Some remarks on "*Remarks*..."
(Wittgenstein's Gödelian Argument)

Virgil Drăghici
Babeş-Bolyai University Cluj-Napoca

Abstract: The purpose of this paper is to make a critical analyse of the Wittgenstein's argument regarding his rejection of the existence of a true but unprovable sentence in *PA*. The section 1 does make explicit the way in which Gödel's first incompleteness theorem is analogous to paradoxical constructions (Richard's paradox and the Liar), *via* Cantor's diagonalization (1.1), and a review of proofs of Gödel's theorems (1.2). It is argued that Wittgenstein's rejection of Gödel's results is essentially based on his finitism (constructivism) conjugated with the thesis that the meaning of an expression is given by its use in a calculus (2.1). The consequences of this stance are the identification of "true" with "proved" and the rejection of the existence of meta-mathematics. These are finally responsible for Wittgenstein's rejection, in § 8 of his *Remarks...*, of the existence of a true but not provable sentence in *PA* (2.2). In the section 3 we are looking for a sense in which a constructive notion of truth can be given. This is considered in terms of Turing computability, but whose consequence, concerning Gödel's sentence, is that the idea of constructive truth does not coincide with the idea of truth given computationally.

Keywords: paradoxes, diagonalization, Gödel's theorems, Wittgenstein, T-computability, Kleene, constructive truth.

1. Gödelian Incompleteness

1.1. Cantor, Richard, Liar Sentence and Gödel

At the beginning of this paper let us review the content of this remarkable discovery of the 20th century. Gödel's sentence *G*, asserting its own unprovability, is a sentence of the form "*G* is not provable". If the system *S* in view is correct, i.e. it does not prove false sentences, then *G* is an example of a *true* but *not provable* sentence of *S*. For let us suppose that *G* were false. This implies, according to its own meaning that *G* is provable, contradicting the correctness of *S*. So *G* is true. On the other hand, if *G* were provable then it will be false and hence not provable in *S* (by the correctness of *S*). So *G* is not provable in *S* (by *reductio*). *G* being true, its negation is false, so ¬*G* is also not provable in *S*. Therefore *G* is an example of an undecidable sentence of *S*.

As Gödel said, "[t]the analogy of this argument with the Richard antinomy leaps to the eye. It is closely related to the "Liar" too; [fn omitted]."[1]

Let us see, more closely, in what this analogy consists. First of all, the sentence *G* is of self-referential kind. Its *formal* counterpart can be achieved by using the *diagonalization* method, due to Cantor.[2] He used this method in order to prove the non-denumerability of the set of all infinite binary sequences, in the following way. If *S* is an arbitrary set, let us consider the two-valued (0 and 1) functions defined over it. Now, suppose that, for each $x \in S$, f_x is such a function. Let *g* be the function defined by: $g(x) = f_x(x) + 1 (\mathrm{mod}\, 2)$. *Then g is not one of the functions* f_x, with $x \in S$. For if it were then let *k* be its index, i.e. $g(x) = f_k(x)$. But by definition of *g*, the following equation holds, for all values of *x*: $f_k(x) = f_x(x) + 1 (\mathrm{mod}\, 2)$. Let *x* be *k* (diagonalization step); then the following contradiction is obtained: $f_k(k) = f_k(k) + 1 (\mathrm{mod}\, 2)$.

This argument involves the application of each function with index *x* to its own index.

From the above result it follows *Cantor's Theorem*: For any set *S*, $\overline{\overline{S}} < \overline{\overline{P(S)}}$.

For let *S* be an arbitrary set and *f* a function that associates with every $x \in S$ a subset $f(x) \subseteq S$. Now we define the subset S^* by $x \in S^*$ iff $x \notin f(x)$. From the supposition that there would be a $z \in S$ such that

---

[1] K. Gödel [1931], 149.
[2] G. Cantor [1891].

$S^{*}=f(z)$ a contradiction can be derived. For such an z we have $z \in S^{*}$ iff $z \notin f(z)$ iff $z \notin S^{*}$. What this argument shows is that *a set cannot be equinumerous with its power set*.

By using the Cantor's diagonalization method Richard[3] constructed the respective paradox. Let us review this result and its relevance.

From any finite alphabet of a natural language one can construct a sequence whose members are all finite strings of letters, by using the space as a new symbol. Some strings will be, of course, definitions of real numbers in (0,1). So we separate them by crossing out all those strings that are not definitions of real numbers, and enumerate them in a list *E*, in alphabetical order. Let u_i, $i=1,2,3,...$ be the number defined by the i^{th} definition in the list *E*. In this way we obtained all numbers that are defined by finite strings of letters, written in a definite order. Hence the set of numbers definable by finitely strings of letters is one denumerably infinite. But, as Richard says, a contradiction appears in the following way.

> We can form a number not belonging to this set. "Let *p* be the digit in the *n*th decimal place of the *n*th number of the set *E*; let us form a number having 0 for its integral part and, in *n*th decimal place, *p*+1 if *p* is not 8 or 9, and 1 otherwise." This number *N* does not belong to the set *E*. If it were the *n*th number of the set *E*, the digit in its *n*th decimal place would be the same as the one in the *n*th decimal place of that number, which is not the case.
> I denote by *G* the collection of letters between quotation marks.
> The number *N* is defined by the words of the collection *G*, that is, by finitely many words; hence it should belong to the set *E*. But we have seen that it does not.
> Such is the contradiction.[4]

Let us accommodate Richard's description of his paradox to the Cantor's reasoning above, used to prove the non-denumerability of all

[3] J. Richard [1905].
[4] Citation from van Heijenoort [1967], 143.

infinite binary sequences. u_i is the number defined by the i^{th} definition, and let $f_i(n)$ be the n^{th} digit of the decimal expansion of u_i. If, for example, u_3 is the number 0.201475 then $f_3(5)=7$. Then the number *N*, according to the description (definition) given in the quotation above, is the number $0.g(1)g(2)g(3)...$, where $g(n)=f_n(n)+1$ if $f_n(n)<8$ and $g(n)=1$ otherwise. Certainly, *N* differs from all u_i, but the definition given in quotation marks in Richard's text above (denoted by *G*) shows that such a number can be defined in a natural language (English here). This is the paradox.

According to Richard, this paradox "in only apparent", for the definition of *N* refers to the set *E*, "which has not been defined. Hence I have to cross it out. The collection *G* has meaning only if the set *E* is totally defined, and this is not done except by infinitely many words. *Therefore there is no contradiction*."[5]

Actually, the definition of *N*, i.e. of $g(x)$, is circular: it appeals to the enumeration *E* in which it is an element. For let us suppose that the definition *G* (of the number *N*) has the index *k* in the enumeration *E*. I.e. $g(x)=f_k(x)$. But according to its definition, $g(x)=f_x(x)+1$. So the following equation holds: $f_k(x)=f_x(x)+1$. For $x=k$ we get the contradiction $f_k(k)=f_k(k)+1$, by a reasoning similar to Cantor's above.[6]

Now, from this Richard's result, via Cantor's Theorem, the Liar Paradox can easily be derived.

