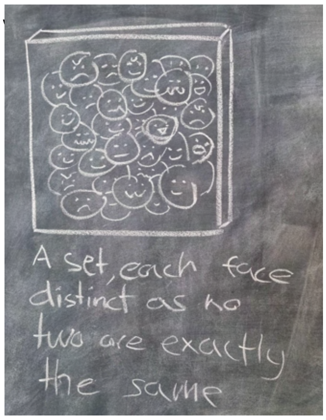
Kurt Gödel’s incompleteness theorems were established after David Hilbert began his “Hilbert program" in which he argued that any incomplete system could be made complete by adding more information. However, Gödel proved that with an advanced enough system of arithmetic, it was essentially impossible to prove completeness within that system. (Obviously this description is mostly just surface level)

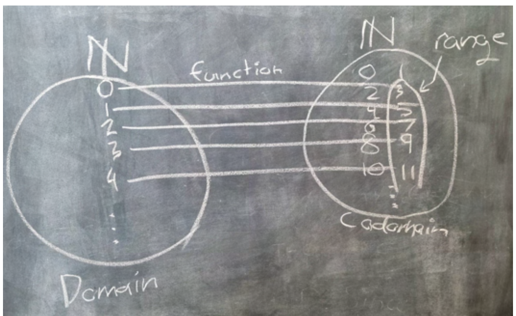
As suggested by its name, this is a “theorem" (in other words, it is a general statement or claim that isn’t self evident, which is proved through a chain of logic). Theorems are extremely useful in maths and logic as they govern basic principles such as addition of **natural numbers**, or $\mathbb{N}$. They extend to possibly the most famous theorem of mathematics, the pythagoras theorem, and to more complicated aspects of mathematics such as calculus or set theory. The collection of all natural numbers, which are also called counting numbers, is given the symbol $\mathbb{N}$, and as you probably guessed, $\mathbb{N}$, the collection of all natural numbers, begins at 0 as the first natural number (or 1 depending on who you talk to), and increases in increments of 1, i.e. 0, 1, 2, 3, and so on. Our first sight of mathematical notation, how exciting! $\mathbb{N}$ is a form of mathematical shorthand, which I referred to as a collection of numbers, but more correctly is called the **set** (A set is defined as a collection of distinct entities, or more mathematically, elements, the set being regarded as a single unit.) of natural numbers, notated as $\{0,1,2,3,\cdots\}$, { } curly brackets enclose all the elements within the set, and the $\cdots$ ellipsis shows the continuation of a pattern, in this case to infinity. Going more foundational, **axioms** are what we start from to create these theorems. An axiom is essentially a statement or rule that is self-evidently true, acting as the basis for further reasoning.
For example, **Robinson arithmetic**[1] has 7 axioms:

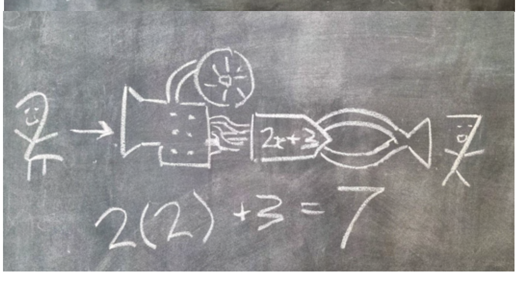


- 0 isn't the **successor** of a number

By successor I mean the number + 1, the number after, e.g. the successor of 3 is 4. In fact, this is the successor **function**, a function simply maps a value to exactly 1 other value. The set of values we map from is called the **domain**, and the set we map to is called the **codomain**. The codomain denotes the set we map into, whereas the range denotes the set of possible values that we can map into.

E.g. if we define a function $f:\mathbb{N}\mapsto\mathbb{N},\ f(x)=2x+3$,
Where f is the function, the : denotes *such that*, and the $\mapsto$ denotes $\mathbb{N}$ *maps to* $\mathbb{N}$, the domain of the function is $\mathbb{N}$, since that's what we map from, and the codomain is also $\mathbb{N}$, as this is the set we map into. However the **range** doesn't include 1 (as an example), as no natural number inputted into our function will ever output 1.





1. Raphael M. Robinson 1911 - 1995. Robinson arithmetic is often written as Q.

For convenience I will start notating the successor of x as S(x), which is another form of mathematical notation.

