Genuinely Mathematical Incompletenesses

N. J. H. Wontner[*][†]

April 13, 2019

1 Introduction

Pick a formal system of mathematics, S. Certain $\mathcal{L}_S$-sentences are made true by S, but unprovable within S. Some examples are outlined in §2, namely Gödel sentences, Goodstein's Theorem, and game-theoretic determinacy.

However, there are reasons to think that Gödel sentences are qualitatively different from those of ordinary mathematics. In §3, I argue that these and other concrete non-examples demonstrate a stable, non-trivial classification 'genuinely mathematical proof' (GMP), and similarly 'genuinely mathematical sentence' (GMS), which are conceptually independent. The classifications are (most likely) contextually determined, dependent on culture and psychology. For example, a necessary condition for both classifications is surveyability. They are intimately connected to a widely agreed upon inter-subjective notion of 'mathematically interesting'.

Isaacson's account of genuine arithmetic is outlined in §4, which ostensibly shows the completeness of PA with respect to genuine arithmetic. In

[*]Christ's College & Faculty of Philosophy, Cambridge

[†]This article adapts and extends a 2018 paper submitted for the Cambridge MPhil philosophy degree. I would like to thank my supervisor, Prof. Michael Potter, and two anonymous reviewers for their many constructive inputs: this article is as much their work as mine. Further thanks are owed to David Milton and Makan Nojoumian for their advice and critique throughout the writing process.

§5, we see Horsten's attempted generalisation, that genuine mathematics is captured by ZFC. However, I argue that the arithmetic case is fundamentally dissimilar from the mathematics case. Overall, Horsten's reasoning is found wanting.

In the final section, §6, I review the game-theoretic results which appear to be the strongest putative examples of genuinely mathematical incompleteness. I argue that these sentences actually include under-determined general mathematical terms. Given that the acceptable precisifications of these under-determined terms yield different truth values for sentences, they are insufficient for genuine mathematics. So, we are able to excise some instances of incompleteness due to lack of precision in our general mathematical terms.

2 Incompletenesses

A suitably strong consistent formal system, S, has (S-)undecidable sentences. This means that a mathematician could be struggling to find an (S-)proof of a particular theorem, where there is none. This is a real worry to some mathematicians today [35]. We outline some examples.

2.1 Gödel Sentences

Theorem 1 (Gödel's First Incompleteness Theorem [14])**.** *Let S be a formal system which interprets PA, and let c be an internal coding of the syntax of S. Then S contains a true sentence, $G_{(S,c)}$, which is not S-provable .*

$G_{(S,c)}$ is exactly the sentence $\forall x F(x) = 0$ for some primitive recursive function, F, which will depend on the specific coding. Hence, the sentence is Π^0_1.

2.2 Goodstein's Theorem

A further example of incompleteness in PA is Goodstein's Theorem:

Theorem 2 (Goodstein's Theorem)**.** *Every Goodstein Sequence terminates.*[1]

Goodstein's Theorem also has problematic metatheoretic properties, as its proof requires ε_0-induction. This is equivalent to proving $\mathrm{Con}_{\mathrm{PA}}$, hence it is PA-undecidable [17] [23].[2]

Theorem 3. *Goodstein's Theorem is independent of* PA*, if* PA *is consistent.*

2.3 Determinacy

Mathematics is stronger than PA, but stronger systems also suffer from 'interesting' incompletenesses ([33] §7.5). Some of these can be found by looking at the theory of games.

In a (Gale-Stewart) game, the two players move successively, choosing from the same (countable) stock of moves, with some winning plays ([29] §6A). The study of strategies on games is clearly mathematically interesting. Such a game can be modelled by its winning set, $W \subseteq \mathbb{N}^{\mathbb{N}}$: $n \in \mathbb{N}$ represents a move; a play is a sequence of moves, so is represented by $\bar{n} \in \mathbb{N}^{\mathbb{N}}$. Then the winning set is defined as $W = \{\bar{n} \in \mathbb{N}^{\mathbb{N}} \mid \bar{n}$ represents a winning play$\}$.

Definition 4 (Determinacy)**.** *A game is determined if and only if at every point there is a winning strategy for one of the players.*

Definition 5 (Borel)**.** *The collection of Borel sets of a space is the smallest collection containing the open sets that is closed under countable unions, and relative complements.*

Theorem 6 (BDT)**.** *Any game on a Borel set in the space of games ($\mathbb{N}^{\mathbb{N}}$) is determined.*[3]

Theorem 7. BDT *is independent of* Z*, if* Z *is consistent.*

[1] A Goodstein Sequence, $(G(n)_k) \in \mathbb{N}^{\mathbb{N}}$ starts with the exponential base 2 representation of $n \in \mathbb{N}$. The next value is obtained by increasing each instance of the current base by 1, then subtracting 1 [17].

[2] Goodstein's Theorem is Π^0_2. The Riemann Hypothesis is a possible natural Π^0_1 example (via some coding) [21].

[3] The topology on $\mathbb{N}^{\mathbb{N}}$ is non-standard.

The proof of BDT essentially relies on repeated uses of Replacement ([11] §1), which is independently controversial ([33] §12, [38] §2.3.10).

Definition 8 (Projective). *The collection of projective sets of a space is the smallest collection containing the Borel sets that is closed under taking continuous images, countable unions, and relative complements.*

Generalisation 9 (PD). *All projective games are determined.*

Theorem 10. PD *is independent of* ZFC*, if* ZFC *is consistent.*

The standard proof of PD relies on a large cardinal axiom ([29] §6H):

Axiom 11 (IW). *There are infinitely many Woodin cardinals.*

Generalisation 12 (AD). *All games are determined.*

Theorem 13. AD *is independent of* ZF*, if* ZF *is consistent.*

The strongest known grounds for AD is that it holds in $L(\mathbb{R})$, given a strengthening of IW ([30] §8.24).[4] Moreover, AD contradicts Choice ([29] §6H), though is consistent with countable Dependent Choice [22].

