ASPECTS OF GÖDEL INCOMPLETENESS

Harvey M. Friedman
Distinguished University Professor
of Mathematics, Philosophy, Computer Science Emeritus
Ohio State University
Columbus, Ohio
https://u.osu.edu/friedman.8/foundational-adventures/downloadable-manuscripts/
https://www.youtube.com/channel/UCdRdeExwKiWndBl4YOxBTEQ

International Workshop on Gödel's Incompleteness Theorems
August 16-20, 2021 ONLINE
delivered August 16, 2021
THIS IS THE THIRD VERSION AFTER THE TALK WAS DELIVERED
8/18/21, 8/19/21, 9/8/21

Abstract. We begin with a discussion of various forms of G1 put into the following general form: If a first theory satisfies one or more adequacy conditions then it has one or more wildness properties. We give a list of familiar adequacy conditions and wildness properties. We propose an investigation into the myriad forms of G1 in this framework. Some such forms of G1 will be well known, some well known to be false, and some yet to be investigated. We expect many will suggest further investigations. We then discuss various new "no interpretation" forms of G2. These are exquisitely simple formulations of G2 in the following sense. The proofs of them from G2 are entirely straightforward applications of G2 and Gödel Completeness. The derivation of G2 is also straightforward and does not rely on any of the ingredients in the known proofs of G2. This "no interpretation G2" evolved from a long series of struggles with G2 which are compiled in section 3. We also give corresponding surprisingly simple characterizations of the consistency statement Con(T) for finitely axiomatized T. We then discuss G2/1-con which is G2 with the strengthened hypothesis of 1-consistency and the weakened conclusion of the unprovability of 1-consistency. We give the long since known, if not well known, proof of G2/1-con which is much simpler than the proof of G2. It is best proved by what we call "transparent diagonalization" which is the kind of informative diagonalization used by Cantor in his proof that given an infinite sequence of subsets of N,

some subset of N is missing. A by product of this proof is the association of a crucially important set of objects to T that gets properly expanded by T + 1-Con(T) - namely the provably recursive functions. Since so much of the philosophical and foundational import of G2 is already present with G2/1-con, we propose that G2/1-con be revisited with the same deep intensity as has G2. We call for a proof of G2 by transparent diagonalization. We then discuss some proofs of G2. First a clean formulation of the more or less original and most common proof, seeking to minimize the mystery. This usual construction is regarded as mysterious since it is used in the Gödel Rosser theorem, and there we have no reasonable understanding of the nature of Rosser sentences under natural syntactic numberings. We then give a non mysterious proof of G2 from (a straightforwardly explicit form of) the fixed point theorem of Hartley Rogers, which is easily derived from the earlier Kleene Second Recursion Theorem (and vice versa). However, the proofs of these Kleene, Rogers theorems can also be regarded as mysterious. We propose a demystification of Kleene, Rogers using the RASP model of computation. This approach suggests that there might be a kind of hidden sequential structure behind Kleene, Rogers and G2. We then give another proof of G2 using (explicitly) remarkable sets. This really is another way of very cleanly pushing the mystery behind proving G2 to a particularly simple recursion theoretic construction. The idea here is that the notion of (explicitly) remarkable might suggest a way of incorporating G2 into category theory and/or type theory, something which has long been sought. Finally, we make a brief presentation of the state of the art in Tangible Incompleteness with the most recent examples of implicitly Π^0_1 sentences provably equivalent to the consistency of certain large cardinal hypotheses (ongoing research).

1. G1. Gödel's first incompleteness theorem.
2. G2. Gödel's second incompleteness theorem.
3. Previous evolving work on G2.
4. Proof of G2/1-con by transparent diagonalization.
5. Towards Demystifying G2.
6. Tangible Incompleteness.

1. G1. GÖDEL'S FIRST INCOMPLETENESS THEOREM

I recently found the article [Ch20]. There is some overlap between what I say here about G1 and this valuable well written article.

By a theory we will always mean a theory T in the usual PC(=), (predicate calculus with equality), which comes with a designated language (of constant, relation, and function symbols).

The most common way to formulate G1 is to assert that any theory T with an "adequacy condition" has a "wildness property". There are several important kinds of adequacy conditions and wildness properties.

Common adequacy conditions on a theory T (with multiple choices):
a. T is consistent.
b. T is (finitely axiomatized, axiomatized by finitely many axiom schemes, recursively axiomatized).
c. T interprets a given theory K, (finitely axiomatized, axiomatized by finitely many axiom schemes, recursively axiomatized).
d. T is consistent with an interpretation of a given theory K, in the same language as T, (finitely axiomatized, axiomatized by finitely many axiom schemes, recursively axiomatized).
e. The language of T is or extends a given language, and T proves a certain theory K (finitely axiomatized, axiomatized by finitely many axiom schemes, recursively axiomatized).