- If $S(x) = S(y)$, then $x = y$

(I use x and y as placeholders for numbers here, as these axioms apply to all the objects we're considering)

- If $x \neq 0$, then x has a successor

(I'm using ≠ as a shorthand for doesn't equal)

- $x + 0 = x$
- $x + S(y) = S(x + y)$

(The brackets around x + y signify the successor of the singular value of x + y)

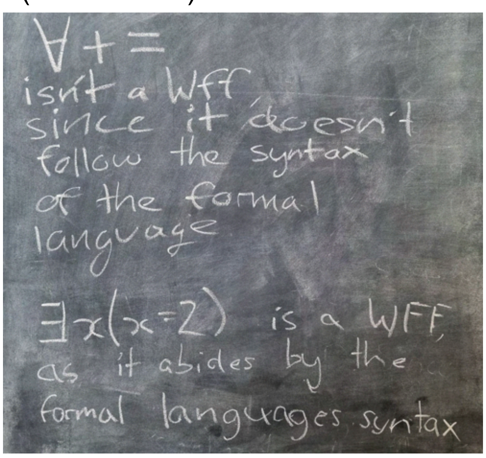
- $x \times 0 = 0$
- $x \times S(y) = (x \times y) + x$

Although it may be obvious, it's always good to define all of your symbols, so + means added to, × means multiplied by, 0 denotes the constant zero, and = means equals. This set of axioms is an example of an advanced enough system of arithmetic that can’t prove its own completeness

(Bear with me here) A formal language in logic has 1. An alphabet, the set of symbols used in that formal language 2. Semantics, or in other words definitions to its symbols, and 3. Syntax, i.e. the allowed order of strings of symbols to create a well-formed formula, its grammar. A well-formed formula (often abbreviated WFF, or just formula) is a finite number of symbols from the alphabet arranged according to the syntax of the formal language (but doesn't necessarily have meaning yet,). And finally, a sentence is a WFF that is either true or false.

Robinson arithmetic (Q) uses its **formal language** to **interpret** multiple **sentences** and simultaneously make them all true. To interpret in this sense is to give a sentence meaning and make it either true or false. It is using the semantics of the formal language to allow us to understand what is being stated in a specific scenario, (there are multiple interpretations for different semantics). In this way, Robinson arithmetic is a **theory**. To remove any ambiguity, a model is an interpretation of (a) sentence(s) that makes the sentence true. And furthermore a theory is the set of sentences (i.e. axioms) which we use a model to *satisfy*.

(Sorry for all the definitions)



Let me introduce you to some friends of mine, some mathematical symbols:

Symbol	Informal meaning	Symbol name
∀	For all/ for every	Universal quantifier
∃	There exists	Existential quantifier
⇒	Implies/ if … then …	Conditional
⇔	If and only if	Biconditional
S(x)	Take the successor of x	Successor function
¬ (or ~)	It is not true that …	Negation
⊢	… proves …	Turnstile
⊨	If the left hand side is true, the right hand side must be true (independent of your interpretation)	Double turnstile
⋀	… and …	Logical conjunction
⋁	… or …	Logical disjunction
⊆	… is a subset of …	
∈	… is an element of …	
>	… is greater than …	Greater-than sign
↦	… maps to …	
: or ∣	… such that …	
⌜…⌝	The Gödel number of …	
:=	Let … equal …	

A slash through a symbol usually denotes the opposite of the unslashed meaning, e.g. ∉ means “not an element of".

The turnstile denotes **semantic consequence**, whereas the double turnstile denotes **syntactic consequence**. True to their name, semantic consequence depends on the semantics of the model you're using, whereas syntactic consequence depends only on the syntax of the model. So if you used ⊢, you would be saying using this and *only* this specific model, I can prove the right from the left. If you used ⊨, you would be saying if the left is true, *regardless* of the model, the right is true. (When I say left and right, I refer to whatever lies to the left and right of the turnstile/double turnstile).

The ⊆ describes a relationship between two sets where the left set can only have elements included in the right set, it could have no elements or every element from the right set. In other words, the left set is a **subset** of the right set.

When I say element in reference to ∈, I mean the left object is included in the right set.

For a better definition of a sentence we need to use the terms free variables and bound variables. Free variables essentially don't have a quantifier (the universal or existential quantifiers listed in the table above), and so the formula they are contained in can't be true or false, because you haven't defined what x you are looking at. Whereas bound symbols (as suggested by their name) are variables that are bound to a specific quantifier, allowing the sentence to be either true or false. In this way, a WFF can have free or bound variables, but a sentence is a WFF with only bound variables, and can therefore be only true or false. WFF can also be written as f(x) for example, where the x denotes a free variable as the "input" of the formula.

For example:
$x=1$ has no meaning, as we have not stated what x we're looking at, whereas
$\exists x \in \mathbb{N}, x=1$ has meaning, as we have stated what we are looking at. The domain here represents the input x we are allowing for, in this case the domain is all natural numbers.

