Universiteit Utrecht

Gödel's Incompleteness Theorems

Linda van Vliet

Bachelor Thesis

Mathematics and Applications

Supervised by Dr. J. van Oosten

January 2020

Contents

Introduction

Gödel's incompleteness theorems are considered to be among the most important results in modern logic. His theorems concern what is now called *Peano Arithmetic*, a widely used and powerful theory of arithmetic, and show its limitations. To appreciate the theorems and their impact, we would benefit from considering the context of Gödel's work.

In the begin of the twentieth century, many mathematicians were concerned with paradoxes related to set theory. In an attempt to save mathematics from any such absurdities, Bertrand Russell and Alfred North Whitehead published Principia Mathatica between 1910 and 1913. They were positive, and made a great attempt to show, that all of mathematics could be derived from logic and a given set of axioms. Certainly, mathematics would be free from contradiction. This final property was a great concern to David Hilbert as well, who posed in 1900 as his second of twenty three open problems in mathematics: the consistency of the axioms of arithmetic (which may be interpreted as the axioms of Peano Arithmetic).

Gödel showed that the visions of Russell, Whitehead and Hilbert were entirely too optimistic. His first incompleteness theorem states that Peano Arithmetic is incomplete; that we can construct a *true* number theoretical statement that is unprovable in it. His second incompleteness theorem states that Peano Arithmetic is unable to prove its own consistency. We shall prove both theorems in this paper.

The idea behind the first incompleteness theorem, in its barest form, reminds us of the Epimenides paradox. This paradox was named after the Cretan Epimenides, who made the statement "All Cretans are liars". A more direct phrasing of the paradox would be "This statement is false". The paradox violates the common intuition that all statements must be either true or false; in the case of this sentence, each option leads to a contradiction. In mathematics however, a well-formed statement cannot violate this dichotomy. So how does the paradox relate to mathematical statements? This was one of Gödel's grandest discoveries: finding a way to apply mathematical reasoning *to* mathematical reasoning itself. The proof of the first theorem relies on constructing mathematical statement that, in some sense, is self-referential. Gödel coded statements and proofs and assigned them a unique natural number. This idea is often called *Gödel-coding* and will discussed in this paper. It is this trick that makes it possible to apply mathematics of natural numbers to statements *about* natural numbers as well since they can be represented by a natural number. We will eventually construct a *Gödel-sentence* which may be interpreted as "This statement is unprovable", reminding us of the Epimenides paradox. Mathematical proofs depend on a defined *system*; in

the case of the Gödel-sentence, this was Peano Arithmetic. We will see though, that the theorem can be applied to any system of number theory meeting certain conditions so that we can find a similar unprovable statement. The first theorem therefore refutes the project of Russell and Whitehead.

The second incompleteness theorem troubles Hilbert's vision. Hilbert proposed his *Programme:* to formulate a logical system in which we construct proofs about, among other things, the natural numbers. He advocated for studying proofs so that using this system, we can prove the consistency of arithmetic. Gödel showed that the hope of proving a system such as Peano Arithmetic, using Peano Arithmetic, is in vain. His second theorem therefore crushes Hilbert's Programme.

Outline of this paper

The paper is structured as follows. The first section of this paper discusses languages and systems. It is included to introduce concepts and definitions important to Gödel's incompleteness theorems. In the second section, we present the theory of Peano Arithmetic and develop some number theory inside it. The third section defines primitive recursive functions. These functions are an important tool for the construction of the Gödel-sentece. We furthermore show how they can be represented in Peano Arithmetic. The first three sections all provide groundwork for proving Gödel's incompleteness theorems. We shall finally prove the first and second theorem in the fourth and fifth section of this paper respectively. Additionally, we provide arguments for the theorems being applicable to arithmetical theories neatly extending Peano Arithmetic, which demonstrates their strength.

The structure and results of this paper are heavily based on ([**3**]). However, some definitions and lemmas are omitted or reformulated to suit this paper. The results of section 4 and 5 take some inspiration of ([**4**]) as well.

1 Formal Theories

The Incompleteness Theorems of Gödel prove the limits of *theories of arithmetic*. Before proving the theorems, we need to consider the notion of a formal theory - or simply a theory - as it is used in the study of mathematical logic. A theory is built in a *formal language* and consists of a set of sentences in that language which are considered the *axioms* of that theory. Finally it has a *deductive system* which establishes how sentences may be derived from the axioms.

1.1 Formal languages

Since we will limit ourselves in this paper to first-order logic, we implicitly refer to *first-order* languages whenever we mention or define languages.

Definition 1.1. A *language* $\mathcal{L}$ is a syntactically defined system of expressions. The language consists of a set of non-logical symbols: *constants, relation symbols* and *function symbols.* Next we consider a set of logical symbols consisting of

- the equality sign $=$;
- the absurdity symbol $\bot$;
- readability symbols such as brackets and commas;
- the binary connectives $\vee, \wedge, \rightarrow$ and the unary connective $\neg$;
- the quantifiers $\forall$ and $\exists$.

Together with formation rules, these non-logical and logical symbols determine what constitutes as terms and formulas of $\mathcal{L}$.

Definition 1.2. Given language $\mathcal{L}$, the set of *terms* is defined as follows. Any constant c of $\mathcal{L}$ and any variable x is a term of $\mathcal{L}$. If $t_1, \ldots, t_n$ is an n-tuple of terms and f is an n-place function of $\mathcal{L}$, then $f(t_1, \ldots, t_n)$ is a term of $\mathcal{L}$. Nothing else is a term.

When a term does not contain any variables, it is called *closed.*

Definition 1.3. Given language $\mathcal{L}$ the set of *formulas* is defined as follows.

- If t_1 and t_2 are terms of $\mathcal{L}$, then $(t_1 = t_2)$ is a formula of $\mathcal{L}$.
- If $t_1, \ldots, t_n$ is an n-tuple of $\mathcal{L}$ and R an n-place relation symbol, then $R(t_1, \ldots, t_n)$ is a formula of $\mathcal{L}$.
- $\bot$ is a formula of $\mathcal{L}$.

– If φ and ψ are formulas of $\mathcal{L}$ and x is a variable, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, $(\neg\varphi)$, $(\forall x\varphi)$ and $(\exists x\varphi)$ are formulas of $\mathcal{L}$.

The first three items define what are considered *atomic* formulas.

Next we consider some definitions applicable to variables and formulas. If a variable x occurs in $\forall x$ or $\exists x$ it is called *bound* in that formula; otherwise it is called *free.* A formula that has no free variables is called a *sentence* or a *statement* of $\mathcal{L}$.

Definition 1.4. Let φ be an $\mathcal{L}$-formula and τ be an $\mathcal{L}$-term. We write $\varphi[\tau/x]$ te denote the formula that is obtained by replacing each occurrence of variable x in φ by the term τ. We are allowed to do this when x is a free variable in φ and all variables in τ remain free in φ. Then $\varphi[\tau/x]$ is called a *substitution.*

1.2 Interpretations

A formal language determines what syntactically constitutes a formula or sentence. By itself, a language does not assign meaning to its formulas. This is where an interpretation is needed. First we will define what is means to be an $\mathcal{L}$-structure for a formal language $\mathcal{L}$.

Definition 1.5. An *$\mathcal{L}$-structure* M consists of a non-empty set together with the following:

– for each n-place function symbol f and each k-place relation symbol R of $\mathcal{L}$, a function f^M and a subset R^M called the interpretations of f and R in M:
$$f^M : M^n \to M \quad \text{and} \quad R^M \subset M^k$$

– for each constant c of $\mathcal{L}$, an element c^M of M called the interpretation of c in M

Given $\mathcal{L}$-structure M, we consider language $\mathcal{L}_M$ that consists of $\mathcal{L}$ and, for each element m of M, and extra constant (also denoted m).

Then for each closed term t of language $\mathcal{L}_M$, the interpretation t^M as element of M is defined by induction. It has already been defined for the case when t is a constant. If t is of the form $f(t_1, \ldots, t_n)$ then $t_1, \ldots, t_n$ are closed terms of $\mathcal{L}_M$ for which, by induction hypothesis, interpretations $t_1^M, \ldots, t_n^M$ have already been defined. Then we let
$$t^M = f^M(t_1^M, \ldots, t_n^M)$$

Definition 1.6. Next we will inductively define what it means for "a formula φ to be true in a structure M" (and the synonymous phrasings "φ holds in M" and

"M satisfies φ"). We use the following notation to express this relation:

$$M \models \varphi$$

For atomic formulas we define:

- $M \models \bot$ never holds
- $M \models (t_1 = t_2)$ iff $t_1^M = t_2^M$
- $M \models R(t_1, \ldots, t_n)$ iff $(t_1^M, \ldots, t_n^M) \in R^M$

where $t_1, \ldots, t_n$ are closed terms.

For formulas constructed by logical connectives and quantifiers we define:

- $M \models (\varphi \wedge \psi)$ iff $M \models \varphi$ and $M \models \psi$
- $M \models (\varphi \vee \psi)$ iff $M \models \varphi$ or $M \models \psi$
- $M \models (\varphi \to \psi)$ iff $M \models \psi$ whenever $M \models \varphi$
- $M \models \neg\varphi$ iff $M \not\models \varphi$
- $M \models \forall x\varphi$ iff $M \models \varphi[n/x]$ for all $n \in M$
- $M \models \exists x\varphi$ iff $M \models \varphi[n/x]$ for some $n \in M$

In the last two clauses, $\varphi[n/x]$ is the result of substituting variable x by a new constant n as stated in definition 1.5.

1.3 Formalized theories

Definition 1.7. A *theory* Γ in language $\mathcal{L}$, or an $\mathcal{L}$*-theory*, is a set Γ of $\mathcal{L}$-sentences called the *axioms* of the theory.

If it can be effectively decided whether a given formula is an axiom, we call Γ an *axiomatized theory.*

Definition 1.8. A *model* M of Γ is n $\mathcal{L}$-structure such that $M \models \varphi$ for every $\varphi \in \Gamma$. We will use notation $\Gamma \models \varphi$ if $M \models \varphi$ for every model M of Γ.

Definition 1.9. Given a $\mathcal{L}$-sentence φ, we write

$$\Gamma \vdash \varphi$$

to denote the relation "φ is *provable* in Γ", i.e. there exits a Γ-proof of φ.

We abrreviate $\{\varphi\} \vdash \psi$ as $\varphi \vdash \psi$, $\emptyset \vdash \psi$ as $\vdash \psi$ and $\Gamma \cup \{\varphi\} \vdash \psi$ as $\Gamma, \varphi \vdash \psi$.

Remark 1.10. Given theories Γ and Δ, if $\Gamma \subseteq \Delta$ and $\Gamma \vdash \varphi$, then $\Delta \vdash \varphi$.

Remark 1.11. Formulas are proven in a theory Γ when they are derived from premisses and axioms by the correct application of *inference rules.* These rules depend on the deductive system for which Γ is defined. Unfortunately, it is beyond the scope of this paper to define and discuss any deductive system in particular. We will assume that our 'ordinary' mathematical reasoning as used in this paper is equivalent - in the sense of being able to prove precisely the same formulas from a given set of premisses - to common deductive systems such as natural deduction and Hilbert-style systems. The assumption includes that our reasoning is truth-preserving.

We refer the reader to van Oosten ([**3**], p.11-16) and Mendelson ([**2**], p. 35-36) for detailed overviews of natural deduction and a Hilbert-style system respectively. Furthermore, these texts formally define *proofs* according to the inference rules of their introduced deductive systems.

Definition 1.12. An $\mathcal{L}$-theory Γ is *consistent* if it does not prove any contradiction; for each $\mathcal{L}$-sentence φ, $\Gamma \nvdash \varphi$ or $\Gamma \nvdash \neg\varphi$, or both.

Definition 1.13. An $\mathcal{L}$-theory Γ is *complete* if for every $\mathcal{L}$-sentence φ we have that $\Gamma \models \varphi$ or $\Gamma \models \neg\varphi$. If there is a sentence ψ such that $\Gamma \not\models \psi$ and $\Gamma \not\models \neg\psi$ then Γ is *incomplete* and ψ is called *independent of* Γ.

2 Peano Arithmetic

We consider the system of Peano Arithetic, an important theory of arithmetic. The system of Peano Arithmetic, or PA, is a theory in the language

$$\mathcal{L}_{PA} = \{0, 1, +, \cdot\}.$$

Here 0 and 1 are its constants and $+$ and $\cdot$ its binary relation symbols. For the purpose of readability we will write $(x+y)$ and $(x \cdot y)$ to denote $+(x, y)$ and $\cdot(x, y)$ respectively.

