Gödel and the metamathematical tradition*

Jeremy Avigad

July 25, 2007

Abstract

The metamathematical tradition that developed from Hilbert's program is based on syntactic characterizations of mathematics and the use of explicit, finitary methods in the metatheory. Although Gödel's work in logic fits squarely in that tradition, one often finds him curiously at odds with the associated methodological orientation. This essay explores that tension and what lies behind it.

1 Introduction

While I am honored to have been asked to deliver a lecture in honor of the Kurt Gödel centennial, I agreed to do so with some hesitations. For one thing, I am not a historian, so if you are expecting late-breaking revelations from the Gödel *Nachlass* you will be disappointed. A more pressing concern is that I am a poor representative of Gödel's views. As a proof theorist by training and disposition, I take myself to be working in the metamathematical tradition that emerged from Hilbert's program; while I will point out, in this essay, that Gödel's work in logic falls squarely in this tradition, one often senses in Gödel a dissatisfaction with that methodological orientation that makes me uneasy. This is by no means to deny Gödel's significance; von Neumann once characterized him as the most important logician since Aristotle, and I will not dispute that characterization here. But admiration doesn't always translate to a sense of affinity, and I sometimes have a hard time identifying with Gödel's outlook.

*This essay is only slightly modified from the text of a lecture presented at the spring meeting of the Association for Symbolic Logic in Montreal in May, 2006. Citations to Gödel refer to his *Collected Works* [11], where extensive editorial notes and full bibliographic details can be found. I am grateful for comments, corrections, and suggestions from Mark van Atten, Solomon Feferman, Neil Tennant, Richard Zach, and a number of people at the meeting.

I decided to take the invitation to speak about Gödel as an opportunity to work through this ambivalence by reading and thinking about his work. This essay is largely a report on the outcome. In more objective terms, my goal will be to characterize the metamathematical tradition that originated with Hilbert's program, and explore some of the ways that Gödel shaped and reacted to that tradition. But I hope you will forgive me for adopting a personal tone; what I am really doing is discussing aspects of Gödel's work that are of interest to me, as a working logician, in the hope that you will find them interesting too.

Gödel's work can be divided into four categories:[1]

- Early metamathematical work
  - The completeness and compactness theorems for first-order logic (1929)
  - The incompleteness theorems (1931)
  - Decidability and undecidability for restricted fragments of first-order logic (1932, 1933)
  - Properties of intuitionistic logic, and the double-negation translation (1932, 1933)
  - The provability interpretation of intuitionistic logic (1933)
  - The Dialectica interpretation (1941/1958)
- Set theory
  - The relative consistency of the axiom of choice and the continuum hypothesis (1938)
- Foundations and philosophy of physics
  - Rotating models of the field equations (1949)
- Philosophy of mathematics (ongoing)

Akihiro Kanamori and Sy Friedman discuss Gödel's work in set theory in their contributions to this volume, and the contributions by Steve Awodey, John Burgess, and William Tait discuss philosophical aspects of Gödel's

[1] The dates indicated generally correspond to the first relevant publication. For a more detailed overview of Gödel's work, see Feferman's introduction to Gödel's *Collected Works*, [11, volume I].

work. I am in no position to discuss his work on the foundations of physics, and so I will focus here almost exclusively on the first group of results.

This outline of this essay is as follows. In Section 2, I will characterize what I take to be the core methodological components of the metamathematical tradition that stems from Hilbert's program. In Section 3, I will survey Gödel's work in logic from this perspective. In Section 4, I will digress from my narrative to discuss Gödel's proof of the completeness theorem, since it is a lovely proof, and one that is, unfortunately, not well known today. In Section 5, I will consider a number of Gödel's remarks that show him to be curiously at odds with the metamathematical tradition in which he played such a central role. Finally, in Section 6, I will describe the attitude that I take to lie behind these critical remarks, and argue that recognizing this attitude is important to appreciating Gödel's contributions.[2]

2 Hilbert's program and metamathematics

Although Hilbert's program, in its mature formulation, did not appear until 1922, Hilbert's interest in logic and foundational issues began much earlier. His landmark *Grundlagen der Geometrie* of 1899 provided not just an informal axiomatic basis for Euclidean geometry, but also an extensive metamathematical study of interpretations of the axioms. (He used this, for example, to prove their independence.) A year later, he presented his famous list of twenty-three problems to the Second International Congress of Mathematicians. Three of these had a distinctly foundational character, having to do with Cantor's continuum problem, the consistency of arithmetic, and a mathematical treatment of the axioms of physics. In 1904, he presented a partial and flawed attempt to treat the consistency problem in syntactic terms. He did not publicly address foundational issues again until 1922, save for a talk on axiomatic thought in 1917; but lecture notes and other evidence show that he was actively engaged in the issues for much of the intervening period.[3]

By 1922, the *Grundlagenstreit* resulting from Brouwer's intuitionistic challenge was gathering steam. It was then that Hilbert presented his pro-

[2] In his essay, *Kurt Gödel: Conviction and caution* [9], Feferman addresses the closely related issue of the relationship between Gödel's use of formal methods and his objectivist, or platonist, views of mathematics. There, he assesses some of the same data that I consider in Section 5. Although his analysis and conclusions differ from mine, our views are not incompatible, and provide complementary perspectives on Gödel's later remarks.

