THE GÖDEL INCOMPLETENESS THEOREMS

R. W. Knight

Hilary 2021

ABOUT THE LECTURE NOTES

Under normal circumstances, these notes would be my script for in-person lectures, which would be fleshed out by ad libs, questions and comments from the audience, extra diagrams, digressions and so forth. This year none of that is possible. I'm trying to supply some of that extra material in the videos. The biggest thing missing is you, the audience. As a very distant second-best, you can email me comments, questions, requests for clarification etc., and I will try to respond as quickly as I can (but I'm afraid that sometimes, that may not be very).

The videos will not be a read-through of the notes. I intend in the videos to concentrate on the parts that I think are most difficult, passing over some of the easier stuff. The tone of the videos will also be less formal than that of the notes, and they will concentrate on trying to get over the intuitions and some of the motivation.

It looks as though they are going to end up taking less time than the lectures would have done, possibly partly because I would have said absolutely everything at least once and probably as much as three times (before writing it out, while writing it, and afterwards), and would in addition have talked while waiting for people to finish writing up whatever they were writing.

I hope you find these notes, and the videos, useful. Please, as I say, get in touch if you have any questions.

I will continue to edit this document throughout the term. (I ended last year's lectures wanting to make substantial changes. Then the pandemic hit…I hope to make at least some changes to improve the notes, all the same.)

Latest edit: 18th January 2021.

Everything in the lectures or on the problem sheets is on the syllabus and examinable, unless otherwise indicated.*

Prerequisites: an introductory course in logic is assumed.

* Anything in the footnotes is not on the syllabus.

0. Introduction

We will usually assume that the semiring $(\mathbb{N}, 0, 1, +, \cdot)$ exists and has the usual familiar properties (from which it will follow that various axiom systems for it are consistent, since they have a model).

I assume familiarity with the Completeness Theorem of first-order logic; so when I prove a statement such as $S \vdash \phi$ (ϕ is formally provable from assumptions S) I will on the whole not provide a formal proof of ϕ from S; I will instead argue that such a formal proof exists (which is quite different and much easier). I also assume some skill in distinguishing language from metalanguage and theorems from metatheorems.

These lectures are based on lecture notes by Dan Isaacson, and on Raymond Smullyan's book *Gödel's Incompleteness Theorems* (OUP, 1992). However, I sometimes depart (in notation or in other respects) from both sources.

1. A formal language for arithmetic

1.1. The language itself

We choose a formal language to make Gödel numbering more straightforward.

Definition 1.1.1. *The* symbols *of the language $\mathscr{L}_E$ are:*

$$\overline{0} \;{}^{+}\; v \; f \; {}' \; (\;) \; \neg \; \rightarrow \; \forall \; = \; \leq \; \#$$

An expression *in $\mathscr{L}_E$ is any finite, non-empty sequence of symbols of $\mathscr{L}$ that does not begin with $^+$.*

The rules of syntax are as follows.

Definition 1.1.2. *The* terms *of $\mathscr{L}_E$ are defined as follows.*

$\overline{0}$ is a numeral term*, and if σ is a numeral term, then so is σ^+. We will write $\overline{0}$ followed by n $^+$'s as $\overline{n}$.*

v is a variable term*, and if τ is a variable term, then so is τ'. If n is a natural number (including zero), then we'll write v_n for v followed by n $'$'s.*

The function labels *are f, f', and f''.*

A term *is a numeral term, a variable term, an expression σ^+ where σ is a term, or an expression $(\tau\,\sigma\,\upsilon)$, where σ is a function label and τ and υ are terms.*

Definition 1.1.3. *$\mathscr{L}_E$ contains the following formulae.*

An atomic formula *is an expression $\sigma = \tau$ or $\sigma \leq \tau$, where σ and τ are terms.*

Other formulae are: $\neg\phi$, $(\phi \rightarrow \psi)$, $\forall x\,\phi$, where ϕ and ψ are formulae, and x is a variable term.

We sometimes write $(\phi \vee \psi)$ for $(\neg\phi \rightarrow \psi)$, $(\phi \wedge \psi)$ for $\neg(\neg\phi \vee \neg\psi)$, $(\phi \leftrightarrow \psi)$ for $((\phi \rightarrow \psi) \wedge (\psi \rightarrow \phi))$, and $\exists x\,\phi$ for $\neg\forall v_i\,\neg\phi$, where ϕ and ψ are formulae, and v_i is a variable term.

Definition 1.1.4. *$\mathscr{L}$ is the sublanguage of $\mathscr{L}_E$ containing no occurrences of f''.*

1.2. Logical rules

Given the Completeness Theorem of first-order Predicate Calculus, it does not much matter which system of axioms and logical rules we use. We choose to use one which makes it easier to prove the (meta)theorems we want to use (but which is also difficult to use for constructing formal proofs).

So, we use the following axiom schemes:

DEFINITION 1.2.1. *The logical axioms are all instances of the following schemata, where* ϕ, χ *and* ψ *may be any formulae:*

(A1) $(\phi \to (\chi \to \phi))$

(A2) $((\phi \to (\chi \to \psi)) \to ((\phi \to \chi) \to (\phi \to \psi)))$

(A3) $((\neg\phi \to \neg\chi) \to (\chi \to \phi))$

(A4) $(\forall v_i\, \phi(v_i) \to \phi(t))$, *where* v_i *is a variable letter and* t *is a term which can be sensibly substituted for* v_i, *that is, it contains no variable letter* v_j *such that* v_i *occurs free in* ϕ *in the scope of a quantifier* $\forall v_j$,

(A5) $(\forall v_i\, (\phi \to \chi) \to (\phi \to \forall v_i\, \chi))$, *for* v_i *a variable letter, provided* v_i *does not occur free in* ϕ.

(A6) $\forall v_i\, (v_i = v_i)$

(A7) If F *and* G *are atomic formulae, where* G *is obtained from* F *by replacing some, but not necessarily all, occurrences of* v_i *by* v_j, *then* $((v_i = v_j) \to (F \to G))$.

DEFINITION 1.2.2. *The rules of inference are the following, where* ϕ *and* χ *are any formulae and* x *is any variable letter:*

(MP) Modus Ponens*: that is, from* ϕ *and* $\phi \to \chi$ *deduce* χ

(Gen) Generalisation*: from* ϕ *deduce* $\forall v_i\, \phi$.

DEFINITION 1.2.3. *If* Γ *is a (possibly empty) set of formulae, and* ϕ *is a formula, we say that* ϕ *can be proved from* Γ, *and write* $\Gamma \vdash \phi$, *if and only if there exists a finite sequence* $\phi_1, \ldots, \phi_n$ *of formulae such that* $\phi_n = \phi$, *and for each* i, ϕ_i *is an element of* Γ, *or a logical axiom, or else it is deduced from previous members of the sequence using a rule of inference.*

We will need to refer to the details of the system occasionally.

1.3. Interpretation

We will usually interpret $\mathscr{L}$ as applying to the semiring $(\mathbb{N}, 0, 1, +, \cdot)$, where $\overline{0}$ is interpreted as referring to 0, $^+$ as referring to the successor function $n \mapsto n+1$ (so that $\overline{n}$ refers to n), and the function symbols f, f' as referring, respectively, to addition and multiplication; and we interpret $\mathscr{L}_E$ as referring to the expansion obtained by adding the exponentiation operation, when f'' will refer to exponentiation.

For terms σ and τ, we normally rewrite $(\sigma\ f\ \tau)$ as $\sigma + \tau$, $(\sigma\ f'\ \tau)$ as $\sigma.\tau$, and $(\sigma\ f''\ \tau)$ as σ^τ.

We will normally define truth with respect to this interpretation, though we will sometimes remember to say "true in $\mathbb{N}$" to make this a little clearer. We will occasionally refer to other interpretations.

DEFINITION 1.3.1. *A subset* A *of* $\mathbb{N}^k$ *is* definable *if and only if there exists a formula* $\phi(v_1, \ldots, v_k)$ *with only* $v_1, \ldots, v_k$ *free, such that* $\phi(\overline{n_1}, \ldots, \overline{n_k})$ *is true if and only if*

$(n_1, \dots, n_k) \in A$. *We say that* A *is* provably definable *from a set of assumptions* S *if* $S \vdash \phi(\overline{n_1}, \dots, \overline{n_k})$ *if* $(n_1, \dots, n_k) \in A$, *and* $S \vdash \neg\phi(\overline{n_1}, \dots, \overline{n_k})$ *if* $(n_1, \dots, n_k) \notin A$.

A function $g : \mathbb{N}^k \to \mathbb{N}$ *is* definable *if and only if the set* $A = \{(n_1, \dots, n_k, g(n_1, \dots, n_k)) : n_1, \dots, n_k \in \mathbb{N}\}$ *is definable, and is* weakly provably definable *from a set of assumptions* S *if* A *is provably definable.* f *is* provably definable *from* S *if for all* $n_1, \dots, n_k \in \mathbb{N}$, $S \vdash \forall v_1\, (\phi(\overline{n_1}, \dots, \overline{n_k}, v_1) \leftrightarrow v_1 = \overline{g(n_1, \dots, n_k)})$, *where* ϕ *is the formula defining* A.

2. Peano arithmetic

2.1. The Peano axioms

We will be considering a number of axiom schemes for arithmetic on $\mathbb{N}$ of different strengths. The most famous, and most commonly used, is:

Definition 2.1.1. *We will denote by* PA *(Peano Arithmetic) the following list of statements (all of which are expressible in* $\mathscr{L}$*):*

1. $\forall v_i\, \neg {v_i}^+ = \overline{0}$*;* $\forall v_i \forall v_j\, ({v_i}^+ = {v_j}^+ \to v_i = v_j)$.

($n \mapsto n^+$ *is an injection from* $\mathbb{N} \leftrightarrow \mathbb{N} \setminus \{0\}$*).*

2. $\forall v_i\, v_i + \overline{0} = v_i$ *and* $\forall v_i\, v_i.0 = 0$.

3. $\forall v_i \forall v_j\, v_i + {v_j}^+ = (v_i + v_j)^+$ *and* $\forall v_i \forall v_j\, v_i.{v_j}^+ = (v_i.v_j) + v_i$.

4. $\forall v_i\, \overline{0} \le v_i$*;* $\forall v_i \forall v_j\, (v_i \le v_j \leftrightarrow (v_i = v_j \lor {v_i}^+ \le v_j))$*;* $\forall v_i\, v_i \le v_i$*;* $\forall v_i \forall v_j\, (v_i \le v_j \lor v_j \le v_i)$*;* $\forall v_i \forall v_j \forall v_k\, ((v_i \le v_j \land v_j \le v_k) \to (v_i \le v_k))$.

($\le$ *is a total order, with initial element* 0*, and* n^+ *is the immediate successor of* n*. Antisymmetry can be proved using the second clause and the induction schema, which follows.)*

5. (Induction Schema): For any formula $\phi(v_1)$ *of* $\mathscr{L}$*, the following is an axiom: if* $\phi(0)$*, and if for all* n*,* $\phi(n)$ *implies* $\phi(n^+)$*, then* $\forall n\, \phi(n)$.

Formally:

$$((\phi(\overline{0}) \land (\forall v_1\, \phi(v_1) \to \phi({v_1}^+))) \to \forall v_1\, \phi(v_1)).$$

Exercise 2.1.2. *The successor function* $n \mapsto n^+$ *is onto* $\mathbb{N} \setminus \{0\}$.

Exercise 2.1.3. $m \le n$ *iff* $\exists k\, m + k = m$.

The strongest axiom set for arithmetic we'll be using is the following.

Definition 2.1.4. *We will denote by* PAE *(Peano Arithmetic with Exponentiation) PA, augmented by the following statements:*

$2'$*:* $\forall v_i\, {v_i}^{\overline{0}} = \overline{1}$.

$3'$*:* $\forall v_i, v_j\, {v_i}^{{v_j}^+} = {v_i}^{v_j}.v_i$.

$5'$*: instances of the induction schema involving formulae* $\phi(v_1)$ *belonging to* $\mathscr{L}_E$ *but not to* $\mathscr{L}$*.*

2.2. Gödel numbering

NOTATION 2.2.1. *We will from time to time write numbers in base 13. When we do that, we will use the symbol A to refer to ten, B to refer to eleven, and C to refer to twelve.*

When confusion is likely, we'll use a subscript $_{13}$ to indicate that a number is to be read in base 13, and $_{10}$ to indicate that it should be read in base 10.

The alphabet of $\mathscr{L}$ has thirteen symbols, and we will assign numbers 0 to 12 to them. A string has a Gödel number, which is got by replacing each symbol by a digit in the set $0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C$, and then interpreting the result as a number in base 13.

(Thirteen is convenient partly because it's prime, and partly because with more symbols, it's easier to work out how to write stuff. We could get away with two symbols, by representing each of the above thirteen symbols by a different string of four 0's and 1's.)

More formally:

DEFINITION 2.2.2. Gödel numbers *are assigned to the symbols of the language $\mathscr{L}_E$ as follows:*

$^+$	$\overline{0}$	(	)	f	$'$	v	$\neg$	$\rightarrow$	$\forall$	$=$	$\leq$	#
0	1	2	3	4	5	6	7	8	9	A	B	C

—where all numbers are to be red in base 13. If s is a symbol of $\mathscr{L}_E$, then its Gödel number may be written as $\ulcorner s \urcorner$.