Theorem 14. ZFC $\vdash \neg$ AD

3 GMP & Interest

Something seems to be off about the candidate examples of incompleteness. The first class of candidates were the Gödel sentences. But these seem unnatural, and their proof methods are distinctively *unlike* ordinary mathematics ([33] §13.1). Their meta-theoretic content can only be interpreted via specific arbitrary codings. So, the meta-claim of undecidability is encoded by $G_{(S,c)}$, but $G_{(S,c)}$ itself is a 40,000+ character long arithmetic sentence.[5] This seems to be the kind of arithmetic sentence which (a) mathematicians do not want to prove, and (b) tells us nothing about undecidability. It seems, then, that $G_{(S,c)}$ *is not genuinely mathematical.*

[4] $\mathrm{AD}^{L(\mathbb{R})}$ is also motivated by suitable applications of a width reflection principle ([18]:169).

[5] Where the character count includes indices and primes [26].

The term 'genuinely mathematical' is difficult to unpack in full generality. Hence we focus on GMP, for classifying grounds (or proofs), and GMS, for sentences (or problems, or questions). In §3.3, I argue that these are distinct.

These classifications are somewhat aesthetic. They also appear to be contextually determined. They vary by culture: the Romans were principally interested in mathematical applications ([15]:116), whilst contemporary interest extends much further. Like other cultural tastes, e.g. those in art, some intuitions are more valuable. In our case, the intuitions of mathematicians are most valuable. Moreover, our tastes are tutored, in this case by mathematical educators, and are partially based on psychology and biology.[6]

Rather than total relativism, there is widespread inter-subjective agreement on what is interesting (hence people read each others' papers, work together, etc.), allowing us to classify cases as *not genuinely mathematically interesting*, relative to our standards.[7]

3.1 Sociological Information

The aim is to sharpen the approximate notion of GMP held by many. One attempt at capturing this sociological information is to characterise GMP (and GMS) as 'what is done in mathematics departments'. But this is overly simplistic. Mathematics departments contain "foundations of mathematics provocateurs" (Friedman [12]), e.g. certain set theorists, whose inclusion seems unsuitable ([16] §5.2). For example, the higher reaches of set theory are only tenuously linked with mathematics proper, especially given their segmentation into incompatible systems. Questions like "does this set have this property?" are sometimes unanswerable, as set theories and theorists disagree. This is distinctively unlike canonical areas of mathematics.

[6]There may be an informative objective substratum underpinning the classification 'mathematically interesting' which GMP approximates, like saltiness in (gustatory) taste.

[7]We need not accept conventionalism, as our mathematical interests may be dictated by *how we are*, but indexed by context.

The sociological information[8] is better captured by a suitably tutored opinion poll. The equivalent GMS test seems plausible also.

Test (Reaction Test)**.** *A proof is GMP if having explained to a mathematician, they assert that it is mathematical/mathematically interesting.*

This test is naturalistically justified by the intuitive authority mathematicians have to *identify* mathematics (though they may not have the authority to *explicate* it) [31]. For the reasons above, care is needed on who counts as a mathematician.[9]

3.2 Non-Triviality

To show that the classification 'GMP' is non-trivial, we can consider some paradigmatic examples and non-examples (the same can be given for GMS sentences).

We start with paradigmatic examples of GMP proofs. Ordinary proofs of arithmetic or algebra are unambiguously GMP, e.g. proof that $\forall r \in \mathbb{R}(r^2 \geq 0)$. There are also sociological examples, e.g. the content of today's A level mathematics syllabus.

A plausible non-example is Goodstein's Theorem. Any given instance, "$(G(n)_k)$ terminates", *is* a PA-theorem. However, some of these PA-proofs are incomprehensibly long. For example, $G(51)$ terminates after around $10^{15,151,337}$ steps [33]. If we limit ourselves to only the structures legitimated by PA (i.e. disallowing ε_0-insights), we simply cannot engage with the content of this proof: it essentially relies on a sequence of increments beyond our capacity for understanding. The same phenomenon occurs with Boolos' Sentence in [2], whose shortest PA_1-proof is simply monumental.

These sentences are only problematic if we constrain ourselves to PA, whilst we may be able to give GMP proofs using a stronger proof theory, e.g. Z. But a similar phenomenon occurs in 'everyday' mathematics, where there are no obvious proofs in legitimate stronger formal theories.

[8]This build on ideas from Potter, Friedman, and Feferman [9].

[9]One candidate solution would be a generic generalization of 'mathematician'.

Current proofs of the four-colour theorem depend on computational assistance [1], they distinguish cases and computationally check each case. But the checking process, again, is too long for human comprehension. These proofs are unfeasible, no person could comprehend one at once.[10] The four-colour theorem's *status* is interesting, as are understandable proof *subsections*. But given that no one can *engage* with the proof's content in full, any interest is surely only in connected material, e.g. sections of the proof or implications of the theorem.

