A Chalet on Mount Everest: Interpretations of Wittgenstein's Remarks on Gödel

MSc Thesis *(Afstudeerscriptie)*

written by

Ásgeir Berg Matthíasson
(born November 29th, 1985 in Reykjavík, Iceland)

under the supervision of **Prof Martin Stokhof**, and submitted to the Board of Examiners in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam.*

Date of the public defense:	**Members of the Thesis Committee:**
October 2nd, 2013	Prof Benedikt Löwe
	Prof Jeroen Groenendijk
	Prof Frank Veltman

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis offers a critical overview of the debate on Wittgenstein's remarks on Gödel. These remarks—which have since their publication been the source of much controversy—were said by early commentators to show that (a) Wittgenstein did not understand the role consistency plays in the proof of Gödel's theorem and logic in general, and that (b) he mistakenly thought that a natural language interpretation of the Gödel sentence was necessary to establish the incompleteness of *Principia Mathematica*.

Later commentators have been more sympathetic to Wittgenstein's remarks and have with few exceptions denied both claims. We examine their interpretations and evaluate their cogency and exegetical value.

Acknowledgements

First of all I would like to thank my supervisor, Martin Stokhof, for his insightful comments and discussions on the topic of this thesis. I would also like to thank him for his vital support and guidance throughout my career at the ILLC, and for nurturing my budding interest in Wittgenstein.

I would also like to express my gratitude to Mikael M. Karlsson, my old supervisor at the University of Iceland, and Erlendur Jónsson, professor at the same institution, for their encouragement and help. Without them it is unlikely that I ever would have pursued a degree in logic.

Finally, I would like to thank my friends in Amsterdam and elsewhere for their constant companionship and support over the past two years, as well as my family. They know who they are. I am however especially indebted to Julia Ilin, Fenner Tanswell, Nathaniel Forde, Will Frager, Ryan Nefdt, Arngrímur Vídalín, Guðmundur D. Haraldsson, Siggeir F. Ævarson, Maja Jaakson, Vésteinn Snæbjarnarson, Vlasta Sikimić and Callum Sida-Murray for giving me a helping hand in writing this thesis, each in their own way.

“It is as if someone had extracted from certain principles about natural forms and architectural style the idea that on Mount Everest, where no one can live, there belonged a châlet in the Baroque style.”

—Ludwig Wittgenstein

Contents

Abbreviations of works by L. Wittgenstein

AWL *Wittgenstein's Lectures: Cambridge, 1932–1935*, ed. A. Ambrose, Chigaco University Press, Chicago, 1982.

BB *The Blue and Brown Books*, Basil Blackwell, Oxford, 1958.

LFM *Wittgenstein's Lectures on the Foundations of Mathematics: Cambridge, 1939*, ed. C. Diamond, Cornell University Press, Ithaca, 1976.

LWL *Wittgenstein's Lectures: Cambridge, 1930–1932*, ed. D. Lee, Chigaco University Press, Chicago, 1982.

NB *Notebooks 1914–1916*, eds. G.H. von Wright and G.E.M. Anscombe, trans. G.E.M. Anscombe, 2nd edition, Blackwell Publishing, Oxford, 1979.

PG *Philosophical Grammar*, ed. R. Rhees, trans. A. Kenny, Basil Blackwell, Oxford, 1974.

PI *Philosophical Investigations*, trans. G.E.M.Anscombe, 2nd edition, Basil Blackwell, Oxford, 1958.

PR *Philosophical Remarks*, ed. R. Rhees, trans. R. Hargreaves and R. White, Basil Blackwell, Oxford, 1975.

RFM *Remarks on the Foundations of Mathematics*, eds. G. H. von Wright, R. Rhees and G. E. M. Anscombe, trans. G. E. M. Anscombe, 3rd edition, Basil Blackwell, Oxford, 1978.

TLP *Tractatus Logico–Philosophicus*, trans. C.K. Odgen, Routledge & Kegan Paul, Ltd., London, 1922. English translation reprinted with corrections 1933.

WVC *Wittgenstein and the Vienna Circle: Conversations Recorded by Friderich Waismann*, ed. B.F. McGuinness, trans. J. Schulte and B.F. McGuinness, Basil Blackwell, Oxford, 1979.

Introduction

In his much praised biography of Ludwig Wittgenstein, Ray Monk recounts the story of how, in 1944, Wittgenstein was given the opportunity to make corrections to a short paragraph which was intended for publication in a philosophical dictionary. The only correction he insisted on was the inclusion of the following remark:[1]

> Wittgenstein's chief contribution has been in the philosophy of mathematics.

If this story is true, it certainly is striking that Wittgenstein's own view of his work was so radically out of step with how subsequent philosophers have come to see it, as his philosophy of mathematics is, again in the words of Monk, "that most neglected and maligned aspect of his work".[2]

This indifferent, if not cold reception of Wittgenstein's philosophy of mathematics can be partly explained, without a doubt, by the form in which it has reached us. Much of what he wrote on the subject—in what is often called the later period—was published in 1956 as the *Remarks on the Foundations of Mathematics*. The volume is a selection of remarks taken from many different manuscripts and typescripts, some of which had not been prepared for publication by its author at all. The result is a work which is very difficult to read and even more difficult to interpret. The following passage, taken from Michael Dummett's influential review of the *Remarks*, expresses what is probably the most widely held view:

> Many of the thoughts are expressed in a manner which the author recognised as inaccurate or obscure; some passages contradict others; some are quite inconclusive; some raise objections to ideas which Wittgenstein held or had held which are not themselves stated clearly in the volume; other passages again, particularly those on consistency and on Gödel's theorem, are of poor quality or contain definite errors.[3]

And indeed, it is precisely those remarks on Gödel's theorem and consistency which have raised the most ire of mathematically minded philosophers and logicians, especially the remarks contained in Appendix III to Part I. One prominent example, and whose reaction is typical, is Gödel himself, who wrote in a letter to Menger:

[1] See Monk 1991, p. 466.

[2] Monk 2007, p. 269. This paper contains a good overview of how Wittgenstein's philosophy of mathematics has been received.

[3] Dummett 1959, p. 324.

> As far as my theorem about undecidable propositions is concerned it is indeed clear from the passages you cite that Wittgenstein did not understand it (or pretended not to understand it). He interprets it as a kind of logical paradox, while in fact it is just the opposite, namely a mathematical theorem within an absolutely uncontroversial part of mathematics (finitary number theory or combinatorics). Incidentally, the whole passage you cite seems nonsense to me. See, e.g. the 'superstitious fear of mathematicians of contradictions'.[4]

Other early reviewers of the *Remarks on the Foundations of Mathematics* expressed a similar view. According to them, Wittgenstein's remarks on Gödel were e.g. said to "throw no light on Gödel's work", that the arguments found therein are "wild", that Wittgenstein confused "truth with provability", and perhaps most commonly, that Wittgenstein did not understand the role which the premise of consistency played in Gödel's proof. The early consensus seems to have been that Wittgenstein misunderstood Gödel's proof, and mistakenly thought that he could refute his results by simply restating the view that provability and truth in the system of *Principia Mathematica* are synonymous notions.

After these initial reviews, Wittgenstein's remarks on Gödel received very little attention in the literature, until Stuart Shanker published a long and sympathetic treatment of his remarks,[5] putting them in the context of Wittgenstein's earlier philosophy of mathematics and his opposition to Hilbert's programme, and since 1995, many different interpretations have been published, and in this century the debate has been quite lively. These interpretations, despite perhaps having in common a certain sympathy for Wittgenstein's cause, are wildly different and offer at times completely opposite readings of his remarks.

The aim of this thesis therefore is to give a critical overview of this relatively new debate with an eye to examine—on the one hand—whether or not the harsh verdict of the early commentators is justified and if this movement in the opposite direction has any merit, on the other. We shall start by giving a short introduction to the early debate and what Wittgenstein wrote on Gödel's theorem, and then proceed to critically examine each of the major interpretations that have been offered, starting with Shanker's 1988 paper, *Wittgenstein's Remarks on the Significance of Gödel's Theorem*—the first paper to offer a sympathetic interpretation of Wittgenstein's remarks, and ending with the very interesting interpretation Francesco Berto gives in his book, *There's Something About Gödel*, published in 2009, which gives a dialetheist viewpoint.[6]

[4] The letter appears in (Wang 1987, p. 49).
[5] See (Shanker 1988).
[6] See (Berto 2009b).

Other interpretations considered will be Juliet Floyd's interpretation, which draws a connection between Wittgenstein's remarks on Gödel and the discussion on meaning and understanding in the *Philosophical Investigations*, then another, quite different, interpretation by Floyd—put forth in a paper co-authored with Hilary Putnam—which sees Wittgenstein as making a highly technical point about non-standard models of arithmetic, and finally Victor Rodych's interpretation, who, despite being quite sympathetic to Wittgenstein, is in agreement with the early commentators on the subject of Wittgenstein's alleged mistake.

Unfortunately, for time and space constraints, we must presuppose some familiarity both with Wittgenstein's work and Gödel's proof. The appendix of this thesis contains the text of (RFM I, App. III) and the reader is highly encouraged to give it at least a quick skim before proceeding to the main body of the text.

The "Notorious Paragraph" and early commentators

Early discussion of Wittgenstein's remarks on Gödel were mostly focused on Wittgenstein's remarks on the role of consistency in Gödel's proof and his perceived criticism of the correct interpretation of the Gödel sentence. In this chapter we will give a short introduction to this debate, as it will be beneficial to know the context in which the later, more positive interpretations were written.

Most of the attention of the early commentators was focused on what Juliet Floyd has called "the notorious paragraph"[7] and the remarks surrounding it, where Wittgenstein discusses these issues, particularly the role of consistency and contradiction in Gödel's proof. The following is the remark in its entirety:

> I imagine someone asking my advice; he says: "I have constructed a proposition (I will use 'P' to designate it) in Russell's symbolism, and by means of certain definitions and transformations it can be so interpreted that it says: 'P is not provable in Russell's system'. Must I not say that this proposition on the one hand is true, and on the other hand is unprovable? For suppose it were false; then it is true that it is provable. And that surely cannot be! And if it is proved, then it is proved that it is not provable. Thus it can only be true, but unprovable."
>
> Just as we ask, " 'Provable in what system?", so we must also ask: " 'True in what system?" 'True in Russell's system' means, as was said: proved in Russell's system; and 'false in Russell's system' means: the opposite has been proved in Russell's system.—Now what does your "suppose it is false" mean? *In the Russell sense* it means, 'suppose the opposite is proved in Russell's system'; *if that is your assumption* you will now presumably give up the interpretation that it is unprovable. And by 'this interpretation' I understand the translation into this English sentence.—If you assume that the proposition is provable in Russell's system, that means it is true *in the Russell sense*, and the interpretation "P is not provable" again has to be given up. If you assume that the proposition is true in the Russell sense, *the same* thing follows. Further: if the proposition is supposed to be false in some other than the Russell sense, then it does not contradict this for it to be proved in Russell's system (What is called "losing" in chess may constitute winning in another game.)

Most of the early commentators, as we will see, were not very sympathetic to

[7]See (Floyd 2001, p. 302) and (Floyd and Putnam 2000).

what Wittgenstein was trying to say in this paragraph, and the ones following it, and suppose that Wittgenstein is trying to show that there is some kind of mistake inherent in Gödel's proof, and that this betrays his poor understanding of Gödel's work.

One of the earliest commentators on the *Remarks on the Foundations of Mathematics*, and one who certainly has this view, was G. Kreisel, who had been one of Wittgenstein's students at Cambridge, and described by him as "the most able philosopher he had ever met who was also a mathematician".[8] Kreisel had attended Wittgenstein's lectures on the foundations of mathematics and claims to have had many discussions with him on the topic, but after reading the *Remarks* realised that the issues they discussed "were far from his [Wittgenstein's] interests".[9] In his influential 1958 review of them, he famously stated that the book was a "surprisingly insignificant product of a sparkling mind" and that he had "not enjoyed reading" it.[10]

In view of Kreisel's low opinion of Wittgenstein's philosophy of mathematics, it should then perhaps not be very surprising that his views regarding Wittgenstein's remarks on Gödel are equally damning. He calls Wittgenstein's arguments "wild" and is especially bothered by those in which Wittgenstein discusses consistency and the correct way of interpreting the Gödel sentence (i.e. the notorious paragraph), and claims that Wittgenstein doesn't seem to take seriously Gödel's condition of consistency of the theory of arithmetic in question in order for his proof to work. He says,

> Even if an inconsistency didn't 'matter', one cannot hope to discuss significantly on this basis a result which explicitly supposes consistency of the system.[11]

He then adds that, indeed, given the assumption of consistency, the translation of the Gödel sentence is correct.[12] From his consequent discussion of how the truth of the Gödel sentence is established by semantic considerations and "separated from the question of truth in arithmetic", he seems to be implicitly criticising Wittgenstein for not properly distinguishing between a theory and its metatheory, and simply not understanding Gödel's proof (earlier in the essay he says: "Wittgenstein's views on mathematical logic are not worth much because he knew very little...").[13]

Much later, in 1998, Kreisel claimed that Wittgenstein had asked him in the

[8] Monk 1991, p. 498.
[9] Kreisel 1958, p. 157.
[10] Ibid., p. 158. Surprisingly, Kreisel is much more positive overall in his 1978 review of *The Lectures on the Foundations of Mathematics*. See (Kreisel 1978).
[11] Ibid., p. 153.
[12] Ibid.
[13] Ibid., p. 143.

1940's to explain to him Gödel's proof. According to Kreisel, Wittgenstein had before that never read Gödel's original paper, because he had "been put off by the introduction".[14] However, in 1983, Kreisel told the following story, which was supposed to have happened in the mid-1930's.

> One day after several brief, reasonable explanations of the Gödel incompleteness proof Wittgenstein spoke with complete enthusiasm. Gödel must indeed be an extraordinarily original mathematician, for he had derived from utterly banal arithmetical statements—implicitly [Wittgenstein meant] metamathematical sentences—qualities such as consistency. Gödel had discovered a brand new proof method.[15]

These claims are somewhat inconsistent, and evaluating their truth, especially since they were made so long after the fact, is difficult.

Another early critic who also accuses Wittgenstein of not properly heeding Gödel's assumption of consistency, and thus not understanding the proof, is Paul Bernays. Bernays' review of the *Remarks*, which first appeared in German in 1959, is quite substantial, and despite the author's obvious disagreement with Wittgenstein, also quite sympathetic.[16] According to Bernays, however,

> ...[Wittgenstein's] discussion of Gödel's theorem of non-derivability and its proof, in particular, suffers from the defect that Gödel's quite explicit premiss concerning the consistency of the formal system under consideration is ignored.

He does on the other hand, think that Wittgenstein's comparison in (RFM I, App III, §14) of Gödel's proof and certain impossibility results in mathematics is quite apt. Here Wittgenstein compares Gödel's results to geometric impossibility proofs, and points out that such proofs contain a prediction, since in consequence of such a proof one would stop trying to to certain things in mathematics (for instance, squaring the circle or deriving the Gödel sentence), but he puzzles over Wittgenstein's remark that "a contradiction is unusable as such a prediction" and points out that "such impossibility proofs usually proceed via the derivation of a contradiction".[17] Here, Bernays is of course correct, but possibly misinterprets Wittgenstein's intentions. The remark will be discussed later in the thesis in various chapters (it is for instance a large part of Juliet Floyd's interpretation).

Another early commentator we should consider is R.L. Goodstein. He had been at Cambridge in the 1930's and seems to have discussed philosophy and mathematics with Wittgenstein while he was there. He wrote a review of the

[14]Kreisel 1998, p. 119.
[15](Kreisel 1983, p. 300) quoted in (Floyd 2001, p. 285).
[16]See Bernays 1959.
[17]Bernays 1959, p. 23.

Remarks, which appeared in *Mind* in 1957,[18] and was, unlike most of the others, quite positive towards Wittgenstein's philosophy of mathematics overall, calling many of his remarks "illuminating" and concluding that the book "contains many gems of insight and wisdom".[19]

His opinion on Wittgenstein's remarks on Gödel is, however, not so positive. Apart from §7 of the Appendix, which he claims contains a "very important insight", his overall judgement is that the appendix "is unimportant and throws no light on Gödel's work". Following this, he puzzles over the fact that "what Wittgenstein said on the subject in 1935 was far in advance of his standpoint three years later".[20] What Wittgenstein had suggested to Goodstein was that Gödel's proof indirectly showed that there were non-standard models of arithmetic. In his original 1957 review of the *Remarks*, Goodstein wrote that

> Wittgenstein with remarkable insight said in the early thirties that Gödel's results showed that the notion of a finite cardinal could not be expressed in an axiomatic system and that formal number variables must necessarily take values other than natural numbers; a view which, following Skolem's 1934 publication, of which Wittgenstein was unaware, is now generally accepted.[21]

And in 1972 he wrote,

> I do not think Wittgenstein heard of Gödel's discovery before 1935; on hearing about this his immediate reaction, with I think truly remarkable insight, was to observe that it showed that the formalisation of arithmetic with mathematical induction and the substitution of numerals for variables fails to capture the concept of natural number, and the variables must admit values which are not natural numbers.[22]

This obviously contradicts Kreisel's claims somwehat, according to which Wittgenstein hadn't read Gödel's paper until at least the 1940's.

Juliet Floyd at least takes Goodstein's testimony as a good reason to believe that Wittgenstein understood Gödel's theorem, and even Dummett agrees, but he has less faith in Wittgenstein's ability in keeping that understanding. He says that Goodstein's testimony indicates that Wittgenstein had "rather a good understanding of the incompleteness theorem." He then immediately adds: "He surely lost it".[23] How plausible that is, is hard to see.

[18]See Goodstein 1957.
[19]Goodstein 1957, p. 553.
[20]Ibid., p. 551.
[21]Ibid. Goodstein must have meant that non-standard models exist, not that no standard models exist.
[22]Goodstein 1972, p. 279.
[23]Dummett 1997, p. 363.

The last of the major early commentators we shall mention is Alan Ross Anderson.[24] His review is largely negative, and on the questions of consistency and Gödel's theorem quite hostile. On these matters, he declares that

> ... Wittgenstein misunderstood both content of and the motivation for a number of the results he discusses at length.[25]

He accuses Wittgenstein of ignoring that Gödel's proof is constructive and explicitly based on the assumption of consistency.[26] He further accuses Wittgenstein of missing the point when he claims that "if the proposition is supposed to be false in some other sense than Russell's sense",[27] that would not entail a contradiction if it were true in Russell's sense, since the problem is to determine whether or not all semantically false sentences are provably false in the formal system, or

> in Wittgenstein's eccentric terminology, to determine whether all falsehoods in "some other than Russell's sense" are also falsehoods in the alleged "Russell sense."[28]

His response is that this Gödel showed that this is not the case, and accuses Wittgenstein how trying to skirt the problem by simply equating truth with provability.[29] He also claims that Wittgenstein misses the point of Gödel's theorem more generally. In §7, Wittgenstein suggests that there is nothing surprising by the fact that there are true but unprovable propositions in some system, and says,

> Certainly, why should there not be such propositions; or rather: why should not propositions—of physics, e.g.–be written in Russell's symbolism.[30]

This is trivial, claims Anderson. While the language of PM is of course not expressive enough to be able to express many physical statements, there is nothing surprising about the fact that PM would not be able to prove even the most elementary statement of physics, were it extended in the proper way. The real point of Gödel's theorem is that Gödel showed that PM was unable to prove truths that were expected to be provable in it—he showed that no formal system can capture all *arithmetical* truths. Anderson's conclusion is harsh:

[24] We will not cover Dummett, as his dismissal of Wittgenstein's remarks, despite their influence, comprise of no more than what has already been cited in the introduction.
[25] Anderson 1958, p. 451.
[26] Ibid.
[27] RFM I, App. III, §8.
[28] Anderson 1958, p. 453.
[29] Ibid. We will see in a later section that Mark Steiner makes a similar point, albeit from a more technical standpoint.
[30] RFM I, App. III, §7.

> It is hard to avoid the conclusion that Wittgenstein failed to understand clearly the problems in which workers in the foundations have been concerned

This conclusion of Anderson's was the dominant (and perhaps the only) view until Stuart Shanker published his interpretation in 1988, in the paper *Wittgenstein's Remarks on the Significance of Gödel's Theorem.* In the next chapter we will take a critical look at his interpretation.

Shanker's interpretation

In the preface to his book, *Wittgenstein and the Turning-Point in the Philosophy of Mathematics*, S. G. Shanker remarks that in order to fully understand Wittgenstein's remarks on Gödel, one must understand Wittgenstein's attack on Hilbert's Programme and his conception of metamathematics, and to appreciate them, one must in turn understand Wittgenstein's views on the nature of mathematical proof and propositions.[31] This is what he attempts to do in the book. Despite this, the book itself, somewhat ironically, contains no discussion of Wittgenstein's remarks on Gödel. The "self-imposed obligation", however, of covering those remarks was discharged in a later paper, *Wittgenstein's Remarks on the Significance of Gödel's Theorem*.[32] This paper was one of the first attempts to give a sympathetic interpretation of Wittgenstein's remarks on Gödel and will be the first interpretation we examine in this thesis.

In the first section of this chapter we will cover some of the aspects of Wittgenstein's philosophy of mathematics which Shanker sees as necessary for the latter's criticisms of Gödel, viz. Wittgenstein's claim that the meaning of a mathematical proposition is determined by its position in a calculus, and some of the consequences of this position, e.g. his rejection of Hilbert's conception of metamathematics. When these things are out of the way, we will be in a position to see what Shanker's interpretation of Wittgenstein's remarks, set forth in the latter paper, consists in. This we will do in the second section. The last section will contain an evaluation and criticism of Shanker's interpretation.

Metamathematics and the meaning of mathematical propositions

To give a whole overview of Shanker's position on Wittgenstein's philosophy of mathematics is unfortunately beyond the scope of this thesis, as his book covers vastly more than we need for our purposes, and his 86 page paper on Wittgenstein's remarks on Gödel contains everything from musings on what makes a mathematical theorem important, to a discussion of the revolutionary merit of Beethoven's *Eroica* and Zen Buddhism in the work of Hofstadter. In this section we will rather focus on giving an overview of the relevant aspects of Shanker's account of Wittgenstein's philosophy of mathematics: the meaning of a mathematical statements and Wittgenstein's views on the role of metamathematics.

First of all, Shanker argues that for a correct interpretation of Wittgenstein's philosophy of mathematics, a close reading of *Philosophical Remarks* and

[31]Shanker 1987, p. vii.
[32]Shanker 1988, p. 242f.

Philosophical Grammar is necessary, works written in the years 1929–1930 and 1930–1932, respectively. The reason for this, Shanker says, is that these works contain a fuller treatment of topics not mentioned, or only alluded to, in the *Remarks on the Foundations of Mathematics*.[33] Thus, reading them becomes required for understanding of the latter work, written between 1937 and 1944. Reading these works will, in Shanker's eyes, help us in

> the task of clarifying exactly what themes Wittgenstein was objecting to in the philosophy of mathematics and what position he was arguing for becomes far more straightforward.[34]

In order to understand the later work, one should therefore follow the steps Wittgenstein himself actually took in developing the views found therein, else his true position is bound to be misunderstood.[35] Of course, we cannot trace these steps fully in this account of Shanker's view here, but the following quick outline should give some idea of what it consists in.

For Shanker, Wittgenstein's approach to the philosophy of mathematics has its genesis in the problems facing him when he returned to philosophy in 1929. The immediate difficulty was that in the *Tractatus*, Wittgenstein had claimed that because we can infer from statements such as "A is red" other statements such as "A is not green" and "A is not blue" (and that we consider "A is both red and green" a contradiction), it follows from the independence of elementary propositions that "A is red" is itself not an elementary proposition, and thus not fully analysed.[36] However, a full analysis remained elusive, and Wittgenstein came to think that "A is red" was in fact elementary.[37]

In order to solve this problem, Wittgenstein abandoned the view that language had one underlying logic or calculus. He now believed, says Shanker, that it consisted in "a complex network of interlocking calculi: autonomous 'propositional systems' each of which constitutes a distinct 'logical' space."[38] Wittgenstein called these 'calculi' *Satzsysteme* and we can follow Shanker in calling this the '*Satzsysteme* conception of language',[39] or simply the 'calculus conception'. On this view, the meaning of a name was no longer the object that it denoted, but rather its position in the calculus it belongs to and in order to understand a proposition, one must now understand the whole *Satzsystem* of which it is a part—the inference "A is red, therefore it is not blue", for instance, becomes possible because it is a part of the grammar of the word "red" that it must be

[33] Shanker 1987, p. ix.
[34] Ibid., p. 5.
[35] Ibid., p. 6.
[36] See (TLP, 6.3751).
[37] Shanker 1987, p. 6.
[38] Ibid., p. 7.
[39] Ibid. This is of course a simplification of the matter, but enough for our purposes.

so.[40] Shanker summarises his new view thus:

> Wittgenstein now argued that a word only has meaning in the context of its propositional system, and that the meaning of a word is the totality of rules governing its use in that system.[41]

With this came a 'principle of verification' where the meaning of a proposition is shown by its method of verification, "in so far as the method of verification manifests to which *Satzsystem* a proposition belongs."[42]

This conception of language is mirrored in Wittgenstein's conception of mathematics, Shanker claims. The meaning of mathematical propositions too is determined by their place in a calculus, rather than some kind of connection between our language and a mathematical reality. Mathematical propositions are thus not *about* anything. Rather, they are "norms of representation" that "fix the use of concepts in empirical propositions".[43] Shanker explains:

> Mathematics consists of a network of calculi, and the meaning of any particular mathematical proposition is *Satzsystem*-specific, and thus must also be given by its method of verification, which in this case of mathematics is given by its method of proof.