Let *Nat* be the set of natural numbers and *E* an enumeration of all definitions of sets of natural numbers. A set *S* of natural numbers is formally definable iff there is a formula α(x), with x free, such that for every *n* holds:

$n \in S$ iff $\alpha(\boldsymbol{n})$ is true,

---

[5] van Heijenoort [1967], 143.

[6] Today the paradox is taken as solved either by introducing the distinction of language levels (i.e. the definition of $g(x)$ does belong to the meta-language) or by considering the partial functions (i.e. in the above diagonalization step, $x=k$, the function $f_k(k)$ is undefined, and the contradiction disappears).

where $\boldsymbol{n}$ is the numeral for n. If $Tr(\mathrm{x})$ is the formula expressing the truth predicate, then the above (meta-) equivalence becomes

1. $n \in S$ iff $Tr(\ulcorner\alpha(\boldsymbol{n})\urcorner)$, respective $n \notin S$ iff $\neg Tr(\ulcorner\alpha(\boldsymbol{n})\urcorner)$,

where $\ulcorner\alpha(\boldsymbol{n})\urcorner$ is the name for the formula $\alpha(\boldsymbol{n})$. $\alpha(\boldsymbol{n})$ is the fixed point of the formula $Tr(\mathrm{x})$. By the Diagonal Lemma the following equivalence holds:

2. $\alpha(\boldsymbol{n}) \equiv Tr(\ulcorner\alpha(\boldsymbol{n})\urcorner)$.

Hence $n \in S$ iff $\alpha(\boldsymbol{n})$.

Let us consider E^* an enumeration of all formulas of the form $\alpha(\mathrm{x})$, corresponding to the definitions in E. Let us define the set K thus:

$n \in K$ iff "n does not belong to the set defined
by the formula with the index n"

that is,

3. $n \in K$ iff $\neg Tr(\ulcorner\alpha_n(\boldsymbol{n})\urcorner)$.

If $\alpha_k(\mathrm{x})$ is the formula defining K, then for all n:

$n \in K$ iff $\alpha_k(\boldsymbol{n})$. But $\alpha_k(\boldsymbol{n}) \equiv \neg Tr(\ulcorner\alpha_k(\boldsymbol{n})\urcorner)$, by 2, and $Tr(\ulcorner\alpha_k(\boldsymbol{n})\urcorner) \equiv \neg Tr(\ulcorner\alpha_n(\boldsymbol{n})\urcorner)$, by 3. From where, by setting $n = k$, the *Liar Paradox* follows:

4. $Tr(\ulcorner\alpha_k(\boldsymbol{k})\urcorner) \equiv \neg Tr(\ulcorner\alpha_k(\boldsymbol{k})\urcorner)$,

or, by 2

5. $\alpha_k(\boldsymbol{k}) \equiv \neg Tr(\ulcorner\alpha_k(\boldsymbol{k})\urcorner)$,

just the *Liar sentence*.

What shows the above argument is that the set K is not definable by any formula in E^*.

As we saw above, the semantical predicates "definable" and "true" lead to paradoxical constructions. Gödel's sentences G replaces the predicate "true" with the syntactical predicate "provable", obtaining a sentence asserting its own unprovability: *G is not provable*. In terms of the above symbolism, 5 becomes

5^*. $\alpha_k(\boldsymbol{k}) \equiv \neg Bew(\ulcorner\alpha_k(\boldsymbol{k})\urcorner)$,

where $Bew(\mathrm{x})$ is the formula expressing the provability predicate in a formal language. By a simple reasoning we can see that the assumption of provability of $\alpha_k(\boldsymbol{k})$ or of its negation leads to a contradiction. But in that case the conclusion is: neither *G*, i.e. the sentence $\alpha_k(\boldsymbol{k})$, nor its negation is provable. Therefore, *G is true but not provable* in *S*.

1.2. Gödel's Incompleteness Theorems

The formulation and proof of these theorems are based on the two fundamental facts: a) the arithmetization of syntax, a process by which the sentences and proofs about the system *PA* (of Peano Arithmetic) are "translated" in recursive functions and relations, and b) the representability and expressibility of these functions, respective relations in formal *PA*.

By using this idea of arithmetization Gödel defines the following relations:

$$xBy = Bw(x)\ \&\ [l(x)]Glx = y\ ,$$

meaning that *x* is a proof ($Bw(x)$) whose last formula is *y*. I.e. "*xBy*" means that *x is a proof of y*. It is usually rendered by $Pf(x, y)$.

$$Bew(x) \equiv \exists y yBx\ ,$$

meaning that *x is a provable formula*.[7]

By the *diagonalization* of a formula α(x) containing x free, is meant the formula obtained from α(x) by substituting the numeral of the Gödel number of α(x) for x in α, i.e. $\alpha(\boldsymbol{n})$, where $\boldsymbol{n}$ is the respective numeral.[8]

Correspondingly, a *diagonal function* $\delta(n)$ can be defined in the following way: if *n* is the Gödel number of a formula α(x), with x free, then $\delta(n)$ is the Gödel number of its diagonalization, $\alpha(\boldsymbol{n})$.

If by *Sub*(*y,z,w*) is meant "Gödel number of the formula resulting from the formula with Gödel number *y* by substituting the term with Gödel number *z* for all free occurrences of the variable with Gödel number *w*", then $\delta(n)$ can be defined as:

$$\delta(n) = Sub(n, Num(n), 2^a)\ ,$$

where $Num(n)$ is the Gödel number of the numeral for *n* and 2^a is the Gödel number of the free variable of the formula with Gödel number *n*.

---

[7] The definitions 45 and 46, respectively, in Gödel's [1931] paper.

[8] The diagonalization of α(x) can also be rendered as $\exists x(x=\ulcorner\alpha\urcorner \wedge \alpha)$, where $\ulcorner\alpha\urcorner$ is the name (numeral of Gödel number) of α.

Definition 1. *A number-theoretic relation* $R(x_1,...,x_n)$ *is expressible in PA iff there is a formula* α(x₁,...,xₙ) *of PA with n free variables such that for any natural numbers* $k_1,...,k_n$ *the following hold:*

1. *If* $R(k_1,...,k_n)$ *is true, then* $\vdash \alpha(\boldsymbol{k}_1,...,\boldsymbol{k}_n)$.