Common wildness properties of a theory T (with multiple choices):
A. T is incomplete in the sense that there is a sentence in the language of T that is neither provable nor refutable in T.
B. T is essentially incomplete in the sense that no consistent extension of T by finitely many sentences is complete.
C. The set of theorems of (T, any finite extension of T, any recursive extension of T) is (complete r.e., not recursive, not primitive recursive, not elementary recursive, not polytime computable).
D. The set of theorems of T and the set of refutables of T are recursively inseparable.
E. Assuming the language of T is or extends a given language, A-D restricted to sentences in a given sublanguage.

There are likely some other interesting adequacy conditions and wildness properties that should be considered in such a systematic investigation.

TEMPLATE FOR G1. Let T obey a chosen one or more (parts) of a-e. Then T has a chosen one or more (parts) of properties A-E.

SYSTEMATIC G1 INVESTIGATION. Determine relationships between various instances of the Template for G1, including their correctness for various K.

The most elemental form of G1 involving only the most rudimentary of notions, is arguably the following.

PURE G1 (finite). There is a consistent finitely axiomatized theory K such that any consistent finitely axiomatized theory T interpreting K is incomplete.

Robinson's Q is most commonly used for pure G1, as well as its many natural "variants" in the sense of being mutually interpretable with Q. There is no known natural system K for this pure G1 that does not interpret Q.

PURE G1 (schematic). There is a consistent theory K with finitely many axiom schemes, such that any consistent theory T axiomatized with finitely many axiom schemes, interpreting K, is incomplete.

Here there is a natural infinitely axiomatized system R, interpretable in Q, but where Q is not interpretable in R, that we can use. But we would like to say that R is "very recursive". However, as "schematic" as the system R looks, it is not officially given by finitely many axiom schemes. So we need to either expand the notion of axiom scheme to allow R, or we need to modify R to fit into the usual notion of schemes. This should be investigated.

What is missing is insight into the special status of Q and R and perhaps variants of Q and R, for G1.

Furthermore, as we vary the wildness properties we seek for T, how does that affect the choices of K that we can use in the adequacy conditions?

There is also an attractive simplicity investigation here. There are some reasonably natural measures of the complexity of presentations of finitely axiomatized axiom systems in PC(=). E,g., one can count the number of occurrences of symbols other than parentheses and commas, each occurrence of a variable counted as 1. We can seek information on the smallest complexity of a K supporting Pure G1 or other instances of the G1 Template. The language of arithmetic would not be a good choice for this. The language of set theory would be much better, through the system AS of Adjunctive Set Theory, as well as theories of strings. AS,Q are mutually interpretable.

CONJECTURE. Any finitely axiomatized system K usable for pure G1 and variants of G1, of complexity at most that of AS, interprets AS.

The most common languages used to formulate versions of G1 are arithmetic (with and without exponentiation, with and without primitive recursive function symbols, with and without <), set theory with membership, and string theory concatenation. There are some important special classes of formulas, most notably Π^0_1, Σ^0_1, and Σ^0_1 using polynomial equations. Here G1 meets Hilbert's Tenth Problem. See. e.g., [Je16].

I was gratified to hear from Albert Visser during my talk that he just came up with a seemingly interesting new G1 question that fits nicely into my G1 Template. Can we find a K' that does for G1 what K does, or some of what K does, where the set of theorems of K' is recursive? Perhaps this is an indication of the usefulness of the G1 Template for generating new research.

2. G2. GÖDEL'S SECOND INCOMPLETENESS THEOREM

After my talk at the conference, I proved Theorems 2.1-2.5 below, using ideas for related results that I wrote about from 2003 through 2017. I include some references to this earlier work in section 3. We use EFA = exponential function arithmetic, which is well known to be finitely axiomatizable.

THEOREM 2.1. No consistent extension T of PA, in any language, is interpretable in any theorem of T in the language of PA.

NOTE: Here we can take the language of PA to be 0,S,+,• or any extension of 0,S,+,• using any set of primitive recursive function symbols, with the usual axioms for successor, and the usual primitive recursive defining axioms. Also here PA uses induction for only arithmetic formulas as usual, and not induction for formulas involving other symbols.

Proof: Let T prove PA + φ, φ a sentence in L(PA), where T is interpretable in φ. We show that T is inconsistent.

It is well known that $\varphi \rightarrow \mathrm{Con}(\varphi)$ is provable in PA using partial truth definitions and cut elimination. Let T' be a finitely axiomatized fragment of T which proves EFA and $\varphi \rightarrow \mathrm{Con}(\varphi)$.

Therefore T' proves Con(φ). Then T' is interpretable in φ and Con(T') makes sense. From this interpretation, EFA proves Con(φ) → Con(T'). Hence T' proves Con(T'), and therefore by G2, T' ⊆ T is inconsistent. QED

THEOREM 2.2. No consistent extension T of PA[n], in any language, is interpretable in any Σ^0_{n+2} theorem of T.

Proof: Let T prove φ, where T is interpretable in φ and φ is Σ^0_{n+2}. We show that T is inconsistent.

