Cambridge University Press
978-0-521-76144-4 - Kurt Godel and the Foundations of Mathematics: Horizons of Truth
Edited by Matthias Baaz, Christos H. Papadimitriou, Hilary W. Putnam, Dana S. Scott and Charles L. Harper
Excerpt
More information

PART ONE

Historical Context: Gödel's Contributions and Accomplishments

Gödel's Historical, Philosophical, and Scientific Work

CHAPTER 1

The Impact of Gödel's Incompleteness Theorems on Mathematics

Angus Macintyre

The Incompleteness Theorems are best known for interpretations put on them beyond – sometimes far beyond – mathematics, as in the Turing Test, or Gödel's belated claims in the Gibbs Lecture, or in Penrose's more recent work. In this chapter, the emphasis is much narrower, in contrast to the title of the 2006 Gödel Centenary Conference in Vienna – "Horizons of Truth: Logics, Foundations of Mathematics, and the Quest for Understanding the Nature of Knowledge" – that preceded the development of this volume. What is discussed here is almost exclusively the impact on pure mathematics.

That the 1931 paper had a broad impact on popular culture is clear. In contrast, the impact on mathematics beyond mathematical logic has been so restricted that it is feasible to survey the areas of mathematics in which ideas coming from Gödel have some relevance. My original purpose in my presentation at the conference was simply to give such a survey, with a view to increasing resistance to the cult of impotence that persists in the literature around Gödel. After the Vienna meeting, Kreisel persuaded me to write an appendix providing some justification for claims I had made concerning formalizing in First-Order Peano Arithmetic (PA) Wiles' proof of Fermat's Last Theorem. Our discussions on this have, in turn, affected Kreisel's (2008) paper, which provides indispensable proof-theoretic background for the appendix.

The Incompleteness Theorems and their proofs are strikingly original mathematics, with something of the charm of Cantor's first work in set theory. The original presentation is free of any distracting reference to earlier mathematics. Unique factorization is used (not proved), but what is used of the Chinese Remainder Theorem is proved. In fact, the paper would have lost little of its drama if the definability of exponentiation from addition and multiplication had been missed. However, this has always seemed to me the prettiest part of the paper, and nearly forty years passed before this definition was refined in the negative solution of Hilbert's 10th problem.

The grand event in Vienna celebrated many aspects of Gödel's work, but I find it hard to imagine that such a popular event would have happened if Gödel had not proved the Incompleteness Theorems of 1931 or had missed such a snappy proof. This is said not to play down the drama and importance of his and Cohen's work on the continuum hypothesis (or Gödel's work in cosmology). For popular scientific culture, the set-theoretic independence, however troubling or challenging, is less dramatic than the general incompleteness phenomenon. In particular, it is hard to imagine the set-theoretic independence alone giving rise to speculations about creativity. (To my knowledge, the independence results for non-Euclidean geometry did not give rise to such speculations, though they did lead to slogans about breakdown of intuitions.) The classic set-theoretic independence results initiated a great, and still ongoing, creative effort, leading to profound mathematics whose future development we can scarcely guess. On the other hand, the technique of diagonalization, so much admired in popular accounts, has had one infusion of new ideas, with the priority methods that first appeared fifty years ago, but it has now gone rather a long time without startling developments.

1.1 His Contemporaries in Logic

I think that one can get a proper perspective on Gödel only by considering the work of his exact contemporaries in logic because there is certainly a case to be made that others had ideas and results that were to prove more fertile than Gödel's. The relevant points are as follows:

- Presburger's work on the ordered abelian group $\mathbb{Z}$, though neither as original nor as difficult as Gödel's work, is in no way dated. Here we have completeness of an intelligible set of axioms, a useful quantifier elimination, and decidability. The result underlies much of p-adic model theory and model-theoretic motivic integration (Denef and Loeser, 2001). There is extensive literature on the computer science side. Lately, Scowcroft (2006) has refined the definability results to give connections to linear geometry and the work of Weyl.
- Tarski's work on real-closed fields (again, a completeness theorem) has inspired the modern theory of o-minimality (see later), which has brought together model theory and real analytic geometry and led to applications in Lie theory. In contrast, Tarski's work on truth, closer to Gödel's work, has had little effect on mathematics beyond logic.
- Skolem had a broad mathematical range and made diverse, suggestive contributions to logic. In the hands of an imaginative mathematician, the Löwenheim-Skolem theorem can be very powerful. In addition, Skolem's work on the ultrapower construction was of exceptional originality and the origin of many fundamental results in pure model theory and the model theory of arithmetic and set theory.
- Herbrand, before Gödel, and Gentzen, after Gödel, did the first (and still among the best) combinatorial work in the transformation of proofs. At this time, Gödel was not doing combinatorial proof theory but rather dramatic unprovability theory. The ideas of Herbrand and Gentzen have been extensively developed in theoretical computer science. On the other hand, functional interpretations, to which Gödel made major contributions, have been very successful in the unwinding of proofs. Very elementary considerations around growth rates of Herbrand terms, brought out by Kreisel, have proved very efficient and memorable in unwinding bounds for the number of solutions in the classical finiteness theorems (Luckhardt, 1996).
- Herbrand and Skolem made important contributions to the thriving number theory of their time, Herbrand to ramification theory and cohomology and Skolem to p-adic

analytic proofs of finiteness theorems for Diophantine equations. Gödel is known to have attended advanced lectures on class field theory but used in his work no more than the Chinese Remainder Theorem, the most primitive of all local-global principles. Julia Robinson (1999) would need much more, around modern formulations of quadratic reciprocity, in her definition of the integers in the rationals. Of course, Gödel's use of the Chinese Remainder Theorem, to code recursions, remained a central idea in the long evolution of his ideas toward the solution of Hilbert's 10th Problem.

- Ramsey's Theorem, proved in connection with what is now regarded as a very peripheral decision problem, was the starting point for combinatorial work that has flourished for many years and is currently very rewarding in connection with harmonic analysis. From the mid-1950s onward, the technology of indiscernibles has been involved in many of the deepest proofs in mathematical logic. That a recursion-theoretic analysis of Ramsey theory could be rewarding was not seen until the late 1960s (Specker, 1971). Curiously, this analysis seems to have had no influence on the the discovery of "Ramsey Incompleteness."

1.2 The Mathematical Evolution of the Ideas

Despite the awe with which the results are still described, the basic ideas are easy and were quickly adapted to obtain a number of striking consequences:

- Even before Gödel, work was under way on recursive definitions. Gödel provided the central technique for negative results, but Turing gave the subject general interest by providing a rigorous notion of machine and computation and by showing the equivalence of his notion of computability to the rather less natural notions based on formal recursion or lambda recursion. Again, Gödel contributed the main technique for proving noncomputability, from which Church readily proved undecidability of the decision problem for first-order predicate calculus. The arithmetization technique was demystified, leading to such notable results as Turing's on universal machines. Somewhat later, one looked at the finer structure of recursively enumerable sets, developed relative computability, and posed the influential Post problem, whose solution led to the introduction of new techniques of diagonalization.
- Rather more slowly, one moved toward the fine structure of definitions in arithmetic. Here Gödel's coding of recursions by the Chinese Remainder Theorem was central to the repertoire, while being recognized as too weak on its own to yield undecidability of Hilbert's 10th problem. At the level of set theory, Tarski adapted Gödel's diagonal argument to get the classic results on undefinability of truth.
- After a while, one had the very polished treatment of undecidable theories in Tarski (1968). The landmark problems, the word problem for groups and Hilbert's 10th, posed well before Gödel, took longer to solve. In the end, both negative solutions revealed some unforeseen structure.

  In the case of Hilbert's 10th problem, there was certainly no reason to think it might be decidable. At the time of Gödel's first work, the only large-scale decidability was in the area of quadratic forms (Siegel), and it took another thirty-five years before Baker's flexible method gave a variety of effective estimates, including a decision procedure for

