Heijenoort, 1967, 480-484].

[Weyl, 1968] H. Weyl. *Gesammelte Abhandlungen*, volume 1, K. Chandrasekharan, editor.
Springer Verlag, Berlin, 1968.

[Whitehead and Russell, 1910] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol-
ume 1. Cambridge University Press, Cambridge, 1910.

[Whitehead and Russell, 1912] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol-
ume 2. Cambridge University Press, Cambridge, 1912.

[Whitehead and Russell, 1913] A. N. Whitehead and B. Russell. *Principia Mathematica*, vol-
ume 3. Cambridge University Press, Cambridge, 1913.

[Zach, 1999] R. Zach. Completeness before Post: Bernays, Hilbert, and the development of
propositional logic. *Bulletin of Symbolic Logic*, 5(3):331–366, 1999.

[Zach, 2003a] R. Zach. Hilbert's program. In Edward N. Zalta, editor, *The Stanford
Encyclopedia of Philosophy*. Fall 2003. `http://plato.stanford.edu/archives/entries/
hilbert-program/`.

[Zach, 2003b] R. Zach. The practice of finitism. Epsilon calculus and consistency proofs in
Hilbert's Program. *Synthèse*, 137:211–259, 2003.

[Zach, 2004] R. Zach. Hilbert's "Verunglückter Beweis," the first epsilon theorem and consis-
tency proofs. *History and Philosophy of Logic*, 25:79–94, 2004.