ADVENTURES IN 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
September 18, 2023

Abstract. We begin discussing various forms of G1 put into the form: If a first theory satisfies one or more adequacy conditions then it has one or more wildness properties. We continue with the new "no interpretation" forms of G2, which are fundamentally model theoretic formulations. We also give corresponding model theoretic characterizations of the consistency statement Con(T) for finitely axiomatized T. We present the known proof of the 1-Con form by G2 by transparent diagonalization and discuss attempts to do so for G2.

1. INTRODUCTION

We begin with a discussion of various forms of G1 put into the following general form: If a first order 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 fundamental model theoretic 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 from them is also straightforward and does not rely on any of the ingredients in the known proofs of G2. We also give corresponding fundamental model theoretic 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 there are uncountably many infinite sequences of 0’s and 1’s. 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 present two proofs of G2. The first proof is a proof using what we call explicit transparent diagonalization. We don’t view the explicitness there as rising to the level of transparency so we regard that finding a proof of G2 using transparent diagonalization (as we have for G2/1-con) as open. The second proof puts all of the diagonalization related ideas into a basic familiar situation in recursion theory that is a particularly transparent diagonalization of its own. This is the construction of what we call a remarkable set and an EFA effectively remarkable set. Then we take any EFA effectively remarkable set and apply its remarkability to a naturally closely associated set and derive G2 now without any semblance of diagonalization. We have retained [Fr21] in the list of references, because there are ideas we don’t discuss here that may have some future importance.

2. G1. GÖDEL'S FIRST INCOMPLETENESS THEOREM

By a theory we will usually mean a theory T in the usual PC(=), (predicate calculus with equality), which comes with a designated language (of constants, relations, and function symbols). Sometimes it is important to use many sorted logic.

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 elementary form of G1 involving only the most rudimentary of notions is unarguably 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. R was introduced in [TMR53] and is often discussed along with Q as in, e.g., [Vi17]. 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 reasonably natural finitely axiomatized system K usable for pure G1 and variants of G1, of complexity at most that of AS, interprets AS.

This conjecture can be approached via complexity notions like quantifier depth and number of symbols.

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].

3. G2. GÖDEL'S SECOND INCOMPLETENESS THEOREM

This section originated with section 2 of [Fr21].

We use the systems PA (Peano arithmetic), PRA (primitive recursive arithmetic), SEFA (superexponential arithmetic), EFA (exponential function arithmetic), PFA (polynomial function arithmetic). The super exponential is the iterated exponential. SEFA, EFA, PFA are also known as $I\Sigma_0$(superexp), $I\Sigma_0$(exp), $I\Sigma_0$, respectively.

3.1. NO INTERPRETATION VERSIONS

We formulate a purely model theoretic form of G2 which we call No Interpretation G2/PA. Its more natural formulation uses many sorted logic. Many fundamental logical systems such as Z_2 are traditionally formulated using many sorted logic for a variety of metamathematical purposes.

We begin with the listing of the five versions of G2 for five basic systems PA, $I\Sigma_n$, PRA, SEFA, and EFA.

FIVE NO INTERPRETATION VERSIONS OF G2

Below for all many sorted T we require that one of the sorts of T is the arithmetic sort with the primitives 0,S,+,• included.

NO INTERPRETATION G2/PA. NIG2(PA). No consistent many sorted theory T that proves PA is interpretable in any theorem of T in the language of PA. We cannot remove "in the language of PA".

NO INTERPRETATION G2/($I\Sigma_n$). NIG2($I\Sigma_n$). Let $n \geq 1$. No consistent many sorted theory T that proves $I\Sigma_n$ is interpretable in any Σ^0_{n+2} theorem of T. We cannot replace Σ^0_{n+2} by Π^0_{n+2}.

NO INTERPRETATION G2/PRA. NIG2(PRA). No consistent many sorted theory T that proves PRA is interpretable in any Π^0_1 theorem of T in L(PRA). We cannot replace Π^0_1 by Π^0_2.

NO INTERPRETATION G2/SEFA. NIG2(SEFA). No consistent many sorted theory T that proves SEFA is interpretable in any Π^0_1 theorem of T in L(EFA). We cannot replace Π^0_1 by Π^0_2. We cannot replace L(EFA) by L(SEFA).