The Gödel number of an expression of $\mathscr{L}_E$ is obtained by writing the Gödel numbers of the individual symbols in order, and reading the result in base 13; that is, if $\phi = s_0 s_1 \dots s_r$, where the s_i are symbols of $\mathscr{L}$ and s_0 is not $^+$, then:

$$\ulcorner \phi \urcorner = \left(\ulcorner s_0 \urcorner \ulcorner s_1 \urcorner \dots \ulcorner s_r \urcorner \right)_{13}.$$

One can quickly convince oneself that

$$\left(\ulcorner s_0 \urcorner \ulcorner s_1 \urcorner \dots \ulcorner s_r \urcorner \right)_{13} = \ulcorner s_0 \urcorner 13^r + \ulcorner s_1 \urcorner 13^{r-1} + \dots + \ulcorner s_r \urcorner.$$

DEFINITION 2.2.3. *The Gödel number of a term or formula is defined as in the previous definition. The Gödel number of a sequence of terms or formulae is obtained by separating the formulae by #, so that*

$$\ulcorner (\phi_1, \dots, \phi_k) \urcorner = \ulcorner \# \phi_1 \# \phi_2 \# \dots \# \phi_k \# \urcorner.$$

NOTE: All we really require of our system of Gödel numbering is that there should exist a definable (in $\mathscr{L}_E$) function $\circ$ such that $\ulcorner \phi\psi \urcorner = \ulcorner \phi \urcorner \circ \ulcorner \psi \urcorner$, and such that the function $n \mapsto \overline{\ulcorner n \urcorner}$ is definable.

We commit the abuse of using symbols such as x, y, m, n and so forth for v_0, v_1 and so forth.

2.3. The arithmetic hierarchy

We classify formulae in prenex normal form according to the string of quantifiers at the front.

DEFINITION 2.3.1. *A bounded quantifier is of the form $\exists m \leq n$ or $\forall m \leq n$ (strictly not in our language, but $\exists m \leq n\, \phi$ can be expressed by $\exists m\, (m \leq n \wedge \phi)$, and $\forall m \leq n\, \phi$ can be expressed by $\forall m\, (m \leq n \to \phi)$.)*

DEFINITION 2.3.2. *A formula is Σ_0, Π_0, or Δ_0, if it contains no unbounded quantifiers.*

If ϕ is Σ_n, then $\forall m\, \phi$ is Π_{n+1}, and if ϕ is Π_n, then $\exists m\, \phi$ is Σ_{n+1}.

We say that a formula ϕ is provably Σ_n (or) Π_n if there is a formula ϕ' which is respectively Σ_n or Π_n, such that $\phi \leftrightarrow \phi'$ is a theorem. If S is a set of axioms, then we say ϕ is provably Σ_n or Π_n with respect to S if there is a formula ϕ' which is respectively Σ_n or Π_n, such that $S \vdash \phi \leftrightarrow \phi'$. If ϕ is provably Σ_n and provably Π_n, then we say that it is Δ_n; similarly with Δ_n with respect to S.

We often omit the word "provably".

EXAMPLE 2.3.3. *As an example, $(\neg m \leq n \vee \exists k\, m + k = n)$ is not Σ_0, but is provably Σ_0, since it is provably equivalent to $\exists k \leq n\, (\neg m \leq n \vee m + k = n)$.*

2.4. Results concerning expressibility

DEFINITION 2.4.1. *A set T of expressions of $\mathscr{L}_E$ is* definable *if and only if there exists a formula $\phi(x)$ such that $\phi(\overline{\ulcorner \psi \urcorner})$ is true if and only if ψ belongs to T.*

If S is a set of sentences of $\mathscr{L}_E$, and T is a set of expressions, we will say that T is provably definable from *S if for some formula ϕ, $\phi(\overline{\ulcorner \psi \urcorner})$ is provable from S if and only if ψ belongs to T, and $\neg\phi(\overline{\ulcorner \psi \urcorner})$ is provable from S if and only if ψ does not belong to T.*

DEFINITION 2.4.2. *A property ϕ of natural numbers or of finite sequences of natural numbers is* expressible *iff the set $\{\mathrm{n} : \phi(\mathrm{n})\}$ is definable.*

PROPOSITION 2.4.3. *Any finite set of expressions is definable, and is indeed provably definable (from $\varnothing$).*

LEMMA 2.4.4. *Express $m < n$ as $m \leq n \wedge \neg m = n$. This is Σ_0.*

$m \mid n$ is provably Σ_0.

Write $[m, n]$ for $\frac{1}{2}(m+n+1)(m+n)+m$. The function $(m, n) \mapsto [m, n]$ is a bijection between $\mathbb{N} \times \mathbb{N}$ and $\mathbb{N}$, and is provably Σ_0.

LEMMA 2.4.5. *The n alternating blocks of quantifiers in a Σ_n or Π_n formula can be of any length.*

PROOF: Problem sheets. □

LEMMA 2.4.6. *The statement that r is the largest number such that $13^r \leq n$ is expressible in Σ_0.*

PROOF: We express it as follows: $\overline{13}^{r} \leq \overline{n} \wedge \neg \overline{13}^{r^+} \leq \overline{n}$. □

LEMMA 2.4.7. *The statement that n is the result of concatenating k and l is expressible in Σ_0.*

PROOF: We express the statement as follows:

$k \neq 0$; also $l = 0$ and $n = 13.k$, or $l \neq 0$ and there exists $r \leq l$ such that r is the greatest number such that $13^r \leq l$, and $n = 13^{r^+}k + l$. □

It's straightforward to get concatenations of more than two.

LEMMA 2.4.8. *k is an initial part of n: there exists $l \leq n$ such that n is the result of concatenating k and l. We write this as $k\ B\ n$.*

l is a final part of n: there exists $k \leq n$ such that n is the result of concatenating k and l. We write this as $k\ E\ n$.

k is a substring of n: there exists $l \leq n$ such that l is an initial part of n and k is a final part of l. We write this as $k\ P\ n$.

These are all provably Σ_0.

LEMMA 2.4.9. *The first element of the string with Gödel number n has Gödel number m (where n is not zero, and in this case necessarily, m is not zero) can be expressed thus: $0 < m$, $m < 13$, and for some $k \leq n$, n is the result of concatenating m and k.*

The last element of the string with Gödel number n has Gödel number m: $m = 0$ and $13 \mid n$, or $0 < m < 13$ and there exists $k \leq n$ such that n is the result of concatenating k and m.

These statements are provably Σ_0.

LEMMA 2.4.10. *n codes a sequence of formulae or terms, the last member of which is σ: σ contains no #, and either n results from concatenating $\ulcorner \# \urcorner$, $\ulcorner \sigma \urcorner$, and $\ulcorner \# \urcorner$, or there exists $a \leq n$ such that a has first and last characters #, the string ## does not occur in a, and n results from concatenating $\ulcorner a \urcorner$, $\ulcorner \sigma \urcorner$ and $\ulcorner \# \urcorner$.*

This is Σ_0.

2.5. Defining provability

LEMMA 2.5.1. *"$n = \ulcorner \overline{m} \urcorner$" is expressible in Σ_0.*

PROOF: $n = 13^m \overline{\ulcorner 0 \urcorner} + \sum_{i=0}^{m-1} 13^i \ulcorner + \urcorner = 13^m \overline{\ulcorner 0 \urcorner}$ since $\ulcorner + \urcorner = 0$. □

COROLLARY 2.5.2. *"n is the Gödel number of a numeral term" is expressible in Σ_0.*

PROOF: We say: there exists $m \leq n$ such that $n = \overline{\ulcorner m \urcorner}$. □

LEMMA 2.5.3. *"$n = \ulcorner v_m \urcorner$" is expressible in Σ_0.*

PROOF: The following are all expressible in Σ_0: the first character of the expression that n codes for, is v, and all the other characters are $'$; and m is largest such that $13^m \leq n$.

COROLLARY 2.5.4. *"n is the Gödel number of some variable term" is expressible in Σ_0.*

PROOF: We say: there exists $m \leq n$ such that $n = \ulcorner v_m \urcorner$. □

The following looks overcomplicated to me and I'll try and streamline it before Monday.

LEMMA 2.5.5. *"n is the Gödel number of a term" is expressible in Δ_1.*

PROOF: The basic strategy is to define a way of expressing the following notion in Σ_0: "k is the Gödel number of a string of terms, each of which is a numeral term, or a variable term, or is derived from previous terms in the sequence using the successor function, addition, multiplication, or exponentiation"; and then quantifying over values k such that n occurs, or doesn't, in the associated string.

Let $\Phi(k, n)$ express the conjunction of the following, in which I will indicate from time to time that certain numbers are $\leq k$, in order to emphasise that this can be expressed in Σ_0.

1. $\ulcorner \# \# \urcorner$ does not occur as a part of k.

2. The first digit and last digit of k are both $\#$.

3. The number obtained by concatenating $\ulcorner \# \urcorner$, n and $\ulcorner \# \urcorner$ in that order occurs as a part of n.

4. Suppose that $l \leq k$ is an initial part of k, and that the concatenation of $\ulcorner \# \urcorner$, $r \leq k$ and $\ulcorner \# \urcorner$ is a final part of l, where $\ulcorner \# \urcorner$ is not a part of r.

Then one of the following must be true.

4i. r is the Gödel number of a numeral term.

4ii. r is the Gödel number of a variable term.

4iii. There exists a number $\leq k$ which is a part of l, and is the concatenation of $\ulcorner \# \urcorner$, $s \leq k$ and $\ulcorner \# \urcorner$, where $\ulcorner \# \urcorner$ is not a part of s and $r = 13.s$ (ie. r can be obtained from s by putting $^+$ on the end).

4iv. For $i = 1, 2$, there exists a number $\leq k$ which is a part of l, and is the concatenation of $\ulcorner \# \urcorner$, $s_i \leq k$ and $\ulcorner \# \urcorner$, where $\ulcorner \# \urcorner$ is not a part of s_i, and r is the concatenation of $\ulcorner (\urcorner$ with s_1, one of $\ulcorner f \urcorner$, $\ulcorner f' \urcorner$ or $\ulcorner f'' \urcorner$, s_2, and $\ulcorner) \urcorner$.

Now we can express the notion "n is the Gödel number of a term" as $\exists k \, \Phi(k, n)$, which is Σ_1.

Now we observe that for any given n we can calculate (based on the number of digits in n) a number K such that if there exists k such that $\Phi(k, n)$, then some such k must exist such that $k \leq K$.

We reason as follows. The number of substrings of n is $\frac{1}{2}n(n-1) < n^2$, and they all have length no greater than that of n itself. The length of n is around $\log_{13} n$ and is definitely no greater than n. A total of n^2 strings of length n, with $n^2 + 1$ $\#$'s added, has length $n^2(1 + n) + 1$. A number whose base 13 representation has that length, must be less than $13 \times 13^{n^2(1+n)+1}$. So let $K(n) = 13^{2+n^2(1+n)}$ (which is expressible in Σ_0).

Then we can express "n is the Gödel number of a term" by $\forall m \, (m \geq K(n) \rightarrow \exists k \leq m \, \Phi(k, n)$, which is Π_1. $\square$

LEMMA 2.5.6. *The notion "n is the Gödel number of an atomic formula" is expressible in Δ_1.*

PROOF: We say the following: there exist $k \leq n$ and $l \leq n$ such that k and l are the Gödel numbers of terms, and n is the concatenation of k, one of $\ulcorner = \urcorner$ or $\ulcorner \leq \urcorner$, and l. $\square$

LEMMA 2.5.7. *The notion "n is the Gödel number of a formula" is expressible in Δ_1.*

PROOF: Similar to the corresponding argument for terms. $\square$

LEMMA 2.5.8. *The notion "n is the Gödel number of a logical axiom" is expressible in Δ_1.*

PROOF: We express "n is the Gödel number of an instance of (A1), or of (A2), ..., or of (A7)".

For most of the schemata (Ai), it is fairly straightforward, and we give (A1) as an example.

We express "n is the Gödel number of an instance of (A1)" by: there exist k, $l \leq n$ such that k and l are the Gödel numbers of formulae, and n can be obtained by concatenating, in this order, $\ulcorner(\urcorner$, k, $\ulcorner\rightarrow(\urcorner$, l, $\ulcorner\rightarrow\urcorner$, k, and $\ulcorner))\urcorner$.

The most important exception is (A4). We approach this cautiously.

First, we express the notion "k is the Gödel number of a term t such that whenever v_j occurs in t, v_i never occurs free in the formula ϕ in the scope of a quantifier $\forall v_j$".

We express it as the conjunction of the following two statements:

1. k is the Gödel number of a term.

2. Suppose that $j \leq k$ and there exist m_1 and m_2, which are the Gödel numbers of expressions such that $k = m_1{}^\frown\ulcorner v_j\urcorner{}^\frown m_2$ (here using $^\frown$ to indicate concatenation). Then the following never happens: there exist m_3 and m_4 which are the Gödel numbers of expressions such that $\ulcorner\phi\urcorner = m_3{}^\frown\ulcorner v_i\urcorner{}^\frown m_4$ (so v_i occurs in ϕ at a particular location); there do not exist m_5, m_6, m_7 and m_8 such that $m_3 = m_5{}^\frown m_6$ and $m_4 = m_7{}^\frown m_8$ and $m_6{}^\frown\ulcorner v_i\urcorner{}^\frown m_7$ is the Gödel number of a formula beginning with $\forall v_i$ (so, that occurrence of v_i is free), while there do exist m_9, m_{10}, m_{11} and m_{12} such that $m_3 = m_9{}^\frown m_{10}$, $m_4 = m_{11}{}^\frown m_{12}$, $m_{10}{}^\frown\ulcorner v_i\urcorner{}^\frown m_{11}$ is the Gödel number of a formula beginning $\forall v_j$ (so, that occurrence of v_i is inside the scope of a $\forall v_j$).

Now we define the relation $\Phi(\phi, \psi)$ to hold if and only if for some formula θ, $\phi = \forall v_i\, \theta$, and $\psi = \theta(t)$. We can define this by recursion. We say that $\Phi(\phi, \psi)$ holds in any of the following cases.

