THE STRUCTURE & LIMITS OF FINITISM

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Alexei E. Angelides
August 2014

Preface

In the years since Hilbert announced his now famous proof theory and the years following Gödel's Incompleteness Theorems much important work has been done in the subject that has cemented its fate as one of the pillars of metamathematics. To be sure, many of the directions that proof theory has taken since the 1920s can be found as germs in Hilbert and the Hilbert School's writings from the period — Kreisel's unwinding program a classic example of the metamathematical analysis of a proof to extract its mathematical content, Kohlenbach's program to extract the computational content of a proof using proof-theoretic techniques, Feferman's conservativity results, a spectacular example that shows the proof-theoretic relationships between formal theories, and so many more. But the one direction that precipitated the whole of proof theory — namely, to provide an epistemic foundation for a mathematical theory by providing it with a consistency proof from the point of view of a theory strictly weaker than it, a *finitistic* theory in the Hilbert School's language — has remained mostly unearthed. It's the goal in what follows to excavate some of these old stones.

One thing I have learned in writing a dissertation is that there are those who help in its writing, and there are those who *help* in its writing. Of the former conjunct, there are, of course, many: Paolo Mancosu, Hans Sluga, Grisha Mints, and Michael Friedman were especially encouraging along the way. My exchanges with Wilfried Sieg and his genuine love for the subject historical and mathematical are inspiring. Mic Detlefsen's graciousness as host at the ENS was crucial for cementing some of the final ideas. To Krista Lawlor and Alexis Burgess I owe all the sound philosophical direction they could muster and I've failed to exhibit, so it should be noted that in finding things wanting readers should conclude it's probably not their fault. Shout out to the Up All Night Logic Crew in my cohort. Of both the former and the latter there are fewer: Solomon Feferman, my advisor at Stanford for most of my time there, has a special place in my heart as both teacher and friend — there is no equal. A knowing nod to Jesse Alama, my good friend and one of the best logicians ever. And Thomas Ryckman's mastery of the history, philosophy, and mathematics in this subject and constant encouragement in my writing is one of the only reasons this document exists today. Of the latter conjunct alone, there is just one: Marie. Without her unbounded patience and love I honestly would have never made it through. It is to her that I dedicate this dissertation.

Contents

Chapter 1

Introduction

1.1 Introduction

Hilbert's Program has a bad rap. Not long after Gödel's announcement in the 1930s of his first and second incompleteness theorems, most of those working in metamathematics thought that if the first theorem didn't portend Hilbert's Program's death, then Gödel's second theorem certainly did. Indeed, how could it be otherwise? Since lore has it that Hilbert's Program (HP) is an attempt to use mathematics to prove that formalized mathematical theories are consistent, as it is usually glossed, and the first theorem produces the existence of the troubling true but unprovable sentence in the theory, it's the second that seems to do the real damage. For if we code that sentence up in the theory to say that the theory itself is consistent, then it seems to follow from the first theorem that said sentence is true but unprovable. In plain English, Gödel's first and second theorems taken together seem to state a simple and, to the Hilbertian, devastating fact: the consistency of a mathematical theory is unprovable by that theory. It's precisely that fundamental fact that seems to dash all hopes and the history of metamathematics since the late 1930s seems to bear this out. Current interest in metamathematics centers primarily around competing metamathematical disciplines that gained steam after Gödel's result — the dominance of set theory and category theory in metamathematics are outstanding examples. On the other hand, the direction of classical proof theory turned — Kreisel's unwinding program and Feferman's conservativity results famously — and any inkling of rehabilitating, mining, or reworking Hilbert's consistency program along its original lines has, for the most part, vanished.

Hilbert's Program was intended to be a one-two punch. The consistency proof was part of a program of providing epistemic foundations for classical mathematics. Set-theoretic paradoxes showed that the general classical mathematics once thought to be an irrefrageable science stood on shaky ground — what Hermann Weyl called in his *Das Kontinuum* a house "to a large degree built on sand" (Weyl (1918), 1). Hilbert's thought went that if it were possible to clean up the classical concepts by formalizing them in a regimented theory and proving mathematically that the theory could not produce a contradiction, then such a proof would constitute grounds for believing that theory. For if a theory is consistent, then it follows that there's a model of it in which all of its axioms are true. Hence, if we believe the axioms, the axioms are true, and the consistency proof justifies our belief that the axioms are true, then on most classical accounts it follows that we have knowledge of the axioms. But in light of the traditional interpretation of the incompleteness theorems, using a consistency proof as justification is impossible. But might there have been a different way to salvage the epistemic portion of Hilbert's Program? Consensus seems to have settled on the negative. In the years following, many well-known philosophers took shots at dismantling even that — Kitcher's famous 1976 essay and Maddy's continued offense against the very idea of proof-theoretic foundations for mathematics are resounding. In more recent years, there's been no shortage of prominent philosophers to lend their ink to the claim that no such epistemic rehabilitation effort is plausible. For whom it tolls indeed.

Because two distinctive types of pressure have been applied to Hilbert's Program since the incompleteness theorems, most commentators believe that there is no meaningful way to rehabilitate its central theme of consistency on its own terms from either epistemological or mathematical points of view. That is, it is said that Hilbert's Program (HP) fails both for "underlying" principled philosophical reasons and for general technical mathematical ones, and hence, the consensus seems to be that there is simply no hope. It is, of course, the goal of this dissertation to demonstrate that the consensus is incorrect on both counts.[1]

Before we launch into sketches of the two main claims of this dissertation, let's begin with the two main claims that we'll be arguing against throughout. (It will take until the end of the last chapter to bear out our counterargument, so there are no free lunches in what follows.) The first claim to be argued against is found in Benacerraf (1973), an article that set the stage for nearly every discussion that followed about the epistemology of mathematics. In it he argues that on one hand, the semantics for an account of mathematical truth must parallel the semantics for the remainder of natural language. On the other hand, any account of mathematical truth must also explain how its sentences are known by providing a reasonable explanation for why beliefs about those sentences are justified or reliable. Benacerraf calls views that satisfy the first desideratum the "standard view," views for which truth-conditions for propositions such as:

(1) there are at least three perfect numbers greater than 17

parallel the truth-conditions for propositions such as

(2) there are at least three large cities older than New York

in that the terms in both refer to objects, properties, and relations, the latter cities and sizes, the former abstract objects and properties. Views that satisfy the second desideratum are dubbed

[1] Of course there are exceptions. See, e.g., Simpson (1999) and Franks (2009).

"combinatorial views," views in which mathematical truth is identified with provability in some formal system, and proofs are, of course, finite and cognitively accessible artifacts. Let us call Benacerraf's Dilemma (BD) the claim that no single account can consistently satisfy both desiderata.

The second set of claims that we'll take up are inferences from Gödel's (1931) famous first and second incompleteness theorems. Gödel's first incompleteness theorem (G1) states that for any effectively enumerable consistent formal theory T that axiomatizes a sufficient amount of arithmetic there is a formula ϕ in the language of T that is true in its model but neither provable nor refutable by T. Gödel's second incompleteness theorem (G2) states that for a consistent T it is possible to produce a formula ϕ that formalizes the assertion that T is consistent such that ϕ is true in its model, but again neither provable nor refutable by T. (G1) is a theorem about the existence of an unprovable sentence in the language of T but it does not specify what that sentence might be. By contrast (G2) specifies an inherently metamathematical sentence that was and is central to the finitist's goal of constructing a finitistic consistency proof for arithmetic. With very few exceptions philosophers and logicians following Gödel (1931) have claimed that (G2) is the death knell for (HP) since it appears to entail that for any T satisfying the above conditions, T fails to prove its own consistency, and no consistency proof of T is formalizable in T. Let's distinguish between these claims. In particular refer to the second incompleteness theorem found in Gödel (1931) as:

$$\text{(G2)} \qquad T \nvdash Con(T)$$

and distinguish it from:

(1) that T does not prove its own consistency; and

(2) that no consistency proof of T is formalizable in T.

Hence, while (G2) is itself a mathematical assertion, it is inferences (1) and (2) from (G2) by which the consensus concludes that (HP) is unviable.

Though there is some precedent, in what follows our intention is neither to argue that (BD) is incoherent nor that (1) and (2) are false — and certainly not to argue that (G2) is anything other than exactly what it is, a proven, established mathematical theorem. On the contrary, if we take (BD) to be a straightforwardly problematic dilemma for accounts of mathematical knowledge, and if we take (1) and (2) at face value, then the general goal of this dissertation can be distilled to the thought that (HP) provides a reasonable mathematical epistemology and a mathematical program that survives (the inferences from) Gödel's incompleteness theorems, specifically the second. More precisely, our main claim is that there is a reasonable epistemic reconstruction of finitism under which the inferences from (G2) fail and that is epistemically neutral with respect to (BD).

1.1.1 Finitism and the Benacerraf Dilemma

Clearly the main claim is a conjunction of two subclaims, so let's spell them out for a moment. The first claim is that:

($*$) our reconstruction of (HP) is neutral with respect to (BD),

and the second claim in this dissertation is that:

($**$) on our reconstruction inferences (1) and (2) fail.

For the moment let's focus on ($*$). Hilbert isolated a class of processes and a class of judgments that at various points he calls "intuitive," "finitely verifiable," or "completely surveyable." The general thought was to isolate a class of judgments and the processes for recognizing them as in need of no further justification — they are basic, bedrock. This class was to be sharply distinguished from the class of judgements concerning closed infinite sets such as the set of integers $\mathbb{Z}$. Hilbert understood the latter, infinitary part of mathematics as primarily a means to derive facts about the former, finitary part of mathematics, but because of the classical antinomies in set theory, infinitary mathematics was seen to be epistemically unreliable. Hence, were it possible to isolate the reliable part of classical mathematics in order to prove that the unreliable part is consistent, then because consistency implies truth in this context, one has justified, true beliefs about infinitary mathematics, and so, a clear claim to knowledge.

Before we go any further there are some important distinctions to make. It's useful to view (HP) as having two distinct tiers. At the first, pre-theoretic, tier lies the distinction between finitary and infinitary mathematics. This is the "school arithmetic" tier in which we make judgments involving equalities and inequalities between numerals and learn to construct and decompose these numerals intuitively by iteration. It's at this level that the finitary class of judgments can be called "intuitive" or "finitely verifiable." Because this is a pre-theoretic and unanalyzed tier, the distinction between finitary and infinitary arithmetic remains vague. But at the second, theoretic tier, one views arithmetic as a formal theory regimented in a formal language. Then, those judgments that were intuitively understood to be finitary at the pre-theoretic level form a class of sentences that Hilbert calls "real," to be contrasted with those sentences that are infinitary at the pre-theoretic level and that now form a class called "ideal." However, it is not simply that one views arithmetic in this setting as a formal theory. Rather, one regiments a set of arithmetical judgments into a formal language in order to provide a rigorous *analysis* of the judgments, processes, and distinctions that does not affect practice at the pre-theoretic tier. It's an important, but often overlooked point. For although Benacerraf and others following him have understood Hilbert's epistemological view to land squarely on the combinatorial side, as will become clear in chapter two Hilbert understood his program as one that neither presupposed nor entailed any philosophical supplementation. Indeed one of the goals of the consistency proof was to eradicate a kind of philosophical skepticism about the reliability of arithmetic that had produced theory-laden "solutions" to the classical antinomies such as Frege's platonism and Brouwer's constructivism.

At the pre-theoretic tier it is possible, but unnecessary, to adopt whatever position one wants towards arithmetic — platonist, constructivist, whatever — though in adopting a philosophical position towards arithmetic before its analysis one obviously puts the cart before the horse. Hilbert's idea was to leave the pre-theoretic tier of arithmetic alone, to view it in the most natural way when we approach it from the standpoint of someone practicing arithmetic. For Hilbert this means that there is a natural point of view when we operate with numbers in which the iterative manipulation of finite numbers and finite sequences (and, for that matter, the intuitive use of induction in proofs of simple arithmetical theorems) is bedrock and basic.[2] It is this tier and this point of view towards arithmetic and mathematics generally that informs the epistemological position that Hilbert dubbed finitism. However, it's at the second tier that Hilbert's Program is to be understood as a philosophical description of mathematical activity whose goal is to explain how the finitist at the first tier can regard the practice of infinitary mathematics as consistent with his rational standards. It's at this second tier where one deploys real sentences in order to explain how ideal sentences are not at odds with them (read: not inconsistent with), just in case such sentences "save" or preserve our most natural way of approaching arithmetic (and mathematics more generally). One might, per the platonist or nominalist, nonetheless cite our epistemic situation towards the objects of arithmetic and insist on the literal truth or falsity of ideal sentences, but it's unnecessary baggage.

Hence, the sense in which (HP) is neutral with respect to (BD) is that if a consistency proof is possible from the finitist standpoint (what Hilbert calls from 1922 onwards the *finit Standpunkt*), then one need not commit to either the standard or combinatorialist positions any more than what's warranted by our natural attitude towards the practice of arithmetic, which is to say, not much. Of course, (G2) appears to present a serious hurdle to claim ($*$). However, in what follows we shall show that on one interpretation of (G2), inferences (1) and (2) fail, but on a more plausible interpretation inferences (1) and (2) follow but only for certain choices of T and for certain interpretations of the consistency predicate. So while (G2) presents numerous hurdles to (HP), we intend to show that they are distinct from what the consensus believes them to be and, in fact, our reconstruction shows that it is possible to carry out (HP) for certain non-trivial fragments of arithmetic. If we then conjoin that to the claim ($*$), then the main claim we'll pursue throughout this dissertation is that (HP) is neutral with respect to (BD) for those fragments of arithmetic that remain untouched by (G2). One of the main questions that we pursue throughout chapter two is which theories those are, and later we shall argue that for both historical and epistemic reasons it's at least those theories whose upper bound is $\omega^{\omega^{\omega}}$, but that there is good reason to think that even more is possible. Our claim is that such theories constitute the upper bound for the mathematics that can serve as a solution to (BD), but that philosophical defenses of mathematics that go beyond must answer it head on.

[2] In the next chapter, one goal is to understand the limit of this "natural" point of view. For example, if primitive recursion is admissible as finitistic, what about nested recursions? We'll return to these questions then.

1.2 Finitism and Quasi-Empiricism

In the last section we argued that since it is possible to construe finitism either platonistically or nominalistically, and hence it is neutral with respect to the Benacerraf dilemma, finitism deserves a special place in the philosophy of mathematics. The idea is that the finitistic standpoint provides an epistemological point of view and an associated mathematical program that, if successfully carried out, renders the question of whether the objects of a mathematical theory are abstracta or concreta irrelevant to the analysis of a mathematical theory. At the outset we mentioned that these two questions ought to be separated for purposes of analysis. On the one hand, there is the question of whether or not Hilbert's Program can be successfully carried out — that question we leave to chapter three. On the other hand, there is the question as to what constitutes the epistemological point of view underlying the finitistic standpoint — that question is tackled in chapter two. Only if we understand the question about its epistemological point of view can we answer whether or not (HP) might be successful and, if so, to what extent. It is the task of this section to outline the epistemological point of view reconstructed throughout the next chapters. Our basic goal here is provide a guide to those chapters and, hence, the epistemological reconstruction of the finitistic standpoint found in the dissertation.

1.2.1 The Finitist Hierarchy

Our claim is that adopting the finitistic standpoint gives rise to an implicit analysis of mathematical judgments. For finitists there are three levels of judgments and each level possesses its own type of justification. At the lowest level are variable-free judgments about finite sequences of the form:

$$|||| = ||||$$

indicating the parity or disparity of a representation of what Hilbert labeled "concrete" sequences of strokes. At this level of judgments and representations we have the finitist's basic data. In the middle of the hierarchy we find free-variable judgments involving variables intended to "stand in" for individual but arbitrary numerals of the form:

$$a + b = b + a$$

such that the judgment is true or false upon replacing the variables a and b with concrete numerals. Though such judgments go beyond the scope of empirical justification per se, in chapter two we shall argue that the judgments in this level of the hierarchy admit of "quasi-empirical" justification. At the highest and final level of the hierarchy are judgments involving "abstract" reasoning whose mathematical representations are universally quantified statements of the form:

$$\forall x \forall y (x + y = y + x)$$

Hilbert and his School claimed that these judgments presuppose the existence of infinite sets — e.g., the set of natural numbers over which the universal quantifiers range. Since infinite sets are connected to the classical paradoxes, the Hilbert School claimed that the justification for this level must proceed through a consistency proof that only uses the resources of the previous two levels. One thereby grounds "abstract" mathematical reasoning in its "concrete" foundation.

Let's dig a little deeper. Because it is the level that finitists take to be indisputably justifiable from their standpoint, call the bottom level of claims the "finitistic base." That category includes claims involving closed term equations, inequalities, and bounded quantifications, and in chapter two we shall see that the reconstruction of Hilbert and his School's views entails the claim that the judgments included in the finitistic base are *a priori* justified. More specifically, what we're calling the finitistic base is a category that consists of judgements of parity and disparity between representations of sequences of strokes that are generated by finite constructions and decompositions. What makes them and the associated representations *a priori* is neither that the finite sequences are ontologically "mind-independent" nor that their representations are delivered through rational intuition. We'll argue that both of these approaches are irrelevant to the finitistic analysis of mathematics. Instead, the approach that is defended throughout the dissertation is that what makes our judgments about finite sequences *a priori* is that they possess the highest degree of justification relative to other judgments. Such judgments possess a high degree of justification because it is possible to use, *ceteris paribus*, a myriad of resources — empirical, rational, representational, etc. — to verify them. In short, judgements at the finitistic base possess a high degree of justification because of their strong evidential relationships to basic practices such as writing finitely-many figures down, counting discrete objects off forwards and backwards, picturing them to ourselves, and so on. In the sense in which these practices are "empirical" is the same sense in which the finitistic base is too. One of the tasks of this dissertation is to explain how a variable-free judgment about a mathematical proposition can be both *a priori* and empirical.

Following Sieg we'll call the next level of the hierarchy that includes free-variable claims "quasi-empirical," since the evidential relationships in which they stand go beyond the scope of those at the finitistic base. Our claim is that finitists can accept certain claims involving free-variables just in case they have evidence that the claim is consistent (in some sense) with how justifications for claims at the base are generated. The core ideas are that finitists must have a procedure that demonstrates *that* the function picked out by the claim terminates — or can be reduced to a finite number — and it must be possible to show *how* the procedure is evidentially related to claims at the finitistic base. Hence, claims at this level of the hierarchy have a lower degree of justification than claims at the finitistic base because they are *pro tanto* justified — based on the fact that the degree of justification for a free-variable claim depends on the degree of justification for variable-free claims. Hence, we have a high degree of justification for claims such as:

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

because the absence of a defeater can be verified via finite constructions and decompositions. We have a relatively lower degree for claims such as:

$$a + b = b + a$$

because it is justified just in case each instance at the finitistic base is not a defeater. But from the finitist's point of view it is not possible to know that all of its instances are not defeaters, hence quasi-empirical judgments must have a lower degree of justification than those at the finitistic base.

At the top of the hierarchy are the types of judgments that genuine mathematical practice trades on — universally quantified statements — that constitute the mathematical theories for which finitists must provide a consistency proof. It's the finitist's goal to explain the success of the class of these judgments or statements in deriving theorems using standard mathematical practices such as writing out a proof of a theorem from axioms and established theorems via inference rules. For the finitist explaining the success of a theory amounts to using a strict subset of the theory no stronger than the class of "quasi-empirical" judgments to show that the theoretical statements are "consistent." Hence, one uses classes of judgments with higher degrees of justification to ground classes with fewer relative degrees of justification. The idea is that, if a mathematical theory can be proved to be consistent — in some as-of-yet unanalyzed sense of that word — using reasoning at the finitistic base, then that theory adequately explains the facts at the finitistic base. Hence, as we'll argue in later chapters, the consistency proof through Hilbert's Program is a criterion of mathematical adequacy where a mathematical theory is adequate just in case it cannot be shown to be inconsistent relative to the set of claims about finite numbers that we take to be "self-evident," just in case it "saves the phenomena."

1.2.2 Finitism and the Adequacy Relation

Let's put aside for the moment questions about the finitistic base, quasi-empirical judgments and what belongs where in the hierarchy and hone in on the claim about mathematical adequacy. In contrast to the kind of mathematical realisms and naturalisms that claim that mathematics aims to give us a valid explanation or a literal picture of in what the mathematical world consists, the general view that's defended in this dissertation is anti-realist, with one important caveat. While we argue that mathematics aims at truth with respect to the finite base, it does *not* aim at truth concerning the kinds of quantified statements that genuine mathematical practice trades on. The basic ideas are simple. At the finitistic base, the dissertation lays out a kind of empiricism as the epistemic foundation of the finitistic standpoint. Then, we explain mathematical theories that go beyond the finitistic base as mathematically adequate to explain the facts at the finitistic base just in case the theory is "provably" consistent with those facts. On the view argued for in this dissertation, as far as epistemic belief is concerned, one accepts a mathematical theory not because its claims are all true, but because it is mathematically adequate in the sense gestured at above. And a mathematical theory is adequate just in case it cannot be shown to be inconsistent relative to the set of claims about finite numbers.

Hence, for example, if an extension of arithmetic can be shown to be consistent, in a sense of that word that we'll clarify later, then per the finitist's standpoint, it is mathematically adequate and its proof theory provides an explanation of how the claims involving finite numbers are derived from its theoretical sentences *without* requiring appeal to the literal truth (or falsity) of the axioms. If, on the other hand, a mathematical theory is not adequate, then its explanation requires appeal to the literal truth of the axioms and hence some theoretical explanation of truth. In this context, finitism succeeds if it is possible for the finitist to show which mathematical theories have theoretical sentences that refer to concepts that might be more theoretically (or ontologically or metaphysically) loaded, such as truth, and which require explicit appeal only to their own proof theory. One important consequence of our analysis is that if a theory is shown to be adequate, then it does not follow that theories that depend upon non-proof-theoretic concepts must be rejected. While the debate is often, historically, framed as one in which one is either a constructivist or else platonistic, our intention is only to show that an adequate theory is one in which the theoretical statements explain our school arithmetic experience, and that theories that go beyond this might require other explanations. But that does not preclude that explanation be a mathematical one if there is one to be had.

In what does adequacy consist? For us it's a relation between closed term equations — the "observables" or phenomena to be explained — and the generalizations from which one deduces those equations — the theoretical sentences of the theory. From the finitist's point of view, if all of the theoretical statements involving quantifiers explain for us why we believe that the grounded statements or statements at the finite base are highly justified, then that theory is adequate, and it's the consistency criterion that gives us a meaningful measure of a theory's adequacy. Hence, acceptance of a theory is epistemic insofar as one forms a belief that the theory presents its phenomena in the right way. What's novel about the approach here is that it is framed in terms of degrees of justification. If, for example, it can be shown that certain sub-theories of Peano Arithmetic can be proven to be consistent then that constitutes proof that the finitist's statements have a high degree of justification. If what's meant by consistency in such proof is that it is provable that no two sentences of some theory are negations of one another, such an interpretation of consistency yields an extremely high degree of justification. On the other hand, if a weaker sense of consistency is employed, then it yields a perhaps lower degree of justification. Hence, part of what is meant by high degree of justification gets fleshed out by what's meant by consistency. And, *mutatis mutandis*, what's meant by adequacy is dependent on the sense of consistency. Perhaps one theory is highly adequate but, from the epistemic point of view, unreasonably so in the sense that it requires too high a standard for normal mathematical practice. It's the task of the final chapter of this dissertation to tackle such questions.

Clearly our point of reference for the kind of quasi-empiricism about mathematics argued for in this dissertation is van Fraassen's constructive empiricism. On his view empirical adequacy is a relation between observable appearances and the models of a scientific theory. Roughly speaking, a theory is empirically adequate just in case its appearances are isomorphic to the empirical substructures. In such a case, the theory represents the phenomena "in the right way," and hence, if a theory is adequate, one ought to accept it — one forms the belief that the theory in question represents the phenomena and provides an explanation for why it does so. The key point is that the belief formed upon acceptance is *not* that the theory is true, but that it is adequate. Similarly, as we shall see in the last two chapters of this dissertation, perhaps we should accept a theory in which the consistency predicate differs from the usual consistency predicate because perhaps such a theory grounds a more fruitful research program, or explains the addition of ideal sentences better, and so on. Hence, while at bottom we shall argue that one accepts a mathematical theory because it is mathematically adequate, because it explains the phenomena that it is intended to explain, one's acceptance of that theory is a commitment to, as van Fraassen remarks of constructive empiricism, "the further confrontation of new phenomena within the framework of that theory, a commitment to a research programme, and a wager that all relevant phenomena can be accounted for without giving up that theory" (van Fraassen (1980), 88).

1.2.3 Looking Forward

In what remains of this chapter we turn to one of Hilbert's first programmatic attempts to analyze mathematics using his then-new axiomatic method in the Geometry (1899). The goal here is to flesh out some of the concepts above in a bit more detail by turning to one of Hilbert's most important works that predates what came to be known as Hilbert's Program in the 1920s. Indeed, above we claimed that (HP) is neutral with respect to (BD) and tried to suggest what that might mean, drawing a parallel with van Fraassen's constructive empiricism and the idea that one might accept a mathematical theory not because its axioms are true, but because it is mathematically adequate, just in case it "saves the phenomena" it represents. In the remaining sections our goal is to show that this basic thought is already present in those early works of Hilbert's that focus on a kind of proto-metamathematics of axiomatic geometry. Specifically we'll look at the concepts of completeness (*Vollständigkeit*) and consistency (*Widerspruchslosigkeit*) — the two concepts to emerge as central for (HP) in the 1920s — to see what role they play in the analysis of a mathematical theory. Our claim is that the idea that a mathematical theory is adequate to the phenomena it represents forms the backbone to Hilbert's earliest thoughts on metamathematics.

1.3 The Geometric Prologue

In late 1899, Hilbert announced the formation of what he took to be a novel approach to the foundations of mathematics, his axiomatic method. Starting around 1891, Hilbert began exploring the approach to systematically treating the informal reasoning processes and patterns implicit in the practice of mathematical theorem proving. It is both a *logical* method of investigation that aims to explicate the techniques of reasoning and the mathematical relations implicit in an informal mathematical theory, and equally important to Hilbert, the axiomatic method is an *epistemological* method of investigation, the aim of which is to explicate the premisses and proof methods — the necessary and sufficient conditions — called upon in the justification of a mathematical assertion. In explicating the conditions by which a mathematical assertion is provable, Hilbert's axiomatic method aims to show that the axioms and inference rules are, in some sense, epistemically justified. Hilbert's most mature deployment of the method during his "geometric period" is his *Grundlagen der Geometrie*, written as a *Festschrift* for the unveiling of the Gauss-Weber monument in Göttingen.[3] In the introduction, Hilbert announces that his "novel approach" consists in finding a "*simple* [*einfaches*] and *complete* [*vollständiges*] system of mutually *independent* [*unabhängiger*] axioms," that illuminate the meaning and significance of the most important geometric assertions ((1899), 1).

In chapter two of (1899) Hilbert demonstrates that his system (HG) is consistent [*widerspruchslos*] relative to the consistency of the arithmetic of the reals. In his debate with Frege, Hilbert writes that if the "arbitrarily given axioms do not contradict one another with all their consequences, then they are true and the things defined by the axioms exist" (Frege (1980), 39). Philosophical accounts of Hilbert's philosophy of mathematics stress the consistency criterion and take it to legitimate an inference from consistency to truth, and from truth to the existence of the points, lines, and planes over which HG's (German language) quantifiers range. But it might come as a surprise to learn that consistency is *not* listed in the introduction to (1899) as a goal of the axiomatic method applied to geometry, and the majority of chapter two is devoted to independence proofs for the five axiom groups.[4] What explains the incongruity between the narrow focus on consistency by contemporary readers, and the much wider range of conditions that Hilbert states? One of the reasons is the influence of Hilbert's later consistency program. But a deeper reason is that of the following four conditions — simplicity, completeness, consistency, and independence — that Hilbert lists, it is assumed that the only condition that admits of precise treatment *and* carries epistemic significance is the consistency condition.

But Hilbert claims that if all four conditions are satisfied, if a set of axioms is "compatible [*verträglich*]," roughly meaning that the set "coheres," then the mathematician is permitted to believe that the informal mathematical theory presented by the set of axioms is epistemically justified. In what follows we show that implicit in this concept of coherence are the particular metamathematical ingredients for at least three important concepts of the later proof-theoretic Hilbert Program.

[3] In what follows, we refer to the *Grundlagen der Geometrie* as (1899); likewise by date for the other works of the geometric period studied herein, all of which are found in Hallett and Majer (2004); translations from Hallett and Majer (2004) are my own.

[4] The axiom groups are: axioms of incidence (group I: 1-7); axioms of order (group II: 1-5); axiom of parallels, or "Euclid's Axiom" (group III); axioms of congruence (group IV: 1-6); and the axiom of continuity [*Stetigkeit*], or "Archimedes' Axiom" (group V).

In the next section, an analysis of Hilbert's proof, in chapter five of (1899), of Desargues' Theorem shows that nascent in his use of points and lines at infinity is the concept of conservative extensions. Section five analyzes Hilbert's concept of semantic completeness by contrasting it to Dedekind's approach. Section six develops a different concept of completeness, "input-completeness," against the background of Hilbert's appropriation of Hertz' (1894). Through an analysis of some of the central metamathematics that emerge from Hilbert's early geometric program, our goal is to suggest ways in which these properties might be better understood as they are deployed in the proof-theoretic consistency and completeness program that Hilbert proposes in the early 1920s. Hence, rather than read the finitistic consistency program back in to the geometric program, our claim is that metamathematics is better understood in terms of its emergence in Hilbert's geometric program.

1.4 Extension, Preservation, and Consistency

Consistency and independence are complements of one another. If a proposition ϕ and its negation are unprovable from (and so, independent of) a set of propositions Φ, then Φ is consistent with ϕ and consistent with $\neg\phi$. On the other hand, if Φ is consistent with ϕ and consistent with $\neg\phi$, then ϕ is independent of Φ. For Hilbert one of the most important features of the axiomatic method is that it makes it possible to present and explicate, through the axiomatic provability relation, the informal provability interrelations that obtain or fail to obtain between various kinds of mathematical beliefs or propositions. In other words, it makes it possible to distinguish between those propositions that are independent of one another, and so those whose justification is not inferential, and those propositions that are derived from the basic ones. This section presents an analysis of the consistency and independence proofs found in chapter V of (1899), Hilbert's axiomatic analysis of Desargues' Theorem, and shows that implicit in Hilbert's use of the simplicity condition in (1899) is found the nascent concept of an extension, both a proto-model-theoretic concept of domain extension and a proto-proof-theoretic concept of a conservative extension. Our analysis not only provides a clear meaning to Hilbert's demand that the axioms be "mutually independent" and "simple" but also shows how two of the metamathematical virtues that are often taken to be merely aesthetic play an objective role in the context of Hilbert's axiomatic method.

1.4.1 Hilbert (1899) and Desargues' Theorem

Projective geometry owes its birth to Renaissance efforts to represent three dimensional scenes on two dimensional canvases, and Desargues' Theorem owes its origin to Desargues' efforts to characterize the invariant properties of two dimensional representations when three dimensional scenes are spied from a different perspective. When his work was rediscovered in the 19th Century, it was found that his theorem characterized the central concept of *perspectivity*. Desargues' Theorem asserts:

(DT) if two triangles are perspective from a point, then they are perspective from a line.[5]
(Fig. 1)

In chapter V of (1899) Hilbert employs the axiomatic method to study the provability relations between (DT) and the incidence axioms. Theorem 32 is a proof that (DT) "holds [*Gültigkeit*]" under the assumption that the planar incidence axioms (I: 1-2), the full order axioms (II), Hilbert's axiom of parallels (III), and the full congruence axioms (IV) hold.[6] Theorem 33 is a proof that when the above axioms are assumed — with the *exception* of SAS congruence (IV: 6) — (DT) is "unprovable [*Nichtbeweisbarkeit*]." Both proofs proceed by constructing interpretations of the terms that, under the assumptions, either satisfy or fail to satisfy (DT). Thus, theorem 32 demonstrates that (DT) is consistent with the plane axioms of HG, whereas theorem 33 demonstrates that the negation of (DT) is consistent with those axioms; jointly, that it is independent of incidence geometry.

From the rediscovery of projective geometry in the 19th Century emerged the projective dualities, pairs of theorems one of which is obtained from the other by applying a scheme of uniform substitution — the duality theorems — for geometric terms: *point* is substitutable by *line*; *lie on* by *pass through*; *collinear* by *concurrent*; and *intersect* by *joins.* In projective plane geometry, for example, by applying the duality theorems to:

(I: 1) for any two points there is a unique line passing through both,

one obtains:

(I:1)′ for any two distinct lines, there is an unique point incident with both.

In order to obtain (I: 1)′ by the duality scheme, to a Euclidean 2-space one must add *ideal* points, points that are at an "infinite distance" in order for there to be a point at which any two lines meet, in particular for the special case when the two distinct lines in question are parallel.[7] Extending the real affine plane by adding ideal points in the interpretation makes it possible to prove theorems of projective plane geometry. But though (I: 1)′ holds in the extension, (I: 1) does *not* hold, since two distinct ideal points in the projective plane do not uniquely determine a straight line, but a set of parallels. In order for both (I: 1) and (I: 1)′ to hold, one must extend again by adding the *ideal* line, so that every Euclidean straight line meets the ideal line in exactly one point. The interpretation that contains ideal points and the ideal line is the projective completion of the affine plane.

Consider Hilbert's proof of theorem 32. On the assumption that (I: 1-2), (II), (III), and (IV) hold, Hilbert obtains the synthetic, "intuitive" interpretation of the non-logical terms. Working in the informal metatheory, the Euclidean plane, the goal is to extend the interpretation such that the

[5] Triangles $\triangle ABC$ and $\triangle A'B'C'$ are in *perspective* from a line if the extensions of the three points of corresponding sides meet in collinear points. We consider only the planar version here. Note its Π_2 form.

[6] See Hallett and Majer (2004), p. 420 for a list of changes between the two editions.

[7] The interpretation of the terms occurring in the incidence axioms that is extended by ideal points has the property that any two lines meet (the elliptic parallel property), distinct from the usual affine plane in which one works (which has the Euclidean parallel property), where some lines, i.e., parallels, do not meet.

extension satisfies (DT). Hilbert maps line segments into algebraic numbers, thereby mapping the synthetic interpretation into its algebraic coordinate interpretation, assigning points in the plane to pairs of coordinates for line segments (x, y) in the analytic plane Ω; and lines in the Euclidean plane to ordered sets of ratios of three segments $(u : v : w)$ in Ω such that either $v \neq 0$ or $u \neq 0$, defined as the set of all pairs $((x, y), (u : v : w))$ for which:

$$ux + vy + w = 0,$$

the condition for incidence of point and line. He extends the interpretation to an algebraic coordinatization of the real projective plane. Reinterpreting "point" as an ordered triple of three segments (x, y, z) in Ω that are not all zero, and reinterpreting "plane" as an ordered set of ratios of four segments $(u : v : w : r)$ in Ω, where u, v, and w are not all zero, and "line" is reinterpreted as the set of all (x, y, z) with common solutions to two linear equations, or as the "intersection" of two "planes." Hence:

$$ux + vy + wz + r = 0$$

is the reinterpretation of "incidence" of point and plane. Hilbert's coordinatization of the projective plane satisfies (DT), where ideal points and the ideal line are associated with any irrational algebraic numbers that satisfy the second equation. Our only remaining step is to demonstrate that (DT) is satisfied in the algebraic coordinatization of the affine plane. Hilbert observes that by setting $z = 0$ in the second equation, the first equation is recovered, thereby completing the proof (see fig. 2).

First, note that Hilbert does not assume that the spatial incidence axioms (I: 3-7) hold as hypotheses. In the usual affine plane interpretation, it is not possible to pick points outside of that plane and construct another plane intersecting it. Second, note that what makes it possible to pick a point outside of the original plane is an enrichment of the interpretation, the projection of an ideal point not lying in the Euclidean plane, but in the projective plane. Since, by hypothesis (IV: 6) holds, one then constructs a triangle in the projective plane that shares at least one of its vertices with a triangle in the Euclidean plane, and is congruent to the latter triangle. By adding the ideal line, one projects "infinitely far" the sides of the triangle lying in the new plane, and those of one of the triangles lying in the Euclidean plane, such that the lines meet in ideal points collinear on the ideal line, showing that (DT) is consistent with the axioms in force. By assuming that ideal points exist, one extends the domain and increases the semantic range of the interpretation, thereby making it possible to prove (DT) from the usual set of incidence axioms, here (I: 1-2) and (II)-(IV), and increasing the range of provability in the axiomatic theory. However, the addition of ideal points alone has the effect of invalidating (I:1), thereby decreasing the range of provability in the axiomatic theory. Hence, in order to recover the full range of provability for the set of axioms, one must assume that the ideal line exists in the interpretation.[8]

[8] In effect, to the metatheory one adds a set of propositions of Σ_2 form, in order to prove, in the axiomatic theory,

Hilbert's proof of (DT) in chapter five of (1899) illustrates the interplay between syntactic provability in the axiomatic theory and semantic extensions of the domain of the interpretation. Because of the presence of ideal points and the ideal line, the converse of (DT) holds, since one need only apply the duality scheme and uniformly substitute "point" for "line." One preserves the validity of the set of axioms and obtains two proofs for the price of one. In section 24 of chapter five, Hilbert assumes that (DT) holds in conjunction with the planar incidence axioms, the order axioms, Euclid's axiom, but absent the congruence axioms. By reinterpreting segments as numbers, Hilbert proves that the laws of arithmetic introduced in chapter three of (1899) are satisfied in a geometric interpretation. In this interpretation, commutative, associative, and distributive laws hold for addition, while for multiplication only the associative and distributive laws hold. Hence, the axioms in force implicitly define a "desarguean" plane geometry, which is then reinterpreted as what Hilbert baptizes a "desarguean number system," a proper substructure of the complex field, namely the substructure in which commutativity fails. Hence, the consistency of the aforementioned axioms implicitly defines an interpretation, a concrete geometry that is then reinterpreted as a concrete substructure of the complex number field. When balanced such delicate extension and contraction procedures increase the interrelations between the basic and complex beliefs captured by an axiomatic theory, making it a presentation of those beliefs that is as efficient as it is powerful.

1.4.2 *Grundthatsache* and the Ideal Method

Hilbert's distinction between basic beliefs, the *Grundthatsache*, and complex beliefs precedes but tracks his later distinction between "real" and "ideal" propositions made for the finitistic consistency program. In (1922), Hilbert calls the method of adding ideal propositions to real propositions in order to prove the consistency of extensions of the formal system the "ideal method." Much of the literature on the consistency program describes Hilbert as an instrumentalist, that ideal propositions are "meaningless in themselves" in the sense that the only propositions that are meaningful are the real propositions, and that ideal propositions are meaningful only if they shorten deductions between the real propositions.[9] In (1898c), for example, Hilbert writes that the use of ideal points and the ideal line "is entirely methodological" ((1898c), 167). One might think, then, that some evidence for this reading is to be found in Hilbert's geometric period. But this misses the central point of Hilbert's conception of the role of ideal elements. It is not, as he writes only lines later, a question of whether ideal elements exist, and hence, not a question about whether the propositions in which such terms occur are true independently of the theory in which they are asserted.[10] From Hilbert's point of view, such questions are "completely idle," for the addition of ideal elements makes it possible "to develop analytic geometry in its entire extent" (ibid.). In other words, by adding ideals one

a set of propositions of Π_2 form.

[9] See, for example, Field (1980), Kitcher (1976), and Detlefsen (1986).

[10] Zach (2005), 19*ff* argues for this point with respect to Hilbert's treatment of ideals during the later finitistic program. Also see, for a thorough exposition of the point vis-à-vis Hilbert's geometric period, Hallett (1990), 239*ff*.

completes the real line, in the sense that to each real number there corresponds an ideal point. By extending the domain, one preserves provability in the axiomatic theory: the extension ensures that no theorem provable in analytic geometry is unprovable in synthetic geometry.

Consider the proof of theorem 33 in (1899), the proof that (DT) is *not* a consequence of (I: 1-2), (II)-(III), (IV: 1-5), and (V). The missing axiom is (IV: 6), the SAS congruence axiom.[11] Hilbert's goal is to prove that (DT) is unprovable without the aid of all of the congruence axioms, in particular, the axiom that binds angles to sides in triangle constructions. Hilbert begins the construction in the ordinary Cartesian plane. First one takes an arbitrary line in the plane as an axis and assigns to it an origin O. Hilbert fixes the axis to be the horizontal x-axis and introduces half-planes relative to it, the plane above the axis and the plane below the axis. He then constructs an ellipse with semi-axes at $[1, 0]$ and $[0, 1/2]$. Fixing another point on the x-axis at $[3/2, 0]$, Hilbert constructs circles from that point that intersect the ellipse in at most two real points. Then Hilbert reinterprets the Euclidean points as positions (x, y) on the Cartesian plane. So, Euclidean lines are preserved with one exception: if a Euclidean line intersects the ellipse in two Euclidean points, then the line segment between those two points is reinterpreted as the arc of a circle through $[3/2, 0]$, where the ellipse now functions "intuitively" like a convex fish-eyed lens for lines passing through it. Consider now the following: the line through O, the point $[3/5, 2/5]$, the x-axis, and the y-axis. In the ordinary Cartesian plane, two triangles with vertices on these three lines are perspective from a line. But here linear and angle structures are not preserved, so (DT) and SAS fail (see Fig. 3).

