David Hilbert's contributions to logical theory

CURTIS FRANKS

1. A mathematician's cast of mind

Charles Sanders Peirce famously declared that "no two things could be more directly opposite than the cast of mind of the logician and that of the mathematician" (*Peirce 1976*, p. 595), and one who would take his word for it could only ascribe to David Hilbert that mindset opposed to the thought of his contemporaries, Frege, Gentzen, Gödel, Heyting, Łukasiewicz, and Skolem. *They* were the logicians *par excellence* of a generation that saw Hilbert seated at the helm of German mathematical research. Of Hilbert's numerous scientific achievements, not one properly belongs to the domain of logic. In fact several of the great logical discoveries of the 20th century revealed deep errors in Hilbert's intuitions—exemplifying, one might say, Peirce's bald generalization.

Yet to Peirce's addendum that "[i]t is almost inconceivable that a man should be great in both ways" (*Ibid.*), Hilbert stands as perhaps history's principle counter-example. It is to Hilbert that we owe the fundamental ideas and goals (indeed, even the name) of proof theory, the first systematic development and application of the methods (even if the field would be named only half a century later) of model theory, and the statement of the first definitive problem in recursion theory. And he did more. Beyond giving shape to the various sub-disciplines of modern logic, Hilbert brought them each under the umbrella of mainstream mathematical activity, so that for the first time in history teams of researchers shared a common sense of logic's open problems, key concepts, and central techniques. It is not possible to reduce Hilbert's contributions to logical theory to questions of authorship and discovery, for the work of numerous colleagues was made possible precisely by Hilbert's influence as Europe's preeminent mathematician together with his insistence that various logical conundra easily relegated to the margins of scientific activity belonged at the center of the attention of the mathematical community.

In the following examination of how model theory, proof theory, and the modern concept of logical completeness each emerged from Hilbert's thought, one theme recurs as a unifying motif: Hilbert everywhere wished to supplant philosophical musings with definite mathematical problems and in doing so made choices, not evidently necessitated by the questions themselves, about how to frame investigations "so that," as he emphasized in *1922*, "an unambiguous answer must result." This motif and the wild success it enjoyed are what make Hilbert the chief architect of *mathematical logic* as well as what continue to inspire misgivings from several philosophical camps about logic's modern guise.

2. Model theory

Hilbert's early mathematical work stands out in its push for ever more general solutions and its articulation of new, increasingly abstract frameworks within which problems could be posed and solved in the sought generality. Because of the deviation from traditional methods brought about by these ambitions, Hilbert's work engendered praise for its scope and originality as well as criticism for its apparent lack of concreteness and failure to exhibit the signs of what was customarily thought of as proper mathematics. His fundamental solution of Gordan's problem about the bases of binary forms exemplifies the trend: Hilbert extended the problem to include a much wider range of algebraic structures, and his proof of the existence of finite bases for each such structure was "non-constructive" in that it did not provide a recipe for specifying the generating elements in any particular case. When Hilbert submitted the paper for publication in *Mathematische Annalen* Gordan himself rejected it, adding the remark "this is not mathematics; this is theology" (Reid 1996, p. 34) The work was recognized as significant by other eminent mathematicians, however, and it was eventually published in an expanded form. Felix Klein wrote to Hilbert about the paper, "Without doubt this is the most important work on general algebra that the *Annalen* has ever published" (Rowe 1989, p. 195).

The split reception of Hilbert's early work foreshadows an infamous ideological debate that ensued with constructively minded mathematicians L. E. J. Brouwer and Herman Weyl in the 1920s—a debate which resurfaces in section 3 of this chapter. For now it will do to focus on the pattern of abstraction and generality itself, as it arose in Hilbert's logical study of the foundations of geometry.

The question of the relationship of Euclid's "parallel postulate" (PP) to the other principles in his *Elements* had mobilized scholars since antiquity until in the 19th century Gauss, Lobachevski, Riemann, Beltrami, and others offered examples of well-defined mathematical spaces in which Euclid's principles each are true except for PP. Whereas the classical ambition had been to derive PP from the other principles, showing it to be redundant, such spaces made evident the *independence* of PP from the others: Because one of these spaces can be read as an interpretation of the principles in the *Elements*, the truth of the other postulates does not guarantee the truth of PP; neither does the truth of the other postulates guarantee that PP is false, for the familiar Euclidean plane can be viewed as an interpretation in which PP, alongside the other principles, is true.

One burden shouldered by advocates of independence proofs of this sort was to demonstrate that the several interpretations described are in fact coherent mathematical structures—indeed, a favorite tactic among medieval and Renaissance thinkers was to argue indirectly for the derivability of PP by showing that any such interpretation would be self-contradictory. This burden proved to be a heavy one because of a prevailing view that a "space" should be responsible to human intuition or visual experience—quite often writers would cite the content of their spatial perception to undermine "deviant" interpretations of the geometric postulates. 19th century advances in algebraic geometry helped sheild independence proofs from this sort of objection, but some confusion persisted about whether the results were refinements of the concept of space or purely logical observations about the relationships among mathematical principles. Thus von Hemholtz said both that Riemann's work "has the peculiar advantage that all its operations consist in pure calculation of quantities, which quite obviates the danger of habitual perceptions being taken for necessities of thought," but also that Riemann's geometries should be considered as "forms of intuition transcendentally given ... into which any empirical content whatever will fit".