It is known that φ → Con(φ) is provable in PA[n] using partial truth definitions and cut elimination, as proved in [Le83] (also see [Be97], [Be05]). Let T' be a finitely axiomatized fragment of T which proves PA[n] and φ, and therefore Con(φ). Then T' is interpretable in φ and Con(T') makes sense. From this interpretation, EFA proves Con(φ) → Con(T'). Hence T' proves Con(T'), and therefore by G2, T' ⊆ T is inconsistent. QED

THEOREM 2.3. No consistent extension T of PRA, in any language, is interpretable in any Π^0_1 theorem of T in L(PRA).

Proof: Let T prove φ, φ being Π^0_1 in L(PRA), where T is interpretable in φ. We show that T is inconsistent.

It is known that φ → Con(φ) is provable in PRA since we have Herbrand's theorem available in PRA and induction applied to bounded formulas in the primitive recursive function symbols used in φ. Use of Herbrand here involves iteration of the underlying functions. Let T' be a finitely axiomatized fragment of T which proves EFA and φ, and therefore Con(φ). Then T' is interpretable in φ and Con(T') makes sense. From this interpretation, EFA proves Con(φ) → Con(T'). Hence T' proves Con(T'), and therefore by G2, T' ⊆ T is inconsistent. QED

SEFA is super exponential function arithmetic, and EFA is exponential function arithmetic.

THEOREM 2.4. No consistent extension T of SEFA, in any language, is interpretable in any Π^0_1 theorem of T in L(EFA).

Proof: Let T prove φ, φ being Π^0_1 in L(EFA), where T is interpretable in φ. We show that T is inconsistent.

It is known that $\varphi \rightarrow \mathrm{Con}(\varphi)$ is provable in SEFA. To see this, assume φ is refutable, and apply Herbrand's theorem, available in SEFA. This creates indefinite iterations of addition and multiplication and exponentiation, and the associated truth definitions are handled appropriately by SEFA. Let T' be a finitely axiomatized fragment of T which proves EFA and φ, and therefore $\mathrm{Con}(\varphi)$. Then T' is interpretable in φ and Con(T') makes sense. From this interpretation, EFA proves $\mathrm{Con}(\varphi) \rightarrow \mathrm{Con}(T')$. Hence T' proves Con(T'), and therefore by G2, $T' \subseteq T$ is inconsistent. QED

PFA is polynomial function arithmetic, also known as bounded arithmetic.

THEOREM 2.5. No consistent extension T of EFA, in any language, is interpretable in any Π^0_1 theorem of T in L(PFA).

Proof: Let T prove φ, φ being Π^0_1 in L(EFA), where T is interpretable in φ. It suffices to prove that T is inconsistent.

It is known that $\varphi \rightarrow \mathrm{WCon}(\varphi)$ is provable in EFA, where WCon is the weakened form of Con also referred to as cut free consistency. Since we have Herbrand's theorem available in EFA for specific complexity, and we can use it here with indefinite iteration of addition and multiplication, we obtain $\varphi \rightarrow \mathrm{WCon}(\varphi)$ in EFA. Let T' be a finitely axiomatized fragment of T which proves EFA and φ, and therefore $\mathrm{WCon}(\varphi)$. Then T' is interpretable in φ and Con(T') makes sense. From this interpretation, EFA proves $\mathrm{WCon}(\varphi) \rightarrow \mathrm{WCon}(T')$. Hence T' proves WCon(T'), and therefore by G2, $T' \subseteq T$ is inconsistent. NOTE: G2 is well known to hold for WCon. QED

Note that we have derived Theorems 2.1 - 2.5 from G2 applied to the r.e. presented theories extending PA, PA[n], PRA, SEFA, EFA, respectively.

Now we want to obtain equivalence. The PRINCIPAL POINT here is that this equivalence is very transparent and does not even remotely involve any of the ideas used to prove G2.

THEOREM 2.6. Theorems 2.1 - 2.5 are demonstrably equivalent to G2 for r.e. presented theories in any language, where the axioms extend PA, PA[n], PRA, SEFA, EFA, respectively.

Proof: We have already obtained the derivations from G2. We now want to derive G2. Now suppose Theorem 2.1 and let T be a consistent r.e. presented extension of PA, PA[n], PRA, SEFA, EFA, respectively. For G2, let T prove Con(T), where Con(T) is formulated as a Π^0_1 sentence in L(PFA). Now in each of the five cases, T is interpretable in Con(T) with some infrastructure needed to properly use Con(T). EFA easily serves as this infrastructure. So using Theorems 2.1 - 2.5, we see that T is inconsistent, establishing G2 in each case. QED

We now characterize the Con statement for finitely axiomatized theories (single sentences). We first characterize the Con statement up to provable equivalence over PA.

THEOREM 2.7. Let A be any sentence. For all arithmetic B, PA + B interprets A if and only if PA + B proves Con(A). For a given sentence A, Con(A) is the unique arithmetic sentence with this property up to PA provability.