NO INTERPRETATION G2/EFA. NIG2(EFA). No consistent many sorted T that proves EFA is interpretable in any Π^0_1 theorem of T in L(PFA). We cannot replace Π^0_1 by Π^0_2. We cannot replace L(PFA) by L(EFA).

FIVE RESTRICTED NO INTERPRETATION THEOREMS

Read "RNI" as "restricted no interpretation". The simplification amounts to just requiring that the theory T be in the same language as PA, $I\Sigma_n$, PRA, SEFA, EFA, respectively.

RNI(PA). No consistent theory T in L(PA) that proves PA is interpretable in any theorem of T.

RNI($I\Sigma_n$). No consistent theory T in L(PA) that proves $I\Sigma_n$ is interpretable in any Σ^0_{n+2} theorem of T. We cannot replace Σ^0_{n+2} by Π^0_{n+2}.

RNI(PRA). No consistent theory T in L(PRA) that proves PRA is interpretable in any Π^0_1 theorem of T. We cannot replace Π^0_1 by Π^0_2.

RNI(SEFA). No consistent theory T in L(SEFA) that proves SEFA is interpretable in any Π^0_1 theorem of T in L(EFA). We cannot replace Π^0_1 by Π^0_2. We cannot replace L(EFA) by L(SEFA).

RNI(EFA). No consistent theory T in L(EFA) that proves EFA is interpretable in any Π^0_1 theorem of T in L(PFA). We cannot replace Π^0_1 by Π^0_2. We cannot replace L(PFA) by L(EFA).

EASY DERIVATIONS OF THE RNI FROM THE NIG2

In each case, we have merely restricted the language of the theory T. We now directly prove the negative parts of the RNI's except RNI(PA), and the negative part of NIG2(PA), and therefore the negative parts of the NIG2's.

negative NIG2(PA). Set T = ACA_0. Then T proves PA, T proves T, and T is interpretable in T (ACA_0 is finitely axiomatized).

negative RNI($I\Sigma_n$). Set T = $I\Sigma_n$ and T proves the Π^0_{n+2} sentence T.

negative RNI(PRA). Let T be PRA + the Π^0_2 sentence A = "EFA + 'the usual algorithms for the primitive recursively defined functions terminate at every argument'". Then T is a consistent theory in L(PRA) that proves PRA + A. Also T is interpretable in A by interpreting each primitive recursive function using its standard algorithm.

negative RNI(SEFA). Let T = SEFA + the sentence A = "the usual algorithm for the super exponential terminates at every argument." Then T is consistent, is in L(SEFA), proves SEFA, and is interpretable in the Π^0_2 sentence A in L(EFA). For the second negative claim, set T = SEFA and note that SEFA itself is Π^0_2 in L(SEFA).

negative RNI(EFA). Let T = EFA + the sentence A = "the usual algorithm for the exponential terminates at every argument." Then T is consistent, is in L(PFA), proves EFA, and is interpretable in the Π^0_2 sentence A in L(PFA). For the second negative claim, set T = EFA and note that EFA itself is Π^0_2 in L(EFA).

EASY DERIVATIONS OF THE NIG2 FROM THE RNI

In each case, we apply RNI to the set of sentences in the language used that is provable in the given T. We have already proved the negative claims in the NIG2 by proving the negative claims in the corresponding RNI.

RNI(PA) $\rightarrow$ NIG2(PA). Let T be a many sorted theory that proves PA, T proves A in L(PA), and T is interpretable in A. Let T* be the set of theorems of T that are in L(PA). Then T* is a theory in L(PA) that proves PA and is interpretable in A, where A is a theorem of T*. Hence T* is inconsistent. Therefore T is inconsistent.

RNI($I\Sigma_n$) → NIG2($I\Sigma_n$). Let T be a many sorted theory that proves $I\Sigma_n$, T proves Σ^0_{n+2} sentence A, and T is interpretable in A. Let T* be the set of theorems of T that are in L(PA). Then T* is in L(PA), T* proves Σ^0_{n+2} sentence A, and T* is interpretable in A. Hence T* is inconsistent. Therefore T is inconsistent.