The proof of theorem 32 demonstrates that (DT) is consistent with the aforementioned subset of HG, and theorem 33 demonstrates that the negation of (DT) is consistent with the aforementioned subset of HG. Since the spatial axioms are necessary and sufficient conditions for its proof, (DT) has "spatial content," the significance of which is that (DT) is a theorem of *plane* projective geometry ((1899), 70-1).[12] Moreover, its independence shows that any theorem provable with any of the spatial axioms is provable with (DT). Herein lies an important distinction. For Hilbert, consistency and independence are *inferential* conditions on the provability relation for a set of axioms, in the sense that if a set of axioms is inconsistent, then it is possible, *ex falso quodlibet*, to prove anything from them. Likewise, if an axiom is independent of another set of axioms, then it is not possible to prove that axiom (or its negation) from the set. On the other hand, domain extensions and the preservation of axiomatic provability are *explanatory* conditions, in the sense that if axiomatic provability is preserved come what may, the provability relation provides less of an explanation of those relations between the *Grundthatsachen* and the complex beliefs. On the other hand, the more elements by which the domain is extended tends to increase the explanatory power of the axioms,

[11] Hilbert takes SAS as an axiom, whereas Euclid attempts to prove it by a superposition argument that "lifts" the triangle out of the plane in which it lies and "moves" it on top of the given triangle. Euclid's hidden assumptions include the spatial reasoning used in lifting the triangle, and the assumption that rigid bodies are invariant under motions, i.e., that they do not stretch, or bend, when in motion.

[12] For discussion of its spatial content against the background of Hilbert's remark, in the conclusion to (1899), about the "purity of method," see Hallett (2008), 309-16.

but also tends to decrease the extent to which the syntactic provability relation is preserved. Hence, though contemporary metamathematicians often assimilate them, in Hilbert's geometric period, a conservative extension plays a distinctively different theoretical role than axiomatic consistency.

During his geometric period, Hilbert understands axiomatic mathematical theories in analogy with scientific theories. Showing that an axiomatic mathematical theory is consistent is intended to provide the mathematician with justification for believing that the axioms and their consequences are true. But his method of domain extensions to the semantic interpretation and the preservation of the provability relation has the role of providing the mathematician with an explanation of the relationship between the basic beliefs codified by the axioms and the more complex beliefs that are. or ought to be, derivable. For Hilbert, the balance between extensions of the domain, by ideals, and the preservation of provability in the base theory, the *Grundthatsache*, does not necessarily permit the mathematician to infer that because extensions by ideal elements do not invalidate any one of the axioms in the base theory, the extended theory is dispensable, that he can merely work in, or be justified in believing, the simpler set of axioms as opposed to the more complex ones. Hence, it seems that instrumentalist interpretations of the geometric and later programs, insofar as they take the conservativity of the extension over the base theory to show that the extension is in principle dispensable, are too quick. In the next two sections, our goal is to disentangle Hilbert's conceptions of completeness [*Vollständigkeit*] so that a better understanding of the range of possibilities emerges.

1.5 *Vollständigkeit*, Continuity, and Categoricity

In the last section we discussed domain extensions and the preservation of axiomatic provability and argued that unlike consistency these methods are explanatory rather than inferential. In this section we turn to Hilbert's addition, in the first French edition (1900b) and the second German edition (1903), of the *Axiom der Vollständigkeit.* The literal translation of "vollständig" is 'complete', and that of "Vollständigkeit" is 'completeness'. The *Axiom der Vollständigkeit* (VA) is absent from (1899). When it does appear in subsequent additions, it appears as (V: 2) alongside Archimedes' Axiom (V: 1) as an axiom of continuity. Intuitively, in (1899) Archimedes' Axiom ensures that to each point on the line there corresponds a real number. But still there may be real numbers that satisfy the axioms but that do not correspond to points on the line. In (1899) such points must be constructed "piecemeal," by the addition of ideal elements or by the construction of limit points, additions that provide for the "local" completeness of the constructions that require them. But the addition of (VA) in (1900b) assuages that need, since it ensures that to each real number there exists a point, and so ensures "globally" that there are sufficiently many points on the line that correspond to real numbers. In the next two sections, we argue that there are at least two senses in which Hilbert uses the concept of completeness. In this section, we isolate completeness in its "model-theoretic" sense. In the next section, we isolate its "proof-theoretic" sense.

In order to be clear about the difference between the proto-model-theoretic and proto-proof-theoretic senses of "completeness," here we'll quote the French and German versions of (VA). In (1900b), the *Axiom d'Intégrité* asserts that to the "système des points, droites et plans,"

> il est impossible d'adjoindre d'autres êtres de manière que le système ainsi généralisé forme une nouvelle géométrie où les axiomes des cinq groupes I-V soint tous vérifiés; en d'autres termes: les elements de la Géométrie forment un système d'êtres qui, si l'on conserve tous les axiomes, n'est susceptible d'aucune extension. ((1900b), p. 123)

In (1903), axiom V: 2, the *Axiom der Vollständigkeit*, is a more or less direct translation of the French, and asserts:

> Elemente (Punkte, Geraden, Ebenen) der Geometrie bilden ein System von Dingen, welches bei Aufrechterhaltung sämtlicher genannten Axiome keiner Erweiterung mehr fähig ist, d.h.: zu dem System der Punkte, Geraden, Ebenen ist es nicht möglich, ein anderes System von Dingen hinzuzufg̈en, so dass in dem durch Zusammensetzung entstehenden System sm̈tlich aufgeführ-ten Axiome I-IV, V 1 erfüllt sind. ((1903), p. 16)

The most natural and straightforward translation of (1903) takes "Erweiterung" as 'extension', and "erfüllt" as 'fulfilled' or 'satisfied', yielding:

> the elements (points, lines, planes) of geometry form a system of things that are capable of no extension [*Erweiterung*] while preserving all of the listed axioms; in other words, to the system of points, lines, and planes it is not possible to adjoin [*hinzufügen*] another system of things, so that in the resulting composite system [*Zusammensetzung entstehenden System*] the entire presentation of axioms I-IV, V: 1 are fulfilled [*erfüllt*].

Note that, at least on the surface, the axiom appears to refer to the intended interpretation, all other possible interpretations of the terms occurring in HG, and to the axioms themselves.

1.5.1 Digression on Dedekind

(VA) enforces the global completeness of all constructions carried out in the interpretation. It ensures that the real line is complete for any geometric construction, and it is this, intuitive sense of "completeness" that is Hilbert's primary concern and his primary motivation for adding (VA) to editions after (1899). In conjunction with Archimedes' Axiom, (VA) ensures that any interpretation of the axioms in the real numbers is isomorphic to the intended, intuitive geometric interpretation, and in this proto-model-theoretic sense of completeness, (VA) asserts that the axioms are categorical. In the respect in which (VA) (along with Archimedes' Axiom (AA)) asserts the continuity and completeness of the real line and the categoricity of the set of axioms, there is a conspicuous connection to Dedekind (1872). Dedekind takes line segments PQ with endpoints P, Q to consist in linearly ordered sets the elements of which are points ordered from left to right. In Dedekind's view, on a straight line it is possible to select an arbitrary point of origin O, a unit of measurement, and points such that for arbitrary line segments QR, there is a point S such that $lgth(OS) = 1$. It follows that the length of every line segment commensurate with QR corresponds to some rational number $lgth(OP)$ for some point P. Letting O be the rational number 0, by laying off line segments to the left and right, one constructs lengths of line segments, each one of which corresponds to a rational number (Dedekind (1963), 8). However, because there exist lengths that are incommensurable with a given unit of length, it is possible to obtain segments that do not correspond to any rational numbers, for example, the length of the diagonal of a unit square.

Dedekind argues that since the "straight line L is infinitely richer in point individuals than the domain R of rational numbers in number individuals," one must "improve" the set of rationals by "the creation of new numbers" such that the extended set gains the same "completeness, or as we may say at once, the same *continuity*, as the straight line" (Dedekind (1963), 9). In this context, that of comparing the set of rational numbers with the pre-theoretic concept of the straight line, Dedekind introduces his famous continuity principle. He writes that the "comparison of the domain R of rational numbers with a straight line has led to the recognition of the existence of gaps, of a certain incompleteness or discontinuity of the former, while we ascribe to the straight line completeness, absence of gaps, or continuity" (Dedekind (1963), 10). His solution is to adopt a principle that captures the "essence of continuity" that is based on his method of defining reals as "cuts," points on straight lines that partition the line into two sets of points such that every point of one set lies to the left, and every point of the other set lies to the right. Lines are continuous when:

> if all points of the straight line fall into two classes such that every point of the first class lies to the left of every point of the second class, then there exists one and only one point which produces this division of all points into two classes, this severing of the straight line into two portions. (Dedekind (1963), p. 11)

The principle ensures that the real line contains sufficiently many irrational points to satisfy our pre-theoretic conception of completeness. Dedekind argues that the continuity principle for the line holds regardless of the nature of physical space *per se*. If space is real, it need not be continuous, and if it is discovered to be discontinuous, "there would be nothing to prevent us, in case we so desired, from filling up its gaps, in thought, and thus making it continuous" (Dedekind (1963), 12). According to him, nothing about the nature of space itself forces us to adopt continuity. Rather, Dedekind regards continuity as "conceptually necessary," forced upon us by thought itself, as an explication of our pre-theoretic understanding of the completeness or continuity of the real line.

On this point, Hilbert and Dedekind appear to agree. In (1894), of the relation between irrational or ideal points on the line and (positive) real numbers, Hilbert argues that it is unnecessary to complete the line uniformly. Instead one can adjoin new points piecemeal to the "originally given system of points" such that, for the specific construction or proof in which ideals are added piecemeal, every real number corresponds to an ideal point ((1894), 35-6). The procedure outlined by Hilbert implies, for example, that every point has a square, and thus may be understood as a means of idealizing local completeness for straight edge and compass constructions. But only pages later Hilbert goes on to add a full continuity axiom that evinces an obvious point of contact with Dedekind's approach. Hilbert labels it a "continuity axiom," rather than completeness axiom:

> [i]f one has infinitely many points $P_1, P_2, \ldots, P_3$ in an ordered sequence, and all points lie to one side of a point A, then there exists one and only one point P such that every point of the sequence lies to one side of P, and no points lie between P and any other point of the sequence. P is a called a limit point. ((1894), 38)

So stated, the axiom recalls a principle that Dedekind (1872) proves to be equivalent to a form of Cauchy continuity, because of its reference to limit points. Four years later, in (1898c), Hilbert introduces three versions of the Axiom of Archimedes. The first version is the one that appears as Archimedes' Axiom in (1899); the second is a projective version of the first; but the third is a continuity principle precisely like the one above. Of the third, Hilbert writes that "in the set theorist's language" it ensures the existence of limit points in infinite point-sets, and immediately he notes the analogy between it and Dedekind's cuts [*Theorie der Dedekind'schen Schnitte*] ((1898c), 140-1). Thus, Hilbert recognizes explicitly that the third version of Archimedes' Axiom in (1898c), a descendent of the continuity axiom found in (1894), is equivalent to Bolzano-Dedekind continuity.[13]

But there seems to be a further touchstone between Dedekind and Hilbert. Following a discussion of ideal elements and the application of the projective duality theorems, in (1894) Hilbert discusses the relationship between the geometric concepts defined by the "Schema der Axiome" and the spatial properties of empirical bodies. He suggests that while we may come to know the geometric concepts experientially, axiomatic frameworks present only conceptual relationships between the "Schema der Begriffe," as opposed to empirically contingent relations ((1894), 60). In (1898c) Hilbert writes that because the strong continuity axiom had not been officially adopted, it does not follow that every real corresponds to a point.[14] To rectify the situation, he suggests one introduce ideal or irrational points, the axiom for which he calls "Cantor's Axiom" ((1898c), 166), which is clearly related to his earlier analogy to Dedekind cuts. Then he argues that the question of the "existence" of ideal points is entirely idle, and that their introduction is not forced upon us by "our empirical knowledge of the spatial properties of things" ((1898c), 167). Hence, in both (1894) and (1898c), Hilbert suggests that while continuity or completeness for the real line, and in general concepts defined by the geometric axioms themselves, might arise from knowledge of the spatial properties of empirical bodies, accepting an axiom that defines a concept is not justified by experience but by the logical relations in which the concepts stand, an obvious relation to Dedekind's point.

---

[13] See also Sieg (1997) for an analysis of the relation between Hilbert and Dedekind.

[14] Leo Corry (2006) has argued that Hilbert understands geometry as an essentially empirical science. Although it is likely that Hilbert believes that the *source* of geometric concepts is the physical world, the fact that, for Hilbert, empirical intuition plays no role in its *justification* puts Corry's conclusion in some jeopardy.

1.5.2 Some Initial Conclusions

In the literature on Hilbert's geometric programs, the relationship between Dedekind's and Hilbert's conceptions of continuity and completeness has led some to claim that, despite the fact that it superficially appears to be a metamathematical axiom (VA) is a concealed version of Dedekind continuity. (VA) appears to refer to all interpretations of HG, but Ferreirós claims that taking that to be evidence that it is metamathematical is "an anachronistic reading, too strongly informed by model-theoretic ideas that would only start to crystallize from about 1915" (Ferreirós (2008), 15).[15] Ferreirós argues that in its first incarnation in (1900a) as an axiom for the arithmetic of the real numbers, the word "erfüllt" is dispensable. He then claims that because the conjunction of the 17 axioms of (1900a) implicitly defines the set predicate *X is an Archimedean ordered field*, (VA) asserts that the set of real numbers is *maximal* with respect to a given Archimedean ordered field. According to Ferreirós, if (VA) is read as maximality condition, it parallels Dedekind's chain axiom that asserts that the set of natural numbers is the chain of a singleton closed under the successor map, which itself is as assertion of *minimality*. From the alleged parallel Ferreirós concludes that Hilbert, like Dedekind, presupposes that full "set theory belongs to the underlying logic in which the axiom system is formulated" (Ferreirós (2008), 17). In fact, Hilbert uses "Systeme" in reference to interpretations, and seems to understand axioms to be something like comprehension principles that provide sets for every concept defined by an axiom. But Ferreirós' conclusion obscures deeper differences between Hilbert and Dedekind, one of which is historical, and one of which is conceptual.

First, the historical point. In (1894) Hilbert's use of "Systeme" seems to be related to Dedekind's use. Hilbert writes that in order to find the necessary, sufficient, and independent conditions:

> one must present in a system of things [*System von Dingen*], so that to each property of things there corresponds a geometric fact and conversely, and so that for the most central of systems of things a *complete description* [*vollständiges Beschreiben*] and *ordering* of all geometric facts is possible. ((1894), 8)

Hilbert claims that the goal of an axiomatic treatment of geometry is to present a structure, in Dedekind's sense, that is isomorphic to the one that represents the standard geometric facts. But unlike Dedekind that goal is subordinate to finding the correct means of presenting in the system the provability conditions for the intuitive geometric facts. Hilbert's claim bears a closer relation to Hertz' (1894) conception of the relation between axiomatic theories and the phenomena presented by them. Hilbert read and was deeply influenced by Hertz during the same year in which he gave the lectures that form (1894). In the footnote that follows the above quote, Hilbert writes that, given the mentioned correspondence, the "system becomes a complete and simple picture [*Bild*] of geometric reality" ((1894), p. 9), an almost verbatim reproduction of the central theme of the introduction

[15] Ferreirós' point follows out a suggested reconstruction of Sieg's that Sieg calls a "reasoned though by no means unproblematic" reconstruction (Sieg (2002), 36).

to Hertz (1894).[16] At best, in (1894) Hilbert equivocates on his use of "System", using it in one and the same sentence in Dedekind's sense of a structure *for* a group of axioms and principles, and in Hertz's sense of an axiomatic presentation *of* a group of pre-theoretic phenomena, the Hertzean sense of which Hilbert calls later in (1894) his "Schema der Begriffe".

But though Hilbert equivocates in (1894), by the time he gives the (1898b) lectures, the basis for the (1898c) *Ausarbeitung*, the latter itself the basis for (1899), he is explicit about the contrast between his and Dedekind's and Cantor's approaches. After Hilbert presents three different versions of Archimedes' Axiom in (1898c), he remarks that the version that imposes Dedekind continuity "in the language of set-theorists expresses the existence of limit points for infinite point-sets" ((1898c), 141). While he draws an analogy between his and Dedekind's set-theoretic approach, nevertheless Hilbert does not align the axiomatic method with the set-theoretic approach. In the passages following this remark, Hilbert declares that the main task is to prove that the new axiom(s) are independent of all of the other axioms, and that such a proof proceeds by constructing a geometry in which the latter axioms hold, but the three versions of Archimedes' Axiom fail, thus constructing a non-Archimedean ordered field geometry. In (1899) of the following year, a continuity axiom like (VA) and the third version in (1898c) is absent, and first appears in the (1900a) lecture on the arithmetic of the reals. What explains the historical data? Our suggestion is that Hilbert was searching for a means of expressing geometric continuity that distinguishes the axiomatic method from the Dedekind-Cantor set-theoretic method in which the ontological category of set is fundamental, one that complements Hilbert's metamathematical point of view. In fact, this is precisely how it worked out, since in (1900a), (1900b), and subsequent editions of (1899), (VA) and (AA) are independent of one another, and it is *only* under their conjunction that one obtains Bolzano-Dedekind continuity.

Now the conceptual point. Recall that Dedekind's analysis of the pre-theoretic continuity of the geometric line takes the irrational points on the line to be cuts that partition the line into two strictly separate parts. On that point Hilbert and Dedekind agree. However, they do not agree about the further precisification of the principle that Dedekind sets out. For Dedekind, the continuity of the line is to be analyzed in terms of Dedekind cuts in an ordered set, according to which a pair (A_1, A_2) is a cut in a linearly ordered set.[17] Dedekind's continuity principle for a line L asserts that, given a cut (A_1, A_2) in L, either there is an $x_1 \in A_1$ such that for all $y_1 \in A_1$, $y_1 < x_1$, or there is a $x_2 \in A_2$ such that for all $y_2 \in A_2$, $x_2 < y_2$. Either A_1 contains a greatest element, or A_2 contains a least element. Real numbers are explained as a linearly ordered set $(\mathbb{R}, <)$ that contains $\mathbb{Q}$ as a densely ordered subset and that satisfies his continuity principle. Hence, Dedekind cuts serve as an *analytic* interpretation of the continuity of the geometric line to which all other, possibly synthetic, interpretations are isomorphic. Like Dedekind's analysis, Hilbert's (VA)

[16] In the next section, we'll argue explicitly that the contrast between Dedekind and Hilbert is borne out by the deeper, proto-proof-theoretic sense of completeness, what we call "input-completeness," what Sieg (pc) calls "quasi-empirical completeness," and what Hertz calls "correctness," found in Hertz (1894).

[17] More precisely, a pair (A_1, A_2) is a *cut* in a linearly order set $(S, <)$ *iff* A_1 and A_2 are non-empty, $A_1 \subseteq S$, $A_2 \subseteq S$, $(A_1 \cup A_2) = S$, and for all $x \in A_1, y \in A_2,\ x < y$.

ensures categoricity. But for Hilbert preserving the *synthetic* properties of continuity is essential. Hence, (VA) asserts that every interpretation that satisfies the other axioms is isomorphic to the standard synthetic interpretation of Euclidean geometry. While Dedekind's and Hilbert's analyses of continuity ensure that the continuity requirement is categorical, in that any interpretation for it is unique (up to isomorphism), Dedekind's fails to be complete in Hilbert's extended sense because the proposed interpretation fails to explain the "geometric facts." For Hilbert, continuity and completeness are interdependent, but separate, concepts.

Both the historical and conceptual points explain why the first version of (VA) in (1900a) appears against Hilbert's famous contrast between the axiomatic and Dedekind's "genetic" methods. Hilbert is searching for a means of expressing the completeness of the reals that makes essential reference to the axioms such that they provide a correct characterization or presentation of pre-theoretic concept of continuity. In (1898c), for example, Hilbert declares that "we are faced with the task of discussing the logical dependencies of the above axioms," and that the "task furnishes many new and fruitful ideas for the investigation of the principles of arithmetic" ((1900a), p. 1094). Hilbert goes on to deduce the "existence" of the numbers 0 and 1, which are implicitly defined by axioms I: 3 and I: 6, by pointing out that the axioms are consequences of other axioms.[18] Of course, the characterizations given by axiom I: 3 and axiom I: 6 are unique because the addition of (VA) ensures that the axioms are categorical. Although (VA) and (AA) are independent of one another, and "make no statement about the concept of convergence or about the existence of limits," they nevertheless imply Bolzano-Dedekind continuity. For Hilbert, the axiomatic presentation of completeness allows us to recognize "the agreement of our number-system with the usual system of real numbers" ((1900a), 1095). In this sense, completeness is not just the completeness of the reals, but a "hidden" higher-order categoricity condition that Hilbert elevates to an axiomatic principle.[19] In the next section, we explore the proof-theoretic sense of completeness.

1.6 *Vollständigkeit* and the Input Condition

Up to now, we have argued that Hilbert distances his axiomatic and metamathematical approach from the Dedekind-Cantor set-theoretic approach. The distance is especially evident in the contrast between his and their treatments of the concept of completeness. But what explains the motivation for his distinctive approach to that concept? In this section it is suggested that the answer lies in Hilbert's reading and appropriation of Hertz. At some point during early 1894, Hilbert read Hertz (1894) and was deeply influenced by some of its core theoretical and philosophical concerns. Most

[18] Our point puts one that Ferreirós makes in jeopardy. He argues that for Hilbert, Dedekind, Cantor, and Frege consistency is the consistency of the concept being defined (Ferreirós (2008), p. 26). Hence, according to him, that for Hilbert consistency entails existence means only that $\exists s \forall x (x \in s \leftrightarrow Con(\Phi) \wedge \Phi(x))$, where Φ is a concept. But Hilbert refers directly to the axioms and only indirectly to the concepts.

[19] "Hidden" is borrowed from Isaacson (1996) who argues that first-order arithmetical truth presupposes hidden higher-order concepts. Kreisel (1967) makes a similar point.

explicitly Hilbert appropriates Hertz' concept of a *Bild*, using the word in various places throughout his geometric period, and referring to Hertz (1894) in his own (1894), (1898c), and (1902b). Each reference to Hertz occurs in the context of Hilbert's discussion of "completeness," each instance of which Hilbert calls the *Hauptfrage* of axiomatic investigations. In (1898c), for example, Hilbert writes that with the "aid of an expression of Hertz'," the *Hauptfrage* consists in finding:

> the necessary, sufficient and mutually independent conditions which must be imposed upon a system of things, so that to every property of these things a geometric fact corresponds and conversely, in such a way that these things form a complete and simple "picture [*Bild*]" of geometric reality [*geometrischen Wirklichkeit*]. ((1898c), 2)

In order to understand Hilbert, let's turn to a brief exploration relationships between the two, beginning with a short digression on Hertz.

1.6.1 Hertz, Hilbert, *Bilder*

The introduction to Hertz (1894) is his last, posthumously published effort to spell out the relationship between empirical theories and empirical phenomena. Hertz claims that empirical theories are sets of "symbols [*innere Scheinbilder oder Symbole*]" or sentences constructed in order to predict future events from the observed past. For Hertz the sentences of an empirical theory stand in a "picturing" relationship with empirical phenomena, where a picture, or *Bild*, is a physical, or mathematical interpretation of the sentences of the empirical theory.[20] Hertz holds that the sentences — theoretical, observational, and other principles — of an empirical theory admit many different interpretations, and no one *Bild* is unique. One might assign to the sentences of an empirical theory a mathematical *Bild* consisting in a representation of a set of numbers and the relevant relations (e.g., a parallelogram to additive velocities in a material body). One might assign a physical *Bild* consisting in a representation of material bodies and the relevant relations (e.g., the observed motions of the planets to Newton's laws). Or, one might assign a mix. Since *Bilder* are representations rather than phenomena, what is the relation between theory and phenomena? Hertz proposes a criterion of "empirical correctness." He writes that the form given to the sentences "is such that the necessary consequents of the pictures in thought are always the pictures of the necessary consequents in nature of the objects pictured" (Hertz (1894), 1). Observational sentences of an empirical theory must be interpretable by *Bilder* of observed material bodies. Of course, this is to say nothing at all about the relation between an empirical theory, its *Bilder*, and the phenomena themselves other than that there *is* a relation.

Hertz goes on to transform that point into a stronger one. It is not necessary for material bodies to conform to empirical theories in any respect other than satisfying that relation. Rather:

[20] In our discussion, we employ the German word for "picture" throughout.

> [a]s a matter of fact, we do not know, nor have we any means of knowing, whether our representations [*Vorstellungen*] of things are in conformity with [the things themselves] in any other than this *one* fundamental respect. (Hertz (1894), p. 2)

Majer claims that this suggests Hertz is blocking the rejoinder, raised by realists against anti-realist objections, that in the long run empirical theories and theory-independent reality converge because, as methods of detection and measurement improve, empirical theories better track truth (Majer (1998), 237). But this is not quite right. Hertz' point is that we have no way of *knowing*, beyond knowing that there exists some relation between theory and experience, whether the correct interpretation of that relation is best construed along realist or anti-realist lines. What suggests Majer's reading is that Hertz is also clear that observation fixes the range of available *Bilder* for observational sentences. But both realists and anti-realists agree about the epistemological status of observational sentences, yet differ in their assessments of theoretical sentences, sentences that putatively describe unobservables. Realist accounts of scientific knowledge construe both observational and theoretical sentences as true, while anti-realist empiricists construe only observational sentences as true. Hertz, by contrast, takes observational sentences not as true, but as epistemically acceptable *given* our representations of material bodies (see Hertz (1894), 36). Theoretical sentences are epistemically acceptable relative to the observational sentences. Hence, Hertz's claim is that an adequate empirical theory is one that explains the phenomena and remains philosophically neutral.

Eschewing realist or anti-realist criteria for evaluating sentences, Hertz proposes three conditions that he claims are necessary and sufficient. In order for the sentences of an empirical theory to be epistemically acceptable, its *Bilder* must be (i) "logically permissible," (ii) "correct," and (iii) "appropriate" (Hertz (1894), 2). By (i) Hertz means logical consistency. Condition (iii) is the conjunction of two other conditions: "distinctiveness," and "simplicity." To Hertz, a *Bild* is *distinct* in proportion to the number of properties of material bodies that are represented in it. *Bilder* are distinct if every property of material bodies represented in them has a consequence for the observational sentences of the theory. *Bilder* are *simpler* if the number of properties represented that corresponds to properties of material bodies is greater than the number of properties represented that do *not* correspond to properties of material bodies. *Bilder* are simplest if no superfluous properties for the observational sentences are represented.[21] Hertz takes the simplicity and distinctiveness of an empirical theory's *Bilder* to be conditions for the appropriateness of the theory because they work together: the simpler an empirical theory's *Bilder* are, the less distinct they tend to be, and contrapositively. Hertz illustrates by deriving Newton's second law using his own "fundamental law" rather than Hamilton's principle of least action. For while the latter derivation is equivalent to the former, unlike Hertz', Hamilton's derivation of Newton's second law is invalid for contiguous actions,

[21] See Majer (1998), 238 for a similar characterization. Roughly, the difference between his characterization of simplicity and distinctiveness and ours is that he construes both in terms of elements in a *Bild*, rather than the properties represented by or in it, and this difference reflects, minimally, the difference between Majer's reading of Hertz as a proto-anti-realist, and our, so to speak, deflationary reading.

unless one introduces unobservable "hidden masses" (Hertz (1894), 39).[22] Hence, an empirical theory is appropriate when its *Bilder* are both maximally simple and distinct.

On the other hand, condition (ii) is a condition that governs how a scientific theory presents its observational phenomena. Of the three conditions only (iii) might appear to have little reason for being called a necessary condition for the specifically epistemic justification of a theory, since on its surface it appears to be an aesthetic criterion of evaluation. But insofar as "aesthetic" is associated with matters of taste, "merely" subjective, or secondary qualities, such an appearance ought to be resisted. First note that condition (iii) constrains the admissible extensions and contractions to the *Bilder* for theoretical sentences, but not the observational sentences, since *Bilder* for the latter are fixed by meeting condition (ii). Next note that condition (iii) provides a procedure for introducing and eliminating sentences about unobservables in relation to its explanatory power. Theoretical sentences are admissible into an empirical theory if their introduction explains but does not alter the consequence relation among the observational sentences. It is dispensable if its elimination does not increase the explanatory power of the theory and its absence does not alter the consequence relation among the observational sentences.[23] Recall from section two that Hilbert extends the domain by introducing ideal points at infinity, "mathematical unobservables," in order to increase the range of the axiomatic provability relation. One then introduces the ideal line at infinity to ensure that no other axiom is rendered invalid, thereby striking a balance.

Of the relation between an empirical theory, its *Bilder*, and the empirical phenomena, Hertz claims that no more may be said than there is a relation. In describing (ii), Hertz provides an at first negative characterization. He writes that the *Bilder* for an empirical theory are incorrect "if their essential relations contradict the relations of external things, *i.e.*, if they do not satisfy our first fundamental requirement," namely that observational sentences are interpreted by *Bilder* of observed material bodies given the "state of our present experience [...] and permitting an appeal to later and riper experience" (Hertz (1894), 2-3). Hertz' claim seems to be that the observational sentences of an empirical theory are correct if for each property and relation represented in a *Bild*, there is a corresponding observed property and relation among material bodies. Note that Hertz does not address the worry about the source of the input for our representations. The most Hertz claims is that a *Bild* is "built up" as a "result of experience" (Hertz (1894), 3). Hertz' claims should be understood as follows: *Bilder*, or the contents of our observational beliefs, result from experience, in the sense that external stimuli cause, or "build up," *Bilder*. But while the contents of the beliefs that result from this process are caused by the properties of material bodies, they are justified by the existence, if there is one, of an isomorphism between pictures whose every property corresponds to some property of the empirical phenomena. If all routes from sensory input to representation are

[22] See, for a technical discussion of this point, Simon Saunders (1998), 132*ff*.

[23] It might be objected that the interesting cases of admissibility are the borderline cases such as Hamilton's. But this objection ignores the implication that Hertz understands condition (iii) as a means of evaluation (and preference ranking) within a single theory *and* across two or more empirical theories. See Hertz (1894), 3.

isomorphic, then the empirical theory meets condition (ii).

The relation between Hertz' description of "correctness" and Hilbert's conception of what we might call "input-completeness," ought to be clear. But let's consider our three data points chronologically. In (1894) Hilbert introduces his geometry by claiming that "with respect to the appearances [*Erscheinungen*] or facts of experience [*Erfahrungsthatsachen*]" that are studied in natural science there is a distinguished class of facts that constitute the "outer form of things," and which are the special concern of geometry ((1894), 7). He says that the goal of a geometric investigation is to give order to these facts, describe the appearances that constitute this distinguished class by setting up a "Fachwerk der Begriffe" whose intended interpretation represents the "Grundthatsache," such that they stand in relation to one another through the laws of logic [*Gesetze der Logik*] (ibid.). Axiomatic theories must represent the "geometric facts," whose epistemic source lies in empirical phenomena, but whose justification does not ((1894), 8). Now the problem is to find:

> the *necessary* and *sufficient* and mutually *independent* conditions that must be imposed upon a system of things, so that to every property of these things a geometric fact *corresponds* and conversely, [in which case] a *complete description* and *ordering* of all geometric facts is possible. ((1894), 8-9)

Hilbert's marginal note makes implicit reference to Hertz. He writes that if there is a correspondence between properties represented by the system of things and the geometric facts, it follows that the axiomatic theory presents "a complete and simple *Bild* of geometric reality" (ibid.).

In (1898c) the distinction between propositions (axioms and theorems), their interpretations, or *Bilder*, and the phenomena that propositions present through the *Bilder* has crystallized.[24] He repeats the description of input-completeness in (1894) word for word, although now refers to Hertz in the text body, describing geometry as the "the most perfect natural science," and announcing that his goal is to explicate the provability relations between Euclidean, projective, and analytic geometry ((1898c), 1-2). Significant changes between this introduction and (1894) suggest a deeper connection to Hertz. First, he describes his presentation of Euclidean geometry as a "system of sentences," some of which are provable, others not. Of the latter the problem consists in finding the conditions under which extensions to the definite domain [*definierten Bereich*] that the usual Euclidean axioms describe are acceptable (ibid.). Reading these comments in light of Hertz, it appears that Hilbert's conception of the Euclidean axioms is akin to Hertz' conception of observation sentences. The problem is one of finding the conditions under which one accepts theoretical sentences that may refer to "mathematical unobservables," such as ideal points, the ideal line, or the continuity of the line. Second, Hilbert writes that his investigation is a "logical analysis of the scope of our intuition," but that nevertheless the question as to whether the origin of spatial intuition is *a priori* or empirical

[24] Hilbert's introduction to (1902b) is worded in almost precisely the same way as (1898c). We mention only that he begins (1902b) by explicitly listing the three conditions any satisfactory "Fachwerk von Begriffen" must meet: (1) completeness; (2) independence; and (3) consistency ((1902b), 1).

must "remain unanswered" (ibid.). Hence, just as for Hertz, it seems as though Hilbert's axiomatic method is an attempt to provide an adequate theory of "mathematical phenomena" that remains as neutral as possible towards philosophical problems, such as what entities are the fundamental ones, that might arise for the Dedekind-Cantor set-theoretic approach.

Above we discussed the sense in which Hilbert's use of the concept of completeness expresses the categoricity of the axioms and the continuity of the line. Our analysis suggests that, in addition, Hilbert's geometric period carves out a distinct concept of completeness. Awodey and Reck (2002) have argued that this concept is one of "relative completeness." If S is a set of sentences in a language $\mathcal{L}$, where HG is in $\mathcal{L}$, then their proposal is that:

> HG is complete relative to S if every $\phi \in S$ is provable from HG.

But this definition fails to take into account the relation between the formal theory and the intuitive facts that it represents. If, for example, in informal or intuitive geometry congruence relations between geometric figures are provably reflexive, then the axiomatic provability relation must reflect this fact in that congruence must be provably reflexive in the axiomatic presentation. Strictly speaking, then, we need two languages, the intuitive $\mathcal{L}$ and formal $\mathcal{L}'$, two theories, the intuitive EG in $\mathcal{L}$ and the axiomatic HG in $\mathcal{L}'$, and some translation $*$ from the intuitive set of sentences S and formal sentences S'. Then call HG "input-complete" if for every $\phi \in S$ that "follows from" EG, there is a $\phi* \in S'$ derivable in HG. Our analysis shows that an axiomatic theory is fully complete when its provability relation is such that every property represented in its interpretations is isomorphic to properties expressed by provable propositions in the standard interpretation, "geometric reality," a fact that explains Hilbert's requirement that completeness express categoricity as well.

1.7 Conclusion

In section four we studied domain extension and contraction procedures in the context of Hilbert's axiomatic analysis of Desargues' Theorem, and claimed that such procedures play an explanatory role in addition to an inferential role. Section five isolated the model-theoretic constraints imposed on HG by Hilbert's *Vollständigkeitsaxiom* by contrasting his approach with Dedekind's set-theoretic approach, and argued that Hilbert's analysis makes explicit hidden higher-order categoricity that the Dedekind-Cantor approach leave implicit. Section six argued that Hilbert's early proof-theoretic understanding of completeness, input-completeness, is a constraint that the provability relation in a formal system must meet in order to faithfully represent the properties and relations occurring at the intuitive or informal level. It was suggested that, like Hertz, Hilbert believes that the satisfaction of input-completeness permits mathematicians to remain as neutral as possible towards philosophical problems concerning the nature of the axiomatic theory's ontology. Each section presented a central concept of Hilbert's geometric period that, it was claimed in the introduction, bare a nascent and essential relation to modern metamathematical concepts employed in Hilbert's Program during the 1920's and beyond, such as conservativity, consistency, categoricity, and completeness. In the conclusion, we discuss some of these relationships briefly.

In (1929) Hilbert discusses four unsolved problems for finitistic proof theory, his 1920's proposal for the foundations of mathematics. Problem one discusses consistency proofs and proof-theoretic reductions for different types of mathematical theories. Problem two discusses consistency proofs and reductions for formal systems of functionals of higher types. Problem three discusses the completeness problem for formal systems of arithmetic and analysis. Problem four discusses the completeness problem for formal systems of logic. Hilbert's discussion of problem one makes it clear that reductions are to be carried out by conservativity proofs. But Hilbert does not specify whether he has in mind model-theoretic conservation, in which it is shown that given two models one can always be contracted to another, or proof-theoretic reduction, in which it is shown that given two formal theories every formula provable in one is provable in the other. In his discussion of the third problem, Hilbert points out that "the usual argument with which one shows that any two realizations of the axiom system for number theory, respectively, of analysis, must be isomorphic, do not satisfy the demands of finitistic rigor" ((1929), 231). He claims that the problem is to "transform the usual proof of isomorphism finitistically" (ibid.) in order to show that if a statement in formal arithmetic is consistent then it is provable. In his debate with Frege, Hilbert claims that consistency of a statement implies its intuitive truth, and the latter problem implies the need to show that if a statement is true, then it is provable. Hence, on the face of it there is a strand that runs through his work.

Let's reflect further on the question of the relationship between Hilbert's earlier works and the later finitist program. We've argued, especially through our analysis of Hertz and Hilbert, that in Hilbert's geometric period there is a clear sense in which an axiomatic theory contains two core features. One, a set of axioms in Hilbert's geometric period is uninterpreted with the caveat that one need not appeal to the intuitive truth of that set of axioms. The other is that although the theory is uninterpreted, it still must present its target class of phenomena correctly. In contrast to Euclidean geometry, for example, in which the axioms are specifically assumed to be "about" space or have spatial content, that a set of axioms is uninterpreted for Hilbert means, during this period of Hilbert's thought, that the axioms are not assumed to have spatial content. That a set of axioms must nonetheless present its target class correctly means, as we've argued, that the points, lines, and relations are assumed to exist and that the "correct" presentation of a set of axioms involves those axioms tracking their relationships, though that is not fixed in advance. Hence, there appear to be two "moments" to Hilbert's geometric period. One, begin with a set of uninterpreted axioms. But, two, construct the set such that its presentation allows it to reflect and be given an interpretation in the structure of the original objects and concepts. These steps are what Hilbert understands later on as the development of a formal axiomatic theory, in contrast to what would become the development of contentual finitist theories alongside them. Ravaglia describes the difference as follows:

> [t]he notion of a contentual axiomatic theory [...] is a descriptive notion. To call an

> axiomatic theory contentual is not to endorse its semantic interpretation. It is just to say that the theory comes with an interpretation. In this sense, the notion of contentual (without further qualification) is metaphysically neutral. (Ravaglia (2003), 7)

By contrast for him, a formal axiomatic theory has two features: one, an abstraction away from intuitive content; and two, an existential formulation. The latter coincides with our discussion just above of the two moments to Hilbert's geometric period, so it's now possible to characterize Hilbert's later work in terms of our argument in this chapter and Ravaglia's distinction between contentual and axiomatic theories.

Hilbert grappled with three main currents in the foundations of mathematics during the early periods of his career. One, as we have shown, was the axiomatic method as he encountered it through Dedekind. Two, Kronecker's restrictive constructivism is a clear influence (both positive and negative). And three, his encounter with Russell and Whitehead's regimented logical calculus. A full treatment of the latter two influences upon Hilbert's thought would take us far afield from this dissertation, as it is rich enough to require an independent treatment. So it must suffice to say for the moment that the development of proof theory, and in particular, finitist theory, can be linked to all three. First, his encounter with the formal axiomatic method *via* Dedekind and his own development of the subject through his consistency and completeness proofs codified his conception that mathematics as an enterprise stood on sound, if still largely unexplored, territory. Second, his encounter with Kronecker lead him to a kind of philosophical skepticism and the belief that one's analysis ought to be grounded in the *Grundthatsache* or contentual. And finally, his encounter with Russell and Whitehead led him to regiment all of it into a logical calculi in order to show that classical formal axiomatics rest on the contentual.[25] The introduction of ideal points and the ideal line increases, we showed, both the efficiency and power of the axiomatic theory, and helps explain the relationship between the *Grundthatsache* of an intuitive theory and its more complex beliefs. We might think that a conservativity proof does not show that one can "dispense with" the theory being shown conservative, as Field and Detlefsen seem to think, but that it is intended to be an explanation of the relationship between the two theories. If we understand Hilbert's later proposal for proving the completeness of a formal system in light of the dual senses of the concept that emerged from sections five and six, then it opens up the possibility of reconsidering the hidden higher-order concepts that are, or might be, implicit in the practice of proof theory. In this chapter, we hope to have shown how an analysis and better understanding of the origin of some central modern metamathematical concepts employed in Hilbert's Program in the 1920's and its post-Gödelian successors is a useful way to reconsider their contemporary philosophical and proof-theoretic roles. It's to the development of proof theory during the 1920s and beyond that we now turn.

[25] Again Ravaglia puts the point nicely: "[t]he axiomatic method under Dedekind's conception appealed to Hilbert's belief in the soundness of existing mathematics. The requirement for elementary foundations underlying Kronecker's constructive approach to mathematics appealed to Hilbert's philosophical skepticism. The logical developments of Whitehead and Russell (among others) allowed Hilbert to reconcile these beliefs" (Ravaglia (2003), 6).