elliptic curves over $\mathbb{Q}$. It still seems to me extremely optimistic to have expected that all recursively enumerable sets should be Diophantine, and miraculous that it turned out so. That said, it remains disappointing that so little has followed from the result. Putnam's neat observation that every Diophantine set is the set of positive values of a polynomial gives the startling result that the set of primes is the set of positive values of a polynomial (and one has been able to spell out such polynomials). But I think it fair to say that the example remains merely striking, and there is no hint of any underlying theory of such representations (e.g., in geometric terms). Again, one would have hoped that the hardcore theory of recursively enumerable sets might by now have told us something quite new and suggestive about Diophantine sets, but this has not happened, and it is notable that no known method for the undecidability of Hilbert's 10th problem gives any information about the corresponding problem for the field of rationals. This makes it all the more encouraging that Poonen (2009) has recently made the first serious progress in that area for many years by showing that the integers have a universal-existential definition in the rational field (improving the old result of Julia Robinson).
- Sarnak's (2006) recent Rademacher lectures on solving equations in primes have an interesting point of contact with the Putnam trick mentioned earlier. It turns out that to get the right formulation of the rather deep theorems of these lectures, he has to allow primes to be positive or negative. Forcing them to be positive would bring the pitfalls of undecidability too close.

1.3 How the Number Theorists React to the Gödel Phenomenon

There has certainly been no cult of impotence. In the last thirty-five years, number theory has made sensational progress, and the Gödel phenomenon has surely seemed irrelevant. The number theorists are keenly aware of issues of effectivity (Hindry and Silverman, 2000) and, indeed, of relative effectivity. Moreover, each of the classical finiteness theorems is currently lacking expected effective information. However, there is not the slightest shred of evidence of some deep-rooted ineffectivity. Some key points are the following:

- The dimensions (of varieties) where unsolvability of Hilbert's 10th problem sets in are far beyond those where current, theory-driven research in number theory takes place or concern varieties with no discernible structure. Lang says somewhere that no undecidability is to be expected for abelian varieties. The sense (and plausibility) of this is pretty clear, even if one can contrive undecidability results about abelian varieties, for example, by considering period lattices with nonrecursive generators.
- The equations whose unsolvability is equivalent (after Gödel decoding) to consistency statements have no visible structure and thus no special interest. Recall Gauss's dismissive remarks about the ad hoc nature of the Fermat equation. What we now appreciate in Wiles' great work is not that a specific family of equations has no nontrivial solution but the link to modularity of elliptic curves over $\mathbb{Q}$ and the profound structure underlying modularity. This has been followed, in an amazingly short time, by fundamental results about Galois representations and the Langlands program (i.e., Sato-Tate conjecture, Serre modularity conjecture).

- The problem of deciding whether curves over $\mathbb{Q}$ have integer points is not yet known to be decidable, but there is a bodyguard of theory, quite independent of logical considerations and by now heavily supported by numerical evidence, that implies that undecidability is not to be expected. Indeed, it is to be dreaded (and this is Shafarevich's "gloomy joke") because of its implications for high theory. There are several proofs of Siegel's theorem, each with at least one region of current ineffectivity (Roth's theorem or the Mordell-Weil theorem). In the case of the Mordell-Weil theorem (needed for abelian varieties in the standard proof of Siegel's theorem), Manin (see Hindry and Silverman, 2000, 463) has shown that two dominant analytic conjectures imply effective bounds for generators of the Mordell-Weil group and thereby remove one of the two regions of ineffectivity. The first conjecture concerns the L-series of abelian varieties A over $\mathbb{Q}$ and says, first, that the series has an analytic continuation to the whole complex plane and, second, that it satisfies a standard functional equation involving the conductor N. By modularity considerations, the conjecture is known to be true for elliptic curves over Q. In general, if one assumes the analytic continuation, one then has the second conjecture, of Birch and Swinnerton-Dyer, relating the order of vanishing at $s = 1$ to the rank of the Mordell-Weil group and giving the leading term of the expansion around $s = 1$. Manin's proof involves showing how these two conjectures give a bound for the regulator and thereby for generators of the Mordell-Weil group. Thus dominant analytic conjectures reduce the ineffectivity to that coming from Roth's theorem.

