



Logic and Logical Philosophy
Volume 6 (1998), 173–185

Dirk Ullrich

**A GENERAL PRINCIPLE FOR
PURELY MODEL-THEORETICAL
PROOFS OF GÖDEL'S SECOND
INCOMPLETENESS THEOREM**

Abstract. By generalizing Kreisel's proof of the Second Incompleteness Theorem of Gödel I extract a general principle which can also be used for other purely model-theoretical proofs of that theorem.

CONTENTS

1. Preliminaries, p. 175
2. The Arithmetized Completeness Theorem and the Model Chain Lemma, p. 180
3. A proof of the Second G.I.T. following Kreisel, p. 184

Received October 2, 1998

Concerning the tools used there are three types of proofs for the Second Gödelian Incompleteness Theorem (the last three words I will shorten to “G.I.T.”): purely proof-theoretical ones, purely model-theoretical proofs and such of mixed type.

A purely proof-theoretical proof of the Second G.I.T. normally takes a part of a proof for the First G.I.T. and formalizes it within the theory Φ just considered. To carry this through, some special properties of the formula representing the provability predicate \vdash_{Φ} of Φ have to be established. These special properties are usually expressed as the well-known Derivability Conditions; and a purely proof-theoretical proof only uses proof-theoretical means to establish these conditions. The classical proof of the Second G.I.T. is a typical example of this kind.¹

A proof of the Second G.I.T. of mixed type runs similar to a purely proof-theoretical one but also uses model-theoretical tools. By doing this the proof becomes more intuitive — at least for someone (like me) not used to think within formal systems. For instance, it is possible to follow the classical proof but to establish the Derivability Conditions model-theoretically.²

Last but not least, a purely model-theoretical proof uses (almost) only model-theoretical means — especially it works with a representation of satisfaction or truth in some or all models of Φ instead of a representation of \vdash_{Φ} . Georg Kreisel’s proof of the Second G.I.T. is such a proof.³

In this paper I extract a general principle used in the proof of the Second G.I.T. of Kreisel stated below as Model Chain Lemma. With the help of this general principle it is possible to get purely model-theoretical proofs of the Second G.I.T. by a model-theoretical “translation” of that part of a proof for the first G.I.T. which leads to the sentence undecidable in the respective theory. In this sense Kreisel’s proof of the Second G.I.T. is such a translation of the classical proof of the first G.I.T., and in [8] I give a proof of the Second G.I.T. translating Boolos’ proof of the first G.I.T. resting on Berry’s paradox.

The rest of the paper is divided into three sections. Section 1 will only fix terminology and recall some general results. The following section 2 deals with the Arithmetized Completeness Theorem and the Model Chain Lemma,

¹ For complete proofs following classical lines see e.g., [2], p. 15–50, or [3].

² Such a proof you can find, e.g., in [4], p. 163f., or in [6].

³ For instance, [1], p. 192–194, [6] and [7], p. 862f., give this proof.

and the final section 3 reformulates Kreisel's proof of the Second G.I.T. by using the Model Chain Lemma. For the sake of simplicity I will restrict myself to Peano Arithmetic PA, but the following results are generalizable to more general settings in a direct manner.

Acknowledgement. Many thanks to my friends Friederike Körner and Jörg Ackermann. Both helped me to translate my 'private English' into a more commonly used form. And Friederike toiled to understand the first, rather confused, version of this paper and helped me to improve it.

1. Preliminaries

This section only reminds of some well-known facts.

\mathcal{L}_{PA} is the first-order language of arithmetic, possessing the non-logical constants $\bar{0}$, \mathbf{s} (successor), $+$, \cdot , $=$ and $<$, the logical operators \wedge , \vee , \rightarrow , \leftrightarrow , \forall and \exists , the variables \mathbf{v}_i ($i \in \mathbb{N}$, where \mathbb{N} is the set of all natural numbers) and the brackets $(,)$. For the sake of later convenience all these \mathcal{L}_{PA} -symbols should be appropriate finite sequences of length one; then the \mathcal{L}_{PA} -strings can be chosen as concatenations of finitely many \mathcal{L}_{PA} -symbols. Since I only deal with the language \mathcal{L}_{PA} here I usually suppress its mentioning. For any (\mathcal{L}_{PA} -) term t , $t(t_0, \dots, t_{k-1})$ is the result of substituting the term t_i for all free occurrences of \mathbf{v}_i in t for all $i < k$. Similarly $\phi(t_0, \dots, t_{k-1})$ is used for formulas ϕ . For every $n \in \mathbb{N}$, \bar{n} is $\underbrace{\mathbf{s} \cdots \mathbf{s}}_{n \text{ times}} \bar{0}$.