In his *Foundations of Geometry* of 1899, Hilbert provided a general setting for proving the independence of geometric principles and greatly sharpened the discussion of mathematical axioms. Along the way he offered new independence proofs of PP and of the Archimedean principle and constructed interpretations of assemblages of axioms never before thought of together. But the impact of the work lay not in any one of the independence proofs it contained nor in the sheer number of them all, so much as in the articulation of an abstract setting in which the compatibility and dependence of geometric principles could be investigated methodically and in full generality. "It immediately became apparent," explained his close collaborator Paul Bernays, "that this mode of consideration had nothing to do with the question of the epistemic character of the axioms, which had, after all, formerly been considered as the only significant feature of the axiomatic method. Accordingly, the necessity of a clear separation between the mathematical and the epistemological problems of axiomatics ensued" (Bernays p. 191).

Hilbert managed this separation by dispensing with the stricture that the models used to interpret sets of geometric principles be viewed as spaces in any traditional way. In the abstract setting of *Foundations of Geometry*, Euclid's principles were recast as collections of formal axioms (as theories, $\mathbf{T}_1$, $\mathbf{T}_2, \ldots$). Though they contained words like "point" and "line," these axioms no longer had any meaning associated with them. Further, Hilbert exploded the distinction between the mathematical principles being studied and the structures used to interpret collections of them. In *Foundations of Geometry*, there are only theories, and an interpretation of a collection of geometric axioms $\mathbf{T}$ is carried out in an algebraic theory (typically a field over the integers or complex numbers) $\mathbf{S}$, so that each axiom in $\mathbf{T}$ can be translated back into a theorem of $\mathbf{S}$.

From this point of view, Hilbert was able to articulate precisely the sense in which his demonstrations established the consistency of collections of axioms or the independence of one axiom from others. In each case the question of the consistency of $\mathbf{T}$ is reduced to that of the simpler or more perspicuous theory $\mathbf{S}$ used in the interpretation, demonstrating the "relative consistency" of $\mathbf{T}$ with respect to $\mathbf{S}$. For if $\mathbf{T}$ were inconsistent, in the sense that its axioms logically imply a contradiction, then because logic is blind to what meaning we ascribe to words like "point" and "line" this same implication holds also under the reinterpretation, so that a collection of theorems of $\mathbf{S}$ itself implies a contradiction. Because theorems of $\mathbf{S}$ are implied by the axioms of $\mathbf{S}$, in this case $\mathbf{S}$ is inconsistent. So the interpretation shows that $\mathbf{T}$ is inconsistent only if $\mathbf{S}$ is. Similarly, the independence of an axiom like PP from a collection of axioms $\mathbf{C}$ can be demonstrated relative to the consistency of (typically two different) theories $\mathbf{S}_1$ and $\mathbf{S}_2$: One constructs theories $\mathbf{T}_1 = \mathbf{C} \cup \text{PP}$, $\mathbf{T}_2 = \mathbf{C} \cup \neg\text{PP}$ and demonstrates the consistency of $\mathbf{T}_1$ relative to that of $\mathbf{S}_1$ and of $\mathbf{T}_2$ relative to that of $\mathbf{S}_2$.

Of course the generality of Hilbert's methods is suited to investigations unrelated to geometry, to *metatheoretical* questions about axiom systems in general. And while some contemporaries, notably Gottlob Frege, demurred from the purely logical conception of consistency on offer, Hilbert's techniques won the day. "The important thing," Bernays remarked, "about Hilbert's *Foundations of Geometry* was that here, from the beginning and for the first time, in the laying down of the axiom system, the separation of the mathematical and logical from the spatial-intuitive, and with it from the epistemological foundation of geometry, was completely carried out and expressed with complete rigor" (Bernays 192). Indeed the abstract point of view Hilbert introduced, together with the idea of using interpretations of this sort to study the logical relationships among the axioms of familiar mathematical theories are the basic setting and tool of contemporary model theory.

Amid the revolutionary turn of thought displayed in Hilbert's work on geometry are two notorious doctrines worth attention because of their influence in logic. The first is Hilbert's insistence that mathematical existence amounts to nothing more than the consistency of a system of axioms. "If the arbitrarily given axioms do not contradict one another," he wrote to Frege, "then they are true, and the things defined by the axioms exist" (Frege 1980, p. 40). As a doctrine of mathematical existence, this idea is doubly dubious: It would later be clear from discoveries of Skolem and Gödel that a consistent theory is typically not "categorical"—its several models are not isomorphic—so the sense in which it is supposed to implicitly (partially?) define the terms that appear in its axioms is not clear. Further, the inference from the compatibility of a collection of axioms to the existence of a structure that models them is *an inference*. As Gödel would emphasize, it is careless *to define* existence in this way, because the validity of that inference depends on the completeness of the underlying logic. Among the reasons that a contradiction might be underivable from a set of axioms is the possibility that the logic used is too meager to fully capture the semantic entailment relation. In the case of first-order theories, consistency does indeed imply the existence of a model, but the incompleteness of higher-order logic with respect to its standard semantics leaves open the possibility of consistent theories that are not satisfied by any structure at all.

The second doctrine is Hilbert's idea that his "axiomatic method" would do more than allow a general setting for consistency and independence results but in fact provide significant advances in the ordinary mathematical theories that were subjected to axiomatization. At times, Hilbert even expressed an ambition that axiomatics would open the door to the solution of all mathematical problems, perhaps even by rendering unsolved mathematical conjectures to systematic resolution in the abstract setting of formal, uninterpreted sentences subject to combinatorial tests of derivability. Just articulating this idea generated significant interest in the *decision problem*: The question whether the truth or falsity of any given sentence in the language of a formal theory can be effectively determined. By the work of Church and Turing it became known that even first-order quantification theory is undecidable, for although there is an algorithm for discovering of any valid formula of quantification theory that it is valid, there is no corresponding procedure for discovering of a formula that it is not valid (or, equivalently, that its negation has a model) in the event that it isn't. Still worse, according to Gödel's first incompleteness theorem, no axiomatic theory $\mathbf{T}$ whatsoever in the signature of even basic arithmetic is "syntactically complete" in the sense that for any sentence ϕ in that signature, either $\mathbf{T} \vdash \phi$ or $\mathbf{T} \vdash \neg\phi$. One cannot, as Hilbert had hoped to do, provide a full axiomatization of number theory.