The link with psychology explains the importance of grasping or understanding; the mathematical content of a proof can only be interesting if it *could* be understood by a mathematician in its entirety.[11] This means that the test puts a limit on the *length* of GMP proofs. It must have a *surveyable* or *comprehensive* form, no longer than e.g. a few articles. This is necessary, but possibly not *constitutive* of GMP.[12]

However, GMP does not coincide with surveyability. This can be seen from the case of $G_{(S,c)}$. *If* $G_{(S,c)}$ were surveyable, we would *still* excluded it. We can draw this out by considering the community reaction to the acts of proving different theorems. Spending a lifetime proving that $\forall x F(x) = 0$ for some random primitive recursive function, F would not engage the community, whilst proving Goldbach's conjecture would. Even excluding facts about form, we do not (inter-subjectively) value every primitive recursive function. This gives us reason to there may be *no* GMP proof of $G_{(S,c)}$.

One candidate explanation for *contemporary* proofs of $G_{(S,c)}$ not being GMP is their use of unorthodox methods. Consider a theorem picked at random from a mathematics journal. Such a proof is overwhelmingly likely to purport to be an unconditional proof, outside of a formal system. So, the essential use of "formal languages, ... axiomatic systems and their models relative to a language" (Feferman, [8]) disqualify some proofs of

[10] Another example may be the current proofs which make up the classification of simple groups [32].

[11] We exempt 'unconstrained' mathematicians, e.g. god.

[12] $G_{(S,c)}$ or $\overbrace{\neg ... \neg}^{2n} \phi$, for some large n and GMS ϕ, are possibly non-GMS, as they are unsurveyable.

undecidability.[13]

As with the use of unorthodox methods, the use of non-mathematical *assumptions* might disqualify a proof from being GMP. Proofs relying on large cardinal axioms are dubious, because the axioms themselves are not currently socially accepted, nor are they likely to be ([9] §1). This classifies the proofs of AD and PD as non-GMP, given that they rely on large cardinal assumptions.

3.3 GMP vs. GMS

A possible pretheoretic hypothesis is that GMP and GMS are conceptually codependent.[14] If a proof of ϕ is GMP, we naturally conclude that ϕ is GMS, reasoning that each line in a GMP proof is GMS, and that ϕ is a line of the proof. Conversely, one might think that mathematics is closed under proofs, i.e. a proof of a GMS sentence (in a suitable system) just is GMP.

However, the claim of dependency is unjustifiable. Firstly, one cannot generally ascertain proof complexity from sentential complexity. One example is PA, where speed-up theorems remain short, whilst their proof length (in lines) becomes unfeasibly long ([33] §13.8). The lack of proof of $G_{(S,c)}$ is something like a limiting case, where the proof is 'infinite'. So just because a sentence is GMS, it does not mean that it will have a GMP proof in any reasonable formal system.

Secondly, proof complexity essentially depends on the strength of the proof system. Adding set-rank levels permits both more theorems *and* shorter proofs. The short ε_0-induction proof of Goodstein's Theorem is an example ([33] §13.1). Another is Boolos' Sentence in [2], which has a comprehensible 3 page PA_2-proof, unlike in PA_1. More generally, there are Π^0_1 $\mathcal{L}_{\mathrm{PA}}$-sentences which are first provable high up the set-theoretic hierarchy. These are plausibly GMS (or even arithmetical), but they have non-GMP grounds, as their justifications require large cardinals. Whilst sen-

[13] These techniques *are* commonly distinguished from normal mathematics, being classified as metamathematical, foundational, etc.

[14] Perhaps based on Wittgenstein's view that the meaning of $\phi \in Sent(\mathcal{L}_{\mathrm{M}})$ is its proof [34].

tences can be classified as GMS independent of *any* formal system, GMP grounds are entangled with the salient system.

So there may be GMS sentences with no GMP solutions. A possible natural example is PD (see §6). Like BDT, this is the kind of question that interests mathematicians, i.e. it is GMS. But its proof relies on IW, suggesting that it is non-GMP.

4 Isaacson & Arithmetic

The question then is whether these classifications coincide with some explanatory property which is informative about the alleged genuinely mathematical incompletenesses. To begin, we limit our discussion to arithmetic, where Isaacson argues for an understanding of '(genuinely) arithmetical' which excludes the putative arithmetical incompletenesses [19].

Isaacson conceives of arithmetic as a stable "natural type within the space of mathematical knowledge" [20]. With Isaacson, I take it that PA naturally captures this conception, as the weakest categorical first-order analysis of natural number ([19] §2), but this does not seem essential for an Isaacson's account.[15] The more general thought is that only proofs which ultimately rely exclusively on *arithmetical grounds* are arithmetical. Arithmetic grounding is roughly an (epistemic) *direct grasping* of the truth of a sentence based only on the articulation of our understanding of the natural numbers ([19] §5). So, for the language of arithmetic, $\mathcal{L}_{\mathrm{A}}$:

Isaacson's Hypothesis. *$\phi \in Sent(\mathcal{L}_{\mathrm{A}})$ is genuinely arithmetical iff the proof of ϕ is (ultimately) justified by truths directly grasped from our understanding of the natural numbers.*

Like GMP, this classification is sociological, based on *our* grasp of the natural number structure. Examples seem to me to lend this account credibility. Our understanding of this structure *is* such that we immediately see that e.g. the proof that there are infinitely many primes is arithmetically grounded, but it gives us no direct evidence for $G_{(\mathrm{PA},c)}$.

[15] The inspiration seems to be Tait's account of PRA and finitistic mathematics ([20] §VII).

He then argues that all PA-undecidable $\mathcal{L}_{\mathrm{A}}$-sentences contain "hidden higher order content", so they are "not about natural numbers, but 'about' the statements themselves" ([19] §5). They are *only* $\mathcal{L}_{\mathrm{PA}_1}$-expressible through unnatural codings ([19] §6).