We are ready to write our first well-formed formula!
Here are 2 examples:

$\forall x \exists y (\neg(x=0) \Longleftrightarrow (S(y)=x))$ $x, y \in \mathbb{N}$
For all x there exists y such that x doesn't equal zero if and only if the successor of y equals x. Where = means equals. The "such that" here comes from the brackets after the quantifiers.

$\forall x \forall y (x+y=y+x)$ $x, y \in \mathbb{N}$
[For all x and y, x + y = y + x][2]

x and y are both natural numbers as stated by the $x, y \in \mathbb{N}$

2. This is known as **commutativity**, where x + y = y + x

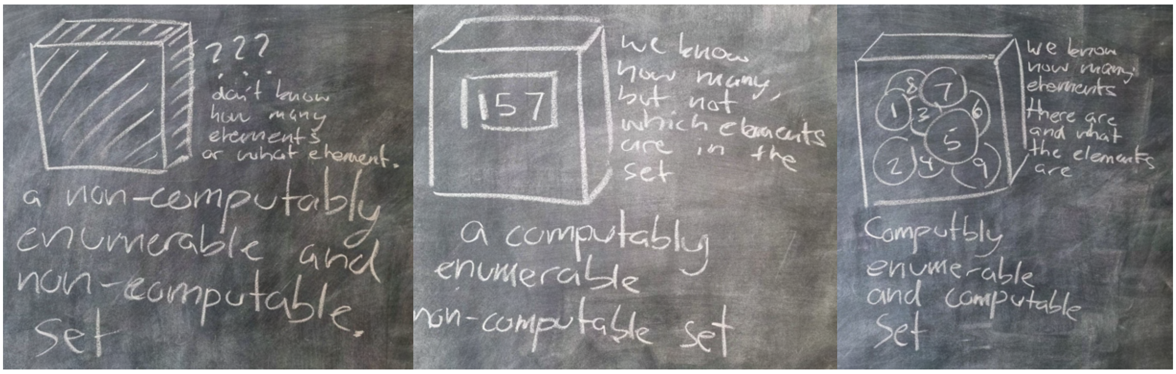
Let me begin by defining what we mean by **incomplete**: when we consider every possible WFF we could create within the axiomatized system, if every possible sentence can be proven either true or false, the **system**[3] is complete, and, what a surprise, if you can't prove or disprove *every* sentence within the system, it is incomplete.

So how do we prove incompleteness? A tough question. I will *attempt* to show you **an outline** of a proof:

Let us build on what Gödel numbers are, and say that we will assign every formula its own Gödel number in an **effective way** (in a finite number of exact steps). I will notate the Gödel number of a formula, lets say φ, as $\ulcorner \varphi \urcorner$. Gödel numbers are a way of **encoding** formulas. Additionally, if $\ulcorner \varphi \urcorner = c$, we can notate this as φ_c. Then we can also assign Gödel numbers to finite sequences of formulas (this will become useful later).

A set is **computable** if all its elements are solvable with an effective procedure.

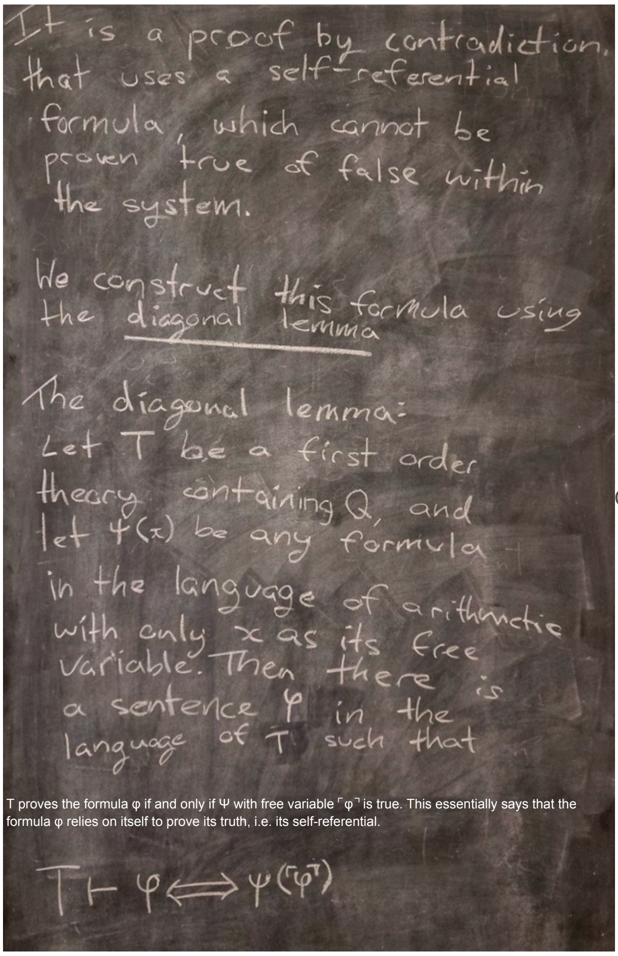
A set is **computably enumerable** if there is an **algorithm** (a finite sequence of mathematically rigorous instructions) that can **index** (label/number uniquely) any element within a finite time, (even if the set is infinite, if any element can be indexed in a finite time the set is computably enumerable).