RNI(PRA) → NIG2(PRA). Let T be a many sorted theory that proves PRA, T proves Π^0_1 sentence A, and T is interpretable in A. Let T* be the set of theorems of T that are in L(PRA). Then T* is in L(PRA), T* proves PRA proves Π^0_1 sentence A, and T is interpretable in A. Hence T* is inconsistent. Therefore T is inconsistent.

RNI(SEFA) → NIG2(SEFA). Let T be a many sorted theory that proves SEFA, T proves Π^0_1 sentence A in L(EFA), and T is interpretable in A. Let T* be the set of theorems of T that are in L(SEFA). Then T is in L(SEFA), proves SEFA, is interpretable in the Π^0_1 theorem A of T, A in L(EFA). Then T* is inconsistent. Therefore T is inconsistent.

RNI(EFA) → NIG2(EFA). Let T be a many sorted theory that proves EFA, T proves Π^0_1 sentence A in L(PFA), and T is interpretable in A. Let T* be the set of theorems of T that are in L(EFA). Then T is in L(EFA), proves EFA, is interpretable in the Π^0_1 theorem A of T, A in L(PFA). Then T* is inconsistent. Therefore T is inconsistent.

PROOFS OF THE RNI

RNI(PA). Let T be in L(PA), T proves PA + A, T is interpretable in A. We show that T is inconsistent.

It is known that A → Con(A) is provable in PA using partial truth definitions and cut elimination. (This comes under reflection principles and reflexivity. See, e.g., [Ber90], Theorem 2.6). Since T proves PA + A, we have T proves Con(A). Let T′ ⊆ T be finite, T′ proves Con(A) + EFA. Since T′ is interpretable in A, we have EFA + Con(A) implies Con(T′). Hence EFA + T′ proves Con(EFA + T′). By G2, EFA + T′ is inconsistent. Hence T′,T are inconsistent. QED

RNI($I\Sigma_n$). Let T be in L(PA), T proves $I\Sigma_n$ + A, A is Σ^0_{n+2}, T is interpretable in A. We show that T is inconsistent.

It is known that A → Con(A) is provable in $I\Sigma_n$ using partial truth definitions and cut elimination, as proved in [Le83] (also see [Be97], [Be05]). Since T proves $I\Sigma_n$ + A, we have T proves Con(A). Let T′ ⊆ T be finite, T′ proves Con(A) + EFA (using $n \geq 1$). Since T′ is interpretable in A, we have EFA + Con(A) implies Con(T′). Hence EFA + T′ proves Con(EFA + T′). By G2, EFA + T′ is inconsistent. Hence T′ and T are inconsistent. QED

RNI(PRA). Let T be in L(PRA), T proves PRA + A, A is Π^0_1, T is interpretable in A. We show that T is inconsistent.

It is known that A → Con(A) is provable in PRA for Π^0_1 A in L(PRA), since we have Herbrand's theorem available in PRA and induction applied to bounded formulas in the primitive recursive function symbols used in A. Use of Herbrand here involves iteration of the underlying functions, afforded by PRA. Since T proves PRA + A, we have T proves Con(A). Let T′ ⊆ T be finite, T′ proves Con(A) + EFA. Since T′ is interpretable in A, we have EFA + Con(A) implies Con(T′). Hence EFA + T′ proves Con(EFA + T′). By G2, EFA + T′ is inconsistent. Hence T′ and T are inconsistent. QED

RNI(SEFA). Let T be in L(SEFA), T proves SEFA + A, A in Π^0_1 in L(EFA), T is interpretable in A. We show that T is inconsistent.

It is known that A → Con(A) is provable in SEFA for Π^0_1 A in L(SEFA). To see this, assume A 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. Since T proves SEFA + A, we have T proves Con(A). Let T′ ⊆ T be finite, T′ proves Con(A) + EFA. Since T′ is interpretable in A, we have EFA + Con(A) implies Con(T′). Hence EFA + T′ proves Con(EFA + T′). By G2, EFA + T′ is inconsistent. Hence T′ and T are inconsistent. QED

RNI(EFA). Let T be in L(EFA), T proves EFA + A, A is Π^0_1 in L(PFA), T is interpretable in A. We show that T is inconsistent.

It is known that φ → WCon(φ) 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 φ → WCon(φ) in EFA.