Isaacson's arguments for the exclusion of putative examples are heterogeneous. The ε_0-proof of Goodstein's Theorem is excluded since to see the truth of ε_0-induction requires second-order non-arithmetical insight into e.g. well-ordering ([19] §5), despite being $\mathcal{L}_{\mathrm{PA}_1}$-expressible. These grounds are non-arithmetical, hence disqualifying this proof from arithmeticality. Likewise, $G_{(\mathrm{PA},c)}$ has no directly comprehensible arithmetical content, such content can only be assigned through coding.

The problematic propositions seemingly have *no* arithmetical grounds, so are not *arithmetical* truths, and are not arithmetical incompletenesses of PA. So Isaacson can maintain that PA is complete with respect to arithmetical statements [28].

5 Horsten & ZFC

Given the credibility of Isaacson's account of arithmetic, one might hope to provide a more general account which describes mathematics *tout court*. Perhaps some system captures mathematics, and the putative incompletenesses also use 'clever coding', with no *mathematical* grounds, so are not genuinely mathematical. Enter Horsten, who explicitly generalises Isaacson, arguing that the 'mathematical truths' are exactly the ZFC theorems [16]. I contend that this account is unsustainable. There are specific issues with Horsten's argument and a general disanalogy between number theory and set theory. This suggests that GMP grounds do not coincide with ZFC proofs, i.e. ZFC does not capture genuine mathematics.

ZFC is an obvious candidate for such a system. It is commonly described as a foundation for mathematics ([24] §I.4), and there are easy and canonical ZFC-formalisations for much mathematics. Horsten also quotes the mathematician Shelah, who professes to interpreting 'true' as 'provable in ZFC' ([36], xi), as direct naturalistic evidence in support of his position.

Horsten argues indirectly that ZFC captures mathematics, by rejecting any *alternative* system ([16] §4). Systems stronger than ZFC are rejected on the naturalistic grounds that mathematicians value proofs in extensions of ZFC *less* than 'outright' ZFC-proofs ([16] §5.1). So, certain problems which can be expressed in mathematical language (e.g. CH, Suslin's Hypothesis [18]:16) are not questions of mathematics proper. On pain of inconsistency, Horsten cannot *definitively* exclude notions that cannot be formalised in ZFC ([16] §4.3). Instead, he gestures towards heterogeneous explications of each. The most developed is a brief comment excluding category theory as there is "too much disagreements among constructivists today for us to see a solid body of mathematics there" ([16] §5.1).

Against the plausibility of weaker systems, he suggests that ZFC most accurately describes the amount of set theory necessary to capture the "general principles" of mathematics (rather than a specification of the set-theoretic hierarchy). Further, he claims that results requiring Replacement, e.g. BDT, justify ZFC over Z ([16] fn. 38, [11]:339).

His explanation depends upon a distinction between *mathematical* and *philosophical* grounds. He persuasively argues that the proofs of, e.g., $G_{(\mathrm{ZFC},c)}$ necessarily involves philosophical grounds ([16] §4.1). Such proofs are not ZFC-formalisable. To formalise these, the (informal) language of mathematics, $\mathcal{L}_{\mathrm{M}}$, is extended by a truth predicate, and we add (e.g.) Tarski's truth axioms. But then mathematicians' intuitions about the set-theoretic universe are too weak to "produce the conviction that all the axioms of ZFC are true" in this structure ([16] §4.1). Moreover, the mere *notion* of a truth predicate is inherently philosophical and non-mathematical, hence classifying $G_{(\mathrm{ZFC},c)}$ as not genuinely mathematical. This accords with our explanation for why $G_{(S,c)}$ is non-GMP.

5.1 Specific Issues

Horsten's account is untenable (though the distinction between philosophical and mathematical grounds is credible). Firstly, it does not have the intended naturalistic support [17]. He justifies ZFC as his candidate on the grounds that it formalises much mathematics, but almost all mathematics is ZC-formalisable as well as ZFC-formalisable. Horsten's inclusion of

Replacement is suspect, given its dearth of use [6]. The regressive justification for Replacement is also limited, as (1) certain versions of ZC interprets the ordinals (in particular, Potter's ZC [33] §12), and (2) we might have reason to think that BDT is not a sentence of genuine mathematics. More generally, mathematics happens in an *informal* system, never mentioning any set-theoretic axioms at all.

A concrete difference between the two accounts is the treatment of putative counterexamples in Isaacson and Horsten. To exclude a proof in an extension of ZFC, Horsten must "show that the mathematical community does not accept a certain statement as true" (Incurvati [17]). Incurvati points out that this is distinctively different and harder than the equivalent for Isaacson, who need only show that "to see the truth of [PA-undecidable $\mathcal{L}_{\mathrm{A}}$-sentences], we need to employ concepts that go beyond our basic grasp of the structure of the natural numbers", which could still be *mathematical* structures.

This suggests a vicious circularity. Isaacson is in a position to distinguish truths established by looking to "some more comprehensive mathematical theory" ([19] §7), and those using only arithmetic, i.e. he classifies proofs as arithmetical-mathematical vs. non-arithmetical-mathematical. However, Horsten seemingly wants to characterise 'mathematical' by generalising a pre-existing characterisation *which already includes the term 'mathematical'*, in his distinction between mathematical and philosophical grounds. The analogous claim for Horsten would be to distinguish ZFC-mathematical truths from non-ZFC truths established in a more comprehensive mathematical theory. But this does not explain mathematicality. In general, it is dubious to suppose that, given that there are models of ZFC, discussion of those models would be necessarily non-mathematical.

