

Deciding the Undecidable: Wrestling with Hilbert's Problems

Solomon Feferman*

In the year 1900, the German mathematician David Hilbert gave a dramatic address in Paris, at the meeting of the 2nd International Congress of Mathematicians—an address which was to have lasting fame and importance. Hilbert was at that point a rapidly rising star, if not superstar, in mathematics, and before long he was to be ranked with Henri Poincaré as one of the two greatest and most influential mathematicians of the era. Like Poincaré, Hilbert worked in an exceptional variety of areas. He had already made fundamental contributions to algebra, number theory, geometry and analysis. After 1900 he would expand his researches further in analysis, then move on to mathematical physics and finally turn to mathematical logic. In his work, Hilbert demonstrated an unusual combination of direct intuition and concern for absolute rigor. With exceptional technical power at his command, he would tackle outstanding problems, usually with great originality of approach.

The title of Hilbert's lecture in Paris was simply, "Mathematical Problems". In it he emphasized the importance of taking on challenging problems for maintaining the progress and vitality of mathematics. And with this, he expressed a remarkable conviction in the solvability of all mathematical problems, which he even called an axiom. To quote from his lecture:

Is the axiom of the solvability of every problem a peculiar characteristic of mathematical thought alone, or is it possibly a general law inherent in the nature of the mind, that all questions which it asks must be answerable?...This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no *ignoramus*¹.

To be sure, one problem after another had been vanquished in the past by mathematicians, though sometimes only after considerable effort and only over a period of many years. And Hilbert's own experience was that he could eventually solve any problem he turned to. But it was rather daring to assert that there are *no* limits to the power of human thought,

*Text of a lecture to a general audience at Stanford University, May 13, 1994.

¹For a fuller extract from Hilbert's lecture, see the Appendix.

at least in mathematics. And it is just this that has been put in question by some of the results in logic that I want to tell you about today.

Among mathematicians, what Hilbert's Paris lecture is mainly famous for is the list he proposed of twenty-three problems at the then leading edge of research. These practically ran the gamut of the fields of mathematics of his day, from the pure to the applied, and from the most general to the most specific. Naturally, the choice of these problems was to some extent subjective, restricted by Hilbert's own knowledge and interests, broad as those were. But the work on them led to an extraordinary amount of important mathematics in the 20th century, and individual mathematicians would become famous for having solved one or another of Hilbert's problems. In 1975, a conference was held under the title, "Mathematical Developments Arising from Hilbert Problems", which summarized the advances made on each of them to date. In many cases, the solutions obtained thus far led to still further problems which were being pursued vigorously, though no longer with the cachet of having Hilbert's name attached.

The solution of three of Hilbert's problems were to involve mathematical logic and the foundations of mathematics in an essential way, and it is these I want to tell you something about in this lecture. They are the problems numbered 1, 2 and 10 in his list, but for reasons that you'll see, I want to discuss them in reverse order.²

Problem 10 called for an *algorithm* to determine of any given *Diophantine equation* whether or not it has any integer solutions. I'll explain to you what is meant by these specialized terms later. For the time being, just think of an algorithm as some sort of mathematical recipe or computational procedure. Algorithms are ubiquitous; we use them automatically in our daily arithmetical calculations. And algorithms are built into the myriad kinds of electronic devices all around us and in the software we use to put our computers through their special paces. By Diophantine equations are meant equations expressed entirely in terms of integers and operations on integers, whose unknowns are also to be solved for integers. And by *integers*, of course, we mean the whole numbers 1, 2, 3, . . . extended to include 0 and the negative integers $-1, -2, -3, \dots$. The integers are one of the basic number systems of mathematics; others that I'll also have to say something about are the *rational number system*, which are simply ratios of integers, in other words fractions, and the *real number system*, which are the numbers used for measuring arbitrary lengths to any degree of precision. The most famous Diophantine equation is that addressed in the so-called Fermat's Last Theorem.

Contrary to Hilbert's expectations, Problem 10 was eventually solved in the negative. This was accomplished in 1970 by a young Russian mathematician, Yuri Matiyasevich, who built on earlier work in the 1950s and 1960s by the American logicians, Martin Davis, Hilary Putnam and Julia Robinson. (Incidentally, Robinson was just finishing her Ph.D. work in Berkeley with Alfred Tarski, one of the leading logicians of our time, when I began graduate studies there, and Tarski was to become my teacher as well. Because of her contributions in the following years to logic and recursive function theory, Robinson was eventually elected to the mathematical section of the National Academy of Sciences and later became President of the American Mathematical Society; moreover, she was the first woman to be honored in each of these ways.) The result of the Davis-Putnam-Robinson-Matiyasevich work, as

²Hilbert's statements of these problems are reproduced in full in the Appendix below.

we describe it nowadays, is that *the general problem of the existence of integer solutions of Diophantine equations is algorithmically undecidable*. Now the word ‘undecidable’ here is being used in a very special technical sense that I’ll explain to you later; there is also a second technical meaning of the term which we’ll come to shortly. Neither of these has to do with the kinds of minor and major *indecisions* that we face in our daily lives—e. g. , whether to repaint the bathroom “Whisper White” or “Decorator White”, or what to do about Bosnia.

To return to the undecidability result that was just stated: it’s quite definite—there’s no question about it—but that’s by no means the end of the story. For, number theorists have been working on decision procedures for special classes of Diophantine equations, and there’s a sizeable gap between what’s known to be decidable by their work, and what’s established to be undecidable by the work of Davis, Putnam, Robinson and Matiyasevich. Moreover, Hilbert’s 10th Problem is just one of a host of decision problems that have been settled one way or the other, and even where the result is positive there remain significant open questions; these turn out to lie on the borderline of logic, mathematics and computer science. So this is the first thing I’ll want to take up in more detail, after introducing you to the other two problems.

Next comes Hilbert’s Second Problem, which called for a proof of consistency of the arithmetical axioms. Now, in 1900 Hilbert was a bit vague in stating just which axioms he had in mind in this problem. But when he took up logic full scale in the 1920s, he made quite specific what axioms were to be considered. Moreover, in order not to beg the question, he placed strong restrictions on the methods to be applied in consistency proofs of these and other axiom systems for mathematics: namely, these methods were to be completely *finitary* in character. The proposal to obtain finitary consistency proofs of axiom systems for mathematics came to be called *Hilbert’s Program* for the foundations of mathematics; to avoid confusion with his list of problems, I will refer to this as *Hilbert’s Consistency Program*. Hilbert himself initiated specific work in the 1920s on his formulation of Problem 2. Here again, contrary to Hilbert’s expectations, there was a negative solution, namely through the stunning results of the young Austrian logician Kurt Gödel, whose incompleteness theorems in 1931 have become among the most famous in mathematical logic. The title of Gödel’s paper was, “On formally undecidable propositions of *Principia Mathematica* and related systems.” *Principia Mathematica* was the landmark work of Russell and Whitehead whose aim was to give a formal axiomatic basis for all of mathematics. Gödel showed that for any such system (and even much more elementary ones), there will always be individual propositions in its language that are undecidable by the system, i. e., which can neither be proved nor disproved from its axioms provided it is consistent. Even more, Gödel showed that the consistency of such a system can’t be proved within the system itself, and so finitary methods can’t suffice. Though these results apparently knocked down Hilbert’s Consistency Program, several questions remain which reach to the very foundation of our subject: Is incompleteness an essential barrier to the process of discovery in mathematics, or is there some way that it can be overcome? In answer to that, Gödel himself proposed one way forward, in fact a way that connects up with Hilbert’s 1st Problem, as we’ll see. Later, the English logician and proto-computer-scientist Alan Turing proposed a quite different way to overcome incompleteness that will be explained at greater length. A second important

question is whether Hilbert's Consistency Program is still viable in any way, either by restricting its scope or by somehow enlarging the methods of proof to be admitted. There's been considerable research in recent years on both aspects of this question, and toward the conclusion of my lecture I'll tell you something about the current state of progress in that respect.