However overreaching Hilbert's ambitions may have been, his more modest prediction that, through axiomatics, symbolic logic would facilitate advances in ordinary mathematics was confirmed. In the first decade of the 20th century, Hilbert himself applied his model-theoretical techniques to problems in algebra, geometry, and mathematical physics with considerable success: already in *Foundations of Geometry* one finds the description of non-Archimedean geometries, a topological definition of the plane, and new results about continuous functions. With the further maturation of model theory, Mal'tsev, Tarksi, Robinson, and others successfully proved results in group theory and the theories of algebraic classes defined via interpretations of the sort found in *Foundations of Geometry* by applying the metatheorems of classical logic (e.g., the compactness theorem) to these domains. Robinson's statement of the significance of this breakthrough can be read as an acknowledgment that Hilbert's vision is being realized even if the dream that logic could answer all mathematical questions has been refuted: "[The] concrete examples produced in the present paper will have shown that contemporary symbolic logic can produce useful tools—though by no means omnipotent ones—for the development of actual mathematics, more particularly for the development of algebra and, it would appear, of algebraic geometry. This is the realization of an ambition which was expressed by Leibniz in a letter to Huygens as long ago as 1679." Hrushovski's solution to the Modell-Lang conjecture and other recent results indicate that model theory not only continues to have its uses in mathematical discovery but in fact has proved to be a setting for understanding ordinary mathematics at a fundamental level.

3. Proof theory

Whereas the consistency of a geometrical theory in which the axiom of parallels fails was a pressing open question to mathematicians in the 19th century, only the severest skeptics could doubt the consistency of basic arithmetic or even of mathematical analysis. Though Kronecker had earned scorn for his repudiation of higher-mathematics, Hilbert noted that "it was far from his practice to think further" about what he did accept, "about the integer itself." "Poincáre," too, "was from the start convinced of the impossibility of a proof of the axioms of arithmetic" because of his belief that mathematical "induction is a property of the mind." This conviction, like Cantor's insistence that "a 'proof' of their 'consistency' cannot be given" because "the fact of the 'consistency' of finite multiplicities is a simple, unprovable truth" (937) struck Hilbert as short-sighted. Why should one infer, from the fact that the consistency of a set of principles is not legitimately in doubt, the belief that the question of their consistency can not be meaningfully posed? Opposed to this way of thinking, Hilbert proposed that a definite mathematical problem can be formulated about the consistency of any axiomatic system, and at the dawn of the 20th century he set out to show just this.

A new tactic would be needed for the task, for the relative consistency proofs of Hilbert's earlier work appear to be unavailable in the case of arithmetic. The consistency of various geometric theories had been proved relative to that of arithmetical ones, but relative to what could the consistency of arithmetic be meaningfully established? "Recourse to another fundamental discipline," Hilbert remarked, "does not seem to be allowed when the foundations of arithmetic are at issue" (1904, p. 130). The consistency of arithmetical theories must be established in some sense "directly."

His wish to design direct consistency proofs led Hilbert to return once more to the fundamental insight of the metatheoretical perspective: the fact that the axioms of a formalized theory could be viewed as meaningless inscriptions. Rather than, as before, using this fact to construct *reinterpretations* of the axioms by changing the meanings of the terms they contain, Hilbert now proposed that the theory be left *uninterpreted*, so that each axiom, and indeed each formal derivation from the axioms, could be treated as an object for the mathematician to reason about. The production of mathematical proofs is an activity infused with meaning, but a proof in a formal axiomatic theory is "a concrete object surveyable in all its parts" (Bernays 195). This groundbreaking idea, which gives shape to the whole enterprise of proof theory, first appeared in Hilbert's 1917 talk at Zürich:

> ... we must—this is my conviction—take the concept of the specifically mathematical proof as an object of investigation, just as the astronomer has to consider the movement of his position, the physicist must study the theory of his apparatus, and the philosopher criticizes reason itself.

The question of the consistency of a system of axioms can be posed formally as the question whether there are proofs in the system of two contradictory results. It is easy to see that this formulation is equivalent to the question whether there is a proof in the system of some one predesignated evident falsehood. In the case of an arithmetical theory $\mathbf{T}$, a direct consistency proof would thus amount to an (informal mathematical) argument that it is not possible to derive the expression $1 \neq 1$ from axioms of $\mathbf{T}$ by means of the rules of inference designated for $\mathbf{T}$.

Before describing how such an argument might unfold, a few words about the philosophical debate surrounding the whole program are due. Hilbert emphasized in many of his early papers that the informal mathematical arguments comprising his consistency proofs do not involve reasoning more "complex" or "dubious" than the principles of reasoning encoded in the axioms of the theory about which one is reasoning. It is not hard to see why: If one draws from complex principles in order to show that a relatively simple theory is consistent, then it is not clear that one has demonstrated anything, for in even asking about the consistency of a theory (even if one does not literally harbor any doubts), one has presumably assumed a position from which that theory's own principles are not all available. To use them, and especially to use principles stronger than them, would seemingly be to drop the question one meant to be asking. Indeed, the proofs issued by Hilbert and his colleagues Ackermann and von Neumann were criticized on precisely these grounds, their insistence that they had avoided any such circularity notwithstanding. Interesting and heated debates ensued, fueled in large part by Brouwer's and Weyl's allegiance to constructivist principles violated by the systems Hilbert aimed to prove consistent. By Gödel's second incompleteness theorem, it is known that no proof of the consistency of any consistent arithmetical theory of the sort Hilbert studied can be formalized in that very theory, that the background theory in which a consistency proof is carried out must be in at least some ways stronger. The debate about the sense in which Hilbert's methods are circular as well as the debate about whether the epistemological gains of a consistency proof were actually of central importance to Hilbert continue to this day. To the latter issue, Kreisel's report that "Hilbert was asked (before his stroke) if his claims for the ideal of consistency should be taken literally" and that Hilbert "laughed and quipped that the claims served only to attract the attention of mathematicians to the potential of proof theory" is noteworthy (p. 43). In any case, the mathematical and logical interest of Hilbert's style of consistency proof is completely unscathed by any defect in the epistemological motivations of Hilbert or anyone else.