The relation of a proof to its proposition is internal and *creates* the meaning of the mathematical proposition, i.e. the role of proof is not to merely convince its reader of the truth of the proved proposition (which would be an *external* relation on this picture) but is necessary to establish the very meaning of the proposition being proved—a proof is thus an *essential* part of the proposition it proves.[44]

This picture is what leads to Wittgenstein's rejection of Hilbert's conception of metamathematics. Since mathematical propositions do not refer to anything,[45]

---

[40]Wittgenstein explained his new view to Waismann saying: "The assertions telling me the length of of an object form a system, a system of propositions. It's such a whole system which is compared with reality, not a single proposition. If, for instance, I say such and such a point in the visual field is *blue*, I not only know that, I also know the point isn't green, isn't red, isn't yellow, etc." (PR, p. 317).

[41]Shanker 1987, p. 7.

[42]Ibid., p. 41. Shanker is very careful in making it clear that this is not a 'principle of verification' in the positivist sense, i.e. there is no epistemological aspect to Wittgenstein's principle. It is simply logical matter.

[43]Shanker 1988, p. 211.

[44]This of course immediately raises the following problem: If the meaning of a mathematical *proposition* is dependent on its proof, a mathematical *conjecture* changes its meaning when it has been proven. It then follows that a mathematical conjecture can *never* be proven (since the proposition proven is not the same as the one conjectured). For Shanker, conjectures are strictly speaking meaningless but provide a 'stimulus' for the mathematician to come up with a proof, and thereby a new calculus. See (PG, p. 371) and (Shanker 1987, p. 104–). We will not discuss the matter here.

[45]Except perhaps its own proof. See for instance (PG, p. 455): "We could say: a mathematical proposition is an allusion to a proof." and (PR, §122): "A mathematical proposition is an allusion to a proof."

the idea that the "objects of investigation" of metamathematics are proofs is for Shanker "yet another fresh confusion on the basis of the entrenched platonist assumption that mathematical propositions are descriptive."[46] This, however, does not mean that metamathematics is completely illegitimate. For Shanker's Wittgenstein, metamathematics can be "incorporated into the family of mathematics", provided one is clear that its proofs and propositions do not refer to anything in particular, not even mathematical proofs.[47] Metamathematics has thus no special status—it is merely another calculus. Shanker cites Wittgenstein's conversations with Waismann and Schlick, where he said,

> What Hilbert does is mathematics and not metamathematics. It is another calculus just like any other.[48]

and the following passage from *Philosophical Grammar*:

> I can play with chessmen, according to certain rules. But I can also invent a game in which I play with the rules themselves. The pieces of my game are now the rules of chess, and the rules of the game are, say, the laws of logic. In that case I have yet another game and not a metagame.[49]

Wittgenstein also used this analogy with chess in his conversation with Waissman. Waismann had objected that just like there is a theory of chess in which we can learn about the possibilities of certain positions in chess, there is a theory of arithmetic in which we "obtain material information about the possibilities of this game". He further added: "This theory is Hilbert's metamathematics".[50] Wittgenstein's reply is of importance to Shanker:

> What is know as the 'theory of chess' isn't a theory describing something, it's a kind of geometry. It is of course in its turn a calculus and not a theory. [. . .] [I]f I establish in the 'theory' that such and such possibilities are present, I am again moving about within the game, not within a metagame. Every *step* in the calculus corresponds to a move in the game, and the whole difference consists only in the physical movement of pieces of wood.[51]

In other words, what is essential in the playing of chess is not the moving of little wooden pieces across a wooden board, but the "totality of rules that constitutes playing chess"[52] and it is a mistake to think that some kind of

[46]Shanker 1988, p. 211. Cf. (PG, 290): "Since mathematics is a calculus and hence isn't really *about* anything, there isn't any metamathematics."
[47]Ibid., p. 212.
[48]WVC, p. 121. See also (WVC, p. 136).
[49]PR, §319.
[50]PR, p. 326.
[51]PR, p. 327.
[52]Shanker 1988, p. 211.

symbolic description of the moves in a game of chess is fundamentally different from moving the pieces on the board. And since it follows from the meaning-as-rule conception of mathematical propositions that what might seem on the surface as the same proposition occurring in two different *Satzsysteme*, is in fact two distinct propositions, Shanker's conclusion is that what is supposedly a meta-game is either a completely different game[53] or the same game played with different signs "standing for the same pieces".[54]

This applies especially to Hilbert's metamathematics. When Hilbert looks for a metamathematical proof, Shanker says, that arithmetic is consistent, his metamathematical expressions are merely new expressions in a new calculus he constructed, i.e. an expression such as "$0 \neq 0$ is not a theorem" is not really a metamathematical expression but a rule in a different calculus Hilbert has made especially for the purpose of prohibiting such a construction. In one of his conversations with Waismann, Wittgenstein remarked:

> If someone were to describe the introduction of irrational numbers by saying he had discovered that between the rational points on a line there were yet more points, we would reply: 'Of course you haven't discovered new points between the old ones: you have constructed new points. So you have a new calculus before you.' That's what we must say to Hilbert when he believes it to be a discovery that mathematics is consistent. In reality the situation is that Hilbert doesn't establish something, he lays it down. When Hilbert says $0 \neq 0$ is not to occur as a provable formula, he defines a calculus by permission and prohibition.[55]

Wittgenstein's objections to Hilbert's programme do thus not arise from any 'finitist' misgivings,[56] as some commentators have had it, but from his conception of mathematics as a collection of autonomous calculi, which leads to the view that no mathematical argument or proof can lead to a solution of philosophical problems.[57]

This account of mathematics gave rise to two reasons why Wittgenstein

[53]Cf. (PR §152): "The system of rules determining a calculus thereby determines the 'meaning' of its signs too. Put more strictly: the form and the rules of syntax are equivalent. So if I change the rules — seemingly supplement them, say — then I change the form, the meaning."

[54]Shanker 1988, p. 214.

[55]PR, p. 339.

[56]This is for instance Victor Rodych's view.

[57]Shanker 1988, p. 214. Shanker cites (PR, p. 320): "If the contradictions in mathematics arise through an unclarity, I can *never dispel this unclarity by a proof.* The proof only proves what it proves. But it can't lift the fog." But see also (PG, p. 296): "No calculus can decide a mathematical problem. A calculus cannot give us information about the foundations of mathematics." and (PG, p. 369): "In mathematics there can only be mathematical troubles, there can't be philosophical ones."

"reacted so strongly to Gödel's theorem", Shanker claims.[58] The former, more general one, is the philosophical confusions Wittgenstein saw underpinning the way Gödel's theorem was being used and interpreted by philosophers, more specifically they way it was used to argue for platonism in the philosophy of mathematics.[59] The other, more specific, reason is that while Gödel's Second Theorem (which shows that PA cannot prove its own consistency) shows that Hilbert's programme cannot be carried out, Wittgenstein was against the sceptical considerations that made such a proof necessary in the first place, and the way of interpreting Gödel's theorem he was against depends on accepting the premises of the sceptic. For him, there simply was never any foundation crisis in mathematics, as mathematics is simply a network of rule-bound calculi where the meaning of propositions is established by those rules (and thus scepticism becomes not only false, but meaningless).

Thus for Shanker, the real force of Wittgenstein's critique of Gödel's theorem lies in how Hilbert's "blending of philosophy and mathematics"[60] leads to a platonist interpretation of the theorems. However, when the former has been removed we are still left with, in the words of Shanker,

> a 'piece of mathematics' which – stripped of its metaphysical associations – lies waiting to be elucidated. And elucidated it must be if the lingering epistemological pull which shapes Hilbert's programme is to be avoided, and thence the platonist interpretation of Gödel's theorem.[61]

Thus for Shanker, Wittgenstein never intended to attack Gödel's theorem as a piece of mathematics, but rather a certain philosophical interpretation of the theorem he believed led to a mistaken platonism. In the next section we will see more specifically what Shanker believes this attack consists in.

Shanker's account of Wittgenstein's criticisms of Gödel

For Shanker, Gödel's arithmetization of syntax has two aspects in his proof. The former, which involves establishing a one-to-one correspondence between the expressions of a given formal system with a set of numbers, Shanker has no problem with. For him, Gödel has simply set up new calculus, which we can call the calculus of Gödel numbers, and way of mapping expressions from one to the other.[62] The latter use, however, is more objectionable. This is what Shanker

[58] Shanker 1988, p. 180.
[59] Ibid., p. 171. As we will see, this is a common theme amongst the various interpretations of Wittgenstein's remarks on Gödel.
[60] Ibid., p. 183.
[61] Ibid.
[62] Ibid.

calls “arithmetization of meta-mathematics”.[63] Here, the idea is, Shanker claims, that a metamathematical statement about the expressions in the object theory can at the same time be read as a statement about arithmetical relations between the Gödel numbers of the very same expressions. This, he claims, is misguided, in the light of Wittgenstein’s conception of mathematics covered in the last section: The meta-mathematical statement which is supposedly *about* the expressions in the object language, and therefore about arithmetical relations in Gödel number calculus must either just be the old expression in a different guise or a new expression in a new, completely independent calculus.[64]

The importance of this point, Shanker claims, can be seen from the fact that in Gödel’s interpretation of his proof, he needs to show that metamathematical statements about the “structural properties of the expressions within a calculus” can be mirrored by those same expressions, and also that the arithmetical relations between numbers in the Gödel number calculus can be interpreted as

> reflecting meta-mathematical propositions about the logical relations between these statements and metamathematical issues can be pursued by examining these arithmetical relations themselves.[65]

An example is how Shanker treats an example from Nagel and Newman.[66] Nagel and Newman discuss a certain meta-mathematical statement used by Gödel in his proof: “The sequence of formulas with Gödel number x is a proof of the formula with Gödel number z”.[67] This sentence is true if and only if x and z stand in a certain arithmetical relationship we can call after Nagel and Newman, Dem.

Now, “[a]ccordingly,” Nagel and Newman say, “to establish the truth or falsity of the meta-mathematical statement under discussion, we need concern ourselves only with the question wether the relation Dem holds between two numbers”.[68] This way of putting the matter, Shanker finds highly misleading, since the way of constructing the meta-mathematical statement depends essentially on “stipulating rules for mapping ‘object’ onto ‘GN’ expressions”,[69] and it is thus a necessary, mathematical statement (bound by the rules of a calculus, on Wittgenstein’s view). Yet it also viewed as almost contingent, since the metamathematical statement does not stand in an internal relation to its proof (since if it did, it would be a different proposition than the one in the object calculus, which it is not supposed to be). For Shanker, Gödel’s interpretation depends on going between

[63] Shanker 1988, p. 216.
[64] Ibid.
[65] Ibid., p. 217.
[66] See Nagel and Newman 1958.
[67] Nagel and Newman 1958, p. 79. quoted in Shanker 1988, p. 218.
[68] Nagel and Newman 1958, p. 79–80. quoted in Shanker 1988, p. 218.
[69] Shanker 1988, p. 216.

these two ways of viewing meta-mathematical statements. Shanker explains:

> We are asked to accept that a meta-mathemtical statement can be mirrored by a formula in the GN-calculus which depicts a purely arithmetical relation between x and z. Gödel's metamathematical propositions are thus invariably treated as a species of – necessarily true – mathematical proposition; but like ordinary empirical propositions they are supposed to be *about* the logical relations holding in the 'object calculus' and are contingently true or false.[70]

More specifically, Gödel's step from showing that P is undecidable to showing that it must therefore be true, depends on viewing P in two different calculi at the same time, the level of the object language, and the metalanguage. But since the relation of proof to its the proposition is internal, this *eo ipso* cannot be.[71]

That is to say, on Gödel's conception, P is the same proposition in both the object calculus and the metacalculus, and if this is the case, on Wittgenstein's view, it both has an internal proof relation (where its meaning is established by the rules in the object calculus) and an external one in the metalanguage (since if it was internal in the latter case, that would show that they were different). Shanker's conclusion, in short, is that

> Gödel was barred by virtue of the logical grammar of *mathematical proposition* from claiming that he had constructed identical versions of the same mathematical proposition in two different systems.[72]

And this, for Shanker, is what Wittgenstein meant in (RFM I, App. III, §6) when Wittgenstein asks

> "Under what circumstances is a proposition asserted in Russell's game?" the answer is: at the end of one of his proofs, or as a 'fundamental law' (Pp.). There is no other way in this system of employing asserted propositions in Russell's symbolism.[73]

Namely, to be a mathematical proposition is to belong to a system of proof, and those cannot be "transposed from one system to another"[74] because "the meaning

[70]Shanker 1988, p. 218. Shanker doesn't say so, but perhaps the following quote from Gödel's original paper expresses the same sentiment, except approvingly: "Contrary to appearance, such a proposition involves no faulty circularity, for initially it [only] asserts that a certain well-defined formula (namely, the one obtained from the qth formula in the lexicographical order by a certain substitution) is unprovable. Only subsequently (and so to speak by chance) does it turn out that this formula is precisely the one by which the proposition itself was expressed." (Gödel 1976, p. 589. Footnote 15.)

[71]Shanker 1988, p. 228.

[72]Ibid., p. 229.

[73]RFM I, App III, §6.

[74]Shanker 1988, p. 228.

of a mathematical proposition is strictly determined by the rules governing its use in a specific system."[75] At most, Shanker says, Gödel could have concluded that he had constructed *parallel* but different mathematical statements, and in that case, there would not be anything remarkable about one of them being unprovable in its calculus, while the other was true in its.[76]

In other words, Shanker's view is that since the meaning of a mathematical proposition is solely determined by its place in a calculus and thus determined by its proof, the meaning of P in the object calculus and the meaning of P in the metacalculus cannot be the same, and therefore it is wrong to say that P is a true but unprovable proposition in the system of *Principia Mathematica*. At most could it be said that there are two propositions, both called P, one of which is undecidable in *Principia Mathematica* and another which is true (and provable) in the metalanguage. Furthermore, since P is not provable in the object calculus, it has no meaning there.

This has negative consequences for Hilbert's programme, Shanker claims, but not Gödel's proof as a mathematical proof. Wittgenstein's real target was the "two pages of prose prefaced to the proof" which contained "Wittgenstein's real quarrel with Gödel".[77] On Gödel's interpretation of his theorem, it shows the relation of a proposition to its proof is external, or what comes to the same, that provability and truth are not synonymous. Wittgenstein however emphasised the opposite point, and thus for him, the theorems become a *reductio ad absurdum* of the philosophical assumptions of Hilbert's Programme, not for the reasons often supposed, however, but because it shows that a mathematical statement cannot be true "outside of the set of rules that create its meaning" and that a mathematical theorem cannot "intelligible *prior to* or *independently of* the construction that determines its meaning."[78]

This, Shanker says, is the real philosophical problem of Gödel's interpretation,[79] for only someone who already holds the *Bedeutungskörper* conception of mathematical meaning (i.e. the idea that meaning is an object, abstract or otherwise, that propositions and words stand for) could think that a true but unprovable sentence could be mechanically generated by a formal system.[80] Seen in this light, Wittgenstein's motivations are the same in the philosophy of mathematics, as they are in the philosophy of language, namely to remove this mistaken view of meaning. Gödel's theorem was simply particularly apt to lead one astray and adopt this conception.

[75] Shanker 1988, p. 229.
[76] Ibid.
[77] Ibid., p. 233–234.
[78] Ibid., p. 234.
[79] Ibid., p. 240.
[80] Ibid., p. 237.

In the next section we will provide some criticism of Shanker's reading of Wittgenstein's remarks.

Criticism of Shanker's view

As seen from the previous section, Shanker's interpretation of Wittgenstein's remarks on Gödel makes heavy use of what we called the calculus (or *Satzsysteme*) conception of mathematics. In Shanker's view, Wittgenstein sees mathematics as a system of autonomous calculi, where the meaning of mathematical statements is determined by the rules of the calculus it inhabits. This was also Wittgenstein's view of language in the intermediary period. There is little doubt that Wittgenstein held something close to the calculus view when he wrote *Philosophical Grammar* and *Philosophical Remarks*, as can be seen from the plethora of citations from those works that can be found in Shanker's paper. It is however highly doubtful that he held this view in the later period, the period of the *Philosophical Investigations* and the *Remarks on the Foundations of Mathematics*. In those works, the notions of rule and calculus have been replaced by an emphasis on human practice and forms of life.

Perhaps surprisingly, Shanker agrees. Despite constantly emphasising Wittgenstein's calculus conception of meaning in his paper and motivating his whole interpretation on it, he seems to be well aware of the inadequacy of this way of viewing Wittgenstein's philosophy of mathematics in the latter period. Towards the end of his paper, he even claims that the *Philosophical Investigations* is a "sustained attack on the calculus conception of meaning"[81] and his interpretation of §§19–20 correctly states that the reason Wittgenstein viewed P as meaningless is that it cannot be given any application, neither inside mathematics, nor in any extra-systematic sense (i.e. not in counting, weighing, measuring etc.). Shanker says:

> What we must 'remember is that in mathematics we are convinced of *grammatical* propositions; so the expression, the result of our being convinced is that we *accept a rule*' It is not the rules of inference and the construction rules of a system which, on their own as it were, determine what shall count as a mathematical proposition; it is the use to which we do or can put such rules in the transformation of empirical propositions.

This is nothing less than a complete denial of the calculus conception. If this is indeed Shanker's view, it is hard to see why he spent the first 70 pages of his paper trying to convince the reader that Wittgenstein held the calculus

[81] Shanker 1988, p. 239.

conception of mathematics, which then serves as a motivation behind his whole interpretation, only to later abandon it in the same paper. It is of course not implausible that the conclusion and spirit of Shanker's interpretation, whose major claim is that P cannot be unprovable in the object language and true in the metalanguage because a correct view of meaning in mathematics would show that P would be two different propositions, is true, i.e. it is still possible that Wittgenstein held this view and that the language-games conception entails a similar conclusion, but if so, Shanker has done nothing to support it.[82]

We will not further pursue the matter of which conception is the correct one, since it is clear that Shanker agrees that Wittgenstein did not hold the calculus conception in the latter period, and this is enough to establish that his position is incoherent or unmotivated: either he has himself swept the rug from under his own interpretation, or he hasn't given any motivation for it at all (since it was all based on the calculus interpretation, which is false)![83]

Shanker's interpretation has other problems, however. He rarely cites evidence in the text of (RFM I, App. III) for his exposition (indeed, his paper is dominated by references to the *Philosophical Remarks* and *Philosophical Grammar* rather than to the very remarks he is explaining), and in those cases where he does try to give textual evidence for his interpretation, it is often very difficult to fit what he has to say onto the actual text. Some of his interpretations of Wittgenstein's specific remarks seem at best irrelevant, and at worst taken out of context to fit his interpretation.

An example of this is how he interprets §17 as grappling with the question of how Gödel's theorem should be interpreted,[84] when it is obvious from reading the remark that the word "interpretation" refers specifically to the interpretation of the *sentence* P,[85] and reading that as referring to how to best interpret Gödel's theorem seems misguided, since in earlier sections Wittgenstein had repeatedly written as a certain natural language interpretation of P was needed for Gödel's proof.[86] The only thing Shanker's interpretation and the obvious reading have in common is the word "interpretation".

Another example, are Shanker's claims, on the basis of §15, that Wittgenstein was worried that the P might "serve as a paradigm of synthetic *a priori*

[82]§7 might be interpreted such that Wittgenstein is saying that P has a different meaning in the two language games, but §8, which Shanker ignores, might point in another direction, since it seems to be a continuation of §7 and there Wittgenstein seems to be focused on avoiding Gödel's result by abandoning a natural language interpretation of P.

[83]A good overview over the differences in these two periods of Wittgenstein's philosophy of mathematics can be found in (Gerrard 1991).

[84]"The problem here is not so much how one arrives at Gödel's theorem as how much such a theorem should be interpreted (RFM I App III §)" (Shanker 1988, p. 236.)

[85]For instance, Wittgenstein writes: "...and it must now come out how this interpretation of the symbols of P collides with the fact of the proof..." and "—When the interpretation "P is unprovable" was given to P..."

[86]For more on this, see the chapter on Victor Rodych's interpretation.

knowledge". Such a reading of §15 is highly unlikely, as it contains reflections on whether or not P could serve as a prediction of unprovability (and is as such a direct continuation of §14).[87]

The main difficulty with Shanker's interpretation is, however, not the problems just outlined. Many of his claims are quite plausible in themselves (for instance that Wittgenstein disliked a platonist interpretation of Gödel's theorem or objected to Hilbert's Programme). It is rather that Shanker's account of Wittgenstein's remarks is so specific and detailed—and lacking in actual references to the text—that the text itself is completely unable to support it. Wittgenstein simply never said most of these things Shanker says that he did. (Shanker himself notes that nowhere in (RFM I, App. III) is there a detail examination of Gödel's proof—yet Shanker's interpretation consists of little else.)

Of course, Shanker's interpretation doesn't come from nothing, and it very possible to debate whether the position and arguments Shanker attributes to Wittgenstein are compatible with other things the latter had to say, but being consistent with Wittgenstein's philosophy in general is at most a necessary condition of an interpretation of his remarks, and not a sufficient one.

Shanker has simply not given enough motivation for his interpretation, other than roughly being in the spirit in which he himself sees Wittgenstein's philosophy. One might then do well to consider Shanker's own words in the concluding paragraphs of his paper, where he excuses himself for not covering Wittgenstein's criticisms of the notions of 'effective decidability' and 'mechanical calculation' for the reason that they lead "dangerously into that area of exegesis where accusations of critical revisionism are most easy to prosecute".[88] This is precisely what he is guilty of in his interpretation of Wittgenstein's remarks on Gödel, and is, as we will see, a fairly common vice of those who have tried to interpret Wittgenstein's remarks on Gödel.

In the next chapter we will take a look at Juliet Floyd's interpretation, which is, in many ways, quite different from Shanker's.

[87] We will look closer at §14 in the chapter on Juliet Floyd's interpretation.
[88] Shanker 1988, p. 201.

What Wittgenstein said or should have said: Floyd vs. Steiner

After Shanker's paper from 1988, not much was written on Wittgenstein's remarks on Gödel, and indeed it seems to have been largely ignored. In her 1995 paper *On Saying What You Really Want to Say: Wittgenstein, Gödel and the Trisection of an Angle*, Juliet Floyd published an interpretation which was in many ways at odds with Shanker's. There, she interprets Wittgenstein as using Gödel's theorem to show that the concept of understanding mathematical propositions is unclear and needs to be elucidated. It spurred some debate, with a notable example being Mark Steiner's rather late reply to her in 2001. Later that year Floyd replied to Steiner in another paper, *Prose versus Proof: Wittgenstein on Gödel, Tarski and Truth*. In this chapter, we will give an overview of their respective interpretations, starting with Floyd, and then critically evaluate them.

Gödel's theorem as an impossibility result: Floyd's interpretation

Floyd's starting point in her interpretation is (RFM VII, §19): "My task is, not to talk about (e.g.) Gödel's proof, but to by-pass it". She interprets this passage to mean that Wittgenstein's intention in his remarks on Gödel were not to argue *against* Gödel's proof, like many have thought, particularly the early commentators, but rather "deflate the *apparent* significance of Gödel's theorem".[89] What Floyd means by this is not that Wittgenstein believed that Gödel's proof was unimportant as a proof *in* mathematics—its importance as a mathematical proof is undeniable—but rather that it does not have the philosophical importance as some have liked to think.

Prima facie, one might that Floyd is overstating her case, after all "talk about" does not mean "argue against". However, the context suggests otherwise, as we can read earlier in the same remark:

> It is my task, not to attack Russell's logic from within, but from without.
>
> That is to say: not to attack it mathematically—otherwise I should be doing mathematics—but its position, its office.

Furthermore, there are almost countless examples from almost any text Wittgenstein wrote on the philosophy of mathematics that show that he was very much

[89] Floyd 1995, p. 375.

concerned with *not* criticising the work of mathematicians *as* mathematics.[90] This should not be surprising, as Wittgenstein saw the role of philosophy as to "leave everything as it is"[91], and meddling in the affairs of mathematicians goes explicitly against this maxim:

> It also leaves mathematics as it is, and no mathematical discovery can advance it. A "leading problem of mathematical logic" is for us a problem of mathematics like any other. (PI, §124)

This of course raises the question of why Wittgenstein concerned himself with Gödel's proof in the first place—it has certainly seemed to many that he is trying to legislate what mathematicians can and cannot do.

A somewhat plausible answer to this question would be that Wittgenstein did in fact view the activity of giving a proof as necessary for us to understand mathematical statements in the first place, and that he thought that the notion of 'calculus' was central in the giving of mathematical proofs. This view of Wittgenstein's philosophy of mathematics, especially what has been called the 'intermediary phase', is not implausible.[92] Gödel's proof, however, showed that mathematical truth cannot be equated with a proof in a formal system and thus, as many commentators have assumed, Wittgenstein was—by somehow pointing out flaws in Gödel's proof—mainly concerned with defending the idea that truth in mathematics is to be equated with a formal proof.[93] This reading of Wittgenstein's remarks, which is superficially plausible, is rejected by Floyd.

Her reasons are that Wittgenstein never believed in Russell and Frege's logicist programme in the first place and criticised it at least since the time of the *Notebooks 1914–1916* and the *Tractatus*. This rejection can also be found extensively in the *Remarks on the Foundations of Mathematics* where Wittgenstein often criticises heavily "the idea that *Principia Mathematica* exhibits an underlying logical structure which forms the basis, or essence, of mathematics."[94] For instance, in (RFM III, §46), Wittgenstein says,

> I should like to say: mathematics is a MOTLEY of techniques of proof:—And upon this is based its manifold applicability and its importance...[95]

[90]Cf. (RFM VII, §19).
[91]PI, §124.
[92]See, e.g. the chapter on Shanker for examples of this.
[93]Whether this interpretation of Gödel's proof actually contradicts Wittgenstein's views is however not clear—there is no reason to think that Wittgenstein's notion of a 'calculus' is confined to the formal kind and many reasons to suppose the opposite, at least in the later period when the *Remarks on the Foundations of Mathematics* were written. See the chapters on Shanker and Rodych for further discussion.
[94]Floyd 1995, p. 378.
[95]RFM III, §46.

and:

> I should like to say: Russell's foundation of mathematics postpones the introduction of new techniques—until finally you believe that this is no longer necessary at all.