Moreover, the arguments for the exclusion of ZFC-uninterpretable notions are unconvincing. For one, his 'disagreement' argument against category theory is simply false. Mac Lane & Moerdijk's topos-theoretic interpretation of ZC [27], and its plausible enlargements to ZFC ([38] §2.3.10), does all the foundational work asked of Z(F)C and more besides, so could just as easily work as a foundation for mathematics. Further, categorically-founded algebra is reliant on category theory, using techniques which are increasingly remote from ZFC [39], suggesting that ZFC *is* insufficient as an analysis of mathematicality. There are ordinary bits of mathematics (which are increasingly standard) which are only captured with categorial theories, not with ZFC.

On a similar note is the issue of randomness. Randomness is simply not ZFC-interpretable [25], but is apparently a paradigm notion of mathematics. Horsten's only possible response is to deny that (genuine) randomness is a part of mathematics proper. However, Horsten presents *no* such argument against the mathematicality of randomness, nor does this seem plausible. There's a sense in which this is expected: if we think of ZFC as fundamentally capturing a notion of *set*, rather than being deliberately designed as a foundation for all mathematics, then it is unsurprising that not all mathematical notions can be captured within it.

5.2 Disanalogy

Horsten alludes to one disanalogy between Isaacson's account and his own, that adopting new set-theoretic axioms is more palatable than adopting new (first-order) arithmetical axioms ([16] §4.4). But the differences run deeper, as ZFC has different intuitive reactions and independent dimensions of indeterminacy.

In number theory, the hunch is that for a suitable $\mathcal{L}_{\mathrm{PA}}$-sentence, *whatever* works is a GMP proof. For example, a hypothetical proof of Goldbach's Conjecture using $\mathrm{PA} + \mathrm{Con}_{\mathrm{PA}}$ would seem to be GMP, just as we take Wiles' non-PA proof of Fermat's Last Theorem to be both GMP and to be an outright proof of the theorem.

But assuming $\mathrm{PA} + \mathrm{Con}_{\mathrm{PA}}$ is fairly tame, indeed this is Z-provable. When we consider *set-theoretic* additional assumptions, the intuitions change. Suppose instead Goldbach's Conjecture was proved using Z(FC) + IW. Here, the astonishment at the use of large cardinals and strength of the intuition that IW is a set-theoretic claim would overwhelm the 'whatever works' approach to grounds. It seems we would describe this as a proof that "*IW* $\rightarrow$ *Goldbach's Conjecture*", rather than a proof of Goldbach's Conjecture. These *conditional* proofs of *conditional* theorems might well be GMP and GMS respectively, entirely independent of whether the outright proof and theorems are accepted as mathematical. It seems that arithmetical assumptions just causes different intuitions to high consistency strength mathematical assumption.

Intuitions on conditional proofs might differ. A more theoretically grounded disanalogy is that there are two independent *kinds* of indeterminacy of Z(FC) [9].

1. Height: almost any number of levels in the set-theoretic hierarchy is consistent with Z(FC).
2. Width: above V_ω, (for CH), Z(FC) does not determine the richness of its levels.[16]

These kinds of indeterminacy seem necessarily distinct. A clear way to see this is that certain notions of maximisation make height maximisation inconsistent with width maximisation ([18]:172). They are different beasts, and so we should not generally expect properties of one to apply to the other.

Meanwhile, roughly speaking, PA is only indeterminately high. An analogue of Isaacson's argument classifies large cardinal (i.e. height) techniques as non-GMP, but an independent argument is necessary to show that this analogy applies to techniques specifying the width of the universe. Without this, the project of generalising from arithmetic to mathematics *tout court* may be unworkable. As it stands, it might well be that axioms which controlling the *width* of the hierarchy all lead to GMP proofs.

These problems with Horsten's account and the issues in arguing analogously from arithmetic to mathematics suggest that to the best of our knowledge there is no reason to maintain that either inclusion between GMP and ZFC, i.e. ZFC does not capture genuine mathematics.

6 Incompletenesses Again

If Isaacson's argument for the completeness of arithmetic cannot be generalised, one might worry that our putative examples demonstrate GMP

[16]This indeterminacy represents the different ways we might understand "subset" ([18]:162, 167). See §6.1.

incompleteness. However, I maintain that these examples are either unproblematic, or they suffer from under-determination, rather than incompleteness. This is justified by analysing certain concepts in mathematics as under-determined.

(Non-foundational) pure mathematicians would find the Z-proof of Goodstein's Theorem distantly recognisable, and conceivably interesting. Perhaps this is because Z is widely accepted, and the statement involves only elements which every mathematician is accustomed to, e.g. $\mathbb{N}^{\mathbb{N}}$, basic operations. So it passes the Reaction test.

However, this seems unlike the Gödel sentence case. $G_{(\mathrm{PA},c)}$ is non-arithmetical, but it also seems as though proofs of $G_{(S,c)}$ are non-GMP. The standard proof relies on notions like truth predicates, or truth-like predicates. These seem to go outside the bounds of mathematics.

Moreover, $G_{(S,c)}$ is unsurveyable, and reaction-wise, "many mathematicians' gut feeling is that they are *not* mathematical truths" for they lack "concrete mathematical content" (Horsten, [16] §3, fn 8.). Hence it is presumably non-GMS. But one natural hypothesis is that *any* justification for a non-GMS sentence must involve non-GMP grounds, which disqualifies $G_{(S,c)}$ ([17] §3).

One might suggest that the discussion here is too focused the standard form of $G_{(S,c)}$, and that the Gödelain sentences can be converted into other forms which are much more likely to have genuinely mathematical content. Perhaps the strongest candidate is the transformation of Gödel's theorem into a proof that there is an undecidable Diophantine, $S(\bar{n})$, the MRDP theorem. $S(\bar{n})$ states that there is an integer solution to an integer coefficient equation in 11 variables ([5]:51), and is clearly surveyable.[17] Moreover, many Diophantine equations and their solutions are of interest to mathematics at large, indeed being the topic of Hilbert's tenth problem ([10]:70), suggesting that this might have mathematical content.