In the vicinity of these issues, however, one finds Hilbert describing the logic of his proof theory as "finitist." As others had objected in principle to the use of non-constructive existence proofs and uses of the law of excluded middle on infinite totalities like those featured in Hilbert's solution of Gordan's problem, now Hilbert himself banned these techniques. But Hilbert's stance was certainly not based on a principled opposition to classical logic: On the contrary, his aim was to demonstrate that mathematical theories laden with non-constructive principles are consistent and conservative extensions of the finitary ones used to reason about their proofs. Again, a convincing interpretation of Hilbert's several remarks about the significance of finitism has proved to be elusive: Was he hoping only to show that infinitary mathematics and non-constructive techniques are safe and efficient, though in fact meaningless, tools for discovering facts about the "real," finitary realm? Or was he rather fully in defense of the meaningfulness of ordinary mathematics, adopting finitist restrictions in his proof theory in an attempt to avoid circularity and thereby arrive at meaningful consistency arguments? A third reading, closer to the attitude of contemporary proof theorists, is that the stipulation of restrictions in one's logic was simply a response to the constructive nature of proof transformations. Each interpretation has its textual support, but whatever Hilbert's motives were, it is undeniable that to him the logic of proof theory and the logic of ordinary mathematics are importantly different. For Hilbert there is no single "true logic"; rather, the logic appropriate for a particular investigation is derived *a posteriori* from the details of that investigation—a position that foreshadows the contemporary notion of "logical pluralism."

In roughest outline, the reasoning in a direct consistency proof is the following *reductio ad absurdum*: Assume that a proof in the theory $\mathbf{T}$ of $1 \neq 1$ has been constructed (so that $\mathbf{T}$ is evidently inconsistent). This object will be a finite list of formulas, the last of which is $1 \neq 1$, each of which is an axiom of $\mathbf{T}$ or the result of applying one of the rules of inference of $\mathbf{T}$ to formulas that appear earlier in the list. Following a general algorithm, first transform this object into another proof in $\mathbf{T}$ of $1 \neq 1$ containing only closed formulas, typically in some perspicuous normal form. Then transform the resulting proof into a third object (again this will be a finite list of formulas, but not necessarily a proof) consisting entirely of formulas containing only numerals, propositional connectives, and the equality sign. A recursive argument can then be used to verify that every formula (beginning with those that emerged from axioms, continuing to those that emerged from formulas that were arrived at in one inference, etc.) in this list is "correct" according to a purely syntactic criterion (for this, Hilbert stipulated that numerals were finite strings of the symbol "1" concatenated.) This is the birth of a common proof-theoretical technique: In contemporary parlance, one has shown that the axioms each have this property and that the inference rules "preserve" the property. After a finite number of steps, one will have verified that the formula $1 \neq 1$ is "correct" (although the way Hilbert defined this notion it, of course, is not). From this contradiction, one concludes that no proof of $1 \neq 1$ can be constructed in $\mathbf{T}$, so that $\mathbf{T}$ is consistent.

A special case to consider is the treatment of quantifiers $\forall$, $\exists$ in this algorithm. To facilitate the "elimination" of quantifiers along the way to construct a list of purely numerical formulas of the sort just described, Hilbert rendered his arithmetical theories in the ϵ-calculus. This is an extension of quantifier-free number theory with a function symbol ϵ that operates on formulas $A(a)$ to form terms $\epsilon_a A(a)$, with the intuitive meaning "that which satisfies the predicate A if anything does." (If this construction seems peculiar, bear in mind that $\exists x(\exists y A(y) \supset A(x))$ and $\exists x(A(x) \supset \forall y A(y))$ are first-order logical truths, the first corresponding to the ϵ term, and the second corresponding to its dual τ term—"that which satisfies A only if everything does.") Hilbert's arithmetical theories included the *transfinite axioms*:

1. $A(a) \supset A(\epsilon_a A(a))$ and
2. $\epsilon_a A(a) \neq 0 \supset \neg A(\delta\epsilon_a A(a))$ (δ is the predecessor function)

It is not hard to see that these axioms allow one to derive the usual axioms for the quantifiers and induction.

To accommodate the ϵ-terms that appear in a proof, the algorithm stipulates how to substitute numerical terms for each appearance of the ϵ-terms: In the case that only one ϵ-term appears in the proof, one simply replaces each of its occurrences with the numeral 0 and each occurrence of $\delta\epsilon_a A(a)$ with the least n for which $A(n)$ is true. When more than one ϵ-term appears in the proof and especially when there are nested ϵ-terms, the substitution becomes far more complicated, and the algorithm loops in response to these complications.