  Kreisel has made me aware of the curious impact on Weil of issues around incompleteness and undecidability. In 1929, before Gödel, as described in Weil (1980, 526), he confronted issues of effectivity for elliptic curves and made little progress. Much later, in the notes (Weil, 1980) provided for his *Collected Works*, Weil speculates about inherent ineffectivity for the decision problem for curves and expresses the hope that progress in mathematical logic will bring these problems within reach. It seems that he was unaware of how devastating such ineffectivity would be for the edifice of conjectures in modern arithmetic.
- Serre gives another perspective (1989, 99, remark 3). He derives Siegel's theorem from Mordell-Weil and an approximation theorem on abelian varieties. The latter is naturally deduced from Mordell-Weil and Roth. Serre updates some observations of Weil and remarks that from effective Mordell-Weil, one can dispense with Roth in favor of a known abelian analogue of Baker's lower bounds for linear forms in logarithms. Thus it seems that the two analytic conjectures suffice to get the decidability of the decision problem for curves.
- A propos of the decision problem for curves, the natural logical parameters of the problem, such as number of variables and degree of polynomials involved, obscure the geometrical notions (notably genus but also projective nonsingularity and Jacobian varieties) that have proved indispensable to much research since Siegel's work of 1929. (I say "much" because some methods of Diophantine approximation are not naturally described as geometric.) If one's formalism obscures key ideas of the subject, one can hardly expect logic alone to contribute much.
- There is no hint of incompleteness linked to classical analytic number theory. It was, of course, natural to raise the issue of such incompleteness at the time when the prime number theorem did not yet have an "elementary" proof. However, if one is familiar with the standard analytic proofs, using, for example, analytic continuation, Kreisel's

sketches (1951, 1952a, 1952b) from the 1950s are totally convincing. Note, too, that *sketch* here is anything but snide. Sketches are the appropriate medium for this kind of mathematics. The whole enterprise is futile for those who do not know the analytic proofs, whereas once one knows them, it typically requires little effort to reorganize the proof into a weaker system. I recommend Kreisel's brief account (2008).

- There is no hint of incompleteness (in the sense that one may need set theory) in the work of Deligne on the Weil conjectures or Faltings on the Mordell conjecture. At the Vienna meeting, I expressed confidence that the Wiles proof of Fermat's last theorem fits into PA. If the discussions on FOM (the "Foundations of Mathematics" moderated e-mail list)[1] are to be taken as representative of opinion among logicians, my claim remains controversial. In the appendix to this chapter, I attempt to open some informed discussion on the matter.
- The appendix is mainly concerned with the role of set theory in the formulation and/or proof of modularity conjectures such as the restricted one used by Wiles in his proof of Fermat's last theorem. One goal is to show that the conjectures themselves are Π_1^0. For this, it is crucial to use the input of Weil to the modularity conjecture, giving precise formulations in terms of *conductors* (see Weil, 1999). Beyond this issue of formulation, one has to survey proofs of modularity and show that each component can be arithmetized. This is not at all a trivial enterprise, even if one aims only at an overview of the proof.

  We are fortunate to have the guide (Cornell et al., 1997) to the original proof. This reveals a number of distinct regions of the proof, almost all hitherto unvisited by proof theorists. These include Galois representations, distribution of primes, étale cohomology, modular forms and functions, elliptic curves and their L-series, the analysis behind Langlands's theory (e.g., the proof of Langlands-Tunnell), and advanced commutative algebra of noetherian rings. Even prior to all this, there is basic p-adic and adelic analysis. Some but certainly not all this territory is known to some model theorists who have learned something about definitions and uniformities therein. Moreover, this group of people typically know a fair amount of complex analysis and geometry, and some are accustomed to keeping track of the complexity of inductive arguments.

  It is not clear to me what further useful proof-theoretic tool kit one has for the preceding enterprise. I claim that there is no need for a proof of Wiles' theorem (or beyond, e.g., to the modularity theorem) of the use of strong second-order axioms with existential quantifiers involved. The claim leaves open the question of the utility of second-order *formalism* and of metatheorems (such as conservation results) formulated in functional terms. It seems to me likely that these will be useful a little further down the road, if enough logicians begin to pay attention to the structure of large-scale proofs of the type of Wiles (Colin McLarty has recently circulated an interesting paper in this area.) In addition, my own experience suggests that such a global structural analysis will bring to light many local issues of model-theoretic and proof-theoretic interest. For my own purposes, in getting the appendix written, I have relied (perhaps too exclusively) on rules of thumb (and tricks of the trade) from model-theoretic algebra and the model theory of PA and related systems. For a compact account of the basic proof-theoretic

[1] See http://www.cs.nyu.edu/mailman/listinfo/fom/.

issues, informed by over sixty years of case studies, one should consult Kreisel (2008). Thus I choose to give arithmetic interpretations of, for example, those parts of real, complex, or p-adic analysis (or topology) used in the proof.