Hilbert's First Problem is in a way the most technical of the three, which is why I leave it for last. I will also have to content myself with giving an indication of its character and significance and what's been done about it. Here Hilbert called for a proof of Cantor's conjecture on the cardinal number of the continuum of real numbers, the so-called *Continuum Hypothesis*. I'll not even try to explain the details of this here. Let me just say that this is part of a subject called *set theory* and that all presently generally accepted facts in set theory have been derived from principles which have been codified in a specific system of axioms for this subject, called the *Zermelo-Fraenkel axioms* for set theory including (what's called) the *Axiom of Choice*. So naturally, one would also seek to decide the Continuum Hypothesis on the basis of these axioms, i. e. either prove or disprove it from them. The latter possibility was shown to be excluded by Gödel's second outstanding result, in 1938: he showed that *the Zermelo-Fraenkel axioms cannot disprove the Continuum Hypothesis*. There the matter rested until 1963 when Paul Cohen, of our own [Stanford] Mathematics Department, obtained the major result that *these same axioms cannot prove the Continuum Hypothesis*. In other words, *Cantor's conjecture is undecidable on the basis of currently accepted principles for set theory*, provided, of course, that the Zermelo-Fraenkel axioms are consistent. There was also a second part to Hilbert's Problem #1; in that he called for the construction of a specific well-ordering of the continuum. For those of you who know the meaning of this concept, let me just take a moment to explain how matters turned out in that respect, too. As you know, the Axiom of Choice, which I'm here counting among the Zermelo-Fraenkel axioms, implies that for any set there *exists* a well-ordering of that set, but it doesn't tell you how to construct one. In fact, I was able to show using Cohen's methods that it is consistent with the Zermelo-Fraenkel axioms plus the Continuum Hypothesis, that there is *no* definable well-ordering of the continuum—again contrary to Hilbert's expectations.

But now to return to Cantor's continuum problem itself, the first question to ask following the undecidability results of Gödel and Cohen is whether that situation could change by adding further axioms for set theory in some reasonable way. As it turns out, their results apply to *all* plausible (and even not so plausible) such extensions that have been considered so far. There are sharply divergent views as to whether it is still hopeful to obtain a reasonable extension of the Zermelo-Fraenkel axioms which will settle the Continuum Hypothesis. And, as in any subject, we have both optimists and pessimists: the direction one leans may be a matter of basic differences of temperament, but in this case I believe it really comes down to basic differences in one's philosophy of mathematics, namely, as to whether one thinks mathematics in general, and set theory in particular, is about some independently existing abstract, "Platonic" reality, or whether it is somehow the objective part of human conceptions and constructions. The Platonists say the Continuum Hypothesis must have a definite answer and so for them it is still hopeful to find that out by some means or other. I, for one, am a pessimist or, better, anti-Platonist about the Continuum Hypothesis: I think that the problem is an inherently vague or indefinite one, as are the propositions of

higher set theory more generally. On the other hand, I'm an optimist of sorts concerning Hilbert's Second Problem, so it's not just a matter of temperament. Anyhow, if you agree with me that the Continuum Hypothesis does not constitute a *genuine* definite mathematical problem, then its undecidability relative to any given axioms ceases to be an issue with which to struggle; it simply evaporates as a problem. Because of the limitations of time, and the technicality of this subject, I'll have to leave the discussion there, and won't try to say any more about Hilbert's First Problem in this lecture.

In each of these cases, what Hilbert apparently took to be a fairly definite problem for which he expected a positive solution, the outcome not only led in the opposite direction, but also to the very foundations of our subject. More broadly speaking, these results connect with the question whether there are any essential limits to the power of human reasoning. I don't pretend to have an answer to this question. My purpose here is just to try to explain what the attacks on the Hilbert problems have led to thus far, and why and how the struggle with them is still continuing.

So now, let's get down to work by returning to Hilbert's 10th Problem, which called for an algorithm for determining of any given Diophantine equation whether or not it has any integer solutions. The idea of an algorithm (as I said earlier) is that of a step-by-step procedure or sequence of rules to go from the data of any specific problem of a certain type to its solution. The data could be a number, or an expression, or a sequence of numbers and symbols, and so on. The word, 'algorithm' is derived from the name of the 9th century Persian mathematician, Mohammed Al-Khowârizmi who wrote a book of rules for adding, subtracting, multiplying and dividing numbers in our familiar base 10 notation. However, algorithms themselves are as old as mathematics; one of the most famous is called Euclid's algorithm for computing the greatest common divisor of two integers, which is much faster than the more obvious algorithm of factoring each number completely in order to combine all the common prime factors.

The word 'Diophantine' stems from the 3rd century A.D. Greek mathematician Diophantus, who was the first to study solutions of equations in integers, more or less systematically. The subject of Diophantine equations lapsed for a long time after that but was revived in the 17th century by Pierre de Fermat, and this subject has received steady attention by number-theorists ever since. Though Fermat made many important and lasting discoveries, which have been verified in one way or another, the so-called Fermat Last Theorem has challenged mathematicians to the present day. While the specific proposition stated in the theorem is simple enough to explain, I don't want to distract attention from our main line to spell it out, but I do want to tell something about the circumstances because they provide an interesting cautionary tale. What Fermat had done is write in the margin of his copy of the *Arithmetica* by Diophantus that he had found a truly marvelous demonstration of this proposition, which the margin was too narrow to contain. Well, did he or didn't he? In all other cases, we know from Fermat's correspondence and other evidence that he had done what he claimed to have done. But in this case there is no such further evidence, and we suspect he was simply mistaken—though we have good reason to believe that he saw how to handle certain specific cases. At any rate, many mathematicians worked on Fermat's Last Theorem with only limited partial success in the 300 following years, up to the present. Most of you are probably aware through the news reports in the summer of 1993 that the

Princeton mathematician Andrew Wiles announced in a lecture for specialists that he had finally found a proof of the Fermat Last Theorem. This proof is so advanced and difficult technically, that only a few experts in the world can work through it to help verify its correctness. As it happens, in the process of going over his 200-page manuscript, Wiles himself found a gap in his proof, and so it's up in the air right now whether that gap can be filled. It's quite possible that it can't, and though there will still be much of technical value from his proof, it will leave the Fermat problem as a continuing challenge to mathematicians. Probably we'll know within a year one way or the other. At any rate, the history of all this shows that problems about even very simple Diophantine equations are notoriously difficult. Surely, Hilbert was aware of this even back in 1900, and his incautious optimism in the face of that in his statement of the 10th Problem is surprising.