Now, the strength of the metatheory needed to conduct such reasoning is determined by the problem of verifying that the algorithm for proof transformation eventually comes to a halt. The transformed proofs are typically much larger than the objects with which one begins, and in order to rule out the possibility that *any one* of the proofs in $\mathbf{T}$ ends in $1 \neq 1$, one in effect considers the transformation of *every one* of the proofs of $\mathbf{T}$. For this, because of the spiraling nature of the algorithm, one must use multiply embedded inductive arguments. This comes to light especially in the consideration given to the treatment of ϵ-terms. When several terms are present, some falling within the scope of others, these must be indexed, and an ordering defined on the indices, in order to keep track of how the proof transformation proceeds. One then observes that verifying that the algorithm eventually halts involves "transfinite induction" through the ordinal number used to order the indices of terms. Already in *1924*, Ackermann was explicit that induction to $\omega^{\omega^{\omega}}$ was used in the proof of the consistency of a theory he considered. He wrote: "The disassembling of functionals by reduction does not occur in the sense that a finite ordinal is decreased each time an outermost function symbol is eliminated [as in an ordinary inductive proof]. Rather, to each functional corresponds as it were a transfinite ordinal number as its rank, and the theorem that a constant functional is reduced to a numeral after carrying out finitely many operations corresponds to [a previously established fact]." (13). This is the sense in which today one speaks of ordinals associated with mathematical theories: The "consistency strength" of a theory is measured by the ordinal used to track the induction used to prove its consistency.

In the continuation of the passage just quoted, Ackermann claimed that the use of transfinite induction did not violate Hilbert's "finitist standpoint." This posture was later emulated by Gentzen, who presented his masterful consistency proof of first-order Peano Arithmetic (PA) together with a statement that although one "might be inclined to doubt the finitist character of the 'transfinite' induction" through

$$\epsilon_0 = \sup_{n<\infty} \omega_1^{\omega_2^{\cdot^{\cdot^{\omega_n}}}}$$

used in his proof "even if only because of its suspect name" it is important to consider

> that the reliability of the transfinite numbers required for the consistency proof compares with that of the firstinitial segments . . . in the same way as the reliability of a numerical calculation extending over a hundred pages with that of a calculation of a few fines: it is merely a considerably vaster undertaking to convince oneself of this certainty . . . . (p. 286)

The debate about the relationship between Hilbert's wish to provide a finitist consistency proof of arithmetic, Gödel's theorem to the effect that any consistency proof would have to extend the principles encoded in the theory one is proving to be consistent, and Gentzen's proof (which is carried out in the relatively weak theory PRA together with the relatively strong principle of transfinite induction up to ϵ_0) is not likely to resolve any time soon. From the point of view of logic, however, this debate is a distraction from what seems to have been Hilbert's main purpose: to show that the question of the consistency even of elementary mathematical theories could be formulated as a mathematical problem requiring new perspectives and techniques for its solution and ushering in rewarding insights along the way.

An example of such an insight can be extracted from Gentzen's achievement: By showing that transfinite induction to the ordinal ϵ_0 can be used to prove the consistency of PA, Gentzen demonstrated that this principle is unprovable in PA (This follows immediately from Gödel's "second incompleteness theorem," mentioned above). But he also showed that transfinite induction to any ordinal below ϵ_0 (any ordinal $\omega_1^{\omega_2^{\cdot^{\cdot^{\cdot^{\omega_n}}}}}$ for $n \in \mathbb{N}$) is provable in PA. This is the sense in which ϵ_0 is sometimes described as *the* ordinal of PA: no smaller ordinal will suffice. But Gentzen did more. PA has as an axiom a principle of mathematical induction over all formulas in its signature. One can also consider fragments of PA defined by restricting induction to formulas with a maximum quantifier complexity (call these the theory's class of inductive formulas). Gentzen showed that the size of the least ordinal sufficient for a proof of the consistency of such a fragment corresponds with the quantifier complexity of that theory's class of inductive formulas. In effect he established a correspondence between the number of quantifiers of formulas in the inductive class and the number of exponentials needed to express the ordinal that measures the theory's consistency strength. This correspondence "one quantifier equals one exponential" has been called the central fact of the logic of number theory and is rightly seen as the maturation of Hilbert's technique of quantifier elimination, the realization of Hilbert's idea that consistency proofs can be used to analyze quantifiers.

Though observing, in 1922, that "the importance of our question about the consistency of the axioms is well-recognized by the philosophers," Hilbert strove to distinguish his interests from theirs. "But in this literature," he continued, "I do not find anywhere a clear demand for the solution of the problem in the mathematical sense." That, we have seen, is what Hilbert demanded, and if Gentzen can be credited with providing the solution, Hilbert must be credited for formulating the question in purely mathematical terms and for articulating a setting for its investigation.

As for the philosophical significance of the result, one must appreciate that the "mathematical sense" of the consistency question, unprecedented and unpopular at the time that Hilbert first put it forward, is the one that commands the interest of mathematicians today as well as the one that brought logical investigations once again, in a yet different way, into the mainstream of mathematical activity. Anticipating this revolution, Bernays remarked that "Mathematics here creates a court of arbitration for itself, before which all fundamental questions can be settled in a specifically mathematical way, without having to rack one's brain about subtle logical dilemmas such as whether judgments of a certain form have a meaning or not" (222). In Hilbert's proof theory, he wrote elsewhere, "mathematics takes over the role of the discipline formerly called mathematical natural philosophy" ().

4. Logical completeness

In the description of a "symbolic calculus" with which he began his treatise on *Trigonometry and Double Algebra*, Augustus de Morgan listed three ways in which a formal system, even one whose "given rules of operation be necessary consequences of the given meanings as applied to the given symbols," could nevertheless be "imperfect." The last sort of imperfection he considered is that the system "may be incomplete in its rules of operation." He explained: "This incompleteness may amount either to an absolute privation of results, or only to the imposition of more trouble than, with completeness, would be required. Every rule the want of which would be a privation of results, may be called *primary*: all which might be dispensed with, except for the trouble that the want of them would give, may be treated merely as consequences of the primary rules, and called *secondary*."