[3] Bibliographic data on the works mentioned here can be found in any of [8, 11, 12]. For more on Hilbert's program, see [3, 14, 17, 21].

gram to secure the methods of modern mathematics, and hence to "settle the problem of foundations once and for all." The general features of the program are by now well known: one was to characterize the methods of modern, infinitary reasoning using formal axiomatic systems, and then prove those systems consistent, using secure, "finitary" methods. This program is often taken to presuppose a "formalist" position, whereby mathematics is taken to be nothing more than a game of symbols, with no meaning beyond that given by the prescribed rules. One finds such characterizations of formalism, for example, in Brouwer's inaugural address to the University of Amsterdam as early as 1912 [6], and in Ramsey's "Mathematical Logic" of 1926 [15]. When Hilbert is emphasizing the syntactic nature of his program, his language sometimes suggests such a view, but I think it is silly to take this position to characterize his attitudes towards mathematics in general. When one ignores the rhetoric and puts the remarks in the relevant context, one is left with two simple observations: first, with a syntactic characterization of infinitary mathematical reasoning in hand, the question of consistency becomes a purely *mathematical* question; and, second, a consistency proof using a restricted, trusted body of methods would provide solid reassurance to anyone concerned that infinitary methods might be unsound. Thus, I take the core methodological orientation of Hilbert's program to be embodied in the following claims:

- Formal axiomatic systems provide faithful representations of mathematical argumentation.

- With these representations, at least *some* foundational and epistemological questions can be formulated in mathematical terms.

- A finitary, syntactic perspective makes it possible to address such questions without presupposing substantial portions of the body of mathematics under investigation.

In particular, the formal axiomatic method makes it possible to use mathematical methods to address the question as to the consistency of infinitary reasoning, without presupposing the existence of infinitary objects in the analysis.

I, personally, subscribe to these views, and find them eminently reasonable. Taken together, they allow one to use mathematical methods to address epistemological questions, resulting in clear and concrete philosophical gains. Since Hilbert's day, there has been an explosion of interest in computational and symbolic methods in the sciences, while, at the same time, important branches of mathematics have developed methods that are increasingly abstract and removed from computational interpretation. For that reason, I take this broader construal of Hilbert's program to be as important today as it was in Hilbert's time, if not even more so.

3 Gödel and the metamathematical tradition

With this characterization of the metamathematical tradition in hand, let us now turn to Gödel's work in logic. In his 1929 doctoral dissertation, Gödel proved the completeness theorem for first-order logic, clarifying a relationship between semantic and syntactic notions of logical consequence that had bedeviled early logicians [5]. The dissertation makes frequent references to Hilbert and Ackermann's 1928 *Grundzüge der theoretischen Logik*, where the problem of proving completeness for first-order logic was articulated clearly. The compactness theorem, which was ancillary to Gödel's proof, is undeniably model theory's most important tool.

In contrast to the completeness theorem, the incompleteness theorems of 1930 are negative results in Hilbert's metamathematical program. The first incompleteness theorem shows the impossibility of obtaining a complete axiomatization of arithmetic, contrary to what Hilbert had proposed in 1929 [13]. Of course, the second incompleteness theorem, which shows that no reasonable theory of mathematics can prove its own consistency, was a much bigger blow, since it indicates that the central goal of Hilbert's metamathematical program cannot be attained.

Over the next few years, Gödel issued a remarkable stream of striking and seminal results. His double-negation interpretation of classical arithmetic (as well as classical logic) in intuitionistic arithmetic (resp. intuitionistic logic), discovered independently by Gerhard Gentzen, clarified the relationship between the two forms of mathematical reasoning that had been the subject of intense discussion. His interpretation of intuitionistic propositional logic in a modal logic with a "provability" operator helped clarify the relationship between provability and the intuitionistic connectives. His results on the decidability and undecidability of various fragments of first-order logic are fundamental, and are close to being optimal and exhaustive for the first-order setting.