In the 1920s, other decision problems in algebra and logic emerged which looked equally difficult, and the feeling began to grow that for some of these, *no possible algorithm could work*, or—as we say nowadays—that the problem is *algorithmically undecidable*. Now, if someone comes along with a proposed algorithm to settle a given decision problem in a positive way, one can check to see that it does the required work (or at least, try to do so), without inquiring into the general nature of what constitutes an algorithm. But if it is to be shown that the problem is undecidable, one has to have a precise explanation of what algorithms can compute in general. Analogous situations had arisen earlier in mathematics, for example to show that there is no possible construction by straight edge and compass which will trisect any given angle, or that there is no possible explicit formula for finding the roots of any polynomial equation (in one unknown) of degree 5 or higher (we do have such formulas for degrees 1, 2, 3 and 4). In each of these cases, one first needed a precise characterization of everything that can be constructed or defined in the specified way, before showing that certain problems cannot be solved by such constructions or definitions.

Similarly, in order to establish undecidability results, one first needed to have a precise characterization of what in general can be computed by an algorithm. Several very different looking answers to this were offered by logicians—including Kurt Gödel, Alonzo Church and Alan Turing—in the mid 1930s, but they were eventually all shown to be equivalent. The most familiar explanation is that due to Turing, who described what can be done by an ideal computer if no restrictions are placed on how much time or memory space is required to carry out a given computation. These are called *Turing Machines* nowadays, and Turing's conception is the foundation of theoretical computer science, at least of *what can be computed in principle*. Both Church and Turing applied their characterization of what is algorithmically computable, to show that *the decision problem for (what is called) first-order logic is undecidable*, i. e. that no possible algorithm can be used to determine of any given proposition formulated in the symbolism of that logic whether or not it is universally valid. In the years following that first work of Church and Turing, many other logical and mathematical problems were shown to be undecidable. But it took a long time to settle Hilbert's 10th Problem. As I said earlier, significant progress was made on this in the 1950s and 60s by Davis, Putnam and Robinson, but it was not until 1970 that Matiyasevich was able to take the final crucial step to show that there is no possible decision procedure which will determine of an arbitrary Diophantine equation whether or not it has any integer solutions. Thus, at least in the terms that Hilbert posed the 10th problem, the answer is

definitely negative.

So what's unsettled here? Well, the best that the logicians working on this have been able to show is that there is no decision procedure for equations in 9 or more unknowns. But number-theorists have been working on comparatively special classes of Diophantine equations, and the best *they* have done is to obtain decision procedures for certain classes of equations in two unknowns. So, there is here a sizable gap between the positive cases and the negative cases, and it is really up in the air how this will be filled, if at all.

The vast body of higher mathematics deals with problems where it's not even appropriate to ask whether they fall under a mechanical decision procedure, i. e. whether they are algorithmically decidable or undecidable. Whatever one's expectations, the problems that I've mentioned about logical validity and about Diophantine equations *were* cases where it was appropriate to ask that kind of question but where as we have seen, the answers turned out to be negative. These results tend to confirm our everyday impressions about the inherent difficulty of such general problems. It is thus surprising to see how far logicians were able to establish the *decidability* of various non-trivial classes of algebraic problems. For example, in the 1930s, before he immigrated to the United States, Alfred Tarski established a famous decision procedure for the algebra of real numbers. This allows one to determine, among other things, of any finite system of equations and inequalities in any number of unknowns whether or not it has some solution in real numbers. Tarski's procedure has both theoretical and practical applications. However, when questions of actual feasibility of computation came to the fore with the advent of high-speed electronic computers in the 1950s, such decision procedures began to be re-examined. So here we face the new question of *computability in practice* instead of computability in principle.

In the mid-1970s, it was shown by Michael Fischer and Michael Rabin that *no* algorithm for the algebra of real numbers can work faster than exponential rate in general. That is, given a problem expressed with n symbols, it will, in general, take on the order of 2^n steps (i. e. $2 \times 2 \times 2 \times \dots, n$ times) to settle it. For $n = 50$, taking 2^{50} steps is beyond the limits of actual computers—it would require about 35 years full time, and for $n = 60$ or beyond, it would take over 30,000 years full time. Fischer and Rabin also obtained similar results for other cases where the decision problem had been settled positively. Remarkably, the decision problem for the first-order theory of integers under addition alone requires 2^{2^n} steps for an input of length n , which becomes prohibitive for $n = 6$ or beyond, and for multiplication it goes even one exponential level higher. All these might be considered examples of *undeciding the decidable*.

There are some specific problems concerning multiplication which are of great practical importance, since modern cryptographic systems are based on them. These are the problems of determining of an integer whether or not it is a prime number, and the related problem of factoring an integer into prime parts. It is believed that the factorization problem is not, in general, computationally feasible for large integers, though there is no proof of that. On the other hand, there are relatively quick methods of determining primality. And—if one is willing to give up absolute certainty—there are even much faster methods that work within seconds using non-deterministic procedures, by making successive random guesses; one of these is due to Michael Rabin, and another is due to Robert Solovay and Volker Strassen. So, if you start looking at the question, what can be decided within probability, say, 99%,

and within practical time limits, *all* problems such as those indicated above have to be re-examined.

You may have noticed that I hedged one statement above, namely that if someone comes to you with a proposed algorithm, you can check whether it does the required work—to which I added: ‘or, at least, try to do so’. The point is that there is *no* mechanical method to verify that an algorithm does what it is supposed to do, or even that it always terminates with an answer. Again, our picture of this has changed considerably with the advent of high-speed computers. Nowadays, algorithms are fed to machines in the form of programs, and many programs in actual practice are very long and complicated, sometimes requiring thousands of lines. It is thus of great practical importance to have good design principles for programs that allow one to break them into manageable, understandable parts and to have usable methods to prove their termination and correctness. This is really a problem in logic, and in recent years the theory of proofs originally developed for Hilbert’s Second Problem (that I’m going to tell you about next), has turned out to be one of the major tools to deal with these problems in computer science.

My own involvement in applications of proof theory to questions of termination and correctness of computer programs has been relatively recent, starting about five years ago. But I’ve been making good use in this of a general approach to systems of reasoning for various forms of constructive mathematics that I initiated some twenty years ago. I’m very happy to be participating in this way in the enormous ferment that has been taking place between logic and computer science, especially during the last decade.

Before getting into Hilbert’s Second Problem, I have to say something in very broad terms about its historical background going back to antiquity, and the historical tension between the process of discovery and invention in mathematics on the one hand and its systematization and verification on the other. Greek mathematics was dominated by geometry; the best known names from that period are those of Euclid and Archimedes. In Euclid’s *Elements*, circa 300 B. C. , geometry was developed axiomatically, and for many years this exposition was considered the ideal of what mathematics, if not all logical argument, is or should be like. Not much is known about Euclid personally, but what is known is that he was not himself one of the main creators of Greek geometry, but rather a systematizer of what was known by his day. Bertrand Russell said that Euclid’s *Elements* is certainly one of the greatest books ever written; a recent commentator added that it is also one of the dullest. From Euclid you get no idea about how mathematics is actually discovered, how one arrives at the constructions, in many cases ingenious, that lead from the data to the conclusion; one can only go through his proofs step by step to see that they are indeed correct. What has come down to us from Archimedes is much closer in spirit and practice to modern mathematics and, in fact, some of his methods anticipate the modern calculus and can’t be reduced to Euclidean geometry.