An ω-inconsistent theory, can prove for some formula, F(x), F(x) is true for all $\mathbb{N}$, while also being able to prove it isn't true for some $\mathbb{N}$. Therefore in an **ω-consistent** theory there doesn't exist F(x) for which this occurs. The definition of a **consistent** theory (not the definition of *ω-consistent*) is that a theory cannot prove something is both true and false.[4]

3. When I say system, I refer to a formal **axiomatic system**, which is essentially a framework from which, with the axioms of the system, we derive theorems.
4. The **principle of explosion** is the idea that you can prove anything from a contradiction, i.e. proving something is both true and false.

Outline of Gödel's first incompleteness theorem proof



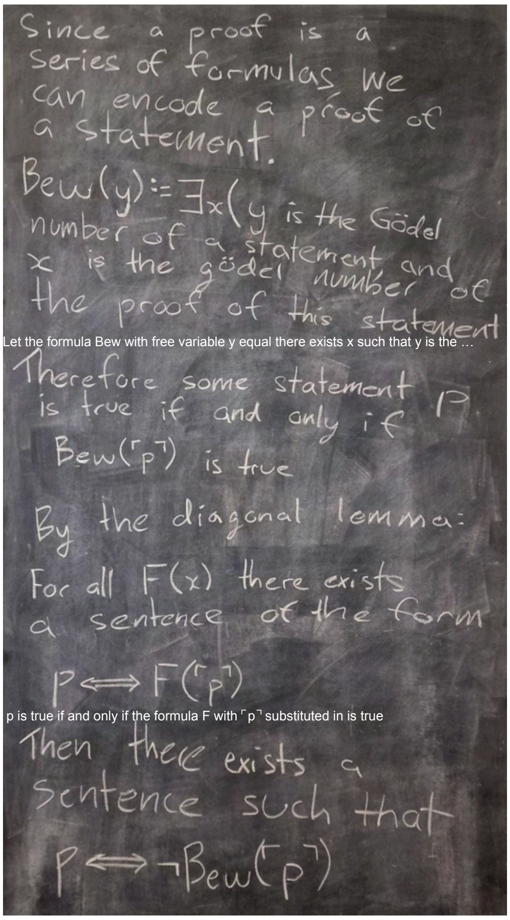
T proves the formula φ if and only if Ψ with free variable ⌜φ⌝ is true. This essentially says that the formula φ relies on itself to prove its truth, i.e. its self-referential.

Self-referential means referring to itself, in this case a formula which refers to itself

A lemma is a small theorem

First order is a long definition, to find out more sources are listed at the bottom of this document.
Crucially, T is also ω-consistent

Gödel numbers are useful as they prevent a formula "going on forever"[5] if we want it to be self-referential.



Let the formula Bew with free variable y equal there exists x such that y is the …

p is true if and only if the formula F with ⌜p⌝ substituted in is true

Bew(y) is a formula with one free variable

5. We call this property of "going on forever" in reference to a formula, where it keeps plugging itself into itself, an **infinite regress**.

I.e. p is true if and only if the negation of Bew with ⌜p⌝ substituted in is true

Intuitively, we have just proved that given a first order theory that contains Q, we can use the diagonal lemma to prove that p is true if and only if there doesn't exist a Gödel number for the proof of p, i.e. p is true if and only if p doesn't have a proof within T. So if p is true then p is unprovable. And if p is false then Bew(⌜p⌝) is true since this is the definition of Bew, this is essentially stating that if p is false, then there exists a proof for p, which breaks the ω-consistency of T. Therefore p cannot be provable, and therefore T is incomplete.

Sources:

- Gödel's Incompleteness Theorem
- https://en.wikipedia.org/wiki/Diagonal_lemma
- https://en.wikipedia.org/wiki/G%C3%B6del%27s_incompleteness_theorems#Proof_sketch_for_the_first_theorem
- https://plato.stanford.edu/entries/goedel-incompleteness/
- Math stack exchange with numerous different questions
- Oxford english dictionary
- Many more wikipedia articles to confirm definitions

First-order theory:

- First-order Model Theory
- First-Order Logic