Although Gödel's *Dialectica* interpretation of arithmetic was not published until 1958, he obtained the results much earlier, and lectured on them at Yale in 1941. The interpretation amounts to a translation of intuitionistic arithmetic (and, via the double-negation translation, classical arithmetic) in a quantifier-free theory of primitive recursive functionals of higher-type. This reduces induction for formulas that quantify over the infinite domain of natural numbers to explicit, quantifier-free, induction, modulo a computational scheme of primitive recursion in the higher types.

It is worth mentioning that Gödel's contributions to the study of computability are not only fundamental to computer science today, but firmly in the tradition of the Hilbert school. These include his work on formal notions of computability, which was an important by-product of his work on incompleteness, and the study of primitive recursion in the higher types that accompanies the Dialectica interpretation.

While reading up on Gödel for this essay, I was struck by a remarkable fact: all of Gödel's results in logic, except the completeness theorem, are syntactic in nature.[4] That is to say, every theorem has to do with either provability in a formal system, a translation between formal systems, or the existence of an algorithm for determining whether or not something is provable. Moreover, all the proofs, except for the proof of the completeness theorem, are explicitly finitary, and can be formalized straightforwardly in primitive recursive arithmetic. This is true even of his work in set theory, as he is careful to point out in every statement of the results. For example, in his abstract announcing the relative consistency of the axiom of choice and the continuum hypothesis in 1938, he emphasizes:

> The proof of the above theorems is constructive in the sense that, if a contradiction were obtained in the enlarged system, a contradiction in T could actually be exhibited.[5] (Gödel 1938, II, p. 26)

The exception is the proof of the completeness theorem, which, of course, is nonconstructive. The introduction to his dissertation closes with the following comments:

> In conclusion, let me make a remark about the *means of proof* used in what follows. Concerning them, no restriction whatsoever has been made. In particular, essential use is made of the principle of the excluded middle for infinite collections (the non-denumerable infinite, however, is not used in the main proof).

[4]There is another small exception, namely, a short note on the satisfiability of uncountable sets of sentences in the propositional calculus (Gödel 1932c, I, pp. 238–241).

[5]In other words, we have a finitary proof that if set theory without the additional principles is consistent, then it remains so when the new principles are added as axioms. The double-negation interpretation of classical arithmetic in intuitionistic arithmetic has a similar character; see, for example, the last paragraph of (Gödel 1933e, I, p. 295).

> It might perhaps appear that this would invalidate the entire completeness proof. (Gödel 1929, I, p. 63)

What follows this passage is a discussion of the relevant epistemological issues, and the sense in which the completeness theorem is informative. I don't want to go into the anticipated criticism of the results, or Gödel's response. Rather, I wish to highlight Gödel's remarkable sensitivity to the question as to what *metamathematical* methods are necessary to obtain the requisite results, and the impact these methods have on the epistemological consequences.

Gödel's proof is not often presented these days, which is a shame, because it is interesting and informative. I will therefore break from my narrative, briefly, to share it with you now.

4 Gödel's proof of the completeness theorem

I will take some liberties in describing the proof. For example, I will use contemporary terminology and notation throughout, and rearrange the order in which some of the ideas are presented. A historian will be able to point out all the ways in which my modern gloss ignores interesting and important historical nuances.[6] But one does not have to be a historian to read and enjoy Gödel's original article; what is striking is the extent to which such a

[6]One issue that I have set aside is the influence of Skolem's work on Gödel. In papers published in 1920 and 1923, Skolem gave two clarified and improved proofs of Löwenheim's 1915 theorem, both of which are reprinted in [12]. The normal form used by Gödel below is taken from Skolem's 1920 paper, which is acknowledged in Gödel's 1929 dissertation and in the version of the proof published in 1930. In fact, if one replaces satisfiability by consistency in Skolem's 1923 paper, the result is essentially Gödel's proof. Gödel later acknowledged this fact (this is the context of the quote on the "blindness of logicians" in Section 5, below), but claimed he did not know of Skolem's 1923 proof when he wrote the dissertation. Mark van Atten [1] has combed through the Gödel *Nachlass* and has discovered library slips showing that Gödel requested the volume with Skolem's paper on three separate occasions, but each time the library reported that it was unable to secure the volume.

Another interesting issue has to do with the use of what we now call König's lemma, which was used in papers by König in 1926 and 1927. Gödel seems to be unaware of this, since he does not cite König in either the dissertation or the paper. Gödel gave a quick proof of the lemma in the dissertation, and in the paper said only that the desired truth assignment can be obtained "by familiar arguments." It is worth noting that Skolem also provided a proof of the lemma in his 1923 paper.

These issues are well covered in Dreben and van Heijenoort's introductory notes to (Gödel 1929, 1930, and 1930a) in volume I of the *Collected Works*.

young researcher in a new subject could produce such a clear and mature presentation.