Peano Arithmetic has the following axioms:

(A1) $\forall x \neg(x + 1 = 0)$

(A2) $\forall xy(x + 1 = y + 1 \rightarrow x = y)$

(A3) $\forall x(x + 0 = x)$

(A4) $\forall xy(x + (y + 1) = (x + y) + 1)$

(A5) $\forall x(x \cdot 0 = 0)$

(A6) $\forall xy(x \cdot (y + 1) = (x \cdot y) + x)$

(A7) $\forall \vec{x}[(\varphi(0, \vec{x}) \wedge \forall y(\varphi(y, \vec{x}) \rightarrow \varphi(y + 1, \vec{x}))) \rightarrow \forall y \varphi(y, \vec{x})]$

The seventh item is to be read as an axiom scheme for what we call the *induction axioms*; so there is an induction axiom for every formula $\varphi(y, \vec{x})$. We see that PA is an *axiomatized* theory: it can be effectively decided whether a formula is among the first six axioms or whether it is an instance of the induction schema, despite PA having infinitely many axioms.

As the formation rules of terms have been defined for formal languages in general we can apply them to determine the terms of $\mathcal{L}_{PA}$. Any variable, as well as the constants 0 and 1, are terms. If σ and τ are terms, then so are $(\sigma + \tau)$ and $(\sigma \cdot \tau)$. Finally, as expected, nothing else is a term.

Since there are different structures in which the axioms above are true, there are multiple models of the theory. The set of natural numbers, including 0 and 1, together with its usual multiplication and addition, is one model of $\mathcal{L}_{PA}$. This model is referred to as the *standard model* of PA. The standard model is denotated by $\mathcal{N}$.

Definition 2.1. Even though constants for all natural numbers are not included in $\mathcal{L}_{PA}$, it will be useful to abbreviate the terms that represent the natural numbers in $\mathcal{N}$. To this end, we define for every $n \in \mathbb{N}$ a term $\overline{n}$ of $\mathcal{L}_{PA}$ which we call the

standard numeral. The standard numerals are defined as follows: $\overline{0} = 0$ and for $n \in \mathbb{N}$, $\overline{n+1} = \overline{n} + 1$. For example, the term ((0+1)+1) is equal to the term $\overline{2}$.

Peano arithmetic is able to express and prove many properties of elementary number theory as will be shown in later chapters. For now, we will prove some basic properties of addition.

Proposition 2.2. *The following statements hold.*

(P1) $\text{PA} \vdash \forall xyz(x + (y + z) = (x + y) + z)$

(P2) $\text{PA} \vdash \forall xy(x + y = y + x)$

(P3) $\text{PA} \vdash \forall x(x = 0 \vee \exists y(x = y + 1))$

(P4) $\text{PA} \vdash \forall xy\exists z(x + z = y \vee x = y + z)$

(P5) $\text{PA} \vdash \forall xyz(x + z = y + z \to x = y)$

Proof. To prove the properties listed above we will use induction in PA.

For *(P1)* we let $\varphi(z)$ be $\forall xy(x+(y+z) = (x+y)+z)$. It immediately follows from (A3) that $\text{PA} \vdash \varphi(0)$. Assuming $\varphi(z)$, we see that $\forall xy(x + (y + z) = (x + y) + z)$ holds in PA. Using (A4), we can deduct that $x+(y+(z+1)) = x+((y+z)+1) = (x + (y + z)) + 1 = ((x + y) + z) + 1 = (x + y) + (z + 1)$ so the conclusion $\varphi(z + 1)$ holds. Therefore $\text{PA} \vdash \forall z\varphi(z)$.

To prove *(P2)* we set $\varphi(x) \equiv \forall y(x + y = y + x)$ so that we need to prove $\text{PA} \vdash \forall x\varphi(x)$. First we prove $\text{PA} \vdash \varphi(0)$, equivalently $\text{PA} \vdash \forall y(0 + y = y + 0)$, using induction on y. Set $\psi_0(y) \equiv (0 + y = y + 0)$. Immediately we notice that $\text{PA} \vdash 0 + 0 = 0 + 0$ so $\text{PA} \vdash \psi_0(0)$. Next we assume $\text{PA} \vdash \psi_0(y)$ and derive $\text{PA} \vdash \psi_0(y+1)$. By our assumption, PA proves that $0+y = y+0$; by the axioms and *(P1)*, we have that

$$0 + (y + 1) = (0 + y) + 1 = (y + 0) + 1 = y + 1 = (y + 1) + 0$$

so $\text{PA} \vdash 0 + (y + 1) = (y + 1) + 0$. This proves that $\text{PA} \vdash \psi_0(y + 1)$ and finally we have $\text{PA} \vdash \forall y\psi_0(y)$, so inded $\text{PA} \vdash \varphi(0)$.

Next we let $\psi_1(y)$ be $1 + y = y + 1$ and prove $\text{PA} \vdash \forall y\psi_1(y)$. Clearly $\text{PA} \vdash \psi_1(0)$. We assume $\text{PA} \vdash \psi_1(y)$, i.e. PA proves $1 + y = y + 1$. We see that $1 + (y + 1) = (1 + y) + 1 = (y + 1) + 1$ by *(P1)* so $\text{PA} \vdash \psi_1(y + 1)$. We conclude $\text{PA} \vdash \forall y\psi_1(y)$. Now we have all we need to prove $\text{PA} \vdash \forall x\varphi(x)$. We assume that $\text{PA} \vdash \varphi(x)$, i.e. that PA proves $x + y = y + x$. Using the induction hypothesis, as well as the axioms and properties of PA, we have

$$(x + 1) + y = x + (1 + y) = x + (y + 1) = (x + y) + 1 = (y + x) + 1 = y + (x + 1).$$

It follows that PA $\vdash \forall x(\varphi(x) \to \varphi(x{+}1))$. Finally we have PA $\vdash \forall xy(x{+}y = y{+}x)$ as desired.

The proof of *(P3)* is quite straightforward. For $x = 0$, we trivially have $x = 0 \vee \exists y(x = y + 1)$. Assuming for an arbitrary x that $x = 0 \vee \exists y(x = y + 1)$ holds in PA, we see that $x + 1 = 0 \vee \exists y(x + 1 = y + 1)$ holds as well. So we have PA $\vdash \forall x(x = 0 \vee \exists y(x = y + 1))$.

For *(P4)* we let $\varphi(x) \equiv \forall y \exists z(x + z = y \vee x = y + z)$. Considering $\varphi(0)$, we take $z = y$ and see that $0 + z = y$ follows from the axioms and *(P2)*. Therefore we have PA $\vdash \varphi(0)$. Next we assume PA $\vdash \varphi(x)$: for an arbitrary x we have that $\forall y \exists z(x + z = y \vee x = y + z)$. We consider two cases: $z = 0$ and $\neg(z = 0)$ and see if $\exists v((x + 1) + v = y \vee x + 1 = y + v)$ holds in PA in order to conclude $\varphi(x + 1)$.

When $z = 0$, our assumption leads to $(x{+}0 = y \vee x = y{+}0)$ or simply $x = y$. Let v be 1 and we see that the formula $x + 1 = y + v$ is satisfied since $x + 1 = y + 1$. So $\varphi(x + 1)$ is satisfied as well.

When $\neg(z = 0)$, there is a variable w such that $z = w{+}1$ by *(P3)*. Our assumption then leads to $x + (w + 1) = y$ or $x = y + (w + 1)$. In the first case we let $v = w$ and apply *(P1)* and *(P2)* to arrive at $(x + 1) + v = y$. In the second case we let $v = (w + 1) + 1$ and have $x = y + (w + 1)$, so $x + 1 = y + (w + 1) + 1$ and $x + 1 = y + v$. In all cases we have $\exists v((x + 1) + v = y \vee x + 1 = y + v)$, so $\varphi(x{+}1)$ holds in PA whenever $\varphi(x)$ holds in PA. Therefore we may conclude PA $\vdash \forall x \varphi(x)$.

(P5). We let $\varphi(x, y, z)$ be the formula $x + z = y + z \to x = y$. Clearly PA $\vdash x + 0 = y + 0 \to x = y$ from the axioms so PA $\vdash \varphi(x, y, 0)$ holds. Next we assume $varphi(x, y, z)$ holds in PA. Furthermore, we assume $x{+}(z{+}1) = y{+}(z{+}1)$. By *(P1)*, $(x{+}z){+}1 = (y{+}z){+}1$ and by (A2), $x{+}z = y{+}z$. From the induction hypothesis we have that $x = y$. We see that PA $\forall xy(\varphi(x, y, 0) \wedge \forall z(\varphi(x, y, z) \to \varphi(x, y, z + 1)))$. By induction we finally see that PA $\vdash \forall xyz(x + z = y + z \to x = y)$. ■

We can express the *less-than* relation in $\mathcal{L}_{PA}$ by the formula $\exists z(x + (z + 1) = y)$. We will prove this formula defines a discrete linear order in PA. Additionally, this order has a least element and satisfies the least number principle as we will soon see. Since the order will be used often, we introduce the notation $x < y$ to abbreviate $\exists z(x + (z + 1) = y)$.

We introduce some more common abbreviations. We will use the natural notation $x \leq y$ for $x < y \vee x = y$ and $x \neq y$ for $\neg(x = y)$. Furthermore, we will use the abbreviations $\forall x < y\varphi$ and $\exists x < y\varphi$ for $\forall x(x < y \to \varphi)$ and $\exists x(x < y \wedge \varphi)$ respectively.

Proposition 2.3. *In* PA, $<$ *satisfies the following.*

1. *It is a discrete linear order:*
   - *(i)* $\text{PA} \vdash \forall x(x < x+1)$
   - *(ii)* $\text{PA} \vdash \forall x \neg(x < x)$
   - *(iii)* $\text{PA} \vdash \forall xyz(x < y \wedge y < z \to x < z)$
   - *(iv)* $\text{PA} \vdash \forall xy(x < y \vee x = y \vee y < x)$
   - *(v)* $\text{PA} \vdash \forall xy(x < y \to x + 1 \leq y)$
2. *It has a least element:*

   $$\text{PA} \vdash \forall x(x = 0 \vee 0 < x)$$
3. *It satisfies the least number principle: for all formulas* φ*,*

   $$\text{PA} \vdash \exists x\varphi(x) \to \exists y(\varphi(y) \wedge \forall w < y \neg\varphi(w))$$

Proof. 2. We will consider first the proof of the order having a least element in PA. It follows immediately from *(P3)* that for all x, $x = 0$ or $\exists y(y + 1 = x)$. In the second case, $\exists y(0 + (y + 1) = x)$ and we see that $0 < x$ holds by definition.

1. New we prove $<$ is a linear order, beginning with item *(i)*. Notice that $x < x+1$ is an abbreviation for $\exists z(x + (z + 1) = x + 1)$. We let z be 0 and we are done. For property *(ii)* we assume the opposite. That is: there is an x such that $x < x$ holds. For that x there is a z such that $x + (z + 1) = x$, which is equivalent to $x + (z + 1) = x + 0$. By *(P5)*, it follows that $z + 1 = 0$, which contradicts the axioms of PA. Therefore $\text{PA} \vdash \forall x \neg(x < x)$.

For the third property, reason within PA and assume $x < y$ and $y < z$ for arbitrary x, y, z. Then there exist u, v such that $x + (u + 1) = y$ and $y + (v + 1) = z$. By substituting y we get $(x + (u + 1)) + (v + 1) = z$ which is equivalent to $x + (((u + 1) + v) + 1 = z$. Evidently this may be written as $x < z$.

Next we prove property *(iv)*. It follows from *(P4)* that given x, y in PA, there is a z such that $x + z = y$ or $x = y + z$. If that z is equal to 0, in both cases $x = y$ holds, therefore $x < y \vee x = y \vee y > x$ holds as well. If that z is greater than 0, we define an integer v such that $v + 1 = z$. Then in the case of $x + z = y$ it follows that $x + (v + 1) = y$ which, by definition, is equivalent to $x < y$. In the case of $x = y + z$ we have that $y + (v + 1) = x$ therefore $y > x$. We have proven that in all cases, either $x < y$, $x = y$ or $y > x$ holds. So $\text{PA} \vdash \forall xy(x < y \vee x = y \vee y < x)$.

(v) For the final item we assume in PA $x < y$; that is, $\exists z(x + (z + 1) = y)$. Since PA has a least number, we know that $z = 0$ or $0 < z$. If $z = 0$ we have that

$x + (z + 1) = x + 1 = y$. If $0 < z$, we have $\exists v(0 + (v + 1) = z)$. We substitute z in the assumed formula $\exists z(x + (z + 1) = y)$ to get $\exists v(x + ((v + 1) + 1) = y)$. Rewriting gives us $\exists v(x + 1 + (v + 1) = y)$ which is abbreviated by $x + 1 < y$. So in both cases we have $x + 1 \leq y$.