2. *If* $R(k_1,...,k_n)$ *is false, then* $\vdash \alpha(\boldsymbol{k}_1,...,\boldsymbol{k}_n)$.

Definition 2. *The function* $f(x)$ *is representable in PA if there is a formula* α(x,y), *with* x, y *free, such that whenever* $f(k_1) = k_2$ *the following holds:* $PA \vdash \forall y(\alpha(\boldsymbol{k}_1, y) \equiv y = \boldsymbol{k}_2)$.

Finally, the following facts are theorems of formal number theory:

Fact 1. *A number-theoretic relation* $R(x_1,...,x_n)$ *is recursive iff it is expressible in PA.*

Fact 2. *A number-theoretic relation* $R(x_1,...,x_n)$ *is recursive iff it is representable in PA.*

Let now $R(k, y)$ be the following relation: "*k* is the Gödel number of a formula α(x) containing x free and *y* is the Gödel number of a proof in *PA* of its diagonalization, $\alpha(\boldsymbol{k})$.

This relation is primitive recursive, for it is equivalent to the expression: $Form(k) \wedge Fr(k, 2^a) \wedge Pf(y, Sub(k, Num(k), 2^a)$, where all predicates occurring in it: $Form(y)$: "*y* is the Gödel number of a formula", $Fr(y, 2^a)$: "the formula with Gödel number *y* contains the free variable with Gödel number 2^a", $Pf(y, x)$: "*y* is the Gödel number of a proof of the formula with Gödel number *x*", $Num(y)$: "the Gödel number of the numeral for *y* (i.e. ***y***) are primitive recursive, and the function $Sub(x, y, z)$: "the Gödel number of the expression resulting from the expression with Gödel number *x* by substituting the term with Gödel number *y* for all free occurrences of the variable with Gödel number *z*" is primitive recursive.

Being primitive recursive, $R(k, y)$ is expressible in *PA* by a formula with two free variables, say β(x,y). That is, the conditionals 1 and 2 above hold.

Let now F be the formula $\forall y \neg\beta(x,y)$. Since β expresses in PA the relation R, its intuitive meaning is: "For all y, y is not a (Gödel number of a) proof of the diagonalization of the formula (with Gödel number) x", or simply, "the diagonalization of x is not provable". Let k be the Gödel number of F. Let now construct the formula G: $\forall y \neg\beta(\boldsymbol{k},y)$, i.e. the diagonalization of F. Its intuitive meaning is: "the diagonalization of k is not provable". But the diagonalization of k is just the formula G. Therefore, G is a self-referential construction, asserting its own unprovability.

Gödel's first incompleteness theorem

1. *If PA is consistent, then G is not provable.*

2. *If PA is* ω*-consistent,*[9] *then* $\neg G$ *is not provable.*

Proof. 1 (*Reductio*). Suppose that PA is consistent and G is provable in PA. Hence there is a number n the Gödel number of a proof of G. This means that $R(k,n)$ holds. By expressibility of R in PA by β it follows that $\beta(\boldsymbol{k},\boldsymbol{n})$ is provable in PA. But from the supposition of the provability of G: $\forall y \neg\beta(\boldsymbol{k},y)$ results that $\neg\beta(\boldsymbol{k},\boldsymbol{n})$ is also provable, contradicting the consistency of PA.

2. (*Reductio*). Suppose that PA is ω-consistent and $\neg G$: $\exists y \beta(\boldsymbol{k},y)$ is provable. PA is consistent (it follows from ω-consistency). So there is no m such that m is the Gödel number of a proof of G in PA. That means that $R(k,m)$ is false for every m, and consequently $\neg\beta(\boldsymbol{k},\boldsymbol{m})$ is provable for every m, and this contradicts the assumption of ω-consistency of PA.

Therefore this formula G is an example of an undecidable sentence of PA. *How do we know that G is true*?

By considering the meaning of R and the expressibility of R in PA by the formula β, the intuitive meaning of G is that there is no number y such that y is the Gödel number of a proof of G in PA, i.e. $R(k,y)$ is false for every y, equivalent G asserts its own unprovability. By Gödel's theorem 1, if PA is consistent, then G is really unprovable. So G is true.

---

[9] A system S is ω-consistent iff for every formula $\alpha(x)$ of S if $\alpha(\boldsymbol{n})$ is provable for every n, then $\exists x \neg\alpha(x)$ is not provable in S.

Gödel's second incompleteness theorem

As we saw above, the relation $Pf(y,x)$ means "*y* is a proof of *x*". Now let *c* be the Gödel number of the formula $\boldsymbol{0=1}$.[10] If $Pf(y,x)$ is expressible in *PA* by the formula $\gamma(y,x)$, then the formula $\forall y\neg\gamma(y,\boldsymbol{c})$ expresses the consistency of *PA*. Let *Con* be this formula.

Gödel's second incompleteness theorem

If PA is consistent, then $PA\nvdash Con$.

I.e. if *PA* is a consistent theory, then the formula *Con*, expressing its consistency, is not provable in *PA*.

The proof of this theorem is simple. First, the equivalence $G\equiv Con$ is provable in *PA*, for $Con\supset G$ is just the formalization of the first incompleteness theorem, where *G* is the undecidable sentence $\forall y\neg\beta(\boldsymbol{k},y)$. Second, the converse $G\supset Con$ is trivial, since if there is a sentence not provable in *PA*, then *PA* does not prove the sentence $\boldsymbol{0=1}$, i.e. *PA* is consistent. By this equivalence $G\equiv Con$, the second incompleteness theorem follows from the first.

Gödel's first incomplenetess theorem (via Diagonal Lemma)

Diagonal Lemma. *If T is a theory extending Q (Robinson), then for any formula* $\alpha(x)$ *there is a sentence S such that* $T\vdash S\equiv\alpha(\boldsymbol{s})$, *where s is the Gödel number of S.*

Now, let $\gamma(y,x)$ be the above formula expressing the relation $Pf(y,x)$ in *PA*. Let $\alpha(x)$ from Diagonal Lemma be the formula $\forall y\neg\gamma(y,x)$ containing x free. According to this lemma, there is a sentence *G* such that:

$$PA\vdash G\equiv\forall y\neg\gamma(y,\boldsymbol{g}),$$

where *g* is the Gödel number of Gödel sentence *G*.

A sentence is provable in *PA* iff there exists a proof of it. Formally, the construction $\exists y\gamma(y,x)$ expresses the existence of a proof of the formula whose Gödel number is *x*, i.e. this formula is provable. Let *Bew*(x),

[10] Actually, the negation of any theorem can be set instead of this formula.

equivalent to $\exists y\gamma(y,x)$, be the formula expressing the provability predicate for *PA*. So, the provable equivalence by Diagonal Lemma becomes $G\equiv\neg Bew(\ulcorner G\urcorner)$, where $\ulcorner G\urcorner$ is the formal name of *G*.