An (\mathcal{L}_{PA} -) *theory* is a set of (\mathcal{L}_{PA} -) sentences, an (\mathcal{L}_{PA} -) *semi-theory* is a set of (\mathcal{L}_{PA} -) formulas; and $\vdash_{\Phi} \phi$ means that the formula ϕ is provable within the (semi-) theory Φ .

Δ_0^{PA} is the set of all formulas ϕ with $\vdash_{\text{PA}} \phi \leftrightarrow \psi$ for some formula ψ containing no or only bounded quantifier prefixes $\forall(x < t)$ or $\exists(x < t)$. Σ_1^{PA} exactly consists of those formulas ϕ with $\vdash_{\text{PA}} \phi \leftrightarrow \exists x \psi$ for some Δ_0^{PA} -formula ψ . And Δ_1^{PA} contains all Σ_1^{PA} -formulas ϕ with $\neg \phi \in \Sigma_1^{\text{PA}}$, too. All these sets possess some nice and well-known closure properties: With $\phi, \psi \in \Delta_0^{\text{PA}}$ ($\in \Delta_1^{\text{PA}}$ respectively) $\neg \phi$, $\phi \wedge \psi$, $\phi \vee \psi$, $\phi \rightarrow \psi$, $\phi \leftrightarrow \psi$, $\forall(x < t)\phi$, $\exists(x < t)\phi \in \Delta_0^{\text{PA}}$ ($\in \Delta_1^{\text{PA}}$ respectively); and from $\phi, \psi \in \Sigma_1^{\text{PA}}$ follows that $\phi \wedge \psi$, $\phi \vee \psi$, $\forall(x < t)\phi$, $\exists x \phi \in \Sigma_1^{\text{PA}}$ (and $\phi \rightarrow \psi \in \Sigma_1^{\text{PA}}$, if even $\phi \in \Delta_1^{\text{PA}}$ holds). Furthermore $\Delta_0^{\text{PA}} \subseteq \Delta_1^{\text{PA}} \subseteq \Sigma_1^{\text{PA}}$ is true.

Let \mathcal{S} be an (\mathcal{L}_{PA} -) structure. Then $|\mathcal{S}|$ is the universe of \mathcal{S} . For every term t containing at most $\mathbf{v}_0, \mathbf{v}_1, \dots, \mathbf{v}_{k-1}$ free, $t^{\mathcal{S}}[s_0, \dots, s_{k-1}]$ is the value of t when assigning $s_i \in |\mathcal{S}|$ to \mathbf{v}_i for all $i < k$, and, for any formula ϕ ,

$\mathcal{S} \models \phi[s_0, \dots, s_{k-1}]$ means that ϕ is always true in \mathcal{S} when assigning $s_i \in |\mathcal{S}|$ to v_i for all $i < k$, and assigning arbitrary $s \in |\mathcal{S}|$ to all other variables occurring free in ϕ . $\underline{\mathbb{N}}$ is the standard model of PA; hence $|\underline{\mathbb{N}}| = \mathbb{N}$ holds. For any theory Φ a structure \mathcal{M} is a *model of Φ* iff $\mathcal{M} \models \phi$ is true for all $\phi \in \Phi$.

