Hilbert's Finitism: Historical, Philosophical, and Metamathematical Perspectives

by

Richard Zach

Dipl.-Ing. (University of Technology, Vienna, Austria) 1993
M.A. (University of California, Berkeley) 1997
C.Phil. (University of California, Berkeley) 1997

A dissertation submitted in partial satisfaction of the
requirements for the degree of
Doctor of Philosophy

in

Logic and the Methodology of Science

in the

GRADUATE DIVISION
of the
UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Paolo Mancosu, Cochair
Professor Jack H. Silver, Cochair
Professor Barry Stroud

Spring 2001

The dissertation of Richard Zach is approved:

Cochair Date

Cochair Date

Date

University of California, Berkeley

Spring 2001

Hilbert's Finitism: Historical, Philosophical, and Metamathematical Perspectives

Abstract

Hilbert's Finitism: Historical, Philosophical,
and Metamathematical Perspectives

by

Richard Zach
Doctor of Philosophy in Logic and the Methodology of Science

University of California, Berkeley

Professor Paolo Mancosu, Cochair
Professor Jack H. Silver, Cochair

In the 1920s, David Hilbert proposed a research program with the aim of providing mathematics with a secure foundation. This was to be accomplished by first formalizing logic and mathematics in their entirety, and then showing—using only so-called finitistic principles—that these formalizations are free of contradictions.

In the area of logic, the Hilbert school accomplished major advances both in introducing new systems of logic, and in developing central metalogical notions, such as completeness and decidability. The analysis of unpublished material presented in Chapter 2 shows that a completeness proof for propositional logic was found by Hilbert and his assistant Paul Bernays already in 1917–18, and that Bernays's contribution was much greater than is commonly acknowledged. Aside from logic, the main technical contribution of Hilbert's Program are the development of formal mathematical theories and proof-theoretical investigations thereof, in particular, consistency proofs. In this respect Wilhelm Ackermann's 1924 dissertation is a milestone both in the development of the Program and in proof theory in general. Ackermann gives a consistency proof for a second-order version of primitive recursive arithmetic which, surprisingly, explicitly uses a finitistic version of transfinite induction up to $\omega^{\omega^{\omega}}$. He also gave a faulty consistency proof for a system of second-order arithmetic based on Hilbert's ε-substitution method. Detailed analyses of both proofs in

Chapter 3 shed light on the development of finitism and proof theory in the 1920s as practiced in Hilbert's school.

In a series of papers, Charles Parsons has attempted to map out a notion of mathematical intuition which he also brings to bear on Hilbert's finitism. According to him, mathematical intuition fails to be able to underwrite the kind of intuitive knowledge Hilbert thought was attainable by the finitist. It is argued in Chapter 4 that the extent of finitistic knowledge which intuition can provide is broader than Parsons supposes. According to another influential analysis of finitism due to W. W. Tait, finitist reasoning coincides with primitive recursive reasoning. The acceptance of non-primitive recursive methods in Ackermann's dissertation presented in Chapter 3, together with additional textual evidence presented in Chapter 4, shows that this identification is untenable as far as Hilbert's conception of finitism is concerned. Tait's conception, however, differs from Hilbert's in important respects, yet it is also open to criticisms leading to the conclusion that finitism encompasses more than just primitive recursive reasoning.

Es ist gar nicht richtig, daß der Forscher der Wahrheit nachstellt, sie stellt ihm nach. Er erleidet sie.

— Robert Musil

Contents

Acknowledgements

The first debt of gratitude I must acknowledge is to my parents Wilhelm and Waltraud and my sister Ruth, without whose support and encouragement I probably would not chosen an academic career, let alone come this far.

My teachers at Vienna and Berkeley have shaped my thinking and helped me improve my work more than anyone else. I am indebted in particular to my teachers at Berkeley, foremost among them Charles Chihara, Christos Papadimitriou, Hans Sluga, Barry Stroud, and Hugh Woodin. But above all I must thank my advisors Paolo Mancosu and Jack Silver, without whose support, criticism, praise, and insight this thesis would not be what it is now. Among my teachers at the University of Technology, Vienna, I have to mention Matthias Baaz, Chris Fermüller, Georg Gottlob and Alex Leitsch, who have taught me logic, and that philosophical and methodological reflection are as important as technical sophistication. I am also grateful to Solomon Feferman and Grisha Mints, current colleagues of sorts at Stanford, who have played a not insignificant part in my graduate studies, and my student colleagues at Berkeley and Vienna, specifically Eddie Cushman, Johannes Hafner, Ben Hansen, Christine Heitsch, Georg Moser, Chris Pincock, and Helmut Veith. Special thanks goes to Catalina Córdoba Boniffaccini, my mom away from home.

Chapter 2 appeared as “Completeness before Post: Bernays, Hilbert, and the development of propositional logic,” *Bulletin of Symbolic Logic* **5** (1999) 331–366, and is ©1999 by the Association of Symbolic Logic. It is reprinted here with the kind permission of the ASL.

Versions of this paper were presented at a symposium on *The Development of Modern Logic* in Helsinki, June 1998, at the Townsend Center Working Group on History and Philosophy of Logic and Mathematics at the University of California, Berkeley, at the Stanford Logic Lunch, and at the 1999 Spring Meeting of the ASL. I would like to thank in particular Jeremy Avigad, Martin Davis, Solomon Feferman, Georg Kreisel, Grigori Mints, Volker Peckhaus, Chris Pincock, Wilfried Sieg, and the referee for the BSL for their helpful comments. My thanks also goes to the Wissenschaftshistorische Sammlungen of the ETH Zürich Library, the Handschriftenabteilung of the Niedersächsische Staats- und Universitätsbibliothek, the library of the Institute of Mathematics at the University of Göttingen, and to the curators of the Heinrich Behmann Papers in Erlangen, Christian Thiel and Volker Peckhaus, for their help in my research and permission to quote from unpublished material.

Parts of Chapter 3 were presented to the Townsend Center Working Group in History and Philosophy of Logic and Mathematics at the University of Caliornia, Berkeley, December 1999, at a workshop on Hilbert at the Institut d'Histoire des Sciences et Techniques of the Université Paris I (partially funded by the Townsend Center through an International Working Group grant), May 2000, at the *International Symposium on the History of Logic* at the University of Helsinki, June 2000, at the Department of Logic and Philosophy of Science of the University of California, Irvine, and at the Philosophy of Mathematics Workshop at UCLA in 2001. I am indebted to Hans Richard Ackermann for providing me with copies of Wilhelm Ackermann's correspondence, and to Wilfried Sieg for providing me with Kneser's notes to the 1921–22 and 1922–23 lectures (Hilbert 1922a, Hilbert and Bernays 1923a). I would also like to thank in particular Aldo Antonelli, Solomon Feferman, Ivor Grattan-Guinness, David Kaplan, Penelope Maddy, Tony Martin, Yiannis Moschovakis, and Grigori Mints for valuable suggestions. My understanding of the ε-substitution method in general and Ackermann's proof in particular has benefitted greatly from conversations with and writings of Grigori Mints, Georg Moser, and W. W. Tait. Any errors in my presentation are, of course, mine.

Parts of Chapter 4 appeared as "Numbers and functions in Hilbert's finitism," *Taiwanese Journal for Philosophy and History of Science* **10** (1998) 33–60, and are ©1998 by the Taiwanese Journal for Philosophy and History of Science. They are reprinted here with the kind permission of the editors of the Taiwanese Journal for Philosophy and History of Science.

That paper grew out of a term paper for Charles Chihara's course on philosophy of mathematics, taught at Berkeley in Spring 1996. Versions of it were presented at Stanford–Berkeley Philosophy Conference at Stanford and at the Collegium Logicum of the Kurt Gödel Society, Vienna, in Spring 1997. I am indebted to Charles Chihara for many insightful comments and suggestions, and for the invitation to publish the paper in the *Taiwanese Journal*. My thanks also goes to Thomas Hofweber, who commented on the paper in the Stanford–Berkeley Conference. Versions of Chapter 4 were subsequently also presented at the Colloquium Logico-Philosophicum, Universität Erlangen–Nürnberg in 1998, at the workshop *Collegium Logicum 1999: Proof Theory* in Vienna, at Stanford University, the University of Chicago, and Oxford University in 2000. I would like to thank in particular Lanier Anderson, Solomon Feferman, Chris Fermüller, Daniel Isaacson, Georg Kreisel, Grigori Mints, Volker Peckhaus, Bill Tait, and Gaisi Takeuti for their comments.

My graduate studies and the writing of this dissertation have received financial support from the Group in Logic and the Methodology of Science, the University of California through a Berkeley Fellowship for Graduate Study, the Austrian Ministry for Science and Research through a Kurt Gödel Fellowship, and the Mabelle McLeod Lewis Fund through a McLeod Lewis Memorial Fellowship. Travel to various conferences to present this work and for archival research was made possible by grants from the Graduate Division, the Department of Mathematics, and the Townsend Center for the Humanities at Berkeley. I gratefully acknowledge their generous support.

Shoutouts

Anna, Hermione, Yoni, Jill, Heidi, Static, Jendon, Zai and everyone at Etoy, Paul, Johnny Mads, Evany Evany Thomas, Lisa, Mozilla, Liz, Schatz, Lib, Jason, Eddie, Sunny, Arye, Ulli, Miry, Rochus, Homay, Barb, Max, Evi, Cheu, Brandy, Dalva, Halliday, New Wave City, Cat, the 500, Matmos, Amy, Gayle, Zimmer, Alison and everyone at the Uptown, Helmut, Georg, Vincent, Caroline, Jessica Louise, Ben, Jason, Marie, Sandy, Matt, Christine, Mandana, Sibylle, Brian, Dan, Justine, Chris, Jennifer, Dave, Tina, Val, Anthony, Laura, Betty, Lynn, Justine, Elaine, everyone at the SFMOMA, Mike, and the Grady Sisters.

Chapter 1

Introduction

1.1 David Hilbert and the Foundations of Mathematics

David Hilbert was born on January 23, 1862 in Königsberg (now Kaliningrad). He studied at Königsberg from 1880 to 1884, receiving his doctorate in 1885 under Friedrich Lindemann. He received the *venia docendi* in 1886, in 1892 was appointed *Ausserordentlicher Professor* and finally became *Ordinarius* in 1893. In 1895, he accepted a chair at the University of Göttingen, where he remained until his retirement in 1930. He died in 1943 in Göttingen.

Freudenthal (1973) distinguishes six parts in Hilbert's mathematical work. Until 1893 Hilbert worked on the theory of algebraic forms, 1894–99 on algebraic number theory, 1899–1903 on the foundations of geometry, 1904–1909 on analysis, 1912–1914 on theoretical physics, and after 1918 on the foundations of mathematics. This division, however, obscures the fact that Hilbert was interested in foundational questions almost continuously from the mid-1890s onwards. His work on axiomatic geometry, culminating in the publication of the extremely influential *Grundlagen der Geometrie* (1899), marked the beginning of an area of interest, and a research program, which he continued to pursue till the end of his career: the axiomatic method. In contrast to the Euclidean conception of axiomatics, according to which the axioms expressed intuitive truths about space, Hilbert took the aim of axiomatics to provide implicit definitions of the geometrical concepts.[1] As such, the main value of axiomatics is that it lays bare the logical relationships between the concepts so defined and between the various axioms. Since for Hilbert, the antecedently given intended model of the axiomatic theory loses relevance for the axiomatic development, the questions of completeness, independence, and consistency of the axioms are of primary importance. For geometry, the problem of consistency can be solved by providing an arithmetical (in fact, analytical) interpretation of the axioms, thus reducing the question of consistency to the consistency of the arithmetic of the reals. The consistency of arithmetic (in this general sense) is then the ultimate problem for the foundation of mathematics.

It is not surprising then that the question of providing a non-reductive consistency proof for arithmetic was the second problem Hilbert posed as one of his catalog of open problems in mathematics, presented at the International Congress of Mathematicians in Paris in 1900 (1900a). And indeed his interest in the axiomatic method, and in foundational questions such as consistency and completeness never subsided. It was only overshadowed, as it were, by more specific interests such as theoretical physics. But even there, the methodology was very much in line with his foundational outlook. Not only did Hilbert himself consider complete, consistent axiomatizations the crowning achievements of scientific disciplines, he actively encouraged work in this direction by his students. In particular, he took an active interest in the work of Zermelo—who taught in Göttingen between 1899 and 1910—on the axiomatization of set theory, and in the axiomatization of physics.[2]

Hilbert's work on the foundations of geometry, the debate with Frege and the development of set theory and the discovery of the set-theoretical paradoxes made it clear to Hilbert that a closer investigation of the foundations of mathematics was needed. The first order of business was, of course, the consistency proof for arithmetic. Aside from the "Problems" lecture, Hilbert had explicitly formulated this as a desideratum in his (1900b). After the impact of the paradoxes, however, it became clear that such a proof required a deeper understanding of logic. In 1904, he writes:

> It is my opinion that all the difficulties touched upon [i.e., the paradoxes] can be overcome and that we can provide a rigorous and completely satisfying foundation for the notion of number, and in fact by a method hat I would call *axiomatic* and whose fundamental idea I wish to develop briefly in what follows.
>
> Arithmetic is often considered to be part of logic, and the traditional fundamental logical notions are usually presupposed when it is a question of es-

> tablishing 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 arithmetic is required if paradoxes are to be avoided.[3]

Hilbert continues in this paper to give a first formal calculus for number theory which represents numbers as strings of 1's, and sketches a direct consistency proof for it. A first attempt at a development of logic can be found in a lecture course the following year (Hilbert 1905a) (See Chapter 2). Subsequently he 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.[4] He also continued to teach classes on issues related to foundational questions. The list of courses Hilbert gave in Göttingen up to 1917 include the following:[5]

1. Logische Principien des mathematischen Denkens (Summer 1905) (Hilbert 1905a, 1905b)
2. Zahlbegriff und Prinzipienfragen der Mathematik (Summer 1908) (Hilbert 1908)
3. Prinzipien der Mathematik (Seminar, Winter 1908/09)
4. Elemente und Prinzipienfragen der Mathematik (Summer 1910) (Hilbert 1910)
5. Logische Grundlagen der Mathematik (Winter 1911/12)[6]
6. Grundlagen der Mathematik und Physik (Summer 1913) (Hilbert 1913)
7. Probleme und Prinzipien der Mathematik (Winter 14/15)
8. Mengenlehre (Winter 1916/17)
9. Mengenlehre (Summer 1917) (Hilbert 1917)

While foundational questions were clearly on Hilbert's mind between 1905 and 1917,[7] his published work in those years centered mostly on integral equations and theoretical physics (from 1912 onwards). It was not until 1917 that he returned to his foundational investigations with full force. Weyl (1944) put the motivation for this shift thus:

> One hears a lout rumbling of the old problems in his Zürich address, *Axiomatisches Denken* [(Hilbert 1918a, 1918b)], of 1917. Meanwhile the difficulties concerning the foundations of mathematics had reached a critical stage, and the situation cried for repair. Under the impact of undeniable antinomies in set theory, Dedekind and Frege had revoked their own work on the nature of numbers and arithmetical propositions, Bertrand Russell had pointed out the hierarchy of types which, unless one decides to "reduce" them by sheer force, undermine the arithmetical theory of the continuum; and finally L. E. J. Brouwer by his intuitionism had opened our eyes and made us see how far generally accepted mathematics goes beyond such statements as can claim real meaning and truth founded on evidence. [. . .]
>
> Hilbert was not willing to make the heavy sacrifices which Brouwer's standpoint demanded, and he saw, at least in outline, a way by which the cruel mutilation could be avoided. At the same time he was alarmed by signs of wavering loyalty within the ranks of mathematicians, some of whom openly sided with Brouwer. My own article on the *Grundlagenkrise* [Weyl (1921)], written in the excitement of the first postwar years in Europe, is indicative of the mood. Thus Hilbert returns to the problem of foundations in earnest. He is convinced that complete certainty can be restored without "committing treason to our science."

Hilbert's reaction, and his work on the foundations of mathematics after 1917 can be divided into two parts: The first part, executed between 1917 and 1924, was a thorough investigation of logic. The second part is Hilbert's Program proper: thwarting the intuitionist *Putsch* by carrying out his goals of 1900–05 of giving a direct consistency proof of arithmetic.

1.2 Hilbert, Bernays, and Logic

It is clear that no attempt at giving rigorous proofs of consistency or completeness of axiomatic systems can succeed without a clear understanding of the nature and status of logical inference. For, to say that something is or is not derivable from the axioms, the language and the rules of inference have to be laid out. When Hilbert began his work on axiomatics, formal logical systems in the tradition of Frege–Peano–Russell were in their infancy and certainly not widely accepted in mathematical circles. So in *Foundations of Geometry*, while making significant conceptual advances in axiomatic geometry, Hilbert formulated axioms and theorems in German. A first interest in logical notation arises in his work of 1904 and 1905; the notations and systems used there however stand in the algebraic tradition (in particular, Schröder was a main influence). These systems were hardly adequate for Hilbert's aims.[8] It was not until Russell and Whitehead's *Principia* (1910, 1913) became known in Göttingen around 1914 that a formal framework suited to Hilbert's aims was available.[9]

Hilbert adopted Russell's framework and notation in his lectures on *Prinzipien der Mathematik* during the Winter Semester of 1917–18. The preceding Fall, Hilbert had been joined in Göttingen by Paul Bernays as his assistant. With Hilbert's metatheoretical outlook already in place since his early work on axiomatics, it was only natural to treat the logical axiomatics of *Principia* like any other axiomatic system. And, as for any other axiomatic system, the prime concerns were consistency and completeness. Russell, as did his predecessors Frege and Peano, considered formal logic to be a language which, free from the imprecision of natural language, allowed the logically perfect formulation of mathematics. As it was seen as a universal language, however, it was not a consideration for Russell to treat the system itself as an object of investigation of the kind that Hilbert desired. It was thus a significant break with and advance over the logical tradition to even pose questions such as completeness for a logical system. The formulation of the question of completeness for the logical formalism of *Principia* thus marks a milestone in the development of modern logic. It fell to Bernays to solve the problem thus posed, to prove the completeness theorem for propositional logic.

In fact, two different notions of completeness were involved. One was the syntactic notion which now is known as Post completeness: a system is complete, if the addition of any underivable sentence renders the system inconsistent.[10] Bernays took this a step further and saw that the propositional calculus admits a very simple interpretation—a semantics—based on truth values. The corresponding notion of completeness is then the usual one: any sentence which is true under every truth-value assignment (every *valid* sentence) is derivable from the axioms.

In Chapter 2 below, I trace the development of Hilbert's logical thinking from its beginnings in 1905 through the study of propositional calculi in Hilbert's school in the 1920s. The main contribution is certainly the one just mentioned: truth-value semantics for propositional logic and the proof of completeness. Abrusci (1989), Moore (1997), and Sieg (1999) have already pointed out the importance of the 1917–18 lecture notes in the development of logic in Hilbert's school and in the development of first-order logic in the 1920s in general, and Sieg in particular has shown how these lectures were the basis of Hilbert and Ackermann's *Grundzüge der Theoretischen Logik* of 1928. My contribution in Chapter 2 consists mainly in giving a more detailed account of the results on propositional logic obtained by Hilbert and Bernays, in particular in Bernays's *Habilitationsschrift* of 1918, but goes beyond earlier work both in terms of the depth with which the material is covered and in scope. 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 in the syntactic sense does not hold for first-order logic (an early result due to Ackermann), a development of model theory of first-order logic was needed first. The decision problem, one of Hilbert's main aims for metamathematics in the 1920s, was already at 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 Ackermann (1928b), Behmann (1921), and Bernays and Schönfinkel (1928) working on special cases throughout the 1920s.[11] In this line of research of the Hilbert school, Bernays deserves credit for proving for the first time, in his (1918), the decidability of an axiomatic logical calculus, viz., the propositional fragment of *Principia*.[12]

When Hilbert came to propose, in 1921 (Hilbert 1922c), his program for a proof-theoretic foundation of mathematics, the beginnings of a logical calculus suitable for the formalization of mathematics and a framework for metamathematical investigations thereof were already in place. Two things needed to be done, still: a separation in the axiomatic system of the finitistically unproblematical connectives (conjunction, disjunction, implication) from the problematic negation, and an extension to first- and higher-order systems. The former developments were carried out by Bernays (and are outlined in Chapter 2), the latter resulted in the development of the ε-formalism. This development is traced in the first half of Chapter 3, where I show how, on the one hand, the rejection of Russell's logicist approach to the foundations of mathematics by Hilbert around 1920 and the influence of Brouwer's criticisms of the law of the excluded middle and the axiom of choice (in its classical meaning) for infinite totalities lead Hilbert to include a logical version of the choice principle in his systems. This logical choice principle is encoded in the ε-calculus: $\varepsilon_x A(x)$ chooses a witness x for which $A(x)$ is true. This development was crucial both for proof theory, as systems based on the ε-calculus became the testing ground for Hilbert's quest for consistency proofs. The main advantage of the ε-calculus is that it not only allows definition of the quantifiers (and thus extends the systems to higher orders) in an almost logic-free manner (i.e., the quantifiers are represented by terms), but it also allows the formulation of the principle of induction in a way that was amenable to Hilbert's proof-theoretic approach.

1.3 Hilbert's Proof-Theoretical Program

Hilbert's proposal for a proof-theoretic grounding of mathematics was prompted, as already noted, by Brouwer's attacks on classical mathematics and Weyl's subsequent conversion to intuitionism.[13] The force with which Weyl, Hilbert's star student, condemned classical mathematics and endorsed intuitionism is conveyed clearest in the opening paragraph of (1921):

> The antinomies of set theory are usually treated as border conflicts concerning only the most remote provinces of the mathematical realm, and as in no way endangering the inner soundness and security of the realm and its proper core provinces. The statements on these disturbances of the peace that authoritative sources have given (with the intention to deny or mediate) mostly do not have the character of a conviction born out of thoroughly investigated evidence that rests firmly on itself. Rather, they belong to the sort of one-half to three-quarters honest attempts of self-delusion that are so common in political and philosophical thought. Indeed, any sincere and honest reflection has to lead to the conclusion that these inadequacies in the border provinces of mathematics must be counted as symptoms. They reveal what is hidden by the outwardly shining and frictionless operation in the center: namely, the inner groundlessness of the foundations upon which rests the superstructure of the realm.

Weyl's and Brouwer's proposed solutions to this "groundlessness" of mathematics, however, were out of the question for Hilbert, who famously characterized it as a "dictatorship of prohibitions" and suggested that restricting mathematics in the intuitionistic fashion was akin to denying a boxer the use of his fists. Hilbert's proposed alternative—which should "clear the good name of mathematics once and for all"—was a combination of the axiomatic method he had developed over the preceding 30 years, and an intuitive (in the sense of Kant) metamathematics, free from objectionable (impredicative or infinitistic) assumptions which would serve as the basis for consistency proofs of the axiomatic systems.

The old aim of Hilbert's foundational thought, formal axiomatics, was thus supplemented with a philosophical foundation which was to make the direct consistency proofs for arithmetic and analysis possible. This philosophical view is the "finitist standpoint". Hilbert first outlined his basic position in (1922c):

> As we saw, the abstract operation with general concept-scopes and contents has proved to be inadequate and uncertain. Instead, as a precondition for the application of logical inference and for the activation of logical operations, something must already be given in representation [*in der Vorstellung*]: certain 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 that cannot be reduced to something else. Because I take this standpoint, the objects [*Gegenstände*] of number theory are for me—in direct contrast to Dedekind and Frege—the signs themselves, whose shape [*Gestalt*] can be generally and certainly recognized by us—independently of space and time, of the special conditions of the production of the sign, and of insignificant differences in the finished product [footnote:In this sense, I call signs of the same shape "the same sign" for short.] The solid philosophical attitude that I think is required for the grounding of pure mathematics—as well as for all scientific thought, understanding, and communication—is this: *In the beginning was the sign.*

It was applied by Hilbert and his students in the program of finding consistency proofs for mathematics. The distinction between the finitistic, contentual metamathematics and the formal, axiomatic development of mathematics itself is explained later in the same paper:

> [W]e can achieve an analogous point of view if we move to a higher level of contemplation, from which the axioms, formulae, and proofs of the math-

> ematical theory are themselves the objects of a contentual investigation. But for this purpose the usual contentual ideas of the mathematical theory must be replaced by formulae and rules, and imitated by formalisms. In other words, we need to have a strict formalization of the entire mathematical theory, inclusive of its proofs, so that—following the example of the logical calculus—the mathematical inferences and definitions become a formal part of the edifice of mathematics. The axioms, formulae, and proofs that make up this formal edifice are precisely what the number-signs were in the construction of elementary number theory which I described earlier; and with them alone, as with the number-signs in number theory, contentual thought takes place—i.e., only with them is actual thought practices. In this way the contentual thoughts (which of course we can never wholly do without or eliminate) are removed elsewhere—to a higher plane, as it were; and at the same time it becomes possible to draw a sharp and systematic distinction in mathematics between the formulae and formal proofs on the one hand, and the contentual ideas on the other.
>
> In the present paper my task is to show how this basic idea can be carried out in a rigorous and unobjectionable manner, and to show that our problem of proving the consistency of the axioms of arithmetic and analysis are thereby solved.

The program of formalization of mathematics and of consistency proofs was developed in a number of courses in the early 1920s, in particular:

1. Probleme der mathematischen Logik (Summer 1920) (Hilbert 1920b)
2. Grundlagen der Mathematik (Winter 1921–22) (Hilbert 1922b, 1922a)
3. Logische Grundlagen der Mathematik (Winter 1922–23), co-taught with Bernays (Hilbert and Bernays 1923b, 1923a)

In these courses, both the preferred axiomatics (using the ε-calculus) and the preferred method of consistency proofs (the ε-substitution method) were laid out. The work begun in there was carried on by Wilhelm Ackermann, who attempted to extend Hilbert's ideas for a consistency proof to all of analysis in his 1924 dissertation, "Grounding on the tertium non datur using Hilbert's theory of consistency" (1924a, 1924b).

Chapter 3 contains a thorough study of both the developments of axiomatic systems for arithmetic and analysis as they were carried out by Hilbert and Bernays in these courses, in particular the development of the ε-calculus and its use in axiomatizing mathematics, but also an in-depth evaluation of Ackermann's work. The ε-substitution method had been presented in a very basic form by Hilbert in the courses of 1921–22 and 1922–23 as well as in (1923), but it fell to Ackermann to make it work for the general case. The application of Hilbert's finitist standpoint in Ackermann's consistency proofs provides a deeper insight into what methods Hilbert considered to be available to the finitist. In his published writings, Hilbert never goes beyond the basic idea of numerals as sequences of strokes and the primitive ways of reasoning about them. It is only in the practice of finitism that the extent of finitist, contentual mathematics becomes evident. In the second half of Chapter 3, I show how Ackermann used a form of transfinite induction, which was nevertheless accepted as finitist, to prove the consistency of a second-order version of what is now known as primitive recursive arithmetic PRA. This theory permits formalization of the Ackermann function, the prime example of a non-primitive recursive function. Together with an analysis of passages in later writings of Bernays on the nature of finitistic reasoning, this shows that the Ackermann function and other, non-primitive recursive notions and methods of proof, fell within the finitist standpoint as originally conceived by Hilbert. This is interesting in the light of the characterization of finitism advanced by (Tait 1981), according to which finitism *coincides* with primitive recursive reasoning, and provides evidence that finitism as conceived by Hilbert is not restricted to primitive recursive methods.

The last section of Chapter 3 is devoted to an analysis of Ackermann's attempted consistency proof for analysis. This proof is based on Hilbert's ε-substitution method. The idea of this method is, roughly, that the ε-terms $\varepsilon_x A(x)$ occurring in a formal proof—which represent numerical witnesses and do the duty of quantifiers—are replaced by actual numbers, resulting in an essentially quantifier-free proof. The method by which the numerical substitutions for the ε-terms is found is simple enough for the basic cases considered by Hilbert, but soon becomes extremely complex. It is thus not surprising that Ackermann's proof only succeeds for a rather restricted fragment of arithmetic: in modern terminology, first-order arithmetic with induction restricted to open formulas. Ackermann claimed, however, that the proof goes through for a much larger fragment of analysis (essentially, elementary analysis with arithmetic comprehension). With the work of von Neumann (1927) and a study of the proof by Bernays, who was preparing a manuscript that was to become the monumental *Grundlagen der Mathematik* (Hilbert and Bernays 1934, 1939), gaps in the proof soon become evident. It is worth noting that Ackermann and Bernays considered the proof to be correct for the entire first-order fragment of arithmetic up until Gödel's incompleteness results became known in 1930 (Gödel 1931). These results showed that the program of consistency proofs cannot be carried out in the way Hilbert had envisaged.

1.4 Finitism

In the final chapter, I turn to an investigation of the finitist standpoint itself. This is the most obscure, yet also most philosophically influential part of Hilbert's position. The idea that there is a way in which knowledge of numerical equations is a privileged sort of mathematical knowledge resonates with many mathematicians and philosophers of mathematics, and so it is not surprising that finitism has remained an area of interest both for proof theorists and philosophers of mathematics.

Even though Gödel's results show that there is no way to give a consistency proof even for first-order arithmetic with the means available to the finitist, it has been claimed that *relative* consistency proofs using finitistic principles (in practice, primitive recursive principles) have philosophical import. Such projects of relative consistency proofs are the aim of work in so-called relativized Hilbert programs, examples of which abound in the recent proof-theoretic literature.[14] There is thus ample motivation for assessing the possible worth for the current debate on the philosophy of mathematics of a finitistic position.

Such an assessment has been attempted by a number of philosophers, foremost among them Kreisel (1960, 1958, 1965, 1970), Tait (1968, 1981, 2000) and Parsons (1979–80, 1994, 1998a).[15] All three approaches attempt a reconstruction of finitism, taking Hilbert's remarks as their starting point, and yield more or less definite characterizations of the strength of finitist reasoning. In Chapter 4, I focus on the approaches of Parsons and Tait. For Parsons, the defining characteristic of finitism is its intuitive nature. Using his own account of mathematical intuition, Parsons criticizes what he calls Hilbert's Thesis, that proofs according to the finitist method, in particular, primitive recursive proofs, yield intuitive knowledge of the theorems thus proved. His conclusion is mostly negative: intuition does not suffice to ground primitive recursive reasoning. I hope to have shown that Parsons' arguments for this negative result are at best inconclusive. It is in any case doubtful that an intuitive foundation is necessary for establishing the epistemic priority of finitist reasoning required by its application in relativized Hilbert programs. Tait has argued in fact, that Hilbert was simply wrong to believe that finitist reasoning corresponds to any kind of intuitive knowledge about numbers. For him, the special character of finitism lies in that its objects, constructions, and methods of proof are distinguished by being implicit in the notion of number itself and thus presupposed by all non-trivial mathematical reasoning about numbers. His thesis is that these constructions and methods of proof are exactly those which can be effected in primitive recursive arithmetic. I argue that this conception of finitism, although weaker than Parsons' and perhaps Hilbert's own, are all that is required for finitism to do the work it is supposed to do when it is invoked in recent proof-theoretical work, but argue against Tait that finitism surpasses primitive recursive reasoning. My argument is based both on the historical evidence from finitistic practice (as presented in the analysis of Ackermann's consistency proofs in Chapter 3, and textual evidence from Hilbert and Bernays presented in Chapter 4), and more conceptual arguments in favor of accepting nested recursion as finitistic.

Notes

1. Hilbert's view on axiomatics spurned an interesting debate with Frege, who held contra Hilbert that the axioms of geometry are propositional, i.e., express determinate truths. This of course was in conflict with Hilbert's view, famously expressed in his remark that "it must be possible to replace in all geometric statements the words *point*, *line*, *plane* with *table*, *chair*, *beermug*." For an analysis of the Frege-Hilbert debate, see Kambartel (1975) and Resnik (1974, 1980).

2. See Moore (1978) and (1982) for Zermelo's work on set theory and Corry (1997) for Hilbert's role in the axiomatization of physical theories.

3. Hilbert (1905c, p. 131).

4. On Hilbert's foundational interests before 1917, and his engagement for Husserl, Zermelo, and Nelson, see Peckhaus (1990).

5. Not all these courses were on genuinely foundational questions (logic, axiomatics). For instance, (1908) (and the first three quarters of (1910)) dealt with questions of constructibility with

compass and straightedge, and the quadrature of the circle.

6. No lecture notes for this course, or for those listed in (7) and (8) are available, but the classes were listed in the *Vorlesungsverzeichnisse* of the University.

7. See Sieg (1999), especially Part A, for more on the development of Hilbert's foundational interests in this period.

8. But compare the use Zermelo made of them in his (1908).

9. The influence of Russell's work was not merely in the area of notation. For the impact on Hilbert's philosophical views, mediated especially through his student Heinrich Behmann, see Mancosu (1999a, 200?).

10. The first occurrence of something like this notion was in Hilbert (1905c, p. 182).

11. The results of Bernays and Schönfinkel (1928) were already found several years before their publication. Schönfinkel was in Göttingen in the early part of the 1920s. A draft of the paper in Schönfinkel's hand is contained in the Bernays Nachlaß, ETH Zürich Library, WHS, Hs. 974.282. The manuscript lists a number of courses on the first page; these courses were given in the Winter Semester of 1923–23. It this respect it might also be worth pointing out that Schönfinkel (1924) was originally written in 1920—Schönfinkel spoke on the results to the Göttingen Mathematical Society on December 7, 1920—and prepared for publication by Behmann. A draft of the paper is contained in Behmann's *Nachlaß.*

12. Löwenheim (1915) showed that satisfiability of (a system equivalent to) monadic first-order logic is decidable; the result was not formulated, however, as a decidability result. Moreover, what was at issue for Hilbert was decidability of the question of *derivability* of a theorem in an axiomatic system.

13. Brouwer's views of the time were presented in his papers (1919, 1921), which were presented in 1919 and 1920, respectively. Weyl's conversion to intuitionism occurred in 1920, when he gave the talk on which (1921) was based at the University of Hamburg, July 28–30. The intuitionistic and predicativistic foundations of mathematics were widely discussed in Göttingen: already in 1917 (on November 20), Bernays reported on Weyl's *Das Kontinuum* (Weyl 1918), on May 11, 1920, Weyl himself spoke on the subject, and on February 1 and 8, 1921, Courant and Bernays gave talks on "the new arithmetical theories of Brouwer and Weyl." Two weeks later, on February 21 and 22, 1921, Hilbert presented his program to the mathematical audience in Göttingen under the title "Eine neue Grundlegung des Zahlbegriffs [A new grounding of the number concept]." (Announcements of these talks can be found in the *Jahresberichte der Deutschen Mathemaiker-Vereinigung*,

2. Abteilung.).

14. See in particular Sieg (1988, 1990), Feferman (1988, 1993b, 1993b) and Simpson (1988).

15. Detlefsen (1986) is of course a classic, but his aims are largely orthogonal to mine.

Chapter 2

Completeness before Post: Bernays, Hilbert, and the Development of Propositional Logic

2.1 Introduction

Paul Bernays is best known today for being Hilbert's primary collaborator on foundational matters in the Göttingen of the 1920s. He both shaped and helped execute the research project now known as Hilbert's program. The *Grundlagenbuch* (Hilbert and Bernays 1934, 1939), the decidability of the so-called Bernays-Schönfinkel class of first-order formulas (Bernays and Schönfinkel 1928), and his work on axiomatic set theory (Bernays 1958) are considered to be his major contributions to the foundations of mathematics. Bernays is also the author of a number of influential papers on philosophy of mathematics, and the details and refinements of Hilbert's mature philosophical views certainly owe much to him. His mathematical work in the early 1920s however, is little known and even less appreciated.

Bernays came to Göttingen in the Fall of 1917, at Hilbert's invitation.[1] For the following 17–odd years, Bernays worked in Göttingen as his assistant. His main task was to collaborate with Hilbert in his foundational work, in particular, to assist in the preparation of Hilbert's lecture courses and in preparing polished typescripts of these lectures.

Many of these lecture notes are preserved at the library of the Department of Mathematics at the University of Göttingen, and in Hilbert's *Nachlaß* at the Niedersächsische Staats- und Universitätsbibliothek. Hilbert's lectures have recently received much attention, since they provide a much more nuanced and detailed way of understanding the development not only of Hilbert's views on the foundations of mathematics, but on the development of first-order logic in the 20s. Moore (1997) and Sieg (1999) discuss, inter alia, the lecture notes for the course on the "Principles of Mathematics" (Hilbert 1918c).[2] I, too, want to focus on these notes, and on Bernays's *Habilitationsschrift* (Bernays 1918), of which only parts were published (1926). My central concern, however, shall be the results on propositional logic contained therein. These results include: explicit semantics for propositional logic using truth values, decidability of the set of valid propositional formulas, completeness of the axiom systems considered relative to that semantics, as well as what is now called Post completeness, consistency and independence results, general three- and four-valued matrices, and rule-based derivation systems.

All these results were obtained independently of logicians to whom they are usually credited (notably Pierce, Wittgenstein, Post, and Łukasiewicz).[3] Far be it from me to dispute their priority. After all, Hilbert and Bernays's work remained unpublished, and in some respects the work by those other logicians investigates the questions at hand more deeply or is more precise than Hilbert and Bernays's. I do think, however, that a detailed exposition of the results may provide clues to the development of logic in the 1920s, in particular in the Hilbert school.

While I believe that all of the results on propositional logic in question are interesting in their own right, some of my discussion also has significant bearing on the understanding of the development of first-order logic and Hilbert's foundational program as a whole. For instance, one of the conclusions of a close look at the historical record will be that the seminal early results on propositional *and first-order* logic were in large part due to Bernays.

About his *Habilitationsschrift* of 1918, Bernays said:

> [It] was certainly of a mathematical character. But the opinion at the time was that foundational investigations connected to mathematical logic were not taken seriously. They were considered amusing, playful. I had a similar ten-

> dency, and so did not take it seriously either. I was not very ambitious to get it published in time, and it appeared only much later, and then only in part [...] And so some of what I had achieved there was not duly recognized in the expositions of the development of mathematical logic.[4]

The present paper is in part an attempt to answer this complaint.

In §2, I give an exposition of the ideas contained in the lecture notes and in the *Habilitationsschrift* concerning semantics and completeness. Since there is significant overlap between Hilbert's lecture and Bernays's Habilitationsschrift, a discussion of the issue of authorship of the relevant passages is in order. This is the topic of §3. In §4, I present the parts of the Habilitationsschrift dealing with dependence and independence of axioms. §5 deals with Bernays's efforts to provide an axiomatization of propositional logic based on rules as opposed to axioms, an approach influencing later axiomatic developments and also Gentzen's sequent calculus. In §6, I try to provide several hints as to how this early work by Bernays and Hilbert influenced the further direction that logical investigations took in the Göttingen of the 1920s.

2.2 Semantics, Normal Forms, Completeness

2.2.1 Prehistory: Hilbert's Lectures on *Logical Principles of Mathematical Thought* 1905

In the Summer semester of 1905, Hilbert holds a course on "Logical principles of mathematical thought" (Hilbert 1905a). A detailed exposition of the lectures and their historical context is given by Peckhaus (1990), to whom much of the discussion in this section is indebted (see also his 1994, 1995). These lectures are highly interesting, for they contain developments of axiom systems not only for arithmetic and geometry, but also thermodynamics and probability theory. In them, Hilbert first discusses set theory and the paradoxes. In Chapter V ("The logical calculus"), we then read: "The paradoxes we have just introduced show sufficiently that an examination and redevelopment of the foundations of mathematics and logic is urgently necessary."[5]

Following a discussion of the purpose of logic and of the significance of contradictions, Hilbert develops propositional logic algebraically, using ideas from his first Heidelberg lecture given the year before (1905c). Hilbert lays down the following axioms:

> Axiom I. If $X \equiv Y$[6] then one can always replace X by Y and Y by X.
> Axiom II. From 2 propositions X, Y a new one results ("additively")
>
> $$Z \equiv X + Y$$
>
> Axiom III. From 2 propositions X, Y a new one results in a different way ("multiplicatively")
>
> $$Z \equiv X \cdot Y$$
>
> The following identities hold for these "operations":
>
> $$\begin{array}{ll} \text{IV. } X+Y \equiv Y+X & \text{VI. } X \cdot Y \equiv Y \cdot X \\ \text{V. } X+(Y+Z) \equiv (X+Y)+Z & \text{VII. } X \cdot (Y \cdot Z) \equiv (X \cdot Y) \cdot Z \end{array}$$
> $$\text{VIII. } X \cdot (Y+Z) \equiv X \cdot Y + X \cdot Z$$
>
> [...] There are 2 definite propositions 0, 1, and for each proposition X a different proposition $\overline{X}$ is defined, so that the following identities hold:
>
> $$\begin{array}{llll} \text{IX.} & X+\overline{X} \equiv 1 & \text{X.} & X \cdot \overline{X} \equiv 0 \\ \text{XI.} & 1+1 \equiv 1 & \text{XII.} & 1 \cdot X \equiv X \end{array}$$[7]

Hilbert's intuitive explanations make clear that X, Y, and Z stand for propositions, $+$ for conjunction, $\cdot$ for disjunction, $\overline{\cdot}$ for negation, 1 for falsity, and 0 for truth.[8] The axioms are followed by a discussion of the system from an algebraic standpoint. Hilbert points out how the axioms with the exception of (XI) also apply to arithmetic, and discusses the correspondence between negation and subtraction. Then he poses the main metatheoretical questions:

> It would now have to be investigated in how far the axioms are dependent and independent of one another [...] What would be most important here, however, is the proof that the 12 axioms do not contradict each other, i.e., that using the process defined one cannot obtain a proposition which contradicts the axioms, say, $X + \overline{X} = 0$. These are only hints which have not been carried out completely as of yet, and one still has free reign in the details; generally speaking this whole section supplies material for the ultimate solution of the interesting questions, rather than give the ultimate solution.[9]

These questions are to be solved 12 years later in the lectures from 1917–18 and in Bernay's *Habilitationsschrift*. It is interesting to note that Hilbert has all the tools in hand to give the solution already in 1905. We even find a nonderivability proof using an arithmetical interpretation of the axioms on p. 233: The axioms (XI) and (XII) are not derivable from the other axioms together with $X + 0 \equiv X$ and $X \cdot 0 \equiv 0$ (interpret $+$ and $\cdot$ as ordinary sum and product of reals, and take $\overline{X}$ to be $1 - X$.)

Hilbert proceeds to establish a number of consequences of the axioms in the style of algebraic proofs, in particular, de Morgan's laws. There is no distinction between consequence and the material conditional, $X \mid Y$[10] "Y follows from X [*aus X folgt Y*]" is defined by $\overline{X} \cdot Y \equiv 0$. Given this definition, it seems problematic to use nested conditionals, but subsequent examples indicate that $X \mid Y$ is intended also as an abbreviation for $\overline{X} \cdot Y$ not only for the equation $\overline{X} \cdot Y = 0$.

Hilbert then proves that every propositional formula can be brought into one of two normal forms. First one uses DeMorgan's laws repeatedly to see that every sentence can be written as sums and products of primitive propositions and their negations. Using the distributive law, this can be rewritten as a sum of products. Hilbert then uses a number of ways to simplify these, and claims (erroneously) that the resulting conjunctive normal form is unique up to reordering of conjuncts.[11] Using duality, it is then proved that every expression can also be brought into a disjunctive normal form.

Hilbert also discusses consequence at length. The system of propositional logic is intended as a background framework for other axiomatic theories. The axioms of those theories are interpreted as "correct" propositions, and the calculus is intended to make clear which propositions follow from the axioms according to the definition of consequence: Y follows from X if $\overline{X} \cdot Y = 0$. Hilbert proves the following about this notion of consequence:

> A proposition Y follows from another proposition X if and only if it is of the form $A \cdot X$, where A is some proposition. To deduce is to multiply correct propositions with arbitrary propositions.[12]

This theorem leads Hilbert to identify proofs with such factors A. The normal form theorem then provides the first proof of decidability of the propositional calculus. In the lecture on mathematical problems (Hilbert 1900a, p. 262), Hilbert discussed the issue of the decidability of every mathematical problem and proclaimed that "in mathematics there is no ignorabimus." The decidability of the propositional calculus is an example of what Hilbert is looking for:

> I now want to point out what is probably the most important application of the normal form of a proposition and its uniqueness. We will—and this is a restriction we have to impose for the time being—take a finite number of propositions *a, b, c, . . .* (axioms about the things considered or proper names) as given. Then there can be only a finite number of propositions (that is, propositions built up from these basic propositions), for every one can be brought into the form of a sum of products [conjunction of disjunctions] in basically a unique way. Every basic proposition appears in any summand [conjunct] only in the first dimension and any product [disjunction] appears only once as a summand [conjunct]. Every correct proposition must follow from the sum of the axioms $a+b+\cdots$ by multiplication with a certain factor A (proof) and for this A there are only finitely many [possible] forms by what has just been said. So it turns out that for every theorem there are only *finitely many possibilities of proof*, and thus we have solved, in the most primitive case at hand, the old problem that it must be possible to achieve any correct result by a *finite proof*. This problem was the original starting point of all my investigations in our field, and the solution to this problem in the most general case[,] the proof that there can be no "ignorabimus" in mathematics, has to remain the ultimate goal.[13]

There are many difficulties with this passage. First of all, if one takes the axioms of a theory to be a finite set of unanalyzed propositions a, b, c, $\ldots$, the propositional consequences of such a theory will not cover any significant number of their logical consequences. Taking the passage at face value, what we get is essentially a decision procedure for the propositional consequences of a set of variables. The argument can, however, easily be modified to apply to consequences of a finite set of propositional formulas.[14] This would not get us too far either, but Hilbert after all acknowledges that we are here dealing only with "a most primitive case." The next difficulty arises from Hilbert's earlier error of claiming that the normal form for a given formula is unique. For Hilbert's procedure to work, we would not only have to be able to enumerate all possible proofs A, but also be able to check if $A\cdot(a+b+\cdots)=Y$. This would presumably have to be done by comparing normal forms, since no other method—e.g., truth tables—is available. But normal forms are not unique, so there is no guarantee that the left and right side will result in the same one.[15] Lastly, the worry about the existence of a finite proof of any correct proposition is puzzling. It is not that the proof itself has to be finite what is important, but that there are only finitely many possibilities for a proof; we may decide, after finitely many steps, whether there is a proof or not.

All these difficulties aside, the main point is still notable. Here, in 1905, one of Hilbert's aims in the foundations of mathematics is made almost explicit, namely the aim to provide decision procedures for logic on the one hand, and particular systems of mathematics and science, e.g., arithmetic, on the other.