Gödel states the completeness theorem in the following form: "if a first-order sentence φ is not refutable, then it has a model." He also considers the stronger, infinitary version, where φ is replaced by a set of sentences, Γ. He restricts attention to countable first-order languages, so Γ is at worst countably infinite.

Step 1: If a propositional formula φ is not refutable, it has a satisfying truth assignment. This was proved by Post and Bernays, independently, around 1918. One way to prove it is to simply simulate the method of checking each line of the truth table, in the relevant deductive system.

Step 2: If a set Γ of propositional formulas is not refutable, it has a satisfying truth assignment. Write $\Gamma = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}$. Build a finitely branching tree where the nodes at level one are all the truth assignments to variables of φ_0 that make φ_0 true; the nodes at level two are all the truth assignments to variables of $\varphi_0 \wedge \varphi_1$ that make *that* formula true; and so on. (The descendants of a node are all the truth assignments that extend it.) If, at some level k, there is no satisfying assignment to $\varphi_0 \wedge \varphi_1 \wedge \ldots \wedge \varphi_{k-1}$, then, by step 1, Γ is refutable. Otherwise, by König's lemma, there is a path through the tree, which corresponds to a satisfying truth assignment for Γ.

Step 3: Now consider a first-order sentence φ of the form $\forall \bar{x}\ \exists \bar{y}\ \psi(\bar{x}, \bar{y})$, where ψ is quantifier-free in a language with neither function symbols nor the equality symbol. We can prove that φ is either refutable or has a model, as follows. Add countably many fresh constants to the language, let $\vec{c}^{\,i}$ denote an enumeration of all the tuples of constants that can be substituted for $\bar{x}$, and define a sequence of sentences θ_i recursively by taking

$$\theta_i \equiv \psi(\vec{c}^{\,i}, c_k, c_{k+1}, \ldots, c_{k+l})$$

where $c_k, c_{k+1}, \ldots, c_{k+l}$ do not appear in $\theta_0, \ldots, \theta_{i-1}$. The idea is that we are trying to build a model of φ whose universe consists of the constant symbols, so at each stage i we choose new constants to witness the truth of $\exists \bar{y}\ \psi(\vec{c}^{\,i}, \bar{y})$. Now treat the atomic sentences in the language, which are of the form $R(c_{i_0}, \ldots, c_{i_m})$, as propositional variables. By step 2, either some finite subset of $\{\theta_i \mid i \in \mathbb{N}\}$ is propositionally refutable, or there is a satisfying truth assignment. In the second case, we get a model of φ by taking the universe to be the set of constant symbols and using the truth assignment to determine which relations hold of which tuples. In the first case, φ is refutable: from a refutation of

$$\Delta \cup \{\psi(\bar{c}^i, c_k, c_{k+1}, \ldots, c_{k+l})\},$$

where $c_{k+1}, \ldots, c_{k+l}$ do not occur in the finite set Δ, it is easy to obtain a refutation of

$$\Delta \cup \{\forall \bar{x}\ \exists \bar{y}\ \psi(\bar{x}, \bar{y})\},$$

and this move can be iterated until all the formulas in Δ are replaced by $\forall \bar{x}\ \exists \bar{y}\ \psi(\bar{x}, \bar{y})$.

Note that the same method works for infinite sets of $\forall\exists$ sentences, with only slightly more elaborate bookkeeping.

Step 4: Now let φ be an arbitrary first-order sentence in a language without equality or function symbols. The idea is that φ is "equivalent" to a sentence $\exists \bar{R}\ \varphi'$, where φ' is $\forall\exists$. For example, a formula of the form $\forall \bar{x}\ \exists \bar{y}\ \alpha(\bar{x}, \bar{y})$ is "equivalent" to

$$\exists R\ (\forall \bar{x}\ \exists \bar{y}\ R(\bar{x}, \bar{y}) \wedge \forall \bar{x}\ \forall \bar{y}\ (R(\bar{x}, \bar{y}) \rightarrow \alpha(\bar{x}, \bar{y}))). \tag{1}$$

To see this, note that formula (1) clearly implies $\forall \bar{x}\ \exists \bar{y}\ \alpha(\bar{x}, \bar{y})$, and the converse is obtained by taking $R(\bar{x}, \bar{y})$ to be $\alpha(\bar{x}, \bar{y})$. But now note that if α is a prenex formula with $k+1$ $\forall\exists$-blocks of quantifiers, the formula after $\exists R$ in (1) can be put in prenex form, with only k such blocks. Iterating this, we end up with a formula in the desired normal form.