In (RMF V, §24), he goes on to say,

> The harmful thing about logical technique is that it makes us forget the special mathematical technique. Whereas logical technique is only an auxiliary technique in mathematics. For example it sets up certain connexions between different techniques. It is almost if one tried to say that cabinet-making consisted in glueing.[96]

These passages, and others, show that Wittgenstein thought that a logicist 'reduction' of mathematics presented a too narrow a picture of how mathematics really is.

He grants that mathematical logic has an important role to play in the way we happen to do mathematics, namely setting up connections between different techniques, but denied that it could somehow show us what the essence of mathematics really is, as Frege and Russell insisted—this would be similar to claiming that the essence of cabinet-making is glueing. It is of course no surprise that logic cannot reveal the essence of mathematics, on Wittgenstein's view, since indeed mathematics has no such essence ("mathematics is a MOTLEY of techniques of proof"). The picture Frege and Russell offer is thus misleading, it could at most capture a very narrow application of mathematics (compare this with Wittgenstein's discussion of the Augustinian picture of language in the *Philosophical Investigations*: it is not that it is *false*, but it is *misleading*).

This, Floyd argues, is a good reason to think that Wittgenstein could not have viewed Gödel's theorem as "decisive for our notions of 'mathematically true' and 'mathematically provable' ",[97] because in order to do so, it is necessary for him to already believe that the notion of a formal system (perhaps Frege's or Russell's) is of supreme importance in clarifying (or perhaps providing a foundation for) those very notions.[98] And as the above shows, Wittgenstein did not believe that.

Now, it might be argued against Floyd that in some way Wittgenstein did believe that mathematics has an essence—mathematics is a collection of different techniques, sure, but it is a collection of different techniques of *proof*, and as Gödel showed, mathematical truth and mathematical proof do not coincide. Hence there would still be reason for Wittgenstein to argue against Gödel's theorem. But this is too quick. Gödel's proof only shows that there are true sentences which

[96]RFM III, §46.
[97]Floyd 1995, p. 380.
[98]Ibid.

are not decidable in a certain class of formal systems. This is what Gödel *proved* and the undecidable sentence in question is *proven* by Gödel to be true—but it is not a formal proof (in the sense of a derivation from axioms).[99] On Floyd's view, Wittgenstein's notion of proof perfectly accommodates this, as his notion of proof is much wider than the notion of formal proof. Hence, it would seem quite plausible, as Floyd claims, that Wittgenstein's remarks should not be taken to be criticism of Gödel's proof as a mathematical proof.[100]

Instead, according to Floyd, Wittgenstein simply viewed Gödel's proof as a valid impossibility proof—in a similar sense of the algebraic proof of the impossibility of trisecting an arbitrary angle, and what bothered him about the proof was not its status as a mathematical proof, but that it is "likely to mislead (some people) philosophically".[101] In the next two sections, we'll take a closer look Floyd's reading, first by explaining the context of her reading of Wittgenstein's remarks on Gödel in relation to her understanding of his remarks on trisection, and then we will see how she reads the relevant sections in (RFM I, App. III).

The trisection of an angle

Wittgenstein usually discussed the example of the trisection of the angle when he was concerned with the question of what it is to find (or try to find) a proof of a mathematical conjecture.[102] According to Floyd, this question is, for Wittgenstein, closely connected to a host of intentional notions, such as thinking, understanding, meaning, etc. — notions which Wittgenstein is very much concerned with throughout is later work.[103] The first place in the *Philosophical Investigations* where Wittgenstein mentions trisection is (PI, §334). Here Wittgenstein is discussing the use of phrases like "So you really wanted to say…". He points out that we sometimes use these kinds of phrases to "lead someone from one form of expression to another" but that there is a temptation to think, when one uses phrases of this kind (also "he meant…"), that what is meant is already present in the speakers mind before he said it. Next Wittgenstein says,

> To understand this, it is useful to consider the relation in which the situations of mathematical problems stand to the context and origin of their formulation. The concept 'trisection of the angle with ruler and compass', when people are trying to do it, and, on the other

[99]Gödel himself claims that the truth of the undecidable sentence was established by metamathematical means (Gödel 1976, p. 599.), but for Wittgenstein there are no metamathematics, only more mathematics (at least in the intermediary period).
[100]This is also Shanker's view. See above.
[101]Floyd 1995, p. 375.
[102]Ibid.
[103]Ibid., p. 381.

> hand, when it has been proved that there is no such thing.[104]

Here, Floyd says, Wittgenstein is granting that there is a sense in which the phrase "So you really wanted to say..." *is* used when it is natural to suppose that the meaning is "present in the mind of the speaker" but that this notion can also be misleading.[105] One example of such a case where this supposition is natural is for instance when a speaker obviously misspeaks and is corrected by his listener by her saying "You surely meant...".[106] However, another important use of the phrase is also used in context where the intention of the speaker is in fact unclear. A teacher, to use Floyd's example, uses it to lead his student into a better (from the teacher's point of view) way of behaving, by offering another expression and "urge its adoption" instead.[107] Floyd explains:

> "So you really wanted to say..." can be used to secure the application of logic: in the course of presenting an argument, when one traces out the implications of a thought, one may be led from one step to the next by use of such a phrase.[108]

For Floyd, this comes out clearly in the case of mathematical proofs, especially impossibility proofs, where one is brought to see, through the steps of the proof, that what once seemed to be a perfectly plausible (and perhaps conjectured to be true) mathematical statement, is in fact contradictory. The example of trisection, is used here, Floyd says, to make us see that "the "picture" of something clearly present to the mind ahead of (or apart from) its expression is both a useful picture, and at the same time one whose application is limited, appropriate only in a contextual sense, when it does apply."[109]

After all, is there anything clearer—as an expression of thought, Floyd might say—than the statement that a general method of trisecting an arbitrary angle in Euclidian geometry exists? We have the example of bisection before us, and perfectly understand what trisecting is—understanding what such a general method would consist in seems easy. And indeed, people have sought such a proof for almost 2.500 years. It is perfectly natural for us to say that what the trisectors were trying to find was "present in somewhere their minds", and yet, this is exactly what the impossibility proof *seems* to rule out (how can something *incoherent* be present in someone's mind, something that quite strictly *cannot* exist?).

[104]See (PI, §334). In (PI, §329), Wittgenstein had remarked that "[w]hen I think in language, there aren't meanings going through my mind in addition to the verbal expressions: the language is itself the vehicle of thought. (PI, §329)"
[105]Floyd 1995, p. 381.
[106]Think about the old joke: "A Freudian slip is when you mean one thing, but say your mother."
[107]Ibid., p. 383.
[108]Ibid.
[109]Ibid., p. 382. See also (BB, p. 41).

Nevertheless, when we have the impossibility proof before us, we see that what the unfortunate trisectors were looking for not only doesn't exist, but cannot possibly exist. And isn't this insisting that people were engaged, with considerable effort, in something which didn't make sense and they didn't want to be engaged in? Floyd summarises the situation thus,

> So that once the proof has been accepted, there can *seem* to be a conflict between wanting to grant full and determinate meaning to the (former) conjecture that "A trisection construction exists"; and yet wanting, as a result of proof, to deny that this claim *really* makes any sense at all, to insist that no one *really*, ultimately, wants - or ever wanted - to say such a thing.[110]

Wittgenstein seems to be making the same point in the following passage from the notes G.E. Moore took when he attended Wittgenstein's lectures in the early 1930's. According to Moore, Wittgenstein said

> that 'looking for' a trisection by rule and compasses is not like 'looking for' a unicorn, since 'There are unicorns' has sense, although there are in fact no unicorns, whereas 'There are animals which show the on their foreheads a construction by rule and compasses of the trisection of an angle' is just nonsense like 'There are animals with three horns but also with only one horn': it does not give a description of any possible animal.[111]

However, things are not so simple in Wittgenstein's later work, Floyd claims, despite his tendency of viewing mathematical falsehoods as incoherent. For her, Wittgenstein is "questioning philosophical (pre-)conceptions which are the sources of debates about sense and senselessness."[112] We are perfectly reasonable when we think that no one can actually believe a contradiction but there is also a reason to think that we *can*—as the example of trisection shows: "We can apparently inquire into something which is "contradictory" ".[113] Floyd sees Wittgenstein making precisely this point in (RMF V, §28):

> The difficulty which is felt in connexion with *reduction ad absurdum* in mathematics is this: what goes on in this proof? Something mathematically absurd, and hence unmathematical? How—one would like to ask—can one so much as assume the mathematically absurd at all? That I can assume what is physically false and reduce it *ad*

[110] Floyd 1995, p. 382.
[111] Moore 1959, p. 304.
[112] Floyd 1995, p. 383.
[113] Ibid., p. 383.

absurdum gives me no difficulty. But how to think the—so to speak—unthinkable?

What an indirect proof says, however, is: "If you want *this* then you cannot assume that: for only the opposite of what you do not want to abandon would be compatible with *that*"[114]

This, Floyd claims, shows that what Wittgenstein is trying to do is "wean us away from a certain tempting conception of rationality", since only someone with a certain idea of what proof is, "according to which appreciation of the true logical basis of a judgment is essential to (fully) understanding it", could have this worry.[115]

Floyd thinks therefore that there *is* a sense in which the trisectors didn't fully understand what they were in search of, but this does not mean that their more sceptical fellows were on "firmer ground", since neither a proof of the conjecture had been found nor of its negation.[116] The position of the trisectors is for her analogous to the position of a chess player who thinks he can find a way to force a checkmate with only a king and a knight. It can be proven, with methods which are just as mathematical as the methods used to prove that trisection is impossible, that this cannot be done. This chess player does not see what the rules of chess preclude, and thus entertains an incoherent notion. Yet, he (we can assume) fully understands the rules of chess and indeed must do so if he were ever to come to understand why they preclude such a forced checkmate (and one might add, if only players who fully see what the rules of chess preclude are to be said to understand the rules, no one in the history of the world has ever understood the rules of chess).

This kind of situation is for Floyd characteristic of conjectures in mathematics, and some commentators on Wittgenstein's philosophy of mathematics have claimed that these considerations show that Wittgenstein's view was that mathematical conjectures are without meaning and that we do not understand them until we have the proof in hand.[117] Floyd denies this, for her, there is a shift in understanding when the conjecture is proven but it is not a shift where we move from having *no* concept of e.g. trisecting, and nothing can be meaningfully said about trisection, to fully having such a concept. An important passage for Floyd is the following remark from *Philosophical Grammar*, written some fifteen years before *Philosophical Investigations*:

[114] RFM V, §28.

[115] Floyd 1995, p. 384. Cf. (RFM V, §46): "Understanding a mathematical proposition'— that is a very vague concept."

[116] Floyd 1995, p. 384.

[117] See the section on Shanker above.

> We might say: in Euclidian geometry we can't look for the trisection of an angle, because there is no such thing.
>
> In the world of Euclid's Elements I can no more ask for the trisection of an angle than I can search for it. It just isn't mentioned.
>
> (I can locate the problem of the trisection within a larger system but can't ask within the system of Euclidian geometry whether it's solvable. In what *language* should I ask this? In the Euclidian? - But neither can I ask in Euclidian language about the possibility of bisecting an angle within the Euclidian system. For in that language that would boil down to a question about absolute possibility, which is always nonsense.)
>
> ...A question makes sense only in a calculus which gives us a method for its solution; and a calculus may well give us a method for answering the one question without giving us a method of answering the other. For instance, Euclid doesn't show us how to look for the solutions to his problems; he gives them to us and then proves that they are solutions. And this isn't a psychological or pedagogical matter, but a mathematical one. That is, the *calculus* (the one he gives us) doesn't enable us to look for the construction. A calculus which does enable us to do that is a *different* one. (Compare methods of integration with methods of differentiation, etc.)[118]

For Floyd, Wittgenstein's uses of the word 'calculus' and 'system' here are "loose" and perhaps not what one would expect.[119] She considers it an 'overstatement' to think that it is Wittgenstein's view that mathematics is completely 'algorithmic' or that only conjectures for which we already have a method of solution have meaning, as one might be lead to believe from the quoted passage. Rather, she claims, a 'calculus' or a 'system' for Wittgenstein is "a practice of characteristic linguistic action involving more or less specific techniques" and by a 'conjecture' or 'mathematical question' he means only those for which we can make a systematic search for the answer.[120]

It is in this sense that Wittgenstein means that neither the conjecture that there is a method of bisection in Euclid's system *nor* that there is a method of trisection are possible question within Euclid's system. The former is not a conjecture, says Floyd, because I cannot systematically search for what I already have, namely the proof that there is such a bisection: "once I accept the proof, I cannot conjecture its outcome".[121] Conjecturing that there might be a method

[118]PG, p. 387.
[119]Floyd 1995, p. 392.
[120]Ibid., p. 392–393. This is exactly opposite to Shanker view. It is likely that Floyd is wrong here, but what she says does apply to the late period, however.
[121]Floyd 1995, p. 393.

of trisection is, as we saw above, not properly a question whose answer has a definite truth–value—and how this is mirrored in Wittgenstein's remarks on Gödel we will see anon—but rather a demand for "clarification of the notion of a possible construction",[122] and this cannot be found within Euclid's system, using his methods — a new way is needed to interpret the question (a way which was found in the 19th century algebraic proof). This is for Floyd what Wittgenstein meant in the following passage from *Philosophical Remarks*:

> It is a genuine question if we ask whether it's possible to trisect an angle? And what sort is the proposition and its proof that it's impossible with ruler and compass alone?
>
> We might say, since it's impossible, people could never even have tried to look for a construction.
>
> Until I can *see* the larger system encompassing them both, I can't try to solve the higher problem.
>
> I can't ask whether an angle can be trisected with ruler and compass until I can see the system "Ruler and Compasses" as embedded in a larger one, where the problem is solvable; or better, where the problem is a problem, where this question has a sense.
>
> This is also shown by the fact that you *must* step outside the Euclidian system for a proof of the impossibility.
>
> A system is, so to speak, a world.
>
> Therefore we can't search for a system: What we can search for is the expression for a system that is given me in unwritten symbols.[123]

In other words, conjectures in mathematics, despite looking like propositions with truth–values, are more like "linguistic stimuli", acting as demands for clarification. And the proper answer to such a 'question' is an action, namely to produce a proof.[124] However, before we find the proof, we do not know exactly what it is that we are looking for, and yet to be able carry out such a search meaningfully, we must be able to recognise what would satisfy those demands, that is to say, to recognise what it would be for the answer to be a mathematical solution.[125] This point is brought out by Wittgenstein, Floyd says, in his Lectures in Cambridge in 1934–1935:

> [Mathematical conjectures] are like the problem set by the king in the fairy tale who told the princess to come neither naked nor dressed,

[122]Floyd 1995, p. 393.
[123]PR, 177–178.
[124]In a similar way, a proper answer to a question such as "Can you open the window?" is not to say 'yes' or 'no', but to open the window.
[125]Floyd 1995, p. 393. Wittgenstein sometimes uses the analogy of 'groping about', trying to 'wiggle one's ears' or 'willing an object to move across a room'. See for instance, (PR XIII), (PG, p. 393) and (WCV, pp. 34 and 136).

> and she came wearing fish net. That might have been called not naked and yet not dressed either. He did not really know what he wanted her to do, but when she came thus he was forced to accept it. The problem was of the form, Do something which I shall be inclined to call neither naked nor dressed. It is the same with a mathematical problem.[126]

For Floyd then, a mathematical conjecture is not a proposition which has a definite truth-value. It is rather a request for the clarification of the concepts involved, and in the case of trisection, the impossibility proof shows that our concepts prior to accepting it were confused (we were looking for the impossible). By using this example, Wittgenstein is trying to draw our attention to that our notion of what it is to understand a mathematical proposition is not as simple and clear-cut as we had first supposed, in particular, that the mathematical problem isn't clearly 'present' in our mind, in any precise sense, before we find its solution. In the next section we will see how this relates to Wittgenstein's discussion of Gödel.

Floyd's reading of (RFM I, App. III)

According to Floyd, Wittgenstein's aim in his remarks on Gödel is to make the concepts of 'mathematical conjecture' and 'mathematical proof' clearer. In this section we will take a closer look at her reading of (RFM I, App. III) in order to show how she believes Wittgenstein to achieve these goals.

The first three sections of (RFM I, App. III) do not directly relate to mathematics nor logic, rather Wittgenstein discusses the role of assertion and assumption, and as Floyd points out, were originally placed after a discussion of the use of expressions such as "I *can*" and "I *believe* I can" in the *Frühversion* of the *Philosophical Investigations*.[127] In §1, Wittgenstein says,

> It is easy to think of a language in which there is not a form for questions, or commands, but question and command are expressed in the form of statements, e.g. in forms corresponding to our: "I should like to know if..." and "My wish is that...".
>
> No one would say of a question (e.g. whether it is raining outside) that it was true or false. Of course it is English to say so of such a sentence as "I want to know whether...". But suppose this form were always used instead of the question?—[128]

[126]AWL, 185–186.
[127]Floyd 1995, p. 395.
[128]RFM I, App. III, §1.

The lesson Floyd wants us to draw from this passage is that even though it is perfectly good grammar to say of a question formulated as a statement (e.g. "I should like to know if it's raining outside") that it can be true or false, its role in our speech is not (or very rarely) to inform others of our mental states, but rather to *ask* a question. And this shows that we should be careful not be fooled in to thinking that every sentence that has the form of a proposition (or a statement) is true or false in the same way.[129] Wittgenstein's point here, on Floyd's interpretation, is that *if* we would suppose that this form were always used, we would see that while the grammatical form of questions would change (we would be able to say whether they are true or false, in a sense), their role in the language would not—they are still questions and would require answers.

This of course mirrors Floyd's discussion we saw above about mathematical conjectures having the grammatical form of propositions while not strictly playing that role in their use. And so, says Floyd, a seemingly mathematical statement such as "It is possible to trisect an arbitrary angle in Euclid's system", used in a context where a proof of this very impossibility has not been found, could be said to be true or false ("if one wishes"), but its real function is to act as

> a demand for clarification, the announcement that one is going to try to prove something, to change the circumstances of the "statement's" utterance. It says (like a command), "Go out and make a mathematical search!"[130]

In §2 and 3, Wittgenstein points out that assertion is not something that is "that get's added to the proposition, but an essential feature of the game we play with it"[131] and is, Floyd says, trying to get us to abandon the Fregean idea that the propositional content of a proposition is separate from its force. In §4, and the subsequent remarks, Wittgenstein brings this to bear on mathematical propositions:

> Might we not do arithmetic without having the idea of uttering arithmetical *propositions*, and without ever having been struck by the similarity between a multiplication and a proposition?
>
> Should we not shake our heads, though, when someone shewed us a multiplication done wrong, as we do when someone tells us it is raining?—Yes; and here is a point of connexion. But we also make gestures to stop our dog, e.g. when he behaves as we do not wish.

[129] Floyd 1995, p. 396.

[130] Ibid., p. 396. And here Floyd reminds the reader that in the *Tractatus*, Wittgenstein doesn't view mathematical propositions as real propositions at all—a view she maintains he never abandoned.

[131] RFM I, App. III, §2.

> We are used to saying "2 times 2 is 4" and the verb "is" makes this into a proposition and apparently establishes a close kinship with everything we call a 'proposition'. Whereas it is a matter only of a very superficial relationship.[132]

His purpose, as in the example of trisection, is to show us that using the notion of 'proposition' to describe mathematical statements is highly misleading. For Floyd, the result of this realisation is that one has "already broken away from the idea that we have a clear grasp of the concepts "mathematically true" and "mathematically provable" ".[133] And it is this idea—which Wittgenstein thinks is the result of the "disastrous invasion of mathematics by logic"[134]— he is really against: the idea that we have a clear grasp of such concepts as "provable", "true", "proposition" or even "mathematics". As a result, Floyd argues, only those who already think that these concepts are clearly demarcated would place philosophical importance on Gödel's proof that mathematical truth and mathematical proof cannot be equated.[135]

But this was never Wittgenstein's view. For him, Gödel's theorem is not a problem, it is merely an impossibility proof "which when misconstructed gives rise to a way of talking about mathematics he abhors."[136] And it is an impossibility proof in the sense that it shows that there are certain sentences (and their negation) in the symbolism of *Principia Mathematica* which are not derivable from the axioms of *Principia Mathematica*, just as the proof of the impossibility of trisecting an arbitrary angle shows that a certain construction is not possible in Euclid. And when Wittgenstein's interlocutor in §5 asks "Are their true propositions in Russell's system, which cannot be proved in his system"[137] his question should be understood as being asked *before* accepting Gödel's proof, and for Wittgenstein as fundamentally vague. He is, as Floyd puts it,

> concerned that the interlocutor thinks that a system-independent notion of "true proposition" is not only available, but required for an understanding of the interlocutory question.[138]

And by his answer "—What is called a true proposition in Russell's system, then?"[139] Wittgenstein tries to make clear—with some irony, Floyd claims—that what is needed is to understand what would count as an answer to this question and how Gödel's proof serves as such an answer in a mathematical way.[140] And

[132] RFM I, App. III, §4.
[133] Floyd 1995, p. 398.
[134] RFM V, §24.
[135] Floyd 1995, p. 400.
[136] Ibid.
[137] RFM I, App. III, §5.
[138] Floyd 1995, p. 401.
[139] RFM I, App. III, §5.
[140] Floyd 1995, p. 401.

Wittgenstein's answer is simple, "For what does a proposition's '*being true*' mean? '*p' is true* $=$ *p*. (That is the answer.)".[141] And as Floyd emphasises, this does not necessarily need to be read as a general analysis of the notion of truth, but rather as only applying in this context,[142] as Wittgenstein goes on to analyse how we assert propositions in 'Russell's system':

> So we want to ask something like: under what circumstances do we assert a proposition? Or: How is the assertion of the proposition used in the language game? And the 'assertion of the proposition' is here contrasted with the utterance of the sentence, e.g. as practice in elocution,— or as *part* of another proposition, and so on.
>
> If, then, we ask in this sense: "Under what circumstances is a proposition asserted in Russell's game?" the answer is: at the end of one of his proofs, or as a 'fundamental law' (Pp.) There is no other way in this system of employing asserted propositions in Russell's symbolism.[143]

For Floyd, Wittgenstein is trying to point out how we use the propositions of *Principia Mathematica*, or what the interlocutor called 'Russell's system' (and here Wittgenstein's use of the word 'language-game' is essential). What we should see is that in this particular language-game, *before* Gödel's proof, we would only have been inclined to assert a sentence p (or to say that it is true) if it stood at the end of a formal derivation (or was taken as an axiom). In other words, what is here called 'Russell's system' is a "a particular activity or language-game" in which the activity of 'proving' convinces us of the truth of what was proved.[144] 'Russell's system' is thus not merely the axioms and rules of *Principia Mathematica*, but also the activity of using those axioms and rules and applying them for certain purposes.

This is for Floyd 'precisely analogous' to the way Euclid's system is used to construct geometrical figures.[145] Wittgenstein's answer prompts the interlocutor to restate his question from §5 in a more precise way: "But may there not be true propositions which are written in this symbolism, but are not provable in Russell's system?".[146] Wittgenstein's answer again is a further clarification, he says,

> 'True propositions', hence propositions which are true in another system, i.e. can rightly be asserted in *another* game. Certainly; why

[141]RFM I, App. III, §6.
[142]Floyd 1995, p. 401.
[143]RFM I, App. III, §6.
[144]Floyd 1995, p. 402.
[145]Ibid.
[146]RFM I, App. III, §7.