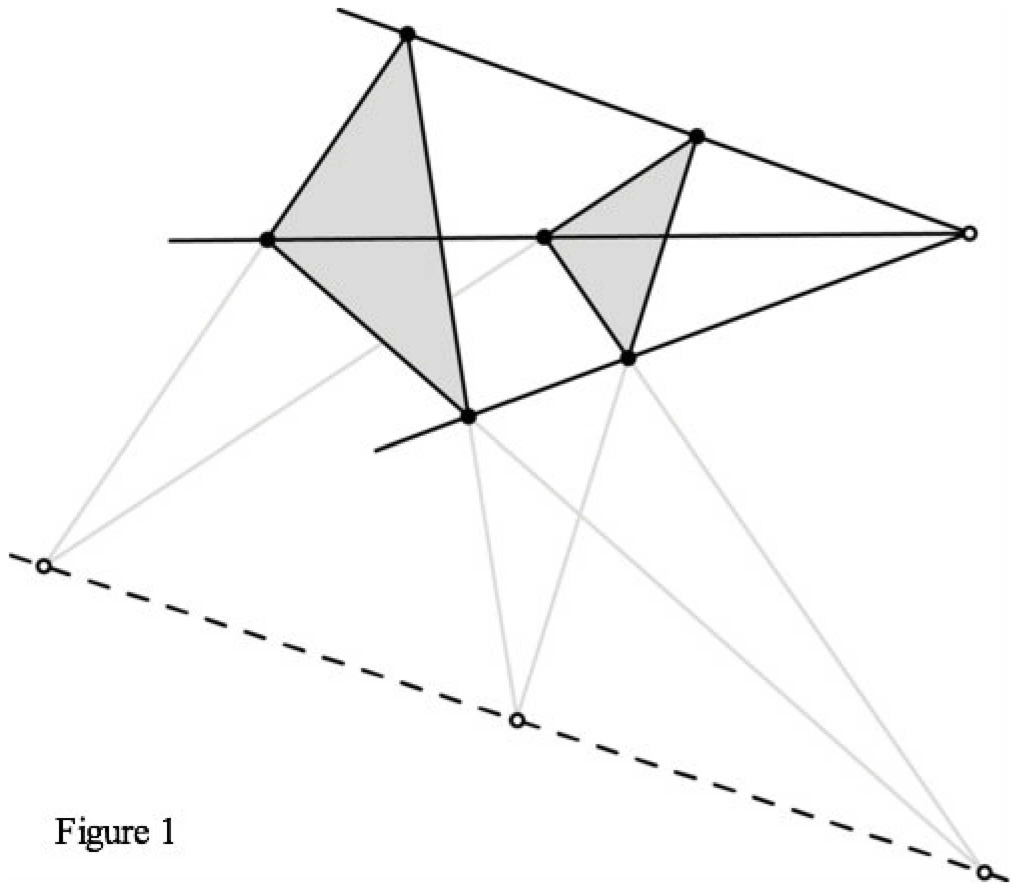


Figure 1

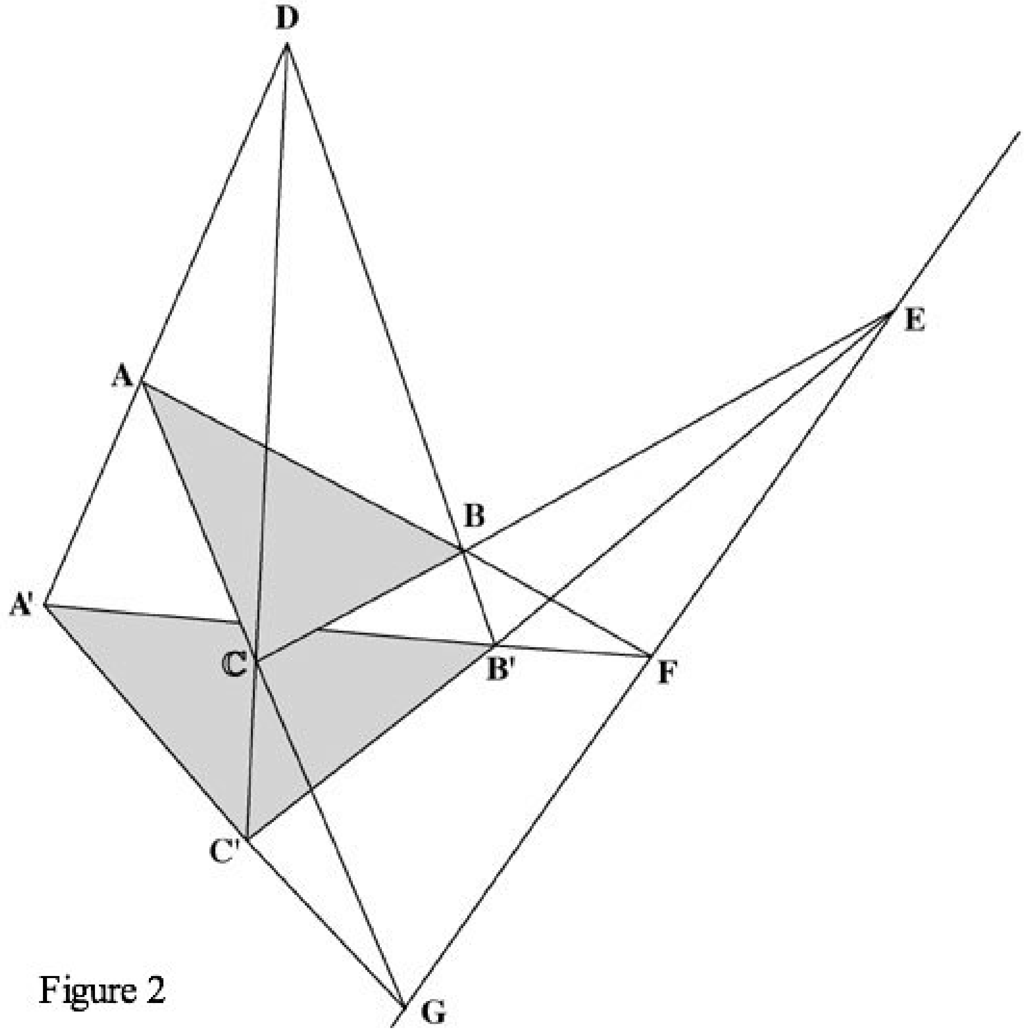
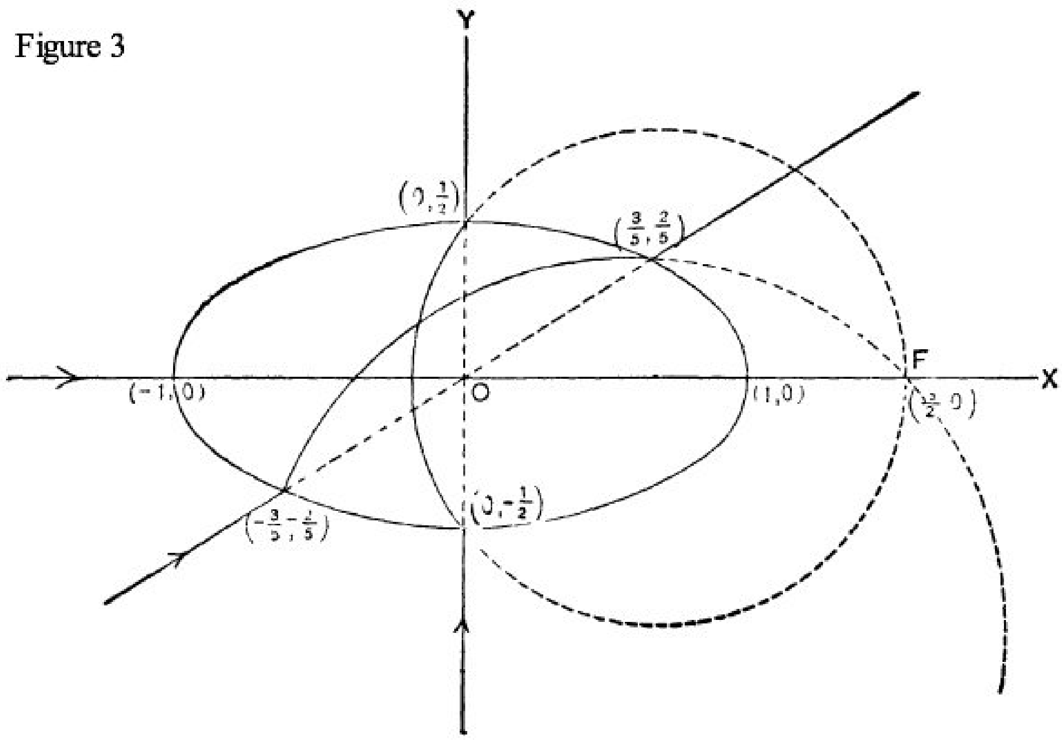


Figure 2

Figure 3



Chapter 2

Structure & Limits

2.1 Introduction

In the last chapter we argued for two main points concerning Hilbert's Program. First, by a close analysis of its inception in Hilbert's Geometric Period we argued that Hilbert's Program is meant to be neutral towards the ontological consequences in a given theory and, hence, neutral with respect to the Benaceraff Dilemma. Second, by a close analysis and reconstruction of Hilbert's early proof-theoretic practice, we argued that proof-theoretic practice explicates hidden higher order concepts and that this practice entails a kind of quasi-empiricism about extensions of the finitistic base beyond finite sequences. But of the two conclusions from the last chapter — namely, ontological neutrality and quasi-empiricism about knowledge of numbers — neither of them speak to the epistemic pedigree that finitism claims for itself. Why should we think that the finitist's proof methods are more reliable and enjoy a higher epistemic pedigree than, say, set-theoretic methods? As has become lore in the contemporary literature Hilbert's argument towards this end isolates an epistemically unquestionable fragment of mathematics, the finitistic fragment, adopted in order to answer questions about the structure of purportedly epistemically dubious fragments of mathematics. In the case of real analysis, for example, the finitist might adopt the standpoint in order to prove, using only finitistic methods, that because the set of formal axioms for non-finitistic real analysis is consistent, the beliefs codified by the axioms are justified. One might, of course, prove that a set of non-finitistic axioms is consistent from non-finitistic principles, but if the proof uses principles that are equally or more powerful than the axioms being proved consistent, then it is not clear that the proof is epistemologically significant.[1] Hence, one might adopt the finitistic standpoint because finitistic analyses of mathematical knowledge restrict the methods and principles employed to *doxa*

[1] Researchers in Hilbert's Program during the 1920's pursued a consistency proof for classical mathematics. See Kreisel (1983) and Zach (2001) for an overview. See also Zach (2006) and Feferman (1998) for an overview after Gödel's incompleteness theorems.

that are more "secure," methods of justification that are more "reliable," the beliefs of which are "presupposed by all thought and communication." Finitism enjoys its epistemic pedigree because its methods of justification satisfy constraints that are, in the finitist's view, less prone to skepticism.

2.1.1 Setting the Stage

In (1922) Hilbert stated these constraints in a well-known paragraph, one that he would repeat almost verbatim from the early 1920s onward. From the finitistic standpoint, the precondition for logical inference is that "something must be already given in representation [*Vorstellung*], certain extra-logical discrete objects" that are "intuitive," "immediate," and prior to discursive thought. Moreover, he continues, if logical inference is to be:

> certain [*sicher*], these objects must be capable of being completely surveyed in all their parts, and their presentation, their difference, their concatenation (like the objects themselves) must be given to us immediately and intuitively, as something that cannot be reduced to something else. ((1922), 202)

Finitists, Hilbert claims, represent in intuition "concrete" sequences of strokes of the form:

$$|||| = ||||$$

by representing a single stroke (or a sequence of strokes) and, by concatenations and decompositions of strokes over given representations, one generates, or "constructs," other representations. Hilbert is not much more precise about the exact structure of finitistic reasoning, but at times he characterizes its epistemic pedigree as resulting from the fact that these strokes and the elementary concatenation and decomposition operations are "representable in intuition," or "presupposed by all thought," and at other times that they constitute a distinct "source of *a priori* knowledge." Indeed it's likely that his belief that the consistency proof would be successful led him to believe that, once complete, whatever philosophical imprecisions existed in the account would be rendered inconsequential.

Owing to Hilbert's opacity assessments of the success or failure of these epistemic constraints vary widely in the literature. Kreisel (1983) argues that the defining feature of finitistic reasoning is that it is "completely surveyable," but his analysis has not been endorsed. Kitcher (1976) argues that the central concept is "representability in intuition," but rejects finitism as a viable epistemology of mathematics since infinite sequences of numerals are not representable in intuition. Like Kitcher, Tait (1980) rejects "representability in intuition" as epistemologically significant, and concludes that there is no basis for Hilbert's distinction between finitistic and non-finitistic reasoning. He argues that the distinguishing feature of finitistic reasoning is that it is "presupposed by thought" in that "it forms a nontrivial part of mathematics presupposed by any treatment of number" (Tait (1980), 540). But the moves that Tait and Kitcher make in rejecting finitistic epistemology are too quick. It might be the case that Hilbert's account of representability in intuition is unsuccessful, but epistemologists might still want to know *why* finitistic reasoning forms the "non-trivial part" of mathematical reasoning that is claimed for it. In order to be unambiguous about what is being accepted or rejected about the finitist's account, we must be able to provide an explanation for how the epistemic structure of finitistic justification delivers the kind of epistemic properties that finitists claim for it. At the same time, we must also know which, if any at all, formal arithmetical theories best characterize finitistic knowledge, and given an answer to that, which sets of formal proofs best characterize the epistemic concept of finitistic proof. For in order to know to which mathematical theories the finitist's "non-trivial" but "epistemically secure" methods of justification apply, one must be able to pinpoint both its lower and upper bounds.

2.1.2 Outline & Goals

In this chapter the goals are twofold. Our first goal is to flesh out our claim that the epistemic structure of finitist reasoning delivers methods with higher degrees of justification than those methods to which finitist reasoning is applied. In (slightly more) plain English, the successful analysis of a non-finitist arithmetic using a finitist arithmetic is epistemically significant in that it has the effect of *a priori* grounding that non-finitist arithmetic from the point of view of the finitist arithmetic. Our second goal is to attempt to spell out the upper and lower bounds for finitist arithmetic entailed by our epistemic account. In section (2.2) we argue that Hilbert and his School aligned their conception of finitistic intuition with a kind of quasi-empirical but *a priori* intuition in that it is defeasible and delivers graded or variable degrees of justification but for which there is an important sense in which it is also independent of experience. In section (2.3) we discuss lower bounds for finitist arithmetic and argue for the claim that all consistent and complete extensions of each proper subset of the theory Q constitute its lower bounds.[2] In sections (2.4) and (2.5) we deepen the epistemology and search for upper bounds for finitist arithmetic. Section (2.4) discusses the most well-known post-Hilbertian proposal to formally characterize finitism in the literature — namely, Tait's characterization of finitistic functions as exactly the primitive recursive functions — and shows that it is not epistemically coherent. Section (2.5) discusses proposals in the historical literature for higher recursive functions as finitistic — including Ackermann (1924), Hilbert (1931a), and Hilbert and Bernays (1934). In section (2.6) we exploit a method found in Hilbert and Bernays (1934) that bootstraps us through the higher recursive functions implicit in a metatheory. We conclude by arguing that while lower bounds for finitist arithmetic are extensions of subtheories of Q, its upper bound is a proper extension of Q equivalent to theories whose ordinal upper bounds are at least $\omega^{\omega^{\omega}}$, hence with bounds above PRA but below PA.

[2] Since Q itself shares the same proof-theoretic ordinal as PRA and PRA is typically identified in the literature as the lower bound for finitist arithmetic (indeed, on the standard account its upper bound as well), our claim boils down to the claim that lower bounds for finitist arithmetic are proof-theoretically weaker than typically thought.

2.2 The Epistemic Structure of the Finitistic Base

In (1926) Hilbert provides two perspicuous examples of claims that are justified from the finitistic standpoint. The first, $3 > 2$, Hilbert writes, "serves to communicate the fact that the sign 3 (that is, |||) extends beyond the sign 2 (that is, ||), or that the latter sign is a proper segment of the former" ((1926), 377). If understood as an indication of synonymy, Hilbert's "that is" is a slight mistake. He means that the proposition expressed by the claim "$3 > 2$" serves to communicate the fact constructed in intuition that one representation of strokes is a proper segment of the other representation of strokes. It is the representation that serves to justify the content of the linguistic claim formulated using Arabic numerals and the sign for inequality. Now the second example he provides is the claim that given a prime number p, there exists some prime q such that:

$$p + 1 \leq q \leq p! + 1.$$

Hilbert claims that it serves to abbreviate a disjunction of finite length:

$$q = p + 1 \text{ or } q = p + 2 \text{ or } \ \ldots \ \text{ or } q = p! + 1,$$

where q is a prime. However, the assertion that there is some prime q, such that $q > p$ is *not* finitistically justified, since "no matter how harmless the transition appears to be, there is a leap into the transfinite when this partial proposition [...] is stated as an independent assertion" ((1926), 378). First, note that for the finitist the propositional content of a belief expressed by a claim involving relations between constants has logical structure, and hence, a truth-value. But the representations that serve to justify her claims are constructed in intuition, and so are "extra-logical," or "non-conceptual." Hence, claims are truth-apt, but representations are not, since the latter cannot serve as the propositional content of a belief because they have no propositional structure. Of course, this does not entail that they have no epistemic structure. Second, note the category of beliefs the finitist characterizes as *a priori* justified. On one hand, the first example is an inequality between constants. We might expect that similar beliefs, e.g., that $5 > 4$, are also *a priori* justified, since there seems to be no principled distinction between the manner in which the two are justified. On the other hand, the second example is a claim involving an existential quantifier. But there is a numerical bound on the quantifier, so the finitist claims that it is equivalent to a finite disjunction of claims involving only numerals. Such examples suggest a natural category distinction between closed term equations, inequalities, and bounded quantificational claims, and all other claims, in particular claims that involve potentially infinite sequences or completed infinite sets. For finitists, the category of claims including closed term equations, inequalities, and bounded quantifications is unequivocally finitistically justifiable. Call that category the "finitistic base."

But what kind of intuition is at play that serves to justify the finitist's beliefs? Our options are *prima facie* twofold. Either finitistic intuition is sensuous, in the sense that some kind of perception delivers the process. Or finitistic intuition is intellectual, in the sense that some kind of rational insight delivers it. Kitcher (1976) rejects the latter view of finitistic intuition based on the familiar objection that it is utterly opaque.[3] In his view, the only option is that it is sensuous. But if so, then finitistic intuition has no hope of delivering a paradigmatically mathematical justification. For, Kitcher claims, the finitist must assume that the principle of mathematical induction is finitistic because the search for a consistency proof for formal systems requires that she believe that no formula such as $0 = 1$ can occur at any step in a proof. That is, it cannot occur at the first step, and if it does not occur at step n, then it does not occur at step $n + 1$. But if this claim is justified through intuitions of sequences such as:

$$| \cdots |,$$

where the ellipsis indicates that the sequence is indefinite, then it will no longer be possible to "survey the array" (Kitcher (1976), 112). Since it is not possible for a representation of a finite sequence to justify it, Kitcher concludes that "Hilbert's epistemology is doomed — even before we get to arithmetic" and regardless of how Hilbert's proposed consistency proof is effected by Gödel's second incompleteness theorem (Kitcher (1976), 114).[4] Grant Kitcher the claim that no finite sequence represented in intuition suffices to justify the principle of induction. Does it follow that the finitist's epistemology is doomed? No. First, if the principle of mathematical induction is finitistic, and therefore justified from the finitistic standpoint, then the justification for believing it and the reason for taking it to be finitistic must be explained independently of whatever use to which it is put. Second, while it may be true that no finite sequence constructed and represented in intuition justifies the principle of mathematical induction, it might be the case that the finitist has some other means of justifying general claims, including the principle of mathematical induction. We'll return to this thought in section (2.3). Return to the claim that finitistic representations are either sensuous or intellectual. Kitcher's claim that if it is intellectual then it must be opaque, and if it is sensuous then it is hopeless is, again, too quick. Finitistic intuition might be a kind of "hybrid" mechanism of justification that has the epistemic virtues of both sensuous and intellectual intuition but is not saddled with their vagaries.[5] It might be "rationalist" or "intellectual" in the sense that the recognition of a sequence and its relation to other sequences is immediate and discursively irreducible, but "sensuous" in the sense that it involves finite and completely surveyable processes. In the remainder of this section, let's take a closer look.

2.2.1 Representation and the Finitistic Base

Consider the following stroke-sequence:

[3] See Kitcher (1976), Wright (2004), and also Boghossian (2003), who reject the idea that "rationalist" or "intellectual" intuition can deliver any kind of justification.

[4] Also see Kitcher (1976), footnote 20.

[5] See Parsons (2008), chapter 5 and pp. 322, for a similar description. See Mancosu (1997), pp. 157-75 for a discussion of the historical development of intuition in Hilbert.

$$||||.$$

Finitists claim that its representation is "immediate," in that no further evidence justifies its individuation. One might object that there are different individuations of the sequence. One individuates each stroke, while another might individuate the leftmost and rightmost strokes as a single stroke. It might be argued, then, that its "immediacy" fails to distinguish between representations of two strokes (|| concatenated with ||), and representations of four (| concatenated with |, then |, then |). But the finitist need not claim that the representation's immediacy individuates uniquely, but that it only need deliver *some* process of concatenation or decomposition for her target class of beliefs. One might arrive at that sequence by once decomposing:

$$|||||$$

when || is considered as a single stroke, or by twice decomposing it when | is considered as a single stroke. Finitistic representations of stroke-sequences need only be invariant under some process that is sufficiently strong to deliver the relevant justification. Since some representations are invariant under concatenations and decompositions of strokes, then that class of representations should deliver the relevant justification for the target beliefs.

Recall, now, the distinction between the content of a belief and the representation, where the former possesses logical structure, while the latter possesses epistemic structure. That finitistic intuition is discursively irreducible means that representations of stroke-sequences cannot function as premises or conclusions in arguments. Hence, this kind of intuition is, in Sosa's words, a "generation faculty" and not a "transmission faculty," since rather than leading "to beliefs from beliefs already formed," it leads "to beliefs but not from beliefs" (Sosa (1991), 225). It is not possible to conclude from a belief to a finitistic representation, nor to use a finitistic representation to derive conclusions. It is *de re*, not *de dicto*. Though we might think the lack of propositional structure implies that the finitist's representations cannot play a justificatory role, that thought is too quick. It is not possible for a representation to have semantic content, but it is possible for it to be veridical. If so, then the relation between the representation in intuition and what is represented, for the finitist, need only be that it *seems* to her that her representation is sufficiently similar to the sequence represented, that there be some mapping between token representations of stroke-sequences and the stroke-sequences. Non-factive but veridical representations can play a justificatory role for the finitist because her method of justification relies on an evidential relation between representation and what's represented. Note that this implies that finitistic intuition might be defeasible.[6]

Unlike immediacy and irreducibility, surveyability seems to be a property of the construction process. Consider the class of claims characterized by the finitistic base. Left unspecified is the size of the constants involved. Suppose the finitist represents to himself a very large but finite sequence

[6] Until recently, many authors have rejected this idea, often based on the premise that mathematical propositions are necessary. But the history of what Feferman calls "monsters" suggests that it is plausible. See Feferman (2000), but also the classic Lakatos (1976), and more recently Leitgeb (2009), section three.

denoted by the numeral $10^{10^{1000}}$. Is its representation in finitistic intuition completely surveyable? If so, then the sense of possibility expressed by "surveyable" cannot be physical possibility, since no physical process can construct a sequence of that size.[7] Bernays considers this question and concludes that it is representable. He argues that if we represent the sequence:

$$||||||||||,$$

take the figure z to stand for what's represented by the sequence, and replace every $|$ by z, then we obtain "a number figure that for the purpose of communication is denoted by '$10 \times z$'" (Bernays (1930), 249).[8] The process can then be converted into a process for exponentiation for given finite sequences. The finitist can represent a finite sequence a as above and let the sequence that represents 10 correspond to the first stroke-symbol $|$ in a and apply the process to each sequence affixed to a until the stroke-sequence is exhausted. His example suggests that the surveyability of a process depends, at least in part, on an ability to recognize the identity of the processes that constitute the construction, even though there may be other processes that seem sufficiently similar but that have different results.[9]

In this section we have described the basic features of finitistic justification, and the target class of beliefs whose justification it is intended to deliver. We described how immediacy, discursive irreducibility, and complete surveyability provide finitistic intuition with the epistemic role that is claimed for it, and tried to take some of the pressure off of the charge that, as a faculty, its cognitive role is opaque. By virtue of finitistic intuition, the finitist can provide the skeptic with a process of comparing and contrasting representations constructed by elementary operations on finite sequences that serves to deliver the relevant justifications for her target class of beliefs, the finitistic base. Of course, under this description, the target class of beliefs that finitistic justification delivers is small, since it only includes claims about closed term equations, inequalities, and bounded quantifications. But now recall the claim that, because finitistic intuition is *de re* and not *de dicto*, it is possible for it to deliver non-veridical representations of stroke-sequences, and so is defeasible. In the next section, we'll take a closer look at the structure of finitistic justification, and by arguing that it is (i) weakly defeasible, and (ii) graded, show that it survives an objection from epistemology and provides a solution to a dilemma that accounts of *a priori* justification often face.

---

[7] Nor can it be practical decidability, since there may be some processes finitists characterize as completely surveyable, but that are impossible to carry out on a computer. Note that authors who restrict "surveyable" to physical surveyability are known as ultrafinitists. Authors who restrict it to practical decidability are strict finitists. For the former, see Kornai (2003); for the latter, see van Dantzig (1956), and Wright (1982).

[8] Is the finitist justified in believing that, for any n, there are n representations of sequences whose size is less than $10^{10^{1000}}$? Dummett (1975) argues that the finitist is not, and without an explanation of how the finitist moves beyond the finitistic base, his point might affect our claim in this section. See, then, section four.

[9] A related thought is to argue that by using Bernaysian descriptions involving concatenation, decomposition, and substitution to justify beliefs, finitists are implicitly committed to unbound variables. See Tait (1980).

2.2.2 Justification and the Finitistic Base

Recall Kitcher's two objections. The first objection attempts to undermine the claim that finitistic intuition can provide the requisite justification for general claims like the principle of induction, since to Kitcher finitists are limited to justifications that make use of their visual powers. If we attempt to escape the first objection by claiming that finitistic intuition is intellectual, rather than sensuous, then Kitcher claims that "the clarity of [Hilbert's] account would be entirely lost," based on the familiar objection that rationalist or "intellectual" conceptions of justification are utterly opaque (Kitcher (1976), 113). Above we characterized finitistic justification as a "hybrid" account in that although the finitist is committed to finite processes of concatenation and decomposition in her representations, those representations are immediate and discursively irreducible. In this section, we defend that claim by arguing that Hilbert and his School's general understanding of justification from the finitistic standpoint shares virtues of both accounts but few of their faults. Our strategy is to place this claim against the background of two strands of thought in mainstream epistemology. One strand argues that *a priori* justification is viable. But the other strand denies it. Our claim is that finitistic justification provides a defense against the latter objection, and a solution to a dilemma from the former.

One tradition, following Quine, has it that the justification of a mathematical claim must be "on a par" with the kind of justification available in natural science, so that mathematical claims, however simple, must be able to face the "tribunal of experience." One strand of this tradition, recently Kitcher (2000), has it that mathematical justification is "tradition-dependent," in that the mathematician's historical and social experiences of justifying his beliefs are the only reasons he has for justifying those beliefs (Kitcher (2000), 81-2).[10] Kitcher suggests that we might take this to be *a priori* justification, but doing so would render it so weak that it would make the appellation worthless. Other conceptions of *a priori* justification Kitcher rejects. In particular, he identifies two. The first:

(S) S's belief that p is *a priori* justified just in case S's belief that p is not justified by experience and it cannot be defeated by experience,

Kitcher calls the "strong view."[11] The second:

(W) S's belief that p is *a priori* justified just in case S's belief that p is not justified by experience,

Kitcher calls the "weak view," a view that appears to meet some of the objections to the strong view by dropping the indefeasibility condition (Kitcher (2000), 67-9). Above we raised the possibility that finitistic justification is defeasible, and hence falls under Kitcher's description of (W).

[10] Kitcher formulates his arguments, both for his view and against traditional conceptions of the *a priori*, in terms of process reliabilism rather than justification *per se*, but it does not seem as though anything essential hinges on it. See Kitcher (2000), 66.

[11] Kitcher associates the strong view with Kant and Frege.

Kitcher argues that the primary problem with (W) is that it does not support a version of *a priori* justification strong enough to be "tradition independent" such that it can also serve to settle debates in the foundations of mathematics (Kitcher (2000), 82). If, he argues, it were (W) that had been endorsed in these debates, then "there are possible lives, given which processes that would normally warrant belief in various mathematical propositions would fail to do so" (ibid.). But then two possibilities follow. Either our predecessors "respond to the subversive experiences by explicitly identifying certain processes as unreliable [...] or they do not" (ibid.). If the latter, then there is no way to distinguish our community from a community of clairvoyants who refuse to regard their methods of justification as unreliable. But if it is the former, then we heirs are no more justified in believing a mathematical claim than someone who has a justified belief that his perceptual beliefs have potential defeaters. Hence, heirs of a tradition are no more justified in believing simple mathematical claims than someone who knows about mirages is justified by his perception of an oasis in believing that there is an oasis in the distance, regardless of whether or not there is one. In the latter case, *a priori* justification is tradition dependent but irrationally so. In the former case, it is no "better" than *a posteriori* justification. Now the finitistic standpoint is, historically and conceptually, a response to the former, so let's pursue that line of thought.

In (1930) Hilbert recognizes that the problem with the traditional understanding of the *a priori* is that it is understood as indefeasible. He argues that the application of Riemannian geometry to physical space and the failure of the classical law of the parallelogram of velocities in relativistic physics show that "much of the knowledge that earlier was counted as *a priori* is today recognized as being not even correct" ((1930), 1162). Adopting the finitistic standpoint, Hilbert argues, is to adopt an "*a priori* attitude [*Einstellung*]" towards a class of beliefs, one that does not rule out the possibility that some might have potential defeaters. Suppose, for example, our finitist has a justified belief that he suffers from an extreme condition, "double intuition," which is like double vision in every respect except the fact that the source of the justification is distinct. In his effort to justify the belief that $1 + 1 = 2$, he has the double intuition:

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

and knows that he has the double intuition.[12] In such a case, it seems that his justified belief that he has double intuition defeats the justification for his belief that $1 + 1 = 2$. Since his justified belief that he suffers from double intuition is not an *a priori* justified belief — that is, he knows he suffers from double intuition because he experiences it — we might conclude, with Kitcher, that our unfortunate finitist has at his disposal a means of justification whose epistemic pedigree is no better than perception.

However, just because the justification for our finitist's belief is defeated by his justified true belief that he suffers from double intuition does not entail that his belief cannot be justified by

[12] Hilbert ((1926), 380) uses $2 = 3$ as a claim whose capability of being truth-apt seems to support the view that finitistic intuition might misfire in the sense described.

means *other* than finitistic intuition. For example, in his effort to justify his belief, he might rely on the testimony of other mathematicians, or he might simply count collections of objects. But this suggests an important analogy between the source of *a priori* justified belief and sources such as perception and memory. One might, for example, see two books on the dining room table and know that one suffers from double vision, but nevertheless justify the belief that two books lie on the table by remembering that one placed two books on the table. Call the defeat engendered by the finitist's knowledge that he suffers from double intuition an *undermining defeater*. Undermining defeaters depend on the source of the process of justification, in the sense that while the faculty of intuition might misfire and undermine the *a priori* justification for his belief, it does not entail that it is not possible to justify the belief by some other source of justification, such as testimony, perception, memory, or imagination. Just as the finitist's knowledge that he suffers from double intuition is an undermining defeater for the justification of beliefs formed through finitistic intuition, double vision can be an undermining defeater for the justification of beliefs formed through ordinary perception.

But between perception and what he calls the "deliverances of pure reflection," Kitcher sees only dissimilarity. He writes that the history:

> of mathematics reveals that the deliverances of pure reflection [...] do not have an impressive track record [...] instead of thinking that beliefs can be produced by processes whose normal power to warrant is well-sustained, as in the standard perceptual cases, we should view our situation with respect to the apriorist's favored knowledge-generating processes as one in which there is ample antecedent reason for doubt. (Kitcher (1997), 403)

Kitcher's claim seems to be that the history of mathematics — e.g., the parallel postulate, Russell's paradox — provides evidence that casts doubt on the claim that processes like finitistic intuition are reliable. But he makes two specious assumptions. First, he assumes that the standards for *a priori* justification need be higher than the standards whose sources lie in experience. If so, then it seems that Kitcher is implicitly holding the proponent of (W) to the standards for *a priori* justification claimed by proponents of (S). Second, Kitcher assumes that the possibility of undermining defeaters for *a priori* justification casts doubt on its ability to be a process that serves to justify belief, a claim that holds just as well for perception or memory. Hence, either Kitcher tacitly assumes that *a priori* justification must be of greater degree than empirical justification, a claim the proponent of (W) need not accept, or he shows only that the standards for *a priori* justification must satisfy the same conditions a general theory of justification satisfies.[13]

On our reconstruction of his argument, it follows that Kitcher's case against our account of finitistic justification fails. In order to show that there is reason to doubt that finitistic intuitions

[13] Compare the argument here with Casullo (2003), chapter two, esp. 42-8. The terminology of "undermining" and "overriding" is adapted from there. But the original taxonomy distinguishing between these types of defeaters is found in Pollock (1970).

serve to justify beliefs at the finitistic base, he would need to show that the standards for it are unjustified, full stop. Not only must he show that undermining defeaters are possible, but that there exist *overriding defeaters*, defeaters that serve to justify the negations of claims at the finitistic base that are justified by representations of finite sequences, so that the method of justification does not satisfy the conditions that a general theory of justification meets. Kitcher, however, doesn't entertain this possibility.[14] But a different tradition, following "rationalists" such as Descartes and Leibniz, claims that the possibility of *a priori* justification gives rise to the possibility that one might have conflicting beliefs that are *a priori* justified. BonJour argues that beliefs are *a priori* justified just in case one has "rational insight" or "rational intuition" into the necessity of the content of the belief, where an act of rational insight is direct, immediate, and non-discursive, but also rational or reason governed, "anything but brute in character" (BonJour (1995), 50). One important feature of BonJour's account, one that our account seems to share, is that rational insight need only be apparent, and not genuine, in which case an agent might be mistaken in an act of insight. It's to that that we now turn.

2.2.3 Quasi-Empiricism & the Finitistic Base

For BonJour's account as well as ours, *a priori* justification is defeasible. But the similarities seem to end there. In order to better understand the virtues of the finitistic standpoint, let's focus on just one of the problems with so-called rationalist accounts. In brief, one agent might be *a priori* justified in believing some proposition because he has the required act of insight, while another fails to have the insight, and hence withholds belief. Cases of the latter type, what BonJour calls cases of "mere disparity," are distinguished from cases of the former type, "actual conflicts." Such "disparities of insight" constitute one of the main objections to rationalism, since they demand that the rationalist show "how to resolve such conflicts in a rational way and what to say about cases where they cannot thus be resolved" (BonJour (1995), 69). BonJour claims that actual conflicts pose a more "serious challenge" to his account, since each member of the dispute knows "(i) that at least one of them is mistaken [...]; and (ii) that each of them has what seems to him a compelling reason for thinking that it is the other person that is mistaken" (ibid.). In his view, (i) and (ii) constitute empirical reasons for each member to believe that he may be mistaken, which tends to defeat the "justificatory force" of their own insights. Hence, the "correct result seems to be that neither of the competing claims is justified" (ibid.). Grant him the claim that (i) and (ii) are empirical reasons to defeat the beliefs arrived at by rational insight. Nonetheless, BonJour's resolution of actual conflict is still inadequate. On one hand, he admits that it is possible for both parties to retrench, insofar as it seems to both that their insight is insight into necessity, despite their knowledge of (i). On the other hand, because on BonJour's account insight is insight into necessity, he needs to explain how these

[14] In fact, Kitcher need only show that such defeaters justify withholding belief about such claims. However, as mentioned above, for beliefs whose contents are contained in the finitistic base, Kitcher thinks that the finitistic method of justification is perfectly fine, although he gives no argument for the claim. See Kitcher (1976), 110.

kinds of cases are resolved. For from third person points of view, actual conflicts are defeaters for the belief that at least one member of the dispute has a rational insight, but from the disputant's point of view actual conflicts are inherently irrational epistemic situations.

What tacit premises lead to BonJour's description and resolution of actual conflicts? Consider the following. Suppose that sometime prior to writing to Frege in 1902, Russell is considering Frege's Basic Law V. Russell, like Frege, is *a priori* justified in believing it, then begins to consider its implications, and soon through a chain of inferences each one of which is itself *a priori* justified, arrives at the set of all sets that is not self-membered if and only if it is. Russell then forms a belief, justified through (an apparent insight into the necessity of) the chain of inferences, in the negation of the naïve comprehension scheme. If each "step" is a rational insight, under BonJour's description Frege and Russell are in actual conflict. But we might think that Russell's belief has a greater degree of justification than Frege's, just as someone who has gone through a proof has a greater degree of justification than someone who learns by testimony that there simply exists a proof. It seems that BonJour's description of actual conflict follows from the tacit assumption that the degree of *a priori* justification for each belief must be equivalent. If that assumption is denied, then it does not follow that neither claim in a dispute is justified, since then it is possible for one *a priori* claim to be justified to a greater degree than another.[15] If this is correct, then it aligns our reconstruction of Hilbert and his School's general conception of mathematical, or more precisely finitistic, justification with a general empirical conception of justification. On its face at least, such a view fits nicely with the Hilbert School's metamathematical practice of experimentation with proof-theoretic constructions in pursuit of the consistency proof. Distinctions between finitistic and non-finitistic methods of justification arise not because non-finitistic methods are not justified, while finitistic methods are. Rather, the distinction arises because the likelihood of the existence of overriding defeaters for non-finitistic methods of justification is higher than that for finitistic methods — something vividly illustrated by, among other things, Russell's counterexample to Basic Law V.

If this description of the graded structure of finitistic justification is correct, then finitistic justification automatically induces an internal hierarchical ranking on the finitist's beliefs. Hilbert writes that besides the claims at the base "we encountered finitary propositions of problematic character, for example, those that were not decomposable," into partial propositions possessing a finite upper bound that are then analyzable into finite truth-functional expressions ((1926), 380). He calls these claims "problematic" and illustrates with the claim that, if $\mathfrak{a}$ and $\mathfrak{b}$ are numerals, then:

$$\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a},$$

the justification of which is lower in degree than for claims at the finitistic base since finitists cannot verify the property for every numeral. Hence, the hierarchy falls out naturally: while beliefs such as $0 \neq 1$ possess a higher degree of *a priori* justification because the likelihood of the existence of

[15] Compare this with Goldman (1999), and Fantl (2003).

overriding defeaters is lower, beliefs such as $10^9 \neq 10^{10}$ possess a lower degree of justification than $0 \neq 1$, although perhaps negligibly so. In turn, problematic claims possess a still lower degree of justification, since the finitist recognizes that her means of justification cannot rule out the possibility of the existence of an overriding defeater, although it might be highly unlikely. Note, though, that because finitistic justification is graded, she is never in a position to know that all and only the class of claims she has considered are finitistically justified. If the finitist's beliefs are ranked in terms of their degrees of justification, then while she knows that the limits of her reasoning are finite, she does not know *how* finite. She is epistemically ignorant of the bounds of finitistic justification. In sum, finitists need not be committed to Kitcher's assumption that the degree of *a priori* justification must be higher than experientially justified belief, nor to BonJour's claim that all *a priori* justification admits of the same degree. It is possible for there to be a type of justification that is *a priori*, defeasible, and admits of degrees. In the next section, these claims are exploited to show how we might pass beyond the finitistic base and pin down general lower bounds for finitist arithmetics.

2.3 Lower Bounds of Finitist Arithmetic

In section (2.2), we picked out the target class of beliefs (the finitistic base) that finitists regard as unequivocally justified and described the basic epistemic properties of finitistic justification (immediacy, discursive irreducibility, and complete surveyability). Then we derived two claims about finitistic justification. One argument shows that it need only satisfy the same conditions that accounts of experiential justification satisfy. The other argument shows that it also satisfies some of the epistemic conditions that the traditional understanding of *a priori* justification does because it is graded or variable. On one hand, claims at the base that express finitistic beliefs are the least likely to have overriding defeaters, and hence "more secure." On the other hand, problematic claims have a lower degree of finitistic justification, and claims involving principles such as mathematical induction still lower degrees of justification, if they do at all. But from the point of view of a practicing proof theorist, claims at the finitistic base are the least interesting, since such claims yield very little information about the mathematical theory that's being analyzed. In this section, we argue that given our account of finitistic justification, it is possible to pass beyond claims at the finitistic base from the finitist standpoint. The goal is to show that while finite representations might not provide the required justification for claims beyond the finitistic base, this does not rule out the possibility that the finitist has at her disposal some other epistemic means of justifying those claims. The recurring theme in this section and in sections (2.4) and (2.5) is that it is possible to employ "bootstrapping" methods in order to pass to lower bounds and beyond to upper bounds insofar as it is possible to justify those methods without appealing to patently non-finitist methods.