The predicate *Bew*(x) does satisfy the HBL[11]-derivability conditions: 1. If $\vdash S$, then $\vdash Bew(\ulcorner S\urcorner)$, 2. $\vdash Bew(\ulcorner(S\supset T)\urcorner)\supset(Bew(\ulcorner S\urcorner)\supset Bew(\ulcorner T\urcorner))$ and 3. $\vdash Bew(\ulcorner S\urcorner)\supset Bew(\ulcorner Bew(\ulcorner S\urcorner)\urcorner)$, for all sentences *S* and *T* of *PA*.

An important means for the study of provability is modal logic, in the form of the axiomatic system *GL*. The correlation between arithmetic and modal logic is based on a metalinguistic translation of modal sentences in the language of *PA*. A translation (or realization) is a function (*) that associates to each atomic sentence of the modal system *GL* a sentence of *PA*, that commutes with the boolean connectives and $(\ \alpha)^*=Bew(\alpha^*)$. By the following theorem the modal system *GL* is proved to be arithmetical complete.

Solovay's theorem. $GL\vdash\alpha$ *iff for every arithmetic translation* (*) $PA\vdash\alpha^*$.[12]

Now, the proof of first incompleteness theorem is the following. By diagonal lemma

1. $PA\vdash G\equiv\neg Bew(\ulcorner G\urcorner)$
2. *Supp* $PA\vdash G$.
3. $PA\vdash Bew(\ulcorner G\urcorner)$, by first derivability condition.
4. $PA\vdash\neg G$, by (1) and (3).

So *PA* is inconsistent. Therefore $PA\nvdash G$.

1. *Supp* $PA\vdash\neg G$.
2. $PA\vdash Bew(\ulcorner\neg G\urcorner)$, by first derivability condition.
3. $PA\vdash Bew(\ulcorner G\urcorner)$, by diag. lemma.
4. $PA\vdash Bew(\ulcorner G\urcorner)\wedge Bew(\ulcorner\neg G\urcorner)$.
5. $GL\vdash(\ p\wedge\ \neg p)\supset\ \bot$.
6. $PA\vdash((\ p\wedge\ \neg p)\supset\ \bot)^*$, by Solovay's theorem.

[11] Hilbert-Bernays-Löb.
[12] R. Solovay [1976].

7. $PA \vdash Bew(\ulcorner\bot\urcorner)$, 4, 6.

But if $PA \nvdash Bew(\ulcorner\bot\urcorner)$, i.e. if PA is consistent, then $PA \nvdash \neg G$. So G is undecidable in PA.

2. Wittgenstein's *Remarks...*

2.1. Wittgenstein's finitism

By Gödel's first theorem, if PA is consistent then there is a sentence G true but not provable in PA. In his *Appendix* on Gödel's theorem,[13] Wittgenstein rejected this important result. According to his anti-descriptivism, arithmetic does not describe any kind of reality, so the idea that arithmetical truth transcends the formal derivability in an axiomatic system must be rejected. And with it the mathematical Platonism associated with Gödel's theorems must also be rejected. Let us see closely the main features of Wittgenstein's finitism.

On Wittgenstein's view the meaning of an expression is given by its use. As regards a mathematical proposition, it is meaningful only within a given calculus *and* where there is an effective procedure whose application allows us to decide it.[14] So the universal or existential quantified propositions, $\forall x\alpha(x)$ and $\exists x\alpha(x)$, with $\alpha(x)$ a decidable predicate, are meaningful only if the quantifiers are *bounded*, for only in these cases the respective proposition is *algoritmically decidable*. Since "[e]very proposition in mathematics must belong to a calculus of mathematics"[15] it is either provable or refutable. So the *tertium non datur* holds for any such a proposition. However, from the fact that *tertium* does apply to any meaningful mathematical proposition we cannot derive that such a proposition is either true or false. For on his anti-descriptivist view, the mathematics is an invention, not a discovery, hence by using a deduction procedure *we make* the respective proposition either true or false. In Wittgenstein's terms "true" is equated to "proved"/"provable" and "false" is

---

[13] L. Wittgenstein [1964], Part I, Appendix I.

[14] L. Wittgenstein [1975] § 152; [1974], 451: "[i]f there is no method provided for deciding whether the proposition is true or false, then it is senseless".

[15] L. Wittgenstein [1974], 376.

equivalent to "refuted". So :true" or "false" have a meaning with reference to mathematical proposition, only in a formal (syntactical) sense.[16]

2.2. The argument from *Remarks*...

Since a true proposition is a proved proposition " "*p*" is true" is just the assertion of "*p*".[17] So a proposition can be asserted in Russell's system only "at the end of one of his proofs, or as a 'fundamental law' (Pp)".[18] If this is the case, then the following equivalences hold: "*p*" is asserted = *p* ="*p*" is proved. And, consequently, "a true but unprovable proposition" is just a contradiction in terms. How can this conclusion be motivated? Essentially, in the following way.

On the one side, this idea has to do with Wittgenstein's rejection of Hilbertian program of meta-mathematics. For a mathematical proposition is meaningful iff it is algorithmically decidable and its decidability is related to a formal system, say Russell's. And "[u]nderstanding *p* means understanding its system. If *p* appears to go over from one system into another, then *p* has, in reality, changed its sense".[19] On the other side, the idea that meaning of an expression is determined by its use enters the picture. So the above conclusion is the result of Wittgenstein's rejection of meta-mathematics, as a direct consequence of his strict finitism, conjugated with the thesis *meaning is use*. Hence it is not the case that the *same* proposition *p* be expressible in a system and undecidable in that system, but provable (and then true) in a different system (in the metatheory). For "[...] a proposition which cannot be proved in Russell's system is "true" or "false" in a different sense from a proposition of *Principia Mathematica*".[20]

An additional feature of Wittgenstein's interpretation of Gödel's theorem is given in his § 8 of *Remarks*...

---

[16] In contrast with the genuine *contingent* propositions which are true/false by virtue of correspondence to facts (comp. *Tractatus*, 4.25, 4.062).

[17] Cf. [1964], § 6, Appendix I: "For what does a proposition's 'being true' mean? '*p*' is true = *p*. (That is the answer)."

[18] Cf. § 6.

[19] L. Wittgenstein [1975], § 3.

[20] L. Wittgenstein [1964] Appendix I, § 7.

> 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.)