More precisely, the sense in which φ is equivalent to $\exists \bar{R}\ \varphi'$ is this:

- $\varphi' \rightarrow \varphi$ is provable in first-order logic, so that if φ' has a model, so does φ; and
- if φ' is refutable, then so is φ.

So the statement "if φ is not refutable, then it has a model" is reduced to the corresponding statement for the $\forall\exists$ formula φ', and we have already handled that case in step 3. Once again, the proof extends straightforwardly to infinite sets of sentences.

Step 5: The result can be extended to languages with function symbols, by interpreting functions in terms of relations; and to languages with equality, in the usual way, by adding equality axioms and then taking a quotient structure.

What do I like about the proof? First, it is extremely modular. Each step turns on one key idea, and the proof clearly dictates the requisite properties of the deductive system: the role of the propositional axioms and rules is clear in step 1; the quantifier rules and axioms are used in steps 3 and 4; the equality axioms only come in at step 5.

Second, the constructive content is clear; the only nonconstructive element is the use of König's lemma in step 2. This observation lies at the heart of recursion-theoretic and proof-theoretic analyses of the completeness theorem.

Third, the proof shows something stronger: if a formula φ doesn't have a model, there is a refutation of φ that involves only propositional combinations of *subformulas* of φ. In particular, if a formula θ is provable in first-order logic (and so $\neg\theta$ does not have a model), then there is a proof of θ involving only formulas with a quantifier complexity that is roughly the same as that of θ, or lower. This is an important proof-theoretic fact that is usually obtained as a consequence of the cut-elimination theorem, and so I was surprised to find it implicit in Gödel's original proof.

Finally, there is the choice of normal form. Skolem functions can be used to reduce the satisfiability of a sentence to the satisfiability of a universal sentence, its *Skolem normal form.* This fact is used often in proof theory and automated reasoning today. But there are messy technical difficulties involved with eliminating Skolem axioms; since Skolem functions are really choice functions, this is closely related to mathematicians' dislike of noncanonical choices in an ordinary mathematical proof. Gödel uses the fact that the satisfiability of any first-order formula can be reduced to the satisfiability of an $\exists\forall$ sentence in a purely relational language. The quantified relations are really choice-free "Skolem multifunctions," making the reduction technically smoother and much more satisfying.

5 Gödel's remarks on finitism

I would like to return to the relationship between Gödel and Hilbert, and, by way of contrast, briefly consider the relationship between Hilbert and *his* predecessor, Kronecker. Hilbert's early work in algebraic geometry and algebraic number theory was strongly influenced by that of Kronecker, though, as is well-known, Hilbert was critical of Kronecker's methodological proscriptions for mathematics. Indeed, Hilbert's program can be seen as an attempt to do battle with Kronecker, on Kronecker's own terms. In his obituary for Hilbert, Weyl colorfully described the situation as follows:

> When one inquires into the dominant influences acting upon Hilbert in his formative years one is puzzled by the peculiarly ambivalent character of his relationship to Kronecker: dependent on him, he rebels against him. Kronecker's work is undoubtedly of paramount importance for Hilbert in his algebraic period. But the old gentleman in Berlin, so it seemed to Hilbert, used his power and authority to stretch mathematics upon the Procrustean bed of arbitrary philosophical principles and to suppress such developments as did not conform: Kronecker insisted that existence theorems should be proved by explicit construction, in terms of integers, while Hilbert was an early champion of Georg Cantor's general set-theoretic ideas.... A late echo of this old feud is the polemic against Brouwer's intuitionism with which the sexagenarian Hilbert opens his first article on "Neubegründung der Mathematik" (1922): Hilbert's slashing blows are aimed at Kronecker's ghost whom he sees rising from the grave. But inescapable ambivalence even here — while he fights him, he follows him: reasoning along strictly intuitionistic lines is found necessary by him to safeguard non-intuitionistic mathematics. (Weyl [20], p. 613)

The relationship between Gödel and Hilbert is not nearly as dramatic. I have characterized Gödel's work as being firmly in the tradition that Hilbert established, much of it devoted to answering questions that Hilbert himself posed. In that regard, Gödel gives credit where it is due, and does not in any way deny Hilbert's influence or play down the importance of his contributions. In fact, his 1931 paper on the incompleteness theorems ends with a remarkably charitable and optimistic assessment of Hilbert's program:

> I wish to note expressly that [the statements of the second incompleteness theorem for the formal systems under consideration] do not contradict Hilbert's formalistic viewpoint. For this viewpoint presupposes only the existence of a consistency proof in which nothing but finitary means of proof is used, and it is conceivable that there exist finitary proofs that *cannot* be expressed [in the relevant formalisms]. (Gödel 1931, I, p. 195)

Within a few years, however, he had abandoned this view.[7] In his lecture at Zilsel's seminar in 1938, he was much more critical of attempts to salvage

[7] See, for example, Gödel *1933o in [11, volume II].