2.2.2 The Structure of *Prinzipien der Mathematik*

In the years following 1905, Hilbert's interest in the foundations of mathematics seems to have subsided. He does not follow up his groundbreaking ideas of 1905 until around 1917, when he returns with full force to his work on axiomatics.[16] In September 1917, Hilbert delivers his lecture on "axiomatic thought" in Zürich, and invites Bernays to come to Göttingen as his assistant. In the Winter semester 1917–18 Hilbert teaches a course on the "Principles of mathematics." The lecture notes to that course are preserved in the library of the Department of Mathematics at the University of Göttingen.[17] They are divided into two parts: Part A (62 pages) on the axiomatic method contains an exposition of axiomatic geometry; Part B (pp. 63–246, 184 pages) deals with mathematical logic. The material in Part B is new and interesting. It starts out with a discussion of propositional calculus in the style of algebraic logic in Section 1 (pp. 63–80). The propositional calculus is extended to a calculus of classes in Section 2 (pp. 81–107), and a theory of syllogisms is developed. In Section 3, the limitations of the class calculus are used to motivate the introduction of the calculus of functions, i.e., first-order logic with quantifiers (pp. 108–129). This calculus of functions is formally introduced and studied in Section 4 (pp. 129–187). Section 5 (pp. 188–246) deals with the extended calculus of functions (i.e., second-order logic), as well as with induction, the definition of identity, the paradoxes, type theory, and the axiom of reducibility.[18]

From a historical point of view, the last two sections of Part B are the most interesting ones. The development of geometry in Part A is standard, and overlaps both with the *Foundations of Geometry* (Hilbert 1903) and the material presented in numerous courses on axiomatic geometry taught by Hilbert at Göttingen. The propositional calculus presented in Section 1 of Part B is exactly the same as the one developed in Hilbert's 1905 course. There are two notable differences in the presentation. The 1917–18 notes contain an independence proof similar to the one in (Hilbert 1905a), as well as a proof of consistency of the axioms of propositional logic. In contrast to the independence proof, which uses an arithmetic interpretation, consistency is proved by restricting the range to only the propositions 0 and 1, and defining sum and product case-by-case:

> Restrict the domain of propositions by allowing only the propositions 0 and 1, and interpret the equations in accordance with this as proper identities. Furthermore, define sum and product by the 8 equations
>
> $0+0=0 \quad 0\times 0=0$
> $0+1=1 \quad 0\times 1=0$
> $1+0=1 \quad 1\times 0=0$
> $1+1=1 \quad 1\times 1=1$
>
> which are characterized by turning into correct arithmetical equations, if one replaces the symbolic sum by the maximum of the summands and the symbolic product by the minimum of the factors. Declare the proposition 1 to be the negation of the proposition 0 and the proposition 0 to be the negation of 1.
>
> These definitions in any case do not lead to a contradiction, for each one of them defines a new symbol. On the other hand, one can establish by finitely many tries that all the axioms I–XII are satisfied by these definitions. These axioms therefore cannot result in a contradiction either. Thus the question of consistency of our calculus can be completely resolved.[19]

What is interesting here is that, while Hilbert thought that an arithmetical interpretation is good enough to establish independence results, something more basic is needed to show consistency. The first sentence in the last paragraph just quoted indicates that Hilbert had scruples regarding the use of arithmetic correctness of equations to establish consistency. He simply wanted to avoid appeal to infinite structures at this point.

The second difference is a much more elaborate discussion of consequence. The definition is the same as in 1905 (only the symbol for implication changes to $\rightarrow$), but now a number of properties are proved that one would expect of a system of logic: For any X and Y, $X \rightarrow X$, $X+Y \rightarrow X$, if $X \rightarrow Y$ then $\overline{Y} \rightarrow \overline{X}$, and others. A discussion of "proofs as multiplication" and of decidability is missing, however.

Taking this notion of consequence as a starting point, Hilbert takes on an investigation of how much of mathematical reasoning can be accommodated in the propositional calculus. In Section 2 (Predicate calculus and class calculus), the propositional calculus is reinterpreted as first, a calculus of predicates, and second, a calculus of classes (extensions of predicates). These reinterpretations are then used to account for the Aristotelian syllogisms in the framework of the calculus. Naturally it is ultimately found (in Section 3: Transition to the calculus of functions) to be insufficient for a foundation of mathematics, for it is unable to deal with relations between individuals or with nested quantifiers. This leads Hilbert to introduce the function calculus, first by example (the difference between convergence and uniform convergence), and then finally, as an axiom system.

Section 4, entitled “Systematic presentation of the function calculus,” contains a presentation of the function calculus, i.e., first-order logic, organized as follows:

4.1. Axioms of the function calculus (pp. 129–140)

4.2. The system of logical propositional formulas (pp. 140–153)

4.3. The complete system of logical formulas (pp. 154–179)

4.4. Examples of applications of the function calculus (pp. 180–187)

Section 5 of Part B of the lecture notes discusses the extended function calculus, i.e., higher-order logic. It includes discussions of definitions of number, set theory, paradoxes and type theory.

Let me now turn to a discussion of the propositional fragment of the function calculus as developed in 4.1 and 4.2. For discussion of the full first-order logic and the later parts of the lecture notes, the interested reader is referred to the papers by Moore (1997) and Sieg (1999).

2.2.3 The Propositional Calculus

The propositional fragment of the function calculus is investigated separately in Subsection 2 of Section 4. Syntax and axioms are modeled after the propositional fragment of *Principia Mathematica* (Whitehead and Russell 1910). The language consists of propositional variables [Aussage-Zeichen] $X, Y, Z, \ldots$, as well as signs for particular propositions, and the connectives $\overline{\cdot}$ (negation) and $\times$ (disjunction). The conditional, conjunction, and equivalence are introduced as abbreviations. Expressions are defined by recursion:

1. Every propositional variable is an expression.
2. If α is an expression, so is $\overline{\alpha}$.

3. If α and β are expressions, so are $\alpha \times \beta$, $\alpha \rightarrow \beta$, $\alpha + \beta$ and $\alpha = \beta$.[20]

Hilbert introduces a number of conventions, e.g., that $X \times Y$ may be abbreviated to XY, and the usual conventions for precedence of the connectives. Finally, the logical axioms are introduced. Group I of the axioms of the function calculus gives the formal axioms for the propositional fragment (unabbreviated forms are given on the right, recall that XY is "X or Y"):

1.	$XX \rightarrow X$	$\overline{X}\,\overline{X}X$
2.	$X \rightarrow XY$	$\overline{X}(XY)$
3.	$XY \rightarrow YX$	$\overline{XY}(YX)$
4.	$X(YZ) \rightarrow (XY)Z$	$\overline{X(YZ)}((XY)Z)$
5.	$(X \rightarrow Y) \rightarrow (ZX \rightarrow ZY)$	$\overline{\overline{X}Y}(\overline{ZX}(ZY))$

The formal axioms are postulated as correct formulas [*richtige Formel*], and we have the following two rules of derivation ("contentual axioms"):

a. Substitution: From a correct formula another one is obtained by replacing all occurrences of a propositional variable with an expression.

b. If α and $\alpha \rightarrow \beta$ are correct formulas, then β is also correct.

Although the calculus is very close to the one given in *Principia Mathematica*, there are some important differences. Russell uses (2′) $X \rightarrow YX$ and (4′) $X(YZ) \rightarrow Y(XZ)$ instead of (2) and (4). *Principia* also does not have an explicit substitution rule.[21] The fact that Hilbert realizes that such a rule must be included in the calculus illustrates how Hilbert's axiomatic method makes the presentation of logic in 1917–18 much clearer than Schröder's algebra of logic and much closer to the modern conception of logic as calculus than Russell's *Principia*. But the division between syntax and semantics is not quite complete. The calculus is not regarded as concerned with uninterpreted formulas; it is not separated from its interpretation. (This is also true of the first-order part, see Sieg (1999), B3.) Also, the notion of a "correct formula" which occurs in the presentation of the calculus is intended not as a concept defined, as it were, by the calculus (as we would nowadays define the term "provable formula" for instance), but rather should be read as a semantic stipulation: The axioms are true, and from true formulas we arrive at more true formulas using the rules of inference.[22] Read this way, the statement of modus ponens is not that much clearer than the one given in *Principia:* "Everything implied by a true proposition is true." (*1.1)

Hilbert goes on to give a number of derivations and proves additional rules. These serve as stepping stones for more complicated derivations. First, however, he proves a normal form theorem, just as he did in the 1905 lectures, to establish decidability and completeness. In the new propositional calculus, however, Hilbert needs to prove that arbitrary subformulas can be replaced by equivalent formulas, that is, that the rule of replacement is a dependent rule.[23] He does so by establishing the admissibility of rule (c): If $\varphi(\alpha)$, $\alpha \rightarrow \beta$, and $\beta \rightarrow \alpha$ are provable, then so is $\varphi(\beta)$.[24] With that, the admissibility of using commutativity, associativity, distributivity, and duality inside formulas is quickly established, and Hilbert obtains the normal form theorem just as he did for the first propositional calculus in the 1905 lectures. Normal forms again play an important role in proofs of decidability and now also completeness.

2.2.4 Consistency and Completeness

"This system of axioms would have to be called inconsistent if it were to derive two formulas from it which stand in the relation of negation to one another."[25] That the system of axioms is not inconsistent in this sense is proved, again, using an arithmetical interpretation. The propositional variables are interpreted as ranging over the numbers 0 and 1, $\times$ is just multiplication and $\overline{X}$ is just $1 - X$. One sees that the five axioms represent functions which are constant equal to 0, and that the two rules preserve that property. Now if α is derivable, $\overline{\alpha}$ represents a function constant equal to 1, and thus is underivable.

Why did Hilbert not use this straightforward arithmetical interpretation to prove consistency for the first propositional calculus in 1905 or earlier in the lectures (Section 4.1)? If it was his concern that an infinite interpretation should not be used to establish consistency of such a basic system as that of propositional logic, then the numbers 0 and 1 alone would do just as well. One possible explanation is that up until the introduction of the new propositional calculus based on the *Principia* system, conjunction and disjunction were both primitives. Giving an arithmetical interpretation for these systems would thus have required finding an interpretation which also satisfies $1 + 1 = 1$. Simply taking congruences modulo 2 does not do the trick here. Only when $+$ is taken as a defined symbol can one take the congruences modulo 2 as an interpretation of the axioms. Compared to the consistency proof in Section 4.1 using true and false propositions, the arithmetical interpretation is further away again from truth-value semantics for propositional logic.

> Let us now turn to the question of *completeness*. We want to call the system of axioms under consideration complete if we always obtain an inconsistent system of axioms by adding a formula which is so far not derivable to the system of basic formulas.[26]

This is the first time that completeness is formulated as a precise mathematical question to be answered for a system of axioms. Before this, Hilbert (1905a, p. 13) had formulated completeness as the question of whether the axioms suffice to prove all "facts" of the theory in question. Aside from that, completeness had always been *postulated* as one of the axioms. In the *Foundations of Geometry*, for instance, we find axiom V(2), stating that it is not possible to extend the system of points, lines, and planes by adding new entities so that the other axioms are still satisfied. In (Hilbert 1905a), such an axiom is also postulated for the real numbers. Following its formulation, we read:

> This last axiom is of a general kind and has to be added to every axiom system whatsoever in some form. It is of special importance in this case, as we shall see. Following this axiom, the system of numbers has to be so that whenever new elements are added contradictions arise, regardless of the stipulations made about them. If there are things which can be adjoined to the system without contradiction, then in truth they already belong to the system.[27]

We see here that the formulation of completeness of the axioms arises directly out of the completeness axioms of Hilbert's earlier axiomatic systems, only that this time completeness is a theorem *about* the system. I shall return to this issue in the final section.

The completeness proof in the 1917–18 lectures itself is an ingenious application of the normal form theorem: Every formula is interderivable with a conjunctive normal form. As has been proven earlier, a conjunction is provable if and only if each of its conjuncts is provable. A disjunction of propositional variables and negations of propositional variables is provable only if it represents a function which is constant equal to 0, as the consistency proof shows. A disjunction of this kind is equal to 0 if and only if it contains a variable and its negation, and conversely, every such disjunction is provable. So a formula is provable if and only if every conjunct in its normal form contains a variable and its negation. Now suppose that α is an underivable formula. Its conjunctive normal form β is also underivable, so it must contain a conjunct γ where every variable occurs only negated or unnegated but not both. If α were added as a new axiom, then β and γ would also be derivable. By substituting X for every unnegated variable and $\overline{X}$ for every negated variable in γ, we would obtain X as a derivable formula (after some simplification), and the system would be inconsistent.

In a footnote, the result is used to establish the converse of the characterization of provable formulas used for the consistency proof: every formula representing a function which is constant equal to 0 is provable. For, supposing there were such a function which was not provable, following the consistency proof above, adding this formula to the axioms would not make the system inconsistent, and this would contradict syntactic completeness (Hilbert 1918c, p. 153).

2.2.5 The Contribution of Bernays's *Habilitationsschrift*

We have seen that the lecture notes to *Principles of Mathematics* 1917–18 contain consistency and completeness proofs (relative to a syntactic completeness concept) for the propositional calculus of *Principia Mathematica*. They also implicitly contain the familiar truth-value semantics and a proof of semantic soundness and completeness. In his *Habilitationsschrift* (Bernays 1918), Bernays fills in the last gaps between these remarks and a completely modern presentation of propositional logic.

Bernays introduces the propositional calculus in a purely formal manner. The concept of a formula is defined and the axioms and rules of derivation are laid out almost exactly as done in the lecture notes. §2 of (Bernays 1918) is entitled "Logical interpretation of the calculus. Consistency and completeness." Here Bernays first gives the interpretation of the propositional calculus, which is the motivation for the calculi in Hilbert's earlier lectures (Hilbert 1905a, 1918c). The reversal of the presentation—first calculus, then its interpretation—makes it clear that Bernays is fully aware of a distinction between syntax and semantics, a distinction not made precise in Hilbert's earlier writings.[28] There, the calculi were always introduced with the logical interpretation built in, as it were. Bernays writes:

> The axiom system we set up would not be of particular interest, were it not capable of an important contentual interpretation.
>
> Such an interpretation results in the following way:
>
> The variables are taken as symbols for *propositions* (sentences).
>
> That propositions are either true or false, and not both simultaneously, shall be viewed as their characteristic property.
>
> The symbolic product shall be interpreted as the connection of two propositions by "or," where this connection should not be understood in the sense of a proper disjunction, which excludes the case of both propositions holding jointly, but rather so that "X or Y" holds (i.e., is true) if and only if at least one of the two propositions X, Y holds.[29]

Similar truth-functional interpretations of the other connectives are given as well. Bernays then defines what a provable and what a valid formula is, thus making the syntax-semantics distinction explicit:

> The importance of our axiom system for logic rests on the following fact: If by a "provable" formula we mean a formula which can be shown to be correct according to the axioms [footnote in text: It seems to me to be necessary to introduce the concept of a provable formula in addition to that of a correct formula (which is not completely delimited) in order to avoid a circle], and by a "valid" formula one that yields a true proposition according to the interpretation given for any arbitrary choice of propositions to substitute for the variables (for arbitrary "values" of the variables), then the following theorem holds:
>
> *Every provable formula is a valid formula and conversely.*
>
> The first half of this claim may be justified as follows: First one verifies that all basic formulas are valid. For this one only needs to consider finitely many cases, for the expressions of the calculus are all of such a kind that in their logical interpretation their truth or falsehood is determined uniquely when it is determined of each of the propositions to be substituted for the variables whether it is true or false. The content of these propositions is immaterial, so one only needs to consider truth and falsity as values of the variables.[30]

Everything one would expect of a modern discussion of propositional logic is here: A formal system, a semantics in terms of truth values, soundness and completeness relative to that semantics. As Bernays points out, the consistency of the calculus, of course, follows from its soundness. Lest the reader—recall that the intended readership includes Hilbert and his colleagues among the Göttingen faculty—have reservations about the "logical interpetation," Bernays points out that the interpretation of the variables by truth values is of no consequence, the same results could be obtained by an arithmetical interpretation using 0 and 1.

The semantic completeness of the calculus is proved in §3, along the lines of the footnote in (Hilbert 1918c) mentioned above. What may be pointed out here is that the formulation of syntactic completeness given by Bernays is slightly different from the lectures and independent of the presence of a negation sign: it is impossible to add an unprovable formula to the axioms without thus making all formulas provable. Bernays sketches the proof of syntactic completeness along the lines of Hilbert's lectures, but leaves out the details of the derivations.

Bernays also addresses the question of decidability. Decidability was not addressed at all in the lecture notes, even though Hilbert had posed it as one of the fundamental problems in the investigation of the calculus of logic. In his talk in Zürich in 1917, he said that an axiomatization of logic cannot be satisfactory until the question of decidability by a finite number of operations is understood and solved (Hilbert 1918a, p. 413, ¶41–42). Bernays gives this solution for the propositional calculus by observing that

> [t]his consideration does not only contain the proof for the completeness of our axiom system, but also provides a uniform method by which one can decide after finitely many applications of the axioms whether an expression of the calculus is a provable formula or not. To decide this, one need only determine a normal form of the expression in question and see whether at least one variable occurs negated and unnegated as a factor in each simple product. If this is the case, then the expression considered is a provable formula, otherwise it is not. The calculus therefore can be completely trivialized.[31]

2.2.6 A Brief Comparison with Post's thesis

Emil L. Post's dissertation of 1920 (Post 1921) is the locus classicus for all of the basic metatheoretical results about the propositional calculus.[32] It contains an explicit account of the truth table method, and the fundamental theorem that a formula is provable from the axioms of *Principia*[33] if and only if it defines a truth function which is always equal to '+' (true). From the fundamental theorem, Post deduces a number of consequences. Among them are, for instance, that the truth table method provides a decision procedure for derivability in the propositional calculus and that the addition of any unprovable formula yields an inconsistent system (inconsistency is understood here alternatively as proving both a formula and its negation, and as proving every formula). Post uses the term “closed” for systems which are such that the addition of an unprovable formula makes all formulas provable (p. 177).

Post’s paper contains a number of other contributions. These are, on the one hand, a discussion of truth-functional completeness, and on the other, the introduction of many-valued logics. We will see later that Bernays’s approach to proving independence of the axioms involved something very much like many-valued logics. It might also be pointed out that some of the discussion of truth-functional completeness can also be found in Bernays. On pp. 16–19 of (Bernays 1918), Bernays makes a number of remarks which are relevant here. For instance, there we find the claim that “all relationships between truth and falsity of propositions can be expressed using conjunction (‘and’), disjunction (exlusive ‘or’) and negation, so and thus also using the symbolism of our calculus.”[34] Another remark concerns the equivalence of formulas in propositional logic. Two formulas are defined to be equivalent if $\alpha \sim \beta$ is provable (‘$\sim$’ is the *Principia* notation for the biconditional; Hilbert uses ‘$=$’). By the completeness theorem, this is the case if and only if $\alpha \sim \beta$ is valid. From this, Bernays shows that any formula is equivalent to one containing only negation and disjunction, or only negation and conjunction, or only negation and implication, and that corresponding claims for negation and equivalence or conjunction and disjunction do not hold. What we do not find, however, is a *proof* that every truth function can be represented by, say, negation and disjunction. A proof of this can be found in lecture notes to a course by Hilbert given in 1920 (Hilbert 1920a, pp. 18–19), the same year that Post submitted his dissertation.

The discussion of the fragment without negation leads Bernays to pose the question of whether there might be an axiom system in which all and only the provable negation-free propositional formulas are derivable. He claims that this can in fact be done, but does not give an axiomatization. We shall return to this question in Section 2.5.

2.3 Hilbert or Bernays?

It is well known that Bernays played an important role in the development of Hilbert's program in the 1920s, and that he wrote the monumental *Grundlagen der Mathematik* (Hilbert and Bernays 1934, 1939) essentially alone, of course using Hilbert's ideas.[35] Mancosu (1998b, p. 175) stresses Bernays's contributions to the program in giving "more explicit discussion of the central philosophical topics surrounding Hilbert's program," and in clarifying Hilbert's views. Of course, there are also Bernays's published contributions to the program, for instance the work on the Entscheidungsproblem with Schönfinkel (Bernays and Schönfinkel 1928), and the investigations of the propositional calculus in the Habilitationsschrift. Through contact with his colleagues in Göttingen, Bernays had great influence on technical developments, and his contributions and suggestions are acknowledged not only by Hilbert himself. I would like to argue here that Bernays was in fact instrumental already for the technical advances made in 1917–18, and that the development of propositional and first-order logic in (Hilbert 1918c) is at least as much due to Bernays as it is to Hilbert. Moore (1997) and Sieg (1999) point out that these advances are not only the formulation of calculi for propositional and first-order logic, but in particular the investigation of meta-logical questions about these calculi: consistency, completeness, decidability. These are the questions that Hilbert (1918a) emphasized as important questions to be answered for the calculus of logic. Their solution is in large part due to Bernays.

The winter term of 1917–18 was Bernays's first semester as Hilbert's assistant in Göttingen. Bernays characterized his duties as assistant as follows: "So I was [Hilbert's] assistant. That job was not like what assistants usually do here [in Zürich], helping the students with exercises and such. I had nothing to do with that. On the one hand, we discussed foundational questions, and on the other I helped with the preparation of his lectures and prepared lecture notes."[36] Bernays held an appointment as *außerplanmäßiger Assistent*, which meant that he did not have a regular position which carried a salary, but that he relied on stipends. Hilbert urged Bernays to obtain the *venia legendi* so that he would be able to teach courses. Bernays submitted his application for the *Habilitation* on 9 July 1918, it included the *Habilitationsschrift* (Bernays 1918).[37] Bernays gave his *Probevorlesung* on 23 December 1918, and the dean of the Faculty of Philosophy granted the *venia legendi* on

14 January 1919.[38] The *Habilitationsschrift* contains page references to the lecture notes for the 1917–18 lectures, so the lecture notes must have been finished by the time Bernays submitted the thesis in July 1918. The winter term lasted from 1 October 1917 to 2 February 1918—approximately 15 weeks of classes. The course on *Prinzipien der Mathematik* was given on Thursdays, 9–11 am. Bernays's own shorthand notes, which he took during the lecture, survive in his *Nachlaß* in Zürich.[39] Bernays marked the end of each lecture with a horizontal line, and thus a comparison with the lecture notes makes it possible to ascertain which parts of the lecture were given when. Approximately, we find the following: The first seven lectures correspond the Part A on geometry. The first version of the propositional calculus is developed in the next three lectures, corresponding to pp. 63–80 of (Hilbert 1918c). The predicate and class calculi are discussed in the next two lectures, corresponding to pp. 81–129. Already it is remarkable that the typewritten notes contain a lot of material that is not contained in Bernays's notes, e.g., the extended discussion of syllogisms on pp. 99–105. The last three lectures cover the following: the axioms and rules of the restricted function calculus (corresponding to pp. 129–135); application to inferences with a singular premise (pp. 180–181), the extended function calculus, definition of identity, number, sets (pp. 188–194), paradoxes (pp. 209–218); and paradoxes continued. We see that key parts of the lecture notes were apparently not covered in the lecture: the sections on derivations of theorems and rules in the propositional calculus (pp. 140–179) including consistency and completeness are completely missing from Bernays's notes, as is the last section on type theory (pp. 219–245); the sections on the extended function calculus, set theory and the paradoxes were only briefly sketched. In total, 117 pages—almost the latter half of the lecture notes—correspond to three two-hour lectures, 8 pages of shorthand notes out of 55. Not surprisingly, while the typescript keeps very closely to the structure of the lectures for the first one hundred pages or so (half-empty pages where a lecture ended, references to subjects discussed "the last time"), these last 117 read more like a monograph than like lecture notes.

The documentary record thus strongly suggests the following: The important results on the propositional and the restricted function calculus were obtained after the lectures were given, approximately in the period February–May 1918, when Bernays elaborated his notes to the lecture. The *Habilitationsschrift* was written after the lecture notes were completed, in the Spring of 1918. Some additional circumstantial evidence can be adduced for the thesis that the additional parts of the lecture notes, including the important results, are due in large part to Bernays. For one, the completeness proof is referred to a number of times by members of the Hilbert school. Bernays mentions it in the introduction of the *Habilitationsschrift*, where he states that proofs of consistency and completeness can be found in the 1917–18 lecture notes, before he gives the proof itself. We have seen that these proofs were not given in the actual lectures, and so these remarks must be understood as merely pointing the reader to the details of the normal form theorem (which was not proved in the *Habilitationsschrift*) rather than crediting Hilbert with the results. The published version (Bernays 1926) does not mention Hilbert's lectures at all. Behmann (1922a) presents the decision procedure based on the completeness proof and refers in this connection only to Bernays (1918), although Behmann is certainly aware of the 1917–18 lectures (they are quoted on p. 165, and he almost certainly took the class). In the notes to a course on mathematical logic given in Göttingen in the summer term 1922, Behmann writes:

> These questions [of independence] concerning the axiomatics of elementary propositions were treated a few years ago by the Göttingen mathematician Bernays (Habilitationsschrift, unfortunately not published), and, one may well say, given a complete and satisfactory answer. [. . .] Bernays also rigorously proved completeness, i.e., has shown that every universally valid elementary proposition can indeed be derived from the basic formulas according to the basic rules.[40]

The most convincing piece of evidence may be the following remark by Bernays:

> My knowledge [of logic] was very incomplete at the time, in 1917. Before Hilbert took up the [investigation of the foundations of mathematics] directly again, which he had started much earlier [in (Hilbert 1905c)], he did not immediately lecture on that, but he gave a course on mathematical logic. And I was in charge of writing up [*ausarbeiten*] that lecture course, and I did this in such a way that I avoided free variables. I had looked at Russell a little bit, and first I found it too broad and did not like it in all respects, but in particular I did not understand what it means to say "for all x, $F(x)$, then $F(y)$ follows." In fact, the application of free variables is something technical. These are two ways to represent generality. One has generality on the one hand through bound variables and on the other through free variables. There is no such difference in natural langauge. So I avoided free variables at first. This is a possible

> way of approach, and later others have also done it this way. So that was a lecture course, which was written up, and then was filed in the library of the Mathematical Institute.[41]

Bernays's testimony here clearly indicates that the formulation of the quantifier axioms in (Hilbert 1918c) is due to him. It explains the particular form of these axioms, and why they differ so much from the corresponding postulates of *Principia Mathematica* and from later presentations (e.g., in Hilbert and Ackermann (1928),[42] which are otherwise based closely on the lecture notes from 1917–18). We may infer from this that the extent of Bernays's influence on the formulation and presentation of the results in the lecture notes from 1917–18 goes far beyond merely typing up what Hilbert said. Some of the results may or may not be due to Bernays. For instance, it is possible that Hilbert simply did not have enough time to present the completeness proof, but told Bernays to include it in the typescript. Given the amount of material that was not covered in the lecture, and the character of Hilbert and Bernays's working relationship, it is clear that a large amount of that material must have been worked out by Bernays alone. The fact that something as central as the formulation of the quantifier axioms is due to Bernays shows that it is very likely that he was the author of the parts of the lecture notes not covered in the lecture itself, and even that much of the material that was covered is in fact due to him. Be that as it may, the insights and results that are certainly due to Bernays—a clear syntax-semantics distinction, formulation of semantic completeness, independence results—are important enough to earn Bernays a prominent place in the history of the subject.

Why did Bernays not claim the results as his? A possible explanation may be his pronounced modesty.[43] (By the same token, if the results were exclusively Hilbert's, Bernays would have made a point of noting that when he presented the proofs, e.g., in Bernays (1926) and Bernays and Schönfinkel (1928).) Also recall the then prevalent tendency described by Bernays in a quote in §1 above, not to take mathematical logic seriously—at the time, he may well have thought of the results as not worth mentioning.

2.4 Dependence and Independence

Consistency and independence are the requirements that Hilbert laid down for axiom systems of mathematics time and again. Consistency was established—but the "contributions to the axiomatic treatment" of propositional logic could not be complete without a proof that the axioms investigated are independent. In fact, however, the axiom system for the propositional calculus, slightly modified from the postulates in (*1) of *Principia Mathematica*, is not independent. Axiom 4 is provable from the other axioms. Bernays devotes §4 of the *Habilitationsschrift* to give the derivation, and also the inter-derivability of the original axioms of *Principia* (2′) and (4′) with the modified versions (2) and (4) in presence of the other axioms. Together the derivations also establish the dependence of (*1.5) from the other propositional postulates in *Principia*.

Independence is of course more challenging. The method Bernays uses is not new, but it is applied masterfully. Hilbert had already used arithmetical interpretations in Hilbert (1905a) to show that some axioms are independent of the others. The idea was the same as that originally used to show the independence of the parallel postulate in Euclidean geometry: To show that an axiom α is independent, give a model in which all axioms but α are true, the inference rules are sound, but α is false. Schröder was the first to apply that method to logic. §12 of his *Algebra of Logic* (Schröder 1890) gives a proof that one direction of the distributive law is independent of the axioms of logic introduced up to that point. The interpretation he gives is that of the "calculus of algorithms," developed in detail in Appendix 4. Bernays combines Schröder's idea with Hilbert's arithmetical interpretation and the idea of the consistency proof for the first propositional calculus in Hilbert (1918c) (interpreting the variables as ranging over a certain finite number of propositions, and defining the connectives by tables). He gives six "systems" to show that each of the five axioms (and a number of other formulas) is independent of the others. The systems are, in effect, finite matrices. He introduces the method as follows:

> In each of the following independence proofs, the calculus will be reduced to a finite system (a finite group in the wider sense of the word [footnote: that is, without assuming the associative law or the unique invertability of composition]), where for each element a composition ("symbolic product") and a "negation" is defined. The reduction is given by letting the variables of the

> calculus refer to elements of the system as their values. The "correct formulas" are characterized in each case as those formulas which only assume values from a certain subsystem T for arbitrary values of the variables occurring in it.[44]

In the published version (Bernays 1926), the elements of the subsystem T are called *ausgezeichnete Werte*—designated values. The term is commonly used today.

I shall not go into the details of the derivations and independence proofs.[45] Let me just say that Bernays's method was of some importance in the investigation of alternative logics. For instance, Heyting (1930) used it to prove the independence of his axiom system for intuitionistic logic and Gödel (1932) was influenced by it when he defined a sequence of sentences F_n so that each F_n is independent of intuitionistic propositional calculus together with all F_i, $i > n$.[46] The many-valued logics Gödel used to show this are now called Gödel logics. It may be debated whether Bernays's systems can properly be called many-valued *logics*, but they certainly had the distinction of being useful in proving independence results in logic, an achievement considered important.

2.5 Axioms and Rules

In the final section of his *Habilitationsschrift*, Bernays considers the question of whether some of the axioms of the propositional calculus may be replaced by rules. This seems like a natural question, given the relationship between inference and implication: For instance, axiom 5 suggests the following rule of inference: (Recall that $\alpha\beta$ is Hilbert's notation for the disjunction of α and β. See §2.3 for a list of the axioms and rules.)

$$\frac{\alpha \to \beta}{\gamma\alpha \to \gamma\beta}\,\mathrm{c}$$

which Bernays used earlier as a derived rule. Indeed, axiom 5 is in turn derivable using this rule and the other axioms and rules. Bernays considers a number of possible rules

$$\frac{\begin{matrix}\alpha \to \beta \\ \beta \to \gamma\end{matrix}}{\alpha \to \gamma}\,\mathrm{d} \qquad \frac{\alpha\alpha}{\alpha}\,\mathrm{r}_1 \qquad \frac{\alpha}{\alpha\beta}\,\mathrm{r}_2 \qquad \frac{\alpha\beta}{\beta\alpha}\,\mathrm{r}_3 \qquad \frac{\alpha(\beta\gamma)}{(\alpha\beta)\gamma}\,\mathrm{r}_4$$

$$\frac{\varphi(\alpha\alpha)}{\varphi(\alpha)}\,\mathrm{R}_1 \qquad \frac{\varphi(\alpha\beta)}{\varphi(\beta\alpha)}\,\mathrm{R}_3$$

and shows that the following sets of axioms and rules are equivalent (and hence, complete for propositional logic):

1. Axioms: 1, 2, 3, 5; rules: a, b

2. Axioms: 1, 2, 3; rules: a, b, c

3. Axioms: 2, 3; rules: a, b, c, r_1

4. Axioms: 2; rules: a, b, c, r_1, R_3

5. Axioms: $\overline{X}X$[47]; rules: a, b, c, r_1, r_2, r_3, r_4

Bernays also shows, using the same method as before, that these axiom systems are independent, and also the following independence results:[48]

6. Rule c is independent of axioms: 1, 2, 3; rules: a, b, d (showing that in (2), rule c cannot in turn be replaced by d);

7. Rule r_2 is independent of axioms: 1, 3, 5; rules: a, b, (thus showing that in (1) and (2), axiom 2 cannot be replaced by rule r_2);

8. Rule r_3 is independent of axioms: 1, 2; rules: a, b, c (showing similarly, that in (1) and (2), rule r_3 cannot replace axiom 3);

9. Rule R_3 is independent of axioms: $\overline{X}X$, 3; rules: a, b (showing that R_3 is stronger than r_3, since 3 is provable from R_3 and $\overline{X}X$);

10. Rule R_1 is independent of axioms: $\overline{X}X$, 1; rules: a, b (showing that R_1 is stronger than r_1, since 1 is provable from $\overline{X}X$ and R_1);

11. Axiom 2 is independent of axioms: $\overline{X}X$, 1, 3, 5; rules: a, b, and

12. Axiom 2 is independent of axioms: $\overline{X}X$; rules: a, b, c, r_1, R_3 (showing that in (5), $\overline{X}X$ together with r_2 is weaker than axiom 2).

The detailed study exhibits, in particular, a sensitivity to the special status of rules like R_3, where subformulas have to be substituted. The discussion foreshadows developments of formal language theory in the 1960s. Bernays also mentions that a rule (corresponding to axiom 1), allowing inference of $\varphi(\alpha)$ from $\varphi(\alpha\beta)$ would be incorrect (and hence, "there is no such generalization of r_1").

Bernays's discussion of axioms and rules, together with his discussion of expressibility in the "Supplementary remarks to §2–3" (discussed above at the end of Section 2.2.6), shows his acute sensitivity for subtle questions regarding logical calculi. His remarks are quite opposed to the then-prevalent tendency (e.g., Sheffer and Nicod) to find systems with fewer and fewer axioms, and foreshadow investigations of relative strength of various axioms and rules of inference, e.g., of Lewis's modal systems, or more recently of the various systems of substructural logics.

At the end of the "Supplementary remarks," Bernays isolates the positive fragment of propositional logic (i.e., the provable formulas not containing negation; here $+$ and $\rightarrow$ are considered primitives) and claimed that he had an axiomatization of it. He did not give an axiom system, but stated that it is possible to choose a finite number of provable sentences as axioms so that completeness follows by a method exactly analogous to the proof given in §3. The remark suggests that Bernays was aware that the completeness proof is actually a proof schema, in the following sense. Whenever a system of axioms is given, one only has to verify that all the equivalences necessary to transform a formula into conjunctive normal form are theorems of that system. Then completeness follows just as it does for the axioms of *Principia*.

In his next set of lectures on the "logical calculus" given in the Winter semester of 1920,[49] Hilbert makes use of the fact that that these equivalences are the important prerequisite for completeness. The propositional calculus we find there is markedly different from the one in Hilbert (1918c) and Bernays (1918), but the influences are clearly visible. The connectives are all primitive, not defined, this time. The sole axiom is $\overline{X}X$, and the rules of inference are:

$$\frac{X}{XY}\,\text{b2} \qquad \frac{\begin{matrix}X\\Y\end{matrix}}{X+Y}\,\text{b3}$$

plus the rule (b4), stating: "Every formula resulting from a correct formula by transformation is correct." "Transformation" is meant as transformation according to the equivalences needed for normal forms: commutativity, associativity, de Morgan's laws, $\overline{\overline{X}}$ and X, and the definitions of $\rightarrow$ and $=$ (biconditional). These transformations work in both directions, and also on subformulas of formulas (as did R_1 and R_3 above).[50] One equivalence corresponding to modus ponens must be added, it is: $(X+\overline{X})Y$ is intersubstitutable with Y.

Anyone familiar with the work done on propositional logic elsewhere might be puzzled by this seemingly unwieldy axiom system. It would seem that the system in Hilbert (1920a) is a step backward from the elegance and simplicity of the *Principia* axioms. Adjustments, if they are to be made at all, it would seem, should go in the direction of even more simplicity, reducing the number of primitives (as Sheffer did) and the number of axioms (as in the work of Nicod and later Łukasiewicz). Hilbert is motivated by different concerns. He was not only interested in the simplicity of his axioms, but in their efficiency. Decidability, in particular, supersedes considerations of independence and elegance. The presentation in Hilbert (1920a) is designed to provide a decision procedure which is not only efficient, but also more intuitive to use for a mathematician trained in algebraic methods. Bernays's study of inference rules made clear, on the other hand, that such an approach can in principle be reduced to the axiomatics of *Principia*. One may ask whether the truth table method is not just as efficient a decision procedure. As any computer scientist working in automated theorem proving knows, truth tables are the worst possible decision procedure for propositional logic—exponential not only in the worst case, but in *every* case. In a similar vein, the subsequent work on the decision problem is not strictly axiomatic, but uses transformation rules and normal forms. The rationale is formulated by Behmann:

> The form of presentation will not be axiomatic, rather, the needs of practical calculation shall be in the foreground. The aim is thus not to reduce everything to a number (as small as possible) of logically independent formulas and rules; on the contrary, I will give as many rules with as wide an application as possible, as I consider appropriate to the practical need. The logical dependence of rules will not concern us, insofar as they are merely of independent practical importance. [...] Of course, this is not to say that an axiomatic development is of no value, nor does the approach taken here preempt such a development. I just found it advisable not to burden an investigation whose aim is in large part the exhibition of new results with such requirements, as can

> later be met easily by a systematic treatment of the entire field.[51]

Such a systematic treatment, of course, was necessary if Hilbert's ideas regarding his logic and foundation of mathematics were to find followers. Starting in (1922c) and (1923), Hilbert presents the logical calculus not in the form of *Principia*, but by grouping the axioms governing the different connectives. In (1922c), we find the "axioms of logical consequence," in (1923), "axioms of negation." The first occurrence of axioms for conjunction and disjunction seems to be in a class taught jointly by Hilbert and Bernays during Winter 1922–23, and in print in Ackermann's dissertation (Ackermann 1924b). The project of replacing the artificial axioms of *Principia* with more intuitive axioms grouped by the connectives they govern, and the related idea of considering subsystems such as the positive fragment, is Bernays's. In 1918, he had already noted that one could refrain from taking $+$ and $\rightarrow$ as defined symbols and consider the problem of finding a complete axiom system for the positive fragment. The notes to the lecture course from 1922–23 (Hilbert and Bernays 1923a, p. 17) indicate that the material in question was presented by Bernays. In 1923, he gives a talk entitled "The role of negation in propositional logic:"

> In axiomatizing the propositional calculus, the predominant tendency is to reduce the number of basic connectives and therewith the number of axioms. One can also, on the other hand, sharply distinguish the various connectives; in particular, it would be of interest to investigate the role of negation.[52]

The emphasis of separating negation from the other connectives is of course necessitated by Hilbert's considerations on finitism as well.[53]

Full presentations of the axioms of propositional logic are also to be found in (Hilbert 1928a), and in slightly modified form in a course on logic taught by Bernays in 1929–30. The axiom system we find there is almost exactly the one later included in (Hilbert and Bernays 1934).

I. $A \rightarrow (B \rightarrow A)$

$(A \rightarrow (A \rightarrow B)) \rightarrow (A \rightarrow B)$

$(A \rightarrow (B \rightarrow C)) \rightarrow (B \rightarrow (A \rightarrow C))$

$(B \rightarrow C) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$

II. $A \& B \to A$

$A \& B \to B$

$(A \to B) \to ((A \to C) \to (A \to B \& C))$

III. $A \to A \vee B$

$B \to A \vee B$

$(B \to A) \to ((C \to A) \to (B \vee C \to A))$

IV. $(A \sim B) \to (A \to B)$

$(A \sim B) \to (B \to A)$

$(A \to B) \to ((B \to A) \to (A \sim B))$

V. $(A \to B) \to (\overline{B} \to \overline{A})$

$(A \to \overline{A}) \to \overline{A}$

$A \to \overline{\overline{A}}$

$\overline{\overline{A}} \to A$[54]

The algebraic perspective, evident only a few years earlier by the adoption of associativity, commutativity, and distributivity as axioms in some way or other, is completely lacking here. On the other hand, the influence of Frege is palpable in groups I, IV, and V. In (1927), Bernays claims that the axioms in groups I–IV provide an axiomatization of the positive fragment, and raises the question of a decision procedure. This is where he first follows up on his his claim in (1918) that such an axiomatization is possible.

2.6 Lasting Influences

Let me now summarize the advances made by Bernays and Hilbert and try to put them in the historical context of the development of mathematical logic and the foundations of mathematics. The most important of these contributions are certainly the distinction between syntax and semantics, the formulation of syntactic and semantic completeness, the proof of completeness for the propositional calculus, and the proof of decidability.

The history of the the concept(s) of completeness of an axiomatic system has yet to be written. The need for such a history, however, is apparent; completeness is the most fundamental property—alongside consistency—that an axiom system can have, and proofs of completeness and incompleteness of some kind or another count among the most celebrated results of mathematical logic. One need only mention the names of Gödel and Tarski in this connection to illustrate its importance. Although I cannot undertake the task of providing this history here, I want to indicate some of its milestones, since the work of Hilbert and Bernays I have been discussing is probably among the most important.

As we have seen, one of the roots of completeness as a property of axiom systems is the completeness axiom that Hilbert introduced in (1900b). The axiom was not present in the first edition of *Foundations of Geometry*, but was included in the French translation of 1900, and then in the second German edition of 1903.[55] In the lectures from 1905 and again in "Axiomatic thought" (Hilbert 1918a) the axiom was formulated as the requirement that the addition of entities (numbers) to a model of the axioms would result in inconsistencies. In 1906, writing in Göttingen, Johannes Mollerup discusses Hilbert's axiomatization of the reals, and—without explicitly criticizing Hilbert on this issue—shifts the focus from completeness as something to be stipulated to something to be proven. He writes: "So we have two requirements for an axiom system, namely *first* an arithmetical requirement of consistency, and *second* a set-theoretic requirement of completeness" (Mollerup 1907, p. 237). König (1914, p. 209) also criticizes Hilbert's use of the completeness axiom, stating that "the 'completeness axiom' is an intuition we should come to have of a completed thought-system; 'completeness' is an assumption which cannot even be formulated as an 'axiom' in our synthesis; just as the assumption of consistency cannot be so formulated."[56]

Hilbert did not address or acknowledge these criticism explicitly, and the completeness axiom survives in subsequent editions of the *Foundations of Geometry*. In the lectures from 1917–18, however, completeness is first formulated as a property of the propositional calculus in the form: whenever a hitherto non-derivable formula is added to the system, the system becomes inconsistent. The shift from talking about adding elements to talking about adding formulas (new axioms) may be explained in one of two ways: Possibly Hilbert and

Bernays agreed that completeness should not be formulated as an axiom, but should be a property which one should prove about the system, and then the formulation corresponding to Post completeness seems to be the straightforward adaptation. On the other hand, if we take into account that the "elements" described by an axiom for propositional logic *are* propositions, then Post completeness says about propositions exactly the same thing that the completeness axiom says about the reals. Such an interpretation of propositions as the "things" that an axiom system is about is actually hinted at by Hilbert (1905a, pp. 257–58), and is supported by Hilbert's comparison of Russell's axiom of reducibility to the completeness axiom in 1918 (reported in Mancosu (1999a, §8)).

By 1921 at least, Hilbert is well aware of the difference between the requirement expressed by the completeness axiom and completeness of axiom system in the syntactic sense, which is equivalent to the requirement that it proves or refutes every formula of the language.[57] The latter requirement is obviously closely related to Hilbert's "no ignorabimus," the conviction that every well-posed mathematical question can be answered positively or negatively. Where and how does the shift from the completeness axiom to the question of completeness *of* the axioms occur? Was it the recognition that in the context of logic the two amount to the same, and that syntactic completeness also makes the informal question of completeness ("We will have to require that all other facts of the area in question are consequences of the axioms."[58]) more precise? I cannot give an answer to this interesting and important question here. The issues are complicated enough to warrant their own extended treatment.[59]

The question of semantic completeness arises only when one makes a clear distinction between syntax and semantics. In 1917, Hilbert is still heavily influenced by Russell and Whitehead's *Principia*, and the influence is clearly visible in the lecture notes from 1917–18. But already there, Hilbert brings his view of axiomatics to bear: Derivation rules are formulated with more care, the expressions of the system are defined recursively, and we find metatheoretical results stated and proved which Russell and Whitehead considered misplaced because they could not be formulated within the system. But, as Sieg (1999) points out, the axiom systems still come with a built-in interpretation, as it were. Bernays (1918) makes the division between syntax and semantics complete.[60] The axioms and rules are stated purely formally—the study of the axiom system would be idle, were it not possible to give a "logical interpretation." This logical interpretation is precisely truth-value semantics for the propositional calculus. Now the semantical concept of completeness arises naturally: every valid formula is formally derivable.

The main application of the completeness proof, besides establishing that the propositional calculus provides an adequate formalization of the domain of propositional logic, is that of its decidability. The decision problem is vaguely formulated in (Hilbert 1900b) and (1918a), but in Bernays's *Habilitationsschrift* we find a model example of what a decision procedure looks like. This procedure serves as the model for subsequent attacks on the *Entscheidungsproblem*. For these attacks, however, axiomatics is put aside in favor of semantic methods. Behmann seems to be the first to state the decision problem explicitly:

> A general [set of] instructions shall be exhibited, according to which the *correctness or falsity* of an arbitrary given claim, which can be formulated with logical means, can be decided after a finite number of steps; this aim shall be realized at least within the bounds—which are to be determined exactly—within which its realization is in fact possible.[61]

The decision problem was, of course, another great problem on which Hilbert's students were working fervently in the 1920s. We have seen how the early work by Bernays and Hilbert in 1917–18 provides a paradigm for the solution. A decision procedure should be a determinate method to answer, in a finite number of steps, whether a logical formula is provable. But one should not forget that Bernays's decision procedure not only provides a model for what *kind* of result was to be proved, but *how* it should be proved. The method of transformation to normal forms, which was used by Behmann, Schönfinkel, and ultimately Gödel, can be traced back to Bernays's *Habilitationsschrift* (Bernays 1918) and Hilbert's 1905 lectures. With the semantic completeness and the work of Behmann (1922a), a shift towards semantic methods occurred, which was foreshadowed by semantic procedures for deciding validity and equivalence of propositional formulas in Hilbert (1920a).