Proof: Let A be a sentence and B be an arithmetic sentence. If PA + B proves Con(A) then obviously PA + B interprets A. Now suppose PA + B interprets A. Let PA[n] + B interpret A. Then EFA proves Con(PA[n] + B) $\rightarrow$ Con(A). Now PA + B proves Con(PA[n] + B). Hence PA + B proves Con(A).

Now let C be an arithmetic sentence such that for all arithmetic sentences B, PA + B interprets A if and only if PA + B proves C. Then for all arithmetic sentences B, PA + B proves C if and only if PA + B proves Con(A). Setting B = C we get PA proves C $\rightarrow$ Con(A), and by setting B = Con(A), we get PA proves Con(A) $\rightarrow$ C. Hence PA proves C $\leftrightarrow$ Con(A). QED

Now we characterize the Con statement up to provable equivalence over PRA[n].

THEOREM 2.8. Let A be any sentence and n $\geq$ 1. For all Σ^0_{n+2} sentences B, PA + B interprets A if and only if PA[n] + B proves Con(A). For a given sentence A, Con(A) is the unique Σ^0_{n+2} sentence with this property up to PA[n] provability.

Proof: Let A be a sentence and B be Σ^0_{n+2}. If PA[n] + B proves Con(A) then obviously PA[n] + B interprets A. Now suppose PA[n] + B interprets A. Then EFA proves Con(PA[n] + B) $\rightarrow$ Con(A). It is known that PA[n] + B proves Con(PA[n] + B) using partial truth definitions and cut elimination, as proved in [Le83] (also see

[Be97], [Be05]). Hence PA[n] + B proves Con(A). Note that PA[n] gets absorbed in B.

Now let C be a Σ^0_{n+2} sentence such that for all arithmetic sentences B, PA + B interprets A if and only if PA + B proves C. Then for all Σ^0_{n+2} sentences B, PA[n] + B proves C if and only if PA + B proves Con(A). Setting B = C we get PA[n] proves C $\rightarrow$ Con(A), and by setting B = Con(A), we get PA[n] proves Con(A) $\rightarrow$ C. Hence PA proves C $\leftrightarrow$ Con(A). QED

Next we characterize the Con statement up to provable equivalence over PRA.

THEOREM 2.9. Let A be any sentence. For all Π^0_1 sentences B in L(PRA), PRA + B interprets A if and only if PRA + B proves Con(A). For a given sentence A, Con(A) is the unique Π^0_1 sentence in L(PRA) with this property up to PRA provability.

Proof: Let A be a sentence and B be Π^0_1 in L(PRA). If PRA + B proves Con(A) then by standard techniques, PRA + B interprets A. Now suppose PRA + B interprets A. Let PRA' + B interpret A, where PRA' is a finite fragment of PRA. Then EFA proves Con(PRA' + B) $\rightarrow$ Con(A). Now PRA + B proves Con(PRA' + B) by standard techniques using that B is Π^0_1 in L(PRA). Hence PRA + B proves Con(A).

Now let C be a Π^0_1 sentence in L(PRA) such that for all Π^0_1 B in L(PRA), PRA + B interprets A if and only if PRA + B proves C. Then for all Π^0_1 B in L(PRA), PRA + B proves C if and only if PRA + B proves Con(A). Setting B = C we get PRA proves C $\rightarrow$ Con(A), and by setting B = Con(A), we get PRA proves Con(A) $\rightarrow$ C. Hence PRA proves C $\leftrightarrow$ Con(A). QED

We next characterize the Con statement up to equivalence over SEFA.

THEOREM 2.10. Let A be any sentence. For all Π^0_1 sentences B in L(EFA), SEFA + B interprets A if and only if SEFA + B proves Con(A). For a given sentence A, Con(A) is the unique Π^0_1 sentence in L(EFA) with this property up to SEFA provability.

Proof: Let A be a sentence and B be a Π^0_1 sentence in L(EFA). If SEFA + B proves Con(A) then by standard techniques, SEFA + B interprets A. Now suppose SEFA + B interprets A. Then EFA proves

Con(PRA' + B) $\rightarrow$ Con(A). Now SEFA + B proves Con(SEFA + B) by standard techniques using that B is Π^0_1 in L(EFA), with Herbrand's theorem, available in SEFA, and the handling of compound terms in L(EFA) within EFA. Hence SEFA + B proves Con(A).

Now let C be a Π^0_1 sentence in L(EFA) such that for all Π^0_1 B in L(EFA), SEFA + B interprets A if and only if SEFA + B proves C. Then for all Π^0_1 B in L(EFA), SEFA + B proves C if and only if SEFA + B proves Con(A). Setting B = C we get SEFA proves C $\rightarrow$ Con(A), and by setting B = Con(A), we get SEFA proves Con(A) $\rightarrow$ C. Hence SEFA proves C $\leftrightarrow$ Con(A). QED

Next, we characterize the WCon statement up to equivalence over EFA.

THEOREM 2.11. Let A be any sentence. For all Π^0_1 sentences B in L(PFA), EFA + B interprets A if and only if EFA + B proves WCon(A). For a given sentence A, WCon(A) is the unique Π^0_1 sentence in L(PFA) with this property up to EFA provability.