The first part of this paragraph is somehow the "official" argumentation of the existence of a *true* but *unprovable* proposition in Russell's system. In both cases the inference made by his interlocutor, is by *reductio*. The argument as such is concise, so we think what Wittgenstein intended to argue is the following. The ingredients of proof are

1. Russell's system is correct, i.e. it does not prove false propositions.
2. The proposition *G* ("*P*" in his notation) is: "*G* is unprovable".

What is argued by *reductio* is

a) *G is true*, for if it were false, then according to its meaning *G* will be provable, contradicting 1. This is why "that surely cannot be".

b) *G is unprovable*. For if it were provable, by 2 *G* is false and by 1 it is not provable. Hence *G* is not provable (by *reductio*). Then by 2 *G* is true.

Being true, its negation, ¬*G*, is false and also unprovable. This is just an intuitive account of the reasoning behind Gödel's construction.

The second part of this paragraph is Wittgenstein's reply to the above argument. Again the identifications "true" = "proved" and "false" = "the opposite is proved" arise. But the novelty consists in the fact that with these identifications the interpretation of *G* as "*G* is unprovable" must be given up; that is, its interpretation ("translation") into English language must be cut out. Of course, Wittgenstein's reasoning is clear: if *G* is true and "true" means provable, then if *G* is provable then it is not unprovable. So the interpretation of *G* as "*G* is unprovable" cannot be kept. So the things are simple: under the above identifications "true" and "unprovable" are not compatible. Also "true and unprovable" is again a contradiction in terms.

What is the source of Wittgenstein's rejection of the existence of a true but not provable proposition in Russell's system?

I think that by "*if that is your assumption*", with reference to the above identification, Wittgenstein misinterprets the Gödel's result given by his first theorem, even in the intuitive fashion given above. For the second part of his § 8 is *not* a reply to what is given in the first part. If the first part is the Gödel's reasoning, then by his first theorem: *If Russell's system S is consistent, then G is not provable in S*. So by this theorem "true" and "provable" do *not* mean the same thing. Hence Wittgenstein's rejection of the interpretation of *G* as "*G* is unprovable" is not rejected by Gödel's result, but by Wittgenstein's identification of "true" with "provable".

Besides, the truth of the proposition *G* is argued *via* the arithmetization of the relation *R* and its expressibility in *PA* by a formula β. So by its very construction the *intuitive* meaning of *G* cannot be other than "*G* is not provable".[21] Of course, the argument that *G* is *true* cannot be formalized *in PA*. If "true" is equated with "proof" and the distinction

[21] Cf. 1.2 above.

"inside" – "outside" with reference to *PA* is omitted, then the results are just those encountered in this § 8 of Wittgenstein's *Remarks*.

The same omission is responsible for the argument given in § 10 of Appendix I:

> "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*"?

Here, of course, Wittgenstein is right in asserting that the interpretation of *G* as "*G* is unprovable" is not compatible with the assertion of provability of *G*. *But* this incompatibility only holds with reference to the *same* language. As such it is perfectly compatible with Gödel's incompleteness theorem and under the assumptions of this theorem.

In the same ton § 11 of *Appendix I* says:

> 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?

Here the *proof of the unprovability of G*, as above, does generate an inconsistency in the form that we proved at once that "[*G*] belonged and did not belong to Russell's system". True, the proof is this:

1. $G \equiv \neg Bew(\ulcorner G \urcorner)$, by the construction of *G* or by Diagonal Lemma.
2. $\vdash G$; assumption.
3. $\vdash Bew(\ulcorner G \urcorner)$; by first deriv. condition.
4. $\vdash \neg G$; 1, 3.
5. $\vdash \neg Bew(\ulcorner G \urcorner)$; 1, 2.

2 and 4 say that the system is inconsistent, 3 and 5 that *G* belongs and does not belong to the respective system. Is this result relevant in any way? In Gödel's terms it means, simply, that *G* is *not* provable in *S*. In Wittgenstein's terms it means that the idea "proof of an unprovable proposition" is inconsistent. Both answers are right. But if we decide to not operate with inconsistencies, then only the Gödel's proposal is the right one. Simply!

On Rodyck's view[22] Wittgenstein's mistake is just the idea that Gödel's proof does not admit the interpretation in English language of *G* as "*G* is not provable in Russell's system" (§ 8). For Gödel's result is just a number-theoretic one and does not depend on such a natural language translation. Of course, Rodyck is right in saying that "an actual proof of [*G*] would enable us to *calculate* the relevant Gödel numbers and thereby arrive at ["¬*G*"] by existential generalization".[23] Detailed, the argument is this: *G* is the proposition $\forall y \neg \beta(\boldsymbol{k},y)$, i.e. the diagonalization of the formula with Gödel number *k*: $\forall y \neg \beta(x,y)$, and $\beta(x,y)$ formally expresses the relation $R(x,y)$: "*y* is a proof of the diagonalization of *x*" (cf. 1.2 above). Now if *G* were provable, then there would be an *n* the Gödel number of its proof (and *n* can be effectively calculated). So $R(k,n)$ holds and, consequently, $\beta(\boldsymbol{k},\boldsymbol{n})$ would be provable in *PA*. From where $\exists y \beta(\boldsymbol{k},y)$ would also be provable (by existential generalization), i.e. $\neg G$ would be provable.

Of course, this is a number-theoretic result. So the proof of the unprovability of *G* is a *mathematical* fact. But if we want to go *beyond* the formal apparatus implied in Gödel's construction, then *metamathematically* what *G* is supposed to *mean*? All we have to do is to go back, *via* arithmetization, to the intuitive relation corresponding to *G* and then to see that it is just an (self-referential) assertion of the form "I am not provable". So we ask the following question: can we retain the mathematical facts related to Gödel's results and reject entirely as *irrelevant* the interpretation

[22] V. Rodyck [1999], 182-183.
[23] V. Rodyck [1999], 182.

in the natural language? By considerations of 1.1 above, this is not the case.[24]

Then how do we explain Wittgenstein's assertion that by Gödel's theorem *G* cannot be translated in English as "*G* is unprovable"? Again the answer is simple. This assertion is part of the same paragraph 8 and occurs immediately after the identification of "truth" with "proved". By suppressing the distinction object-language / meta-language, this identification leads directly to a contradiction in terms. Or, by keeping this distinction and the supposition of the consistency of Russell's system, it follows that "truth" and "provable" cannot be equated. So two very different conceptions about mathematics confront each other: the mathematics as strictly a syntactic affaire (Wittgenstein's) and mathematics as syntax + semantics (Gödel's).

3. Truth à la Tarski vs. constructive truth

In his seminal paper Tarski[25] defines the concept of truth in formalized language via satisfaction relation. Let $R(x_1,...,x_n)$ be an atomic formula of the L_P (the language of first-order predicate logic). Let $M = \langle D,i \rangle$ be a model of L_P. Let μ be an assignment of $x_1,...,x_n$ in *D*.