It is not until 1928 that completeness resurfaces. At the Congress of Mathematicians in Bologna, Hilbert poses the syntactical completeness of arithmetic and the semantic completeness of first order logic as problems of the foundations of mathematics (Hilbert 1928b, 1929).[62] (In the 1917–18 lectures, it was already conjectured that the function calculus was not Post complete. This was subsequently proved by Ackermann.) Completeness of first-order logic was also posed as a question in the book with Ackermann (Hilbert and Ackermann 1928). The question is solved a year later in Gödel's dissertation (Gödel 1929, Gödel 1930).[63]

The metalogical investigations of Bernays on independence and axioms versus rules in 1918 laid the groundwork for several later developments. On the one hand, they provided a rigorous justification for the "algebraic" methods of manipulating formulas (e.g., of "applying" the law of associativity to subformulas) that were used as the official formulation of propositional logic until about 1923. At that time the strictures of Hilbert's developing finitism made it clear that distinctions must be made between the unproblematic connectives (disjunction, conjunction, and in particular the conditional, or "consequence"), and the problematic part, namely negation and the quantifiers. Here, too, Bernays's investigations helped satisfy these strictures by separating the axioms for the unproblematic notions from those for the problematic ones. In the notation of *Prinicipia*, this would not have been possible: there, the unproblematic notion of consequence was even defined in terms of negation (and disjunction).

The development of clear and intuitive axioms for propositional logic, and the investigations of the extent to which axioms can be replaced by rules undoubtedly also had great influence on Gentzen's development of natural deduction and the sequent calculus. Bernays was still teaching in Göttingen at the time when Gentzen was preparing his thesis (Gentzen 1934), and in all likelihood was in close contact with him. Bernays was working closely with Paul Hertz throughout the 1920s, and Hertz's work on axiom systems is commonly acknowledged to be one of Gentzen's main sources.[64] The picture is far from complete, however, and it seems well worth filling in the details. In the course of this, in particular in a reexamination of Hertz's work on logic, it may well be that further important contributions by Bernays may come to light.

Notes

1. Hilbert issued the invitation in September 1917 at the occasion of Hilbert's talk on axiomatic thought in Zürich. Reid (1970, p. 151) reports that the invitation was made in the Spring of 1917. Bernays, however, reported the former version in (Bernays 1976b), in an interview on 25 July 1977

(Bernays 1977), and even in a letter to Reid (27 November 1968, Bernays Nachlaß, WHS, ETH Zürich, Hs 975.3775). As far as I can see, there is no evidence for Reid's version.

2. See also Abrusci (1989).

3. Pierce, Wittgenstein, and Post are commonly credited with the truth-table method of determining propositional validity; Post for the completeness of propositional calculus; and Pierce, Post, and Łukasiewicz for the invention of many-valued logics. The method of using many-valued matrices for independence proofs was also discovered independently by Łukasiewicz and Tarski.

4. "[Sie] hatte zwar durchaus mathematischen Character, aber so die damalige Auffassung war die, dass man diese Grundlagenuntersuchungen, die an die mathematische Logik anknüpften, dass man die mathematisch nicht für voll genommen hat, nicht wahr, ja, das ist ja so ganz nett, das ist so halb spielerisch, nicht wahr, [...] und ich war auch so in der Tendenz [...] und habe das sozusagen auch nicht so ganz für voll genommen, und da [...] hatte ich keinen solchen Eifer, das rechtzeitig zu publizieren, und das ist erst sehr viel später, und doch eigentlich nicht ganz vollständig, sondern bloss mit gewissen Partien herausgekommen [...] so ist das, ist Manches zum Beispiel in den Darstellungen der Entwicklung der mathematischen Logik ist das zum Teil nicht, nicht wahr, entsprechend zum Ausdruck gekommen, was ich da in dieser Arbeit hatte." Interview, 25 July 1977 (Bernays 1977); also reported by Specker (1979). All translations are mine except where English translations are noted in the bibliography. It might be interesting to list some historical accounts and how they treat Bernays. Jørgensen (1931), who in other respects provides a very comprehensive account of the developments in symbolic logic up to 1930, mentions neither Post nor Bernays in connection with completeness or independence results. Kneale and Kneale (1962) treat Bernays's independence proofs in depth and give his completeness proof, but credit it to Post. Bocheński (1956) mentions Post in connection with the decision procedure for propositional logic, but does not mention Bernays. Church (1956) cites Bernays's results on dependence and independence, but does not mention him in connection with consistency, completeness, or decidabilty. Surma (1973) makes no mention of Bernays at all.

5. "Die Paradoxien, die wir im voranstehenden kennen gelernt haben, zeigen zur Genüge, dass eine Prüfung und Neuaufführung der Grundlagen der Mathematik und Logik unbedingt nötig ist." (Hilbert 1905a), p. 215

6. A marginal note on p. 224 instructs: "write more simply = 'equal' ".

7. (Hilbert 1905a), pp. 225–228.

8. "We may think of 0—if we want to proceed intuitively—as the proposition which 'expres-

ses nothing' and which therefore is the ideally correct one; we may call every proposition identical to 0 a *correct* [richtige] or maybe better *non-contradictory* [widerspruchslose] proposition ..." (Hilbert 1905a), p. 226

9. "Es müßte nun untersucht werden, wie weit die Axiome von einander unabhängig sind [...] Das Wichtigste aber wäre hier der Nachweis, dass die 12 Axiome sich nicht widersprechen, d.h. daß man aus ihnen durch die festgelegten Proceße [sic] keine Aussage herleiten kann, die den Axiomen widerspricht, $X+\overline{X}\equiv 0$ etwa. Das sind alles hier nur Andeutungen, die noch keineswegs vollkommen durchgeführt sind, und man hat in Einzelheiten noch sehr viel freie Hand; überhaupt liefert dieser ganze Abschnitt vorläufig eigentlich mehr Materialien zu einer endgültigen Lösung der interessierenden Fragen, als eine endgültige Lösung von ihnen." (Hilbert 1905a), pp. 230–31

10. The notation $X \mid Y$ was introduced in (Hilbert 1905c), this is changed to $X \rightarrow Y$ in a marginal note on p. 236. The influence of Frege is obvious here: "b follows from a" is motivated as excluding the second of the four possibilities: $a+b$, $a+\overline{b}$, $\overline{a}+b$, $\overline{a}+\overline{b}$, compare *Begriffsschrift* Frege (1879), §5.

11. When proving a similar normal form theorem for the calculus in (Hilbert 1918c, p. 149), the fact that normal forms are not unique is pointed out in a footnote. Even if the procedure outlined by Hilbert were deterministic and would thus produce unique normal forms for every formula, different formulas may still have different normal forms, a fact which will become important below.

12. "Eine Aussage Y folgt aus einer andern X dann und nur dann, wenn sie von der From $A\cdot X$ ist, wo A irgend eine Aussage ist. Schliessen heisst richtige Aussagen mit irgend welchen Aussagen multiplizieren." (Hilbert 1905a), p. 246.

13. "Ich will hier noch auf eine, wohl die wichtigste Anwendung der Normalform einer Aussage und ihrer Eindeutigkeit hinweisen. Wir wollen—und darauf müssen wir und zunächst beschränken—eine endliche Anzahl von Aussagen $a, b, c\ldots$ (Axiome über die behandelten Dinge oder Eigennamen) zu Grund legen. Dann kann es überhaupt nur endlich viele Aussagen darüber (d.h. aus diesen Grundaussagen zusammengesetzte Aussagen) geben; denn jede läßt sich auf eine Summe von Produkten im wesentlichen eindeutig bringen, wo in jedem Summand dieselbe Grundaussage nur in der ersten Dimension erscheinen und dasselbe Produkt auch nur einmal als Summand auftreten kann. Jede richtige Aussage muß aus der Summe der Axiome $a+b+\cdots$ durch einen gewissen Multiplikator A folgen (Beweis), und für dieses A gibt es nach dem gesagten auch nur endlich viele Formen. So ergibt sich hier, daß für jeden Satz nur *endlich viele Beweismöglichkeiten* existieren, und wir haben damit in dem vorliegenden primitivsten Falle das alte Problem gelöst, daß jedes rich-

tige Resultat sich durch einen *endlichen Beweis* erzielen lassen muß. Dies Problem war eigentlich der Ausgangspunkt aller meiner Untersuchungen auf unserem Gebiete und die Erledigung dieses Problemes im allerallgemeinsten Falle der Beweis, daß es in der Mathematik kein "Ignorabimus" geben kann, muß auch das letzte Ziel bleiben." (Hilbert 1905a), pp. 248–9.

14. This correction is made by Hilbert later in the lectures (p. 257), see also Peckhaus (1990, pp. 70–72).

15. Of course it would be enough to know that there are only finitely many normal forms, and we can check all of these. But Hilbert does not have a deterministic and finite procedure to produce all these.

16. I do not mean to suggest that Hilbert was not interested in the foundations of mathematics during this period. Sieg (1999) has pointed out that Hilbert lectured a number of times on foundations of mathematics and physics during that time. These lectures, however, contain far less of logical interest than those of 1905 or those after 1917; most of them were courses on "elementary mathematics from a higher standpoint," a topic on which Klein had also often lectured. Even though Hilbert may not himself have worked much on the subject, there is a lot of activity in foundations of mathematics in Göttingen at the time, as the list of lectures in the Mathematical Society published in the *Jahresberichte der Deutschen Mathematiker-Vereinigung* shows. Mancosu (1999a) gives a survey of the developments going on in the early 1910s. He stresses in particular the role of Heinrich Behmann in introducing the mathematicians in Göttingen to the *Principia Mathematica*.

17. (Hilbert 1918c), call number 6817a.44a

18. Heinrich Behmann was completing his dissertation entitled "The antinomy of transfinite number and its solution by Russell and Whitehead" [*Die Antinomie der transfiniten Zahl und ihre Auflösung durch Russell und Whitehead*] under Hilbert in the Spring of 1918 (see Mancosu (1999a)); it would be interesting to compare it with the presentation of the paradoxes and type theory in the 1917–18 lectures.

19. "Man beschränke den Bereich der Aussagen, indem man überhaupt nur die beiden Aussagen 0 und 1 zulässt, und deute dementsprechend die Gleichungen als eigentliche Identitäten. Ferner definiere man Summe und Produkt durch die 8 Gleichungen [...] welche dadurch characterisiert sind, dass sie in richtige arithmetische Gleichungen übergehen, sofern man die symbolische Summe durch den Maximalwert der Summanden und das symbolische Produkt durch den Minimalwert der Faktoren ersetzt. Als Gegenteil der Aussage 0 erkläre man die Aussage 1 und als Gegenteil von 1 die Aussage 0.

Diese Definitionen führen jedenfalls zu keinem Widerspruch, da in jeder von ihnen ein neues Zeichen erklärt wird. Andererseits kann man durch endlich viele Versuche feststellen, dass bei den getroffenen Festsetzungen allen Axiomen I–XII Genüge geleistet wird. Diese Axiome können daher gleichfalls keinen Widerspruch ergeben. So lässt sich für unseren Kalkül die Frage der Widerspruchslosigkeit vollkommen zur Entscheidung bringen." (Hilbert 1918c), p. 70

20. Where Hilbert (1918c) uses '=', Bernays (1918) uses the Russellian '$\sim$'.

21. The use of substitution is indicated at the beginning of *2. A substitution rule was explicitly included in the system of Russell (1906), and Russell also acknowledged its necessity later (e.g., in the introduction to the second edition of *Principia*). For a discussion of the origin of the propositional calculus of *Principia* and the tacit inference rules used there, see O'Leary (1988).

22. This becomes clear from Bernays (1918), who makes a point of distinguishing between correct and provable formulas, in order "to avoid a circle." In (Hilbert 1920a, p. 8), we read: "It is now the first task of logic to find those combinations of propositions, which are always, i.e., without regard for the content of the basic propositions, *correct*."

23. This rule is tacitly used in *Principia*, but Russell's view that logic is universal prevented him from formulating it as a rule. Replacement "can be proved in each separate case, but not generally [...]" (Whitehead and Russell 1910, p. 115).

24. (Hilbert 1918c), p. 144. There is actually a gap in the proof. Hilbert argues that since multiple substitutions can be reduced to successive single substitutions, only the cases where $\varphi(\alpha)$ is $\overline{\alpha}$, $\alpha\gamma$ and $\gamma\alpha$ need to be considered. Somewhere, however, induction has to play a role. What should be done is to prove that whenever $\alpha \to \beta$ and $\beta \to \alpha$ is provable then so are $\overline{\alpha} \to \overline{\beta}$, $\overline{\beta} \to \overline{\alpha}$, $\alpha\gamma \to \gamma\alpha$ and $\gamma\alpha \to \gamma\beta$, and then argue by induction on the depth of the occurrence of α in φ. Compare in this regard Post's (1921, p. 170) proof of essentially the same result; his proof uses induction on the complexity of formulas.

25. "Dieses System von Axiomen wäre als widerspruchsvoll zu bezeichen, falls sich daraus zwei Formeln ableiten liessen, die zueinander in der Beziehung des Gegemteils stehen." (Hilbert 1918c), p. 150.

26. "Wenden wir uns nun zu der Frage der *Vollständigkeit*. Wir wollen das vorgelegte Axiomen-System vollständig nennen, falls durch die Hinzufügung einer bisher nicht ableitbaren Formel zu dem System der Grundformeln stets ein widerspruchsvolles Axiomensystem entsteht." (Hilbert 1918c, p. 152).

27. "Dieses letzte Axiom trägt einen durchaus allgemeinen Character und ist in jedem Axio-

mensystem irgendwelcher Art in gewisser Form anzufügen; hier ist es, wie wir sehen werden, von ganz besonderer Bedeutung. Das Zahlensystem soll nach ihm so beschaffen sein, daß bei jeder Anfügung neuer Elemente Widersprüche auftreten, was für Festsetzungen man auch über sie treffe; lassen sich Dinge angeben die sich widerspruchslos anfügen lassen, so müssen sie dem Systeme in Wahrheit schon angehören." (Hilbert 1905a), p. 17

28. The possibility for such a move was of course already implicit in Hilbert's earlier writings on the foundations of geometry.

29. "Das aufgestellte Axiomen-System könnte kein besonderes Interesse beanspruchen, wenn es nicht einer bedeutsamen inhaltlichen Interpretation fähig wäre.

Eine solche Interpretation ergibt sich auf folgende Art:

Die Variablen fasse man als Symbole für *Aussagen* (Sätze) auf.

Als characteristische Eigenschaft der Aussagen soll angesehen werden, dass sie entweder wahr oder falsch und nicht beides zugleich sind.

Das symbolische Produkt deute man als die Verknüpfung zweier Aussagen durch 'oder,' wobei diese Verknüpfung nicht im Sinne der eigentlichen Disjunktion zu verstehen ist, welche das Zusammenbestehen der beiden Aussagen ausschliesst, sondern vielmer derart, dass 'X oder Y' dann und nur dann zutrifft (d.h. wahr ist), wenn mindestens eine der beiden Aussagen X, Y zutrifft." (Bernays 1918), pp. 3–4.

30. "Die Bedeutsamkeit unseres Axiomen-Systems für die Logik beruht nun auf folgender Tatsache: Versteht man unter einer 'beweisbaren' Formel eine solche, die sich gemäss den Axiomen als richtige Formel erweisen lässt [footnote: Den Begriff der beweisbaren Formel neben dem der richtigen Formel (welcher nicht vollständig abgegrenzt ist) einzuführen, erscheint mir zur Vermeidung eines Zirkels als notwendig.], und unter einer 'allgemeingültigen' Formel eine solche, die im Sinne der angegebenen Deutung bei beliebiger Wahl der für die variablen einzusetztenden Aussagen (also für beliebige 'Werte' der Variablen) stets eine wahre Aussage ergibt, so gilt der Satz:

Jede beweisbare Formel ist eine allgemeingültige Formel und umgekehrt.

Was zunächst die erste Hälfte dieser Behauptung betrifft, so lässt sie sich folgendermassen begründen: Man verifiziert zuerst, dass sämtliche Grundformeln allgemeingültige Formeln sind. Hierzu hat man nur endlich viele Fälle auszuprobieren, denn die Ausdrücke des Kalküls sind alle von der Art, dass bei der logischen Interpretation ihre Wahrheit und Falschheit eindeutig bestimmt ist, wenn von jeder der für die Variablen einzusetzenden Aussagen feststeht, ob sie wahr oder falsch ist, während im übrigen der Inhalt dieser Aussagen gleichgültig ist, sodass man als Wert der Variablen anstatt der Aussagen nur Wahrheit und Falschheit zu betrachten braucht." (Bernays 1918), p. 6.

31. "Diese Betrachtung enthält nicht allein den Beweis für die Vollständigkeit unseres Axiomen-Systems, sondern sie liefert uns überdies noch ein einheitliches Verfahren, durch welches man bei jedem Ausdruck des Kalküls nach endlich vielen Anwendungen der Axiome entscheiden kann, ob er eine beweisbare Formel ist oder nicht. Zum Zweck dieser Entscheidung braucht man nur für den betreffenden Ausdruck eine Normalform zu bestimmen und nachzusehen, ob darin bei jedem der einfachen Produkte mindestens eine Variable sowohl unüberstrichen wie überstrichen als Glied vorkommt. Trifft dies zu, so ist der untersuchte Ausdruck eine beweisbare Formel, andernfalls ist er es nicht. Der Kalkül lässt sich demnach vollkommen trivialisieren." (Bernays 1918), pp. 15–16.

32. For biographical information on Post and his influences, see Davis (1994). Davis (1995) points out that some of the clarifications that Hilbert and Bernays achieved, e.g., the distinction between syntax and semantics, correct and provable formulas, and between theorems about the calculus and theorems in the calculus, were also seen by Lewis (1918), who strongly influenced Post. In fact, the last of the distinctions just mentioned is emphasized by Post.

33. Recall that the axioms investigated by Hilbert and Bernays are not precisely the axioms of *Principia*. While Hilbert and Bernays augment the axiom system with an unrestricted substitution rule, Post's substitution rule allows only substitution of formulas containing one connective.

34. "In der Tat lassen sich ja alle Beziehungen zwischen Wahrheit und Falschheit von Aussagen mit Hülfe der Konjunktion ('und'), der Disjunktion (ausschliessendes 'oder') und der Negation, also auch durch die Symbolik unseres Kalküls zum Ausdruck bringen, und sofern solche Beziehungen für beliebige Aussagen gelten, müssen die ihnen entsprechenden symbolischen Ausdrücke in dem definierten Sinne allgemeingültige Formeln sein." (Bernays 1918), p. 16.

35. "Hilbert hat ja da [an den *Grundlagen der Mathematik*] eigentlich nicht mitgearbeitet, was da benutzt wurde waren sehr viele Gedanken von Hilbert, aber an der Ausgestaltung hat er eigentlich nicht mitgearbeitet, auch schon eigentlich beim ersten Band nicht und beim zweiten schon gar nicht." Interview, 27 August 1977 (Bernays 1977).

36. "[A]lso da war ich damals [Hilberts] Assistent [. . .] [die] Assistentenbeschäftigung, das ist nicht so eine Beschäftigung, wie sie hier [in Zürich] im allgemeinen die Assistenten haben, die den Studenten helfen bei den Übungen, damit hatte ich gar nichts zu tun, sondern das war ganz privatim bei Hilbert, also daß wir einerseits diskutierten über die grundsätzlichen Fragen und dann auch, daß ich ihm für seine Vorlesungen zum Teil half, bei den Vorbereitungen mithalf und Ausarbeitungen machte." Interview, 25 July 1977 (Bernays 1977).

37. Habilitationsakte Paul Bernays, Gemeinsames Prüfungsamt der mathematisch-naturwissenschaftlichen Fakultäten, Universität Göttingen.

38. Bernays Papers, ETH Library/WHS, Hs 976.3.

39. Bernays, "Kolleg von Hilbert über Grundlagen. Zu meiner Göttinger Habilitationsschrift." Unpublished manuscript. Bernays Nachlaß, ETH Zürich Library/WHS, Hs 973.184

40. "Diese Fragen, die die Axiomatik der elementaren Verknüpfungsaussagen betreffen, sind vor wenigen Jahren von dem Göttinger Mathemathiker Bernays behandelt worden (Habilitationsschrift, leider nicht gedruckt) und man kann wohl sagen zu einem vollständigen und befriedigendem Abschluß gebracht worden. [...] Bernays hat die Vollständigkeit ebenfalls streng bewiesen, also gezeigt, daß jede allgemeingültige elementare Verknüpfungsaussage tatsächlich aus den Grundformeln nach den Grundregeln abgeleitet werden kann [...]. " (Behmann 1922b), p. 97.

41. "Damals zum Beispiel ja, auch meine eigenen Kentnisse waren da noch sehr sehr unvollständing, als ich da zunächst, zum Beispiel, 1917, damals. Wie gesagt, zuerst bevor Hilbert direkt an seine Sache wieder ging, die er ja schon viel früher angefangen hatte, da war, da hat er noch nicht gleich darüber gelesen, sondern er hat über mathematische Logik gelesen, eine Vorlesung. Und die hab ich auch ausgearbeitet und die hab ich, nicht wahr, und zwar in solcher Weise, daß ich die freien, das hab ich Ihnen glaub ich erzählt, daß ich die freien Variablen vermieden habe. Bei Russell hatte ich mir so ein bißchen einiges angekuckt aber erstens war mir das überhaupt zu breit diese Art der Behandlung, sagte mir nicht in jeder Hinsicht zu, aber insbesondere hab ich das nicht recht verstanden, was das heisst für alle x, $F(x)$ dann folgt $F(y)$. Tatsächlich ist ja auch die Anwendung der freien Variablen, das ist etwas Technisches, nicht wahr. Es sind eigentlich zwei Arten der Darstellung der Allgemeinheit. Man hat die Allgemeinheit eben einerseits durch die gebunde Variable und andererseits durch die freie Variable. Solch einen Unterschied gibt es nicht in der gewöhnlichen Sprache, nicht wahr. Nun hab ich also da zunächst die freien Variablen vermieden. Das ist ein mögliches Verfahren, das ist auch später wieder von anderen manchmal gemacht worden so. Das ist also eine Vorlesung, die ist ausgearbeitet worden und hat auch nachher da im Hilbertschen, da im Lesezimmer vom Institut gestanden." Interview, 27 August 1977 (Bernays 1977)

42. The use of free variables was also avoided in lectures on the *Logik-Kalkül* (Hilbert 1920a). Free variables are first used in Hilbert's talks of 1922 (Hilbert 1922c, 1923) and in lectures taught by Hilbert and Bernays in 1922–23 (Hilbert and Bernays 1923b, 1923a).

43. See, e.g., Lauener's (1978) testimony.

44. "Es wird also bei jedem der folgenden Unabhängigkeits-Beweise der Kalkül auf ein endli-

ches System (eine endliche Gruppe im weiteren Sinne des Wortes [footnote in text: Das heisst ohne Voraussetzung des assoziativen Gesetzes und der eindeutigen Umkehrbarkeit der Komposition]) zurückgeführt, für dessen Elemente eine Komposition ('symbolisches Produkt') und eine 'Negation' definiert ist, und diese Zurückführung findet in der Weise statt, dass die Variablen des Kalküls auf die Elemente jenes Systems als ihre Werte bezogen werden. Die 'richtigen Formeln' sollen jedesmal dadurch characterisiert sein, dass sie für beliebige Werte der vorkommenden Variablen nur Werte eines gewissen Teilsystems T annehmen." (Bernays 1926), pp. 27–28.

45. The interested reader may consult Kneale and Kneale (1962), pp. 689–694, and, of course, Bernays (1926). The method was discovered independently by Łukasiewicz (1924), who announced results simlar to those of Bernays. Let me remark in passing that Bernays's first system defines Łukasiewicz's 3-valued implication.

46. Gödel (1932) quotes the independence proofs given by Hilbert (1928a).

47. $\overline{X}X$, of course, is the principle of the excluded middle, and is synonymous in the system with $X \to X$.

48. These results extend the method of the previous sections insofar as the independence of rules is also proved. To do this, it is shown that an instance of the premise(s) of a rule always takes designated values, but the corresponding instance of the conclusion does not. This extension of the matrix method for proving independence was later rediscovered by Huntington (1935).

49. According to the *Verzeichnis der Vorlesungen* for the semester, the course was announced under the title "Formal logic and its epistemological value [Formale Logik und ihr erkenntnistheoretischer Wert]." The term lasted 5 January 1920–31 March 1920. Lecture notes by Bernays survive at the library of the Institute of Mathematics at the University of Göttigen (Hilbert 1920a).

50. This is not stated explicitly, but is evident from the derivation on p. 11.

51. Behmann (1922a), p. 167.

52. "Bei der Axiomatisierung des Aussagenkalküls herrscht die Tendenz vor, die Anzahl der Grundverknüpfungen und damit die der Axiome zu reduzieren. Man kann aber andererseits auch die Rolle der verschiedenen Verknüpfungen scharf voneinander sondern; insbesondere ist es von Interesse, die Rolle der Negation zu untersuchen." Talk given in the Mathematical Society at Göttingen, 20 February 1923, as reported in *Jahresbericht der Deutschen Mathematiker-Vereinigung, 2. Abteilung*, vol. 32 (1922), p. 22.

53. Compare, e.g., the logical axioms in (Hilbert 1922c) and (Hilbert 1923). In the latter paper, Hilbert notes: "In (Hilbert 1922c) I had still avoided [the negation sign]; as it turned out, the sign

for ‘not’ can be used in the present, slightly modified presentation of my theory without danger.” (Hilbert 1923, p. 152) He could not have avoided the negation sign if the whole calculus was based on it.

54. Paul Bernays, notes to “Mathematische Logik,” lecture course held Winter semester 1929–30, Universität Göttingen. Unpublished shorthand manuscript. Bernays Nachlaß, WHS, ETH Zürich, Hs 973.212. The signs ‘&’ and ‘$\vee$’ were is first used as signs for conjunction and disjunction in (Hilbert and Bernays 1923b). The third axiom of group I and the second axiom of group V are missing from the system given in (Hilbert and Bernays 1934). The first (*Simp*), third (*Comm*), and fourth axiom (*Syll*) of group I are investigated in the published version of the *Habilitationsschrift* (Bernays 1926), but not in the original version (1918).

55. For a discussion of the history of Hilbert’s *Foundations of Geometry*, and in particular of the completeness axiom, see Toepell (1986, pp. 254–256) and Birkhoff and Bennett (1987).

56. “Ebenso ist das ‘Vollständigkeitsaxiom’ eine an dem fertigen Denkbereiche zu erlangende Anschauung; die ‘Vollständigkeit’ ist ein Postulat, das in unserer Synthese überhaupt nicht als ‘Axiom’ des Denkbereichs gefaßt werden kann; ebensowenig wie das Postulat der Widerspruchslosigkeit.” (König 1914), p. 209. König’s book was known in Göttingen: Felix Bernstein reported on it in the Mathematical Society on 16 February 1915. Compare also Baldus’ (1928) critique of Hilbert’s axiom of completeness.

57. See Hilbert (1922b), pp. 18–19, where both the distinction and the equivalence are pointed out.

58. “Weiters interessiert uns die *Vollständigkeit* des Axiomensystems. Wir werden verlangen müssen, dass alle übrigen Thatsachen des vorgelegten Wissensbereiches Folgerungen aus den Axiomen sind,” (Hilbert 1905a), p. 13.

59. I would like to just mention as two more possible influences the work of the American postulate theorists (Huntington, Veblen) on categoricity (see Corcoran (1980, 1981) and Scanlan (1991)), and the exchange between Husserl and Hilbert on completeness in 1901, recently analyzed by Majer (1997) and Hill (1995).

60. The pivotal role that Bernays (1918) played in the shift from syntactic to semantic completeness is stressed by Moore (1997).

61. “Es soll eine ganz bestimmte allgemeine Vorschrift angegeben werden, die über die Richtigkeit oder Falschheit einer beliebig vorgelegten mit rein logischen Mitteln darstellbaren Behauptung nach einer endlichen Anzahl von Schritten zu entscheiden gestattet, oder zum mindesten dieses

Ziel innerhalb derjenigen — genau festzulegenden — Grenzen verwirklicht werden, innerhalb deren seine Verwirklichung tatsächlich möglich ist.” (Behmann 1922a), p. 166, emphasis mine. This was Behmann’s *Habilitationsschrift*, he received his *venia legendi* in July 1921. Behmann spoke on his results to the mathematical society in Göttingen on 10 May 1921, the talk was entitled “Das Entscheidungsproblem der mathematischen Logik” (*Jahresberichte der Deutschen Mathematiker-Vereinigung, 2. Abteilung*, vol. 30 (1921), p. 47). The manuscript of the talk survives in the Behmann Papers in Erlangen. This seems to be the first documented use of the expression “Entscheidungsproblem.” Behmann had requested leave from his teaching duties in late September 1920 to work on his Habilitation, the problem was probably formulated in its full generality sometime in early to mid-1920.

62. The notion of syntactic completeness of a theory is closely related to what we now call “complete theories.,” i.e., theories which either prove or refute every sentence of the language. Hilbert (1929) proposes the proof of syntactic completeness of arithmetic as a finitistic analog of the proof of completeness in the sense of categoricity. That completeness and categoricity are not the same was realized only with Skolem’s discovery of nonstandard models.

63. See Moore (1997) and Dreben and van Heijenoort (1986) for a discussion of Gödel’s motivations and influences.

64. See in this regard the introduction to Gentzen (1969), and Došen (1993).

Chapter 3

The Practice of Finitism: Epsilon Calculus and Consistency Proofs in Hilbert's Program

3.1 Introduction

Hilbert first presented his philosophical ideas based on the axiomatic method and consistency proofs in the years 1904 and 1905, following his exchange with Frege on the nature of axiomatic systems and the publication of Russell's Paradox. In the text of Hilbert's address to the International Congress of Mathematicians in Heidelberg, we read:

> Arithmetic is often considered to be part of logic, and the traditional fundamental logical notions are usually presupposed when it is a question of establishing a foundation of 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 arithmetic is required if paradoxes are to be avoided.[1]

When Hilbert returned to his foundational work with full force in 1917, he seems at first to have been impressed with Russell's and Whitehead's work in the *Principia*, which—they thought—succeeded in developing large parts of mathematics without using sets. By

1920, however, Hilbert returned to his earlier conviction that a reduction of mathematics to logic is not likely to succeed. Instead, he takes Zermelo's axiomatic set theory as a suitable framework for developing mathematics. He localizes the failure of Russell's logicism in its inability to provide the existence results necessary for analysis:

> The axiomatic method used by Zermelo is unimpeachable and indispensable. The question whether the axioms include a contradiction, however, remains open. Furthermore the question poses itself if and in how far this axiom system can be deduced from logic. [... T]he attempt to reduce set theory to logic seems promising because sets, which are the objects of Zermelo's axiomatics, are closely related to the predicates of logic. Specifically, sets can be reduced to predicates.
>
> This idea is the starting point for Frege's, Russell's, and Weyl's investigations into the foundations of mathematics.[2]

The logicist project runs into a difficulty when, given a second-order predicate S to which a set of sets is reduced, we want to know that there is a predicate to which the union of the sets reduces. This predicate would be $(\exists P)(P(x) \mathbin{\&} S(P))$—$x$ is in the union of the sets in S if there is a set P of which x is a member and which is a member of S.

> We have to ask ourselves, what "there is a predicate P" is supposed to mean. In axiomatic set theory "there is" always refers to a basic domain $\mathfrak{B}$. In logic we could also think of the predicates comprising a domain, but this domain of predicates cannot be seen as something given at the outset, but the predicates must be formed through logical operations, and the rules of construction determine the domain of predicates only afterwards.
>
> From this we see that in the rules of logical construction of predicates reference to the domain of predicates cannot be allowed. For otherwise a *circulus vitiosus* would result.[3]

Here Hilbert is echoing the predicativist worries of Poincaré and Weyl. However, Hilbert rejects Weyl's answer to the problem, viz., restricting mathematics to predicatively acceptable constructions and inferences, as unacceptable in that it amounts to "a return to the prohibition policies of Kronecker." Russell's proposed solution, on the other hand, amounts to giving up the aim of reduction to logic:

> Russell starts with the idea that it suffices to replace the predicate needed for the definition of the union set by one that is extensionally equivalent, and

> which is not open to the same objections. He is unable, however, to exhibit such a predicate, but sees it as obvious that such a predicate exists. It is in this sense that he postulates the "axiom of reducibility," which states approximately the following: "For each predicate, which is formed by referring (once or multiple times) to the domain of predicates, there is an extensionally equivalent predicate, which does not make such reference.
>
> With this, however, Russell returns from constructive logic to the axiomatic standpoint. [. . .]
>
> 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.[4]

With this, Hilbert rejects the logicist position as failed. At the same time, he rejects the restrictive positions of Brower, Weyl, and Kronecker. The axiomatic method provides a framework which can accommodate the positive contributions of Brouwer and Weyl, without destroying mathematics through a Kroneckerian "politics of prohibitions." For Hilbert, the unfettered progress of mathematics, and science in general, is a prime concern. This is a position that Hilbert had already stressed in his lectures before the 1900 and 1904 International Congresses of Mathematics, and which is again of paramount importance for him with the conversion of Weyl to Brouwer's intuitionism.

Naturally, the greater freedom comes with a price attached: the axiomatic method, in contrast to a foundation based on logical principles alone, does not itself guarantee consistency. Thus, a proof of consistency is needed.

3.2 Early Consistency Proofs

Ever since his work on geometry in the 1890s, Hilbert had an interest in consistency proofs. The approaches he used prior to the foundational program of the 1920s were almost always relative consistency proofs. Various axiomatic systems, from geometry to physics, were shown to be consistent by giving arithmetical (in a broad sense, including arithmetic of the reals) interpretations for these systems, with one exception—a prototype of a finitistic consistency proof for a weak arithmetical system in Hilbert (1905c). This was Hilbert's first attempt at a "direct" consistency proof for arithmetic, i.e., one not based on a reduction to another system, which he had posed as the second of his famous list of problems (Hilbert 1900a).

When Hilbert once again started working on foundational issues following the war, the first order of business was a formulation of logic. This was accomplished in collaboration with Bernays between 1917 and 1920 (see Sieg (1999) and Chapter 2 (Zach 1999)), included the establishment of metatheoretical results like completeness, decidability, and consistency for propositional logic in 1917/18, and was followed by ever more nuanced axiom systems for propositional and predicate logic. This first work in purely logical axiomatics was soon extended to include mathematics. Here Hilbert followed his own proposal, made first in 1905,[5] to develop mathematics and logic simultaneously. The extent of this simultaneous development is nowhere clearer than in Hilbert's lecture course of 1921/22, where the ε-operator is first used as both a logical notion, representing the quantifiers, and an arithmetical notion, representing induction in the form of the least number principle. Hilbert realized then that a consistency proof for all of mathematics is a difficult undertaking, best attempted in stages:

> Considering the great variety of connectives and interdependencies exhibited by arithmetic, it is obvious from the start that we will the not be able to solve the problem of proving consistency in one fell swoop. We will instead first consider the simplest connectives, and then proceed to ever higher operations and inference methods, whereby consistency has to be established for each extension of the system of signs and inference rules, so that these extensions do not endanger the consistency [result] established in the preceding stage.
>
> Another important aspect is that, following our plan for the complete formalization of arithmetic, we have to develop the proper mathematical formalism in connection with the formalism of the logical operations, so that—as I have expressed it—a simultaneous construction of mathematics and logic is executed.[6]

Hilbert had rather clear ideas, once the basic tools both of proof and of formalization were in place, of what the stages should be. In an addendum to the lecture course on *Grundlagen der Mathematik*, taught by Hilbert and Bernays in 1922–23,[7] he outlined them. The first stage had already been accomplished: Hilbert gave consistency proofs for calculi of propositional logic in his 1917/18 lectures. Stage II consist in the elementary calculus of free variables, plus equality axioms and axioms for successor and predecessor. The axioms are:

1.	$A \to B \to A$	2.	$(A \to A \to B) \to A \to B$
3.	$(A \to B \to C) \to (B \to A \to C)$	4.	$(B \to C) \to (A \to B) \to A \to C$
5.	$A \& B \to A$	6.	$A \& B \to B$
7.	$A \to B \to A \& B$	8.	$A \to A \vee B$
9.	$B \to A \vee B$	10.	$(A \to C) \to (B \to C) \to A \vee B \to C$
11.	$A \to \overline{A} \to B$	12.	$(A \to B) \to (\overline{A} \to B) \to B$
13.	$a = a$	14.	$a = b \to A(a) \to A(b)$
15.	$a + 1 \neq 0$	16.	$\delta(a+1) = a$[8]

In Hilbert's systems, Latin letters are variables; in particular, *a*, *b*, *c*, ..., are individual variables and *A*, *B*, *C*, ..., are formula variables. The rules of inference are modus ponens and substitution for individual and formula variables.

Hilbert envisaged his foundational project as a stepwise "simultaneous development of logic and mathematics," in which axiomatic systems for logic, arithmetic, analysis, and finally set theory would be developed. Each stage would require a proof of consistency before the next stage is developed. In a handwritten supplement to the typescript of the 1922–23 lecture notes on the foundations of arithmetic, Hilbert presents a rough overview of what these steps might be:

> **Outline**. Stage II was elementary calculation, axioms 1–16.
> Stage III. Now elementary number theory
> Schema for definition of functions by recursion and modus ponens
> will add the schema of induction to modus ponens
> even if this coincides in substance with the results of intuitively obtained number theory, we are now dealing with formulas, e.g, $a + b = b + a$.
> Stage IIII. Transfinite inferences and parts of analysis
> Stage V. Higher-order variables and set theory. Axiom of choice.
> Stage VI. Numbers of the 2nd number class, full transfinite induction. Higher types. Continuum problem, transfinite induction for numbers in the 2nd number class.
> Stage VII. (1) Replacement of infinitely many definitional schemata by one axiom. (2) Analysis and set theory. At level 4, again the full theorem of the least upper bound.
> Stage VIII. Formalization of well ordering.[9]

3.2.1 The Propositional Calculus and the Calculus of Elementary Computation

Step I had been achieved in 1917–18. Already in the lectures from the Winter term 1917/18, Hilbert and Bernays had proved that the propositional calculus is consistent. This was done first by providing an arithmetical interpretation, where they stressed that only finitely many numbers had to be used as "values" (0 and 1). The proof is essentially a modern proof of the soundness of propositional logic: A truth value semantics is introduced by associating which each formula of the propositional calculus a truth function mapping tuples of 0 and 1 (the values of the propositional variables) to 0 or 1 (the truth value of the formula under the corresponing valuation). A formula is called *correct* if it corresponds to a truth function which always takes the value 1. It is then showed that the axioms are correct, and that modus ponens preserves correctness. So every formula derivable in the propositional calculus is correct. Since A and $\overline{A}$ cannot both be correct, they cannot both be derivable, and so the propositional calculus is consistent.

It was very important for Hilbert that the model for the propositional calculus thus provided by $\{0,1\}$ was finite. As such, its existence, and the admissibility of the consistency proof was beyond question. This lead him to consider the consistency proof for the propositional calculus to be the prime example for for a consistency proof *by exhibition* in his 1921/22 lectures on the foundations of mathematics. The consistency problem in the form of a demand for a consistency proof for an axiomatic system which neither proceeds by exhibiting a model, nor by reducing consistency of a system to the consistency of another, but by providing a metamathematical proof that no derivation of a contradiction is possible, is first formulated in lectures in the Summer term of 1920. Here we find a first formulation of an arithmetical system and a proof of consistency. The system consists of the axioms

$$\begin{aligned} 1 &= 1 \\ (a=b) &\to (a+1=b+1) \\ (a+1=b+1) &\to (a=b) \\ (a=b) &\to ((a=c)\to(b=c)). \end{aligned}$$

The notes contain a proof that these four axioms, together with modus ponens, do not allow the derivation of the formula

$$a+1=1.$$

The proof itself is not too interesting, and I will not reproduce it here.[10] The system considered is quite weak. It does not even contain all of propositional logic: negation only appears as inequality, and only formulas with at most two '$\rightarrow$' signs are derivable. Not even $a=a$ is derivable. It is here, nevertheless, that we find the first statement of the most important ingredient of Hilbert's project, namely, proof theory:

> Thus we are led to make the proofs themselves the object of our investigation; we are urged towards a *proof theory*, which operates with the proofs themselves as objects.
>
> For the way of thinking of ordinary number theory the numbers are then objectively exhibitable, and the proofs about the numbers already belong to the area of thought. In our study, the proof itself is something which can be exhibited, and by thinking about the proof we arrive at the solution of our problem.
>
> Just as the physicist examines his apparatus, the astronomer his position, just as the philosopher engages in critique of reason, so the mathematician needs his proof theory, in order to secure each mathematical theorem by proof critique.[11]

This project is developed in earnest in two more lecture courses in 1921–22 and 1922–23. These lectures are important in two respects. First, it is here that the axiomatic systems whose consistency is to be proven are developed. This is of particular interest for an understanding of the relationship of Hilbert to Russell's project in the *Principia* and the influence of Russell's work both on Hilbert's philosophy and on the development of axiomatic systems for mathematics.[12] Sieg (1999) has argued that, in fact, Hilbert was a logicist for a brief period around the time of his paper "Axiomatic Thought" (Hilbert 1918a). However, as noted in Section 3.1, Hilbert soon became critical of Russell's type theory, in particular of the axiom of reducibility. Instead of taking the system of *Principia* as the adequate formalization of mathematics the consistency of which was to be shown, Hilbert proposed a new system. The guiding principle of this system was the "simultaneous development of logic and mathematics"—as opposed to a development of mathematics out of logic—which he had already proposed in Hilbert (1905c, p. 176). The cornerstone of this development is the ε-calculus. The second major contribution of the 1921–22 and 1922–23 lectures are the consistency proofs themselves, including the *Hilbertsche Ansatz* for the ε-substitution method, which were the direct precursors to Ackermann's dissertation of 1924.

In contrast to the first systems of 1920, here Hilbert uses a system based on full propositional logic with axioms for equality, i.e., the elementary calculus of free variables:

I. Logical axioms
a) Axioms of consequence
1) $A \to B \to A$
2) $(A \to A \to B) \to A \to B$
3) $(A \to B \to C) \to B \to A \to C$
4) $(B \to C) \to (A \to B) \to A \to C$

b) Axioms of negation
5) $A \to \overline{A} \to B$
6) $(A \to B) \to (\overline{A} \to B) \to B$
II. Arithmetical axioms
a) Axioms of equality
7) $a = a$
8) $a = b \to Aa \to Ab$
b) Axioms of number
9) $a + 1 \neq 0$
10) $\delta(a+1) = a$[13]

Here, '$+1$' is a unary function symbol. The rules of inference are substitution (for individual and formula variables) and modus ponens.

Hilbert's idea for how a finitistic consistency proof should be carried out is first presented here. The idea is this: suppose a proof of a contradiction is available. We may assume that the end formula of this proof is $0 \neq 0$.

1. *Resolution into proof threads.* First, we observe that by duplicating part of the proof and leaving out steps, we can transform the derivation to one where each formula (except the end formula) is used exactly once as the premise of an inference. Hence, the proof is in tree form.

2. *Elimination of variables.* We transform the proof so that it contains no free variables. This is accomplished by proceeding backwards from the end formula: The end formula contains no free variables. If a formula is the conclusion of a substitution rule,

the inference is removed. If a formula is the conclusion of modus ponens it is of the form

$$\frac{\mathfrak{A} \quad \mathfrak{A} \to \mathfrak{B}}{\mathfrak{B}'}$$

where $\mathfrak{B}'$ results from $\mathfrak{B}$ by substituting terms for free variables. If these variables also occur in $\mathfrak{A}$, we substitute the same terms for them. Variables in $\mathfrak{A}$ which do not occur in $\mathfrak{B}$ are replaced with 0. This yields a formula $\mathfrak{A}'$ not containing variables.[14] The inference is replaced by

$$\frac{\mathfrak{A}' \quad \mathfrak{A}' \to \mathfrak{B}'}{\mathfrak{B}'}$$

3. *Reduction of functionals.* The remaining derivation contains a number of terms (*functionals* in Hilbert's parlance) which now have to be reduced to numerical terms (i.e., standard numerals of the form $(\dots(0+1)+\cdots)+1)$. In this case, this is done easily by rewriting innermost subterms of the form $\delta(0)$ by 0 and $\delta(\mathfrak{n}+1)$ by $\mathfrak{n}$. In later stages, the set of terms is extended by function symbols introduced by recursion, and the reduction of functionals there proceeds by calculating the function for given numerical arguments according to the recursive definition. This will be discussed in the next section.

In order to establish the consistency of the axiom system, Hilbert suggests, we have to find a decidable (*konkret feststellbar*) property of formulas so that every formula in a derivation which has been transformed using the above steps has the property, and the formula $0 \neq 0$ lacks it. The property Hilbert proposes to use is *correctness*. This is not to be understood as truth in a model. The formulas still occurring in the derivation after the transformation are all Boolean combinations of equations between numerals. An equation between numerals $\mathfrak{n} = \mathfrak{m}$ is *correct* if $\mathfrak{n}$ and $\mathfrak{m}$ are syntactically equal, and the negation of an equality is correct of $\mathfrak{m}$ and $\mathfrak{n}$ are not syntactically equal.

> If we call a formula which does not contain variables or functionals other than numerals an "*explicit [numerical] formula*", then we can express the result obtained thus: Every provable explicit [numerical] formula is end formula of a proof all the formulas of which are explicit formulas.
>
> This would have to hold in particular of the formula $0 \neq 0$, if it were provable. The required proof of consistency is thus completed if we show that there can be no proof of the formula which consists of only explicit formulas.