3. Finally we move on to the *least number principle* as stated in item 3. To prove the principle we reason in PA, assume the opposite of the principle and derive a contradiction. So let $\varphi(x)$ hold for a certain variable x and certain formula φ. Now let there be no least y such that $\varphi(y)$. So:

$$\exists x\varphi(x) \wedge \forall y(\varphi(y) \rightarrow \exists z < y\varphi(z)).$$

Consider the formula $\forall y(\varphi(y) \rightarrow z < y)$ which we will denote with $\psi(z)$. We use induction to prove $\forall z\psi(z)$ which will ultimately lead to our desired contradiction. For $z = 0$, the formula $\forall y(\varphi(y) \rightarrow 0 < y)$ is true since by assumption there is no least element for which φ holds, φ cannot hold for 0. So indeed for all y, if $\varphi(y)$ holds then $0 < y$. The induction hypothesis is $\psi(z)$: $\forall y(\varphi(y) \rightarrow z < y)$. We want to derive $\forall y(\varphi(y) \rightarrow z + 1 < y)$.

Next we assume that $\varphi(z + 1)$ holds. Since we assumed that there is no least element for which φ holds, there has to be a $w < z + 1$ such that $\varphi(w)$ holds as well. We have $w \leq z$; however, it follows from the induction hypothesis of ψ that $z < w$ follows from $\varphi(w)$. This clearly contradicts $w \leq z$, so our assumption of $\varphi(z + 1)$ holding in PA is false. Then $\psi(z + 1)$ holds in PA trivially, so we may conclude $\mathrm{PA} \vdash \forall yz(\varphi(y) \rightarrow z < y)$. From this statement we can derive in PA that $\varphi(x) \rightarrow x < x$. We assumed in the beginning of our proof that $varphi(x)$ holds so we must indeed conclude $x < x$. This of course contradicts the irreflexivity of the order $<$. Finally we have reached a contradiction from the assumption of there being no least number for which a formula holds. We may conclude that the least number principle holds for $<$. ■

2.1 Properties of multiplication and division in PA

This section serves two main purposes. First, the strength of PA is demonstrated as it is shown how some of the most fundamental theorems of elementary number theory can be proven in PA. Second of all, the section will show how sequences of numbers may be coded in PA *using* the elementary properties of multiplication and division. As will become clear in the next chapter, this coding is what allows us to capture many other functions in PA.

We begin by stating some fundamental arithmetical properties that hold in PA. They are similar to those in proposition 2.2 and since their proofs are similar as well, proofs will be omitted here.

Proposition 2.4. *The following statements hold for PA:*

1. PA $\vdash \forall xyz((x \cdot y) \cdot z = x \cdot (y \cdot z))$
2. PA $\vdash \forall xyz(x \cdot y = y \cdot x)$
3. PA $\vdash \forall xyz(x \cdot (y + z) = x \cdot y + x \cdot z)$
4. PA $\vdash \forall xyz(z \neq 0 \wedge x \cdot z = y \cdot z \rightarrow x = y)$

As the properties in proposition 2.2 and 2.4 are natural, we will use them in the forthcoming proofs without mentioning.

Theorem 2.5 (Division with remainder)**.**

$$\text{PA} \ \vdash \forall xy(y \neq 0 \rightarrow \exists ab(x = a \cdot y + b \wedge 0 \leq b < y))$$

Furthermore, PA *proves that such a and b are unique.*

Proof. We will use induction on x. For $x = 0$, clearly $0 = 0 \cdot y + 0$ is satisfied in PA for all y. We assume the equality holds for x: for all y, there are a and b such that $x = a \cdot y + b$ and $0 \leq b < y$. Now we consider $x + 1 = a \cdot y + b + 1$ and notice that since $b < y$ we have $b+1 \leq y$. If $b+1 < y$, we let $b' = b+1$ and have $x+1 = a \cdot y + b'$ with $0 \leq b' < y$. If $b + 1 = y$, we have $x + 1 = a \cdot y + b + 1 = a \cdot y + y = (a + 1) \cdot y$. We let $a' = a + 1$ and $b' = 0$ and have $x = a' \cdot y + b'$ where b' satisfies $0 \leq b < y$ as well. Our induction proof is completed.

Next we prove uniqueness of a and b. Suppose there are x and y such that $x = a \cdot y + b = a' \cdot y + b'$ with $0 \leq b, b' < y$. First we suppose $a < a'$. It follows that $a + 1 \leq a'$ and

$$a \cdot y + b < a \cdot y + y \leq a' \cdot y \leq a' \cdot y + b'$$

So $a \cdot y + b < a' \cdot y + b'$, but this contradicts our assumption. We conclude that $a \geq a'$, though it is easy to see that an assumption of $a > a'$ leads to $a \leq a'$ similarly. Therefore we have $a = a'$, which means that a is unique. We then have $x = a \cdot y + b = a \cdot y + b'$ and see that $b = b'$ as well. ■

We will introduce some useful notation:

$$\begin{aligned} x|y &\equiv \exists z(x \cdot z = y) \\ \text{prime}(x) &\equiv x > 1 \wedge \forall yz(x|(y \cdot z) \rightarrow x|y \vee x|z) \end{aligned}$$

The following proposition is stated without proof as it may easily be checked by the reader.

Proposition 2.6. $\text{PA} \vdash \forall x(x > 1 \to \exists y(\text{prime}(y) \wedge y|x))$

Theorem 2.5 allows us to define two more important relations. For $x, y \geq 1$ we know that $x|(x \cdot y)$ and $y|(x \cdot y)$. Then by the least number principle, there is a unique smallest $w > 0$ such that $x|w$ and $y|w$. We denote this w by $lcm(x, y)$ and call it the *least common multiple* of x and y. Clearly it is the case that $lcm(x, y) \leq x \cdot y$.

We then write $x \cdot y = a \cdot lcm(x, y) + b$ for certain a, b where $0 \leq b < lcm(x, y)$. It follows that $x|b$ and $y|b$. If we assume that $0 < b < lcm(x, y)$, the minimality of $lcm(x, y)$ is contradicted. Therefore $b = 0$ and we have $x \cdot y = a \cdot lcm(x, y)$ for a certain unique a. This a is denoted by $gcd(x, y)$ and is called the *greates common divisor* of x and y. Since $x|lcm(x, y)$ we write $x = lcm(x, y) \cdot c$ for some c. We substitute this in equation $x \cdot y = gcd(x, y) \cdot lcm(x, y)$ and get $x \cdot y = gcd(x, y) \cdot x \cdot c$; therefore $y = gcd(x, y) \cdot c$. We see that $gcd(x, y)|y$ and that a similar proof exists for $gcd(x, y)|x$.

To confirm that $gcd(x, y)$ is, in fact, the *greatest* common divisor, we argue the following. Write $gcd(x, y)$ as g and $lcm(x, y)$ as l so that by definition, $x \cdot y = g \cdot l$ and $g|x \wedge g|y$. Assume that there is a $d > g$ such that $d|x \wedge d|y$ to derive a contradiction. Then $x = d \cdot x'$ and $y = d \cdot y'$ for certain x' and y'. The following holds.

$$x \cdot d \cdot y' = x \cdot y = x' \cdot d \cdot y$$

Therefore $x \cdot y' = x' \cdot y$ and we see that both $x|(x' \cdot y)$ and $y|(x' \cdot y)$ hold. Then $(x' \cdot y)$ is a common multiple of x and y. As we defined $g|x$, there is a w such that $x = g \cdot w$. We have $x = d \cdot x'$ as well and since $d > g$, it must be the case that $x' < w$. We have

$$g \cdot l = x \cdot y = w \cdot g \cdot y$$

from which it follows that $l = w \cdot y$. Then $x' \cdot y < w \cdot y = l$. So $x' \cdot y$ is a common multiple smaller than l, which contradicts the definition of $lcm(x, y)$. So it cannot be the case that there is a common divisor d of x and y greater than $gcd(x, y)$.

Theorem 2.7 (Bézout's theorem for PA)**.**

$$\text{PA} \vdash \forall xy \geq 1 \exists a \leq y, b \leq x(\text{gcd}(x, y) = a \cdot x - b \cdot y)$$

Proof. We prove the theorem by induction on x. For $x = 1$ we take $a = 1$ and $b = 0$. Clearly $gcd(x, y) = 1$ for all y and the equality holds.

Next we consider the case of $x > 1$. We write $y = u \cdot x + v$ for certain u, v. Dividing both sides by $gcd(x, y)$ gives $y' = u \cdot x' + v'$ where $y = y' \cdot gcd(x, y)$ and $x = x' \cdot gcd(x, y)$. We have $gcd(x', v') = 1$. By induction, we assume the identity

holds for x' and v' and prove it holds for x and y. By assumption then, there exist $c \leq v'$ and $d \leq x'$ such that

$$1 = c \cdot x' - d \cdot v'$$

We multiply both sides with $gcd(x, y)$ and get $gcd(x, y) = c \cdot x - d \cdot v$. By definition, $v = y - u \cdot x$, so we have $gcd(x, y) = c \cdot x - d \cdot (u - u \cdot v)$, therefore

$$gcd(x, y) = (c + d \cdot u) \cdot x - d \cdot y$$

Using this equality, we will construct $a \leq y$ and $b \leq x$ such that $gcd(x, y) = a \cdot x - b \cdot y$. Take the term $(c + d \cdot u)$ and write $(c + d \cdot u) = z \cdot y + a$ for certain z and $0 \leq a < y$. We have

$$\begin{aligned} gcd(x, y) &= (c + d \cdot u) \cdot x - d \cdot y \\ &= (z \cdot y + a) \cdot x - d \cdot y \\ &= a \cdot x + z \cdot y \cdot x - d \cdot y \\ &= a \cdot x - (d - z \cdot x) \cdot y \end{aligned}$$

We write $b := (d - z \cdot x)$ and see that $b \cdot y \leq a \cdot x < x \cdot y$ from which it follows that $b < x$. We have $gcd(x, y) = a \cdot x - b \cdot y$ as desired and our induction proof is completed. ■

2.2 Coding sequences using Gödel's β-function

We introduce another abbreviation: $rm(x, y)$ denotes the remainder of x when dividing by y. Finally we say that numbers x and y are *coprime* if $gcd(x, y) = 1$. Next we present a version of the Chinese Remainder theorem which we will thereafter use to code sequences of numbers in PA.

Theorem 2.8 (Chinese Remainder Theorem)**.** *Given a sequence of numbers $x_0, \ldots, x_k$ and pairwise coprime positive integers $n_0, \ldots, n_k$. If for all $0 \leq i \leq k$ it is the case that $x_i < n_i$, then there exists an integer a such that:*

$$x_i = rm(a, n_i)$$

for all $0 \leq i \leq k$.

Proof. For each $0 \leq i \leq k$, we define N_i as the product of all n_j, $0 \leq j \leq k$ except for n_i. Note that N_i and n_i are coprime as well, so $gcd(N_i, n_i) = 1$ and there are c_i, d_i such that $c_i \cdot N_i - d_i \cdot n_i = 1$ by Bézout's theorem (theorem 2.7).

We construct $a = \sum_{j=0}^{k} x_j \cdot c_j \cdot N_j$ and consider the remainder of dividing by n_i. For all $0 \leq j \leq k$, if $i \neq j$, each term $x_j \cdot c_j \cdot N_j$ is divisible by n_i so there is no

remainder. That leaves the remainder of the term $x_i \cdot c_i \cdot N_i$. Substitution gives $x_i \cdot (d_i \cdot n_i + 1) = x_i \cdot d_i \cdot n_i + x_i$. Since $x_i < n_i$, dividing the term by n_i clearly leaves remainder x_i. We see that for all $0 \leq i \leq k$, $x_i = rm(a, n_i)$ as desired. ■

Gödel showed that these are the tools we need to code sequences of numbers in PA. Given a sequence of natural numbers $x_0, \ldots, x_{k-1}$, using only multiplication and addition, we can construct a pair (a, m) that code the sequence by Gödel's three-place β-function:

$$\beta(a, m, i) = x_i \text{ for all } 0 \leq i < k.$$

First we need to construct a sequence of k pairwise coprime numbers. Let $m = max(x_0, \ldots, x_{k-1}, k)!$ (note that the max-function can be expressed in $\mathcal{L}_{PA}$). We show that for all $0 \leq j < i < k$ the numbers $m(i+1)+1$ and $m(j+1)+1$ are coprime. Assume the opposite, there being a prime p that divides both numbers, to reach a contradiction. If p divides both numbers, it divides their difference, $(m(i+1)+1) - (m(j+1)+1) = m(i-j)$, as well. So $p|m(i-j)$ and since p is prime, either $p|m$ or $p|(i-j)$. We notice that $k!|m$ and since $0 < i-j < k$ $(i-j)|k!$, so $(i-j)|m$. Therefore if $p|(i-j)$ we have $p|m$ as well. So in both cases $p|m$, which leads to $p|m(i+1)$. By assumption we have $p|m(i+1)+1$ as well so p must divide 1 as well, which contradicts p being prime. So for all $0 \leq j < i < k$, the numbers $m(i+1)+1$ and $m(j+1)+1$ are coprime. Since $x_i < m(i+1)+1$ for all i, we apply the Chinese Remainder Theorem to find an a such that for all $0 \leq i < k$, $x_i = rm(a, m(i+1)+1)$. We define our coding function by

$$\beta(a, m, i) = rm(a, m(i+1)+1.$$

Theorem 2.9. *The following properties hold when coding sequences in* PA*:*

1. PA $\vdash \forall x \exists am(\beta(a, m, 0) = x)$
2. PA $\vdash \forall yxam \exists bn(\forall i < y(\beta(a, m, i) = \beta(b, n, i)) \wedge \beta(b, n, y) = x)$
3. PA $\vdash \forall ami(\beta(a, m, i) \leq a)$

Proof. See theorem 4.9 in ([**3**]). ■

The properties stated in theorem 2.9 of the β-function show that for any x, there is a sequence starting with x and that any sequence can be extended. Gödel used these properties to represent and prove facts about all primitive recursive functions in PA, not just addition and multiplication. We will go over the proof in the next paragraph, but the idea of the proof might be clear already. For primitive recursive functions, the function value of a number depends on the

function value of the previous number. Therefore, by considering the function values to be a sequence of numbers that each determine the next, we are able to express primitive recursive functions in PA.