2.3.1 Problematic Claims and Lower Bounds

Let's return to problematic claims. Hilbert writes that problematic claims are "from the finitist point of view *incapable of being negated*" ((1926), 378). Hilbert argues that finitists ought to regard problematic claims as hypothetical judgments, since unlike closed term equations, they are incapable of being true or false. Suppose, for example, that the commutativity claim in section (2.2.3) is capable of being negated. Then, like closed term equations, if it is true, it must be satisfied by every numeral, and if it is false, then there must exist a counterexample. But that follows from unrestricted excluded middle when the domain of the property instantiated by the principle is the set of natural numbers (or rather, the set of stroke-sequences isomorphic to the natural numbers), the content of which is *not* finitistically justified. Hence, if the finitist is to pass beyond the finitistic base, he must be in a position to explain how problematic claims admit of finitistic justification. Niebergall and Schirn argue that no such justification exists. They write that if:

> there are only finitely many numerals, [$\mathfrak{a} + 1 = 1 + \mathfrak{a}$] can be transformed, without alteration of meaning, into a finite conjunction containing no quantifiers. (Niebergall and Schirn (1998), 292)

They conclude that "the claim that [$\mathfrak{a} + 1 = 1 + \mathfrak{a}$] is incapable of being negated can be sustained only if infinitely many numerals do exist" (ibid.). But this point seems to confuse the question of the domain of finitist statements with their justification. Finitely many numerals may exist, but from the finitist standpoint the issue is whether she is justified in believing her claims. Even if infinitely many numerals exist, the finitist cannot represent sequences of strokes whose cardinality is equipollent to the natural numbers. She might fail to believe the negation of a problematic claim because she does not have a degree of justification sufficient to believe it, regardless of the size of the domain satisfying the claim. The two questions are independent.

How, then, should the finitist explain her justification for problematic claims? Again, finitists cannot represent to themselves infinite sequences of strokes in order to justify problematic claims, a point granted to Kitcher above. But it does not thereby follow that finitists have *no* means by which to justify problematic claims when attempting to pass beyond the finitistic base. Recall that the issue even arises for finite numbers. Bernays' procedure permits finitists to justify claims about finite numbers such as:

$$10^{10^{10}} = 10^{10^{10}}$$

not by a single finitistic representation *simpliciter*, but by a kind of reflection and recombination of representations. If so, then it seems possible for her to compare the degree of justification she has for the above with the degree of justification had by claims with a high degree of justification, such as:

$$1 = 1,$$

by forming the belief that if k is a numeral, then $k-1$ applications of modus ponens are sufficient to find her way back to a belief that has a higher degree of justification. Nothing prohibits the finitist from reflecting on the graded structure of her justification, and coming to know that her belief is justified to one degree or another, then passing from it to other more complex beliefs via a series of inferences justified from the finitist standpoint. Finitistic justifications have an "internal" structure that permit finitists to pass from one, less complex belief to other more complex beliefs and vice-versa, regardless of the size of the domain satisfying the claim. Reflection on this method permits finitists to generalize from iterations and decompositions of the finite numerals to conditionals like the one above.

2.3.2 Bootstrapping to Lower Bounds

Let's now imagine a face-off between the finitist and non-finitist. Our non-finitist claims that the finitist's degree of justification for his set of finite numbers is irrelevant, since it hobbles his ability to justify mathematically interesting claims. Suppose our non-finitist writes down the claim:

$$Sx \neq 0,$$

where x is free, and challenges the finitist to provide a justification for it. Our finitist begins by considering its instances ranked in terms of her degree of justification for each one. She considers $1 \neq 0$, and knows that her degree of justification for it is greater than, although negligibly so, her degree of justification for $2 \neq 1$. Suppose, then, that she continues, successively representing to herself increasingly complex sequences of strokes that serve to justify increasingly complex beliefs. Passing to some larger but finite numeral k, she recognizes that her degree of justification for:

$$Sk \neq 0,$$

is lower than her degree of justification for:

$$S(k-1) \neq 0.$$

Is her degree of justification for $Sk \neq 0$ given her degree of justification for $S(k-1) \neq 0$ sufficient to justify the belief? Since there's no reason to believe that the finitist doesn't have second-order knowledge about her beliefs, she might reason as follows. If it's not justified, then because she was justified in believing $S(k-1) \neq 0$ there must be an overriding defeater for that given $S(k-2) \neq 0$. But if there was an overriding defeater for $S(k-1) \neq 0$, then there must be one for $S(k-2) \neq 0$ given $S(k-3) \neq 0$, etc. Hence, either $1 \neq 0$ is not justified or the existence of an overriding defeater — a counterexample — is inconsistent with the justification for $Sk \neq 0$.

Since neither of these two options is satisfactory, and granting that she has a high degree of justification for $Sk \neq 0$ given $S(k-1) \neq 0$, how do we pass from her justification for the one to her justification for the free-variable claim $Sx \neq 0$ for any x. We must provide some criterion of justification for the free-variable claim that doesn't depend on the criteria for which any of its instances are justified. Let's consider the following. Hilbert's finitist knows that the degree of justification she has for $Sk \neq 0$ for some k does not *eo ipso* provide her with a high degree of justification for $Sx \neq 0$ for any x. Again these are claims against finitist epistemology that we've granted for the moment. But our finitist is in a position to justify $Sx \neq 0$ by arguing that the free-variable claim *coheres* with claims for which she has a high degree of justification for believing, claims at the finitistic base. In other words, she is *pro tanto* justified in believing it because the degree of justification she has for its instances fails to provide her with a high degree of justification for its negation. In short, her free-variable claims *pro tanto* cohere with the claims at the finitistic base that have a very high degree of justification. If the finitist is able to bootstrap to $Sx \neq 0$ in the manner described, then it's reasonable to think that she'll be able to use a similar style of reflection in order to bootstrap to each basic arithmetical axiom:

$$\begin{gathered} Sx = Sy \rightarrow x = y \\ y = 0 \vee \exists x\,(Sx = y) \\ x + 0 = x \\ x + Sy = S(x + y) \\ x \cdot 0 = 0 \\ x \cdot Sy = (x \cdot y) + x. \end{gathered}$$

Conjoining these six axioms together with $Sx \neq 0$ results in the formal theory of arithmetic known as Q (Robinson's Arithmetic).

Our claim is that while it is possible for finitists to bootstrap to the conjunction of six of the above axioms, it is not possible for them to bootstrap to the conjunction of all seven, that is, to Q itself. First note that Q is incomplete and undecidable in Gödel's sense, but that extensions of each of its proper subtheories formed from the conjunction of some six of the axioms are complete and consistent.[16] Since Q is incomplete, it follows from Gödel's Second Incompleteness Theorem that it is only possible to show that it is consistent, and hence "justify" it in one of Hilbert's senses, from an extension of Q strictly stronger than Q. Hence in order to claim that *all* seven axioms are justified, one would need to see this fact from the point of view of a theory stronger than Q. Of course the obvious objection is that such a justification is patently circular — at best it's unreasonable to justify a theory by citing a theory that contains the theory whose justification is at issue. For a more specific example of the phenomenon we're describing, consider the following function of two arguments, $A(m, n)$, defined by:

$$A(m,n) = \begin{cases} n+1 & \text{if } m = 0 \\ A(m-1, 1) & \text{if } m > 0 \text{ and } n = 0 \\ A(m-1, A(m, n-1)) & \text{if } m, n > 0 \end{cases}$$

[16] See Tarski, Mostowski, and Robinson (1953), 51*ff*.

Suppose we want to calculate the value of this, the Ackermann function, for some small argument values. Let $m = 1$ and $n = 1$, then $A(1, 1)$ equals:

$$\begin{aligned} A(1-1, A(1, 1-1)) &= A(0, A(1, 0)) \\ &= A(0, A(0, 1)) \\ &= A(0, 2) \\ &= 3. \end{aligned}$$

But if $m = 4$ and $n = 1$, then $A(4, 1) = 65533$. For other arguments its value becomes extremely large. Ackermann's function is an example of a general but not primitive recursive function. Computations of its value terminate for any two arguments, and it is provably total by induction on pairs in a lexicographic ordering, i.e., induction up to the ordinal ω^2. One can "see" this by noticing that when $m > 0$ and $n = 0$, or m and $n > 0$, the computation results in a pair that is prior to m and n in the ordering. Thus, only a "double induction" is needed to show that the function is total.

Richard Zach claims that:

> 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. (Zach (2001), 129)

In Zach's view, general recursive functions are justified if primitive recursive functions are. On his account if she can recognize that the class of functions satisfies a description of functions using just iteration and substitution, then the finitist may pass to and then beyond the seven claims listed above. But from this, Zach concludes that *every* function defined by k-fold nested recursion is finitistic. But, as we have shown, from a criterion for justifying a single function it does not follow that the finitist is in possession of a criterion for justifying the class of functions of which it is a member. We might be *pro tanto* justified that a particular function has a value but unable to infer from that to the claim that *every* such function is finitistic. Suppose, for example, that the finitist is considering a sequence of strokes whose upper bound is a finite number less than the ordinal ω^2 at which Ackermann's function is provably total. Though she is not finitistically justified in believing that the function is total, and hence cannot know that the function has a value for all arguments, it is possible for her to consider instances of the function for given numerals m and n. Given a pair of numerals as arguments, nested recursions permit her to recognize that in calculating the value, pairs occur that are prior to the given pair of numerals in the lexicographic ordering. By arguing that her degree of justification for believing induction on pairs of numerals is higher than her degree of justification for believing that there is a defeater, she concludes that her belief that induction on pairs of numerals (or induction up to ω^2) is justified only insofar as it coheres with beliefs that have a higher degree of justification.[17] In upcoming sections, we'll argue through a close look at Hilbert and Bernays (1934) and (1939) that there is no single method of passing to higher recursions, and that because of this, the upper bounds of finitist arithmetic are vague.

2.4 Upper Limits for Finitists

Up to now we've drawn two conclusions about the structure of finitistic reasoning. First, by drawing on historical and conceptual considerations, it was argued that finitistic justification possesses a structure that provides variable degrees of justification for different claims. Hence, for example, claims about "small" finite numbers might possess a higher degree of justification relative to claims about "large" finite numbers. Second, it was argued that it is nonetheless possible for the finitist to pass beyond variable-free claims by taking claims that involve free-variables as *pro tanto* justified relative to his degree of justification for the set of his variable-free beliefs. In this chapter we further explore the scope of finitistic reasoning, for in order for Hilbert's dream of a consistency proof of classical formal systems to be identifiable as a *finitistic* consistency proof, it must be possible to obtain traction on the formal limits of the epistemic concepts that underlie the standpoint. Should we want to know, positively, whether a classical formal theory T is an extension of F, where F is a formal theory of finitistic reasoning, then we need a lower bound for finitistic reasoning, something we explored in the last section. On the other hand, should we want to know, negatively, whether classical formal theories do not prove the negation of any finitistic propositions, then we need an upper bound for finitistic reasoning. Our question is then which, if any at all, formal arithmetical theories best characterize finitistic knowledge, and given an answer to that, which sets of formal proofs best characterize the epistemic concept of finitistic proof?

Perhaps the best known proposal for finitism has been provided by Tait (1981). Tait argues that the finitistic functions are precisely the primitive recursive functions, and that, as a corollary, the class of finitistic proofs consists of exactly those proofs of theorems derivable in Skolem's primitive recursive arithmetic (PRA).[18] For Tait the informal concepts of finitistic proof and provability are made formally precise by the class of primitive recursive functions and the class of theorems derivable in (PRA). Expressed in terms of ordinals, Tait holds that the existence of each ordinal $\alpha < \omega^\omega$ is finitistically justified. Tait's analysis begins from the claim that "the 'finite' in 'finitism' means precisely that all reference to infinite totalities should be rejected" (Tait (1981), 524), and by making the concept of an "arbitrary finitistic function" precise, he concludes that each ordinal below ω^ω is finitistically justified. On the other hand, Tait argues that it is not possible for finitists to

[17] Tait (2005) rejoins Zach's argument and claims that only particular functions may be considered, not classes of all functions of that type. But it is not clear that Tait's argument hits its mark, since it seems to depend on the ontological type of object considered rather than on how the finitist might introduce functions "by description," so to speak, as Zach's does. In this respect, at least, our point is in agreement with Zach's approach.

[18] Skolem (1923) presents two formal systems, S and S_1, where S axiomatizes recursion equations for the primitive recursive functions, and S_1 is S plus quantifier-free induction.

justify the existence of *all* ordinals, thereby passing to a justification for the existence of the ordinal constituting the bound. To do so requires a means of justifying the existence not just of individual ordinals, but of all of them at once. As we saw in the last section above, however, the universal claim is just what is in question.

2.4.1 Outline & Goals

Section (2.4) presents Tait's analysis of finitistic reasoning and argues that it presupposes two different principles. In section (2.5), we argue that the two principles that Tait's analysis presupposes are incompatible with the claim that finitistic reasoning is bounded above. We then argue that of the three principles, one is false and suggest some consequences for our argument. In section (2.6) we discuss evidence for what Hilbert and Bernays call "extensions to the finitistic standpoint [*finit Standpunkt*]" in three guises. First we return to Ackermann's dissertation in (1924) for passing to each ordinal α such that $\alpha < \varepsilon_0$. Then we look at evidence from Hilbert and Bernays (1934) for passing to ordinals α such that $\alpha < \omega^{\omega^{\omega}}$. Finally we analyze Hilbert's proposal for an extension by the ω-rule from (1931a). Our claim is that the only successful proposal for an extension to finitistic reasoning is the argument contained in Hilbert and Bernays (1934). From this it follows that it is possible for finitists to justify the introduction of some ordinals α beyond ω^{ω} but not all ordinals up to ε_0. Finally we present a formal characterization of finitistic reasoning. In section (2.7), our conclusion, the significance of this claim is discussed in the context of the degree structure of finitistic epistemology.

2.4.2 Tait, Finitism, and PRA

Up to now we've seen that the basic class of claims that finitists are justified in believing are closed-term equations constructed by iterating the successor and predecessor operations over finite sequences. Moreover, it is possible for finitists to iterate those operations (and operations constructed from them) over finite sequences obtained by previous constructions. Hence, as we saw, if a finitist has the justified belief that some closed-term equation is finitistic, then the equation arrived at by taking the successor or predecessor of (one or both of) those terms is also finitistic. Call that conditional "iterated reflection" since the basic idea is that if the finitist has a description of a procedure for generating a numeral given a numeral at some stage i using the procedures at stages $k \leq i$, then the description of a procedure for generating a numeral at stage $i+1$ is also finitistic. In this section it is also argued that Tait ascribes to the finitist an epistemic principle that asserts that for some functions, if they are finitistic, then the finitist has sufficient evidence to be in a position to know that the function is finitistic. Call that conditional "luminosity" since the basic idea is that if a function meets the condition of being finitistic, then *eo ipso* the finitist has transparent justification for believing that it is. Let's label the first condition (R), and the second condition

(L).[19] Our claim in this section is that Tait is committed to (R), (L), and the claim that finitistic reasoning is bounded above.

Tait (1981) introduces the most widely-accepted argument for both upper and lower bounds of finitistic reasoning. He claims that the central question that faces finitists is the means by which they are to prove, in the sense of "justify," simple arithmetical claims such as:

$$\forall x \forall y\, (x + y = y + x)$$

without presupposing knowledge of the totality of natural numbers, the putative domain over which the formal quantifiers range. In his view, the most obvious attempts at proving such formulae involve either an infinite regress or are patently vicious circularity. Then rejecting Hilbert's claim that the epistemic superiority of finitistic reasoning lies, in part, in its "security [*Sicherheit*]," Tait claims that what is distinctive about the epistemology of finitistic reasoning is that it is "indubitable" in the "Cartesian sense" that neither is there a preferred standpoint nor is one available or possible upon which to criticize mathematics ((1981), 525). Moreover, in his view, any reconstruction of finitistic reasoning must meet the challenge of showing how finitists might prove Π_1 sentences such as the one above when the only available means are numerals and k-tuple sequences of numerals for fixed values of k.[20] In his view, finitistic proof is to be analyzed in terms of the concept of a finitistic function, a concept that is epistemically grounded in what he calls "Number," or the "form" of a finite sequence ((1981), 529), where "Number" is understood in its ordinal sense.

Tait claims that finitists cannot "recognize" the concept of a function, *a fortiori*, the concept of a finitistic function. His claim seems to be that in order for the finitist to apply a function as such to a concrete argument in order to yield a concrete value, he must be in a position to know that it is defined for all arguments. That is, the finitist must be in a position to prove a statement, the logical form of which is $\forall x \exists y\, F(x, y)$, that asserts that for each argument the function $f : A \to B$ defined by the matrix has a value. Note that the previous formula has the logical form of a Π_2 sentence. Hence, if our basic challenge is to show how finitists might prove Π_1 sentences, sentences strictly lower in the arithmetical hierarchy, and finitists are limited to consideration of numerals and constructions of finite sequences therefrom, then a finitistic proof of a Π_2 sentence is question begging. Nonetheless, although the finitist cannot recognize that f is a function from A to B, Tait argues that it is possible for the finitist to "understand it as recording the fact that he has given a specific procedure for *defining a B from an arbitrary A* or, we shall say, of *constructing* a B from an arbitrary A" ((1981), 528). Which constructions are then finitistically admissible? Tait claims that, because finite sequences are obtained from the null sequence "by iterating the operation of taking one element extensions," it is possible for finitists to obtain a numeral k by iterating

[19] Our terminology, of a condition being "luminous" or "transparent" is lifted directly from the literature in mainstream epistemology. See especially Williamson (2000), chapter four. It seems there is no reason not to use the concepts in discussions of mathematical epistemology, so the decision to do so won't be defended here.

[20] One reason that finitists must be in a position to prove Π_1 sentences is that one of the goals of a finitistic analysis of a formal classical mathematical theory is to prove that it is consistent, the assertion of which has, at least *prima facie*, the logical form of a Π_1 sentence.

the operation of taking one element extensions k times starting from the null sequence. But such an operation commits the finitist to taking, as finitistically justified, the constant function $\mathbf{0}$, the projection function, and the successor function. Hence, Tait concludes, the constructions "implicit in Number" are just those functions that are constructible from the constant function $\mathbf{0}$, the projection function, and the successor function by means of composition, pairing, and primitive recursion, i.e., the primitive recursive functions ((1981), 533). Let's take a closer look at Tait's argument for how finitists construct such functions.

In general, from the finitistic standpoint it is possible for us to construct a numeral k (or a sequence of strokes, or more generally, a sequence of elements taken from a finite alphabet) from an arbitrarily given numeral j by iterating the successor operation $k-j$ times. No matter how we spell out our conception of finitistic justification (i.e., Tait's "Number," or Hilbert's "representability in intuition"), it is clear that it must provide us with a means of justifying at least the successor function and the constant function, since we need both a starting point and a means of obtaining larger numerals from that starting point. More precisely, for an arbitrarily given numeral x_k, finitists can read:

(1) $\mathbf{0}(x_k) = 0$

(2) $S(x_k) = x_k + 1$,

(3) $U_i^k(x_1, \ldots, x_k) = x_i$,

with $1 \leq i \leq k$, as descriptions, respectively, for constructing the initial finite ordinal from a given finite ordinal, constructing the successor of a finite ordinal from a given finite ordinal, and (re)constructing the ith finite ordinal from a list of k numerals. First note that these "descriptions" are listed in order of dependence. One begins from the initial ordinal by (1), constructs the kth ordinal by (2), and by (3) ignores all the ordinals in the list of length k except the ith ordinal. Second note that the order of dependence implies that if one has justification for believing that (1) is finitistic, then since (2) is constructed by appealing to (1), (2) is finitistic. Likewise, one's justification for believing that (1) and (2) are finitistic implies that (3) is finitistic, since (3) is constructed by appealing only to (1) and (2). As finitists, we begin with evidence or justification for (1), and in virtue of it construct (2). Likewise, then we construct (3) in virtue of our evidence or justification for (1) and (2). Hence, it is possible for us to construct operations of greater complexity from operations of lesser complexity, when and only when we are "given" the operations of lesser complexity.

Tait argues that by "taking account of iteration, we obtain a finitist construction which is essential for nontrivial mathematics" ((1981), 531). Suppose we have a high degree of justification for the belief that a function g is finitistic, that we have constructed a numeral x_k, and, following Tait, that we are also given an arbitrary numeral x_n. We want to construct a function f by iterating previously constructed functions that we have a sufficiently high degree of justification for taking as finitistic. If we begin with x_k, apply g, then iterate our application of g:

$$x_k, g(x_k), g(g(x_k)), \ldots, f(x_n) = g \ldots (g(g(x_k))),$$

then we can express the construction by means of:

$$f(0) = x_k \qquad f(S(x_n)) = g(f(x_n)).$$

Such an expression is a description of a procedure for constructing $f(x_n)$ from x_k and $f(0)$ from iterations of g over itself. It gives rise to the description (or definition) of primitive recursive procedures. Hence, in this example, f is defined by a constant function, successor, and a previously given function. From our standpoint, it means that if we possess justifications for given functions, then the function produced from the justified functions is also finitistic. In general, Tait argues, finitists may read:

(4) $h(x_k, 0) \simeq f(x_k) \qquad h(x_k, S(x_n)) \simeq g(x_k, h(x_k, x_n))$

as the description of a procedure that specifies how to construct $h(x_k, x_n)$ from $f(x_k)$ using g. For example, suppose we have high degrees of justification for believing that functions (1) and (2) are finitistic, then by (4) we may construct addition, $h(x_j, x_k) = x_j + x_k$, by the functions:

$$f(x_j) = x_j \qquad g(x_j, x_k, z) = S(z).$$

That is, we describe a procedure for specifying values for:

$$x_j + 0 = x_j \qquad x_j + S(x_k) = S(x_j + x_k),$$

and thereby describe a procedure for constructing a finitistic function h from functions S and $\mathbf{0}$ that already have a high degree of finitistic justification.

Hence, for Tait, it is possible for a finitely long list of descriptions of primitive recursive procedures for which the finitist has sufficient evidence to imply that the description of a procedure that contains no descriptions other than the descriptions on his list is, in fact, a finitistic procedure. It is in this sense that Tait is committed to imputing to the finitist condition (R). If the finitist knows that all the descriptions on his list produce values when he is given arbitrary values for their arguments, and hence satisfy Tait's gloss on finitistic functions, then any description produced from those descriptions itself produces a value when given arbitrary values, and hence, is also finitistic. On Tait's view, when the finitist knows that some condition obtains, that of being a description of a procedure for producing a numeral when given an arbitrary numeral, then the description of a procedure for producing a numeral constructed from just those descriptions on his finite list obtains. But what about condition (L)? Is it true that, for Tait's finitist, if a condition obtains, that of being a description of a procedure for producing a numeral when given an arbitrary numeral, then the finitist *eo ipso* knows that it obtains? From an epistemological point of view, what we want to know is whether or not Tait's analysis of finitistic knowledge commits him to imputing to the finitist the claim that finitistic conditions, such as being a description of a procedure for producing the successor of an arbitrary numeral, are *luminous*. Is the finitist's construction of a description of a procedure sufficient evidence to justify his belief that the procedure produces a numeral when given an arbitrary numeral? In what remains we shall focus on Tait's analysis of what he calls "Number".

Tait distinguishes between the "form" of a finite sequence and the sequence itself. Examples of the latter include peals of a bell, words on a page, and cars in rush hour traffic. He claims that we "not only discern such sequences but we see them *as* sequences, i.e., as having the *form* of finite sequences" ((1981), 529). It is the form of a finite sequence that Tait calls "Number". Tait seems to believe that Number is "conceptually prior" to the content of the sequences, that one "recognizes" a finite sequence as a finite sequence and only then is it possible to determine, in the sense of count, its (ordinal or cardinal) number. If one recognizes, for example, that the columns of congested cars on the highway is finite, then is it possible to determine the number of cars in the gridlock. Hence, for Tait, if

$$||||||||$$

is recognized as a finite sequence, then it is possible to determine that there are eight strokes in the sequence. Tait's argument rests on two claims. First, he suggests that the form of a finite sequence is conceptually prior to its content (assuming that is the intended contrast). Second, he suggests that both finitists and non-finitists "recognize," or "discern" the form of a finite sequence prior to counting out its number. For the moment, let's ignore the possible ontological implications of the claim about the conceptual priority of Number over number, for in any case above we have seen what "procedures" Tait holds to be implicit in that concept. Instead, in what follows we'll focus on what Tait might mean by the claim that one "recognizes" the form of a finite sequence prior to counting out its elements. Our question, in short, is whether Tait assumes that Number — the condition that obtains when a sequence has the form of a finite sequence — is a luminous condition.

Recall that just above we saw that Tait claims that "recognizing" a sequence as a finite sequence — that is, recognizing a sequence's Number — is a sufficient condition for finitists to have justified beliefs about (descriptions of) procedures involving the constant function $\mathbf{0}$, the successor function S, composition, projection, and iteration (the latter of which leads to definition (or description) by primitive recursion). Since Tait does not fill out the description of what it might mean to recognize a finite sequence as a finite sequence except to claim that it is primitive and immediate, and since he eschews Hilbertian intuition, recognizing a finite sequence cannot be "immediate" in the sense that is often intended in conjunction with the epistemological use of intuition. Positively, Tait writes that Number is a condition under which we recognize instances of finite sequences via our recognition of their general form, and that:

> [w]e discern finite sequences as such in our everyday experience and this is what gives meaning to Number in the broad sense: it is the source of our ability to apply the number

> concept. But Number also has a purely formal content, independent of our experiences. ((1981), 531)

Since it is independent of our experiences, by Number's "purely formal content" Tait must mean the functions and procedures that are implicit in the concept, namely, the constant function, successor, composition, projection and those functions obtained by the schema of primitive recursion. But in the "broad sense" Tait must mean that feature of Number by which *we* recognize that a given procedure yields a number given a number. It seems, then, that recognizing a sequence as a finite sequence is "immediate" in the same sense that recognizing pain is immediate. One has the experience of pain, and irrespective of the mental and physical faculties by which we recognize it, our everyday experiences are the source of our abilities to apply the concept to the experience; that is, to know that we are in pain.

Tait (1986) is more explicit. He rejects the "myth of the model-in-the-sky," that, he argues, is the source of the debate between platonists and anti-platonists. Models are mathematical, and hence cannot serve as external standpoints from which to criticize or defend philosophical theses about mathematics. He writes:

> [t]he fact is there are no such models [external to mathematics]. Our grasp of such a model *presupposes* that we understand the relevant mathematical propositions and can determine the truth of at least some of them — for example, those whose truth is presupposed in the very definition of the model. ((1986), 158)

His positive claim is that one begins with an "intuitive" or pre-theoretic conception of truth and *from* that conception one develops a conception of proof. In set theory, for example, one begins with a stock of intuitive truths, brute facts — that sets are extensional, that unions, intersections, and complements of sets are sets — and these truths imply axiomatic principles that are the source of the set-theoretic conception of proof. In the case under discussion, the idea is that one begins with a set of intuitive and recognizable conditions that the numerals must meet — that there is a least element, that the successor of a numeral is a numeral, and so on — and if those conditions obtain, then one is in a position to know that the procedures described by those conditions yield a numeral given a numeral. In "being in a position to know" one only need be in a position to prove — in the sense of uniformly justify — the claim that the procedure yields a numeral given a numeral.[21] If it is then possible for finitists to prove, given a numeral, that the procedure yields a numeral, then for Tait it is finitistic knowledge.

That one begins with an intuitive conception of mathematical truth and from it one develops or derives one's conception of proof, axiomatic or intuitive, commits Tait to the claim that if a finitistic proposition is true, that is, if the condition the proposition describes obtains, then we are

[21] To "uniformly justify" a claim means that the finitist may use only methods that preserve the same degree of justification across all contexts. So, here a uniform justification for a claim is a deductive "proof," albeit limited to the methods that the finitist has at his disposal for a given procedure. It follows that his reasoning must be monotonic.

in a position to know that the condition obtains, that is, we are in a position to prove, or justify, that the condition obtains. In the case of finitistic knowledge, then, reflection upon the description of a procedure for generating a numeral given a numeral, on Tait's view, alone provides finitists with sufficient evidence to justify the belief that the procedure is finitistic when it is. In other words, Tait is committed to the claim that finitistic knowledge is luminous. But it is not as though his commitment ought to strike us as straightforwardly counterintuitive, since (L) at least might *prima facie* seem to be obvious. Suppose, for example, that the condition that the description of a procedure for producing a numeral given a numeral obtains. Then, since finitists are the subjects who produce the descriptions such that such conditions do or do not obtain, it might seem that it ought to follow that, if the procedure, in fact, produces a numeral given a numeral then the finitist must be in a position to prove — again in the sense of uniformly justify — that it does. Suppose, on the other hand, that the condition obtains but that the finitists is *not* in a position to prove that it does. Then there might be procedures that, under Tait's characterization, turn out to be finitistic procedures, but for which the finitist is not in a position to prove the claim; that is, for which the finitist is not in a position to verify that the computation of the function for a given numeral terminates in a numeral.

In this section we have seen that Tait's characterization of finitistic knowledge imputes to the finitist one principle for constructing procedures from procedures known to yield numerals, and another for obtaining knowledge about his procedures if they yield values. The first principle, (R), asserts that if the finitist knows that some procedure produces a numeral given a numeral, and perhaps, has sufficient evidence to justify this belief to a high degree for other arguments, then procedures constructed from that procedure also yield a numeral given a numeral. It is important to recognize that (R) is not a general principle but a description of a specific stage, namely the ith stage, wherein the finitist obtains sufficient evidence to justify his belief that his procedures yield numerals when given a numeral. By contrast the second principle, luminosity, is a general principle that asserts that if the condition of the description of a procedure that produces a numeral given a numeral obtains, then the finitist must be in a position to prove — in the sense of uniformly justify — the claim. Here, for example, the idea is that for a function f and a given numeral $n \neq 0$, if the construction of $f(n)$ yields a value, then the finitist must be in a position to prove — by reducing $f(n)$ to $f(n-1)$, reducing $f(n-1)$ to $f(n-2)$, and so forth — that it terminates in 0 ((1981), 532). Our goal in this section was to motivate the fact that Tait presupposes both principles in his characterization of finitistic reasoning along with the claim that it is bounded above, and, in part, to motivate the apparent plausibility of the principles independently of Tait's analysis. In section (2.5), below, it is shown that these principles are jointly incompatible.

2.5 Epistemic Limitations

In the last section we argued that global analyses of finitistic reasoning, such as Tait's, Kreisel's, and to some extent Gödel's, demand that the finitist meet conditions (R) and (L). In section (2.5.1) it is shown that these two conditions are jointly incompatible. It follows from this that it is unreasonable to insist that the finitist meet both conditions: finitists may either meet condition (R) or meet condition (L), but not both conditions. Our question then becomes, of the two conditions, which one is more reasonable to impute to the finitist? In section (2.5.2) it is argued that condition (L) is unreasonable as a condition on mathematical knowledge in general, and more specifically, finitistic knowledge, and that under certain restrictions condition (R) is an essential feature of finitistic reasoning. Section (2.5.3) looks at possible objections to the arguments in the two previous sections. In brief, the argument of (2.5.1) is that if one imputes both conditions to the finitist, then it is possible for him to construct every ordinal up to and including the ordinal that serves as the strict upper bound. In language sometimes used in the literature, conditions (R) and (L) together make it possible for him to "see" the limits of finitism "from the outside," a claim that each person in the debate denies. Then, the argument of section (2.5.2) shows that finitistic constructions need not be luminous or transparent. That is, the demand that the finitist always have a proof that a given function is a total function, from condition (L), is too strong. It implies, along with the more plausible principle, contained in condition (R), that finitists are always in a position to know that all the functions considered are total.

2.5.1 Luminosity and Reflection

Many philosophers, and many philosophers of mathematics, believe that there exists a special class of mathematical phenomena — functions, proofs, and so forth — that are transparent or luminous, in the sense, outlined above, that once one is in possession of a construction for a function or a proof, then the construction itself is all the justification required to conclude that the function is known. More precisely, if a proof is a kind of justification (whether defeasible or indefeasible is irrelevant for the purposes at hand), and a proof, at the intuitive level, shows that the proposition of which it is a proof is true, then the classical account of knowledge suggests that the construction of a proof for a proposition is sufficient to know the proposition. For the proposals for finitism discussed above (such as Tait's, Kreisel's, and so forth) the connection lies in the fact that each proposal demand that every finitistic function be transparent. Tait's proposal suggests that if the finitist "has the Number concept" then that is sufficient justification for knowing that all of the functions he constructs (up to all of the primitive recursive functions) are total. Likewise, Kreisel's proposal suggests that if the finitist can "visualize" the construction, then that is sufficient to know that all of the functions he constructs (up to all of the arithmetical functions) are total. It is the goal of this section to show that no account of finitistic reasoning can jointly endorse the claim that finitistic reasoning is transparent and endorse the claim that finitists generate their functions recursively.

Following the discussion above, we shall say that a finitistic condition obtains or fails to obtain in certain cases, where a case is individuated by an individual subject who is in a position to know that he has constructed an arithmetical function. To be in a position to know requires that the subject has gathered all of the right kinds of evidence, that the subject has done everything needed to believe the putative piece of knowledge, but it does not require that the subject be physically or psychologically capable of knowing. That is, the point of being in a position to know is not that the finitist must have learned all of the constructions for arithmetical functions, but rather that if he puts himself in a position to know that he has constructed a function by doing everything necessary to construct the function, including learning how to do so, and if there are no other (psychological or physical) obstacles in his way, then the finitist knows that he has constructed such a function. Williamson puts it as follows: "being in a position to know, like knowing and unlike being physically and psychologically capable of knowing, is factive: if one is in a position to know p, then p is true" (Williamson (2000), 95). Hence, if it came to pass that the finitist has a finitistic construction for a function, and if finitistic constructions are transparent, then it follows that he is in a position to know that he has a finitistic construction for a function. Since being in a position to know is factive, it follows that the constructed function is finitistic. Then, the converse of factivity, the luminosity principle for finitistic constructions, is that if in every case when the construction for a function is finitistic, then one is in a position to know that one has a finitistic construction for a function. Let the following conditional define luminous finitistic constructions:

(L) If the construction of a function is finitistic, then the finitist is in a position to know that the construction of the function is finitistic.

Our question is whether finitistic constructions are luminous in this sense.

Suppose, then, that a finitist is considering a sequence of functions $\phi_1, \phi_2, \phi_3, \ldots, \phi_n$, and that for some m such that $m \leq n$ in the sequence, the constructions for the functions cease to be finitistic, where the type of construction might conform to any of the analyses of finitism considered herein. Suppose further that the finitist begins to construct the functions from left to right, as it were, beginning with the zero function, $\mathbf{0}$, then constructing the successor function S, then the projection function U^k_i. As he begins to construct other functions in the sequence, his confidence in the claim that his constructions are finitistic begins to diminish. We have seen, for example, how Parsons doubts that finitists can construct functions beyond addition, how Tait doubts that finitists can construct functions beyond the primitive recursive functions, and how Kreisel doubts that finitists can construct functions beyond the functions definable in first-order arithmetic. Whatever the case may be, the point is to make it plausible that the finitist's confidence in his constructions being finitistic may vary such that he has a high degree of confidence that his construction of the zero function, $\mathbf{0}$, is finitistic, a slightly lower degree of confidence that his construction for successor, S, is finitistic, a slightly lower degree of confidence that his construction for projection, U^k_i, is finitistic, and so forth. It is as if the finitist is climbing up the Grzegorczyk hierarchy, and in scaling each rung of the ladder upwards he has slightly lower confidence that the rung he is on is finitistic than that for the lower rungs. If there is a fact of the matter — a point in the sequence at which the functions cease to be finitistic — then irrespective of where that point is (all primitive recursive functions, all first-order arithmetically definable functions, etc.), the finitist gradually changes from being in a position to know that his constructions for functions in the sequence are finitistic to not being in a position to know that his constructions for functions in the sequence are finitistic.

Consider, now, a function ϕ_i between ϕ_1 and ϕ_n and let ϕ_i be a finitistically constructed function. Suppose that at the ith stage of construction, the finitist considering the sequence knows that his construction for ϕ_i is finitistic. Since we are assuming that he *knows* that the construction is finitistic, he must have a high degree of confidence that the construction is finitistic and he must have a sufficient amount of evidence to justify his high degree of confidence that his construction for ϕ_i is finitistic. Consider, now, the $i+1$st stage of construction. At this stage of the construction, too, the finitist must have a high degree of confidence that the construction of the ϕ_{i+1st} function is finitistic. It need not be as high as his confidence that the ϕ_{ith} function is finitistic. But it cannot be the case that he has a very low degree of confidence. First, since the kind of evidence that justifies his belief for the ith stage must be at least reasonably similar to the kind of evidence that justifies his belief for the next stage, the process by which he came to know that the ith stage is finitistic must be similar enough to the process by which the $i+1$st stage is constructed to come out finitistic to warrant a reasonably high degree of belief that it is finitistic. Second, suppose that the $i+1$st stage of the construction is *not* finitistic. Then the finitist's confidence that the ith stage of construction is finitistic is not warranted, since his confidence in the construction being finitistic at the next stage is incorrect. But the kind of evidence justifying his belief for the $i+1$st stage is the same (or similar enough to) the kind of evidence justifying his belief for the ith stage. In order for the finitist to have a high degree of belief that the ith stage is finitistic, the $i+1$st stage must also be finitistic.[22] It follows that the finitist's confidence that the ith stage of construction is finitistic is evidentially based only if the $i+1$st stage is finitistic. That is:

(R_i) If one knows that the ith function is finitistic, then the $i+1$st function is finitistic.

Here (R_i) is a schema describing stages from the first function to the $m-1$st function in the process of constructing functions from other functions, and not a general principle about finitism. If, as has been argued, it is possible for finitists to construct functions of greater complexity from functions of lesser complexity — if, that is, functions may be constructed by explicit definition and by recursion — then at any point in the construction from ϕ_1 to ϕ_n in which a finitist knows that his function is finitistic it must follow that the function constructed at the next stage from functions known to be

[22] Here we need not assume that the $i+1$st stage is actually finitistic, since we are only trying to establish the conditional that if he *knows* that the ith stage is, then the $i+1$st stage is. This point forestalls the objection that the $i+1$st stage might be a function that ceases to be finitistic, but we shall return to the objection in more detail below.

finitistic at previous stages are also finitistic.

We're now in a position to ask whether finitistic constructions are luminous. Let's assume that they are luminous, that they satisfy (L). If they are, then in cases in which the function(s) constructed are finitistic, the finitist is in a position to know that his construction is finitistic. Our assumption, that is, implies that because the finitist is constructing a specific function at a specific stage, and if the function constructed at the ith stage is finitistic, then the finitist is in a position to know that the function constructed at the ith stage is finitistic. The construction of the function using finitistic means is itself the justification for believing that the function is finitistic. Hence, we have the following instance of the principle (L) for luminous constructions.

> (1_i) If the ith stage of the construction is finitistic, then the finitist is in a position to know that the ith stage is finitistic.

Now assume that:

> (2_i) At the ith stage, the function ϕ_i is finitistic.

Hence, by (1_i) and (2_i) we have the following.

> (3_i) The finitist is in a position to know that ϕ_i is finitistic.

By (R_i) and (3_i) we have that:

> (2_{i+1}) At the $i+1$st stage, the function ϕ_{i+1} is finitistic.

Suppose that the first function constructed, ϕ_1, is the function that returns zero, **0**, for each input. Nobody in the debate doubts that **0** is finitistic, if any functions are. Then, if we iterate the argument from (2_i) to (2_{i+1}) m times for values of i beginning at 1 and increasing to $m-1$ we have:

> (2_m) At the mth stage, the function ϕ_m is finitistic.

But recall that we assumed that for some $m \leq n$, the functions in the sequence cease to be finitistic. Hence, it cannot be the case that at the mth stage, the function ϕ_m is finitistic. Hence, the argument above is invalid, and it follows that either not all of $(R_1), \ldots, (R_{m-1})$ and $(1_1), \ldots, (1_{m-1})$ are true together or that the constant function **0** is not finitistic. Since the constant function is, as everyone in the debate agrees, finitistic, and because (R_i) and (1_i) are representative instances of conditions (R) and (L) above, it follows that conditions (R) and (L) are jointly incompatible. Insofar as Tait and Kreisel impute these conditions to finitistic reasoning, as we argued in section two, it follows that their analyses of finitism are inconsistent.

2.5.2 Against Luminosity

In the last section we saw that conditions (R) and (L), the underlying conditions that global analyses of finitistic epistemology impute (albeit implicitly) to the finitist, are jointly incompatible. It follows that it is unreasonable to demand that finitists meet both conditions. We might meet (R) or we might meet (L) but not both together. For if both are imputed to the finitist, then it is possible for him to construct every function (or ordinal) up to and *including* the function(s) that each member of the debate over the limits of finitism denies is constructible. That is, it would be possible for finitists to "step outside their own skin," to use a Fregean turn of phrase, and see the limits of their knowledge "from the outside." In this section we argue that finitistic knowledge does not meet (L) but that (R) is one of its core features. It is possible for a finitist to have the construction of a finitistic function but fail to be in a position to know that he has the construction of a finitistic function, and hence fail to know that he has a finitistic function. Such constructions are not luminous. Our basic goal is to show that even trivial logico-mathematical conditions can fail to be luminous, where "trivial" means to denote a class of conditions that, if they obtain, are typically thought to obtain in every possible world. If, for example, a function is constructed, then the function must have an output value for each input since that is just the definition of a function. Hence, our claim is that it is possible to "have" a finitistic construction that is defined for all arguments, but fail to know that it is a finitistic function. The section concludes with a discussion of why (R) is a core feature of finitistic knowledge.