> To see that this is impossible it suffices to find a concretely determinable [*konkret feststellbar*] property, which first of all holds of all explicit formulas which result from an axiom by substitution, which furthermore transfers from premises to end formula in an inference, which however does not apply to the formula $0 \neq 0$.[15]

Hilbert now defines the notion of a (conjunctive) normal form and gives a procedure to transform a formula into such a normal form. He then provides the wanted property:

> With the help of the notion of a normal form we are now in a position to exhipit a property which distinguishes the formula $0 \neq 0$ from the provable explicit formulas.
>
> We divide the explicit formulas into "*correct*" and "*incorrect*." The explicit atomic formulas are equations with *numerals* on either side [of the equality symbol]. We call such an *equation correct*, if the numerals on either side *coincide*, otherwise we call it *incorrect*. We call an *inequality* with numerals on either side *correct* if the two numerals are *different*, otherwise we call it *incorrect*.
>
> In the normal form of an arbitrary explicit formula, each disjunct has the form of an equation or an inequality with numerals on either side.
>
> We now call a *general explicit formula correct* if in the corresponding normal form each disjunction which occurs as a conjunct (or which constitutes the normal form) contains a correct equation or a correct inequality. Otherwise we call the formula *incorrect*. [...]
>
> According to this definition, the question of whether an explicit formula is correct or incorrect is *concretely decidable* in every case. Thus the "tertium non datur" holds here...[16]

This use in the 1921–22 lectures of the conjunctive normal form of a propositional formula to define correctness of Boolean combinations of equalities between numerals goes back to the 1917–18 lecture notes,[17] where transformation into conjunctive normal form and testing whether each conjunct contains both A and $\overline{A}$ was proposed as a test for propositional validity. Similarly, here a formula is *correct* if each conjunct in its conjunctive normal form contains a correct equation or a correct inequality.[18] In the 1922–23 lectures, the definition involving conjunctive normal forms is replaced by the usual inductive definition of propositional truth and falsehood by truth tables (Hilbert and Bernays 1923a, p. 21). Armed with the definition of correct formula, Hilbert can prove that the derivation resulting from a proof by transforming it according to (1)–(3) above contains only correct formulas.

Since $0 \neq 0$ is plainly not correct, there can be no proof of $0 \neq 0$ in the system consisting of axioms (1)–(10). The proof is a standard induction on the length of the derivation: the formulas resulting from the axioms by elimination of variables and reduction of functionals are all correct, and modus ponens preserves correctness.[19]

3.2.2 Elementary Number Theory with Recursion and Induction Rule

The system of stage III consists of the basic system of the elementary calculus of free variables and the successor function, extended by the schema of defining functions by primitive recursion and the induction rule.[20] A primitive recursive definition is a pair of axioms of the form

$$\begin{aligned} \varphi(0, b_1, \ldots, b_n) &= \mathfrak{a}(b_1, \ldots, b_n) \\ \varphi(a+1, b_1, \ldots, b_n) &= \mathfrak{b}(a, \varphi(a), b_1, \ldots, b_n) \end{aligned}$$

where $\mathfrak{a}(b_1, \ldots, b_n)$ contains only the variables $b_1, \ldots, b_n$, and $\mathfrak{b}(a, c, b_1, \ldots, b_n)$ contains only the variables a, c, $b_1, \ldots, b_n$. Neither contains the function symbol φ or any function symbols which have not yet been defined.

The introduction of primitive recursive definitions and the induction rule serves, first of all, the purpose of expressivity. Surely any decent axiom system for arithmetic must provide the means of expressing basic number-theoretic states of affairs, and this includes addition, subtraction, multiplication, division, greatest common divisor, etc. The general schema of primitive recursion is already mentioned in the Kneser notes for 1921–22 (Hilbert 1922a, Heft II, p. 29), and is discussed in some detail in the notes for the lectures of the following year (Hilbert and Bernays 1923a, pp. 26–30).

It may be interesting to note that in the 1922–23 lectures, there are no axioms for addition or multiplication given before the general schema for recursive definition. This suggests a change in emphasis during 1922, when Hilbert realized the importance of primitive recursion as an arithmetical concept formation. He later continued to develop the notion, hoping to capture all number theoretic functions using an extended notion of primitive recursion and to solve the continuum problem with it. This can be seen from the attempt at a proof of the continuum hypothesis in (1926), and Ackermann's paper on "Hilbert's construction of the reals" (1928a), which deals with hierarchies of recursive functions. The general outlook in this regard is also markedly different from Skolem's (1923), which is usually credited with the definition of primitive recursive arithmetic.[21]

Hilbert would be remiss if he would not be including induction in his arithmetical axiom systems. As he already indicates in the 1921–22 lectures, however, the induction principle cannot be formulated as an axiom without the help of quantifiers.

> We are still completely missing the axiom of complete induction. One might think it would be
>
> $$\{Z(a) \to (A(a) \to A(a+1))\} \to \{A(1) \to (Z(b) \to A(b))\}$$
>
> That is not it, for take $a = 1$. The hypothesis must hold for *all* a. We have, however, no means to bring the *all* into the hypothesis. Our formalism does not yet suffice to write down the axiom of induction.
>
> But as a schema we can: We extend our methods of proof by the following schema.
>
> $$\frac{\mathfrak{K}(1) \quad \mathfrak{K}(a) \to \mathfrak{K}(a+1)}{Z(a) \to \mathfrak{K}(a)}$$
>
> Now it makes sense to ask whether this schema can lead to a contradiction.[22]

The induction schema is thus necessary in the formulation of the elementary calculus only because quantifiers are not yet available. Subsequently, induction will be subsumed in the ε-calculus.

The consistency proof for stage II is extended to cover also the induction schema and primitive recursive definitions. Both are only sketched: Step (3), reduction of functionals, is extended to cover terms containing primitive recursive functions by recursively computing the value of the innermost term containing only numerals. Both in the 1921–22 and the 1922–23 sets of notes by Kneser, roughly a paragraph is devoted to these cases (the official sets of notes for both lectures do not contain the respective passages).

> How do we proceed for recursions? Suppose a $\varphi(\mathfrak{z})$ occurs. Either [$\mathfrak{z}$ is] 0, then we replace it by $\mathfrak{a}$. Or [it is of the form] $\varphi(\mathfrak{z}+\mathfrak{1})$: [replace it with] $\mathfrak{b}(\mathfrak{z},\varphi(\mathfrak{z}))$. Claim: These substitutions eventually come to an end, if we replace innermost occurrences first.[23]

The claim is not proved, and there is no argument that the process terminates even for terms containing several different, nested primitive recursively defined function symbols.

For the induction schema, Hilbert hints at how the consistency proof must be extended. Combining elimination of variables and reduction of functionals we are to proceed upwards in the proof as before until we arrive at an instance of the induction schema:

$$\frac{\mathfrak{K}(1) \quad \mathfrak{K}(a) \to \mathfrak{K}(a+1)}{Z(\mathfrak{z}) \to \mathfrak{K}'(\mathfrak{z})}$$

By copying the proof ending in the right premise, substituting numerals 1, ..., $\mathfrak{y}$ (where $\mathfrak{z} = \mathfrak{y} + 1$) for a and applying the appropriate substitutions to the other variables in $\mathfrak{K}$ we obtain a proof of $Z(\mathfrak{z}) \to \mathfrak{K}'(\mathfrak{z})$ without the last application of the induction schema.

With the introduction of the ε-calculus, the induction rule is of only minor importance, and its consistency is never proved in detail until Hilbert and Bernays (1934, pp. 298–99).

3.2.3 The ε-Calculus and the Axiomatization of Mathematics

In the spirit of the "simultaneous development of logic and mathematics," Hilbert takes the next step in the axiomatization of arithmetic by employing a principle taken from Zermelo's axiomatization of set theory: the axiom of choice. Hilbert and Bernays had dealt in detail with quantifiers in lectures in 1917–18 and 1920, but they do not directly play a significant role in the axiom systems Hilbert develops for mathematics. Rather, the first- and higher-order calculi for which consistency proofs are proposed, are based instead on choice functions. The first presentation of these ideas can be found in the 1921–22 lecture notes by Kneser (the official notes do not contain these passages). The motivation is that in order to deal with analysis, one has to allow definitions of functions which are not finitary. These concept formations, necessary for the development of mathematics free from intuitionist restrictions, include definition of functions from undecidable properties, by unbounded search, and choice.

> Not finitely (recursively) defined is, e.g., $\varphi(a) = 0$ if there is a b so that $a^5 + ab^3 + 7$ is prime, and $= 1$ otherwise. But only with these numbers and functions the real mathematical interest begins, since the solvability in finitely many steps is not foreseeable. We have the conviction, that such questions, e.g., the value of $\varphi(a)$, are solvable, i.e., that $\varphi(a)$ is also finitely definable. We cannot wait on this, however, we must allow such definitions for otherwise we would restrict the free practice of science. We also need the concept of a function of functions.[24]

The concepts which Hilbert apparently takes to be fundamental for this project are the principle of the excluded middle and the axiom of choice, in the form of second-order functions τ and α. The axioms for these functions are

1. $\tau(f) = 0 \rightarrow (Z(a) \rightarrow f(a) = 1)$
2. $\tau(f) \neq 0 \rightarrow Z(\alpha(f))$
3. $\tau(f) \neq 0 \rightarrow f(\alpha(f)) \neq 1$
4. $\tau(f) \neq 0 \rightarrow \tau(f) = 1$

The intended interpretation is: $\tau(f) = 0$ if f is always 1 and $= 1$ if one can choose an $\alpha(f)$ so that $f(\alpha(f)) \neq 1$.

The introduction of τ and α allows Hilbert to replace universal and existential quantifiers, and also provides the basis for proofs of the axiom of induction and the least upper bound principle. Furthermore, Hilbert claims, the consistency of the resulting system can be seen in the same way used to establish the consistency of stage III (primitive recursive arithmetic). From a proof of a numerical formula using τ's and α's, these terms can be eliminated by finding numerical substitutions which turn the resulting formulas into correct numerical formulas.

These proofs are sketched in the last part of the 1921–22 lecture notes by Kneser (Hilbert 1922a).[25] In particular, the consistency proof contains the entire idea of the *Hilbertsche Ansatz*, the ε-substitution method:

> First we show that we can eliminate all variables, since here also only free variables occur. We look for the innermost τ and α. Below these there are only finitely defined [primitive recursive] functions φ, φ'. Some of these functions can be substituted for f in the axioms in the course of the proof. 1: $\tau(\varphi) = 0 \rightarrow (Z(\mathfrak{a}) \rightarrow \varphi(\mathfrak{a}) = 1)$, where $\mathfrak{a}$ is a functional. If this is *not* used, we set all $\alpha(\varphi)$ and $\tau(\varphi)$ equal to zero. Otherwise we reduce $\mathfrak{a}$ and $\varphi(\mathfrak{a})$ and check whether $Z(\mathfrak{a}) \rightarrow \varphi(\mathfrak{a}) = 1$ is correct everywhere it occurs. If it is correct, we set $\tau[\varphi] = 0$, $\alpha[\varphi] = 0$. If it is incorrect, i.e., if $\mathfrak{a} = \mathfrak{z}$, $\varphi(\mathfrak{z}) \neq 1$, we let $\tau(\varphi) = 1$, $\alpha(\varphi) = \mathfrak{z}$. After this replacement, the proof remains a proof. The formulas which take the place of the axioms are correct.
>
> (The idea is: if a proof is given, we can extract an argument from it for which $\varphi = 1$.) In this way we eliminate the τ and α and applications of [axioms]

(1)–(4) and obtain a proof of $1 \neq 1$ from I–V and correct formulas, i.e., from I–V,

$$\begin{array}{rcl} \tau(f,b)=0 & \rightarrow & \{Z(a) \rightarrow f(a,b)=1\} \\ \tau(f,b) \neq 0 & \rightarrow & Z(f(\alpha,b)) \\ \tau(f,b) \neq 0 & \rightarrow & f(\alpha(f,b),b) \neq 1 \\ \tau(f,b) \neq 0 & \rightarrow & \tau(f,b)=1 \end{array}$$[26]

Although not formulated as precisely as subsequent presentations, all the ingredients of Hilbert's ε-substitution method are here. The only changes that are made en route to the final presentation of Hilbert's sketch of the case for one ε and Ackermann's are mostly notational. In (Hilbert 1923), a talk given in September 1922, the two functions τ and α are merged to one function (also denoted τ), which in addition provides the *least* witness for $\tau(f(a)) \neq 1$. There the τ function is also applied directly to formulas. In fact, $\tau_a A(a)$ is the primary notion, denoting the least witness a for which $A(a)$ is false; $\tau(f)$ is defined as $\tau_a(f(a)=0)$. Interestingly enough, the sketch given there for the substitution method is for the τ-function for functions, not formulas, just as it was in the 1921–22 lectures.

The most elaborate discussion of the ε-calculus can be found in Hilbert's and Bernays's course of 1922–23. Here, again, the motivation for the ε-function is Zermelo's axiom of choice:

> What are we missing?
>
> 1. As far as logic is concerned: we have had the propositional calculus extended by free variables, i.e., variables for which arbitrary functionals may be substituted. Operating with "all" and "there is" is still missing.
> 2. We have added the induction schema, but without consistency proof and also on a provisional basis, with the intention of removing it.
> 3. So far only the arithmetical axioms which refer to whole numbers. The above shortcomings prevent us from building up analysis (limit concept, irrational number).
>
> These 3 points already give us a plan and goals for the following.
>
> We turn to (1). It is clear that a logic without "all'—"there is" would be incomplete, I only recall how the application of these concepts and of the so-called transfinite inferences has brought about major problems. We have not yet addressed the question of the applicability of these concepts to infinite totalities. Now we could proceed as we did with the propositional calculus:

> Formulate a few simple [principles] as axioms, from which all others follow. Then the consistency proof would have to be carried out—according to our general program: with the attitude that a proof is a figure given to us. Significant obstacles to the consistency proof because of the bound variables. The deeper investigation, however, shows that the real core of the problem lies at a different point, to which one usually only pays attention later, and which also has only been noticed in the literature of late.[27]

At the corresponding place in the Kneser *Mitschrift*, Hilbert continues:

> [This core lies] in Zermelo's *axiom of choice*. [. . .] The objections [of Brouwer and Weyl] are directed against the choice principle. But they should likewise be directed against "all" and "there is", which are based on the same basic idea.
>
> We want to extend the axiom of choice. To each proposition with a variable $A(a)$ we assign an object for which the proposition holds only if is holds in general. So, a counterexample, if one exists.
>
> $\varepsilon(A)$, an individual logical function. [. . .] ε satisfies the *transfinite axiom*:
>
> (16) $A(\varepsilon A) \rightarrow Aa$
>
> e.g., Aa means: a is corrupt. εA is Aristides.[28]

Hilbert goes on to show how quantifiers can be replaced by ε-terms. The corresponding definitional axioms are already included in Hilbert (1923), i.e., $A(\varepsilon A) \equiv (a)A(a)$ and $A(\varepsilon\overline{A}) \equiv (\exists a)A(a)$. Next, Hilbert outlines a derivation of the induction axioms using the ε-axioms. For this, it is necessary to require that the choice function takes the minimal value, which is expressed by the additional axiom

$$\varepsilon A \neq 0 \rightarrow A(\delta(\varepsilon A)).$$

With this addition, Hilbert combined the κ function of (Hilbert 1922c) and the μ function of (1923) with the ε function. Both κ ("k" for *Kleinstes*, least) and μ had been introduced there as functions of functions giving the least value for which the function differs from 0. In (1923, pp. 161–162), Hilbert indicates that the axiom of induction can be derived using the μ function, and credits this to Dedekind (1888).

The third issue Hilbert addresses is that of dealing with real numbers, and extending the calculus to analysis. A first step can be carried out at stage IV by considering a real number as a function defining an infinite binary expansion. A sequence of reals can then be given by a function with two arguments. Already in Hilbert (1923) we find a sketch of the proof of the least upper bound principle for such a sequence of reals, using the π function:

$$\pi A(a) = \begin{cases} 0 & \text{if } (a)A(a) \\ 1 & \text{otherwise} \end{cases}$$

The general case of sets of reals needs function variables and second-order ε and π. These are briefly introduced as $\varepsilon_f A$ with the axioms

$$\begin{aligned} A\varepsilon_f A &\to Af \\ (f)Af &\to \pi_f Af = 0 \\ \overline{(f)}Af &\to \pi_f Af = 1 \end{aligned}$$

The last two lectures transcribed in (Hilbert and Bernays 1923a) are devoted to a sketch of the ε substitution method. The proof is adapted from Hilbert (1923), replacing εf with εA, also deals with π, and covers the induction axiom in its form for the ε-calculus.[29] During the last lecture, Bernays also extends the proof to second-order ε's.

> If we have a *function variable*:
>
> $$A\varepsilon_f Af \to Af$$
>
> [...] Suppose ε only occurs with $\mathfrak{A}$ (e.g., $f0 = 0$, $ff0 = 0$). How will we eliminate the function variables? We simply replace fc by c. This does not apply to *bound* variables. For those we take some fixed function, e.g., δ and carry out the reduction with it. Then we are left with, e.g., $\mathfrak{A}\delta \to \mathfrak{A}\varphi$. This, when reduced, is either correct or incorrect. In the latter case, $\mathfrak{A}\varphi$ is incorrect. Then we substitute φ everywhere for $\varepsilon_f \mathfrak{A} f$. Then we end up with $\mathfrak{A}\varphi \to \mathfrak{A}\psi$. That is certainly correct, since $\mathfrak{A}\varphi$ is incorrect.[30]

The last development regarding the ε-calculus before Ackermann's dissertation is the switch to the dual notation. Both (Hilbert 1923) and (Hilbert and Bernays 1923a) use εA as denoting a counterexample for A, whereas at least from Ackermann's dissertation onwards, εA denotes a witness. Correspondingly, Ackermann uses the dual axiom $A(a) \to A(\varepsilon_a A(a))$. Although it is relatively clear that the supplement to the 1922–23 lectures (Hilbert and Bernays 1923a)—24 sheets in Hilbert's hand—are Hilbert's notes based on which he and partly Bernays presented the 1922–23 lectures, parts of it seem to have been altered or written after the conclusion of the course. Sheets 12–14 contain a sketch of the proof of the axiom of induction from the standard, dual ε axioms; the same proof for the original axioms can be found on sheets 8–11.

This concludes the development of mathematical systems using the ε-calculus and consistency proofs for them presented by Hilbert himself. We now turn to the more advanced and detailed treatment in Wilhelm Ackermann's (1924b) dissertation.

3.3 Ackermann's Dissertation

Wilhelm Ackermann was born in 1896 in Westphalia. He studied mathematics, physics, and philosophy in Göttingen between 1914 and 1924, serving in the army in World War I from 1915–1919. He completed his studies in 1924 with a dissertation, written under Hilbert, entitled "Begründung des 'tertium non datur' mittels der Hilbertschen Theorie der Widerspruchsfreiheit" (Ackermann 1924a, 1924b), the first major contribution to proof theory and Hilbert's Program. In 1927 he decided for a career as a high school teacher rather than a career in academia, but remained scientifically active. His major contributions to logic include the function which carries his name—an example of a recursive but not primitive recursive function (Ackermann 1928a), the consistency proof for arithmetic using the ε-substitution method (Ackermann 1940), and his work on the decision problem (Ackermann 1928b, 1954). He served as co-author, with Hilbert, of the influential logic textbook *Grundzüge der theoretischen Logik* (Hilbert and Ackermann 1928). He died in 1962.[31]

Ackermann's 1924 dissertation is of particular interest since it is the first non-trivial example of what Hilbert considered to be a finitistic consistency proof. Von Neumann's paper of 1927 does not entirely fit into the tradition of the Hilbert school, and we have no evidence of the extent of Hilbert's involvement in its writing. Later consistency proofs, in particular those by Gentzen and Kalmár, were written after Gödel's incompleteness results were already well-known and their implications understood by proof theorists. Ackermann's work, on the other hand, arose entirely out of Hilbert's research project, and there is ample evidence that Hilbert was aware of the range and details of the proof. Hilbert was Ackermann's dissertation advisor, approved the thesis, was editor of *Mathematische*

Annalen, where the thesis was published, and corresponded with Ackermann on corrections and extensions of the result. Ackermann was also in close contact with Paul Bernays, Hilbert's assistant and close collaborator in foundational matters. Ackermann spent the first half of 1925 in Cambridge, supported by a fellowship from the International Education Board (founded by John D. Rockefeller, Jr., in 1923). In his letter of recommendation for Ackermann, Hilbert writes:

> In his thesis "Foundation of the 'tertium non datur' using Hilbert's theory of consistency," Ackermann has shown in the most general case that the use of the words "all" and "there is," of the "tertium non datur," is free from contradiction. The proof uses exclusively primitive and finite inference methods. Everything is demonstratded, as it were, directly on the mathematical formalism.
>
> Ackermann has here surmounted considerable mathematical difficulties and solved a problem which is of first importance to the modern efforts directed at providing a new foundation for mathematics.[32]

Further testimony of Hilbert's high esteem for Ackermann can be found in the draft of a letter to Russell asking for a letter of support to the International Education Board, where he writes that "Ackermann has taken my classes on foundations of mathematics in recent semesters and is currently one of the best masters of the theory which I have developed here."[33]

Ackermann's work provides insight into two important issues relating to Hilbert's program as it concerns finitistic consistency proofs. First, it provides historical insight into the aims and development of Hilbert's Program: The first part of the program called for an axiomatization of mathematics. These axiomatizations were then the objects of metamathematical investigations: the aim was to find finitistic consistency proofs for them. Which areas of mathematics were supposed to be covered by the consistency proofs, how were they axiomatized, what is the strength of the systems so axiomatized? We have already seen what Hilbert's roadmap for the project of axiomatization was. Ackermann's dissertation provides the earliest example of a formal system stronger than elementary arithmetic. The second aim, the metamathematical investigation of the formal systems obtained, also poses historical questions: When did Ackermann, and other collaborators of Hilbert (in particular, Bernays and von Neumann) achieve the results they sought? Was Ackermann's proof correct, and if not, what parts of it can be made to work?

The other information we can extract from an analysis of Ackermann's work is what methods were used or presupposed in the consistency proofs that were given, and thus, what methods were sanctioned by Hilbert himself as falling under the finitist standpoint. Such an analysis of the methods used are of a deeper, conceptual interest. There is a fundamental division between Hilbert's philosophical remarks on finitism on the one hand, and the professed goals of the program on the other. In these comments, rather little is said about the concept formations and proof methods that a finitist, according to Hilbert, is permitted to use. In fact, most of Hilbert's remarks deal with the objects of finitism, and leave the finitistically admissible forms of definition and proof to the side. These, however, are the questions at issue in contemporary conceptual analyses of finitism. Hilbert's relative silence on the matter is responsible for the widespread—and largely correct—opinion that Hilbert was too vague on the question of what constitutes finitism to unequivocally define the notion, and therefore later commentators have had enough leeway to disagree widely on the strength of the finitist standpoint while still claiming to have explicated Hilbert's own concept.

3.3.1 Second-order Primitive Recursive Arithmetic

In Ackermann (1924b), the system of stage III is extended by second-order variables for functions. The schema of recursive definition is then extended to include terms containing such variables. In the following outline, I shall follow Ackermann and adopt the notation of subscripting function symbols and terms by variables to indicate that these variables do not occur freely but rather as placeholders for functions. For instance, $\mathfrak{a}_a(f(a))$ indicates that the term $\mathfrak{a}$ does not contain the variable a free, but rather that the function $f(a)$ appears as a functional argument, i.e., that the term is of the form $\mathfrak{a}(\lambda a.f(a))$. The schema of second-order primitive recursion is the following:

$$\begin{aligned} \varphi_{\vec{b}_i}(0, \vec{f}(\vec{b}_i), \vec{c}) &= \mathfrak{a}_{\vec{b}_i}(\vec{f}(\vec{b}_i), \vec{c}) \\ \varphi_{\vec{b}_i}(a+1, \vec{f}(\vec{b}_i), \vec{c}) &= \mathfrak{b}_{\vec{b}_i}(a, \varphi_{\vec{d}_i}(a, \vec{f}(\vec{d}_i), \vec{c}), \vec{f}(\vec{b}_i)) \end{aligned}$$

To clarify the subscript notation, compare this with the schema of second-order primitive recursion using λ-abstraction notation:

$$\begin{aligned}\varphi(0,\lambda\vec{b}_i.\vec{f}(\vec{b}_i),\vec{c}) &= \mathfrak{a}(\lambda\vec{b}_i.\vec{f}(\vec{b}_i),\vec{c})\\ \varphi(a+1,\lambda\vec{b}_i.\vec{f}(\vec{b}_i),\vec{c}) &= \mathfrak{b}(a,\varphi(a,\lambda\vec{d}_i.\vec{f}(\vec{d}_i),\vec{c}),\lambda\vec{b}_i.\vec{f}(\vec{b}_i))\end{aligned}$$

Using this schema, it is possible to define the Ackermann function. This was already pointed out in Hilbert (1926), although it was not until Ackermann (1928a) that it was shown that the function so defined cannot be defined by primitive recursion without function variables. Ackermann (1928a) defines the function as follows. First it is observed that the iteration function

$$\rho_c(a,f(c),b) = \underbrace{f(\dots f(f}_{a\ f\text{'s}}(b))\dots)$$

can be defined by second-order primitive recursion:

$$\begin{aligned}\rho_c(0,f(c),b) &= b\\ \rho_c(a+1,f(c),b) &= f(\rho_c(a,f(c),b))\end{aligned}$$

Furthermore, we have two auxiliary functions

$$\iota(a,b) = \begin{cases}1 & \text{if } a=b\\ 0 & \text{if } a\neq b\end{cases} \quad\text{and}\quad \lambda(a,b) = \begin{cases}0 & \text{if } a\neq b\\ 1 & \text{if } a=b\end{cases}$$

which are primitive recursive, as well as addition and multiplication. The term $\mathfrak{a}(a,b)$ is short for $\iota(a,1)\cdot\iota(a,0)\cdot b+\lambda(a,1)$; we then have

$$\mathfrak{a}(a,b) = \begin{cases}0 & \text{if } a=0\\ 1 & \text{if } a=1\\ b & \text{otherwise}\end{cases}$$

The Ackermann function is defined by

$$\begin{aligned}\varphi(0,b,c) &= b+c\\ \varphi(a+1,b,c) &= \rho_d(c,\varphi(a,b,d),\mathfrak{a}(a,b)).\end{aligned}$$

In more suggestive terms,

$$\begin{aligned}\varphi(0,b,c) &= b+c\\ \varphi(1,b,c) &= b\cdot c\\ \varphi(a+1,b,c) &= \underbrace{\varphi(a,b,\varphi(a,b,\ldots\varphi(a,b,b)\cdots))}_{c \text{ times}}\end{aligned}$$

The system of second-order primitive recursive arithmetic 2PRA^- used in Ackermann (1924b) consists of axioms (1)–(15) of Hilbert and Bernays (1923b, see Section 3.2), axiom (16) was replaced by

16. $a \neq 0 \to a = \delta(a)+1$,

plus defining equations for both first- and second-order primitive recursive functions. There is no induction rule (which is usually included in systems of primitive recursive arithmetic), although the consistency proof given by Ackermann can easily be extended to cover it.

3.3.2 The Consistency Proof for the System without Epsilons

Allowing primitive recursion axioms for functions which contain function variables is a natural extension of the basic calculus of stages III and IIII, and is necessary in order to be able to introduce sufficiently complex functions. Hilbert seems to have thought that by extending primitive recursion in this way, or at least by building an infinite hierarchy of levels of primitive recursions using variables of higher types, he could account for *all* the number theoretic functions, and hence for all real numbers (represented as decimal expansions). In the spirit of the stage-by-stage development of systems of mathematics and consistency proofs, it is of course necessary to show the consistency of the system of stage IIII, which is the system presented by Ackermann. As before, it makes perfect sense to first establish the consistency for the fragment of stage IIII not containing the transfinite ε and π functions. In Section 4 of his dissertation, Ackermann undertakes precisely this aim.

The proof is a direct extension of the consistency proof of stage III, the elementary calculus of free variables with basic arithmetical axioms and primitive recursive definitions, i.e., PRA. This proof had already been presented in Hilbert's lectures in 1921–22 and 1922–23. The idea here is the same: put a given, purported proof of $0 \neq 0$ into tree form, eliminate variables, and reduce functionals. The remaining figure consists entirely of correct formulas, where correctness of a formula is a syntactically defined and easily decidable property. The only complication for the case where function variables are also admitted is the reduction of functionals. It must be shown that every functional, i.e., every term of the language, can be reduced to a numeral on the basis of the defining recursion equations. For the original case, this could be done by a relatively simple inductive proof. For the case of $2\mathrm{PRA}^-$, it is not so obvious.

Ackermann locates the difficulty in the following complication. Suppose you have a functional $\varphi_b(2, \mathfrak{b}(b))$, where φ is defined by

$$\begin{aligned} \varphi_b(0, f(b)) &= f(1) + f(2) \\ \varphi_b(a+1, f(b)) &= \varphi_b(a, f(b)) + f(a) \cdot f(a+1) \end{aligned}$$

Here, $\mathfrak{b}(b)$ is a term which denotes a function, and so there is no way to replace the variable b with a numeral before evaluating the entire term. In effect, the variable b is bound (in modern notation, the term might be more suggestively written $\varphi(2, \lambda b.\mathfrak{b}(b))$.) In order to reduce this term, we apply the recursion equations for φ twice and end up with a term like

$$\mathfrak{b}(1) + \mathfrak{b}(2) + \mathfrak{b}(0) \cdot \mathfrak{b}(1) + \mathfrak{b}(1) \cdot \mathfrak{b}(2).$$

The remaining $\mathfrak{b}$'s might in turn contain φ, e.g., $\mathfrak{b}(b)$ might be $\varphi_c(b, \delta(c))$, in which case the above expression would be

$$\varphi_c(1, \delta(c)) + \varphi_c(2, \delta(c)) + \varphi_c(0, \delta(c)) \cdot \varphi_c(1, \delta(c)) + \varphi_c(1, \delta(c)) \cdot \varphi_c(2, \delta(c)).$$

By contrast, reducing a term $\psi(\mathfrak{z})$ where ψ is defined by first-order primitive recursion results in a term which does not contain ψ, but only the function symbols occurring on the right-hand side of the defining equations for ψ.[34]

To show that the reduction indeed comes to an end if innermost subterms are reduced first, Ackermann proposes to assign indices to terms and show that each reduction reduces this index. The indices are, essentially, ordinal notations $< \omega^{\omega^\omega}$. Since this is probably the first proof using ordinal notations, it may be of some interest to repeat and analyze it in some detail here. In my presentation, I stay close to Ackermann's argument and only change the notation for ranks, indices, and orders: Where Ackermann uses sequences of natural numbers, I will use the more perspicuous ordinal notations. Note again, however, that Ackermann does not explicitly use ordinal notations.

Suppose the primitive recursive functions are arranged in a linear order according to the order of definition. We write $\varphi < \psi$ if φ occurs before ψ in the order of definition, i.e., ψ cannot be used in the defining equations for φ. Suppose further that we are given a specific term t. The notion of *subordination* is defined as follows: an occurrence of a function symbol ξ in t is subordinate to an occurrence of φ, if φ is the outermost function symbol of a subterm s, the occurrence of ξ is in s, and the subterm of s with outermost function symbol ξ contains a bound variable b in whose scope the occurrence of φ is (this includes the case where b happens to be bound by φ itself).[35] In other words, $\mathfrak{t}$ is of the form

$$\mathfrak{t}'(\ldots \varphi_b(\ldots \xi(\ldots b \ldots) \ldots) \ldots)$$

The *rank* $rk(\mathfrak{t}, \varphi)$ of an *occurrence* of a function symbol φ with respect to $\mathfrak{t}$ is defined as follows: If there is no occurrence of $\psi > \varphi$ which is subordinate to φ in $\mathfrak{t}$, then $rk(\mathfrak{t}, \varphi) = 1$. Otherwise,

$$rk(\mathfrak{t}, \varphi) = \max\{rk(\mathfrak{t}, \psi) : \psi > \varphi \text{ is subordinate to } \varphi\} + 1.$$

The rank $r(\mathfrak{t}, \varphi)$ of a term $\mathfrak{t}$ with respect to a function symbol φ is the maximum of the ranks of occurrences of φ or $\psi > \varphi$ in t. (If neither φ nor $\psi > \varphi$ occur in $\mathfrak{t}$, that means $r(\mathfrak{t}, \varphi) = 0$.

Ackermann now goes on to define the indices and orderings on these indices; the proof proceeds by induction on these orderings. The indices correspond to ordinal notations in modern terminology, and the orderings imposed are order-isomorphic to well-orderings of type $\omega^{\omega^{\omega}}$. Ackermann does of course not use ordinals to define these indices; he stresses that he is only dealing with finite sequences of numbers, on which an elaborate order is imposed. Rather than appeal to the well-orderedness of $\omega^{\omega^{\omega}}$, he gives a more direct argument that by repeatedly proceeding to indices which are smaller in the imposed order one eventually has to reach the index which consists of all 0. To appreciate the flavor of Ackermann's definitions, consider the following quote where he defines the rank and index of a term:

> Each of the functionals out of which the given functional is constructed has a definite rank with respect to the last, the next-to-last, etc., until the first recursive function. Each such combination of ranks is characterized by [a sequence of] $\mathfrak{n}$ ordered numbers [the order of $\mathfrak{t}$ in my notation]. We now want to order all these finitely many rank combinations. Taking two different rank combinations, we write the corresponding numbers on top of one another, i.e, first the rank with respect to the last, then those with respect to the next-to-last function, etc. At some point the numbers are different for the first time. We now call that rank combination higher which has the greater number at this point. In this manner we order all the finitely many rank combinations occuring in the given functional. For each rank combination we then write down, how many functionals of that kind occur in the given functional. We will call the totality of these numbers the index of the functional.[36]

We assign to a subterm $\mathfrak{s}$ of $\mathfrak{t}$ a sequence of ranks of $\psi_n, \ldots, \psi_0$ with respect to $\mathfrak{s}$, where $\psi_0 < \cdots < \psi_n$ are all function symbols occurring in $\mathfrak{t}$. This is the *order* [*Rangkombination*] of $\mathfrak{s}$:

$$o(\mathfrak{s}) = \langle r(\mathfrak{s}, \psi_n), \ldots, r(\mathfrak{s}, \psi_0) \rangle$$

In modern notation, we may think of this as an ordinal notation corresponding to an ordinal $< \omega^\omega$, specifically, $o(\mathfrak{s})$ corresponds to

$$\omega^n \cdot r(\mathfrak{s}, \psi_n) \cdot + \cdots + \omega \cdot r(\mathfrak{s}, \psi_1) \cdot + r(\mathfrak{s}, \psi_0)$$

Now consider the set of all distinct subterms of $\mathfrak{t}$ of a given order o which are not numerals. The *degree* $d(\mathfrak{t}, o)$ of o in $\mathfrak{t}$ is the cardinality of that set. The *index* $j(\mathfrak{t})$ of $\mathfrak{t}$ is the sequence of degrees ordered in the same way as the orders, i.e.,

$$j(\mathfrak{t}) = \langle o : d(\mathfrak{s}, o) \rangle$$

where o ranges over all orders of subterms of $\mathfrak{t}$. In modern notation, this can be seen as an ordinal notation such an index corresponding to an ordinal of the form

$$\sum_o \omega^{\hat{o}} \cdot d(\mathfrak{t}, o)$$

where the sum again ranges over the orders o of subterms of $\mathfrak{t}$, and $\hat{o}$ is the ordinal corresponding to the order o.. Obviously, this is an ordinal $< \omega^{\omega^\omega}$.

Suppose a term $\mathfrak{t}$ not containing ε or π is given. Let $\mathfrak{s}$ be an innermost constant subterm which is not a numeral, we may assume it is of the form $\varphi_{\bar{b}}(\mathfrak{z}_1, \ldots, \mathfrak{z}_n, \mathfrak{u}_1, \ldots, \mathfrak{u}_m)$ where $\mathfrak{u}_i$ is a term with at least one variable bound by φ and which doesn't contain a constant subterm. We have two cases:

(1) $\mathfrak{s}$ does not contain bound variables, i.e., $m = 0$. The order of $\mathfrak{s}$ is a sequence with 1 in the k-th place, and 0 everywhere else (where $\varphi = \psi_k$), which corresponds to ω^k. Evaluating the term $\mathfrak{s}$ by recursion results in a term $\mathfrak{s}'$ in which only function symbols of lower index occur. Hence, the first non-zero component of the order of $\mathfrak{s}'$ is further to the right than i (in the corresponding ordinal, the highest exponent in the order of $\mathfrak{s}'$ is less than i), and so $o(\mathfrak{s}') < o(\mathfrak{s})$. Furthermore, since no variable which is bound in $\mathfrak{t}$ can occur in $\mathfrak{s}'$ (since no such variable occurs in $\mathfrak{s}$), replacing $\mathfrak{s}$ by $\mathfrak{s}'$ in $\mathfrak{t}$ does not result in new occurrences of function symbols which are subordinate to any other. Thus the number of subterms in the term $\mathfrak{t}'$ which results from such a replacement with orders $> o(\mathfrak{s})$ remains the same, while the number of subterms of order $o(\mathfrak{s})$ is reduced by 1. Hence, $j(\mathfrak{t}') < j(\mathfrak{t})$.

(2) $\mathfrak{s}$ does contain bound variables. For simplicity, assume that there is one numeral argument and one functional argument, i.e., $\mathfrak{s}$ is of the form $\varphi_b(\mathfrak{z}, \mathfrak{c}(b))$. In this case, all function symbols occurring in $\mathfrak{c}(b)$ are subordinate to φ, or otherwise $\mathfrak{c}(b)$ would contain a constant subterm.[37] Thus, the rank of $\mathfrak{c}(b)$ in $\mathfrak{t}$ with respect to ψ_i is less that the rank of $\mathfrak{s}$ with respect to ψ_i.

We reduce the subterm $\mathfrak{s}$ to a subterm $\mathfrak{s}'$ by applying the recursion. $\mathfrak{s}'$ does not contain the function symbol φ. We want to show that replacing $\mathfrak{s}$ by $\mathfrak{s}'$ in $\mathfrak{t}$ lowers the index of $\mathfrak{t}$.

First, note that when substituting a term $\mathfrak{a}$ for b in $\mathfrak{c}(b)$, the order of the resulting $\mathfrak{c}(\mathfrak{a})$ with respect to φ is the maximum of the orders of $\mathfrak{c}(b)$ and $\mathfrak{a}$, since none of the occurrences of function symbols in $\mathfrak{a}$ contain bound variables whose scope begins outside of $\mathfrak{a}$, and so none of these variables are subordinate to any function symbols in $\mathfrak{c}(b)$.

Now we prove the claim by induction on $\mathfrak{z}$. Suppose the defining equation for φ is

$$\begin{aligned} \varphi_b(0, f(b)) &= \mathfrak{a}_b(f(b)) \\ \varphi_b(a+1, f(b)) &= \mathfrak{b}_b(\varphi_c(a, f(c)), a, f(b)). \end{aligned}$$

If $\mathfrak{z} = 0$, then $\mathfrak{s}' = \mathfrak{a}_b(\mathfrak{c}(b))$. At a place where $f(b)$ is an argument to a function, $f(b)$ is replaced by $\mathfrak{c}(d)$, and d is not in the scope of any φ (since $\mathfrak{a}$ doesn't contain φ). For instance, $\mathfrak{a}_b(f(b)) = 2 + \psi_d(3, f(d))$. Such a replacement cannot raise the φ-rank of $\mathfrak{s}'$ above that of $\mathfrak{c}(b)$. The term $\mathfrak{c}$ might also be used in places where it is not a functional argument, e.g., if $\mathfrak{a}_b(f(b)) = f(\psi(f(2)))$. By a simple induction on the nesting of f's in $\mathfrak{a}_b(f(b))$ it can be seen that the φ-rank of $\mathfrak{s}'$ is the same as that of $\mathfrak{c}(b)$: For $\mathfrak{c}(\mathfrak{d})$ where $\mathfrak{d}$ does not contain $\mathfrak{c}$, the φ-rank of $\mathfrak{c}(\mathfrak{d})$ equals that of $\mathfrak{c}(b)$ by the note above and the fact that $\mathfrak{d}$ does not contain φ. If $\mathfrak{d}$ does contain a nested $\mathfrak{c}$, then by induction hypothesis and the first case, its φ-rank is the same as that of $\mathfrak{c}(b)$. By the note, again, the entire subterm has the same φ-rank as $\mathfrak{c}(b)$.

The case of $\varphi_b(\mathfrak{z}+1, \mathfrak{c}(b))$ is similar. Here, the first replacement is

$$\mathfrak{b}_b(\varphi_c(\mathfrak{z}, \mathfrak{c}(c)), \mathfrak{z}, \mathfrak{c}(b)).$$

Further recursion replaces $\varphi_c(\mathfrak{z}, \mathfrak{c}(c))$ by another term which, by induction hypothesis, has φ-rank less than or equal to that of $\mathfrak{c}(b)$. The same considerations as in the base case show that the entire term also has a φ-rank no larger than $\mathfrak{c}(b)$.

We have thus shown that eliminating the function symbol φ by recursion from an innermost constant term reduces the φ-rank of the term at least by one and does not increase the ψ_j-ranks of any subterms for any $j > i$.

In terms of ordinals, this shows that at least one subterm of order o was reduced to a subterm of order $o' < o$, all newly introduced subterms have order $< o$, and the order of no old subterm increased. Thus, the index of the entire term was reduced. (In the corresponding ordinal notations, the factor $\omega^{\hat{o}} \cdot n$ changed to $\omega^{\hat{o}} \cdot (n-1)$).

> We started with a given constant function, which we characterized by a determinate index. We replaced a $\varphi_b(\mathfrak{z}, \mathfrak{c}(b))$ within that functional by another functional, where the φ-rank decreased and the rank with respect to functions to the right of φ [i.e., which come after φ in the order of definition] did not increase. Now we apply the same operation to the resulting functional. After finitely many steps we obtain a functional which contains no function symbols at all, i.e., it is a numeral.
>
> We have thus shown: a constant functional, which does not contain ε and π, can be reduced to a numeral in finitely many steps.[38]

3.3.3 Ordinals, Transfinite Induction, and Finitism

It is quite remarkable that the earliest extensive and detailed technical contribution to the finitist project would make use of transfinite induction in a way not dissimilar to Gentzen's later proof by induction up to ε_0. This bears on a number of questions regarding Hilbert's understanding of the strength of finitism. In particular, it is often said that Gentzen's proof is not finitist, because it uses transfinite induction. However, Ackermann's original consistency proof for 2PRA$^-$ also uses transfinite induction, using an index system which is essentially an ordinal notation system, just like Gentzen's. If it is granted that Ackermann's proof is finitistic, but Gentzen's is not, i.e., transfinite induction up to ω^{ω^ω} is finitistic but not up to ε_0, then where—and why—should the line be drawn? Furthermore, the consistency proof of 2PRA$^-$ is in essence a—putatively finitistic—explanation of how to compute second order primitive recursive functions, and a proof that the computation procedure defined by them always terminates. In other words, it is a finitistic proof that second order primitive recursive functions are well defined.[39]

Ackermann was completely aware of the involvement of transfinite induction in this case, but he sees in it no violation of the finitist standpoint.

> The disassembling of functionals by reduction does not occur in the sense that a finite ordinal is decreased each time an outermost function symbol is eliminated. Rather, to each functional corresponds as it were a transfinite ordinal number as its rank, and the theorem, that a constant functional is reduced to a numeral after carrying out finitely many operations, corresponds to the other [theorem], that if one descends from a transfinite ordinal number to ever smaller ordinal numbers, one has to reach zero after a finite number of steps. Now there is naturally no mention of transfinite sets or ordinal numbers in our metamathematical investigations. It is however interesting, that the mentioned theorem about transfinite ordinals can be formulated so that there is nothing transfinite about it.[40]

Without appealing to the well-orderedness of the corresponding ordinals, it remains to argue finitistically that the finite sequences of numbers ordered in the appropriate manner are also well-ordered. Ackermann does not attempt this for the entire class of sequences of sequences of numbers needed in the proof (corresponding to ω^{ω^ω}), but only for ω^2.

> Consider a transfinite ordinal number less than $\omega \cdot \omega$. Each such ordinal number can be written in the form: $\omega \cdot \mathfrak{n} + \mathfrak{m}$, where $\mathfrak{n}$ and $\mathfrak{m}$ are finite num-

> bers. Hence such an ordinal can also be characterized by a pair of finite numbers $(\mathfrak{n}, \mathfrak{m})$, where the order of these numbers is of course significant. To the descent in the series of ordinals corresponds the following operation on the number pair $(\mathfrak{n}, \mathfrak{m})$. Either the first number $\mathfrak{n}$ remains the same, then the number $\mathfrak{m}$ is replaced by a smaller number $\mathfrak{m}'$. Or the first number $\mathfrak{n}$ is made smaller; then I can put an arbitrary number in the second position, which can also be larger than $\mathfrak{m}$. It is clear that one has to reach the number pair $(0,0)$ after finitely many steps. For after at most $\mathfrak{m}+1$ steps I reach a number pair, where the first number is smaller than $\mathfrak{n}$. Let $(\mathfrak{n}', \mathfrak{m}')$ be that pair. After at most $\mathfrak{m}'+1$ steps I reach a number pair in which the first number is again smaller than $\mathfrak{n}'$, etc. After finitely many steps one reaches the number pair $(0,0)$ in this fashion, which corresponds to the ordinal number 0. In this form, the mentioned theorem contains nothing transfinite whatsoever; only considerations which are acceptable in metamathematics are used. The same holds true if one does not use pairs but triples, quadruples, etc. This idea is not only used in the following proof that the reduction of functionals terminates, but will also be used again and again later on, especially in the finiteness proof at the end of the work.[41]

Over ten years later, Ackermann discusses the application of transfinite induction for consistency proofs in correspondence with Bernays. Gentzen's consistency proof had been published (Gentzen 1936), and Gentzen asks, through Bernays,

> whether you [Ackermann] think that the method of proving termination [*Endlichkeitsbeweis*] by transfinite induction can be applied to the consistency proof of your dissertation. I would like it very much, if that were possible.[42]

In his reply, Ackermann recalls his own use of transfinite ordinals in the 1924 dissertation.

> I just realized now, as I am looking at my dissertation, that I operate with transfinite ordinals in a similar fashion as Gentzen.[43]

A year and a half later, Ackermann mentions the transfinite induction used in his dissertation again:

> I do not know, by the way, whether you are aware (I did at the time not consider it as a transgression beyond the narrower finite standpoint), that transfinite inferences are used in my dissertation. (Cf., e.g., the remarks in the last paragraph on page 13 and the following paragraph of my dissertation).[44]

These remarks may be puzzling, since they seem to suggest that Bernays was not familiar with Ackermann's work. This is clearly not the case. Bernays corresponded with Ackerman extensively in the mid-20s about the ε-substitution method and the decision problem, and had clearly studied Ackermann's dissertation. Neither Bernays nor Hilbert are on record objecting to the methods used in Ackermann's dissertation. It can thus be concluded that Ackermann's use of transfinite induction was considered acceptable from the finitist standpoint.