2.3 What PA can prove

Our focus for the remaining of this chapter is to explore further what can be proven in PA. We will prove the important theorem of the Σ_1-completeness of PA that we will use later on. First we will introduce an important definition.

Definition 2.10. We speak of *bounded quantifiers* when they are of the form $\forall x < y$ or $\exists x < y$ for a term y not containing the variable x. We call other quantifiers *unbounded.* An $\mathcal{L}_{PA}$-formula φ is called a Δ_0-*formula* if all of its quantifiers are bounded (if it has quantifiers at all) and we may write $\varphi \in \Delta_0$.

The Σ_1-*formulas* are defined as the formulas of the form $\exists x\varphi$ where $\varphi \in \Delta_0$. Similarly, the Π_1-*formulas* are defined as the formulas of the form $\forall x\varphi$ where $\varphi \in \Delta_0$.

PA being Σ_1-complete means that a Σ_1-sentence is true in $\mathcal{N}$ if and only if it is provable in PA. To arrive at this theorem, we will first prove that PA correctly evaluates all terms. Next we will cover some properties and definitions of formulas and prove that PA is Δ_0-complete. Finally, we use this lemma to prove the Σ_1-completeness of PA. First we mention some properties of formulas.

Lemma 2.11. *For all natural numbers n and m, the following properties hold:*

(i) $\text{PA} \vdash \overline{n} + \overline{m} = \overline{n+m}$

(ii) $\text{PA} \vdash \overline{n} \cdot \overline{m} = \overline{n \cdot m}$

(iii) $\text{PA} \vdash \overline{n} < \overline{m} \Leftrightarrow n < m$

(iv) $\text{PA} \vdash \forall x(x < \overline{n} \leftrightarrow x = \overline{0} \vee \ldots \vee x = \overline{n-1})$ *if* $n > 0$

Proof. (i) We prove $\text{PA} \vdash \overline{n} + \overline{m} = \overline{n+m}$ by induction on n. For $n = 0$ we have to show that $\text{PA} \vdash \overline{0} + \overline{m} = \overline{m}$, which follows from the axioms. Assume we have a proof of *(i)* for a number n and proceed in PA. We have $\overline{n+1} + \overline{m} = \overline{n} + 1 + \overline{m} = \overline{n+m} + 1 = \overline{n+1+m}$ by associativity, so the property holds for $n+1$ whenever it holds for n. The proof of *(ii)* is similar.

For *(iii)*, we suppose first that $\text{PA} \vdash \overline{n} < \overline{m}$ so $\text{PA} \vdash \exists z(\overline{n} + (z+1) = \overline{m})$. It follows that $\mathcal{N} \models (n + (z+1) = m)$ which means that $n < m$ holds. Next suppose that $n < m$, then $n + (z+1) = m$ for some $z \in \mathbb{N}$. By *(i)* we have $\text{PA} \vdash \overline{n} + \overline{z+1} = \overline{m}$ so $\text{PA} \vdash \overline{n} + (\overline{z} + 1) = \overline{m}$. This is abbreviated with $\text{PA} \vdash \overline{n} < \overline{m}$.

(iv) We prove the property by induction on n. For $n = 1$ we have to prove PA $\vdash \forall x(x < \overline{1} \leftrightarrow x = \overline{0})$. We reason in PA and let $x < \overline{1}$. It follows that either $x + 1 = \overline{1}$ or $x + 1 < \overline{1}$, though the latter is clearly impossible. But if we have that $x + 1 = \overline{1}$, it follows that $x = \overline{0}$. Now assume $x = \overline{0}$, trivially we have that $x < \overline{1}$, so the property holds for $n = 1$. Next we assume the property holds for n so PA $\vdash \forall x(x < \overline{n} \leftrightarrow x = \overline{0} \vee \ldots \vee x = \overline{n-1})$. To prove it holds for $n + 1$ we first let $x < \overline{n+1}$. Either $x + 1 < \overline{n} + 1$ or $x + 1 = \overline{n} + 1$. The first case leads to $x < \overline{n}$, which by assumption leads to $x = \overline{0} \vee \ldots \vee x = \overline{n-1}$. The second case is equivalent to $x = \overline{n}$. In both cases, it follows that $x = \overline{0} \vee \ldots \vee x = \overline{n-1} \vee x = \overline{n}$. Next we assume the opposite. If $x = \overline{n}$ it follows that $x < \overline{n+1}$; if $x = \overline{0}$, ..., $x = \overline{n-1}$, it follows from our assumption that $x < \overline{n}$ and $x < \overline{n+1}$. Therefore, if $x = \overline{0} \vee \ldots \vee x = \overline{n}$ then $x < \overline{n+1}$ and the property is proven for $n + 1$ if it holds for n. By induction, the property holds for all $n > 0$. ■

As we have just seen, equations correctly evaluating any term of the form $\overline{n} + \overline{m}$ or $\overline{n} \cdot \overline{m}$ can be proved in PA. Since closed terms are constants or the result of repeated applications of the addition or multiplication functions on constants, it is easy to see that PA can evaluate *any* closed term, regardless of its complexity. This is stated in the following lemma.

Lemma 2.12. *If τ is a closed term of $\mathcal{L}_{PA}$ that takes the value t on interpretation in the standard model, then*

$$\text{PA} \vdash \tau = \bar{t}$$

Next we consider some useful properties of formulas.

Lemma 2.13. *For every formula φ, there is:*

- *an equivalent formula φ', called the* negation normal form*, such that each occurrence of the negation symbol only applies to atomic formulas;*
- *an equivalent formula φ'', called the* prenex normal form*, starting with a string of quantifiers, followed by a formula in which no quantifiers occur.*

Proof. We refer the reader to exercise 1.42 and lemma 2.29 of ([**2**]). ■

Example 2.14. The following equivalencies hold for any formulas φ and ψ. They are examples of formulas being written in negation normal form (as in the first

three items) and prenex normal forms (as in the final three items).

$$\begin{aligned}
\neg(\varphi \wedge \psi) &\leftrightarrow \neg\varphi \vee \neg\psi \\
\neg(\varphi \rightarrow \psi) &\leftrightarrow \varphi \wedge \neg\psi \\
\neg\exists x\varphi(x) &\leftrightarrow \forall x\neg\varphi(x) \\
(\forall x\varphi(x)) \rightarrow \psi &\leftrightarrow \exists x(\varphi(x) \rightarrow \psi) \\
\varphi \rightarrow (\forall x\psi(x)) &\leftrightarrow \forall x(\varphi \rightarrow \psi(x))
\end{aligned}$$

Finally, we are able to prove Δ_0- completeness of PA.

Lemma 2.15. *Let $\varphi(x_1, \ldots, x_k)$ be a Δ_0-sentence, then for all $n_1, \ldots, n_k \in \mathbb{N}$:*

$$PA \vdash \varphi(\overline{n_1}, \ldots, \overline{n_k}) \Leftrightarrow \mathcal{N} \models \varphi(n_1, \ldots, n_k).$$

Proof. The proof of the implication from left to right is trivial. Since $\mathcal{N}$ is a model for PA, it follows that if PA $\vdash (\overline{n_1}, \ldots, \overline{n_k})$ then $\mathcal{N} \models \varphi(n_1, \ldots, n_k)$.

For the implication from right to left, assume $\mathcal{N} \models \varphi(n_1, \ldots, n_k)$. We use induction for the number of bounded quantifiers. For a Δ_0-sentence φ with no bounded quantifiers, we use induction as well on the number k of binary connectives that are in φ.

For $k = 0$, φ is of the form $\sigma = \tau$ or $\sigma \neq \tau$. For the latter, it suffices to consider $\sigma < \tau$. As each closed term of $\mathcal{L}_{PA}$ is interpreted by $\mathcal{N}$ as a natural number, $\sigma^{\mathcal{N}}$ and $\tau^{\mathcal{N}}$ are equivalent to certain s and t respectively in $\mathcal{N}$. Since $\mathcal{N} \models \varphi$, either $\sigma^{\mathcal{N}} = \tau^{\mathcal{N}}$ or $\sigma^{\mathcal{N}} < \tau^{\mathcal{N}}$ is true in $\mathcal{N}$. So either $s = t$ or $s < t$ for certain $s, t \in \mathcal{N}$. By lemma 2.11, closed terms are evaluated correctly in PA. So in the first case, PA $\vdash \sigma = \overline{s} = \overline{t} = \tau$. In the case of $s < t$, there is a natural number z such that $s + (z + 1) = t$ so PA $\vdash \exists y(\overline{s} + (y + 1) = \overline{t})$ which is equivalent to PA $\vdash \overline{s} < \overline{t}$ so PA $\vdash \sigma < \tau$. In both cases, PA $\vdash \varphi$.

Now we assume the induction hypothesis that $\mathcal{N} \models \varphi \Rightarrow$ PA $\vdash \varphi$ for a quantifier-free sentence φ having k binary connectives. Furthermore, we assume $\mathcal{N} \models \psi$ with ψ having $k + 1$ binary connectives. We can bring ψ into its equivalent negation normal form ψ' without changing the number of binary connectives. Then $\mathcal{N} \models \psi'$ as well. Moreover, ψ' is of the form $(\psi_1 \wedge \psi_2)$, $(\psi_1 \vee \psi_2)$ or $(\psi_1 \rightarrow \psi_2)$ where ψ_1 and ψ_2 are both formulas with at most k binary connectives.

Considering the first case, $\mathcal{N} \models \psi_1 \wedge \psi_2$ from which it follows that both $\mathcal{N} \models \psi_1$ and $\mathcal{N} \models \psi_2$. By our induction hypothesis, PA $\vdash \psi_1$ and PA $\vdash \psi_2$ thus PA $\vdash \psi_1 \wedge \psi_2$. Similarly for the second case, we have that $\mathcal{N} \models \psi_1 \vee \psi_2$. This means that $\mathcal{N} \models \psi_1$ or $\mathcal{N} \models \psi_2$ which leads, by assumption, to PA $\vdash \psi_1$ or PA $\vdash \psi_2$. Therefore PA $\vdash \psi_1 \vee \psi_2$, as desired. Finally, in the case of $\mathcal{N} \models \psi_1 \rightarrow \psi_2$, we

consider the cases $\mathcal{N} \models \psi_2$ and $\mathcal{N} \models \neg\psi_2$. In the first case we have that PA $\vdash \psi_2$ from which we may conclude PA $\vdash \psi_1 \to \psi_2$. In the case of $\mathcal{N} \models \neg\psi_2$, we know that $\mathcal{N} \models \neg\psi_1$ so PA $\vdash \neg\psi_1$. Then PA $\vdash \psi_1 \to \psi_2$ as well. In all cases we may conclude PA $\vdash \psi'$ from $\mathcal{N} \models \psi$, so the lemma holds for quantifier free Δ_0-formulas.

We assume $\mathcal{N} \models \varphi \Rightarrow$ PA $\vdash \varphi$ for all Δ_0-sentences φ containing j bounded quantifiers. Moreover, we assume $\mathcal{N} \models \psi$ with $\psi \in \Delta_0$ containing $j+1$ bounded quantifiers. We bring ψ in its equivalent prenex normal form, ψ', which does not change its number of quantfiers. So $\mathcal{N} \models \psi'$. Then ψ' is equivalent to either $\forall x < t\chi(x)$ or $\exists x < t\chi(x)$ with χ being a Δ_0-sentence with j bounded quantifiers.

In the first case, $\mathcal{N} \models \forall x < t\chi(x)$. The interpretation of t in $\mathcal{N}$ is an integer, so it follows that $\chi(0), \chi(1), \ldots, \chi(t-1)$ are all true in the standard model. It follows that $\chi(0) \wedge \ldots \wedge \chi(t-1)$ is true as well. Since χ has j bounded quantifiers, it follows from the induction hypothesis that PA $\vdash \chi(\overline{0}) \wedge \ldots \wedge \chi(\overline{t^{\mathcal{N}} - 1})$. Therefore PA $\vdash \forall x < \overline{t^{\mathcal{N}}}\chi(x)$ by item *(iv)* of lemma 2.10 and we may conclude PA $\vdash \psi$.