The importance of Σ_1^{PA} - and Δ_1^{PA} -formulas results from their ‘nice behaviour’ under embeddings of models of PA. An *embedding E of a structure \mathcal{S} into a structure \mathcal{T}* is a 1-1 function $E: |\mathcal{S}| \rightarrow |\mathcal{T}|$ such that, for all $s, t \in |\mathcal{S}|$, $E(\mathbf{s}^{\mathcal{S}}(s)) = \mathbf{s}^{\mathcal{T}}(E(s))$, $E(s \circ^{\mathcal{S}} t) = E(s) \circ^{\mathcal{T}} E(t)$ for $\circ \in \{+, \cdot\}$ and, finally, $s <^{\mathcal{S}} t$ iff $E(s) <^{\mathcal{T}} E(t)$ hold. And an *initial segment A of \mathcal{S}* is an $A \subseteq |\mathcal{S}|$ such that $s <^{\mathcal{S}} t \in A$ implies $s \in A$. Now can be stated:

Lemma 1.1. (embedding lemma) *Let \mathcal{M} and \mathcal{N} be models of PA and E an embedding of \mathcal{M} onto an initial segment of \mathcal{N} . Then, for any Σ_1^{PA} -formula (Δ_1^{PA} -formula) ϕ such that at most v_0, v_1, \dots, v_{k-1} occur free in ϕ ,*

(E) $\mathcal{N} \models \phi[E(s_0), E(s_1), \dots, E(s_{k-1})]$ if (iff) $\mathcal{M} \models \phi[s_0, s_1, \dots, s_{k-1}]$.

holds for all $s_0, s_1, \dots, s_{k-1} \in |\mathcal{S}|$.

Proof. The proof amounts to a straightforward formula induction over ϕ .⁴ \square

An important corollary of this lemma establishes the so-called Σ_1 -completeness of PA. I give a model-theoretical formulation because I will later need this one:

Corollary 1.2. (Σ_1 -completeness of PA) *Let \mathcal{M} be a model of PA and $\phi \in \Sigma_1^{\text{PA}}$ ($\phi \in \Delta_1^{\text{PA}}$) with at most v_0, v_1, \dots, v_{k-1} occurring free in ϕ . Then, for all $n_0, n_1, \dots, n_{k-1} \in \mathbb{N}$, $\mathcal{M} \models \phi(\overline{n_0}, \overline{n_1}, \dots, \overline{n_{k-1}})$ holds if (iff) $\underline{\mathbb{N}} \models \phi[n_0, n_1, \dots, n_{k-1}]$.*

Proof. Using that \mathcal{M} is a model of PA it is not difficult to show that $C_{\mathcal{M}}: \mathbb{N} \rightarrow |\mathcal{M}|$ with $C_{\mathcal{M}}(n) =_{\text{def}} \overline{n}^{\mathcal{M}}$ is an embedding of $\underline{\mathbb{N}}$ into \mathcal{M} and $\text{Rng}(C_{\mathcal{M}}) = \{\overline{n}^{\mathcal{M}} \mid n \in \mathbb{N}\}$ is an initial segment of \mathcal{M} . Thus the claim directly follows from lemma 1.1. \square

For any structure \mathcal{S} , a k -ary relation $R \subseteq |\mathcal{S}|^k$ is *defined by ϕ in \mathcal{S}* iff ϕ is a formula containing exactly v_0, v_1, \dots, v_{k-1} free and for all $s_0, s_1, \dots, s_{k-1} \in |\mathcal{S}|$ the following holds: $R(s_0, s_1, \dots, s_{k-1})$ iff $\mathcal{S} \models \phi[s_0, s_1, \dots, s_{k-1}]$. A k -ary function $F: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ is *defined by ϕ in \mathcal{S}* iff its ‘graph’, i.e. the relation $\{\langle s_0, s_1, \dots, s_k \rangle \mid s_k = F(s_0, s_1, \dots, s_{k-1})\}$, is defined by ϕ in \mathcal{S} .

⁴ Details can be found, for example, in [5], p. 24f.