3.3.4 The ε-Substitution Method

As we have seen above, Hilbert had outlined an idea for a consistency proof for systems involving ε-terms already in early 1922 (Hilbert 1922a), and a little more precisely in his talk of 1922 (Hilbert 1923) and in the 1922–23 lectures (Hilbert and Bernays 1923a). Let us review the *Ansatz* in the notation used in 1924: Suppose a proof involves only one ε term $\varepsilon_a A(a)$ and corresponding *critical formulas*

$$A(\mathfrak{k}_i) \to A(\varepsilon_a A(a)),$$

i.e., substitution instances of the transfinite axiom

$$A(a) \to A(\varepsilon_a A(a)).$$

We replace $\varepsilon_a A(a)$ everywhere with 0, and transform the proof as before by rewriting it in tree form ("dissolution into proof threads"), eliminating free variables and evaluating numerical terms involving primitive recursive functions. Then the critical formulas take the form

$$A(\mathfrak{z}_i) \to A(0),$$

where $\mathfrak{z}_i$ is the numerical term to which $\mathfrak{k}_i$ reduces. A critical formula can now only be false if $A(\mathfrak{z}_i)$ is true and $A(0)$ is false. If that is the case, repeat the procedure, now substituting $\mathfrak{z}_i$ for $\varepsilon_a A(a)$. This yields a proof in which all initial formulas are correct and no ε terms occur.

If critical formulas of the second kind, i.e., substitution instances of the induction axiom,

$$\varepsilon_a A(a) \neq 0 \to \overline{A(\delta \varepsilon_a A(a))},$$

also appear in the proof, the witness $\mathfrak{z}$ has to be replaced with the least $\mathfrak{z}'$ so that $A(\mathfrak{z}')$ is true.

The challenge was to extend this procedure to (a) cover more than one ε-term in the proof, (b) take care of nested ε-terms, and lastly (c) extend it to second-order ε's and terms involving them, i.e, $\varepsilon_f \mathfrak{A}_a(f(a))$. This is what Ackermann set out to do in the last part of his dissertation, and what he and Hilbert thought he had accomplished.[45]

The system for which Ackermann attempted to give a consistency proof consisted of the system of second-order primitive recursive arithmetic (see Section 3.3.1 above) together with the transfinite axioms:

1.	$A(a) \to A(\varepsilon_a A(a))$	$A_a(f(a)) \to A_a((\varepsilon_f A_b(f(b)))(a))$
2.	$A(\varepsilon_a A(a)) \to \pi_a A(a) = 0$	$A_a(\varepsilon_f A_b(f(b))(a)) \to \pi_f A_a(f(a)) = 0$
3.	$\overline{A(\varepsilon_a A(a))} \to \pi_a A(a) = 1$	$\overline{A_a(\varepsilon_f A_b(f(b))(a))} \to \pi_f A_a(f(a)) = 1$
4.	$\varepsilon_a A(a) \neq 0 \to \overline{A(\delta(\varepsilon_a A(a)))}$[46]	

The intuitive interpretation of ε and π, based on these axioms, is obvious: $\varepsilon_a \mathfrak{A}(a)$ is a witness for $\mathfrak{A}(a)$ if one exists, and $\pi_a \mathfrak{A}(a) = 1$ if $\mathfrak{A}(a)$ is false for all a, and $= 0$ otherwise. The π functions are not necessary for the development of mathematics in the axiom system. They do, however, serve a function in the consistency proof, viz., to keep track of whether a value of 0 for $\varepsilon_a \mathfrak{A}(a)$ is a "default value" (i.e., a trial substitution for which $\mathfrak{A}(a)$ may or may not be true) or an actual witness (a value for which $\mathfrak{A}(a)$ has been found to be true).

I shall now attempt to give an outline of the ε-substitution procedure defined by Ackermann. For simplicity, I will leave the case of second-order ε-terms (i.e., those involving ε_f) to the side.

An ε-*term* is an expression of the form $\varepsilon_a \mathfrak{A}(a)$, where a is the only free variable in $\mathfrak{A}$, and similarly for a π-term. For the purposes of the discussion below, we will not specifically refer to π's unless necessary, and most definitions and operations apply equally to ε-terms and π-terms. If a formula $A(a)$ or an ε-term $\varepsilon_a \mathfrak{A}(a)$ contains no variable-free subterms which are not numerals, we call them *canonical*. Canonical formulas and ε-terms are indicated by a tilde: $\varepsilon_a \tilde{\mathfrak{A}}(a)$.

The main notion in Ackermann's proof is that of a *total substitution* S (*Gesamtersetzung*). It is a mapping of canonical ε- and π-terms to numerals and 0 or 1, respectively. When canonical ε-terms in a proof are successively replaced by their values under the mapping, a total substitution reduces the proof to one not containing any ε's. If S maps $\varepsilon_a\tilde{\mathfrak{A}}(a)$ to $\mathfrak{z}$ and $\pi_a\tilde{\mathfrak{A}}(a)$ to i, then we say that $\tilde{\mathfrak{A}}(a)$ receives a $(\mathfrak{z},\mathfrak{i})$ substitution under S and write $S(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z},\mathfrak{i})$.

It is of course not enough to define a mapping from the canonical ε-terms *occuring in the proof* to numerals: The proof may contain, e.g., $\varepsilon_a\mathfrak{A}(a,\varphi(\varepsilon_b\mathfrak{B}(b)))$. To reduce this to a numeral, we first need a value $\mathfrak{z}$ for the term $\varepsilon_b\mathfrak{B}(b)$. Replacing $\varepsilon_b\mathfrak{B}(b)$ by $\mathfrak{z}$, we obtain $\varepsilon_a\mathfrak{A}(a,\varphi(\mathfrak{z}))$. Suppose the value $\varphi(\mathfrak{z})$ is $\mathfrak{z}'$. The total substitution then also has to specify a substitution for $\varepsilon_a\mathfrak{A}(a,\mathfrak{z}')$.

Given a total substitution S, a proof is reduced to an ε-free proof as follows: First all ε-free terms are evaluated. (Such terms contain only numerals and primitive recursive functions; these are computed and the term replaced by the numeral corresponding to the value of the term) Now let $\mathfrak{e}_{11}$, $\mathfrak{e}_{12}$, ... be all the innermost (canonical) ε- or π-terms in the proof, i.e., all ε- or π-terms which do not themselves contain nested ε- or π-terms or constant (variable-fee) subterms which are not numerals. The total substitution specifies a numeral substitution for each of these. Replace each $\mathfrak{e}_{1i}$ by its corresponding numeral. Repeat the procedure until the only remaining terms are numerals. We write $|\mathfrak{e}|_S$ for the result of applying this procedure to the expression (formula or term) $\mathfrak{e}$. Note that $|\mathfrak{e}|_S$ is canonical.

Based on this reduction procedure, Ackermann defines a notion of subordination of canonical formulas. Roughly, a formula $\tilde{\mathfrak{B}}(b)$ is subordinate to $\tilde{\mathfrak{A}}(a)$ if in the process of reducing some formula $\tilde{\mathfrak{A}}(\mathfrak{z})$, an ε-term $\varepsilon_b\tilde{\mathfrak{B}}(b)$ is replaced by a numeral. For instance, $a = b$ is subordinate to $a = \varepsilon_b(a = b)$. Indeed, if $\tilde{\mathfrak{A}}(a)$ is $a = \varepsilon_b(a = b)$, then the reduction of $\tilde{\mathfrak{A}}(\mathfrak{z}) \equiv \mathfrak{z} = \varepsilon_b(\mathfrak{z} = b)$ would use a replacement for the ε-term belonging to $\tilde{\mathfrak{B}}(\mathfrak{z} = b)$.[47] It is easy to see that this definition corresponds to the notion of subordination as defined in Hilbert and Bernays (1939). An ε-*expression* is an expression of the form $\varepsilon_a\mathfrak{A}(a)$. If $\varepsilon_a\mathfrak{A}(a)$ contains no free variables, it is called an ε-*term*. If an ε-term $\varepsilon_b\mathfrak{B}(b)$ occurs in an expression (and is different from it), it is said to be *nested* in it. If an ε-expression $\varepsilon_b\mathfrak{B}(a,b)$ occurs in an expression in the scope of ε_a, then it is *subordinate* to that expression. Accordingly, we can define the degree of an ε-term and the rank of an ε-expression as follows: An ε-term with no nested ε-subterms is of degree 1; otherwise its degree is the maximum of the degrees of its nested ε-subterms $+1$. The rank of an ε-expression with no subordinate ε-expressions is 1; otherwise it is the maximum of the ranks of its subordinate ε-expressions +1. If $\tilde{\mathfrak{B}}(b)$ is subordinate to $\tilde{\mathfrak{A}}(a)$ according to Ackermann's definition, then $\varepsilon_b\tilde{\mathfrak{B}}(b)$ is subordinate in the usual sense to $\varepsilon_a\tilde{\mathfrak{A}}(a)$, and the rank of $\varepsilon_b\tilde{\mathfrak{B}}(b)$ is less than that of $\varepsilon_a\tilde{\mathfrak{A}}(a)$. The notion of degree corresponds to an ordering of canonical formulas used for the reduction according to a total substitution in Ackermann's procedural definition: First all ε-terms of degree 1 (i.e., all innermost ε-terms) are replaced, resulting (after evaluation of primitive recursive functions) in a partially reduced proof. The formulas corresponding to innermost ε-terms now are reducts of ε-terms of degree 2 in the original proof. The canonical formulas corresponding to ε-terms of degree 1 are called the formulas of *level* 1, the canonical formulas corresponing to the innermost ε-terms in the results of the first reduction step are the formulas of level 2, and so forth.

The consistency proof proceeds by constructing a sequence $S_1, S_2, \ldots$ of total substitutions together with bookkeeping functions $f_i(\tilde{\mathfrak{A}}(a), j) \to \{0,1\}$,[48] which eventually results in a *solving substitution*, i.e., a total substitution which reduces the proof to one which contains only correct ε-free formulas. We begin with a total substitution S_1 which assigns $(0,1)$ to all canonical formulas, and set $f_1(\tilde{\mathfrak{A}}(a),1) = 1$ for all $\tilde{\mathfrak{A}}(a)$ for which S_1 assigns a value. If S_i is a solving substitution, the procedure terminates. Otherwise, the next total substitution S_{i+1} is obtained as follows: If S_i is not a solving substitution, at least one of the critical formulas in the proof reduces to an incorrect formula. We have three cases:

1. Either an ε-axiom $\mathfrak{A}(a) \to \mathfrak{A}(\varepsilon_a\mathfrak{A}(a))$ or a π-axiom of the first kind $\mathfrak{A}(\varepsilon_a\mathfrak{A}(a)) \to \pi_a\mathfrak{A}(a) = 0$ reduces to a false formula of the form $\tilde{\mathfrak{A}}(\mathfrak{z}) \to \tilde{\mathfrak{A}}(0)$ or $\tilde{\mathfrak{A}}(0) \to 1 = 0$, and $S_i(\tilde{\mathfrak{A}}(a)) = (0,1)$. Pick one such $\tilde{\mathfrak{A}}(a)$ of lowest level (i.e., $\varepsilon_a\mathfrak{A}(a)$ of lowest degree).

   If $\tilde{\mathfrak{A}}(0) \to 1 = 0$ is incorrect, $\tilde{\mathfrak{A}}(0)$ must be correct; let $S_{i+1}(\tilde{A}(a)) = (0,0)$. Otherwise $\tilde{\mathfrak{A}}(\mathfrak{z}) \to \tilde{\mathfrak{A}}(0)$ is incorrect and hence $\tilde{\mathfrak{A}}(\mathfrak{z})$ must be correct; then let $S_{i+1}(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z},0)$. In either case, set $f_{i+1}(\tilde{\mathfrak{A}}(a), i+1) = 1$.

   For other formulas $\tilde{\mathfrak{B}}(b)$, $S_{i+1}(\tilde{\mathfrak{B}}(b)) = S_j(\tilde{\mathfrak{B}}(b))$ where j is the greatest index $\leq i$ such that $f_i(\tilde{\mathfrak{B}}(b), j) = 1$. $S_{i+1}(\tilde{\mathfrak{B}}(b)) = (0,1)$ if no such j exists (i.e., $\tilde{\mathfrak{B}}(b)$ has never before received an example substitution). Also, let $f_{i+1}(\tilde{\mathfrak{B}}(b), i+1) = 1$. For all canonical formulas $\tilde{\mathfrak{C}}(c)$, let $f_{i+1}(\tilde{\mathfrak{C}}(c), j) = f_i(\tilde{\mathfrak{C}}(c), j)$ for $j \leq i$.

2. Case (1) does not apply, but at least one of the minimality axioms $\varepsilon_a \mathfrak{A}(a) \neq 0 \to \overline{\mathfrak{A}}(\delta(\varepsilon_a \mathfrak{A}(a))$ reduces to a false formula, $\mathfrak{z} \neq 0 \to \overline{\tilde{\mathfrak{A}}}(\mathfrak{z}-1)$. This is only possible if $S_i(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z}, 0)$. Again, pick the one of lowest level, and let $S_{i+1}(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z} - 1, 0)$ and $f_{i+1}(\tilde{\mathfrak{A}}(a), i+1) = 1$. Substitutions for other formulas and bookkeeping functions are defined as in case (1).

3. Neither case (1) nor (2) applies, but some instance of an ε-axiom of the form $\mathfrak{A}(a) \to \mathfrak{A}(\varepsilon_a \mathfrak{A}(a))$ or of a π-axiom of the form $\overline{\mathfrak{A}}(\varepsilon_a \mathfrak{A}(a)) \to \pi_a \mathfrak{A}(a) = 1$, e.g., $\tilde{\mathfrak{A}}(\mathfrak{a}) \to \tilde{\mathfrak{A}}(\mathfrak{z}))$ or $\overline{\tilde{\mathfrak{A}}}(\mathfrak{z}) \to 0 = 1$, reduces to an incorrect formula. We then have $S_i(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z}, 0)$ (since otherwise case (1) would apply). In either case, $|\tilde{\mathfrak{A}}(\mathfrak{z}))|_{S_i}$ must be incorrect. Let j be the least index where $S_j(\tilde{\mathfrak{A}}(a)) = (\mathfrak{z}, 0)$ and $f_i(\tilde{\mathfrak{A}}(a)), j) = 1$. At the preceding total substitution S_{j-1}, $S_j(\tilde{\mathfrak{A}}(a)) = (0, 1)$ or $S_{j-1}(A(a)) = (\mathfrak{z}+1, 0)$, and $|\mathfrak{A}(\mathfrak{z})|_{S_{j-1}}$ is correct. $\tilde{\mathfrak{A}}(\mathfrak{z})$ thus must reduce to different formulas under S_{j-1} and under S_i, which is only possible if a formula subordinate to $\tilde{\mathfrak{A}}$ reduces differently under S_{j-1} and S_i.

   For example, suppose $\tilde{\mathfrak{A}}(a)$ is really $\tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b))$. Then the corresponding ε-axiom would be

$$\tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b)) \to \tilde{\mathfrak{A}}(\underbrace{\varepsilon_a \tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b))}_{\varepsilon_a \tilde{\mathfrak{A}}(a)}, \varepsilon_b \tilde{\mathfrak{B}}(\underbrace{\varepsilon_a \tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b))}_{\varepsilon_a \tilde{\mathfrak{A}}(a)}, b))$$

   An instance thereof would be

$$\tilde{\mathfrak{A}}(\mathfrak{a}, \varepsilon_b \tilde{\mathfrak{B}}(\mathfrak{a}, b) \to \tilde{\mathfrak{A}}(\varepsilon_a \tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b)), \varepsilon_b \tilde{\mathfrak{B}}(\varepsilon_a \tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b)), b)).$$

   This formula, under a total substitution with $S_i(\tilde{\mathfrak{A}}(a, \varepsilon_b \tilde{\mathfrak{B}}(a, b))) = (\mathfrak{z}, 0)$ reduces to

$$\tilde{\mathfrak{A}}(\mathfrak{a}, \varepsilon_b \tilde{\mathfrak{B}}(\mathfrak{a}, b) \to \tilde{\mathfrak{A}}(\mathfrak{z}, \varepsilon_b \tilde{\mathfrak{B}}(\mathfrak{z}, b))$$

   The consequent of this conditional, i.e., $\tilde{\mathfrak{A}}(\mathfrak{z})$, can reduce to different formulas under S_i and S_{j-1} only if $\varepsilon_b \tilde{\mathfrak{B}}(\mathfrak{z}, b)$ receives different substitutions under S_i and S_{j-1}, and $\tilde{\mathfrak{B}}(a, b)$ is subordinate to $\tilde{\mathfrak{A}}(a)$.

   The next substitution is now defined as follows: Pick an innermost formula subordinate to $\tilde{\mathfrak{A}}(a)$ which changes substitutions, say $\tilde{\mathfrak{B}}(b)$. For all formulas $\tilde{\mathfrak{C}}(c)$

which are subordinate to $\tilde{\mathfrak{B}}(b)$ as well as $\tilde{\mathfrak{B}}(b)$ itself, we set $f_{i+1}(\tilde{\mathfrak{C}}(c),k) = 1$ for $j \leq k \leq i+1$ and $f_{i+1}(\tilde{\mathfrak{C}}(c),k) = 0$ for all other formulas. For $k < j$ we set $f_{i+1}(\tilde{\mathfrak{C}}(c),n) = f_i(\tilde{\mathfrak{C}}(c),k)$ for all $\tilde{\mathfrak{C}}(c)$. The next substitution S_{i+1} is now given by $S_{i+1}(\tilde{\mathfrak{C}}(c)) = S_k(\tilde{\mathfrak{C}}(c))$ for k greatest such that $f_{i+1}(\tilde{\mathfrak{C}}(c),k) = 1$ or $= (0,1)$ if no such k exists.

Readers familiar with the substitution method defined in Ackermann (1940) will note the following differences:

a. Ackermann (1940) uses the notion of a *type* of an ε-term and instead of defining total substitutions in terms of numeral substitutions for canonical ε-terms, assigns a function of finite support to ε-types. This change is merely a notational convenience, as these functional substitutions can be recovered from the numeral substitutions for canonical ε-terms. For example, if S assigns the substitutions to the canonical terms on the left, then a total substitution in the sense of Ackermann (1940) would assign the function g on the right to the type $\varepsilon_a\tilde{\mathfrak{A}}(a,b)$:

$$\begin{array}{ll} S(\tilde{\mathfrak{A}}(a,\mathfrak{z}_1) = \mathfrak{z}'_1 & \quad g(\varepsilon_a\tilde{\mathfrak{A}}(a,b))(\mathfrak{z}_1) = \mathfrak{z}'_1 \\ S(\tilde{\mathfrak{A}}(a,\mathfrak{z}_3) = \mathfrak{z}'_2 & \quad g(\varepsilon_a\tilde{\mathfrak{A}}(a,b))(\mathfrak{z}_2) = \mathfrak{z}'_2 \\ S(\tilde{\mathfrak{A}}(a,\mathfrak{z}_2) = \mathfrak{z}'_3 & \quad g(\varepsilon_a\tilde{\mathfrak{A}}(a,b))(\mathfrak{z}_3) = \mathfrak{z}'_3 \end{array}$$

b. In case (2), dealing with the least number (induction) axiom, the next substitution is defined by reducing the substituted numeral $\mathfrak{z}$ by 1, whereas in (Ackermann 1940), we immediately proceed to the least $\mathfrak{z}'$ such that $\tilde{\mathfrak{A}}(\mathfrak{z}')$ is correct. This makes the procedure converge more slowly, but also suggests that in certain cases (depending on which other critical formulas occur in the proof), the solving substitution does not necessarily provide example substitutions which are, in fact, least witnesses.

c. The main difference in the method lies in case (3). Whereas in (1940), example substitutions for all ε-types of rank lower than that of the changed $\varepsilon_a\tilde{\mathfrak{A}}(a)$ are retained, and all others are reset to initial substitutions (functions constant equal to 0), in (1924b), only the substitutions of some ε-terms actually subordinate to $\varepsilon_a\tilde{\mathfrak{A}}(a)$ are retained, while others are not reset to initial substitutions, but to substitutions defined at some previous stage.

3.3.5 Assessment and Complications

A detailed analysis of the method and of the termination proof given in the last part of Ackermann's dissertation must wait for another occasion, if only for lack of space. A preliminary assessment can, however, already be made on the basis of the outline of the substitution process above. Modulo some needed clarification in the definitions, the process is well-defined and terminates at least for proofs containing only least-number axioms (critical formulas corresponding to axiom (4)) of rank 1. The proof that the procedure terminates (§9 of Ackermann (1924b)) is opaque, especially in comparison to the proof by transfinite induction for primitive recursive arithmetic. The definition of a substitution method for second-order ε-terms is insufficient, and in hindsight it is clear that a correct termination proof for this part could not have been given with the methods available.[49]

Leaving aside, for the time being, the issue of what was *actually* proved in Ackermann (1924b), the question remains of what was *believed* to have been proved at the time. The system, as given in (1924b), had two major shortcomings: A footnote, added in proof, states:

> [The formation of ε-terms] is restricted in that a function variable $f(a)$ may not be substituted by a functional $\mathfrak{a}(a)$, in which a occurs in the scope of an ε_f.[50]

This applies in particular to the second-order ε-axioms

$$A_a(f(a)) \rightarrow A_a(\varepsilon_f A_b(f(b))(a)).$$

If we view $\varepsilon_f A_b(f(b))(a)$ as the function "defined by" A, and hence the ε-axiom as the ε-calculus analog of the comprehension axiom, this amounts roughly to a restriction to arithmetic comprehension, and thus a predicative system. This shortcoming, and the fact that the restriction turns the system into a system of predicative mathematics was pointed out by von Neumann (1927).

A second lacunae was the omission of an axiom of ε-extensionality for second-order ε-terms, i.e.,

$$(\forall f)(A(f) \rightleftharpoons B(f)) \rightarrow \varepsilon_f A(f) = \varepsilon_f B(f),$$

which corresponds to the axiom of choice. Both problems were the subject of correspondence with Bernays in 1925.[51] A year later, Ackermann is still trying to extend and correct the proof, now using ε-types:

> I am currently working again on the ε_f-proof and am pushing hard to finish it. I have already told you that the problem can be reduced to one of number theory. To prove the number-theoretic theorem seems to me, however, equally hard as the problem itself. I am now again taking the approach, which I have tried several times previously, to extend the definition of a ground type so that even ε with free function variables receive a substitution. This approach seems to me the most natural, and the equality axioms $(f)(A(f) \rightleftharpoons B(f)) \to \varepsilon_f Af \equiv \varepsilon_f Bf$ would be treated simultaneously. I am hopeful that the obstacles previously encountered with this method can be avoided, if I use the ε_a formalism and use substitutions for the ε_f which may contain ε_a instead of functions defined without ε. I have, however, only thought through some simple special cases.[52]

In 1927, Ackermann developed a second proof of ε-substitution, using some of von Neumann's ideas (in particular, the notion of an ε-type, *Grundtyp*). The proof is unfortunately not preserved in its entirety, but references to it can be found in the correspondence. On April 12, 1927, Bernays writes to Ackermann:

> Finally I have thought through your newer proof for consistency of the ε_a's based on what you have written down for me befor your departure, and believe that I have seen the proof to be correct.[53]

Ackermann also refers to the proof in a letter to Hilbert from 1933:

> As you may recall, I had at the time a second proof for the consistency of the ε_a's. I never published that proof, but communicated it to Prof. Bernays orally, who then verified it. Prof. Bernays wrote to me last year that the result does not seem to harmonize with the work of Gödel.[54]

In Hilbert's address to the International Congress of Mathematicians (Hilbert 1928a), the success of Ackermann's and von Neumann's work on ε-substitution for first-order systems is also taken for granted. Although Hilbert poses the extension of the proof to second-order systems as an open problem, there seems no doubt in his mind that the solution is just around the corner.[55]

It might be worthwhile to mention at this point that at roughly the same time a third attempt to find a satisfactory consistency proof was made. This attempt was based not on ε-substituion, but on Hilbert's so-called unsuccessful proof (*verunglückter Beweis*).

> While working on the *Grundlagenbuch*, I found myself motivated to rethink Hilbert's second consistency proof for the ε-axiom, the so-called "unsuccessful" proof, and it now seems to me that it can be fixed after all.[56]

This proof bears a striking resemblance to the proof the first ε-theorem in (Hilbert and Bernays 1939) and to a seven-page sketch in Bernays's hand of a "consistency proof for the logical axiom of choice" found bound with lecture notes to Hilbert's course on "Elements and principles of mathematics" of 1910.[57] This "unsuccessful" proof seems to me to be another but independent contribution to the development of logic and the ε-calculus, independent of the substituion method. Note that Bernays's proof of Herbrand's theorem in (Hilbert and Bernays 1939) is based on the (second) ε-theorem is the first correct published proof of that important result.

The realization that the consistency proof even for first-order ε's was problematic came only with Gödel's incompleteness results. In a letter dated March 10, 1931, von Neumann presents an example that shows that in the most recent version of Ackermann's proof, the length of the substitution process not only depends on the rank and degree of ε-terms occurring in the proof, but also on numerical values used as substitutions. He concludes:

> I think that this answers the question, which we recently discussed when going through Ackermann's modified proof, namely whether an estimate of the length of the correction process can be made uniformly and independently of numerical substituends, in the negative. At this point the proof of termination of the procedure (for the next higher degree, i.e., 3) has a gap.[58]

There is no doubt that the discussion of the consistency proof was precipitated by Gödel's results, as both von Neumann and Bernays were aware of these results, and at least von Neumann realized the implications for Hilbert's Program and the prospects of a finitistic consistency proof for arithmetic. Bernays corresponded with Gödel on the relevance of Gödel's result for the viability of the project of consistency proofs just before and after von Neumann's counterexample located the difficulty in Ackermann's proof. On January 18, 1931, Bernays writes to Gödel:

> If one, as does von Neumann, assumes as certain that any finite consideration can be formulated in the framework of System *P*—I think, as you do too, that this is not at all obvious—one arrives at the conclusion that a finite proof of consistency of *P* is impossible.

The puzzle, however, remained unresolved for Bernays even after von Neumann's example, as he writes to Gödel just after the exchange with von Neumann, on April 20, 1931:

> The confusion here is probably connected to that about Ackermann's proof for the consistency of number theory (System $\mathfrak{Z}$), which I have not so far been able to clarify.
>
> That proof—on which Hilbert has reported in his Hamburg talk on the "foundations of mathematics".[59] [. . .]—I have repeatedly thought through and found correct. On the basis of your results one must now conclude that this proof cannot be formalized within System $\mathfrak{Z}$; indeed, this must hold even if one restricts the system whose consistency is to be proved by leaving only addition and multiplication as recursive definitions. On the other hand, I do not see which part of Ackermann's proof makes the formalization within $\mathfrak{Z}$ impossible, in particular if the problem is so restricted.[60]

Gödel's results thus led Bernays, and later Ackermann to reexamine the methods used in the consistency proofs. A completion of the project had to wait until 1940, when Ackermann was able to carry through the termination proof based on transfinite induction—following Gentzen (1936)—on ε_0.

3.4 Conclusion

With the preceding exposition and analysis of the development of axiomatizations of logic and mathematics and of Hilbert and Ackermann's consistency proofs I hope to have answered some open questions regarding the historical development of Hilbert's Program. Hilbert's *technical* project and its evolution is without doubt of tremendous importance to the history of logic and the foundations of mathematics in the 20th century. Moreover, an understanding of the technical developments can help to inform an understanding of the history and prospects of the *philosophical* project. The lessons drawn in the discussion, in particular, of Ackermann's use of transfinite induction, raise more questions. The fact that transfinite induction in the form used by Ackermann was so readily accepted as finitist, not just by Ackermann himself, but also by Hilbert and Bernays leaves open two possibilities: either they were simply wrong in taking the finitistic nature of Ackermann's proof for granted and the use of transfinite induction simply cannot be reconciled with the finitist standpoint as characterized by Hilbert and Bernays in other writings, or the common view of what Hilbert thought the finitist standpoint to consist in must be revised. Specifically, it seems that the explanation of why transfinite induction is acceptable stresses one aspect of finitism while downplaying another: the *objects* of finitist reasoning are—finite and—intuitively given, whereas the methods of proof were not required to have the epistemic strength that the finitist standpoint is usually thought to require (i.e., to guarantee, in one sense or another, the intuitive evidence of the resulting theorems). Of course, the question of whether Hilbert can make good on his claims that finitistic reasoning affords this intutive evidence of its theorems is one of the main difficulties in a philosphical assessment of the project (see, e.g., Parsons (1998a)).

I have already hinted at the implications of a study of the practice of finitism for philosophical reconstructions of the finitist view (in note 39). We are of course free to latch on to this or that aspect of Hilbert's ideas (finitude, intuitive evidence, or surveyability) and develop a philosophical view around it. Such an approach can be very fruitful, and have important and inightful results (as, e.g, the example of Tait's (1981) work shows). The question is to what extent such a view should be accepted as a reconstruction of Hilbert's view as long as it makes the practice of the technical project come out off base. Surely rational reconstruction is governed by something like a principle of charity. Hilbert and his students, to the extent possible, should be construed so that what they preached is reflected in their practice. This requires, of course, that we know what the practice was. If nothing else, I hope to have provided some of the necessary data for that.

Notes

1. Hilbert (1905c, p. 131). For a general discussion of Hilbert's views around 1905, see Peckhaus (1990, Chapter 3).

2. „Die von ZERMELO benutze axiomatische Methode ist zwar unanfechtbar und unentbehrlich. Es bleibt dabei doch die Frage offen, ob die aufgestellten Axiome nicht etwa einen Widerspruch

einschliessen. Ferner erhebt sich die Frage, ob und inwieweit sich das Axiomensystem aus der Logik ableiten lässt. [. . . D]er Versuch einer Zurückführung auf die Logik scheint besonders deshalb aussichtsvoll, weil zwischen Mengen, welche ja die Gegenstände in ZERMELOs Axiomatik bilden, und den Prädikaten der Logik ein enger Zusammenhang besteht. Nämlich die Mengen lassen sich auf Prädikate zurückführen.

Diesen Gedanken haben FREGE, RUSSEL[L] und WEYL zum Ausgangspunkt genommen bei ihren Untersuchungen über die Grundlagen der Mathematik.“ Hilbert (1920b, pp. 27–28).

3. „Wir müssen uns nämlich fragen, was es bedeuten soll: „es gibt ein Prädikat *P*.“ In der axiomatischen Mengentheorie bezieht sich das „es gibt“ immer auf den zugrunde gelegten Bereich $\mathfrak{B}$. In der Logik können wir zwar auch die Prädikate zu einem Bereich zusammengefasst denken; aber dieser Bereich der Prädikate kann hier nicht als etwas von vorneherein Gegebenes betrachtet werden, sondern die Prädikate müssen durch logische Operationen gebildet werden, und durch die Regeln der Konstruktion bestimmt sich erst nachträglich der Prädikaten-Bereich.

Hiernach ist ersichtlich, dass bei den Regeln der logischen Konstruktion von Prädikaten die Bezugnahme auf den Prädikaten-Bereich nicht zugelassen werden kann. Denn sonst ergäbe sich ein *circulus vitiosus*.“ (Hilbert 1920b, p. 31).

4. “RUSSELL geht von dem Gedanken aus, dass es genügt, das zur Definition der Vereinigungsmenge unbrauchbare Prädikat durch ein sachlich gleichbedeutendes zu ersetzen, welches nicht dem gleichen Einwande unterliegt. Allerdings vermag er ein solches Prädikat nicht anzugeben, aber er sieht es als ausgemacht an, dass ein solches existiert. In diesem Sinne stellt er sein „Axiom der Reduzierbarkeit“ auf, welches ungefähr folgendes besagt: „Zu jedem Prädikat, welches durch (einoder mehrmalige) Bezugnahme auf den Prädikatenbereich gebildet ist, gibt es ein sachlich gleichbedeutendes Prädikat, welches keine solche Bezugnahme aufweist.

Hiermit kehrt aber RUSSELL von der konstruktiven Logik zu dem axiomatischen Standpunkt zurück. [. . .]

Das Ziel, die Mengenlehre und damit die gebräuchlichen Methoden der Analysis auf die Logik zurückzuführen, ist heute nicht erreicht und ist vielleicht überhaupt nicht erreichbar.” (Hilbert 1920b, pp. 32–33).

5. “. . . zur Vermeidung von Paradoxien ist daher eine teilweise gleichzeitige Entwicklung der Gesetze der Logik und der Arithmetik erforderlich.” (Hilbert 1905c, p. 176).

6. “In Anbetracht der grossen Mannigfaltigkeit von Verknüpfungen und Zusammenhängen, welche die Arithmetik aufweist, ist es von vornherein ersichtlich, dass wir die Aufgabe des Nachweises

der Widerspruchslosigkeit nicht mit einem Schlage lösen können. Wir werden vielmehr so vorgehen, dass wir zunächst nur die einfachsten Verknüpfungen betrachten und dann schrittweise immer höhere Operationen und Schlussweisen hinzunehmen, wobei dann für jede Erweiterung des Systems der Zeichen und der Uebergangsformeln einzeln der Nachweis zu erbringen ist, dass sie die auf der vorherigen Stufe festgestellte Widerspruchsfreiheit nicht aufheben.

Ein weiterer wesentlicher Gesichtspunkt ist, dass wir, gemäss unserem Plan der restlosen Formalisierung der Arithmetik, den eigentlich mathematischen Formalismus im Zusammenhang mit dem Formalismus der logischen Operationen entwickeln müssen, sodass—wie ich es ausgedrückt habe—ein simultaner Aufbau von Mathematik und Logik ausgeführt wird." (Hilbert 1922b, pp. 8a–9a). The passage is not contained in Kneser's notes (Hilbert and Bernays 1923a) to the same course.

7. The notes by Kneser (Hilbert and Bernays 1923a) do not contain the list of systems below. The version of the ε-calculus used in the addendum is the same as that used in Kneser's notes, and differs from the presentation in Ackermann (1924b), submitted February 20, 1924.

8. Hilbert and Bernays (1923b, pp. 17, 19).

9. "**Disposition**. Stufe II war elementares Rechnen Axiome 1–16

Stufe III. Nun elementare Zahlentheorie

Schema für Def. von Funktionen durch Rekursion u. Schlussschema
wollen [?] unser Schlussschema noch das Induktionsschema hinzuziehen
Wenn auch inhaltlich das wesentlich mit den Ergebnissen der anschauliche gewonnenen [?] Zahlenth. übereinstimmt, so doch jetzt Formeln z.B. $a+b=b+a$.

Stufe IIII. Transfinite Schlussweise u. teilweise Analysis

Stufe V. Höhere Variablen-Gattungen u. Mengenlehre. Auswahlaxiom

Stufe VI. Zahlen d[er] 2^{ten} Zahlkl[asse], Volle transfin[ite] Induktion. Höhere Typen. Continuumsproblem, transfin[ite] Induktion für Zahlen der 2^{ten} Zahlkl.

Stufe VII. 1.) Ersetzung der ∞ vielen Definitionsschemata durch ein Axiom. 2.) Analysis u[nd] Mengenlehre. Auf der 4^{ten} Stufe nochmals der volle Satz von der oberen Grenze

Stufe VIII. Formalisierung der Wohlordnung. " (Hilbert and Bernays 1923a; *Ergänzung*, sheet 1).

10. The proof can also be found in Hilbert (1922c, pp. 171–173); cf. Mancosu (1998a, pp. 208–210).

11. "Somit sehen wir uns veranlasst, die Beweise als solche zum Gegenstand der Untersuchung zu machen; wir werden zu einer Art von *Beweistheorie* gedrängt, welche mit den Beweisen selbst als Gegenständen operiert.

Für die Denkweise der gewöhnlichen Zahlentheorie sind die Zahlen das gegenständlich-Aufweisbare, und die Beweise der Sätze über die Zahlen fallen schon in das gedankliche Gebiet. Bei unserer Untersuchung ist der Beweis selbst etwas Aufweisbares, und durch das Denken über den Beweis kommen wir zur Lösung unseres Problems.

Wie der Physiker seinen Apparat, der Astronom seinen Standort untersucht, wie der Philosoph Vernunft-Kritik übt, so braucht der Mathematiker diese Beweistheorie, um jeden mathematischen Satz durch eine Beweis-Kritik sicherstellen zu können." Hilbert (1920b, pp. 39–40). Almost the same passage is found in Hilbert (1922c, pp. 169–170), cf. Mancosu (1998a, p. 208).

12. For a detailed discussion of these influences, see Mancosu (1999a).

13. (Hilbert 1922b, part 2, p. 3). Kneser's *Mitschrift* of these lectures contains a different system which does not include negation. Instead, numerical inequality is a primitive. This system is also found in Hilbert's first talks on the subject in Copenhagen and Hamburg in Spring and Summer of 1921. Hilbert (1923), a talk given in September 1922, and Kneser's notes to the course of Winter Semester 1922–23 (Hilbert and Bernays 1923b) do contain the new system with negation. This suggests that the developments of Hilbert's 1921–22 lectures were not incorporated into the published version of Hilbert's Hamburg talk (1922c). Although (1922c) was published in 1922, and a footnote to the title says "This communication is essentially the content of the talks which I have given in the Spring of this year in Copenhagen [...] and in the Summer in Hamburg [...]," it is clear that the year in question is 1921, when Hilbert addressed the Mathematisches Seminar of the University of Hamburg, July 25–27, 1921. A report of the talks was published by Reidemeister in *Jahrbuch der Deutschen Mathematiker-Vereinigung* **30**, 2. Abt. (1921), 106. Hilbert and Bernays (1923b) also have separate axioms for conjunction and disjunction, while in (1923) it is extended it by quantifiers.

14. The procedure whereby we pass from $\mathfrak{A}$ to $\mathfrak{A}'$ is simple in this case, provided we keep track of which variables are substituted for below the inference. In general, the problem of deciding whether a formula is a substitution instance of another, and to calculate the substitution which would make the latter syntactically identical to the former is known as *matching*. Although not computationally difficult, it is not entirely trivial either.

15. "Nennen wir eine Formel, in der keine Variablen und keine Funktionale ausser Zahlzeichen vorkommen, eine „*explizite [numerische] Formel*", so können wir das gefundene Ergebnis so aussprechen: Jede beweisbare explizite [numerische] Formel ist Endformel eines Beweises, dessen sämtliche Formeln explizite Formeln sind.

Dieses müsste insbesondere von der Formel $0 \neq 0$ gelten, wenn sie beweisbar wäre. Der verlangte Nachweis der Widerspruchsfreiheit ist daher erbracht, wenn wir zeigen, dass es keinen Beweis der Formel geben kann, der aus lauter expliziten Formeln besteht.

Um diese Unmöglichkeit einzusehen, genügt es, eine konkret feststellbare Eigenschaft zu finden, die erstens allen den expliziten Formeln zukommt, welche durch Einsetzung aus einem Axiom entstehen, die ferner bei einem Schluss sich von den Prämissen auf die Endformel überträgt, die dagegen nicht auf die Formel $0 \neq 0$ zutrifft." (Hilbert 1922b, part 2, pp. 27–28).

16. "Wir teilen die expliziten Formeln in „*richtige*" und „*falsche*" ein. Die expliziten Primformeln sind Gleichungen, auf deren beiden Seiten *Zahlzeichen* stehen. Eine solche *Gleichung* nennen wir *richtig*, wenn die beiderseits stehenden Zahlzeichen *übereinstimmen*; andernfalls nennen wir sie *falsch*. Eine *Ungleichung*, auf deren beiden Seiten Zahlzeichen stehen, nennen wir *richtig*, falls die beiden Zahlzeichen *verschieden* sind; sonst nen[n]en wir sie *falsch*.

In der Normalform einer beliebigen expliziten Formel haben alle Disjunktionsglieder die Gestalt von Gleichungen oder Ungleichungen, auf deren beiden Sieten Zahlzeichen stehen.

Wir nennen nun eine *allgemeine explizite Formel richtig*, wenn in der zugehörigen Normalform jede als Konjunktionsglied auftretende (bezw. die ganze Normalform ausmachende) Disjunktion eine richtige Gleichung oder eine richtige Ungleichung als Glied enthält. Andernfalls nennen wir die Formel *falsch*. [. . .]

Nach der gegebenen Definition lässt sich die Frage, ob eine explizierte [sic] Formel richtig oder falsch ist, in jed[e]m Falle *konkret entscheiden*. Hier gilt also das „tertium non datur" [. . .]" (Hilbert 1922b, part 2, p. 33).

17. Hilbert (1918c, pp. 149–150). See also Section 2.2.3 (Zach 1999, Section 2.3).

18. A sketch of the consistency proof is found in the Kneser *Mitschrift* to the 1921–22 lectures (Hilbert 1922a) in Heft II, pp. 23–32 and in the official notes by Bernays (Hilbert 1922b, part 2, pp. 19–38). The earlier Kneser *Mitschrift* leaves out step (1), and instead of eliminating variables introduces the notion of *einsetzungsrichtig* (correctness under substitution, i.e., every substitution instance is correct). These problems were avoided in the official Bernays typescript. The Kneser notes did contain a discussion of recursive definition and induction, which is not included in the official notes; more about these in the next section.

19. In the 1921–22 lectures, it is initially argued that the result of applying transformations (1)–(3) results in a *proof* of the same end formula (if substitutions are added to the initial formulas). Specifically, it is suggested that the result of applying elimination of variables and reduction of

functionals to the axioms results in formulas which are substitution instances of axioms. It was quickly realized that this is not the case. (When Bernays presented the proof in the 1922–23 lectures on December 14, 1922, he comments that the result of the transformation need not be a proof (Hilbert and Bernays 1923b, p. 21). The problem is the axiom of equality

$$a = b \to (A(a) \to A(b)).$$

Taking $A(c)$ to be $\delta(c) = c$, a substitution instance would be

$$0+1+1 = 0 \to (\delta(0+1+1) = 0+1+1 \to \delta(0) = 0)$$

This reduces to

$$0+1+1 = 0 \to (0+1 = 0+1+1 \to 0 = 0)$$

which is not a substitution instance of the equality axiom. The consistency proof itself is not affected by this, since the resulting formula is still correct (in Hilbert's technical sense of the word). The official notes to the 1921–22 lectures contain a 2-page correction in Bernays's hand (Hilbert 1922b, part 2, between pp. 26 and 27).

20. The induction rule is not used in (Ackermann 1924b), since he deals with stage III only in passing and attempts a consistency proof for all of analysis. There, the induction rule is superseded by an ε-based induction axiom. For a consistency proof of stage III alone, an induction rule is needed, since an axiom cannot be formulated without quantifiers (or ε). The induction rule was introduced for stage III in the Kneser notes to the 1921–22 lectures (Hilbert 1922a; Heft II, p. 32) and the 1922–23 lectures (Hilbert and Bernays 1923b, p. 26). It is not discussed in the official notes or the publications from the same period (Hilbert 1922c, 1923).

21. The general tenor, outlook, and aims of Skolem's work are sufficiently different from that of Hilbert to suggest there was no influence either way. Skolem states in his concluding remarks that he wrote the paper in 1919, after reading Russell and Whitehead's *Principia Mathematica*. However, neither Hilbert nor Bernays's papers contain an offprint or manuscript of Skolem's paper, nor correspondence. Skolem is not cited in any of Hilbert's, Bernays's, or Ackermann's papers of the period, although the paper is referenced in (Hilbert and Bernays 1934).

22. "Uns fehlt noch ganz das Ax[iom] der vollst[ändigen] Induktion. Man könnte meinen, es wäre

$$\{Z(a) \to (A(a) \to A(a+1))\} \to \{A(1) \to (Z(b) \to A(b))\}$$

Das ist es nicht; denn man setze $a = 1$. Die Voraussetzung muß für *alle a* gelten. Wir haben aber noch gar kein Mittel, das *Alle* in die Voraussetzung zu bringen. Unser Formalismus reich noch nicht hin, das Ind.ax. aufzuschreiben.

Aber als Schema können wir es: Wir erweitern unsere Beweismethoden durch das nebenstehende Schema

$$\frac{\mathfrak{K}(1) \quad \mathfrak{K}(a) \rightarrow \mathfrak{K}(a+1)}{Z(a) \rightarrow \mathfrak{K}(a)}$$

Jetzt ist es vernünftig, zu fragen, ob dies Schema zum Wspruch führen kann." (Hilbert 1922a, p. 32). Z is the predicate expressing "is a natural number," it disappears from later formulations of the schema.

23. "Wie ist es bei Rekursionen? $\varphi(\mathfrak{z})$ komme vor. Entweder 0, dann setzten wir $\mathfrak{a}$ dafür. Oder $\varphi(\mathfrak{z}+1)$: $\mathfrak{b}(\mathfrak{z}, \varphi\mathfrak{z})$. Beh[auptung]: Das Einsetzen kommt zu einem Abschluß, wenn wir zu innerst anfangen." (Hilbert and Bernays 1923a, p. 29)

24. "Nicht endlich (durch Rek[ursion]) definiert ist z.B. $\varphi(a) = 0$ wenn es ein b gibt, so daß $a^5 + ab^3 + 7$ Primz[ahl] ist sonst $= 1$. Aber erst bei solchen Zahlen und Funktionen beginnt das eigentliche math[ematische] Interesse, weil dort die Lösbarkeit in endlich vielen Schritten nicht vorauszusehen ist. Wir haben die Überzeugung, daß solche Fragen wie nach dem Wert $\varphi(a)$ lösbar, d.h. daß $\varphi(a)$ doch endlich definierbar ist. Darauf können wir aber nicht warten: wir müssen solche Definitionen zulassen, sonst würden wir den freien Betrieb der Wissenschaft einschränken. Auch den Begriff der Funktionenfunktion brauchen wir." (Hilbert 1922a, Heft III, pp. 1–2).

25. A full proof is given by Ackermann (1924b).

26. "Als erstes zeigt man, daß man alle Variablen fortschaffen kann, weil auch hier nur freie Var[iable] vorkommen. Wir suchen die innersten τ und α. Unter diesen stehen nur endlich definierte Funkt[ionen] φ, φ'... Unter diesen können einige im Laufe des Beweises für f in die Ax[iome] eingesetzt sein. 1: $\tau(\varphi) = 0 \rightarrow (Z(\mathfrak{a} \rightarrow \varphi\mathfrak{a} = 1)$ wo $\mathfrak{a}$ ein Funktional ist. Wenn dies *nicht* benutzt wird, setze ich alle $\alpha(\varphi)$ und $\tau(\varphi)$ gleich Null. Sonst reduziere ich $\mathfrak{a}$ und $\varphi(\mathfrak{a})$ und sehe, ob $Z(\mathfrak{a} \rightarrow \varphi(\mathfrak{a}) = 1$ in allen ... wo sie vorkommt, richtig ist. Ist die richtig, so setze ich $\tau = 0$ $\alpha = 0$. Ist sie falsch, d.h. is $\mathfrak{a} = \mathfrak{z}$ $\varphi(\mathfrak{z} \neq 1$, so setzen wir $\tau(\varphi) = 1$, $\alpha(\varphi) = \mathfrak{z}$. Dabei bleibt der Beweis Beweis. Die an Stelle der Axiome gesetzten Formeln sind richtig.