The case of $\mathcal{N} \models \exists x < t\chi(x)$ is approached similarly. We notice that in this case at least one of the Δ_0-sentences $\chi(0), \ldots, \chi(t-1)$ is true in the standard model where each sentence has j bounded quantifiers. Therefore $\chi(0) \vee \ldots \vee \chi(t-1)$ is true as well. So PA $\vdash \chi(\overline{0}) \vee \ldots \vee \chi(\overline{t^{\mathcal{N}} - 1})$, PA $\vdash \exists x < \overline{t^{\mathcal{N}}}\chi(x)$ and PA $\vdash \psi$ hold.

Finally, by induction on both the number of bounded quantifiers and the number of binary logical connectives, we have proven that for all Δ_0-sentences φ,

$$\mathcal{N} \models \varphi \Rightarrow \text{ PA} \vdash \varphi.$$

■

Theorem 2.16 (Σ_1-completeness)**.** *Let $\varphi(x_1, \ldots, x_k)$ be a Σ_1-sentence. Then for all $n_1, \ldots, n_k \in \mathcal{N}$:*

$$PA \vdash \varphi(\overline{n_1}, \ldots, \overline{n_k}) \Leftrightarrow \mathcal{N} \models \varphi(n_1, \ldots, n_k).$$

Proof. Similar to the proof of Δ_0-completeness, the implication from left to right is trivial. To complete the proof, we assume $\mathcal{N} \models \varphi(n_1, \ldots, n_k)$ and derive that PA $\vdash \varphi(\overline{n_1}, \ldots, \overline{n_k})$. Since φ is a Σ_1 formula, it can be written as $\exists x_1, \ldots, x_m \psi(n_1, \ldots, n_k, x_1 \ldots, x_m)$ for a Δ_0-formula ψ. Since $\exists x_1, \ldots, x_m \psi(n_1, \ldots, n_k, x_1 \ldots, x_m)$ is true by assumption, for certain $y_1, \ldots, y_m \in \mathcal{N}$, we know that $\psi(n_1, \ldots, n_k, y_1, \ldots, y_m)$ is true. For $\psi \in \Delta_0$ we may apply lemma 2.15 and conclude that PA $\vdash \psi(\overline{n_1}, \ldots, \overline{n_k}, \overline{y_1}, \ldots, \overline{y_m})$ for certain $y_1, \ldots, y_m$, therefore PA $\vdash \exists x_1, \ldots, x_m \psi(\overline{n_1}, \ldots, \overline{n_k}, x_1, \ldots, x_m)$. Finally we conclude that PA $\vdash \varphi(\overline{n_1}, \ldots, \overline{n_k})$. ■

Remark 2.17. Naturally each theory containing PA is Σ_1-complete.

3 Primitive Recursive Functions

This chapter will describe a large class of numerical functions: *primitive recursive functions*. They are named after the process of recursion, which allows the value of a function to be defined from other values of the same function. We will prove how these functions can be represented in PA. The importance of these functions will become clear in the next chapter, as some important properties of formulas and proofs that we need to prove Gödel's theorems are decided by primitive recursive functions.

Before defining primitive recursive, we will go over a convention of notation we will often employ in this paper. It is not uncommon to encounter expressions such as x^y without a specification of this expression referring to a function or a variable.

The expression x^y could refer to any one of $(x, y) \mapsto x^y$, $(y, x) \mapsto x^y$ or $(x, y, z) \mapsto x^y$, or to a different function altogether. Usually, the context of a text will implicate which function is referred to. However, for the purpose of this paper it is important to clearly and explicitly distinguish between these different meanings. We will use the λ-notation to avoid this ambiguity: the function $(x_1, \ldots, x_n) \mapsto F(x_1, \ldots, x_n)$ will be denoted by $\lambda x_1 \ldots x_n.F(x_1, \ldots, x_n)$.

Example 3.1. The functions $(x, y) \mapsto x^y$, $(y, x) \mapsto x^y$ and $(x, y, z) \mapsto x^y$ will be written as $\lambda xy.x^y$, $\lambda yx.x^y$ and $\lambda xyz.x^y$ respectively.

3.1 Primitive recursive functions

Definition 3.2. The class of *primitive recursive* functions is a subclass of all numerical functions $\mathbb{N}^k \to \mathbb{N}$ (for $k \in \mathbb{N}$), generated by the following clauses:

1. the *zero function* $Z = \lambda x.0$ is primitive recursive;
2. the *successor function* $S = \lambda x.x + 1$ is primitive recursive;
3. the *projections* $\Pi_i^k = \lambda x_1, \ldots, x_k.x_i$ (for $1 \leq i \leq k$) are primitive recursive;
4. for primitive recursive functions $G_1, \ldots G_l : \mathbb{N}^k \mapsto \mathbb{N}$ and $H : \mathbb{N}^l \mapsto \mathbb{N}$, the function
$$\lambda \vec{x}.H(G_1(\vec{x}), \ldots, G_l(\vec{x}))$$
is primtive recursive as well. Furthermore, it is said to be defined from $G_1, \ldots, G_l$ and H by *composition*;
5. for primitive recursive functions $G : \mathbb{N}^k \mapsto \mathbb{N}$ and $H : \mathbb{N}^{k+2} \mapsto \mathbb{N}$, the primitive recursive function $F : \mathbb{N}^{k+1} \to \mathbb{N}$ may be defined from G and H

by *primitive recursion* as follows:

$$F(0, \vec{x}) = G(\vec{x})$$
$$F(y+1, \vec{x}) = H(y, x, F(y, \vec{x}))$$

Example 3.3. $\lambda xy.x + y$. We define a function G by $\Pi^1_1(y)$ which is known to be primitive recursive. Next, we define H by the following composition of the successor and projection function $\lambda uwv.S(\Pi^3_3(u, w, v))$. Then we have $F(0, y) = G(y) = \Pi^1_1(y) = y = 0 + y$, and $F(x + 1, y) = H(x, y, F(x, y)) = H(x, y, x + y) = S(x + y) = (x + y) + 1$. Since F is defined by primitive recursion, it is primitive recursive.

Example 3.4. $\lambda xy.xy$. We take G to be the zero function. Using the result of example 1.3 we know that the addition function is primitive recursive as well. By composition, we define H as $\lambda uwv.\Pi^3_3(u, w, v) + \Pi^3_2(u, w, v)$. Then F is primitive recursive, for $F(0, y) = G(y) = 0 = 0y$ and $F(x + 1, y) = H(x, y, F(x, y)) = H(x, y, xy) = xy + y = (x + 1)y$.

Example 3.5. Consider the *sign function*, sg, and its compliment, $\overline{sg}$:

$$sg(x) = \begin{cases} 1 & x > 0 \\ 0 & \text{else} \end{cases} \qquad \overline{sg}(x) = \begin{cases} 0 & x > 0 \\ 1 & \text{else} \end{cases}$$

The sign function is defined from primitive recursion. Since $sg(0) = 0$ and $sg(x + 1) = 1$, it follows that sg is primitive recursive. A similar proof can be given to show that $\overline{sg}$ is primitive recursive.

Definition 3.6. A *k-ary relation* is a subset of $\mathbb{N}^k$. For its *characteristic function* we will use the following convention. A relation R has a characteristic function $\chi_R : \mathbb{N}^k \to \mathbb{N}$ such that

$$\chi_R(\vec{x}) = \begin{cases} 0 & \vec{x} \in R \\ 1 & \text{else} \end{cases}$$

If the characteristic function of a relation is primitive recursive, we speak of a *primitive recursive relation.*

Proposition 3.7. *If G_1, G_2 and H are primitive recursive functions $\mathbb{N}^k \to \mathbb{N}$ and F is defined by*

$$F(\vec{x}) = \begin{cases} G_1(\vec{x}) & H(\vec{x}) = 0 \\ G_2(\vec{x}) & \textit{else} \end{cases}$$

then F is a primitive recursive function.

Proof. F may be defined by $F(\vec{x}) = \overline{sg}(H(\vec{x}))G_1(\vec{x}) + sg(H(\vec{x}))G_2(\vec{x})$. ■

3.2 Representing primitive recursive functions in PA

This chapter will be concluded by showing that facts about primitive recursive functions and relations can be proven in Peano Arithmetic. Considering the primitive recursive functions constitute such a large class of functions, Peano arithmetic turns out to be a very strong theory, despite its seemingly rudimentary axioms. First we go over some general definitions.

Definition 3.8. Given a theory Γ in the language $\mathcal{L}$, a $\mathcal{L}$-formula $\varphi(x)$ *represents* the numerical property P iff, for all $n \in \mathbb{N}$,

$$\begin{aligned}&\text{if } n \text{ has the property } P, \text{ then } \Gamma \vdash \varphi(\overline{n})\\ &\text{if } n \text{ does not have the property } P, \text{ then } \Gamma \vdash \neg\varphi(\overline{n})\end{aligned}$$

The $\mathcal{L}$-formula $\varphi(x_1, \ldots, x_k)$ of k free variables *represents* k-ary relation R iff, for all $n_1, \ldots, n_k \in \mathbb{N}$,

$$\begin{aligned}&(n_1, \ldots, n_k) \in R \Rightarrow \Gamma \vdash \varphi(\overline{n_1}, \ldots, \overline{n_k}) \quad \text{and}\\ &(n_1, \ldots, n_k) \notin R \Rightarrow \Gamma \vdash \neg\varphi(\overline{n_1}, \ldots, \overline{n_k})\end{aligned}$$

The $\mathcal{L}$-formula $\varphi(x_1, \ldots, x_k, y)$ of $k+1$ free variables *represents* k-ary function $F : \mathbb{N}^k \to \mathbb{N}$ iff, for all $n_1, \ldots, n_k \in \mathbb{N}$,

$$\begin{aligned}&\Gamma \vdash \varphi(\overline{n_1}, \ldots, \overline{n_k}, \overline{F(n_1, \ldots, n_k)}) \quad \text{and}\\ &\Gamma \vdash \exists! y \varphi(\overline{n_1}, \ldots, \overline{n_k}, y)\end{aligned}$$

Definition 3.9. A k-ary function $F : \mathbb{N}^k \to \mathbb{N}$ is *provably recursive* in PA if it is represented by a Σ_1-formula $\varphi(x_1, \ldots, x_k, y)$ such that

$$\text{PA} \vdash \forall x_1 \ldots x_k \exists! y \varphi(x_1, \ldots, x_k, y)$$

Theorem 3.10. *All primitive recursive functions are provably recursive in* PA*.*

Proof. Given a primitive recursive function $F : \mathbb{N}^k \to \mathbb{N}$, we have to prove that there is a Σ_1-formula $\varphi(x_1, \ldots, x_k, y)$ such that for all $n_1, \ldots, n_k \in \mathbb{N}$:

$(i)\, \text{PA} \vdash \varphi(\overline{n_1}, \ldots, \overline{n_k}, \overline{F(n_1, \ldots, n_k)})$ and $(ii)\, \text{PA} \vdash \forall x_1 \ldots x_k \exists! y \varphi(x_1, \ldots, x_k, y)$

The class of primitive recursive functions is generated inductively so we shall prove the theorem using induction accordingly. The formulas for the three initial functions are quite straightforward. For the zero function we can define $\varphi_Z(x, y)$ as $(y = 0)$; for the successor function we can define $\varphi_S(x, y)$ as $(y = x + 1)$; for

the projection function we can define $\varphi_{\Pi_i^k}(x_1, \ldots, x_k)$ as $(y = x_i)$. These are all Δ_0-formulas and therefore Σ_1-formulas as well.

The proofs that these formulas satisfy the properties are fairly uncomplicated so we will only explicitly prove the theorem for the zero function. The other initial functions are handled similarly. Let $n \in \mathbb{N}$ arbitrarily. We have $Z(n) = 0$ and since PA proves that $0 = 0$, PA $\vdash \varphi_Z(\overline{n}, 0)$. So it follows that PA $\vdash \varphi_Z(\overline{n}, \overline{Z(n)})$. Now reasoning within PA, we notice that for any x, there is a unique y such that $y = 0$ holds. Therefore PA $\vdash \forall x \exists! y \varphi_Z(x, y)$ which concludes the proof of both properties.