Let's suppose, for the moment, that it is condition (R) and not (L) that is the offending condition. If (R) is false but (L) is true, then by (L) if the ϕ_{ith} function is finitistic, then the finitist knows that the ϕ_{ith} function is, and since (R) is false, the ϕ_{i+1st} function fails to be finitistic. Recall that if a function is finitistic, then it must have been constructed by specifying how each input yields a unique output. For example, in the last chapter we learned from Bernays that given a sequence of strokes as input:

$$||||||||||$$

the exponentiation function with 10^{10} representing the output is constructed finitistically first by making ten copies of the ten member stroke-sequence in order to represent the product of 10 and 10, and second by iterating the latter process 10 times in order to represent raising 10 to the 10. Let ϕ_i be the exponentiation function above and continue to suppose our finitist is evaluating the function with $x = 10$ and $y = 10$ as values for the independent variables. We saw above that if the finitist knows that ϕ_i is finitistic, then he must have a high degree of confidence that it is. The process by which ϕ_i is constructed for given numerals begins by concatenating then adding each stroke contained in the the ten member stroke-sequence in some way consistent with any of the proposed analyses of finitism (including our own in the previous chapter). Then, by substituting the result for ten copies of itself, we arrive at the function for multiplication with the number of substitutions and the original number of strokes as arguments. By iterating the last process ten times, we arrive at the function for exponentiation with the number of iterations of substitutions and the original number of substitutions as arguments. Since we assumed that (L) is true, it is immediate (by *modus ponens*) from the process just described that the finitist knows that exponentiation, the ϕ_{ith} function, is finitistic.

But now let ϕ_{i+1} be a function with multiple exponents, e.g., x^{y^w}, and suppose that our finitist is considering the values for the independent variables at $x = 10$, $y = 10$, and $w = 10$. At the ϕ_{ith} function, the finitist had a sufficient degree of confidence to ensure that he knows that the ϕ_{ith} function is finitistic. Hence, if the ϕ_{i+1st} function fails to be finitistic, then it must be the result of one of two things. Either the process by which the ϕ_{i+1st} function is constructed is sufficiently distinct from the process by which the ϕ_{ith} function is constructed for it to fail to be finitistic, or the means by which the finitist believes that ϕ_i is finitistic change drastically between the ith and the $i+1$st stage. Immediately we can rule out the first disjunct. For it is possible for the process by which the value $10^{10^{10}}$ is constructed to be an instance of copying and substituting as above. We can therefore find a process sufficiently similar to the process by which the value 10^{10} is constructed that the process of construction cannot be the reason that the ϕ_{i+1st} function fails to be finitistic. Hence, suppose that the means of belief change drastically. Our finitist might, for example, believe that ϕ_i is finitistic on the basis of abstracting from non-inferential representations of values for the function (given different arguments), but believe that ϕ_{i+1} is finitistic on an inferential basis. In such a case, the finitist's false belief at the $i+1$st stage might be consistent with his knowledge at the ith stage. Unlike the argument for (R_i) in section (5.1), here the finitist's reliability at either the ith or the $i+1$st stage cannot be put in question, since his non-inferential and inferential beliefs might be otherwise reliable, yet his inferential beliefs fail to track the relevant process. What we need, in effect, is a means to ensure that changes in the means of belief between stages of construction do not produce false beliefs at different stages. For the moment, let's flag the concern, returning to it in sections (6.1)-(6.3). For now it is enough to note that it is possible to choose a single means of belief (e.g., non-inferential) in which changes are sufficiently gradual to find a counterexample to (L).

To sum up. If we assume for the moment that changes in the means of belief between stages do not produce false beliefs at different stages, and since single means of belief may be sufficiently gradual to ensure that underlying changes in the basis for belief do not entail unreliability at further stages of construction, then it cannot be (R) that is the offending condition. Hence, it must be (L) that is the offending condition. Hence, luminosity fails for finitistic constructions. Having, being in possession of, or being given the construction of a finitistic function is not a sufficient condition for knowing that the function in question is finitistic. It is possible for a function to be finitistic, for the function to be constructed finitistically, but for the finitist who is actively considering the matter to fail to know that the function is finitistic. If a finitist constructs a finitistic function, he is never guaranteed to *eo ipso* know that the function in question is finitistic. His ability to construct finitistic functions outruns his knowledge of them. Of course, it is just this point that suggests that finitists, contra Tait, may be in possession of finitistic constructions that go beyond the primitive recursive functions, yet fail to know that they are finitistic. Non-finitists, looking in to the structure of finitism from the outside, might know that his functions are finitistic while he remains ignorant of the fact. In what remains, we'll discuss this point in the context of an argument for condition (R). Our question is whether, given some means of belief (either inferential or non-inferential), (R) is a core feature of finitistic knowledge. Our goal is to show that it is.

Let's consider, towards an argument for (R_i), a series of increasingly complex examples. First, consider the following stroke-sequence:

$$|||||||||$$

and consider the "next" ten-member sequence:

$$||||||||||$$

produced by the concatenation (or addition) of a single stroke to the nine-member sequence. Now suppose that the finitist *knows* that the nine-member sequence is finitistic. That is, that there are nine strokes in the sequence, that he believes that there are, and that he has sufficient evidence and warrant to judge that there are nine strokes in the sequence, such as having the concept of "Number" (à la Tait), or being able to "survey" the array (à la Kreisel), or having a sufficiently high degree of justification through the successful iteration of elementary operations that are sufficiently similar to one another (our own account). If the finitist knows that the nine-member sequence is finitistic, then since knowledge is factive, it follows that it is. Since the ten-member sequence is obtained by one iteration of concatenating a single stroke to the nine-member sequence, and the nine-member sequence is obtained by nine iterations of the same operation, it seems that the ten-member sequence must be finitistic, if the nine-member sequence is. If it isn't, then as noted above, his degree of confidence is sufficiently high for him to know that the nine-member sequence is finitistic, while his belief that the ten-member sequence is finitistic is mistaken (irrespective of his degree of confidence). In such a case, it follows either that his construction process for the ten-member sequence must not have been sufficiently similar to the construction process for the nine-member sequence or that he is not reliable enough to have counted as knowing that the nine-member sequence is finitistic. Both entailments are, however, contrary to assumption. Hence, in this case, it seems that (R_i) must hold.

The argument above might be granted. Nonetheless it might be objected that the problem is not with cases in which "small" integers are calculated from "small" integers, in particular in cases where the only operation under consideration is a single iteration of the successor operation. Let's consider again, then, a function such as exponentiation above. Suppose that the ϕ_{ith} function is x^y and that the finitist knows that x^y (for $x = 10$ and $y = 10$) is finitistic. Suppose further that the ϕ_{i+1st} function, the next iteration of that function for those values, is $10^{10^{10}}$, and that it is not finitistic. One might be able to construct (or represent to oneself) "small" integers such as 10^{10} but fail to be able to construct much larger ones such as $10^{10^{10}}$ simply due to the difference in size. Though we considered this case briefly above, a similar case is raised by Bernays' in his elucidation of Hilbert's finitist epistemology, so let's look at Bernays' response to this problem in more detail. If one knows that 10^{10} is finitistic, then one iterates, as above the process used to obtain 10^{10} tenfold in order to obtain $10^{10^{10}}$. That the process of construction by which the latter number is produced from the former number cannot fail to be finitistic can be seen in the following discussion introduced by Bernays. He writes that finitist methodology:

> is not concerned with such thresholds for the possibility of the realization [of constructing a number like $10^{10^{10}}$ from 10^{10}]. For these thresholds are accidental from the point of view of the formal consideration. Formal abstraction, as it were, does not find an earlier position for a delimitation in principle than the difference between finite and infinite. (Bernays (1930), 249-50)

In this quote, by "formal abstraction," Bernays has in mind something like the following mathematical argument. Suppose that there is an upper bound, or threshold, on the finite numbers, above which no ordinal construction reaches. Let z be that number. Since z is a finite number, it can be represented by a finite set Z. By Cantor's Theorem, Z has a finite power set 2^Z such that the cardinality of the latter is strictly larger than that of the former. Since 2^Z is larger than Z, Z is the largest and not the largest set. We reach a contradiction (where the sets considered are clearly decidable), and it follows that no such bound exists. But it does not follow (finitistically) that the set of finitistic functions are unbounded. Bernays' claim is that whatever the mathematical upper bounds are, they can be adduced neither by empirical considerations, nor by considering the limits of our powers of representation, nor by arbitrary conceptual stipulations (such as Tait's and Kreisel's), but only by mathematical means. If we know that one function is constructible by one means and is finitistic, then if the "next" function is constructed in a sufficiently similar manner to the previous function, it too must be finitistic. That is, if the evidential bases for the ith and $i+1$st stages are relatively uniform, (R_i) seems to hold.

In effect, we are asking about the kinds of considerations that we are able to adduce for limits on the functions that can be constructed from the finitistic point of view. Bernays' claim is that no empirical reasons can count for those limits since otherwise it would follow that iterations of values for a constructed and known mathematical function can be bounded above by non-mathematical means. In such a case, the iteration of values of a function might be justified as finitistic to a sufficiently high degree for knowledge for some values but for some larger value fail to be finitistic. It is not clear upon what grounds, other than the processes being sufficiently *distinct* or the finitist being insufficiently reliable, one might declare one function that produces a finite number to be finitistic and another function that produces a larger finite number to be non-finitistic. On the other hand, Bernays does not tell us much more in (1930) about the kinds of mathematical considerations finitists are permitted to (or forbidden from) adduce in their constructions of functions. For example, if the ith stage of construction yields each primitive recursive function, and the finitist knows that each primitive recursive functions has been constructed (but not all of them together), and knows that each is finitistic, then if the ϕ_{i+1st} function is the function that outputs each primitive recursive function, a general but not primitive recursive function, is it finitistic or does it fail to be so? In section (6) we look in more detail at some of the procedures proposed by finitists in mathematical practice by which construction processes can be extended and limited. So far our account has shown that (L) and (R_i) are incompatible and that of the two principles implicit in accounts of mathematical, and in particular finitistic, knowledge, (R_i) and not (L) is the more plausible. Our negative claim, in short, is that although finitistic constructions are *a priori*, it is not the case that the construction of a function (or a proof of a procedure being well-defined) alone is sufficient to provide knowledge of that function; that is, that (L) fails such that it is possible to have a proof that a procedure is finitistic but fail to know that it is finitistic. Our positive claim, on the other hand, is that reflection upon the method and means of a finitistic construction yields a sufficiently high degree of justification for the finitist to pass from one stage of construction to another, given that his degree of justification for the predecessor stage is sufficiently high to warrant knowledge. Now let's turn to some objections to the present account.

2.5.3 Further Reflections

In section (5.1) it was argued that a variety of existing analyses of finitistic epistemology tacitly presuppose (R_i) and (L). In section (5.2) it has been argued that (R_i) and (L) are incompatible with one another plus the claim that the limit of finitistic reasoning is generically bounded above — that there exists a precise upper bound to the functions that finitists are permitted to accept. Then, it was argued that of the two principles (L) is implausible and that (R_i) is a core feature of the epistemic structure of finitistic reasoning. On most of the existing analyses of finitistic epistemology if the finitist has a proof, immediately it follows that he knows.[23] Our analysis, by contrast, implies that the existence of a proof, even a finitistic one that shows that a function terminates given a numeral as argument, is insufficient for finitistic knowledge. Our basic claim, distilled, is that one might be in possession of a construction for a function or a proof that it terminates, but not be in a position to know that it does. One might need more than just a proof that a function terminates in order to know that it terminates. Our finitist might need, for example, a proof along with a means of justifying epistemic procedures that permit him to pass to the formalism in which the proof might be carried out. Or he might not have the right kind of extension procedure to provide a sufficient amount of evidence to warrant the claim that the particular function terminates. Or he might need to be able to show that the particular function being evaluated possesses a particular set of properties that are sufficiently similar to functions that have been previously justified. He might, for example, have one kind of method that justifies (to a sufficiently high degree for knowledge) his belief that each α-recursive function such that $\alpha < \omega^2$ is finitistic, but a different method that justifies (to a sufficiently high degree for knowledge) his belief that each β-recursive function such

[23] On Tait's analysis, of course, we saw that he needs not only a proof, but also to have the concept *Number*. But this stipulation seems to do no other work than make it clear that Tait is committed to (L). See section (4) above.

that $\beta < \omega^\omega$ is finitistic. Likewise for functions of higher type. In section (4) we shall explore some of the options for such methods. But first we look at two possible objections to our account.

First, one might object to the argument against (L) on what we have suggested are Cartesian grounds. Standard accounts of mathematical proof more often than not take being in possession of a construction for a function or having a proof that a function terminates as a sufficient condition for being in a position to know it. Having a (formal or informal) proof is transparent to the prover, in the sense that once one goes through each step of a proof and a theorem is derived, then one has produced all the justification required for the belief. If the argument is, moreover, sound, then the conclusion must be true and it follows, on the classical account of knowledge, that the prover has a justified true belief and hence is in a position to know the theorem in question. In response to the above argument against (L), the Cartesian might retrench and claim that while one might have a case against *a posteriori* conditions (such as feeling cold or being in pain) that are claimed to be luminous, it is far more difficult to see how putatively *a priori* conditions (such as having a finitistic construction of a function) might fail to be luminous. In contrast to *a posteriori* claims, *a priori* claims, if justified, appear to be justified independently of the contexts in which proofs or constructions of the claims have been carried out. Mathematical claims are justified deductively and a mathematical proof is all the evidence that is required to establish the claim. Once one has a proof that a function terminates there can be no doubt that it does, since the degree to which a conclusion in mathematics (finitistic or otherwise) is justified is at least as high as the degree to which each of its premises are justified. Because the claim is *a priori* justified, the objection might continue, the justification for the claim obtains in all contexts (or in all possible worlds), so there are no contexts in which one might fail to be in a position to know that it obtains. Hence, the objection concludes, it is not possible for one to have a proof that a function terminates but fail to be in a position to know that it does.[24]

Our argument against (L) from sections (5.1) and (5.2) did not assume any specific account of mathematical knowledge. In section two of this chapter, however, we attempted to make plausible the claim that finitistic justification possesses variable degrees and that its variability is similar to that recognized for *a posteriori* justification. Hence, on our account, it is possible for the conclusion to a proof to be *pro tanto* justified but nonetheless count as a logico-mathematical proof. In fact, we showed that although finitists have high degrees of justification for variable-free claims, and that claims involving free-variables might have slightly lower degrees of justification, he might nonetheless be warranted in believing the latter. Let, for example, f be a recursive proof and $F(n)$ (such that $n \in \mathbb{N}$) be a formula of complexity Π_0. Then, our finitist might have a relatively high degree of justification for the claim that for each arbitrary n, $f(n)$ is a proof of $F(n)$, and a degree of

[24] Note that this formulation, to be in a *position* to know, becomes important in the context of objections to the argument against (L), since were it to *know*, full stop, we'd be able to respond to its defender by pointing out that a proof is insufficient for knowledge. One might, e.g., have a proof but fail to understand how it is a proof of the theorem.

justification lower than that for the premises but that still warrants belief for the claim that f is a recursive proof of $\forall x F(x)$. Despite that possibility, he might draw the conclusion that f is a proof of a Π_1 sentence, since nothing in the formulation of finitist methods requires that his proofs, qua finitist metamathematician, obey the classical account of deductive formal proof. Hence, because the claim that f is a proof of a Π_1 sentence is defined over an infinite set, he might have a high degree of justification for his claim, yet still fall just short of putting him in a position to know. Our claim is that it is possible to have a proof that a function terminates but nonetheless fail to be in a position to know that it does. One might have a high degree of justification for each premise, a high degree of justification for each step in the sequence of inferences, and a degree of justification for the conclusion that warrants outright belief, but still fail to be in a position to *know* the conclusion. Knowledge acquisition for finitists should be hard. About that the majority of analyses agree. But by tacitly presupposing (L), other analyses assume that it is not.

Second, one might object to the argument against (L) on the grounds that "is finitistic" or "being in possession of a finitistic function" or "having a finitistic proof" are vague predicates. Critics of the argument might claim that, on some sharpening of the predicate "is finitistic", it is true that if a function is finitistic at the i stage, then the finitist knows that it is, but that for that i it is not true that if the finitist knows that the ith stage of construction is finitistic, then the $i+1$st stage is. In particular, it is not true if one constructs a limit ordinal at the $i+1$st stage. Fix the sharpening of "is finitistic" to Tait's and consider the following sequence of functions:

$$\phi_0(a,b), \phi_1(a,b), \phi_2(a,b), \ldots, \phi_n(a,b),$$

where:

$$\phi_0(a,b) = a + b$$

$$\phi_1(a,b) = a \times b$$

$$\phi_2(a,b) = a^b$$

and for $n \geq 2$, the function $\phi_{n+1}(a,b)$ is defined by:

$$\phi_{n+1}(a,0) = a$$

$$\phi_{n+1}(a,S(b)) = \phi_n(a,\phi_{n+1}(a,b)).^{25}$$

On Tait's sharpening, it is possible for the finitist to know that for each k, the function $\phi_k(a,b)$ is finitistic, but that it is not true that for each k the $k+1st$ function is finitistic, since if it turned out to be so, then it would be possible for general but not primitive recursive functions to be finitistic, a claim that Tait rejects. But are there such k from the point of view of an epistemic analysis of

[25] Note this sequence defines the primitive recursive functions and hence is general but not primitive recursive.

finitism? If the $k+1$st stage is non-finitistic, then there must be some standard by which it is false that the $k+1$st stage is finitistic. That standard cannot be the generic one that Tait (and others) often cite, that it is not possible for finitists to pass beyond limit ordinals. For the finitist has already passed beyond limit ordinals at the stage at which he constructs $\phi_0(a,b)$, whose recursive ordinal is ω^2. If, on the other hand, that standard is the finitist's reliability at successor stages of construction such as the $k+1$st, then it is possible for him to achieve a sufficiently high degree of justification to warrant knowledge for each kth stage of construction, but for him to have a false belief for $k+1$st stages. But if finitists are required to meet normal rational standards, then we should count as abnormal finitists who *fail* to have a high degree of justification (though possibly lower) that the $k+1$st stage is finitistic if it is known that for each $k, \phi_k(a,b)$ is.[26]

Above we used Tait's sharpening as an example, and we might have just as easily used Kreisel's. But note that we have also shown that any sharpening of the predicate "is finitistic" or sharpening of the class of finitistic functions in which the transparency of a specific class of mathematical functions (and methods for constructing them) is the generic standard for knowing those functions to be finitistic must show why strict upper bounds for finitistic reasoning, if not arbitrary, nonetheless preserve the finitist's rationality. In order to argue that (L) fails, and hence, that there can be no such single class of functions and methods that form a cognitive home for the finitist we need only construct a single counterexample, and the argument from sections (5.1)-(5.2) suffices for that, irrespective of the specific sharpening of what it means to be finitistic. What we need, then, are descriptions of different methods that suffice to show when the finitist is justified in believing a function to be finitistic and when he is not justified in believing a function to be finitistic. For the means and methods with which finitists justify their claims that a particular function terminates (or is finitistic) are tailored or tied to the needs that arise in the evaluation of that particular function. One method might suffice to justify the exponentiation function as finitistic, whereas another method might be needed to justify the Ackermann function as finitistic. Finitistic reasoning is not bounded above because a certain kind of reasoning forms for the finitist a cognitive home. He is homeless, and the limits within which it is possible for him to acquire finitistic knowledge are drawn by mathematics itself, not conceptual stipulations about the "nature" of his reasoning (though that is not to say that conceptual analyses of different methods are irrelevant). In section (6) below, our goal is to make this claim as specific as possible.

2.6 Reflecting Further on Finitism

In section (5) we saw that a variety of existing analyses of finitism presuppose an epistemic principle that guarantees that finitistic functions are always known — (L) — and an epistemic principle that guarantees that if a function is known to be finitistic, then that knowledge is iterative — (R_i).

[26] Note that what has just been provided is an initial epistemic defense of induction for finitists. We return to this below from the point of view of finitistic practice as well.

Then we argued that the two together are incompatible with analyses that propose generic and uniform upper bounds for finitistic reasoning, whatever those bounds might be, and then we argued that of the two principles, it is (L) that is the offending one. However, up to now our argument has been mostly negative and we have relied entirely on conceptual arguments to show that global analyses that presuppose (L) in order to obtain generic and uniform upper bounds on finitistic reasoning run into the aforementioned difficulties. In this section our goal is to explore, through an analysis of methods proposed by practicing finitists, the types of procedures that we argue finitists are entitled to employ in the justification of what in the introduction we noted Hilbert and Bernays call "extensions to the finitistic standpoint." At stake here are extensions to the finitist standpoint that entitle finitists to pass beyond each ordinal $\alpha < \omega^\omega$, the proof-theoretic ordinal for PRA, and vice-versa, limitations to the standpoint such that it is not possible for finitists to pass to all the ordinals up to ϵ_0. In short, we're searching for principled procedures that permit the finitist to justify the introduction of α-recursive functions such that $\omega^\omega < \alpha < \epsilon_0$. In the first subsection, we look at a procedure proposed by Ackermann (1924) and show how it might serve as a means of justifying an extension to the finitist standpoint for all ordinals $\alpha < \omega^{\omega^\omega}$. In the second subsection, we analyze passages from Hilbert and Bernays (1934) and discuss their significance for fixing upper bounds to finitism for ordinals $\alpha < \omega^{\omega^k}$ for some k. In the third subsection, we analyze Hilbert's proposal for extending a formalism by an ω-rule in (1931a), and discuss how it might serve to extend or limit its scope. In the final subsection, we turn to our own proposal for a formal characterization and discuss some of its metamathematical properties.

2.6.1 Higher Recursion and Ackermann's Dissertation

Ackermann's dissertation (1924) is one of the earliest major contribution to Hilbert's finitist proof theory and Hilbert's Program following Hilbert and Bernays's development in 1922 and 1923 of the ϵ-calculus, their formalism for the analysis of arithmetic and analysis. He was a student of Hilbert's, and only four years after his dissertation, with Hilbert, Ackermann co-authored *Grundzüge der theoretischen Logik*. In his dissertation Ackermann develops in succession four types of basic calculi and presents consistency proofs for each one. That Ackermann's consistency proofs were considered, historically, to be finitistic, there can be no doubt. In a letter of recommendation for a fellowship awarded him in 1925, for example, Hilbert reports that:

> 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 demonstrated, as it were, directly on the mathematical formalism.[27]

[27] Hilbert-Nachlass, Niedersächsische Staats- und Universitätsbibliothek, Gottingen. Cod. Ms. Hilbert 458, sheet 6, undated. Quoted in Zach (2000).

For our purposes the important consistency proof is the one that takes the consistency proof given for an elementary calculus of free-variables, quantifier-free axioms, and schemata for generating primitive recursive definitions and extends it to an elementary calculus with free-function-variables. Ackermann calls the first calculus stage III (in keeping with the Hilbert School's "stepwise" and "simultaneous" development of logic and mathematics alongside one another), and calls the second calculus stage IV. Hilbert had already given a consistency proof for a version of stage III (essentially PRA) in lectures from 1921-22. In (1924) Ackermann follows Hilbert's lead from those lectures. First, one must assume, for reductio, that there is a linear derivation of $0 \neq 0$, then transform that derivation into one of tree-form, eliminate variables, and reduce function terms to numerals. At the end of the process one obtains a sequence of correct formulae, where "correct" is syntactically defined and decidable. One then concludes that there can be no derivations of $0 \neq 0$. But since no function variables are admitted at stage III, the verification of the argument is straightforward. But for stage IV it must be shown that every function is reducible to a numeral on the basis of the recursion equations when functions are permitted to take functions as arguments. Herein lies the crux of the problem.

Ackermann notes the problem with the following example.[28] Let $\phi_m(2, \mathfrak{g}(m))$ be a functional of higher type such that ϕ is defined by:

$$\phi_m(0, f(m)) = f(1) + f(2)$$

$$\phi_m(n+1, f(m)) = \phi_m(n, f(m)) + f(n) \times f(n+1)$$

Note that here $\mathfrak{g}(m)$ is a term for a function that occurs within a function. Hence, it is not possible to replace the variable b occurring as its argument before evaluating the entire function. Reduce the functional to the following form by applying the recursion equations:

$$\begin{aligned}
\phi_m(2, \mathfrak{g}(m)) &= \phi_m(1+1, \mathfrak{g}(m)) \\
&= \phi_m(1, \mathfrak{g}(m)) + \mathfrak{g}(1) \times \mathfrak{g}(1+1) \\
&= \phi_m(0+1, \mathfrak{g}(m)) + \mathfrak{g}(1) \times \mathfrak{g}(1+1) \\
&= \phi_m(0, \mathfrak{g}(m)) + \mathfrak{g}(0) \times \mathfrak{g}(0+1) + \mathfrak{g}(1) \times \mathfrak{g}(1+1) \\
&= \mathfrak{g}(1) + \mathfrak{g}(2) + \mathfrak{g}(0) \times \mathfrak{g}(1) + \mathfrak{g}(1) \times \mathfrak{g}(2) + \mathfrak{g}(1) + \mathfrak{g}(2)
\end{aligned}$$

Since $\mathfrak{g}(m)$ is a term for a function, the function ϕ itself might be in its domain. For example, $\mathfrak{g}(m)$ might denote the function $\phi_j(m, \gamma(j))$. But if so, then $\phi_m(2, \mathfrak{g}(m))$ reduces to:

$$\phi_j(1, \gamma(j)) + \phi_j(2, \gamma(j)) + \phi_j(0, \gamma(j))$$

[28] In what follows we follow Ackermann's own presentation from (1924) but change some notation, most notably the notation for ranks, indices, and orderings. Ackermann uses natural numbers in his proof, but we use ordinal notations for perspicuity.

$$\times$$

$$\phi_j(1, \gamma(j)) + \phi_j(1, \gamma(j))$$

$$\times$$

$$\phi_j(2, \gamma(j)) + \phi_j(1, \gamma(j)) + \phi_j(2, \gamma(j)).$$

Hence, it is possible to reduce a functional of higher type to a form in which it occurs as argument. On the other hand, if one reduces a primitive recursive term $\psi(\mathfrak{m})$ to a numeral, where ψ and $\mathfrak{m}$ are given, using the recursion equations for the function, then, unlike the reduction of the above term, ψ will not be contained in the final form of the reduced term.

In other words, we have not shown that the reduction terminates in a numeral. Rather we've shown that it is possible to reduce the given function to another function if functions take themselves as argument. But not only must we show that it terminates. Also we must provide a justification of the associated method that shows that finitists are entitled to employ it. Ackermann proceeds in two steps as follows. First, he assumes that innermost subterms are reduced first, and then assigns indices to terms to show that each reduction reduces the given index. Then he proves that the reduction of functionals of higher type terminates in a numeral. On that basis he concludes that the calculus is, in fact, consistent. We'll proceed with the proof as follows. Definitions precede the statement of the theorem. The inequality $\phi < \psi$ denotes that ϕ occurs before ψ in the order of definition.

Definition 4.0 (subordination): if t is a term containing ϕ and ψ, then ψ is *subordinate* to ϕ, abbreviated $sub(\psi, \phi)$, just in case:

$$t = t'(\ldots \phi(\ldots \psi(\ldots b \ldots) \ldots) \ldots),$$

where b occurs bound in ψ.

Definition 4.1 (rank): if $\psi > \phi$, then the *rank* of the function symbol with respect to its occurrence in t, abbreviated $rk(t, \phi)$, is defined as:

$$rk(t, \phi) = \begin{cases} 1 & \text{if no } \psi \text{ is such that } sub(\psi, \phi) \\ max\{rk(t, \psi) \mid sub(\psi, \phi)\} + 1 & \text{otherwise.} \end{cases}$$

Definition 4.2 (order): if $\mathfrak{s}$ is a subterm of $\mathfrak{t}$ and $\psi_0 < \cdots < \psi_n$, then the *order* of the subterm $\mathfrak{s}$, abbreviated $o(\mathfrak{s})$, is defined by:

$$o(\mathfrak{s}) = \omega^n \cdot rk(\mathfrak{s}, \psi_n) + \cdots + \omega \cdot rk(\mathfrak{s}, \psi_1) + 1 \cdot rk(\mathfrak{s}, \psi_0)$$

Definition 4.3 (degree): if $\mathfrak{s}$ is the set of all distinct subterms of t with an order o where the order is not a numeral, then the *degree* of o in t, abbreviated $d(t,o)$, is that set's cardinality, defined by:

$$d(t,o) = |\,\mathfrak{s}\,|$$

Definition 4.4 (index): if o ranges over all orders of subterms of t, and $\vec{o}$ is the ordinal corresponding to o, then the *index* i of t, abbreviated $i(t)$, is the sequence of degrees ordered identically to the orders, defined by:

$$i(t) = \sum_{o} \omega^{\vec{o}} \cdot d(t,o).$$

Note that if $\psi_0, \ldots, \psi_n$ is a sequence of primitive recursive functions (defined as in (3.3)) then the *order* of $\mathfrak{s}$, $o(\mathfrak{s})$, corresponds to an ordinal α such that $\alpha < \omega^{\omega}$, and the *index* of $\mathfrak{s}$ corresponds to an ordinal $\beta < \omega^{\omega^{\omega}}$.

We are now in a position to prove Ackermann's theorem for formal languages containing free function variables.

Theorem 4.0: Every higher type functional reduces to a numeral — or terminates in a numeral — on the basis of its recursion equations.

Proof: Let $\mathfrak{t}$ be a term not containing function symbols for transfinite recursive functions, and let $\mathfrak{s}$ be an arbitrary constant subterm occurring innermost in the term $\mathfrak{t}$ such that $\mathfrak{s}$ has the form:

$$\mathfrak{s} = \phi_{\vec{b}}(\mathfrak{a}_1, \ldots, \mathfrak{a}_n, \mathfrak{v}_1, \ldots, \mathfrak{v}_m)$$

where the $\mathfrak{v}_i$'s are terms with at least one variable bound by ϕ and don't contain constant subterms. There are two cases to consider.

Case One: $m = 0$, that is, $\mathfrak{s}$ contains no bound variables. If $m = 0$, then since $\psi_{k-1} > \phi_{\vec{b}}$ does not occur in $\mathfrak{t}$ and there's no occurrence of $\psi_k > \phi$ subordinate to ϕ in $\mathfrak{t}$, it follows that:

$$\begin{aligned} o(\mathfrak{s}) &=< rk(\mathfrak{s},\psi_k), \ldots, rk(\mathfrak{s},\psi_1), rk(\mathfrak{s},\psi_0) > \\ &= \omega^k \cdot 1 + \cdots + 0 + 0 \\ &= \omega^k \end{aligned}$$

Using the recursion equations to reduce $\mathfrak{s}$ to a term $\mathfrak{s}'$ it is clear that only function symbols with indices i lower than the index for ψ_k occur. Hence, the highest exponent in the order of $\mathfrak{s}'$ is less than k. Hence, $o(\mathfrak{s}') < o(\mathfrak{s})$. Since, by assumption, no variable which is bound occurs in $\mathfrak{s}$, no variable bound in $\mathfrak{t}$ occurs in $\mathfrak{s}'$. Hence, reducing $\mathfrak{s}$ to $\mathfrak{s}'$ results in no new occurrences of function symbols subordinate to any other function symbols. It follows, when $\mathfrak{t}'$ is the result of the reduction of $\mathfrak{t}$ in this way, that the number of subterms of orders greater than $\mathfrak{s}$ in $\mathfrak{t}'$ remains constant, but that the number of subterms of orders exactly equal to $o(\mathfrak{s})$ in $\mathfrak{t}'$ decreases by 1. Hence:

$$i(\mathfrak{t}') < i(\mathfrak{t}).$$

Case Two: $m > 0$, that is, $\mathfrak{s}$ contains bound variables. Without loss of generality let $\mathfrak{s}$ be of the form:

$$\mathfrak{s} = \phi_b(\mathfrak{z}, \mathfrak{g}(b)),$$

that is, that $\mathfrak{s}$ contains one argument with an arbitrary constant term $\mathfrak{z}$ and one argument with an arbitrary functional $\mathfrak{g}$. All function symbols occurring in $\mathfrak{g}(b)$ are subordinate to ϕ, since otherwise it would contain a constant subterm and this case would collapse to the evaluation of a (first-order) primitive recursive function. Since function symbols occurring in $\mathfrak{g}(b)$ are subordinate to ϕ, it follows that:

$$rk(\mathfrak{t}, \mathfrak{g}(b)) < rk(\mathfrak{t}, \mathfrak{s}).$$

Using the recursion equations to reduce $\mathfrak{s}$ results in a term $\mathfrak{s}'$ that does not contain the symbol ϕ. Note that if one substitutes an arbitrary term $\mathfrak{b}$ for the variable b in $\mathfrak{g}(b)$, then because none of the occurrences of function symbols in $\mathfrak{b}$ contain bound variables whose scope is outside of $\mathfrak{b}$, so that no bound variables are subordinate to any function symbols in $\mathfrak{g}(b)$, it follows that:

$$(*) \quad o(\mathfrak{g}(\mathfrak{b})) = max\{o(\mathfrak{g}(b)), o(\mathfrak{b})\}.$$

We must show that reducing $\mathfrak{s}$ to $\mathfrak{s}'$ in $\mathfrak{t}$ decreases $\mathfrak{t}$'s index. The proof now proceeds by induction on the term $\mathfrak{z}$.

Let the following be the recursion equations for ϕ:

$$\phi_b(0, f(b)) = \mathfrak{a}_b(f(b))$$

$$\phi_b(a+1, f(b)) = \mathfrak{b}_b(\phi_c(a, f(c)), a, f(b)).$$

Base Case: If $\mathfrak{s} = \phi_b(0, \mathfrak{g}(b))$, then $\mathfrak{s}' = \mathfrak{a}_b(\mathfrak{g}(b))$. We must argue that $\mathfrak{s}'$ has the same rank as $\mathfrak{g}(b)$ with respect to ϕ. Since $\mathfrak{a}$ does not contain ϕ, when $f(b)$ is an argument, it's replaced by $\mathfrak{g}(d)$ such that the variable d is not in the domain of ϕ. That shows that $rk(\mathfrak{t}, \mathfrak{s}') \leq rk(\mathfrak{t}, \mathfrak{g}(b))$. In the other direction, with a term $\mathfrak{g}(\mathfrak{d})$ such that $\mathfrak{d}$ does not contain $\mathfrak{g}$, by $(*)$ and the fact that $\mathfrak{d}$ does not contain ϕ, it follows that $rk(\mathfrak{t}, \mathfrak{g}(\mathfrak{d})) = rk(\mathfrak{t}, \mathfrak{g}(b))$. If $\mathfrak{d}$ does contain an occurrence of a nested $\mathfrak{g}$, then by the induction hypothesis and the argument in case one,

$rk(\mathfrak{t}, \mathfrak{d}) = rk(\mathfrak{t}, \mathfrak{g}(b))$.[29] Hence, by $(*)$, $rk(\mathfrak{t}, \mathfrak{s}') \geq rk(\mathfrak{t}, \mathfrak{g}(b))$. Hence:

$$rk(\mathfrak{t}, \mathfrak{s}') = rk(\mathfrak{t}, \mathfrak{g}(b)).$$

Induction: If $\mathfrak{s} = \phi_b(\mathfrak{z} + 1, \mathfrak{g}(b))$, then $\mathfrak{s}' = \mathfrak{b}_b(\phi_c(\mathfrak{z}, \mathfrak{g}(c)), \mathfrak{z}, \mathfrak{g}(b))$. Further applications of the recursion equations replace the first argument, $\phi_c(\mathfrak{z}, \mathfrak{g}(c))$, by a sequence of other terms $\mathfrak{u}'_0, \ldots, \mathfrak{u}'_k$ such that for each $0 \leq i \leq k$, by k applications of the induction hypothesis, we have:

$$rk(\mathfrak{t}, \mathfrak{u}'_i) \leq rk(\mathfrak{t}, \mathfrak{g}(b)).$$

Hence, in this case it too follows that:

$$rk(\mathfrak{t}, \mathfrak{s}) \leq rk(\mathfrak{t}, \mathfrak{g}(b)).$$

Hence, if $\mathfrak{t}'$ results from eliminating ϕ from $\mathfrak{t}$ by recursion, then the argument shows that the rank of $\mathfrak{t}'$ decreases by at least 1 and the recursion does not increase the rank of any subterms ψ_j for $j > i$.

It is important to note that since the index of $\mathfrak{s}$ is an ordinal ω^{ω^k} for some k, a formal argument for the above claim requires transfinite induction on each ordinal $\alpha < \omega^{\omega^\omega}$ in order to show that in a finite number of steps one obtains a functional that contains no function symbols — that is, a numeral. But Ackermann describes the theorem above as showing only that the computation procedure defined by 2-fold nested recursions such as the function given by the sequence $\psi_0, \ldots, \psi_n$ terminates. He writes that:

> [t]o each functional corresponds a transfinite ordinal number as its rank, and the theorem, that a constant functional is reduced to a numeral after finitely many operations, corresponds to [the other theorem] that if one descends from a transfinite ordinal number to ever smaller ordinal numbers, one must reach zero after a finite number of steps. (Ackermann (1924), 13-4)

And he sees nothing in it that violates the finitistic standpoint:

> there is naturally no mention of transfinite sets or ordinal numbers in our metamathematical investigations. It is, however, interesting that the theorem about transfinite ordinals can be formulated so that there is nothing transfinite about it. (Ibid.)

Ackermann proceeds to argue that the finite sequences of numbers in the finitistic version of the theorem are well-ordered. Importantly, however, he does not argue for the well-orderedness of the

[29] Note that because we are dealing with *nested* recursions, the base case includes a (series of) sub-induction(s), and the inductive case includes a series of sub-inductions, one for each evaluation of the function using the recursion equations.

entire class of sequences of sequences (i.e., for each α-recursive function such that $\alpha < \omega^{\omega^\omega}$), but only for that subset of the sequences pertaining to the 2-fold nested recursion that bears his name (i.e., for $\alpha < \omega^2$). We reproduce the argument below.

Ackermann asks us to 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 numbers. 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. (Ackermann (1924), 14)

Once we have a characterization of the ordinals up to ω^2 in terms of a pair of finite numbers, Ackermann characterizes the reduction of the pair to $(0, 0)$ in terms of the operations performed on them. He continues:

> [t]o 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 if 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. (Ibid.)

Ackermann's argument appeals to our ability to record the order of the applications of the recursion equations over a pair of finite numbers. As one considers one or the other number contained in the pair, one need only apply the equations finitely many times over that pair in order to reach the point at which the application of the equations terminates. Note that an argument that shows that a given primitive recursive function terminates need only apply the equations finitely many times over a single arbitrary finite number. Hence, the only difference between Ackermann's argument and one that shows that a primitive recursive function terminates is that his argument proceeds over *pairs* of arbitrary finite numbers and not a single arbitrary finite number. Ackermann appears to take that similarity to be evidence for the claim that 2-fold nested recursion is finitistic.

If we accept the similarity as evidence for the claim, then we obtain a justification for extending the finitistic standpoint from its characterization in a formalism in which individual primitive recursive functions are definable to a formalism in which all the primitive recursive functions are definable by 2-fold nested recursion. In terms of an ordinal characterization, if we accept Ackermann's argument, we obtain a justification for passing from each ordinal $\alpha < \omega^\omega$ to each ordinal

$\alpha < \omega^{\omega^2}$. But is it possible to extend — with justification of equal degree — the claim that each ordinal $\alpha < \omega^{\omega^2}$ is finitistic to each ordinal $\alpha < \omega^{\omega^\omega}$? Ackermann claims that it is:

> [t]he 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. (Ibid.)

There are three points to consider. First, Ackermann's description of the process serves as a set of instructions for reducing recursion on ω^2 to primitive recursion when substitution in the parameters is permitted. Since that process can be reduced to primitive recursion, it shares the same degree of epistemic justification as primitive recursion. Second, the schemata of quantifier-free induction and recursion on ω^k for each k are derivable in PRA instance by instance. If the functions definable in PRA are finitistic (and are by most accounts), and since the 2-fold general recursive function bearing Ackermann's name is a description of that set of functions, then it must be finitistic. Third, and on the other hand, if it is possible to further iterate Ackermann's justification, then his description ought to justify induction on each ordinal up to ω^ω. For "the same holds true" comment seems to claim that the justification for the construction process at one stage, since it warrants knowledge at that stage, warrants knowledge at the next stage. But note that Ackermann purports to provide a finitistic proof of the induction principle on ω^ω, and a proof must appeal to induction on ordinals beyond ω^ω.[30] That is, Ackermann's proof involves a tacit appeal to induction on ordinals whose status as finitistic is not justified by his description.