1. $\phi = \forall v_i\, v_i \leq \overline{0}$, then $\psi = t \leq \overline{0}$; and similarly for the various other atomic formulae.
2. $\phi = \forall v_i \neg\phi_1$ and $\psi = \neg\psi_1$, where $\Phi(\forall v_i\, \phi_1, \psi_1)$ holds.
3. $\phi = \forall v_i\, (\phi_1 \rightarrow \phi_2)$ and $\psi = (\psi_1 \rightarrow \psi_2)$, where $\Phi(\forall v_i\, \phi_1, \psi_1)$ and $\Phi(\forall v_i\, \phi_2, \psi_2)$ hold.
4. $\phi = \forall v_i\, \forall v_i\, \phi_1$ and $\psi = \forall v_i\, \psi_1$, where $\psi_1 = \phi_1$.
5. For $j \neq i$, $\phi = \forall v_i\, \forall v_j\, \phi_1$ and $\psi = \forall v_j\, \psi_1$, where $\Phi(\forall v_i\, \phi_1, \psi_1)$ holds.

Now we say that the sequence $((\theta_1, \chi_1), (\theta_2, \chi_2), \ldots, (\theta_r, \chi_r))$ *witnesses* $\Phi(\theta, \chi)$ if $\theta = \theta_r$, $\chi = \chi_r$, and for each i, $\Phi(\theta_i, \chi_i)$ holds, and either χ_i is atomic, or the recursive definition of the relation Φ for θ_i and χ_i refers to pairs (θ_j, χ_j) earlier in the sequence.

Now we use the same ideas as in the proof of Lemma 2.5.5. We can express the statment "n is the Gödel number of an expression $\#\#\theta_1\#\chi_1\#\#\theta_2\#\chi_2\#\# \cdots \#\#\theta_r\#\chi_r\#\#$ such that the sequence $((\theta_1, \chi_1), \ldots, (\theta_r, \chi_r))$ witnesses $\Phi(\theta, \chi)$" in complexity Δ_1, and because we can provide a bound for the shortest length of a sequence witnessing $\Phi(\theta, \chi)$ in terms of the Gödel numbers of θ and χ, we can express "there exists S such that S codes a sequence witnessing $\Phi(\theta, \chi)$, where $k = \ulcorner\theta\urcorner$ and $l = \ulcorner\chi\urcorner$" in complexity Δ_1.

Now, to express that n is the Gödel number of an instance of (A4), we need to say: "there exist $k, l, m \leq n$ such that k is the Gödel number of a formula $\forall v_i\, \phi$, l is the Gödel number of a formula χ, the relation $\Phi(\forall v_i\, \phi, \chi)$ holds, whenever v_j occurs in t, v_i does not occur free in ϕ in the scope of a quantifier $\forall v_j$, and

$$n = \ulcorner(\urcorner{}^\frown k{}^\frown\ulcorner\rightarrow\urcorner{}^\frown l{}^\frown\ulcorner)\urcorner\text{''}.$$

□

LEMMA 2.5.9. *The statement "n is the Gödel number of a formula, of which m is the Gödel number of a proof from assumptions S" may be expressed by Δ_i formula* $\mathrm{proof}_S(\overline{n}, \overline{m})$*, if S is definable in Δ_i.*

PROOF: We define $\mathrm{proof}_S(\overline{n}, \overline{m})$ as follows.

n is the Gödel number of a formula. m codes a sequence whose last member is n. Whenever c is a member of this sequence, then either c codes an axiom, or an assumption, or there exist earlier members a and b of the sequence coding sequences which are connected to c by a rule of inference. □

LEMMA 2.5.10. *Suppose that S is a set of assumptions definable in Σ_i, where $i \geq 1$.*

Then there is a Σ_i formula $\mathrm{Pr}_S(\overline{n})$ which holds if and only if n is the Gödel number of a formula provable from S.

PROOF: $\mathrm{Pr}_S(\overline{n})$ can be written as $\exists m\ \mathrm{proof}_S(\overline{n}, m)$.

This is Σ_i. □

LEMMA 2.5.11. *Assume that PAE is consistent, and true in* $\mathbb{N}$.

If S is a definable set of assumptions, then $\mathrm{proof}_S(\overline{n}, \overline{m})$ *is provable from PAE if and only if n is the Gödel number of a formula ϕ, and m is the Gödel number of a proof of ϕ from S.*

LEMMA 2.5.12. *Assume that PAE is consistent, and true in* $\mathbb{N}$.

If S is a definable set of assumptions, then $\mathrm{Pr}_S(\overline{n})$ *is provable from PAE if and only if n is the Gödel number of a formula provable from S.*

PROOF: Recall that $\mathrm{Pr}_S(\overline{n})$ can be written as $\exists m\ \mathrm{proof}_S(\overline{n}, m)$.

If $n = \ulcorner \phi \urcorner$, where ϕ is provable from S, then for some m, $\mathrm{proof}_S(\overline{n}, \overline{m})$ is true. If it's true, then we can prove from PAE that m is the Gödel number of a proof of ϕ from S.

Hence we can prove that there exists an m which is the Gödel number of a proof of ϕ from S; that is, we can prove $\mathrm{Pr}_S(\overline{n})$.

Now suppose that $\mathrm{Pr}_S(\overline{n})$ is provable from PAE. From our assumption that PAE is true in $\mathbb{N}$, $\mathrm{Pr}_S(\overline{n})$ is true in $\mathbb{N}$ by the Soundness Theorem. Hence by Lemma 2.5.10, n is the Gödel number of a formula provable from S. □

Note that we cannot in general express $\mathrm{Pr}_S(\overline{n})$ in complexity Δ_i; we do have here an increase in complexity.

DEFINITION 2.5.13. *Suppose that S is a set of sentences of $\mathscr{L}$. Then a* proof predicate *for S is a formula $\mathrm{Pr}_S(x)$ such that for all formulae ϕ of $\mathscr{L}$, ϕ is provable from S if and only if $\mathrm{Pr}_S(\phi)$ is provable from PAE.*

So what we have just proved is that any Δ_i-definable set of formulae has a Σ_i-definable proof predicate.

2.6. PA and PAE are definable

NOTATION 2.6.1. *Suppose $\phi(v_1)$ is a formula, and σ is a term.*

Then we write $\forall v_1\,(v_1 = \sigma \to \phi(v_1))$ as $\phi[\sigma]$, and note that it is provably equivalent to $\phi(\sigma)$.

Note that

$$\ulcorner \phi[\sigma] \urcorner = \ulcorner \forall v_1\,(v_1 = \urcorner\frown\ulcorner\sigma\urcorner\frown\ulcorner\to\urcorner\frown\ulcorner\phi(v_1)\urcorner\frown\ulcorner)\urcorner,$$

which is easily calculated from $\ulcorner\phi(v_1)\urcorner$.

Before looking at how to express PA and PAE, we look at the induction schema:

For all formulae $\phi(x)$ of $\mathscr{L}$, the following is an axiom:

$$\mathrm{Ind}_\phi = (\phi(0) \wedge \forall x(\phi(x) \to \phi(x^+)) \to \forall x\,\phi(x).$$

We note that Ind_ϕ is equivalent to:

$$\mathrm{IND}_\phi = (\phi[0] \wedge \forall y(\phi[y] \to \phi[y^+]) \to \forall y\,\phi[y].$$

Recalling that $\phi[y] = \forall x\,(x = y \to \phi(x))$, we see that the following is the case:

LEMMA 2.6.2. *There exists a Δ_1-formula $F(x, y)$ of $\mathscr{L}$ such that for all formulae ϕ of $\mathscr{L}$, for each formula ϕ, the following statement is provable from PAE:*

For all n, $F(\ulcorner\phi\urcorner, n)$ holds if and only if $n = \ulcorner\mathrm{IND}_\phi\urcorner$.

COROLLARY 2.6.3. *The set of all formulae IND_ϕ, for ϕ a formula, is definable in Δ_1.*

PROOF: This set is defined by: there exists $y \le x$ such that y is the Gödel number of a formula, and $F(y, x)$ holds. $\square$

THEOREM 2.6.4. *PA and PAE are definable in Δ_1.*

PROOF: Note that each of PA and PAE is the union of the induction schema for the appropriate language with a finite set.

The result follows. $\square$

COROLLARY 2.6.5. *Δ_1 formulae $\mathrm{proof}_{\mathrm{PA}}(m, n)$ and $\mathrm{proof}_{\mathrm{PAE}}(m, n)$ expressing that n codes a proof of the statement coded by m, in PA and PAE respectively, can be defined, and so can Σ_1 proof predicates $\mathrm{Pr}_{\mathrm{PA}}(n)$ and $\mathrm{Pr}_{\mathrm{PAE}}(n)$.*

3. Diagonalisation and truth

3.1. Diagonalisation

Definition 3.1.1. *Let E_n be the expression (whatever it is) of Gödel number n, assuming this exists.*

Definition 3.1.2. $d(n) = E_n[\overline{n}]$. *(d for "diagonal".)*

Definition 3.1.3. $D(m, n)$ *is the formula* $n = \overline{\ulcorner d(m) \urcorner}$.

Theorem 3.1.4. *(Diagonal Theorem): given a formula $F(x)$, there exists a formula C such that $C \leftrightarrow F(\overline{\ulcorner C \urcorner})$ is provable in PAE.*

This is a fixed-point theorem.

Proof: We consider the formula $F(\overline{\ulcorner d(y) \urcorner})$.
This is equivalent to $\psi(y) := \forall z \left(D(y, z) \to F(z)\right)$.
Let $k = \ulcorner \psi \urcorner$.
Let $C = \psi[\overline{k}]$.
Now C is $\psi[\overline{k}]$, which is equivalent to $\psi(\overline{k})$, which is equivalent to $F(\overline{\ulcorner d(k) \urcorner})$.
Also $k = \ulcorner \psi \urcorner$, so $C = E_k[\overline{k}] = d(k)$. Hence $F(\overline{\ulcorner d(k) \urcorner}) = F(\overline{\ulcorner C \urcorner})$.
So C is equivalent to $F(\overline{\ulcorner C \urcorner})$. $\square$

3.2. The undefinability of truth

Truth is not expressible.

Theorem 3.2.1. (Tarski's Theorem) *There does not exist a formula* $\mathrm{True}(x)$ *such that* $\mathrm{True}(\overline{\ulcorner \phi \urcorner})$ *is true exactly when ϕ is true in $\mathbb{N}$.*

Proof: Suppose such a formula to exist.

Then by the Diagonal Lemma, there exists a formula C such that C holds if and only if $\neg\,\mathrm{True}(\overline{\ulcorner C \urcorner})$.

But then, $\mathrm{True}(\overline{\ulcorner C \urcorner})$ is true if and only if C is true, if and only if $\neg\,\mathrm{True}(\overline{\ulcorner C \urcorner})$ is true, giving a contradiction. $\square$

4. Recursive functions

In this section we pin down exactly what sets and functions can be described in complexity Δ_1 and Σ_1.

4.1. Recursive functions

Definition 4.1.1. *The primitive recursive functions are the smallest class of functions from finite powers of $\mathbb{N}$ to $\mathbb{N}$ with the following properties.*

1. The constant function $n \mapsto 0$ is primitive recursive.

2. The successor function $n \mapsto n + 1$ is primitive recursive.

3. For any positive integer k, for any $i \leq k$, the projection function $(n_1, \dots, n_k) \mapsto n_i$ is primitive recursive.

4. The function $h(n_1, \dots, n_k) = g(f_1(n_1, \dots, n_k), \dots, f_m(n_1, \dots, n_k))$ is primitive recursive, when g and all f_j are primitive recursive.

5. Primitive recursion: f is primitive recursive, where $f(n_1, \dots, n_k, 0) = g(n_1, \dots, n_k)$, and for all n, $f(n_1, \dots, n_k, n + 1) = h(n_1, \dots, n_k, n, f(n_1, \dots, n_k, n))$, where g and h are primitive recursive.

Example 4.1.2. *The addition function $A : (m, n) \mapsto m + n$ is primitive recursive.*

Proof: Let $h(m, n, k) = k + 1$ (this is primitive recursive, since it is the composition of a projection function with the successor function).

Let $g(m) = m$ (the identity on $\mathbb{N}$ is a projection, so is primitive recursive).

Then for all m, we define A by primitive recursion so that $A(m, 0) = g(m)$, and for all m and n, $A(m, n + 1) = h(m, n, A(m, n))$. $\square$

Example 4.1.3. *The modified subtraction function S defined so that $S(m, n) = m - n$ if $m \geq n$ and $S(m, n) = 0$ if $m < n$, is primitive recursive.*

Multiplication and exponentiation are primitive recursive.

We obtain the recursive partial functions by also using the minimalisation operator, which, given a function g, returns the least n such that $g(n_1, \ldots, n_k, n) = 0$ if there is one, and is undefinable otherwise.

Definition 4.1.4. *The recursive functions are the smallest class of partial functions from finite powers of $\mathbb{N}$ to $\mathbb{N}$ with the following properties.*

1. Any primitive recursive function is recursive.

2. Minimalisation: suppose that $g(n_1, \ldots, n_k, n_{k+1})$ is a primitive recursive function. Then the partial function $f(n_1, \ldots, n_k)$, defined to be the least value of n such that $g(n_1, \ldots, n_k, n) = 0$ if this exists, and undefined if it does not, is recursive.

Example 4.1.5. *Ackerman's function is recursive but not primitive recursive:*

$\psi(0, n) = n + 1$

$\psi(m + 1, 0) = \psi(m, 1)$

$\psi(m + 1, n + 1) = \psi(m, \psi(m + 1, n))$.

It also grows rather fast.

Fact 4.1.6. *The recursive partial functions $f(x_1, \ldots, x_k)$ are precisely those that can in principle be calculated by a computer algorithm (that is, such that there is an algorithm that when presented with input $(a_1, \ldots, a_k)$ for which $f(a_1, \ldots, a_k)$ is defined, outputs $f(a_1, \ldots, a_k)$ after a finite time, and when presented with input for $(a_1, \ldots, a_k)$ for which $f(a_1, \ldots, a_k)$ is undefined, runs for ever without halting).*

Primitive recursion is expressible in $\mathscr{L}_E$.