Nevertheless, it seems that there are some reasons to think that this particular Diophantine sentence might not be GMS. The authors write that the "Although mathematical content is not so readily understood, the sentence has a simple arithmetic form" (([5]:51). It is prima facie odd for the

[17]Many thanks to an anonymous reviewer for pointing this out.

content of a GMS sentence to not wear its content openly, but this is isn't a serious issue. Instead, we might doubt its mathematicality for the reason that we doubted that of the standard $G_{(S,c)}$, namely that $G_{(S,c)}$ says that some otherwise unremarkable primitive recursive function has a root. A proof to this effect would not interest the mathematical community. So too, we might doubt whether solving this otherwise unremarkable Diophantine equation would interest the mathematicians at large. Worse still, we don't know which Diophantine $S(\bar{n})$ is ([5]:54)! In summary, just as one might doubt that all statements of the existence of roots of primitive recursive functions are GMS, we might also doubt that all statements of the existence of solutions to Diophantines are GMS.

In both the $S(\bar{n})$ and $G_{(S,c)}$ cases, it is the metatheoretic properties of the sentences that are of interest: their surface mathematical content is limited. Moreover, the proof of the metatheoretic properties of $S(\bar{n})$ uses unorthodox methods ([5]:53), like axiomatic systems and proof theory, which may disqualify it from GMP. This suggests that any doubts about the mathematicality of $G_{(S,c)}$ are not artefacts of its form, instead other formulations might also be disqualified.

The crux is determinacy. BDT is a contender for GMP and GMS ([33] §13.2), but its status is somewhat controversial, as our standard proof relies on repeated uses of Replacement, which is itself dubious. Even if one accepts Replacement, the same problem occurs for the further generalisations, PD and AD. We noted above that these are ZFC-undecidable and ZF-undecidable respectively ([29] §6). More significantly, their non-ZFC proofs fail the Reaction test. These proofs would seem to leave our mathematician gawking. They are simply too set-theoretic and abstract to be mathetical(ly interesting). The combination of these widespread reactions (though not generally sufficient) and the use of IW in the proofs together give evidence to suggest that these proofs are non-GMP. However, AD and PD seem GMS. So, they are seemingly GMS sentences with non-GMP proofs, ostensibly demonstrating genuinely mathematical incompleteness.

6.1 Under-Determined Terms

I contend that PD and AD are in fact under-determined. They feature an under-determined instance of the term 'game', which disqualifies them from genuine mathematics, which suggests that we should not expect them to have, e.g., a determinate truth value. This can be made clear by considering the similarity of the following two questions:

1. Are all sets well-founded?
2. Are all (projective) games determined?

I take it that it is unsustainable to think that there is a single structural conception of set. Instead, there is a common core, or 'partial conception' [9] with competing, *non-integrable*, legitimate precisifications [37], e.g. well-founded or non-well-founded.

More specifically, Feferman argues that the notion of an 'arbitrary set', notably 'arbitrary subset of $\mathbb{N}$', is under-determined and unsharpenable.[18] The issue is the breadth of competing conceptions of arbitrary subsets, including but not limited to $S(\mathbb{N})$, the set of all subsets of N, $2^{\mathbb{N}}$ the set of all binary sequences of N ([9] §2*d*., [8]).[19] This seems convincing to me: we simply *do* change between different notions of a subset of the natural numbers, coming from all kinds of different concrete investigations and historical contexts. The ambiguity here seems worse that Feferman details, as he glosses $2^{\mathbb{N}}$ as both the binary $\mathbb{N}$-sequences *and* the binary $\mathbb{N}$-trees. Prima facie, we may have different and competing ideas of how these work.

Further ambiguity derives from competing conceptions of the continuum, Feferman's focus. For example, two geometric conceptions, the Hilbertian and the Euclidean, differ wildly in their ontology (the Euclidean line being something beyond the collection of points). The ambiguity in the concept of the continuum only adds to the ambiguity in the notion of an arbitrary subset of the naturals, as the later is meant to subsume the former: we *do* sometimes interpret reals as collections of rationals.

[18] The general notion of a "subset" also requires unpacking ([18]:162), but the specific case of subsets of naturals is the principal concern, as this is the lowest set-rank level where notions of subset diverge.

[19] See also his argument concerning platonism [7].

The next step in the argument is the claim that we cannot sharpen the notion to an "arbitrary constructible subset of $\mathbb{N}$, or any specific relativization thereof, without violating the idea of an arbitrary subset of a set, independent of any means of definition" (Feferman, [7]). Equating 'arbitrary subset' with 'constructible subset' is clearly unsuitable: the whole point of the concept of an arbitrary subset is that it *does not* make claims about constructability. Such an identity requires a serious and controversial philosophical commitment, which we cannot simply stipulate. Nor can we sharpen the notion of an arbitrary subset by equating it to a subset in one *kind* of visualisation (e.g. as subtrees), this again would not be an arbitrary subset at all, it would instead be a subset of a particular type.

We do seem to have good grip on the notion 'natural number'. But we do *not* have such grip on 'arbitrary set of natural numbers',[20] on grounds of unsharpenable ambiguity of use and conception. So we are not in a position to measure $\mathcal{P}(\mathbb{N})$. Hence, he argues that CH is indeterminate and 'inherently vague' [9] [8].