> should there not be such propositions; or rather: why should not propositions—of physics, e.g.—be written in Russell's symbolism? The question is quite analogous to: Can there be true propositions in the language of Euclid, which are not provable in his system, but are true?—Why, there are even propositions which are provable in Euclid's system, but are *false* in another system. May not triangles be—in another system—similar (*very* similar) which do not have equal angles?—"But that's just a joke! For in that case they are not 'similar' to one another in the same sense!"—Of course not; and a proposition which cannot be proved in Russell's system is "true" or "false" in a different sense from a proposition of *Principia Mathematica*[147]

The point here is, Floyd says, that the question asked only has a sense "within a particular practice" and 'Russell's symbolism' could be used in a variety of ways (for instance to contain propositions of physics) but 'Russell's system' is just such a practice.[148] And here Wittgenstein does not mean, Floyd emphasises, that sentences in 'Russell's symbolism' are true or false relative to the system of *Principia Mathematica*, but rather that 'truth', 'provability' and other related notions only have a sense within a specific technique of use and application, and not in the formalism itself.[149]

Moreover, 'Russell's system' is not accorded any special place—it is not a generalisation of mathematics. It is merely a part of mathematics, a language-game among other language-games, and just as there are different notions of what similar triangles are in different systems of geometry, there are different notions of proof at play in mathematics. This, Floyd claims, shows that the notion of *Principia Mathematica* being incomplete is far from being uncomfortable for Wittgenstein and would have been his view even *without* Gödel's proof: the idea of a calculus not deciding every question capable of being asked in its language, and there being some other way of deciding its truth, is natural on this view—just as Euclidian geometry cannot decide the trisection problem, other systems can (this is of course not incompleteness in the exact same sense as Gödel's).[150]

The comparison with geometry plays another role as well. Before the development of alternative geometries in the nineteenth century it was easy to think that Euclidian geometry fully captured and described our "intuitive" notion of space, in the sense that the notion of there being true facts about space not captured by Euclidian geometry (much less false ones!) would not have been

[147]RFM I, Appendix III, §7.
[148]Floyd 1995, p. 403.
[149]Ibid. This is of course analogous to how Wittgenstein treats related notions in the philosophy of language in the *Philosophical Investigations*.
[150]Ibid., p. 404.

considered. We thought we knew, Floyd says, "*what* Euclid's theory was a theory of", but in fact we didn't. We simply never had such a clear notion of space, apart from how we described it in our theories (or rather, theory) and applied it in particular situations.[151] And for Floyd, Wittgenstein is calling into question a similar intuitive picture with his remarks on Gödel. On this picture, Gödel showed that no formal system can completely capture the notion of 'mathematical truth' (that is to say derive all sentences we intuitively call true) and by treating 'Russell's system' as a language-game and analogous to Euclidian geometry, we should see that our notions of 'mathematical truth', 'mathematical proof', 'mathematical proposition', etc. are not "independently and generally clear" and thus that 'general semantic notions of truth and consistency' play not vital role in Gödel's proof, just as they play no such role in the trisection impossibility proof (or indeed any indirect argument).[152]

And this brings us to her interpretation of the notorious paragraph.[153] Floyd admits that on the surface, this paragraph is nonsensical and would betray a misunderstanding of Gödel's proof. However, it is the interlocutor that gives the faulty version of Gödel's proof, one which emphasises the natural language interpretation of the sentence P—what Wittgenstein is doing is trying to

> depict the situation of someone making a mathematical search, preparing to accept the solution of a mathematical conjecture about incompleteness".[154]

The interlocutor's way of putting Gödel's reasoning makes it seem that the concepts "provable in Russell's system", "proposition" and "true" give us a clear interpretation of what this particular well-formed formula means in Russell's symbolism, but for Wittgenstein, no such interpretation exists outside the context of Gödel's proof itself: the sentence P which says of itself that it is provable, only does so in this particular case. Floyd says,

> In this case there is no application of an antecedently clear notion of truth or provability or proposition which is simultaneously a determining of what those notions themselves mean *here*.[155]

In other words, Gödel's proof itself clarifies the use of these terms that the interlocutor was so unclear about, in the same way the proof of the impossibility of trisection clarified the notions involved there. After Gödel's proof, these notions have become clearer and have thus a new sense—we are playing a new game

[151] Floyd 1995, p. 403.
[152] Ibid., p. 404.
[153] RFM I, App. III, §8.
[154] Floyd 1995, p. 409.
[155] Ibid., p. 405.

and we use the words "provable in Russell's system" and "true" in a different sense.[156]

But of course, Floyd says, people will insist that their notions of 'proof' and 'truth' have not changed, just as in the case of the trisection—we feel that our current notion of trisection is the same as the one the trisectors, looking for the impossible, had. But this does not, in her opinion, show that we did in fact *have* a clear idea of these notions before Gödel's proof. The interlocutor's question needs, like the question of whether there exists a general method of trisection, a reinterpretation to be answered, it asks for a mathematical search. And so, rather than "looking at the *Principia* as faulty or inadequate",[157] Floyd's Wittgenstein maintains that Gödel's proof changed the way we understand what it means to prove arithmetical statements and deriving them in a formal system, and it is therefore no wonder he draws parallels between they way we view attempted derivations of P and attempted trisections:

> A proof of unprovability is as it were a geometrical proof; a proof concerning the geometry of proofs. Quite analogous e.g. to a proof that such-and-such a construction is impossible with ruler and compass. Now such a proof contains an element of prediction, a physical element. For in consequence of such a proof we say to a man: "Don't exert yourself to find a construction (of the trisection of an angle, say)—it can be proved that it can't be done". That is to say: it is essential that the proof of unprovability should be capable of being applied in this way. It must —we might say—be a *forcible reason* for giving up the search for a proof (i.e. for a construction of such-and-such kind).
>
> A contradiction is unusable as such a prediction.[158]

The proof of the underivability of the Gödel sentence is then, for Wittgenstein, completely analogous to the proof that it is impossible to find a general method of trisecting an angle — just as we view trisectors who persist in their endeavours to find a proof as strange crackpots, we will not accept it as a reasonable goal to try to find a derivation of the Gödel sentence, and will warn people against trying to find such a derivation ("Don't exert yourself [. . .] —it can't be done"). We will even insist that any suggested derivation must be faulty or not a derivation in the relevant sense — and perhaps most importantly, we will not try to find a single axiomatic system to capture the whole of mathematics.[159]

[156] Floyd 1995, p. 405.
[157] Ibid.
[158] RFM I, App. III, §14. Floyd does not give an analogous close reading of §§9–13.
[159] It speaks volumes about the *very* different approaches different authors have had to Wittgenstein's remarks that in his paper *Mathematics and the "Language Game"*, Alan Ross

Wittgenstein's point with his remarks is then that, just as the proof of the impossibility of trisection changed our willingness to consider something constructible in geometry, Gödel's proof changed "the grounds of our willingness to "call" a certain sentence "unprovable" or "provable." "[160] We see things in a new light, and this

> to Wittgenstein is a way of thinking about the situation which does far better justice to its complexity than the idea that Frege and Russell simply made a mistake in conflating one sharply expressible concept ("mathematical provability") with another sharply expressible concept ("mathematically true"); just as to Wittgenstein his treatment of the problem of trisecting an angle does far better justice to its complexity than the idea that lots of people simply make geometrical mistakes.[161]

So, for Floyd, Wittgenstein's interpretation of Gödel's proof is not at odds with Gödel's own, as he states in his letter to Menger[162]—the proof is *not* a logical paradox and a perfectly uncontroversial piece of mathematics, "yielding a clarification of the question about whether there are "true but unprovable" statements of *Principia Mathematica*"[163]

He disagrees with Gödel, however philosophically, as he does not think that his proof shows anything important about the "underlying forms of our notions" of proof and truth, as to do so, we must already assume that there is such a thing.[164] Wittgenstein is thus trying to attack a certain philosophical interpretation of Gödel's theorem, not the theorem itself, and his aim in discussing it is then the same as in his later philosophy as a whole, namely to argue against the idea that the meaning of our words is either some 'extra thing' standing above or beyond them, or something present in the mind of the speaker, but rather get their meaning from "within the context of some practice, or ongoing system of use".[165]

Anderson reproaches Wittgenstein for *not* viewing Gödel's theorem as an impossibility result "like that one cannot trisect an arbitrary angle with straight-edge and compasses". See (Anderson 1958, p. 456).

[160] Floyd 1995, p. 407.

[161] Ibid., p. 408.

[162] See page 5.

[163] Ibid., p. 409.

[164] Ibid.

[165] Ibid., p. 410.

Wittgenstein as his own worst enemy: Steiner's criticisms

In his paper *Wittgenstein as his Own Worst Enemy*, Mark Steiner criticises Floyd's interpretation of Wittgenstein's remarks. He tries to show that despite his stated view, Wittgenstein misunderstands Gödel and consequently attempts to refute his proof, albeit inadvertently. His remarks are thus in fact attack on the autonomy of mathematics from philosophy. He believes, as Shanker, that Wittgenstein's motivation for writing his remarks on Gödel was that he correctly saw they could be taken as 'strengthening metaphysical Platonism'[166] However, he claims that his remarks are, as they stand, 'indefensible' and rather than attacking Gödel, he could have, given Floyd's analysis, used his proof to strengthen his own position. In this section and the next, I will examine these claims.

Steiner agrees with Floyd that "officially", Wittgenstein should have viewed Gödel's proof as a valid piece of mathematics and that it is his stated view to not interfere with affairs of mathematicians.[167] However, he sees in Wittgenstein's remarks on Gödel something more than simply a criticism of metaphysical assumptions, and it is here his real disagreement with Floyd lies. On her interpretation, the interlocutor's way of putting Gödel's reasoning in the notorious paragraph is supposed to be a version of Gödel's argument that puts it in a philosophically suspect way—superficially correct, but misleading. It is so because if one doesn't have Gödel's actual proof in mind when reading it, one might be misled into thinking that there are *absolute* notions of provability and truth in play, when in fact provability *in* a particular formal system is what is meant and—perhaps more importantly—that the theorem favours one notion of truth in mathematics over another, which Wittgenstein denies, and the subsequent discussion is supposed to show. Floyd puts this point forcefully in her *Prose* versus *Proof*:

> 'There are true but unprovable propositions in mathematics' is misleading prose for the philosopher, according to Wittgenstein. It fools people into thinking that they understand Gödel's theorem simply in virtue of their grasp of the notions of mathematical proof and mathematical truth. And it fools them into thinking that Gödel's theorem supports or requires a particular metaphysical view.[168]

Steiner however, believes that Wittgenstein was confused by Gödel's introductory remarks into thinking that they were a "synopsis of Gödel's argument",

[166]Steiner 2001, p. 258f.
[167]Ibid., p. 258.
[168]Floyd 2001, p. 299.

and therefore that Gödel's whole paper was "philosophy disguised as mathematics".[169] For Steiner then, the first part of the notorious paragraph is supposed to be a restatement of Gödel's proof and not just a statement of a philosophically suspect interpretation of it, an interpretation Wittgenstein thought was being bandied about carelessly, as Floyd and Shanker would have it. The second part of §8 is a refutation of that proof. Steiner paraphrases the statement of the proof and the subsequent refutation thus:

> [W]e have a sentence of mathematics, P, which can be interpreted: P is not provable. If P is false, then we have a provable, but false, sentence, which is impossible; so it must be true, but not provable.[170]

And:

> [T]here is no contradiction in a false, but provable sentence—what is false is context (or 'game') dependent. The very same words might sometimes express a truth and sometimes a falsehood. Thus, Gödel's proof rests on an elementary mistake.[171]

This, Steiner points out differs from Gödel's introductory remarks in an important way, despite the similarities, namely in that in his remarks, Gödel already *supposes* that the undecidability of P has already been established and then goes on to argue for its truth.

Thus, for Steiner, Wittgenstein is criticising an argument Gödel never made. However, as he points out, the argument Wittgenstein offers is very close to being an informal outline of a semantic proof of Gödel's theorem—which can, by using Tarski's definition of truth, be given a very precise mathematical treatment—and since Wittgenstein's supposed refutation is supposed to show that such arguments for incompleteness fail ("what is called 'losing in chess may constitute winning in another game"), he has unwittingly attacked a perfectly good mathematical proof, a mathematical proof that shows that his supposed refutation is misguided. What Wittgenstein failed to realise (or rather didn't anticipate) is that after Tarski we now have a mathematical theorem that states that if something is a theorem in Russell's system, it will be true in Tarski's sense, and furthermore, it is a theorem that P is true in Tarski's sense if and only if it is not provable in Russell's system. This mathematical treatment of truth for arithmetic is not arbitrary, Steiner points out, given some very basic assumptions about how such

[169] Steiner 2001, p. 272. Gödel wrote in is introduction: "The analogy with the Richard antinomy leaps to the eye. It is closely related to the 'Liar' too; for the undecidable proposition [...] states that [it] is not provable. We therefore have before us a proposition that says about it self that it is not provable [in PM]... From the remark that [the undecidable sentence] says of itself that it is not provable follows at once that [it] is true."(Gödel 1976, p. 598.)

[170] Steiner 2001, p. 261.

[171] Ibid.

a theory should look, and when applied to P, does not depend on any natural language interpretation.[172] It then follows, that Wittgenstein's argument cannot work, since "losing in Russell's system implies losing in Tarski's".[173]

To paraphrase Steiner yet again: Wittgenstein's refutation of Gödel's theorem involves arguing that if P is false in some other sense than Russell's, there is no contradiction in saying that the sentence is false in that sense, and yet provable in Russell's system. This is a poor argument because we have a natural extension of truth in mathematics, namely Tarski's, and a mathematical theorem which shows that P is true in that sense if and only if P is not provable is Russell's system. Given this treatment, the interlocutor's argument is perfectly good mathematics, and Wittgenstein has inadvertently criticised a mathematical argument.

For Steiner then, §14 is not evidence of Wittgenstein's insight that Gödel's proof can be viewed as an impossibility result, as Floyd thinks, but rather a mistaken response to interpreting P as a paradoxical sentence—a self-contradiction—much like the Liar-paradox, rather than reducing the assumption that P is provable to absurdity, as Gödel's actual, syntactic proof does.[174] Because of this misunderstanding, says Steiner, Wittgenstein believes that the correct response to Gödel's proof is to abandon the natural language interpretation of P and thus avoid the result, as a "mere interpretation" of a sentence can never make it unprovable in Russell's system. This is for Steiner what Wittgenstein meant by the much misunderstood remark in §14.[175] If this is the case, the correct interpretation of the remark Floyd cites as evidence for Wittgenstein's comparison of Gödel's theorem with geometric impossibility proofs shows the complete opposite. Wittgenstein is not, as Floyd thinks, *favourably* comparing Gödel's proof with other impossibility results, but in fact criticising Gödel's proof for *not* having the same character as a standard impossibility proof.

Steiner however agrees with Floyd's overall characterisation of Wittgenstein's philosophy of mathematics. He believes that Wittgenstein saw mathematical proof and mathematical truth as family resemblance concepts, devoid of "an eternal essence".[176] But unlike Floyd, he doesn't see Wittgenstein's remarks on Gödel as trying to throw light on these concepts, but as an attempted refutation of the theorem. What Wittgenstein should have done, claims Steiner, is to appropriate Gödel's theorem for his own ends. For him, Gödel's theorem could be used to support Wittgenstein's own philosophy. He could have argued in-

[172]Steiner 2001, pp. 277–278. What Steiner has in mind is that we want it to be the case that a "generalisation of φ is true if and only if φ is satisfied by every natural number" and further that this only holds for P if it is true.

[173]Steiner 2001, p. 267.

[174]Ibid., p. 263.

[175](RFM I, App. III, §14): "A contradiction is unusable as such a prediction [of improvability — ÁBM]".

[176]Steiner 2001, p. 260.

stead that since truth cannot equal provability in a formal system, mathematical truth "admits of flexibility";[177] he could have argued that because the concept of number cannot be captured by a recursively enumerable set of axioms, it has no essence, and lastly, he could have claimed that Gödel's theorem shows that the whole of mathematics cannot be formalised, and thus that it is in fact a "'motley' of techniques of proof".[178]

Wittgenstein should have used Gödel's theorem for his own purposes and used it to show how "the academic world overreacted"—going from the one extreme of logicism and formalism, to the other of mathematical platonism.[179] Instead, he misunderstands Gödel and blunders into attacking him, against his own views. For Steiner then, when Wittgenstein discussed Gödel, he was his own worst enemy—and what Floyd thinks he said, he didn't say but should have said.

Evaluation of Floyd's and Steiner's arguments

A central passage for Floyd's reading of (RFM I, App. III) is §14. It is here Wittgenstein mentions the trisection of an angle, and on Floyd's reading, compares it to Gödel's result. However, when giving her reading, Floyd does not explain the final remark of §14 that so annoyed the early commentators: "A contradiction is unusable as such a prediction".[180] This remark is however highly relevant in interpreting §14, as Steiner observed. For him, Wittgenstein sees P as a fundamentally paradoxical sentence, whose interpretation can be withdrawn in order to escape that paradox. And indeed, if §14 read in context, it is clear that Wittgenstein does think that Gödel's proof implies a contradiction *in* Russell's system, and that this contradiction cannot in itself compel us to abandon the system—in §11, Wittgenstein gives an argument that presumably is supposed to mirror Gödelian reasoning, showing what the contradiction is, and in subsequent paragraphs he further discusses it.

This is a problem for Floyd, and shows why she could not give a close reading of §§11–13. These remarks show that the context of the controversial remark in §14 is to show how P is *not* usable as an impossibility proof. In §11, Wittgenstein shows (or thinks he shows) how P leads to a contradiction in Russell's system, and in §§12–13 compares it to a "profitless performance" and derides it as a problem which "grows out of language".[181] In §14, he discusses how impossibility proofs, "of the trisection of an angle, say", contain an "element of prediction" and need to be capable of being applied in a certain way. The remark ends with him

[177] Steiner 2001, p. 261.
[178] Ibid. The phrase comes from (RFM III, §46).
[179] Steiner 2001, p. 273.
[180] RFM I, App. III, §14.
[181] RFM I, App. III, §§12–13.

proclaiming that a contradiction (presumably the one discussed in preceding remarks) is unusable for this purpose.[182] The most natural reading of §14 is therefore, *contra* Floyd, that Gödel's proof is *not* an impossibility result like the proof that there is no general method of trisecting an angle.

Steiner's reading of §14 fares no better, however. On his reading, there seem to be two refutations being offered in (RFM I, App. III): the refutation in §8, the notorious paragraph, which we can call the 'elementary mistake'–argument, summarised above by Steiner, and the 'no mere interpretation'–argument, which ends according to him in §14. However, Steiner seems to confuse these two refutations and mix them up. The former argument doesn't seem to crucially depend on the natural language interpretation of P to work, and yet this interpretation is the main focus of §8, and it is the only place where the natural language interpretation is given any emphasis, and the latter argument, which for Steiner crucially depends on such an interpretation, doesn't in fact need it. For him, the latter argument reaches its conclusion in the final remark of §14, the meaning of which is that the paradoxical interpretation of §8 cannot be used to establish an impossibility result.

But he is wrong in supposing that the roots of the argument being developed in §14 are to be found in §8. Rather, the start of the discussion of which §14 is a part starts at §11 and ends at §18. The overarching theme of the discussion found therein is that of contradiction and consistency—notions which are not mentioned in the preceding remarks. On the other hand, Wittgenstein's talk of 'withdrawing an interpretation' stops at §10 and is not mentioned again. In other words, §11 is a start of a new thread in his discussion of which §14 is only a part, and there is no refutation to be found there of Gödel's theorem and no important mistakes, as Steiner believes.

It is true that it still seems that Wittgenstein is labouring under the belief that Gödel's proof requires P to be interpreted in a certain way, but it doesn't really affect Wittgenstein's point here, and isn't explicitly mentioned. Wittgenstein writes,

> Let us suppose I prove the unprovability (in Russell's system) of P; then by this proof I have proved P. Now if this proof were one in Russell's system—I should in that case have proved at once that it belonged and did not belong to Russell's system.—That is what comes of making up such sentences.— But there is a contradiction here!—Well, then there is a contradiction here. Does it do any harm here?[183]

[182] RFM I, App. III, §14.
[183] RFM 1, App. III, §11.

His argument can be restructured slightly in the following way:

1. Let us suppose there exists a true but unprovable sentence P in Russell's system.

2. From the discussion in §§1–6, we know that a proposition can only be asserted in Russell's system if there is a proof of the proposition—this is how Russell plays his game.

3. From (1) we have that P is true, and thus from (2) that it is provable in Russell's system. But from (1) we also have that it is unprovable in Russell's system. We have then both asserted that P belongs to Russell's system and that it does not belong to Russell's system. Contradiction!

Wittgenstein then suggests that this contradiction is not a matter of concern. This is of course not Gödel's argument, but shouldn't be surprising in the least, since, after all, Gödel proved that if PM is consistent, P cannot be derived in it, and so if P were derived, PM is inconsistent.

Because of his emphasis on the natural language interpretation, however, Wittgenstein does make the mistake to suppose that by proving the unprovability of P in Russell's system, he has proven P in Russell's system. This is not the case, P is never proved in Russell's system (and this is what Gödel's theorem says)—it is rather that the truth of P is established, in Gödel's words, "by metamathematical means". His mistake is of course natural, if he thinks that the natural language interpretation of P is essential for Gödel's proof (and perhaps if he confuses Gödel's *syntactic* proof of the undecidability of P and Gödel's *semantic* proof of the truth of P, as he seems to do), and on this account Steiner is correct.

Wittgenstein is however not mistaken in saying that if a proof of P were to be found, Russell's system would be inconsistent, and thus if Russell insisted that the truth of P showed that it was provable, he would be faced with a contradiction. Wittgenstein's question in §14 is whether this inconsistency is a "forcible reason" to suppose that a proof of P cannot be found in Russell's system, and his answer is no: "A contradiction is unusable as such a prediction".[184]

This reading fits well with Wittgenstein's other remarks on the role of contradictions in formal systems which he believes are never (or very rarely) to be seen as a reason to abandon a formalism.[185] It also fits well with the preceding remarks, §§12–13, which make a similar point. On this reading, it is the contradiction itself that does not have this force, not the 'mere interpretation' of P

[184] Instead of this treatment, it could also be said that if a distinction between theory and metatheory is not made, Russell's system must be inconsistent.

[185] Why Wittgenstein believes this will be explained in the section on Berto and the dialetheist interpretation.

as a paradoxical sentence, as Steiner would have it. Wittgenstein is not trying to refute Gödel's theorem in §14, but questioning its consequences. When he asks what harm the contradiction does, he is trying to show that Gödel's proof doesn't in fact force us (or rather Russell) to abandon the notion that provability equals truth ('Russell's game'), because the contradiction doesn't have this force in of itself—his point is *not* that a mere interpretation of a contradiction doesn't have this force, as Steiner claims. That neither fits well with the text of (RFM I, App. III) nor Wittgenstein's other comments on contradiction.

But what about §8 and the 'elementary mistake'–argument? First of all, it should be noted that Steiner's argument depends on the fact that there is now a mathematical precise way of giving a semantic proof of Gödel's theorem, a way that did not exist when Wittgenstein wrote this paragraph. Wittgenstein writes that something could have been false in this new way, but proved in Russell's, when in fact it is impossible for P to be false and provable in standard model-theoretic semantics. But this way of putting the matter, in now way shows that Floyd's way of reading the paragraph is wrong, and can at most be used to criticise Wittgenstein for not foreseeing the way model-theoretic semantics were to be developed—not a very substantial criticism. Floyd's point all along was that Wittgenstein is trying to show that his interlocutor's words are not clear and he hasn't given mathematical content to his words. But isn't it a good way to show that something is unclear, to make it clear?

Wittgenstein's response to Steiner would presumably be, on Floyd's reading, to say that Tarski's definition of truth was simply more mathematics and that it was only *after* Tarski's work that this way of looking at Gödel's proof became clear. Thus in fact, Wittgenstein's point, on Floyd's reading, is seen even stronger after Tarski: The interlocutor's way of putting Gödel's reasoning is misleading because he hasn't given mathematical content to his words. But when his words *have* been given this new context, it becomes clear that their supposed metaphysical force is gone. The words 'true' and 'false' have been given mathematical content and we see that what the sentence "There exists a true but unprovable sentence" means is just that there is set of axioms in a certain formal language such that a certain sentence (or its negation) is not derivable from the axioms but is true in every model of the axioms. But this sense is *mathematical* and wasn't there *before*, as the intuitive picture would have it.

Steiner's arguments therefore do not show that Floyd's reading is wrong, but this does of course not mean that she is right. We have already seen that it is very unlikely that Wittgenstein wanted to compare Gödel's theorem to an impossibility result, similar to the proof of the impossibility of trisection, and in fact probably wanted to say the opposite. But what about Floyd's contention that he thought that philosophers abused Gödel's theorem for their own metaphysical purposes? It is very likely that he thought something of the sort. In 1935, he advised Moritz Schlick that,

> If you hear someone has proved that there must be improvable propositions in mathematics, there is in this *first of all nothing astonishing*, because you have as yet no idea whatsoever what this apparently utterly clear prose sentence says. You have therefore to go through the proof from A to Z in order to see what it proves. That is, before you have gone over this particular proof down to its last detail, you don't as yet know anything.[186]

It certainly seems that Wittgenstein was concerned with metaphysical abuse of Gödel's theorem.

However, if one reads (RFM I, App. III) and especially §8, it seems unlikely that this is what he wanted to say there. It is true that Wittgenstein wants his interlocutor to ask himself what he means by the words 'false' and 'provable', as Floyd's reading requires, but given his emphasis on the natural language interpretation of P and his subsequent insistence that his interlocutor should withdraw this natural language interpretation, it is hard to escape the reading of the early commentators who thought that Wittgenstein is in fact concerned with showing that Gödel's semantic argument somehow contains a mistake. In key paragraphs for Floyd's interpretation, namely §§8–10, the emphasis is on the abandoning of this natural language interpretation of P, not the meaning of the concepts 'false' and 'provable' (though the reader should be lead to the abandonment of the interpretation by considering their meaning). This is especially pertinent in §10, also ignored by Floyd.[187] The examination of the meanings of those terms seems to be only a way to the abandonment of the Gödelian interpretation of P, and not a goal in itself, as Floyd's reading suggests, and if Wittgenstein really just wanted to point out that Gödel's theorem was being abused by philosophers, it is not easy to see why he constantly emphasised that a certain interpretation of P should be withdrawn. What is certain, however, is that if Wittgenstein wanted to say what Floyd suggests that he does, he could have said it much more clearly.

[186] Nedo and Ranchietti 1983, p. 260.

[187] (RFM I, App. III, §10): " "But surely P cannot be provable, for supposing it were proved, then the proposition that it is not provable would be proved" But if this were now proved, or if I believed—perhaps through an error—that I had proved it, why should I not let the proof stand and say I must withdraw my interpretation?"

Floyd and Putnam's interpretation of the notorious paragraph

In their paper *A Note On Wittgenstein's "Notorious Paragraph" About the Gödel Theorem*, Juliet Floyd and Hilary Putnam give an interpretation of the notorious paragraph which, if they are correct, shows that Wittgenstein not only understood Gödel's proof but also "contains a philosophical claim of great interest".[188] Their claim is that Wittgenstein, with some fairly advanced mathematics in the background, shows that if it is supposed that $\neg P$ is provable, then the proposed interpretation of P as meaning "P is not provable" must be given up. In this section I will examine the whether this interpretation of Wittgenstein's text is plausible, but not the other question, which is somewhat in the background in their paper, of whether Wittgenstein did in fact have a correct understanding of Gödel's theorems.