Theorem 4.1.7. *Every primitive recursive function is definable in complexity Σ_1.*

Proof: We argue by induction on the length of the demonstration that a function is primitive recursive. The only difficult step is when we use primitive recursion. Then we appeal to the following lemma.

Lemma 4.1.8. *Suppose that $g : \mathbb{N}^k \to \mathbb{N}$ and $h : \mathbb{N}^{k+1} \to \mathbb{N}$ are functions that can be defined in complexity Σ_1.*

Then the function f defined by primitive recursion from g and h, that is to say, defined so that:

1. $f(n_1, \ldots, n_k, 0) = g(n_1, \ldots, n_k)$, and

2. $f(n_1, \ldots, n_k, n + 1) = h(n_1, \ldots, n_k, n, f(n_1, \ldots, n_k, n))$,

is definable in complexity Σ_1.

Proof: We express the statement $z = f(n_1, \ldots, n_k, n)$ as follows.

Consider the statement $D(y)$, where y is a natural number: "y is the Gödel number of a sequence, the first element of which is $\overline{[0, g(n_1, \ldots, n_k)]}$, and for each $m \leq y$, if m is a member of the sequence, then for all i, $j \leq y$ such that $m = \overline{[i, j]}$, for all $m' \leq y$, if m' immediately follows m in the sequence, for all i', $j' \leq y$ such that $m' = \overline{[i', j']}$, $i' = i + 1$ and $j = [i, h(n_1, \ldots, n_k, i, j)]$".

This statement is Σ_1, and expresses the idea that y codes a derivation of values of the function f using primitive recursion.

We now express "$z = f(n_1, \ldots, n_k, n)$".as "There exists y such that $D(y)$ holds, and $[n, z]$ occurs in the sequence coded by y."

This is Σ_1. $\square$

THEOREM 4.1.9. *Every recursive partial function is Σ_1-definable, and vice versa.*

PROOF: $\Rightarrow$): Easy once we know primitive recursion is expressible. Minimalisation adds an existential quantifier.

$\Leftarrow$): Let ϕ be Σ_0 such that $y = f(\mathbf{x}) \leftrightarrow \exists z\, \phi(\mathbf{x}, y, z)$.

Roughly speaking, search for y and z—or for $[y, z]$—such that $\phi(\mathbf{x}, y, z)$. If one exists, stop and output y. If not, return $\perp$.

How do we tell if $\phi(\mathbf{x}, y, z)$?

We define primitive recursive h_ψ which tells whether ψ is true or not by recursion on Σ_0 ψ as follows. We will define h_ψ to have k arguments, where k is largest such that v_k occurs free in ψ.

$h_{v_i = v_j}(n_0, \ldots, n_k) = S(1, S(n_i, n_j) + S(n_j, n_i))$, where k is the larger of i and j.

$h_{v_i \leq v_j}(n_0, \ldots, n_k) = S(1, S(n_i, n_j))$ where k is the larger of i and j.

$h_{v_i = \overline{n}}(n_0, \ldots, n_i) = S(1, S(n_i, n) + S(n, n_i))$, and so on through all the other kinds of atomic formula.

$h_{\neg\psi}(n_0, \ldots, n_k) = S(1, h_\psi(n_0, \ldots, n_k))$.

$h_{\phi \to \psi}(n_0, \ldots, n_k) = S(1, h_\phi(n_1, \ldots, n_i).S(1, h_\psi(n_1, \ldots, n_j)))$ where k is the larger of i and j.

$h_{\forall v_i \leq v_j\, \phi}(n_0, \ldots, n_k) = \max_{m \leq n_j} h_\phi(n'_0, \ldots, n'_l)$, where $l = k$ unless $i = k$ and $j < k$, when $l = k - 1$, and $n'_r = n_r$ unless $r = i$, when $n'_r = m$.

$h_{\exists v_i \leq v_j\, \phi}(n_0, \ldots, n_k) = \min_{x \leq y} h_\phi(n'_0, \ldots, n'_l)$, where $l = k$ unless $i = k$ and $j < k$, when $l = k - 1$, and $n'_r = n_r$ unless $r = i$, when $n'_r = m$.

Similarly for formulae beginning $\forall v_i \leq \overline{n}$ and $\exists v_i \leq \overline{n}$.

Then express "$f(\mathbf{x}) = y$" as: "y is the first component of $[y, z]$, where $n = [y, z]$ is least such that $1 - h_\phi(\mathbf{x}, y, z) = 0$". $\square$

DEFINITION 4.1.10. *A set is recursive if its characteristic function χ_A is recursive, and recursively enumerable if the partial function π_A which is 1 on the set and undefined off, is recursive.*

THEOREM 4.1.11. *Equivalently, a set is recursively enumerable iff it is Σ_1, and recursive iff it is Δ_1.*

PROOF: If A is recursively enumerable, then π_A is recursive, and hence Σ_1-definable. Then A is defined by the statement "$\pi_A(\overline{n}) = \overline{1}$", which is Σ_1.

Suppose A is defined by a Σ_1 formula ϕ.

Then we define π_A thus: a pair (n, m) belongs to the graph of π_A if and only if $\phi(\overline{n})$ and $\overline{m} = \overline{1}$. This statement can be expressed in Σ_1.

Suppose A is recursive. Then $\pi_A = \pi_{\{1\}} \circ \chi_A$, which is recursive, and $\pi_{A^c} = \pi_{\{0\}} \circ \chi_A + 1$, which is also recursive.

By the above reasoning, both A and its complement are Σ_1-definable.

Also by the above, if A is Δ_1, then A and its complement are both Σ_1. Let us suppose that A is defined by ϕ and its complement by ψ.

Then we may define χ_A in Σ_1 as follows: The formula $\theta(n, m)$ asserting that (n, m) belongs to the graph of χ_A expresses: "either $\phi(\overline{n})$ and $\overline{m} = \overline{1}$, or $\psi(\overline{n})$ and $\overline{m} = \overline{0}$".

So χ_A is Σ_1-definable, and therefore recursive. □

Corollary 4.1.12. *A set A is recursive if and only if both A and its complement are recursively enumerable.*

Corollary 4.1.13. *A subset A of $\mathbb{N}$ is recursively enumerable if and only if it is the range of some recursive partial function.*

Proof: If A is recursively enumerable, then define $f(n)$ to be $n.\pi_A(n)$. This is recursive, and has range A.

Now suppose that f is a recursive partial function with range A. Since f is recursive, the statement "(n, m) is in the graph of f" is Σ_1-definable. Now the statement "(n, m) is in the graph of π_A" may be expressed as: "there exists k such that $(k, \overline{n})$ is in the graph of f, and $\overline{m} = \overline{1}$". □

4.2. Defining exponentiation

Lemma 4.2.1. *The property of being a power of 13 can be defined in $\mathscr{L}$ in complexity Σ_0.*

Proof: We express "n is a power of 13" as: "$\overline{n} = \overline{1}$, or $\overline{13} \mid \overline{n}$, and if $p \leq n$ and $p \mid \overline{n}$ and $\neg\overline{13} \mid p$, then $p = \overline{1}$". □

Lemma 4.2.2. *The statement "n is the smallest power of 13 greater than m" can be expressed in $\mathscr{L}$ in complexity Σ_0.*

Proof: We express it as "$\overline{n}$ is a power of $\overline{13}$ and $\overline{n}$ is greater than $\overline{m}$ and for all $k \leq \overline{n}$, if k is a power of $\overline{13}$ greater than $\overline{m}$, then $k = \overline{n}$". □

Lemma 4.2.3. *We can express the concatenation operator $(m, n) \mapsto m^\frown n$ in $\mathscr{L}$ in complexity Σ_0.*

Proof: We express "$k = m^\frown n$" as "there exists $l \leq \overline{k}$ such that l is the smallest power of $\overline{13}$ greater than $\overline{n}$, and $\overline{k} = \overline{m}.\overline{l} + \overline{n}$". □

Proposition 4.2.4. *We can express any primitive recursive function in $\mathscr{L}$ in complexity Δ_1, and any recursive partial function in complexity Σ_1.*

Corollary 4.2.5. *The statement $k = m^n$, for m, n and k natural numbers, is definable in $\mathscr{L}$, and is indeed Δ_1.*

Proof: Exponentiation is primitive recursive. □

COROLLARY 4.2.6. *Any formula ϕ of $\mathscr{L}_E$ is provably equivalent from PAE to a formula ϕ' of $\mathscr{L}$.*

Moreover, for $n \geq 1$, if ϕ is Σ_n, then ϕ' can be chosen to be Σ_n, and if ϕ is Π_n, then ϕ' can be chosen to be Π_n.

PROPOSITION 4.2.7. *A formula of $\mathscr{L}$ is provable in PAE if and only if it is provable in PA.*

PROOF: Exponentiation can be defined in $\mathscr{L}$ so that formulae of $\mathscr{L}$ equivalent to the axioms of PAE can be written in $\mathscr{L}$.

In more detail, the axioms of PAE not included in PA fall into two classes.

Firstly, instances of the induction schema written in $\mathscr{L}_E$. These are rewritten as equivalent instances of the induction schema written in $\mathscr{L}$, which are already axioms of PA.

Secondly, the two statements $\forall v_1\, {v_1}^{\overline{0}} = \overline{1}$ and $\forall v_1\, \forall v_2\, ({v_1}^{{v_2}^+} = {v_1}^{v_2}.v_1)$. The replacements of these are theorems of PA. $\square$

COROLLARY 4.2.8. *Any subset of $\mathbb{N}^k$, or any function, that is Σ_n or Π_n in PAE is similarly in PA, for $n \geq 1$.*

COROLLARY 4.2.9. *If n is the Gödel number of a formula of $\mathscr{L}$, then the statements $\mathrm{Pr}_{\mathrm{PA}}(\overline{n})$ and $\mathrm{Pr}_{\mathrm{PAE}}(\overline{n})$ are equivalent.*

5. Provability

5.1. Properties of provability

THEOREM 5.1.1. *Suppose S is a provably definable set of assumptions, and $S \vdash \phi$. Then $\mathrm{PA} \vdash \mathrm{Pr}_S(\phi)$.*

PROOF: Write out a formal proof of ϕ from S. Let n be its Gödel number. Then $\mathrm{PA} \vdash \mathrm{proof}_S(\overline{\ulcorner \phi \urcorner}, \overline{n})$.

In more detail (we will be referring to this in the proof of the next theorem but one), suppose that $(\phi_1, \ldots, \phi_m)$ is a formal proof of ϕ from S, and that $n = \ulcorner(\phi_1, \ldots, \phi_m)\urcorner$. We argue that $\mathrm{PA} \vdash \mathrm{proof}_S(\overline{\ulcorner \phi \urcorner}, \overline{n})$.

Let $\psi_S(x)$ be a formula which provably expresses "x is the Gödel number of a member of S", that is, $\psi_S(x)$ is such that $\mathrm{PA} \vdash \psi_S(\overline{n})$ if n is the Gödel number of an element of S, and $\mathrm{PA} \vdash \neg\psi_S(\overline{n})$ otherwise.

Let $\psi_{\mathrm{ax}}(x)$ be a formula which provably expresses "x is the Gödel number of a logical axiom", let $\psi_{\mathrm{rule}}(x, y, z)$ express "x is the Gödel number of a formula which can be obtained by means of a logical rule (MP or Gen) from the formulae with Gödel numbers y and z". Let $\psi_{\mathrm{last}}(x, y)$ provably express "x is a formula, and is the last member of the sequence of formulae whose Gödel number is y".

Then $\mathrm{proof}_S(x, y)$ expresses the following: "x is the Gödel number of a formula, y is the Gödel number of a sequence of formulae, if ϕ_i occurs in this sequence, then $\psi_S(\overline{\ulcorner \phi_i \urcorner})$ or $\psi_{\mathrm{ax}}(\overline{\ulcorner \phi_i \urcorner})$ or for some earlier members ϕ_j and ϕ_k, $\psi_{\mathrm{rule}}(\overline{\ulcorner \phi_i \urcorner}, \overline{\ulcorner \phi_j \urcorner}, \overline{\ulcorner \phi_k \urcorner})$; and $\psi_{\mathrm{last}}(x, y)$".

Now $\mathrm{proof}_S(\overline{\ulcorner \phi \urcorner}, \overline{\ulcorner(\phi_1, \ldots, \phi_m)\urcorner})$ is true, so we can compile a proof of it from PA by putting together the proofs of the various formulae $\psi_*(x, y)$ that we need. $\square$

THEOREM 5.1.2. *Suppose S is a definable set of assumptions.* $\mathrm{PA} \vdash \mathrm{Pr}_S(\phi \to \psi) \to (\mathrm{Pr}_S\,\phi \to \mathrm{Pr}_S\,\psi)$.

PROOF: We show that $\mathrm{PA} \cup \{\mathrm{Pr}\,\phi, \mathrm{Pr}(\phi \to \psi)\} \vdash \mathrm{Pr}\,\psi$, and deduce the result from that.

Suppose that in a model of PA, $\mathrm{proof}(\overline{\ulcorner\phi\urcorner}, n_1)$ and $\mathrm{proof}(\overline{\ulcorner(\phi \to \psi)\urcorner}, n_2)$.

Then if $n = n_1^\frown n_2^\frown \ulcorner\psi\urcorner^\frown \ulcorner\#\urcorner$, then $\mathrm{proof}(\overline{\ulcorner\psi\urcorner}, n)$ holds. $\square$

THEOREM 5.1.3. *Suppose that S is a provably definable set of assumptions, including PA as a subset.*

Then $\mathrm{PA} \vdash \mathrm{Pr}_S(\overline{\ulcorner\psi\urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner\mathrm{Pr}_S(\overline{\ulcorner\psi\urcorner})\urcorner})$.

PROOF: This is an arithmetised version of the proof of Theorem 5.1.1.