After the Dark Ages in Europe, mathematics was revived in the early Renaissance, first with the need for a more practical arithmetic for commerce, then with the development of algebra in the 15th and 16th centuries. The results of Greek mathematics were transmitted to the West through Arabian sources. In the 17th century, geometry and algebra were married in the so-called analytical geometry of Fermat and Descartes. Then mathematics exploded at the beginning of the 18th century through the creation of the calculus by Newton

and Leibniz, and especially with Newton's applications of calculus and differential equations to physics. Here we have the wholesale use of infinite methods in mathematics, and the use of troublesome concepts such as infinitesimals and infinite sums. In the 18th century and into the 19th century mathematics was mostly developed in a free-wheeling way. The processes of discovery and invention outran careful verification. That only began to receive systematic attention in the 19th century, first with the foundations of the calculus and analysis and then going still deeper with the axiomatic foundations of the number systems basic to mathematics. Richard Dedekind gave axioms for the real number system with a kind of reduction of that to the rational numbers, and those are in turn easily reduced to the positive integers; finally both Dedekind and Giuseppe Peano gave axioms for the system of positive integers. In the latter part of the 19th century, the very basics of logical reasoning in mathematics were analyzed in a new symbolic form by Gottlob Frege, and that work was combined with Peano's symbolism by Bertrand Russell in the early 20th century in his attempt to give a development of all of mathematics in completely symbolic logical form.

Even Euclid's geometry turned out to be lacking full rigor; without realizing it, Euclid had made implicit use of certain hypotheses that were not included in his axioms. Moreover, one of the most startling discoveries of the 19th century was the realization of the existence of non-Euclidean geometries, coming out of the independence of the parallel postulate from the other postulates. One of Hilbert's major pieces of work before 1900 was the development of Euclidean and non-Euclidean geometry in a completely rigorous way. One part of his work shifted attention to the question of *independence* of this or that axiom from the others. That's really a question of *consistency*: statement A is independent from a set of statements S if the negation of A is consistent with S .³ One way to prove consistency of a set of statements is to produce a model for it, and that model has to be defined in terms that we already recognize to be consistent. Hilbert was here evolving what we nowadays call a *metamathematical* point of view: we treat mathematics formally as what can be carried out in an axiom system; then we investigate questions of independence, consistency and completeness for these axioms. For Hilbert, a mathematical concept "exists" if it is determined by a system of axioms. And to be thus determined, the system must be both *consistent* and *complete*, that is it should suffice to prove or disprove every proposition in the subject matter under consideration. What Hilbert did in verifying the consistency of various combinations of the geometrical axioms was to construct models by means of analytic geometry, i. e., built up using the real number system. So, when in 1900 Hilbert raised his second problem, he was calling next for a proof of consistency of axioms for the real numbers. In addition, he raised there the following much more ambitious question. In the latter part of the 19th century, Georg Cantor's theory of sets had introduced revolutionary and surprisingly powerful concepts and methods to mathematics, through a full embrace of the mathematical infinite. Now, that had been shown to lead to inconsistencies when the basic set theoretic principles were assumed

³*NB.* There is another usage of 'independence' in logic whose meaning is close to this but different, namely; statement A is said to be independent of S if neither A nor its negation is provable from S . In the terminology used in the discussion of Hilbert's Problem 1 above, that is the same as saying that A is an undecidable proposition on the basis of S . And, in terms of consistency, that is equivalent to each of A and its negation being consistent with S . The two distinct senses of 'independence' are a source of possible confusion when reading the literature, especially when it is not specified which is being used.

unrestrictedly. In Problem 2, Hilbert also called for a consistency proof of a somehow restricted theory of sets, though he did not say anything about how that was to be regarded axiomatically. In fact, the first axioms for set theory were not introduced until a few years later by Ernst Zermelo, one of Hilbert's protégés.

Later, in the 1920s, when Hilbert turned to logic and metamathematics full scale (with the assistance of Wilhelm Ackermann and Paul Bernays), he proposed much more definite consistency problems to be solved. But overarching these was what he considered to be the general problem of the infinite in mathematics. Even Peano's axioms for the integers involved, according to his view, an implicit appeal to the infinite, and one would also have to demonstrate their consistency and completeness. But if all infinitary methods would have to be justified on a *prior* basis, that could only admit finitary methods, i. e. reasoning about finite combinations of objects confined entirely to such combinations. On Hilbert's view, with mathematics represented in axiomatic systems, and every proposition and proof represented in such a system as a finite sequence of basic symbols, one could hope to prove consistency of such a system by completely finitary methods, by showing that no possible proof could result in a contradiction. This would be done through a theory of proofs and their transformations which Hilbert initiated. Again, he was fully optimistic for his program to secure all of mathematics by finitary consistency proofs, first for number theory with Peano's axioms, then for the real number system and analysis, and finally for set theory.

Hilbert's program thus received a big shock in 1931 when Kurt Gödel showed that Peano's axioms, and moreover, any effectively described extension of those axioms—including analysis and set theory—is *incomplete* if consistent. Moreover, the consistency of such a system cannot be proved within itself, so any such proof would have to use something from outside the system. In particular, any system which incorporates all finitary methods cannot be proved consistent by those methods.

How does all this look to us now? It's generally agreed that Gödel's results definitively undermine Hilbert's program, at least as originally conceived. Incidentally, Hilbert himself never admitted that, so strong were his convictions that his general program was the only way to go: he said that Gödel's results showed it would be more difficult to carry his program through than he originally thought. His co-workers, though, saw that some modification would be necessary, and that in fact is what happened in the subsequent development of proof theory, as I'll indicate below.

But nowadays, the whole preoccupation with consistency proofs has receded as a matter of central concern. On the one hand, hardly anyone doubts that the axiom systems of number theory, analysis and set theory in current use are consistent. The reason is that—whether or not one is a Platonist—they accord with basic informal reasonably coherent conceptions we have of what these systems are supposed to be about, under which their axioms are recognized to be correct. Moreover, in the past, all inconsistencies resulted from very obvious defects of formulation, and those defects have been eliminated. To be sure, there could still be hidden defects we're not aware of, but none has emerged even after a massive amount of detailed investigation of these systems.

Also, Hilbert's idea that mathematical concepts “exist” *only* through axiom systems for them, is accepted by very few. For, given that the systems we use are necessarily incomplete (granted their consistency), no such system can be said to fully determine its subject matter.

So we are led back to philosophical questions about the nature of mathematical concepts and how we come to accept and have our knowledge about them, questions that are just the sort that Hilbert hoped to avoid by his consistency and completeness programs.

While consistency may no longer preoccupy us, Gödel's incompleteness results are still a live matter. For, proofs are the only means we have for final verification of mathematical truths. And proofs must proceed from what is evidently true or already known to be true by incontrovertible steps to what is to be established. And when we analyze this picture, we are led to the representation of proofs in formal axiomatic systems. Every system in use for other than very specialized parts of mathematics contains Peano's Axioms. So, by Gödel's results, it is incomplete (assuming it is consistent). From this line of argument, it appears that there will be sensible mathematical propositions which we can never prove or disprove, and that, contrary to Hilbert's general view: *Ignoramus et ignorabimus* (We do not know and we shall never know—Emil du Bois Reymond). Is this conclusion really justified by Gödel's incompleteness results?