Evidently, De Morgan would fault a system, not only for our inability in principle to prove with it everything we would like to know (all the truths or valid laws in some domain), i.e., for its lacking certain primary rules, but also for being cumbersome. But in distinguishing these two weaknesses, he clearly isolated a property of logical systems converse to the first one he mentioned. To say that all a system's rules are necessary consequences of the given meanings is to say that the system is sound. To say that it has enough rules to derive each such necessary consequence is to say that it is complete.

Some years later in an paper called "On the algebra of logic," Peirce boldly asserted, "I purpose to develope an algebra adequate to the treatment of all problems of deductive logic," but issued this caveat: "I shall not be able to perfect the algebra sufficiently to give facile methods of reaching logical conclusions; I can only give a method by which any legitimate conclusion may be reached and any fallacious one avoided." The concern about efficiency had been dropped. Peirce sought only to present a sound (avoiding any fallacious conclusion) and complete (reaching each legitimate one) logic.

It is entirely mysterious why Peirce felt entitled to claim that his logical system is complete. No argument of any sort to this effect appears in his paper. De Morgan made no such boast. But the two logicians shared an appreciation of what a good logical system would be like. One wonders what the source of this commonality might have been, because outside the writing of these two, one scarcely finds a hint that the question of completeness even occurred.

One exception is Bolzano, who a half century earlier developed logical theory on two fronts. He designed one theory of "ground and consequence" that was supposed to track the dependencies of truths and another of "derivability" that was supposed to allow us to reason from hypotheses to their necessary conclusions. Those dependencies of truths were the things we are supposed to care about, and derivability was merely an analysis of well-regulated reason. Peculiar, though, is the fact that his theory of ground and consequence closely resembles the logical calculi of the modern era, whereas his definition of derivability is a precursor to today's notion of logical consequence. So when Bolzano asked if every ground-consequence relation is in fact derivable, he seems to have our concept of completeness, which we inherited from De Morgan and Peirce, the wrong way around. Evidently, the completeness question, which seems today so perfectly natural and central, is of recent vintage.

Bolzano despaired at not finding a way to formulate his version of the completeness question so that he would know how to answer it, which is noble compared to De Morgan and Peirce's apparent lack even of an attempt at such a formulation. The whole enterprise was rather ill-fated, and by the turn of the 20th century, as logic began to settle into its modern guise, the question of logical completeness simply did not arise. Logicians began either to think of logical systems as primitive encodings of the principles of right reasoning, with no eye towards matters of logical truth, or not to think of them at all, focusing entirely on matters of truth, satisfaction of formulas, and such semantic notions. Those like Gentzen in the first camp could still ask if anything was missing from their systems, but the question was a matter of psychological introspection or possibly an empirical study of the types of inferences that appear in mathematical journals. Skolem and others in second camp had no systems to ask after and pursued instead questions of decidability of classes of formulas.

In this setting, Hilbert stood alone. He was discontent equally with the idea of empirically validating logical systems and of ignoring them altogether. In an address at the Bologna International Congress of Mathematicians, Hilbert remarked: "[T]he question of the completeness of the system of logical rules [of the predicate calculus], put in general form, constitutes a problem of theoretical logic. Up till now we have come to the view that these rules suffice only through experiment" (Hilbert 1929, p. 140). The sentiment first appeared in Hilbert's lectures from the academic year 1917–18, in which he remarked that "whether [the predicate calculus] is complete in the sense that from it all logical formulas that are correct for each domain of individuals can be derived is still an unsolved question," because our knowledge about this at the time was entirely "empirical."

As before, it is not clear that Hilbert harbored any doubts about the completeness of quantification theory. Quite possibly, the empirical evidence that familiar systems could not be improved upon with the addition of new principles was convincing to him. But even if the evidence was conclusive, what the world did not have and what the prevailing attitudes about logic precluded was a mathematical proof of a theorem about these matters. Once again, Hilbert seemed to be motivated to transform a question from philosophy or natural science into a mathematical problem and to see what sorts of mathematical ideas would be generated in the process. The stock of fundamental insights is vast. But Hilbert's own intellectual trajectory to posing the question is subtle, and it is rewarding to trace this history before tallying the spoils.

In 1905 Hilbert defined completeness in two apparently different ways. He asked first (p. 13) whether or not the axioms of a formal theory suffice to prove all the "facts" of the theory in question. On page 17 he formulated a "completeness axiom" that he claimed "is of a general kind and has to be added to every axiom system whatsoever in any form." Such an "axiom" had first appeared in *Foundations of Geometry*. Hilbert explained:

> ...in this case ...the system of numbers has to be so that whenever new elements are

> added contradictions arise, regardless of the stipulations made about them. If there are things that can be adjoined to the system without contradiction, then in truth they already belong to the system.

The first thing to notice is that Hilbert is both times speaking about axiomatic mathematical theories and is not yet asking about the completeness of a logical calculus. But the dissimilarities are also important: The completeness axiom is an axiom *in* a formal theory; the first notion of completeness is a property *of* such a theory. One wonders why the same word would be used in these two different ways.

A partial answer can be found in 1917–18 where Hilbert elided these notions together as he turned his attention to logical calculi. In this passage, he is discussing propositional logic:

> Let us now turn to the question of completeness. We want to call the system of axioms under consideration complete if we always obtain an inconsistent system of axioms by adding a formula which is so far not derivable to the system of basic formulas. (152)

Here, for the first time, the question of the completeness of a logical calculus has been posed as a precise mathematical problem. But although the question is being asked about a system of logic rather than being formulated as a principle within the system, the question bears more structural resemblance to the axiom of completeness than to the primitive question about all "facts" (or tautologies) of propositional logic being proved: The completeness axiom says that the addition of any new element in, for example, an algebraic structure will result in a contradiction; the completeness of the propositional calculus is the conjecture that the addition of any formula to a set of theorems will result in a contradiction.