Consider the statement "m is the Gödel number of a proof of ϕ from S which is l steps long"; write as $\mathrm{proof}^*_S(\overline{\ulcorner\phi\urcorner}, m, l)$. Note that $\mathrm{proof}_S(\overline{\ulcorner\phi\urcorner}, m)$ is equivalent to $\exists l\ \mathrm{proof}^*_S(\overline{\ulcorner\phi\urcorner}, m, l)$. We will argue inductively that for all l, if $\mathrm{proof}^*_S(\overline{\ulcorner\phi\urcorner}, m, l)$ holds, then $\mathrm{Pr}_S(\overline{\ulcorner\mathrm{Pr}_S(\overline{\ulcorner\phi\urcorner})\urcorner})$ holds; the argument will proceed by recursively constructing the Gödel number M of a proof in S of $\mathrm{Pr}_S(\overline{\ulcorner\phi\urcorner})$, and noting that $\mathrm{proof}_S(\overline{\ulcorner\mathrm{Pr}_S(\overline{\ulcorner\phi\urcorner})\urcorner}, M)$ holds.

We will do this in a general model of PA, which means we need to be careful, because in arbitrary models of PA, $\mathrm{Pr}_S(\overline{\ulcorner\psi\urcorner})$ does not necessarily entail $S \vdash \phi$.

Suppose that $\mathfrak{N}$ is a model of PA, and that $\mathfrak{N} \vDash \mathrm{Pr}_S(\overline{\ulcorner\phi\urcorner})$.

Then there exist $m, l \in \mathfrak{N}$ such that $\mathfrak{N} \vDash \mathrm{proof}^*_S(\overline{\ulcorner\phi\urcorner}, m, l)$. In the argument that follows we need to bear in mind that $\mathfrak{N}$ may not be $\mathbb{N}$, and that m and l may not be actual natural numbers; they just have, in $\mathfrak{N}$, some of the first-order properties that natural numbers possess.

We argue using induction on l (this can be formalised in PA, using the induction scheme). We examine the inductive step; the base case is similar, but easier.

Suppose then that in $\mathfrak{N}$, $l > 1$, and that $\mathfrak{N} \vDash \mathrm{proof}^*_S(i, m, l)$. Suppose, using the inductive hypothesis, that if, in $\mathfrak{N}$, j is the Gödel number other than the last one of an element of the sequence whose Gödel number is m, then there exists M_j in $\mathfrak{N}$ such that $\mathfrak{N} \vDash \mathrm{proof}_S(\overline{\ulcorner\mathrm{Pr}_S(i)\urcorner}, M_j)$.

If in $\mathfrak{N}$, i is the Gödel number of a member of S, that is to say, if $\mathfrak{N} \vDash \psi_S(i)$, then we let M_j be the Gödel number of a proof of $\psi_S(i)$.

To justify this step further, recall that in $\mathbb{N}$, $\psi_S(\overline{n})$ is provable if n is the Gödel number of a member of S, and $\neg\psi_S(\overline{n})$ is provable if not. It follows that there is an algorithm which, when presented with input n, outputs the Gödel number of a proof of $\psi_S(\overline{n})$ if $n \in S$, and outputs the Gödel number of a proof of $\neg\psi_S(\overline{n})$ if not. The algorithm goes like this. Examine all formal proofs in S, one by one. (They can be listed in a recursive way that permits us to do this). We will eventually encounter either a proof of $\psi_S(\overline{n})$, in which case we output its Gödel number; or we encounter a proof of $\neg\psi_S(\overline{n})$, in which case we output the Gödel number of that. The existence of this algorithm means that there is a Σ_1-definable function f_S inputting n and outputting the Gödel number of the appropriate proof. Suppose that $\chi(x, y)$ expresses "$y = f_S(x)$". Then if we are in the situation of the previous paragraph, and $\mathfrak{N} \vDash \psi_S(i)$, then there exists $M_i \in \mathfrak{N}$ such that $\mathfrak{N} \vDash \chi(i, N_i)$. Construct M_i by appending N_i and a proof of ψ_{last}.

In a similar way, if in $\mathfrak{N}$, i is the Gödel number of a logical axiom, then let M_i be the Gödel number of a proof of $\psi_{\mathrm{ax}}(i)$.

If i is, in $\mathfrak{N}$, the Gödel number of a formula obtained using an application of a rule to formulae earlier in the sequence coded by n whose Gödel numbers are j and k, then by the inductive hypothesis there exist elements M_j and M_k of $\mathfrak{N}$ such that $\mathfrak{N} \models \mathrm{proof}_S(j, M_j)$ and $\mathfrak{N} \models \mathrm{proof}_S(k, M_k)$. Then we generate M_i by combining M_j and M_k, deleting repeated #'s as necessary. $\square$

THEOREM 5.1.4. *Suppose that S is a set of sentences of $\mathscr{L}$, definable from P.*

For all ϕ, ϕ is provable in S if and only if $\mathrm{Pr}_S(\overline{\ulcorner\phi\urcorner})$ is true in the standard model $\mathbb{N}$.

PROOF: If ϕ is provable from S, then $\mathrm{Pr}(\overline{\ulcorner\phi\urcorner})$ is provable in PA, so $\mathrm{Pr}(\overline{\ulcorner\phi\urcorner})$ is true in $\mathbb{N}$ (since true in all models of PA).

Now suppose that $\mathrm{Pr}(\overline{\ulcorner\phi\urcorner})$ is true in $\mathbb{N}$.

Now $\mathrm{Pr}(\overline{\ulcorner\phi\urcorner})$ is an existential statement saying that there is an n such that n is the Gödel number of a proof whose last line is ϕ. Since it is true in $\mathbb{N}$, there must be a natural number n about which the formula $\mathrm{proof}(\overline{\ulcorner\phi\urcorner}, \overline{n})$ is true. Then n is indeed the Gödel number of a proof in S whose last line is ϕ, and so that proof witnesses that ϕ is provable in S. $\square$

5.2. A limit on the power of proof

THEOREM 5.2.1. (Weak form of the First Incompleteness Theorem) *Suppose that S is a definable set of sentences that is true in $\mathbb{N}$ and includes PA.*

Then there exists a formula G such that G is true in $\mathbb{N}$, but is not provable from S.

PROOF: Using the Diagonal Lemma, find a formula G such that $G \leftrightarrow \neg\,\mathrm{Pr}_S(\overline{\ulcorner G\urcorner})$ is provable from PA.

Suppose that G is provable from S. Then G is true in $\mathbb{N}$, by assumption and by soundness. Also $\mathrm{Pr}_S(\overline{\ulcorner G\urcorner})$ is true by Theorem 5.1.4. But then G is false, contradiction.

So G is not provable from S.

Now suppose that G is false. Then $\mathrm{Pr}_S(\overline{\ulcorner G\urcorner})$ is true, so G is provable from S by Theorem 5.1.4. But S is true in $\mathbb{N}$, so G is true in $\mathbb{N}$ also by soundness, contradiction.

So G is true in $\mathbb{N}$. $\square$

5.3. Grades of completeness

DEFINITION 5.3.1. *A set of axioms S is n-inconsistent if and only if there exists a Σ_n formula $\exists x\,\phi(x)$ such that $\vdash_S \exists x\,\phi(x)$, but for all m, $\vdash_S \neg\phi(\overline{n})$.*

S is n-consistent if and only if it is not n-inconsistent.

S is ω-consistent if and only if it is n-consistent for all n.

DEFINITION 5.3.2. *A set S of statements is Σ_i-complete if and only if all true Σ_i-sentences are provable from S.*

We will say it is Σ_i-sound if and only if all Σ_i-sentences provable from S are true.

DEFINITION 5.3.3. *The axiom scheme Q is the following list of axioms:*

$\forall v_1 \forall v_2\, {v_1}^+ = {v_2}^+ \to v_1 = v_2$.

$\forall v_1\, \neg {v_1}^+ = \overline{0}$.

$\forall v_1\, v_1 + \overline{0} = v_1$.
$\forall v_1 \forall v_2\, v_1 + {v_2}^+ = (v_1 + v_2)^+$.
$\forall v_1\, v_1.\overline{0} = \overline{0}$.
$\forall v_1 \forall v_2\, v_1.{v_2}^+ = v_1.v_2 + v_1$.
$\forall v_1\, v_1 \leq \overline{0} \leftrightarrow v_1 = 0$.
$\forall v_1 \forall v_2\, v_1 \leq {v_2}^+ \leftrightarrow (v_1 \leq v_2 \vee v_1 = {v_2}^+)$.
$\forall v_1 \forall v_2\, v_1 \leq v_2 \vee v_2 \leq v_1$.

This has no induction.

Definition 5.3.4. *The following list of axioms is known as R.*
All sentences $\overline{m} + \overline{n} = \overline{k}$, for which $m + n = k$.
All sentences $\overline{m}.\overline{n} = \overline{k}$, for which $m.n = k$.
All sentences $\overline{m} \neq \overline{n}$, where $m \neq n$.
All sentences $\forall v_1\, v_1 \leq \overline{n} \leftrightarrow (v_1 = \overline{0} \vee \cdots \vee x = \overline{n})$.
All sentences $\forall v_1\, v_1 \leq \overline{n} \vee \overline{n} \leq v_1$.

Proposition 5.3.5. *Q extends R.*

Theorem 5.3.6. *R is Σ_0-complete.*

Proof: The second-to-last schema gives a method of eliminating bounded quantifiers. The other axioms allow us to compute the diagrams of $+$, $.$ and $\leq$. □

Corollary 5.3.7. *Q and PA are Σ_0-complete.*

Theorem 5.3.8. *Any system S that is Σ_0-complete is also Σ_1-complete.*

Proof: Suppose that $\exists v_1\, F(v_1)$ is true. Then $F(\overline{n})$ is true for some n. Then $S \vdash F(\overline{n})$ by Σ_0-completeness. So $S \vdash \exists v_1\, F(v_1)$ as required. □

Corollary 5.3.9. *R, Q, and* PA *are Σ_1-complete.*

Presburger arithmetic is PA with all mention of multiplication erased.

Definition 5.3.10. *The following list of statements, in the sublanguage $\mathscr{L}_P$ of $\mathscr{L}$ containing no uses of the multiplication symbol f', are known as* Presburger arithmetic*:*
1. $\neg\forall v_i\, {v_i}^+ = \overline{0}$; $\forall v_i \forall v_j\, ({v_i}^+ = {v_j}^+ \to v_i = v_j)$.
($n \mapsto n^+$ is an injection from $\mathbb{N} \leftrightarrow \mathbb{N} \setminus \{0\}$).
2. $\forall v_i\, v_i + \overline{0} = v_i$.
3. $\forall v_i \forall v_j\, v_i + {v_j}^+ = (m + n)^+$.
4. $\forall v_i\, \overline{0} \leq v_i$; $\forall v_i \forall v_j\, (v_i \leq v_j \leftrightarrow (v_i = v_j \vee {v_i}^+ \leq v_j))$; $\forall v_i\, v_i \leq v_i$; $\forall v_i \forall v_j\, (v_i \leq v_j \vee v_j \leq v_i)$; $\forall v_i \forall v_j \forall v_k\, ((v_i \leq v_j \wedge v_j \leq v_k) \to (v_i \leq v_k))$; $\forall v_i \forall v_j\, (v_i \leq v_j \vee v_j \leq v_i)$.
($\leq$ is a total order, with initial element 0, and n^+ is the immediate successor of n).
5. (Induction Schema): For any formula $\phi(v_1)$ of $\mathscr{L}_P$, the following is an axiom: if $\phi(0)$, and if for all n, $\phi(n)$ implies $\phi(n^+)$, then $\forall n\, \phi(n)$.
Formally:

$$((\phi(\overline{0}) \wedge (\forall v_1\, \phi(v_1) \to \phi({v_1}^+))) \to \forall v_1\, \phi(v_1)).$$

The following theorem is not examinable for part C or OMMS.

Theorem 5.3.11. *Presburger arithmetic is consistent and complete, and the set of consequences of it is decidable.*

Proof: Rather long, very ingenious, and involving quantifier elimination and modular arithmetic. □

5.4. The first incompleteness theorem

Theorem 5.4.1. (First Incompleteness Theorem) *There exists a Π_1-sentence G such that if PA is consistent, then* $\mathrm{PA} \nvdash G$*, and if in addition PA is* 1*-consistent, then* $\mathrm{PA} \nvdash \neg G$.

Proof: Let a be the Gödel number of "$d(x)$ is not provable". (This is Π_1.)

More formally, $a = \ulcorner \neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner d(x) \urcorner}) \urcorner$.

Let G be $E_a[\overline{a}]$ (which is also Π_1).

Now $E_a(x)$ is $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner d(x) \urcorner})$, so $G = E_a[\overline{a}]$ is provably equivalent to $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner d(a) \urcorner})$, that is, to $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner E_a[\overline{a}] \urcorner})$, that is, to $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$.

Suppose that $\mathrm{PA} \vdash G$.

Then $\mathrm{PA} \vdash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$, by Theorem 5.1.1.

So since $\mathrm{PA} \vdash G$, $\mathrm{PA} \vdash \neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$.

Thus PA is inconsistent, giving a contradiction.

Now suppose that $\mathrm{PA} \vdash \neg G$, and that PA is 1-consistent.

Then $\mathrm{PA} \vdash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$, because G is provably equivalent to $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$.

Now $\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner})$ is the same thing as $\exists x \; \mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner}, x)$, and $\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner}, x)$ is Σ_1.

Write $\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner}, x)$ as $\exists y \, \phi(x, y)$, where $\phi(x, y)$ is Σ_0.

Then $\exists x \, \exists y \, \phi(x, y)$ may be rewritten $\exists z \, \exists x \leq z \, \exists y \leq z \, \phi(x, y)$, which is Σ_1 in the strict sense.

Now $\mathrm{PA} \vdash \exists z \, \exists x \leq z \, \exists y \leq z \phi(x, y)$, so because PA is 1-consistent, there must exist n such that $\mathrm{PA} \nvdash \neg \exists x \leq \overline{n} \, \exists y \leq \overline{n} \, \phi(x, y)$.

But PA is Σ_0-complete by Lemma 5.3.6., so $\mathrm{PA} \vdash \exists x \leq \overline{n} \, \exists y \leq \overline{n} \, \phi(x, y)$.

So this statement is true in $\mathbb{N}$. Let m be such that $\mathbb{N} \models \exists y \leq \overline{n} \, \phi(\overline{m}, y)$.

Then $\mathbb{N} \models \mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G \urcorner}, \overline{m})$.

Hence m is the Gödel number of a proof of G in PA.

So we can read off a proof of G in PA from m, and see that $\mathrm{PA} \vdash G$.

Hence PA is inconsistent, giving a contradiction. □

Corollary 5.4.2. *Assume $\mathbb{N}$ is a model of PA. Then G is true in $\mathbb{N}$ and not provable.*

Proof: If $\mathbb{N}$ is a model of PA, then PA is consistent and 1-consistent.

Hence PA proves neither G nor $\neg G$.

Since PA does not prove G, there is no natural number coding a proof of G, so then $\neg \mathrm{Pr}_{\mathrm{PA}}(\overline{G})$ is true in $\mathbb{N}$, so G is true. □

Theorem 5.4.3. *(Rosser's Theorem) Let S be any definable consistent set of sentences including* PA*. Then there is a sentence G such that S neither proves nor disproves G.*

Proof: Let $H(x)$ be the statement $\exists y \, (\mathrm{proof}_S(\overline{\ulcorner \neg \urcorner ^\frown x}, y) \wedge \forall z \leq y \, \neg \, \mathrm{proof}_S(x, z))$.

(Informally, $H(\overline{\ulcorner \phi \urcorner})$ says "there is a y coding a refutation of ϕ, and no $z \leq y$ codes a proof of ϕ".)

Using the Diagonal Lemma, let G be such that $G \leftrightarrow H(\overline{\ulcorner G \urcorner})$ is provable from PA.

We argue that S neither proves nor refutes G.

Suppose first that $S \vdash G$.

Then there is a proof of G from S. Let n be its Gödel number.

Then PA $\vdash \text{proof}_S(\overline{\ulcorner G \urcorner}, \overline{n})$, and so $S \vdash \text{proof}_S(\overline{\ulcorner G \urcorner}, \overline{n})$.

Now S is consistent, so given that $S \vdash G$, then it is not the case that $S \vdash \neg G$; and so no disproof of G exists.

So no natural number m is the Gödel number of a proof from S of $\neg G$; in particular no natural number $m < n$ is the Gödel number of a proof from S of $\neg G$.

So if $m < n$, PA $\vdash \neg\, \text{proof}_S(\overline{\ulcorner \neg G \urcorner}, \overline{m})$, so that PA $\vdash \forall m < \overline{n}\, \neg\, \text{proof}_S(\overline{\ulcorner \neg G \urcorner}, \overline{m})$.

Now n is the Gödel number of a proof of G from S, so PA $\vdash \forall m \geq n\, \exists x \leq m\, \text{proof}_S(\overline{\ulcorner G \urcorner}, x)$ (the value of x that witnesses this is of course n itself).

Putting these two sentences together, PA $\vdash \forall y \neg(\text{proof}_S(\overline{\ulcorner \neg G \urcorner}, y) \wedge \forall z \leq y\, \neg\, \text{proof}_S(\overline{\ulcorner G \urcorner}, z))$.

That is, PA $\vdash \neg H(\overline{\ulcorner G \urcorner})$. Hence $S \vdash \neg H(\overline{\ulcorner G \urcorner})$.

But $S \vdash G$, so $S \vdash H(\overline{\ulcorner G \urcorner})$, giving a contradiction.

Now suppose that $S \vdash \neg G$.

Then there is a proof of $\neg G$ from S. Let n be the Gödel number of that proof.

Since S is consistent, it is not possible that $S \vdash G$. So there is no proof of G from S. So for all $m \leq n$, m is not the Gödel number of a proof of G from S.

Hence PA $\vdash (\text{proof}(\overline{\ulcorner \neg G \urcorner}, \overline{n}) \wedge \forall m \leq n\, \neg\, \text{proof}_S(G, m))$.

So S proves the same thing.

Hence $S \vdash H(\overline{\ulcorner G \urcorner})$. From this it follows that $S \vdash G$, giving a contradiction. □

5.5. The Second Incompleteness Theorem and Löb's Theorem

THEOREM 5.5.1. (Second Incompleteness Theorem) *If S is a provably definable set of sentences including PA, and if a sentence G has the property that $S \vdash G \leftrightarrow \neg\, \text{Pr}_S(\overline{\ulcorner G \urcorner})$, then $S \vdash \neg\, \text{Pr}_S(\overline{\ulcorner X \urcorner}) \rightarrow \neg\, \text{Pr}_S(\overline{\ulcorner G \urcorner})$.*

PROOF: $(G \rightarrow (\neg G \rightarrow X))$ is a tautology and so a theorem of S.

By hypothesis, $S \vdash (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow \neg G)$.

Hence $S \vdash (G \rightarrow (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X))$.

From Theorem 5.1.1, and the assumption that S extends PA, it follows that $S \vdash \text{Pr}_S(\overline{\ulcorner (G \rightarrow (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X)) \urcorner})$.

From Theorem 5.1.2, we have $S \vdash (\text{Pr}_S(\overline{\ulcorner (G \rightarrow (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X)) \urcorner}) \rightarrow (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X) \urcorner})))$.

Hence $S \vdash (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X) \urcorner}))$.

Also from Theorem 5.1.2, we have $S \vdash (\text{Pr}_S(\overline{\ulcorner (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow X) \urcorner}) \rightarrow (\text{Pr}_S(\overline{\ulcorner \text{Pr}_S(\overline{\ulcorner G \urcorner}) \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner X \urcorner})))$.

Thus $S \vdash (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow (\text{Pr}_S(\overline{\ulcorner \text{Pr}_S(\overline{\ulcorner G \urcorner}) \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner X \urcorner})))$.

Now by Theorem 5.1.3, we have that $S \vdash (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner \text{Pr}_S(\overline{\ulcorner G \urcorner}) \urcorner}))$.

Thus $S \vdash (\text{Pr}_S(\overline{\ulcorner G \urcorner}) \rightarrow \text{Pr}_S(\overline{\ulcorner X \urcorner}))$.

Hence $S \vdash (\neg\, \text{Pr}_S(\overline{\ulcorner X \urcorner}) \rightarrow \neg\, \text{Pr}_S(\overline{\ulcorner G \urcorner}))$, as required. □

COROLLARY 5.5.2. *If X is a sentence, and S is consistent, then $S \nvdash \neg\, \text{Pr}(\overline{\ulcorner X \urcorner})$.*

Proof: If G exists, and $S \vdash \neg \mathrm{Pr}(\overline{\ulcorner X \urcorner})$, then $S \vdash \neg \mathrm{Pr}(\overline{\ulcorner G \urcorner})$, so $S \vdash G$. But then $S \vdash \mathrm{Pr}(\overline{\ulcorner G \urcorner})$. So S is inconsistent. $\square$

Definition 5.5.3. *Suppose that S is a definable set of sentences. We define* Con_S *to be the formula* $\neg \mathrm{Pr}_S(\overline{\ulcorner \neg \overline{0} = \overline{0} \urcorner})$. *We read this as "$S$ is consistent".*

Corollary 5.5.4. *If S is consistent, then it is not the case that $S \vdash \mathrm{Con}_S$.*

Proof: In fact, S does not prove the statement $\neg \mathrm{Pr}_S(\overline{\ulcorner X \urcorner})$ for any formula X. $\square$

Theorem 5.5.5. (Löb's Theorem) *Suppose that S is a provably definable set of sentences extending PA. Then from $S \vdash (P(\overline{\ulcorner \phi \urcorner}) \to \phi)$ we can deduce $S \vdash \phi$.*

Proof: Let L be diagonal for $\mathrm{Pr}_S(\cdot) \to \phi$, ie $S \vdash (L \leftrightarrow (\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \phi))$.

Then by Theorem 5.1.1, $S \vdash \mathrm{Pr}_S(\overline{\ulcorner L \to (\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \phi) \urcorner})$.

By Theorem 5.1.2, $S \vdash \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \mathrm{Pr}_S(\overline{\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \phi})$.

By Theorem 5.1.2, $S \vdash \mathrm{Pr}_S(\overline{\ulcorner \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \urcorner}) \to (\mathrm{Pr}_S(\overline{\ulcorner \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner \phi \urcorner})$, so $S \vdash \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to (\mathrm{Pr}_S(\overline{\ulcorner \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner \phi \urcorner}))$ by HS, so $S \vdash ((\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \urcorner})) \to (\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner \phi \urcorner})))$ by (A2), so $S \vdash (\mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \mathrm{Pr}_S(\overline{\ulcorner \phi \urcorner}))$ by Theorem 5.1.3 and MP.

Using HS, $S \vdash \mathrm{Pr}_S(\overline{\ulcorner L \urcorner}) \to \phi$.

But this is equivalent to L, so $S \vdash L$.

By Theorem 5.1.1, $S \vdash \mathrm{Pr}_S(\overline{\ulcorner L \urcorner})$.

Now by MP, $S \vdash \phi$ as required. $\square$

5.6. A stronger version of Σ_1-completeness

We proved earlier that if ϕ is Σ_1 and true, then it is provable.

In this section we strengthen this result.

Theorem 5.6.1. *If ϕ is Σ_1, then* $\mathrm{PA} \vdash (\phi \to \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \phi \urcorner}))$.

Sketch proof: Messy induction on ϕ, of which the messiest part is when ϕ is atomic or negated atomic.

The following special cases can be done algorithmically using induction:

1. $\overline{n} = \overline{n}$ is a logical axiom.
2. $\neg \overline{m} = \overline{n}$ where $m \neq n$.
3. $\overline{m} + \overline{n} = \overline{m+n}$ and $\overline{m}.\overline{n} = \overline{m.n}$.

Bounded quantifiers can be coped with as follows.

If $\phi = \forall m \leq n \psi(\overline{m})$, then the following procedure can be expressed in $\mathscr{L}$: for each $m \leq n$, write down a proof of $\psi(\overline{m})$, deduce $\bigvee_{m \leq n} \psi(\overline{m})$, and then by induction on n deduce $\forall m \leq n \psi(\overline{m})$; the result is a proof of ϕ.

We treat bounded existential quantifiers in a similar way.

Now $\overline{m} \leq \overline{n}$ is provably equivalent to $\exists k \leq n\, \overline{m} + k = \overline{n}$.

Now any Σ_0 formula is provably equivalent, by standard proofs that can be generated by an algorithms, to a disjunction of conjunctions of statements of the above forms.

So it is possible to see that there is an algorithm which inputs true Σ_0 formulae and outputs proofs for them (and recall that truth for Σ_0 formulae is definable).

Now suppose ϕ is a Σ_1 formula $\exists x\, \psi(x)$.

The process that inputs true formulae $\psi(\overline{n})$ and outputs their proofs can be described by an algorithm, and so is Σ_1-definable. So in effect we have the following statement as a theorem of PA: $\forall x\,(\psi(x) \to \mathrm{Pr}_S(\overline{\ulcorner\psi(\overline{x})\urcorner}))$; and $\exists x\ \mathrm{Pr}_S(\overline{\ulcorner\psi(\overline{x})\urcorner})$ entails $\mathrm{Pr}_S(\overline{\ulcorner\exists x\,\psi(\overline{x})\urcorner})$.

The statement $\exists x\,\psi(x) \to \mathrm{Pr}_S(\overline{\ulcorner\exists x\,\psi(x)\urcorner})$ now follows. $\square$

Theorem 5.6.2. *If $\phi(x)$ is Σ_1, then the statement $\forall x\,(\phi(x) \to \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner\phi(x)\urcorner}))$ is a theorem of PA.*

Sketch proof: This is proved by the same techniques as the previous theorem. $\square$

6. Strengthenings of PA

Given that PA is incomplete, we look around for reasonable strengthenings of it. We could use $\mathrm{PA} \cup \mathrm{Con}_{\mathrm{PA}}$, $\mathrm{PA} \cup \mathrm{Con}_{\mathrm{PA}} \cup \mathrm{Con}_{\mathrm{PA}\cup\mathrm{Con}_{\mathrm{PA}}}$, etc. The next section provides a more systematic possible approach.

6.1. The ω-rule

Definition 6.1.1. *Suppose that S is a set of formulae of $\mathscr{L}$.*

We define S^ω to be the logical system whose axioms are S together with all logical axioms, and whose rules are MP, Gen, and the ω-rule *which allows one to deduce $\forall x\,\phi(x)$ from the entire set of assumptions $\{\phi(\overline{n}) : n \in \mathbb{N}\}$.*

A proof *in S^ω is a sequence $(\phi_\alpha : \alpha < \beta)$, where β is an ordinal, such that each ϕ_α is an element of S or a logical axiom, or else is obtained from previous members of the sequence using a rule.*

ϕ is a theorem *of S^ω, and we write $S^\omega \vdash \phi$, iff there is a proof in S^α of which ϕ is the last element.*

We could, if we wished, insist that all proofs have length $< \omega_1$.

The ω-rule looks reasonable-ish. However there is a big problem with it.

Theorem 6.1.2. *R^ω is complete.*

Proof: We can prove by induction on the complexity of a formula ϕ that $R^\omega \vdash \phi$ or $R^\omega \vdash \neg\phi$.

The ω-rule allows us to eliminate quantifiers.

The case where ϕ is Σ_0 is already done since R is Σ_0-complete.

Now suppose that ϕ is Π_{n+1}; say $\phi = \exists x\,\psi(x)$.

There are two cases. If there exists n such that $R^\omega \vdash \psi(\overline{n})$, then it is certainly true that $R^\omega \vdash \exists x, \psi(x)$, so $R^\omega \vdash \phi$. The alternative is (appealing to the inductive hypotheis) that for all n, $R^\omega \vdash \neg\psi(\overline{n})$. Then by the ω-rule, $R^\omega \vdash \forall x\,\neg\psi(x)$, so $R^\omega \vdash \neg\phi$.

This argument of course also does the case when ϕ is Σ_{n+1}. $\square$

Corollary 6.1.3. PA^ω *is complete.*

Corollary 6.1.4. *Assuming that R^ω and PA^ω are sound with respect to truth in $\mathbb{N}$, then the set of theorems of R^ω or of PA^ω is undefinable and so* a fortiori *not recursively enumerable.*

Proof: Using Tarski's Theorem, and the statement that a set is recursively enumerable iff it is Σ_1-definable. $\square$

The upshot is that since, as human beings, we are limited to what is recursively enumerable, R^ω and PA^ω are of no practical use.

In the next section we look at an adaptation of the ω-rule which may be more useful.

6.2. The uniform reflection principle

The uniform reflection principle is an arithmetised version of the ω-rule, and says "if $\phi(\overline{n})$ is provable for all n, then $\forall x\, \phi(x)$ is true" (which can be said in the language).

DEFINITION 6.2.1. *The* uniform reflection principle *URP is the set of axioms got by adding to PA all instances of the following, where formula $F(v_1)$ is a formula of $\mathscr{L}$:*

$$\forall n\ \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \forall v_1\,(v_1 = \urcorner{}^\frown\ulcorner\overline{n}\urcorner{}^\frown\ulcorner \to F(v_1))\urcorner}) \to \forall n\, F(n).$$

We write this as $\forall n\ \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner F[\dot{\overline{n}}]\urcorner}) \to \forall n\, F(n)$, and refer to it as the reflection principle for *F.*

This is better—we have a definable set of axioms here—so less powerful. How powerful?

THEOREM 6.2.2. *Suppose that G is a sentence such that $\mathrm{PA} \vdash G \leftrightarrow \neg\,\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner})$.*

Then $\mathrm{PA} \vdash \forall n\ \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner\neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, \dot{n})\urcorner})$.

PROOF: Recall that $\mathrm{proof}_{\mathrm{PA}}$ is Δ_1.

Suppose that $\mathfrak{N}$ is a model of PA, and that $n \in \mathfrak{N}$.

Then either $\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$ is true in $\mathfrak{N}$, or $\neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$ is true.

If $\mathfrak{N} \models \neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$ is true, then because $\neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$ is Σ_1, we can deduce by Theorem 5.6.1. that $\mathrm{Pr}_{\mathrm{PA}}(\overline{\neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, \overline{n})})$.

If on the other hand $\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$, then $\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner})$, and so $\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner X\urcorner})$ for all X by the Second Incompleteness Theorem, and so $\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, \overline{n})\urcorner})$ in particular. □

COROLLARY 6.2.3. $\mathrm{URP} \vdash G$.

PROOF: Using the previous theorem and URP, we have $\mathrm{URP} \vdash \forall n\, \neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner G\urcorner}, n)$, that is, $\mathrm{URP} \vdash \neg\,\mathrm{Pr}(\overline{\ulcorner G\urcorner})$, from which we deduce $\mathrm{URP} \vdash G$. □

So URP is stronger than PA. By how much?

THEOREM 6.2.4. *Writing URP_{Π_1} for the axiom system got by adding to PA only instances of the reflection principle for Π_1 formulae, $\mathrm{PA} \cup \mathrm{URP}_{\Pi_1}$ is equivalent to $\mathrm{PA} \cup \{\mathrm{Con}_{\mathrm{PA}}\}$.*

PROOF: Assume $\mathrm{PA} \cup \mathrm{URP}_{\Pi_1}$. We set out to prove $\mathrm{Con}_{\mathrm{PA}}$.

Recall that $\mathrm{Con}_{\mathrm{PA}}$ is $\neg\,\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner\neg\overline{0} = \overline{0}\urcorner})$, which is $\forall x\, \neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner\neg\overline{0} = \overline{0}\urcorner}, x)$.

Recalling that $\mathrm{proof}_{\mathrm{PA}}$ is Δ_1, express $\neg\,\mathrm{proof}_{\mathrm{PA}}(\overline{\ulcorner\neg\overline{0} = \overline{0}\urcorner}, x)$ as $\forall y\, \psi(x, y)$.

So, $\mathrm{Con}_{\mathrm{PA}}$ can be written as $\forall z\, \forall x \le z, \forall y \le z\, \psi(x, y)$.

Now $\neg\overline{0} = \overline{0} \to \forall x \le z\, \forall y \le z\, \psi(x, y)$ is an instance of a tautology.

It follows that if for some z, $\neg\,\mathrm{Pr}_{\mathrm{PA}}(\forall x \le z, \forall y \le z\, \psi(x, y))$ is true in a particular model of PA, then so is $\neg\,\mathrm{Pr}(\overline{\ulcorner\neg\overline{0} = \overline{0}\urcorner})$; that is, $\mathrm{Con}_{\mathrm{PA}}$ is true, as desired.

If, on the other hand, $\forall z\ \mathrm{Pr}_{\mathrm{PA}}(\forall x \le z, \forall y \le z\, \psi(x, y))$ holds in the model, then by URP_{Π_1}, we have $\forall z\, \forall x \le z\, \forall y \le z\, \psi(x, y)$, and thus we again have $\mathrm{Con}_{\mathrm{PA}}$.

Now assume that $\mathrm{PA} \cup \{\mathrm{Con}_{\mathrm{PA}}\}$ is true in a particular structure $\mathfrak{N}$.

Suppose that $F(x)$ is Π_1, and that in $\mathfrak{N}$, $\forall x\ \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner F(x) \urcorner})$ is true.

Suppose that $\mathfrak{N} \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \neg F(\bar{\bar{x}}) \urcorner})$.

Then using 5.1.1. and 5.1.3., and the fact that $(F(x) \to (\neg F(x) \to \neg \bar{0} = \bar{0}))$ is a tautology, we conclude that $\mathfrak{N} \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \neg \bar{0} = \bar{0} \urcorner})$, contradicting the assumption that $\mathrm{Con}_{\mathrm{PA}}$ is true in $\mathfrak{N}$.

Hence $\mathfrak{N} \vDash \forall x \neg \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \neg F(\bar{\bar{x}}) \urcorner})$.

Now $\neg F(x)$ is provably Σ_1, so $\mathrm{PA} \vdash \forall n\, (\neg F(n) \to \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \neg F(\bar{\bar{n}}) \urcorner}))$.

So we deduce that in $\mathfrak{N}$, $\forall n\, F(n)$ holds, as required. $\Box$

7. Gödel-Löb logic

Here we abstract out some of the features of the logic of provability we've been deriving, finding that a surprisingly small part of it is sufficient to give us the Incompleteness Theorems.

7.1. Definitions and basic results

Definition 7.1.1. *Gödel-Löb logic is a system of* modal propositional logic*.*

The symbols are: a countably infinite number of propositional variables p, q, r etc; a logical constant $\perp$, a binary connective $\to$, and a unary operator $\Box$.

The formulae are: all propositional variable letters; the symbol $\perp$; and all strings $(\phi \to \psi)$ and $\Box\,\phi$ where ϕ and ψ are formulae.

The logical axioms are all propositional tautologies (with $\perp$ interpreted as a contradiction), together with all instances of $\Box(\phi \to \psi) \to (\Box\,\phi \to \Box\,\psi)$, and $\Box(\Box\,\phi \to \phi) \to \Box\,\phi$. The rules of inference are modus ponens and necessitation*, by which we mean the rule "if $\vdash \phi$, then $\vdash \Box\,\phi$".*

$\Box\,\phi$ is to be interpreted "ϕ is provable".

The following theorem (whose proof is not examinable) shows that the abstraction process is very successful.

Theorem 7.1.2. *Suppose that Σ is a set of formulae of GL logic and ϕ is a formula of GL logic.*

Then $\Sigma \vdash \phi$ if and only if whenever $\psi \mapsto \psi^*$ is a map from formulae of GL logic to formulae of $\mathscr{L}$ having the properties that $(\neg\psi)^* = \neg\psi^*$, that $(\psi \to \chi)^* = (\psi^* \to \chi^*)$, and that $(\Box\,\psi)^* = \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \psi^* \urcorner})$, $\mathrm{PA} \cup \{\sigma^* : \sigma \in \Sigma\} \vdash \phi^*$.

Proof: The forward direction is relatively easy. The reverse direction involves clever use of what is known as *Kripke frames*, which are not on the syllabus of this course (unfortunately). $\Box$

The following feature of propositional logic carries over.

Proposition 7.1.3. *(Substitution) Suppose that ϕ, χ, ψ and θ are formulae of Gödel-Löb logic, and that θ' is obtained from θ by replacing one or more subformulae of θ that are copies of χ, by copies of ψ.*

Then $\vdash ((\phi \to (\chi \leftrightarrow \psi)) \to (\phi \to (\theta \leftrightarrow \theta')))$.

Proof: Induction on the complexity of θ. $\Box$

PROPOSITION 7.1.4. *(Modalised substitution) Suppose that $X = X(p)$ is a formula in which p only occurs within the scope of $\Box$ operators, and let $X(q)$ be the result of replacing all instances of p in X by q.*

Then $\vdash (\Box(p \leftrightarrow q) \rightarrow (X(p) \leftrightarrow X(q))$.

PROOF: Induction on the complexity of X. $\Box$

7.2. The fixed-point theorem for GL logic

Fixed point theorem; more abstract proof of incompleteness.

THEOREM 7.2.1. *Fixed point theorem: if $A(p)$ is a formula in which p only occurs in the scope of a $\Box$, then there exists a formula X, in which p does not occur and containing only letters from $A(\cdot)$, such that $X \leftrightarrow A(X)$ is provable.*

Moreover, X is unique in the sense that $\vdash ((\Box(p \leftrightarrow A(p)) \wedge \Box(q \leftrightarrow A(q))) \rightarrow \Box(p \leftrightarrow q))$.

LEMMA 7.2.2. *If $B(p)$ is a formula, then there exists a formula X, in which p does not occur and containing only letters from $A(\cdot)$, such that $X \leftrightarrow \Box B(X)$ is provable.*

PROOF: Then the appropriate X is: $\Box B(\top)$, where $\top$ is some tautology (such as $\bot \rightarrow \bot$).

Proof: $\Box B(\top) \rightarrow (\top \leftrightarrow \Box B(\top))$ is a tautology.

Thus, using substitution, so is $\Box B(\top) \rightarrow (\Box B(\top) \leftrightarrow \Box B(\Box B(\top)))$.

So we get $\Box B(\top) \rightarrow \Box B(\Box B(\top))$.

As for the other way round, given $\Box B(\top) \rightarrow (\top \leftrightarrow \Box B(\top))$, we use substitution again to get $\Box B(\top) \rightarrow (B(\top) \leftrightarrow B(\Box B(\top)))$.

It follows by propositional logic that $B(\Box B(\top)) \rightarrow (\Box B(\top) \rightarrow B(\top))$.

By necessitation, $\Box(B(\Box B(\top)) \rightarrow (\Box B(\top) \rightarrow B(\top)))$.

Using the first axiom and MP, $\Box B(\Box B(\top)) \rightarrow \Box(\Box B(\top) \rightarrow B(\top))$.

The second axiom scheme gives us $\Box(\Box B(\top) \rightarrow B(\top)) \rightarrow \Box B(\top)$.

Now by propositional logic, $\Box B(\Box B(\top)) \rightarrow \Box B(\top)$ as required. $\Box$

LEMMA 7.2.3. *Given a set of formulae $C_i(D(p_1, \ldots, p_n))$ $(i \leq n)$, there exist formulae F_i for $i \leq n$ such that $\vdash (F_i \leftrightarrow \Box C_i(D(F_1, \ldots, F_n)))$.*

PROOF: We do induction on n.

The base case was done above. Suppose that any such family of equivalences of size n can be solved, and suppose that we have a family of formulae $C_i(D(p_1, \ldots, p_{n+1}))$ $(i \leq n+1)$.

Then we set C_{n+1} aside for a moment. Let q be some propositional letter we have not yet used. Using the inductive hypothesis, let $G_i(q)$ (for $i \leq n$) be formulae such that

$$\vdash G_i(q) \leftrightarrow \Box C_i(D(G_1(q), \ldots, G_n(q), q))$$

$(i \leq n)$.

Now use the preceding lemma to find F_{n+1} such that

$$\vdash G_{n+1} \leftrightarrow \Box C_{n+1}(D(G_1(F_{n+1}), \ldots, G_n(F_{n+1}), F_{n+1})).$$

Now, for $i \leq n$, let $F_i = G_i(F_{n+1})$. $\Box$

[I've edited the above lemma to introduce the letter D. I believe that the previous version was correct, bar a missing $\Box$ which I've also introduced, but confusing in context.]

LEMMA 7.2.4. *(Existence of the fixed point) If $A(p)$ is a formula in which p only occurs in the scope of a $\Box$, then there exists a formula X, in which p does not occur and containing only letters from $A(\cdot)$, such that $X \leftrightarrow A(X)$ is provable.*

PROOF: Suppose that $A(p)$ has the form $D(\Box C_1(p), \ldots, \Box C_n(p))$. Use the preceding lemma to find F_i equivalent to $\Box C_i(D(F_1, \ldots, F_n))$ for $i \leq n$; then $D(F_1, \ldots, F_n)$ is equivalent to $D(\Box C_1(D(F_1, \ldots, F_n)), \ldots, \Box C_n(D(F_1, \ldots, F_n)))$, that is, to $A(D(F_1, \ldots, F_n))$, which is what we want. $\Box$

LEMMA 7.2.5. *(Uniqueness of the fixed point) Suppose that $A(p)$ is a formula in which p only occurs in the scope of a $\Box$, and that X is a formula in which p does not occur and containing only letters from $A(\cdot)$, such that $X \leftrightarrow A(X)$ is provable.*

Then X is unique in the sense that $\vdash ((\Box(p \leftrightarrow A(p)) \land \Box(q \leftrightarrow A(q))) \to \Box(p \leftrightarrow q))$.