Der Gedanke ist: wenn ein Beweis vorliegt, so kann ich aus ihm ein Argument finden für das $\varphi = 1$ ist). So beseitigt man schrittweise die τ und α aund Anwendungen von 1 2 3 4 und erhält einen Beweis von $1 \neq 1$ aus I–V und richtigen Formeln d.h. aus I–V,

$$\tau(f,b) = 0 \quad \rightarrow \quad \{Z(a) \rightarrow f(a,b) = 1\}$$

$$\tau(f,b) \neq 0 \quad \rightarrow \quad Z(f(\alpha,b))$$
$$\tau(f,b) \neq 0 \quad \rightarrow \quad f(\alpha(f,b),b) \neq 1$$
$$\tau(f,b) \neq 0 \quad \rightarrow \quad \tau(f,b) = 1\text{”}$$

(Hilbert 1922a), Heft III, 3–4. The lecture is dated February 23, 1922.

27. “Was fehlt uns?

1. in logischer Hinsicht. Wir haben nur gehabt den Aussagenkalkül mit der Erweiterung auf freie Variable d.h. solche für die beliebige Funktionale eingesetzt werden konnten. Es fehlt das Operieren mit „alle“ und „es gibt“.

2. Wir haben das Induktionsschema hinzugefügt, ohne W[iderspruchs]-f[reiheits] Beweis und auch nur provisorisch, also in der Absicht, es wegzuschaffen.

3. Bisher nur die arithmethischen Axiome genau [?] die sich auf ganze Zahlen beziehen. Und die obigen Mängel verhindern uns ja natürlich die Analysis aufzubauen (Grenzbegriff, Irrationalzahl).

Diese 3 Punkte liefern schon Disposition und Ziele für das Folgende.

Wir wenden uns zu 1.) Es ist ja an sich klar, dass eine Logik ohne „alle“—„es gibt“ Stückwerk wäre, ich erinnere wie gerade in der Anwendung dieser Begriffe, und den sogenannten transfiniten Schlussweisen die Hauptschwierigkeiten entstanden. Die Frage der Anwendbarkeit dieser Begriffe auf ∞ Gesamtheiten haben wir noch nicht behandelt. Nun könnten wir so verfahren, wie wir es beim Aussagen-Kalkül gemacht haben: einige, möglichst einfache [???] als Axiome zu formalisieren, aus denen sich [sic] dann alle übrigen folgen. Dann müsste der W-f Beweis geführt werden—unserem allgem[einen] Programm gemäss: mit unserer Einstellung, dass Beweis eine vorliegende Figur ist. Für den W-f Beweis grosse Schwierigkeiten wegen der gebundenen Variabeln. Die tiefere Untersuchung zeigt aber, dass der eigentliche Kern der Schwierigkeit an einer anderen Stelle liegt, auf die man gewöhnlich erst später Acht giebt und die auch in der Litteratur erst später wahrgenommen worden ist.”(Hilbert and Bernays 1923b; *Ergänzung*, sheets 3–4).

28. “[Dieser Kern liegt] beim *Auswahlaxiom* von Zermelo. [...] Die Einwände richten sich gegen das Auswahlprinzip. Sie müßten sich aber ebenso gegen „alle“ und „es gibt“ richten, wobei derselbe Grundgedanke zugrunde liegt.

Wir wollen das Auswahlaxiom erweitern. Jeder Aussage mit einer Variablen $A(a)$ ordnen wir ein Ding zu, für das die Aussage nur dann gilt, wenn sie allgemein gilt. Also ein Gegenbeispiel, wenn es existiert.

$\varepsilon(A)$, eine individuelle logische Funktion. [...] ε genüge dem *transfiniten Axiom*:

(16) $A(\varepsilon A) \to Aa$

z.B. Aa heiße: a ist bestechlich. εA ist Aristides." (Hilbert and Bernays 1923a, pp. 30–31). The lecture is dated February 1, 1922, given by Hilbert. The corresponding part of Hilbert's notes for that lecture in (Hilbert and Bernays 1923b; Ergänzung, sheet 4) contains page references to (Hilbert 1923; pp. 152 and 156, paras. 4–6 and 17–19 of the English translation), and indicates the changes made for the lecture, specifically, to replace τ by ε.

29. See section 3.3.4 on the ε-substitution method.

30. "Wenn wir eine *Funktionsvariable* haben:

$$A\varepsilon_f Af \to Af$$

(π fällt fort)? ε komme *nur* mit $\mathfrak{A}$ vor (z.B. $f0 = 0$, $ff0 = 0$). Wie werden wir die Funktionsvariablen ausschalten? Statt fc setzen wir einfach c. Auf die *gebundenen* trifft das nicht zu. Für diese nehmen wir probeweise eine bestimmte Funktion z.B. δa und führen damit die Reduktion durch. Dann steht z.B. $\mathfrak{A}\delta \to \mathfrak{A}\varphi$. Diese reduziert ist r[ichtig] oder f[alsch]. Im letzten Falle is $\mathfrak{A}\varphi$ falsch. Dann setzen wir überall φ für $\varepsilon_f \mathfrak{A} f$. Dann steht $\mathfrak{A}\varphi \to \mathfrak{A}\psi$. Das ist sicher *r[ichtig]* da $\mathfrak{A}\varphi$ f[alsch] ist." (Hilbert and Bernays 1923a, pp. 38–39).

31. For a more detailed survey of Ackermann's scientific contributions, see Hermes (1967). A very informative discussion of Ackermann's scientific correspondence can be found in Ackermann (1983).

32. "In seiner Arbeit „Begründung des „Tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit" hat Ackermann im allgemeinsten Falle gezeigt, dass der Gebrauch der Worte „alle" und „es gibt", des „Tertium non datur" widerspruchsfrei ist. Der Beweis erfolgt unter ausschliesslicher Benutzung primitiver und endlicher Schlussweisen. Es wird alles an dem mathematischen Formalismus sozusagen direkt demonstriert.

Ackermann hat damit unter Ueberwindung erheblicher mathematischer Schwierigkeiten ein Problem gelöst, das bei den modernene auf eine Neubegründung der Mathematik gerichteten Bestrebungen an erster Stelle steht." Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Cod. Ms. Hilbert 458, sheet 6, no date. The three-page letter was evidently written in response to a request by the President of the International Education Board, dated May 1, 1924.

33. "Ich bemerke nur, dass Ackermann meine Vorlesungen über die Grundlagen der Math[ematik] in den letzten Semestern gehört hat und augenblicklich einer der besten Herren der Theorie

ist, die ich hier entwickelt habe." *ibid.*, sheet 2. The draft is dated March 19, 1924, and does not mention Russell by name. Sieg (1999), however, quotes a letter from Russell's wife to Hilbert dated May 20, 1924, which responds to an inquiry by Hilbert concerning Ackermann's stay in Cambridge. Later in the letter, Hilbert expresses his regret that the addressee still has not been able to visit Göttingen. Sieg documents Hilbert's effort in the preceding years to effect a meeting in Göttingen; it is therefore quite likely that the addressee was Russell.

34. See Section 4.3.3 for a discussion of primitive recursion.

35. Ackermann only requires that b be bound by the occurrence of φ, but this is not enough for his proof.

36. "Jedes der Funktionale, aus denen sich unser vorliegendes Funktional aufbaut, hat einen bestimmten Rang bezüglich der letzten, der vorletzten usw. bis ersten Rekursionsfunktion. Jede derartige Rangkombination wird durch $\mathfrak{n}$ geordnete Zahlen gekennzeichnet. Alle diese endlich vielen verschiedenen Rangkombinationen, die bei unserem Funktional auftreten, wollen wir nun ordnen. Bei zwei verschiedenen Rangkombinationen schreiben wir die entsprechenden Zahlen untereinander, also zuerst den Rang bezüglich der letzten, dann der vorletzten Funktion usw. An irgendeiner Stelle sind dann die untereinanderstehenden Zahlen zuerst verschieden. Diejenige Rangkombination heißt nun die höhere, bei der an der betreffenden Stelle die größere Zahl steht. In dieser Weise ordnen wir alle die endlich vielen verschiedenen Rangkombinationen, die bei dem vorliegeneden Funktional auftreten. Zu jeder Rangkombination schreieben wir dann auf, wieviel Funktionale dieser Art in dem vorliegenden vorkommen. Die Gesamtheit dieser Zahlen wollen wir den Index des Funktionals nennen." (Ackermann 1924b, p. 15).

37. According to Ackermann's definition of subordination, this would not be true. A subterm of $\mathfrak{c}(b)$ might contain a bound variable and thus not be a constant subterm, but the variable could be bound by a function symbol in $\mathfrak{t}$ other than the occurrence of φ under consideration. See note 35.

38. Ackermann (1924b, p. 18)

39. Tait (1981) argues that finitism coincides with primitive recursive arithmetic, and that therefore the Ackermann function is not finitistic. Tait does not present this as a historical thesis, and his conceptual analysis remains unaffected by the piece of historical evidence presented here. For further evidence (dating however mostly from after 1931) see Section 4.3.4 (Zach 1998, §5) and Tait's response in (2000).

40. "Der Abbau der Funktional durch Reduktion erfolgt nicht in dem Sinne, daß jedesmal beim Herausschaffen eines äuseren Funktionszeichens sich eine endliche Ordnungszahl, die man einem

Funktional als Rang zuordnen kann, erniedrigt, sondern jedem Funktional entspricht gewissermaßen eine transfinite ordnungszahl als Rang, und der Satz, daß man nach Ausführung von endlich vielen Operationen ein konstantes Funktional auf ein Zahlzeichen erducziert hat, entspricht dem anderen, da's, wenn man von einer transfiniten Ordnungszahl zu immer kleineren Ordnungszahlen zurückgeht, man nach endlich vielen Schritten zur Null kommen muß. Nun ist natürlich bei unseren metamathematischen Überlegungen von tarnsfiniten mengen und Ordnungszahlen keine Rede. Es ist aber interessant, daß der erwähnte Satz über die transfiniten Ordnungszahlen sich in ein Gewand kleiden läßt, in dem ihm vom tarnsfiniten gar nichts mehr anhaftet." (Ackermann 1924b, pp. 13–14).

41. "Betrachten wir etwa eine transfinite Ordnungszahl, die vor $\omega \cdot \omega$ steht. Jede derartige Ordungszahl läst sich in der Form schreiben: $\omega \cdot \mathfrak{n} + \mathfrak{m}$, wo $\mathfrak{n}$ und $\mathfrak{m}$ endliche Zahlen sind. Man kann also eine derartige Ordnungszahl auch durch ein Paar endlicher Zahlen $(\mathfrak{n}, \mathfrak{m})$ characterisieren, wobei es natürlich auf die Reiehnfolge dieser Zahlen ankommt. Dem Zurückgehen in der Reihe der Ordnungszahlen entspricht folgende Operation mit dem Zahlenpaar $(\mathfrak{n}, \mathfrak{m})$. Entweder behalte ich die erste Zahl $\mathfrak{n}$ bei; dann setzte ich an Stelle von $\mathfrak{m}$ eine kleinere Zahl $\mathfrak{m}'$. Oder aber ich erniedrige die erste Zahl $\mathfrak{n}$; dann daf ich an diese zweite Stelle eine beliebige zahl setzten, die also größer sein kann als $\mathfrak{m}$. Es ist klar, daß man so nach endlich vielen Schritten zu dem Zahlenpaar $(0,0)$ kommen muß. Denn nach höchstens $\mathfrak{m}+1$ Schritten komme ich zu einem Zahlenpaar, bei dem die erste Zahl kleiner ist als $\mathfrak{n}$. Es sei dies $(\mathfrak{n}', \mathfrak{m}')$. Nach höchstens $\mathfrak{m}'+1$ Schritten komme ich dann zu einem Zahlenpaar, bei dem die erste Zahl wieder kleiner ist als $\mathfrak{n}'$, usw. Nach endlich vielen Schritten kommt man so zum Zahlenpaar $(0,0)$, das der ordnungszahl 0 entspricht. In dieser Form enthält der genannte Satz durchaus nicht transfinites; es werden nur solche Überlegungen benutzt, wie sie in der Metamathematik zuläsig sind. Analoges gilt, falls man nicht Paare endlicher zahlen, sondern Tripel, Quadrupel usw. benutzt. Dieser Gedanke wird nun nicht nur bei denfolgenden Beweisen dafür, daß man mit der Reduktion der Funktional zu Ende kommt, benutzt, sondern er wird auch später immer wieder angewandt, insbesondere bei dem Endlichkeitsbeweis am Schluß der Arbeit." (Ackermann 1924b, p. 14).

42. "[Gentzen fragt,] ob Sie der Meinung sind, dass sich die Methode des Endlichkeitsbeweises durch transfinite Induktion auf den Wf-Beweis Ihrer Dissertation anwenden lasse. Ich würde es sehr begrüssen, wenn das ginge." Bernays to Ackermann, November 27, 1936, Bernays Papers, ETH Zürich Library/WHS, Hs 975.100.

43. "Mir fällt übrigens jetzt, wo ich gerade meine Dissertation zur Hand nehme, auf, dass dort in ganz ähnlicher Weise mit transfiniten Ordnungszahlen operiert wird wie bei Gentzen." Ackermann

to Bernays, December 5, 1936, Bernays Papers, ETH Zürich Library/WHS, Hs 975.101.

44. “Ich weiss übrigens nicht, ob Ihnen bekannt ist (ich hatte das seiner Zeit nicht als Ueberschreitung des engeren finiten Standpunktes empfunden), dass in meiner Dissertation transfinite Schlüsse benutzt werden. (Vgl. z.B. die Bemerkungen letzter Abschnitt Seite 13 und im nächstfolgenden Abschnitt meiner Dissertation.” Ackermann to Bernays, June 29, 1938, Bernays papers, ETH-Zürich, Hs 975.114. The passage Ackermann refers to is the one quoted above.

45. The ε-substitution method was subsequently refined by von Neumann (1927) and Hilbert and Bernays (1939). Ackermann (1940) gives a consistency proof for first-order arithmetic, using ideas of Gentzen (1936). See also (Tait 1965) and (Mints 1994). Useful introductions to the ε-substitution method of Ackermann (1940) and to the ε-notation in general can be found in Moser (2000) and Leisenring (1969), respectively.

46. Ackermann (1924b, p. 8). The π-functions were already present (Hilbert 1922a) as the τ-function and also occur in (Hilbert and Bernays 1923a). They were dropped from later presentations.

47. It is not clear whether the definition is supposed to apply to the formulas with free variables (i.e., to $a = b$ and $a = \varepsilon_b(a = b)$ in the example) or to the corresponding substituion instances. The proof following the definition on p. 21 of (Ackermann 1924b) suggests the former, however, later in the procedure for defining a sequence of total substitutions it is suggested that the ε-expressions corresponding to formulas subordinate to $\tilde{\mathfrak{A}}(a)$ receive substitutions—but according to the definition of a total substitution only ε-*terms* ($\varepsilon_b(\mathfrak{z} = b)$ in the example) receive substitutions.

48. The bookkeeping functions are introduced here and are not used by Ackermann. The basic idea is that that in case (3), substitutions for some formulas are discarded, and the next substitution is given the “last” total substitution where the substitution for the formula was not yet marked as discarded. Instead of explicit bookkeeping, Ackermann uses the notion of a formula being “remembered” as having its value not discarded.

49. With the restriction on second-order ε-terms imposed by Ackermann, and discussed below, the system for which a consistency proof was claimed is essentially elementary analysis, a predicative system. A consistency proof using the ε-substitution method for this system was given by Mints and Tupailo (1996).

50. Ackermann (1924b, p. 9)

51. Ackermann to Bernays, June 25, 1925, Bernays Papers, ETH Zürich, Hs. 975.96.

52. “Ich habe augenblicklich den ε_f-Beweis wieder vorgenommen, und versuche mit aller Gewalt da zum Abschluß zu kommen. Daß sich das Problem auf ein zahlentheoretisches reduzieren

läßt, hatte ich Ihnen damals ja schon mitgeteilt. Den zahlentheoretischen Satz allgemein zu beweisen scheint mir aber ebenso schwierig wie das ganze Problem. Ich habe nun den schon mehrfach von mir versuchten Weg wieder eingeschlagen, den Begriff des Grundtyps so zu erweitern, das auch die ε mit freien Funktionsvariablen eine Ersetzung bekommen. Dieser Weg scheint ja auch der natürlichste, und die Gleichheitsaxiome $(f)(A(f) \rightleftharpoons B(f)) \to \varepsilon_f Af \equiv \varepsilon_f Bf$ würden dann gleich mitbehandelt. Ich habe einige Hoffnung, daß die sich früher auf diesem Weg einstellenden Schwierigkeiten vermieden werden können, wenn ich den ε_a-Formalismus benutze und statt ohne ε definierte Funktionen, solche zur Ersetzung für die ε_f nehme, die ein ε_a enthalten können. Ich habe mir aber erst einfache Spezialfälle überlegt." Ackermann to Bernays, March 31, 1926. ETH Zürich/WHS, Hs 975.97. Although Ackermann's mention of "ground types" precedes the publication of von Neumann (1927), the latter paper was submitted for publication alread on July 29, 1925.

53. "Letzthin habe ich mir Ihren neueren Beweis der Widerspruchsfr[eiheit] für die ε_a an Hand dessen, was Sie mir vor Ihrer Abreise aufschrieben, genauer überlegt und glaube diesen Beweis als richtig eingesehen zu haben." Bernays to Ackermann, April 12, 1927, in the posession of Hans Richard Ackermann. Bernays continues to remark on specifics of the proof, roughly, that when example substitutions for ε-types are revised (the situation corresponding to case (3) in Ackermann's original proof), the substitutions for types of higher rank have to be reset to the initial substitution. He gives an example that shows that if this is not done, the procedure does not terminate. He also suggests that it would be more elegant to treat all types of the same rank at the same time and gives an improved estimate for the number of steps necessary. Note that the reference to "ε_a's" (as opposed to ε_f) suggest that the proof was only for the first-order case. A brief sketch of the proof is also contained in a letter from Bernays to Weyl, dated January 5, 1928 (ETH Zürich/WHS, Hs. 91.10a).

54. "Wie Sie sich vielleicht erinnern, hatte ich damals einen 2. Beweis für die Widerspruchsfreiheit der ε_a. Dieser Beweis ist von mir nie publiziert worden, sondern nur Herrn Prof. Bernays mündlich mitgeteilt worden, der sich auch damals von seiner Richtigkeit überzeugte. Prof. Bernays schrieb mir nun im vergangenen Jahre, daß das Ergebnis ihm mit der Gödelschen Arbeit nicht zu harmonisieren scheine." Ackermann to Hilbert, August 23, 1933, Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek, Cod. Ms. Hilbert 1. Ackermann did not then locate the difficulty, and even a year and a half later (Ackermann to Bernays, December 8, 1934, ETH Zürich/WHS, Hs 975.98) suggested a way that a finitistic consistency proof of arithmetic could be found based on

work of Herbrand and Bernays's drafts for the second volume of *Grundlagen*.

55. "Problem I. The consistency proof of the ε-axiom for the function variable f. We have the outline of a proof. Ackermann has already carried it out to the extent that the only remaining task consists in the proof of an elementary finiteness theorem that is purely arithmetical." Hilbert (1928a), translated in Mancosu (1999a, p. 229). The extension to ε-extensionality is Problem III.

56. "Anlässlich der Arbeit für das Grundlagenbuch sah ich mich dazu angetrieben, den zweiten Hilbertschen Wf.-Beweis für das ε-Axiom, den sogenannten „verunglückten" Beweis, nochmals zu überlegen, und es scheint mir jetzt, dass dieser sich doch richtig stellen lässt." Bernays to Ackermann, October 16, 1929, in the possession of Hans Richard Ackermann. Bernays continues with a detailed exposition of the proof, but concludes that the proof probably cannot be extended to include induction, for which ε-substitution seems better suited.

57. The sketch bears the title "Wf.-Beweis für das logische Auswahl-Axiom", and is inserted in the front of *Elemente und Prinzipienfragen der Mathematik*, Sommer-Semester 1910. Library of the Mathematisches Institut, Universität Göttingen, 16.206t14. A note in Hilbert's hand says "Einlage in W.S. 1920." However, the ε-Axiom used is the more recent version $Ab \to A\varepsilon_a A_a$ and not the original, dual $A\varepsilon_a Aa \to Ab$. It is thus very likely that the sketch dates from after 1923.

58. "Ich glaube, dass damit die Frage, die wir bei der Durchsprechung des modifizierten Ackermannschen Beweises zuletzt diskutierten, ob nämlich eine Längen-Abschätzung für das Korrigier-Verfahren unabhängig von der Grösse der Zahlen-Substituenden gleichmässig möglich sei, verneinend beantwortet ist. An diesem Punkte ist dann der Nachweis des endlichen Abbrechens dieses Verfahrens (für den nächsten Grad, d.h. 3) jedenfalls lückenhaft." von Neumann to Bernays, March 10, 1931, Bernays Papers, ETH Zürich/WHS, Hs. 975.3328. Von Neumann's example can be found in Hilbert and Bernays (1939, p. 123).

59. Hilbert (1928a)

60. In a letter dated May 3, 1931, Bernays suggests that the problem lies with certain types of recursive definitions The Bernays–Gödel correspondence will shortly be published in Volume IV of Gödel's collected works. For more on the reception of Gödel's results by Bernays and von Neumann, see Dawson (1988) and Mancosu (1999b).

Chapter 4

Finitism and Mathematical Knowledge

4.1 Introduction

The years between the two World Wars was a remarkable time in philosophy and particularly in logic and the foundations of mathematics. It saw the rise of logical positivism in Vienna, Berlin and Prague; it was the time of Wittgenstein of the Tractatus, of the Polish school of logic, of Gödel, Ramsey, Skolem, and Turing. It was also the time of one of the grandest research projects in the philosophy of mathematics: Hilbert's Program. Hilbert's motivation for proposing the project was a range of criticisms brought by other mathematicians. They charged that a number of mathematical principles, such as impredicative definitions, the axiom of choice, and the law of excluded middle for infinite totalities, were contradictory, false, or at least were unfounded assumptions. His proposed solution, in short, was to prove the consistency of formalized mathematics on the basis of what he called "finitist" reasoning.

It is widely believed that Hilbert's project in its original formulation can't be made to work. The main reason for this widespread belief is that Gödel's second incompleteness theorem shows that a consistency proof of a formalization of a reasonably strong part of mathematics must use principles outside of that area. Hence, the principles required for a consistency proof of a non-finitist area of mathematics must themselves include non-finitist principles (as long as all finitistic reasoning can be formalized in the non-finitist theory). So it may seem that the study of Hilbert's program and the surrounding issues is at best of historical interest. One of the points I would like to make here will be that this is not so. I will try to make this point not by arguing directly that this or that part of Hilbert's philosophy of mathematics can be salvaged. Rather, I want to show how historically informed reflection on Hilbert's ideas, and more importantly, more recent *criticisms* of Hilbert's ideas, can contribute to parts of the current debate in the philosophy of mathematics.

The style of doing philosophy of mathematics of the 20s was supplanted in the 50s and 60s by a very different style: Platonism in mathematics became a dominant view, and the debate revolved mainly about the issue of the metaphysical status of mathematical objects. More recently, some of the classics of philosophy of mathematics have found their way back into the literature. For instance, logicism has been resurrected to some extent by the work of Boolos (1998, Part II) and Wright and Hale (2001), and both Tait (1986) and Maddy (1997) have appealed to Wittgenstein in the formulation of their positions. Hilbert's work, of course, is no exception. There are two aspects to his views of mathematics which I think are still highly relevant today: One is the strategy of providing support for an area of mathematics by providing a proof of consistency using epistemologically privileged principles. Such a strategy is still used in so-called revised Hilbert programs of providing proof-theoretic reductions of one area of mathematics to another, where such reductions are carried out by finitist means. Another aspect of Hilbert's work is the notion of finitism itself.

A discussion of the latter aspect raises both historical and conceptual questions. On the one hand, a comprehensive discussion of finitism requires an elucidation of both intent (What characterizes finitistic reasoning? What are the objects of finitistic reasoning?) and extent (What methods of reasoning can be justified by Hilbert as finitist? What methods were accepted as finitist?) of Hilbert had in mind when he spoke of the "finitist standpoint." On the other hand—if, as seems likely, Hilbert did not give a unique characterization of finitism—it is necessary to determine what the best way(s) to characterize finitism from a current standpoint are. In this vein, conceptual reconstructions of finitism have been attempted in more recent writings (in particular by Tait (1981, 2000), Parsons (1979–80, 1994), and Kreisel (1960, 1970)), and these reconstructions differ in both their starting points and their conclusions. For instance, Tait takes it that finitistic reasoning is characterized by being a minimal mode of reasoning underlying all mathematical thought about numbers, Parsons focuses on the notion of finitism as the domain of intuitive mathematical evidence, and Kreisel sees the surveyability of the domain of finitistic objects as characteristic. From these starting points, the analyses of these authors arrive at different extensional characterizations of what is finitistically provable.

I shall first take up the issue of the relevance of the notion of finitistic reasoning and evidence to current issues in the philosophy of mathematics. I shall then attack the historical question of what Hilbert (and his contemporaries) took finitism to consist in. Finally, I shall assess and compare the more recent reconstructions and criticisms of finitism of Parsons and Tait.

4.2 The Significance of Finitism

Finitism, as Hilbert understood it, is a methodological position that a philosopher of mathematics might take for a particular philosophical purpose. For Hilbert, the purpose was to show that formal mathematical reasoning is free from contradictions, and hence enjoys a status of security that had been put into doubt by his predicativist and intuitionist critics. But other purposes can be imagined: rather than attempting to ground all of mathematics in a formalistic fashion, we could imagine someone giving a reductivist account of mathematics, or a broadly instrumentalist or fictionalist account without reference to a particular formalism. As an example, consider Field's (1980) fictionalism: Field attempted to show that mathematics is conservative over a nominalistic formulation of physics. The proof of the conservativity result employed abstract—platonistic—methods. He was promptly criticized on this point. In response, he argued that he succeeded in showing that Platonism (motivated by the indispensability arguments) is led ad absurdum by his result. The proof is acceptable on platonistic grounds, and thus a platonist would have to concede, so Field, that mathematics is not indispensable after all. He could, however, just have bitten the bullet, argued that a version of finitism is nominalistically acceptable, and attempted to show that the conservativity result can be obtained on stricter grounds. (I am not suggesting that it is likely that such a stronger result can be obtained, but as an illustration of my point.)

The scope of the foundational project undertaken also need not necessarily be all of higher mathematics. So-called relativized Hilbert programs are projects of exactly this kind. 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. What is common to these approaches to mathematical foundations is that proof theoretic reductions are given of systems of classical mathematics to more restricted systems. The reduction is carried out using finitist means; and from this fact philosophical significance of these reduction is derived.

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 finitistic framework, such partial foundational reductions are therefore all justified finitistically. 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)

But even 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.[1]

4.3 Finitism in Hilbert

4.3.1 Numbers and Numerals

Hilbert conceived of his finitistic view of numbers in reaction and contrast to the logicist conception of number, which was supposed to yield a reduction of the number concept to broadly logical concepts. In Frege's view, numbers are extensions of certain concepts (For him, extensions of concepts, i.e., classes, are a logical notion). Russell took over this conception and tried to avoid the use of classes in *Principia* by use of his no-class theory.[2] Hilbert found this reduction of numbers to logical notions circular. In 1905, e.g., 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 1905c, p. 131).

Although around 1917 Hilbert was leaning towards Russell's viewpoint, he abandoned logicism a few years later. Bernays (1930, p. 243) made the disagreements between logicism and Hilbert's view very clear, and asserted that Frege's definition of cardinal numbers ("Numbers," *Anzahlen*) conceals the epistemologically essential characteristics of mathematics. Hilbert's finitistic viewpoint is not concerned with an analysis of the Number concept in the line of the Fregean project, but with a methodological program that presents an analysis of a certain minimal mode of numerical reasoning which is epistemologically grounded and serves to secure higher mathematics. Like Wittgenstein's proverbial ladder, the finitistic viewpoint can be abandoned once it provides this secure foundation.[3] Securing higher mathematics through consistency proofs does not and cannot presuppose a general analysis of Number. The Fregean analysis requires concepts and their extensions, entities which the finitist cannot consider; and Russell's solution was rejected because of the use of the axiom of reducibility. Hilbert attempts to account for numerical reasoning in terms of finite sequences, at first introduced as sequences of strokes. After 1923 Hilbert and Bernays are careful to distinguish between these "finitistic numbers" and the general concept of number (in either ordinal or cardinal sense), and usually refer to them as "numerals" (*Ziffern*).

Hilbert is interested in an account of elementary number theoretic reasoning which satisfies certain constraints of immediacy, intuitiveness, and certainty. These requirements of the methodological program translate into requirements on the subject matter of contentual finitistic arithmetic [*inhaltliche finite Arithmetik*]. According to Hilbert, the numerals are "concretely" given and surveyable by us. We have some immediate access to them which allows us to gain knowledge of finitary number-theoretic facts. The interesting question here is: What are the numerals exactly, and how do we have knowledge of them? It seems clear that Hilbert wants some sort of intuition to be the source of knowledge, so we might pose the question the other way around: What sort of intuition is the "primitive arithmetical intuition" and what do we have an intuition of when we engage in the "primitive mode of arithmetical thought"? Recent philosophy of mathematics would put the question in terms such as the following: Are the numerals physical objects? Mental constructions? Token or type? Abstract or concrete?

Some of the most fruitful sources on the topic of Hilbert's conception of finitism are his 1922 and 1926 papers, Bernays's exchange with Müller (Müller 1923, Bernays 1923), as well as the relevant sections in Hilbert and Bernays (1934, 1939). In 1905, Hilbert gives a first finitistic account of number theory in terms of strokes and equality signs. We note that *no identification of certain (sequences of) signs with numbers is made*, rather, the sequences of 1's and =’s are divided into two classes, the class of entities (these are the sequences of the form "$1 \ldots 1 = 1 \ldots 1$" with equal numbers of ‘’s on the left and right) and the class of nonentities; the former are the *true propositions*. Hence we have here a finitistic account, not of numbers, but of numerical truth.[4]

In his programmatic paper "The new grounding of mathematics," (1922c) we read:

> As we saw, the abstract operation with general concept-scopes and contents has proved to be inadequate and uncertain. Instead, as a precondition for the application of logical inference and for the activation of logical operations, something must already be given in representation [*in der Vorstellung*]: certain 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 presenta-

> tion, their difference, their succession (like the objects themselves) must exist for us immediately, intuitively, as something that cannot be reduced to something else. Because I take this standpoint, the objects [*Gegenstände*] of number theory are for me—in direct contrast to Dedekind and Frege—the signs themselves, whose shape [*Gestalt*] can be generally and certainly recognized by us—independently of space and time, of the special conditions of the production of the sign, and of insignificant differences in the finished product [footnote:In this sense, I call signs of the same shape "the same sign" for short.] The solid philosophical attitude that I think is required for the grounding of pure mathematics—as well as for all scientific thought, understanding, and communication—is this: *In the beginning was the sign.*[5]

Hilbert continues with an explicit account of numbers as signs built up from 1's and +'s. He writes:

> The sign 1 is a number.
>
> A sign which begins with 1 and ends with 1, and such that in between + always follows 1 and 1 always follows +, is likewise a number [. . .]
>
> These number-signs, which are numbers and which completely make up the numbers, are themselves the objects of our consideration, but have otherwise no *meaning* of any sort. (Hilbert 1922c, pp. 202–203)

Hilbert's account was criticized by the philosopher Aloys Müller, and some of the points he made were well taken. Following Müller's criticism, Hilbert (and Bernays) change the account slightly.

1. The term "sign" connotates having a meaning, but the number-signs are supposed to have no meaning attached to them. To avoid ambiguity, Hilbert and Bernays subsequently use the term 'numeral' [*Ziffer*] instead of "number-sign." For, I suppose, similar reasons, they also cease to use the word 'number' in this context, i.e., after Hilbert (1922c), no identification of numbers with numerals is made. This is in keeping with my remark above that Hilbert is not after an analysis of the number concept in general. Hilbert does, however, give an account of how numerals *function* as numbers in the sense of cardinal numbers, *Anzahlen*.[6]

2. The particular shape of the signs is immaterial. Bernays clarifies that what is important is that some objects of the same type are put together in a (finite) sequence.

> [...T]he special shapes "1" and "$1+1$" are inessential. If we disregarded the connection to habit, it would even be advisable, in order to emphasize the principle, to take as numerical signs figures of the type
>
> . . . . . .
>
> (which are thus constituted merely of points). And, of course, stars, vertical strokes, circles and other shapes could just as well be chosen instead of points. One could also take a time sequence, say, of similar noises, instead of a spatial sequence.
>
> But it is essential that *specimens of equal shape be joined in the same sort of arrangement* [*Zusammensetzung*]. (Bernays 1923, p. 224)

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 makes 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.[7] Hilbert must certainly hold that the number of strokes in a numeral is at least potentially infinite. 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.[8] Only our seeing or using "1111" as a sequence of 4 strokes as opposed to a sequence of 2 symbols of the form "11" makes "1111" into the numeral that it is. Would two stones lying side by side count as a numeral of the same kind as 11? If yes, then pretty much everything would be a numeral. If no, what decides whether something *is?* The obvious alternative would be that numerals are mental constructions. However, Bernays denies 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, p. 226). Kitcher considers this option as well. If the numerals were mental constructions, he writes,

> it seems that we shall have to accept many 3's (the array of three strokes I am currently contemplating, the array you are currently contemplating, the array I contemplated yesterday, etc.). So our discourse about numbers *simpliciter*

> should be replaced with talk about X's number n at time t. Arithmetical knowledge is immediately vulnerable to all kinds of scepticism. Perhaps 'My 2 at t plus my 2 at t equals my 4 at t' holds for some past t (not all, for I have not always been alive, nor always awake). But what of the future? What of your 2's to which I am forbidden access? And what is the status of arithmetic before anyone ever constructed a stroke-symbol?[9]

Kitcher's alternative is to hold that, whatever the numerals are, the strokes on paper or the stroke sequences I am contemplating represent these 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, p. 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, pp. 438–39) 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. Their *shape*[10] (but not they themselves) are 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 1922c, p. 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 usually considered to be abstract objects, however, 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 thought could be drawn from Hilbert's statements, and that was in turn disavowed by Bernays.[11] 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, pp. 225–26)

This is not to say that it is *incoherent* to consider the numbers as being abstract objects, only that the finitistic 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, p. 226)

Another open question in this regard is exactly what Hilbert meant by "concrete." He very likely does not use it 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.

The considerations outlined so far should have convinced the reader by now that the view is not as simple to make sense of as one might be inclined to think on a cursory reading of "On the infinite." On the one hand, for instance, the numerals are supposed to be objective, not merely created by thought, on the other hand they should not be independent of their intuitive representation. They need to be concrete and surveyable, but they also cannot be physical objects. The situation is not alleviated by the fact that it is not even clear what Hilbert means by “intuition” in this regard. Kitcher argues that it is sensuous intuition, and argues that this kind of intuition cannot meet all of Hilbert’s requirements. Mancosu (1998b) has shown that Hilbert and Bernays at first held the intuition involved to be empirical, but later in the 1920s turned to pure intuition as the source of certainty in elementary contentual arithmetic.

Many of the problems discussed so far arise because Hilbert considers contentual finite mathematics to be about certain entities, the numerals, and we are puzzled by their epistemological and ontological status. Not only that, but Niebergall and Schirn (1998) argue that assumptions of infinity are implicitly made by Hilbert’s finitism. If they are right, then the question of whether numerals are tokens or types is the least of Hilbert’s problems. Their argument depends in part on assuming that Hilbert’s contentual mathematics does have a standard referential semantics; that, say, “$2+2=4$” is true in virtue of the properties of the numerals that “2” and “4” refer to. What if we tried to make sense of finite mathematics without assuming standard semantics, without assuming that there are entities (the numerals “11” and “1111”) that “2” and “4” refer to and which make “$2+2=4$” true?

A promising avenue which has been suggested by Kitcher (1976) is that in order to accommodate all of the finitist’s requirements on numbers, the so-called standard account of mathematical truth must be abandoned. On the standard account, statements involving number terms are supposed to be analyzed in the same way as statements involving physical object terms are, i.e., the terms refer and the truth conditions of sentences in which they occur are given by Tarski’s semantics. Benacerraf (1973) argues that the virtue of the standard account is that it provides a “homogeneous semantical theory in which semantics for the propositions of mathematics parallel the semantics for the rest of the language.” Such a theory is part one of two requirements a theory of mathematical truth must fulfill, the other is that it “mesh with a reasonable epistemology.” Given the foundational character of the finitist viewpoint and its explicit (and only) goal, namely to give an account of truth for (a fragment of) arithmetic which is *secure*, it is reasonable to allow the second requirement to be the more important one.[12]

In many places Hilbert does seem to expound something like the standard account. For instance, in (Hilbert 1926, p. 377) he introduces first the numerals, which are supposed to "have no meaning at all by themselves." Contentual arithmetic, however, requires, besides the numerals, other signs "that mean something and serve to convey information, for example the sign 2 as an abbreviation for the numeral 11." There are also passages that suggest a non-standard reading, both have been mentioned before. In the early (1905c), the numerals are not introduced as an independent notion, but only in the context of identity statements. All we are given there are conditions of when an expression of the form '1...1 = 1...1' should count as a true proposition. In the response to Müller's criticism of Hilbert, Bernays introduces a distinction between the numerals [*Zahlzeichen*] and the notion of cardinal number [*Anzahl*]. He writes:

> It should also be noted that the contentual character of the *Number* [*Anzahl*] concept is indeed compatible with the purely figural character of the number signs. The figures are used as tools for *counting*, and by counting one arrives at the determination of cardinal number[...]
>
> One here has to recognize that the cardinal numbers are only defined in the context of the entire *Number statement.* For example, it will not be explained what "the Number five" is, but only what it means for the Number five to apply to a given totality of things. (Bernays 1923, p. 225)

We find two very interesting ideas here. Bernays suggests a non-standard reading of Number statements quite explicitly. His formulation is very close to Russell's "meaning in use" of integral sign or class abstraction in *Principia Mathematica.* The other idea is that the numerals are tools for counting. Hand (1989, 1990) has presented a non-standard account of finitistic number theory which takes up this idea. Since in this account truth conditions for finitistic statements take center stage, let me now turn to a discussion of statements.

4.3.2 Statements

We can find in Hilbert's writings three distinctions of mathematical (1) contentual vs. formal, (2) finitistic vs. infinitistic, and (3) real vs. ideal.[13] Contentual mathematics is comprised of those statements that we make based on our intuition of concrete objects.

This includes statements we make about numerals, but also those about formulas and formal derivations. These are concrete objects given in intuition, we can come to know about them directly. Formal discourse starts where we leave the grounds of intuition and instead proceed from an axiomatic standpoint. According to Hilbert, all mathematical discourse, "mathematics proper," [*die eigentliche Mathematik*] is to be formalized and thus becomes a "stock of provable formulas" Hilbert (1922c, p. 211). Whereas contentual mathematics deals with those statements and ways of reasoning that are justified on finite grounds, formalized mathematics appeals also to completed infinities, uses unbounded quantification, the unrestricted principle of the excluded middle. The exact nature of the "finitistic standpoint" is subject to debate and historical analysis. Hilbert and Bernays acknowledge that they have not drawn the distinction precisely:

> [W]e have introduced the expression 'finitistic' [*finit*] not as a sharply delineated term, but only as the name of methodical guideline, which enables us to recognize certain kinds of concept-formations and ways of reasoning as definitely finitistic and others as definitely not finitistic. This guideline, however, does not provide us with a precise demarcation between those [concept-formations and ways of reasoning] which accord with the requirements of the finitistic method and those that do not.[14]

We find, however, certain examples in Hilbert's and Bernays's writing of what falls within this standpoint. First there are the basic equalities and inequalities in the contentual sense between numerals; these are finitistically justified statements. We are also allowed to introduce certain computable functions, however, not as functions in the sense of abstract objects, i.e., sets of ordered pairs of numbers, but only as descriptions of methods of construction (see discussion in the next section). As a mode of reasoning, induction on propositions with "elementary intuitive content" is finitistically acceptable.[15] The question then arises to what extent general statements can be finitistically interpreted. The following quote from an unpublished manuscript by Bernays elaborates on the relevant passage in (Hilbert 1926, p. 378):

> We have to distinguish between such statements which express a *diagnosis* (a direct assertion) and those which express an insight, such as: "'$\mathfrak{a}+\mathfrak{b}$' is always the same numeral as '$\mathfrak{b}+\mathfrak{a}$'", or a claim of consistency.
>
> The second kind of statements are not obviously capable of negation, also they cannot be taken as the *hypothesis in a conditional statement*; rather, an

> *assumption* in finite reasoning can only ever refer to a *diagnosis* (like in a physical thought experiment).
>
> This implies that a proof of existence by refutation of a universal judgment cannot (obviously) be transferred into finite reasoning—while the proof of a universal judgment by refutation of an existential assumption can be transferred to the finite immediately by replacing the assumption that a certain thing exists by the assumption that the thing is *given*.[16]

A more detailed discussion of this question can be found in §2 of *Grundlagen* I:

> For the characterization of the finitistic standpoint we may emphasize some general considerations concerning the use of logical forms of judgment in finitist thought [*logische Urteilsformen im finiten Denken*], which we shall exemplify in the case of propositions about *numerals*.[17]
>
> A *universal* judgment about numerals can be interpreted finitistically only in a hypothetical sense, i.e., as a proposition about any given numeral. Such a judgment pronounces a law which must verify itself in each given particular case.
>
> An *existential sentence* about numerals, i.e., a sentence of the form "there is a numeral $\mathfrak{n}$ with the property $\mathfrak{A}(\mathfrak{n})$," is to be understood finitistically as a "partial judgment," i.e., as an incomplete communication of a more specific proposition consisting in either a direct exhibition of a numeral with the property $\mathfrak{A}(\mathfrak{n})$, or the exhibition of a process to obtain such a numeral,—where part of the exhibition of such a process is a determinate bound for the sequence of actions to be performed.
>
> Those judgments combining a universal proposition with an existential assertion have to be finitistically interpreted correspondingly. For instance, a sentence of the form "for each numeral $\mathfrak{k}$ with the property $\mathfrak{A}(\mathfrak{k})$ there is a numeral $\mathfrak{l}$, for which $\mathfrak{B}(\mathfrak{k}, \mathfrak{l})$ holds," has to be finitistically understood as an incomplete communication of a process which makes it possible to find, for each given numeral $\mathfrak{k}$ with the property $\mathfrak{A}(\mathfrak{k})$, a numeral $\mathfrak{l}$ which stands in the relation $\mathfrak{B}(\mathfrak{k}, \mathfrak{l})$ to $\mathfrak{k}$.[18]

It is clear that the negation of a finitistic statement need not be a finitistic statement, more precisely: A finitistic interpretation for a contentual statement about numerals does not, by itself, yield a finitistic interpretation of the negation of the statement. This is what it means for a finitist statement to be "incapable of being negated:" From the inability, or even impossibility, to see that $\mathfrak{A}(\mathfrak{k})$ for each given numeral $\mathfrak{k}$, it does not follow that we have a witness $\mathfrak{l}$, or even a bound on such, for $\neg\mathfrak{A}(\mathfrak{l})$. At this point, Hilbert and Bernays propose the formalization of mathematics and the introduction of ideal elements, i.e., unbounded quantifiers, to retain the simplicity of classical logic. In formalized mathematics, we have real statements (roughly, these are quantifier-free formulas) which admit of a contentual, finitistic interpretation, and ideal statements which round out the theory, but appeal or presuppose infinite totalities, and thus do not have a finite, contentual interpretation.

Now how is the semantics of the basic finitistic statements explicated? Hand (1990) discusses an *iterativistic* tendency in Hilbert's views on the issue. The basic idea is that we have the capacity to count intransitively, i.e., to count without counting any thing (in particular, not the natural numbers). The numerals are, so to speak, a crutch for us to remember when we are supposed to stop counting. An equality between numerals is then, e.g., to be understood as the assertion that in counting the strokes in both strings simultaneously, we will stop at the same point (in contrast to the view that identity is "figural correspondence"); an inequality the statement that we can count further on one numeral than on another (in contrast to the idea that one numeral extends beyond the other).

This view has a certain appeal. After all, as a matter of developmental and historical fact, the concept of number had its origins in the human capacity to *count*. We count fingers, we count cattle, eventually, we just count. The phenomenon of counting without counting any thing, it seems to me, is much better explained as engaging in an iterative mental procedure than as naming, in succession, the members of the natural number sequence (whatever those are). Such a conception, argues Hand (1990), is at work in some places in Hilbert's finitism. He supports his view mainly by textual evidence from Hilbert's (1905c) and (1926). I believe that there are even stronger suggestions in this direction in *Grundlagen*.

For instance, in "On the infinite," the numerals are introduced in a very pictorial way, as sequences of strokes. A few years later, when even Hilbert is becoming convinced that empirical intuition will not do the trick of providing arithmetic with a secure and philosophically acceptable footing, Bernays writes:

> The ordinal number is in and of itself also not determined as object; it is only a place marker [*Stellenzeiger*]. We can objectively standardize it by choosing as a *place marker the simplest structure from those that originate in the form of the succession.* . . . [W]e have an inital thing and a process; the objects are then the initial thing itself and further the objects one obtains,

> beginning with the intial thing, through a single or repeated application of that process. (Bernays 1930, p. 244)

In *Grundlagen* the idea is made more concrete:

> In number theory, we have an initial object and a process of succession [*Prozeß des Fortschreitens*]. For both we must settle on a particular intuitive representation. The particular kind of representation is inessential, but the choice, once made, must be retained throughout the whole of the theory. We choose as initial object the numeral 1 and as process of succession the attachment of 1.
>
> The objects which we obtain from the numeral 1 by applying the process of succession, such as
>
> $$1, 11, 1111,$$
>
> are figures of the following kind: The start with 1, they end in 1; each 1 which does not already form the end of the figure is followed by an attached 1. They are obtained through application of the process of succession, that is by an *assembling* which concretely comes to an end, and this assembling can be undone by a stepwise *disassembling*.[19]

The idea we are given here is that the object of number theory is an iterative process, and this iterative process is representable in intuition. This iterative process is basic, the numerals only help us, so to speak, to keep track of how far the iteration has proceeded. This is necessary since a crucial aspect of the iterative process is that it can be *reversed*. This is the basis for induction and recursion in finitistic mathematics.