Let T′ ⊆ T be finite, T′ proves WCon(A) + EFA. Since T′ is interpretable in A, we have EFA + WCon(A) implies WCon(T′). Hence EFA + T′ proves WCon(EFA + T′). By G2, EFA + T′ is inconsistent. Hence T′ and T are inconsistent. NOTE: G2 is well known to hold for WCon. QED

Note that we have used five special systems here. The question naturally arises as to what systems we can use. The five RNI's and also the five NIG2's can be investigated in this way, in a few directions, seeking exact characterizations. We will merely scratch the surface of this by restricting our attention to NIG2(PA) and RNI(PA).

First we take up RNI(PA), and see what we can replace PA with. For any system S in the language of PA, we consider

RNI(S). No consistent theory T in L(PA) that proves S is interpretable in any theorem of T.

THEOREM 3.1.1. S being mutually interpretable with PA is not sufficient for RNI(S). Use S = {Con(IΣ_n): n ≥ 1} as a counterexample.

Proof: It is known that S and PA are mutually interpretable, as they are recursively axiomatized theories each interpreting every finite fragment of the other. Now let T = {Con(PA)}. Then T proves and is interpretable in Con(PA). QED

DEFINITION 3.1.1. S has property 1) if and only if S is in L(PA), and there is an interpretation π_1 from PA into S and an interpretation π_2 from S into PA such that
i. for all sentences B in L(S), $\pi_1\pi_2$(B) ↔ B is provable in S.
ii. for all sentences B in L(PA), $\pi_2\pi_1$(B) ↔ B is provable in PA.

LEMMA 3.1.2. If S in L(PA) has property 1) then
i. π_2(S) and PA are logically equivalent.
ii. π_1(PA) and S are logically equivalent.

Proof: Let S be as given. Let PA prove B. Then $\pi_2\pi_1$(B) ↔ B is provable in PA. Hence $\pi_2\pi_1$(B) proves B and π_1(B) is provable in S. So π_2(S) logically implies PA. Obviously PA logically implies π_2(S) since π_2 is an interpretation of S into PA. The second claim is by symmetry. QED

THEOREM 3.1.3. Let S have property 1). Then RNI(S).

Proof Let S be as given. Let T be in L(PA), T proves S + A, T interpretable in A. By Lemma 3.1.2, π_2(T) proves π_2(S) and π_2(S) proves PA. Also π_2(T) proves π_2(A). We claim that π_2(T) is interpretable in π_2(A). We first interpret π_2(T) into T by π_1, and then T into A, and then A into π_2(A) by π_2. By NIG2(PA), we see that π_2(T) is inconsistent. Hence $\pi_1\pi_2$(T) is inconsistent. Let T′ $\subseteq$ T be finite, where $\pi_1\pi_2$(T′) is inconsistent. Then T proves $\pi_1\pi_2$(T′) $\leftrightarrow$ T′, and so T proves $\neg$T′. Hence T is inconsistent. QED

We conclude this section by routinely lifting RNI(S) to NIG2(S).

NIG2(S). No consistent theory T in L(PA) that proves S is interpretable in any theorem of T.

THEOREM 3.1.4. Let S be a many sorted theory. RNI(S) and NIG2(S) are equivalent.

Proof: Looking at the proof of the equivalence of RNI(PA) and NIG2(PA), no properties of PA were used. QED

THEOREM 3.1.5. S being mutually interpretable with PA is not sufficient for NIG2(S). Use S = {Con($I\Sigma_n$): n >= 1} as a counterexample. S having property 1) is a sufficient condition for NIG2(S).

Proof: From Theorems 3.1.3, 3.1.4. QED

3.2. EQUIVALENCE WITH CLASSICAL G2

To justify the name "No Interpretation G2" we now consider the relationship between NIG2/PA, NIG2($I\Sigma_n$), NIG2/PRA, NIG2(SEFA), NIG2(EFA), and certain forms of G2. These relationships need to be established without using any of the techniques involved in proving G2.

Note that we have derived Theorems NIG2/PA, NIG2($I\Sigma_n$), NIG2/PRA, NIG2(SEFA), NIG2(EFA) with only the invoking of G2 applied to finitely axiomatized theories extending EFA as the nontrivial step.