The Floyd–Putnam attribution

Floyd and Putnam take the following part of the notorious paragraph as their focus:

> Now what does your "suppose it is false" mean? *In the Russell sense* it means, 'suppose the opposite is proved in Russell's system'; *if that is your assumption* you will now presumably give up the interpretation that it is unprovable. And by 'this interpretation' I understand the translation into this English sentence.[189]

These few lines do not contain much, and in themselves do not say why Wittgenstein thinks that the interpretation must be given up, but Floyd and Putnam reconstruct Wittgenstein's reasoning in the following way:

Assume that Russell's system (which we can just call PM) is consistent and suppose we have a proof of $\neg P$. Then, by Gödel's Theorem, we know that PM is ω-inconsistent. That simply means that there is some formula $\varphi(x)$ such that for every natural number n, $PM \vdash \varphi(\overline{n})$ but $PM \vdash \neg\forall x\varphi(x)$.[190] That is to say, PM proves $\varphi(1)$, $\varphi(2)\ldots$ and so on for each natural number but also proves that it is not the case for all elements in the domain that φ holds. But then it follows that PM can only be true on a model containing elements which are not natural numbers (since $\varphi(x)$ is true of all numbers, but not everything in the domain)—or as Floyd and Putnam put it, assuming that $PM \vdash \neg P$ shows that it "has no model in which the predicate we have been interpreting as 'x

[188] See Floyd and Putnam 2000.
[189] RFM I, App. III, §8.
[190] $\overline{n}$ represents the numeral of n from within PM.

is a natural number' possesses an extension that is isomorphic to the natural numbers."[191]

However, we interpret the Gödel sentence P as 'P is not provable' because we constructed it in such way, as Wittgenstein says, "by means of certain definitions and transformations", and in this construction the notion of the predicate $Prf(m, n)$ is essential, which is supposed to hold between m and n if and only if m is the Gödel number of a proof of the formula with the Gödel number n.[192] It then follows, Floyd and Putnam claim, that our intended translation of P is not tenable after all, since from ω-inconsistency it follows that

> [t]hose predicates of PM (for example, 'Proof(x,t)') whose extensions are provably infinite, and which we believed to be infinite subsets of N (the set of all natural numbers), do not have such extensions in any model. Instead, they have extensions that invariably also contain elements that are not natural numbers.[193]

To see why this is the case remember that the predicate $Prf(m, n)$ can be represented from within PM in such a way that there exists a formula $\varphi(\overline{m}, \overline{n})$ such that it is derivable in PM when $Prf(m, n)$ is true and its negation is derivable in PM when $Prf(m, n)$ is false. In other words, if we let this φ be called **Prf**, and shift our attention to PA, the following holds:[194]

$$\text{If } Prf(m, n) \text{ is true, then } PA \vdash \mathbf{Prf}(\overline{m}, \overline{n})$$
$$\text{If } Prf(m, n) \text{ is not true, then } PA \vdash \neg\mathbf{Prf}(\overline{m}, \overline{n}).$$

In other words, if m is the Gödel number of a proof of n, there is a formula of PM such that this formula can be derived in PM, and if m is not the number of a proof of n, there exists a formula such that its negation can be derived. We can call this numeralwise representability. From numeralwise representability and the assumption that PA is sound (i.e. for the set of natural numbers $\mathbb{N}$, it holds that $\mathbb{N} \vDash PA$), a semantical counterpart of this notion follows, which we can call arithmetic expressibility: For any two numbers n and m, $\mathbb{N} \vDash \mathbf{Prf}(\overline{m}, \overline{n})$ if and only if m is the code of the proof of the sentence which has the code n.

This notion, and the construction of P allows us to make sense of the idea that P 'says' of itself that it is unprovable. Note that it follows from arithmetic expressibility (and the assumption of the soundness of PA) that $PA \nvdash P$ and

[191] Floyd and Putnam 2000, p. 625.
[192] Floyd and Putnam's construction also uses a predicate $NaturalNo(x)$ which is interpreted as holding only for natural numbers. This is however not essential.
[193] Ibid.
[194] Wittgenstein, Floyd and Putnam talk about PM, but for convenience it is better to use PA as an example. Nothing depends on this, and we will use either PM or PA, whichever seems more appropriate.

$\mathbb{N} \models P$.[195] Now let $\mathbb{M}$ be one of the non-standard models that satisfy PA (remember that we concluded from the ω–inconsistency of PA that it is only satisfied by such models). Now suppose $PA \vdash \neg P$. Then $\mathbb{M} \models \neg P$. From the way P is constructed, we can write, $\mathbb{M} \models \exists x \mathbf{Prf}(\overline{x}, \overline{\ulcorner P \urcorner})$. In that case, there is some element $m \in \mathbb{M}$ such that $\mathbb{M} \models \mathbf{Prf}(\overline{m}, \overline{\ulcorner P \urcorner})$. But since $PA \vdash \neg P$, it follows from arithmetical expressibility that $\mathbb{N} \models P$. In that case it holds for each natural number n that $\mathbb{M} \models \neg \mathbf{Prf}(\overline{n}, \overline{\ulcorner P \urcorner})$. Thus, the element m is not a natural number, but rather one of the non-standard elements of $\mathbb{M}$. In that case, the predicate $Prf(m, n)$ cannot be interpreted as m is the code of the proof of the sentence which code is n, as the element m is not a number. But the translation of P as 'P is unprovable' depends on this interpretation.

In other words, the translation of P as 'P is unprovable' relies on being interpreted in a standard model of arithmetic, not containing any 'rogue' elements, and if PM is ω-inconsistent no such models are available. The proposed translation must then be given up, which, Floyd and Putnam claim, is exactly what Wittgenstein proposes. This, they say, is not in any contradiction with Wittgenstein's aim, which is not to refute Gödel's Theorem, but rather "by-pass it"[196] as nothing in Gödel's original proof needs this translation into "ordinary prose". This argument we can after Victor Rodych, call the 'Floyd-Putnam–attribution'.[197]

To labour the point further: They claim that Wittgenstein's intention in the notorious paragraph was to show that if "the opposite [of P, i.e. its negation] is proved in Russell's system", the interpretation that this says that 'P is not provable in Russell's system' must be given up. The reason Wittgenstein held this was that if $\neg P$ is provable, PM is ω-inconsistent, and thus P cannot be constructed in the right way for this interpretation to be possible. Now, as we saw above, some early critics had accused Wittgenstein of not understanding the role of consistency in Gödel's proof. If Floyd and Putnam are right, he not only understood, but correctly saw that if we do not assume the ω-consistency of our system of arithmetic, we cannot be sure that our interpretation of the Gödel sentence is the correct one (and indeed not of its incompleteness, as ω-consistency is, in Gödel's original syntactic proof, used to establish the underivability of $\neg P$).

Is this what Wittgenstein had in mind?

This is all well and good, but the question immediately presents itself whether or not Wittgenstein could have known about these matters and whether this is in fact the correct interpretation of the notorious paragraph. Aren't Floyd and

[195]For details on why this is the case, see (Bays 2004, p. 199.).

[196]Floyd and Putnam 2000, p. 625. Wittgenstein's original remark was (RFM VII, §19): "My task is, not to talk about (e.g.) Gödel's proof, but to by-pass it".

[197]See (Rodych 2003).

Putnam "overly charitable" in supposing that this is Wittgenstein's reason for arguing that if $\neg P$ is assumed to be provable, the interpretation of P as 'P is unprovable in Russell's system' must be given up?

Floyd and Putnam's answer is that Wittgenstein did know about these things. First of all, we have Goodstein's testimony, as we saw in the chapter on the early commentators,[198] that Wittgenstein saw "with remarkable insight" that Gödel's proof shows that non-standard models exist.[199]

The other reason that we know this, Putnam and Floyd claim, is that one of Wittgenstein's students, Alister Watson published a paper in *Mind* on the foundations of mathematics. Watson had attended Wittgenstein's lectures on mathematics and, by his own accounts, discussed Gödel's proof with both Wittgenstein and Alan Turing.[200] In the paper, Watson gives, in the words of Floyd, "a clear presenation of Gödel's result",[201] in addition to an overview of Cantor's diagonal argument and Turing's work on the *Entscheidungsproblem.* Watson's exposition of Gödel's result, despite containing some errors (we will examine Watson's paper a bit more in the next section), is overall fine. He starts by giving a quick overview of how Gödel constructs his unprovable sentence and then gives the following semantic argument for why it must be true:

> If we assume for the moment that this axiomatic system [of *Principia Mathematica*] is indeed a good basis for arithmetic, we shall have to conclude that the formula is not provable, and therefore, since this is just what it says, that it is true. For if it were provable, it would be false, and the system would be incorrect.[202]

This is a bit quick, but not at all incorrect. After sketching the argument, Watson goes on:

> This method of putting the argument, however, obscures rather than illuminates the point. Suppose we assume the falsity of the formula, we cannot of course, derive a contradiction, for this would amount to a proof of the formula. Instead, we reach the following peculiar situation, which is called by Gödel an ω-contradiction (ω is the 'ordinal number' of a sequence). We find that there is a function of a cardinal variable, say $f(n)$, such that (all on the basis of the falsehood of Gödel's formula) $(n)f(n)$ can be disproved, and yet we can convince ourselves we can prove in turn $f(0)$, $f(1)$, $f(2)$ and so

[198]See page 10.
[199]Floyd and Putnam express surprise that Goodstein didn't notice this connection in his original review. But, there is of course only a connection if their interpretation is right, and as we shall see, this is not obvious at all.
[200]Watson 1938, p. 445.
[201]Floyd 2001, p. 283.
[202]Watson 1938, p. 446.

on. In other words, we apply mathematical induction to the proofs of the system, and obtain $f(0)$, and from a proof of $f(n)$ for any particular value of n, a proof for $n+1$. [203]

He then goes on to argue that we would still maintain that the Gödel sentence is true, since if it were false, the ω-contradiction would entail that "the notion of a cardinal variable, i.e. of a number in the everyday sense, is something that cannot be completely expressed in the axiomatic system". In other words, the derivation of an ω-contradiction is a *reductio ad absurdum* of the assumption of the provability of $\neg P$. This, Floyd and Putnam claim, is an argument very similar to the Floyd–Putnam attribution, and since Watson got his knowledge of the theorems from Wittgenstein, this shows that this is what Wittgenstein had in mind.

The problem with this, however, is that the premises of this argument are simply not true. Wittgenstein, according to the Floyd–Putnam attribution, claims that ω-inconsistency forces us to give up the interpretation of P as 'provable in Russell's system' but Watson is using ω-inconsistency as an argument for the truth of the Gödel sentence. The two arguments are almost diametrically opposed. ω-inconsistency and the existence of non–standard models, admittedly, play a central role in both of them, but that in itself does not establish that Wittgenstein had what Floyd and Putnam attribute to him in mind. After all, according to Goodstein, 'Wittgenstein's insight' is from no later than 1935. Watson published his paper in 1938, so there would be plenty of time for him to learn of this result from others, especially since Skolem's original paper on non–standard models came out in 1934.[204]

Another problem with using Watson's paper as evidence for the Floyd–Putnam attribution is that Wittgenstein doesn't explicitly say that ω-inconsistency is the reason that the proposed interpretation of P should be given up, in fact, he doesn't mention ω-consistency at all. This makes Putnam and Floyd's argument very indirect and circumstantial (and perhaps circular, since Watson's paper has nothing to do with the Floyd–Putnam attribution, unless the Floyd–Putnam attribution is correct). The reason for this, Floyd and Putnam argue, is that, since Wittgenstein's remarks were originally intended as notes for himself, and not for publication (or so Floyd and Putnam claim), it is no wonder that they do not contain everything he knew about them. The mere fact that he knew that "variables must necessarily take on values other than the natural numbers" is supposedly enough.

[203]Watson 1938, p. 446–47. There are some technical errors in Watson's argument. Induction on the Gödel numbers of proofs is not used by Gödel, and cannot be used, to establish ω-inconsistency from the assumption of the provability of $\neg P$.
[204]See (Skolem 1934).

This is wonderfully unpersuasive, and a complete *non sequitur* as Wittgenstein doesn't mention non-standard models in the notorious paragraph any more than ω-inconsistency. Wittgenstein's knowledge of non–standard models could at most be a necessary condition for the truth of the Floyd–Putnam attribution, but never a sufficient one. We can see this even better when we keep in mind that the argument sketched above which shows that from ω-inconsistency it follows that the proof predicate cannot be given its standard interpretation is far from being trivial. The step from acknowledging that there exist non-standard models to seeing that if PM proves $\neg P$, then the proof predicate must admit non-standard elements, is simply to big and nothing in Wittgenstein's remark justifies this leap.

It is perfectly plausible that someone might have seen and understood a proof that there exist non–standard models of arithmetic and yet not see how the proof of $\neg P$ implies the ω–inconsistency of PM (and further not see how this entails that the proof predicate must necessarily contain a non–standard element). Showing that Wittgenstein *might* have had an understanding of non–standard models is no argument for him ever having understood why a proof of $\neg P$ entails ω–inconsistency nor of seeing why this makes the standard interpretation of P troublesome. Floyd and Putnam have simply not even tried to show that Wittgenstein had the Floyd-Putnam attribution in mind when he wrote the notorious paragraph—they haven't even shown that he could have understood it.

Furthermore, if Wittgenstein believes that the reason the given interpretation must be given up, on the assumption that $\neg P$ can be proven, is the non–standard interpretation of the proof predicate that follows from ω-inconsistency, why then does he also say that if "you assume that the proposition *is* provable in Russell's system" then "the interpretation *again* has to be given up" (emphasis mine)? There is no analogous argument that shows that if PM proves P, then the proof predicate must be interpreted in a non-standard way (which according to the Floyd-Putnam attribution is the reason the interpretation must be withdrawn), yet Wittgenstein's remark implies that the interpretation must be given up for similar reasons in both cases. This surely speaks strongly against Floyd and Putnam's interpretation.

Does Floyd and Putnam's interpretation make sense?

So far Floyd and Putnam have made no philosophically interesting point. It has all been mathematics. In this section I will explain what Floyd and Putnam believe to have been Wittgenstein's point with the notorious paragraph and examine whether or not it is a good argument that they attribute to Wittgenstein.

Now, as Floyd and Putnam point out, much of the hostility towards Wittgenstein's remarks in the notorious paragraph stems from the following sentence: "'True in Russell's system' means, *as was said*: proved in Russell's system; and 'false in Russell's system' means: the opposite has been proved in Russell's system.—Now what does your "suppose it is false" mean? *In the Russell sense* it means, 'suppose the opposite is proved in Russell's system"' (emphasis added by Floyd and Putnam). To see what Wittgenstein means by this, Floyd and Putnam say, we need to take a look at the preceding paragraph, which 'sets the scene'.

> "But may there not be true propositions which are written in this symbolism, but are not provable in Russell's system?"—'True propositions', hence propositions which are true in another system, i.e. can rightly be asserted in *another* game. Certainly; why should there not be such propositions; or rather: why should not propositions—of physics, e.g.—be written in Russell's symbolism? The question is quite analogous to: Can there be true propositions in the language of Euclid, which are not provable in his system, but are true?—Why, there are even propositions which are provable in Euclid's system, but are *false* in another system. May not triangles be—in another system—similar (*very* similar) which do not have equal angles?—"But that's just a joke! For in that case they are not 'similar' to one another in the same sense!"—Of course not; and a proposition which cannot be proved in Russell's system is "true" or "false" in a different sense from a proposition of *Principia Mathematica*.[205]

Here, according to Floyd and Putnam we need to note two things, first that Wittgenstein is pointing out that PM is a system in the same sense as non–Euclidian geometry is a system of geometry, and in two different systems a sentence can have two different senses, and secondly that this remark does not deny that a sentence that cannot be decided in Russell's system can be 'true' or 'false' in some sense outside the system—it's just that the two senses are different.

The former point, they claim is directed at Frege and Russell, who did not "see themselves as providing a mere notation"[206] into witch the utterances of mathematicians can be translated, but rather as providing an 'ideal language' that replaces natural language, and in the end provides a foundation for mathematics. For Frege and Russell then, ordinary language only serves to 'lead someone into'[207] the ideal language and it would be, as Floyd and Putnam say,

[205] RFM I, Appendix III, §7.
[206] Floyd and Putnam 2000, p. 630.
[207] Ibid.

> utterly foreign to this spirit to explain the truth of a formula of *Principia Mathematica* by merely writing down an English sentence, and saying that this is what it means for P to be true. To confess that this what one has to do would be to abandon the claim for the foundational status of a system such as *Principia Mathematica* entirely.[208]

So, Floyd and Putnam claim, when Wittgenstein writes in (§8) "And by 'this interpretation' I understand the translation into this English sentence", he his denying that this idea makes sense—there can not be an interpretation which can only be indicated in natural language indirectly and is completely independent from natural language. Rather, Wittgenstein is emphasising that a formal system can be free-standing, but not as an ideal language. It is free-standing *as* formal system, but then the only notion of truth we have available is 'being a theorem' ("a proposition which cannot be proved in Russell's system is "true" or "false" in a different sense from a proposition of *Principia Mathematica*").

The latter point is that Wittgenstein does not deny that there can be true but unprovable propositions, but that if someone were to say that when they say that P is a true but unprovable proposition in the language of *Principia Mathematica*, they mean the translation of that sentence into English, Wittgenstein would say that the idea of truth is eliminable here and the only thing she has said is that P is unprovable — which is the very translation into English proposed for P. So, for Floyd and Putnam, Wittgenstein is fully aware that Gödel's proof is a valid mathematical proof, it is rather that Wittgenstein is pointing out to his imagined interlocutor that his way of putting Gödel's argument 'obscures rather than illuminates', borrowing Watson's phrase, the point of Gödel's proof—the point being that Gödel only showed that if PM is consistent, P is unprovable and if PM is ω-consistent, $\neg P$ is unprovable, and the interlocutors way of speaking might make it seem that something more, perhaps metaphysical, had been proven, thus obscuring the point by making it seem more mysterious than it actually is that a sentence can be true yet undecidable.

Wittgenstein's goal in the notorious paragraph then is to demonstrate that the idea that Gödel showed that there is a well defined notion of mathematical truth and that, if PM is consistent, then there are mathematical truths in that sense which PM cannot decide, "is *not* a mathematical result but a metaphysical claim".[209] In other words,

> that if P is provable in PM then PM is inconsistent and if $\neg P$ is provable in PM then PM is ω-inconsistent is precisely the mathemat-

[208]Floyd and Putnam 2000, p. 630.
[209]Ibid., p. 632.

ical claim that Gödel proved. What Wittgenstein is criticising is the philosophical naiveté involved in confusing the two, or thinking that the former follows from the latter. But not because Wittgenstein wants simply to simply to *deny* the metaphysical claim; rather he wants us to see how little sense we have succeeded in giving it."[210]

This argument has a few problems. Floyd and Putnam see Wittgenstein as giving an argument which takes as its starting point a certain interpretation of the sentence P, namely that P itself states that 'P' is unprovable. This interpretation depends on at least two things. First, that the language of PA captures our notion of arithmetic, i.e. '+' is addition, '×' is multiplication, the quantifiers run over the domain of natural numbers, etc., and secondly, that PA is sound, i.e. $\mathbb{N} \models PA$ in more modern model–theoretic terms.[211] If we suppose then that $PA \vdash \neg P$, it follows, as we saw above, that $\mathbb{N} \not\models PA$. According to Floyd and Putnam, then, we must abandon our interpretation of P, since it is based on the assumption that our interpretation of arithmetic captures our intuitive notion of arithmetic, which it cannot do if $\mathbb{N} \not\models PA$.

Timothy Bays points out that Floyd and Putnam are almost certainly wrong about the reaction of mathematicians were a proof of $\neg P$ discovered.[212] It is, as he says, "almost unimaginable" that they would accept non–standard models as their canonical interpretation of the language of arithmetic, nor would they accept PA as an adequate axiomatization of arithmetic (naturally, since it would be provably ω–inconsistent, and therefore, by definition, not describe the natural numbers). It is far more likely that mathematicians would find ways of modifying PA in such a way that the proof of $\neg P$ would be blocked in someway. Those of them studying arithmetic would still mostly be interested in axiomatizations that can satisfy $\mathbb{N}$ (although some might still be interested in PA, but probably with the caveat that its models *are* non–standard), and if PA cannot do this job, it would simply be modified. This, Bays points out, is a good reason to think that Floyd and Putnam's reasoning is flawed.[213] In any case, a proof of $\neg P$ would not force us to abandon the proposed interpretation, other options are far more likely and reasonable.

Bays makes two further points that we should consider. First that it is this *very* interpretation that makes Gödel's theorem interesting in the first place. It shows that a *standard* interpretation of an axiom system of arithmetic is *intrinsically* incomplete, and not just some random selection of axioms. There are in fact many incomplete systems, many of them uninterestingly so.[214] Second

[210] Floyd and Putnam 2000, p. 632.
[211] We keep using PA for our example.
[212] Bays 2004, pp. 202–203.
[213] Ibid.
[214] For instance the theory $\varnothing$ in the language of propositional logic. It is quite trivially and

that the interpretation of P as 'P is not provable' has very close ties to Gödel's own work and his perfectly mathematically respectable, there are for instance perfectly good semantic proofs of the theorem (which Gödel himself knew). To say that there is a true but undecidable sentence of PA is perfectly good mathematics, and so there are no "metaphysical claims on the horizon".[215] In other words, despite that this particular—natural language—interpretation of P is not necessary for Gödel result (even though Wittgenstein might have thought so), it is in fact not mathematically problematic and no metaphysical conclusions could be drawn from it, even if this was the case. Floyd and Putnam's argument shouldn't even get off the ground then—the source of metaphysics surrounding Gödel's proof is not to be found in any particular interpretation of P.[216]

But even if we overlook these problems, if they are indeed problems, it cannot be seen that the argument Floyd and Putnam attribute to Wittgenstein makes much sense. This argument seems to be that if a proof of $\neg P$ were to be found, the translation of P would have to be given up because PM would be ω–inconsistent and therefore the construction of the interpretation that says 'P is not provable' would be illegitimate. This is supposed to show, as we saw above, that the way Wittgenstein's interlocutor puts Gödel's reasoning 'obscures rather than illuminates' the point of his proof. This in turn is supposed to enlighten us to the fact that no metaphysical claims about mathematical truth can be given support by Gödel's proof.

The problem is that the fact that we would have to abandon the interpretation of P does not show that the interlocutor's way of putting Gödel's reasoning obscures the point. While it is perfectly plausible that it is Wittgenstein's position that no metaphysical conclusions can be drawn from Gödel's proof, the notion of the correct interpretation of P plays no role in here. If someone did *draw* metaphysical conclusions from Gödel's proof, Wittgenstein's supposed contention that we need to abandon the interpretation of P if we had a proof of $\neg P$ is unlikely to change their mind, since they already assume that this is impossible in order to draw those conclusions, and if proof was actually found, it would only be true that the *same* metaphysical conclusions could not be drawn (since P would no longer be true and thus the conclusions people want to draw from *that* would no longer be possible.) Those who do draw metaphysical conclusions from Gödel's proof simply assume PM to be ω–consistent—telling them that they couldn't draw the same conclusions as they do if their premises were false is unlikely to

boringly incomplete (in the Gödelian sense), since $\varnothing \nvdash p$ and $\varnothing \nvdash \neg p$.

215 Bays 2004, p. 10

216 Floyd and Putnam give a reply to Bays' paper in (Floyd and Putnam 2006) and he in turn answers them in (Bays 2006). This further debate involves other claims Floyd and Putnam ascribe to Bays, and needs not to concern us here, as Bays' point's mentioned here remain unchallenged.

sway them. To say that one cannot believe q on the basis of p, if p would be false, is saying nothing.

Furthermore, if it was Wittgenstein's intention to use the supposed proof of $\neg P$ to draw attention to the fact that Gödel's proof cannot support metaphysical claims, he doesn't need this argument at all — it would perfectly suffice for him to give Gödel's syntactic proof and then point out that truth is eliminable in the sentence "'P is unprovable' is true", since if P is true, that only means that P is the case, and since we interpret P as meaning that 'P is unprovable', that is the only thing we are left with, and exactly what Gödel's syntactic proof proves. This already establishes what Floyd and Putnam want Wittgenstein to say, namely that it is mathematically fine to say that there are true but unprovable propositions of PM, and yet there is nothing mysterious about this fact.

Floyd and Putnam of course do claim, on the basis of §7,[217] that Wittgenstein thinks that the notion of truth is eliminable here but this is not the same argument as I just sketched (and Floyd and Putnam somewhat give at the end of their paper).[218] On their reading the fact that the proof $\neg P$ forces us to abandon the interpretation of P as 'P is not provable' is essential, and this, as we have seen, is not needed for Wittgenstein to make the point Floyd and Putnam want him to make. These considerations, coupled with the textual problems facing Floyd and Putnam, should suffice to convince the reader that their interpretation of §8 is *not* the correct one.[219]

[217]It should rather be based on §6, however: "For what does a proposition's '*being true*' mean? *'p' is true* $= p$. (That is the answer.)" (RFM I, App. III, §6). §7 doesn't mention this at all.
[218]See Floyd and Putnam 2000, p. 632.
[219]And indeed the strange fact that this interpretation of the notorious paragraph looks nothing like Floyd's other interpretation of Wittgenstein's remarks. It's almost inconceivable that they could both be true.

Wittgenstein's inversion of Gödel: Rodych's interpretation

In his paper, *Wittgenstein's Inversion of Gödel's Theorem*, Victor Rodych argues that the main reason Wittgenstein's remarks on Gödel were so ill-received by the early commentators was not that Wittgenstein rejected the standard interpretation of Gödel's theorem, but rather a certain mistake he made in his discussion of it. This mistake, Rodych claims, is that Wittgenstein often writes as if Gödel's result was primarily based on an "extra-systemic, *natural language interpretation*" of the sentence Gödel proves to be undecidable. This mistake, he says, while unfortunate, does not affect most of what Wittgenstein had to say is about the theorem.

Rodych further tries to show that Wittgenstein had three goals in mind when he wrote his remarks, namely that (a) on his own terms, there cannot be any such thing as a 'true but unprovable' proposition, that (b) it is "highly questionable" whether the Gödel sentence P has any meaning in the first place and (c) that even given the standard interpretation of Gödel's theorem, Gödel's underivable sentence "may or may not be" derivable after all, because Gödel has not proven the consistency of the system in question. In this chapter we will see how Rodych motivates his interpretation, including what Wittgenstein's mistake consisted in, and finally evaluate and criticise it.

Gödel's theorem on Wittgenstein's terms

Like Shanker, Rodych sees Wittgenstein's remarks on Gödel being informed by views only "explicitly articulated in the middle period".[220] The most important feature of his account, which carries over to the later period, according to Rodych, is Wittgenstein's belief that the meaning of a mathematical proposition is solely determined by its position in a calculus.[221] Rodych holds a particularly strong version of this view, as the following shows:

> The central idea of Wittgenstein's view of mathematics is, to the contrary, that everything is syntax, nothing is semantics (PG 468). A true mathematical proposition is calculus-specific and either proved within its calculus, or perhaps provable within its calculus.[222]

On this strong formalist view of Wittgenstein's mathematics, we only call mathematical propositions as such because the have a determined status in a certain

[220] Rodych 1999, p. 174.
[221] This is also Shanker's view. See the section on him for criticism of this view.
[222] Ibid., p. 195. The relevant parts of (PG, p. 468) read: "Mathematics consists entirely of calculations. In mathematics *everything* is algorithm and *nothing* is meaning;".

calculus and can be used to make assertions in that calculus.[223] They don't refer to any external reality or facts, and talk of 'true' and 'false' is eliminable, i.e. a mathematical statement is true if and only if it can be asserted in a given calculus.