Definition 1. $R(x_1,...,x_n)$ *is true in M and an assignment* μ *in M iff* $\langle x_1^{\mu},...,x_n^{\mu} \rangle \in R^i$.

This means that the formula R is true in *M* and μ iff the sequence of *n* elements of *D* are in the relation assigned to R by the interpretation *i* of *M*.

The truth of non-atomic formulas, resulting by applications of propositional connectives, in *M* and an assignment μ in *M* commutes with these connectives.

---

[24] Moreover, the English expression "*G* is unprovable", has a perfectly formal correspondent: $\neg Bew(\ulcorner G \urcorner)$. By Diagonal Lemma $G \equiv \neg Bew(\ulcorner G \urcorner)$ is provable. Then *G* says about itself: "I am not provable".

[25] Cf. A. Tarski [1933].

Definition 2. $\forall x_i R(x_1,...,x_n)$ *is true in M and an assignment* μ *in M iff* R *is true in M and any assignment* ν x_i*-variant of* μ. (ν *is* x_i*-variant of* μ *if it differs from* μ *in at most the assignment made to* x_i).

Definition 3. $R(x_1,...,x_n)$ *is true in M iff* R *is true in any assignment* μ *in M*.

Definition 4. $\forall x_i R(x_1,...,x_n)$ *is true in M iff* R *is true in any assigment* μ *in M*.

If the considered language L_{PA} is that of Reano Arithmetic (*PA*), then if $R(x_1, x_2)$ is the formula $x_1 < x_2$, then R^i is the arithmetic (intuitive) relation $x_1^\mu < x_2^\mu$ and it is true in the standard model $M = \langle N, 0, =, ', +, \cdot \rangle$ iff the number x_1^μ is less than the number x_2^μ.

If "Snow is white" is the formal name for the respective sentence then, according to Definition 1, "Snow is white" is true iff Snow is white.[26]

To be sure, such a definition of truth is not constructive, it does not say anything about the ways the truth of a sentence can be verified. What about a constructive one?

At the beginning let us review some fundamental facts of recursion theory.

First of all, *any recursive predicate R is effectively decidable*. This is so because any recursive function *f* is effectively computable. For if *E* is a system of equations defining it recursively,[27] then the value of *f* can be found for any values of its arguments. The computability of *f* is given by the following theorem.

Normal form theorem.[28] *If* $f(x_1,...,x_n)$ $(n \geq 0)$ *is a recursive function a number e can be found such that:*

1. $(x_1)...(x_n)(Ey)T_n(e, x_1,...,x_n, y)$
2. $f(x_1,...,x_n) = U(\mu y T_n(e, x_1,...,x_n, y))$,

---

[26] This is the sense in which Tarski defines satisfiability of an atomic formula by an assigment *s*.

[27] Such a system *E* always exists (cf. Kleene [1964], § 54, Th. II and § 55.

[28] Cf. Kleene [1964], Th. IX.

where $T_n(z, x_1, ..., x_n, y)$ *is a primitive recursive predicate and* $U(y)$ *is a primitive recursion function.*[29]

In this theorem *e* is the Gödel number of *E* and it is also called the Gödel number of *f*. If an *e* is given such that 1 and 2 in the theorem above hold, then the recursive function *f* is effectively given.

Secondly, if $R(x_1,...,x_n)$ is a predicate, then its characteristic function C_R is defined as follows:

$$C_R(x_1,...,x_n) = \begin{cases} 0 & \text{if} \quad R(x_1,...,x_n) \quad \text{is true} \\ 1 & \text{if} \quad R(x_1,...,x_n) \quad \text{is false} \end{cases}$$

The following equivalences hold:

$R(x_1,...,x_n)$ is recursive iff $C_R(x_1,...,x_n)$ is recursive iff $C_R(x_1,...,x_n)$ is Turing-computable.

Therefore, if $R(x_1,...,x_n)$ is a recursive predicate then for any assignment $\mu = (k_1,...,k_n)$ to $x_1,...,x_n$, its logical value, *true* or *false*, can be determined. And this is possible by computing the value of its characteristic function $C_R(x_1,...,x_n)$ and see whether this is 0 or 1.

In what follows we adopt the following notation: if $R(x_1,...,x_n)$ is a recursive predicate then by "$R(x_1,...,x_n)$ is T^1-computable on the assignment $\langle k_1,...,k_n \rangle$" we understand that its characteristic function C_R takes the value 0 for this assignment, respectively *R* is true in this assignment.[30]

A constructive notion of truth

If by *M* we mean the standard model of *PA*, let *Constr* be the model $\langle N, 0, =, +, \cdot, ' \rangle$, with the same domain *N*, but in which the *satisfaction* (i.e. the truth in an assignment) and the truth are defined in the following way.

---

[29] (*x*) and (*Ex*) are Kleene's notation for the "for all" and "there is", respectively, in their *intuitive* use, and T_n is Kleene's "*T*-predicate".

[30] Correspondingly, T^0-computable means that *R* is false for this assignment.

Now let $R(x_1,...,x_n)$ be a formula of *PA* containing the free variables $x_1,...,x_n$.

Definition 1*. $R(x_1,...,x_n)$ *is true in Constr and an assignment* μ *in Constr iff* $\langle x_1^\mu,...,x_n^\mu\rangle \in R^i$, *i.e., iff the arithmetical relation* $R(x_1,...,x_n)$ *is* T^1*-computable on the assignment* μ.[31]

Definition 2*. $\forall x_i R(x_1,...,x_n)$ *is true in Constr and an assignment* μ *in Constr iff the arithmetical relation* $R(x_1,...,x_n)$ *is* T^1*-computable for any assignment* ν x_i*-variant of* μ.

Definition 3*. $R(x_1,...,x_n)$ *is true in Constr iff* $R(x_1,...,x_n)$ *is* T^1*-computable on any assignment* μ *in Constr*.

Definition 4*. $\forall x_i R(x_1,...,x_n)$ *is true in Constr iff* $R(x_1,...,x_n)$ *is* T^1*-computable on any assignment* μ *in Constr*.

If the formulas considered are the axioms of *PA*, then if an axiom is true in *Constr*, then the respective arithmetical relation is T^1-computable. Similarly, it can be argued that the inference rules, *modus ponens* and *Generalization* (Gen), preserve the truth in *Constr*.

The Gödel's undecidable sentence *G* and *Constr*

As we saw above, the sentence *G* is a universally quantified formula $\forall y\neg\beta(\boldsymbol{k},y)$, where *k* is the Gödel number of the formula $\forall y\neg\beta(x,y)$, with β decidable.

In its original form[32] the undecidable sentence 17*Gen r* is the following construction.