Finally, an *individual* $s \in |\mathcal{S}|$ is *defined by* ϕ in \mathcal{S} iff the (0-ary) function $\emptyset \mapsto s$ is defined by ϕ in \mathcal{S} . Every formula ϕ defines at most one relation, function or individual in \mathcal{S} ; in this case $\phi^{\mathcal{M}}$ is the respective relation, function or individual. Furthermore a formula ϕ defines some k -ary relation in \mathcal{S} iff ϕ contains exactly v_0, v_1, \dots, v_{k-1} free; whereas ϕ defines a k -ary function in \mathcal{S} iff exactly v_0, v_1, \dots, v_k occur free in ϕ and $\mathcal{S} \models \forall v_0 \forall v_1 \dots \forall v_{k-1} \exists v_k \phi$ holds; and ϕ defines some individual in \mathcal{M} iff ϕ contains exactly v_0 free and $\mathcal{S} \models \exists v_0 \phi$ holds. Some useful possibilities to get definable relations or functions from given ones should be mentioned. Piecewise PA-definition of functions is simple: If ϕ, ψ and χ are formulas then there is a formula $\{\phi : \chi/\psi\}$ such that, if ϕ, ψ and χ define $G, H: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ and $R \subseteq |\mathcal{S}|^k$, then $\{\phi : \chi/\psi\}$ defines in \mathcal{S} the function $F: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ with $F(s_0, \dots, s_{k-1}) = G(s_0, \dots, s_{k-1})$ iff $R(s_0, \dots, s_{k-1})$, and $F(s_0, \dots, s_{k-1}) = H(s_0, \dots, s_{k-1})$ otherwise. Furthermore $\{\phi : \chi/\psi\} \in \Sigma_1^{\text{PA}}$ ($\{\phi : \chi/\psi\} \in \Delta_1^{\text{PA}}$) can be chosen if one has $\phi, \psi \in \Sigma_1^{\text{PA}}$ ($\phi, \psi \in \Delta_1^{\text{PA}}$) and $\chi \in \Delta_1^{\text{PA}}$. The next example concerns PA-substitution: For any formulas $\phi, \psi_0, \psi_1, \dots, \psi_{l-1}$ a formula $\phi\{\psi_0, \psi_1, \dots, \psi_{l-1}\}$ exists such that, if $\psi_0, \dots, \psi_{l-1}$ define the functions $G_0: |\mathcal{S}|^k \rightarrow |\mathcal{S}|, \dots, G_{l-1}: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$, then $\phi\{\psi_0, \psi_1, \dots, \psi_{l-1}\}$ defines the relation $R \subseteq |\mathcal{S}|^k$ with $R(s_0, \dots, s_{k-1})$ iff $S(G_0(s_0, \dots, s_{k-1}), \dots, G_{l-1}(s_0, \dots, s_{k-1}))$ in \mathcal{S} if ϕ defines in \mathcal{S} the relation $S \subseteq |\mathcal{S}|^l$; and $\phi\{\psi_0, \psi_1, \dots, \psi_{l-1}\}$ defines the function $F: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ with $F(s_0, \dots, s_{k-1}) = H(G_0(s_0, \dots, s_{k-1}), \dots, G_{l-1}(s_0, \dots, s_{k-1}))$ if ϕ defines in \mathcal{S} the function $H: |\mathcal{S}|^l \rightarrow |\mathcal{S}|$. If $\phi \in \Sigma_1^{\text{PA}}$ ($\phi \in \Delta_1^{\text{PA}}$) and $\psi_0, \dots, \psi_{l-1} \in \Sigma_1^{\text{PA}}$ hold then $\phi\{\psi_0, \psi_1, \dots, \psi_{l-1}\}$ belongs to Σ_1^{PA} (to Δ_1^{PA}) too. The following example is extremely important and provides some kind of PA-recursion: For any two formulas ϕ and ψ there is a formula $\text{rec}\{\phi, \psi\}$ such that $\text{rec}\{\phi, \psi\}$ defines in \mathcal{S} a function $F: |\mathcal{S}|^{k+1} \rightarrow |\mathcal{S}|$ satisfying the two conditions $F(s_0, \dots, s_{k-1}, \bar{0}^{\mathcal{S}}) = G(s_0, \dots, s_{k-1})$ and $F(s_0, \dots, s_{k-1}, \mathbf{s}^{\mathcal{S}}(s)) = H(s_0, \dots, s_{k-1}, s, F(s_0, \dots, s_{k-1}, s))$ if ϕ defines $G: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ and ψ defines $H: |\mathcal{S}|^{k+2} \rightarrow |\mathcal{S}|$ in \mathcal{S} respectively; and $\text{rec}\{\phi, \psi\} \in \Sigma_1^{\text{PA}}$ ($\text{rec}\{\phi, \psi\} \in \Delta_1^{\text{PA}}$) follows from $\phi, \psi \in \Sigma_1^{\text{PA}}$ ($\phi, \psi \in \Delta_1^{\text{PA}}$). Finally, PA-minimization should be mentioned: For any formula ϕ a formula $\mu\{\phi\}$ exists such that $\mu\{\phi\} \in \Sigma_1^{\text{PA}}$ follows from $\phi \in \Delta_1^{\text{PA}}$, and $\mu\{\phi\}$ defines in \mathcal{S} the function $F: |\mathcal{S}|^k \rightarrow |\mathcal{S}|$ with $F(s_0, \dots, s_{k-1}) = \min^{<^{\mathcal{S}}} \{s \mid R(s_0, \dots, s_{k-1}, s)\}$ if ϕ defines in \mathcal{S} the relation $R \subseteq |\mathcal{S}|^{k+1}$ and for every $s_0, \dots, s_{k-1} \in |\mathcal{S}|$ there is a $s \in |\mathcal{S}|$ with $R(s_0, \dots, s_{k-1}, s)$. A further property of Σ_1^{PA} -formulas is important: If a formula $\phi \in \Sigma_1^{\text{PA}}$ defines a function in *every* model of PA then even $\phi \in \Delta_1^{\text{PA}}$ holds.