I submit that (2.) would then be similar: we would not have a firm enough grasp on an arbitrary (projective) game to ask whether it is determined, it would depend on the salient notion of 'game'. A (projective) game correspond to $W \subseteq \mathbb{N}^{\mathbb{N}}$.[21] If we have insufficient grip on arbitrary subsets of $\mathbb{N}$, then we cannot have good grip on arbitrary subsets of sequences of $\mathbb{N}$ (as e.g. the constant sequences would then be subsets of $\mathbb{N}$). *If* Feferman's arguments for the under-determination of CH hold, they also show that 'arbitrary game' is under-determined. However, there still can be a common, agreed-upon *core* to what constitutes a game, as with a set.

We can now see how this affects the three key examples, AD, PD, and BDT, by considering the parallel questions again. The well-founded-ness of a set theory somehow *identifies* which set theory we are considering. I submit that the same would hold for AD: if all games *are* determined, this is a fact about our theory of games, not games themselves.[22] The general

[20]This mirrors the account of Dummett in [4].

[21]These sequences are more than labels: we manipulate the number structure as a stand-in for the game-play topology.

[22]Choice provides a ZF proof of $\neg$AD. However, Choice is to at least some extent con-

notion of game or set does not determine a truth value for either of the questions.

Next, we ask what happens when *restrict* the class of games. When we restrict our class of *sets* to only those which are hereditarily built from the empty set, we can prove that these sets *are* well-founded. In the same way, we should expect that certain subclasses of games can be proved determinate, despite an under-determined notion of 'game'. Even if one doubts Replacement, 'game' seems clear enough to prove the determinacy of a relatively restricted class of games, e.g. there are GMP proofs that open and closed subsets of Baire spaces are determined [13].[23] On the other hand, it seems that there is good evidence to think that projective sets are sufficiently complex to mean that they will vary a lot based on precisifications of the notion of 'game' (which in turn will depend on precisifications of 'set'). This suggests to me that PD is susceptible to the kind of variation that would lead to its under-determination. The open question is whether the notion of game is sufficiently clear to guarantee the determinacy of Borel sets, for BDT.

6.2 Stability

I maintain that under-determination is essentially different from GMP incompleteness. Under-determination is roughly a non-categoricity result. This is to say that there are simply different models of 'set' or 'game' which verify the mathematical common core, but which are not elementarily equivalent. But proofs giving differing truth values cannot be GMP: one of the widest held intuitions about mathematics is that it is characteristically unanimous. Mathematical disagreement shows that some party is in error. There is no space for genuine relativity.

Formally, this justifies a supervaluationist test for a particular *way* to be non-GMP (and there do seem to be other ways to be non-GMP, for example being an argument clearly outside of the domain of mathematics, e.g. in

troversial (though e.g. countable Choice may not be). One explanation is that some precisified conceptions of set verify Choice, but not all. Without the presupposition of such a precisified notion of Choice, AD remains indeterminate.

[23]Thanks to an anonymous reviewer for highlighting this.

history):

Test (Stability Test)**.** *If proofs of $\phi \in Sent(\mathcal{L}_M)$ yield different truth values across (suitable) interpretations of under-determined terms, then:*
(i) the grounds for ϕ are non-GMP,
(ii) ϕ is under-determined.

Such sentences might potentially be GMS, the test is cautiously mute on this. A sentence is not immediately 'seen' to be under-determined, this requires careful analysis of its proofs. Notably, uncontroversial set-theoretic generics (e.g. "unions commute") are not under-determined, as their truth values are stable over interpretations.

PD has a differing truth value for suitable interpretations of 'set', i.e. whether 'set' applies to particular large cardinal pseudo-sets, so is under-determined. Dependent on the status of Choice [9], AD is either under-determined or false. If all acceptable models witness Replacement, then the ZF-proof of BDT would be GMP, but this is controversial.

One possible analysis is that questions like (1.) and (2.) are (equivalently) questions about the usage of mathematics terms, not *in* mathematics. Take the question "what precisification of 'game' is salient?". This mention of the term 'game' seems *meta*-mathematical. It seems to be *what kind of object* we are discussing, rather than one within the discourse of mathematics. A parallel might be with when a physical system is analysed, in which case we might ask "what model best approximates the system?". This does not concern *the properties* of the model once it is decided, it instead tries to *fix* which object (i.e. model) is being discussed. Such questions would be considered outside of mathematics, perhaps in the domain of natural science. Worse still, in the case of 'game', the facts needed to fix which kind of game we are considering seem equivalent to questions about what kind of mathematical universe is in play. But these facts about the mathematical universe are seemingly outside of the domain of mathematics.

Either way, proofs can be stably classified as non-GMP due to their inclusion of an under-determined predicate. This motivates a program of sharpening the under-determined notions with the intent to deliver a GMP proof which removes incompleteness, e.g. "interpreting 'game' by *these kinds* of sets (which model IW) implies PD". A rich source of work would be to consider exactly the status of projective games, or where the cut-off point falls in the projective hierarchy for games, especially whether the indeterminacy in $\mathbb{N}^{\mathbb{N}}$ can be explicitly related to the indeterminacy of the notion of a projective game, and also whether PD is determined by a suitable sharpening of the notion of 'game'. This would, in turn, require some outline as to exactly what we take to be a *suitable* precisification.

In short, we have exchanged some incompleteness for under-determination. "['Incompleteness'] suggests that the formal system in question fails to offer a deduction which it ought to" (Isaacson [19] §8). But, as these sentences are under-determined, we have no obligation to prove them. Even though we cannot capture GMP with ZFC, to remove incompleteness *à la* Isaacson, we can still exclude some instances of putative incompleteness due to the lack of precision of general mathematical terms.

References

[1] APPEL, K.; HAKEN, W. 'Every Planar Map is Four-Colorable', *Contemporary Mathematics*, 98, 1989.