Gödel defines the relation $Q(x,y)$ as $\overline{xB\delta(y)}$, i.e. *Q* means: "*x* is not a proof of the diagonalization of the formula with Gödel number *y*". This is a recursive relation, for it is defined by using the (primitive) recursive relation *xBy* ("*x* is a proof of *y*") and the (primitive) recursive substitution

[31] This means that if μ is the assignment $(k_1,...,k_n)$ then $R(k_1,...,k_n)$ is T^1-computable on this μ.
[32] Cf. K. Gödel [1931], 175.

function (implied in defining the diagonalization of a formula[33]). Now, if $\mathrm{q(x, y)}$ is the formula expressing it in *PA*, whose free variables have the Gödel numbers 17 and 19,[34] then 17*Gen q* is the formula $\forall \mathrm{xq(x, y)}$, whose Gödel number is *p*. Further on, the diagonalization of this formula is $\forall$xq(x,***p***), i.e. 17*Gen r* in Gödel's notation (where *r* is the Gödel number of q(x,***p***)).

The intuitive meaning of the formula $\forall$xq(x,***p***) is therefore the sentence asserting that *the formula resulting by the diagonalization of the formula whose Gödel number is p*, i.e., *just this formula, is not provable in PA*. This formula is a self-referential construction. Accordingly, the formula q(x,***p***) is the formal counterpart of *Q*(*x*,*p*). In our notation *G* is the formula $\forall y \neg \beta(\boldsymbol{k}, y)$ and therefore $\neg \beta(\boldsymbol{k}, y)$ replaces q(x,*p*) with the same intuitive meaning.

As we mentioned above (sec. 1.2), by Solovay's theorem there is a close relationship between the modal system *GL* and Peano Arithmetic (arithmetical completeness of *GL*). But there is also a correspondence between the provability of a formula in *PA* and its computability version.

> Classically, there are two equivalent ways to look at the mathematical notion of proof: logical, as a finite sequence of sentences strictly obeying some axioms and inference rules, and computational, as a specific type of computation. Indeed, from a proof given as a sequence of sentences one can easily construct a Turing-machine producing that sequence as the result of some finite computation and, conversely, given a machine computing a proof we can just point all sentences produced during the computation and arrange them into a sequence.[35]

Let *Alg* be this algorithmic completeness of *PA*:

Alg. $PA \vdash R$ iff *R* is T^1-computable on any *n*.

As we know, Gödel's sentence *G*: $\forall \mathrm{y} \neg \beta(\boldsymbol{k}, \mathrm{y})$ is a sentence of the form $\forall$xR(x) such that if *PA* is consistent, then

[33] Comp. sect. 1.2.
[34] Gödel refers to this formula by using its Gödel number *q*.
[35] C. Calude; E. Calude, S. Marcus [2001], 13.

1. $PA \nvdash \forall x R(x)$.

But, as Göel showed, all the sentences R($\boldsymbol{n}$) are provable in *PA*, that is

2. $PA \vdash R(\boldsymbol{n})$, for every natural number *n*.[36]

And by definitions 1^*-4^* above we have

Def*. $\forall x R(x)$ *is true in Constr iff* R(x) *is true in Constr iff the arithmetical relation R*(*x*) *is* T^1*-computable on any n.*

We add a fundamental fact about recursiveness:

Eq. *The class of recursive functions and the class of T-computable functions are coextensive*.[37]

And, correspondingly, *R* is a recursive relation iff C_R is recursive iff C_R is *T*-computable also holds.

Now, by considering Alg. 1, 2, Def* and Eq the following hold.

a) A universal quantified sentence $\forall x R(x)$ is *constructively* true iff for any *n* a constructive proof of R($\boldsymbol{n}$) can be accomplished. So the Gödel sentence is constructively true, by Def* and Alg.

We have to note that this result, regarding Gödel's sentence, does not contrast with the non-constructive definition of a sentence $\forall x R(x)$ (given in Definition 4). For the Gödel sentence *G* this non-constructive meaning of "$\forall$" also holds. Actually someone who endorses Def* will also endorse Definition 4 (but not reverse!). That this is the case results from the fact that under the assumption of *PA*-consistency all sentences R($\boldsymbol{n}$), being provable, are true. For let us suppose that there is an *m* such that R($\boldsymbol{m}$) though provable is false. In our notation the formula R($\boldsymbol{m}$) is just the formula $\neg\beta(\boldsymbol{k},\boldsymbol{m})$, whose intuitive meaning is "*m* is not a proof of *G*". Being false, $\beta(\boldsymbol{k},\boldsymbol{m})$ is true, that means that "*m* is a proof of *G*". So there exists a proof in *PA* of the sentence *G* (and, consequently, of the consistency of *PA*, by Gödel's second theorem).

---

[36] Comp. sect. 1.2, Proof 2 of Gödel's first incompleteness theorem.

[37] More generally, the equivalence holds between the following notions: "λ-definability" and "general recursiveness" (proved in Kleene [1936] and Church [1936]), "λ-definability" and "Turing computability" (proved in Turing [1937]).

b) G is not provable in PA, by Th. Gödel 1. So G, equivalently R(x), and therefore $R(x)$ is not T^1-computable for any value n of x (by Alg). This is due to the fact that $R(x)$ is a Halting-type relation.

Indeed, as we said above, R(x) is the formula $\neg\beta(\boldsymbol{k},y)$, whose intuitive meaning is "y is not a proof of the diagonalization of k".[38] Let $R^*(k,y)$ be this relation.[39] This is a recursive relation and, by Eq, it is T-computable.

But $R^*(k,y)$ is circular, for the diagonalization of the formula whose Gödel number is k is just the formula G: $\forall y\neg\beta(\boldsymbol{k},y)$. So the Gödel number of G is one of the values of y in $R^*(k,y)$. Hence the domain of the relation $R^*(k,y)$ does include the Gödel number of the formula G.[40]

c) *G is constructively true*. This fact results just by the Gödel's argument used in proof of his theorem. For by first incompleteness theorem G is not provable. So for any n, n is not a proof in PA of the sentence G. And that means that $PA\vdash\neg\beta(\boldsymbol{k},\boldsymbol{0})$, $PA\vdash\neg\beta(\boldsymbol{k},\boldsymbol{1})$, $PA\vdash\neg\beta(\boldsymbol{k},\boldsymbol{2})$,... For a given n, let m_n be the Gödel number of the formula $\neg\beta(\boldsymbol{k},\boldsymbol{n})$. By definition, if $PA\vdash\neg\beta(\boldsymbol{k},\boldsymbol{n})$ then there is a proof of $\neg\beta(\boldsymbol{k},\boldsymbol{n})$ whose Gödel number is d_n. So the following relation holds: $d_n\beta m_n$, where xBy is a primitive recursive relation "x is a proof of y".[41] And this holds for any n. Hence $d_n\beta m_n$ is T^1-computable (by Eq). And this means that G is true in *Constr* (by Def*).