Hilbert's original plan to establish the consistency of mathematics. Commenting on Gentzen's proof of the consistency of arithmetic using transfinite induction up to ε_0, he says:

> I would like to remark by the way that Gentzen sought to give *a "proof"* of this rule of inference and even said that this was the essential part of his consistency proof. In reality, it's not a matter of proof at all, but of an appeal to evidence... I think it makes more sense to formulate an axiom precisely and to say that it is just not further reducible. But here again the drive of Hilbert's pupils to derive something from nothing stands out. (Gödel 1938a, III, pp. 107–109)

In later years, one finds him generally critical of a finitist methodology. For example, in a letter he wrote to Hao Wang in 1967, he blamed the failure of Skolem to extract the completeness theorem from his results of 1923 on the intellectual climate of that time, and to a misplaced commitment to a finitist metatheory:

> This blindness (or prejudice, or whatever you may call it) of logicians is indeed surprising. But I think the explanation is not hard to find. It lies in a widespread lack, at that time, of the required epistemological attitude towards metamathematics and toward non-finitary reasoning.
>
> Non-finitary reasoning in mathematics was widely considered to be meaningful only to the extent to which it can be "interpreted" or "justified" in terms of a finitary metamathematics. (Note that this, for the most part, has turned out to be impossible in consequence of my results and subsequent work.) (Quoted in [18, p. 8], and [19, pp. 240–241])

Despite the fact that almost all of his proofs were explicitly finitary, Gödel went out of his way to emphasize that the "objectivistic conception of mathematics and metamathematics in general, and of transfinite reasoning in particular, was fundamental to my other work in logic." Of course, by representing transfinite methods within explicit formal systems, Gödel can make use of such reasoning while maintaining finitary significance. But it is interesting that here Gödel plays up the importance of the transfinite methods, while downplaying the importance of the finitary metamathematical stance.

A few months later, in a follow-up to that letter, he repeated the claim that it would have been practically impossible to discover his results without an objectivist conception. He then continued:

> I would like to add that there was another reason which hampered logicians in the application to metamathematics, not only of transfinite reasoning, but of mathematical reasoning in general. It consists in the fact that, largely, metamathematics was not considered as a science describing objective mathematical states of affairs, but rather as a theory of the human activity of handling symbols. [18, pp. 9–10]

This last passage indicates a critical attitude towards syntactic characterizations of mathematics that one also finds in an essay that Gödel began preparing in 1953 for the Schilpp volume on Carnap. The essay was titled "Is mathematics the syntax of language?" and was designed to refute this core tenet of logical positivism. Although he never completed a version that he found satisfactory, he did feel that his refutation of Carnap's position was decisive.[8] In 1972, he said to Wang:

> Wittgenstein's negative attitude towards symbolic language is a step backward. Those who, like Carnap, misuse symbolic language want to discredit mathematical logic; they want to prevent the appearance of philosophy. The whole movement of the positivists want to destroy philosophy; for this purpose, they need to destroy mathematical logic as a tool. [19, p. 174]

Although these comments are not directed at Hilbert *per se*, they can be viewed as a criticism of the types of formalism that are often ascribed to Hilbert. Gödel did provide a direct assessment of Hilbert's program in a lecture that he prepared for the American Philosophical Society around 1961 but never delivered. In that lecture, Gödel characterized general philosophical world-views along a spectrum, in which "skepticism, materialism, and positivism stand on one side, spiritualism, idealism, and theology on the other." The tenor of the times, according to Gödel, had led towards a general shift to the skeptical values that he had located on the left side of the spectrum. Mathematics had traditionally been a stronghold for those

[8] Awodey and Carus have shown, however, that the argument is flawed; see [4]. See also Warren Goldfarb's introductory notes to (Gödel *1953/9, III, 324–363) for a detailed discussion of Gödel's essay.

idealistic values on the right. But, according to Gödel, the skeptical attitudes eventually reached the point where they began to influence foundational thinking in mathematics, resulting in concerns about the consistency of mathematical reasoning.