The derivation of G2 uses the formalized completeness theorem as the only nontrivial step as we document now.

THEOREM 3.2.1. NIG2/PA, NIG2($I\Sigma_n$), NIG2/PRA, NIG2(SEFA), NIG2(EFA), each implies G2 for r.e. presented theories in any language, where the axioms extend PA, $I\Sigma_n$, PRA, SEFA, EFA, respectively.

Proof: Suppose NIG2/PA, NIG2($I\Sigma_n$), NIG2/PRA, NIG2(SEFA), NIG2(EFA), and let T be a consistent r.e. presented extension of PA, $I\Sigma_n$, 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). This is the step that uses the formalized completeness theorem. EFA easily serves as this infrastructure. So using Theorems NIG2/PA, NIG2($I\Sigma_n$), NIG2/PRA, NIG2(SEFA), NIG2(EFA), we see that T is inconsistent, establishing G2, where the r.e. axioms extend PA, $I\Sigma_n$, PRA, SEFA, EFA, respectively, in each case. QED

3.3. CHARACTERIZATION OF Con STATEMENTS

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

THEOREM 3.3.1. For all sentences A, (A,Con(A)) obeys the following property P(A,Con(A)): For all arithmetic B, PA + B interprets A if and only if PA + B proves Con(A). For all sentences (A,Con(A)) is the unique arithmetic sentence with P(A,Con(A)) up to PA provable equivalence.

Proof: Let A be a sentence and B be an arithmetic sentence. If PA + B proves Con(A) then obviously PA + B interprets A via the formalized completeness theorem. Now suppose PA + B interprets A. Let $I\Sigma_n$ + B interpret A. for a suitable $n \geq 1$. Then EFA proves Con($I\Sigma_n$ + B) $\rightarrow$ Con(A). Now PA + B proves Con($I\Sigma_n$ + B) by formalized cut elimination and truth definition. Hence PA + B proves Con(A).

Now let C be an arithmetic sentence such that P(A,C). I.e., for all arithmetic sentences B, PA + B interprets A if and only if PA + B proves C. Then by P(A,Con(A)), we have that for all arithmetic sentences B,

*) PA + B proves C if and only if PA + B proves Con(A).

Setting B = C in *), we get PA proves C $\rightarrow$ Con(A). By setting B = Con(A) in *), we get PA proves Con(A) $\rightarrow$ C. Hence PA proves C $\leftrightarrow$ Con(A). QED

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

THEOREM 3.3.2. For all sentences A, (A,Con(A)) obeys the following property P(A,Con(A)): For all Π^0_1 sentences B in L(PRA), PRA + B interprets A if and only if PRA + B proves Con(A). For all sentences A, Con(A) is the unique Π^0_1 sentence in L(PRA) with P(A,Con(A)) up to PRA provable equivalence.

Proof: Let A be a sentence and B be Π^0_1 in L(PRA). If PRA + B proves Con(A) then obviously PRA + B interprets A via the formalized completeness theorem. 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 formalized cut elimination and truth definition, 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).

*) PA + B proves C if and only if PA + B proves Con(A).

Setting B = C in *), we get PRA proves C $\rightarrow$ Con(A). By setting B = Con(A) in *), we get PRA proves Con(A) $\rightarrow$ C. Hence PRA proves C $\leftrightarrow$ Con(A). QED

We leave it to the reader to obtain analogous statements for systems weaker than PRA.

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 the set of infinite sequences of 0’s and 1’s cannot be countable. This diagonalization argument is more direct and straightforward than the self reference argument used in Gödel's original proofs of G1 and G2. (The self reference Lemma goes back to [Ca37]). Those original proofs using the self reference lemma are 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], [KK17] for background information.

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

DEFINITION 4.1. T is adequate if and only if T is a finitely axiomatized theory extending EFA in many sorted logic with finitely many sorts. 1-Con(T) asserts that “every true Σ^0_1 sentence provable in T is true”, formalized in the well known way using a natural enumeration of the Σ^0_1 formulas.

1-Con(T) is also referred to as Σ_1 soundness for T.

G2/1-Con. No 1-consistent adequate theory proves its own 1-consistency. I.e., if T is adequate and 1-consistent, then T does not prove 1-Con(T).