This is for Rodych the background of the first four remarks of (RFM I, App. III). In them he sees Wittgenstein as trying to show that not every sentence with the grammatical form of a statement is to be judged true or false in the same way.[224] In the case of propositions, or statements, proper, we usually use them to make assertions and when we do so, they are said to be either true or false, and the logic of truth–functions applies to them. When we play this game, as it were, we are inclined to say that a statement is true if and only if a fact corresponds to it (which may or may not be a misleading way of speaking).

However, this is not the case with mathematical statements. When we speak of mathematical statements as being either true or false, it is not the case that their truth value is determined by some correspondence with facts (or perhaps more precisely, correspondence with an external reality). It is rather, Rodych claims, that when we say that a mathematical statement is true or false, the truth value is not necessarily determined at the moment of utterance, but it can become so when we've proven or refuted it.[225]

There is however a "point of connexion" between propositions and mathematical statements, namely that our use of the dichotomy between 'true' and 'false' in the former is mirrored in the latter.[226] The connection is a superficial one, because the notion of truth is very different here, and this is what Wittgenstein is showing us with his remark that we could "do arithmetic without having the idea of uttering arithmetical propositions". Wittgenstein makes this point as well at (RFM I, §§143–144). He says,

> If somebody calculates like this must he utter any 'arithmetical *proposition*'? Of course, we teach children the multiplication tables in the form of little *sentences*, but is that essential? Why shouldn't they simply: *learn to calculate*? And when they do so haven't they learnt arithmetic?[227]

In other words, that we do in fact utter arithmetical propositions is inessential to the practice of arithmetic, and this shows that there is a difference between making the contingent statements we usually do when we 'play the truth-function game', where this is essential, and arithmetical statements, where it is not — call-

[223] Rodych 1999, p. 195.
[224] Ibid., p. 177. See page n. or the appendix for a full quotation of these remarks.
[225] Ibid., p. 179.
[226] RFM I, App. III, §4.
[227] RFM I, §144.

ing a statement such as '$5+7=13$' false, is more akin to calling the calculation that lead to it, wrong. All that we need, Rodych says, is

> a dichotomy of '$+$' vs '$-$', 'right' or 'good' (doggy) vs. 'bad' (doggy), since this is simply all there is to the "superficial relationship" between contingent truth and mathematical truth.[228]

There is then nothing more to the relationship between propositions and mathematical statements—we are merely "misled" by the grammatical form of statements such as '2×2 is 4' into thinking that their truth-value is decided in the same way, by correspondence. This, for Rodych, is what Wittgenstein means when he suggested in a lecture that perhaps it would be best to avoid words like 'true' and 'false' completely in these cases, so we will be become clear that "to say that p is true is simply to assert p; and to say that p is false is simply to deny p or to assert $\sim p$."[229]

The questions Wittgenstein poses in §5 should be read in this light. There he asks,

> Are there true propositions in Russell's system, which cannot be proved in his system?—What is called a true proposition in Russell's system, then?[230]

To answer this question, Wittgenstein first points out that in the context of Russell's system, "'p' is true = p". And hence, the question of what a true sentence in Russell's system is, simply becomes a question of what sentences we are prepared to assert in his system. And the answer to this question, Wittgenstein says, is that we only assert sentences in Russell's system if they stand

> at the end of one of his proofs, or as a 'fundamental law' (Pp.). There is no other way in this system of employing asserted propositions in Russell's symbolism.[231]

For Rodych however, Wittgenstein is stressing that *he* believes that there can be no way of asserting a proposition in mathematics except in these ways. To support this claim, Rodych points to a passage in *Philosophical Remarks* where Wittgenstein says,

[228] Rodych 1999, p. 179. The word 'doggy' in parenthesis seems a little bit baffling, but it mirrors Wittgenstein's remark that we can also shake our heads when someone multiplies wrongly as we do when someone makes a wrong statement — "we also make gestures to stop our dog, e.g. when he behaves as we do not wish", but that doesn't make his actions into statements.
[229] LFM, p. 188.
[230] RFM I, App. III, §5.
[231] RFM I, App. III, §6.

> A mathematical *proposition* can only be either a stipulation, or a result worked out from stipulations in accordance with a definite method.[232]

This, for Rodych, obviously mirrors Wittgenstein's remark about how propositions are asserted in Russell's system.

Furthermore, since the later Wittgenstein's still held the view that the meaning of a mathematical proposition is solely determined by the rules of the calculus it belongs to, and since Wittgenstein equates provability with truth, Rodych's conclusion is that Wittgenstein's answer to the question posed in §5 is that, *on is own terms*, 'an unprovable but true sentence' is a contradiction in terms—every true mathematical statement "must be proved (or, at least, provable) in some particular calculus."[233] That is to say, being a true proposition in Russell's system means to be a proved, or provable, proposition in Russell's system. There is therefore no room for anything called a true but unprovable proposition in his system because being true means being provable, and vice versa.

Wittgenstein's mistake

As a result of the discussion recounted in the last section, Wittgenstein's interlocutor puts the question differently, and in a more precise way. He asks: "But may there not be true propositions which are written in this symbolism, but are not provable in Russell's system?"[234] Since Wittgenstein argued in §6 that provability is the same as truth in Russell's system. Wittgenstein's answer, Rodych says, is that if there are in fact true propositions in his system, that cannot be proved in the same system, they must be true in some other system instead.[235] And that means that it can be *proved* in another system. Rodych explains:

> Wittgenstein is now saying, "Well if there *are* true propositions which are not provable in Russell's system, they must be 'true in *another* system', which means they must have been proved (or be provable) in another system".[236]

Rodych uses the following example to illustrate what Wittgenstein means:

> For example, the most obvious way to make the Gödelian proposition 'P' proved (or provable) is to simply add it to the axiom set of Peano

[232]PR, §202.
[233]Rodych 1999, p. 180.
[234]RFM I, App. III, §7.
[235]Rodych 1999, p. 180. See (RFM I, App. III, §7).
[236]Rodych 1999, p. 180–181.

> Arithmetic (PA), such that we create a new calculus, call it "PA + P". Now, 'P' is not provable in PA, but it *is* provable in PA + P.[237]

In this example, says Rodych, it follows that P is now true or false in a different sense than in PA, which is precisely what Wittgenstein meant.[238]

In the next remark, the notorious paragraph, Wittgenstein goes even further, says Rodych, by trying to show that in order to avoid Gödel's result that P is true but unprovable, it is enough to give up "the interpretation that it is unprovable".[239] Wittgenstein's interlocutor asks,

> I imagine someone asking my advice; he says: "I have constructed a proposition (I will use 'P' to designate it) in Russell's symbolism, and by means of certain definitions and transformations it can be so interpreted that it says: 'P is not provable in Russell's system'. Must I not say that this proposition on the one hand is true, and on the other hand is unprovable? For suppose it were false; then it is true that it is provable. And that surely cannot be! And if it is proved, then it is proved that it is not provable. Thus it can only be true, but unprovable."[240]

Here, Wittgenstein's interlocutor runs two *reductio*-arguments in parallel. First he shows that P must be true, since if it were false, it must be provable and hence true, resulting in a contradiction, and then he shows that it must be unprovable, since if it were provable, it would be false, "and that surely cannot be!"[241] It is worth it to repeat Wittgenstein's response yet again:

> Just as we ask, " 'Provable in what system?", so we must also ask: " 'True in what system?" 'True in Russell's system' means, as was said: proved in Russell's system; and 'false in Russell's system' means: the opposite has been proved in Russell's system.—Now what does your "suppose it is false" mean? *In the Russell sense* it means, 'suppose the opposite is proved in Russell's system'; *if that is your assumption* you will now presumably give up the interpretation that it is unprovable. And by 'this interpretation' I understand the translation into this English sentence.—If you assume that the proposition is provable in Russell's system, that means it is true *in the Russell sense*, and the interpretation "P is not provable" again has to be given up. If you assume that the proposition is true in the Russell sense, *the same*

[237]Rodych 1999, p. 181.
[238]Rodych's example is a little bit unfortunate, since P cannot be "true or false in a different sense" in PA+P than it was in PA, since *ex hypothesi*, it wasn't true at all in PA to begin with!
[239]RFM I, App. III, §8.
[240]RFM I, App. III, §8.
[241]RFM I, App. III, §8.

thing follows. Further: if the proposition is supposed to be false in some other than the Russell sense, then it does not contradict this for it to be proved in Russell's system (What is called "losing" in chess may constitute winning in another game.)[242]

Here Wittgenstein makes the mistake that so upset the early commentators. The mistake is of course that Gödel's proof of his theorem does not require any specific, natural language, interpretation to give up in the first place. What Gödel does in his proof is to show that by "certain definitions and transformations", he can construct a purely arithmetical sentence P *within*, say, PA, such that PA is ω-inconsistent, if either P or its negation is derivable. Gödel then shows that P is true if and only if it is underivable, and in establishing this relationship only its number-theoretic properties are needed, or as Rodych puts it,

> To show, meta- mathematically, that 'P' is true if it is unprovable, we need only show that a particular number-theoretic proposition, say [R(q); q], is true iff a particular number-theoretic proposition, say [R(q); q], is unprovable (in Russell's system). It is entirely unnecessary to give [R(q); q] a natural language interpretation to establish the bi-conditional relationship.[243]

Therefore, when Wittgenstein says, "if *that* [i.e. 'suppose P is false' means 'suppose the opposite is proved in Russell's system'] is your assumption, you will now presumably give up your interpretation that it is unprovable",[244] Gödel's immediate reply would be, "Well, yes, I would 'now presumably give up the interpretation that it is unprovable' if that were my assumption – but it *isn't* – it's your assumption".[245] Wittgenstein seems to be labouring under the misconception that he can show, by supposing that if the negation of P meant 'Suppose it is true that it is provable', that a contradiction is not forthcoming if P were supposed to be false—since this would only mean that it were false in some other sense than Russell's. And this is exactly what an orthodox interpreter of Gödel's proof would deny, for him, Gödel's proof exactly shows that the notions of proof in Russell's system and truth in Russell's system cannot be equated.

§8 is not the only remarks where this mistake occurs. Wittgenstein repeats the mistake, Rodych says, at §10, and probably also in §11 and §17. In §10 he wrote:

> "But surely P cannot be provable, for, supposing it were proved, then the proposition that it is not provable would be proved". But if this

[242]RFM I, App. III, §8.
[243]Rodych 1999, p. 182.
[244]RFM I, App. III, §8.
[245]Rodych 1999, p. 182.

were now proved, or if I believed—perhaps through an error—that I had proved it, why should I not let the proof stand and say I must withdraw my interpretation "unprovable"?[246]

And in §17, we find:

> Suppose however that not-P is proved.—Proved *how*? Say by P's being proved directly—for from that follows that it is provable, and hence not-P . What am I to say now, "P" or "not-P"? Why not both? If someone asks me "Which is the case, P, or not-P¿" then I reply: P stands at the end of a Russellian proof, so you write P in the Russellian system; on the other hand, however, it is then provable and this is expressed by not-P , but this proposition does not stand at the end of a Russellian proof, and so does not belong to the Russellian system.[247]

Here, Rodych says, Wittgenstein supposes that P has been proved 'directly', and through the meta-mathematical interpretation P—i.e. the natural language interpretation, this shows that $\neg P$ can be proven, and we are forced to accept a contradiction, and as with §8, Wittgenstein seems to think that the natural language interpretation is essential in Gödel's reasoning.[248] This is without doubt an unfortunate mistake but Rodych emphasises that one should not let it "blind us to whatever merit Wittgenstein's remarks may have" and we now turn to examining what he believes that consists in.

The meaningfulness of P and the consistency assumption

A key passage for Rodych's interpretation is §11, where Wittgenstein wrote,

> Let us suppose I prove the unprovability (in Russell' system of P; then by this proof I have proved P. Now if this proof were one in Russell's system—I should in that case have proved at once that it belonged and did not belong to Russell's system—That's is what comes of making up such sentences.—But there's a contradiction here!—Well, then there is a contradiction here. Does it do any harm here?[249]

[246] RFM I, App. III, §10.
[247] RFM I, App. III, §17.
[248] Rodych speculates that this misunderstanding of Wittgenstein has it roots in Gödel's opening remarks to his paper, where he wrote, "From the remark that [R(q); q] says about itself that it is not provable it follows at once that [R(q);q] is true."(Gödel 1976, p. 598.) According to (Dawson 1988), many readers of Gödel's paper read no further than his introduction and were thus misled by this "informal précis". Wittgenstein might very well have been one of them, since according to Kreisel, Wittgenstein had admitted to him that he hadn't read Gödel's paper before 1942, because he had been "put off by the introduction". See (Kreisel 1998).
[249] RFM I, App. III, §11.

In this short, cryptic remark, Rodych sees Wittgenstein as making essentially three different points, developed further developed in other remarks. The first point we have already seen, and is illuminated by §17 (quoted above). If we suppose that we can prove P 'directly', we thereby, through the natural language interpretation of P, have proven $\neg P$. This is essentially Wittgenstein's mistake. But the two other points do not contain mistakes, Rodych claims, and we shouldn't discount them because of it.

The second point made in §11 has to do with the meaningfulness of mathematical propositions. In the intermediary period, Wittgenstein was a finitist, Rodych claims, and believed that any number-theoretic statement that quantifies over an infinite domain is strictly speaking meaningless,[250] and for Rodych, this is a position he still held in the later period.[251] It is true, Rodych admits, that Wittgenstein never *explicitly* questions the meaningfulness of such propositions in (RFM), but he does in fact do so *implicitly.* In (RFM VI, §13), Wittgenstein writes:

> Now isn't is absurd to say that one doesn't understand the sense of Fermat's theorem?—Well, one might reply: the mathematicians are not *completely* blank and helpless when they are confronted with this proposition. After all, they try certain methods of proving it; and, so far as they try methods, *so far* do they understand the proposition.—But is that correct? Don't they *understand* it as completely as one can understand it?[252]

Since Wittgenstein often used Fermat's Last Theorem as an example of a sentence which is meaningless because it involves quantifying over an infinite domain, this quote shows, Rodych claims, that he had not abandoned that belief.[253]

Similarly, Wittgenstein's remarks on Gödel do not *explicitly* doubt the meaningfulness of such propositions, but do so *implicitly*, which is precisely the second point of §11: When Wittgenstein says that Gödel's proof shows that we have a proposition which does and does not belong to Russell's system that is merely an inaccurate way of expressing that when we can prove the Gödel sentence we can also prove its syntactical negation. Wittgenstein's answer, "That's is what comes of making up such sentences" is supposed be such an implicit denial of the meaningfulness of P (since of course the P is in fact a number-theoretic statement with a universal quantifier). The point, Rodych says, is 'made more forcefully' in (RFM V, §46):

[250]See for instance (PR §126–129) or (PR §173).
[251]Rodych 1999, p. 185.
[252]RFM VI, §13
[253]It should be noted that Wittgenstein only considered Fermat's Last Theorem to be a meaningless proposition because it was yet unproven at the time—for him, an arithmetical proposition with unbounded quantification *stands for* its inductive proof.

> The curse of the invasion of mathematics by mathematical logic is that now any proposition can be represented in mathematical symbolism, and this makes us feel obliged to understand it. Although of course this method of writing is nothing but the translation of vague ordinary prose.[254]

This curse, Rodych says, is that we now use the quantifiers “without any *known* connection to a method of decision”.[255] Wittgenstein’s point is that, just because a sentence is composed of mathematical symbols, that doesn’t mean that it is by that very fact, meaningful.

The point in §11, Rodych says, is then that we should not be surprised that a sentence, involving a universal quantifier over an infinite domain leads to a contradiction: “That is what comes of making up such sentences”.[256] This reading is confirmed in §20:

> Here one needs to remember that the propositions of logic are so constructed as to have no application as *information* in practice. So it could very well be said that they were not *propositions* at all; and one’s writing them down at all stands in the need for justification. Now if we append to these ‘propositions’ a further sentence-like structure of another kind, then we are all the more in the dark about what kind of application this system of sign combinations is supposed to have; for the mere *ring of a sentence* is not enough to give these connexion of signs any meaning.[257]

This of course reminds the reader of an old doctrine of the Tractatus, namely that the propositions of logic are meaningless. Wittgenstein’s point here is not so simple, however, as it is both more general and more specific. The general point, Rodych believes, is that a truth function combining two propositions, such as $\varphi \wedge \psi$ is merely ‘a propositional skeleton’ and cannot be used to make assertions about the world. In order to do so φ and ψ must be replaced with contingent, meaningful propositions.

The more specific point, and Wittgenstein’s main claim, is that if we replace a propositional variable in a logical skeleton such as $\forall x \varphi$ with an arithmetical sentence (which Rodych takes “append to these ‘propositions’ further sentence-like structure of another kind” to mean) we might easily end up with a meaningless sentence. In other words, even though we think that we have a well-formed formula of number theory, it “does not necessarily follow that we have constructed

254RFM V, §46.
255Rodych 1999, p. 186.
256Ibid., p. 187.
257RFM I, App. §20.

a meaningful arithmetical (or mathematical) proposition".[258] The only sense in which a sentence such as P could be given meaning would be for it to *be* proven, and since *eo ipso* Gödel has not (or indeed cannot, by his own theorem) do so, P is meaningless.

The third point found in §11 concerns the consistency assumption of Gödel's theorem. Rodych claims that Wittgenstein, *contra* the early commentators, not only understood this assumption, but emphasised it in his remarks. An important clue is Wittgenstein's remarks in §14. The last sentence of that remark, "A contradiction is unusable as such a prediction." seemed particularly bizarre. For Rodych, however, Wittgenstein does not mean that a contradiction cannot be used in a regular *reductio ad absurdum*-proof. It is rather to point out that Gödel's proof isn't such a proof. In (RFM V, §28) Wittgenstein wrote:

> What an indirect proof says, however, is: "If you want *this* then you cannot assume *that*: for only the opposite of what you do not want to abandon would be combinable with *that*".[259]

In other words, in a standard *reductio*-proof it is concluded that an assumed proposition is false (and thus unprovable if our system is consistent) because "its negation is contained within the axioms of our system".[260] This is not the case with Gödel's theorem, P is not unprovable because it's negation can be derived, but rather proves a conditional which states that if his system is ω-consistent, P cannot be derived.[261]

What has to be had in mind then, Wittgenstein is trying to remind us, is that Gödel hasn't proven the antecedent of his conditional (and indeed cannot by finitistic means, given his second theorem). The real question is, for Rodych, whether, as Wittgenstein states in §15,

> the 'proof of the unprovability of P' is here a forcible reason for the assumption that a proof of P will not be found.[262]

For Rodych, Wittgenstein is suggesting that this is not the case, and that only way Gödel could make us accept his conclusions would be for him to prove the consistency of his system, and perhaps, Rodych suggests, Wittgenstein would say that "all calculi that admit such sentence-constructions [as Gödel's] are syntactically inconsistent!".[263]

[258] Rodych 1999, p. 187.
[259] RFM V, §28.
[260] Rodych 1999, p. 187.
[261] An astute reader might have noticed that Rodych's interpretation depends on the stronger assumption of consistency, not ω-consistency, which was Gödel's. However, this is not a serious mistake on Wittgenstein's part, if indeed it was, since Gödel's result was later strengthened into the one we see here by Rosser.
[262] RFM I, App. III, §15.
[263] Rodych 1999, p. 190.

Evaluation of Rodych's interpretation

In the last section we saw how Rodych interprets Wittgenstein as making four distinct claims in his remarks on Gödel,namely that (a) his proof rests on there being a natural language interpretation of P that can be abandoned, that (b) on his own terms, there cannot be any such thing as a 'true but unprovable' proposition, that (c) since P is an arithmetical sentence with an unbounded universal quantifier (and has no proof), it is meaningless, and lastly, that (d) since Gödel hasn't proven the consistency of his system, he cannot claim that there are true but unprovable sentences in it, because his proof rests on that assumption. In this section we will examine and criticise some of these claims.

The first claim that we will consider here, that Wittgenstein made a mistake in his interpretation, is correct, and since Rodych is obviously an interpreter sympathetic to Wittgenstein, and should therefore be commended for the courage—which none of the other interpreters examined so far have had—of admitting that Wittgenstein seriously misunderstood Gödel's theorem when he wrote the passages that make up (RFM I, App. III). It is simply undeniable for anyone reading those remarks without prejudice that Wittgenstein was under the false impression that Gödel's proof essentially dependent on an natural language interpretation of the sentence P (whether or not he thought this avoided Gödel's result is not so clear).[264]

Rodych's second claim is that Wittgenstein objected to the saying that a sentence could be 'true and unprovable in Russell's system' because being true and provable mean the same in Russell's system. This is not implausible from the way Rodych reads the text. There are however some objections to this view. First of all, if Wittgenstein was insisting that being a true proposition in Russell's system means to be a proved proposition, the obvious Gödelian response would be that this just shows that Russell's conception of truth and provability is untenable—and indeed Wittgenstein's position would be nothing more than a simple restatement of Russell's logicism and the Gödelian reply just sketched the standard refutation of that view. Wittgenstein's instance would again betray a misunderstanding of Gödel's proof and the surrounding debate, a part from being very uninteresting.

I believe Rodych is mistaken however, and his can be seen well from the following remark he gives in a footnote. He says:

> Here [in §7] again Wittgenstein equates 'truth' with 'provability'.

[264]It should perhaps be noted, as Rodych does in a footnote to his paper (Rodych 1999, p.204. Footnote 62.), that Wittgenstein did give a correct number-theoretic interpretation of P at (RFM VII, §22). That remark however, was written 3–6 years after the remarks under consideration here, and therefore cannot be used to establish that Wittgenstein had such an interpretation in mind previously.

> Note that in the second quoted passage from #8, below, Wittgenstein returns to equating "true in Russell's system" with "proved in Russell's system".[265]

The problem is that Rodych equates Wittgenstein's insistence that propositions can only be asserted in 'Russell's system' if they are proven in 'Russell's system', with Wittgenstein's own purported belief that provability equals truth in mathematics in general. But no such equivalence can be found in the text—the only thing Wittgenstein says is that this is what Russell does. It simply does not follow from Wittgenstein describing what he calls Russell's system that this is what he himself believes, and the only thing Rodych does is say that Wittgenstein said something similar years before in another context.

A reason to suppose however, that Wittgenstein is not endorsing 'Russell's system' is that, for him, 'Russell's system' (or game) is more than just a formal system, it is, to borrow a phrase from Floyd, a "practice, or ongoing system of use".[266] In the *Remarks*, Wittgenstein says,

> Now what do we call 'inferences' in Russell or Euclid? Am I to say: the transitions from one proposition to the next on in the proof? But where is this *passage* to be found?—I say that in Russell one proposition follows from another if the one can be derived from the other according to the to the position of both in a proof and the appended signs—when we read the book. For reading this book is a game that needs to be learnt.[267]

It would be very strange for Wittgenstein to endorse Russell's logicism in the passage Rodych cites (because what can an endorsement of "Russell's system" be if not that?) after arguing against it in the *Remarks* as a whole (or indeed as early as the *Notebooks* and the *Tractatus*). After all, Russell's system is *Principia Mathematica*, and Wittgenstein's rejection of it as a foundation for mathematics is well know.

Rodych's third claim, that Wittgenstein thought that P was a meaningless pseudo-proposition, is however hard to deny—as it is hard to read §19–20 in any other way. However, Rodych does not motivate his interpretation well. For him, this is somehow connected with Wittgenstein's finitism and denial that arithmetical propositions with quantifiers that range over infinite domains have meaning. The former claim is contentious at best. It is true that Wittgenstein did not believe in the independent existence of infinite sets as abstract objects, but this simply follows from his rejection of platonism. He doesn't believe that

[265] Rodych 1999, p. 184.
[266] See above, p. 41.
[267] RFM I, §18.

the meaning of *any* mathematical statements is dependent on the existence of such objects, including statements about infinite sets. Furthermore, Wittgenstein is quite explicit in (RFM) that he has no objections with set theory as a subfield of mathematics, but only with its metaphysical interpretation.[268] If the former is enough to be tarred with the brush of finitism, that position seems so broad as to be almost meaningless, as it would only exclude platonists, and if the latter is true, it is hard to see what content the word 'finitism' has at all (at least it would have to involve a claim that reasoning with infinite sets is somehow *mathematically* problematic).

The latter claim is even more problematic. As Rodych himself notes, Wittgenstein never does explicitly say in the *Remarks* that propositions that quantify over an unlimited domain are meaningless and Rodych's evidence is circumstantial at best. Rodych's point is that logical propositions are like 'propositional skeletons' and for them to be meaningful, real, contingent propositions need to be replaced—and when arithmetical propositions are used in such a substitution, the result might be meaningless. The problem is that the remark he bases this on says "if we append to these 'propositions a further sentence-like structure of another kind, then we are all the more in the dark…".[269] Since 'append' does not mean 'replace' or 'substitute', or anything like that, Rodych's interpretation is quite far-fetched.

Similarly, a much better reading of the remark on Fermat's Last Theorem would be to see it in similar terms as Juliet Floyd: If the meaning of a proposition is changed by its proof, a mathematician cannot be said to fully understand a conjecture before it is proven. In any case, the idea that P is meaningless because it contains a universal quantifier which runs over an infinite domain, is problematic from Rodych's own viewpoint, since , it is hard to see how could Wittgenstein have made what Rodych calls 'his mistake', if he was aware that P was in fact an arithmetical proposition. That Wittgenstein both misunderstood the nature of P and thought that it needed a natural language interpretation, and its number-theoretic nature seems very unlikely.