What b) and c) show is the following fact: "True in Constr" and "T^1-computability" do not coincide.

d) Being constructively true, for Gödel's sentence G *tertium non datur* holds. So $\neg G$ is a false sentence. But this also holds for G in the

---

[38] I.e. it is the complement of our relation $R(k,y)$ in sect. 1.2.

[39] In Gödel's notation it is $Q(x,p)$, defined as $\overline{x\beta\delta(p)}$ and expressed in PA by the formula q(x,***p***).

[40] Therefore R^* implicitly refers to itself, a fact resulting just from its definition. In terms of T-computability, this means that there is an n such that the Turing machine computing R^* does not halt.

[41] In our notation it is the relation "*Pf*" (cf. sect. 1.2).

standard model M. For as we have seen, the falsity of G would mean the truth of its negation, $\exists y B(\boldsymbol{k},y)$, i.e. G would be provable in PA, contradicting Gödel's first incompleteness theorem.

Conclusion

Gödel's construction of an undecidable sentence G is based on the Cantor's idea of diagonalization. Essentially, this idea is the binder between Gödel's results and paradoxical construction. As we showed in section 1 by using the diagonalization Richard constructed the respective paradox, from which the Liar paradox can easily be derived. Finally, by replacing the semantic notion "true" in the latter with the syntactic notion "provable" the Gödel sentence G, asserting its own unprovability, follows.

By Gödel's theorems there is an asymmetry between the syntax of a language and its semantics, respective the truth of a sentence is not always identical to its proof. In Gödel's terms if PA is ω-consistent then G is not provable in PA. But G is true in the standard model M of PA. How do we decide that G is true? The answer is: by interpreting it intuitively. As we saw, the meaning of the sentence G: $\forall y \neg \beta(\boldsymbol{k},y)$, is that *it* is not provable in PA, and by Gödel' first theorem, it is proved that G is not provable in PA (if PA is consistent). So for G we have: G is true iff G is not provable.

The existence of such a sentence G is rejected by Wittgenstein in his *Remarks on the Foundations of Mathematics*. Wittgenstein's argument is based on his finitism (constructivism), according to which a mathematical sentence is meaningful only within a given calculus and if it is effectively decidable in this calculus. So a *true* sentence is just a *proved* sentence. By transcending the given calculus (or system) and by inserting the same sentence in other calculus its meaning have changed. So a sentence cannot be true but unprovable in a given calculus. In other words, Gödel's sentence G, true and unprovable in PA, is just a contradiction in terms.

The source of Wittgenstein's rejection of G is his rejection of the meta-mathematics (i.e. of the distinction object language / meta-language), by his identification of the arithmetical truth with the formal derivability, in connection to the thesis that a sentence is meaningful only in the context of its use.

A particular note of Wittgenstein's stance is brought in the Appendix of his *Remarks*, § 8, by saying that Gödel's sentence *G* does not admit the interpretation in English language as "*G* is unprovable". As we showed this mistaken interpretation of Gödel's theorem is also based on the mistaken elimination of the distinction syntax / semantics, a distinction imposed by Gödel's result: if *PA* is consistent then *G* is true iff *G* is not provable, so "true" and "provable" do *not* mean the same thing. Hence, Wittgenstein's rejection of the interpretation of *G* as "*G* is unprovable" is not rejected by Gödel's result, but by Wittgenstein's identification of "true" with "provable", a case in which this interpretation of *G* is not compatible with the assertion of provability of *G*. But the setting up of the distinction language / meta-language and *eo ipso* of the difference between truth and proof, the compatibility of the interpretation of *G* as "*G* is umprovable" is completely established.

As we mentioned in 2.1, on Wittgenstein's view every sentence of mathematics must belong to a calculus and then it is either provable or refutable. So the *tertium non datur* holds for all mathematical sentences. But what can be said if such a sentence in *G*?

According to Goodstein[42] Wittgenstein overlooks the possibility of arguing the truth of the sentence *G* by using *tertium non datur*.

> We may simply appeal to the *tertium non datur* to assert that *one* of $(\forall x)G(x), \exists x \neg G(x)$ is true, and since neither of these sentences is provable (in Gödel's version of *Principia Mathematica*) one of the sentences is both true and unprovable (without committing ourselves to saying which one is true).[43]

Concerning this passage, the following remarks can be made. Firstly, this line of arguing the truth of *G* cannot be taken as an omission in Wittgenstein's view, for according to Wittgenstein, the sentence *G*, being not provable in *PA*, is not completely meaningful. So by what is said above the *tertium* does not apply to it. Secondly, Goodstein's fragment does

---

[42] R.I. Goodstein [1972].
[43] 279-280.

contain an error. For though G and $\neg G$ are not provable in PA, we *can* say that G is true, for, as we showed in 1.2, if G were false, then its negation $\exists y\beta(\boldsymbol{k},y)$ will be true and that means that G is provable in PA, contradicting Gödel's first incompleteness theorem.

Finally, in section 3, we question the idea whether there is a sense in which the truth can *constructively* be defined. More exactly, based on fundamental results of recursion theory and computability, we showed that the truth of a PA-formula can be defined in terms of Turing-computability. But in this interpretation (i.e. in a constructive model) the sentence G is constructively true but not T^1-computable.

References

Ambrose, A. and M. Lazerowitz (eds.), [1972], *Ludwig Wittgenstein: Philosophy and Language*, Allen and Unwin, London.

Benacerraf, P. and H. Putnam, [1964], *Philosophy of Mathematics*, Prentice Hall, Inc.

Calude, C., E. Calude, S. Marcus [2001], *Passages of Proof*, Workshop *Truths and Proofs*, Annual Conference of the Australasian Association of Philosophy, Auckland.

Cantor, G., [1891], "Über eine elementare Frage der Manningfaltigkeitslehre", *Jahresbericht der Deutschen Math. Vereinigung*, vol. I, 75-78.

Church, A., [1936], "An unsolvable problem of elementary number theory", *American Journal of Mathematics*, 58, 345-363.

Goodstein, R.L., [1972], "Wittgenstein's Philosophy of Mathematics", in A. Ambrose and M. Lazerowitz (eds.) [1972], 271-286.

Gödel, K., [1931], "Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I", *Monatshefte für Mathematik und Physik*, 38, 173-198; Engl. transl. in *K. Gödel. Collected Works*, vol. I (ed. Solomon Feferman), Oxford UP, 1986, 144-195.