- We are dealing with very specific functions, such as zeta functions, L-series, and modular forms, that are directly connected to arithmetic. There is little difficulty in developing the basics of complex analysis for these functions, on an arithmetic basis, sufficient for classical arithmetical applications (e.g., Dirichlet's theorem or the prime number theorem) (Kreisel, 1952b), and for me, nothing would be gained by working in a second-order formalism in a very weak system. At best, such systems codify elementary arguments of general applicability. About more sophisticated analysis, as involved in Langlands's theory, I see no difficulty in principle, but one must be aware that this analysis is much more delicate than the early analysis of classical analytic number theory, and so one has simply got to go out and do it in arithmetic style. It seems to me very unlikely that one will encounter in this part of the proof any inductions that somehow involve either unlimitedly many first-order quantifier changes or second-order quantifiers, the hallmarks of incompleteness.
- Given the complexity of the inductive arguments used in the foundations of étale cohomology, one would have to be more careful in that area, *if one were using the theory for general schemes*, to get by with inductions of bounded complexity. Fairly extensive experience on the metamathematics of Frobenius leaves me confident of avoiding trouble in any application (via the Weil conjectures) to varieties over finite fields (and in any case, one does not need very general varieties in Wiles' proof).

  Étale cohomology of schemes can be used to prove the basic facts of the coefficients of zeta functions of abelian varieties over finite fields, but there are more elementary ways to get those results in PA. However, étale cohomology provides a huge variety of Galois representations, and so any arithmetization of the latter theory is likely to need a corresponding arithmetization of the former. Though I stress that such an arithmetization is a distinctly nontrivial matter, I hasten to dissociate the problems therein from the discussions on the FOM online forum suggesting that *higher set theory* (beyond ZFC) might be needed for Wiles' proof because of a perceived need for higher set theory in the foundations of étale cohomology.

  Fortunately, there is an FOM entry by Timothy Chow (2007) quoting an anonymous number theorist on the daftness of such claims about higher set theory and étale cohomology. I urge the reader to seek out these remarks, which I expect to be very useful in the enterprise of detaching the innocent from this cult of impotence.
- Again, existing work on model theory of various complete noetherian rings should give confidence that one will find the right setting for the arguments in the deformation-theoretic part of Wiles' proof. This point and ones related to the last few sections are developed in the appendix.
- I want to stress that it is by no means clear to me right now how much induction is needed for a transcription of the mathematics of Cornell et al. (1997) to get a proof of Wiles' theorem in PA. In particular, I think there is little evidence that bounded arithmetic plus the totality of exponentiation would suffice. It is obvious that the latter proves some of the basics around distribution of primes, but one has so little experience of more modern mathematics in the system that one should be cautious.

- In the appendix, I try to open a road to a proof of Fermat's last theorem that does not use set theory at all. Indeed, I seek to show that the modularity conjecture is itself arithmetical and that its proof needs no set theory. It is not claimed, though I see no reason to doubt it, that PA can reproduce the general theory of étale cohomology of varieties over number fields and the resulting Galois representations. That general theory is not needed for Wiles' proof.