To begin with, let's look more specifically at how Gödel proved those results themselves. What he did was to translate properties of symbols and finite sequences of symbols including formal propositions and proofs, into properties of numbers systematically assigned to those symbols and finite sequences of symbols. Then, for each of the axiom systems S to which his results apply, he was able to construct a number-theoretical proposition A which "says of itself" (under this translation) that it is not provable in S . Next he showed that if S is consistent then the proposition A constructed in that way is indeed not provable in S , but since that's exactly what A asserts of itself, it follows that A is true, though we can't establish it's truth *within* S . Now, if S only proves true statements, as we would want, then certainly the negation of A is also not provable in S , i. e. A is undecidable by S . (In fact, a much weaker hypothesis on S stated by Gödel is sufficient for this second part, and by a variant argument due to Barkley Rosser, it is sufficient for both parts to simply assume that S is consistent.)

Like Heisenberg's Uncertainty Principle, the Big Bang, and black holes, Gödel's incompleteness theorems and their self-referential aspects have attracted much popular attention. Paradoxical self-reference goes back to antiquity, and we still have no generally accepted way of accounting for the so-called Liar Paradox, "I am lying," or "This statement is false", which if true is false, and if false is true. What Gödel did was similar, using provability in S in place of truth, but unlike the liar statement, it is non-paradoxical, since we *do* have a proof of A , though not in S ; that's the escape hatch. I expect a number of you will have seen the popular and very successful (but, at 775 pages, *very* long) book, *Gödel, Escher, Bach*, by Douglas Hofstadter in which he made vivid but loose play with analogous uses of self-reference or self-return or "strange looping" in the art of M. C. Escher and the music of J. S. Bach, as well as in computer programs which refer back to or "call" themselves in the process of computation. This last use of self-reference is more aptly analogous to Gödel's than the others. Since Hofstadter's ideas are all over the place, in my opinion that's *not* where to go for a straight story on Gödel's theorem, but it is entertaining and stimulating and not *too* misleading. Other examples of popular bounce-offs that some of you may have seen are the books by Raymond Smullyan, *What is the Name of this Book?* and *Forever Undecided: A Gödel Puzzle Book*. And, science-fiction aficionados tell me of several recent

works featuring Gödel's theorem in their plots. But I'll bet that few, if any, of you are aware of what I think is the most unusual spin-off from the Gödel incompleteness theorems, namely the 2nd Violin Concerto by Hans Werner Henze, which incorporates a poem by Hans Magnus Enzensberger, entitled, "Homage à Gödel". Here's one of its high points. [This was followed by a passage from a taped performance of the composition.⁴]

But to come back to the Gödel example of an undecidable sentence A , it is disappointing in the following way: as soon as we follow Gödel's argument we are ready to accept A , i. e. , it becomes *decided*, though by an argument which goes beyond those represented in the system S . A second source of disappointment is that A itself has no prior mathematical interest unlike famous unsettled problems in number theory such as the Twin Primes Conjecture or Goldbach's Conjecture.

Now logicians, starting with work by Jeff Paris and Leo Harrington, have been chipping away at this last by coming up with propositions which use only ordinary mathematical notions—i. e. are not obtained by translating logical symbolism into number theory—and which are independent of Peano Arithmetic and many still stronger systems S . Unlike Gödel's A , which is constructed in the same way for each system S , these are more and more complicated the stronger one takes S . But here's the kicker: just like Gödel's statement A , we always know how to decide these more mathematical looking propositions, because they're provably equivalent to a form of consistency of S (unless there is a question about that very consistency).

Returning to Gödel's own proposition A undecidable in S , he also showed that if C is the number-theoretical translation of the statement that S is consistent, then C implies A in S . Hence, also C is not provable in S if S is consistent. Now C is of prior *metamathematical* interest, since it is just this which was the object of Hilbert's Consistency Program, to be established by finitary means. And it is not a "cooked up" statement, like A .

In Gödel's view, the "true reason" for his incompleteness theorems lies in the fact that beyond any system S , we must accept new axioms concerning arbitrary subsets of the universe of objects with which S deals. If S is itself a system of set theory, these new axioms are called "axioms of higher infinity," since the new sets obtained will be infinitely larger in a suitable sense than the sets which can be shown to exist in S . It is indeed the case that by adding such new axioms one is able to establish the consistency of S . Of course, we then obtain a new system S' —which is again incomplete—and then the process of adding new axioms must be repeated, so it must be iterated indefinitely. All this accords with Gödel's underlying belief in the Platonic reality of set theory and that the kind of informal reasoning which led us to accept the Zermelo-Fraenkel axioms, as true of this reality, can be continued to expand these axioms indefinitely to settle hitherto undecided propositions.

If one does not subscribe to Gödel's philosophy, there are much more philosophically conservative ways of overcoming incompleteness to be considered. The first of these was due to Alan Turing in 1939, following his work on the general theory of computation. Namely, starting with any given system S_1 all of whose number-theoretical consequences are true, (for example, Peano's Axioms), one simply adds the statement C_1 of consistency of S_1 as a new (and, in fact, true) axiom to S_1 in order to obtain S_2 , then similarly adds C_2 to S_2 to obtain S_3 and so on. However, doing this an unbounded finite number of times is not the

⁴Unfortunately, this is not available commercially.

end of the story, because the “limit” system (that we denote) S_ω , which is S_1 together with C_1, C_2, C_3 , etc., is still incomplete, so we can start the process all over again. As we say, the procedure of adding consistency statements can be iterated effectively into the transfinite. At any stage S in this iteration process, the consistency statement C for S is added as a new axiom to S in order to obtain a system S' at the next stage, and at limit stages one simply combines all axioms previously obtained to form a new system.

Turing’s main result about this procedure is a completeness theorem for purely universal arithmetical statements, i. e. for statements of the form: that all positive integers have an effectively checkable property. (The consistency statements C and also many number-theoretical statements are of this form, e. g., Goldbach’s conjecture). Turing’s result is that any such statement which is true is provable at some stage in the transfinite iteration process. There is, however, a catch to his result, namely that it is very dependent on *how* the passage into the transfinite is effected, and there is no obvious way in advance to prefer one such passage to another. This is a delicate matter which can’t be explained without fairly technical notions, and so I won’t try to say anymore about it here. But basically, the catch is that *we have traded incompleteness for uncertainty about which path to follow*; in other words, you don’t get something for nothing.

Formal consistency statements are just one kind of statement of trust in the correctness of a system; more general such statements are called *reflection principles*, because they result from reflecting on what has led one to accept that system in the first place. In the early 1960s, I obtained a considerable strengthening of Turing’s completeness result, and showed that for the transfinite iteration of adding a reflection principle as new axioms at each stage, one is able to prove *every* true arithmetical statement at some stage in this transfinite progression of systems. However, this completeness theorem suffers the same catch as Turing’s earlier result, since it is dependent on which paths are taken into the transfinite, the choice of which is not justified in advance.

Well, is there anything that can be rescued from all this that doesn’t require such a trade-off? The answer is, yes, if we’re willing to settle for less. Namely, if we impose a kind of “boot-strap” or autonomous generation condition that requires justifying which path into the transfinite is to be taken before moving down that path. I’ve called what we obtain by imposing this condition, an *autonomous transfinite progression of axiom systems*, and began investigating their properties some years ago. I’ve shown more recently that everything that is proved in such a progression for the case of Peano’s Axioms as S_1 , is in what I call the *reflective closure* of S_1 , which is defined directly without the apparatus of transfinite ordinals.