Proof: Let A be a sentence and B be a Π^0_1 sentence in L(PFA). If EFA + B proves WCon(A) then by standard techniques, EFA + B interprets A. Now suppose EFA + B interprets A. Then EFA proves WCon(PRA' + B) $\rightarrow$ WCon(A). Now EFA + B proves WCon(EFA + B) by standard techniques using that B is Π^0_1 in L(EFA), with schematic Herbrand's theorem, available in EFA, and the handling of compound terms in L(PFA) within EFA. Hence EFA + B proves WCon(A).

Now let C be a Π^0_1 sentence in L(PFA) such that for all Π^0_1 B in L(PFA), EFA + B interprets A if and only if EFA + B proves C. Then for all Π^0_1 B in L(PFA), EFA + B proves C if and only if EFA + B proves WCon(A). Setting B = C we get EFA proves C $\rightarrow$ WCon(A), and by setting B = WCon(A), we get EFA proves WCon(A) $\rightarrow$ C. Hence EFA proves C $\leftrightarrow$ WCon(A). QED

3. PREVIOUS EVOLVING WORK ON G2

I wrote extensively about G2 on the FOM email list from 2003 through 2017. There was a long series of evolving and perfecting ideas.

165: Incompleteness Reformulated 4/29/03 1:42PM http://www.cs.nyu.edu/pipermail/fom/2003-April/006441.html

Godel's Theorems 4/29/03 9:57PM
http://www.cs.nyu.edu/pipermail/fom/2003-April/006443.html

166: Clean Godel Incompleteness 5/6/03 11:06AM
http://www.cs.nyu.edu/pipermail/fom/2003-May/006496.html

167: Incompleteness Reformulated/More 5/6/03 11:57AM
http://www.cs.nyu.edu/pipermail/fom/2003-May/006504.html

168: Incompleteness Reformulated/Again 5/8/03 12:30PM
http://www.cs.nyu.edu/pipermail/fom/2003-May/006518.html

174: Directly Honest Second Incompleteness 6/3/03 1:39PM
http://www.cs.nyu.edu/pipermail/fom/2003-June/006694.html

284: Godel's Second 5/9/06 10:02AM
http://www.cs.nyu.edu/pipermail/fom/2006-May/010524.html

285: Godel's Second/more 5/10/06 5:55PM
http://www.cs.nyu.edu/pipermail/fom/2006-May/010529.html

286: Godel's Second/still more 5/11/06 2:05PM
http://www.cs.nyu.edu/pipermail/fom/2006-May/010532.html

305: Proofs of Godel's Second 12/21/06 11:31AM
http://www.cs.nyu.edu/pipermail/fom/2006-December/011214.html

306: Godel's Second/more 12/23/06 7:39PM
http://www.cs.nyu.edu/pipermail/fom/2006-December/011223.html

307: Formalized Consistency Problem Solved 1/14/07 6:24PM
http://www.cs.nyu.edu/pipermail/fom/2007-January/011282.html

343: Goedel's Second Revisited 1 5/27/09 6:07AM
http://www.cs.nyu.edu/pipermail/fom/2009-May/013753.html

344: Goedel's Second Revisited 2 6/1/09 9:21PM
http://www.cs.nyu.edu/pipermail/fom/2009-May/013778.html

346: Goedel's Second Revisited 3 6/16/09 11:04PM
http://www.cs.nyu.edu/pipermail/fom/2009-June/013828.html

347: Goedel's Second Revisited 4 6/20/09 1:25AM
http://www.cs.nyu.edu/pipermail/fom/2009-June/013837.html

348: Goedel’s Second Revisited 5 6/22/09 11:00AM
http://www.cs.nyu.edu/pipermail/fom/2009-June/013843.html

388: Goedel’s Second Again/definitive? 1/7/10 11:06AM
http://www.cs.nyu.edu/pipermail/fom/2010-January/014290.html

580: Goedel’s Second Revisited 5/29/15 5:52 AM
http://www.cs.nyu.edu/pipermail/fom/2015-May/018750.html

Explaining Some Goedel 4/18/16 11:06PM
http://www.cs.nyu.edu/pipermail/fom/2016-April/019745.html

706: Second Incompleteness/1 9/5/16 2:03AM
707: Second Incompleteness/2 9/8/16 3:37PM
708: Second Incompleteness/3 9/11/16 10:33PM

706-708 are the model theoretic forms of G2 with most proofs given in detail. Most of the ideas are already presented in the 165-580 listed above.

773: Goedel’s Second: Proofs/1 Dec 18 20:31:25 EST 2017
774: Goedel’s Second: Proofs/2 Dec 18 20:36:04 EST 2017
775: Goedel’s Second: Proofs/3 Dec 19 00:48:45 EST 2017
777: Goedel’s Second: Proofs/4 12/28/17 8:02PM
778: Goedel’s Second: Proofs/5 12/30/17 2:40AM

773-778 concern the Remarkable Set proof of G2, which has been refined and perfected recently according to section 5 below.