The definability of relations, functions and individuals in structures is used to formalize a certain portion of logic within PA which is of crucial importance for the Second G.I.T. This very formalization I will call “representation in PA”; I explain it for a k -ary relation $R \subseteq \mathbb{N}^k$ (functions and individuals have to be treated analogously): To represent R in PA it doesn’t suffice to find a formula ϕ that defines R within the standard model $\underline{\mathbb{N}}$; moreover ϕ has to define in every model \mathcal{M} of PA an $R_{\mathcal{M}} \subseteq \mathbb{N}^k$ such that $R(n_0, n_1, \dots, n_{k-1})$ iff $R_{\mathcal{M}}(\overline{n_0}^{\mathcal{M}}, \overline{n_1}^{\mathcal{M}}, \dots, \overline{n_{k-1}}^{\mathcal{M}})$ holds for all $n_0, n_1, \dots, n_{k-1} \in \mathbb{N}$ and $R_{\mathcal{M}}$ ($= \phi^{\mathcal{M}}$) possesses ‘similar properties’ as R ($= \phi^{\underline{\mathbb{N}}}$) ‘in the real world’. (The first of this two requirements is easily achieved by Σ_1^{PA} -completeness if $\phi \in \Delta_1^{\text{PA}}$ can be chosen.)

Let us start with finite sequences of natural numbers. Since only relations, functions and individuals of \mathbb{N} can be represented in PA, we have to code the finite sequences of natural numbers by natural numbers. Therefore fix a 1-1 function $\text{fs}: \{s \mid s \text{ finite sequence in } \mathbb{N}\} \rightarrow \mathbb{N}$. Now choose Δ_1^{PA} -formulas **fseq**, **len**, \square , **memb** and **app** such that **fseq** represents in PA the 1-ary relation expressed by the phrase “ n_0 is the code of a finite sequence” (i.e. there is an $i \in \mathbb{N}$ with $n_0 = \text{fs}(i)$), **len** the 1-ary function expressed by “ n_1 is the length of the finite sequence with the code n_0 ” (i.e. $n_1 = \text{Len}(\text{fs}^{-1}(n_0))$, where $\text{Len}(\alpha)$ is the length of the finite sequence α), \square the individual coding the empty sequence, **memb** the 2-ary function expressed by “ n_2 is the n_0 th member of the finite sequence with code n_1 ” and **app** the 2-ary function expressed by “ n_2 is the code of the finite sequence which results from the finite sequence with code n_0 by appending n_1 as last element”.⁵ These formulas can be chosen such that, for any model \mathcal{M} of PA and $s, e \in |\mathcal{M}|$, $s, e <^{\mathcal{M}} \mathbf{app}^{\mathcal{M}}(s, e)$ holds if $s \in \mathbf{fseq}^{\mathcal{M}}$.⁶ Using these basic formulas all other interesting relations and functions for finite sequences can be defined by PA-substitution, piecewise PA-definition, PA-recursion and PA-minimization. Below formulas \frown and **last** are needed representing in PA the 1-ary function expressed by “ n_1 is the code of the finite sequence resulting from concatenation of all finite sequences being members of the finite sequence with code n_0 ” and the 1-ary function expressed by “ n_1 is the

⁵ I use the following canonical notation to express the connection between a formula ϕ and the relation or function represented by ϕ : n_i is always the $(i + 1)$ th argument of the relation or function just considered, and, for a function F , the n_i with the highest index is the value of F .