Now, in general, while the reflective closure of a system S_1 overcomes the incompleteness of S_1 to a considerable extent, it too is still incomplete. Its significance, as I’ve argued especially in the case of Peano’s Axioms for the initial system S_1 , is instead that it contains everything you *ought* to have accepted if you accepted the concepts and principles in S_1 . In order to go beyond the reflective closure of S_1 , one must use essentially new concepts or essentially new kinds of arguments. This accords with the historical necessity in mathematics to periodically introduce essentially new concepts and principles. However, more work needs to be done (and is waiting to be done) on what is obtained by reflective closure applied to other initial axiom systems, in order to test this general view of its significance.

What, finally, is to be said about the viability of Hilbert’s Program? The part of logic that

was developed to carry this out is called *proof theory*. In order to use this for the ultimate foundations of mathematics in which all vestiges of assumptions about the infinite would be eliminated, Hilbert had insisted on carrying out proof theory by strictly finitary means. Gödel's incompleteness theorems showed that if proof theory is to be applied to stronger and stronger systems, one will have to give up that restriction in one way or another. Indeed, once that was abandoned, proof theory made enormous strides forward, and has led to a number of technical results of very high order of importance. With no false modesty, let me say that I've had the good fortune to be one of the leaders in promoting these technical developments. At the same time, I've been very concerned that they not be pursued simply for their own sake; this requires a new rationale and conceptual framework to take the place of the original Hilbert Program. In my view, a fruitful way of looking at what has been accomplished by this work is as a *relativized* form of Hilbert's program, in which we develop a global picture of *What rests on what?* in mathematics, i. e. what higher-order concepts and principles can be reduced to *prima facie* more basic ones.

But there has also been unexpected progress in recent years in the opposite direction, to which I have also contributed, namely to see what can be done in systems to which strictly finitary proof theory applies. Here, the surprising result is the demonstration that an enormous amount of scientifically applicable mathematics, even using very modern methods, can be carried out in proof-theoretically very weak systems, or as I have put it elsewhere, that: *A little bit goes a long way*. This is of philosophical interest, because it shows that infinitary concepts are not essential to the mathematization of science, all appearances to the contrary. And this also puts into question the view that higher mathematics is justified by science or is somehow embodied in the world, rather than that it is the conceptual edifice raised by mankind in order to make sense of the world. If, then, the search for absolute truth in mathematics is replaced by the progressive clarification and improvement of human understanding, undecidability ceases to be a bug-a-boo.

If Hilbert were to reappear at a meeting of the International Congress of Mathematicians in the year 2000, he would have to take cognizance of all the work I've told you about with respect to these three problems. No doubt he would be shaken in his original naive formulations and his optimism concerning them. But I venture to say that he would come up with a new and more sophisticated version of his basic belief in the solvability of all mathematical problems, something like: *there are no genuine absolutely undecidable problems*. The hard part for him would be to say what constitutes a *genuine* mathematical problem, and by what means it is to be decided. But since he isn't coming back, it will be up to us and those who follow to try to answer these questions.

APPENDIX

Hilbert's Problems 1, 2, and 10

The circumstances surrounding Hilbert's address to the International Congress of Mathematicians' meeting in Paris in 1900 are well described in the biography *Hilbert* by Constance Reid (pp. 69-83). Due to limitations of time, he actually presented only ten out of his list of 23 problems there. The text of his lecture with the full list of problems was first published in German as Hilbert (1900) [see References below], and was followed by an authorized English translation under the title *Mathematical Problems* as Hilbert (1902); a reprinting of the latter is to be found in Browder (1976), pp. 1-34. Prior to the list of problems, the first part of that text (pp. 1-7 op. cit.) consists of a general discussion of the vital role played by mathematical problems in the advancement of mathematics, their varied sources, the necessity of rigorous demonstration in their solutions, and an acknowledgement of the fecundity of good notation and informal geometric reasoning. It concludes with Hilbert's statement of his conviction in the solvability of all definite mathematical problems; that section of his lecture is reproduced below, along with his statements (in full) of the problems numbered 1, 2 and 10.

Special note should be taken of Hilbert's recognition that in certain cases a problem may be "solved" by showing its impossibility of solution by prescribed methods or "hypotheses." Among these he mentions the parallel postulate and equations of the fifth degree. The former is independent of the remaining axioms of geometry, i. e., in our terminology, it is an undecidable proposition on the basis of those axioms. The latter are not solvable by radicals (i. e., their roots are not in general expressible by formulas built up from the coefficients by means of the arithmetic operations and n th roots); this is similar to the situation of algorithmically unsolvable (i. e., undecidable) problems, discussed above in connection with the 10th Problem. However, Hilbert muddies the picture in both cases by speaking of these as "insufficient hypotheses" or of seeking the solution "in an incorrect sense."

* * *

Herewith, the extracts from Hilbert (1902):

Occasionally it happens that we seek the solution [of a problem] under insufficient hypotheses or in an incorrect sense, and for this reason do not succeed. The problem then arises: to show the impossibility of the solution under the given hypotheses, or in the sense contemplated. Such proofs of impossibility were effected by the ancients, for instance when they showed that the ratio of the hypotenuse to the side of an isosceles right triangle is irrational. In later mathematics, the question as to the impossibility of certain solutions plays a preeminent part, and we perceive in this way that old and difficult problems, such as the proof of the axiom of parallels, the squaring of the circle, or the solution of equations of the fifth degree by radicals have finally found fully satisfactory and rigorous solutions, although in another sense than that originally intended. It is probably this important fact along with other philosophical reasons that gives rise to the conviction (which every mathematician shares, but which no one has as yet supported by a proof) that every definite mathematical problem must necessarily

be susceptible of an exact settlement, either in the form of an actual answer to the question asked, or by the proof of the impossibility of its solution and therewith the necessary failure of all attempts. Take any definite unsolved problem, such as the question as to the irrationality of the Euler-Mascheroni constant C , or the existence of an infinite number of prime numbers of the form $2^n + 1$. However unapproachable these problems may seem to us and however helpless we stand before them, we have, nevertheless, the firm conviction that their solution must follow by a finite number of purely logical processes.

Is this axiom of the solvability of every problem a peculiarity characteristic of mathematical thought alone, or is it possibly a general law inherent in the nature of the mind, that all questions which it asks must be answerable? For in other sciences also one meets old problems which have been settled in a manner most satisfactory and most useful to science by the proof of their impossibility. I instance the problem of perpetual motion. After seeking in vain for the construction of a perpetual motion machine, the relations were investigated which must subsist between the forces of nature if such a machine is to be impossible; and this inverted question led to the discovery of the law of the conservation of energy, which, again, explained the impossibility of perpetual motion in the sense originally intended.

This conviction of the solvability of every mathematical problem is a powerful incentive to the worker. We hear within us the perpetual call: There is the problem. Seek its solution. You can find it by pure reason, for in mathematics there is no *ignorabimus*.

...