4. PROOF OF G2/1-con BY TRANSPARENT DIAGONALIZATION

The prime example of what we call Transparent Diagonalization is the usual proof by Cantor that given an infinite sequence of subsets of N, there is a subset of N that is missing. This diagonalization argument is unlike the diagonalization/self reference argument used in Gödel's original proofs.

The diagonalization/self reference argument used in Gödel's original proofs is still considered rather mysterious in light of, for example, Barkley Rosser's use of it in the Gödel/Rosser theorem. To this day we don't have a good understanding of what Rosser sentences are like under "natural" numberings. For a "usual" numbering, we don't know whether any two Rosser sentences are equivalent, and also how the Rosser sentences compare when we use different "natural" numberings. See [GS79], [Bu08] for some background information.

A somewhat well known proof of a modified form of G2 can be proved using a very straightforward Transparent Diagonalization.

We associate the well known set of objects $\Theta(T)$ to each reasonable theory T. Namely the set of all provably recursive functions of T.

By reasonable here we will simply mean that T is presented as an r.e. theory whose language contains 0,S,+,• and whose axioms include PFA (polynomial function arithmetic, or bounded arithmetic). 1-Con(T) is formalized in a natural way using a natural enumeration of the Σ^0_1 formulas. We won't go into more details here.
1-Con(T) is also referred to as $\Sigma 1$ soundness for T.

G2(1-con) asserts that no reasonable theory proves its own 1-consistency. I.e., if T is reasonable then T does not prove 1-Con(T).

We associate an important well known set of objects $\Theta(T)$ to each reasonable theory T, namely $\Theta(T)$ = the set of all provably recursive functions of T. f is a provably recursive function of T if and only if there exists e such that f = φ_e is total, and T proves "φ_e is total". We will prove by Transparent Diagonalization that for reasonable T, $\theta(T)$ is properly included in $\Theta(T + 1\text{-Con}(T))$.

G2/1-CON. Let T be an r.e. presented extension of PFA with language containing L(PFA) = (0,S,+,•). Assume T is 1-consistent. Then T does not prove 1-Con(T).

The obvious proof of this gives the following satisfying information:

G2/1-CON GROWTH. Let T be an r.e. presented extension of PFA with language containing L(PFA) = (0,S,+,•). Assume T is 1-consistent. Then T + 1-Con(T) has properly more provably recursive functions than T. In fact there is a provably recursive function of T + 1-Con(T) that eventually strictly dominates every provably recursive function of T.

Proof: Let T be as given, where T is 1-consistent. Define f(n) by looking at all partial recursive functions for which its index and a proof in T that it is everywhere defined can be found ≤ n, and returning the least nonnegative integer that is greater than all of the values these functions have at n. Since T is 1-consistent, this describes a recursive function. It is clear that this recursive function eventually strictly dominates all provably recursive functions of T. Finally, note that this recursive function is a provably recursive function of T + 1-Con(T). QED

The origins of G2/1-con = G2 for 1-consistency are rather unclear. Lev Beklemishev has a paper in the 1980's about this, but it probably was first proved much earlier, perhaps when the notion of provably recursive functions of a theory first came into common use. That is probably in the 1950s with G. Kreisel. Some of the early proof theorists of that period are good candidates for having known about the directly straightforward proof of G2/1-con that we sketch now. (I don't know if Kreisel had this).

Much of the philosophical force of G2 is already available with G2/1-con. This indicates that it is very worthwhile to investigate G2/1-con with the same intensity and detail as G2 has been investigated.

5. TOWARDS DEMYSTIFYING G2

We do not have a proof of G2 by Transparent Diagonalization like we have for G2/1-con in section 4. It would seem that there ought to be such a proof.

But we do have an Arguable Demystification of G2 by another route. First let us review for the record the usual proof of G2.

USUAL MYSTERIOUS PROOF OF G2

Let us review a common proof of G2. Let T be reasonable and consistent. Form $\{e: e \in W_e\}$ and $\{e: T \text{ proves } e \notin W_e\}$. The first formation is mysterious. The second formation is not mysterious because it is a straightforward and interesting strong kind of complementation of the first formation.

But then we add some additional mystery by fixing n so that $W_n = \{e: T \text{ proves } e \notin W_e\}$. Then we get that T proves

$$1)\ n \in W_n \leftrightarrow T \text{ proves } n \notin W_n$$

The rest of the proof of G2 establishes crucial properties of these constructions, and so we do not regard it as mysterious (in the sense we are concerned with here).

We have that T cannot prove $n \notin W_n$, because otherwise T proves $n \in W_n$, and T is inconsistent.

But T + Con(T) does prove $n \notin W_n$. For arguing in T + Con(T), from $n \in W_n$ we obtain T proves $n \notin W_n$ and T proves $n \in W_n$, which is a contradiction in T + Con(T).