The import the iterativist account has for semantics is that it no longer requires denotations for the terms of arithmetic. The numerals do not stand for anything, they just give us information on how to verify or falsify arithmetical statements. Equalities between numerals are true, not if the numerals are the same or they have the same form, but if they both give the same bound on the process of succession. If we can start counting and both numerals tell us to stop at the same time, they are equal.

4.3.3 Finitistic Functions

The issue is a little more involved when considering more complex arithmetical terms, e.g., sums and products. If numbers are controls on iteration, then how do we get the functions of arithmetic? Well, by iteration of course. A finitistic function, or more precisely, its definition, is a way of communicating a certain procedure which allows us to obtain from certain numerals another numeral. The operations that this procedure may appeal to are assembling and disassembling numerals and iterating (previously defined) operations. This is all that is needed to give truth conditions for equalities between arithmetical terms: Use the definition of the function to obtain an explicit bound on iteration and we have reduced the question to an equality between numerals.

How much of this is present in Hilbert's writings? Hilbert nowhere discusses a general notion of finitistic function explicitly. Discussions of particular functions in the context of the finitistic standpoint are limited to addition and multiplication. In (1922c), addition is not introduced as an operation on numerals, but as part of the definition of numeral: We are told that numerals are sequences of the form 1, $1+1$, $1+1+1$, etc. In (1926), the '+' sign serves to communicate concatenation of numerals. Concatenation is, however, not thought of as a directly as an operation: $111+11 = 11111$ is not intended as the statement that the *operation* $+$ applied to the numerals 111 and 11 yields 11111, but rather that 111 juxtaposed with 11 is the same numeral as 11111. It is not until Bernays's (1930) that addition and multiplication on numerals are explicitly conceived of as *operations* on finitistic objects. However, from about 1922 onwards, primitive recursion is part of the formalized mathematical systems that Hilbert and his students develop and investigate, and consistency proofs for these systems appeal to recursive procedures (e.g., to reduce terms occurring in a proof and containing function symbols for primitive recursive functions to numerals; see Chapter 2). Even in *Grundlagen*, the initial explanation of addition is still much dependent on the visual image of numerals as sequences of strokes, but addition is understood as an operation of concatenation.

> If a numeral $\mathfrak{b}$ corresponds to a part of $\mathfrak{a}$, then the rest is again a numeral $\mathfrak{c}$; thus we obtain the numeral $\mathfrak{a}$ by appending $\mathfrak{c}$ to $\mathfrak{b}$, in the manner in which the 1 which starts $\mathfrak{c}$ is appended to the 1 in which $\mathfrak{b}$ ends according to the process of succession. We call this kind of composition of numerals *addition* and use for it the sign +.[20]

However, the definition of multiplication agrees much more with the idea of definition of a function by iteration. A first discussion of this can be found in (Bernays 1930), where exponentiation is also treated:

> Let us consider, for example, the number $10^{10^{1000}}$. We can arrive at this number in a finitistic way as follows: We start from the number 10 which we represent, according to one of our earlier standardizations, by the figure
>
> 1 1 1 1 1 1 1 1 1 1
>
> Let now z be any number that is represented by a corresponding figure. If in the previous figure we replace every 1 by the figure z, then we obtain again, as we can intuitively make clear to ourselves, a number figure that for the purpose of communication is denoted by "$10 \times z$". In this manner we obtain the process of a decuplication of a number. From this we arrive at the process of transition from a to 10^a, as follows. We let the number 10 correspond to the first 1 in a and to every affixed a we apply the process of decuplication, and we keep going until we exhaust the figure a. The number obtained by means of the last process of decuplication is denoted by 10^a. (Bernays 1930, p. 249)

In *Grundlagen*, Bernays gives a similar characterization of multiplication in general:

> *Multiplication* can be defined as follows: $\mathfrak{a} \cdot \mathfrak{b}$ denotes the numeral obtained from $\mathfrak{b}$ by replacing, during its construction, always the 1 by the numeral $\mathfrak{a}$, so that one first forms $\mathfrak{a}$ and then appends $\mathfrak{a}$ instead of each appending of 1 in the construction of $\mathfrak{b}$.[21]

It is clear, however, that the general notion of finitistic function is based on iteration/recursion, and Hilbert could as well have given an explanation of addition and multiplication in these terms, without appeal to the picture of sequences of strokes and their geometrical manipulation. This is evident from the discussion of recursion:

> [O]ne point still requires a basic discussion, the method of *recursive definition*. Let us see what this method consists in: A new function symbol, say φ, is introduced, and the [corresponding] function is defined by two equations. In the simplest case, these equations are of the form:
>
> $$\begin{aligned} \varphi(1) &= \mathfrak{a} \\ \varphi(\mathfrak{n}+1) &= \psi(\varphi(\mathfrak{n}), \mathfrak{n}). \end{aligned}$$
>
> Here, $\mathfrak{a}$ is a numeral and ψ a function which is formed from previously known functions by composition, so that $\psi(\mathfrak{b}, \mathfrak{c})$ can be computed for given numerals $\mathfrak{b}$, $\mathfrak{c}$ and gives another numeral as value. [...]
>
> It is not immediately clear, which sense may be assigned to this method of definition. For its elucidation we must first make the notion of function precise.

> A *function*, for us, is an intuitive instruction on the basis of which to each given numeral another numeral is assigned. A pair of equations of the above kind—called a "*recursion*"—is to be understood as an *abbreviated communication* of the following instruction:
>
> Let $\mathfrak{m}$ be any numeral. If $\mathfrak{m} = 1$, so let the numeral $\mathfrak{a}$ be assigned to $\mathfrak{m}$. Otherwise, $\mathfrak{m}$ has the form $\mathfrak{b}+1$. One then writes down schematically:
>
> $$\psi(\varphi(\mathfrak{b}), \mathfrak{b}).$$
>
> Now if $\mathfrak{b} = 1$, so one replaces therein $\varphi(\mathfrak{b})$ by $\mathfrak{a}$; otherwise $\mathfrak{b}$ again has the form $\mathfrak{c}+1$, and one then replaces $\varphi(\mathfrak{b})$ by
>
> $$\psi(\varphi(\mathfrak{c}), \mathfrak{c}).$$
>
> Again, either $\mathfrak{c} = 1$ or $\mathfrak{c}$ is of the form $\mathfrak{d}+1$. In the former case one replaces $\varphi(\mathfrak{c})$ by $\mathfrak{a}$, in the latter case by
>
> $$\psi(\varphi(\mathfrak{d}), \mathfrak{d}).$$
>
> Repeating this process in any case terminates. For the numerals
>
> $$\mathfrak{b}, \mathfrak{c}, \mathfrak{d}, \ldots,$$
>
> which we obtain one after the other, develop through the *disassembling of the numeral* $\mathfrak{m}$, and this must terminate just like the assembling of $\mathfrak{m}$ does. When we arrive at 1 in this process of disassembling, then $\varphi(1)$ is replaced by $\mathfrak{a}$; the sign φ does then no longer occur in the resulting figure. Rather, the only function symbol occurring, possibly in multiple superposition, is ψ and the innermost arguments are numerals. Thus we have arrived at a computable expression; for ψ was supposed to be a function already known. This computation must be executed from the inside out, and the numeral thus obtained shall be assigned to the numeral $\mathfrak{m}$.[22]

We see how this conception of computation fits in with the iterativist conception of finitistic truth: An equation between arithmetical terms is evaluated, by iteration, until we arrive at an equation between numerals. The numerals themselves are mnemonic devices, in principle dispensable, for effecting this procedure. This corresponds precisely to Hand's "canonical verifications."

4.3.4 Finitism and PRA

Let me now make some remarks on Tait's claim that the finitistically acceptable functions are exactly the primitive recursive ones. There is no question that the primitive recursive functions are finitistic, since they are all given by recursive definitions of the above kind. But are they *all* the finitistic functions? It seems to me that they cannot be. For any description of an iterative procedure that allows us, given a term involving the function symbols introduced, to arrive at a numeral as "value" of the term, using only iteration and substitution (in particular, no unbounded search), should count as finitistic, if the primitive recursive ones do. Such a function is, e.g., the Ackermann function, which is well known not to be primitive recursive. I believe there is ample evidence in *Grundlagen* that Hilbert considered it to be finitistic.

Tait did consider the issue of the Ackermann function being finitistic, as part of an objection to Kreisel's (1960) characterization of finitistic functions as the provably total functions of first-order Peano arithmetic. The issue is that Kreisel points out[23] that in "On the infinite," Hilbert explicitly discusses the Ackermann function. Tait's argument against the conclusion, based on this fact, that we should regard it as finitistic is that Hilbert introduces the Ackermann function in the context of a theory of functions of higher types, and these higher types are certainly not finitistic. It is true that in Hilbert (1926) introduces the Ackermann function by recursion on higher types, but it can also be introduced by nested recursion on two arguments. If it is correct that all that is needed for a function to be finitistic is that it be given by a process of recursion which allows the computation of the "value" by successive rewriting of numerals, then this definition would certainly make the Ackermann function finitistic. In fact, there is rather explicit evidence that the Ackermann function was considered to be finitistic, in *Grundlagen*. In §7 of the first volume, Hilbert and Bernays discuss (what is now called) primitive recursive arithmetic. This is a formal theory, all of whose statements, however, are *real*, i.e., finitistically meaningful:

> This recursive number theory is close to intuitive number theory as considered in §2, as all of its formulas *admit of a finitistic contentual interpretation.* This contentual interpretability is a result of the verifiability of all derivable formulas of recursive number theory, [a fact] which we have already stated. Indeed, in this area verifiability has the character of a direct contentual interpretation, and it was because of this that the proof of consistency was so easy to give here.
>
> The difference of recursive number theory vis a vis intuitive number theory consists in its formal restrictions; its only method of concept-formation, aside from explicit definition, is the schema of recursion, and also the methods of deduction are strictly limited.

> We may, however, admit certain *extensions of the schema of recursion* as well as of the induction schema, without taking away what is characteristic of the method of recursive number theory. We shall now discuss these briefly.[24]

Hilbert and Bernays then go on to introduce course-of-values recursion, simultaneous recursion (which can both be reduced to primitive recursion), nested recursion [*verschränkte rekursion*], the Ackermann function, and nested induction, and prove that the Ackermann function is not primitive recursive. The most conclusive statements, however, we find in the second volume:

> Certain methods of *finitistic mathematics which go beyond recursive number theory* (in the original sense) have been discussed in §7 [of volume I of *Grundlagen*], namely the introduction of functions by nested recursion [e.g., Ackermann's function] and the more general induction schemata.[25]

A few pages later, we read:

> The original narrow concept of a finitistic proposition amounts in the field of number theory to admitting as finitistic number-theoretic propositions only such propositions which can be expressed in the formalism of [primitive] recursive number theory, possibly *including symbols for certain computable number-theoretic function (of one or more arguments)*, but without use of formula variables, or which admit a stricter interpretation through a formula of such a form.[26]

Elsewhere, we read that "contentual finite arithmetic [is formalized by] recursive number theory."[27] This remark is made in the context of arithmetization of syntax, and I take its force to be that the methods used in arithmetization are primitive recursive and hence finitistic, rather than making a programmatic identification of finitistic with primitive recursive. In a footnote, the reader is referred to the passage from p. 325 of the first volume quoted above, where the relationship between contentual arithmetic and recursive arithmetic is discussed. That passage suggests that Hilbert and Bernays considered finite arithmetic as *partially* but not necessarily completely formalized by primitive recursive arithmetic. It also suggests that the *verifiability* of formulas is the criterion of finitistic meaningfulness. Verifiability here is defined as follows: Every true equality or inequality between numerals is verifiable. Every boolean combination of verifiable formulas is verifiable. A formula containing free individual variables (but no formula variables or bound variables) is verifiable if every instance resulting by substituting numerals for the free variables is verifiable (1939, pp. 229, 238; 1970, pp. 228, 237). A closed formula containing primitive recursive function symbols is verifiable if the formula resulting from calculating the primitive recursive terms occurring in it is verifiable (1939, 1970, p. 297). If the function can be calculated, it is finitistic? It is obvious that some restriction must be placed on the notion of calculability involved here. Without restrictions, every total general recursive function would be finitistic, and any formula containing symbols for total recursive functions would be verifiable in that sense. The restriction would most likely have to do with being able to see that the calculation process comes to an end, and this is precisely the issue in the question of whether the Ackermann function should be considered finitistic. Bernays was aware that there is a substantial difference between primitive and nested recursion in this respect, and the issue comes up when he proves that primitive recursion can be replaced by the μ-operator (Hilbert and Bernays 1934, pp. 421–22; 1968, pp. 430–431).[28] Much later, he took nested recursions (in the sense of *verschränkte Rekursionen* considered in *Grundlagen* I) to be finitist on the grounds that they could be computed by a sequence of replacements of terms, the number of which is bounded. In a letter to Gödel from 1970, he writes:

> These nested recursions [. . .] appear to me to be finite in the same sense as the primitive recursions, i.e., if one regards them as a statement of a computation procedure where one can recognize that the function defined by the respective process satisfies the recursion equations (for every system of numeral values [*Ziffernwerte*] of the arguments). Indeed, the computation of the value of a function according to a nested recursion, when the numeral values of the arguments are given, comes down to the application of several primitive recursions, the number of which is determined by a numeral argument [*Ziffernargument*].[29]

It is consistent with Hilbert's early writings that finitism, as originally conceived in the early 1920s, does not surpass primitive recursive methods. In all likelihood, Hilbert and Bernays did not think they had to address the issue explicitly. The Ackermann function had not been discovered when the finitistic standpoint was first formulated, and in any case it was probably thought initially that primitive recursive methods suffice for metamathematics. Tait (2000) suggests that the remarks in *Grundlagen* have to be read in light of Gödel's incompleteness theorems—Hilbert mentions the necessity to "exploit the finitist standpoint in a sharper form" in the preface to *Grundlagen* I—and that therefore it is unclear whether any discussion of finitism by Hilbert or Bernays after 1931 does not already include an extension of the original view.[30] As I have discussed in Chapter 3, however, even though finitistic principles of definitions of functions were hardly discussed before *Grundlagen*, non-primitive recursive methods were used, and clearly accepted as finitistic, in the consistency proofs given long before Gödel's results, specifically, in Ackermann's 1924 dissertation. In summary, the finitistic standpoint as conceived and applied by Hilbert, Bernays, Ackermann, and others, goes beyond reasoning in primitive recursive arithmetic.

4.4 The Status of Finitism

Since the proof-theoretic reductions sought are finitistic, the significance of the project of foundational reductions thus rests, in part at least, on an account of the status of finitism. Presumably, the restricted methods the finitist has access to should be epistemologically privileged. I will argue that there are two distinct parts to Hilbert's characterization of the finite standpoint. One is, so to speak, from the bottom up. Adopting a Kantian framework, he takes it as a given that our faculty of intuition provides us with the objects of elementary intuitive arithmetic, the stroke symbols. Knowledge of these objects is immediate and unproblematic, since no concepts are involved. Now it is one thing to claim that the objects of finite "contentual" arithmetic are intuitively given to us, and quite another to claim that on the basis of our access to these stroke symbols we can gain knowledge of the rather complex sort of propositions like the consistency statement for Peano arithmetic. So even if we leave aside reservations about the involvement of Kant's notion of intuition or more general concerns about whether and how access to stroke symbols is possible in a way that has bearing on mathematical knowledge, there is still work to be done to explain how intuitive knowledge of stroke symbols extends to knowledge of the interesting foundational claims like the consistency statement. In other words, anyone who wants to invoke the finitist standpoint for foundational purposes will have to give a theory of finitist proof:

what the acceptable methods of proof are, and how they transfer the epistemological status of the intuitive knowledge of stroke symbols and their basic relationships to knowledge of less basic but finitistically provable propositions.

Another way to characterize the finitist viewpoint is from the top down, as that area of mathematical reasoning which is basic to all exercise of mathematical thought. Such a consideration shines through at points in Hilbert's writings. In the oft-cited passage where Hilbert speaks of the "extra-logical concrete objects that exist intuitively as immediate experience before all thought", namely the objects of contentual finite arithmetic, he calls their existence a "condition for the use of logical inferences and and the performance of logical operations" (Hilbert 1926, p. 376). In "The Grounding of Elementary Number Theory" (Hilbert 1931), he calls the finitistic standpoint the "fundamental mode of thought that [he] hold[s] to be necessary for mathematics and for all scientific thought, understanding, and communication, and without which mental activity is not possible at all." It is this consideration that is central to W. W. Tait's analysis and reconstruction of finitism (Tait 1981).

So we have here two fundamentally different approaches to thinking about finitism, or any part of mathematics for which epistemological priority of some sort is claimed: One taking its starting point in a notion of intuition and intuitive knowledge, where the main question is: how far can this intuitive knowledge take us? The other approach proceeds in the reverse direction: Rather than appealing to a notion of intuition which serves as the foundation of finitist knowledge, it is argued that finitist reasoning forms a core of mathematical principles, a Cartesian foundation immune to sceptical doubt about the certainty of mathematics. Both are views taken in characterizations and criticisms of finitism in the literature: the first by Parsons (1998a), the second by Tait (1981).

I should point out that I have nothing to say about *radical* scepticism about mathematics of the sort held by so-called ultra-finitists. Just as in epistemology there is an important difference between scepticism about sense experience and scepticism about what Descartes called "clear and distinct ideas," so it makes sense here to distinguish between taking certain notions of mathematics (e.g., impredicative definitions, strong set existence assumptions, or the existence of the continuum) as prima facie problematic while leaving others to the side (e.g., the existence of a successor to every natural number.)

4.5 Parsons' Criticism of Finitism as Intuitive Knowledge

In a recent paper, Charles Parsons (1998a) attempts an evaluation of the epistemological status of finitist reasoning. He does this in the framework of his own theory of mathematical intuition, which he has elaborated in a series of papers (1979–80, 1998b, 2000). I will not concern myself here with this notion of intuition, although I agree with Parsons that it shares essential features with Hilbert's finitist intuition of stroke symbols. Indeed, Hilbert's writings were a major influence on Parsons. But the intuition both Hilbert and Parsons stress is an intuition *of* objects, while what is at stake here is intuition *that*, or more generally, propositional intuitive knowledge. In order for any broadly finitist project to get off the ground, the finitist has to show that the truths required for carrying out the relevant foundational work can be known, and that this knowledge has epistemic features which make it suitable to provide the foundational work with the philosophical import it is claimed to have. In the context of finitism, this comes down to what Parsons calls Hilbert's Thesis:

> A proof of a proposition according to the finitary method yields intuitive knowledge of that proposition. In particular, this is true of proofs in primitive recursive arithmetic.

I could not find anything close to an explicit formulation of this thesis in Hilbert's writings, and Parsons does not offer any, either. There is indirect support offered for the thesis as a correct interpretation of what Hilbert thought: that Hilbert appealed explicitly to Kant, that such a notion of intuition *that* can be found in Kant, and that Bernays speaks of finitism as the "domain of intuitive evidence." Aside from the matter of Hilbert-exegesis, however, we should ask: is the thesis a requirement for the finitist's goal? But first, we must understand what intuitive propositional knowledge is.

Parsons writes:

> Intuitive knowledge, as I understand it, is knowledge that is based on intuition in an appropriate way. In the case of singular propositional knowledge, it will typically be based on intuition of the objects the proposition is about.

This applies straightforwardly to Hilbert's most basic examples, those of finitary propositions of the form of equalities and inequalities between stroke symbols and sums of stroke symbols. If we follow Parsons' analogy of mathematical intuition with perception, then one would be inclined to agree that a proposition such as that expressed by the equation

$$||| + || = |||||$$

can be known in the relevant sense. So let us suppose that Parsons is right in that a specifically mathematical kind of intuition can provide access to the stroke symbols here in a manner analogous to the access provided by ordinary sense perception to the physical tokens representing them printed on this page. There is are three difficulties for finitist knowledge still to be surpassed:

1. The question raised by practical limitations to the intuitability of strings of unbounded complexity,

2. the question of the status of propositions involving recursively defined functions, and

3. the question of knowledge of general—as opposed to singular—propositions.

I will leave (1) to the side for the time being. It is a question that has to be dealt with when explaining the notion of intuition itself, and I hardly have anything to add to Parsons' discussion on the matter. In fact, it comes down to the "radical scepticism" mentioned—and excluded from discussion—above. (2) and (3), on the other hand, have to be addressed independently. For on the notion of intuition under consideration, it is hard to see how such propositions can be directly grounded in intuition. You may "intuit" that three strokes concatenated with two strokes make five strokes, but it is harder to see how intuition *of* can directly provide knowledge of a finitistic procedure producing a certain result, or of a general claim being true. At some point, finitistic proof has to play a role. For example, the consistency claims Hilbert was considering took something of the following form: for any intuitively given object of a certain kind, viz., a formal derivation, a finite procedure yields a finitistic object of another kind, viz., a derivation which can be seen to have something other than "$0 = 1$" as its last formula. If every derivation could be directly intuited to have this property, there would be no need for proof.

Note also that this goes farther than the issue of a mere practical impossibility for the notion of intuition to produce the required result. At best, I can concede that the conclusion of a bounded induction can be directly intuited, by having intuitions, together or successively, of the individual steps in the induction. But I can see no way to intuit a general finitistic statement in the same way. To do so would require to have an intuition of all infinitely many cases at the same time, and that is hardly within the scope of a notion of intuition that claims to be finitist.[31]

A simple example will illustrate the point: To see that $2 \cdot 3 = 3 \cdot 2$, I can intuit the result of replacing each stroke in the stroke symbol || with ||| on the one hand, and also each stroke in ||| by the entire figure ||, and then see, intuitively, that the results are the same. But to have intuitive knowledge of the general statement $x \cdot y = y \cdot x$ seems to require an infinity of such insights.

This brings us back to the beginning, and to the question what sense we can give to the phrase "intuitive knowledge" as it is used in the context of general finitistic propositions. This is a kind of propositional knowledge, of knowledge *that*. What would make such knowledge qualify as intuitive, specifically, what would qualify it as finitist? Surely it must be finitistically stated, and dealing only with finitistic objects. For instance, we cannot have intuitive knowledge of infinite sets, and we cannot have intuitive knowledge of unrestricted existence claims. But more must be true: it must be arrived at by finitistically acceptable means—by finitistic proof. Parsons takes this to mean that there is a property—"intuitive evidence"—which the basic finitistic propositions have. They have this property because intuition, that is, intuition of the stroke symbols, immediately gives it to us. In order for the finitist to be able to claim intuitive evidence for all finitistically provable propositions he would have to argue that the finitistically acceptable methods of inference preserve intuitive evidence. However, Parsons does not tell us what intuitive evidence is in general, and neither does Hilbert. The strategy that Parsons uses to argue for and against various parts of Hilbert's Thesis, however, suggests the following:

> A proposition is intuitively evident if it is arrived at either directly by intuition or by reasoning which is minimal for any reasoning about intuitively given objects.

In arguing for the sub-thesis that propositional inference preserves intuitive evidence, for instance, he writes:

> [I]t should be clear that what is assumed, the basic logic of identity and minimal propositional logic, is minimal for reasoning about objects.

And for the case of induction he continues:

> We can't make for induction the claim we made for logical reasoning [...], that it is basic for reasoning in general, or for reasoning about objects. It has rather the character of being basic to reasoning about objects in a particular domain, the objects obtained from an initial element by arbitrary finite iteration of a given operation. [...] If we admit some conception of the domain of strings as intuitive, it seems we ought to admit induction as preserving intuitive knowledge. (Parsons 1998a, pp. 258–259)

Parsons goes further in his assessment of primitive recursion. Primitive recursion differs from logical inference and induction in that it is not a principle of reasoning, rather, it is a principle of definition. In short, it is the principle that allows the finitist to introduce new procedures operating on strings by stipulating what the result of the procedure is for an initial string, and explains how to reduce the application of the procedure to a string to an application to its predecessor. In formal terms, we may introduce a functor by explaining what the functor applied to 1 is, and how to reduce an application to $x+1$ to an application to x. The definiens, in each case, can only make reference to functors already introduced.

The question here, again, is whether defining functors by primitive recursion and using functors so defined in finitist proofs preserves intuitive evidence. Parsons, who is committed to a notion of intuitive evidence resting on that of intuition *of*, viz., intuition of stroke symbols, sees a need to provide an answer to this question. And here we are in a similar situation as in that of induction above: At best, it seems, we can have an intuition of the sequence of strings needed to compute the result of a particular primitive recursion. For example, consider exponentiation. The exponentiation function can be recursively defined from multiplication. To see that 2^3 is well defined, we could fall back on the intuitions we have of 2, $2 \cdot 2$, and $2 \cdot 2 \cdot 2$. But to show that in general 2^x is well defined, we can give no singular intuition of a finite sequence of strokes which would provide us with immediate insight that this is the case. We would have to give a proof using induction, but either the statement on which we are inducing is not finitist—it would contain an unbounded quantifier—or we would have to use a non-finitist object, the exponentiation procedure itself. Here we would be assuming already that induction preserves intuitive evidence. And that, as we have seen above, cannot be made intuitively evident other than by arguing that induction is implicit in finitist intuition itself.

In Parsons' discussion of recursion the strain between the bottom-up and the top-down approach becomes apparent. While in the case of logical inference and induction, Parsons is content to fall back on a top-down approach, in the case of recursion he demands a non-circular argument to the effect that exponentiation is well-defined. He claims that such an argument cannot be given, for the following reason. Parikh's Theorem (1971) states that the exponentiation function is not provably recursive in formal theories of arithmetic where induction is restricted to formulas without unbounded quantifiers, and the functions allowed are bounded by polynomials.[32] An independent argument that a function defined by primitive recursion is well-defined, and that its defining equations are intuitively known, should be convertible into a proof in that theory that the defined function is total. Parsons concludes from this that it is not possible to give a finitist proof that exponentiation is well defined using only addition, multiplication, and induction. The cogency of this argument rests on the assumptions that a difference in the strength of two formal theories (in this case, polynomially bounded arithmetic $I\Delta_0$ and stronger theories such as elementary function arithmetic or primitive recursive arithmetic—some theory that proves that exponentiation is total) implies a difference in the intuitive evidence we can have of the theorems of each.

Leaving aside the question of whether or under what conditions a formal theory correctly formalizes a body of informal discourse, it seems natural to grant that if a theory establishes (directly or indirectly, e.g., by explicit definition or coding) a theorem of a prima facie stronger theory, the epistemic strength of the (prima facie) weaker theory extends to the stronger theory. In other words, it would be sufficient to establish the intuitive character of exponentiation if the totality of the exponentiation function (using a suitable translation) were provable in a theory which itself is directly justified by intuition. The question at hand, and which is the issue in the assumption, is whether it is also necessary. This seems to me to require additional argument. It is certainly not the case without a separate assumption. For instance, formal arithmetic without multiplication or induction is weaker than a theory resulting by adding multiplication and its defining axioms. Both theories, however, are justified by intuition in Parsons' sense.

It is not at all obvious to me why addition and multiplication should be the only functions which can be independently seen to be finitistically acceptable functions. It seems quite conceivable that there are “easily” computable functions which can be independently seen to be well-defined, but which also cannot be shown to be total in $I\Delta_0$. Here is a possible example: Parsons conjectures that a finitistic proof that exponentiation is well defined should be formalizable in $I\Delta_0$. The kinds of proofs (e.g., that addition and multiplication are intuitive) proceed by arguing about strings, and replacement of strings by others. $I\Delta_0$ in fact allows coding of strings, and so it would be natural to suppose that to formalize an intuitive proof of the well-definedness of a function, the coding mechanisms available in $I\Delta_0$ would be used. However, $I\Delta_0$ cannot prove that in a string a symbol can be substituted by another string Krajíček (1995, p. 66). But substitution of strings seems to be intuitively well-defined—it is the kind of operation that Parsons uses to motivate that multiplication is well defined.

Parsons suggests that feasible computability plays a role in considerations of whether a given function being well-defined is intuitive. For instance, he argues that polynomial-time computability is a necessary feature of intuitive functions (and that hence exponentiation is not intuitive), and his remarks at the end of Section 6 of his paper suggest that he would be inclined to accept the polynomial-time computable functions as intuitive (provided that they are closed under bounded recursion, an open question). On the other hand, he thinks it is necessary and sufficient for a function to be intuitive that it is Δ_0-definable and provably total in $I\Delta_0$. But those functions are just the linear-time computable ones,[33] and it would be surprising if it turned out that the classes of polynomial-time and of linear-time computable functions coincide. If they do not, then there are plenty more feasible functions which will not be Δ_0-definable and provably total in $I\Delta_0$.

All of this suggests that Parsons’ argument for why no independent justification for the intuitiveness of exponentiation can be given falls short. Even if it were conceded that exponentiation does require independent justification, that leaves the possibility open that there might be principles which, in concert with addition, multiplication, and induction provide such an intuitive justification. A look at the literature on bounded arithmetic shows that such is often the case: slight variations in the formulation of the language or the principles allowed may cause significant differences in the strength of the theories.

Bernays has given finitist accounts of computing the multiplication and the exponentiation functions. Roughly, to multiply 10 by 10, we should imagine a string of 10 strokes, and then replace each stroke by the stroke sequence for 10 itself. Parsons takes this as more or less unproblematic. Similarly, Bernays suggests, we can compute the exponentiation function: to intuitively construct 10^{10}, imagine two sequences of 10 strokes. For each stroke in the first, apply the procedure—which we've just seen can be intuitively constructed—of multiplying by 10 to the second sequence. When we are done, we will have constructed a sequence of 10^{10} strokes. Parsons is not satisfied that this provides an intuitive construction. Let me take a closer look.

Take multiplication again. We are given two stroke sequences, x and y, and want to intuitively construct their product. Bernays explains that this is done by replacing each stroke in x by a copy of x. Now recall that x and y are determined as stroke sequences: their sole and determining characteristic is that they are formed from 1 by appending strokes. So the replacement of the strokes in x has to proceed one-by-one if we are to keep strictly to what is intuitive about x and y. In other words, the procedure is: for the first stroke in x, construct a copy of y, and for each additional stroke in x, append a copy of y to what has been constructed up to that point. In the end we get x concatenated copies of y. But now the difference to the case of exponentiation is only in the kind of operation that has to be performed for each stroke in x: addition of y in one case, multiplication by y in the other. But I don't see what could be special about multiplication by y that makes iteration of it problematic that would not make concatenation of y, or concatenation of the previously constructed string plus y just as problematic. But these procedures define functions which are perfectly acceptable by Parsons' criteria: $x \cdot y$ and $y \cdot (1 + \cdots + x)$. On the other hand, it might be the idea of iteration of a procedure itself is what Parsons thinks causes problems for exponentiation. But if so, I don't see how multiplication can survive them, either.

So if there is anything problematic about the evidence that primitive recursion is well defined, it is the notion of finite iteration itself. But in any specific instance of a primitive recursive calculation, the argument serving as the control on the iteration is intuitively given, and as such is finite. So in any calculation of a primitive recursively defined function with intuitively given arguments, there can be no question that the calculation is finite. The recursive definition itself provides a procedure for obtaining the intuitive construction of the result through all intermediate steps. The fact that a recursively defined function in general is well defined, i.e., that its computation terminates for each stroke string as argument and not just for a particular given string, does never directly enter into finitistic proofs. At most, it does so as a presupposition. But that presupposition is of a kind with the presupposition that every stroke symbol can be extended by appending another stroke. In a similar, although more complex fashion, the *n*-fold iteration of a functor can be extended by adding another functor in front. The supposition that the arbitrary finite iteration of the functor is well defined presupposes only that the functor itself is well defined. But this is not a problem of circularity, only of bootstrapping. If we have intuitive access to the basic functions, as Parsons agrees, we can work our way up through all the primitive recursive functions one at a time. Of course, the proposition that all primitive recursive functions are well defined is not finitist; but it is also not used nor needed by the finitist.

4.6 Tait's Analysis of Finitism

At the outset I outlined my interest in considering the epistemic status of finitism. Finitism has been used to provide reductions of one area of mathematics to another, where the notion of reduction involved is a technical, proof-theoretical notion rather than the usual one of ontological reduction. The epistemic value of such reductions rests in large part on the assumption that finitism is in a sense epistemically privileged. I have discussed one sense in which this assumption can and has been supported. Following Hilbert's writings, Parsons has sketched an account of mathematical intuition which aims to have the features Hilbert claimed for his notion of intuition in the formulations of the finitist standpoint. Parsons is doubtful that such a notion of intuition can serve to imbue all finitist propositions provable by primitive recursive methods with intuitive evidence. I have tried to argue that Parsons doubts are unfounded. If they can indeed be put aside, however, the question still remains whether Parsons' notion of intuition itself can stand up to critical analysis. I have not attempted to provide such an analysis, however, even if Parsons' notion of intuition ultimately cannot be sustained, there is still another way to argue for the special status of finitism; the way I called "top-down" above.

W. W. Tait (1981) proceeded in exactly this way in his explication of finitism. In contrast to Parsons, he does not consider intuitive evidence to be the distinguishing mark of finitism. Rather, he holds that it embodies the principles implicit in any kind of mathematical reasoning about numbers. To be precise: he argues that primitive recursive arithmetic embodies these principles, infers that finitism must at least include primitive recursive arithmetics, and finally offers support for the thesis that nothing exceeding primitive recursion has claim to the title "finitism." As we have seen in the discussion of Parsons, the critical questions of the admissibility of induction, and, I have argued, of iteration and with it primitive recursion, ultimately have to be answered by arguing that they are implicit in the notion of finite sequence.

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. He 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, p. 525)

In terms of the significance of finitism for the projects outlined in Section 4.2, a defense of finitism being fundamental in this sense seems to me to be sufficient. For if there is any mathematical knowledge at all, anything that is fundamental to number-theoretical mathematics in the sense alluded to by Tait would enjoy whatever status is conferred on the items of mathematical knowledge by one's preferred epistemology, either directly by inclusion under those items or indirectly as a condition of mathematical knowledge. The additional thesis that finitism is epistemologically privileged only requires the plausible thesis that some areas of mathematics are more secure, or better justified than others. That thesis is certainly widely held (e.g., by all those who find that it makes sense to distinguish between the foundational frameworks mentioned in Section 4.2) and certainly is supported both by every-day and properly mathematical practice.

The difficulty that Tait faces here is that it is unclear whether finitistic reasoning, or in any case primitive recursive reasoning, is indeed "a minimal kind of reasoning presupposed by all nontrivial mathematical reasoning about numbers." The issue here is both with "presupposed" and "nontrivial." Some areas of mathematics do not rely on the entirety of primitive recursive methods; at least some nontrivial results can be obtained without using the full force primitive recursive reasoning. But one may presuppose certain constructions and principles without actually using them. So even if there are nontrivial results which never appeal to the totality of all numbers, or to the general principles of primitive recursive definitions, the existence of the numbers and the validity of the principles they do appeal to are justified only by the existence of the entire number sequence and the validity of primitive recursive methods in general. This is plausible enough for the first part: a mathematician engaged in such a limited enterprise would surely answer affirmatively when asked: "So you only use and prove things about numbers less than 10^{10}, but 10^{10} is still a number, right?" One might be inclined to simply deny that she is working with numbers in the usual sense if she were to answer in the negative, but rather on elements of a certain finite structure which she denotes by numerals. It is less plausible for the second part: definition schemas can be given for classes of functions which are significantly smaller than the class of primitive recursive functions, and yet allow formulation and proof of nontrivial results. The classes of elementary functions E and the polynomial-time computable functions P are examples of prominent such classes. Formal systems are known which characterize these classes in the same way that PRA characterizes the primitive recursive functions, and these formal systems allow the derivation of nontrivial theorems.

The analogy that Tait draws between his and Church's thesis might suggest that to complain that since it is possible to precisely circumscribe proper subclasses of finitistic functions and theorems as an argument against the thesis is no less fallacious than it would be to complain that general recursiveness is not a characterization of computability because there are well-circumscribed sub-recursive classes of functions. But whereas the computable functions are the maximal class of functions that are, well, computable; the finitistic functions (and finitistic reasoning) answer to a minimality condition: they are the minimal class whose existence is presupposed by the general notion of Number. And so to defuse the objection, it should be specified what it is to be presupposed in that notion, either directly or indirectly by closure conditions. This is exactly what Tait does: For Tait the finitistic functions are those that come with, as it were, the basic finitist type of Number. These functions are the constant 0 function and the successor function. The finitist functions are the class of functions including these basic functions, closed under composition and primitive recursion. Clearly, the weak point here is that it is not beyond doubt that primitive recursion is a construction implicit in the notion of Number, and that hence the finitist functions must be closed under primitive recursion. It is made plausible by the particular elegance and simplicity of the scheme, and the fact that primitive recursion, when it is viewed as n-fold iteration of composition, corresponds exactly to the iteration of the successor which is the defining characteristic of Number. Tait writes,

> Suppose that we have $k{:}B$ and $g{:}B \to B$ and an arbitrary $n{:}N$. n is built up from 0 by iterating $m \mapsto m'$:
>
> $$0, 1, 2, \ldots, n = 0''^{\cdots\prime}$$
>
> By using exactly the same iteration, the same $'' \ldots '$, we may build up $fn{:}B$:
>
> $$k, gk, ggk, \ldots, fn = g \ldots gk$$
>
> We may express this construction by means of the equations
>
> $$f0 \equiv k \qquad fn' \equiv g(fn)$$
>
> and we say that $f{:}N \to B$ is defined from k and g by *primitive recursion.* (Tait 1981, pp. 531–32)

So to make it plausible that a nontrivial and yet less inclusive class of functions is a rival candidate for the title “finitist functions”, it would have to be shown that the closure conditions, i.e., the ways of constructing new functions from initial functions, are presupposed by the notion of Number in a similarly well-motivated way. This seems to be not the case for the examples suggested above, the classes of P and E of polynomial-time and elementary functions. In the case of P, the machine-independent characterization is given by the system PV of Cobham (1965). In PV, the basic type is not Number as constructed by iterated application of the successor function, but by iterated applications of two successor function, essentially resulting in binary representations. Correspondingly, the recursion schema used is (bounded) *recursion on notation.* If the two successors are denoted s_0, s_1, then recursion on notation is the following principle of definition: If functions g, h_0, and h_1 are already defined, then a function f is defined by

$$\begin{aligned} f(0) &= g \\ f(s_0x) &= h_0(f(x)) \\ f(s_1x) &= h_1(f(x)). \end{aligned}$$

So while recursion on notation may well be said to be a construction implicit in the notion of binary notations, it isn't implicit in the same obvious way in Number. Furthermore for *bounded* recursion on notation, for f to be defined it has to satisfy in addition that $fx \leq \ell x$ for all x (ℓ a function already defined). This means that we have only succeeded in defining a function f if f satisfies the bound (in PV, a new function symbol f is introduced only if PV proves that the function defined by the three equations is bounded). But this can hardly be seen to be a restriction implicit in the notion of binary representation. Indeed in the case of finitist functions a circle threatens: Finitist functions were introduced to explain finitist proof, but if the definition of a finitist function requires a (finitist) proof that the function defined satisfies the bound, then finitist function would presuppose and not explain finitist proof. In the case of elementary functions, the basic type is the same as in the case of primitive recursion, but the construction by which we obtain new functions is bounded addition and multiplication. That addition and multiplication are definable using the constructions is a result in need of proof, not something that obviously flows from the notion of Number; the boundedness requirement introduces the same difficulties as above.

What this shows, I think, is that Tait's claim that finitist reasoning is "a minimal kind of reasoning presupposed by all nontrivial mathematical reasoning about numbers" is not a definition, but a consequence of the analysis; the analysis presupposes that any class delimited as the finitist functions (and by extension, the finitist principles of reasoning) must not only be "minimal" and "nontrivial", but also closed under all constructions implicit in the notion of Number. Only this rules out the rival classes which are closed under weaker constructions than primitive recursion. This, however, opens up the possibility that the finitist could come to see more than just primitive recursion as being implicit in Number. We have already seen, in Section 4.3.4, that Hilbert, Bernays, and Ackermann accepted more than primitive recursive functions and principles as finitist. In particular, 2-fold nested recursion, sufficient to define the Ackermann function, was accepted.

The schema of k-fold nested recursion is this:

$$
\begin{aligned}
f(x,0,y_2,\ldots,y_k) &= 0 \\
&\ddots \\
f(x,y_1,\ldots,y_{k-1},0) &= 0 \\
f(x,y_1',\ldots,y_k') &= g(x,y_1,\ldots y_k,t_1,\ldots,t_k), \text{ where} \\
t_i &= f(x,y_1',\ldots,y_{i-1}',y_i,h_1^i(x,y_1,\ldots,y_k,f(x,y_1',\ldots,y_{k-1}',y_k),\ldots \\
&\qquad f_{k-i}^i(x,y_1,\ldots,y_k,f(x,y_1',\ldots,y_{k-1}',y_k))
\end{aligned}
$$

For $k = 1$, this is just primitive recursion. If the reason that we must accept primitive recursion as finitistic is that the computation of a function proceeds in "exactly the same way" as the construction of a numeral, then this is in a way also true of functions defined by k-fold nested recursion. For to compute such a function, it is also only necessary to pass from the computation required for n' to the computation for n, although now we have to keep track of several recursion arguments (but only a fixed number k).

In (2000), Tait addresses an objection of Ignjatovič (1994), that "one cannot rule out the possibility that any basis sufficient to justify what is formalized in [PRA] and which satisfies some necessary closure properties in order to be acceptable as an epistemologically distinguished system of methods, is also sufficient to justify ε_0-induction." Tait's response is that to go beyond PRA would require the introduction of higher types, and the restriction to finitist types is warranted as an "epistemologically distinguished system of methods." It is, however, hard to see in what way k-fold nested recursion introduces higher types; it certainly doesn't mention type variables. Yet the k-fold recursive functions are exactly the ω^{ω^k}-recursive ones.[34] This leads to the conclusion that, at the very least, there is not a single notion of finitist function, but a hierarchy, corresponding to the level of complexity of iterations one is prepared to accept as implicit in the simple iteration required for an understanding on Number.[35]

4.7 Conclusion

There are two fundamentally different kinds of sceptical doubts about number-theoretic reasoning: The doubt that there might not be a potential infinity of numbers, that the successor operation is not total; and the doubt that there are certain inferences or constructions which are unwarranted even once the successor operation (the general notion of Number) is accepted. The finitist in Hilbert's sense is not concerned with the first kind, and it is the second kind that motivates the discussions of acceptable inferences and constructions in the work of Hilbert, Parsons, and Tait. Several different aspects of Hilbert's characterization of finitism have been analyzed in the literature: that finitism is the domain of intuitive evidence by Parsons, that finitism is a minimal kind of reasoning implicit in, or presupposed by, all reasoning about Number by Tait. These two concerns correspond to two different views about what the appropriate way to characterize finitism is: the bottom-up approach, arguing for the existence of a special kind of intuition which applies to finitist objects and argue that this intuition underwrites a notion of intuitive evidence which encompasses all finitistically provable propositions, and the top-down approach, arguing that finitistic principles are basic to mathematical reasoning, that finitism thus constitutes a core of mathematics, and characterizing finitism as that which must be included in that core. The bottom-up approach receives its epistemic status from the notion of intuition; in the top-down approach the epistemic primacy of finitism results from the picture that different parts of mathematics have different epistemic status, and that finitism, as the most basic area of mathematical reasoning about number, is thus also the epistemically most secure.

Notes

1. Feferman (1993a). See also Hofweber (2000) for a discussion of the relation of proof-theoretic reductions to other kinds of reductions, e.g., ontological and theory reduction.

2. For a discussion of this effort, see Chihara (1973, Ch. 1).

3. This is methodological point is made clear in a letter from Bernays to Rosza Péter, probably from around 1940 (Bernays Papers, ETH Library/WHS, Hs. 975:3473).

4. See Hilbert (1905c, p. 131f). In a course at Göttingen, Hilbert went even further in the

development of this idea, see Peckhaus (1990), Chapter 3.

5. Hilbert (1922c, p. 202), repeated almost verbatim in Hilbert (1926, p. 376). This is the text of a talk given in Hamburg, July 25–27, 1921.

6. This account is based on our ability to put finite collections of objects into one-to-one correspondences with the strokes making up a numeral. This ability accounts for the usefulness of contentual number theory. The account is indicated in passing by Bernays (1923, p. 225) and developed in detail by Hilbert and Bernays (1934, pp. 28–29).

7. In particular in Hilbert (1926).

8. "[The numbers do not] exist independently of their *intuitive construction*." (Bernays 1923, p. 226).

9. Kitcher (1976, pp. 107–108). Frege (1884/1980, §27) advanced essentially the same criticism against Schloemilch.

10. "Figures [i.e., numerals] *are* not shapes, they *have* a shape" (Bernays 1923, p. 159)

11. "These objects must be [...] space- and timeless [...]" (Müller 1923, p. 158)

12. Benacerraf and Putnam (1983) also finds a non-standard account in Hilbert's view of mathematics. That account, however, does not concern the contentual mathematics we are interested in, but formalized mathematics. According to Benacerraf, Hilbert's account of formalized mathematics is non-standard since unbounded quantifiers—since they are finitistically meaningless—are not evaluated according to standard semantics, but based on the derivability of sentences containing them from axioms systems that have been shown to be consistent.

13. For a discussion of these distinctions, see Sinaceur (1993).

14. "Es stellt sich nun die Frage ein, ob denn überhaupt die finiten Methoden imstande sind, den bereich der in (Z_μ) formalisierbaren Schlußweisen zu überschreiten.

Diese Frage ist freilich, so wie sie eben formuliert ist, nicht präzise; denn wir haben den Ausdruck „finit" ja nicht als einen scharf abgegrenzten Terminus eingeführt, sondern nur als bezeichnung einer methodischen Richtlinie, die uns zwar ermöglicht, gewisse Arten der Begriffsbildung und des Schließens mit Bestimmtheit als finit, gewisse andere mit Bestimmtheit als nicht finit zu erkennen, die aber dennoch keine genaue Scheidelinie liefert zwischen solchen, die ihnen nicht mehr genügen." Hilbert and Bernays (1939, pp. 347–48; 1970, p. 361).