> Although the nihilistic consequences are very well in accord with the spirit of the time, here a reaction set in—obviously not on the part of philosophy, but rather on that of mathematics, which, by its nature, as I have already said, is very recalcitrant in the face of the *Zeitgeist*. And thus came into being that curious hermaphroditic thing that Hilbert's formalism represents, which sought to do justice both to the spirit of the time and to the nature of mathematics. It consists in the following: on the one hand, in conformity with the ideas prevailing in today's philosophy, it is acknowledged that the truth of the axioms from which mathematics starts out cannot be justified or recognized in any way, and therefore the drawing of consequences from them has meaning only in a hypothetical sense, whereby this drawing of consequences itself (in order to satisfy even further the spirit of the time) is construed as a mere game with symbols according to certain rules, likewise not [supported by] insight. (Gödel 1961/?, III, p. 379)

On the other hand, Gödel went on to explain, Hilbert's program was designed to justify the desired "rightward" view of mathematics, via finitary consistency proofs. The incompleteness theorems, however, show that "it is impossible to rescue the old rightward aspects of mathematics in such a manner" (*ibid.*, p. 381). Thus a more subtle reconciliation of the leftward and rightward views is required:

> The correct attitude appears to me to be that the truth lies in the middle, or consists of a combination of the two conceptions.
>
> Now, in the case of mathematics, Hilbert had of course attempted just such a combination, but one obviously too primitive and tending too strongly in one direction. (*ibid.*)

Gödel is not excessively critical of Hilbert in the lecture. But while he is respectful of Hilbert's attempt to rescue mathematics from the destructive tendencies of materialism and skepticism, he clearly feels that Hilbert's viewpoint is inadequate to the task at hand. (The lecture goes on to suggest that the methods of Husserl's phenomenology provide a more promising approach, but this is not something I can go into now.[9])

6 Conclusions

In the end, what are we to make of Gödel's critical remarks? Most of the comments we have just considered were made toward the end of Gödel's life, and I think it would be a mistake to assume that such views influenced his earlier work. But the remarks do indicate a fundamental aspect of Gödel's outlook that puts it in stark opposition to Hilbert's and which, I believe, was constant throughout Gödel's career.

The fundamental difference between Gödel and Hilbert, as I see it, lies in their views on the relationship between mathematics and philosophy. Hilbert was a consummate mathematician, with an unbounded optimism and faith in the ability of mathematics to solve all problems; at the same time, he was openly disparaging of the contemporary philosophical climate, and skeptical of philosophy's ability to settle epistemological issues on its own terms.[10] Thus, from Hilbert's perspective, progress is only possible insofar as philosophy can be absorbed into mathematics, that is, insofar as one can replace philosophical questions with properly mathematical ones.

What Gödel and Hilbert had in common was an unshakeable faith in rational inquiry. But, in contrast to Hilbert, Gödel was intensely sensitive to the limitations of formal methods, and deemed them insufficient, on their own, to secure our knowledge of transcendent mathematical reality. Thus, for Gödel, important epistemological questions require philosophical methods that go beyond the formal mathematical ones, picking up the slack where mathematical methods fall short:

> The analysis of concepts is central to philosophy. Science only combines concepts and does not analyze concepts. It contributes to the analysis of concepts by being stimulating for real analysis.... Analysis is to arrive at what thinking is based on: the inborn intuitions. [19, p. 273]

This, I take it, explains his disdain for mathematicians and philosophers who expect too much from syntactic methods. They are the ones who ex-

[9] See van Atten and Kennedy [2] for a detailed analysis of Gödel's interest in phenomenology, and further references.

[10] Carnap was even more critical of the metaphysical turn in philosophy, as it evolved from Husserl to Heidegger; see, for example, the discussion in Friedman [10].

pect "to derive something from nothing" while avoiding "the appearance of philosophy."

There is a touch of drama here. Gödel had inherited a powerful metamathematical tradition from Hilbert, and he shared Hilbert's strong desire to save mathematics from destructive skeptical attitudes. But, in the end, he concluded that an overly narrow reading of the metamathematical tradition leaves skepticism with the upper hand. Remember that Hilbert ended his Königsberg lecture, "Naturerkennen und Logik," with the words "wir müssen wissen, wir werden wissen."[11] At the same time, unbeknownst to Hilbert, Gödel was at a conference on epistemology and the exact sciences in that very same city. It is one of the great ironies in the history of ideas that this was the conference where Gödel announced the first incompleteness theorem, just a day before Hilbert gave that famous speech (see [7, 68–71]).

Before I began preparing to write this essay, I would have sided with Hilbert. I take Gödel's most important and enduring contributions to lie in his mathematical work; one cannot deny that the stunning corpus of theorems that he produced extend our knowledge in profound and important ways. His philosophical views on mathematical realism and the nature of our faculties of intuition seem to me to be comparatively thin. To be sure, one can imagine that had his health been better and his life been longer, he might have produced more striking and compelling theorems to fill out the informal views. But this is exactly my point: absent the mathematical analysis, it is hard to say what these views amount to.