These lecture notes contain proofs both of the consistency of a calculus for propositional logic and of the completeness, in the sense just described, of that same system. To prove consistency, Hilbert used the following interpretation strategy (the system under consideration contains connectives only for disjunction and negation): Let the propositional variables range over the numbers 1 and 0, interpret the disjunction symbol as multiplication and the negation symbol as the function $1 - x$. In this interpretation, every formula in the classical propositional calculus (CPC) is a function on 0 and 1 composed of multiplication and $1 - x$. Hilbert observes that the axioms are each interpreted as functions that return the value 0 on any input, and that the rules of inference each preserve this property (so that every derivable formula is constant 0). Furthermore, the negation of any derivable formula is constant 1 and therefore underivable. Thus, no formula is derivable if its negation also is, and so the system is consistent.

The same interpretation figures in the completeness argument: It is known how to associate with every formula ϕ another ϕ_{cnf} in "conjunctive normal form," so that ϕ and ϕ_{cnf} are each derivable from the other in CPC. (A cnf formula is a conjunction of clauses, each of which is a disjunction of propositional variables and negations of propositional variables.) By the previous argument, a formula is provable only if it is constant 0 under the interpretation, which, in the case of a cnf formula, occurs precisely when each of its clauses contains both a positive (unnegated) and negative (negated) occurrence of some propositional variable. Now let ϕ be any unprovable formula. Then its associated formula ϕ_{cnf} must also be underivable and therefore must contain a clause c with no propositional variable appearing both positively and negatively. To show that CPC$+\phi$ (CPC augmented with ϕ as an additional axiom) is inconsistent, let ψ be any formula whatsoever, and label χ the result of substituting, into ϕ_{cnf}, ψ for every variable that occurs positively in c and $\neg\psi$ for every variable that occurs negatively in c. It is easy to show that CPC$+\phi \vdash \phi$, CPC $\vdash \phi \supset \phi_{\text{cnf}}$, CPC $\vdash \phi_{\text{cnf}} \supset \chi$, and CPC $\vdash \chi \supset \psi$. Thus CPC$+\phi \vdash \psi$ for any formula ψ.

Suppose, now, that some formula ϕ were constant 0 under this interpretation but unprovable in CPC. Then the same consistency argument as before would carry through for CPC$+\phi$, contradicting the completeness result just proved. It follows that every formula interpreted as a constant 0 function is a theorem of CPC. This reasoning, reproduced from Hilbert's lectures, establishes that CPC is complete with respect to the functional interpretation and foreshadows the concept of "semantic completeness" familiar today. It might be surprising initially that these two notions of completeness, the first purely syntactic and the second establishing a bridge between a formal system and its interpretive scheme, mesh so nicely for propositional logic. It might be more surprising, still, that the syntactic criterion held the primary role in Hilbert's thought—his school customarily referred to it, not only as the "stronger" sense of completeness (which it is), but also as the "stricter" sense of the word. The influence of the concept of the completeness of mathematical theories is palpable.

The coincidence of these two notions of completeness served the Hilbert school well for their investigations of propositional logic, for even if the primitive notion of logical completeness is the one about the tractability, with one's logical system, of all truths, they always held talk about truth and content at arm's distance for being insufficiently "formal." Hilbert preferred to use interpretations as tools for discovering purely mathematical facts and wished to avoid debates about which interpretations are correct. So if the primitive question of logical completeness could be sharpened into one entirely about formal provability, Hilbert viewed this as progress.

As it happens, however, such was not the fate for the concept of completeness as it arises for quantification theory. By the publication of *Hilbert and Ackermann* in 1928, Ackermann had discovered that unprovable formulas can be consistently added to the classical predicate calculus. To see this, one again begins by verifying an axiom system's consistency with an interpretation: To interpret a formula of first-order quantification theory, first erase all quantifiers. Interpret propositional variables and propositional connectives as before (variables range over $\{0, 1\}$, disjunction is the multiplicative product, etc.) Further, ignore how the argument places of the predicate letters are filled, and interpret these also as variables ranging over $\{0, 1\}$. As before, each axiom gets interpreted as a constant 0 function, and each rule of inference preserves this property. But the negation of any theorem is a constant 1 function and therefore underivable, so the system is consistent. (Unlike the case of propositional logic, however, this interpretation in no way foreshadows a semantic theory. It is *only* a tool for metatheoretical investigations, not an *intended* interpretation.)

The incompleteness, in the strong sense, of the predicate calculus is witnessed by the formula $\exists x F(s) \supset \forall x F(x)$, which gets interpreted as a constant 0 function in the scheme just described. To see that this formula is underivable, one need only (1) observe that in any domain with more than one object, the sentence that results from this formula if we interpret F as any predicate true of only one thing is false and (2) appeal to the soundness of the system under consideration. Hilbert's scruples are evident when he describes this reasoning as merely making the underivability of the formula "plausible"—he proceeds to present a "strictly formal proof" of its underivability with no appeal to the system's soundness or matters of truth and falsity.

In any case this result certainly does not establish that there are logically valid formulas of quantification theory that are unprovable in the predicate calculus. Instead, it drives a wedge between the two senses of completeness that coincide in the case of propositional logic. The demonstration that "any legitimate conclusion may be reached and any fallacious one avoided" so that we are left with no "absolute privation of results" cannot be "sharpened" in this case so as to eliminate all talk of truth.