The origins of G2/1-Con, or G2 for 1-consistency, with its transparent diagonalization proof are rather unclear, but it probably was first proved when the notion of provably recursive functions of a theory first surfaced. 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 below including G. Kreisel. Much more recently, [Be97] presents the straightforward proof of G2/1-Con by transparent diagonalization.

We associate an important well known set of objects $\Theta(T)$ to adequate T.

DEFINITION 4.2. Let T be adequate. $\Theta(T)$ is 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 $f = \varphi_e$ is total, and T proves "φ_e is total".

We prove the following strengthening of G2/1-Con by Transparent Diagonalization.

G2/1-Con GROWTH. Let T be adequate and 1-consistent. Then Θ(T) is a proper subset of Θ(T + 1-Con(T)). There is an enumeration of Θ(T) by a provably recursive function (of two variables) of T + 1-Con(T).

Proof: Let T be as given. 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 the adequate T + 1-Con(T). This is because 1-Con(T) implies that every provable Σ^0_1 sentence is true, and so the relevant values exist to be found. QED

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. We have chosen to simplify matters by requiring that T be finitely axiomatized.

DEFINITION 4.3. T is adequate/re if and only if T is r.e. presented extending EFA in many sorted logic with an r.e. presented set of sorts. Θ(T) is defined using the r.e. presentation.

Note that G2/1-Con more generally for all r.e. presented T extending EFA in many sorted logic whose sorts are r.e. presented, trivially follows from our G2/1-Con. Moreover,

G2/1-Con GROWTH/re. Let T be adequate/re and 1-consistent. Then Θ(T) is a proper subset of Θ(T + 1-Con(T)). There is an enumeration of Θ(T) by a provably recursive function (of two variables) of T + 1-Con(T).

Proof: Adapt the proof of G2/1-Con Growth to adequate/re T. QED

THEOREM 4.1. Let T be adequate/re. The following are equivalent.
i. Θ(T) is not the set of all recursive functions.
ii. Θ(T) has a recursive enumeration.
iii. T is 1-consistent.

Proof: iii → ii → i by G2/1-Con Growth/re. It now suffices to prove i → iii. Suppose T is not 1-consistent. Let $(\exists n)(R(n))$ be provable in T and false, R Δ_0. Let $f:\omega \to \omega$ be a recursive function with index e which is not a provably recursive function of T. Change e to the natural index e′ for computing g(n) = f(n) if f(n) is computed in a number of steps m such that $\neg(\exists n \leq m)(R(n))$; 0 otherwise. Then T proves $\varphi_{e'}$ is total (because $(\exists n)(R(n))$ is provable in T), and the actual $\varphi_{e'} = g$ is recursive, being the same as f (using that $(\exists n)(R(n))$ is false). So f is a provably recursive function of T. QED

THEOREM 4.2. Let T be adequate/re and 1-consistent. Then T + not 1-Con(T) is 1-consistent.

Proof: Let T be as given. Let T + not 1-Con(T) prove $(\exists n)(R(n))$, R Δ_0. We find n such that R(n). Now T proves $(\exists n)(R(n)) \vee$ 1-Con(T). Hence T + $(\forall n)(\neg R(n))$ proves 1-Con(T). Hence T + $(\forall n)$(notR(n)) is not 1-consistent by G2/1-Con. Let T + $(\forall n)(\neg(R(n))$ prove $(\exists m)(S(m))$, S Δ_0, $(\forall m)(\neg S(m))$. Then T proves $(\exists n)(R(n)) \vee (\exists m)(S(m))$. Since T is 1-consistent, let R(n) or S(m) hold. Since S(m) fails, we must have R(n). QED

5. PROOF OF G2 BY TRANSPARENT DIAGONALIATION

We recognize that the transparent diagonalization we use here falls short of the kind of transparent diagonalization we used in section 4. (We will offer a valiant attempt in Theorem 6.3 in section 6).

For specificity we work with G2 for PA. The discussion immediately generalizes.

We start with the following well known fact proved by transparent diagonalization.

THEOREM 5.1. There is a Π^0_1 sentence A such that not(A ↔ A is provable in PA). This is provable in PA.