On this point Tait concludes, correctly in our view, that Ackermann's description of the process does not serve to justify as finitistic the ordinals past ω^2 and only serves as a finitistic justification of the Ackermann function if it reduces to primitive recursion. Against Zach (2001), who argues that there is not a single notion of finitist function but a hierarchy corresponding to the complexity of iterations one is prepared to accept, Tait writes that the "notion of an iteration here has become unclear: an arbitrary number X represents a *single* arbitrary iteration [which allows] no sense to be made of speaking of the 'complexity of iterations' " (Tait (2001), 60). But since his argument against Zach and Ackermann involves an appeal to *his* analysis of the concept of Number in terms of iteration, it begs the question. On the other hand, based on the analysis in Zach (2001) of Ackermann's description and the fact that k-fold nested recursion reduces *inter-alia* to recursion on ordinals ω^{ω^k} for each k, Zach's claim that there is a hierarchy of concepts of finitist functions overshoots its mark. Since the description that Ackermann offers fails to take into account the fact that with increasing iterations through the ordinals, the complexity of finitistic justification increases and, since the complexity of the justification is inversely proportional to its epistemic degree, the

[30] Such a justification might be straightforward if (L), (R_i), and the claim that finitistic reasoning is bounded above are jointly true. In section (5) we argued that these claims cannot be jointly true, and that (L) fails, so that the justification is not straightforward.

epistemic degree of justification decreases.[31] Hence, it cannot be the case that "the same holds true" for ordinals beyond ω^{ω^2}. The process Ackermann describes that takes us from each primitive recursive function to a nested 2-fold recursive function, if it warrants knowledge, cannot be extended past the 2-fold recursive functions where that extension enjoys an equal degree of justification.[32] In the next section, we explore some analyses of extensions to the finitist standpoint in Hilbert and Bernays (1934) in order to understand if, and how, their analyses fare any better.

2.6.2 Higher Recursion and *Grundlagen* I

From our analysis in the last section it is clear that Ackermann's description provides a justification for recursion on ω^2 but fails to iterate further and hence fails to provide a justification for recursion on each ordinal up to ω^ω. In sections (2) and (7) of Hilbert and Bernays' *Grundlagen der Mathematik* (GMI) and (GMII) the authors discuss extensions of the finitist standpoint to functions beyond the primitive recursive functions. Most of the discussion of (GMI) and (GMII) in the secondary literature on the scope of finitism has focused on section (7) (see, e.g., Zach (1998), Tait (2000)), since therein is the most extended discussion of the Ackermann function and the enumeration function for the primitive recursive functions. Of course, Tait concludes from their discussion that the authors fail to prove that general but non-primitive recursive functions are finitistic. However, in section (7) the authors refer back to section (2), in which they introduce what they call *finitist number theory* (FNT), a contentual non-formal theory of primitive arithmetic the authors consider to be finitistic, in contrast to what they call *recursive number theory* (RNT), a formal theory of arithmetic one version of which includes a quantifier-free induction scheme (and is equivalent to PRA). In contrast to Ackermann's purported proof that each ordinal up to ω^{ω^ω} to be finitistic, in (GMI) Hilbert and Bernays offer a means weaker than proof — what they call a "verification" — that they suggest justifies the extension of the finitistic standpoint from the primitive recursive functions to general recursive but non-primitive recursive functions. In this section we analyze section (2) of (GMI), then section (7) of (GMI).

Section (2) of (GMI) is an extended discussion of (FNT) with respect to its admissible operations. Some operations, claim Hilbert and Bernays, are *independently* justified as finitistic, while other operations are *dependently* justified as finitistic. The operations that follow from the basic operation of constructing a numeral by iterating over given sequences in representation are independent, while those that stand in need of some sort of justification — such as the principle of induction, or definitions by recursion — are dependent ((GMI), 23). Recall, from chapter two, that an arbitrary

[31] Recall, again, that in chapter two we claimed that the ratio captures the common idea that the longer a proof becomes, the more difficult it is to understand. Finitists interpret the idea as the claim that "larger" finite numbers (and functions of higher type) are more difficult to construct so that the justification for claiming that they are finitistic decreases.

[32] In terms of ordinal notation systems, this entails a justification for each ordinal up to ω^{ω^2} but not for each ordinal up to ω^{ω^ω}. In other words, Ackermann's proof (executed above) and his description of recursion on ω^2 (quoted above) fail to count as finitistic justifications for recursive functions that are *not* reducible to primitive recursive functions.

finite numeral consists in a representation of finite sequences constructed by iterating over a given representation such as $|$. Hence, these representations are of the form $| \cdots |$ and constitute (FNT)'s domain ((GMI), 21). It is now possible to state the language for which its "communication signs" are defined. Let $\mathcal{L}_{FNT}$ consist (at least partially) in the following:

- *constants*: '2' abbreviates '11', '3' abbreviates '111', etc., and '11', '111', etc., refer to the represented stroke-sequence numerals $||$, $|||$, etc.;
- *variables*: '$\mathfrak{a}$', '$\mathfrak{b}$', etc., refer to arbitrary stroke-sequences of numerals;
- *operations*: '$+$' refers to the operation of addition over numerals, '$\cdot$' refers to the operation of multiplication over numerals, etc.;
- *constructions*: 'ϕ', 'ψ', etc., refer to defined (function) constructions;
- *relations*: '$=$' and '$\neq$' refer to numerical identity and difference.

Hilbert and Bernays are explicit about the fact that the language is *not* closed. For as one develops (FNT) new communication signs (for abbreviations and the introduction of functions) might be needed (ibid.). In other words, if it is possible to finitistically derive, based on previously derived or primitively given facts or rules, other facts or rules, then the derived facts or rules may be taken as new facts or rules in (FNT). The authors provide a simple example. Suppose the stroke-sequence denoted by a numeral $\mathfrak{a}$ "matches" the sequence denoted by a numeral $\mathfrak{b}$. Then one of the two sequences is an initial segment of the other or they are identical. Hilbert and Bernays claim that trichotomy follows from the justification above, and that similar considerations justify transitivity ((GMI), 22). Hence, it is legitimate to extend $\mathcal{L}_{FNT}$ by the signs '$<$' and '$>$' and to extend (FNT) by rules for inequalities between numerals (interpreted in (FNT)'s domain).

After introducing addition (whose justification rests on our ability to append stroke-sequences to stroke-sequences) into (FNT), Hilbert and Bernays claim that associativity for addition is independently justified based on considerations involving stroke-sequences. But the authors then argue that commutativity for addition is not independently justified because it requires an appeal to the "proof method of complete induction [*die Beweismethode der vollständigen Induktion*]" ((GMI), 23). Hilbert and Bernays are discussing mathematical induction, or more precisely, quantifier-free induction over the numerals, and induction over the numerals, they argue, requires a finitistic justification in order to be introduced into (FNT). Since its character is reflected in their subsequent discussion of recursive functions, let's briefly look at their justification for it. Hilbert and Bernays ask us to consider a predicate $\mathfrak{P}$ that one wants to conclude holds of an arbitrary numeral $\mathfrak{a}$. In order to conclude that $\mathfrak{P}$ holds for arbitrary $\mathfrak{a}$, one must derive:

(i) $\mathfrak{P}(1)$;

(ii) $\mathfrak{P}(\mathfrak{n}) \rightarrow \mathfrak{P}(\mathfrak{n}+1)$.

Hilbert and Bernays argue that if $\mathfrak{a}$ is an arbitrary numeral, then it has a construction procedure that begins with a representation of a sequence of strokes to which the basic iteration and concatenation operations are applied finitely-many times. In order to derive (i) and (ii), one verifies that $\mathfrak{P}$ holds for $\mathfrak{a}$ in tandem with the means by which $\mathfrak{a}$ is constructed. Hence, (i) and (ii) are derived co-temporally with $\mathfrak{a}$'s construction (ibid.).[33] According to the authors, the above is a sufficient justification for introducing quantifier-free induction into (FNT). Moreover, it is now possible to prove, using the introduced principle, the commutativity of addition.[34]

Theorem 4.1: For any arbitrary numerals $\mathfrak{a}, \mathfrak{b}$: $\mathfrak{a} + \mathfrak{b} = \mathfrak{b} + \mathfrak{a}$.

Proof: The proof proceeds by induction on numerals $\mathfrak{a}$ and $\mathfrak{b}$.

Base Case: Since the term $1 + 1$ is syntactically equivalent to the term $1 + 1$ it follows that $1 + 1 = 1 + 1$ immediately.

Induction: Let $\mathfrak{n}$ and $\mathfrak{m}$ be arbitrary numerals and assume:

$$\mathfrak{n} + \mathfrak{m} = \mathfrak{m} + \mathfrak{n}.$$

Syntactic identity between terms gives us immediately:

$$\mathfrak{n} + (1 + \mathfrak{m} + 1) = \mathfrak{n} + (1 + \mathfrak{m} + 1).$$

By the associativity of addition, it follows that:

$$\mathfrak{n} + (1 + \mathfrak{m} + 1) = \mathfrak{n} + (1 + \mathfrak{m}) + 1.$$

By commutativity for arbitrary numeral $\mathfrak{a} + 1$, it follows that:

$$\mathfrak{n} + (1 + \mathfrak{m} + 1) = \mathfrak{n} + (\mathfrak{m} + 1) + 1.$$

By the associativity of addition, it follows that:

$$\mathfrak{n} + (1 + \mathfrak{m} + 1) = (\mathfrak{n} + \mathfrak{m}) + 1 + 1.$$

[33] Recall that Hilbert (1931a) argues that the above construction is the "intuitive-contentual" justification of induction. One difference between (1931a) and (1934) is that in (1934) the authors are clear that the *proofs* of (i) and (ii), the *verification* that $\mathfrak{P}$ holds of $\mathfrak{a}$, and the *construction* of the numeral $\mathfrak{a}$ occur co-temporally.

[34] The proof below depends, in three steps, on a lemma that establishes that addition is commutative for the case in which $\mathfrak{a}+1 = 1+\mathfrak{a}$. Hilbert and Bernays do not derive commutativity for addition in (1934) at all, but mention that it is derivable, and discuss the mentioned lemma. On the other hand, in (1921/22), Hilbert derives the theorem 4.1, but proceeds by reducing numerals on either side of the identity side to obtain the identity $0 = 0$. Hence, Hilbert's proof in (1921/22) demonstrates that the procedure of commuting over addition terminates, and hence does not depend on the lemma.

> By the inductive hypothesis and substitution, it follows that:
>
> $$\mathfrak{n} + (1 + \mathfrak{m} + 1) = (\mathfrak{m} + \mathfrak{n}) + 1 + 1.$$
>
> By the associativity of addition, it follows that:
>
> $$\mathfrak{n} + (1 + \mathfrak{m} + 1) = \mathfrak{m} + (\mathfrak{n} + 1) + 1.$$
>
> By commutativity for arbitrary numeral $\mathfrak{a} + 1$, it follows that:
>
> $$\mathfrak{n} + (1 + \mathfrak{m} + 1) = \mathfrak{m} + (1 + \mathfrak{n}) + 1.$$
>
> By the associativity of addition, it follows that:
>
> $$\mathfrak{n} + (1 + \mathfrak{m} + 1) = (\mathfrak{m} + 1) + (\mathfrak{n} + 1).$$
>
> By commutativity for arbitrary numeral $\mathfrak{a} + 1$, it follows that:
>
> $$\mathfrak{n} + (1 + \mathfrak{m} + 1) = (1 + \mathfrak{m}) + (1 + \mathfrak{n}).$$
>
> By the associativity of addition, it follows that:
>
> $$(\mathfrak{n} + 1) + (\mathfrak{m} + 1) = (1 + \mathfrak{m}) + (1 + \mathfrak{n}).$$
>
> Since the claim holds for the base case and the inductive case, then by the quantifier-free induction principle it follows that commutativity of addition for arbitrary $\mathfrak{a}, \mathfrak{b}$ is derivable in (FNT). □

We derived theorem 4.1 in order to provide an example of how, in (FNT), it is possible to justify new rules and facts based on previously justified rules and facts. By the argument from our section (3), the degree to which theorem 4.1 is justified as finitistic depends upon the degree to which induction is finitistic, in the sense that if it is known that induction is finitistic then the degree of confidence that proofs based on induction are finitistic must be high enough to warrant knowledge that induction is. Following the discussion of commutativity and some properties of multiplication, Hilbert and Bernays discuss the method of introducing a function sign via definition by recursion. Hilbert and Bernays write:

> [a] new function-sign, such as ϕ, is introduced and its definition is presented through two equations, which in the simplest cases have the form [*im einfachsten Falle die Form haben*]:

$$\phi(1) = \mathfrak{a}$$
$$\phi(\mathfrak{n} + 1) = \psi(\phi(\mathfrak{n}), \mathfrak{n}).$$

> Here $\mathfrak{a}$ is a numeral and ψ is a previously known function constructed according to previously known rules so that it is possible to calculate $\psi(\mathfrak{b}, \mathfrak{c})$ given numerals $\mathfrak{b}, \mathfrak{c}$ such that its calculation results in a numeral as value. ((GMI), 25-6)

In the next section, we shall see how a crucial ambiguity in the description of function signs given by Hilbert and Bernays makes it possible to understand how finitists might ascend to and claim that higher recursive functions are finitistic. It is to this that we now turn.

2.6.3 Higher Recursion & Ravaglia's k-folds

Ravaglia (2003) argues that there are two distinct readings of the passages in (GM) in which Hilbert and Bernays discuss recursive functions: the *broad* construal and the *narrow* construal (Ravaglia (2003), 77*ff.*). Under the narrow construal of the above passage one takes the pair of equations as defining a specific recursive function. Under the broad construal one takes the pair of equations as a definitional principle for specifying the class of multiply recursive functions. Of the broad reading of the pair of recursion equations introduced in the previous section, Ravaglia writes that:

> what Hilbert and Bernays are really doing is defining a calculation procedure and explaining how when applied to a primitive recursive definition (with a specification of numerals for its arguments) this procedure results in a numeral. (Ravaglia (2003), 78)

On this reading $\mathfrak{a}$ is a parameter ranging over the numerals and ψ ranges over the definitions for the primitive recursive functions and the passages in which Hilbert and Bernays introduce the function signs are understood to be rules for obtaining specific definitions of functions from the general sketch of calculation procedures for those functions. On the other hand, under the narrow construal of the passage $\mathfrak{a}$ denotes a specific number, and ψ is defined by a specific function. Clearly the broad construal entails the narrow construal, but not the other way around, since it is possible under the broad construal to substitute a concrete function for the definitional schema, but it is not possible to infer from the narrow construal of the passage that all functions can be calculated using the given procedure. Hilbert and Bernays appear to be aware of the ambiguity. After providing an example of a function, the authors continue:

> [i]t is not entirely clear which sense this method of definition possesses. In order to clarify it one must make the concept of a function precise. By a *function* we understand an intuitive instruction [*anschauliche Anweisung*] into which numerals are input, for example, a pair, a triple, ..., of numerals, and which results in a numeral. ((GMI), 26)

Consider the fact that in their "definition" of a function, the authors claim that *any* numeral may be substituted. If so, then $\mathfrak{a}$ must range over the numerals and ψ must range over the primitive recursive functions. Ravaglia argues that although this provides evidence that Hilbert and Bernays endorse the broad construal, there are two main reasons to call this construal into question. First, on the broad construal it is possible to define the entire class of primitive recursive functions. Second, the finitist can define a "canonical calculation procedure for evaluating signs defined by primitive recursive definitions so long as these expressions are given along with the recursion equations defining these signs" (Ravaglia (2003), 81).

From a conceptual point of view we want to know whether Hilbert and Bernays are justified in claiming that we ought to read the equations as a definitional principle specifying the class of multiply recursive functions. The authors justify the claim with the following argument, continued after the definition of the concept of a function:

> A pair of equations such as above, which we call a "*recursion*," are interpreted as an *abbreviated communication* [*abgekürtze Mitteilung*] of the following instructions:
>
> Let $\mathfrak{m}$ be an arbitrary numeral. If $\mathfrak{m} = 1$, then $\mathfrak{m}$ is assigned the numeral $\mathfrak{a}$. Otherwise $\mathfrak{m}$ has the form $\mathfrak{b} + 1$. Then one writes down schematically:
>
> $$\psi(\phi(\mathfrak{b}), \mathfrak{b}).$$
>
> If $\mathfrak{b} = 1$, then one replaces $\phi(\mathfrak{b})$ with $\mathfrak{a}$; otherwise $\mathfrak{b}$ has the form $\mathfrak{c} + 1$ and one replaces $\phi(\mathfrak{b})$ with:
>
> $$\psi(\phi(\mathfrak{c}), \mathfrak{c}).$$
>
> Either $\mathfrak{c} = 1$ or $\mathfrak{c}$ has the form $\mathfrak{d} + 1$. In the first case one replaces $\phi(\mathfrak{c})$ with $\mathfrak{a}$, and in the second one replaces it with:
>
> $$\psi(\phi(\mathfrak{d}), \mathfrak{d}).$$
>
> The iteration of this method always terminates. For numerals:
>
> $$\mathfrak{b}, \mathfrak{c}, \mathfrak{d}, \ldots$$
>
> obtained sequentially arise from the *decomposition* [*Abbau*] *of the numeral* $\mathfrak{m}$ and it, like the construction [*Aufbau*] of $\mathfrak{m}$, must terminate. [...] From the content of these instructions one sees first that in every case they can be executed for any given numeral $\mathfrak{m}$ and that the result is unique. (Ibid.)

Ravaglia points out that Hilbert and Bernays do not require a proof, in the classical sense, that the procedure terminates from within the finitistic standpoint (Ravaglia (2003), 82). Once one has shown that it is possible to eliminate ϕ by rewriting it in terms of previously known functions and the numerals occurring in it, the authors conclude that we "have then obtained a computable expression" ((GMI), 27). That conclusion must be immediate from the data when it is understood in terms of the broad construal of their discussion.

In order to support the last claim consider the analogies between the decomposition [*Abbau*] of an arbitrary numeral, the construction [*Aufbau*] of an arbitrary numeral, and the inductive derivation that a statement $\mathfrak{P}$ holds for arbitrary numerals. Since quantifier-free induction proceeds in tandem with the construction of an arbitrary numeral, and since the decomposition of the numeral according to the above instructions is sufficiently similar to these procedures, then insofar as quantifier-free induction is finitistic, it must be legitimate to introduce these instructions (as rules) into (FNT).[35] In section (7) of (GMI) Hilbert and Bernays discuss sharpening up the earlier results. In section (7) they write that in section (2) recursions:

> serve as an abbreviated communication of a procedure through which from one or more numerals one determines a numeral. We can reproduce the procedure in the formalism by permitting the general introduction of function-signs in connection with recursion equations. ((GMI), 287)

Later on the authors expand upon what precisely the reproduction of the argument in (RNT) from section (2) amounts to by claiming that:

> [r]ecursive number theory is equivalent to intuitive number theory [*anschauliche Zahlentheorie* (FNT)], as we studied it in section 2, just in case its formulae have an interpretation consonant with finitistic contentual meaning. ((GMI), 330)

Hilbert and Bernays' subsequent discussion of the Ackermann function (formulated both as a sequential enumeration of two-place functions and as a nested recursion) makes it clear that finitists must be in a position to argue that the formal process for evaluating functions in a formalism (such as (RNT)) and the contentual instructions for evaluating procedures (in (FNT)) are equivalent in order to provide a higher degree of justification for the claim that the function is finitistic. In other words, the finitist must be able to provide a finitistic proof of the equivalence in order to justify the claim that functions beyond primitive recursive functions are finitistic. Below is a sharpened form of the set of claims made by Hilbert and Bernays in section (7).

Theorem 4.2: Let F be a formalism extending the predicate calculus by identity and axioms for the primitive recursive functions. If:

$$f(a, \ldots, k, z)$$

[35] Note that if we formalize the instructions, the resulting function is a 2-fold nested recursive function, and the argument that it terminates is, if formalized, an argument whose complexity is at least that of a 3-fold recursion. We return to the point below.

is a recursively defined function in F, then there is a proof in F of:

$$\mathfrak{f}(\mathfrak{a}, \ldots, \mathfrak{k}, \mathfrak{z}) = \mathfrak{l}$$

given an assignment of numerals, $\mathfrak{a}, \ldots, \mathfrak{k}, \mathfrak{z}$, for variables, $a, \ldots, k, z$, such that $\mathfrak{l}$ is the interpretation of the $\mathcal{L}_{FNT}$-expression $\mathfrak{f}(\mathfrak{a}, \ldots, \mathfrak{k}, \mathfrak{z})$ that is obtained by a calculation procedure justified in (FNT).

But consider again the class of multiply recursive functions on Ravaglia's broad construal. It is enumerable through an enumeration schema for k-fold recursions for arbitrary k. If one defines a function $\mathfrak{f}(\mathfrak{m}, \mathfrak{n}_1, \ldots, \mathfrak{n}_k)$ of higher type than each function in the class that enumerates the functions of this class, where $\mathfrak{m}$ is an index of the multiply recursive functions and $\mathfrak{n}_1, \ldots, \mathfrak{n}_k$ are argument parameters for said functions, then since each $\mathfrak{m}$ can take arbitrary k as argument, there is no finite upper bound on the arity. It follows that $\mathfrak{f}$ has no finite upper bound. On their own the considerations in (GMI) do not support the claim that functions with infinitely many arguments are finitistic. Hence, one might object, as Tait does, that a proof in the classical sense that the enumeration schema for k-fold recursion terminates involves quantification over the class of (signs for) k-fold recursive functions. Since that involves transfinite induction over a class of functions (signs) and not mere induction over numerals, it might be thought to be non-finitistic. Ravaglia notes that Hilbert and Bernays "provide no additional argument for" the conclusion that they have obtained a computable expression (Ravaglia (2003), 83). For "one can thus think of the canonical calculation procedure as involving a replacement procedure followed by an appeal to the canonical calculation procedure applied to previously defined signs" and the proof that such a procedure terminates is taken for granted at the informal level of the exposition (ibid.). Hence, in response to such an objection, Ravaglia argues that Hilbert and Bernays' argument permits us to conclude that each function up to $\omega^{\omega^{\omega}}$ is a finitist function but, again not all of them. We are only licensed to conclude that *that* k-fold nested recursion is finitistic for that k. In other words, unlike Ackermann's generalization, the analysis of (GMI) here suggests that for each function (beyond the primitive recursive functions) whose finitistic status is in question, one must have a "proof" such as the one above.[36] Insofar as we have argued, but from an epistemic point of view, that finitists are licensed to take functions up to $\omega^{\omega^{\omega}}$ as finitistic, our argument coincides with Ravaglia's.

But the natural question is whether it is possible to consider functions beyond $\omega^{\omega^{\omega}}$ as finitistic. Ravaglia seems to conclude that on their own his analyses of (GMI) do not suggest such a conclusion. For functions beyond $\omega^{\omega^{\omega}}$, in general there is no finite bound on the arity of the functions. Hence, it seems that we must conclude that there are some functions beyond those formalizable in (PRA)

[36] In fact, Hilbert and Bernays (in (GMII)) require *more*. One must also have a proof that the formal system is "verifiable [*verifizierbar*]," a concept introduced and defined in section (6) ((GMI), 237). If it is verifiable, then one may introduce an axiom asserting that the function is total. In effect, the concept defined may be taken as an instance of a reflection principle.

that count as finitistic, but that these considerations do not license us to conclude that there are any functions beyond ω^{ω^ω} that should so count. In order to make the conclusion to the argument in this section as explicit as possible, consider the following. From an ontological point of view, there exists no finite bound on the class of multiply recursive functions, and that entails that there exist infinitely many such functions. But that does *not* entail that, from an epistemic point of view, the finitist need be committed to the claim that there are infinitely many multiply recursive functions, if the notion of "commitment" at work itself entails knowledge of that claim. In fact, the argument from (GMI) above entails, contra Tait, that it is possible for the finitist to know that each multiply recursive function is finitistic without having a proof that the evaluation function for this class is itself a function. Our finitist's evidence for his knowledge claims might arise from multiple mathematical sources, and a classical proof that the enumeration function for this class terminates might outrun his reach, while we have argued that the methods outlined above do not. Our argument for (R_i) and against (L) in section (3) supports this conclusion insofar as knowing that each multiply recursive function is finitistic does not entail that the finitist must provide a proof that the enumeration function for that class terminates, and that being in possession of a proof that each function terminates does not entail that one knows that each function terminates. But the arguments in (GMI) suggest methods that can be utilized to overcome those difficulties such that for the k-fold recursive function under consideration, the finitist must write down an unique warrant for a proof in a formalism that the interpretation of the defined function has a calculation procedure (in (FNT)). From an epistemic point of view, it is not possible to write down infinitely many such proofs utilizing the methods suggested in (GMI), and hence, it is not possible to transcend the claim there are α-recursive functions beyond ω^{ω^ω}. Hence, from the argument and analysis in this section, it seems to follow that the methods suggested in Hilbert and Bernays' (GMI) license taking some k-fold recursive functions as finitistic for $k > 1$. But the main difference between Ravaglia's and our own is that our analysis of "bootstrapping" above leaves it open whether there are systematic reasons for finitists to be able to claim that functions beyond ω^{ω^ω} are also finitistic. Our main task in the next section is to analyze defenses of the ω-rule in order to see if methods found therein justify the claim that there are α-recursive functions beyond ω^{ω^ω} that are finitistic.

2.6.4 Higher Recursion and Hilbert's ω-Rule

In section (4.1) we saw that Ackermann's argument for methods justifying higher recursion fails. In section (4.2) we saw that the methods suggested by sections (2) and (7) of (GMI) justify some higher recursions up to ω^{ω^ω} but support nothing further. In this section we look at whether (recent and historical) defenses of recursion beyond that have finitistic justifications. Hilbert introduces the ω-rule in (1931a) with the claim that proofs for two theorems are the most important task for the finitist. These theorems are:

1. If a statement can be shown to be consistent, then it is also provable; and furthermore,

2. If for some statement $\mathfrak{S}$ the consistency with the axioms of number theory can be established, then it is impossible to also prove for $\bar{\mathfrak{S}}$ [i.e., $\neg\mathfrak{S}$] the consistency with those axioms. ((1931a), 1154)

Hilbert claims that for certain simple cases of formal theories he has successfully found finitistic proofs for (1) and (2). He argues that success had been made possible for these cases by the introduction of a new rule. Let's call this Hilbert's Rule (HR) and, following Hilbert, state it as follows:

(**HR**): Let $\mathfrak{A}(\mathfrak{z})$ be a formula in the language of (FNT), with $\mathfrak{z}$ arbitrary. Then, if $\mathfrak{A}(\mathfrak{z})$ is proved as correct for each numerical instance, then it is legitimate to introduce $\forall x A(x)$ as an axiom.

(HR) should not be viewed as a standard inference rule stated in a formal theory.[37] Instead, one ought to understand it as permitting the introduction of a universally quantified formula into a formal system if one has provided a finitistic justification for each of its instances. Hilbert claims that he obtained proofs for (1) and (2) based on "adding to the already given rules of inference (substitution and inference schema) the following equally finite new rule" (ibid.). From this it is clear that he believes that the use of this rule in finitistic consistency proofs is finitistically justified. Hence, in this section, our goal is to analyze the arguments for (or against) taking (HR) to be finitistic that have been made in historical and recent discussions.

In Hilbert (1931a) there is not much provided in the way of a defense of the status of (HR). Hilbert reminds us that "the statement $\forall x A(x)$ extends far wider than the formula $\mathfrak{A}(\mathfrak{z})$, where $\mathfrak{z}$ is an arbitrary given numeral. For in the former case not merely a numeral, but any expression of our formalism having a numerical character can be substituted for x in $A(x)$; moreover, the negation can be formed in accordance with the logical calculus" (ibid.). In addition, he provides a proof of theorems (1) and (2) for $\mathfrak{S}$ restricted to Π_1 formulae and proves (2) for $\mathfrak{S}$ restricted to Σ_1 for a system F of elementary arithmetic plus the new "rule" (HR).[38] Moreover, Sieg (2010) writes that we ought to consider the putative finitistic character of the proofs in (1931a) in contrast to methods that, in a talk given in Bologna two years earlier, Hilbert claims are non-finitistic. Sieg writes that in the Bologna talk Hilbert points out that completeness for arithmetic and analysis is usually assumed to hold and that these claims are grounded in their (second-order) categoricity. But Hilbert then claims that arguments for categoricity do not meet the demands of finitist rigor. Hence, "[Hilbert] suggests, as a next step, transforming the standard categoricity proof for number theory into a finitist argument that would establish [theorems (1) and (2)]" (Sieg (2010), 6). In (1931) Hilbert first proves that $F + HR$ is (relatively) consistent, taking for granted that Ackermann and von Neumann have proved the consistency of first-order arithmetic. He argues that if the extended

[37] With this claim we follow the interpretation of the ω-rule given by Sieg (2010). This point becomes relevant below where we discuss interpretations of (HR) in the literature.

[38] With this claim we follow the interpretation of the formal system in (1931a) given by Sieg (2010) and Ignjatović (1994). Again it is important to our discussion below of (HR).

system is inconsistent, then there is a "proof-figure" Φ obtained from the addition of (HR) whose end-formula is a contradiction. Since Φ consists in initial formulae (axioms) or formulae obtained from initial formulae by inference rules, its end-formula must be obtained from one of those. Since F is consistent, the end-formula must be obtained from (HR) and be of the form $\forall x A(x)$. By instantiation, it follows that $\mathfrak{A}(\mathfrak{a})$ for a determinate numeral. But by the antecedent of (HR) $\mathfrak{A}(\mathfrak{z})$ has been shown to be correct for arbitrary $\mathfrak{z}$. Hence, contradiction. Hence, there is no such proof-figure. Hence, $F + HR$ is consistent. Hilbert turns to the proof of (1) and (2) for "certain simple" cases:

Theorem 4.3: For each Π_1 formula, $F + HR$ satisfies (1) and (2).

Proof: Let $\mathfrak{S}$ be of the form $\forall x A(x)$, where x is the only bound variable, and let $\forall x A(x)$ be consistent with $F + HR$. If the formula $\mathfrak{A}(\mathfrak{a})$ obtained from $\forall x A(x)$ by substitution is not correct and hence not provable, then the formula $\neg\mathfrak{A}(\mathfrak{a})$ is provable and hence correct. Hence $\exists x \neg A(x)$ is provable. But then, contrary to hypothesis, $\forall x A(x)$ is inconsistent with $F + HR$. Hence, each Π_1 formula of the form $\forall x A(x)$ containing no variable other than x is provable, and (1) follows. Since $\mathfrak{S}$ is Π_1, theorem (2) follows from theorem (1) ((1931a), 1155). $\Box$

It is important to note at this point that, whatever $F + HR$ may be, it is *not* a standard formal system in Hilbert's (earlier) more traditional sense. Nor is it like (FNT) or (RNT) as understood in Hilbert and Bernays (1934). Rather, due to the fact that (HR) is formulated such that its antecedent itself requires finitistic justification that, when given, permits (by its consequent) the introduction of universally quantified statements as axioms (initial formulae) in a formal system, we might think of it as a semi-formal system.[39] Recall, from chapter one, that one of Hilbert's goals was the step-by-step construction of mathematics through consistency proofs where for each formal system proved consistent, its axioms and rules are then taken as starting points for the metamathematics of the extended system. In (1931a) Hilbert notes that Ackermann's and von Neumann's consistency proofs for arithmetic (both for systems equivalent to (PA)) demonstrate that "the transfinite modes of inference — in particular the mode of inference of *tertium non datur* — can be seen to be admissible in the domain of elementary number theory" ((1931a), 1154). Moreover, Hilbert describes the axioms as axioms for his proof theory. Hence, contra Niebergall and Schirn's (2001) claim that $F + HR$ is the object theory, it seems a more reasonable conclusion that Hilbert takes $F + HR$ to be a (semi-formal) enrichment of the metamathematical proof-theory and that his task (through the proof of consistency) is to show it is consistent in order to apply it to an object theory.

If we accept this conclusion, then the question arises as to the precise interpretation of Hilbert's extension. Consider the following. Hilbert's proof of theorems (1) and (2) for Π_1 formulae appears to be a proof that $F + HR$ is Π_1-complete, in the sense that every true Π_1-formula possesses a proof

[39] Feferman (1986) describes the system in Hilbert (1931a) as a "semi-formal" system. Here we follow that description. Though we must keep in mind in the following that it raises some interpretive objections. For a contrast see Niebergall and Schirn (2001).

in $F + HR$. There are (at least) four plausible interpretations of (HR) under which it can be made more precise. Either (i) (HR) specifies the set of all Π_1 formulae whose matrices are Π_0-formulae ϕ such that some finitistic theory F proves that $\phi(\mathfrak{n})$ is provable for each $\mathfrak{n}$. Or (ii) (HR) is interpreted as specifying the set of Π_1 formulae whose matrices are Π_0-formulae ϕ such that for each $\mathfrak{n}$ some finitistic theory F proves $\phi(\mathfrak{n})$. More precisely:

(i) $(\text{HR})_1 := \{\forall x\phi(x) : \phi \in \Pi_0 \text{ and } F \vdash \forall x\, Pr_F(\ulcorner \phi(x) \urcorner)\}$;

(ii) $(\text{HR})_2 := \{\forall x\phi(x) : \phi \in \Pi_0 \text{ and } F \vdash \phi(n) \text{ for each } n\}$.

From (i) and (ii) one obtains (i′), (ii′), (i″), and (ii″) by observing that either one takes only a *single* application of (i′) $(\text{HR})_1$ or (ii′) $(\text{HR})_2$ in the extension, or one takes *all* applications of (i″) $(\text{HR})_1$ or (ii″) $(\text{HR})_2$ in the extension. Note that we do not have a precise mathematical understanding of either $(\text{HR})_1$ or $(\text{HR})_2$ insofar as we have not specified to what F amounts. On the one hand, if F is PRA (or one of its conservative extensions), then $F + (HR)_1$ is an axiomatizable subtheory of (PA), since (PA) proves uniform reflection for PRA, and (i′) is provably equivalent to (i″).[40] It follows that $F + (HR)_1$ is not Π_1-complete as Hilbert's proof seems to demonstrate. On the other hand, if again F is PRA, then $F + (HR)_2$, if it is closed under the addition of $(\text{HR})_2$, then (ii′) is provably equivalent to (ii″). But then the extension $F + (HR)_2$ is proof-theoretically equivalent to $(QF - IA) + T(\phi)$, where $T(\phi)$ is a formal truth-definition for formulae ϕ such that ϕ are Σ_2. Since a set is recursively enumerable if and only if it is Σ_1, it follows that $F + (HR)_2$ is *not* recursively enumerable. But that seems to clash with finitist epistemology.[41] Hence, if we take F to be (PRA), then the interpretations of (HR) (i′) and (i″) pose an historical obstacle to Hilbert's textual claim, while (ii′) and (ii″) seem to pose a conceptual obstacle for finitists.

But a different situation arises if we interpret F as (PA). Under that assumption $F + (HR)_2$ for both (ii′) and (ii″) are provably equivalent, and since that $F + (HR)_2$ is provably equivalent to $(PA) + T(\phi)$ where $T(\phi)$ is a formal truth-definition for formulae ϕ such that ϕ are Π_1, $F + (HR)_2$ is Π_1-complete just as Hilbert appears to demonstrate. In fact, Hilbert seems to claim that, pace Ackermann's and von Neumann's consistency proofs for (PA), it is legitimate to extend (PRA) — what we might take as the base — in the step-by-step way outlined by his program all the way to (PA) and then extend (PA) with (HR). However, although this interpretation agrees with Hilbert's claim to have demonstrated what appears to be the Π_1-completeness of the semi-formal system, just because of this it is, again, not recursively enumerable. Hence, on this interpretation of $F + (HR)$ we have a genuine trade-off. On the one hand, since $F + (HR)$ is Π_1-complete, and since the

[40] Although it may seem anachronistic to use PRA here, it ought to suffice to note that it is possible to replace (PRA) with (QF-IA) or $I\Delta_0$, both of which are proof-theoretic conservative extensions of PRA containing the usual arithmetic axioms plus first-order logic. The definition of $F + HR$ as $(QF - IA) + (HR)_1$ above is presented in a slightly different but equivalent form (and then defended) in Ignjatović (1994).

[41] Detlefsen (1979), 310, denies that the metamathematical framework needs to be recursively enumerable. We come back to his claim below. Note that if we adopt the strategy of interpreting F as PRA and (HR) as $(\text{HR})_2$, then $F + (HR)_2$ proves the consistency of theories as strong as PA and even Zermelo-Fraenkel (ZF), if PA and ZF are consistent.

consistency formula for F is an instance of a Π_1 formula, it is possible to prove in $F + (HR)$ that F is consistent. On the other hand, since $F + (HR)$ is not recursively enumerable, the argument for F's consistency cannot be formalized in F. Hence, we obtain traction on one of HP's central goals, but no longer have traction on to what the metamathematics amounts. Sieg has found evidence that suggests that Hilbert was aware of Gödel's results and that, in part, Hilbert's (1931a) is a response to the perceived difficulties (Sieg (2010), 8). Hence, it seems reasonable to think that some working finitists concluded that such arguments count as finitistic, despite the fact that they might not be formalizable.

One cannot find many arguments for the above conclusion in the historical literature. But in the contemporary literature there are essentially two different types of arguments. One strategy, represented best by Detlefsen (1979) takes a pragmatic attitude towards the finitist's acceptance of (HR). Another strategy, best represented by Niebergall and Schirn (2001), considers the acceptance of (HR) from a conceptual point of view. But both strategies are unsatisfactory. On Detlefsen's view, since one of the main goals of HP is to prove that classical theories are consistent, the finitist is licensed in accepting those instances of Π_1 formulae that express the consistency of the theory (Detlefsen (1979), 289). Hence, in the case above, the consistency of F becomes provable in F plus the instance of (HR) that is the consistency formula for F. Since (HR) is accepted as finitistic on these grounds, Detlefsen claims that one obtains a finitistic proof of the consistency of classical arithmetic. He then admits that such additions produce theories that are not recursively enumerable, but that he sees "nothing in Hilbert's program which suggests that such formalizability is an essential or important feature of it" (ibid.). But what we must show is *not* that there is no evidence for the claim that the metatheory must be formalizable, but rather that there is positive evidence that the metatheory ought not be.[42] Niebergall and Schirn suggest that one might accept (HR) based on the fact that one instance of the rule is no more:

> problematic than to make the assumption that one can conclude from the PA-provability of '$\forall x(0 \leq x)$' to the PA-provability of '$0 \leq n$' for every n. Both cases require that *modus ponens* be applied infinitely many times, where the sequence of the prooflines has order type ω. (Niebergall and Schirn (2001), 140)

The comparison invoked is between one application of (HR) and Gentzen's consistency proof for classical arithmetic, which appeals to transfinite induction on sequences of ordinals up to ϵ_0. But the problem is that Gentzen's proof appeals to sequences of ordinals built up from below, and hence might have some claim to being finitistic, while one instance of (HR) added to F interpreted as

[42] Others have objected to Detlefsen's argument on different, often technical, grounds. Ignjatović (1994), for example, shows that in order for Detlefsen's proposal to work, one must presuppose the consistency of the theory being proved consistent. Hence, one assumes that such a theory is consistent (an assumption that might not be finitistic), and only then do we get a finitistic proof that it is. Moreover, if we follow Detlefsen's approach in (1979), it is possible to prove along these lines that *any* theory, including (PA), (ZF), and even stronger theories, is consistent. See Ignjatović (1994), 326-9.

(PA) is not recursively enumerable.[43] In sum, neither strategy for a defense of (HR) in the existing literature seems to be adequate.

It was Gentzen who recognized the deeper difficulty in Hilbert's approach to (HR). Recall that, as stated in Hilbert (1931a), (HR) contains an unanalyzed concept — that of a predicate being verified as "correct" for each numeral $\mathfrak{n}$ — and that is of course one of the main hurdles in the analysis of its status. In (1936) Gentzen proves the consistency of (PA) by replacing the concept of correctness (Gentzen refers to it as "contentual correctness" [*inhaltlicher Richtigkeitsbegriff*]) with what he calls the "stability of a reduction rule" for derivations that occur in the formal theory to be analyzed. Each derivation is assigned an ordinal number that measures its complexity, and Gentzen shows that each reduction step decreases the complexity of the derivation. Once this is shown, Gentzen proves that for each reduction step, the ordinal number of a given derivation decreases. Hence, by iterating reduction steps on a given derivation, one arrives at its end-sequent in finitely many steps. Once Gentzen has shown that only finitely-many steps are required for all derivations with ordinal numbers less than some ordinal α, it follows, by his proof for of this fact for transfinite induction, that this property holds for α itself and hence holds for *all* derivations in (PA) of arbitrary ordinal number. Gentzen claims that the universal quantifier in question must be taken as finitistic. He writes that "in each case we are dealing with a totality [the set of PA-derivations] for which a constructive rule for generating all elements is given" (Gentzen (1936), 560). Since he replaces Hilbert's informally-defined concept of correctness with his own precisely defined concept of stability, Gentzen's proof of the consistency of (PA) replaces a finitistic concept with a well-defined syntactic concept. Hence, unlike Hilbert's proof using (HR), Gentzen's need not appeal to informal finitistic concepts.[44]

In fact, Gentzen identifies precisely the problem in Hilbert's proof. On a Hilbert-style interpretation of (HR) in which the concept of correctness is replaced by a truth-definition, Gentzen writes that one might:

> proceeding from these considerations, develop a purely formal consistency proof for this part of number theory. But such a proof would have little value, for in the proof itself one would have to use transfinite statements and the accompanying modes of inference that one wants to "ground" by the proof. So the proof would not be a proper reduction, but rather a confirmation of the finitist character of the formalized rules of inference. But one must already be clear in advance what counts as finitist (in order then to be able to carry out the consistency proof itself with finitist means of proof). (Gentzen,

[43] Niebergall and Schirn recognize this, claiming that "[q]uestions concerning the strength of assumptions of infinity ought not to be conflated with questions concerning recursion-theoretic complexity" (Niebergall and Schirn (2001), 141). But in the absence of a means to measure the strength of "assumptions of infinity" the point is vacuous.