⁶ Here the notation for definable relations, functions and individuals introduced above is used: $\mathbf{app}^{\mathcal{M}}$ is the function defined by **app** in \mathcal{M} ; thus $\mathbf{app}^{\mathcal{M}}$ has two arguments and not three!

last member of the finite sequence with code n_0 ". Note that \frown , **last** $\in \Delta_1^{\text{PA}}$ are possible. Moreover, by using \square and **app** define recursively for every $k \in \mathbb{N}$ a formula $[\dots]_k \in \Delta_1^{\text{PA}}$ containing exactly v_0, v_1, \dots, v_k free such that, for every model \mathcal{M} of PA and $m_0, m_1, \dots, m_{k+1} \in |\mathcal{M}|$, $[\]_0^{\mathcal{M}} = \square^{\mathcal{M}}$ and $[m_0, m_1, \dots, m_{k+1}]_{k+1}^{\mathcal{M}} = \mathbf{app}^{\mathcal{M}}([m_0, m_1, \dots, m_k]_k^{\mathcal{M}}, m_{k+1})$ hold.

Genuinely logical notions mostly concern relations, functions or individuals of strings; therefore a formalization of them in PA needs a suitable Gödel numbering, i.e. a 1-1 function $\text{gn}: \{\alpha \mid \alpha \text{ } \mathcal{L}_{\text{PA}}\text{-string}\} \rightarrow \mathbb{N}$. Above the (\mathcal{L}_{PA} -) strings were chosen to be certain finite sequences; hence we should take a gn such that we have, for all models \mathcal{M} of PA and strings α , $\underline{\alpha}^{\mathcal{M}} \in \mathbf{fseq}^{\mathcal{M}}$, $\mathbf{len}^{\mathcal{M}}(\underline{\alpha}^{\mathcal{M}}) = \overline{\text{Len}(\alpha)}^{\mathcal{M}}$ and $\underline{\alpha}^{\mathcal{M}} = \frown[\alpha(0), \alpha(1), \dots, \alpha(\text{Len}(\alpha) - 1)]_{\text{Len}(\alpha)}^{\mathcal{M}}$, where $\underline{\alpha} =_{\text{def}} \overline{\text{gn}(\alpha)}$. (Choose, e.g., $\text{gn}(\alpha) \in \mathbf{fseq}^{\mathbb{N}}$ with $\mathbf{len}^{\mathbb{N}}(\alpha) = 1$ for all \mathcal{L}_{PA} -symbols α and set $\text{gn}(\alpha) = \frown^{\mathbb{N}}([\text{gn}(\alpha(0)), \text{gn}(\alpha(1)), \dots, \text{gn}(\text{Len}(\alpha) - 1)]_{\text{Len}(\alpha)}^{\mathbb{N}})$ for all strings α with $\text{Len}(\alpha) > 1$. Because of **fseq**, $\mathbf{len}[\dots]_k \in \Delta_1^{\text{PA}}$ we get what we want.) Now it is not difficult to proceed: Fix suitable Δ_1^{PA} -formulas **form** and **sent** to represent in PA the 1-ary relation expressed by " n_0 is the Gödel number of a formula" and the 1-ary relation expressed by " n_0 is the Gödel number of a sentence"; choose both formulas such that $\mathcal{M} \models \mathbf{form} \rightarrow \mathbf{fseq}$ for every model \mathcal{M} of PA holds.⁷