1.4 Group Theory

The evolution here is of particular interest. After the original proof by Novikov of the unsolvability of the word problem, one had (in the Adjan-Rabin theorem) confirmation that most classical decision problems about finitely presented groups are unsolvable. Positive results were very rare in those days, mainly around the intricate small cancellation theory, which had a geometrical interpretation (Lyndon and Schupp, 2004). Rather later, in the 1980s, the important isomorphism problem for finitely generated nilpotent groups was shown to be decidable by the essential use of the geometric and number-theoretic ideas of Siegel (Grunewald and Segal, 1980a, 1980b). But for a while, the emphasis was mainly on undecidability, and the tools from Gödel and his followers were appropriate. In 1960, Higman (1961) gave a delightful shift of emphasis. He changed the logical part of undecidability into something positive and versatile. In effect, he took the neat observation from the 1940s on finite axiomatizability by extra predicates and put it into group theory, by essential use of amalgamated products and other basics of the subject to prove that the finitely generated subgroups of finitely presented groups are exactly the finitely generated, recursively presented groups. This became a powerful tool for constructing pathological groups, and it yielded the existence of universal finitely presented groups (in analogy with universal Turing machines). It readily yielded characterizations of groups with solvable word problems, in terms free of recursion theory. Yet Higman's work was virtually the end of the importance of Gödel's ideas in combinatorial group theory. Serre's work on trees gave a new perspective on amalgamation and HNN extensions. By the 1980s, the deep ideas of Gromov around metrical and hyperbolic aspects of group theory had begun to penetrate the subject. These relate to the geometry of small cancellation theory but are much more conceptual. That the class of hyperbolic groups is quite restricted may be a defect from the point of view of logic, but it is not at all from the standpoint of those who connect geometry, topology, and group theory. For masterly accounts of how the subject looks now, one should consult (Bridson, 2002; Bridson and Haefliger, 1999) and Bridson's recent ICM lecture (Bridson, 2007). Machines have not quite disappeared from the scene, inasmuch as finite automata are basic.

My point here is that the ideas from Gödel were very important in establishing the limitations of algorithmic methods for general finitely presented groups and got a beautiful final version in Higman's work but are now of little relevance. They delineate the boundaries beyond which there are monsters (the monster itself is a highly structured entity, connected to many currently central parts of mathematics). But one knows now that the monsters are far away, and one has every confidence in the cordon sanitaire that keeps them away from current research. There is no geometry visible in the pathology.

In contrast, classically posed problems, such as that on the elementary theory of free groups, are currently being analyzed positively through the essential use of geometric or arboreal, as opposed to finite combinatorial, ideas (Sela, 2002).

1.5 Geometry and Dynamical Systems

There are two intriguing recent developments that appear to link a part of mainstream geometry to classical Gödelian ideology. Both involve rather difficult mathematics and give formulations in terms of classical recursion theory.

The first is the work of Nabutovsky and Weinberger (2000) (and see also the related logical results in (Soare, 2004), which interprets surprisingly much of the working of a Turing machine in the variational theory of Riemannian metrics on compact manifolds of dimension at least five. Here, as Soare remarks in his useful account, one is contributing to a very natural geometric classification problem using refined notions of recursion theory. It is worth noting that no fuss is made about unprovability, although one could surely torture the proof into an incompleteness proof. There is, however, a remarkable example of a statement in differential geometry, with no reference to recursion theory, whose only known proof uses recursion theory (Soare, 2004). Soare chooses to stress the computability aspect, whereas it seems to me here that it is really the delicate nature of the recursive construction that is important. (It is notable that an early version of the work used the Sacks density theorem.) It is surely natural to explore the possibilities of the techniques in lower dimensions.

The second is the work of Braverman and Yampolsky (2006), showing, within the ideology of computable analysis, that there are quadratic Julia sets that are noncomputable relative to an oracle for the coefficients of the underlying quadratic polynomial. (There are quadratic polynomials for which one has computability.) It would have been startling had the problem turned out to be computable, but the proof, like most undecidability results for mainstream mathematical structures, is intricate and demands detailed knowledge of central work in the specific subject matter. As of now, there is less evidence of coding as refined as one sees in Nabutovsky and Weinberger (2000). In contrast to the latter, the emphasis here is on noncomputability. In neither case do we yet have a shift of emphasis as decisive as that for Higman's theorem, identifying a natural notion in recursion-theoretic terms.

1.6 Set Theory

Here there is no doubt that Gödel's influence remains, and probably will remain, major. The mathematical facts are quite evident. He gave a beautifully focused development of the inner model L and, by an argument certainly not remote from that for Löwenheim-Skolem, established that GCH and AC hold in L whether they hold in V. As with the work on incompleteness, one is struck by the freshness of the work. ZF settles for L, with ease, the issue of AC and CH (issues not to be clarified for ZF for another twenty-five years). Yet Gödel missed some things about L that are important and were not far away. As usual, I collect a few points: