# beyond formalism: Gödel's incompleteness

| | |
|---|---|
| Canonical name | BeyondFormalismGodelsIncompleteness |
| Date of creation | 2013-03-22 18:26:04 |
| Last modified on | 2013-03-22 18:26:04 |
| Owner | gribskoff (21395) |
| Last modified by | gribskoff (21395) |
| Numerical id | 28 |
| Author | gribskoff (21395) |
| Entry type | Topic |
| Classification | msc 03B10 |
| Classification | msc 03A05 |
| Classification | msc 03-01 |
| Synonym | Gödel's Theorem |
| Synonym | Incompleteness Theorem |
| Synonym | Tarski's theorem |
| Related topic | FOUNDATIONSOFMATHEMATICSOVERVIEW |
| Related topic | AnOutlineOfHilbertsProgramme |
| Related topic | GodelNumbering |
| Related topic | GodelsIncompletenessTheorems |
| Related topic | FormalLogicsAndMetaMathematics |
| Related topic | AlgebraicCategoryOfLMnLogicAlgebras |
| Related topic | FromHilbertsTenthProblemToGodelsTrichotomy |
| Related topic | GodelsBetaFunction |
| Related topic | Mathematical |
| Defines | Incompleteness |
| Defines | Omega-Consistency |
| Defines | Arithmetic truth |

# Contents

# 1 Overview

In the literature on the foundations of mathematics the term "formalism" occurs with three different configurations.

The first (chronologically) was refuted by Frege in both *Grundagen der Arithmetik*, on the issue of the existence of the complex number system, and in *Grundgesetze der Arithmetik* (Sec. 86 *et seq.*). In its oldest configuration the formalist doctrine is to be found in the writings of Frege's contemporaries Heine, Thomae and von Helmholz and it is essentially composed out of two fundamental thesis.

According to the first thesis mathematical propositions are merely sequences of symbols, the interpretation of which may be helpful but not relevant. Thus mathematical propositions can be said to have form but not content, since this is only given by means of an interpretation. In today's terminology mathematics would then consist of a language with a fixed syntax but without semantics.

To this thesis is usually associated the well known metaphor that mathematics is a game like a board or a card game, for which rules to move the pieces or to play the cards are provided, without having to stipulate the meaning that the game will have to yield. A paradigm case is a game of chess, as described in the usual notation.

The second thesis of this first configuration is the equally repeated ontological doctrine that the existence of a mathematical object is secured by the consistency proof of the theory in which the object occurs. Under these circumstances we can secure the existence of everything that does not lead to a contradiction and the associated slogan is that the criterion of existence is consistency.

In a second configuration the term "formalism" is frequently used to refer to the set of proposals known as Hilbert's Programme. It is actually a terminological inaptitude, since Hilbert was not a formalist in the first sense. About Hilbert's proposals as to how to preserve the meaning of mathematical propositions and his hopes of a consistency proof for arithmetic the reader can use the PM entry "An Outline of Hilbert's Programme" where the issue is developed with some detail.

Modern formalism is the object of a very successful refutation by Georg

Kreisel, who coined the complex term "formalist-positivist doctrine". The formalist-positivist doctrine set foot in the philosophy of mathematics after the successful formalization of some mathematical theories, i.e., theories with a positive solution to the decision problem. The formalist-positivist doctrine rejects the validity of our knowledge of abstract objects. These are, for this theory, no more than purely verbal extrapolations of the true knowledge of concrete objects or facts.

The main *fazit* of the formalist-positivist doctrine is the elimination of the use of abstract concepts, defined as being of second or higher order, and their substitution by concepts that can be subject to the control of a formal system. In this sense formalist mathematics is reduced to concepts, for the understanding of which it is sufficient to have a list of formal rules which describes them completely.

In epistemology the formalist-positivist doctrine maintains that formal rules are not only qualitatively different from the abstract concepts used in classical mathematics, but also that the knowledge provided by them has a degree of reliability higher than the knowledge obtained by the use of abstract objects or by the use of our intuition of mathematical reality. It is argued that the use of our intuition caused the paradoxes and the foundational crisis of the late XIX century, and that this fact is evidence against the reliability of our intuition and therefore favours the need to subject such use to the control of formal systems.

On the greater reliability of formal rules and mechanical control Kreisel argues that it is a fact of experience, acknowledge even by Bourbaki, that this mechanical or formal control is hardly ever carried out in full. From this it follows that this increased reliability is no more that a matter of faith, because if the formalization is not fully carried out and the mechanical control is not verified then on what grounds are we supposed to trust their reassuring value?

It remains the so called historical fact that the paradoxes prove the unreliability of our intuition and of the use of abstract concepts. It is also a fact of experience that the paradoxes do not affect our confidence in the use of our intuition more than debugging a program affects our confidence in mechanical computation. Our intuition of mathematical reality has an homological image in our perception of physical reality. The paradoxes shake our confidence in that intuition just as much as errors of perception shake our confidence in our knowledge of the external world.

In its third configuration formalism articulates again the issues of Frege's

contemporaries (using the benefit of hindsight) and is associated with the names of Rudolph Carnap, Haskell Curry and more recently S. Mac Lane.

The position is better described as syntacticist and Gödel refuted it in his "Gibb's Lecture" and in "Is mathematics syntax of language?". The argument is based on his theorems to be discussed bellow and can be essentially reproduced as follows: I can not propose a formal system for a mathematical theory and claim with consistency that I can with mathematical rigor have the insight that $i$) the axioms of the system and the rules of inference are correct and $ii$) they express all true statements of the theory. Because if I can see that the axioms and the rules are correct then I can also see that the axiom system is consistent. But the consistency of the axioms is not provable within the system (Theorem 2 bellow) and so I claim to have an insight into something that can not be proved in the system. But that also makes it impossible to claim that the system contains all true statements of the theory (Theorem 1 bellow).

## 2   Arithmetic

The object of study of arithmetic is not only the set of natural numbers but also other sets of objects that one can define categorically, like the integers or the rational numbers, so that a theory about one of these sets of objects is also called an arithmetic. In general the objects studied are taken to be individuals, in the sense that they can not be further analyzed as being composed out of other objects. A departure from this restriction is sometimes tolerated, as it is the case when the basic properties of the positive rational numbers are explained using a representation of them in terms of pairs of natural numbers.

In contrast with the term "number theory", the word "arithmetic" is also used to denote the study of the properties of some particular functions over the integers like sum, multiplication and related concepts.

An obvious extension of this vocabulary takes place when one speaks of arithmetic to denote the sum of non-denumerable sets, in contexts like "the arithmetic of transfinite cardinals".

# 3 From the number sequence to formal arithmetic

In spite of the fact that the philosophical reflection on the concept of number is as old as philosophy itself, the scientific treatment of this reflection only took place at the dawn of the XX century, with the work of Dedekind (1831-1916) and Frege (1848-1925). It is important to notice that the new scientific outlook raised, rather than lowered, the philosophical significance of the whole subject, as we can see in the abundant literature on the theorems of Löwenheim, Gödel and Tarski, thereby refuting the positivistic claim that progress in science makes philosophical reflection obsolete.

The first modern characterization of the concept of number was formulated by Dedekind in 1901 by characterizing the number sequence with the following axioms:

1. 0 is a number.

2. For every number $x$ there is a unary function $f$ defined by $f(x) = x+1$, the value of which is the number that immediately follows $x$ in the sequence, called the successor of $x$.

3. 0 is not a successor.

4. The function $f(x)$ is 1-1.

5. If $P$ is an arithmetical property such that $P(0)$ and for all $x$ $P(x)$ implies $P(x + 1)$, then for all $x$, $P(x)$.

Dedekind's propositions together with the theory of sets allow the derivation not only of the theory of natural numbers but also of other numbers systems, rational, real, and complex.

But these propositions can not be considered as an axiomatic system, in the sense of a formal axiomatic system, because it contains informal concepts like "property".

Formal arithmetic is the syntactical version of Dedekind's axioms and we will be using a simplified version of the treatment proposed by Hilbert and Bernays in *Grundlagen der Mathematik*.

# 4    The formal theory Z

The formal theory to be used is a first order theory with equality and it will be called **Z** (the first letter of the german word *Zahl*, number). It has a denumerable number of individual variables, one binary predicative symbol $I(a,b)$ which will be abbreviate to $a = b$. There are three functional symbols, unary $f$, binary $g$ and binary $h$. They will be abbreviated to $f(a)$, $a + b$ and $a \cdot b$. The only individual constant is Dedekind's initial object 0.

The proper axioms of **Z** are the following:

- **Z1**: $(a = b) \rightarrow [(a = c) \rightarrow (b = c)]$

- **Z2**: $(a = b) \rightarrow [f(a) = f(b)]$

- **Z3**: $(\forall x)\neg[0 = f(x)]$

- **Z4**: $[f(a) = f(b)] \rightarrow (a = b)$

- **Z5**: $a + 0 = a$

- **Z6**: $a + f(b) = f(a + b)$

- **Z7**: $a \cdot 0 = 0$

- **Z8**: $a \cdot f(b) = (a \cdot b) + a$

- **Z9**: If $\Phi(x)$ is a well formed formula of Z then:

$$\Phi(0) \rightarrow \{[(\forall x)[\Phi(x) \rightarrow \Phi(f(x))] \rightarrow (\forall x)\Phi(x)\}.$$

Axioms **Z1** and **Z2** state the properties of equality between objects and their successors. **Z3** and **Z4** correspond to Dedekind's axioms 3 and 4. Dedekind's axioms 1 and 2 are represented by the individual constant and the unary function of **Z**.

It is to be noticed that **Z9** differs from the other axioms **Z1**, ..., **Z8**, since it is an axiomatic scheme. It does not correspond exactly to Dedekinds axiom 5, the principle of mathematical induction, since this refers to a non-denumerable number of properties of the number sequence. Instead **Z9** refers only to a denumerable number of properties, precisely those properties that can be defined in **Z** by means of well-formed formulas of **Z**.

To prove by induction in **Z** that $(\forall x)\ \Phi(x)$ one uses the premises $\Phi(0)$ and $(\forall x)[\Phi(x) \to \Phi(f(x))]$. Using **Z9** as starting formula two applications of *modus ponens* give $(\forall x)\Phi(x)$.

Axioms **Z5** to **Z8** allow the proof in **Z** of all known results of addition, product and divisibility of numbers, the existence and the uniqueness of the quotient and rest.

Order can be defined in **Z** and thereby the principle of complete induction, by saying that $t$ is less than $s$ if and only if there is a number $m$ different from 0 such that $t + m = s$.

The underlying logic allows the derivation of the usual results about order. If we call **Z9** the rule of induction we can use it to prove the principle of complete induction:

If $P$ is a property such $P(x)$ and for every $x$ if $P$ is satisfied by all numbers
less than $x$ then $P(x)$,
then for every $x$ $P(x)$.

The rule of induction allows the demonstration of the principle of complete induction and the least number principle as theorems of **Z**.

# 5    A model for Z

It is now obvious that a model $M$ for **Z** has to satisfy the following conditions:

1. The domain of the model is the set of natural numbers.

2. The natural number 0 is the interpretation of the individual constant of **Z**.

3. The unary function $f(x)$ is interpreted as "the successor of $x$".

4. The binary function $g(x, y)$ is the addition of $x$ and $y$.

5. The binary function $h(x, y)$ is the product of $x$ and $y$.

6. The only predicative letter is equality.

The terms of **Z** are the initial object 0 and the result of prefixing and reiterating $f$. We shall call these terms numerals and represent them by:

$$0,\ f(0),\ f(f(0)),\ \ldots$$

We allow the typographical convention of representing $f(0)$ by $1^*$ and in general if $n$ is a natural number $n^*$ is the corresponding numeral.

This interpretation is known as the standard model for **Z**. A model that is not isomorphic to $M$ is called non-standard.

If one accepts $M$ as a model for **Z**, then from the semantic point of view **Z** is consistent. To see this one verifies that the axioms of **Z** are true in $M$ and that the rules of inference preserve truth. Therefore the theorems of **Z** are also true in $M$.

As a student of Hilbert's Programme Gödel was interested in the problem of proving the consistency of a theory like **Z**, using only concepts definable in **Z**. He found a negative solution to this problem and in the same work he also proved the existence of propositions true in $M$ but not provable in **Z**, thereby showing, contrary to Hilbert's expectations, that the concepts of truth and formal proof are not equivalent. We turn now to that work.

# 6   Gödel numbers

We begin by defining a 1-1 function $g$, whose domain is the union of the set of all symbols, expressions and finite sequences of expressions of **Z** and whose range is a subset of the set of positive integers. We want $g$ to satisfy the following clauses:

1. $g$ is computable.

2. There is a procedure that given a integer $m$ it can be verified if $m$ is in the range of $g$.

3. If $m$ is an integer then the procedure finds the object $x$ such that $m$ is the image of $x$ by $g$.

The values of $g$ for the symbols of **Z** are the following:

- $g(0) = 1$

- $g(f) = 3$

- $g(+) = 5$

- $g(\cdot) = 7$

- $g(=) = 9$

- $g(() = 11$

- $g()) = 13$

- $g(,) = 15$

- $g(\neg) = 17$

- $g(\forall) = 19$

- $g(\rightarrow) = 21$

- $g(x_k) = 15 + 8 \cdot k$

- $g(a_k) = 17 + 8 \cdot k$

These values are called the Gödel numbers of the symbols.

Since $g$ is 1-1 different symbols have different Gödel numbers and every Gödel number of a symbol is an odd positive integer.

An expression of **Z** is a sequence of symbols of **Z**, $s_1$, ..., $s_n$, and we can define a unique representation of $g(s_1, \ldots, s_n)$ by defining:

$$g(s_1, \ldots, s_n) = 2^{g(s_1)} \cdot 3^{g(s_2)} \cdot \ldots \cdot p_{n-1}^{g(s_n)}$$

where $p_i$ is the ith prime number and $p_0 = 2$.

Thus different expressions of **Z** have different Gödel numbers. Expressions and symbols have different Gödel numbers since the Gödel number of an expression is an even positive integer.

A proof in **Z** is a sequence of expressions of **Z** and we can define a unique representation of $g(e_1, \ldots, e_n)$ by defining:

$$g(e_1, \ldots, e_n) = 2^{g(e_1)} \cdot 3^{g(e_2)} \cdot \ldots \cdot p_{n-1}^{g(e_n)}$$

where $p_i$ is the ith prime number and $p_0 = 2$.

Different sequences of expressions and thus different proofs in **Z** have different Gödel numbers.

Gödel numbering is the tool that will allow our formal theory to refer to its own symbols, expressions and proofs, thereby expressing the metamathematics of the theory inside the theory.

# 7 Expressing relations and representing functions in Z

It would be completely surprising if we could give an account of Gödel's theorems without a certain amount of terminology and conceptual complexity. In this sense we begin by introducing a metamathematical binary predicate $P(y, x)$ the meaning of which can be expressed by the sentence:

"$y$ is the Gödel number of a proof of a formula whose Gödel number is $x$."

In particular in our theory **Z** this predicate can take either:

- i) The form $P^+(g, y)$ with the following interpretation:

  "$g$ is the Gödel number of a well formed formula $\Phi(x_1)$ where the occurrence of $x_1$ is free and $y$ is the Gödel number of a proof of $\Phi(g*)$";

  or else it can take the form:

- ii) $P^\neg(g, y)$ with the following interpretation:

  "$g$ is the Gödel number of a well formed formula $\Phi(x_1)$ where the occurrence of $x_1$ is free and $y$ is the Gödel number of a proof of the formula $\neg\Phi(g^*)$."

It will be useful for future reference to discuss now the status of the formulas of **Z** and for that purpose we will separate relations from functions.

If $R(x_1, \ldots, x_n)$ is an arithmetical relation we say that it is expressible in **Z** if and only if there is a well formed formula $\Phi$ of **Z** with $n$ free variables and such that for any $n$-tuple of natural numbers $k_1, \ldots, k_n$ the following two conditions are satisfied:

1. If $R(k_1, \ldots, k_n)$ is true then there is a proof in **Z** of $\Phi(k_1^*, \ldots, k_n^*)$.

2. If $R$ is false then there is a proof in **Z** of $\neg\Phi(k_1^*, \ldots, k_n^*)$.

In the case of an arithmetical function $f(x_1, \ldots, x_n)$ one speaks of its representation in **Z**. Thus an arithmetical function of $n$ arguments is representable in **Z** if and only if there is a well formed formula of **Z** with $n + 1$ free variables such that for any $k_{n+1}$-tuple of natural numbers the following two conditions are satisfied:

1. If $f(k_1, \ldots, k_n) = k_{n+1}$ then there is in **Z** a proof of $\Phi(k_1^*, \ldots, k_{n+1}^*)$.

2. There is in **Z** a proof $(\exists! \; x_{n+1}) \; \Phi(k_1^*, \ldots, k_n^*, x_{n+1})$ where $\exists!$ means uniqueness.

The relations between the concepts of expression, representation and the formal theory **Z** are regulated by two main theorems which we will use later without proving them now:

**Theorem 1.: An arithmetical relation is expressible in Z if and only if it is recursive**.

**Theorem 2.: An arithmetical function is representable in Z if and only if it is recursive**.

# 8 Omega-consistency and metamathematical predicates

We turn now to the concept of $\Omega$-consistency which plays an essential role in both proofs.

The concept was first discovered by Tarski and we begin by defining a theory as being $\Omega$-inconsistent if and only if we have a well formed formula $\Phi(x)$ such that we have for every natural number $n$ a proof in **Z** of $\Phi(n^*)$ and a proof in **Z** of $(\exists x)\neg\Phi(x)$. If on the contrary it is not possible to derive in **Z** for every $n$ a proof of $\Phi(n^*)$ and a proof of $(\exists x)\neg\Phi(x)$ then **Z** is $\Omega$-consistent.

A simple argument shows that if **Z** is $\Omega$-consistent then it is also simply consistent. To see this one allows the formula $\Phi(x)$ to be the well formed formula of **Z**:

$$(\forall x)(x = x) \rightarrow (x = x).$$

Of course one has in **Z** for every $n$ a proof of:

$$(n^* = n^*) \rightarrow (n^* = n^*).$$

Therefore there is no proof in **Z** of:

$$(\exists x)\neg[(x = x) \rightarrow (x = x)].$$

Therefore **Z** is simply consistent.

If we take the semantic point of view we also see that if **Z** is interpreted in the standard model it turns out $\Omega$-consistent.

We are now going to take advantage of the use of Gödel numbers to formulate in **Z** metamathematical predicates.

Paradigm cases are the predicates "provable" and "refutable".

"Provable" can be formulated as:

"there is a number $y$ such that $y$ is the Gödel number of a proof of a formula with the Gödel number $m$."

"Refutable" can be formulated as:

"there is a number $y$ such that $y$ is the Gödel number of a proof of the negation of a formula with the Gödel number $m$."

# 9  How to build Gödel's undecidable statement

In order to achieve a reasonably clear view we break down the construction process in several steps.

1. Let $\Phi(x_1)$ be a well formed formula of **Z**, with a free occurrence of $x_1$, and let $g$ be the Gödel number of $\Phi(x_1)$.

2. From this formula we can obtain by insertion for $x_1$ the formula $\Phi(g^*)$. Let $y$ the Gödel number of the formula $\Phi(g^*)$.

3. We can now form the predicate $D^+(g, y)$, already mentioned, which is a recursive relation and therefore expressible in **Z** by a well formed formula $\Delta(x_1, x_2)$ where both variables have free occurrences.

4. By the definition of expression of a relation we have that if the relation is true and therefore $D(k_1, k_2)$ is true, then we have a proof in **Z** of $\Delta(k_1^*, k_2^*)$.

5. If the relation is false and therefore $\neg D(k_1, k_2)$ we have in $\mathbf{Z}$ a proof of $\neg\Delta(k_1^*, k_2^*)$.

6. If we consider now the case in which the relation is false and we have therefore in $\mathbf{Z}$ a proof of $\neg\Delta(k_1^*, k_2^*)$ we can, by quantifier-introduction, build the formula:

$$(\Theta) \qquad (\forall x_2)\neg\Delta(x_1, x_2)$$

   where $x_1$ still has a free occurrence.

7. Let $m$ be the Gödel number of the formula:

$$(\Theta) \qquad (\forall x_2)\neg\Delta(x_1, x_2).$$

8. We interpret this formula as saying that no number $x_2$ is the Gödel number of a proof of the formula with the Gödel number $x_1$. And this is the same as saying that the formula with the Gödel number $x_1$ is unprovable.

9. Since $x_1$ has a free occurrence we can insert for $x_1$ the numeral for the Gödel number of the formula $(\Theta)$ and obtain the closed well formed formula:

$$(\heartsuit) \qquad (\forall x_2)\neg\Delta(m^*, x_2).$$

10. By 1-3 above $D^+(g, y)$ is satisfied if and only if $g$ is the Gödel number of a well formed formula $\Phi(x_1)$ and $y$ the Gödel number of a proof of $\Phi(g^*)$.

11. Since $(\heartsuit)$ comes from $(\Theta)$ by substitution of $m^*$ for $x_1$ we are led to the following conclusion:

$$(*) \qquad D^+(m, y)$$

    is satisfied if and only if $y$ is the Gödel number of a proof of $(\heartsuit)$.

## 10   Gödel's first theorem

The statement of the theorem is the following:

"If **Z** is consistent then the formula ($\heartsuit$) is not provable in **Z** and if **Z** is

$\Omega$-consistent

then the formula $\neg(\heartsuit)$ is not provable in **Z**."

We split the proof in two halves, one using the consistency assumption and the other the $\Omega$-consistency. We prove both by reductio.

**Part I**.

We assume consistency and the reductio hypothesis is that we have a proof in **Z** of:

$$(\heartsuit) \qquad\qquad (\forall x_2)\neg\Delta(m^*, x_2).$$

Let $k$ be the Gödel number of a proof in **Z** of ($\heartsuit$). Then by (*) we have $D^+(m, k)$. As $\Delta$ expresses $D^+$ in **Z** one has a proof in **Z** of $\Delta(m^*, k^*)$.

But $(\forall x_2)\neg\Delta(m^*, x_2) \to \neg\Delta(m^*, k^*)$. This formula and the reduction hypothesis give us a proof in **Z** of $\neg\Delta(m^*, k^*)$. This proves the inconsistency of **Z** contrary to hypothesis of the theorem. Therefore if **Z** is consistent ($\heartsuit$) does not have a proof in **Z**.

**Part II**.

We assume the $\Omega$-consistency of **Z** and the reductio hypothesis is that we have in **Z** a proof of:

$$\neg(\forall x_2)\neg\Delta(m^*, x_2).$$

Now since **Z** is $\Omega$-consistent it is also simply consistent. Therefore we do not have in **Z** a proof of ($\heartsuit$). This means that for every $n$, $n$ is not the Gödel number of a proof in **Z** of ($\heartsuit$). Thus by (*) $\qquad (\forall n)D^+(m, n)$ is false. This implies a proof in **Z** of $\neg\Delta(m^*, n^*)$.

Now in the definition of $\Omega$-consistency we let $\Phi(x)$ be the formula $\neg\Delta(m^*, x_2)$. Then there is no proof in **Z** of $(\exists x_2)\neg\neg\Delta(m^*, x_2)$ and by double negation there is no proof in **Z** of $(\exists x_2)\Delta(m^*, x_2)$.

By the reductio hypothesis there is in **Z** a proof of $\neg(\forall x_2)\neg\Delta(m^*, x_2)$ and this formula is by predicate calculus equivalent to $(\exists x_2)\Delta(m^*, x_2)$. Therefore we have in **Z** a proof of $(\exists x_2)\Delta(m^*, x_2)$. This proves the inconsistency of **Z** contrary to the hypothesis. Thus if **Z** is $\Omega$-consistent the formula $\neg(\heartsuit)$ is not provable in **Z**.

We have proven that neither the formula ($\heartsuit$) nor the formula $\neg(\heartsuit)$ is provable in **Z**. A formula such as ($\heartsuit$) is called an undecidable formula.

# 11 The interpretation of the undecidable formula

We have seen that the predicate $\Delta$ expresses the relation $D^+$ in **Z**. Thus the interpretation of the undecidable proposition ($\heartsuit$) in the standard model results in the assertion that $D^+(m, x_2)$ is false for every natural number $x_2$.

But by proposition (*) this means that in **Z** there is no proof of the closed well formed formula ($\heartsuit$). Therefore the formula $(\forall x_2)\neg\Delta(m^*, x_2)$ asserts its own unprovability.

Now since we assume **Z** to be consistent and the theorem proves that there is no proof in **Z** of the undecidable formula, then the undecidable formula is true and unprovable in **Z**. This shows that the set of proofs of **Z** does not contain all propositions true in the standard model.

If we define a formal theory as complete if for every well formed formula $\Phi$ either there is a proof of $\Phi$ or a proof of $\neg\Phi$, then the formal theory **Z** is incomplete and the theorem proves that incompleteness.

It would seem *prima facie* that the reason for this incompleteness lies in the fact that **Z** does not have enough axioms to allow the derivation of the undecidable proposition. Of course we can add to **Z** the undecidable proposition and obtain a theory $\mathbf{Z}^+$ and since the undecidable propostion is true it would be derivable in $\mathbf{Z}^+$.

But all recursive functions that are representable in **Z** are representable in $\mathbf{Z}^+$, just like all predicates, relations and metamathematical functions. These will also be recursive in $\mathbf{Z}^+$.

But if $\mathbf{Z}^+$ is $\Omega$-consistent it will also contain an undecidable proposition which will differ from ($\heartsuit$). And so the undecidability of ($\heartsuit$), the incompleteness of the formal theory **Z** and the impossibility of finding in **Z** an equivalent to the concept of truth do not depend on the number of axioms of **Z**.

# 12 Gödel's second theorem

For the significance of the consistency problem within the framework of Hilbert's Programme it can be useful to read the PM encyclopaedia entry "An Outline of Hilbert's Programme".

We have seen that in the first theorem the argument can be expressed by the implication:

If **Z** is consistent then ($\heartsuit$) is unprovable in **Z**.

This means then that if we add to this implication a proof of the consistency of **Z** we get again the unprovability of Gödel's undecidable proposition. We want now to build a sentence that represents the consistency of **Z** and which can be expressed in **Z**.

We begin by introducing the binary predicate $P(y, x)$ interpreted as meaning "$y$ is the Gödel number of a proof of a well formed formula with the Gödel number $x$".

Now this predicate is recursive and thereby is expressible in **Z** by a well formed formula $\Pi(x_1, x_2)$.

Let $x$ be the Gödel number of a well formed formula and Ng $(x)$ the Gödel number of its negation. This function is primitive recursive and therefore representable in **Z** by a well formed formula $\nu(x_1, x_2)$.

Then the well formed formula:

($\spadesuit$)  $\qquad\qquad (\forall x_1)(\forall x_2)(\forall x_3)(\forall x_4)\neg[\Pi(x_1, x_3) \wedge \Pi(x_2, x_4) \wedge \nu(x_3, x_4)].$

expresses the consistency of **Z**.

The statement of the first theorem is obviously equivalent to the formula:

($\diamondsuit$)  $\qquad\qquad$ If {**Z** is consistent} then {($\heartsuit$) is unprovable}.

By using Gödel's numbering the whole proof of the formula ($\diamondsuit$) can be written in **Z**. For { **Z** is consistent } we insert the formula ($\spadesuit$). For { ($\heartsuit$) is unprovable } we insert the formula ($\heartsuit$) itself, since this formula asserts its own unprovability. We are then led to a proof in **Z** of:

($\clubsuit$)  $\qquad\qquad$ ($\spadesuit$) $\rightarrow$ ($\heartsuit$).

The statement of Gödel's second theorem is the following:

If **Z** is consistent then the formula ($\spadesuit$) is not provable in **Z**.

We prove the theorem by *reductio ad absurdum* assuming the consistency of **Z** and the provability of ($\spadesuit$) as the reductio hypothesis.

Now ($\clubsuit$) by we have in **Z** a proof that the consistency statement implies the undecidable proposition. Then using *modus ponens* we can have in **Z** a proof of the undecidable proposition. But this entails the inconsistency of **Z** since by the first theorem we can not have in **Z** a proof of the undecidable proposition. Therefore the reductio hypothesis is false and we can not have in **Z** a proof of the consistency statemnent ($\spadesuit$).

15

# 13    The interpretation of the second theorem

The interpretation of the theorem is that if **Z** is consistent then there is no proof of the consistency of **Z** by means of concepts that can be formalized in **Z**.

The consistency hypothesis used is needed simply because if **Z** is not consistent then any formula can be proved in **Z**, in particular also our formula (♠).

It is interesting to notice that an inconsistent system allows the proof of its own consistency.

Gödel's second theorem can be described as a negative solution to the problem of finding the consistency proof of an axiomatic system by means of concepts more elementary than those formalized in the system.

It was an essential point of Hilbert's Programme that the justification of principles of deduction, like the *tertium non datur* when applied to infinite domains, could only be realized by a consistency proof of the formal system in which such principles could be represented.

However in the consistency proof one could only use principles of deduction more evident than the infinitary machinery, whose elimination should be made possible by the formalization.

Hilbert's conception was that the evident principles were the finitist evident principles, and they were only a part of the whole of classical reasoning.

Therefore for the consistency proof of **Z** the concepts needed were only a part of all the concepts that can be formalized in **Z**.

Gödel's second theorem shows that this goal is unattainable, since the consistency proof is not possible even if one uses all the principles of **Z**, the more and the less evident. Thus *a fortiori* it can not be achieved by using only finitist evident principles.

# 14    The concept of truth

We begin by defining a variant of the already introduced metamathematical function substitution, which we shall call monadic substitution.

**Definition 1**:

> *Let g be the Gödel number of a well formed formula*
> $$\alpha(x_1)$$

*where $x_1$ has a free occurrence. Then the monadic function*
$$S(g)$$
*denotes the Gödel number of the well formed formula*
$$\alpha(g^*),$$
*where $g^\star$ is the numeral which denotes g.*

**Proposition 1**:

*The function $S(g)$ is $\phi$-representable.*

Proof:

We reduce monadic substitution to the already defined case of triadic substitutuion. Thus the reduction formula is:

$$S(g) = \mathrm{Sub}[g, \mathrm{Num}(g), 2^g(x_1)].$$

Now since the triadic function is recursive the monadic function is recursive too. But if it is recursive it is also representable.

This means that the result of the substitution in the formula with Gödel number $g$, for the free occurrence of $x_1$ the numeral for $g$, is still a primitive recursive function which is representable.

We turn now to a 1st order theory with equality $\Omega$, which has the same symbols as the already known theory **Z**. The notation

$$\{T_\Omega\}$$

denotes the set of all Gödel numbers of the $\Omega$-theorems. Thus if $x$ is the Gödel number of a $\Omega$-theorem then this can be denoted by

$$x \in \{T_\Omega\}.$$

We want to prove that the membership relation to $\{T_\Omega\}$ can not be expressed in $\Omega$.

**Proposition 2**:

*If $\Omega$ is consistent and the monadic function $S$ is representable in $\Omega$, then the membership relation to $\{T_\Omega\}$ can not be expressed in $\Omega$.*

Proof:

1. We assume that $S$ is representable in $\Omega$ and that

2. The membership relation to $\{T_\Omega\}$ is expressible in $\Omega$.

3. From these assumptions we get well formed formulas:

$$\sum(x_1, x_2) \text{ and}$$
$$\theta(x_2)$$

   with the following four properties:

   - i)     If $S(k) = j$ then

   $$\vdash_\Omega \sum(k^*, j^*).$$

   - ii)     $\vdash_\Omega (\exists! x_2) \sum(k^*, x_2).$
   - iii)     If $k \in \{T_\Omega\}$ then

   $$\vdash_\Omega \theta(k^*).$$

   - iv)     If $\neg[k \in \{T_\Omega\}]$ then

   $$\vdash_\Omega \neg\theta(k^*).$$

4. Let $\alpha(x_1)$ be the well formed formula:

   $$(\star) \qquad\qquad (\forall x_2)[\sum(x_1, x_2) \to \neg\alpha(x_2)],$$

   with a free occurrence of $x_1$.

5. Let $m$ be the Gödel number of $(\star)$.

6. Then we can build $\alpha(m^*)$ with the following configuration:

   $$(\star\star) \qquad\qquad (\forall x_2)[\sum(m^*, x_2) \to \neg\alpha(x_2)].$$

7. Let $n$ be the Gödel number of $(\star\star)$.

8. This proves that:

$$S(m) = n.$$

**Lemma 1**:

$$\neg[n \in \{T_\Omega\}].$$

Proof:

1. By 8. and by 3.iii) (above) we have:

$$\vdash_\Omega \sum(m^*, n^*).$$

2. Now in $\Omega$ either:

$$\vdash_\Omega \alpha(m^*)$$

or else:

$$\neg \vdash_\Omega (\alpha^*)$$

3. We first assume the second disjunct, then:

$$\neg(n \in \{T_\Omega\}).$$

4. Thus by 3.iv):

$$\vdash_\Omega \neg\alpha(n^*).$$

5. If we assume the first disjunct, then:

$$\vdash_\Omega (\forall x_2)[\sum(m^*, x_2) \rightarrow \neg\alpha(x_2)].$$

6. From 5. by *dictum de omni* we have:

$$\vdash_\Omega \sum(m^*, n^*) \rightarrow \neg\alpha(n^*).$$

19

7. From 1., 6. and *modus ponens* we get:

$$\vdash_\Omega \neg\alpha(n^*).$$

8. From 4.,7. and $\vee$-Elimination we have:

$$\neg\alpha(n^*)$$

and that means that

$$\neg(n \in \{T_\Omega\}).$$

**Lemma 2**:

$$n \in \{T_\Omega\}.$$

Proof:

1. From 1. (Lemma 1) we have in $k$:

$$\vdash_\Omega \sum(m^*, n^*)$$

and by 3.ii) we deduce:

$$\vdash_\Omega \sum(m^*, x_2) \to x_2 = n^*.$$

2. Since from Lemma 1 $\vdash_\Omega \neg\alpha(n^*)$ we deduce that:

$$\vdash_\Omega (x_2 = n^*) \to \neg\alpha(x_2).$$

3. From 1. and 2. we have:

$$\vdash_\Omega \sum(m^*, x_2) \to \neg\alpha(x_2).$$

4. By $\forall$-Introduction:

$$\vdash_\Omega (\forall x_2)[\textstyle\sum(m^*, x_2) \to \neg\alpha(x_2)].$$

5. But that is the formula $\alpha(m^*)$.

6. Since $n$ is the Gödel number of this formula we have:

$$n \in \{T_\Omega\}$$

as desired.

7. This means that by 3.iii):

$$\vdash_\Omega \alpha(n^*).$$

With 7. from Lemma 1 and 7. from Lemma 2 we prove the inconsistency of $\Omega$, contrary to our hypotesis. And that establishes Proposition 2.

**Proposition 3**:

*If $\Omega$ is consistent then the set $\{T_\Omega\}$ is not recursive.*

Proof:

1. The function $S(g)$ is primitive recursive and therefore $\Omega$-representable.

2. But by Proposition 2, the membership relation to $\{T_\Omega\}$ can not be expressed in $\Omega$.

3. This implies that its characteristic function:

$$K_{\{T_\Omega\}}$$

is not $\Omega$-representable. But then the function:

$$K_{\{T_\Omega\}}$$

is not recursive and therefore the set $\{T_\Omega\}$ can not be recursive.

**Definition 2**:

*A 1st order theory $\Omega$ with equality is said to be recursively undecidable if*
*and only if the set*
$$\{T_\Omega\}$$
*is not recursive.*

**Definition 3**:

*A 1st order theory $\Omega$ with equality is said to be recursively and essentially*
*undecidable if and only if*
*every consistent extension of $\Omega$ is recursively undecidable.*

If we now assume Church's Thesis we can then say that recursive undecidability is equivalent to effective undecidability. And this implies that there is no algorithm to decide the theoremhood predicate.

**Proposition 4**:

*If $\mathbf{Z}$ is consistent then it is recursively and essentially undecidable.*

Proof:
Let $\Omega$ be a consistent extension of $\mathbf{Z}$.
Since every recursive function is $\mathbf{Z}$-representable it will also be $\Omega$-representable.
Therefore the membership relation to the set $\{T_\Omega\}$ can not be expressed in $\Omega$ and this entails that the set $\{T_\Omega\}$ is not recursive.
By Definition 2 this proves that $\mathbf{Z}$ is recursively and essentially undecidable.

**Definition 4**:

*Let $\{W\}$ denote the set of Gödel numbers of the well formed formulas that*
*are true in the standard model.*
*The set $\{W\}$ is said to be arithmetic if and only if there is a well formed*
*formula $\alpha(x)$ in $\mathbf{Z}$ such that $\{W\}$ is the set of numbers m for which $\alpha(m^*)$*
*is true in the standard model.*

**Proposition 5**: [**Tarski's Theorem**]

*The set $\{W\}$ is not arithmetic.*

Proof:

1. Let $\Omega$ be an extension of **Z**.

2. We build the axioms of $\Omega$ by using all the well formed formulas that are true in the standard model.

3. Thus $\{T_\Omega\} = \{W\}$.

4. Since the standard model is a model of $\Omega$, $\Omega$ is consistent.

5. Every recursive function is $\Omega$-representable and therefore the membership relation to $\{W\}$ is not expressible in $\Omega$.

6. Now $\{T_\Omega\} = \{W\}$ implies that the membership relation to $\{T_\Omega\}$ is not expressible in $\Omega$.

7. Since a relation is expressible in $\Omega$ if and only if it is the standard interpretation of a well formed formula of **Z**, this shows that there is no well formed formula $\alpha(x)$ in **Z** satisfying the definition 4.

8. This proves that $\{W\}$ is not arithmetic.

# References

[1] Carnap, C., *The logical syntax of language*, London, 1959.

[2] Crossley, J., *What is mathematical logic?*, OUP, London, 1972.

[3] Curry, H., *Outlines of a formalist philosophy of mathematics*, North-Holland, Amsterdam, 1958.

[4] Dummett, M., "The philosophical significance of Gödel's theorem", in *Truth and other enigmas*, Duckworth, 1977.

[5] Gödel, K., *Collected works*, ed. S. Feferman, Oxford, 1987-2003.

[6] Hilbert, D., Bernays, P., *Grundlagen der mathematik*, 2.Auflage, Springer-Verlag, Berlin, 1968.

[7] Kleene, S., *Introduction to metamathematics*, North-Holland, Amsterdam, 1964.

[8] Kreisel, G., "Hilbert's Programme", *Dialectica* 12, 1958.

[9] Kreisel, G., "Die formalistisch-positivistische Doktrin der mathematischen Präzision im Lichte der Erfahrung", *Zentrallblatt für Mathematik und ihre Grenzgebiete*, 196, pgs. 65-122, 1970.

# 15   Q & A

**Reply to Dr. Roger Lipsett**.

Thank you so much for your message. I don't know if the concepts you mention are already available in PM. But at the risk of being redundant I am going to define them.

**Part I**.

A *formal theory* like **Z** is a countable set of distinct symbols and a finite sequence of such symbols is an expression of **Z**. A subset of these expressions are *the well-formed-formulas* of **Z** and there is an algorithm to determine whether a given expression is well-formed.

Thus if $f$ is a symbol for a monadic function of **Z** the expression $f(a,b)$ is not well-formed.

We set aside some wffs of **Z** as being *the axioms* of **Z** and there is also an algorithm to decide whether a wff is an axiom of **Z**. There is a finite set of relations among the wffs of **Z** called *the rules of inference*. For each $R_i$ there is a unique integer $k$ such that for every set of $k$-wffs and each wff $\phi$ one can decide in a finite number of steps whether the given $k$-wffs stand in the relation $R_i$ to $\phi$, in which case one says that $\phi$ is a *direct consequence* of the given wffs by an application of $R_i$.

*A proof in* **Z** is a sequence of wffs $f_1, \ldots, f_n$ such that for each $i$ either $f_i$ is axiom of **Z** or $f_i$ is a direct consequence of some of the previous wffs by the application of some rule of inference. In that case *a theorem of* **Z** is a wff $T$ of **Z** such that there is a proof whose last term is $T$.

Wffs only have *meaning* if an interpretation is provided.

An interpretation $I$ is a non-empty set $D$ called *the domain* of $I$, which is given and a mapping $\theta$ such that:

- **i)** $\theta$ maps a predicate symbol of **Z** to a relation of $D$;

- **ii)** $\theta$ maps each functional symbol of **Z** to an operation in $D$, i.e.,

$$\theta : D^n \to D;$$

- **iii)** $\theta$ maps each individual constant of $\mathbf{Z}$ to a fixed object of $D$.

A sequence of objects in $D$, $a_1, a_2, \ldots$, *satisfies in $I$* a wff $\phi$ if and only if the proposition that results from the insertion of a symbol representing $a_i$ in all free occurrences of $x_i$ in $\phi$ is true in $I$.

Thus $\phi$ *is true under $I$* if and only if every sequence in $D$ satisfies $\phi$ and is false if no sequence satisfies $\phi$.

*An interpretation is a Model* for a set of wffs if and only if every wff is true under the interpretation.

In particular an interpretation of $\mathbf{Z}$ in which:

- **i)** $D = \mathbb{N}$;

- **ii)** $0 \in \mathbb{N}$ is the interpretation of "0";

- **iii)** The successor function is the interpretation of "$f$";

- **iv)** Addition and product in $\mathbb{N}$ are the interpretation of "$+$" and "$\cdot$".

- **v)** Identity in $\mathbb{N}$ is the interpretation of "$=$", is called *the standard model for* $\mathbf{Z}$, *usually denoted by $M$*.

## Part II.
**"Am I to think of Z as the integers and $M$ as a set of formal axioms for Z?"**

The answers is no. Neither $\mathbf{Z}$ is the integers nor $M$ a set of formal axioms for $\mathbf{Z}$. $\mathbf{Z}$ is a formal theory as defined above and one interpretation of the symbols of that theory is $M$, the standard Model. So $M$ is not a set of axioms but you can interpret in $M$ an axiom. To do this one lets the individual variables range over $\mathbb{N}$ and one interprets the other symbols as specified above.

Example: Axiom **Z2**: $(a = b) \rightarrow [f(a) = f(b)]$. The variables stand for natural numbers, the equals sign for identity, $f$ for the successor function and the meaning of **Z2** is then that if two natural numbers are the same they have the same successor, a proposition that is true in $\mathbb{N}$.

## Part III.
**"In addition the truth of the sentence starting" To see this...**

Now if one interprets all 9 axioms of $\mathbf{Z}$ in the standard Model $M$ we only get true propositions (about the natural numbers). The formal theory $\mathbf{Z}$ has

as underlying logic the first order predicate calculus. The rules of inference of this calculus are all truth-preserving. So a theorem of **Z**, as a sequence of formulas that are either axioms (which we have seen to be true) or derived from the axioms by rules of inference that are truth-preserving, will be a true proposition. The inconsistency of **Z** will only arise when one has *two theorems*, in which one is the negation of the other. But the negation of a true theorem can not be true.

## Part IV.

### "Using only the concepts definable in Z".

The sentence (above) refers to the concept of truth as not being arithmetically definable.

The situation was identified by Tarski and has essentially the following content:

If we call $\{W\}$ the set of Gödel numbers of wffs of **Z** *which are true in M* we will realize that this set is not arithmetical. That is, we can not build up a formula $\phi(x)$ in **Z** such that $\{W\}$ contains the Gödel numbers $k$ of formulas $\phi(k^*)$ *which are true in M*.

One can paraphrase this by saying that truth in **Z** is not **Z**-definable, or in other words, the concept of arithmetic truth is not arithmetically definable.

I believe that this also answers the second question "How is that different from $M$?".

I report on Tarski's Theorem in the last section of "Beyond Formalism" where a more formal approach is used.

## Part V.

### "What concepts are definable in Z?"

A concept or a relation is definable in **Z** if it is numeralwise expressible in **Z**. A function is definable in **Z** if it is numeralwise representable or numeralwise $\phi$-representable in **Z**.

On the whole we have a theorem about the equivalence of a function being recursive and being numeralwise representable in **Z**. Above one can find formal definitions of the concepts of expression, resp. representation in **Z**, as they were first formulated in Hilbert and Bernays *Grundlagen*.

I thank you for giving me the opportunity to try to be clear and remain yours sincerely,

M.S. Lourenço.