If however we weaken Rodych's claim a little bit and focus on the argument that the use of quantifiers might sometimes make vague sentences in mathematics seem precise, and this gives a feeling of understanding which is perhaps not justified, things don't fare much better. If this really is Wittgenstein's view on P, he is again mistaken. P does contain quantifiers (and other logical symbols besides) but there is nothing unclear about its meaning—it is given precisely by the standard semantics of first-order logic and the axioms of formal arithmetic.

Lastly, if Rodych's interpretation is the correct one, his last claim really

[268]Cf. (RFM II, §58–62).
[269]RFM I, App. III, §20.

boils down to not more than an insistence that since Gödel hasn't proven the consistency of PM, he hasn't proven the truth and undecidability of P.[270] If this is the case, Wittgenstein hasn't really said anything not already contained in the standard interpretation of Gödel's theorem apart from a point of emphasis—Gödel of course knew better than anyone that his proof rested on the assumption of consistency, and in his remarks Wittgenstein hasn't done anything to dislodge this assumption. Thus on Rodych's interpretation, Wittgenstein's remarks again become very uninteresting—unless of course he could offer some argument for why this is actually the case. In the next, and last chapter, we will take a look at an interpretation which involves exactly that.

[270]This point is due to Timm Lampert. See (Lampert 2006).

Berto and the dialetheist perspective

Paraconsistent logicians and dialetheists consider Wittgenstein a forerunner of their approach to logic. This is not without foundation. The following remark from Wittgenstein's conversations with Waismann in the early 1930's, for instance, is widely quoted:

> Indeed, even at this stage I predict a time when there will be mathematical investigations of calculi containing contradictions, and people will actually be proud of having emancipated themselves even from consistency.[271]

Other remarks by Wittgenstein lend themselves equally to the interpretation that Wittgenstein would have at least agreed with much of what the dialetheists and the paraconsistent logicians have said. A prominent example, is for instance,

> And suppose [Russell's] contradiction had been discovered but we were not excited about it, and had settled e.g. that no conclusions were to be drawn from it. (As no one does draw conclusions from the 'Liar'.) Would this have been an obvious mistake?[272]

Wittgenstein also discusses contradictions in his remarks on Gödel, more precisely in §§11–14. Those sections are the starting point for Fransesco Berto's very interesting interpretation of Wittgenstein's remarks, set forth in is book, *There's Something About Gödel*.[273]

Berto agrees, as the others, that it is a "common trait" in Wittgenstein's various remarks on Gödel, that he wants to separate the actual mathematical result from the philosophical consequences philosophers have wanted to draw from them, particularly the idea that Gödel's theorem shows that platonism in mathematics is true.[274] More specifically, however, Berto sees in Wittgenstein's remarks certain insights which, despite having being panned by the early critics, have been made more plausible by the relatively recent advent of paraconsistent logic. In this section and the next, we will see what these insights are and how they have been rendered more credible.

Berto's starting point for his interpretation is (RFM I, App. III, §11) where Wittgenstein seems to offer something resembling a Gödelian argument, ending with a contradiction, and then indicating that the contradiction is harmless:

[271]PR. p. 332. A prominent example is Graham Priest's influential paper, *Logic of Paradox*, which has it as a kind of motto.
[272]RFM VII, §15.
[273]See (Berto 2009b, chapter 12.). Berto also has a paper based on this chapter (or vice versa). See (Berto 2009a).
[274]Berto 2009b, pp. 190–191.

> Let us suppose I prove the unprovability (in Russell's system) of P; then by this proof I have proved *P*. Now if this proof were one in Russell's system—I should in that case have proved at once that it belonged and did not belong to Russell's system.—That is what comes of making up such sentences.— But there is a contradiction here!—Well, then there is a contradiction here. Does it do any harm here?[275]

As Berto notes, Wittgenstein, seems to understand *P* as a paradoxical sentence similar to the Liar Paradox, leading to an inconsistency.[276] He points out that other prominent logicians at the time, most notably Zermelo and possibly Russell, made a similar mistake in their interpretation of Gödel's proof.[277] The mistake, he says, consists in not properly heeding the distinction between theory and metatheory, or between syntax and semantics.[278] This confusion makes it difficult to see the difference between the provability predicate of formal arithmetic, which is weakly expressible in the theory, and the truth predicate, which is inexpressible in a consistent theory. This reading of the above paragraph is consistent with Gödel's own interpretation, as expressed in his letter to Karl Menger, quoted above in the introduction.[279]

Berto however, claims that this is no mistake on Wittgenstein's behalf. He knowingly rejects such a distinction. He cites the following remark from *Philosophical Grammar*:

> 4. I said earlier "calculus is not a mathematical concept"; in other words, the word "calculus" is not a chess piece that belongs to mathematics.
>
> There is no need for it to occur in mathematics.—If it is used in a calculus nonetheless, that doesn't make the calculus into a metacalculus; in such a case the word is just a chessman like all the others.
>
> Logic isn't metamathematics either; that is, work within the logical calculus can't bring to light essential truths about mathematics. Cf. here the "decision problem" and similar topics in modern mathematical logic.

[275] RFM 1, App. III, §11.

[276] Berto 2009b, p. 192.

[277] Another example is Charles Perelman, who claimed that Gödel had in fact discovered a paradox. His claim was widely dismissed as displaying "a rather obvious conflation of object- and metalanguage" (Dawson 1988, pp. 84–85.). Later in this section, we will see that the dialetheist treatment of Gödel's theorem has certain affinities with this. See (Dawson 1988) for details.

[278] Berto 2009b, p. 192.

[279] However, Berto favourably mentions Rodych's claim, based on his reading of unpublished manuscripts, that Wittgenstein "correctly understood the number-theoretic nature of Gödel's proposition". See (Rodych 2002, p. 380.) Rodych's conclusion is too strong: The remarks he cites date from 1939, two years after the remarks of (RFM I, App. III) were written.

...

(Hilbert sets up rules of a particular calculus as rules of meta-mathematics.)[280]

That is to say, what is called 'metamathematics' has, for Wittgenstein, no special status and there is no such thing as a 'metacalculus', simply another calculus among calculi.[281] Thus, Berto's Wittgenstein sees mathematics as a family of independent calculi, who—to use Shanker's phrase—"stand on equal footing".[282] Furthermore, the propositions contained in each of these calculi get their meaning from their place in them—"its net of theorems and rules".[283] This makes the way in which the truth of the undecidable sentence P is established suspect. Gödel had shown by means of a *syntactic* proof that the sentence P was undecidable in the system of PM, and then established its truth *outside* PM, through what he himself described as "metamathematical considerations".[284] But if the sense of a proposition is derived from its place in a calculus, the sense of P *inside* PM cannot be the same as the sense of P *outside* PM. This is probably what Wittgenstein had in mind, Berto claims, when he wrote the following remark in the *Philosophical Remarks*:

> What is a proof of provability? It's different from the proof of proposition.
>
> And is a proof of provability perhaps the proof that a proposition makes sense? But then, such a proof would have to rest on *entirely different* principles from those on which the proof of the proposition rests. There cannot be a hierarchy of proofs!
>
> On the other hands there can't in any fundamental sense be such a thing as meta-mathematics. Everything must be of one type (or, what comes to the same thing, not of a type...
>
> Thus, it isn't enough to say that p is provable, what we must say is: provable according to a particular system.
>
> Further, the proposition doesn't assert that p is provable in the system S, but in *its own* system, the system of p. That p belongs to the system S cannot be asserted, but must show itself.
>
> You can't say p belongs to the system S; you can't ask which system p belongs to; you can't search for the system of p. Understanding p means understanding it's system. If p appears to go over from one system into another, then p has, in reality, changed its sense.[285]

[280]PG p. 296–297.
[281]Berto 2009b, p. 195.
[282]Shanker 1988, p. 213.
[283]Berto 2009b, p. 195.
[284]See (Gödel 1976, p. 599.).
[285]PR, p. 180. Whether or not Wittgenstein had this in mind when he wrote the paragraph,

Thus for Berto, Wittgenstein rejects metamathematics, and holds that the meaning of a proposition is determined by the system to which it belongs. It is then impossible that the very same sentence could be shown to be false in one calculus and true in another—the sentence would by that very fact be a different sentence (and a mathematical proposition without a calculus is even more inconceivable on this view, one should hasten to add). As a consequence of this "bold general doctrine", Berto claims, Wittgenstein is forced to reject the idea that a mathematical system can be incomplete, since if the meaning of a mathematical proposition is established by a proof, an incomplete system would mean that the true proposition it cannot prove had an incomplete meaning.[286] The following remarks from *Philosophical Grammar* supposedly confirm this:

> The edifice of rules must be *complete*, if we are to work with a concept at all — we cannot make any discoveries in syntax.—For, only the group of rules defines the sense of our signs, and any alteration (e.g. supplementation) of the rules means an alteration of the sense.[287]

and

> Mathematics cannot be incomplete; any more than a sense can be incomplete.[288]

As a consequence, Berto takes Wittgenstein's identification of mathematical truth with assertability in §§1–6 to come with an implicit understanding that the formal counterpart of assertability is provability. This is for him confirmed in §7.[289] This rejection of the incompleteness of a mathematical calculus, Berto (clearly under the influence of Rodych's interpretation) associates with Wittgenstein's supposed strong finitism and his insistence that any mathematical question is decidable. Berto notes of course that if Wittgenstein is rejecting that a system can be incomplete, so strongly identifying provability and truth, and rejecting the notion of metamathematics, he is of course opposing some very established results of contemporary logic and mathematics. Here Berto joins the ranks of commentators who believe that Wittgenstein goes against his own principles:

> This speaks against Wittgenstein's own claim, according to which "it is my task, not to attack Russell's logic from within, but from

it should perhaps still be noted that the sentence p above is just supposed to be a propositional variable, and not the specific sentence P from (RFM I, App. III). This remark was written approximately 7 years before.

[286]Berto 2009b, p. 198.

[287]PR, p. 182.

[288]PG, p. 188.

[289]Berto 2009b, p. 200 §§1–7 can be found in the appendix.

> without," and "my task is not to talk about (e.g.) Gödel's proof, but to pass it by."[290]

However, unlike them, he believes that Wittgenstein's remarks can be thus interpreted as to give it "an unexpected plausibility precisely from the point of view of modern *non*-classical logic."[291] This standpoint is of course paraconsistent logic, and its accompanying philosophical view, dialetheism. In the next section we will see what exactly this interpretation consists in.

The 'single argument'

Berto's paraconsistent interpretation of Wittgenstein's remarks uses an argument first propounded by Graham Priest and Richard Routley in several different essays.[292] In this section, we will first explicate the argument, called the 'the single argument' by Berto, and then we can see how Berto uses it to interpret Wittgenstein's remark. In the following we will mostly follow the argument as it is put forth in Graham Priest's book, *In Contradiction.* The argument aims to show that when Gödel's proof method is applied to what Priest calls our "naïve notion of proof", Gödel's theorem should properly be viewed as a paradox.[293] This naïve notion of proof is for Priest simply ordinary mathematical practice—the activity mathematicians are engaged in when they are busy proving their theorems. It is defined by Priest thus:

> Proof, as understood by mathematicians (not logicians), is that process of deductive argumentation by which we establish certain mathematical claims to be true.[294]

This notion of proof is not merely a syntactic object, as the logicians would have it, but simply they way mathematicians reason—it is the method mathematicians use to know that what they prove is true. This informal way of reasoning of course has its own rules and conventions, but they are informal. Much like a formal system, however, a proposition proved by a mathematician engaged in mathematics is proven from things already known, and must therefore start from what Priest calls "basic statements". These basic statements could for instance be something like Euclid's axioms or the fact that every number has a successor. These things are of course analogous to the axioms of a formal system (and its rules of inference analogous to the rules of inference of a formal system) but they are not in themselves formal.

[290] Berto 2009b, p. 200. The quotes from Wittgenstein are from (RFM VII, §19).
[291] Berto 2009b, p. 200.
[292] See (Priest 1979), (Priest 1984), (Priest 2006) and (Routley 1979).
[293] Priest 1979, p. 40.
[294] Priest 2006, p. 40.

This informal notion of mathematics, however, could be formalised (as mathematicians are wont to say), and for Priest, this allows us to talk as if it was a formal theory.[295] The resulting formal theory would, he claims, satisfy the conditions of Gödel's theorem, i.e. it would be strong enough to represent all primitive recursive functions and its proof relation would be decidable. The former claim should not be very controversial, since our naïve notion formalised would contain everything a formal arithmetic does (and indeed the latter is even attempt to capture our informal notions). The latter claim, however, is more controversial. Priest argues that it is nonetheless true, since the whole concept of proof is to be a method by which a statement can be recognised as being true or not—therefore only something can be easily recognised as a proof, is a proof.[296] Priest quotes Alonzo Church on the matter:

> [C]onsider the situation which arises if the notion of proof is non-effective. There is then no certain means by which, when a sequence of formulas has been put forward as a proof, the auditor may determine whether it is in fact a proof. Therefore he may fairly demand a proof, in any given case, that the sequence of formulas put forward is a proof; and until the supplementary proof is provided, he may refuse to be convinced that the alleged theorem is proved. This supplementary proof ought to be regarded, it seems, as part of the whole proof of the theorem. . . [297]

And if the notion of proof is effectively recognisable, then by Church's thesis, Priest says, it must be recursive.[298] This can furthermore explain how we can learn arithmetic in the first place:

> We appear to obtain our grasp of arithmetic by learning a set of basic and effective procedures for counting, adding, etc.; in other words, by knowledge encoded in a decidable set of axioms. If this is right, then arithmetic truth would seem to be just what is determined by these procedures. It must therefore be axiomatic. If it is not, the situation is very puzzling. The only real alternative seems to be Platonism, together with the possession of some kind of sixth sense, "mathematical intuition".

A Platonist, of course, would have no problem with accepting this consequence of an undecidable proof relation (Gödel himself didn't believe that it was), and

[295] Priest 2006, p. 41.
[296] Ibid.
[297] Church 1956, p. 53.
[298] Priest 2006, p. 41.

indeed Gödel's theorem itself could be used as an argument against the decidability of the naïve proof relation (but not without begging the question against the dialetheist, as we will see). But what is the paradox?[299] Consider the following:

Let T be the aforementioned formalisation of our naïve mathematical theory. Since we have assumed that T fulfils all the conditions of Gödel's theorem, we know that if T is consistent, there exists a sentence φ which is not provable in T. This φ can be shown by a naïve proof to be true, and thus is provable. φ is therefore both provable and not provable in T, and T is inconsistent.[300]

The real point, however, Berto points out, is that when the sentence P is interpreted in terms of our naïve, informal way of doing mathematics (i.e. as saying "This sentence is not demonstrably true"), Gödel's proof becomes the derivation of a very real paradox.[301] Priest explains:

> In fact, in this context the Gödel sentence becomes a recognisably paradoxical sentence. In informal terms, the paradox is this. Consider the sentence "This sentence is not provably true.". Suppose the sentence is false. Then it is provably true, and hence true. By *reductio* it is true. Moreover, we have just proved this. Hence it is provably true. And since it is true, it is not provably true. Contradiction.[302]

And this, Berto claims, is what Wittgenstein had in mind in §11, when he said:

> Let us suppose I prove the provability (in Russell's system) of P; then by this proof I have proved P. Now if this proof were one in Russell's system—I should in that case have proved at once that it belonged and did not belong to Russell's system.[303]

This shouldn't be surprising, since if Wittgenstein denies that there is any distinction between a theory and its metatheory, and everything is done on the same level, it should follow that the theory T is semantically closed (as it would coincide with the natural language in which is expressed, which in turn has no theory/metatheory-distinction either, or so one could at least argue).[304] But since T is inconsistent, this seems to be big problem for mathematics, and thus Wittgenstein's position, since surely an inconsistent theory would be useless. In the next section, we will examine how Wittgenstein could have responded to this challenge.

---

[299] For our purposes we can just assume that the relation *is* decidable (and that the whole plan could be carried out), since the purpose of this chapter is to assess Berto's interpretation, and not the merits of dialetheism.

[300] See Priest 2006, p. 44 for details.

[301] Berto 2009b, p. 205.

[302] Priest 2006, p. 46.

[303] RFM I, App. III, §11.

[304] Of course one could argue that the inconsistency means that Gödel's theorem doesn't apply in the first place, but that would be *because* the theory T is inconsistent, and the paradox is not really avoided.

Wittgenstein and inconsistency

As Berto says, if Wittgenstein's real point in his remarks on Gödel was to turn his proof into a sort of logical paradox, it would mean that inconsistency were a central feature of our mathematical practices—our mathematics would in its very essence be inconsistent.[305] This is of course a striking consequence and would perhaps in the mind of many be taken as a complete *reductio ad absurdum* of Wittgenstein's view. This was of course the line taken by many of the early commentators, as we saw in an earlier section, some of which accused Wittgenstein of not understanding the technical implications of allowing for a contradiction. What they had in mind, of course, was that in classical logic, everything follows from a contradiction, and the system under consideration would collapse into triviality—it would make every sentence true.

However, Wittgenstein was of course fully aware of this elementary result (or rule) in propositional logic, and his response was very much in the same vein as one a modern paraconsistent logician would give: Simply abandon classical logic and adopt a paraconsistent one where the rules of inference have been changed in an appropriate way as to block the inference of arbitrary propositions from a contradiction. In the *Lectures on the Foundations of Mathematics*, Wittgenstein said,

> One may say, "From a contradiction everything would follow." The reply to that is: Well then, don't draw any conclusions from a contradiction; make that a rule.[306]

Wittgenstein makes similar remarks in other lectures, for instance, he also compares the logician's way of viewing contradiction to a "germ" which shows that the "whole body is diseased" and the deriving of a contradiction to "a place from which you can go in every direction".[307] These remarks of course show that Wittgenstein was fully aware of the unfortunate consequences of a contradiction in classical logic.

However, in the defence of the early commentators, who didn't have access to this material, Wittgenstein does not make this point in the *Remarks*, the very work they were reviewing. If it were not kept in mind that Wittgenstein would—in the face of contradiction—advocate the abandonment of the classical *ex falso*–rule, many of his pronouncements would seem absurd, even for the most mediocre student of logic. An example would be Wittgenstein's question in (RFM I, App. III, §11) of what harm a contradiction in a formal system would do.[308] From a classical, monist perspective, the answer is obvious: The

[305] Berto 2009b, p. 207.
[306] LFM XXI, p. 209.
[307] See (LFM XIV, p. 138.) and (LFM XXIII, p. 224).
[308] RFM I, App. III, §11.

contradiction would render the system trivial by making every sentence true.[309] But keeping this in mind is not enough. There are other serious objections to Wittgenstein's point, some of which were raised in the lectures. I will consider two of the most serious here, both of which were raised by Alan Turing, who attended them.

Turing points out that if it were made a rule that it was not allowed to draw conclusions from a contradiction, one would still be able derive an arbitrary sentence "without actually going through the contradiction".[310] What he had in mind is presumably something like the following natural deduction:[311]

(1)	$\varphi \wedge \neg\varphi$	Premise
(2)	φ	$\wedge$–elimination from (1)
(3)	$\varphi \vee \psi$	$\vee$–introduction from (2)
(4)	$\neg\varphi$	$\wedge$–elimination from (1)
(5)	ψ	$\vee$–elimination from (3) and (4)

This would be an example of deriving an arbitrary formula from a contradiction, without inferring it directly (and could be made even more indirect by supposing that φ and $\neg\varphi$ never in fact appear on the same line—the overall deduction would however be similar). Charles Chihara fiercely criticised Wittgenstein for not understanding this point (he correctly notes that Wittgenstein never answered Turing's objection):

> Turing's point was a simple one, intelligible, one would think, to anyone with an understanding of elementary logic; yet there are reasons for thinking that Wittgenstein failed to grasp it. In the first place, Wittgenstein made no attempt to answer it.[312]

This is however not as big of a problem for Wittgenstein as Chihara thinks. In any paraconsistent logic worth its name, derivations like this would also need to be blocked, and this is in fact easy to do. One prominent example is Priest's Logic of Paradox, in which $\neg\varphi, \varphi \vee \psi \vDash \psi$ is not a valid semantic consequence, and it would therefore not be possible to carry out our derivation in a suitable proof system for it.[313] Pointing to a real-life example of a non-trivial paraconsistent

[309] By 'monist' I mean someone holding the doctrine that there is only one logic which is the correct logic.
[310] LFM XXIII, p. 220.
[311] Similar derivations are possible in an axiomatic system, but are harder to demonstrate without actually giving the axioms.
[312] Chihara 1977, p. 372.
[313] Priest 1979, §III.

logic would of course not be a possible answer for Wittgenstein, since none had been invented at the time of Turing's objection, nor can we know that he would have said something similar, but the now obvious possibility should however be enough to vindicate Wittgenstein's position somewhat.

The other objection raised by Turing we might call the "falling bridges"-objection.[314] The objection goes as follows. Engineers are building a bridge and use for this purpose an arithmetical calculus. The building of this bridge is a fairly complicated affair and the calculations of the engineers are very long. Now it happens that the calculus is inconsistent, and on one line of their calculations the engineers derive the proposition $12 \times 12 = 144$ and on another line they derive the a proposition equivalent to $12 \times 12 \neq 144$. They have tried to follow Wittgenstein's advice of not deriving anything from a contradiction, but their calculations were just so long and complicated that they didn't notice that they had derived one. They accidentally derive an arbitrary proposition from their contradiction, this results in a mistake in building the bridge, and it falls down.

Wittgenstein offers two reasons for a bridge might fall down:

> There seems to me to be an enormous mistake there. For your calculus gives certain results and you want the bridge not to break down. I'd say things can go wrong in two ways: either the bridge breaks down or you have made a mistake in your calculation–for example, you multiplied wrongly. But you seem to think there may be a third thing wrong: the calculus is wrong.[315]

For Chihara however, Wittgenstein has indeed forgotten the third possibility: "(3) the logical system they used was unsound and led them to make invalid inferences (that is, they followed the rules of derivation correctly, but their calculus was wrong)".[316] He then claims that in this case, the collapse of the bridge was not due to wrong calculations nor faulty data. He goes on:

> In fact, as I have described the situation, if the engineers were to recheck their data and retest their empirical theories, they would find everything in order. Hopefully, there would be some nonWittgensteinian logicians around to discover the unsoundness of their logical system.[317]

[314] LFM XXII, p. 212.

[315] LFM XXII, p. 218. What Wittgenstein means by the first possibility is that it was simply badly made—a mistake was made in its construction independent of any calculation.

[316] Chihara 1977, p. 377. It is strange however, that Chihara describes this as forgetfulness on Wittgenstein's part. He obviously didn't forget, he simply rejects the possibility.

[317] Ibid. Chihara's example is somewhat different than mine. There a computer is used to check the work of the engineers. Now, if that were the case, the computer would obviously be programmed to use a paraconsistent logic. The computer would simply never make such a mistake as to derive something from a contradiction, were it properly programmed and in

But this is in fact not the case. The way the example is set up is exactly so that using a contradiction falls under a miscalculation (that is what Wittgenstein's abolishment of the *ex falso*-rule means). And this is what Wittgenstein replied to Turing when he raises the exact same objection. Wittgenstein says,

> The trouble described is something you get into if you apply the calculation in a way that leads to something breaking. This you can do with *any* calculation, contradiction or no contradiction.
>
> What is the criterion for a contradiction *leading* you into trouble? Is it specially *liable* to lead you into trouble?[318]

Chihara doesn't consider this objection by Wittgenstein, possibly because he phrased it somewhat obscurely, and he didn't have time to elaborate on it.

The point is, I believe, this: If the engineers had *made* it a rule in their calculus that one should not draw conclusions from a contradiction (and made it in an appropriate way), they *would* find their mistake if they rechecked their calculations rigorously enough—there is no essential difference between checking whether one rule of inference was broken or another, and if the engineers had made some *other* calculation mistake (say, accidentally written $13 \times 13 = 144$ on one line), they would also find their data and empirical theories in order. Yet Chihara wouldn't say that this showed that their usual way of multiplying was wrong. Making a mistake in the calculation is what led them into trouble, not the contradiction.

The only thing that Chihara's example shows is that paraconsistent systems might be too unwieldy to use to build bridges, because it would be difficult for the all too human engineers to follow the rules. If the calculation of the engineers was, say, five lines long, it would be easy for them to spot a calculation error, and it would also be easy for them to see whether they violated their rule of not using a contradiction. However, if the calculation was very long, maybe 100.000 lines, it would be difficult for them to check whether they multiplied wrongly in one place, and it would be difficult for them to know whether they followed their own rules. It would certainly be *easier* to see whether one of them made a multiplication error in such a long calculation than to see whether they accidentally derived a false statement from a contradiction, but this doesn't point to a fundamental difference between the two mistakes. It is also easier for a human engineer (without a computer, of course) to check his calculations in an arithmetic which uses, for instance, base ten, rather than base three—from the point of view of logic, they are however of course equivalent.

---

good working order. This in itself should be enough to show that given a well chosen calculus, any derivation from a contradiction of an arbitrary proposition would simply be a calculation mistake.

[318] LFM XXII, p. 219.

There still seems to be something difficult to accept in the idea of an inconsistent arithmetic. For Wittgenstein however, there is nothing intrinsically problematic about such a thing: if a contradiction were to be found in a formal system of arithmetic—and we assume then that the formal system mirrored they way we actually do arithmetic—that would only show that an inconsistent arithmetic is perfectly fine, since it is used with good results in practice all the time. Wittgenstein explains:

> I mean: if a contradiction were now actually found in arithmetic—that would only prove that an arithmetic with *such* a contradiction in it could render very good service; and it will be better for us to modify our concept of the certainty required, than to say that it would really not have been a proper arithmetic.[319]

Furthermore, as Berto points out, inconsistent arithmetics have some nice properties that fit well with what he takes to be "Wittgensteinian intuitions".[320] First of all, there exist inconsistent arithmetics such that they can be shown to be both complete and decidable, and therefore fit well with those aspects of what Berto perceives to be Wittgenstein's philosophy of mathematics, namely his insistence that no mathematical proposition can be undecidable and no system is incomplete. Further, it can be shown that there exist an inconsistent theory $\mathcal{M}$ such that $\mathcal{M}$ includes all sentences true in the standard model of arithmetic, proves its own Gödel sentence (and its negation), and is true in a *finite* inconsistent arithmetical model (in fact, any arithmetical theory can be shown to have a finite, inconsistent model).[321] This rhymes well with Wittgenstein's finitism, one of the tenets of Wittgenstein's philosophy of mathematics, according to Berto.