15. *Aussagen mit elementar anschaulichem Inhalt*. I propose to read this as: propositions which permit a finitistic interpretation (see below). See Bernays (1922, p. 221) for the distinction between the form of induction discussed here, "the narrower form of induction," and the full schema of in-

duction on arbitrary formulas. This distinction is essential for the rebuttal, by Hilbert, of Poincaré's and Becker's charge of circularity in Hilbert's theory. For this, see Mancosu (1998b).

16. "Wir haben bei den Aussagen zu unterscheiden zwischen solchen, die einen *Befund* (eine direkte Feststellung) ausdrücken und solchen die eine Einsicht zum Ausdruck bringen, wie: „$\mathfrak{a}+\mathfrak{b}$" ist stehts dasselbe Zahlzeichen wie $\mathfrak{b}+\mathfrak{a}$", oder z.B. die Behauptung einer Widerspruchsfreiheit.

Die Aussagen der zweiten Art sind nicht ohne weiteres negationsfähig; auch können sie nicht in die *Voaussetzung eine Bedingungssatzes* genommen werden; vielmehr kann sich im finiten Schliessen eine *Annahme* immer nur auf einen *Befund* beziehen (entsprechend wie bei einem physikalischen Gedanken-Experiment).

Hieraus ergibt sich, dass ein Beweis einer Existenz durch Widerlegung eines allgemeinen Urteils sich nicht (ohne weiteres) ins finite Schliessen übertragen lässt—während der Beweis eines allgemeinen Satzes durch Widerlegung einer Existenz-Annahme sofort ins Finite zu übertragen ist, indem man die Annahme dass ein gewisses Ding existiere, durch die Annahme ersetzt, dass das Ding *vorgelegt* ist." Paul Bernays, "Zur finiten Einstellung," Manuscript, 1 p., no date but written after 1925. Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Cod. Ms. Hilbert 685:9.

17. I take the word "exemplify" to imply that the same forms of judgment also apply to other finitistically acceptable concept-formations, e.g., functions defined by recursion.

18. "Zur Characterisierung des finiten Standpunktes seien noch einige allgemeine Gesichtspunkte hervorgehoben, betreffend den Gebrauch der logischen Urteilsformen im finiten Denken, wobei wir zur Exemplifizierung Assagen über *Ziffern* betrachten wollen.

Ein *allgemeines* Urteil über Ziffern kann finit nur im hypothetischen Sinn gedeutet werden, d. h. als eine Aussage über jedwede vorgelegte Ziffer. Ein solches Urteil spricht ein Gesetz aus, das sich an jedem vorliegenden Einzelfall verifizieren muß.

Ein *Existenzsatz* über Ziffern, also ein Satz von der Form „es gibt eine Ziffer $\mathfrak{n}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{n})$", ist finit aufzufassen als ein „Partialurteil", d. h. als eine unvollständige Mitteilung einer genauer bestimmten Aussage, welche entweder in der direkten Angabe einer Ziffer von der Eigenschaft $\mathfrak{A}(\mathfrak{n})$ oder der Angabe eines Verfahrens zur Gewinnung einer solchen Ziffer besteht,—wobei zur Angabe eines Verfahrens gehört, daß für die Reihe der auszuführenden Handlungen eine bestimmte Grenze aufgewiesen wird.

In entsprechender Weise sind diejenigen Urteile finit zu interpretieren, in denen eine allgemeine Aussage mit einer Existenzbehauptung verknüpft ist. So hat man z. B. einen Satz von der Form „zu jeder Ziffer $\mathfrak{k}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{k})$ gibt es eine Ziffer $\mathfrak{l}$, für welche $\mathfrak{B}(\mathfrak{k},\mathfrak{l})$ gilt“, finit aufzufassen als unvollständige Mitteilung von einem Verfahren, welches gestattet, zu jeder vorgelegten Ziffer $\mathfrak{k}$ von der Eigenschaft $\mathfrak{A}(\mathfrak{k})$ eine Ziffer $\mathfrak{l}$ zu finden, welche zu $\mathfrak{k}$ in der Beziehung $\mathfrak{B}(\mathfrak{k},\mathfrak{l})$ steht.“ (Hilbert and Bernays 1934, 1968, pp. 32–33).

19. “In der Zahlentheorie haben wir ein Ausgangsobjekt und einen Prozeß des Fortschreitens. Beides müssen wir in bestimmter Weise anschaulich festlegen. Die besondere Art der Festlegung ist dabei unwesentlich, nur muß die einmal getroffene Wahl für die ganze Theorie beibehalten werden. Wir wählen als Ausgangsding die Ziffer 1 und als Prozeß des Fortschreitens das Anhängen von 1.

Die Dinge, die wir, ausgehend von der Ziffer 1, durch Anwendung des Fortschreitungsprozesses erhalten, wie z. B.

$$1, 11, 1111,$$

sind Figuren von folgender Art: sie beginnen mit 1, sie enden mit 1; auf jede 1, die nicht schon das Ende der Figur bildet, folgt ine angehängte 1. Sie werden durch Anwendung des Fortschreitungsprozesses, also durch einen konkret zum Abschluß kommenden *Aufbau* erhalten, und dieser Aufbau läßt sich daher auch durch einen schrittweisen *Abbau* rückgängig machen.” (Hilbert and Bernays 1934, 1968, p. 20–21).

20. “Wenn eine Ziffer $\mathfrak{b}$ mit einem Teilstück von $\mathfrak{a}$ übereinstimmt, so ist das Reststück wiederum eine Ziffer $\mathfrak{c}$; man erhält also die Ziffer $\mathfrak{a}$, indem man $\mathfrak{c}$ an $\mathfrak{b}$ ansetzt, in der Weise, daß die 1, mit welcher $\mathfrak{c}$ beginnt, an die 1, mit welcher $\mathfrak{b}$ endigt, nach der Art des Fortschreitungsprozesses angehängt wird. Diese Art der Zusammensetzung von Ziffern bezeichnen wir als *Addition* und wenden dafür das Zeichen $+$ an.” (Hilbert and Bernays 1934, 1968, p. 22).

21. “Die *Multiplikation* kann folgendermaßen definiert werden: $\mathfrak{a}\cdot\mathfrak{b}$ bedeutet die Ziffer, die man aus der Ziffer $\mathfrak{b}$ erhält, indem man beim Aufbau immer die 1 durch die Ziffer $\mathfrak{a}$ ersetzt, so daß man also zunächst $\mathfrak{a}$ bildet und anstatt jedes in der Bildung von $\mathfrak{b}$ vorkommenden Anfügens von 1 das Ansetzen von $\mathfrak{b}$ ausführt.” (Hilbert and Bernays 1934, 1968, p. 24).

22. “[N]ur noch ein Punkt bedarf hier der grundsätzlichen Erörterung, das Verfahren der *rekursiven Definition*. Vergegenwärtigen wir uns, worin dieses Verfahren besteht: Ein neues Funktionszeichen, etwa φ wird eingeführt, und die Definition der Funktion geschieht durch zwei Gleichungen, welche im einfachsten Falle die Form haben:

$$\begin{aligned} \varphi(1) &= \mathfrak{a} \\ \varphi(\mathfrak{n}+1) &= \psi(\varphi(\mathfrak{n}),\mathfrak{n}). \end{aligned}$$

Hierbei ist $\mathfrak{a}$ eine Ziffer und ψ eine Funktion, die aus bereits bekannten Funktionen durch Zusammensetzung gebildet ist, so daß $\psi(\mathfrak{b},\mathfrak{c})$ für gegebene Ziffern $\mathfrak{b}$, $\mathfrak{c}$ berechnet werden kann und als Wert wieder eine Ziffer liefert.

Es ist nicht ohne weiteres klar, welcher Sinn diesem Definitionsverfahren zukommt. Zur Erklärung ist zunächst der Funktionsbegriff zu präzisieren. Unter einer *Funktion* verstehen wir hier eine anschauliche Anweisung, auf grund deren einer vorgelegten Ziffer, bzw. einem Paar, einem Tripel,... von Ziffern, wieder eine Ziffer zugeordnet wird. Ein Gleichungspaar der obigen Art—wir nennen ein solches eine „*Rekursion*"—haben wir anzusehen als eine *abgekürzte Mitteilung* folgender Anweisung:

Es sei $\mathfrak{m}$ irgendeine Ziffer. Wenn $\mathfrak{m} = 1$ ist, so werde $\mathfrak{m}$ die Ziffer $\mathfrak{a}$ zugeordnet. Andernfalls hat $\mathfrak{m}$ die Form $\mathfrak{b}+1$. Man schreibe dann zunächst schematisch auf:

$$\psi(\varphi(\mathfrak{b}),\mathfrak{b}).$$

Ist nun $\mathfrak{b} = 1$ so ersetze man hierin $\varphi(\mathfrak{b})$ durch $\mathfrak{a}$; andernfalls hat wieder $\mathfrak{b}$ die Form $\mathfrak{c}+1$, und man ersetze dann $\varphi(\mathfrak{b})$ durch

$$\psi(\varphi(\mathfrak{c}),\mathfrak{c}).$$

Nun ist wieder entweder $\mathfrak{c} = 1$ oder $\mathfrak{c}$ von der Form $\mathfrak{d}+1$. Im ersten Fall ersetze man $\varphi(\mathfrak{c})$ durch $\mathfrak{a}$, im zweiten Fall durch

$$\psi(\varphi(\mathfrak{d}),\mathfrak{d}).$$

Die Fortsetzung dieses Verfahrens führt jedenfalls zu einem Abschluß. Denn die Ziffern

$$\mathfrak{b},\mathfrak{c},\mathfrak{d},\ldots,$$

welche wir der Reihe nach erhalten, entstehen durch den *Abbau der Ziffer* $\mathfrak{m}$, und dieser muß ebenso wie der Aufbau von $\mathfrak{m}$ zum Abschluß gelangen. Wenn wir beim Abbau bis zu 1 gekommen sind, dann wird $\varphi(1)$ durch $\mathfrak{a}$ ersetzt; das Zeichen φ kommt dann in der entstehenden Figur nicht mehr vor, vielmehr tritt als Funktionszeichen nur ψ, eventuell in mehrmaliger Überlagerung, auf, und die innersten Argumente sind Ziffern. Damit sind wir zu einem berechenbaren Ausdruck gelangt; denn ψ soll ja eine bereits bekannte Funktion sein. Diese Berechnung hat man nun von innen her auszuführen, und die dadurch gewonnene Ziffer soll der Ziffer $\mathfrak{m}$ zugeordnet werden." (Hilbert and Bernays 1934, 1968, pp. 25–26).

23. "For instance, Tait refers to (Hilbert 1926) as a source concerning Hilbert's notion of a finitist proof, goes on to say 'it is difficult perhaps to determine what Hilbert really had in mind' and

argues that Ackermann's enumeration of the primitive recursive functions is not finitist. But whatever else may be in doubt, Hilbert's own notion as *used in* (1926) certainly includes Ackermann's function since it is explicitly mentioned!" (Kreisel 1970, p. 514, n. 43)

24. "Diese rekursive Zahlentheorie steht insoferne der anschaulichen Zalentheorie, wie wir si im §2 betrachtet haben, nahe, als ihre Formeln sämtlich *einer inhaltlichen Deutung fähig* sind. Diese inhaltliche Deutbarkeit ergibt sich aus der bereits festgestellten Verifizierbarkeit aller ableitbaren Formeln der rekursiven Zahlentheorie. In der Tat hat in diesem Gebiet die Verifizierbarkeit den Character einer direkten inhaltlichen Interpretation, und der Nachweis der Widerspruchsfreiheit war daher auch hier so leicht zu erbringen.

Der Unterschied der rekursiven Zahlentheorie gegenüber der anschaulichen Zahlentheorie besteht in ihrer formalen Gebundenheit; sie hat als einzige Methode der Begriffsbildung, außer der expliziten Definition, das Rekursionsschema zur Verfügung, und auch die Methoden der Ableitung sind fest umgrenzt.

Allerdings können wir, ohne der rekursiven Zahlentheorie das Charakteristische ihrer Methode zu nehmen, gewisse *Erweiterungen des Schemas der Rekursion* sowie auch des Induktionsschemas zulassen. Auf diese wollen wir noch kurz zu sprechen kommen." (Hilbert and Bernays 1934, p. 325; 1968, p. 330). I am reading this passage so that "what is characteristic of the method of recursive number theory" to be the availability of a direct contentual, i.e., finitistic interpretation. Tait (2000), by contrast, takes this phrase to refer to the kinds of definitions and rules of inference available in it, so that the scope of "certain extensions of the schema of recursion" would be only those extensions which can be reduced to primitive recursion, and not nested recursion. Since in the ensuing discussion, Bernays keeps stressing that the extensions discussed all have the character of a "recursion by which a procedure for stepwise successive computation of one or more functions is formalized" (p. 334) and that the introduction of the Ackermann function has the "required property of a formalized computation procedure" (p. 335) it seems to me that *that* is what Bernays takes to be the issue here. And as such, the Ackermann function does not violate "what is characteristic of the method of of recursive number theory." In any case, Tait concedes that the Ackermann function is not rejectd as finitist.

25. "Gewisse über die rekursive Zahlentheorie (im ursprünglichen Sinn) hinausgehende Verfahren der finiten Mathematik haben wir bereits im §7 besprochen, nämlich die Einführung von Funktionen durch verschränkte Rekursionen und die allgemeineren Induktionsschemata." (Hilbert and Bernays 1939, p. 340; 1970, p. 354). Emphasis mine.

26. “Der ursprüngliche engere Begriff der der finiten Aussage kommt im Gebiet der Zahlentheorie darauf hinaus, daß als finite zahlenthoretische Aussagen nur solche Aussagen zugelassen sind, die sich im Formalismus der rekursiven Zahlentheorie, eventuell unter Hinzunahme von Symbolen für gewisse berechenbare zahlentheoretische Funktionen (von einem oder mehreren Argumenten), jedoch ohne Benutzung von Formelvariablen darstellen oder die eine verschärfte Interpretation durch eine Aussage von dieser Form gestatten.” (Hilbert and Bernays 1939, p. 348; 1970, p. 362). Emphasis mine. The “original concept of finitism” refers to the conception of finitistic meaningfulness first introduced, in contrast to some slight extensions that are discussed subsequently, in particular, admission of implications with a universal antecedent and inductions with premises of such a form. The passage occurs in the context of considering the question of whether there are finitistic principles which go beyond number theory Z.

27. Hilbert and Bernays (1939, p. 214; 1970, p. 224). This passage was pointed out by Tait (2000).

28. For a discussion of nested recursion and the issues coming up in computing functions defined by nested recursion, see Tait (1961).

29. Bernays to Gödel, January 7, 1970. Bernays Papers, ETH Zürich/WHS, Hs. 975:1745.

30. Tait also mentions a passage from the preface to the second volume, where Bernays speaks of an extension of the finite standpoint. That passage, however, discusses specifically the methods used in Gentzen’s consistency proof, i.e., transfinite induction up to ε_0.

31. Compare this with Kant’s notion of intuition and geometrical knowledge: In that case we could conceivably buy the idea of intuition directly producing knowledge of a general geometrical theorem: The requisite geometrical proof is one diagram, and constructing the diagram in intuition provides knowledge of the theorem. But the case of finitist number theory is different in that there is no generic numeral which can take the place of the diagram.

32. Parikh’s Theorem is this: If $\varphi(x,y)$ represents a recursive function $f(x)$ which is provably total in $I\Delta_0$, i.e., $I\Delta_0 \vdash (\forall x)(\exists y)\varphi(x,y)$, then there is a polynomial p such that $f(x) \leq p(x)$ for all x.

33. See Hájek and Pudlák (1993, p. 320) and Parikh (1971).

34. This follows from Tait’s Theorem (Tait 1961), see also Rose (1984, Chapter 3).

35. I believe that Hilbert would have agreed to such a view, or at least not have drawn the line at $k = 1$, the primitive recursive functions. It is conceivable that stronger recursion schemata exist which take us all the way through the $< \varepsilon_0$-recursive functions. Even if all these were accepted as finitistic, as Kreisel (1960, 1970) does, still a qualitatively different extension of proof methods

would be required to attain a system strong enough to prove the consistency of arithmetic.

Bibliography

Abrusci, Vito Michele. (1989). David Hilbert's *Vorlesungen* on logic and the foundations of mathematics. In Giovanna Corsi, Corrado Mangione, and Massimo Mugnai (Eds.), *Atti del Convegno Internazionale di Storia della Logica, Le teorie delle modalità. 5–8 December 1987, San Gimignano* (pp. 333–338). Bologna: CLUEB.

Ackermann, Hans Richard. (1983). Aus dem Briefwechsel Wilhelm Ackermanns. *History and Philosophy of Logic* **4**, 181–202.

Ackermann, Wilhelm. (1924a). *Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit.* Dissertation, Universität Göttingen.

Ackermann, Wilhelm. (1924b). Begründung des "tertium non datur" mittels der Hilbertschen Theorie der Widerspruchsfreiheit. *Mathematische Annalen* **93**, 1–36.

Ackermann, Wilhelm. (1928a). Zum Hilbertschen Aufbau der reellen Zahlen. *Mathematische Annalen* **99**, 118–133.

Ackermann, Wilhelm. (1928b). Über die Erfüllbarkeit gewisser Zählausdrücke. *Mathematische Annalen* **100**, 638–649.

Ackermann, Wilhelm. (1940). Zur Widerspruchsfreiheit der Zahlentheorie. *Mathematische Annalen* **117**, 162–194.

Ackermann, Wilhelm. (1954). *Solvable Cases of the Decision Problem.* Amsterdam: North-Holland.

Baldus, Richard. (1928). Zur Axiomatik der Geometrie I: Über Hilberts Vollständigkeitsaxiom. *Mathematische Annalen* **100**, 321–333.

Behmann, Heinrich. (1921). *Entscheidungsproblem und Algebra der Logik.* (unpublished manuscript, dated May 10, 1921. Behmann Archive, Erlangen)

Behmann, Heinrich. (1922a). Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem. *Mathematische Annalen* **86**, 163–229.

Behmann, Heinrich. (1922b). *Mathematische Logik.* (Universität Göttingen, Sommer-Semester 1922. Unpublished lecture notes. Behmann Nachlaß, Institut für Philosophie, Universität Erlangen)

Benacerraf, Paul. (1973). Mathematical truth. *Journal of Philosophy* **70**, 661–680.

Benacerraf, Paul, and Putnam, Hilary (Eds.). (1983). *Philosophy of Mathematics* (2nd ed.). Cambridge: Cambridge University Press.

Bernays, Paul. (1918). *Beiträge zur axiomatischen Behandlung des Logik-Kalküls.* Habilitationsschrift, Universität Göttingen. (Bernays Nachlaß, WHS, Bibliothek, ETH Zürich, Hs 973.192)

Bernays, Paul. (1922). Über Hilberts Gedanken zur Grundlegung der Arithmetik [On Hilbert's thoughts concerning the grounding of arithmetic]. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **31**, 10–19. (English translation in Mancosu (1998a), pp. 215–222)

Bernays, Paul. (1923). Erwiderung auf die Note von Herrn Aloys Müller: Über Zahlen als Zeichen [Reply to the note by Mr. Aloys Müller, "On numbers as signs"]. *Mathematische Annalen* **90**, 159–63. (English translation in Mancosu (1998a), pp. 223–226)

Bernays, Paul. (1926). Axiomatische Untersuchungen des Aussagen-Kalkuls der „Principia Mathematica". *Mathematische Zeitschrift* **25**, 305-20.

Bernays, Paul. (1927). Probleme der theoretischen Logik. *Unterrichtsblätter für Mathematik und Naturwissenschaften* **33**, 369–77.

Bernays, Paul. (1930). Die Philosophie der Mathematik und die Hilbertsche Beweistheorie [The philosophy of mathematics and Hilbert's proof theory]. *Blätter für deutsche*

Philosophie **4**, 326–67. (Reprinted in Bernays (1976a), pp. 17–61. English translation in Mancosu (1998a), pp. 234–265)

Bernays, Paul. (1958). *Axiomatic Set Theory.* Amsterdam: North-Holland.

Bernays, Paul. (1976a). *Abhandlungen zur Philosophie der Mathematik.* Darmstadt: Wissenschaftliche Buchgesellschaft.

Bernays, Paul. (1976b). A short biography. In Gert H. Müller (Ed.), *Sets and Classes* (pp. xi–xiii). Amsterdam: North-Holland.

Bernays, Paul. (1977). *Interviews with J.-P. Sydler and E. Clavadetscher.* (Bernays Nachlaß, WHS, ETH Zürich, T 1285)

Bernays, Paul, and Schönfinkel, Moses. (1928). Zum Entscheidungsproblem der mathematischen Logik. *Mathematische Annalen* **99**, 342–372.

Birkhoff, Garrett, and Bennett, Mary Katherine. (1987). Hilbert's "Grundlagen der Geometrie". *Rendiconti del Circolo Matematico di Palermo, Serie II* **36**, 343–389.

Bocheński, I. M. (1956). *Formale Logik.* Freiburg: Alber.

Boolos, George. (1998). *Logic, Logic, and Logic.* Cambrdige, Mass.: Harvard University Press.

Brouwer, L. E. J. (1919). Intuitionistische Mengenlehre [Intuitionist set theory]. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **28**, 203–208. (Translated in: Mancosu (1998a), pp. 23–27)

Brouwer, L. E. J. (1921). Besitzt jede reelle Zahl eine Dezimalbruchentwicklung? [Does every real number have a decimal expansion?]. *KNAW Verslagen* **29**, 803–12. (Translated in: Mancosu (1998a), pp. 28–35)

Chihara, Charles S. (1973). *Ontology and the Vicious-Circle Principle.* Ithaca, N.Y., and London: Cornell University Press.

Church, Alonzo. (1956). *Introduction to Mathematical Logic.* Princeton, N.J.: Princeton University Press.

Cobham, A. (1965). The intrinsic computational difficulty of functions. In Y. Bar-Hillel (Ed.), *Proceedings of the International Congress on Logic, Methodology and Philosophy of Science* (pp. 24–30). Amsterdam: North-Holland.

Corcoran, John. (1980). Categoricity. *History and Philosophy of Logic* **1**, 187–207.

Corcoran, John. (1981). From categoricity to completeness. *History and Philosophy of Logic* **2**, 113–119.

Corry, Leo. (1997). David Hilbert and the axiomatization of physics (1894–1905). *Archive for the History of the Exact Sciences* **51**, 83–198.

Davis, Martin. (1994). Emil L. Post: His life and work. In Martin Davis (Ed.), *Solvability, Provability, Definability: The Collected Works of Emil L. Post* (pp. xi–xxviii). Boston: Birkhäuser.

Davis, Martin. (1995). American logic in the 1920s. *Bulletin of Symbolic Logic* **1**, 273–278.

Dawson, John W. (1988). The reception of Gödel's incompleteness theorems. In Stuart G. Shanker (Ed.), *Gödel's Theorem in Focus* (pp. 74–95). London and New York: Routledge.

Dedekind, Richard. (1888). *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg. (English translation in Dedekind (1901))

Dedekind, Richard. (1901). *Essays on the Theory of Number.* New York: Open Court.

Detlefsen, Michael. (1986). *Hilbert's Program.* Dordrecht: Reidel.

Došen, Kosta. (1993). A historical introduction to substructural logics. In Peter Schröder-Heister and Kosta Došen (Eds.), *Substructural Logics* (pp. 1–30). Oxford: Oxford University Press.

Dreben, Burton, and van Heijenoort, Jean. (1986). Introductory note to Gödel 1929, 1930, and 1930a. In Solomon Feferman et al. (Eds.), *Kurt Gödel. Collected Works* (Vol. 1, pp. 44–59). Oxford: Oxford University Press.

Ewald, William Bragg (Ed.). (1996). *From Kant to Hilbert. A Source Book in the Foundations of Mathematics* (Vol. 2). Oxford: Oxford University Press.

Feferman, Solomon. (1988). Hilbert's Program relativized: Proof-theoretial and fondational reductions. *Journal of Symbolic Logic* **53**(2), 364–284.

Feferman, Solomon. (1993a). What rests on what? The proof-theoretic analysis of mathematics. In Johannes Czermak (Ed.), *Philosophy of Mathematics. Proceedings of the Fifteenth International Wittgenstein-Symposium, Part 1* (pp. 147–171). Vienna: Hölder-Pichler-Tempsky.

Feferman, Solomon. (1993b). Why a little bit goes a long way: Logical foundations of scientifically applicable mathematics. *PSA 1992* **2**, 442–455. (Reprinted in Feferman (1998), Ch. 14, pp. 284–298)

Feferman, Solomon. (1998). *In the Light of Logic*. Oxford: Oxford University Press.

Feferman, Solomon, et al. (Eds.). (1986). *Kurt Gödel. Collected Works* (Vol. 1). Oxford: Oxford University Press.

Field, Hartry. (1980). *Science without Numbers. A Defence of Nominalism.* Princeton: Princeton University Press.

Frege, Gottlob. (1879). *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens.* Halle: Nebert. (Reprinted in Frege (1993), translated in van Heijenoort (1967), pp. 1–82)

Frege, Gottlob. (1980). *The Foundations of Arithmetic* (J. L. Austin, Trans.). Evanston, Ill.: Northwestern University Press. (Original work published 1884)

Frege, Gottlob. (1993). *Begriffsschrift und andere Aufsätze* (2nd ed.). Hildesheim: Olms.

Freudenthal, Hans. (1973). David Hilbert. In Charles Coulston Gillispie et al. (Eds.), *Dictionary of Scientific Biography* (Vol. 6, pp. 388–95). New York: Scribner. (14 vols.)

Gentzen, Gerhard. (1934). Untersuchungen über das logische Schließen I–II. *Mathematische Zeitschrift* **39**, 176–210, 405–431.

Gentzen, Gerhard. (1936). Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen* **112**, 493–565.

Gentzen, Gerhard. (1969). *The Collected Papers of Gerhard Gentzen.* Amsterdam: North-Holland.

Gödel, Kurt. (1929). *Über die Vollständigkeit des Logikkalküls.* Dissertation, Universität Wien. (Reprinted and translated in Feferman et al. (1986), pp. 60–101)

Gödel, Kurt. (1930). Die Vollständigkeit der Axiome des logischen Funktionenkalküls. *Monatshefte für Mathematik und Physik* **37**, 349–360. (Reprinted and translated in Feferman et al. (1986), pp. 102–123)

Gödel, Kurt. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme I. *Monatshefte für Mathematik und Physik* **38**, 173–198.

Gödel, Kurt. (1932). Zum intuitionistischen Aussagenkalkül [On the intuitionistic propositional calculus]. *Anzeiger der Akademie der Wissenschaften in Wien* **69**, 65–66. (Reprinted and translated in Feferman et al. (1986), pp. 222–225)

Hájek, Petr, and Pudlák, Pavel. (1993). *Metamathematics of First-order Arithmetic.* Berlin: Springer.

Hand, Michael. (1989). A number is the exponent of an operation. *Synthese* **81**, 243–265.

Hand, Michael. (1990). Hilbert's iterativistic tendencies. *History and Philosophy of Logic* **11**, 185–192.

Hart, W. D. (Ed.). (1996). *The Philosophy of Mathematics.* Oxford: Oxford University Press.

Hermes, Hans. (1967). In memoriam Wilhelm Ackermann, 1896–1962. *Notre Dame Journal of Formal Logic* **8**, 1–8.

Heyting, Arend. (1930). Die formalen Regeln der intuitionistischen Logik [The formal rules of intuitionistic logic]. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*, 42–56. (Translated in Mancosu (1998a), pp. 311–334)

Hilbert, David. (1899). Grundlagen der Geometrie. In *Festschrift zur Feier der Enthüllung des Gauss-Weber-Denkmals in Göttingen* (pp. 1–92). Leipzig: Teubner.

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

Hilbert, David. (1900b). Über den Zahlbegriff. *Jahresbericht der Deutschen Mathematiker-Vereinigung* **8**, 180–84.

Hilbert, David. (1903). *Grundlagen der Geometrie* (2nd ed.). Leipzig: Teubner.

Hilbert, David. (1905a). *Logische Principien des mathematischen Denkens.* (Vorlesung, Sommer-Semester 1905. Lecture notes by Ernst Hellinger. Unpublished manuscript, 277 pp. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1905b). *Logische Principien des mathematischen Denkens.* (Vorlesung, Sommer-Semester 1905. Lecture notes by Max Born. Unpublished manuscript, 188 pp. Cod. Ms. Hilbert 558a. Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek Göttingen)

Hilbert, David. (1905c). Über die Grundlagen der Logik und der Arithmetik [On the foundations of logic and arithmetic]. In A. Krazer (Ed.), *Verhandlungen des dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904* (p. 174-85). Leipzig: Teubner. (English translation in van Heijenoort (1967), pp. 129–38)

Hilbert, David. (1908). *Prinzipien der Mathematik.* (Vorlesung, Sommer-Semester 1908. Lecture notes. Unpublished manuscript, 206 pp. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1910). *Elemente und Prinzipienfragen der Mathematik.* (Vorlesung, Sommer-Semester 1910. Lecture notes by Richard Courant. Unpublished manuscript, 163 pp. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1913). *Einige Abschnitte aus der Vorlesung über die Grundlagen er Mathematik und Physik.* (Vorlesung, Sommer-Semester 1913. Lecture notes by Bernhard Baule. Unpublished manuscript. Cod. Ms. Hilbert 559. Hilbert-Nachlaß, Niedersächsische Staats- und Universitätsbibliothek Göttingen)

Hilbert, David. (1917). *Mengenlehre.* (Lecture notes by Margarethe Loeb. Sommer-Semester 1917. Unpublished manuscript. Bibliothek, Mathematisches Institut, Universität Göttingen)

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

Hilbert, David. (1918b). Axiomatisches Denken. In *Verhandlungen der Schweizerischen Naturforschenden Gesellschaft. 99. Jahresversammlung vom 9.–12. September 1917 in Zürich. I. Teil* (pp. 129–130). Aarau: Sauerländer.

Hilbert, David. (1918c). *Prinzipien der Mathematik.* (Lecture notes by Paul Bernays. Winter-Semester 1917–18. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1920a). *Logik-Kalkül.* (Vorlesung, Winter-Semester 1920. Lecture notes by Paul Bernays. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1920b). *Probleme der mathematischen Logik.* (Vorlesung, Sommer-Semester 1920. Lecture notes by Paul Bernays and Moses Schönfinkel. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen)

Hilbert, David. (1922a). *Grundlagen der Mathematik.* (Vorlesung, Winter-Semester 1921–22. Lecture notes by Helmut Kneser. Unpublished manuscript, three notebooks)

Hilbert, David. (1922b). *Grundlagen der Mathematik.* (Vorlesung, Winter-Semester 1921–22. Lecture notes by Paul Bernays. Unpublished typescript. Bibliothek, Mathematisches Institut, Universität Göttingen)

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

Hilbert, David. (1923). Die logischen Grundlagen der Mathematik [The logical foundations of mathematics]. *Mathematische Annalen* **88**, 151-165. (Lecture given at the Deutsche Naturforscher-Gesellschaft, September 1922. Reprinted in Hilbert (1935), pp. 178–191. English translation in Ewald (1996), pp. 1134–1148)

Hilbert, David. (1926). Über das Unendliche [On the infinite]. *Mathematische Annalen* **95**, 161–90. (Lecture given Münster, 4 June 1925. English translation in van Heijenoort (1967), pp. 367–392)

Hilbert, David. (1928a). Die Grundlagen der Mathematik [The foundations of mathematics]. *Abhandlungen aus dem Seminar der Hamburgischen Universität* **6**, 65-85. (English translation in van Heijenoort (1967), pp. 464-479)

Hilbert, David. (1928b). Probleme der Grundlegung der Mathematik. In Nicola Zanichelli (Ed.), *Atti del Congresso Internazionale dei Matematici. 3–10 September 1928, Bologna* (pp. 135–141).

Hilbert, David. (1929). Probleme der Grundlegung der Mathematik [Problems of the grounding of mathematics]. *Mathematische Annalen* **102**, 1–9. (Lecture given at the International Congress of Mathematicians, 3 September 1928. English translation in Mancosu (1998a), pp. 266–273)

Hilbert, David. (1931). Die Grundlegung der elementaren Zahlenlehre [The grounding of elementary number theory]. *Mathematische Annalen* **104**, 485–94. (Reprinted in Hilbert (1935), pp. 192–195. English translation in Ewald (1996), pp. 1148–1157.)

Hilbert, David. (1935). *Gesammelte Abhandlungen* (Vol. 3). Berlin: Springer.

Hilbert, David, and Ackermann, Wilhelm. (1928). *Grundzüge der theoretischen Logik*. Berlin: Springer.

Hilbert, David, and Bernays, Paul. (1923a). *Logische Grundlagen der Mathematik.* (Winter-Semester 1922–23. Lecture notes by Helmut Kneser. Unpublished manuscript.)

Hilbert, David, and Bernays, Paul. (1923b). *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)

Hilbert, David, and Bernays, Paul. (1934). *Grundlagen der Mathematik* (Vol. 1). Berlin: Springer.

Hilbert, David, and Bernays, Paul. (1939). *Grundlagen der Mathematik* (Vol. 2). Berlin: Springer.

Hilbert, David, and Bernays, Paul. (1968). *Grundlagen der Mathematik* (Vol. 1, 2nd ed.). Berlin: Springer.

Hilbert, David, and Bernays, Paul. (1970). *Grundlagen der Mathematik* (Vol. 2, 2nd ed.). Berlin: Springer.

Hill, Claire Ortiz. (1995). Husserl and Hilbert on completeness. In Jaakko Hintikka (Ed.), *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics* (pp. 143–163). Amsterdam: Kluwer.

Hofweber, Thomas. (2000). Proof-theoretic reduction as a philosopher's tool. *Erkenntnis* **53**, 127–146.

Huntington, Edward V. (1935). The inter-deducibility of the new Hilbert-Bernays theory and *Principia Mathematica*. *Annals of Mathematics* **36**, 313–324.

Ignjatovič, Aleksandar. (1994). Hilbert's program and the omega rule. *Journal of Symbolic Logic* **59**, 322–343.

Jørgensen, Jørgen. (1931). *A Treatise of Formal Logic*. Copenhagen: Levin & Munksgaard.

Kambartel, Friedrich. (1975). Frege und die axiomatische Methode. Zur Kritik mathematikhistorischer Legitimationsversuche der formalistischen Ideologie. In Christian Thiel (Ed.), *Frege und die moderne Grundlagenforschung* (pp. 77–89). Meisenheim a. Glan: Hain.

Kitcher, Philip. (1976). Hilbert's epistemology. *Philosophy of Science* **43**, 99–115.

Kneale, William, and Kneale, Martha. (1962). *The Development of Logic*. Oxford: Oxford University Press.

König, Julius. (1914). *Neue Grundlagen der Logik, Arithmetik und Mengenlehre.* Veit.

Krajíček, Jan. (1995). *Bounded Arithmetic, Propositional Logic, and Complexity Theory.* Cambridge: Cambridge University Press.

Kreisel, Georg. (1958). Hilbert's programme. *Dialectica* **12**, 346–372. (Reprinted as (Kreisel 1983).)

Kreisel, Georg. (1960). Ordinal logics and the characterization of informal notions of proof. In J. A. Todd (Ed.), *Proceedings of the International Congress of Mathematicians. Edinburgh, 14–21 August 1958* (pp. 289–299). Cambridge: Cambridge University Press.

Kreisel, Georg. (1965). Mathematical logic. In T. L. Saaty (Ed.), *Lectures on Modern Mathematics* (Vol. 3, pp. 95–195). New York: Wiles.

Kreisel, Georg. (1970). Principles of proof and ordinals implicit in given concepts. In A. Kino, J. Myhill, and R. E. Veseley (Eds.), *Intuitionism and Proof Theory.* Amsterdam: North-Holland.

Kreisel, Georg. (1983). Hilbert's programme. In Paul Benacerraf and Hilary Putnam (Eds.), *Philosophy of Mathematics* (2nd ed., pp. 207–238). Cambridge: Cambridge University Press.

Lauener, Henri. (1978). Paul Bernays (1888–1977). *Zeitschrift für allgemeine Wissenschaftstheorie* **9**, 13–20.

Leisenring, A. C. (1969). *Mathematical Logic and Hilbert's* ε*-Symbol.* London: MacDonald.

Lewis, C. I. (1918). *A Survey of Symbolic Logic.* Berkeley: University of California Press.

Löwenheim, Leopold. (1915). Über Möglichkeiten im Relativkalkül. *Mathematische Annalen*, 447–470. (Translated in: van Heijenoort (1967), pp. 228–251)

Łukasiewicz, Jan. (1924). Démonstration de la compatibilité des axiomes de la théorie de la déduction (Abstract). *Annales de la Société Polonaise de Mathématique* **3**, 149. (Talk given 13 June 1924)

Maddy, Penelope. (1997). *Naturalism in Mathematics.* Oxford: Oxford University Press.

Majer, Ulrich. (1997). Husserl and Hilbert on completeness. *Synthese* **110**, 37–56.

Mancosu, Paolo (Ed.). (1998a). *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920s.* Oxford: Oxford University Press.

Mancosu, Paolo. (1998b). Hilbert and Bernays on metamathematics. In Paolo Mancosu (Ed.), *From Brouwer to Hilbert* (pp. 149–188). Oxford: Oxford University Press.

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

Mancosu, Paolo. (1999b). Between Vienna and Berlin: The immediate reception of Gödel's incompleteness theorems. *History and Philosophy of Logic* **20**, 33–45.

Mancosu, Paolo. (200?). The Russellian influence on Hilbert and his school. *Synthese*. (forthcoming)

Mints, Grigori, and Tupailo, Sergei. (1996). Epsilon substitution method for elementary analysis. *Archive for Mathematical Logic* **35**, 103–130.

Mints, Grigori E. (1994). Gentzen-type systems and Hilbert's epsilon substitution method. I. In Dag Prawitz, Brian Skyrms, and Dag Westerståhl (Eds.), *Logic, Methodology and Philosophy of Science IX* (pp. 91–122). Amsterdam: Elsevier.

Mollerup, Johannes. (1907). Die Definition des Mengenbegriffs. *Mathematische Annalen* **64**, 231–238.

Moore, Gregory H. (1978). The origins of Zermelo's axiomatization of set theory. *Journal of Philosophical Logic* **7**, 307–329.

Moore, Gregory H. (1982). *Zermelo's Axiom of Choice. Its Origins, Development and Influence.* New York and Heidelberg: Springer.

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

Moser, Georg. (2000). *The Epsilon Substitution Method.* Master's thesis, University of Leeds.

Müller, Aloys. (1923). Über Zahlen als Zeichen. *Mathematische Annalen* **90**, 153–158; 163.

Niebergall, Karl-Georg, and Schirn, Matthias. (1998). Hilbert's finitism and the notion of infinity. In Matthias Schirn (Ed.), *The Philosophy of Mathematics Today* (pp. 271–305). Oxford: Oxford University Press.

O'Leary, Daniel J. (1988). The propositional logic of *Principia Mathematica* and some of its forerunners. *Russell* **8**, 92–115.

Parikh, Rohit. (1971). Existence and feasibility in arithmetic. *Journal of Symbolic Logic* **36**, 494–508.

Parsons, Charles. (1979–80). Mathematical intuition. *Proceedings of the Aristotelian Society* **80**, 145–68. (Reprinted in Hart (1996), 95–113)

Parsons, Charles. (1994). Intuition and number. In Alexander George (Ed.), *Mathematics and Mind* (pp. 141–157). Oxford: Oxford University Press.

Parsons, Charles. (1998a). Finitism and intuitive knowledge. In Matthias Schirn (Ed.), *The Philosophy of Mathematics Today* (pp. 249–270). Oxford: Oxford University Press.

Parsons, Charles. (1998b). Intuition and the abstract. In Marcelo Stamm (Ed.), *Philosophie in synthetischer Absicht. Synthesis in Mind* (pp. 155–187). Stuttgart: Klett-Cotta.

Parsons, Charles. (2000). Reason and intuition. *Synthese* **125**, 299–315.

Peckhaus, Volker. (1990). *Hilbertprogramm und Kritische Philosophie.* Göttingen: Vandenhoeck und Ruprecht.

Peckhaus, Volker. (1994). Logic in transition: The logical calculi of Hilbert (1905) and Zermelo (1908). In Dag Prawitz and Dag Westerståhl (Eds.), *Logic and Philosophy of Science in Uppsala* (pp. 311–323). Dordrecht: Kluwer.

Peckhaus, Volker. (1995). Hilberts Logik: Von der Axiomatik zur Beweistheorie. *Internationale Zeitschrift für Geschichte und Ethik der Naturwissenschaften, Technik und Medizin* **3**, 65–86.

Post, Emil L. (1921). Introduction to a general theory of elementary propositions. *American Journal of Mathematics* **43**, 163–185.

Reid, Constance. (1970). *Hilbert.* New York: Springer.

Resnik, Michael. (1980). *Frege and the Philosophy of Mathematics.* Ithaca: Cornell University Press.

Resnik, Michael D. (1974). The Frege-Hilbert controversy. *Philosophy and Phenomenological Research* **34**, 386–403.

Rose, H. E. (1984). *Subrecursion. Functions and Hierarchies.* Oxford University Press.

Russell, Bertrand. (1906). The theory of implication. *American Journal of Mathematics* **28**, 159–202.

Scanlan, Michael. (1991). Who were the American Postulate Theorists? *Journal of Symbolic Logic* **56**, 981–1002.

Schirn, Matthias (Ed.). (1998). *The Philosophy of Mathematics Today.* Oxford: Oxford University Press.

Schönfinkel, Moses. (1924). Über die Bausteine der mathematischen Logik. *Mathematische Annalen* **92**, 305–316. (English translation in van Heijenoort (1967), pp. 355–366)

Schröder, Ernst. (1890). *Vorlesungen über die Algebra der Logik* (Vol. 1). Leipzig: Teubner.

Sieg, Wilfried. (1988). Hilbert's program sixty years later. *Journal of Symbolic Logic* **53**(2), 338–348.

Sieg, Wilfrid. (1990). Reflections on Hilbert's program. In Wilfrid Sieg (Ed.), *Acting and Reflecting* (pp. 171–82). Dordrecht: Kluwer.

Sieg, Wilfried. (1999). Hilbert's programs: 1917–1922. *Bulletin of Symbolic Logic* **5**(1), 1–44.

Simpson, Stephen G. (1988). Partial realizations of Hilbert's program. *Journal of Symbolic Logic* **53**(2), 349–363.

Sinaçeur, Hourya. (1993). Du formalisme à la constructivité: le finitisme. *Revue Internationale de Philosophie* **47**(4), 251–283.

Skolem, Thoralf. (1923). Begründung der elementaren Arithmetik durch die rekurrierende Denkweise ohne Anwendung scheinbarer Veränderlichen mit unendlichem Ausdehnungsbereich. *Videnskapsselskapets skrifter I. Matematisk-naturvidenskabelig klasse* **6**. (English translation in van Heijenoort (1967), pp. 302–333)

Specker, Ernst. (1979). Paul Bernays. In M. Boffa, D. van Dalen, and K. McAloon (Eds.), *Logic Colloquium '78* (pp. 381–89). Amsterdam: North-Holland.

Surma, Stansław J. (1973). A historical survey of the significant methods of proving Post's theorem about the completeness of the classical propositional calculus. In Stanisław J. Surma (Ed.), *Studies in the History of Mathematical Logic* (pp. 19–32). Wrocław: Polish Academy of Sciences, Institute of Philosophy and Sociology.

Tait, W. W. (1961). Nested recursion. *Mathematische Annalen* **143**, 236–250.

Tait, W. W. (1965). The substitution method. *Journal of Symbolic Logic* **30**, 175-192.

Tait, W. W. (1968). Constructive reasoning. In B. Van Rootselaar and J. F. Staal (Eds.), *Logic, Methodology and Philosophy of Science III. Amsterdam 1967* (p. 185-199). Amsterdam: North-Holland.

Tait, W. W. (1981). Finitism. *Journal of Philosophy* **78**, 524-546.

Tait, W. W. (1986). Truth and proof: The platonism of mathematics. *Synthese* **69**, 341–370. (Reprinted in Hart (1996), pp142–167)

Tait, W. W. (2000). Remarks on finitism. In Wilfried Sieg, Richard Sommer, and Carolyn Talcott (Eds.), *Reflections. A Festschrift Honoring Solomon Feferman.* (forthcoming)

Toepell, Michael-Markus. (1986). *Über die Entstehung von Hilbert's "Grundlagen der Geometrie"*. Göttingen: Vandenhoek und Ruprecht.

van Heijenoort, Jean (Ed.). (1967). *From Frege to Gödel. A Source Book in Mathematical Logic, 1897–1931.* Cambridge, Mass.: Harvard University Press.

von Neumann, Johann. (1927). Zur Hilbertschen Beweistheorie. *Mathematische Zeitschrift* **26**, 1–46.

Weyl, Hermann. (1918). *Das Kontinuum.* Leipzig: Veit. (English translation: Weyl (1918/1994))

Weyl, Hermann. (1921). Über die neue Grundlagenkrise der Mathematik [On the new foundational crisis of mathematics]. *Mathematische Zeitschrift* **10**, 37–79. (English translation in: Mancosu (1998a), pp. 86–118)

Weyl, Hermann. (1944). David Hilbert and his mathematical work. *Bulletin of the American Mathematical Society* **50**, 612–654.

Weyl, Hermann. (1994). *The Continuum.* New York: Dover. (Original work published 1918)

Whitehead, Alfred North, and Russell, Bertrand. (1910). *Principia Mathematica* (Vol. 1). Cambridge: Cambridge University Press. (Quotes from the 1978 reprint of the second edition, 1927)

Whitehead, Alfred North, and Russell, Bertrand. (1913). *Principia Mathematica* (Vol. 2). Cambridge: Cambridge University Press.

Wright, Crispin, and Hale, Bob. (2001). *The Reason's Proper Study. Essays towards a Neo-Fregean Philosophy of Mathematics.* Oxford: Oxford University Press.

Zach, Richard. (1998). Numbers and functions in Hilbert's finitism. *Taiwanese Journal for Philosophy and History of Science* **10**, 33–60.

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

Zermelo, Ernst. (1908). *Mathematische Logik.* (Vorlesung gehalten von Prof. Dr E. Zermelo zu Göttingen im S.S. 1908. Lecture notes by Kurt Grelling. Nachlaß Zermelo, Kapsel 4, Universitätsbibliothek Freiburg im Breisgau)