1. CANTOR'S PROBLEM OF THE CARDINAL NUMBER OF THE CONTINUUM.

Two systems, i. e., two assemblages of ordinary real numbers or points, are said to be (according to Cantor) equivalent or of equal *cardinal number*, if they can be brought into a relation to one another such that to every number of the one assemblage corresponds one and only one definite number of the other. The investigations of Cantor on such assemblages of points suggest a very plausible theorem, which nevertheless, in spite of the most strenuous efforts, no one has succeeded in proving. This is the theorem:

Every system of infinitely many real numbers, *i. e.*, every assemblage of numbers (or points), is either equivalent to the assemblage of natural integers, 1, 2, 3, ... or to the assemblage of all real numbers and therefore to the continuum, that is, to the points of a line; *as regards equivalence there are, therefore, only two assemblages of numbers, the countable assemblage and the continuum.*

From this theorem it would follow at once that the continuum has the next cardinal number beyond that of the countable assemblage; the proof of this theorem would, therefore, form a new bridge between the countable assemblage and the continuum.

Let me mention another very remarkable statement of Cantor's which stands in the closest connection with the theorem mentioned and which, perhaps, offers the key to its proof. Any system of real numbers is said to be ordered, if for every two numbers of the system it is determined which one is the earlier and which is the later, and if at the same time this determination is of such a kind that, if a is before b and b is before

c , then a always comes before c . The natural arrangement of numbers of a system is defined to be that in which the smaller precedes the larger. But there are, as is easily seen, infinitely many other ways in which the numbers of a system may be arranged.

If we think of a definite arrangement of numbers and select from them a particular system of these numbers, a so-called partial system or assemblage, this partial system will also prove to be ordered. Now Cantor considers a particular kind of ordered assemblage which he designates as a well ordered assemblage and which is characterized in this way, that not only in the assemblage itself but also in every partial assemblage there exists a first number. The system of integers 1, 2, 3, . . . in their natural order is evidently a well ordered assemblage. On the other hand, the system of all real numbers, *i. e.*, the continuum in its natural order, is evidently not well ordered. For, if we think of the points of a segment of a straight line, with its initial point excluded, as our partial assemblage, it will have no first element.

The question now arises whether the totality of all numbers may not be arranged in another manner so that every partial assemblage may have a first element, *i. e.*, whether the continuum cannot be considered as a well ordered assemblage—a question which Cantor thinks must be answered in the affirmative. It appears to me most desirable to obtain a direct proof of this remarkable statement of Cantor's, perhaps by actually giving an arrangement of numbers such that in every partial system a first number can be pointed out.

2. THE COMPATIBILITY OF THE ARITHMETICAL AXIOMS.

When we are engaged in investigating the foundations of a science, we must set up a system of axioms which contains an exact and complete description of the relations subsisting between the elementary ideas of that science. The axioms so set up are at the same time the definitions of those elementary ideas; and no statement within the realm of the science whose foundation we are testing is held to be correct unless it can be derived from those axioms by means of a finite number of logical steps. Upon closer consideration the question arises: *Whether, in any way, certain statements of single axioms depend upon one another, and whether the axioms may not therefore contain certain parts in common, which must be isolated if one wishes to arrive at a system of axioms that shall be altogether independent of one another.*

But above all I wish to designate the following as the most important among the numerous questions which can be asked with regard to the axioms: *To prove that they are not contradictory, that is, that a finite number of logical steps based upon them can never lead to contradictory results.*

In geometry, the proof of the compatibility of the axioms can be effected by constructing a suitable field of numbers, such that analogous relations between the numbers of this field correspond to the geometrical axioms. Any contradiction in the deductions from the geometrical axioms must thereupon be recognizable in the arithmetic of this field of numbers. In this way, the desired proof for the compatibility of the geometrical axioms is made to depend upon the theorem of the compatibility of the arithmetical axioms.

On the other hand a direct method is needed for the proof of the compatibility of the arithmetical axioms. The axioms of arithmetic are essentially nothing else than the known rules of calculation, with the addition of the axiom of continuity. I recently collected them* and in so doing replaced the axiom of continuity by two simpler axioms, namely, the well-known axiom of Archimedes, and a new axiom essentially as follows: that numbers form a system of things which is capable of no further extension, as long as all the other axioms hold (axiom of completeness). I am convinced that it must be possible to find a direct proof for the compatibility of the arithmetical axioms, by means of a careful study and suitable modification of the known methods of reasoning in the theory of irrational numbers.

To show the significance of the problem from another point of view, I add the following observation: If contradictory attributes be assigned to a concept, I say, that *mathematically the concept does not exist*. So, for example, a real number whose square is -1 does not exist mathematically. But if it can be proved that the attributes assigned to the concept can never lead to a contradiction by the application of a finite number of logical processes, I say that the mathematical existence of the concept (for example, of a number or a function which satisfies certain conditions) is thereby proved. In the case before us, where we are concerned with the axioms of real numbers in arithmetic, the proof of the compatibility of the axioms is at the same time the proof of the mathematical existence of the complete system of real numbers or of the continuum. Indeed, when the proof for the compatibility of the axioms shall be fully accomplished, the doubts which have been expressed occasionally as to the existence of the complete system of real numbers will become totally groundless. The totality of real numbers, *i. e.*, the continuum according to the point of view just indicated, is not the totality of all possible series in decimal fractions, or of all possible laws according to which the elements of a fundamental sequence may proceed. It is rather a system of things whose mutual relations are governed by the axioms set up and for which all propositions, and only those, are true which can be derived from the axioms by a finite number of logical processes. In my opinion, the concept of the continuum is strictly logically tenable in this sense only. It seems to me, indeed, that this corresponds best also to what experience and intuition tell us. The concept of the continuum or even that of the system of all functions exists, then, in exactly the same sense as the system of integral, rational numbers, for example, or as Cantor's higher classes of numbers and cardinal numbers. For I am convinced that the existence of the latter, just as that of the continuum, can be proved in the sense I have described; unlike the system of *all* cardinal numbers or of *all* Cantor's alephs, for which, as may be shown, a system of axioms, compatible in my sense, cannot be set up. Either of these systems is, therefore, according to my terminology, mathematically non-existent.

* *Jahresbericht der Deutschen Mathematiker-Vereinigung* vol. 8 (1900), p. 180.

10. DETERMINATION OF THE SOLVABILITY OF A DIOPHANTINE EQUATION

Given a diophantine equation with any number of unknown quantities and with rational integral numerical coefficients: *To devise a process according to which it can be determined by a finite number of operations whether the equation is solvable in rational integers.*

SELECTED REFERENCES

Biographical

- Constance Reid, *Hilbert*, (Springer-Verlag, Berlin) 1970.
- Andrew Hodges, *Alan Turing: The Enigma*, (Simon and Schuster, New York) 1983.
- John W. Dawson, Jr., Kurt Gödel in sharper focus, *The Mathematical Intelligencer* **6**, no. 4 (1984), 9–17. [A full scale biography of Gödel by Dawson is in preparation].

Assorted Non-Technical References

- Keith Devlin, *Mathematics: The New Golden Age*, (Penguin Books, London) 1983. [See especially Chs. 2, 6, 8, and 11]
- Douglas Hofstadter, *Gödel, Escher, Bach: An Eternal Golden Braid*, (Basic Books, New York) 1979.
- Roger Penrose, *The Emperor's New Mind*, (Oxford University Press, Oxford) 1989.
- Raymond M. Smullyan, *Forever Undecided: A Puzzle Guide to Gödel*, (Alfred A. Knopf, New York) 1987.