When this is all taken together, the decidability and completeness of inconsistent arithmetics, their property of always having a finite model and Wittgenstein's nonchalant attitude towards contradiction, it certainly seems plausible that Wittgenstein would have had much sympathy with the dialetheist cause. However, this is not enough to show that this is in fact what Wittgenstein meant when he wrote his remarks on Gödel. In the next section we will take a look at this question.

Assessment of Berto's interpretation

In the last two sections, we saw how Berto interprets Wittgenstein's remarks in (RFM I, App. III) in the light Priest's and Routley's 'single argument', were

[319] RFM VII, §36.
[320] Berto 2009b, p. 209.
[321] For proof of the decidability and completeness results, see (Priest 1994). For proof of the latter claim, see (Bremer 2005).

Gödel's theorem recast as applying to naive proof concept. In this section, we will offer some criticisms of this idea.

In his interpretation, Berto makes some assumptions about Wittgenstein's philosophy of mathematics which are inspired by the interpretations of Shanker and Rodych. These include the idea that Wittgenstein held a strong calculus conception of language, that he was a strong finitist and that he held that every mathematical question was decidable. In the sections on Shanker and Rodych we saw that there are reasons to doubt this. Before we go on to criticise Berto, we should note that this is not a severe problem for him, as his main claims, namely that Wittgenstein saw Gödel's theorem leading to a paradox when put in the context of an absolute notion of 'provability', and that he was fine with this paradox, do not depend on these assumptions. At most he loses those aspects of his interpretation that are directly relevant to this, namely that Wittgenstein advocated an inconsistent arithmetic because it would be finite and complete, and it is very easy for him to change his interpretation in the light of similar criticisms as raised against Shanker and Rodych—that Wittgenstein had this in mind is very unlikely anyway, as the relevant work in paraconsistent models had not been done. Berto's main claims, however, do have other serious problems.

First of all, just a superficial reading of the text makes it very implausible that Wittgenstein had anything like a naïve notion of proof in mind when he wrote §11. There, as in other remarks, Wittgenstein talks what would be the consequences of proving the unprovability of P in 'Russell's system'. Russell's system, by which Wittgenstein of course means the system of *Principia Mathematica*, is a formal system, and not at all comparable to Priest's and Routley's naïve notion of proof—and even if it were, Wittgenstein would have phrased it differently, had he meant anything of the sort. Wittgenstein is thus highly unlikely to have meant any such thing by his remark.

Furthermore, while it is true that Priest's and Routley's notion of naive proof excludes there being a language/metalanguage-distinction, and it is somewhat plausible that Wittgenstein rejected such a distinction as well, other elements of his philosophy make it difficult to subscribe to him a view similar to Priest and Routley. On the calculus conception of mathematics, Wittgenstein would have, much like Berto says, rejected there being any calculus that is considered a metacalculus of any other, but he still would have seen the two supposed levels at work in Gödel's proof as being two different calculi, and therefore no contradiction would be derivable, as there would just be two different sentences in two different calculi.[322] Of course, as we said above, this view of Wittgenstein's philosophy of mathematics, is not essential for Berto.

On the language-game conception, however, things don't fare much better.

[322]See the section on Shanker above on this point.

Priest's and Routley's naïve conception of truth has a certain unity to it, and as Priest notes, the very possibility of it being formalised is essential for their argument. Wittgenstein's view of mathematics in the late period was completely opposite to this picture, for him there was no essence to our notion of proof, and the idea that the whole of mathematics could be formalised in one formal system would surely be anathema to him. It is therefore very unlikely that he would have wanted to say anything like Berto attributes to him.

Lastly, the point of Priest's and Routley's argument is to show that mathematics is in fact inconsistent. That is to say, it is their claim that mathematics (and logic) is or should be inconsistent. Despite Wittgenstein's indifferent attitude towards contradiction, it is unlikely that he would have wanted to make such strong claims. The dialetheists champion contradiction, Wittgenstein merely wanted to show that it are not as important as many philosophers believe. He consistently argued that if a contradiction were to be found, this would not mean a that a calculus was useless:

> I want to object to the *bugbear* of contradiction, the superstitious fear that takes the discovery of a contradiction to mean the destruction of the calculus.[323]

This is not the same as saying that there are in fact important, inconsistent calucli, and that this should be so. Consider for instance (RFM VII, §15):

> 'Then are you in favour of contradiction?' Not at all; any more than soft rulers[324]

Berto, however, is completely right in that the dialetheist program is very much in Wittgenstein's spirit, and it is not unlikely that he would have been very sympathetic towards it.

This is however just speculation and is in no way enough to show that this is what Wittgenstein had in mind. Berto's reading is however highly original and interesting, and he could of course argue that he never intended to argue that this was in fact what Wittgenstein had in mind, but rather that in the light of further developments in logic and philosophy, Wittgenstein's remarks do not seem as crazy and foolish as the early commentators thought. This is not at all implausible, and if this is his aim, he is certainly correct, but it does not show that Wittgenstein's remarks were particularly interesting, nor that he was right—at most Berto could claim Wittgenstein as some kind of authority in his dialetheism, but in light of the severe criticism Wittgenstein has suffered

[323](WVC, p. 196) See as well Wittgenstein's remarks in (RFM I, App. III) where he declares contradictions to be harmless and only of interest because the bother people, or (RFM III, §82): "My aim is to alter the *attitude* to contradiction and consistency".

[324]RFM VII, §15.

for such views (as witnessed by the reaction of the early commentators), it is unlikely that this would be of much help to him.

Concluding remarks

As we saw in the introduction and the first chapter of this thesis, the early debate on Wittgenstein's remarks on Gödel was mostly negative and characterized by the claim that Wittgenstein misunderstood Gödel's theorem in essentially two ways: that (a) he did not see how consistency was a necessary premise of Gödel's proof, or simply did not understand what consequences inconsistency would have on a formal system, and (b) the claim that he mistakenly thought that Gödel's proof used some kind of natural language interpretation of the sentence P in order to show that it was true, but unprovable, thus by abandoning this interpretation, Gödel's result could be avoided. It is almost certain that the widespread acceptance of these claims severely damaged Wittgenstein's reputation as a philosopher of mathematics.

The later debate, however, offered a much more nuanced and broader view, more sensitive to both Wittgenstein's other philosophical writings and the overall spirit in which he wrote. Despite this, they are not at all without their shortcomings. Many of them almost seem made especially for the purpose of vindicating Wittgenstein and absolving the great philosopher from all charges of technical incompetence and misunderstanding. Others have been much more elaborate in their exegesis of Wittgenstein's remarks than the text could possibly support and are thus very unlikely to persuade anyone not already convinced of the truth of their conclusion.

In this thesis a fairly comprehensive critical overview of this debate has been given. Unfortunately, the conclusion of this discussion however is undeniably a certain *aporia*—it is still difficult to give a precise account of what Wittgenstein actually did say in his infamous remarks on Gödel, and given their cryptic and unpolished nature, this will most likely always be the case. However, a few things can be gleaned from the preceding discussion.

First of all, we can see from the discussion on Berto and the dialetheists that Wittgenstein's remarks on consistency are far from being as outrageous as they were originally seen, and that the charge of technical incompetence on Wittgenstein's part was overstated by the early commentators. Of course the philosophical position of dialetheism is far from being the accepted position in mainstream analytic philosophy, but at least it taken seriously nowadays and more logicians are interested in paraconsistent logic than ever before. This should vindicate Wittgenstein's remarks somewhat in the mind of modern readers, even though they are perhaps not ready to accept his claims from a philosophical standpoint.

On the other hand, the interpretations which have tried to interpret Wittgenstein's remarks in such a way as to deny claim (b), namely the interpretations of

Shanker, Floyd and Floyd and Putnam, must—when all is considered—be taken to have failed in this task. This reading is simply the most natural one, and they only manage to avoid this conclusion by building very elaborate interpretations with many implausible assumptions and textual problems. In this respect the unequivocal conclusion must be that the early commentators (and Victor Rodych, of course) were simply right: When Wittgenstein wrote the remarks contained in (RFM I, App. III) he did not have good understanding of Gödel's proof and this led him to say wrong things about it.

This mistake, on Wittgenstein's part, is however not necessarily so severe, as to justify the damage it has done to his reputation as a philosopher of mathematics. It is undoubtedly true that Gödel's proof causes, as it is often seen, serious problems for the positions of logicism and formalism in the philosophy of mathematics, if it is not simply a refutation of those views. The matter is however far from being clear whether or not Wittgenstein's philosophy of mathematics entails such a position, as is often assumed, and many reasons to suppose that this is not the case—at least it is not what one would expect from the author of the *Philosophical Investigations.*

In that case, there is every reason to try to bring Wittgenstein's philosophy of mathematics out of the small and isolated circle of Wittgenstein scholars it currently occupies and further into the mainstream of the philosophy of mathematics, if not only to see whether it would turn out to be as fruitful as his other, more widely read, work. This task will unfortunately not be undertaken in this thesis.

Appendix

This appendix contains the full text of Appendix III to Part I of the *Remarks on the Foundations of Mathematics*, abbreviated in the text as (RFM I, App. III).

1. It is easy to think of a language in which there is not a form for questions, or commands, but question and command are expressed in the form of statements, e.g. in forms corresponding to our "I should like to know if..." and "My wish is that...".

   No one would say of a question (e.g. whether it is raining outside that it was true or false. Of course it is English to say so of such a sentence as "I want to know whether..." But suppose this form were always used instead of the question?—

2. The great majority of sentences we speak, write and read, are statement sentences.

   And—you say—these sentences are true or false. OR, as I might also say, the game of truth-functions is played with them. For assertion is not something that gets added to the proposition, but an essential feature of the game we play with it. Comparable, say, to the that characteristic of chess by which there is winning and losing in it, the winner being the one who takes the other's king. Of course, there could be a game in a certain sense very near akin to chess, consisting in making the chess moves, but without there being any winning and losing in it; or with different conditions for winning.

3. Imagine it were said: A command consists of a proposal ('assumption') and the commanding of the thing proposed.

4. Might we not do arithmetic without having the idea of uttering arithmetical propositions, and without ever having been struck by the similarity between a multiplication and a proposition?

   Should we not shake our heads, though, when someone shewed us a multiplication done wrong, as we do when someone tells us it is raining, if it is not raining?—Yes; and there is a point of connexion. But we also make gestures to stop our dog, e.g. when he behaves as we do not wish.

   We are used to saying "2 times 2 is 4" and the verb "is" makes this into a proposition, and apparently establishes a close kinship with everything we call a 'proposition'. Whereas it is only a matter of a very superficial relationship.

5. Are there true propositions in Russell's system, which cannot be proved in his system?—What is called a true proposition in Russell's system, then?

6. For what does a proposition's '*being true*' mean? *'p' is true* $=$ *p*. (That is the answer.)

   So we want to ask something like: under what circumstances do we assert a proposition? Or: How is the assertion of the proposition used in the language game? And the 'assertion of the proposition' is here contrasted with the utterance of the sentence, e.g. as practice in elocution,— or as *part* of another proposition, and so on.

   If, then, we ask in this sense: "Under what circumstances is a proposition asserted in Russell's game?" the answer is: at the end of one of his proofs, or as a 'fundamental law' (Pp.) There is no other way in this system of employing asserted propositions in Russell's symbolism.

7. "But may there not be true propositions which are written in this symbolism, but are not provable in Russell's system?"—'True propositions', hence propositions which are true in another system, i.e. can rightly be asserted in *another* game. Certainly; why should there not be such propositions; or rather: why should not propositions—of physics, e.g.—be written in Russell's symbolism? The question is quite analogous to: Can there be true propositions in the language of Euclid, which are not provable in his system, but are true?—Why, there are even propositions which are provable in Euclid's system, but are *false* in another system. May not triangles be—in another system—similar (*very* similar) which do not have equal angles?—"But that's just a joke! For in that case they are not 'similar' to one another in the same sense!"—Of course not; and a proposition which cannot be proved in Russell's system is "true" or "false" in a different sense from a proposition of *Principia Mathematica.*

8. I imagine someone asking my advice; he says: "I have constructed a proposition (I will use 'P' to designate it) in Russell's symbolism, and by means of certain definitions and transformations it can be so interpreted that it says: 'P is not provable in Russell's system'. Must I not say that this proposition on the one hand is true, and on the other hand is unprovable? For suppose it were false; then it is true that it is provable. And that surely cannot be! And if it is proved, then it is proved that it is not provable. Thus it can only be true, but unprovable."

   Just as we ask, " 'Provable in what system?", so we must also ask: " 'True in what system?" 'True in Russell's system' means, as was said: proved in Russell's system; and 'false in Russell's system' means: the opposite has been proved in Russell's system.—Now what does your "suppose it is false" mean? *In the Russell sense* it means, 'suppose the opposite is proved in Russell's system'; *if that is your assumption* you will now presumably give up the interpretation that it is unprovable. And by 'this interpretation' I understand the translation into this English sentence.—If you assume that the proposition is provable in Russell's system, that means it is true *in the Russell sense*, and the interpretation "P is not provable" again has to be given up. If you assume that the proposition is true in the Russell sense, *the same* thing follows. Further: if the proposition is supposed to be false in some other than the Russell sense, then it does not contradict this for it to be proved in Russell's system (What is called "losing" in chess may constitute winning in another game.)

9. For what does it mean to say that P and "P is unprovable" are the same proposition? It means that these *two* English sentences have a single expression in such-and-such a notation.

10. "But surely P cannot be provable, for, supposing it were proved, then the proposition that it is not provable would be proved." But if this were now proved, or if I believed—perhaps through an error—that I had proved it, why should I not let the proof stand and say I must withdraw my interpretation "*unprovable*"?

11. Let us suppose I prove the unprovability (in Russell's system) of P; then by this proof I have proved P. Now if this proof were one in Russell's system—I should in that case have proved at once that it belonged and did not belong to Russell's system.—That is what comes of making up such sentences.—But there is a contradiction here!—Well, then there is a contradiction here. Does it do any harm here?

12. Is there harm in the contradiction that arises when someone says: "I am lying.—So I am not lying.—So I am lying.—etc."? I mean: does it make our language less usable if in this case, according to the ordinary rules, a proposition yields its contradictory, and vice versa?—the proposition itself is unusable, and these inferences equally; but why should they not be made?—It is a profitless performance!—It is a language-game with some similarity to the game of thumb-catching.

13. Such a contradiction is of interest only because it has tormented people, and because this shews both how tormenting problems can grow out of language, and what kind of things can torment us.

14. A proof of unprovability is as it were a geometrical proof; a proof concerning the geometry of proofs. Quite analogous e.g. to a proof that such-and-such a construction is impossible with ruler and compass. Now such a proof contains an element of prediction, a physical element. For in consequence of such a proof we say to a man: "Don't exert yourself to find a construction (of the trisection of an angle, say)—it can be proved that it can't be done". That is to say: it is essential that the proof of unprovability should be capable of being applied in this way. It must—we might say—be a forcible reason for giving up the search for a proof (i.e. for a construction of such-and-such a kind).

    A contradiction is unusable as such a prediction.

15. Whether something is rightly called the proposition "X is unprovable" depends on how we prove this proposition. The proof alone shews what counts as the criterion of unprovability. The proof is part of the system of operations, of the game, in which the proposition is used, and shews us its 'sense'.

    Thus the question is whether the 'proof of the unprovability of P' is here a forcible reason for the assumption that a proof of P will not be found.

16. The proposition "P is unprovable" has a different sense afterwards—from before it was proved. If it is proved, then it is the terminal pattern in the proof of unprovability.—If it is unproved, then what is to count as a criterion of its truth is not yet clear, and—we can say—its sense is still veiled.

17. Now how am I to take P as having been proved? By a proof of unprovability? Or in some other way? Suppose it is by a proof of unprovability. Now, in order to see *what* has been proved, look at the proof. Perhaps it has here been proved that such-and-such forms of proof do not lead to P.—Or, suppose P has been proved in a direct way—as I should like to put it—and so in that case there follows the proposition "P is unprovable", and it must now come out how this interpretation of the symbols of P collides with the fact of the proof, and why it has to be given up here.

    Suppose however that not-P is proved.—Proved how? Say by P's being proved directly—for from that follows that it is provable, and hence not-P. What am I to say now, "P"" or "not-P"? Why not both? If someone asks me "Which is the case, P, or not-P?" then I reply: P stands at the end of a Russellian proof, so you write P in the Russellian system; on the other hand, however, it is then provable and this is expressed by not-P, but this proposition does not stand at the end of a Russellian proof, and so does not belong to the Russellian system.

—When the interpretation "P is unprovable" was given to P, this proof of P was not known, and so one cannot say that P says: *this* proof did not exist.— Once the proof has been constructed, this has created a *new situation*: and now we have to decide whether we will call *this* a proof (a *further* proof), or whether we will still call *this* the statement of unprovability.

Suppose not-P is directly proved; it is therefore proved that P can be directly proved! So this is once more a question of interpretation—unless we now also have a direct proof of P. If it were like that, well, that is how it would be.

(The superstitious dread and veneration by mathematicians in face of contradiction.)

18. "But suppose, now, that the proposition were *false*—and hence provable?"— Why do you call it 'false'? Because you can see a proof?—Or for other reasons? For in that case it doesn't matter. For one can quite well call the Law of Contradiction false, on the grounds that we very often make good sense by answering a question "Yes and no". And the same for the proposition '$\sim\sim p = p$' because we employ double negation as a *strengthening* of the negation and not merely as its cancellation.

19. You say: "..., so P is true and unprovable". That presumably means: "Therefore P". That is all right with me–but for what purpose do you write down this 'assertion'? (It is as if someone had extracted from certain principles about natural forms and architectural style the idea that on Mount Everest, where no one can live, there belonged a châlet in the Baroque style. And how could you make the truth of the assertion plausible to me, since you can make no use of it except to do these bits of legerdemain?

20. Here one needs to remember that the propositions of logic are so constructed as to have *no* application as *information* in practice. So it could very well be said that they were not *propositions* at all; and one's writing them down at all stands in need of justification. Now if we append to these 'propositions' a further sentence-like structure of another kind, then we are all the more in the dark about what kind of application this system of sign-combinations is supposed to have; for the mere *ring of a sentence* is not enough to give these connexions of signs any meaning.

References

Anderson, A. R. (1958). "Mathematics and the 'Language Game' ". *The Review of Metaphysics 11*(3), 446–458.

Bays, T. (2004). "On Floyd and Putnam on Wittgenstein on Gödel". *Journal of Philosophy 101*(4), 197–210.

Bays, T. (2006). "Floyd, Putnam, Bays, Steiner, Wittgenstein, Gödel, Etc." Unpublished draft. Available at: http://www3.nd.edu/ tbays/papers/wnp2.pdf.

Benacerraf, P. and H. Putnam (Eds.) (1964). *Philosophy of Mathematics: Selected Readings, first edition.* Prentice-Hall, New Jersey.

Bernays, P. (1959). "Betrachtungen zu Ludwig Wittgensteins 'Bemerkungen über die Grundlagen der Mathematik' ". *Ratio II*(1), 1–22. Appeared in English translation in (Benacerraf and Putnam 1964) as "Comments on Ludwig Wittgenstein's *Remarks on the Foundations of Mathematics*".

Berto, F. (2009a). "The Gödel Paradox and Wittgenstein's Reasons". *Philosophia Mathematica 3*(17), 208–219.

Berto, F. (2009b). *There's Something About Gödel: The Complete Guide to The Incompleteness Theorem.* Wiley–Blackwell.

Bremer, M. (2005). *An Introduction to Paraconsistent Logics.* Lang, Frankfurt am Main.

Chihara, C. S. (1977). "Wittgenstein's Analysis of the Paradoxes in His Lectures on the Foundations of Mathematics". *Philosophical Review 86*(365–381).

Church, A. (1956). *Introduction to Mathematical Logic.* Princeton University Press, Princeton, New Jersey.

Dawson, J. W. (1988). "The Reception of Gödel's Incompleteness Theorems". Printed in S. Shanker (Ed.), *Gödel's Theorem in Focus.* Croom Helm.

Dummett, M. (1959). "Wittgenstein's Philosophy of Mathematics". *The Philosophical Review 68*(3), 324–348.

Dummett, M. (1997). “Review: Wittgenstein’s Philosophy of Mathematics by Pasquale Frascolla; Wittgenstein on Philosophy and Mathematics by Christoffer Gefwert”. *The Journal of Philosophy 94*(7), 359–374.

Floyd, J. (1995). “On Saying What You Really Want to Say: Wittgenstein, Gödel and the Trisection of the Angle”. Printed in J. Hintikka (Ed.), *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, pp. 373–425. Kluwer Academic Publishers.

Floyd, J. (2001). “Prose versus Proof: Wittgenstein on Gödel, Tarski and Truth”. *Philosophia Mathematica 9*(3), 280–307.

Floyd, J. and H. Putnam (2000). “A Note on Wittgenstein’s “Notorious Paragraph” about the Gödel Theorem”. *The Journal of Philosophy 97*(11), 624–632.

Floyd, J. and H. Putnam (2006). “Bays, Steiner, and Wittgenstein’s “Notorious Paragraph” about the Gödel Theorem”. *The Journal of Philosophy 103*(2).

Gerrard, S. (1991). “Wittgenstein’s Philosophies of Mathematics”. *Synthese 87*(1), 125–142.

Gödel, K. (1931/1976). “On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I”. Printed in J. van Heijenoort (Ed.), *From Frege to Gödel: A Sourcebook in Mathematical Logic, 1879–1931*. Harvard University Press, Cambridge, Mass.

Goodstein, R. L. (1957). “Critical Notice of Remarks on the Foundations of Mathematics”. *Mind 66*(264), 549–553.

Goodstein, R. L. (1972). “Wittgenstein’s Philosophy of Mathematics”. Printed in A. Ambrose and M. Lazerowitz (Eds.), *Ludwig Wittgenstein: Philosophy and Language*, pp. 271–286. Allen & Unwin, London.

Kreisel, G. (1958). “Wittgenstein’s Remarks on the Foundations of Mathematics”. *The British Journal for the Philosophy of Science 9*(34), 135–158.

Kreisel, G. (1978). “Book Review: Wittgenstein’s Lectures on the Foundations of Mathematics”. *Bulletin of the American Mathematical Society 84*(1), 79–90.

Kreisel, G. (1983). “Einige Erläuterungen zu Wittgensteins Kummer mit Hilbert und Gödel”. Printed in P. Weingartner and J. Czermak (Eds.), *Epistemology and Philosophy of Science. Proceedings of the 7th International Wittgenstein Symposium*, pp. 295–303. Höldner-Pichler-Tempsky Verlag.

Kreisel, G. (1998). “Second Thoughts around Some of Gödel’s Writings: “A Non-Academic Option” ”. *Synthese 115*(1, Gödel I), 99–160.

Lampert, T. (2006). "Wittgenstein's "notorious paragraph" about the Gödel Theorem". Printed in G. Gasser, C. Kanzian, and E. Runggaldier (Eds.), *Papers of the 29th International Wittgenstein Symposium in Kirchberg am Weschel*.

Monk, R. (1991). *Wittgenstein: The Duty of Genius*. Penguin Books, New York.

Monk, R. (2007). Bourgeois, Bolshevist or Anarchist? The Reception of Wittgenstein's Philosophy of Mathematics. Printed in G. Kahane, E. Kanterian, and O. Kuusela (Eds.), *Wittgenstein and his Interpreters*. Blackwell Publishing.

Moore, G. (1959). "Wittgenstein's Lectures in 1930–1933". Printed in *Philosophical Papers*. Allen & Unwin, London.

Nagel, E. and J. R. Newman (1958). *Gödel's Proof*. New York University Press, New York.

Nedo, M. and M. Ranchietti (Eds.) (1983). *Wittgenstein: Sein Leben in Bildern und Texten*. Suhrkamp Verlag, Frankfurt am Main.

Priest, G. (1979). "The Logic of Paradox". *Journal of Philosophical Logic 8*, 219–241.

Priest, G. (1984). "Logic of Paradox Revisited". *Journal of Philosophical Logic 13*, 153–179.

Priest, G. (1994). "Is Arithmetic Inconsistent?". *Mind 103*(411), 337–349.

Priest, G. (2006). *In Contradiction, second edition*. Oxford University Press, Oxford.

Rodych, V. (1999). "Wittgenstein's Inversion of Gödel's Theorem". *Erkenntnis 51*, 173–206.

Rodych, V. (2002). "Wittgenstein on Gödel: The Newly Published Remarks". *Erkenntnis 56*, 379–97.

Rodych, V. (2003). "Misunderstanding Gödel: New Arguments about Wittgenstein and New Remarks by Wittgenstein". *Dialectica 57*(3), 279–313.

Routley, R. (1979). "Dialectical Logic, Semantics and Metamathematics". *Erkenntnis 14*, 301–331.

Shanker, S. G. (1987). *Wittgenstein and the Turning-Point in the Philosophy of Mathematics*. Routledge.

Shanker, S. G. (1988). "Wittgenstein's Remarks on the Significance of Gödel's Theorem". Printed in S. G. Shanker (Ed.), *Gödel's Theorem in Focus*. Croom Helm, New York.

Skolem, T. (1934). “Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen”. *Fundamenta Mathematicae 23*(1), 150–161.

Steiner, M. (2001). “Wittgenstein as his Own Worst Enemy: The Case of Gödel’s Theorem”. *Philosophia Mathematica 9*(3), 257–279.

Wang, H. (1987). *Reflections on Kurt Gödel.* MIT Press, Cambridge, Mass.

Watson, A. (1938). “Mathematics and Its Foundations”. *Mind 47*(188), 440–451.