So T + Con(T) proves $n \notin W_n$ yet T does not prove $n \notin W_n$, and so T cannot prove Con(T).

FROM ROGER'S FIXED POINT THEOREM TO G2

Rogers Fixed Point Theorem asserts the following.

RFPT. For any recursive $f:N \to N$ there exists e such that $\varphi_e = \varphi_{f(e)}$.

RFPT is really an immediate consequence of the earlier Kleene Second Recursion Theorem.

KSRT. For any partial recursive $F:N^2 \to N$ there exists e such that $(\forall n)(\varphi_e(n) \cong F(e,n))$.

Proof: Let H be a recursive function such that $(\forall m,n)(F(m,n) \cong \varphi_{H(m)}(n))$. Set e such that $(\forall n)(\varphi_e(n) \cong \varphi_{H(e)}(n))$. Then $(\forall n)(\varphi_e(n) \cong \varphi_{H(e)}(n) \cong F(e,n))$. QED

Proof of RFPT: Define $F(e,n) \cong \varphi_{f(e)}(n)$. By KRST, there exists e such that $(\forall n)(\varphi_e(n) \cong F(e,n))$. Hence $(\forall n)(\varphi_e(n) \cong \varphi_{f(e)}(n))$. QED

RFPT is not quite enough for us to derive G2. We need the following.

EXPLICIT RFPT. For any recursive $f:N \to N$ there exists e such that EFA proves $\varphi_e = \varphi_{f(e)}$.

This is of course immediately obvious from the proof we gave of KRST and RRT.

We use EFA = exponential function arithmetic here for convenience and with some care surely it can be reduced considerably.

So this proof of RFPT resembles the above proof of G2, up through 1), that we consider mysterious. However there is a difference. This proof is directly based on COMPUTATION, whereas 1) and its derivation is based on TRUTH VALUES. What we really have for G2 is SENTENCES THAT TALK ABOUT THEMSELVES, where for KRST we have PROGRAMS THAT USE THEMSELVES. Somehow, the first seems more mysterious than the second, because people and mathematicians don't really use directly self referential sentences, which so easily leads to the notorious LIAR'S PARADOX and other catastrophes. Whereas programs that use themselves - well because there is no requirement that executions terminate, this doesn't seem really scary. In fact, there is the so called RASP model of computation which explicitly supports in the most natural way imaginable, use of a program itself as data during execution (although the further step of allowing modification of the program during execution of itself is more problematic and fortunately not relevant here). We exploit this point in order to give an arguably demystified proof of explicit KRST below.

Proof of G2: Define recursive $f:N \to N$ as follows. Let f(e) be natural such that each $\varphi_{f(e)}(n)$ looks for a proof in T that no $\varphi_e(n)$ halts. Arguing in T, let $\varphi_e = \varphi_{f(e)}$. Then T proves

$\varphi_e(0)$ halts if and only if
there is a proof in T that $\varphi_e(0)$ does not halt

Then as in the original proof of G2 above, since T is consistent, T does not prove $\varphi_e(0)$ diverges. Also T + Con(T) proves $\varphi_e(0)$ diverges. Hence T does not prove Con(T). QED

RASP APPLIED TO RFPT

We offer a proof using Random Access Stored Program model of computation. This is the same as the RAM architecture but where the program itself is stored in the first register, and it can be accessed as data during execution just like any other data. However, we do not allow the program to be altered during execution. This has also been investigated, but we don't need this here.

We can define φe in the usual way according to the RASP model and get an admissible enumeration of the partial recursive functions. Recall that all admissible enumerations of the partial recursive functions are recursively isomorphic.

We then get an immediate proof of RFPT by writing a RASP program that calls for computing f at the stored program from the program register, and then restarting by placing this value of f in the program register.

Because f is total, this program will have the same outcome either by using the stored program normally, or by instead by just using the restarts.

So this proof seems different than the proof of RFPT above. Perhaps this difference can manifest itself in new kinds of information about KSRT, RFPT, and G2.

REMARKABLE AND EXPLICITLY REMARKABE SETS

We finally turn to a slightly novel proof of G2 that can be construed as being suggestively organized - rather than radically new.

The idea is to use the notion of REMARKABLE SET to push all of the work that can be construed as diagonalization or mysterious into recursion theory. Actually it is rather invisible also as recursion theory, almost unnoticeable. So what diagonalization remains is particularly friendly.

DEFINITION 5.1. A is remarkable if and only if A is an r.e. set which agrees somewhere with every r.e. set. I.e., for every r.e. set B, there exists e such that $e \in A \leftrightarrow e \in B$.

It is very easy to see that this notion looks to be intriguing, but is really rather pedestrian. For what does it mean to NOT be remarkable? Just that A is r.e. and disagrees everywhere with some r.e. set. But that just means that A is r.e. with an r.e. complement. I.e., A is not remarkable if and only if A is not recursive.

I.e., we have shown the following.

THEOREM 5.1. A is remarkable if and only if A is r.e. and not recursive.