The next step concerns formalization of theories and provability. Let \mathcal{M} be a model of PA. A *semi-theory formula with k parameters in \mathcal{M}* is a formula ϕ such that exactly v_0, v_1, \dots, v_k occur free in ϕ and $\mathcal{M} \models \phi \rightarrow \mathbf{form}$ holds. Similarly, ϕ is a *theory formula with k parameters in \mathcal{M}* iff exactly v_0, v_1, \dots, v_k occur free in ϕ and $\mathcal{M} \models \phi \rightarrow \mathbf{sent}$ holds. If ϕ is a (semi-) theory formula with k parameters in every model \mathcal{M} of PA then for every $n_0, n_1, \dots, n_{k-1} \in \mathbb{N}$ the set $\Phi(n_0, n_1, \dots, n_{k-1}) =_{\text{def}} \{\chi \text{ } \mathcal{L}_{\text{PA}}\text{-formula} \mid \mathbb{N} \models \phi(\underline{\chi}, \overline{n_0}, \overline{n_1}, \dots, \overline{n_{k-1}})\}$ is a (semi-) theory. A well-known result tells us that in this case there is a formula **prov** $_{\phi}$ having the same free variables as ϕ and a formula **con** $_{\phi}$ with the same free variables as ϕ except v_0 such that **prov** $_{\phi}, \neg \mathbf{con}_{\phi} \in \Sigma_1^{\text{PA}}$ follows from $\phi \in \Sigma_1^{\text{PA}}$ and, for every $n_0, n_1, \dots, n_{k-1} \in \mathbb{N}$, **prov** $_{\phi}(v_0, \overline{n_0}, \overline{n_1}, \dots, \overline{n_{k-1}})$ represents the 1-ary relation expressed by " n_0 is the Gödel number of a formula provable in $\Phi(n_0, n_1, \dots, n_{k-1})$ " and **con** $_{\phi}(\overline{n_0}, \overline{n_1}, \dots, \overline{n_{k-1}})$ represents the 0-ary relation expressed by " $\Phi(n_0, n_1, \dots, n_{k-1})$ is consistent in PA". There is an important connection between theories 'in the real world' and theory for-

⁷ For details of representing gödelized versions of syntactical notions, consult, for instance, [3], [4], ch. 0 and ch. 1, or [5], ch. 9.

mulas: If Φ is a recursively enumerable (\mathcal{L}_{PA} -) theory (i.e. the class $\{\text{gn}(\phi) \mid \phi \in \Phi\}$ is recursively enumerable) then there is a Σ_1^{PA} -formula χ defining $\{\text{gn}(\phi) \mid \phi \in \Phi\}$ in $\underline{\mathbb{N}}$. Take the shortest such χ and set $\text{th}_\Phi =_{\text{def}} \chi \wedge \mathbf{sent}$; then th_Φ defines $\{\text{gn}(\phi) \mid \phi \in \Phi\}$ in $\underline{\mathbb{N}}$ too, is a Σ_1^{PA} -formula and a theory formula without parameters for every model \mathcal{M} of PA.

It is time to turn to the Model Chain Lemma.