PROOF: We prove, using modalised substitution, that $\vdash \Box(p \leftrightarrow q) \to (A(p) \leftrightarrow A(q))$.

By propositional logic, $\vdash (((p \leftrightarrow A(p)) \land (q \leftrightarrow A(q))) \to (\Box(p \leftrightarrow q) \to (p \leftrightarrow q)))$.

Doing stuff with $\Box$, we get $\vdash ((\Box(p \leftrightarrow A(p)) \land \Box(q \leftrightarrow A(q))) \to \Box(\Box(p \leftrightarrow q) \to (p \leftrightarrow q)))$.

Then, using an axiom, $\vdash (\Box(p \leftrightarrow A(p)) \land \Box(q \leftrightarrow A(q))) \to \Box(p \leftrightarrow q)$. $\Box$

As an example of silly things happen if p is not boxed, let A be the identity.

7.3. The incompleteness theorems in GL logic

THEOREM 7.3.1. (GL version of the First Incompleteness Theorem). *There exists a formula G such that $\vdash (G \leftrightarrow \neg \Box G)$.*

PROOF: Define G to be a/the fixed point of $\neg \Box p$; that is, $\vdash G \leftrightarrow \neg \Box G$. $\Box$

THEOREM 7.3.2. (GL version of the Second Incompleteness Theorem). *For any formulae A and B, $\vdash \Box \neg \Box A \to \Box B$.*

PROOF: Now $\vdash (\neg \Box A \to (\Box A \to A))$, so $\vdash (\Box \neg \Box A \to \Box(\Box A \to A))$, so $\vdash (\Box \neg \Box A \to \Box A)$. So since, by Theorem 7.1.2., $\vdash (\Box A \to \Box \Box A)$, $\vdash (\Box \neg \Box A \to \Box \Box A)$.

Now $(\neg \Box A \to (\Box A \to B))$ by propositional calculus.

Hence, using Necessitation, the scheme $(\Box(\phi \to \psi) \to (\Box \phi \to \Box \psi))$, and MP,

$$\vdash (\Box \neg \Box A \to (\Box \Box A \to \Box B)).$$

Then using propositional calculus,

$$\vdash (\Box \neg \Box A \to \Box B)$$

as required. $\Box$

The formula $\Box \neg \Box A \to \Box B$ expresses the idea that if anything is provably unprovable, then the system is inconsistent.

8. Constructing models of PA inside other models of PA

8.1. Revision of Henkin's proof of the Completeness Theorem

Recall how the now standard proof of the Completeness Theorem goes. Given a countable language L of first-order predicate calculus, we close the language under the addition of constant symbols c_ϕ for formulae ϕ, and add to PA (or whatever other theory we may be interested in) extra axioms $(\exists v_1\, \phi \to \phi(c_\phi))$. We then extend our theory to a complete consistent set of sentences, and a new model for the theory is then constructed from the closed terms (sc. terms containing no variable letters), in such a way that if $\phi(\sigma_1, \ldots, \sigma_n)$ is a member of the complete consistent set, then $\phi(\sigma_1, \ldots, \sigma_n)$ is true in that model.

So, once we have a process for deciding which sentences are members of our complete consistent set, the construction of the model is routine.

We focus on this process of identifying members of the complete consistent set.

First, we add the constant symbols to our language $\mathscr{L}$. We can do this in any number of ways. The method in the following definition relies on the fact that the string $\overline{0}'$ never occurs in formulae of $\mathscr{L}$.

Definition 8.1.1. *We define a language $\mathscr{L}^*$ with the same alphabet as $\mathscr{L}$, with its set of terms and its set of formulae being the smallest sets which have the following properties.*

(1) Every term of $\mathscr{L}$ is a term of $\mathscr{L}^*$, and every formula of $\mathscr{L}$ is a formula of $\mathscr{L}^*$.
(2) If ϕ is a formula of $\mathscr{L}^*$, then $\overline{0}'(\phi)$ is a term of $\mathscr{L}^*$ (which we will write as c_ϕ).
(3) If σ and τ are terms of $\mathscr{L}^*$, then so are σ^+, $f(\sigma\tau)$ and $f'(\sigma\tau)$.
(4) If σ and τ are terms of $\mathscr{L}^*$, then $\sigma = \tau$ and $\sigma \leq \tau$ are formulae of $\mathscr{L}^*$.
(5) If ϕ and ψ are formulae of $\mathscr{L}^*$, then so are $\neg\phi$, $(\phi \to \psi)$, and $\forall v_i \phi$.

We now add witnesses to existential statements to any theory of $\mathscr{L}$.

Definition 8.1.2. *If T is any set of sentences of $\mathscr{L}$, then we define T^* to be the result of adding to T all formulae $(\exists v_1\, \phi(v_1) \to \phi[c_\phi])$.*

Theorem 8.1.3. *If T is Σ_n-definable for $n \geq 1$, then T^* is also Σ_n-definable. Similarly if T is Π_n-definable for $n \geq 1$, then T^* is also Π_n-definable.*

Proof: The set of Gödel numbers of the extra axioms in T^* is Δ_1-definable. $\Box$

We now take a closer look at the process of deriving the complete consistent set.

Theorem 8.1.4. *Suppose that T has a Δ_n-definable proof predicate Pr_T. Then there is a Δ_n-definable function $\mathfrak{H}_T(n)$ such that if n is the Gödel number of a sentence ϕ of $\mathscr{L}^*$, then in a model $\mathfrak{N}$ of PA, $\mathfrak{H}_T(\phi) = 1$ if, defining θ_n to be the formula*

$$\theta_n = \bigwedge\{\psi : \ulcorner\psi\urcorner < n \wedge \mathfrak{H}_T(\ulcorner\psi\urcorner) = 1\} \wedge \bigwedge\{\neg\psi : \ulcorner\psi\urcorner < n \wedge \mathfrak{H}_T(\ulcorner\psi\urcorner) = 0\},$$

$\mathrm{Pr}_T(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$ is true, and $\mathfrak{H}_T(n) = 0$ otherwise.

We describe $\mathfrak{H}_T$ as successful *if there exists n such that $\mathfrak{H}_T(n) = 0$.*

Note that $\mathfrak{H}_T$ is successful in a model $\mathfrak{N}$ of PA if and only if $\mathfrak{N} \vDash \mathrm{Con}_T$.

Proof: To see that $\mathfrak{H}_T$ is Δ_n-definable, note that the definition of $\mathfrak{H}_T$ is a definition by recursion using a Δ_n-formula. $\Box$

If $\mathfrak{H}_T$ is successful in a model $\mathfrak{N}$, then we think of $\mathfrak{H}_T$ as describing a model of T inside $\mathfrak{N}$. It is certainly the case that from a successful function $\mathfrak{H}_T$, we can construct a model of T.

8.2. Comparing a model and a model inside that model

When we start to talk about models of arithmetic sitting inside models of arithmetic, questions about language and metalanguage, and problems posed by non-standard elements coding formulae and proofs, become complicated. For some results in this section, we restrict our attention to models of arithmetic sitting inside $\mathbb{N}$.

THEOREM 8.2.1. *If $\mathfrak{N}$ is a model of PA defined in $\mathbb{N}$ by a formula $\mathfrak{H}_T$, where T extends PA, then $\mathfrak{N}$ is not elementarily equivalent to $\mathbb{N}$ (that is, there is a sentence which is true in $\mathfrak{N}$ and false in $\mathbb{N}$).*

PROOF: This follows from Tarski's Theorem. If $\mathfrak{N}$ and $\mathbb{N}$ were elementarily equivalent, then $\mathfrak{H}_T$ would define truth in $\mathbb{N}$, which is impossible. $\square$

This theorem can be extended, with extreme caution, to other models of PA.

THEOREM 8.2.2. *If ϕ is a sentence of $\mathscr{L}$, and $\mathfrak{N}$ is a model of PA, then there is a model $\mathfrak{N}'$ of PA inside $\mathfrak{N}$ satisfying ϕ if and only if $\mathfrak{N} \vDash \neg\,\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \neg\phi \urcorner})$, and every model $\mathfrak{N}'$ of PA inside $\mathfrak{N}$ satisfies ϕ if and only if $\mathfrak{N} \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner \phi \urcorner})$.*

Restricting our attention to $\mathfrak{H}_{\mathrm{PA}}$, we have the following.

THEOREM 8.2.3. *There is a Δ_2-sentence K such that if $\mathfrak{N}$ is a model of PA and $\mathfrak{N}'$ is constructed inside $\mathfrak{N}'$ using $\mathfrak{H}_{\mathrm{PA}}$, then K is true in $\mathfrak{N}$ if and only if it is false in $\mathfrak{N}'$.*

PROOF: Use the Diagonal Lemma to find K such that $PA \vdash (K \leftrightarrow \neg\,\mathfrak{H}_{\mathrm{PA}}(\overline{\ulcorner K \urcorner}))$.

We can regard K as being Δ_2 because $\mathfrak{H}_{\mathrm{PA}}$ is. $\square$

COROLLARY 8.2.4. *Any chain $\mathfrak{N}_i$ of models of PA, where $\mathfrak{N}_{i+1}$ is constructed inside $\mathfrak{N}_i$ using $\mathfrak{H}_{\mathrm{PA}}$, is finite.*

PROOF: We order functions from $\mathbb{N}$ to $\{0,1\}$ lexicographically, that is, so that $f \leq g$ iff either $f = g$, or there exists n such that for all $m < n$, $f(m) = g(m)$, and $f(n) < g(n)$.

We define f_i so that for all $n \in \mathbb{N}$, $f_i(n) = 1$ if and only if n is the Gödel number of a sentence ϕ such that $\mathfrak{N}_i \vDash \phi$.

Let $k = \overline{\ulcorner K \urcorner}$. Then for all i, $f_i(k) \neq f_{i+1}(k)$.

We now note that if $i > 1$, then $f_i \leq f_{i+1}$. For, suppose that n is least such that $f_i(n) \neq f_{i+1}(n)$. Then in one of $\mathfrak{N}_{i-1}$ and $\mathfrak{N}_i$ but not the other, $\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$, where ϕ is the formula whose Gödel number is n.

Then if $m = \ulcorner \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})\urcorner$, then $\mathfrak{N}_{i-1} \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_m \to \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner}))\urcorner})$, so that $\mathfrak{N}_{i-1} \vDash \mathfrak{H}_{\mathrm{PA}}(\overline{\ulcorner \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})\urcorner})$, so that $\mathfrak{N}_i \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$, and then it follows that $\mathfrak{N}_i \vDash \mathfrak{H}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$, giving a contradiction.

So it must be the case that $\mathfrak{N}_{i-1} \vDash \neg\,\mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$ while $\mathfrak{N}_i \vDash \mathrm{Pr}_{\mathrm{PA}}(\overline{\ulcorner(\theta_n \to \phi)\urcorner})$, so that $f_i(n) = 0$ and $f_{i+1}(n) = 1$.

We also note two other facts. Firstly, we must have $n \leq k$, since $f_i(k) \neq f_{i+1}(k)$. Also, we must have that for all $j \geq i$, $f_j(n) = 1$.

So, $1 < i < j$ implies that $f_i < f_j$. Moreover, there exists $n \leq k$ such that $f_i(n) = 0$ and $f_j(n) = 1$. So the functions $f_i \restriction \{n : n \leq k\}$ are all different, for $i > 1$, so there are only finitely many of them. $\square$

9. Stuff off the end of the course

9.1. Kripke semantics

Write L for the axiom scheme $(\Box(\Box\phi \to \phi) \to \Box\phi)$.

Proposition 9.1.1. *L gives that the accessibility relation is transitive.*

Proof: Suppose $w\, R\, w'\, R\, w''$, but $w \not R\, w''$.

Declare p to be false at these three worlds and true at all others.

Then $w' \vDash \neg\Box p$, so $w' \vDash (\Box p \to p)$.

Then $w \vDash \neg\Box p$, and $w \vDash \Box(\Box p \to p)$ (since at all other worlds accessible from w, p is true, so $(\Box p \to p)$ is true). Also, $w \vDash \neg\Box p$. $\square$

Proposition 9.1.2. *L gives that the accessibility relation is reverse-well-founded.*

Proof: Suppose for all $i \in \mathbb{N}$, $w_i\, R\, w_{i+1}$.

Declare p to be false at all w_i, and true at all other worlds.

Then for all i, $w_i \vDash \neg\Box p$.

Thus at every world, $(\Box p \to p)$ is true, either because p is true, or because $\Box p$ is false.

Thus $w_0 \vDash \Box(\Box p \to p)$, and $w_0 \vDash \neg\Box p$. $\square$

Proposition 9.1.3. *If the accessibility relation is transitive and reverse-well-founded, then L is true.*

Proof: Easy induction on the rank of a world. $\square$

Proposition 9.1.4. $\vdash_{\mathrm{GL}} (\Box X \to \Box\Box X)$.

Proof: $\vdash (X \to ((\Box\Box X \wedge \Box X) \to (\Box X \wedge X)))$: propositional tautology.

$\vdash (X \to (\Box(\Box X \wedge X) \to (\Box X \wedge X)))$: standard stuff.

$\vdash (\Box X \to \Box(\Box(\Box X \wedge X) \to (\Box X \wedge X)))$: necessitation plus standard stuff.

$(\Box(\Box(\Box X \wedge X) \to (\Box X \wedge X)) \to \Box(\Box X \wedge X))$: instance of L.

$\Box(\Box X \to \Box(\Box X \wedge X))$.

$(\Box X \to \Box\Box X)$. $\square$