Next we suppose that $F : \mathbb{N}^k \to \mathbb{N}$ is defined by composition of primitive recursive functions $H : \mathbb{N}^m \to \mathbb{N}$ and $G_1, \ldots, G_m : \mathbb{N}^k \to \mathbb{N}$. So we have $F(\vec{x}) = H(G_1(\vec{x}), \ldots, G_m(\vec{x}))$. By the induction hypothesis we have Σ_1-formulas $\varphi_H, \varphi_{G_1}, \ldots, \varphi_{G_m}$ that each satisfy the two properties. Then we define $\varphi_F(\vec{x}, y)$ as

$$\exists z_1 \ldots z_m (\varphi_{G_1}(\vec{x}, z_1) \wedge \ldots \wedge \varphi_{G_m}(\vec{x}, z_m) \wedge \varphi_H(z_1, \ldots, z_m, y)).$$

Clearly φ_F is a Σ_1-formula as well. Take $n_1, \ldots, n_k \in \mathbb{N}$ arbitrarily. We have $F(n_1, \ldots, n_k) = H(G_1(n_1, \ldots, n_k), \ldots, G_m(n_1, \ldots, n_k))$ thus there are $l_1, \ldots, l_m$ such that for all $1 \leq i \leq m$, $G_i(n_1, \ldots, n_k) = l_i$ and $H(l_1, \ldots, l_m) = F(n_1, \ldots, n_k)$.

By hypothesis then, PA $\vdash \varphi_{G_1}(\overline{n_1}, \ldots, \overline{n_k}, \overline{l_1})$, ..., PA $\vdash \varphi_{G_m}(\overline{n_1}, \ldots, \overline{n_k}, \overline{l_m})$ and PA $\vdash \varphi_H(\overline{l_1}, \ldots, \overline{l_m}, \overline{F(n_1, \ldots, n_k)})$. It follows that

$$\begin{aligned} \text{PA } \vdash \exists z_1 \ldots z_m (&\varphi_{G_1}(\overline{n_1}, \ldots, \overline{n_k}, z_1) \wedge \ldots \wedge \varphi_{G_m}(\overline{n_1}, \ldots, \overline{n_k}, z_m) \\ &\wedge \varphi_H(z_1, \ldots, z_m, \overline{F(n_1, \ldots, n_k)})) \end{aligned}$$

This satisfies that for all $n_1, \ldots, n_k \in \mathbb{N}$, PA $\vdash \varphi_F(\overline{n_1}, \ldots, \overline{n_k}, \overline{F(n_1, \ldots, n_k)})$ and property *(i)* is proven. Next we prove property *(ii)*. Clearly it follows from the assumption that PA proves that for all $\vec{x}$, there are unique $z_1, \ldots, z_m$ such that $\varphi_{G_1}(\vec{x}, z_1), \ldots \varphi_{G_m}(\vec{x}, z_m)$ hold in PA. Then for those $z_1, \ldots, z_m$ PA proves there is a unique y such that $\varphi_H(z_1, \ldots, z_m, y)$ holds in PA. From this we may conclude that

$$\text{PA } \vdash \forall x_1 \ldots x_m \exists! y \varphi_F(x_1, \ldots, x_k, y).$$

Indeed, the second property holds for F as well and F is provably recursive in PA.

Finally we prove the theorem for a primitive recursive function F that is defined by recursion from primitive recursive G and H. That is:

$$\begin{aligned} F(0, \vec{x}) &= G(\vec{x}) \\ F(z+1, \vec{x}) &= H(z, F(z, \vec{x}), x) \end{aligned}$$

By the induction hypothesis we have Σ_1-formulas $\varphi_G(\vec{x}, y)$ and $\varphi_H(z, u, \vec{x}, y)$ that satisfy the two properties. We define $\varphi_F(z, \vec{x}, y)$ by

$$\exists am(\varphi_G(\vec{x}, \beta(a, m, 0)) \wedge \forall i < z \varphi_H(i, \beta(a, m, i), \vec{x}, \beta(a, m, i+1)) \wedge y = \beta(a, m, z),$$

where the pair (a, m) encodes the sequence $F(0, \vec{x}), \ldots, F(z, \vec{x})$. While $\beta(a, m, i)$ itself is not a term of $\mathcal{L}_{PA}$, it is an abbreviation for something that can be written in $\mathcal{L}_{PA}$. For example, $\varphi_G(\vec{x}, \beta(a, m, 0))$ denotes

$$\exists k, l < a(a = k \cdot (m+1) + l \ \wedge \ 0 \leq l < m+1 \ \wedge \ \varphi_G(\vec{x}, l)).$$

Here we have used property 3 of lemma 2.9 to see that the remainder is indeed bounded. Considering this, as well as the assumption of φ_G and φ_H being Σ_1-formulas, it follows that φ_F is a Σ_1-formula. We will use induction on z to prove the formula satisfies the two properties as state in the theorem.

For $z = 0$ we notice that φ_F equates to $\exists am(\varphi_G(\vec{x}, \beta(a, m, 0)) \wedge y = \underline{\beta(a, m, 0)}$. By hypothesis we have that for all $n_1, \ldots, n_k \in \mathbb{N}$, $\text{PA} \vdash \varphi_G(\overline{n_1}, \ldots, \overline{n_k}, \underline{G(n_1, \ldots, n_k)})$. We know that $F(0, n_1, \ldots, n_k) = G(n_1, \ldots, n_k)$ and by lemma theorem 2.9 we know that there are a, m such that $\beta(a, m, 0) = \overline{G(n_1, \ldots, n_k)} = \overline{F(0, n_1, \ldots, n_k)}$. Naturally we have for all $n_1, \ldots, n_k \in \mathbb{N}$ that $\text{PA} \vdash \exists am(\varphi_G(\overline{n_1}, \ldots, \overline{n_k}, \beta(a, m, 0))$. Therefore $\text{PA} \vdash \varphi_F(0, \overline{n_1}, \ldots, \overline{n_k}, \overline{F(0, n_1, \ldots, n_k)})$ which proves the first property. For any $\vec{x}$, there is exactly one z such that $\varphi_G(\vec{x}, z)$ holds in PA. Likewise there is a unique y such that $\exists am(\varphi_G(\vec{x}, \beta(a, m, 0)) \wedge \beta(a, m, 0))$ from which the second property follows as well.

We assume furthermore that properties *(i)* and *(ii)* hold for $\varphi_F(z, \vec{x}, y)$. Let $n_1, \ldots, n_k \in \mathbb{N}$, then there are a, m such that

$$\begin{aligned}\text{PA} \ \vdash & \varphi_G(\overline{n_1}, \ldots, \overline{n_k}, \beta(a, m, 0)) \ \wedge \ \forall i < \overline{z} \varphi_H(i, \beta(a, m, i), \overline{n_1}, \ldots, \overline{n_k}, \beta(a, m, i+1)) \\ & \wedge \ \overline{F(z, n_1, \ldots, n_k)} = \beta(a, m, \overline{z}).\end{aligned}$$

By the induction hypothesis, there is a unique w such that $\varphi_H(\overline{z}, \beta(a, m, \overline{z}), \overline{n_1}, \ldots, \overline{n_k}, w)$ holds in PA. Then by the second property of theorem 2.9, we can find b, h that satisfy $\forall i < (\overline{z}+1)\beta(a, m, i) = \underline{\beta(b, h, i)}$ and $\beta(b, h, \overline{z}+1) = w$. We have that $\text{PA} \vdash \varphi_F(\overline{z+1}, \overline{n_1}, \ldots, \overline{n_k}, \underline{F(z+1, n_1, \ldots, n_k)})$, from which we conclude that property *(i)* holds for φ_F.

To prove the second property we reason within PA. Let $\vec{x}$ be given as well and by the induction hypothesis there is a unique y for which $\varphi_F(z, \vec{x}, y)$. By assumption, $\varphi_G(\vec{x}, \beta(a, m, 0))$ guarantees that $\beta(a, m, 0)$ is unique and that for all $i < z$, $\varphi_H(i, \beta(a, m, i), \vec{x}, \beta(a, m, i+1))$ makes $\beta(a, m, i+1)$ unique as well. Then $\beta(a, m, z+1)$ in $\varphi_H(z, \beta(a, m, z), \vec{x}, \beta(a, m, z+1))$ is unique. Therefore, a unique y exists which satisfies the following in PA for certain a, m:

$$\varphi_G(\vec{x}, \beta(a, m, 0)) \wedge \forall i < (z+1) \varphi_H(i, \beta(a, m, i), \vec{x}, \beta(a, m, i+1)) \wedge y = \beta(a, m, z+1).$$

Induction on z has proven that indeed

$$\text{PA} \vdash \forall z\, x_1 \ldots x_k \exists! y \varphi_F(z, x_1, \ldots, x_k, y).$$

We have concluded the proof that all primitive recursive functions are provably recursive in PA. ■

4 Gödel's First Incompleteness Theorem

In this chapter we shall prove the first incompleteness Theorem of Gödel. We will construct a $\mathcal{L}_{PA}$-sentence of which we can prove its independence of Peano Arithmetic.

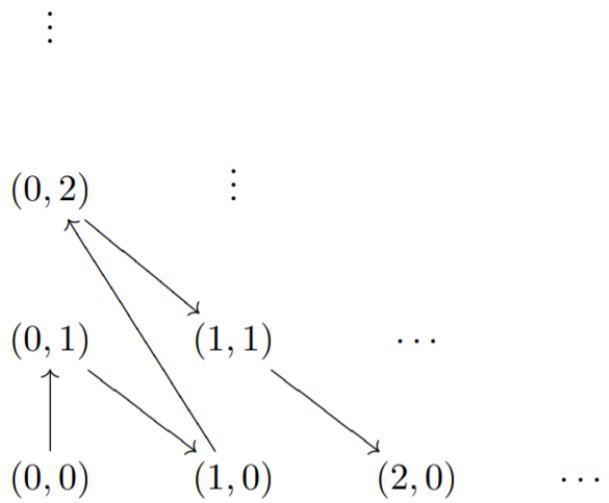
4.1 Coding formulas and proofs

The proof of the incompleteness theorems of Gödel rests PA being able to make statements *about* its own theorems. Gödel's idea was to code formulas and proofs, so that certain predicates can be interpreted as numerical relations, some of which being primitive recursive. And as we have proven, PA is able to prove properties of these primitive recursive relations. First we develop a way of coding formulas. We use the same coding as van Oosten does in chapters 3 and 5 of ([**3**]).

There are many bijections from $\mathbb{N} \times \mathbb{N} \to \mathbb{N}$. We call a function $f : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ a paring function and say that $f(x, y)$ codes the pair (x, y). We consider the following pairing function:

$$j(n, m) = \frac{1}{2}(n + m)(n + m + 1) + n$$

Function j enumerates elements of $\mathbb{N} \times \mathbb{N}$ as follows:



Since j is a composition of addition and multiplication, it is a primitive recursive function. Furthermore, j is a bijection, which means that there exists functions j_1 and j_2 such that

$$j(j_1(z), j_2(z)) = z.$$

Since functions j_1 and j_2 are primitive recursive, we can define bijections from $\mathbb{N}^n$ to $\mathbb{N}$ and projection functions from $\mathbb{N}$ to $\mathbb{N}$ in the following manner.

Definition 4.1. The bijections $j^n : \mathbb{N}^n \to \mathbb{N}$ are defined by:

$$j^1(x) = x$$
$$j^{n+1}(x_1, \ldots, x_{n+1}) = j(j^n(x_1, \ldots, x_n), x_{n+1})$$

There are *projection* functions $j_i^n : \mathbb{N} \to \mathbb{N}$ for $1 \leq i \leq n$, satisfying

$$j^n(j_1^n(z), \ldots, j_n^n(z)) = z$$

for all $z \in \mathbb{N}$, and are defined by:

$$j_1^1(z) = z$$
$$j_i^{n+1} = \begin{cases} j_i^n(j_1(z)) & \text{if } 1 \leq i \leq n \\ j_2(z) & \text{if } i = n+1 \end{cases}$$

Considering that j, j_1 and j_2 are all primitive recursive, it can be seen that the functions j^n and j_i^n are primitive recursive as well for all $n \in \mathbb{N}$ and $1 \leq i \leq n$.

We are able to code finite sequences using these functions. A *sequence* $(x_1, \ldots, x_n)$ is an element of $\mathbb{N}^n$; we consider the unique element $(-)$ of the set $\mathbb{N}^0$ as the *empty sequence*. We will define a function below that, given a sequence, returns the *code of that sequence*. We will use notation $\langle x_1, \ldots, x_n \rangle$ and $\langle\rangle$ to denote the codes of sequences $(x_1, \ldots, x_n)$ and $(-)$ respectively.

Definition 4.2.

$$\langle\rangle = 0$$
$$\langle x_0, \ldots, x_{n-1} = j(n-1, j^n(x_0, \ldots, x_{n-1})) + 1 \quad \text{if } m > 0$$

Finally we need to define some important operations of sequences. Given a sequence, we want to be able to take its length, take its i-th element and concatenate it with another sequence. Again, we have taken the following functions from van Oosten ([**3**]).