[2] BOOLOS, G. 'A Curious Inference', *J. Phil. Logic*, 16, 1987.

[3] BUTTON, T.; WALSH, S. *Philosophy and Model Theory* (OUP), 2018.

[4] DUMMETT, M. 'The Philosophical Significance of Gödel's Theorem', *Ratio*, 5(2):140, 1963.

[5] DYSON, V.H.; JONES, J.P.; SHEPHERDSON, J.C. 'Some Diophantine Forms of Gödel's Theorem', *Archive for Mathematical Logic*, 22(1):51-60, 1980.

[6] ELWES, R. *Concrete Incompleteness 1*, retrieved 2018.
`richardelwes.co.uk/2010/10/21/concrete-incompleteness-1`

[7] FEFERMAN, S., 'My Responses', in Hendricks & Leitgeb (eds.), *Philosophy of Mathematics: 5 Questions* (Copenhagen: Automatic Press/VIP), 2007.

[8] ———, *The Continuum Hypothesis is neither a definite mathematical problem not a definite logical problem*, 2011.
`math.stanford.edu/~feferman/papers/CH_is_Indefinite.pdf`

[9] ———, *Is the Continuum Hypothesis a definite mathematical problem?*, 2011.
`logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf`

[10] FRANZÉN, T., *Gödel's Theorem. An Incomplete Guide to its Use and Abuse* (A K Peters), 2005.

[11] FRIEDMAN, H.M. 'Higher Set Theory and Mathematical Practice', *Annals of Mathematical Logic*, 2(3), 325-357, 1970.

[12] ———, *Adventures in Incompleteness*, 2017.
`cpb-us-w2.wpmucdn.com/u.osu.edu/dist/1/1952/files/2014/01/AdvIncomp03011712pt-1k0z6to.pdf`

[13] GALE, D.; STEWART, F.M. 'Infinite Games with Perfect Information' *Contributions to the Theory of Games*, 2:245-66, 1952.

[14] GÖDEL, K., Über Formal Unentscheidbare Sätze der Principia Mathematica und Verwandter Systeme I, *Monatshefte für Mathematik und Physik*, 38:1, 173–198, 1931. Translated in Gödel 1986.

[15] GOODMAN, M. J. K. *An Introduction to the Early Development of Mathematics* (Hoboken, New Jersey : Wiley), 2016.

[16] HORSTEN, L. 'Platonic Formalism', *Erkenntnis*, 54, 2001.

[17] INCURVATI, L. 'Too Naturalist and Not Naturalist Enough: Reply to Horsten', *Erkenntnis*, 69, 2008.

[18] ———, 'Maximality Principles in Set Theory', *Philosophia Mathematica*, 25(2):159-93, 2017.

[19] ISAACSON, D. 'Arithmetical Truth and Hidden Higher-Order Concepts', *Logic Colloquium '85* (Amsterdam: North-Holland Publishing Company), 147-169, 1987.

[20] ———, 'Some Considerations on Arithmetical Truth and the ω-rule', in Detlefsen (ed.), *Proof, Logic and Formalisation* (London: Routledge), 1992.

[21] KAVEH *Is the Riemann Hypothesis equivalent to a Π_1 sentence?*, retrieved 2016. `mathoverflow.net/q/31846`

[22] KECHRIS, A. S. 'The Axiom of Determinacy Implies Dependent Choices in $L(\mathbb{R})$', *J. Symb. Log.*, 49(1), 161-173, 1984.

[23] KIRBY, L.; PARIS, J. 'Accessible Independence Results for Peano Arithmetic', *Bull. London Math. Soc.*, 14, 1982.

[24] KUNEN, K. *The Foundations of Mathematics* (London: College Publications), 2007.

[25] VAN LAMBALGEN, M. 'Independence, randomness and the axiom of choice', *The Journal of Symbolic Logic*, 57(4):1274-304, 1992.

[26] LEE, S. *The Undecidable Gödel Sentence*, retrieved 2016. `tachyos.org/godel/Godel_statement.html?=$tag?%3E`

[27] MAC LANE, S.; MOERDIJK, I. *Sheaves in Geometry and Logic* (Springer-Verlag New York), 1992.

[28] MOON, S. *Isaacson's Thesis and Wilkie's Theorem*, UvA Thesis, 2017.

[29] MOSCHOVAKIS, Y.N. *Descriptive Set Theory* (Amsterdam: North-Holland Publishing Company), 1980.

[30] NEEMAN, I. 'Determinacy in $L(\mathbb{R})$', in Foreman & Kanamori (eds.), *Handbook of Set Theory* (Amsterdam: Springer Netherlands), 2007.

[31] PASEAU, A. 'Naturalism in Mathematics and the Authority of Philosophy', *BJPS*, 56, 2005.

[32] ———, 'Knowledge of Mathematics without Proof', *BJPS*, 66 (4:1), 775–799, 2015.

[33] POTTER, M. *Set Theory and its Philosophy* (OUP), 2004.

[34] RODYCH, V. 'Wittgenstein's Philosophy of Mathematics', in Zalta (ed.), *SEP*, 2018. `plato.stanford.edu/entries/wittgenstein-mathematics`

[35] DU SAUTOY, M. *What We Cannot Know* (London: Fourth Estate), 2016.

[36] SHELAH, S. *Cardinal Arithmetic* (OUP), 1994.

[37] WONTNER, N.J.H. *Legitimate Conceptions of Set*, unpublished manuscript, 2017.

[38] ———, *Non-Set Theoretic Foundations of Concrete Mathematics*, MMath-Phil Thesis, 2018.

[39] ZHANG, S. *Most striking applications of category theory?*, 2017. `mathoverflow.net/q/19401`