It is well-known that Gödel proved the semantic completeness of the predicate calculus in his 1929 thesis. Less well-known is that the question he answered in the process was not, despite his remarks, one that "arises immediately" to everyone who thinks about logic. It could not arise less than a century earlier, when the distinction between logical systems and the truths they are meant to track was completely reversed in the work of Bolzano, and it could barely figure into the thought of most logicians as a respectable problem in the ensuing years, because it had not been shown that such a question could be framed in precise mathematical terms. Things changed rapidly after Hilbert turned his attention to logic. If his attempts to treat the question of quantificational completeness did not materialize, the conceptual clarification ushered in by those attempts paved the way for Gödel's success.

There is little doubt, however, that Hilbert's distrust of semantic notions hindered his research on quantificational completeness as much as it helped him in the case of propositional logic. There is no "strictly formal proof" of the completeness of quantification theory precisely because the necessary semantics are ineliminably infinitary and therefore not accessible by the methods in which Hilbert aimed to cast all of his metatheoretical investigations. This is witnessed doubly by the existence, on the one hand, of quantificational formulas that can be true but only in an infinite domain and, on the other hand, by the prominent role of the law of excluded middle applied to an infinite domain in Gödel's completeness proof.

Again, it remains a matter of some debate exactly what Hilbert's reasons were for requiring all metatheoretical investigations to be finitary precisely as his most celebrated work in ordinary mathematics is not. But the seeds of this conviction were sown at the very beginning of Hilbert's logical investigations: The "implicit definition" doctrine that mathematical existence amounts to no more than the consistency of a set of axioms, Gödel pointed out, is hard to reconcile with Hilbert's insistence that the completeness of the predicate calculus should not be believed based on evidence but requires mathematical proof. For the completeness of the predicate calculus is equivalent to the claim that every consistent set of first-order formulas has a model. To see the implication in one direction, assume that if a collection of formulas is consistent, then they are all true in some model. By contraposition, if a formula is false in every model, then a contradiction can be derived from it. Let ϕ be a first-order validity (a formula true in every model). Then $\neg\phi$ is false in every model. Therefore, a contradiction can be derived from $\neg\phi$. So $\neg\neg\phi$ is derivable, as is ϕ by double negation elimination.

The hackneyed story about how Hilbert's logical program was undermined by Gödel's results must be toned down. To begin with, we have seen that Hilbert's investigations far extend the search for finitistic consistency proofs of mathematical theories and that the maturation of proof theory, in the work of Gentzen, and of the connections between proof theory and model theory, in the work of Gödel, are actually realizations of Hilbert's broader aims. As for Gödel's arithmetical incompleteness theorems, which would seem to refute at least some of Hilbert's more popular statements, their appropriation by the Hilbert school is instructive.

In short time, Hilbert and Bernays applied Gödel's own techniques to the problem of logical completeness and showed that if ϕ is a first-order quantificational formula that is not refutable in the predicate calculus, then in any first-order arithmetical theory $\mathbf{T}$ there is a true interpretation of ϕ. By Gödel's own completeness theorem, it follows that a first-order formula ϕ is valid if, and only if, every sentence of $\mathbf{T}$ that can be obtained by substituting predicates of $\mathbf{T}$ for predicate letters in ϕ is true. "The evident philosophical advantage of resting with this substitutional definition, and not broaching model theory," wrote Quine, "is that we save on ontology. Sentences suffice, sentences even of the object language, instead of a universe of sets specifiable and unspecifiable" (Quine 1986, p. 55). Whether a distrust of modern semantics such as Quine's motivated Hilbert can probably not be known. But it is a truly Hilbertian final word on the matter when one observes not only that what Hilbert would have deemed an "informal notion" is here reduced to a matter of arithmetical truth, but also that an exact relationship to the "stricter sense" of completeness is hereby salvaged: As a corollary to this arithmetical completeness theorem Hilbert and Bernays showed that the addition of any unprovable formula of quantification theory as a new axiom causes the formal arithmetical theory PA based on the predicate calculus to become ω-inconsistent.

5. Conclusion

In remarks published alongside papers that he, Heyting, and von Neumann wrote to articulate the philosophical positions know as logicism, intuitionism, and formalism, Rudolf Carnap distinguished the outlook of a typical logician, for whom "every sign of the language ... must have a definite, specifiable meaning," with that of the mathematician. The attitude of the latter, he thought, was exemplified by Hilbert when he said, "We feel no obligation to be held accountable for the meaning of mathematical signs; we demand the right to operate axiomatically in freedom, i.e. to set up axioms and operational specifications for a mathematical field and then to find the consequences formalistically" (141). History has shown that the latter cast of mind, though surely not destined to supplant the former in all investigations, has its place in the advancement of logical theory.

Problems

1. Show that the following are equivalent ways of formulating the consistency of CPC:

   (a) For no ϕ is $\vdash \phi$ and $\vdash \neg\phi$

   (b) For some $\phi, \nvdash \phi$

   (c) $\nvdash p \& \neg p$

2. Show that $\exists x A(x) \equiv A(\epsilon_x(A(x)))$.

3. Determine of each formula whether it is unsatisfiable, satisfiable in a finite domain, or unsatisfaibale in every finite domain but satisfiable:

   (a) $\forall x \forall z \exists y (R(x, y) \& \neg R(x, x) \& (R(y, z) \supset R(x, z)))$

   (b) $\forall x \exists y \forall z (R(x, y) \& \neg R(x, x) \& (R(y, z) \supset R(x, z)))$

   (c) $\exists y \forall x \forall z (R(x, y) \& \neg R(x, x) \& (R(y, z) \supset R(x, z)))$

4. Verify the four claims used in the final step of the proof of strong completeness for CPC.

5. Explain the other direction of the equivalence between “Every first-order validity is provable” and “Every consistent collection of first order formulas has a model”; i.e., show that if every validity is provable, then every formula from which no contradiction can be derived has a model.

6. Why does the underivability of $\exists x F(s) \supset \forall x F(x)$ establish the incompleteness, in the strong sense, of the classical predicate calculus?