Assorted Technical References and Source Collections

- Martin Davis (ed.), *The Undecidable: Basic Papers on Undecidable Propositions, Unsolv-able Problems and Computable Functions* (Raven Press, Hewlett, N. Y.) 1965.
- Jean van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic*, (Harvard University Press, Cambridge) 1967.
- Jon Barwise (ed.), *Handbook of Mathematical Logic*, (North-Holland Pub. Co., Amsterdam) 1977.
- Kurt Gödel, *Collected Works, Vol. I. Publications 1929-1936*, and *Vol. II. Publications 1938-1974* (S. Feferman, J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay and J. Van Heijenoort, eds., Oxford University Press, New York) 1986/1990.

Hilbert's Problems

- David Hilbert, Mathematische Probleme. Vortrag, gehalten auf dem internationalen Mathematiker-Kongress zu Paris 1900, *Göttinger Nachrichten* 1900, 253–297, reprinted in *Archiv der Mathematik und Physik*, 3d ser., **1** (1901), 44–63 and 213–237.

_____ Mathematical Problems. Lecture delivered before the International Congress of Mathematicians at Paris in 1900, English translation of Hilbert (1900) by Mary W. Newson, *Bull. Amer. Math. Soc.* **8** (1902), 437–479; reprinted in Browder (1976), 1–34.

Felix E. Browder (ed.), *Mathematical Developments Arising from Hilbert Problems*, Proc. Symposia Pure Maths. **28**, (American Math. Soc., Providence, R. I.) 1976.

Hilbert's First Problem and Cantor's Continuum Hypothesis

Donald A. Martin, Hilbert's first problem: The continuum hypothesis, in Browder (1976), 81–92.

Kurt Gödel, *The Consistency of the Axiom of Choice and of the Generalized Continuum Hypothesis with the Axioms of Set Theory*, (Princeton University Press, Princeton) 1940; reprinted in Gödel *Collected Works* Vol. II, 33–101. (See also: Introductory note to 1938, 1939, 1939a and 1940, by Robert M. Solovay, op. cit. 1–25).

_____ What is Cantor's continuum problem?, *American Math. Monthly* **54** (1947), 515–525, errata v. 55, 151; reprinted in Gödel *Collected Works*, Vol. II, 176–187.

_____ What is Cantor's continuum problem?, Revised and expanded version of Gödel (1947), in P. Benaceraff and H. Putnam (eds.) *Philosophy of Mathematics: Selected Readings*; reprinted in Gödel *Collected Works* Vol. II, 254–270. (See also: Introductory note to 1947 and 1964 by Gregory H. Moore, op. cit. 154–175).

Paul J. Cohen, *Set Theory and the Continuum Hypothesis*, (Benjamin Pub. Co., New York) 1966.

Solomon Feferman, Some applications of the notions of forcing and generic sets, *Fundamenta Mathematicae* **56** (1965), 325–345.

_____ Infinity in Mathematics: is Cantor necessary?, in *L'infinito nella scienza (Infinity in Science)*, 151–209, (Istituto della Enciclopedia Italiana, Rome) 1987.

Hilbert's Second Problem, Hilbert's Consistency Program and the Incompleteness Theorems

Georg Kreisel, What have we learned from Hilbert's second problem?, in Browder (1976), 93–130.

David Hilbert and Paul Bernays, *Grundlagen der Mathematik*, Vol. II, (Springer-Verlag, Berlin) 1939.

Kurt Gödel, Über formal unentscheidbare Sätze der Principia Mathematica und verwandter system I, *Monatshefte für Math. und Physik* **38** (1931), 173–198; English translation as, On formally undecidable propositions of *Principia Mathematica* and related systems I, in van Heijenoort (1967), 590–616; reprinted in the original and in translation in Gödel *Collected Works* vol. I, 144–195. (See also: Introductory note to *1930b*, *1931* and *1932b*, by Stephen C. Kleene, op. cit. 126–141.)

Craig Smorynski, The incompleteness theorems, in Barwise (1977), 821–865.

Raymond M. Smullyan, *Gödel's Incompleteness Theorems*, (Oxford University Press, New York) 1992.

Jeff Paris and Leo Harrington, A mathematical incompleteness in Peano Arithmetic, in Barwise (1977), 1133–1142.

Harvey Friedman, Necessary uses of abstract set-theory in finite mathematics, in *Advances in Mathematics* **60** (1986), 92–122.

[See also references under “*Overcoming Incompleteness*” and “*Proof Theory*” below.]

Hilbert's Tenth Problem on Diophantine Equations

Martin Davis, Yuri Matiyasevich and Julia Robinson, Hilbert's tenth problem. Diophantine equations: positive aspects of a negative solution, in Browder (1976), 323–378.

Yuri Matiyasevich, *Hilbert's 10th Problem*, (M.I.T. Press, Cambridge) 1993.

Algorithms and Decision Problems

David Harel, *Algorithmics. The Spirit of Computing*, (Addison-Wesley Pub. Co., Reading MA), 1987.

Nigel Cutland, *Computability. An Introduction to Recursive Function Theory*, (Cambridge University Press, Cambridge) 1980.

Michael O. Rabin, Decidable theories, in Barwise (1977), 595–629.

Martin Davis, Unsolvability problems, in Barwise (1977), 567–594.

Logics of programs

Patrick Cousot, Methods and logics for proving programs, in *Handbook of Theoretical Computer Science, Vol. B* (J. van Leeuwen, ed., Elsevier Science Publishers, Amsterdam) 1990, 843–993.

Raymond Turner, *Constructive Foundations for Functional Languages*, (McGraw-Hill Book Company (UK) Ltd., Maidenhead), 1991.

Solomon Feferman, Logics for termination and correctness of functional programs, in *Logic from Computer Science* MSRI Pubs. **21**, (Y. Moschovakis, ed., Springer-Verlag, New York) 1992, 95–127.

Overcoming Incompleteness

Alan Turing, Systems of logic based on ordinals, *Proc. London Math. Soc.* (2) **45**, (1939), 161–228; reprinted in Davis (1965), 155–222.

Solomon Feferman, Transfinite recursive progressions of axiomatic theories, *J. Symbolic Logic* **27** (1962), 259–316.

_____ Turing in the land of $O(z)$, in *The Universal Turing Machine. A Half-century Survey*, (R. Herken, ed., Oxford University Press, Oxford) 1988, 113–147.

_____ Reflecting on incompleteness, *J. Symbolic Logic* **56** (1991), 1–49.

Proof Theory

Gaisi Takeuti, *Proof Theory*, 2nd edition, (North-Holland Pub. Co., Amsterdam) 1987.

Solomon Feferman, Proof theory: a personal report, Appendix to Takeuti (1987), 447–485.

_____ Hilbert’s program relativized: Proof-theoretical and foundational reductions, *J. Symbolic Logic* **53** (1988), 364–384.

_____ What rests on what? The proof-theoretic analysis of mathematics, in *Philosophy of Mathematics, Part I*, Proc. 15th International Wittgenstein Symposium (Verlag Hölder-Pichler-Tempsky, Vienna), 1993, 147–171.

_____ Why a little bit goes a long way. Logical foundations of scientifically applicable mathematics, in *PSA 1992, Vol. II*, (Philos. of Science Assoc., East Lansing, MI) 1993, 442–455.