Proof: By transparent diagonalization, the set of true Π^0_1 sentences is not r.e. In particular, the set of true Π^0_1 sentences cannot coincide with the set of Π^0_1 sentences provable in PA. A counterexample to this coinciding is the A we are claiming to exist. QED

A very common theme throughout mathematics (and science for that matter) is to find a specific witness after having proved an existential statement. I.e., to seek a particular object that serves as a witness.

This inexorably leads to the following.

THEOREM 5.2. There is a Π^0_1 sentence A such that PA proves not(A $\leftrightarrow$ A is provable in PA).

The usual way we obtain Theorem 5.2 is to turn to a called self reference lemma, which we consider more mysterious than transparent, at least in its present forms. Notice that no use of self reference is used in section 4. However, we think that there is something missing here that is more general and less technical. Of course, the existence of a constructive proof of an existential statement is sufficient, but that begs the question of when constructive proofs can be given. We won't pursue this further here.

Now using Theorem 5.2, we get a rather vivid proof of G2 for PA. Fix A as given by Theorem 4.2.

1. PA proves not(A iff A is provable in PA).
2. PA proves ((A and A is not provable in PA) or (notA and A is provable in PA)).
3. PA + Con(PA) refutes (notA and A is provable in PA).
4. PA + Con(PA) proves (A and A is not provable in PA).
5. Assume PA proves Con(PA). Then PA proves (A and A is not provable in PA)
6. Assume PA proves Con(PA). Then PA proves A, PA proves A is provable in PA.
7. Assume PA proves Con(PA). Then PA is inconsistent.
8. PA does not prove Con(PA).

6. PROOF OF G2 VIA REMARKABLE 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 6.1. A is remarkable if and only if A is an r.e. subset of ω which agrees somewhere with every r.e. subset of ω. 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., we have shown the following.

THEOREM 6.1. 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.

DEFINITION 6.2. A is EFA remarkable if and only if for all r.e. sets B, there exists e such that EFA proves that A and B agree at e.

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

THEOREM 6.3. 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, not just the special {e: e ∈ We}, forming an obviously interesting and natural set B related to A, apply EFA remarkability to A and B, and then argue without any trace of diagonalization or mystery.

THEOREM 6.4. G2.

Proof: We will use n* for the canonical closed term whose value is n.

Let T be adequate (Definition 3.1), and consistent. Also assume T proves Con(T). We obtain a contradiction as follows.

Let A be EFA remarkable. If we could apply EFA remarkable to A and {e: e ∉ A} then we would have an obvious contradiction (as these two sets agree nowhere). But we can't since {e: e ∉ A} is not r.e. So instead we apply EFA remarkability to A and {e: T proves e* ∉ A}, which is r.e.

By the EFA remarkability of A, fix n such that

1) n* ∈ A ↔ 'T proves n* ∉ A'

is provable in EFA. Arguing in T, if n* ∈ A then T proves n* ∈ A, and also T proves n* ∉ A, using 1). Therefore T is inconsistent. Thus in T, we have proved n* ∈ A → T is inconsistent, and so by hypothesis, T proves n* ∉ A. Then by 1), T proves n* ∈ A. Hence T is inconsistent, which is again a contradiction. QED

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.

[Ber90] A. Berarducci, "The interpretability of arithmetic", JSL 55 (1990), 1059-1089

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

[Ca37] R. Carnap, The Logical Syntax of Language, 1937 (English translation from 1934, by Amethe Smeaton).

[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.

[Fr21] H. Friedman, Aspects of Gödel Incompleteness, 9/8/21, 21 pages. This is the third revised lecture notes for the talk delivered 8/16/21 at the Online International Workshop on Gödel's Incompleteness Theorems, https://u.osu.edu/friedman.8/foundational-adventures/downloadable-lecture-notes-2/, #78.

[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.

[KK17] M. Kikuchi and T. Kurahashi, Universal Rosser Predicates, JSL, vol. 82, no. 1, pp 292-230, 2017.

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

[TMR53] A. Tarski, A. Mostowski, and R.M. Robinson. Undecidable theories. North-Holland,Amsterdam, 1953.

[Vi17] A. Visser, On Q. *Soft Comput* **21**, 39-56 (2017). https://doi.org/10.1007/s00500-016-2341-5