Now we introduce a natural strengthening of remarkable using the weak system EFA of exponential function arithmetic. Other weak systems can be used.

Coming back to the definition of remarkable, it is a very common move in mathematics to take a notion, which asserts existence, and simply ask that one be very explicit about an example. Thus we are led quickly to the following notion.

DEIFNITION 5.2. A is EFA remarkable if and only if for all r.e. sets B, there exists e such that EFA proves that A,B agree at e.

Here we just use EFA = exponential function arithmetic, as a convenient way of making things very explicit.

THEOREM 5.2. There is an explicitly remarkable set A.

Proof: This kind of thing is very much present in recursion theory where one has extra effectivity. We can use a familiar natural complete r.e. set A. We can effectively find a place of agreement for any r.e. set B from the r.e. index of B. NAMELY THE INDEX OF B! So this is NOT EVEN REALLY A DIAGONAL ARGUMENT. Set $A = \{e: e \in W_e\}$. Let $B = W_r$. Then $r \in A \leftrightarrow r \in B$, which is obviously provable in EFA. QED

So the only real hint of a diagonal argument so far is just the definition of $A = \{e: e \in W_e\}$, a very familiar construction in elementary recursion theory.

We now prove G2 by starting with any EFA remarkable A, forming an obviously interesting and natural set B related to A, apply EFA remarkability to A,B, and then argue without any trace of diagonalization or mystery.

THEOREM 3.3. G2.

Proof: Let T appropriately extend EFA, and prove Con(T). We will obtain a contradiction in T.

Let A be EFA remarkable. Apply remarkable to A and $\{e: T \text{ proves } e^* \notin A\}$.

Fix n* such that

$$n^* \in A \leftrightarrow T \text{ proves } n^* \notin A$$

is provable in EFA. Arguing in T, if $n^* \in A$ then T proves $n^* \in A$, and also T proves $n^* \notin A$, and therefore T is inconsistent. But still arguing in T, T is consistent (using hypothesis).

Hence $n^* \notin A$. So we have obtained a proof of $n^* \notin A$ within T. So still arguing in T, we have the right side, and so we have $n^* \in A$. This means that we have also obtained a proof of $n^* \in A$ within T. Thus we have a contradiction in T. QED

We can look at it this way. Obviously we cannot apply any kind of remarkability to A and N\A, for N\A is not r.e. We would get there exists n such that $n \in A \leftrightarrow n \notin A$, which is complete nonsense. However, the next "best" thing would be to apply remarkability to A and a "negative form of A". Now what is a good negative form A to use? A legal one of course. So it has to be r.e., Obviously

T proves $n^* \notin A$

is the most obvious negative form of A. And that is exactly what we did use.

So for all of this effort to remove mysterious diagonalization from the proof of G2, it would be nice to get some ***good solid additional proof theoretic information***. At this point I do not have such a thing.

I have been finding out that both the category theory community and the type theory community both are quite wired into at least some forms of G1. But they both have no clue as to the meaning of G2 for their subject.

I have been hoping that this Remarkable Set proof is so basic as to be suggestive of what G2 means for category theory and type theory.

6. TANGIBLE INCOMPLETENESS

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square. nice undergraduate exercise

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square which is also order invariant. seriously false

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square which is order invariant over Z[0,n]. close to truth but not quite - refutable

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square which is order invariant over Z[1,n]. nice theorem of countable combinatorics

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square whose <1 sections are order invariant over Z[1,n]. tangible incompleteness

This last statement is provably equivalent to Con(SRP) over WKL_0. We are writing up that reversal now. I will talk on it after I have a full blown manuscript because of the heavy details.

There is also the stronger statement, already proved from WKL_0 + Con(SRP) in lectures:

Every order invariant subset of $Q[0,n]^{2k}$ has a maximal square whose <i sections, i < n, are order invariant over Z[i,n]. tangible incompleteness

It is also provably equivalent to Con(SRP) over WKL_0.

REFERENCES

[Be97] L. D. Beklemishev, "Induction rules, reflection principles, and provably recursive functions", *Ann. Pure Appl. Logic*, **85**:3 (1997), 193–242.

[Be05] L. D. Beklemishev, "Reflection principles and provability algebras in formal arithmetic", *Russian Math. Surveys*, **60**:2 (2005), 197–268.

[Bu08] Christopher von Bulow, A remark on equivalent Rosser sentences, Annals of Pure and Applied Logic 151 (2008) 62–67

[Ch20] Yong Cheng, Finding the Limit of Incompleteness 1, The Bulletin of Symbolic Logic, September-December 2020, vol. 26, no. 3/4, pp. 267-286.

[GS79] D. Guaspari, R.M. Solovay, Rosser sentences, Annals of Mathematical Logic 16 (1) (1979) 81–99.

[Je16] Emil Jerabek, Division by zero, Archive for Mathematical Logic 55 (2016), no. 7, pp. 997-1013.

[Le83] D. Leivant, The optimality of induction as an axiomatization of arithmetic. The Journal of Symbolic Logic, 48:182–184, 1983.