[44] Here we do not attempt to answer the question whether Gentzen's proof that (PA) is consistent using (PA) plus transfinite induction on ordinals up to ϵ_0 is finitistically justified. Gentzen claimed that it is. But our main focus here is his criticism of (HR). In (GMII) Hilbert and Bernays recognize Gentzen's objection below and make remarks suggesting that his "purely syntactic" proof overcomes the objection ((1939), 390*ff*).

(1936), 529)

His charge is twofold. First, Gentzen objects to its circularity. One must use the formal system whose consistency is at issue in the consistency proof. Second, Gentzen's deeper point is that by using (HR) one is committed to a truth-definition for the very formulae whose truth is at issue in the consistency proof. Hence, Gentzen's objection is neither that the addition of (HR) produces a set that is not recursively enumerable nor that the addition of (HR) permits methods beyond those recognizable as finitist. Both objections beg the question. For the metamathematics might not necessarily be recursively enumerable from an epistemic point of view, and we must first understand what is recognizable as finitist in order to claim that some method is not recognizable as such. Gentzen's objection is fatal to (HR) because he shows that *no* non-circular methods justify it.[45]

2.6.5 Even Further Reflections

In section (6.1) we argued that Ackermann's methods do not justify the claim that each α-recursive function up to $\omega^{\omega^{\omega}}$ is finitistic. In section (6.2) we argued that the methods suggested in Hilbert and Bernays (1934) license taking some α-recursive functions such that $\alpha < \omega^{\omega^k}$ for some k as finitistic. In section (6.3) we argued that the proposal in Hilbert (1931a) to take (HR) as finitistic fails by Gentzen's objection. Ackermann's justification fails because the procedure he outlines does not generalize to the cases beyond the function that bears his name. Hilbert's (HR) fails because it is not possible to justify it without circularity. On the other hand, we argued that some methods in Hilbert and Bernays (1934) are successful because the authors suggest that for each recursive but not primitive recursive function (i.e., for each α-recursive function beyond ω^{ω}) one must be in possession of a unique warrant that justifies taking it to be finitistic. Hence, up to now, we have argued for the claim that there are α-recursive functions such that $\alpha < \omega^{\omega^k}$ for some k are finitistic (theorem 4.2 demonstrates it for $k = 2$). In this section we look at this claim in more detail and discuss how it relates to the claims made in sections (5). Our goal is to make a first step towards providing a formal characterization of the set of finitistic theorems and proofs by making explicit the positive arguments contained in sections (5) and (6.2).

In section (5) it was argued that three conditions are mutually incompatible. That is, (L), (R_i), and the claim that finitistic reasoning is bounded above are incompatible when taken together. It was further argued that, of the three, (L) fails and that the other two conditions are its essential features. Consider (L). If it fails, then it might be the case that one is in possession of a finitistic function, but one does not know that it is total (or terminates). Finitistic constructions need not be luminous, in the sense that it might be true that the function considered is finitistic without it being known that the function is finitistic. If to know that a function is finitistic entails to have a proof that it is (as in Tait), then the failure of (L) implies that one might be in possession of a

[45] Sieg writes that Hilbert's (HR), "not clearly respecting the line between syntactic and semantic considerations, has been made blindingly clear by Gentzen" (Sieg (2010), 38).

finitistic function but it might not be provably so. Hence, if we are looking for a characterization of this condition in a formal system for finitistic reasoning, then it seems that we might be looking for formal theories that contain implicit evidence for constructing functions for which there might not be a proof that it is defined for all values in that theory. Consider, now, (R_i). If it holds, then it is the case that if at one stage of construction one knows that a function is finitistic, then the degree of justification for the claim that the next stage of construction is finitistic must be high enough to warrant knowledge at the previous stage. If, again, to know is to have a proof, then that (R_i) holds implies that if one has a proof that a function constructed at some stage is finitistic, then there must be sufficient evidence to warrant the claim (but perhaps without proof) that the next stage is finitistic. Hence, if we are looking for a characterization of this condition in a formal system for finitistic reasoning, then it seems that we are looking for formal theories of arithmetic that, when characterized as finitistic for some classes of functions, also entail that functions not included in those classes are finitistic, though not necessarily definable as such.

Our interpretation of the positive argument for (R_i) in section (3) is as follows. There, we understood it to be a means for ensuring that there are not irrational lapses in our description of finitist knowledge such that it requires of the finitist that he find radically different means of justification from the proof of one function to the proof of another function. It was also argued that in this context we ought to expect the proof relation for a theory to get stronger, in the sense that it is able to prove more the more functions that we come to justify as finitistic. Given the success of the argument from Hilbert and Bernays' theorem (4.2), then, we ought to require a unique warrant for passing from one stage to the next, and such a warrant, if it exists, ought to be a principle that allows passage from a predecessor theory to a stronger successor theory in a sequence of finitistic theories. Consider the following sequence of formal systems:

$$F^0, F^1, F^2, \ldots, F^k, \ldots$$

where k indexes the k-fold recursive functions. Reasoning, for example, within F^1 proves only that each of the primitive recursive functions is total. In order to prove that all primitive recursive functions are total, we're required to have an enumeration function that is *not* explicitly definable in F^1 but is definable in F^2. Hence, we require a means of passing from F^1 to F^2 that restricts our passage just to F^2 and blocks passage to the higher recursive (i.e., 3-fold, 4-fold, etc.) functions.[46] But the argument in section (6.2) is intended to show that there is nothing in principle obstructing us from writing down proofs that each primitive recursive function terminates, then transforming those proofs into a warrant or evidence for the claim that the 2-fold recursive derivation of theorem 4.2 above is also finitistic. But how is it possible to codify this informal discussion into one to be used as a formal proof principle? Such a principle would have to express the idea that a proof

[46] Indeed, the objection to Ackermann's approach in previous sections of this chapter was just that it allows passage to all higher recursive functions and, hence, fails as a justification in this context.

that each function is total in a given formalism permits us to infer that it is possible to define an enumeration function for the entire class. Our conjecture is that it is possible to find and formalize such a principle but, for the moment, we leave further discussion to later research.[47]

We suspect, however, that the formalization of the argument for (R_i) and against (L) in the context of our discussion of finitistic proof ought to lead to the addition of restricted reflection principles specific to the bootstrapping argument given in section (6.2) and that such theories ought to be presented in a way that reflects the fact that (L) fails. If this is correct, then we conjecture that such theories would fail to be recursively enumerable. But *that* ought to vindicate a central feature of our conception of finitistic knowledge: its limits are *inexact*. Our claim is that these limits encode facts not about the ontology of finitism, but about its means of justification. From the finitistic standpoint, while it might be possible to recursively enumerate each F^k, it might not be possible for him to put himself in a position to know that there is no recursive enumeration of all of them. Indeed, this is as it should be, for in the absence of a unique warrant for passing beyond it, our own analysis entails that it is possible for finitists to reason to ordinals α such that $\alpha < \omega^{\omega^k}$ for some k. But we've left it open if other means get us farther and, hence, the bound is inexact.

2.7 Conclusion

In section (2) we argued that Tait's view of finitism presupposes two conditions along with the claim that finitistic reasoning is bounded above. In section (3) we showed that these three conditions are jointly incompatible as a characterization of finitism. Section (4) spelled out three means of passing beyond the ordinal ω^ω and concluded that only one — the method described in Hilbert and Bernays (1934) — is a suitable method for extending the finitistic standpoint. That argument implied that an inexact upper bound to finitism is ω^{ω^ω}. But since the concepts of evidence and justification are variable, there may exist kinds of finitistic evidence that permit us to pass beyond that ordinal. Hence, in our view, the upper bound for finitistic reasoning is necessarily inexact, in that it is relative to the kind of evidence that permits us to bootstrap from one class to another in the hierarchy of recursive functions. In this respect, the view of finitism that has been defended here resembles not only Hilbert and Bernays (1934) but Gödel's early (1931) view as well. In (1931) Gödel seems to claim that finitistic reasoning is open-ended because the concept of a finitistic proof is informal and hence not well-defined (Gödel (1931), 198). By contrast our argument suggests that finitistic knowledge is open-ended or inexact because the conditions of finitistic knowledge are such that they fail to permit perfect transparency, from the construction of the finite ordinals all the way to higher recursions. One might have a well-defined formalism in which there are proofs that k-fold recursive functions terminate but for each one there might yet exist a $k+1$-fold recursive function enumerating the k-fold functions not definable in that formalism which is nonetheless well-defined.

[47] Note that a *finitistic proof* in such a formal system would either be a normal derivation within that formal system, or a derivation involving the use of the formalized version of such a rule.

One of the central claims in this chapter has been that as one ascends the hierarchy of recursive functions, from the initial functions to the k-fold recursions, the degree of justification a finitist has for the claim that a given function is finitistic decreases. In the conclusion we discuss the proposal above as it relates to the degrees of justification finitists have for recognizing them as finitistic.

To be clear, the claim here is not that finitistic reasoning is unbounded. As we shall see in what follows, there is a bound, albeit a fuzzy one, but that bound does not arise because we have arbitrarily stipulated that the finitist cannot reason in certain ways due to the "nature" of finite sequences or because, like Tait's Cartesian, a certain set of procedures is "indubitable." Upper bounds to finitistic reasoning arise due to the finitist's methods of justification. Roughly speaking, as he ascends the hierarchy of recursive functions, the degree to which his claims count as finitistic decreases, a ratio that corresponds to the intuitive idea that as his arguments and proofs grow increasingly complex, the degree to which he is justified in believing them to be finitistic decreases. (And this idea corresponds to the intuitive idea that it is more difficult to understand how longer, more computationally complex proofs justify the theorems of which they are proofs.) More precisely, let j represent the justification one has for a belief b; let greek lower case letters $\phi, \psi, \ldots$ (with or without subscripts) stand for the contents of beliefs. Then $d(j(b(\phi)))$ is the degree of justification of the belief that ϕ where:

$$d \in \mathbb{R} \text{ and } 0 \leq d \leq 1.$$

Let's say that a justification j for the belief b that ϕ is *weakly* finitistic just in case $d \to 0$. Call a justification *strongly* finitistic just in case $d \to 1$. On the other hand, since $k \in \omega$ measures the level of the hierarchy of recursive functions, and gives rise to the concept of levels of finitistic proofs in the guise of k-finitistic provability from definitions 4.5 and 4.6, then as k increases d decreases, and as k decreases d increases. Hence, as $k \to \omega$, $d \to 0$, in which case the method defended in section (4) justifies the claim that $F+$ is finitistically the weakest relative to that method of justification.

On our account, with each k-fold recursive function that counts the level of the hierarchy of recursive functions is associated a real number r that measures the degree of justification for the claim at that level. If $k = 0$ for some function defined by a formula ϕ_k, then d comes as close to 1 as one likes in the limit but it is never the case that $d = 1$. For that implies that the finitist has perfect justification for his claim regardless of the evidence he has for it, and as such is an instance of (L). Moreover, the situation is similar as k increases. It is in this sense that finitism is open-ended. For there is never a point at which the finitist has no justification whatsoever for the claim that some function is finitistic. He might have an exceedingly small degree of justification, but it is not a strict cut-off, since he might find evidence that makes it possible for him to justify formalisms beyond $F+$ that is unlike the kind of evidence that permits him to justify the functions definable in F^k for each k as above. Hence, it ought to come as no surprise that our epistemic account and its formal characterization entails that the set of all sentences that are provable k-finitistically is not recursively enumerable. For as the finitist ascends the hierarchy of k-fold recursive functions he loses the power to provide a high enough degree of justification to warrant full knowledge for his claims. As he ascends the hierarchy of recursive functions, the proof relation for the formalism in which he works becomes stronger. But since his degree of justification decreases for it, he loses the power to apply it as widely as at lower levels of the hierarchy. Hence, the set of sentences *logically* entailed by his bootstrapping procedures outruns the set of sentences that are *epistemically* possible for him.

Now, it might be objected that we have failed to provide a strict upper bound for finitistic reasoning. Our response to this objection is that our account provides us with a principled explanation as to why finitists never specified *the* principles of finitistic reasoning. Such principles unfold as one develops further methods of justification for passing beyond the methods justified in a single formalism.[48] One lays down a set of formal axioms and rules in which to carry out the metamathematics of another system and, as one develops proofs in that formalism one extracts the principles implicit but not explicitly provable in that formalism. Our claim is that such inexact limitations reflect the deep structure of the finitist's epistemology. Finally, let's contrast our argument with Tait's once more. In Tait's view, what underlies the finitist's means of knowledge acquisition is the concept of Number, by which he means the concept of an arbitrary iteration over a finite alphabet. Our argument from this chapter shows that, strictly speaking, iteration over a finite alphabet never yields any principles of mathematical interest but only yields finite numbers $k < \omega$. On the other hand, *arbitrary* iteration over a finite alphabet yields any procedure that generates a numeral given a numeral, and it is this that leads to Tait's view.[49] But if we admit this, we have already admitted both iteration over a finite alphabet and the possibility of reflecting on iterations over a finite alphabet. We have admitted something like a reflection principle for finite sequences, which entails that we have admitted procedures that generate procedures that generate numerals given numerals. Hence, we have admitted the possibility of generating k-fold recursive functions given a specification of each $k-1$-fold recursive function. It is this, what Sieg (2009) calls the finitist's "balancing act" between his self-imposed restriction to the finite and his evidential capabilities to reflect further on what is implicit in the finite, that leads to the view defended in this chapter, that there are methods that justify the claim that there are α-recursive functions such that $\alpha < \omega^{\omega^k}$, for some k, that are finitistic and that there might be "bootstrapping" methods utilizing reflection principles that permit passage beyond ω^{ω^ω}. One specifies the base by considering concrete sequences alone, then one specifies extensions by reflecting over what is implicit in those sequences. In our final chapter, this point comes into focus a bit more as we explore how our view is situated with respect to a finitistic proof of the consistency of classical formal theories of arithmetic in the face of (G2).

[48] One question that arises here is whether the so-called Bar Rule provides a means of passing to any of the formalisms beyond (PRA) but below (PA). If not, then if finitists are to reason in formalisms whose ordinals are above ω^{ω^ω} but below ε_0 then the bootstrapping principle involved cannot be the Bar Rule, and such a theorem could be further evidence for the view defended in this chapter.

[49] See also Tait (2010). In it he defends his original claim anew against Hilbert and Bernays (1934) and especially Gödel's views of finitism. In particular, he claims that Gödel simply doesn't understand the "finite" in finitism. But both claim that arbitrary iterations ground finitistic reasoning. The only difference is that Gödel seems to admit some arbitrary procedures of higher type.

Chapter 3

Conclusion

3.1 Introduction

In chapter two we provided an interpretation and defense of finitistic epistemology. Our claim in that chapter is that the type of *a priori* justification that supports finitistic epistemology is "quasi-empirical" and admits of degrees. We argued that different types of procedures permit finitists to pass from the base — variable-free claims — to one of the six subtheories of Q — with claims involving free-variables — obtained by omitting one of its seven axioms. Then we deepened the analysis in chapter two by isolating three claims that authors in the literature on finitism seem to adopt in its analysis — (L), (R_i), and the claim that finitistic reasoning is strictly bounded above — and it was argued that of these three (L) fails. Our claim is that different types of finitistic evidence justify different types of extensions to the finitistic standpoint. We considered three types of extension procedures from the history of finitism and proof-theory, and concluded that the strongest is the "bootstrapping" procedure found in Hilbert and Bernays (1934). Then, we proposed a formal characterization for finitism that captures our analysis that entails, in contrast to Tait's and Kreisel's views, that each α-recursive function such that $\alpha < \omega^{\omega^{\omega}}$ is finitistically justified. Hence, up to now we have provided an interpretation and defense of finitistic epistemology, and a formal characterization for its lower and what we claimed are "inexact" upper bounds. But, as we'll recall from our discussion in both preceding chapters, one of the main goals for the finitist is to provide a finitistic consistency proof for classical mathematics. Our remaining task in this our concluding chapter, then, is to discuss whether this is even possible in the face of Gödel's second incompleteness theorem (G2).

3.1.1 Setting the Stage

Gödel's first incompleteness theorem (G1) states that for any consistent effectively enumerable formal theory T that axiomatizes a sufficient amount of arithmetic there is a formula ϕ in the language of T that is true in its model but neither provable nor refutable using T's axioms and inference rules. (G2) states that, if T is consistent, it is possible to produce for ϕ a formalization of the assertion that T is consistent such that the formula is true in its model, but again neither provable nor refutable in T. In contrast to (G1), (G2) concerns an inherently metamathematical statement that is central to the finitist's goal of constructing a finitistic consistency proof for classical arithmetic. In the tradition following Gödel's first and second incompleteness theorems, with few exceptions philosophers and logicians have claimed that (G2) entails that for any T satisfying the above conditions T does not prove its own consistency, and that, again, no consistency proof of T is formalizable in T. Hence, for example, if T is (PRA), and (PRA) is thought to capture all of finitistic knowledge, then it is claimed that (G2) entails that no finitistic theory can prove its own consistency, and that no finitistic consistency proof of T can be formalized in T. Though the theorem in Gödel (1931) is somewhat different, for the moment, let's refer to (G2) as found therein as:

$$\text{(G2)} \qquad T \nvdash Con(T)$$

and distinguish it from the claims:

1. that T does not prove its own consistency; and
2. that no consistency proof of T is formalizable in T.[1]

One task in this chapter is to understand if (G2) justifies claims (1) or (2). Once we do, we are in a better position to see how (G2) affects (HP).

But let's back up somewhat. Suppose, for the moment and contrary to fact, that (G2) had never been proved or discovered, that nobody had claimed (1) or (2), and that researchers in the Hilbert School had met their goal in constructing a finitistic consistency proof for classical arithmetic and analysis. Recall that one of Hilbert's original motivations for a consistency proof was epistemological. Hilbert's Program is designed to "save" classical mathematics from its intuitionist critique. If one has a finitistic consistency proof for classical mathematics, then one is justified in believing the axioms shown to be consistent to the same degree as one believes the finitistic claims. But how does a proof of consistency legitimate belief in the axioms? Not only should our beliefs be consistent, but in addition they ought to track *truth*. Of course, under certain assumptions, a consistency proof entails that the axioms (the so-called ideal formulae) do not prove the negation of any finitistically true assertions (the so-called real formulae) such as $0 = 0$. But while this might explain why the axioms are sound for claims that the finitist takes to be justified, it does not explain why the classical

[1] Franks (2009) calls (1) and (2) the "Gödelian inferences" since in (1934) Gödel himself was the first to draw them like so. But since most philosophers and logicians post-1931 make them, we'll refer to them as above.

axioms are true of the phenomena that they purport to describe.[2] We desire an explanation of how the classical axioms track truth, or failing that, how consistency is sufficient from an epistemological point of view. Hence, our second task in this chapter is to understand how, if at all, consistency is a sufficient condition for belief. Once we do, we are in a better position to understand how (HP) provides a feasible approach to the Benacerraf dilemma.

3.1.2 Outline & Goals

Our strategy is as follows. In section (2) we look at claims (1) and (2) as interpretations of (G2). It is argued that on an extensional interpretation of (G2) and its generalizations, (1) and (2) do not follow. On an intensional interpretation, claims (1) and (2) are immediate. But intensional interpretations raise the issue of the intensional correctness of the consistency formula, since on such an interpretation, T must recognize the consistency formula as *the* formula that expresses its own consistency. Our question then concerns the correct conditions under which T recognizes its own consistency. Our main claim is that the conditions for capturing the intensional correctness of a consistency formula proposed in Feferman (1960) are insufficient. Since this opens the door to alternative conditions, in section (3) we explore some of these, and since this raises the question of their significance, in the conclusion we discuss the epistemological significance of consistency. It's here that the discussions from chapters one and two and the previous sections of this chapter converge. We argue that our "epistemic" defense of finitism and (HP) is a tractable alternative to current paradigms. In particular, we have argued that it is possible for finitists to define functions in an informal finitistic theory that, if formalized, is proof-theoretically stronger than the formalism that captures his working concept of proof, and that this amounts to a partial overlap between the concepts of provability and truth. What it is possible to know and what the sentences are true of overlap, in the sense that such a formalism is able to recognize that its provable sentences (the "known" sentences) form a subclass of the sentences true (the "assertible" sentences) of each complete extension of itself. But this leads to a natural problem: find the proof-theoretically weakest formal theory that fails to recognize the properties true of complete extensions of itself for such a theory constitutes the upper bound for the mathematics that can serve as a solution to Benacerraf's Dilemma. Since the semantics outruns the proof-theory, what the formalism is true of ontologically diverges from what is provable from an epistemological point of view. Hence, philosophical explanations of mathematics that go beyond such a theory must answer (BD) head on.

[2] In order to see why in more detail, let T be a first-order theory of the arithmetic of the natural numbers. If T is consistent, then by the completeness theorem for first-order logic T has a model. By the upwards Löwenheim-Skolem Theorem, T has an uncountable model, containing, e.g., real numbers. Since T's axioms are intended to describe only the natural numbers, but there is a model of T in the real numbers, it might be argued that T provably does not track the truth.

3.2 Interpreting (G2)

In this section we look at claims (1) and (2) above as consequences of interpretations of (G2). The basic problem is as follows. Feferman (1960) claims that, unlike (G1), interpreters of (G2) are confronted with a unique problem. Now (G1) states that for any formal system T, there is a formula true in a model of T but not provable using T's axioms and rules of inference. If formalized this statement only expresses the fact that *some* sentence is true but formally underivable in T, but does not tell us which sentence, nor if T need recognize the sentence that is formally underivable. Insofar as T contains some primitive recursive arithmetic, then it is possible to recursively arithmetize a sentence for T that satisfies (G1). But for the proof of (G2) T must recognize that the sentence that is formally underivable is a sentence that expresses, or "says" that T is consistent. Feferman writes that:

> [i]n broad terms, the applications of the method [of arithmetization] can be classified as being *extensional* if essentially only numerically correct definitions are needed, or *intensional* if the definitions must more fully *express* the notions involved, so that various of the general properties of these notions can be formally derived. (Feferman (1960), 35)

For examples of the extensional type Feferman lists (G1), the undefinability of truth (and other predicates) in formal theories, and the undecidability and degrees of unsolvability results for various theories. For examples of the intensional type he lists (G2), results of relative consistency strengths between theories, and logics for ordinal analysis. For results of the first type one need only know that arithmetization of the metamathematical concept picks out a unique numerical class. But for results of the second type one must know that the numerical class picked out correctly expresses the metamathematical concept. Since our analysis of this distinction will aid us in understanding how, if at all, claims (1) and (2) are consequences of interpretations of (G2), in what follows we sharpen it.[3]

3.2.1 Intension and Extension

In order to sharpen up the distinction between extensional and intensional applications of arithmetization, let's first consider a few concrete examples. Gödel (1931) states the theorem for the ω-consistency of a formal system. For the system that Gödel studied, *Principia Mathematica* (PM), it is said to be ω-consistent just in case when (PM) proves $\phi(n)$ for all n, where ϕ is an arithmetical formula, then $\phi(n)$ is true in the standard model of arithmetic for each n. After Gödel, Rosser (1936) strengthened Gödel's result and along the way found a predicate distinct from the one that expresses the proof relation "x is a proof in T of y" for a theory T to obtain extensionally improved theorems concerning provability in metamathematics. Let:

[3] By "picking out a numerical class" is meant either bi-numerate or numerate. See Feferman (1960), 51.

$$Prf_T(x, y)$$

be a numerically correct definition of the proof relation for T. Then the Rosser predicate is the relation $Prf_T^R(x, y)$ defined by:

$$Prf_T(x, y) \wedge \neg\exists z\, [z \leq x \wedge Prf_T(z, neg(y))].$$

Under the assumption that T is consistent, then $Prf_T^R(x, y)$ and $Prf_T(x, y)$ are extensionally equivalent. But the Rosser proof predicate says that x is a proof of y in T such that there are no proofs z in T shorter than x that prove the negation of y, and therefore fails to express the proof relation in T, which is instead expressed by $Prf_T(x, y)$. Hence, while $Prf_T^R(x, y)$ and $Prf_T(x, y)$ are extensionally equivalent, they are not intensionally equivalent. Here is, then, an example in which it is possible to use predicates with "non-standard" intensions in order to obtain extensional improvements (in the sense of a wider scope) upon existing theorems.

In a similar direction, let T be a formal system containing some arithmetic, let $Prf_T(x, y)$ be a numerically correct definition of the proof relation as above, and consider the relation $Prf_T^C(x, y)$ defined by:

$$\begin{gathered}\Big(Prf_T(x, y) \\ \wedge \\ \neg\exists u \exists w \exists z \big((u \leq x \wedge w \leq x \wedge z \leq x) \\ \wedge \\ (Prf_T(u, z) \wedge Prf_T(w, neg(z)))\big)\Big).\end{gathered}$$

Again if T is consistent, then $Prf_T^C(x, y)$ is extensionally equivalent to $Prf_T(x, y)$. If we now define a consistency predicate, $Con^C(T)$, as:

$$\neg\exists x \exists y \exists z \big(Prf_T^C(x, z) \wedge Prf_T^C(y, neg(z))\big)$$

then, since the above is an instance of a first-order validity, it is possible to prove $Con^C(T)$ in T. But does that sentence express the consistency of T? Feferman claims that it does not. In his view the predicate $Prf_T^C(x, y)$ is "intensionally incorrect, so we can ascribe no clear intensional meaning to the result." Moreover, in contrast to Rosser's manipulation of the intension of the definition of provability in order to obtain an improved extensional result, with this definition "we cannot formally derive other properties of provability in terms of the definition [$Prf_T^C(x, y)$]," and hence "we see no results of extensional interest which follow from the proof of [$Con^C(T)$]" (Feferman (1960), 37). Note, though, that there are two claims implicit in Feferman's argument. First, he suggests that it is we who ascribe the intensional meaning to a metamathematical result. Second, he suggests that if no fruitful consequences of "extensional interest" result from the use of predicates with "non-standard" intensions, then that counts as evidence against believing that the predicate expresses a meaningful concept. Let's return to these in a moment, once we've further sharpened the problem.

3.2.2 Sharpening the Problem

Franks (2009) makes the distinction between extensional and intensional interpretations of metamathematics by arguing that on an extensional interpretation, statements about formal systems are "theory-independent" in the sense that such statements are independent of what the formal system counts amongst its theorems and proofs. Hence, in the case of a consistency sentence, it is possible for there to be, from our "theory-independent" point of view, various extensionally equivalent means of expressing the consistency of a formalism regardless of whether the formalism itself proves them to be equivalent. But on an intensional interpretation of metamathematics, "statements about mathematics [formal systems of mathematics] are always part of a mathematical theory," or "theory-dependent," in the sense that statements about the system depend entirely upon what is provable in it (Franks (2009), 7). Hence, on an intensional interpretation, a function is not defined unless the formalism proves that it satisfies the relevant existence and uniqueness conditions, two formulae cannot be counted as equivalent unless the formal system proves it, and a formal system proves a formula just in case *that* formal system proves the "correct" formalization of the statement that it proves the formula. But note that both extensional and intensional interpretations presuppose standards of correctness. From the extensional point of view, that standard is tied to whether a metamathematical predicate picks out extensionally identical numerical classes, independent of the details of the formal system for which it is defined. From the intensional point of view, it is tied to how a formal system proves (or fails to prove) that a formula is a correct expression of the metamathematical concept *for* that particular system, and hence, the standard varies dependent upon the details of the formal system for which the predicate is defined.

For example, Hilbert and Bernays (1939) set out three conditions aimed to furnish standards for the intensional correctness of a proof predicate by isolating the conditions that the proof predicate must satisfy in order to derive the second incompleteness theorem. In order for a proof predicate to be intensionally correct, T must satisfy:

(HB1) if $\vdash_T \phi$ then $\vdash_T Prov_T(\ulcorner \phi \urcorner)$;

(HB2) $\vdash_T Prov_T(\ulcorner \phi \to \psi \urcorner) \to Prov_T(\ulcorner \phi \urcorner) \to Prov_T(\ulcorner \psi \urcorner)$;

(HB3) $\vdash_T Prov_T(\ulcorner \phi \urcorner) \to Prov_T(\ulcorner Prov_T(\ulcorner \phi \urcorner) \urcorner)$.[4]

While (HB1) ensures that T recognizes that its proofs are recursively enumerable, (HB2) ensures that T recognizes that all its proofs are closed under *modus ponens*, and (HB3) ensures that T recognizes that all its provable theorems are provably provable. Hilbert and Bernays show that these three conditions are sufficient for the proof of (G2). Hence, one need only verify that a formal system satisfies these conditions in order to prove (G2) without being required to provide an explicit

[4] In Hilbert and Bernays (1939) the derivability conditions stated are slightly different than the three conditions stated above, which are the improved conditions found in Löb (1955). For the original reference see Hilbert and Bernays (1939), 295*ff*.

definition of the proof predicate. However, there are at least two problems with their approach. First, their motivation is to find the conditions sufficient for (G2). But if the possibility of a consistency proof is in question, then it cannot be the case that the standards for intensional correctness are just those conditions that permit the proof of (G2), since that begs the question. Second, while it is true that given a formal system that satisfies these conditions one may assume that (G2) is provable without an explicit definition of proof, it is not true that given an explicit definition of proof for a formal system, there is a general method for testing whether it satisfies these conditions. Hence, these conditions are instrumental for having the resources required to prove (G2), but fail to fix the standard for correctness according to how it varies with the details of the particular formal system for which it is defined.

In response to this, Feferman (1960) proposes a general approach to intensionality that follows Hilbert-Bernays, but attempts to overcome its limitations. Feferman identifies a set of conditions that the proof predicate must meet such that it correctly expresses the concept for any formal system for which it is defined. Given a formal system T, one associates with it the class of all formulae $\tau(x)$ that numerically define the set of axioms of T. Then, by formalizing the concept of *logical* proof, one associates with each τ a formula $Prf_\tau(x, y)$ and a sentence Con_τ. Hence, on his approach, the proof relation is fixed by the set of non-logical axioms of a formal system, and the proof and consistency predicates are built out of the concept of logical proof underlying those axioms. He writes that on his approach "whenever a formula $\tau(x)$ can be recognized to express correctly that x is an axiom of T, the associated sentence Con_τ will be recognized to be a correct expression of the proposition that T is consistent." Due to the generality of the approach, "all intensionally correct statements of consistency for familiar theories can be obtained as special cases" (Feferman (1960), 38). Any changes made to the formula must be from the "inside," that is, made by varying the proof definition based on choosing subclasses τ' of the class of formulae τ. Feferman contrasts his approach with approaches like Rosser's that, he claims, are changes made from the "outside," a designation that is intended to suggest that, while useful in some respects, such changes are artificial because they involve changes in the concept of logical proof.

3.2.3 Problems for Interpreting (G2)

Thus far we have looked at some examples of how the extension and intension of a proof predicate come apart, and looked at two examples of approaches to providing conditions that allow for some generalizations of (G2). Our question is now whether the extensional or intensional approaches entail claim (1) or (2) or both. Hence, the basic questions are whether the extensionalist is able to secure the inference from his interpretation of (G2) to claims (1) and (2), and whether the intensional approach defended by Feferman (1960) is able to secure the inference from his interpretation of (G2) to claims (1) and (2). For the extensionalist, the task is to find general versions of (G2) such that the conditions that it must meet in order to be derived are met by all correct formalizations of consistency. Detlefsen (1986) calls this the "stability problem," and writes that the extensionalist must "show that every set of properties sufficient to make a formula a fit expression of T's consistency is also sufficient to make that formula unprovable in T (if T is consistent)" (Detlefsen (1986), 81). Hence, if some correct formalization of consistency does not meet the conditions, then the inference from (G2) to (1) or (2) is "unstable." On the other hand, the intensionalist must find a set of necessary and sufficient conditions for metamathematical predicates such that a T-proof of those conditions amounts to a demonstration of the intensional correctness of those conditions in T. But here the problem is slightly different, since the intensionalist must, in addition, explain how the proposed conditions are constitutive of the metamathematical concept.

How does the extensionalist face his task? Suppose, for the moment, that his standard of correctness is that the predicates must be numerically correct in the sense that they pick out identical numerical classes. Consider again the proof predicate $Prf_T^C(x, y)$. Above we mentioned that it is numerically correct, and that T need not recognize (in the sense of prove) that the two predicates are equivalent but that it suffices for us, from our theory-independent point of view, to be able to recognize the equivalence. If one builds the "usual" consistency statement out of $Prf_T(x, y)$:

$$Con(T) := \neg\exists x Prf_T(x, \perp)$$

then while T fails to prove $Con(T)$, it does prove $Con^C(T)$. That is, while (G2) goes through for $Con(T)$ it does not go through for $Con^C(T)$. Hence, since there is a T-proof of $Con^C(T)$ and since it is formalizable in T, the inference from (G2) to (1) and (2) fails if the extensionalist's standard is that a proof predicate and consistency statement *merely* be numerically correct. It seems that there are two ways to rectify the situation. First, he might argue for the claim that there is a theory-independent concept of consistency expressed by $Con(T)$ that $Con^C(T)$ fails to express. But this approach requires that he have in hand a theory of meaning or content that supports his claim. Hence, accepting or rejecting $Con(T)$ as expressing uniquely the consistency of T will depend on our inclination to accept or reject the underlying theory of meaning. Or, he might argue that some standard other than numerical correctness governs the arithmetization of metamathematical concepts. But this approach changes the standard because of (G2), and so begs the question in that it assumes that the inference from (G2) to (1) or (2) is valid rather than establish it on independent grounds.

It might be the case that there are other means for the extensionalist to defend the inference from (G2) to (1) or (2), but the prospects do not seem encouraging. On the other hand, on the intensional interpretation of (G2), the inference from it to (1) or (2) is immediate. To understand why, consider the following. Gentzen's *Hauptsatz* asserts that for every classical proof there is a corresponding "cut-free" proof that does not use classical indirect proof methods (that might be longer, but combinatorially simpler). It is possible, then, to build a "Gentzenian" proof predicate $Prf_T^G(x, y)$ that is extensionally equivalent to $Prf_T(x, y)$. Hence, if the cut-rule is admissible for T, then it is possible to build two formulae, $Con(T)$ and $Con^G(T)$, that are extensionally equivalent.

However, depending upon our choice of formalism, T might be proof-theoretically rich enough to arithmetize $Con(T)$, but not be rich enough to prove the formalization of the *Hauptsatz*, and hence, T might not prove that $Con(T)$ and $Con^G(T)$ are equivalent. Hence, from the point of view of some formalisms, $Con(T)$ and $Con^G(T)$ do not express the same concept. It follows that at most one of the two predicates is intensionally correct for those formalisms. Hence, if such a formalism does not prove one of the two formulae, and recognizes the one not provable as an expression of its own consistency, then (1) follows immediately from (G2) for this formalism. Moreover, for the intensionalist there is no other option than to prove in the formalism that the predicate that expresses that formalism's consistency is unprovable, and so (2) follows from (1). It seems, then, that on an intensional understanding of metamathematical predicates, (1) and (2) follow immediately from (G2) for specific formalisms. But our questions are now twofold. First, is it possible to obtain results that do not depend upon the details of specific formalisms? Second, what are the standards that constitute a formula's intensional correctness?

In Feferman's view, it is possible to answer these two questions at the same time. He argues that the goal is to find conditions that are constitutive of metamathematical concepts that permit generalizations of (G2) such that one need not restrict the arithmetization of provability to specific formalisms. Predicates that meet said conditions permit the inference from (G2) to (1) and (2). Let $Proof_T(x)$ represent the unary predicate "x is a proof in T;" let $Prf_T(x, y)$ represent the binary predicate "x is a proof in T of y;" and let $Pr_T(y) := \exists x Prf_T(x, y)$ represent the predicate "y is a theorem of T". In order for a provability predicate for T to be intensionally correct, Feferman claims that it must satisfy the following conditions:

(i) $\vdash_T \forall\phi\forall\psi\,(Pr_T(\ulcorner\phi\urcorner) \wedge Pr_T(\ulcorner\phi \rightarrow \psi\urcorner) \rightarrow Pr_T(\ulcorner\psi\urcorner));$

(ii) $\vdash_T \forall\phi\,(Pr_T(\ulcorner\phi\urcorner) \rightarrow Pr_T(\ulcorner Pr_T(\ulcorner\phi\urcorner)\urcorner));$

(iii) $\vdash_T \forall x(\ulcorner Proof_T(x)\urcorner \rightarrow Pr_T(\ulcorner Proof_T(x)\urcorner));$

(iv) $\vdash_T \forall\phi\forall x(\ulcorner Prf_T(x, \ulcorner\phi\urcorner) \rightarrow \ulcorner Pr_T(\ulcorner Prf_T(x, \ulcorner\phi\urcorner)\urcorner)\urcorner)\urcorner).$[5]

For each condition, if T satisfies it, then T expresses a concept constitutive of provability. Condition (i) expresses the concept that theoremhood is closed under *modus ponens*. Condition (ii) expresses the concept that provability is idempotent, or more suggestively, transparent. Condition (iii) expresses the concept that all T-proofs are formalizable in T. Condition (iv) expresses the concept that all T-proofs of closed formulae are formalizable in T. Feferman claims that a "minimal" standard governs the choice of conditions on intensional correctness. At the least, such conditions must preserve the logical — in contrast to the mathematical or finitistic — steps in a derivation. Hence, for Feferman, provability in T reduces to logical provability.

[5] Note that the list above is incomplete, but captures Feferman's most important conditions that a formal system must meet in order for its provability predicate to be intensionally correct. For the original list and further reference, see Feferman (1960), 60*ff*. Also note that, for simplicity, reference to Gödel numbering in the predicates is omitted.

For example, Feferman derives an important consequence from his approach that shows how conditions (i)-(iv) preserve a derivation's logical steps. Theorem 5.9 (Feferman (1960), 68) demonstrates that it is possible to choose a subformula $\tau*$ of the class of formulae τ that strongly represent the axioms T of a consistent recursive extension $\mathcal{T}$ of (PA) such that:

$$\vdash_{PA} Con^{\tau*}(\mathcal{T}).$$

In other words, there is a formula that is a numerically correct (strong) representation of the consistency of a set of axioms extending (PA) such that (PA) proves the extension consistent. On the face of it, theorem 5.9 (and corollary 5.10) appears to contradict (G2). But Feferman argues that the appearance is just that, and that his approach makes the problem clear. He writes that "one particular conclusion we can draw is that the formula [$\tau*$], although it extensionally corresponds to [T], does not properly express membership in [$\mathcal{T}$]" (Feferman (1960), 69). That is, $\tau*$ is extensionally correct but intensionally incorrect because it fails one of the conditions (i)-(iv), specifically condition (ii). Hence, we have the following:

Theorem 2.0: If $\vdash_{PA} Con^{\tau*}(\mathcal{T})$, then $\nvdash_{PA} Pr_{PA}(\tau*) \rightarrow Pr_{PA}(Pr_{PA}(\tau*))$.

Moreover, since theorem 2.0 is formalizable in (PA) by conditions (iii) and (iv), and (PA) is complete for the proof-predicate, we have the following:

Corollary 2.1: $\vdash_{PA} \Big(Con^{\tau*}(\mathcal{T}) \rightarrow \big(Pr_{PA}(\tau*) \wedge \neg Pr_{PA}(Pr_{PA}(\tau*))\big)\Big)$.

Both say that (PA) itself recognizes that there are intensionally incorrect formulations of provability that are nonetheless extensionally correct. That is, if it proves its own consistency, then it *knows* that it has failed to produce a predicate that expresses the correct concept. For Feferman, it follows that it knows that it has failed to preserve the logical steps in its derivations.[6]

There are at least three points in Feferman's analysis to which one might apply some pressure. Franks (2009) identifies two. He claims that one might demand a defense of Feferman's conditions (i)-(iv) on the "grounds that they seem from one point of view rather strong and from another point of view too weak" (Franks (2009), 122). Feferman's conditions are "too weak," for Franks, because "certain basic properties about provability do not appear in Feferman's list, and in fact cannot" (ibid.). He cites reflection principles:

$$(Ref) \qquad Pr_T(\ulcorner \phi \urcorner) \rightarrow \phi$$

that express the concept that all provable formulae are true and that, by Löb's theorem, are unprovable in formalisms satisfying conditions (i)-(iv). But the objection misfires because Feferman makes no claim for his conditions governing all possible properties, but rather just those

[6] Feferman's approach clearly permits the inferences to (1) and (2). Feferman pursues arithmetization as it appears within formalisms irrespective of a predicate's numerical correctness. Hence, it is possible to claim that if a consistency predicate for a formalism is intensionally correct, then it is unprovable in that formalism, from which (1) follows; and that there are consistency proofs formalizable in T only if the consistency predicate's formalization is intensionally incorrect, from which (2) follows.

properties that express the logical concept of proof. By contrast, (i)-(iv) are "too strong," for Franks, because if ϕ is provable, by condition (ii) there exist infinitely many (Gödel) numbers $\#\phi, \#Pr_T(\phi), \#Pr_T(Pr_T(\phi)), \ldots$, and that "seems like an ontological assumption very far removed from the notion of ϕ's provability" (Franks (2009), 123). But this objection misfires as well. The formulae obtained by iterating condition (ii), given ϕ's provability, are syntactic, and hence, make no claims about the ontology of numbers. At best, such formulae only make claims in a model-theoretic interpretation of the formalism, and even then, it is unclear how formalisms are committed to ontology. Hence, Franks is not entitled to conclude that "Feferman's proposal is incomplete as a method for the explicit arithmetizations needed for a fully mathematical treatment of metatheory" (ibid.). In what follows, we begin to develop our proposal alongside Feferman's third pressure point.