But I have come to realize that this way of separating Gödel's mathematical work from his philosophical views is misleading. For, what is most striking about Gödel's mathematical work is the extent to which it is firmly rooted in philosophical inquiry. We never find Gödel making up mathematical puzzles just for the sake of solving them, or developing a body of mathematical techniques just for the sake of doing so. Rather, he viewed mathematical logic as a sustained reflection on the nature of mathematical knowledge, providing a powerful means of addressing core epistemological issues. Gödel kept his focus on fundamental questions, and had the remarkable ability to to advance our philosophical understanding with concrete and deeply satisfying answers.

[11] "We must know, we will know." A four-minute excerpt from the speech was later broadcast by radio. The text of the excerpt and a translation by James T. Smith can be found online, together with a link to an audio recording of the broadcast:

http://math.sfsu.edu/smith/Documents/HilbertRadio/HilbertRadio.pdf

The final pronouncement is also Hilbert's epitaph; see [16].

When one considers the history of science and philosophy in broad terms, it becomes clear that the sharp separation between the two disciplines that we see today is a recent development, and an unfortunate one. In contrast, Gödel saw mathematics and philosophy as partners, rather than opponents, working together in the pursuit of knowledge. This conception of logic, I believe, is Gödel's most important legacy to the metamathematical tradition, and one we should be thankful for.

References

[1] Mark van Atten. On Gödel's awareness of Skolem's Helsinki lecture. *History and Philosophy of Logic*, 26:321–326, 2005.

[2] Mark van Atten and Juliette Kennedy. On the Philosophical Development of Kurt Gödel. *Bulletin of Symbolic Logic*, 9:425–476, 2003.

[3] Jeremy Avigad and Erich H. Reck. "Clarifying the nature of the infinite": the development of metamathematics and proof theory. Technical Report CMU-PHIL-120, Carnegie Mellon University, 2001.

[4] Steve Awodey and André W. Carus. How Carnap could have replied to Gödel. In Steve Awodey and Carsten Klein, editors, *Carnap brought home: The view from Jena*, Open Court, Chicago, 2004, pages 203–223.

[5] Steve Awodey and Erich H. Reck. Completeness and categoricity, part I: 19th century axiomatics to 20th century metalogic. *History and Philosophy of Logic* 23:1–30, 2002.

[6] L. E. J. Brouwer. Intuitionism and formalism, 1912. An English translation by Arnold Dresden appears in the *Bulletin of the American Mathematical Society*, 20:81–96, 1913–1914, and is reprinted in Ronald Callinger, ed., *Classics of Mathematics*, Prentice Hall, Englewood, 1982, pages 734–740.

[7] John W. Dawson, Jr. *Logical dilemmas: The life and work of Kurt Gödel*. A K Peters Ltd., Wellesley, MA, 2005.

[8] William Ewald, editor. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*. Clarendon Press, Oxford, 1996. Volumes 1 and 2.

[9] Solomon Feferman. Kurt Gödel: Conviction and caution. *Philosophia Naturalis*, 21: 546–562, 1984. Reprinted in Solomon Feferman, *In the Light of Logic*, Oxford University Press, New York, 1998, pages 150–164.

[10] Michael Friedman. *A parting of the ways: Carnap, Cassirer, and Heidegger*. Open Court, Chicago, 2000.

[11] Kurt Gödel. *Collected works*. Oxford University Press, New York, 1986–2003. Edited by Solomon Feferman et al. Volumes I–V.

[12] Jean van Heijenoort. *From Frege to Gödel: A sourcebook in mathematical logic, 1879-1931*. Harvard University Press, 1967.

[13] David Hilbert. Probleme der Grundlegung der Mathematik. *Mathematische Annalen*, 102:1–9, 1929.

[14] Paolo Mancosu, editor. *From Brouwer to Hilbert: The debate on the foundations of mathematics in the 1920's*. Oxford University Press, Oxford, 1998.

[15] Frank Plumpton Ramsey. Mathematical logic. *Mathematical Gazette*, 13:185–194, 1926. Reprinted in F. P. Ramsey, *Philosophical papers*, Cambridge University Press, Cambridge, 1990, pages 225-244.

[16] Constance Reid. *Hilbert*. Springer, Berlin, 1970.

[17] Wilfried Sieg. Hilbert's programs: 1917–1922. *Bulletin of Symbolic Logic*, 5:1–44, 1999.

[18] Hao Wang. *From mathematics to philosophy*. Routledge & Kegan Paul, London, 1974.

[19] Hao Wang. *A logical journey: From Gödel to philosophy*. MIT Press, Cambridge, MA, 1996.

[20] Hermann Weyl. David Hilbert and his mathematical work. *Bulletin of the American Mathematical Society*, pages 612–654, 1944.

[21] Richard Zach. Hilbert's program. In the *Stanford encyclopedia of philosophy*, http://plato.stanford.edu/entries/hilbert-program/.