Definition 4.3. Consider two sequences that are coded by numbers x and y. Define function $\mathrm{lh}(x)$ that returns the *length* of sequence with code x by:

$$\mathrm{lh}(x) = \begin{cases} 0 & \text{if } x = 0 \\ j_1(s-1) + 1 & \text{if } x > 0 \end{cases}$$

Define functions $(x)_i$, returning the i-th element of the sequence coded by x as long as $0 \leq i \leq \mathrm{lh}(x)$, and 0 otherwise, by

$$(x)_i = \begin{cases} j_{i+1}^{\mathrm{lh}(x)}(j_2(x-1)) & \text{if } x > 0 \text{ and } 0 \leq i < \mathrm{lh}(x) \\ 0 & \text{else} \end{cases}$$

Define the *concatenation* function $x \frown y$ by

$$\langle\rangle \frown y = y$$
$$x \frown \langle\rangle = x$$
$$\langle (x)_0, \ldots, (x)_{\mathrm{lh}(x)-1} \rangle \frown \langle (y)_0, \ldots, (y)_{\mathrm{lh}(y)-1} \rangle = \langle (x)_0, \ldots, (x)_{\mathrm{lh}(x)-1}, (y)_0, \ldots, (y)_{\mathrm{lh}(y)-1} \rangle$$

The ability to code sequences is an essential tool for coding formulas and proofs. Given a formula φ, we will assign it a code number. The code number is written as $\ulcorner \varphi \urcorner$ and called the *Gödel number* of φ.

Each symbol of $\mathcal{L}_{PA}$ (from now on we include $<$) is assigned a code number:

0	1	v	$+$	$\cdot$	$=$	$<$	$\wedge$	$\vee$	$\rightarrow$	$\neg$	$\forall$	$\exists$
0	1	2	3	4	5	6	7	8	9	10	11	12

We assume that the variables of $\mathcal{L}_{PA}$ are numbered as $v_0, v_1, \ldots$. Then we code the terms of $\mathcal{L}_{PA}$ by recursion. So $\ulcorner 0 \urcorner = \langle 0 \rangle$, $\ulcorner 1 \urcorner = \langle 1 \rangle$ and $\ulcorner v_i \urcorner = \langle 2, i \rangle$. Given terms t and s we define $\ulcorner t + s \urcorner = \langle 3, \ulcorner t \urcorner, \ulcorner s \urcorner \rangle$ and $\ulcorner t \cdot s \urcorner = \langle 4, \ulcorner t \urcorner, \ulcorner s \urcorner \rangle$. Likewise we code formulas by recursion, examples being $\ulcorner t < s \urcorner = \langle 6, \ulcorner t \urcorner, \ulcorner s \urcorner \rangle$, $\ulcorner \varphi \vee \psi \urcorner = \langle 8, \ulcorner \varphi \urcorner, \ulcorner \psi \urcorner \rangle$, and $\ulcorner \exists v_i \varphi \urcorner = \langle 12, i, \ulcorner \varphi \urcorner \rangle$.

Immediately we have that properties such as "x is the code of a term", "the variable v_i occurs freely in the formula coded by x" and so forth, are primitive recursive in their arguments.

Likewise, we can recursively code proof trees. Similarly to how we assigned code numbers to $\mathcal{L}_{PA}$-symbols, we can do so as well for each inference rule of our deductive system (to see this worked out, we again refer to ([**3**]), this time to section 5.2). As a consequence, we are able to formulate many properties of proofs that are primitive recursive in their arguments as well. This includes the important property that may be translated in ordinary language to "y is the Gödel number of a correct proof in PA for the formula with Gödel number x". We denote this primitive recursive precidate by $\mathrm{Prf}(y, x)$.

4.2 Diagonalization

The proof of Gödel's first incompleteness theorem relies heavily on the diagonlization lemma. For this, we define the following primitive recursive *subsitution function*:

$$\mathrm{Sub}(x,y,i) = \begin{cases} \ulcorner \varphi[s/v_i] \urcorner & \text{if "}y\text{ codes a formula }\varphi\text{" and "}x\text{ codes a term }s\text{"} \\ 0 & \text{else} \end{cases}$$

Lemma 4.4 (Diagonalization Lemma)**.** *Given an $\mathcal{L}_{PA}$-formula φ with free variable v_0, there is a $\mathcal{L}_{PA}$-formula ψ with the same free variables as φ except for v_0, such that*

$$\mathrm{PA} \ \vdash \psi \leftrightarrow \varphi[\overline{\ulcorner \psi \urcorner}/v_0].$$

Proof. The function $\mathrm{Sub}(x,y,i)$ is a primitive recursive function which makes $\lambda xy.Sub(x,y,0)$ a primitive recursive function as well. By theorem 3.10, the function is provably recursive and there is a Σ_1-formula S that represents it in PA. Furthermore, we consider the function that, given natural number n, returns $\ulcorner \overline{n} \urcorner$, the Gödel number of $\overline{n}$ in PA. Let C be the Σ_1-formula representing this primitive recursive function.

By theorem 3.10, the following statements hold for S and C. For all $n, m \in \mathbb{N}$:

$$\mathrm{PA} \vdash S(\overline{n}, \overline{m}, \overline{\mathrm{Sub}(n,m,0)}) \tag{1}$$
$$\mathrm{PA} \vdash \forall xy \exists! z S(x,y,z) \tag{2}$$
$$\mathrm{PA} \vdash C(\overline{n}, \overline{\ulcorner \overline{n} \urcorner}) \tag{3}$$
$$\mathrm{PA} \vdash \forall x \exists! z C(x,z) \tag{4}$$

Now let φ be a formula that has free variable v_0. Furthermore, define a formula D by

$$\forall xy(C(v_0,x) \wedge S(x,v_0,y) \rightarrow \varphi[y/v_0])$$

and define ψ by

$$D[\overline{\ulcorner D \urcorner}/v_0] \ \equiv \ \forall xy(C(\overline{\ulcorner D \urcorner}, x) \wedge S(x, \overline{\ulcorner D \urcorner}, y) \rightarrow \varphi[y/v_0]).$$

We will prove that this formula satisfies the diagonalization lemma. We reason in PA and prove the implication from left to right first. So assume ψ holds. That is,

$$\forall xy(C(\overline{\ulcorner D \urcorner}, x) \wedge S(x, \overline{\ulcorner D \urcorner}, y) \rightarrow \varphi[y/v_0]).$$

Notice that $\mathrm{Sub}(\overline{\ulcorner \ulcorner D \urcorner \urcorner}, \overline{\ulcorner D \urcorner}, 0) = \ulcorner D[\overline{\ulcorner D \urcorner}/v_0] \urcorner$. It follows that

$$C(\overline{\ulcorner D \urcorner}, \overline{\ulcorner \ulcorner D \urcorner \urcorner})$$

and

$$S(\overline{\ulcorner \ulcorner D \urcorner \urcorner}, \overline{\ulcorner D \urcorner}, \overline{\ulcorner D[\overline{\ulcorner D \urcorner}/v_0] \urcorner})$$

hold in PA. It follows from ψ that we may infer $\varphi[\overline{\ulcorner D[\overline{\ulcorner D\urcorner}/v_0]\urcorner}/v_0]$ which, by definition of ψ, is equivalent to $\varphi[\overline{\ulcorner\psi\urcorner}/v_0]$. So the implication $\psi \to \varphi[\overline{\ulcorner\psi\urcorner}/v_0]$ holds in PA.

Next we assume that $\varphi[\overline{\ulcorner\psi\urcorner}/v_0]$ holds in PA and want to prove ψ. We have to consider all x and y such that

$$C(v_0, x) \wedge S(x, v_0, y)$$

and check whether they imply $\varphi[y/v_0]$. We know by theorem 3.10 that the formula is satisfied for $x = \overline{\ulcorner\overline{\ulcorner D\urcorner}\urcorner}$ and $y = \overline{\ulcorner D[\overline{\ulcorner D\urcorner}/v_0]} = \overline{\ulcorner\psi\urcorner}$. Moreover, these x and y are unique. By assumption we have $\varphi[\overline{\ulcorner\psi\urcorner}/v_0]$ so finally we conclude

$$\forall xy(C(\overline{\ulcorner D\urcorner}, x) \wedge S(x, \overline{\ulcorner D\urcorner}, y) \to \varphi[y/v_0]),$$

which is exactly how we defined ψ. Therefore, the implication holds from right to left as well in PA. ■

4.3 PA is incomplete

We have all the tools we need to prove the first incompleteness theorem. We will begin by proving that the system of Peano Arithmetic is incomplete. Shortly thereafter we will demonstrate the strength of Gödel's theorem by proving that any theory extending PA is incomplete. First we will state an important lemma.

Lemma 4.5. *The proof predicate* $\mathrm{Prf}(y, x)$ *that states that "y is the Gödel number of a proof in PA for the sentence with Gödel number x", is primitive recursive. It can therefore be represented by a* Σ_1*-formula in* PA*, denoted by* $\overline{\mathrm{Prf}}(y, x)$.

We then define our *provability* predicate as follows:

$$\mathrm{Prov}(x) \equiv \exists y\overline{\mathrm{Prf}}(y, x)$$

Notice that $\mathrm{Prov}(x)$ is a Σ_1-formula as well.

Proposition 4.6. *If* $\mathrm{PA} \vdash \varphi$*, then* $\mathrm{PA} \vdash \mathrm{Prov}(\overline{\ulcorner\varphi\urcorner})$

Proof. If $\mathrm{PA} \vdash \varphi$, then $\mathrm{Prov}(\ulcorner\varphi\urcorner)$ is a true Σ_1-formula. It follows from Σ_1-completeness that $\mathrm{PA} \vdash \mathrm{Prov}(\overline{\ulcorner\varphi\urcorner})$.

■

Theorem 4.7 (Gödel's First Incompleteness Theorem)**.** *There is an* $\mathcal{L}_{PA}$*-sentence G such that, if* PA *is consistent, it is independent of* PA.

Proof. We will construct a $\mathcal{L}_{PA}$-sentence G such that PA $\nvdash G$ and PA $\nvdash \neg G$.

We apply the diagonalization lemma to $\neg\text{Prov}(x)$ to construct a sentence G such that

$$\text{PA} \vdash G \leftrightarrow \neg\text{Prov}(\overline{\ulcorner G \urcorner}).$$

First we assume PA $\vdash G$ and derive a contradiction. The assumption leads to the equivalent PA $\vdash \neg\text{Prov}(\overline{\ulcorner G \urcorner})$. However, if PA $\vdash G$, it follows from proposition 4.5 that PA $\vdash \text{Prov}(\overline{\ulcorner G \urcorner})$. This contradicts the assumption that PA is consistent and we conclude that PA $\nvdash G$.

Next we assume PA $\vdash \neg G$. By the construction of G we have PA $\vdash \text{Prov}(\overline{\ulcorner G \urcorner})$. Since $\mathcal{N}$ is a model we have $\mathcal{N} \models \exists x \text{Prf}(x, \ulcorner G \urcorner)$. In other words, it is true (in the standard model) that there is a proof in PA of G. This is of course what is represented by PA $\vdash G$. We have reached a contradiction and conclude PA $\nvdash \neg G$ as well. ■

The sentence we have constructed translates in informal terms to "this sentence is unprovable in PA", and is called the *Gödel sentence*. Though G is independent of the system of Peano Arithmetic, we can still deduce that it is true in the standard model.

Proposition 4.8. *For the Gödel sentence G, $\mathcal{N} \models G$.*

Proof. Assume the contrary: $\neg G$ is true in $\mathcal{N}$. Since $\neg G$ is equivalent to the Σ_1-sentence $\exists y \overline{\text{Prf}}(y, \overline{\ulcorner G \urcorner})$ and PA is Σ_1-complete, it follows that $\neg G$ is provable in PA which contradicts that G is independent of PA. So G must be true in $\mathcal{N}$. ■

4.4 Generalizing the argument

Let us consider for a moment what we have proved so far. We have shown that if PA is consistent, it is incomplete. One might hope at this point that the theory may be "patched up" by adding the Gödel sentence to the axioms. The new theory *will* be able to prove the Gödel sentence; however, Gödel's argument may be taken even further and we will shortly hereafter see that the new theory remains incomplete. First we introduce a convenient definition.

Definition 4.9. A theory is called *arithmetically sound*, or simply called *sound*, if it is true in the standard model $\mathcal{N}$.

When constructing the new theory, $\text{PA}^+ = \text{PA} + \{G\}$, we immediately notice that it remains consistent and sound. However, for PA^+ we can define a new primitive recursive provability predicate to construct a new Gödel sentence G_+

that is independent of the theory. This argument is generalized in the following theorem.

Theorem 4.10. *Any axiomatized, consistent and sound theory T extending* PA *has a Gödel sentence G_T that is independent of T.*

Remark 4.11. We specify that a theory needs to be axiomatized since the provability predicate of the theory is not primitive recursive if it cannot be effectively decided whether a certain formula is among the axioms.

It follows that the truths of arithmetic cannot be effectively axiomatized by a formal theory.

4.4.1 The syntactic argument for incompleteness

In our proof of the first incompleteness theorem we have used the fact that PA is sound (and may be extended by sound theories). But if we define a different extension of PA, for example

$$\mathrm{PA}^\dagger = \mathrm{PA} + \neg G,$$

what can we say about $\mathrm{PA}^\dagger$? Since $\mathrm{PA} \nvdash G$, it is consistent; however, it is clearly not a true theory in the standard model. Gödel proved that the semantic condition of soundness is not required to establish incompleteness, though he had to introduce a different, somewhat inconvenient condition[1].