3.3 Inside Consistency

Thus far, we introduced the distinction between (G2) and what it "says," that is, claims (1) and (2). Then, we analyzed the distinction between extensional and intensional interpretations of metamathematics as it arises in discussions of inferring claim (1) or (2) from (G2). In section (2.2) it was argued that the prospects for inferring claims (1) and (2) from (G2) on an extensional interpretation of metamathematics are poor. Then it was argued that on an intensional interpretation of metamathematics, the prospects are much better, but that the intensionalist must, in addition, explain how the chosen conditions constitute the metamathematical concept being arithmetized. Then we showed that two objections by Franks failed to undercut two pressure points in Feferman's approach, and suggested that there is at least a third point. In section (2.1) we saw that Feferman suggests that if no fruitful consequences of "extensional interest" result from the use of predicates with "non-standard" intensions, then that counts as evidence against believing that the predicate expresses the correct concept. But note that his suggestion opens up a route to the following alternative approach to arithmetization. Are there changes in the logical concept of proof, i.e., "changes from the outside," that might be warranted by how formalisms generate "fruitful consequences" relative to our choice of provability predicate? In this section our goal is to begin to develop an approach to arithmetization the main feature of which is that one is warranted to believe that a consistency predicate expresses the correct concept, even if it skirts (G2), on the condition that its use generates fruitful consequences. Our claim is that these consequences constitute a unique kind of "internal" mathematical evidence for intensional conditions that express different concepts of proof.

3.3.1 Evidence and Arithmetization

Let's return, for the moment, to Feferman's claim that the standard for the choice of conditions that govern a predicate's intensional correctness is that they must express the logical concept of proof. Other conditions might express a concept, but in Feferman's view, unless it meets all four conditions listed above, that concept fails to be the logical one. Feferman (1960) does not discuss in much depth the reason behind his claim that the concept expressed must be the logical concept, nor does he discuss why conditions (i)-(iv) jointly express it. However, in the formalisms that Feferman studies every proof predicate is either a "change in the logical concept of proof," or it is provable, in the formalism, that it is equivalent to the "standard concept," and hence, in Feferman's view intensionally correct. For these formal arithmetics, the "standard concept" is the concept of a formal deduction: a sequence of formulae that are instances of axioms or obtained from the axioms by applying the inference rules finitely many times. Moreover, because the arithmetics Feferman studies are strong arithmetics such as (PA), variations in the formalism's proof theory are insignificant.[7] Implicitly, then, it seems that Feferman is committed to the claim that the concept of proof that is expressed by a formalism need not capture differences in the formulation of the proof theory nor needs to be responsive to the particular details of the proof-theoretic capabilities — its strengths, weaknesses, and presentation — of a given formalism.

In part, this implicit commitment is the means by which Feferman ensures that his characterization of the conditions constitutive of the concept of proof (for formalisms) remains fully general. But, as noted above, there are formalisms that fail to satisfy conditions (i)-(iv). How should we treat such formalisms? Feferman's response might be to argue that since such formalisms fail his conditions, the proof predicates for such formalisms fail to express the logical concept of proof, and hence, fail to reflect or present its own logic. But then Feferman seems to lose traction on the claim that his characterization of the conditions constitutive of the concept of proof hold in general. There are at least two responses to the problem of formalisms that fail to satisfy Feferman's conditions. One might, as Franks suggests, argue on *a priori* grounds that his choice of conditions are not constitutive of the concept of proof. Or one might, as suggested above, argue that the failure of an arithmetization to meet one (or more) of Feferman's conditions yet produce consequences of "extensional interest" constitutes evidence against that condition for the particular formalism. On the first route, one assumes, "monistically," that the conditions a formalism must meet in order to express *the* concept of proof are identical for each formalism irrespective of its proof-theoretic capacities. On the second route, one denies that the monistic approach carries any weight and assumes, "pluralistically," that each formalism implicitly expresses its own concept of proof. If we pursue the first route, then we must explain why our choice of alternative conditions are *a priori* constitutive of the concept of proof but Feferman's are not. But if we pursue the second route, then we must find means to extract conditions for each member of a (possibly) countable set of formalisms and show that such conditions constitute counterexamples to Feferman's conditions.

Above we suggested that, because one must explain why the choice of conditions does not vary

[7] In fact, Feferman (1989) shows that his approach to arithmetization may also be utilized for weaker formal arithmetics such as (PRA) and its conservative extensions. There, the solution to the problem of the inference from (G2) to (1) and (2) is solved through what he calls a "finitary inductively presented logic," where if a formalism satisfies a set of inductive conditions, then it satisfies (1) and (2). However, there are still weaker formalisms for which the problem resurfaces. See the discussion below.

with details of the formalism — that is, why the representation of a formalism's metatheory within the formalism must meet the same conditions irrespective of its specific mathematical and proof-theoretic capacities — pursuing the monistic route is unlikely to yield much fruit. Let's pursue the second, pluralistic, route for a moment. Consider the following. Recall that, above, we mentioned that if T is as strong as (PA), then it proves the *Hauptsatz* and hence proves the equivalence between $Pr_T(y)$ and $Pr^G_T(y)$. However, if T is not strong enough to prove the *Hauptsatz*, then T fails to recognize the equivalence, and hence, from T's standpoint only one of the two provability relations is intensionally correct. Feferman seems to imply that in such cases the "standard" construction is the correct one. But that one arithmetization is correct while another is incorrect is a claim that it is possible to make only from outside of T's standpoint, i.e., from the point of view of a proof-theoretically "richer" theory in which *we* know that the two formula are equivalent but that T fails to prove this. From the point of view of T itself, neither formula has more or less claim to correctness. It might be the case that, from within T, we want a formula that expresses T's provability relative to a proof-theory that is syntactically simpler because it does not contain the cut-rule. In such a case, the "natural" choice is the Genztenian predicate. Or, it might be the case that we want a formula that expresses T's provability relative to a standard classical proof-theory containing the cut-rule. In such a case, the "natural" choice is the standard predicate. Hence, if one is unwilling to conclude that such formalisms fail to express a proof concept, then for formalisms that fail to meet conditions (i)-(iv) and that fail to prove the equivalence between non-standard and standard metamathematical concepts, one *must* find the means to formulate such concepts by varying the concept of proof "from the outside."

What variations are permissible? Franks (2009) proposes an approach to the arithmetization of metamathematical concepts that he contrasts with Feferman's "logical" approach by claiming that it is "fully mathematical." First, some background. An equation of the form:

$$f(\vec{x}) = 0$$

where the unknowns $\vec{x} = (x_1, \ldots, x_n)$ are integers and f is a function of the integers is a *Diophantine equation*, and a Diophantine *problem* asks whether or not a given Diophantine equation has a solution in the integers. In Herbrand (1930) it is shown that metamathematical questions such as "is ϕ a theorem of the formalism S?" are equivalent to particular Diophantine problems.[8] Herbrand's Theorem asserts that a formula (possibly with quantifiers) is provable in the predicate calculus if and only if there exists a tautological disjunction (its Herbrand expansion) in which all variables are replaced by closed numerical terms. Kreisel (1951) and (1952) claimed that Herbrand's theorem provides a means of extracting "constructive content" from metamathematical questions. If we ask

[8] As is well known, Hilbert (1900) asked, in his famous Tenth Problem, whether in general there is an algorithm that tells us whether or not a given Diophantine equation has a solution. Matiyasevich (1970) answered Hilbert's question negatively, so that there is no general method for determining if so or not.

if a Diophantine problem has a solution by effectively substituting closed numerical terms for the unknowns, then the metamathematical question that corresponds to the Diophantine problem can be answered in any formalism that proves those terms to be total functions. Franks claims that this approach provides a better solution to the problem of an intensionally correct consistency predicate: a Diophantine equation produces for every formalism T (the metamathematics) an intensionally correct formulation of the consistency of S (the formalism) just in case the equation has a solution in T's provably total functions. Hence, on his approach, the consistency statement for a formalism varies with the functions the metamathematics proves total, and the result of arithmetizing the question of whether an equation has a solution in T's terms is a formula that corresponds to "the statement of S's consistency 'as', one might say, 'T thinks about the question'" (Franks (2009), 10).

3.3.2 Franks on Consistency

Let's dig in to the details of the view for a moment. Suppose that there exists a solution in T's provably total functions for a Diophantine equation. Then, by Herbrand's Theorem, we may construct a *Herbrand disjunction* that states that one (or more) of T's provably total functions is a solution to that equation such that S proves the disjunction. Since S proves the disjunction, it is possible to construct, in T, a consistency predicate for S from the Herbrand disjunction that is accurate for it from T's point of view. However, since it is possible that T might be proof-theoretically stronger than S, the question arises as to whether it is possible for S to construct its consistency predicate on its own. If so, then since S proves the Herbrand disjunction and the consistency predicate is built out of the Herbrand disjunction, then it ought to follow that S proves its own "consistency," where the arithmetization of consistency does not result in the standard consistency predicate, $Con(S)$, but rather $Con^H(S)$, its Herbrand-consistency. Franks answers this in the affirmative. More precisely, let $Pr^H_T(\ulcorner\phi\urcorner)$ be the Herbrand-provability predicate for T, read as "there is a Herbrand proof of ϕ," and which says that for a finite set T' of the axioms of T there is a Herbrand-disjunction of ϕ proved in T' in which closed numerical terms and functions are substituted for variables such that the resulting quantifier-free disjunction is a propositional tautology. Then, the Herbrand consistency of T, $Con^H(T)$, is the assertion that no finite set T' proves a Herbrand-disjunction with a contradiction (or $\perp$) as the end-formula of the proof. Note that there are two features of this approach that indicate that it is a change "from the outside" in the concept of proof. First, in the arithmetization of provability, and hence in the construction of the consistency statement, one must arithmetize the concept of proof using a predicate that expresses it for propositional logic. Second, in formalisms that do not prove Herbrand's Theorem, $Pr_T(\ulcorner\phi\urcorner)$ and $Pr^H_T(\ulcorner\phi\urcorner)$ are not equivalent, and neither are their consistency statements. Hence, as discussed above, for such formalisms one is faced with a choice: either stick with the standard statement $Con(T)$ or change the concept of proof from the outside.

For concreteness, consider a kind of formalism that we shall call, after Buss (1986), Bounded

Arithmetic (BA).[9] Let Con_β and Con_β^H be two consistency formulae for (BA) such that β is a recursively enumerable arithmetization of the axioms of (BA), where the former is the standard consistency predicate while the latter is its Herbrand consistency. No member of this class proves Herbrand's Theorem. Hence, it follows that:

$$(BA) \nvdash Con_\beta \leftrightarrow Con_\beta^H. \quad (*)$$

On an intensional interpretation, only one of the two formulae can be accurate from (BA)'s point of view. Since Con_β is not provable in (BA), whereas Con_β^H is, Feferman might claim that this speaks against the intensional correctness of Con_β^H. But Franks draws an orthogonal conclusion. He argues that $(*)$ provides an analysis of (G2) for (BA). Though the quantified formula that is the result of arithmetizing the standard consistency statement for (BA) is true in the sense that there exist numerical terms that can be substituted for the variables, such terms do not belong to the class of functions that (BA) proves to be total, and hence, do not belong to its provably total recursive functions. "Thus," he writes, "the unprovability of the standard consistency statement in bounded arithmetic appears merely to be a consequence of the fact that there are function symbols in these theories' languages that they do not prove to be functions" (Franks (2009), 150). For Franks, the unprovability of Con_β in (BA) follows from the fact that (BA) contains expressions for functions that (BA) does not prove to be total. But the fact that a formalism contains redundant expressions seems to be a poor reason to believe that Con_β expresses the consistency of (BA), and Franks takes it to be evidence against the intensional correctness of Con_β. Hence, he concludes, here's a case in which a "non-standard" concept of consistency expressed by (BA)'s Herbrand-consistency appears to be preferable.

Let's reflect on this approach for a moment. While we, like Franks, have argued that the idea that the evidence for or against the choice of an intensional predicate for a formal system ought to arise from the constraints of that system itself, Franks' choice of Herbrand-consistency as an alternative to the standard formulation raises some questions. First, in general the existence of a solution given a Diophantine equation is undecidable, though cases of it are decidable. Jones (1980) shows that, given a Diophantine equation, if there exists a decidable algorithm for its solution, then the degree of the equation must be strictly less than four. Hence, a formal system in which its Herbrand-consistency is provable must have a set of provably total recursive functions whose solutions are in a Diophantine equation of less than four degrees. Hence, and in plain(er) English, formal systems in which Herbrand-consistency is preferable to the standard formulation of consistency and provable where the standard is not have extremely limited mathematical application.[10] Second, Franks might

[9] Buss (1986) investigates the computational properties of a class of formalisms, the "bounded arithmetics," which he denotes as S_k^i, and which include Q, a quantifier-free schema of induction, plus an axiom asserting that exponentiation is a total function.

[10] Indeed the limit, and hence the extent to which a system may prove its own Herbrand-consistency, is fixed to just the so-called set of bounded arithmetics studied in Buss (1986).

counter that it might be possible to prove the Herbrand-consistency of a formal system stronger than the bounded arithmetics, just in case there exists a solution to a Diophantine equation of degree greater than or equal to four in that formal system's provably total recursive functions. Then the existence of a consistency proof for a formalism depends upon open mathematical problems involving the solvability of Diophantine equations, many of which are unknown and whose general problem, as we remarked above, is undecidable. Hence, the question of a system S's consistency depends upon the existence of a solution to a Diophantine equation in S's provably total recursive functions. It seems, then, that we've foisted the problem of a consistency proof for a formal system onto the solvability of open mathematical problems, and that direction of dependence is in conflict with Franks' claim that an intensional proof-theoretic analysis ought to contribute to mathematics.

But the problems in Franks' approach go deeper and appear to affect many approaches to the problem of consistency when it's formulated as a problem of arithmetization. That is, either we tinker with the arithmetization of consistency for particular formalisms in order to skirt (G2) or we preserve and generalize (G2) and stick with its canonical arithmetization.[11] In other words, when our approach begins with the question of which arithmetization is the proper one we are faced with a dilemma. Either an arithmetization of consistency is suited, a la Franks, to the particular formalism for which it's formulated but its applicability is not fully general, or an arithmetization is the canonical one, a la Feferman, and its applicability fully general but unsuited to the particular means with which a given formalism is formulated. On the first approach a formalism can be said to more correctly express its own consistency but then its application is extremely limited, while on the second approach arithmetization for specific formalisms is absorbed into the canonical expression of consistency but its application is fully general. Hence, we seem to have a dilemma. Either the inference from (G2) to claims (1) and (2) fails for extremely weak theories with limited applicability, or the inference holds but we lose the explanation for why it holds. In part, this is because Franks primarily puts pressure on the question of the arithmetization of the consistency predicate, while in fact the issue goes deeper than that. Indeed, the issue here lies with how the provability predicate expresses the concept of proof, and the main question, as we have seen, is when is a change in the concept of proof for a given formal system warranted and when is it not? That is, how should a proof predicate and its associated provability predicate be formalized as to encode facts about, specifically, finitism, if of course what we're after is a finitistic consistency proof.

3.4 Conclusion

In chapter one we argued that proof-theoretic practice explicates hidden higher-order concepts and that such a practice entails what we called, following Sieg, quasi-empiricism about claims that go beyond our claims about finite numbers. We showed this through a close analysis of how to view

[11] Cf. Solomon Feferman (2012).

metamathematical concepts such as completeness and, more importantly here, consistency, in terms of how well a formal system presents the mathematical and metamathematical phenomena it is intended to describe. In chapter two we argued both that the analysis of a non-finitist arithmetic via a finitist arithmetic has epistemic consequences for the former, and that it was possible to bootstrap our way to higher recursive functions as finitistic functions. We found that though Q constitutes its lower bounds, its upper bounds are inexact due to the fact that reasoning via the hidden higher order concepts implicit in a given theory allow us to bootstrap further. But we then estimated its upper bounds as a proper extension of Q that is, importantly, *not* in general closed under provability, since on our analysis the closure of finitistic reasoning under provability implies that it is not possible to reflect on iterations of finite sequences. And, of course, it is just that we have argued that is the heart of finitistic justification. Finally, in this chapter so far we have shown that it is possible to put some pressure on the claim that an arithmetization of the consistency predicate must be the canonical arithmetization. But, we have also seen, through Feferman's examples, and Franks' analysis, that other options appear to be wanting in various ways. In what remains of this chapter, we shall begin to develop an approach to arithmetization and consistency along our own lines, and then in our final remarks, develop some of its epistemic consequences for (HP) as a program in the contemporary scene of the philosophy of mathematics.

Recall that in section (2.1) we looked at Feferman's suggestion that if no consequences of extensional interest result from the use of predicates with intensions that fail to meet one of conditions (i)-(iv), then that should count as evidence against believing that such a predicate expresses the correct concept. Our claim there was that the suggestion opened up a third pressure point in our discussion of arithmetization — namely, the idea that if an arithmetization with a non-standard intension of provability does produce a consequence with extensional interest, then that ought to count as evidence for that non-standard intension. That point, in fact, is one for which we have paved the way in preceding chapters. Recall, for example, that in chapter one one of our main claims was that Hilbert's proof that Desargues' Theorem is independent of Euclidean Geometry generates, as a consequence, an affine "Desarguean" plane and in its reinterpretation by Hilbert, a substructure of the complex field in which commutativity fails to hold. Here is a consequence of extensional interest generated by the proof-theoretic analysis of a set of axioms. Then, in chapter two, we began to extract an informal bootstrapping principle for iterative reflection over finitistic theories specifically by conjecturing that in a finitistic theory the use of a rule expresses the idea that for each k-fold recursive function, and for a formal system in which it is possible to define the kth recursive function, such a rule implicitly contains a means to define the $k+1$st recursive function. If we assume that it is arithmetizable, then it might be possible to produce a result of extensional interest by varying the concept of an F-proof and its associated concept of provability. Of particular interest are cases in which the arithmetization of the consistency statement for the theory is an open question. Hence, it seems, we might have evidence for varying the concept of proof from the finitistic point of view such that it expresses a non-standard intension because such variations *both* faithfully express the concept *and* because it might lead to results of extensional interest.

Our claim has been that a special type of justification is involved in finitistic reflection. When we begin to pick it apart, it appears that one of the primary means of characterizing finitistic justification arises through reflection principles. That, in effect, was the issue with (R_i) in chapter two. Indeed, reflection principles express a form of soundness and as such their use within a base theory intended as a finitistic theory that's applied to a proof-theoretically stronger theory appears to be just what's desired. For a reflection principle, when localized to finitistic reasoning, simply expresses the concept that all theorems derived via a finitistic proof are true, and that, of course, is precisely what we have argued for in chapters one and two. Hence, if it's possible for F to encode a reflection principle for itself, then it might also be possible to construct a proof, given the right kind of arithmetization in addition, that F fails to prove that it is inconsistent, and hence, does not know that it is not. Our conjecture is that if it is possible to isolate the finitistic concept of proof *via* the account given in this dissertation, then we are on the road to showing how for such theories the inferences from (G2) to claims (1) and (2) fail without being subject to the problems with, for example, Franks' account on this score. But immediately at least two natural questions arise. First, are there natural formalisms in which in which Feferman's derivability condition (ii) fails such that one can pin down a similar bootstrapping procedure and in terms of the arithmetization of consistency? Second, and more importantly, when a reflection principle appears on the list of derivability conditions but derivability condition (ii) fails, it follows that the set of axioms for which said principle holds are *not* recursively enumerable. Is this too large a concession for such an apparently small payoff? On the first question our hunch is that there is, and on the second our hunch is that it is not, but we leave both to future research. Our claim here is not that ours constitutes the only approach. Rather, where Feferman's analysis of the concept of proof reduces to what he calls the "'logical" concept, and where Franks believes to have isolated the purely "mathematical" concept of proof, our claim is that we have begun to isolate a specifically *finitistic* concept of proof. Whatever the proposed conditions, they must at least be compatible with the epistemic sense of the formula's intensional correctness and capture how the finitist justifies the consistency formula as a correct expression of finitistic consistency for the particular formalism under consideration. But in the pursuit of consistency through Hilbert's Program using finitistic means, this seems to be as it should. In this dissertation we only hope to have shown that (HP) still has much to say with respect to (BD) in the philosophy of mathematics and that its proof-theoretic resources with respect to Hilbert's finitist and consistency programs are, contrary to what much of the mainstream argues to be the case, still largely unexplored.

Bibliography

— Awodey, Steven; Reck, Erich

(2002) "Completeness and Categoricity, part I: Nineteenth Century Axiomatics to Twentieth Century MetaLogic" *History and Philosophy of Logic*, vol. 23, pp. 1-30.

— Barwise, Jon, ed.

(1982) *Handbook of Mathematical Logic* (Amsterdam: North-Holland).

— Benacerraf, Paul

(1965) "What Numbers Could Not Be," in Benacerraf & Putnam (1983).

(1973) "Mathematical Truth," in Hart (1996).

— Benacerraf, Paul; Putnam, Hilary, eds.

(1983) *The Philosophy of Mathematics* (Cambridge: Cambridge University Press).

— Bernays, Paul

(1922a) "Hilbert's Significance for the Philosophy of Mathematics," in Mancosu (1998).

(1922b) "On Hilbert's Thoughts Concerning the Grounding of Arithmetic," in Mancosu (1998).

(1930) "The Philosophy of Mathematics and Hilbert's Proof Theory," in Mancosu (1998).

(1934) "On Platonism in Mathematics," in Benacerraf and Putnam (1983), pp. 258-71.

— Blanchette, Patricia

(1996) "Frege and Hilbert on Consistency," *Journal of Philosophy*, vol. 93, no. 7: pp. 317-36.

(2007) "Frege on Consistency and Conceptual Analysis," *Philosophia Mathematica*, vol. 15, no. 3: pp. 321-46.

— Boghossian, Paul

(2003) "Blind Reasoning," *Proceedings of the Aristotelian Society*, no. 77, pp. 225-48.

— Boghossian, Paul; Peacocke, Christopher, eds.
(2000) *New Essays on the A Priori* (New York: Oxford University Press).

— BonJour, Laurence
(1995) "Toward a Moderate Rationalism," *Philosophical Topics*, vol. 23, pp. 47-78.
(1998) *In Defense of Pure Reason* (London: Cambridge University Press).

— Brouwer, L.E.J.
(1913) "Intuitionism and Formalism," in Benacerraf and Putnam (1983), pp. 77-96.

— Casullo, Albert
(2003) *A Priori Justification* (New York: Oxford University Press).

— Corry, Leo
(2006) "Axiomatics, Empiricism, and *Anschauung* in Hilbert's Conception of Geometry: Between Arithmetic and General Relativity," in *The Architecture of Modern Mathematics: Essays in History and Philosophy*, eds. Jeremy Gray and José Ferreirós (Oxford: Oxford University Press), pp. 155-76.

— van Dantzig, David
(1956) "Is $10^{10^{10}}$ a Finite Number?," *Dialectica*, vol. 19, pp. 273-7.

— Dedekind, Richard
(1963) *Essays on the Theory of Numbers* (New York: Dover Publications).

— Detlefsen, Michael
(1986) *Hilbert's Program: An Essay on Mathematical Instrumentalism* (Boston: Dordrecht Reidel Publishing).
(1990) "Brouwerian Intuitionism," *Mind*, vol. 99: pp. 501-34.
(1992) *Proof, Logic, and Formalization*, ed., (New York: Routledge).
(1992*) *Proof and Knowledge in Mathematics*, ed., (New York: Routledge).
(2001) "What Does Gödel's Second Theorem Say?," *Philosophia Mathematica*, vol. 9, no. 1, pp. 37-71.

— Dummett, Michael
(1975) "Wang's Paradox," *Synthese*, vol. 30, pp. 301-24.

— Enderton, Herbert B.
(2001) *A Mathematical Introduction to Logic* (Los Angeles: Elsevier Academic Press).

— Euclid
(1956) *The Thirteen Books of the Elements*, vol. I, ed. Sir Thomas Heath (New York: Dover Publications).

— Ewald, William
(1996) *From Kant to Hilbert: vol. II* (New York: Oxford University Press).

— Fantl, Jeremy
(2003) "An Analysis of the *A Priori* and *A Posteriori*," *Acta Analytica*, vol. 18, nos. 30-1, pp. 43-69.

— Feferman, Solomon
(1960) "Arithmetization of Metamathematics in a General Setting," *Fundamenta Mathematicae* vol. 49, pp. 35-92.
(1988) "Hilbert's Program Relativized: Proof-Theoretical and Foundational Reductions," *Journal of Symbolic Logic*, vol. 53, pp. 364-84.
(1991) "Reflecting on Incompleteness," *The Journal of Symbolic Logic*, no. 56, pp. 1-49.
(1996) "Gödel's Program for New Axioms: Why, Where, How and What?," in *Gödel '96*, Lecture Notes in Logic, ed. Petr Hájek (ASL Publications), pp. 3-22.
(1998) *In the Light of Logic* (New York: Oxford University Press).
(2000) "Mathematical Intuition vs. Mathematical Monsters," *Synthese*, vol. 125, pp. 317-32.
(2012) "Review of Curtis Franks' *The Autonomy of Mathematical Knowledge: Hilbert's Program Revisited*," *Philosophia Mathematica*, series III, 20, no.3, pp. 387-400.

— Feferman, Solomon; Strahm, Thomas
(2000) "The Unfolding of Non-Finitist Arithmetic," *Annals of Pure and Applied Logic*, vol. 104, pp. 75-96.
(2010) "The Unfolding of Finitist Arithmetic," *Review of Symbolic Logic* vol. 3, pp. 665-89.

— Ferreirós, José
(1999) *Labyrinth of Thought: A History of Set Theory and its Role in Modern Mathematics* (Berlin: Birkhäuser-Verlag).

— Field, Hartry
(1980) *Science Without Numbers* (New Jersey: Princeton University Press).

— van Fraassen, Bas
(1980) *The Scientific Image* (Oxford: Oxford University Press).

— Franks, Curtis
(2009) "The Gödelian Inferences," *History and Philosophy of Logic*, vol. 30, no. 3, pp. 241-56.
(2009) *The Autonomy of Mathematical Knowledge: Hilbert's Program Revisited* (Cambridge: Cambridge University Press).

— Frege, Gottlob
(1978) *The Foundations of Arithmetic* (Oxford: Basil Blackwell), ed. and trans. J.L. Austin.
(1979) *Posthumous Writings* (Oxford: Basil Blackwell), eds. Hans Hermes et al.
(1980) *Philosophical and Mathematical Correspondence*, ed. Brian McGuiness, trans. Hans Kaal (Chicago: University of Chicago Press).
(1998) *Begriffsschrift* (Hildesheim: Georg Olms Verlag).

— Freudenthal, Hans
(1962) "The Main Trends in the Foundations of Geometry in the 19th Century," in *Logic, Methodology, and Philosophy of Science*, eds. Ernst Nagel, et al. (Stanford: Stanford University Press).

— Gödel, Kurt
(1931) "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, I," *Monatshefte für Mathematik und Physik* no. 38, pp. 173-98.
(1964) "What is Cantor's Continuum Problem?," in Benacerraf/Putnam (1983).

— Goldman, Alvin
(1999) "A Priori Warrant and Naturalistic Epistemology," *Philosophical Perspectives*, vol. 13, pp. 1-28.

— Hallett, Michael
(1990) "Physicalism, Reductionism and Hilbert," in *Physicalism in Mathematics*, ed. Andrew Irvine (Dordrecht: D. Reidel Publishing), pp. 182-256.
(1994) "Hilbert's Axiomatic Method and the Laws of Thought," in *Mathematics and Mind*, ed. A. George (New York: Oxford University Press), pp. 158-200.

(1995) "Hilbert and Logic," in *Qubec Studies in the Philosophy of Science, Part 1: Logic, Mathematics, Physics and the History of Science*, eds. Mathieu Marion; Robert Cohen (Dordrecht: Kluwer Publishing), pp. 135-87.
(2008) "Reflections on The Purity of Method in Hilbert's *Grundlagen der Geometrie*," in Mancosu (2008).

— Hallett, Michael; Majer, Ulrich, eds.
(2004) *David Hilbert's Lectures on the Foundations of Geometry: 1891-1902* (Berlin: Springer Verlag).

— Hart. W.D.
(1996) *The Philosophy of Mathematics* (New York: Oxford University Press).

— van Heijenoort, Jean
(1967) *From Frege to Gödel: A Source Book in Mathematical Logic* (Cambridge: Harvard University Press).

— Hertz, Heinrich
(1894) *Die Prinzipien der Mechanik* (Leipzig: Barth-Verlag).

— Hilbert, David
(1891) "Projektive Geometrie," in Hallett (2004).
(1894) "Die Grundlagen der Geometrie," in Hallett (2004).
(1898b) "Grundlagen der Geometrie," in Hallett (2004).
(1898c) "Elemente der Euklidischen Geometrie," in Hallett (2004).
(1899) *Grundlagen der Geometrie*, in Hallett (2004).
(1900) "Mathematische Probleme," *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen, Math.-Phys. Klasse*, pp. 253-97. In Ewald (1996), pp. 1096-105.
(1900a) "Über den Zahlbegriff," *Jahresbericht der Deutschen Mathematiker-Vereinigung*, vol. 8: pp. 180-84. Also in Ewald (1996), pp. 1089-96.
(1900b) excerpt from French translation of (1899), in Hallett (2004).
(1902a) excerpt from English translation of (1899), in Hallett (2004).
(1902b) "Grundlagen der Geometrie," in Hallett (2004).
(1903) (1899), second German edition.
(1905a) "Über die Grundlagen der Logik und der Arithmetik," in *Verhandlungen des dritten Internationalen Mathematiker-Kongresses in Heidelberg vom 8. bis 13. August 1904*, ed. A. Krazer (Leipzig: Tuebner Verlag), pp. 174-85. In van Heijenoort (1967), pp. 129-38.

(1921) "Neubegründung der Mathematik: Erste Mitteilung," *Abhandlungen aus dem Seminar der Hamburgischen Universität*, vol. 1: pp. 157-77. In Ewald (1996), pp. 1115-34.

(1922) "Neubegründung der Mathematik: Erste Mitteilung," *Abhandlungen aus dem Seminar der Hamburgischen Universität*, vol. 1, pp. 157-77. Translated as "The New Grounding of Mathematics: First Report," in Mancosu (1998), pp. 198-214.

(1923) "Die logischen Grundlagen der Mathematik," *Mathematische Annalen*, vol. 88, pp. 151-65. Trans. as "The Logical Foundations of Mathematics," in Ewald (1996), pp. 1134-48.

(1926) "Über das Unendliche," *Mathematische Annalen*, vol. 95, pp. 161-90. Translated as "On the Infinite," in van Heijenoort (1967), pp. 369-92.

(1927) "The Foundations of Mathematics," in van Heijenoort (1967), pp. 465-79.

(1930) "Logic and the Knowledge of Nature," in Ewald (1996), pp. 1157-65.

(1931) "The Grounding of Elementary Number Theory," in Ewald (1996), pp. 1149-57.

(1931a) "Beweis des *tertium non datur*," Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse, pp. 120-5.

(1935) *Gesammelte Abhandlung* (Berlin: Springer Verlag).

— Hilbert, David; Bernays, Paul

(1934/1968) *Grundlagen der Mathematik: I* (Berlin: Springer-Verlag).

(1939/1970) *Grundlagen der Mathematik: II* (Berlin: Springer-Verlag).

— Katz, Jerrold

(1994) "What Mathematical Knowledge Could Be," *Mind*, vol. 104, pp. 491-522.

— Kennedy, Hubert C.

(1972) "The Origins of Modern Axiomatics: Pasch to Peano," *American Mathematical Monthly*, vol. 79, no. 2: pp. 133-6.

— Kitcher, Philip

(1976) "Hilbert's Epistemology," *Philosophy of Science*, vol. 43, no. 1, pp. 99-115.

(1985) *The Nature of Mathematical Knowledge* (New York: Oxford University Press).

(1997) "A Prioristic Yearnings: A Criticial Study of Michael Friedman's *Kant and the Exact Sciences*," *Erkenntnis*, vol. 44, no. 3, pp. 397-416.

(2000) "A Priori Knowledge Revisited," in Boghossian and Peacocke (2000), pp. 65-91.

— Kleene, Stephen

(1943) "Recursive Predicates and Quantifiers," *Transactions of the American Mathematical Society*, vol. 53, no. 1, pp. 41-73.

(1967) *Introduction to Metamathematics* (New York: Elsevier Science Publishing).

— Kornai, András
(2003) "Explicit Finitism," *International Journal of Theoretical Physics*, vol. 42, no. 2, pp. 301-7.

— Kreisel, Georg
(1960) "Ordinal Logics and the Characterization of Informal Concepts of Proof," in *Proceedings of the International Congress of Mathematicians 1958*, ed. John Todd (Cambridge: Cambridge University Press).
(1983) "Hilbert's Programme," in Benacerraf and Putnam (1983), pp. 207-38.

— Lakatos, Imre
(1976) *Proofs and Refutations* (Cambridge: Cambridge University Press).

— Leitgeb, Hannes
(2009) "On Formal and Informal Provability," in *New Waves in Philosophy of Mathematics*, eds. O. Linnebo and O. Bueno (New York: Palgrave MacMillan Publishing).

— Majer, Ulrich
(1993) "Different Forms of Finitism," in *Proceedings of the 15th International Wittgenstein Symposium*, ed. Johannes Czermak (Vienna: Hölder, Pichler, Tempsky), pp. 185-94.
(1995) "Hilbert's Criticism of Poincaré's Conventionalism," in *Henri Poincaré: Science et Philosophie*, eds. J.L. Greffe et. al (Berlin: Akademie Verlag), pp. 355-64.
(1998) "Knowledge By 'Symbolic Constructions'," in *Philosophy and the Many Faces of Science*, eds. D. Anapolitanos et. al (New York: Rowman and Littlefield), pp. 40-64.
(1998) "Heinrich Hertz's Picture-Conception of Theories: Its Elaboration by Hilbert, Weyl, and Ramsey," in *Heinrich Hertz: Classical Physicist, Modern Philosopher*, eds. Davis Baird et al. (Boston: Kluwer Academic Press), pp. 225-42.

— Mancosu, Paolo
(1998) *From Brouwer to Hilbert. The Debate on the Foundations of Mathematics in the 1920's* (Oxford: Oxford University Press).
(1998*) "Hilbert and Bernays on Metamathematics," in Mancosu (1998), pp. 149-88.
(2008) *The Philosophy of Mathematical Practice* (Oxford: Oxford University Press).

— Niebergall, Karl-Georg; Schirn, Matthias

(1998) "Hilbert's Finitism and the Notion of Infinity," in Schirn (2003), pp. 271-305.

— Parsons, Charles

(1979) "Mathematical Intuition," *Proceedings of the Aristotelian Society*, vol. 80, pp. 145-68.

(1998) "Finitism and Intuitive Knowledge," in Schirn (2003), pp. 249-70.

(2008) *Mathematical Thought and Its Objects* (New York: Cambridge University Press).

— Pasch, Moritz

(1882) *Vorlesung über neue Geometrie* (Leipzig).

— Peano, Guiseppe

(1889) *Arithmetices Principia, Nova Methodo Exposita*, in *Selected Works of Guiseppe Peano*, ed. H.C. Kennedy (London: Allen and Unwin, 1973).

— Pollock, John

(1970) "The Structure of Epistemic Justification," *American Philosophical Quarterly*, vol. 4, pp. 62-78.

(1974) *Knowledge and Justification* (Princeton: Princeton University Press).

(2001) "Defeasible Reasoning with Variable Degrees of Justification," *Artificial Intelligence* vol. 133, pp. 233-82.

— Quine, Willard van Orman

(1953) *From a Logical Point of View* (Cambridge: Harvard University Press), pp. 20-46.

— Ravaglia, Mark

(2003) *Explicating the Finitist Standpoint* (Carnegie Mellon University: Ph.D. Dissertation).

— Reid, Constance

(1996) *Hilbert* (New York: Copernicus Publishing).

— Resnik, Michael

(1974) "On the Philosophical Significance of Consistency Proofs," *Journal of Philosophical Logic*, vol. 3: pp. 133-47.

(1998) "Holistic Mathematics," in *The Philosophy of Mathematics Today*, ed. Matthias Schirn (New York: Oxford University Press), pp. 227-46.

— Rosser, John

(1936) "Extensions of Some Theorems of Gödel and Church," *Journal of Symbolic Logic* vol. 1, pp. 87-91.

(1939) "An Informal Exposition of Proofs of Gödel's Theorem and Church's Theorem," *Journal of Symbolic Logic*, vol. 4, pp. 53-60.

— Saunders, Simon

(1998) "Hertz' *Principles*," in *Heinrich Hertz: Classical Physicist, Modern Philosopher*, eds. Davis Baird et al., pp. 123-54.

— Schirn, Matthias

(2003) *Philosophy of Mathematics Today* (New York: Oxford University Press).

— Shapiro, Stewart

(1983) "Conservativeness & Incompleteness," *Journal of Philosophy*, vol. 80, no. 9, pp. 521-31.

(2000) *Thinking About Mathematics* (New York: Oxford University Press).

— Sieg, Wilfried

(1988) "Hilbert's Program Sixty Years Later," *Journal of Symbolic Logic*, vol. 53, no. 2, pp. 338-48.

(1990) "Reflections on Hilbert's Program," in *Acting and Reflecting*, ed. Wilfried Sieg (Dordrecht: Kluwer Academic Press), pp. 171-82.

(1999) "Hilbert's Programs: 1917-1922," *Bulletin of Symbolic Logic*, vol. 5, no. 1: pp. 1-44.

(2002/9) "Beyond Hilbert's Reach?," in *Logicism, Inuitionism, and Formalism*, eds. Sten Lindström, et al. (Berlin: Springer-Verlag), pp. 449-84.

— Simpson, Stephen G.

(1999) *Subsystems of Second-Order Arithmetic* (Berlin: Springer Verlag).

— Smorynski, C.

(1982) "The Incompleteness Theorems," in Barwise (1982), pp. 821-66.

— Sosa, Ernest

(1991) *Knowledge in Perspective* (Cambridge: Cambridge University Press).

— Tait, William W.

(1961) "Nested Recursion," *Mathematische Annalen*, vol. 143, pp. 236-50.

(1981) "Finitism," *Journal of Philosophy*, vol. 78, pp. 524-46.

(2000) "Remarks on Finitism," in *Reflections on the Foundations of Mathematics: Essays in Honor of Solomon Feferman*, eds. Wilfried Sieg et. al (ASL Publications), pp. 1-11.

(2005) *The Provenance of Pure Reason* (New York: Oxford University Press).

(2010) "Gödel on Intuition and on Hilbert's Finitism," in *Kurt Gödel: Essays for His Centennial* eds. Solomon Feferman et. al. (Cambridge: Association for Symbolic Logic).

— Tarski, Alfred; et. al

(1953) *Undecidable Theories* (Amsterdam: North-Holland Publishing).

— Troelstra, A.S.; van Dalen, Dirk

(1988) *Constructivism in Mathematics*, vol. I (Amsterdam: North-Holland Publishing).

— Weyl, Hermann

(1918) *Das Kontinuum* (Leipzig: Veit and Co.)

(1921) "Über die neue Grunlagenkrise der Mathematik," *Mathematische Zeitschrift*, vol. 10, pp. 37-79. Translation in Mancosu (1998), pp. 86-118.

(1927) "Diskussionsbemerkungen zu dem zweiten Hilbertschen Vortrag über die Grundlagen der Mathematik," *Abhandlungen aus dem Mathematischen Seminar der Universität Hamburg: 6*, pp. 86-8. In van Heijenoort (1967), pp. 480-4.

(1949) *Philosophy of Mathematics and Natural Science* (New Jersey: Princeton University Press).

— Williamson, Timothy

(2000) *Knowledge and Its Limits* (Oxford: Oxford University Press).

— Wright, Crispin

(1982) "Strict Finitism," *Synthese*, vol. 51, pp. 203-82.

(2004) "Intuition, Entitlement, and the Epistemology of Logical Laws," *Dialectica*, vol. 58, no. 1, pp. 155-75.

— Zach, Richard

(2001) *Hilbert's Finitism* (U.C. Berkeley: Ph.D. Dissertation).

(2003) "The Practice of Finitism: Epsilon Calculus and Consistency Proofs in Hilbert's Program," *Synthese*, vol. 137: pp. 211-59.

(2005) "Hilbert's Program Then and Now," in *Philosophy of Logic*, ed. Dale Jacquette (Amsterdam: Elsevier Science Press).