Chapter 6

Gödel's 2^{nd} Incompleteness Theorem

6.1 Peano Arithmetic and $I\Sigma_1^0$

To prove Gödel's 2^{nd} incompleteness theorem we need to work in a theory slightly more expressive than $\mathcal{R}ob.$ and slightly less than the theory of *Peano arithmetic* that we know introduce.

Definition 1.1: Peano arithmetic

$\mathcal{P}$*eano arithmetic* is a theory based on the same language as *Robinson arithmetic* : $\mathcal{L}_{\mathcal{A}} = \{0, S, +, \cdot\}$. But contrary to *Robinson arithmetic*, it has infinitely many axioms:

axiom 1. $\forall x\ Sx \neq 0$

axiom 2. $\forall x\ \exists y\ (x \neq 0 \rightarrow Sy = x)$

axiom 3. $\forall x\ \forall y\ (Sx = Sy \rightarrow x = y)$

axiom 4. $\forall x\ x{+}0 = x$

axiom 5. $\forall x\ \forall y\ \Big(x{+}Sy = S(x{+}y)\Big)$

axiom 6. $\forall x\ x{\cdot}0 = 0$

axiom 7. $\forall x\ \forall y\ \Big(x{\cdot}Sy = (x{\cdot}y){+}x\Big)$

axiom schema (induction) for any formula $\varphi_{[x_0,x_1,\ldots,x_n]}$ [a],

$$\forall x_1\ \ldots\ \forall x_n \left(\left(\varphi_{[0/x_0,x_1,\ldots,x_n]} \wedge \forall x_0\ \left(\varphi_{[x_0,x_1,\ldots,x_n]} \rightarrow \varphi_{[Sx_0/x_0,x_1,\ldots,x_n]}\right)\right) \rightarrow \forall x_0\ \varphi_{[x_0,x_1,\ldots,x_n]}\right)$$

[a] the notation $\varphi_{[x_0,x_1,\ldots,x_n]}$ means that the free variable of φ are all among $x_0, x_1, \ldots, x_n$.

So we see that $\mathcal{P}eano$ is nothing but $\mathcal{R}ob.$ augmented with the induction schema for all formulas constructed on the language of arithmetic. In fact we will not need to work within $\mathcal{P}eano$ but only a fragment of it obtained by restricting the induction schema to the sole Σ^0_1-formulas (see next section). This theory is called $\mathcal{R}ob.{+}I\Sigma^0_1$

Example 1.1

We saw that $\mathcal{R}ob.$ does not prove that the addition is commutative. We want to prove, here, that within $\mathcal{R}ob.{+}I\Sigma^0_1$ the addition becomes commutative. For this purpose we make use of several instances of the induction schema.

(1) We first show that

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x \;\; x{+}0 = 0{+}x.$$

Indeed we have both

- $\vdash_c 0{+}0 = 0{+}0$
- $\mathcal{R}ob. \vdash_c \forall x \;\; \Big(\big(x{+}0 = 0{+}x\big) \rightarrow \big(Sx{+}0 = 0{+}Sx\big)\Big)$ because we have by

$$\boxed{④ \;\; \forall x \; x{+}0 = x} \quad \textit{and} \quad \boxed{⑤ \;\; \forall x \; \forall y \; \big(x{+}Sy = S(x{+}y)\big)}$$

$$\mathcal{R}ob. \vdash_c Sx{+}0 = Sx \;\wedge\; 0{+}Sx = S(0{+}x)$$

hence

$$\mathcal{R}ob., \;\; x{+}0 = 0{+}x \vdash_c Sx{+}0 = Sx \;\wedge\; 0{+}Sx = S(x{+}0) \;\wedge\; S(x{+}0) = Sx$$

So by applying the induction schema to the Δ^0_0-formula $x{+}0 = 0{+}x$ we obtain the result.

(2) We then show that

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x \; \forall y \;\; x{+}Sy = Sx{+}y.$$

Indeed we have both

- $\mathcal{R}ob. \vdash_c \forall x \;\; x{+}S0 = Sx{+}0$ by

$$\boxed{④ \;\; \forall x \; x{+}0 = x} \quad \textit{and} \quad \boxed{⑤ \;\; \forall x \; \forall y \; \big(x{+}Sy = S(x{+}y)\big)}$$

- $\mathcal{R}ob. \vdash_c \forall x \ \Big((x+Sy = Sx+y) \to (x+SSy = Sx+Sy)\Big)$ because we have by

$$\boxed{⑤ \quad \forall x \ \forall y \ \big(x+Sy = S(x+y)\big)}$$

$$\mathcal{R}ob. \vdash_c x+SSy = S(x+Sy)$$

hence

$$\mathcal{R}ob., \quad x+Sy = Sx+y \ \vdash_c \ S(x+Sy) = S(Sx+y) = Sx+Sy.$$

So by applying the induction schema to the Δ_0^0-formula $x+Sy = Sx+y$ we obtain the result.

(3) Finally we show that

$$\mathcal{R}ob.+I\Sigma_1^0 \ \vdash_c \ \forall x \ \forall y \ \ x+y = y+x.$$

Indeed we have both

- $\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x \ \ x+0 = 0+x$ for this was what we established in case (1).
- $\mathcal{R}ob. \vdash_c \forall x \ \Big((x+y = y+x) \to (x+Sy = Sy+x)\Big)$ because by

$$\boxed{⑤ \quad \forall x \ \forall y \ \big(x+Sy = S(x+y)\big)}$$

we have

$$\mathcal{R}ob., \quad x+y = y+x \ \vdash_c \ x+Sy = S(x+y) = S(y+x) = y+Sx$$

hence by applying case (2) we obtain

$$\mathcal{R}ob.+I\Sigma_1^0, \quad x+y = y+x \ \vdash_c \ x+Sy = Sy+x.$$

So, in the end, by applying the induction schema to the Δ_0^0-formula $x+y = y+x$ we obtain the result.

Example 1.2

We saw that $\mathcal{R}ob.$ proves that every integer (standard or non-standard) is always comparable with any standard integer:

$$\boxed{4.11 \quad \mathcal{R}ob. \vdash_c \forall x \ \ (x \leqslant n \ \vee \ n \leqslant x)}$$

Now we establish that $\mathcal{R}ob.+I\Sigma_1^0$ proves that any two integers are always comparable:

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x\ \forall y\ \big(x \leqslant y \ \vee\ y \leqslant x\big).$$

We recall that $x \leqslant y$ stands for $\exists z\ z+x = y$. So we consider the instance of the axiom schema for the Σ_1^0-formula

$$\varphi[x,y] := \exists z\ z+x = y \ \vee\ \exists z\ z+y = x.$$

$$\forall y\ \left(\left(\begin{array}{c} (\exists z\ z+0 = y \ \vee\ \exists z\ z+y = 0) \\ \wedge \\ \forall x\ \big((\exists z\ z+x = y \ \vee\ \exists z\ z+y = x) \rightarrow (\exists z\ z+Sx = y \ \vee\ \exists z\ z+y = Sx)\big) \end{array}\right) \rightarrow \forall x \left(\begin{array}{c} \exists z\ z+x = y \\ \vee \\ \exists z\ z+y = x \end{array}\right)\right)$$

(1) By $\boxed{4 \quad \forall x\ x+0 = x}$ have

$$\mathcal{R}ob. \vdash_c \exists z\ z+0 = y$$

which takes care of the first part: $\exists z\ z+0 = y \ \vee\ \exists z\ z+y = 0$.

(2) For the second part we need to distinguish between two cases:

if $\exists \mathbf{z}\ \mathbf{z+y = x}$ we distinguish between $z = 0$ and $z \neq 0$

if $\mathbf{0+y = x}$ by Example 1.1, we have

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c 0+y = x \rightarrow x = y$$

and

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c x = y \rightarrow S0+y = 0+Sy = Sy = Sx$$

if $\exists \mathbf{z \neq 0}\ \mathbf{z+x = y}$ then

$$\mathcal{R}ob.,\ \exists z \neq 0\ z+x = y \vdash_c \exists z'\ Sz'+x = y$$

By Example 1.1, we obtain what we need:

$$\mathcal{R}ob.+I\Sigma_1^0,\ \exists z \neq 0\ z+x = y \vdash_c \exists z'\ z'+Sx = y$$

if $\exists \mathbf{z}\ \mathbf{z+x = y}$ By $\boxed{5 \quad \forall x\ \forall y\ \big(x+Sy = S(x+y)\big)}$ we have

$$\mathcal{R}ob.,\ z+x = y \vdash_c S(z+y) = z+Sy = Sx$$

and by Example 1.1, we have

$$\mathcal{R}ob.{+}I\Sigma_1^0 \ \vdash_c \ z{+}Sy = Sx \to Sz{+}y = Sx$$

which gives the result we need:

$$\mathcal{R}ob.{+}I\Sigma_1^0, \ \ \exists z \ z{+}x = y \ \vdash_c \ \exists z \ z{+}y = Sx$$

So, in the end, by applying the induction schema to the Σ_1^0-formula

$$\varphi[x, y] := \exists z \ z{+}x = y \ \ \vee \ \ \exists z \ z{+}y = x$$

we obtain the result.

Example 1.3

We saw that $\mathcal{R}ob.$ does not prove that the addition is associative. We show here that $\mathcal{R}ob.{+}I\Sigma_1^0$ proves that the addition is associative:

$$\mathcal{R}ob.{+}I\Sigma_1^0 \ \vdash_c \ \forall x \ \forall y \ \forall z \ \ (x{+}y){+}z = x{+}(y{+}z).$$

(1) We first show that

$$\mathcal{R}ob. \ \vdash_c \ \forall x \ \forall y \ \ (x{+}y){+}0 = x{+}(y{+}0).$$

Indeed by $\boxed{4 \quad \forall x \ x{+}0 = x}$ we have

$$\mathcal{R}ob. \ \vdash_c \ \ (x{+}y){+}0 = x{+}y = x{+}(y{+}0).$$

(2) We then show that

$$\mathcal{R}ob. \ \vdash_c \ \forall x \ \forall y \ \forall z \ \ \Big((x{+}y){+}z = x{+}(y{+}z)\Big) \to \Big((x{+}y){+}Sz = x{+}(y{+}Sz)\Big).$$

by $\boxed{5 \quad \forall x \ \forall y \ \big(x{+}Sy = S(x{+}y)\big)}$ we have

$$\mathcal{R}ob. \ \vdash_c \ (x{+}y){+}Sz = S\big((x{+}y){+}z\big)$$

and also

$$\mathcal{R}ob. \ \vdash_c \ S\big(x{+}(y{+}z)\big) = x{+}S(y{+}z) = x{+}(y{+}Sz)$$

therefore we obtain

$$\mathcal{R}ob., \ \ (x{+}y){+}z = x{+}(y{+}z) \ \vdash_c \ (x{+}y){+}Sz = x{+}(y{+}Sz).$$

(3) Finally, by applying the induction schema to the Δ^0_0-formula $(x+y)+z = x+(y+z)$ we obtain the result.

Example 1.4

We saw that $\mathcal{R}ob.$ does not prove that the multiplication is commutative. We show here $\mathcal{R}ob.+I\Sigma^0_1$ proves that the addition is commutative.

(1) We first show that
$$\mathcal{R}ob.+I\Sigma^0_1 \vdash_c \forall x \;\; x\cdot 0 = 0\cdot x.$$
Indeed we have both

- $\vdash_c 0\cdot 0 = 0\cdot 0$
- $\mathcal{R}ob. \vdash_c \forall x \;\; \Big((x\cdot 0 = 0\cdot x) \to (Sx\cdot 0 = 0\cdot Sx)\Big)$ because we have by

$$\boxed{⑥ \quad \forall x \; x\cdot 0 = 0} \quad \textit{and} \quad \boxed{⑦ \quad \forall x \; \forall y \; \big(x\cdot Sy = (x\cdot y)+x\big)}$$

$$\mathcal{R}ob. \vdash_c Sx\cdot 0 = 0 \;\wedge\; 0\cdot Sx = (0\cdot x)+0$$

hence
$$\mathcal{R}ob., \;\; x\cdot 0 = 0\cdot x \vdash_c Sx\cdot 0 = 0 = 0+0 = (0\cdot x)+0.$$

So by applying the induction schema to the Δ^0_0-formula $x\cdot 0 = 0\cdot x$ we obtain the result.

(2) We then show that
$$\mathcal{R}ob.+I\Sigma^0_1 \vdash_c \forall x \; \forall y \;\; Sx\cdot y = (x\cdot y)+y.$$

Indeed we have both

- $\mathcal{R}ob. \vdash_c \forall x \;\; Sx\cdot 0 = (x\cdot 0)+0$ by a simple application of

$$\boxed{④ \quad \forall x \; x+0 = x} \quad \textit{and} \quad \boxed{⑥ \quad \forall x \; x\cdot 0 = 0}$$

- $\mathcal{R}ob. \vdash_c \forall x \;\; \Big((Sx\cdot y = (x\cdot y)+y) \to (Sx\cdot Sy = (x\cdot Sy)+Sy)\Big)$ because we have by

$$\boxed{⑦ \quad \forall x \; \forall y \; \big(x\cdot Sy = (x\cdot y)+x\big)}$$

$$\mathcal{R}ob. \vdash_c Sx\cdot Sy = (Sx\cdot y)+Sx$$

hence

$$\mathcal{R}ob., \quad Sx \cdot y = (x \cdot y) + y \ \vdash_c \ Sx \cdot Sy = ((x \cdot y) + y) + Sx.$$

but we also know that the addition is associative and commutative, thus we have

$$\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ x \cdot y + (y + Sx) = x \cdot y + (y + Sx) = x \cdot y + (Sy + x)$$

and

$$\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ x \cdot y + (Sy + x) = x \cdot y + (x + Sy) = (x \cdot y + x) + Sy = x \cdot Sy + Sy.$$

So by applying the induction schema to the Δ_0^0-formula $Sx \cdot y = (x \cdot y) + y$ we obtain the result.

(3) Finally we show that

$$\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ \forall x \ \forall y \ \ x \cdot y = y \cdot x.$$

Indeed we have both

- $\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ \forall x \ \ x \cdot 0 = 0 \cdot x$ for this was what we established in case (1).
- $\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ \forall x \ \ \big((x \cdot y = y \cdot x) \rightarrow (x \cdot Sy = Sy \cdot x)\big)$ because by

$$\boxed{7 \quad \forall x \ \forall y \ (x \cdot Sy = (x \cdot y) + x)}$$

we have

$$\mathcal{R}ob., \quad x \cdot y = y \cdot x \ \vdash_c \ x \cdot Sy = (x \cdot y) + x = (y \cdot x) + x$$

by case (2) we have

$$\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ (y \cdot x) + x = Sy \cdot x$$

which leads to

$$\mathcal{R}ob. + I\Sigma_1^0, \quad x \cdot y = y \cdot x \ \vdash_c \ x \cdot Sy = Sy \cdot x$$

So, in the end, by applying the induction schema to the Δ_0^0-formula $x \cdot y = y \cdot x$ we obtain the result.

Example 1.5

We show here that $\mathcal{R}ob. + I\Sigma_1^0$ proves that the multiplication distributes over the addition:

$$\mathcal{R}ob. + I\Sigma_1^0 \ \vdash_c \ \forall x \ \forall y \ \forall z \ \ x \cdot (y + z) = (x \cdot y) + (x \cdot z).$$

(1) We first show that

$$\mathcal{R}ob. \vdash_c \forall x\ \forall y\ \ x\cdot(y+0) = (x\cdot y)+(x\cdot 0).$$

which is immediate by

$$\boxed{④ \quad \forall x\ x+0 = x} \quad \textit{and} \quad \boxed{⑥ \quad \forall x\ x\cdot 0 = 0}$$

(2) We then show that

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x\ \forall y\ \forall z\ \ \big(x\cdot(y+z) = (x\cdot y)+(x\cdot z)\big) \rightarrow \big(x\cdot(y+Sz) = (x\cdot y)+(x\cdot Sz)\big).$$

by $\boxed{⑦ \quad \forall x\ \forall y\ \big(x\cdot Sy = (x\cdot y)+x\big)}$ we have

$$\mathcal{R}ob. \vdash_c (x\cdot y)+(x\cdot Sz) = (x\cdot y)+\big((x\cdot z)+x\big)$$

and

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c (x\cdot y)+\big((x\cdot z)+x\big) = \big((x\cdot y)+(x\cdot z)\big)+x$$

So that we obtain

$$\mathcal{R}ob.+I\Sigma_1^0,\ x\cdot(y+z) = (x\cdot y)+(x\cdot z) \vdash_c (x\cdot y)+(x\cdot Sz) = \big(x\cdot(y+z)\big)+x.$$

At last, by

$$\boxed{⑤ \quad \forall x\ \forall y\ \big(x+Sy = S(x+y)\big)} \quad \textit{and} \quad \boxed{⑦ \quad \forall x\ \forall y\ \big(x\cdot Sy = (x\cdot y)+x\big)}$$

we have

$$\mathcal{R}ob. \vdash_c \big(x\cdot(y+z)\big)+x = \big(x\cdot S(y+z)\big) = \big(x\cdot(y+Sz)\big)$$

which terminates this proof.

(3) Finally, by applying the induction schema to the Δ_0^0-formula $x\cdot(y+z) = (x\cdot y)+(x\cdot z)$ we obtain the result.

Example 1.6

We show that $\mathcal{R}ob.+I\Sigma_1^0$ proves that the multiplication is associative:

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x\ \forall y\ \forall z\ \ (x\cdot y)\cdot z = x\cdot(y\cdot z).$$

(1) We first show that

$$\mathcal{R}ob. \vdash_c \forall x\ \forall y\ \ (x\cdot y)\cdot 0 = x\cdot(y\cdot 0).$$

Indeed by ⑥ $\forall x\ x\cdot 0 = 0$ we have

$$\mathcal{R}ob.\ \vdash_c\ (x\cdot y)\cdot 0 = 0$$

and

$$\mathcal{R}ob.\ \vdash_c\ x\cdot(y\cdot 0) = x\cdot 0 = 0$$

(2) We then show that

$$\mathcal{R}ob.\ \vdash_c\ \forall x\ \forall y\ \forall z\ \ \big((x\cdot y)\cdot z = x\cdot(y\cdot z)\big) \to \big((x\cdot y)\cdot Sz = x\cdot(y\cdot Sz)\big).$$

by ⑦ $\forall x\ \forall y\ \big(x\cdot Sy = (x\cdot y)+x\big)$ we have

$$\mathcal{R}ob.\ \vdash_c\ x\cdot(y\cdot Sz) = x\cdot\big((y\cdot z)+y\big)$$

and by Example 1.5 we also have

$$\mathcal{R}ob.{+}I\Sigma_1^0\ \vdash_c\ x\cdot\big((y\cdot z)+y\big) = \big(x\cdot(y\cdot z)\big)+(x\cdot y)$$

so that we obtain

$$\mathcal{R}ob.{+}I\Sigma_1^0\ ,\ (x\cdot y)\cdot z = x\cdot(y\cdot z)\ \vdash_c\ \big(x\cdot(y\cdot z)\big)+(x\cdot y) = \big((x\cdot y)\cdot z\big)+(x\cdot y) = (x\cdot y)\cdot Sz$$

which gives the result.

(3) Finally, by applying the induction schema to the Δ_0^0-formula $(x\cdot y)\cdot z = x\cdot(y\cdot z)$ we obtain the result.

Proposition 1.1

The theory $\mathcal{R}ob.{+}I\Sigma_1^0$ proves that

(1) $+$ is commutative;

(2) $+$ is associative;

(3) $\cdot$ is commutative;

(4) $\cdot$ is associative;

(5) $\cdot$ distributes over $+$.

Proof of Proposition 1.1:

(1) "$+$ is commutative" is Example 1.1,

(2) "$+$ is associative" is Example 1.3,

(3) "$\cdot$ is commutative" is Example 1.4,

(4) "$\cdot$ is associative" is Example 1.6,

(5) "$\cdot$ distributes over $+$" is Example 1.5.

$\square$

6.2 The Arithmetical Hierarchy

For the purpose of defining the arithmetical hierarchy we add a binary symbol "$<$" to our language but essentially for the purpose of denoting bounded formulas such as $\exists y \leqslant t \;\; \varphi$ and $\forall y \leqslant t \; \varphi$. In a sense, this differs from the use of this same symbol inside Robinson arithmetic (see page 114) where it was an abbreviation for "$\exists y \; \big(y+x = z \;\wedge\; x \neq z\big)$". For the reason that in what follows we will have

- "$\exists y \; \big(y+x = z \;\wedge\; x \neq z\big)$" is a Σ^0_1-formula, and
- "$\exists x \leqslant z \;\; x \neq z$" is a Δ^0_0-formula.

We will be working with $\mathcal{R}ob.+I\Sigma^0_1$ so for every x and y we will have both

$$x \leqslant y \;\vee\; y \leqslant x$$

and

$$\exists z \;\; z+x = y \quad\Longleftrightarrow\quad \exists z \;\; x+z = y.$$

Definition 2.1: Δ^0_0-formulas

The set of Δ^0_0-formulas is the least that

(1) contains all atomic formulas: $t_0 = t_1$

(2) is closed under conjunctions, disjunctions and negations

(3) is closed under bounded quantifications. i.e.,

if $\varphi \in \Delta_0^0$ and t is a term, then $\forall x < t\ \varphi$ and $\exists x < t\ \ \varphi$ both belong to Δ_0^0.

Definition 2.2: arithmetical hierarchy

The hierarchy of formulas from arithmetic is defined by induction on $n \in \mathbb{N}$:

(1) $\Sigma_0^0 = \Pi_0^0 = \Delta_0^0$

(2) Σ_{n+1}^0 is the set of all formulas that are (logically equivalent to formulas of) the form $\exists x_1 \dots \exists x_k\ \varphi$ where $\varphi \in \Pi_n^0$.

(3) Π_{n+1}^0 is the set of all formulas that are (logically equivalent to formulas of) the form $\forall x_1 \dots \forall x_k \varphi$ where $\varphi \in \Sigma_n^0$.

(4) $\Delta_{n+1}^0 = \Sigma_{n+1}^0 \cap \Pi_{n+1}^0$

Notice that all the classes defined above are closed under (finite) conjunctions and disjonctions.

Example 2.1

(1) $x_0 < S(x_2 \cdot S x_1) \longrightarrow \forall y \leqslant x_3 \quad y + x_0 = x_3 \quad \in \ \Delta_0^0$

(2) $\forall x\ \forall y\ \Big(x \cdot S y = (x \cdot y) + x\Big) \quad \in \ \Pi_1^0$

(3) $\forall x\ \exists y\ (x \neq 0 \rightarrow S y = x) \quad \in \ \Pi_2^0$

Proposition 2.1

Given any $n \in \mathbb{N}$ and any Σ_1^0-formula $\varphi := \exists x_0\ \exists x_1 \dots \exists x_n \psi$ where ψ is Δ_0^0, there exists some Δ_0^0-formula ψ' such that

$$\mathcal{R}ob. \vdash_c \exists x_0\ \exists x_1 \dots \exists x_n \psi \longleftrightarrow \exists x \psi'.$$

Proof of Proposition 2.1:

We set

$$\psi' = \exists x_0 \leqslant x\ \exists x_1 \leqslant x \dots \exists x_n \leqslant x\ \Big(\alpha_{n+1}\big(x_0, x_1, \dots, x_n\big) = x\ \wedge\ \psi\Big)$$

where $\alpha_{n+1}(x_0, x_1, \dots, x_n) = x$ denotes the Δ^0_0-formula defined by induction on $n > 0$ by

- "$\alpha_2(x_0, x_1) = x$" is "$\exists y \leqslant x\ (x_0+x_1)\cdot(x_0+x_1+S0) = y+y\ \wedge\ y+x_1 = x$"
- "$\alpha_{n+1}(x_0, x_1, \dots, x_n) = x$" is "$\exists z \leqslant x\ \Big(\alpha_2(x_0, x_1) = z\ \wedge\ \alpha_n(x_0, x_1, \dots, x_{n-1}, z) = x\Big)$".

$\square$

Proposition 2.2

Given any $n, k_1, \dots, k_n \in \mathbb{N}$ and any Σ^0_{n+1}-formula and Π^0_{n+1}-formula respectively

$$\varphi := \exists x^n_1\ \exists x^n_2 \dots \exists x^n_{k_n} \forall x^{n-1}_1\ \dots \forall x^{n-1}_{k_{n-1}} \exists x^{n-2}_1 \dots \exists^{n-2}_{k_{n-2}} \forall x^{n-3}_1\ \dots\dots Q x^1_1 \dots Q x^1_{k_1}\ \psi$$

where Q is either $\forall$ or $\exists$ depending on the parity of n, and ψ is Δ^0_0, and

$$\theta := \forall x^n_1\ \forall x^n_2 \dots \forall x^n_{k_n} \exists x^{n-1}_1\ \dots \exists x^{n-1}_{k_{n-1}} \forall x^{n-2}_1 \dots \forall x^{n-2}_{k_{n-2}} \exists x^{n-3}_1\ \dots\dots \bar{Q} x^1_1 \dots \bar{Q} x^1_{k_1}\ \gamma$$

where $\bar{Q}$ is either $\forall$ or $\exists$ depending on the parity of n, and γ is Δ^0_0,

there exists Δ^0_0-formulas ψ' and γ' such that

- $\mathcal{R}ob. \vdash_c \varphi \longleftrightarrow \exists y_n\ \forall y_{n-1}\ \exists y_{n-2} \dots Q y_1 \psi'$, and
- $\mathcal{R}ob. \vdash_c \theta \longleftrightarrow \forall y_n\ \exists y_{n-1}\ \forall y_{n-2} \dots \bar{Q} y_1 \gamma'$.

Proof of Proposition 2.2:

This is an easy exercise based on the same idea as the one used to prove Proposition 2.1. $\square$

From now on, Γ stands for any of the classes Σ^0_{n+1}, Π^0_{n+1}, Δ^0_n (any $n \in \mathbb{N}$).

Proposition 2.3

Given any formula ψ,

- $\vdash_c\ \exists y \leqslant x\ \exists z\ \psi \longleftrightarrow \exists z\ \exists y \leqslant x\ \psi$

- $\vdash_c \ \forall y \leqslant x\ \forall z\ \psi \longleftrightarrow \forall z\ \forall y \leqslant x\ \psi$

Proof of Proposition 2.3:

Immediate. □

Proposition 2.4

Given any $n \in \mathbb{N}$, any term t that does not contain the variable z, and any formula $\psi \in \Gamma$ there exists $\psi' \in \Gamma$,

(1) $\mathcal{R}ob. \vdash_c \ \exists y \leqslant t\ \forall z \ \ \psi \longleftrightarrow \forall z\ \exists y \leqslant t \ \ \psi'$

(2) $\mathcal{R}ob. \vdash_c \ \forall y \leqslant t\ \exists z \ \ \psi \longleftrightarrow \exists z\ \forall y \leqslant t \ \ \psi'$.

Proof of Proposition 2.4:

(1) The idea is to have z encode a sequence of t many integers, and to consider all such sequences. For this we set

$$\psi' = \Big(\varphi_\beta(x_0, y, z', z'') \ \wedge \ \psi_{[x_0/z]}\Big)$$

where $\varphi_\beta(x_0, y, z', z'')$ represents the primitive recursive β-function that was introduced on page 123 in the proof of Lemma 2.4 as a consequence of the Chinese Remainder Theorem (page 122). Indeed, on Lemma 2.4 on page 123 we proved that there exists some function $\beta \in \mathbb{N}^{(\mathbb{N}^3)}$ which is both representable and $\mathcal{P}rim.\ \mathcal{R}ec.$ such that for all $k \in \mathbb{N}$ and every sequence $n_0, n_1, \ldots, n_k$ there exist $a, b \in \mathbb{N}$ such that

$$\begin{cases} \beta(0,a,b) & = & n_0 \\ \beta(1,a,b) & = & n_1 \\ & \vdots & \\ \beta(k,a,b) & = & n_k. \end{cases}$$

We recall that $\varphi_\beta(x_0, y, z', z'')$ stands for

$$x_0 < S(z' \cdot Sy) \ \wedge \ \exists \theta \leqslant z'' \quad \Big(\theta \cdot S(z' \cdot Sy)\Big) + x_0 = z''$$

so that we obtain

$$\mathcal{R}ob. \vdash_c \exists y \leqslant t \, \forall z \quad \psi \longleftrightarrow \forall z' \, \forall z'' \, \exists y \leqslant t \quad \psi'$$

from where we apply Proposition 2.2 to get the result.

(2) *Mutatis mutandis*, the same idea works fine.

□

We are now able to state a stronger version of Theorem 2.2:

Theorem 2.1

Every total recursive function is representable by some Σ^0_1-formula.

Proof of Theorem 2.1:

It is enough to go through the proofs of Examples 2.1, 2.2, 2.3 and Lemmas 2.1 , 2.3, 2.5, and notice that all formulas we defined were Σ^0_1-formulas. □

6.3 A First Glance at Gödel's 2nd Incompleteness Theorem

We first recall that by Theorem 3.1 the set below is

$$\left\{ (\ulcorner P \urcorner, \ulcorner \varphi \urcorner) \in \mathbb{N}^2 \mid P \textit{ is a proof of } T \vdash_c \varphi \right\}$$

- primitive recursive if T is primitive recursive,
- recursive if T is recursive.

We consider any recursive theory $T \supseteq \mathcal{R}ob.$ and consider some Σ^0_1-formula $\varphi_{proof_T}(x_1, x_2)$ which represents the set above. This means that for all $i_1, i_2 \in \mathbb{N}$ we have:

- if $(i_1, i_2) \in \left\{ (\ulcorner P \urcorner, \ulcorner \varphi \urcorner) \in \mathbb{N}^2 \mid P \textit{ is a proof of } T \vdash_c \varphi \right\}$, then $\mathcal{R}ob. \vdash_c \varphi_{proof_T}(i_1, i_2)$;
- if $(i_1, i_2) \notin \left\{ (\ulcorner P \urcorner, \ulcorner \varphi \urcorner) \in \mathbb{N}^2 \mid P \textit{ is a proof of } T \vdash_c \varphi \right\}$, then $\mathcal{R}ob. \vdash_c \neg\varphi_{proof_T}(i_1, i_2)$.

so in particular if T is consistent, we have

$$P \text{ is a proof of } T \vdash_c \varphi \iff \mathcal{R}ob. \vdash_c \varphi_{proof_T}(\lceil \ulcorner P \urcorner \rceil, \lceil \ulcorner \varphi \urcorner \rceil).$$

We consider the following primitive recursive function $diag : \mathbb{N} \longrightarrow \mathbb{N}$.

$$diag(n) = \begin{cases} \ulcorner \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]} \urcorner & \textit{if} \quad n = \ulcorner \varphi \urcorner \in \mathcal{F}_{\checkmark x_0 \,!free} \\ 0 & \textit{otherwise} \end{cases}$$

together with any Σ_1^0-formula $\varphi_{diag}(x_0, x_1)$ that represents $diag$. This means we have for all $n \in \mathbb{N}$

$$\mathcal{R}ob. \vdash_c \forall x_0 \left(diag(n) = x_0 \longleftrightarrow \varphi_{diag}(x_0, n) \right).$$

We define the Σ_1^0-formula $\Xi(x_0)$ by

$$\Xi(x_0) := \exists x_1 \, \exists x_2 \, \left(\varphi_{proof_T}(x_1, x_2) \wedge \varphi_{diag}(x_2, x_0)\right).$$

Proposition 3.1

For every integer n we have

$$\mathbb{N} \models \Xi(n) \iff \mathcal{R}ob. \vdash_c \Xi(n).$$

Proof of Proposition 3.1:

(1) if $n = \ulcorner \varphi \urcorner \in \mathcal{F}_{\checkmark x_0 \,!free}$ and there is a proof P of $T \vdash_c \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]}$ we have both

$$\mathcal{R}ob. \vdash_c \varphi_{diag}(\lceil \ulcorner \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]} \urcorner \rceil, n)$$

and

$$\mathcal{R}ob. \vdash_c \varphi_{proof_T}(\lceil \ulcorner P \urcorner \rceil, \lceil \ulcorner \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]} \urcorner \rceil)$$

therefore

$$\mathcal{R}ob. \vdash_c \exists x_1 \, \exists x_2 \, \left(\varphi_{proof_T}(x_1, x_2) \wedge \varphi_{diag}(x_2, n)\right)$$

which is

$$\mathcal{R}ob. \vdash_c \Xi(n).$$

(2) if $n = \ulcorner \varphi \urcorner \in \mathcal{F}_{\checkmark x_0 \,!free}$ and there is no proof P of $T \vdash_c \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]}$ we have for all proofs P

$$\mathcal{R}ob. \vdash_c \forall x_2 \left(\varphi_{diag}(x_2, n) \longleftrightarrow x_2 = \lceil \ulcorner \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]} \urcorner \rceil \right)$$

and

$$\mathcal{R}ob. \vdash_c \neg \varphi_{proof_T}(\lceil \ulcorner P \urcorner \rceil, \lceil \ulcorner \varphi_{[\lceil \ulcorner \varphi \urcorner \rceil / x_0]} \urcorner \rceil)$$

and furthermore for every integer i

$$\mathcal{R}ob.\ \vdash_c\ \neg\varphi_{proof_T}(i, \ulcorner \ulcorner \varphi_{[\ulcorner\ulcorner\varphi\urcorner\urcorner/x_0]} \urcorner \urcorner)$$

therefore, since $\mathbb{N} \models \varphi_{\mathcal{R}ob.}$, by the soundness theorem we have

$$\mathcal{R}ob. \nvdash_c \Xi(n).$$

(3) if $n \notin \mathcal{F}_{\checkmark x_0\ !free}$, then for every integer i_1,

$$\mathcal{R}ob.\ \vdash_c\ \neg\varphi_{proof_T}(i_1, diag(n))$$

for the reason that for all integer i_1

$$(i_1, 0) \notin \left\{ (\ulcorner P \urcorner, \ulcorner \varphi \urcorner) \in \mathbb{N}^2 \mid P \textit{ is a proof of } T\ \vdash_c\ \varphi \right\}$$

because 0 is never the code of a formula. Hence, by application of the soundness theorem we have

$$\mathcal{R}ob. \nvdash_c \Xi(n).$$

□

So to speak, $\mathbb{N} \models \Xi(n)$ asserts that there exists a proof that there is a 1 on position (i_n, i_n) in the array on page 174, where n is the integer that codes the formula φ_{i_n}.

We now consider the formula $\neg\Xi(x_0)$ – that we write $\neg\Xi$ – together with the term that represents its code $\ulcorner\neg\Xi\urcorner$ and the term that represents the code of the formula $\neg\Xi(x_0)$ which "eats up" its own code $\ulcorner\neg\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner$

Claim 3.1

$$\mathcal{R}ob.\ \vdash_c\ \Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]} \longleftrightarrow \exists x_1 \varphi_{proof_T}(x_1, \ulcorner\ulcorner\neg\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner)$$

which is precisely

$$\mathcal{R}ob.\ \vdash_c\ \exists x_1\ \exists x_2\ \left(\varphi_{proof_T}(x_1, x_2)\ \wedge\ \varphi_{diag}(x_2, \ulcorner\ulcorner\neg\Xi\urcorner\urcorner)\right) \longleftrightarrow \exists x_1 \varphi_{proof_T}(x_1, \ulcorner\ulcorner\neg\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner).$$

Proof of Claim 3.1:

($\Leftarrow$) By the very definition of the function *diag* and that φ_{diag} represents that function we have

$$\mathcal{R}ob.\ \vdash_c\ \varphi_{diag}(\ulcorner\ulcorner\neg\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner, \ulcorner\ulcorner\neg\Xi\urcorner\urcorner))$$

thus

$$\mathcal{R}ob. \vdash_c \exists x_1 \varphi_{proof_T}(x_1, \lceil \ulcorner \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]} \urcorner \rceil) \longrightarrow \exists x_1 \exists x_2 \ \left(\varphi_{proof_T}(x_1, x_2) \ \wedge \varphi_{diag}(x_2, \lceil \ulcorner \neg \Xi \urcorner \rceil)\right).$$

(⇒) Since φ_{diag} represents the function *diag* we have

$$\mathcal{R}ob. \vdash_c \forall x_2 \left(\varphi_{diag}(x_2, \lceil \ulcorner \neg \Xi \urcorner \rceil)\right) \longleftrightarrow x_2 = \lceil \ulcorner \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]} \urcorner \rceil \Big)$$

therefore

$$\mathcal{R}ob. \vdash_c \exists x_1 \exists x_2 \ \left(\varphi_{proof_T}(x_1, x_2) \ \wedge \varphi_{diag}(x_2, \lceil \ulcorner \neg \Xi \urcorner \rceil)\right) \rightarrow \exists x_1 \varphi_{proof_T}(x_1, \lceil \ulcorner \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]} \urcorner \rceil) \ .$$

□

Claim 3.2

$$T \nvdash_c \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]}.$$

Proof of Claim 3.2:

Towards a contradiction, we assume that

$$T \ \vdash_c \ \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]}.$$

It follows that there exists an integer $\ulcorner P \urcorner$ such that

$$(\ulcorner P \urcorner \ , \ \lceil \ulcorner \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]} \urcorner \rceil) \in \left\{ (\ulcorner Q \urcorner, \ulcorner \varphi \urcorner) \in \mathbb{N}^2 \mid Q \ is\ a\ proof\ of\ T \ \vdash_c \ \varphi \right\}.$$

Therefore, since φ_{proof_T} represents the set above, we have

$$\mathcal{R}ob. \ \vdash_c \ \varphi_{proof_T}(\lceil \ulcorner P \urcorner \rceil, \lceil \ulcorner \neg \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]} \urcorner \rceil)$$

and by Claim 3.1 we obtain

$$\mathcal{R}ob. \ \vdash_c \ \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]}.$$

Since $\mathcal{R}ob. \subseteq T$ we obtain

$$T \ \vdash_c \ \Xi_{[\lceil \ulcorner \neg \Xi \urcorner \rceil / x_0]}$$

which contradicts the fact that T is consistent for we obtain both

$$T \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \quad \text{and} \quad T \vdash_c \neg\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}.$$

□

Claim 3.3

$$\left.\begin{array}{c} \mathcal{R}\text{ob.} \\ \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \longrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \lceil\ulcorner\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil) \end{array}\right] \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \longrightarrow \neg cons(T).$$

Where $\neg cons(T)$ stands for the formula[1]

$$\exists \ulcorner\varphi\urcorner \left(\exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\varphi\urcorner) \ \wedge\ \exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\neg\varphi\urcorner)\right)$$

Proof of Claim 3.3:

From Claim 3.1 we obtain

$$\mathcal{R}ob. \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \exists x_1 \varphi_{proof_T}(x_1, \lceil\ulcorner\neg\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil).$$

Thus we have both

- $\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \lceil\ulcorner\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil) \ \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \lceil\ulcorner\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil)$
- $\mathcal{R}ob. \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \exists x_1 \varphi_{proof_T}(x_1, \lceil\ulcorner\neg\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil)$

which leads to

$$\left.\begin{array}{c} \mathcal{R}\text{ob.}, \\ \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \lceil\ulcorner\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil) \end{array}\right\} \vdash_c \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \rightarrow \begin{pmatrix} \exists x_1\ \varphi_{proof_{\mathcal{R}ob.}}(x_1, \lceil\ulcorner\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil) \\ \bigwedge \\ \exists x_1\ \varphi_{proof_T}(x_1, \lceil\ulcorner\neg\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}\urcorner\rceil) \end{pmatrix}$$

By the very definition[a] of φ_{proof_T} and $\varphi_{proof_{\mathcal{R}ob.}}$ we have

- $\mathcal{R}ob. \vdash_c \forall x_0\ \forall x_1\ \left(\varphi_{proof_{\mathcal{R}ob.}}(x_0, x_1) \longrightarrow \varphi_{proof_T}(x_0, x_1)\right).$

[1]We recall that we write $\exists \ulcorner\varphi\urcorner \ldots$ for $\exists x\ (\varphi_{\mathcal{F}}(x)\ \wedge\ \ldots$.

Therefore we obtain

$$\left.\begin{array}{c}\mathcal{R}\text{ob.},\\ \Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\rightarrow\exists x_1\varphi_{proof_{\mathcal{R}ob.}}\left(x_1,\ulcorner\ulcorner\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right)\end{array}\right\}\vdash_c\ \Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\rightarrow\left(\begin{array}{c}\exists x_1\ \varphi_{proof_T}\left(x_1,\ulcorner\ulcorner\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right)\\ \bigwedge\\ \exists x_1\ \varphi_{proof_T}\left(x_1,\ulcorner\ulcorner\neg\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right)\end{array}\right)$$

which yields the result.

□

[a] this means "if we choose wisely the Σ^0_1-formulas φ_{proof_T} and $\varphi_{proof_{\mathcal{R}ob.}}$ that represent the two recursive sets

$$\left\{\left(\ulcorner P\urcorner,\ulcorner\varphi\urcorner\right)\in\mathbb{N}^2\mid P\textit{ is a proof of }T\ \vdash_c\ \varphi\right\}\quad\text{and}\quad\left\{\left(\ulcorner P\urcorner,\ulcorner\varphi\urcorner\right)\in\mathbb{N}^2\mid P\textit{ is a proof of }\mathcal{R}ob.\ \vdash_c\ \varphi\right\}.$$

Lemma 3.1

Let $T\supseteq\mathcal{R}ob.$ be any consistent recursive theory.
If

$$T\ \vdash_c\ \Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\longrightarrow\exists x_1\varphi_{proof_{\mathcal{R}ob.}}\left(x_1,\ulcorner\ulcorner\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right),$$

then

$$T\nvdash_c cons(T).$$

Proof of Lemma 3.1:

Follows immediately from Claims 3.2 and 3.3.

□

So we are left with the problem of characterising the consistent theories that both extend Robinson arithmetic and prove this very strange formula:

$$\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\longrightarrow\exists x_1\varphi_{proof_{\mathcal{R}ob.}}\left(x_1,\ulcorner\ulcorner\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right).$$

Ultimately we will show that $\mathcal{R}ob.+I\Sigma^0_1$ is a good candidate. Indeed, we will show

$$\mathcal{R}ob.+I\Sigma^0_1\ \vdash_c\ \Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\longrightarrow\exists x_1\varphi_{proof_{\mathcal{R}ob.}}\left(x_1,\ulcorner\ulcorner\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}\urcorner\urcorner\right).$$

In order to show that this result holds — and because it can easily be seen that the formula $\Xi_{[\ulcorner\ulcorner\neg\Xi\urcorner\urcorner/x_0]}$ is Σ^0_1 — we will rather show that the following result holds for every Σ^0_1-formula :

$$\mathcal{R}ob.+I\Sigma^0_1\ \vdash_c\ \varphi\longrightarrow\exists x_1\varphi_{proof_{\mathcal{R}ob.}}\left(x_1,\ulcorner\varphi\urcorner\right).$$

6.4 The Core of the Proof

We are going to prove that given any closed Σ_1^0-formula φ, the following holds:

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi \longrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner).$$

There are two different misapprehensions that one must avoid:

(1) The assumption that φ is some closed Σ_1^0-formula is crucial. Indeed, the result does not hold for any closed formula, not even closed Π_1^0-formula. Indeed, we saw in Example 1.4 that $\mathcal{R}ob.+I\Sigma_1^0$ proves the commutativity of the multiplication. We also saw in Example 1.2 that there exists some model of $\mathcal{R}ob.$ that does not satisfy the commutativity of the multiplication, hence, we have both

- $\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x\, \forall y\; x{\cdot}y = y{\cdot}x$
- $\mathcal{R}ob. \nvdash_c \forall x\, \forall y\; x{\cdot}y = y{\cdot}x$.

So, if Lemma 4.6 were to hold for Π_1^0-formulas, then we would have

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\forall x\, \forall y\; x{\cdot}y = y{\cdot}x\urcorner).$$

Hence, by the completeness theorem, we would also have

$$\mathcal{R}ob.+I\Sigma_1^0 \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\forall x\, \forall y\; x{\cdot}y = y{\cdot}x\urcorner);$$

and in particular

$$\mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\forall x\, \forall y\; x{\cdot}y = y{\cdot}x\urcorner).$$

But from an integer that codes a proof of the formula $\forall x\, \forall y\; x{\cdot}y = y{\cdot}x$ from $\mathcal{R}ob.$, we would get a proof of the sequent $\mathcal{R}ob. \vdash \forall x\, \forall y\; x{\cdot}y = y{\cdot}x$, contradicting the fact that $\mathcal{R}ob. \nvdash_c \forall x\, \forall y\; x{\cdot}y = y{\cdot}x$.

(2) At first glance, one may think that

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi \longrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner) \tag{6.1}$$

is equivalent to

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi \implies \mathcal{R}ob.+I\Sigma_1^0 \vdash_c \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner). \tag{6.2}$$

To see this, notice that when $\mathcal{R}ob.+I\Sigma_1^0 \nvdash_c \varphi$ holds, then 6.2 trivially holds but 6.1 is far more involved. Indeed,

- if $\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \neg\varphi$, then proving 6.1 is easy since it is equivalent to proving

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \neg\varphi \vee \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner).$$

- But if $\mathcal{R}ob.{+}I\Sigma_1^0 \nvdash_c \neg\varphi$, then one must show 6.1 on the basis that $\mathcal{R}ob.{+}I\Sigma_1^0$ proves neither φ nor $\neg\varphi$.

On our way to proving Lemma 4.6, we will prove 6.1 which will require the following result:

Proposition 4.1

Let φ be any closed Σ_1^0-formula.

$$\mathbb{N} \models \Big(\varphi \longleftrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\Big).$$

Proof of Proposition 4.1:

we distinguish between the two directions of "$\longleftrightarrow$".

(←) We assume $\mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)$. So, there exists some standard integer that codes a proof of φ from Robinson arithmetic. From this standard integer, we recover a proof of $\mathcal{R}ob. \vdash \varphi$, which shows that φ holds in all models of $\mathcal{R}ob.$. So, φ holds in particular inside the standard model ($\mathbb{N}$).

(→) We assume $\mathbb{N} \models \varphi$. We then show, by induction on the height of φ,

$$\mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\,.$$

We can easily[a] show the following:

(1) for every closed terms t_1 and t_2,

$$\mathbb{N} \models t_1 = t_2 \quad \Rightarrow \quad \mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner t_1 = t_2\urcorner)\,.$$

(This is done by induction on the height of t_1 and t_2.)

(2) For every closed Δ_0^0-formula φ,

$$\mathbb{N} \models \varphi \quad \Rightarrow \quad \mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\,.$$

This is done by induction on the height of φ with the following statement taking care of bounded quantifications:

4.10 $$\mathcal{R}ob. \vdash_c \forall x \ \big[x \leqslant n \longleftrightarrow (x = 0 \ \vee \ x = S0 \ \vee \ \ldots \ \vee \ x = n)\big]\,.$$

(3) for every closed Σ^0_1-formula of the form $\exists x_1 \ \ldots \exists x_n \ \varphi$ where φ is some Δ^0_0-formula,

$$\mathbb{N} \models \exists x_1 \ \ldots \exists x_n \ \varphi \quad \Rightarrow \quad \mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner \exists x_1 \ \ldots \exists x_n \ \varphi \urcorner)\,.$$

This holds because

$$\begin{aligned} \mathbb{N} \models \exists x_1 \ \ldots \exists x_n \ \varphi & \Longrightarrow \ \textit{for some } k_1, \ldots, k_n \in \mathbb{N},\ \varphi_{[k_1/x_1, \ldots, k_n/x_n]} \textit{ holds in } \mathbb{N} \\ & \Longrightarrow \ \mathbb{N} \models \varphi_{[\mathbf{k_1}/x_1, \ldots, \mathbf{k_n}/x_n]} \\ \textit{(by ind. hyp.)} & \Longrightarrow \ \mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner \varphi_{[\mathbf{k_1}/x_1, \ldots, \mathbf{k_n}/x_n]} \urcorner) \\ & \Longrightarrow \ \mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner \exists x_1 \ \ldots \exists x_n \ \varphi \urcorner)\,. \end{aligned}$$

□

[a] the whole proof involves many cases. It is tedious but straightforward.

This proposition yields an easy but amazing corollary. We mention a result about the Goldbach conjecture, but the same remark holds for all conjectures in arithmetic that can be expressed by some Π^0_1-formula.

Corollary 4.1

If the Goldbach conjecture[a] is neither provable nor disprovable, then it holds true in the standard model $\mathbb{N}$.

[a] Goldbach conjecture is: "*every even integer strictly greater than 2 is the sum of two prime numbers*".

Proof of Corollary 4.1:

The Goldbach conjecture is some Π^0_1 statement. If the negation of the Goldbach conjecture were true in $\mathbb{N}$, then by Proposition 4.1 it would be provable. □

We introduce a few definitions in order to characterize the models of $\mathcal{R}ob.$ as final extensions of the standard model of arithmetic.

Definition 4.1

Let $\mathcal{M}$ and $\mathcal{N}$ be two models of $\mathcal{R}ob.$, such that $\mathcal{M}$ is a substructure [a] of $\mathcal{N}$.

$$\mathcal{N} \textit{ is a final extension of } \mathcal{M}$$

$$\Longleftrightarrow$$

for every $a \in |\mathcal{M}|$ and $b \in |\mathcal{N}|$ we have both

- if $\mathcal{N} \models b \leqslant a$, then $b \in |\mathcal{M}|$
- if $b \notin |\mathcal{M}|$, then $\mathcal{N} \models a \leqslant b$.

[a] $\mathcal{M}$ is a substructure of $\mathcal{N}$ iff

- $|\mathcal{M}| \subseteq |\mathcal{N}|$
- for every constant symbol c: $c^{\mathcal{M}} = c^{\mathcal{N}}$
- for every function symbol f whose arity is n: $f^{\mathcal{M}} = f^{\mathcal{N}} \restriction |\mathcal{M}|^n$
- for every relation symbol R whose arity is n: $R^{\mathcal{M}} = R^{\mathcal{N}} \cap |\mathcal{M}|^n$.

Lemma 4.1

If $\mathcal{N}$ is a model of $\mathcal{R}ob.$, then the substructure $\mathcal{M}$ whose domain is

$$|\mathcal{M}| = \{n^{\mathcal{N}} \mid n \in \mathbb{N}\}$$

is isomorphic to the standard model $\mathbb{N}$.

Proof of Lemma 4.1:

Left as a very easy exercise.

□

Lemma 4.2

Up to isomorphism, every model $\mathcal{N}$ of $\mathcal{R}ob.$ is a final extension of the standard model $\mathbb{N}$.

Proof of Lemma 4.2:

The mapping $f : \mathbb{N} \longrightarrow |\mathcal{N}|$ defined by $f(n) = n^{\mathcal{N}}$ is an injective homomorphism that satisfies for every $n \in \mathbb{N}$ and $b \in |\mathcal{N}|$:

(1)

$$\mathcal{N} \models b \leqslant n \implies \mathcal{N} \models b = 0 \lor b = S0 \lor \ldots \lor b = n \implies f^{-1}(b) \in \mathbb{N}.$$

This is by

$$\text{4.10} \quad \mathcal{R}ob. \vdash_c \forall x \left[x \leqslant n \longleftrightarrow (x = 0 \lor x = S0 \lor \ldots \lor x = n)\right].$$

(2)

$$b \notin f[\mathbb{N}] \implies \mathcal{N} \models n \leqslant b.$$

This is by

$$\text{4.11} \quad \mathcal{R}ob. \vdash_c \forall x \ (x \leqslant n \vee n \leqslant x)$$

and

$$\text{4.10} \quad \mathcal{R}ob. \vdash_c \forall x \ [x \leqslant n \longleftrightarrow (x = 0 \vee x = S0 \vee \ \ldots \ \vee x = n)] .$$

This shows that $\mathcal{N}$ is a final extension of its substructure induced by $f[\mathbb{N}]$ which is isomorphic to the standard model.

□

Now that we know that models of Robinson arithmetic are final extensions of the standard model, we can prove the result that follows.

Lemma 4.3

Let φ be any closed Σ_1^0-formula.

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi \implies \mathcal{R}ob.+I\Sigma_1^0 \vdash_c \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner) .$$

Proof of Lemma 4.3:

We make use of the completeness theorem and of the fact that the standard model ($\mathbb{N}$) is a model of $\mathcal{R}ob.+I\Sigma_1^0$.

($\Rightarrow$) Since

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi,$$

the formula φ holds in all models that satisfy $\mathcal{R}ob.+I\Sigma_1^0$. So in particular we have

$$\mathbb{N} \models \varphi.$$

By Proposition 4.1, this leads to

$$\mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner) .$$

Therefore, there exists some (standard integer) n that codes a proof of φ in Robinson arithmetic. i.e.,

$$\mathbb{N} \models \varphi_{proof_{\mathcal{R}ob.}}(n, \ulcorner\varphi\urcorner) .$$

Then, we consider any model $\mathcal{M}$ such that satisfies

$$\mathcal{M} \models \mathcal{R}ob.{+}I\Sigma_1^0.$$

From Lemma 4.2 $\mathcal{M}$ is a final extension of (a structure isomorphic to) $\mathbb{N}$. Now, $\varphi_{proof_{\mathcal{R}ob.}}(n, \ulcorner\varphi\urcorner)$ is some closed Δ_0^0-formula, all the various bounded quantifications are bounded by (terms that depict) standard integers. From this fact, it is as usual tedious but straightforward to show by induction on the height of $\varphi_{proof_{\mathcal{R}ob.}}(n, \ulcorner\varphi\urcorner)$ that

$$\mathcal{M} \models \varphi_{proof_{\mathcal{R}ob.}}(n, \ulcorner\varphi\urcorner)\,.$$

From where we obtain

$$\mathcal{M} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\,.$$

($\Leftarrow$) Since

$$\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)$$

holds, we also have

$$\mathbb{N} \models \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\,.$$

Therefore, there exists some (standard integer) n that codes a proof of φ from $\mathcal{R}ob.$.

$$\mathbb{N} \models \varphi_{proof_{\mathcal{R}ob.}}(n, \ulcorner\varphi\urcorner)\,.$$

Such a proof is also some proof of φ from the theory $\mathcal{R}ob.{+}I\Sigma_1^0$, which witnesses that we have

$$\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \varphi.$$

□

Before we come to the proof of Lemma 4.6 – which will immediately yield Gödel's 2nd incompleteness theorem – we need to take care of some humongous preliminary work.

Lemma 4.4

Let $t_{[x_1,\dots,x_n]}$ be any $\mathcal{L}_{\mathcal{A}}$-term (where $\mathcal{L}_{\mathcal{A}} = \{0, S, +, \cdot\}$).

$$\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \dots \forall x_n \forall x_{n+1}\Big(t_{[x_1,\dots,x_n]} = x_{n+1} \rightarrow \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner t_{[x_1,\dots,x_n]} = x_{n+1}\urcorner)\Big)^{a}.$$

[a] Where $\ulcorner t_{[x_1,\dots,x_n]} = x_{n+1}\urcorner$ stands for the formula $\ulcorner t_{[x_{k_1},\dots,x_{k_n}]} = x_{k_{n+1}}\urcorner[\ulcorner x_1\urcorner/x_{k_1},\dots,\ulcorner x_n\urcorner/x_{k_n},\ulcorner x_{n+1}\urcorner/x_{k_{n+1}}]$ meaning that $\ulcorner t_{[x_{k_1},\dots,x_{k_n}]} = x_{k_{n+1}}\urcorner$ is a formula whose $n+1$ (necessarily free) variables are $x_{k_1},\dots,x_{k_n},x_{k_{n+1}}$ and $\ulcorner t_{[x_1,\dots,x_n]} = x_{n+1}\urcorner$ is this term after the subsequent substitutions have taken place: $\mathcal{S}^{\mathcal{F}}_{ub.}\Big(\dots\mathcal{S}^{\mathcal{F}}_{ub.}\Big(\mathcal{S}^{\mathcal{F}}_{ub.}\Big(\ulcorner t_{[x_{k_1},\dots,x_{k_n}]} = x_{k_{n+1}}\urcorner, \ulcorner x_1\urcorner, k_1\Big), \ulcorner x_2\urcorner, k_2\Big),\dots, \ulcorner x_{n+1}\urcorner, k_{n+1}\Big)$.

Proof of Lemma 4.4:

We prove the result by induction on the height of the term $t_{[x_1,\ldots,x_n]}$.

$\mathbf{ht(t) = 0}$ we have three cases:

$\mathbf{t = 0}$ we need to show

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_{n+1} \left(0 = x_{n+1} \longrightarrow \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner 0 = x_{n+1} \urcorner)\right)$$

which is

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_{n+1} \left(0 \neq x_{n+1} \vee \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner 0 = x_{n+1} \urcorner)\right)$$

which comes down to proving

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner 0 = 0 \urcorner)\Big).$$

The code of the following proof is just what is needed:

$$\dfrac{\dfrac{}{0 = 0 \vdash 0 = 0}\,ax}{\vdash 0 = 0}\,Ref$$

$\mathbf{t = x_{n+1}}$ we need to show

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_{n+1} \left(x_{n+1} = x_{n+1} \longrightarrow \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner x_{n+1} = x_{n+1} \urcorner)\right)$$

in which case the code of the following proof is what is needed

$$\dfrac{\dfrac{}{x_{n+1} = x_{n+1} \vdash x_{n+1} = x_{n+1}}\,ax}{\vdash x_{n+1} = x_{n+1}}\,Ref$$

$\mathbf{t = x_i}\ \ (i \neq n+1)$ we need to show

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_i\, \forall x_{n+1} \left(x_i = x_{n+1} \longrightarrow \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner x_i = x_{n+1} \urcorner)\right)$$

which is also

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_i\, \forall x_{n+1} \left(x_i \neq x_{n+1} \vee \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner x_i = x_i \urcorner)\right)$$

which comes down to proving

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_i \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner x_i = x_i \urcorner)\,.$$

in which case the code of the following proof is what is needed

$$\dfrac{\dfrac{}{x_i = x_i \vdash x_i = x_i}\,{\scriptstyle ax}}{\vdash x_i = x_i}\,{\scriptstyle Ref}$$

$\underline{\mathbf{ht(t) = k+1}}$ we have three cases:

$\underline{\mathbf{t = Su}}$ We need to show

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ \ldots \forall x_n \ \forall x_{n+1} \ \Big(Su_{[x_1,\ldots,x_n]} = x_{n+1} \ \rightarrow \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner Su_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner)\Big).$$

We proceed by induction on x_{n+1}, which means we need to show

(1) $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ \ldots \forall x_n \ \Big(Su_{[x_1,\ldots,x_n]} = 0 \ \longrightarrow \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner Su_{[x_1,\ldots,x_n]} = 0 \urcorner)\Big).$

The result follows immediately by $\boxed{1 \quad \forall x \ Sx \neq 0}$.

(2) And assuming that

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ \ldots \forall x_n \ \Big(Su_{[x_1,\ldots,x_n]} = x_{n+1} \ \longrightarrow \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner Su_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner)\Big)$$

holds, we need to show

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ \ldots \forall x_n \ \Big(Su_{[x_1,\ldots,x_n]} = Sx_{n+1} \ \longrightarrow \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner Su_{[x_1,\ldots,x_n]} = Sx_{n+1} \urcorner)\Big).$$

By $\boxed{3 \quad \forall x \ \forall y \ (Sx = Sy \rightarrow x = y)}$ together with the induction hypothesis we obtain

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ \ldots \forall x_n \ \Big(Su_{[x_1,\ldots,x_n]} = Sx_{n+1} \ \longrightarrow \ \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner u_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner)\Big).$$

We notice then that for any term a, b we have the following proof:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Sa = Sx\ {}_{[b/x]} \vdash Sa = Sx\ {}_{[b/x]}}\,{\scriptstyle ax.}}{a = b, Sa = Sx\ {}_{[b/x]} \vdash Sa = Sx\ {}_{[b/x]}}\,{\scriptstyle wkn_l}}{a = b, Sa = Sx\ {}_{[b/x]}, Sa = Sx\ {}_{[a/x]} \vdash Sa = Sx\ {}_{[b/x]}}\,{\scriptstyle wkn_l}}{a = b, Sa = Sa \vdash Sa = Sb}\,{\scriptstyle Rep}}{a = b \vdash Sa = Sb}\,{\scriptstyle Ref}$$

Or simply

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Sa = Sb \vdash Sa = Sb}\ ax.}{a = b, Sa = Sb \vdash Sa = Sb}\ wkn_l}{a = b, Sa = Sb, Sa = Sa \vdash Sa = Sb}\ wkn_l}{a = b, Sa = Sa \vdash Sa = Sb}\ Rep}{a = b \vdash Sa = Sb}\ Ref$$

Then, we consider one application of the cut rule to get a proof of $\mathcal{R}ob. \vdash Sa = Sb$ assuming a proof of $\mathcal{R}ob. \vdash a = b$

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash a = b} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Sa = Sb \vdash Sa = Sb}\ ax.}{a = b, Sa = Sb \vdash Sa = Sb}\ wkn_l}{a = b, Sa = Sb, Sa = Sa \vdash Sa = Sb}\ wkn_l}{a = b, Sa = Sa \vdash Sa = Sb}\ Rep}{a = b \vdash Sa = Sb}\ Ref}{\mathcal{R}ob. \vdash Sa = Sb}\ cut$$

So, replacing a by $u_{[x_1,\ldots,x_n]}$, and b by x_{n+1} we obtain

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash u_{[x_1,\ldots,x_n]} = x_{n+1}} \qquad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{}{Su_{[x_1,\ldots,x_n]} = Sx_{n+1} \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ ax.}{u_{[x_1,\ldots,x_n]} = x_{n+1}, Su_{[x_1,\ldots,x_n]} = Sx_{n+1} \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ wkn_l}{u_{[x_1,\ldots,x_n]} = x_{n+1}, Su_{[x_1,\ldots,x_n]} = Sx_{n+1}, Su_{[x_1,\ldots,x_n]} = Su_{[x_1,\ldots,x_n]} \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ wkn_l}{u_{[x_1,\ldots,x_n]} = x_{n+1}, Su_{[x_1,\ldots,x_n]} = Su_{[x_1,\ldots,x_n]} \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ Rep}{u_{[x_1,\ldots,x_n]} = x_{n+1} \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ Ref}{\mathcal{R}ob. \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}}\ cut$$

So, from the code of a “proof” of $\mathcal{R}ob. \vdash u_{[x_1,\ldots,x_n]} = x_{n+1}$ we easily obtain a “proof” of $\mathcal{R}ob. \vdash Su_{[x_1,\ldots,x_n]} = Sx_{n+1}$.

We then make use of the fact that $u_{[x_1,\ldots,x_n]} = x_{n+1}$ is Δ^0_0 to obtain

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_{n+1} \left(Su_{[x_1,\ldots,x_n]} = x_{n+1} \longrightarrow \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner Su_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner) \right).$$

<u>$\mathbf{t = u{+}v}$</u> We need to show

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_{n+1} \left((u{+}v)_{[x_1,\ldots,x_n]} = x_{n+1} \rightarrow \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u{+}v)_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner) \right).$$

The proof goes by induction on v.

<u>$\mathbf{v = 0}$</u> We need to show

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_{n+1} \left((u{+}0)_{[x_1,\ldots,x_n]} = x_{n+1} \rightarrow \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u{+}0)_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner) \right)$$

Since by $\boxed{④ \quad \forall x\ x+0 = x}$ we have

$$\mathcal{R}ob.\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \ (u+0)_{[x_1,\ldots,x_n]} = u_{[x_1,\ldots,x_n]}$$

Since our proof is by induction on the complexity of the term t and u is less complicated than $t = u+v$, we have

$$\mathcal{R}ob.+I\Sigma_1^0\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \Big(u_{[x_1,\ldots,x_n]} = x_{n+1}\ \rightarrow\ \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner u_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner)\Big)$$

The "code" of the following "proof" yields the result.

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash u = x_{n+1}} \qquad \dfrac{\dfrac{\dfrac{}{u+0 = u \vdash u+0 = u}\,{\scriptstyle ax.}}{\forall x_0\ \ x_0+0 = x_0 \vdash u+0 = u}\,{\scriptstyle \forall_l} \qquad \dfrac{\dfrac{\dfrac{\dfrac{}{u+0 = x_{n+1} \vdash u+0 = x_{n+1}}\,{\scriptstyle ax.}}{u = x_{n+1}, u+0 = x_{n+1} \vdash u+0 = x_{n+1}}\,{\scriptstyle wkn_l}}{u = u+0, u = x_{n+1}, u+0 = x_{n+1} \vdash u+0 = x_{n+1}}\,{\scriptstyle wkn_l}}{u+0 = u, u = x_{n+1} \vdash u+0 = x_{n+1}}\,{\scriptstyle Rep}}{\forall x_0\ \ x_0+0 = x_0, u = x_{n+1} \vdash u+0 = x_{n+1}}\,{\scriptstyle cut}}{\mathcal{R}ob. \vdash u+0 = x_{n+1}}\,{\scriptstyle cut}$$

<u>$\mathbf{v = x_i}$</u> We need to show

$$\mathcal{R}ob.+I\Sigma_1^0\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \Big((u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1}\ \rightarrow\ \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner)\Big)$$

For this purpose we use the fact $(u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1}$ is Δ_0^0 and proceed by induction on x_{n+1}:

(1) The initial case is $x_i = 0$, which we already considered.

(2) Assuming

$$\mathcal{R}ob.+I\Sigma_1^0\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \Big((u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1}\ \rightarrow\ \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner)\Big)$$

$$\mathcal{R}ob.+I\Sigma_1^0\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = x_{n+1}\ \rightarrow\ \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+Sx_i)_{[x_1,\ldots,x_n]} = x_{n+1}\urcorner)\Big)$$

Once again, we proceed by induction on x_{n+1}:

(a) The initial case is $x_{n+1} = 0$. We need to show

$$\mathcal{R}ob.+I\Sigma_1^0\ \vdash_c\ \forall x_1\ \ldots\forall x_{n+1}\ \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = 0\ \rightarrow\ \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+Sx_i)_{[x_1,\ldots,x_n]} = 0\urcorner)\Big)$$

which follows immediately by $\boxed{⑤ \quad \forall x\ \forall y\ \big(x+Sy = S(x+y)\big)}$ and $\boxed{① \quad \forall x\ Sx \neq 0}$.

(b) We assume

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_{n+1} \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = x_{n+1} \rightarrow \exists x_0 \, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+Sx_i)_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner) \Big)$$

and we need to show

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_{n+1} \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = Sx_{n+1} \rightarrow \exists x_0 \, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+Sx_i)_{[x_1,\ldots,x_n]} = Sx_{n+1} \urcorner) \Big)$$

By

$$\boxed{⑤ \quad \forall x \; \forall y \; \big(x+Sy = S(x+y)\big)}$$

and

$$\boxed{③ \quad \forall x \; \forall y \; (Sx = Sy \rightarrow x = y)}$$

we have

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_{n+1} \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = Sx_{n+1} \rightarrow (u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1} \Big)$$

By the previous induction hypothesis, we have

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_{n+1} \Big((u+Sx_i)_{[x_1,\ldots,x_n]} = Sx_{n+1} \rightarrow \exists x_0 \, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u+x_i)_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner) \Big)$$

For this consider the "code" of the following "proof" where a, b, c are replaced respectively by $u_{[x_1,\ldots,x_n]}$, $x_{i[x_1,\ldots,x_n]}$, x_{n+1}.

$$
\begin{prooftree}
\AxiomC{}
\RightLabel{\scriptsize ax.}
\UnaryInfC{$a+Sb = S(a+b) \vdash a+Sb = S(a+b)$}
\RightLabel{\scriptsize $\forall_l$}
\UnaryInfC{$\forall x_1 \; a+Sx_1 = S(a+x_1) \vdash a+Sb = S(a+b)$}
\RightLabel{\scriptsize $\forall_l$}
\UnaryInfC{$\forall x_0 \; \forall x_1 \; x_0+Sx_1 = S(x_0+x_1) \vdash a+Sb = S(a+b)$}
\AxiomC{$\vdots$}
\UnaryInfC{$\mathcal{R}ob. \vdash a+b = c$}
\AxiomC{}
\RightLabel{\scriptsize ax.}
\UnaryInfC{$S(a+b) = Sc \vdash S(a+b) = Sc$}
\RightLabel{\scriptsize wkn_l}
\UnaryInfC{$(a+b) = c, S(a+b) = Sc \vdash S(a+b) = Sc$}
\RightLabel{\scriptsize wkn_l}
\UnaryInfC{$(a+b) = c, S(a+b) = Sc, S(a+b) = S(a+b) \vdash S(a+b) = Sc$}
\RightLabel{\scriptsize Rep}
\UnaryInfC{$(a+b) = c, S(a+b) = S(a+b) \vdash S(a+b) = Sc$}
\RightLabel{\scriptsize Ref}
\UnaryInfC{$(a+b) = c \vdash S(a+b) = Sc$}
\RightLabel{\scriptsize cut}
\BinaryInfC{$\mathcal{R}ob. \vdash S(a+b) = Sc$}
\AxiomC{}
\RightLabel{\scriptsize ax.}
\UnaryInfC{$a+Sb = Sc \vdash a+Sb = Sc$}
\RightLabel{\scriptsize wkn_l}
\UnaryInfC{$a+Sb = Sc, S(a+b) = Sc \vdash a+Sb = Sc$}
\RightLabel{\scriptsize wkn_l}
\UnaryInfC{$S(a+b) = a+Sb, a+Sb = Sc, S(a+b) = Sc \vdash a+Sb = Sc$}
\RightLabel{\scriptsize Ref}
\UnaryInfC{$a+Sb = S(a+b), S(a+b) = Sc \vdash a+Sb = Sc$}
\RightLabel{\scriptsize cut}
\BinaryInfC{$\mathcal{R}ob., a+Sb = S(a+b) \vdash a+Sb = Sc$}
\RightLabel{\scriptsize cut}
\BinaryInfC{$\mathcal{R}ob. \vdash a+Sb = Sc$}
\end{prooftree}
$$

This terminates the proof by induction on x_{n+1}. Hence, we obtain precisely the formula that we needed to complete the proof by induction on x_i. Therefore, the whole result is proved.

$\underline{\mathbf{v} = \mathbf{v_0} + \mathbf{v_1}}$ is left as a tedious exercise.

$\underline{\mathbf{v} = \mathbf{v_0} \cdot \mathbf{v_1}}$ is left as a tedious exercise as well.

$\underline{\mathbf{t = u{\cdot}v}}$ We need to show

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_{n+1} \Big((u{\cdot}v)_{[x_1,\ldots,x_n]} = x_{n+1} \to \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (u{\cdot}v)_{[x_1,\ldots,x_n]} = x_{n+1} \urcorner) \Big).$$

The proof is similar to the case of the addition, and we leave it as a long and tedious exercise.

□

We took care of atomic formulas. The next result will — almost — take care of Δ^0_0-formula (notice that the negation is missing in what follows).

Lemma 4.5

The set of all formulas φ that satisfy

$$\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \Big(\varphi_{[x_1,\ldots,x_n]} \longrightarrow \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \varphi_{[x_1,\ldots,x_n]} \urcorner) \Big).$$

is closed under

- conjunction
- disjunction
- bounded universal quantification
- existential quantification.

Proof of Lemma 4.5:

Conjunction: if $\varphi := (\psi \wedge \theta)$, and

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \Big(\psi_{[x_1,\ldots,x_n]} \to \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \psi_{[x_1,\ldots,x_n]} \urcorner) \Big)$
- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \Big(\theta_{[x_1,\ldots,x_n]} \to \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \theta_{[x_1,\ldots,x_n]} \urcorner) \Big).$

We consider the "code" of the following "proof":

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash \psi} \qquad \dfrac{\vdots}{\mathcal{R}ob. \vdash \theta}}{\mathcal{R}ob. \vdash \psi \wedge \theta}\ {\scriptstyle \wedge_r \,+\, ctr\ l}$$

which yields

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \Big((\psi \wedge \theta)_{[x_1,\ldots,x_n]} \to \exists x_0\, \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (\psi \wedge \theta)_{[x_1,\ldots,x_n]} \urcorner) \Big).$

Disjunction: if $\varphi := (\psi \vee \theta)$, and

- $\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_n \ \Big(\psi_{[x_1,\ldots,x_n]} \rightarrow \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \psi_{[x_1,\ldots,x_n]} \urcorner)\Big)$

We consider the "code" of the following "proof":

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash \psi}}{\mathcal{R}ob. \vdash \psi \vee \theta}\ \vee_r$$

which yields

- $\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_n \ \Big((\psi \vee \theta)_{[x_1,\ldots,x_n]} \rightarrow \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner (\psi \vee \theta)_{[x_1,\ldots,x_n]} \urcorner)\Big)$.

Existential quantification: we consider $\exists y \ \varphi$, assuming that the following holds:

- $\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_n \ \forall y \ \Big(\varphi_{[x_1,\ldots,x_n,y]} \rightarrow \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \varphi_{[x_1,\ldots,x_n,y]} \urcorner)\Big)$.

by considering the "code" of the following "proof"

$$\dfrac{\dfrac{\vdots}{\mathcal{R}ob. \vdash \varphi_{[x_1,\ldots,x_n,y]}}}{\mathcal{R}ob. \vdash \exists y \ \varphi_{[x_1,\ldots,x_n]}}\ \exists_r$$

we obtain

- $\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_n \ \forall y \ \Big(\varphi_{[x_1,\ldots,x_n,y]} \rightarrow \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \exists y \ \varphi_{[x_1,\ldots,x_n]} \urcorner)\Big)$

Now given any formula ψ in which y does not occur freely, we have

- $\vdash_c \forall x_1 \ldots \forall x_n \ \forall y \ \Big(\varphi_{[x_1,\ldots,x_n,y]} \rightarrow \psi\Big) \longleftrightarrow \forall x_1 \ldots \forall x_n \ \Big(\exists y \varphi_{[x_1,\ldots,x_n]} \rightarrow \psi\Big)$.

Since y does not occur freely in the formula $\exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \exists y \ \varphi_{[x_1,\ldots,x_n]} \urcorner)$, we obtain

- $\mathcal{R}ob.{+}I\Sigma_1^0 \vdash_c \forall x_1 \ldots \forall x_n \ \Big(\exists y \ \varphi_{[x_1,\ldots,x_n]} \rightarrow \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner \exists y \ \varphi_{[x_1,\ldots,x_n]} \urcorner)\Big)$.

Bounded universal quantification: without loss of generality we consider formulas of the form $\forall y < x_m \ \varphi$ since for every term t the formula $\forall y < t \ \varphi$ we have

$$\vdash_c \ \forall y < t \ \varphi \ \longleftrightarrow \ \forall y < x_m \ (x_m = t \wedge \psi).$$

We assume that the following holds:

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall y \ \Big(\varphi_{[x_1,\ldots,x_n,y]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\varphi_{[x_1,\ldots,x_n,y]}\urcorner)\Big).$

Strictly speaking, the formula $\forall y < x_m \ \varphi$ stands for $\forall y \ \big(y < x_m \longrightarrow \varphi\big)$ which is nothing but $\forall y \ \big(\exists z \ (z \neq 0 \wedge z{+}y = x_m) \longrightarrow \varphi\big)$ where y does not occur (freely) in φ and we have

$$\vdash_c \forall y \ \big(\exists z \ (z \neq 0 \wedge z{+}y = x_m) \to \varphi\big) \longleftrightarrow \forall y \ \forall z \ \big((z \neq 0 \wedge z{+}y = x_m) \to \varphi\big)$$

We need to prove

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall x_m \ \Big(\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]}\urcorner)\Big)$

We prove the result by induction on x_m, which means that we make use of an instance of the induction schema. For this, we need to show that

(1) the case holds for $x_m = 0$,

(2) and that it also holds for Sx_m, assuming the case holds for x_m.

(1) the initial case is

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \ \Big(\forall y < 0 \ \varphi_{[x_1,\ldots,x_n]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\forall y < 0 \ \varphi_{[x_1,\ldots,x_n]}\urcorner)\Big)$

which is easy since the following holds.

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall y \ \neg \ y < 0$

(2) we assume

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall x_m \ \Big(\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]}\urcorner)\Big)$

and we show

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall x_m \ \Big(\forall y < Sx_m \ \varphi_{[x_1,\ldots,x_n]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\forall y < Sx_m \ \varphi_{[x_1,\ldots,x_n]}\urcorner)\Big)$

Firstly, notice that we have

- $\mathcal{R}ob. \vdash_c \forall y \ \forall x_m \ \Big(y < Sx_m \longleftrightarrow \big(y < x_m \vee y = x_m\big)\Big)$

Secondly, we have both

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall x_m \ \Big(\varphi_{[x_1,\ldots,x_n,x_m]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\varphi_{[x_1,\ldots,x_n,x_m]}\urcorner)\Big)$
- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \forall x_m \ \Big(\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]} \to \exists x_0 \ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner\forall y < x_m \ \varphi_{[x_1,\ldots,x_n]}\urcorner)\Big)$

By mixing accordingly the two “proofs” we get what we need.

$\square$

We finally come to the proof of the main lemma.

Lemma 4.6

Let φ be any closed Σ_1^0-formula.

$$\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \varphi \longrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}(x_1, \ulcorner\varphi\urcorner)\,.$$

Proof of Lemma 4.6:

From previous lemmas we already know that this result holds for the class $\mathbf{\Sigma}$ that contains all atomic formulas (Lemma 4.4) and is closed under (Lemma 4.5)

- conjunction
- disjunction
- bounded universal quantification
- existential quantification.

We recall that Δ_0^0 was described as the least class that

(1) contains all atomic formulas: $t_0 = t_1$

(2) is closed under conjunctions, disjunctions and negations

(3) is closed under bounded quantifications.

Notice that Δ_0^0 can equivalently be defined as the least class that

(1) contains all atomic formulas: $t_0 = t_1$

(2) contains all negations of atomic formulas: $t_0 \neq t_1$

(3) is closed under conjunctions and disjunctions

(4) is closed under bounded quantifications.

So, notice that in order to show that $\mathbf{\Sigma} = \Sigma_1^0$, it is enough to show that $\mathbf{\Sigma}$ contains all negations of atomic formulas ($t_0 \neq t_1$).

To see establish this, we notice that every formula of the form $t \neq u$ (where t and u are terms) satisfies

- $\mathcal{R}ob.+I\Sigma_1^0 \vdash_c \forall x_1 \ \ldots \forall x_n \ \Big(t \neq u \longleftrightarrow \big(\exists x_0 \ t+Sx_0 = u \ \vee \ \exists x_0 \ u+Sx_0 = t\big)\Big).$

Since the formula $\forall x_1 \ \ldots \forall x_n \ \big(\exists x_0 \ t+Sx_0 = u \ \vee \ \exists x_0 \ u+Sx_0 = t\big)$ belongs to $\mathbf{\Sigma}$, we have

- $\mathcal{R}ob.{+}I\Sigma^0_1 \vdash_c \forall x_1 \ldots \forall x_n \ \Big(\big(\exists x_0\ t{+}Sx_0 = u \ \vee\ \exists x_0\ u{+}Sx_0 = t\big) \to \exists x_0\ \varphi_{proof_{\mathcal{R}ob.}}(x_0, \ulcorner(\exists x_0\ t{+}Sx_0 = u \ \vee\ \exists x_0\ u{+}Sx_0 = t)\urcorner)\Big)$

then, from the "code" of a "proof" of $\big(\exists x_0\ t{+}Sx_0 = u \ \vee\ \exists x_0\ u{+}Sx_0 = t\big)$ we recover the "code" of a "proof" of $t \neq u$.

□

We finally have everything we need in order to prove:

Theorem 4.1: Gödel's 2nd incompleteness theorem

If $T \supseteq \mathcal{R}ob.{+}I\Sigma^0_1$ is any consistent recursive theory, then

$$T \nvdash_c cons(T).$$

Proof of Theorem 4.1:

Follows immediately from Lemmas 3.1 and 4.6. Because the condition required by Lemma 3.1 on a theory $T \supseteq \mathcal{R}ob.$ to satisfy the conditions of the second incompleteness theorem was that

$$T \ \vdash_c \ \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \longrightarrow \exists x_1 \varphi_{proof_{\mathcal{R}ob.}}\big(x_1, \lceil \ulcorner \Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]} \urcorner \rceil\big)\,.$$

And it turns out that the formula $\Xi_{[[\ulcorner\neg\Xi\urcorner]/x_0]}$ is both closed and Σ^0_1 since it is

$$\exists x_1\ \exists x_2 \ \ \big(\varphi_{proof_T}(x_1, x_2) \ \wedge\ \varphi_{diag}(x_2, \lceil\ulcorner\neg\Xi\urcorner\rceil)\big)$$

where both $\varphi_{proof_T}(x_1, x_2)$ and $\varphi_{diag}(x_2, \lceil\ulcorner\neg\Xi\urcorner\rceil)$ are formulas that represent primitive recursive relations, hence Σ^0_1.

□

Gödel second incompleteness Theorem yields some strange consequence.

Corollary 4.2

Let $T \supseteq \mathcal{R}ob.{+}I\Sigma^0_1$ be any recursive theory.

$$T \text{ is consistent}$$

$$\Longleftrightarrow$$

$$\text{there exists } \mathcal{M} \text{ s. t. } \mathcal{M} \models T \cup \Big\{\exists \ulcorner\varphi\urcorner \ \big(\exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\varphi\urcorner) \ \wedge\ \exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\neg\varphi\urcorner)\big)\Big\}.$$

Proof of Corollary 4.2:

($\Rightarrow$) If $T \supseteq \mathcal{R}ob. + I\Sigma_1^0$ is any consistent recursive theory, then $T \nvdash_c cons(T)$ holds by Theorem 4.1. But by the completeness theorem, this is equivalent to $T \not\models cons(T)$ which means that there exists some model $\mathcal{M}$ such that $\mathcal{M} \models T$ and $\mathcal{M} \not\models cons(T)$. Therefore, $\mathcal{M} \models \neg cons(T)$, hence

$$\mathcal{M} \models \exists \ulcorner\varphi\urcorner \left(\exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\varphi\urcorner) \ \wedge\ \exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\neg\varphi\urcorner)\right).$$

($\Leftarrow$) is immediate.

$\square$

Notice that the model $\mathcal{M}$ in Corollary 4.2 cannot be the standard model ($\mathbb{N}$) otherwise we would have

$$\mathbb{N} \models \exists \ulcorner\varphi\urcorner \left(\exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\varphi\urcorner) \ \wedge\ \exists x_0\ \varphi_{proof_T}(x_0, \ulcorner\neg\varphi\urcorner)\right).$$

Because, otherwise we would get standard integers that code both a proof of some formula φ and a proof of $\neg\varphi$, hence from these standard integers we would be able to recover both a proof of φ and a proof of $\neg\varphi$, which would lead to a proof of $\perp$.

So $\mathcal{M}$ is some non-standard model which contains two integers (necessarily at least one of them is non-standard) r_φ and $r_{\neg\varphi}$ which code a proof of φ from T and a proof of $\neg\varphi$ from T, respectively. Then, of course, from both r_φ and $r_{\neg\varphi}$ one may easily construct an integer $r_\perp$ which codes a proof of $\perp$ from T. However, since $r_\perp$ is some non-standard integer, it does not help in providing a proof of a contradiction from T.

Notice also that given any consistent recursive theory $T \supseteq \mathcal{R}ob. + I\Sigma_1^0$, it is easy to find some consistent recursive extension T' of T that satisfies $T' \vdash_c cons(T)$. This can be done, for instance, by simply taking $T' = T \cup cons(T)$. Indeed, T' is clearly recursive, and it is consistent because

$$T, cons(T) \vdash_c \perp \quad \Longleftrightarrow \quad T' \vdash_c \neg cons(T) \quad \Longleftrightarrow \quad T \vdash_c \perp.$$

By induction on $n \in \mathbb{N}$, we construct a family $(T_n)_{n\in\mathbb{N}}$ of theories ordered by inclusion:

- $T_0 = T$
- $T_{n+1} = T_n \cup \{cons(T_n)\}$.

Clearly, every theory T_n is both consistent and recursive. To see that it is consistent, we proceed by induction. Assuming that T_n is consistent, it follows that if T_{n+1} is inconsistent, then we have

$$T_n, cons(T_n) \vdash_c \perp \quad \Longleftrightarrow \quad T_n \vdash_c \neg cons(T_n) \quad \Longleftrightarrow \quad T_n \vdash_c \perp.$$

Now we have for each integer n:

$$T_{n+1} \vdash_c cons(T_n) \quad but \quad T_{n+1} \nvdash_c cons(T_{n+1}).$$

Let T_ω be the theory $T_\omega = \bigcup_{n\in\mathbb{N}} T_n$, T_ω is consistent since otherwise, by compactness $T_\omega \vdash_c \bot$ would yield $T_n \vdash_c \bot$ for some large enough integer n.
We see that this procedure can be extended all along the recursive countable ordinals.

Definition 4.2

A countable ordinal α is recursive if there exists some well ordering $(\mathbb{N}, \prec)$ such that

(1) $\{(a,b) \in \mathbb{N}\times\mathbb{N} \mid a \prec b\}$ is recursive, and

(2) α is order-isomorphic to $(\mathbb{N}, \prec)$.

It is not difficult to show that the class of recursive ordinals is some countable initial segment of the class of all ordinals. This justifies the following definition.

Definition 4.3

The Church-Kleene ordinal ω_1^{CK} is the least ordinal which is not recursive.

Since there are only countably many recursive subsets of $\mathbb{N}\times\mathbb{N}$, it follows that $\omega_1^{\mathrm{CK}} < \omega_1$.

- $T_0 = T$
- $T_{\alpha+1} = T_\alpha \cup \{cons(T_\alpha)\}$.
- $T_\lambda = \bigcup_{\alpha<\lambda} T_\alpha \qquad (\lambda \text{ limit and } \lambda < \omega_1^{\mathrm{CK}})$.

To show that every T_α is consistent (any $\alpha < \omega_1^{\mathrm{CK}}$), it remains to show that T_λ is consistent when λ is a limit ordinal. This is immediate, since

$$T_\lambda = \bigcup_{\alpha<\lambda} T_\alpha \text{ and } T_\lambda \vdash_c \bot \text{ implies } T_\alpha \vdash_c \bot \text{ holds for some } \alpha < \lambda.$$

It follows that every T_α is consistent and does not prove its own consistency, but the consistency of all T_β for $\beta < \alpha$.