For our proof, we will not use Gödel's condition. Instead we turn to the important improvement of J. Barkley Rosser, who showed that an extension of PA merely needs to be consistent in order to prove that it is incomplete ([**4**], p. 185-190). We rely on the fact that consistent extensions of PA are Σ_1-complete and are therefore able to prove primitive recursive properties. We will construct a Rosser sentence R that informally says "if this sentence is provable, there exists a shorter proof of its negation". For a theory T, define a new provability predicate as follows:

$$\mathrm{RProv}_T(x) \equiv \exists y(\overline{\mathrm{Prf}_T}(y, x) \wedge \forall z < y(\neg\overline{\mathrm{Prf}_T}(z, \langle 10, x\rangle))).$$

Naturally we may apply the diagonalization lemma to consistent extensions T of PA. We may therefore construct the sentence R_T such that

$$T \vdash R_T \leftrightarrow \neg\mathrm{RProv}_T(\overline{\ulcorner R \urcorner})$$

We will prove that R_T is independent of T.

[1] Gödel introduced the concept of *ω-consistency* in his original proof. He showed that for an axiomatized theory T, if T satisfies the condition of ω-consistency, we may conclude $T \nvdash \neg G$ regardless of whether T's axioms are true. See section 4.2 of ([**5**]) for an overview of this proof.

Theorem 4.12 (The Gödel-Rosser Theorem)**.** *Any axiomatized, consistent extension T of* PA *is incomplete.*

Proof. We assume $T \vdash R_T$ and derive a contradiction. By consistency of T, $T \nvdash \neg R_T$ so it is the case that a proof of R_T exists in T and that there is no smaller proof of $\neg R_T$. Therefore, the sentence

$$\exists y(\overline{\mathrm{Prf}_T}(y, \overline{\ulcorner R_T \urcorner}) \wedge \forall z < y(\neg\overline{\mathrm{Prf}_T}(z, \overline{\ulcorner \neg R_T \urcorner})))$$

is a true Σ_1-sentence that is in fact equivalent to $\mathrm{RProv}_T(\overline{\ulcorner R_T \urcorner})$. Then by Σ_1-completeness, T proves the sentence. We have that $T \vdash \mathrm{RProv}_T(\overline{\ulcorner R_T \urcorner})$, equivalently $T \vdash \neg R_T$, which contradicts $T \nvdash \neg R_T$. It follows then that $T \nvdash R_T$.

Conversely we assume $T \vdash \neg R_T$. By consistency, $T \nvdash R_T$. Then the Σ_1-sentences $\overline{\mathrm{Prf}_T}(n, \ulcorner \neg R_T \urcorner)$ and $\forall m < n \neg\overline{\mathrm{Prf}_T}(m, \ulcorner R_T \urcorner)$ are true for some $n \in \mathbb{N}$. By Σ_1-completeness, $T \vdash \overline{\mathrm{Prf}_T}(\overline{n}, \overline{\ulcorner \neg R_T \urcorner})$ and $T \vdash \forall y < \overline{n} \neg\overline{\mathrm{Prf}_T}(y, \overline{\ulcorner R_T \urcorner})$. From this, it follows that

$$T \vdash \forall y(\overline{\mathrm{Prf}_T}(y, \overline{\ulcorner R_T \urcorner}) \to \exists z < y(\overline{\mathrm{Prf}_T}(z, \overline{\ulcorner \neg R_T \urcorner}))),$$

which is equivalent to $T \vdash \neg\mathrm{RProv}_T(\overline{\ulcorner R_T \urcorner})$. We have that $T \vdash R_T$ by definition, though this contradicts the consistency of T. Therefore $T \nvdash \neg R_T$. ∎

5 Gödel's Second Incompleteness Theorem

We now turn to the second incompleteness theorem, which, in informal terms, states that PA is unable to prove its own consistency.

To improve readability, we introduce some new notation: we will write $\Box_T\varphi$ to abbreviate $\mathrm{Prov}_T(\overline{\ulcorner\varphi\urcorner})$ for a certain theory T. We will often omit the subscript of the box symbol in practice as the context supplies it. To given an example, for the Gödel sentence of PA we have

$$\mathrm{PA} \vdash G \leftrightarrow \neg\Box G.$$

A consistent system is one that does not prove any contradiction. This leads us to the following notation.

Definition 5.1. Given theory T, we define the $\mathcal{L}_T$-statement $\mathrm{Con}(T)$ as

$$\mathrm{Con}(T) \equiv \neg\Box_T \bot$$

Definition 5.2. The proof of the second incompleteness theorem relies on some properties of the provability predicate, called the *derivability conditions:*

(1) If $T \vdash \varphi$, then $T \vdash \Box\varphi$;

(2) $T \vdash \Box\varphi \wedge \Box(\varphi \to \psi) \to \Box\psi$;

(3) $T \vdash \Box\varphi \to \Box\Box\varphi$

Proposition 5.3. *The derivability conditions hold for PA.*

Proof. We have proven the first condition already, in proposition 4.6 of the previous section.

For the second condition, we reason in PA. Suppose $\Box\varphi$ and $\Box(\varphi \to \psi)$; that is, there are x and y that code proofs of φ and $\varphi \to \psi$ respectively. We are able to combine these two proofs using $\to$-elimination and construct a new proof for ψ, coded by a certain z. Then $\overline{\mathrm{Prf}}(z, \overline{\ulcorner\psi\urcorner})$ holds, which is equivalent to $\Box\psi$.

Condition (3) is a consequence of a theorem more general, called *formalized* Σ_1-*completeness*. This theorem asserts that for any Σ_1-formula ψ, it is the case that $\mathrm{PA} \vdash \psi \to \Box\psi$. Since the provability predicate is Σ_1, we can apply the theorem to $\Box\varphi$ for any φ and the third derivability condition holds for PA. For a proof of formalized Σ_1-completeness of PA we refer to theorem 5.7 of ([**3**]). ■

These conditions are the tools we need to prove the second incompleteness theorem. We first use them to derive some more useful properties of the provability predicate in PA.

Proposition 5.4. *For any $\mathcal{L}_{PA}$-formulas φ and ψ, the following hold:*

(1) If $\text{PA} \vdash \varphi \to \psi$, then $\text{PA} \vdash \Box\varphi \to \Box\psi$

(2) $\text{PA} \vdash \Box(\varphi \wedge \psi) \leftrightarrow \Box\varphi \wedge \Box\psi$

Proof. (1) We assume $\text{PA} \vdash \varphi \to \psi$. By the first condition we have that $\text{PA} \vdash \Box(\varphi \to \psi)$. Combining this with the second condition, which states that $\text{PA} \vdash \Box\varphi \wedge \Box(\varphi \to \psi) \to \Box\psi$, we get $\text{PA} \vdash \Box\varphi \to \Box\psi$.

(2) We first prove the implication from left to right. We know by what we have just proven, that $\text{PA} \vdash \Box(\varphi \wedge \psi) \to \Box\varphi$ and $\text{PA} \vdash \Box(\varphi \wedge \psi) \to \Box\psi$ since $\text{PA} \vdash \varphi \wedge \psi \to \varphi$ and $\text{PA} \vdash \varphi \wedge \psi \to \psi$ hold. From this we may conclude that

$$\text{PA} \vdash \Box(\varphi \wedge \psi) \to \Box\varphi \wedge \Box\psi.$$

For the converse, we consider first the fact $\text{PA} \vdash \varphi \to (\psi \to \varphi \wedge \psi)$. We again apply our previously proved property to derive $\text{PA} \vdash \Box\varphi \to \Box(\psi \to \varphi \wedge \psi)$. It follows from this and the derivability conditions that

$$\text{PA} \vdash \Box\varphi \wedge \Box\psi \to \Box(\psi \to \varphi \wedge \psi) \wedge \Box\psi \to \Box(\varphi \wedge \psi),$$

which indeed leads to

$$\text{PA} \vdash \Box\varphi \wedge \Box\psi \to \Box(\varphi \wedge \psi).$$

■

Theorem 5.5. *Given an $\mathcal{L}_{PA}$ formula φ such that $\text{PA} \vdash \varphi \leftrightarrow \neg\Box\varphi$, we have that*

$$\text{PA} \vdash \varphi \leftrightarrow \neg\Box \bot .$$

Proof. We know that $\text{PA} \vdash \bot \to \varphi$ for any φ. By proposition 5.4 then, it follows that $\text{PA} \vdash \Box \bot \to \Box\varphi$. Following from the assumption of φ, we have

$$\text{PA} \vdash \varphi \to \neg\Box\varphi \to \neg\Box \bot .$$

So indeed $\text{PA} \vdash \varphi \to \neg\Box \bot$.

For the converse, we begin by deriving $\text{PA} \vdash \Box\varphi \to \Box\neg\Box\varphi$ from $\text{PA} \vdash \varphi \to \neg\Box\varphi$. Furthermore, it follows from the third derivability condition that $\text{PA} \vdash \Box\varphi \to$

$\Box\Box\varphi$. Combining those gives $\mathrm{PA} \vdash \Box\varphi \to (\Box\neg\Box\varphi \wedge \Box\Box\varphi)$. By proposition 5.4 then, $\mathrm{PA} \vdash \Box\varphi \to \Box(\neg\Box\varphi \wedge \Box\varphi)$ which is equivalent to $\mathrm{PA} \vdash \Box\varphi \to \Box \bot$. Using the assumption on φ and taking the contraposition, we get

$$\mathrm{PA} \vdash \neg\Box \bot \to \neg\Box\varphi \to \varphi.$$

■

Theorem 5.6 (Gödel's Second Incompleteness Theorem for PA)**.**

$$\text{If PA is consistent, } \mathrm{PA} \nvdash \mathrm{Con(PA)}$$

Proof. We have defined Con(PA) to be equivalent to $\neg\Box \bot$. We apply theorem 5.5 to the Gödel sentence to derive

$$\mathrm{PA} \vdash G \leftrightarrow \mathrm{Con(PA)}.$$

It follows from the first incompleteness theorem that, if PA is consistent, G is independent of PA. Therefore Con(PA) is independent of PA as well. So indeed, $\mathrm{PA} \nvdash \mathrm{Con(PA)}$ if PA is consistent. ■

5.1 Generalizing the argument

Similarly to how we proved the first incompleteness theorem applies to axiomatized extensions of PA, we can do the same with the second theorem. Let T be an axiomatized consistent theory extending PA. As we have seen in theorem 4.9, we can construct a Gödel sentence G_T such that

$$T \nvdash G_T \quad \text{and} \quad T \nvdash \neg G_T.$$

It is clear that the derivability conditions hold for T, so we apply theorem 5.5 to derive

$$T \vdash G_T \leftrightarrow \neg\Box_T \bot \leftrightarrow \mathrm{Con}(T).$$

Clearly T is unable to prove its own consistency as well. This is stated in the following theorem.

Theorem 5.7. *Any axiomatized, consistent theory T extending* PA *is unable to prove its own consistency.*

5.2 T's ignorance about what it cannot prove

Next we consider an interesting result of the second incompleteness theorem. Let T again be an axiomatized theory extending PA. As we have already seen, if $T \vdash \varphi$ then $T \vdash \Box\varphi$. So in a certain sense, T 'knows' that it can prove φ. However, in this same sense, T is ignorant of what it *cannot* prove. As we will soon see, even if $T \nvdash \varphi$, we will not get $T \vdash \neg\Box\varphi$.

As we have seen in the proof of theorem 5.5, for any $\mathcal{L}_T$-sentence φ, we have $T \vdash \neg\Box\varphi \to \mathrm{Con}(T)$. Therefore, if T would be able to prove its inability to prove a sentence, it would prove its consistency. By Gödel's second incompleteness theorem however, we know that this cannot be the case. We have arrived at T's ignorance about what it cannot prove, as stated in the theorem below.

Theorem 5.8. *If T is an axiomatized and consistent theory extending* PA*, then for no $\mathcal{L}_{PA}$-sentence φ do we have*

$$T \vdash \neg\Box\varphi.$$

References

[1] Hájek, P., and P. Pudlák. 1998. *Metamathematics of First-Order Arithmetic.* Perspectives in Mathematical Logic. Berlin: Springer-Verlag.

[2] Mendelson, E. 1964. *Introduction to Mathematical Logic.* London: Chapman and Hall.

[3] van Oosten, J. 2015. *Gödel's Incompleteness Theorems.* Lecture notes, Utrecht University.

[4] Smith, P. 2013. *An Introduction to Gödel's Theorems.* Cambirdge Introductions to Philosophy. Cambridge: University Press.

[5] Smorynski, C. 1977. *Gödel's Incompleteness Theorems.* Handbook of Mathematical Logic. Amsterdam: North-Holland.