2. The Arithmetized Completeness Theorem and the Model Chain Lemma

To state the Model Chain Lemma properly I need a special version of the Arithmetized Completeness Theorem, i.e. of a formalization of the ordinary Completeness Theorem within PA, and some more terminology. $\text{Frm}: \mathbb{N} \rightarrow \{\phi \mid \phi \text{ formula}\}$ is the function recursively defined by $\text{Frm}(0) =_{\text{def}} \min^{\prec_{\text{gn}}} \{\phi \mid \phi \text{ formula}\}$ and $\text{Frm}(n+1) =_{\text{def}} \min^{\prec_{\text{gn}}} \{\phi \mid \phi \text{ formula and } \text{Frm}(n) \prec_{\text{gn}} \phi\}$ where $\phi \prec_{\text{gn}} \psi$ iff $\text{gn}(\phi) < \text{gn}(\psi)$; Frm is 1-1 with $\text{Rng}(\text{Frm}) = \{\phi \mid \phi \text{ formula}\}$. A *formula path* is a finite sequence Γ with $\Gamma(i) = \text{Frm}(i)$ or $\Gamma(i) = \neg \text{Frm}(i)$ for all $i < \text{Len}(\Gamma)$; if Γ and Δ are formula paths then $\Delta \prec_{\neg} \Gamma$ iff $\text{Len}(\Delta) \leq \text{Len}(\Gamma)$ and there is a $i < \text{Len}(\Delta)$ such that $\Delta(i) = \neg \Gamma(i)$ and $\Delta(j) = \Gamma(j)$ for all $j < i$. Using PA-recursion and PA-minimization one can obtain a formula $\mathbf{frm} \in \Delta_1^{\text{PA}}$ representing the 1-ary function expressed by “ n_1 is the Gödel number of $\text{Frm}(n_0)$ ”: chose, for instance, $\mathbf{frm} =_{\text{def}} \text{rec}\{\mu\{\mathbf{form}\}(v_0), \mu\{\mathbf{form}(v_2) \wedge v_1 < v_2\}\}$. With the help of \mathbf{frm} formulas $\mathbf{fpath}, \prec_{\neg} \in \Delta_1^{\text{PA}}$ can be defined representing in PA the relations expressed by “ n_0 is the code of a formula path” and “ n_0 is the code of the formula path Δ , n_1 is the code of the formula path Γ and $\Delta \prec_{\neg} \Gamma$ holds”. If ϕ is any formula, set $\phi^{\cup} =_{\text{def}} \phi \vee (\mathbf{form}(v_0) \wedge \exists v_2 \mathbf{memb}(v_2, v_1, v_0))$. For every model \mathcal{M} of PA ϕ^{\cup} is a semi-theory formula with one parameter in \mathcal{M} (and $\phi^{\cup} \in \Sigma_1^{\text{PA}}$) if ϕ is a semi-theory formula without parameters in \mathcal{M} (and $\phi \in \Sigma_1^{\text{PA}}$). For any model \mathcal{M} of PA and every semi-theory formula ϕ in \mathcal{M} without parameters p is a ϕ -leftmost formula path in \mathcal{M} iff $p \in \mathbf{fpath}^{\mathcal{M}}$ and $\mathbf{con}_{\phi^{\cup}}^{\mathcal{M}}(p)$ holds such that there is no $q \in \mathbf{fpath}^{\mathcal{M}}$ with $q \prec_{\neg}^{\mathcal{M}} p$ and $\mathbf{con}_{\phi^{\cup}}^{\mathcal{M}}(q)$. Finally, a model \mathcal{N} of an (\mathcal{L}_{PA} -) theory Φ is strongly defined in \mathcal{M} by $\rho, [\psi_i]_{i < 5}, \phi$ iff $|\mathcal{N}|$ is defined in \mathcal{M} by $\rho, \bar{0}^{\mathcal{N}}, \mathbf{s}^{\mathcal{N}}, +^{\mathcal{N}}, \cdot^{\mathcal{N}}, <^{\mathcal{N}}$ are defined by ψ_0, \dots, ψ_5 in \mathcal{M} , ϕ contains exactly v_0 free and

(D) $\mathcal{N} \models \chi$ iff $\mathcal{M} \models \phi(\underline{\chi})$ for every sentence χ

holds.

Now the Arithmetized Completeness Theorem can be stated:

Lemma 2.1. (Arithmetized Completeness Theorem) *Let Φ be a recursively enumerable theory. Then there are formulas $\mathbf{henk}_\Phi \in \Sigma_1^{\text{PA}}$, \mathbf{univ}_Φ , \mathbf{int}_{Φ_i} ($i < 5$) and \mathbf{true}_Φ such that for every model \mathcal{M} of PA the following holds: \mathbf{henk}_Φ and \mathbf{true}_Φ are semi-theory formulas without parameters in \mathcal{M} with $\mathcal{M} \models \text{th}_\Phi \rightarrow \mathbf{henk}_\Phi$,*

(L) $\mathbf{true}_\Phi^{\mathcal{M}}(a)$ iff $a \in \mathbf{form}^{\mathcal{M}}$ and there exists a \mathbf{henk}_Φ -leftmost formula path p in \mathcal{M} with $\mathbf{last}^{\mathcal{M}}(p) = a$;

and if $\mathcal{M} \models \mathbf{con}_{\text{th}_\Phi}$ holds then $\mathbf{univ}_\Phi, [\mathbf{int}_{\Phi_i}]_{i < 5}, \mathbf{true}_\Phi$ strongly define a model \mathcal{N} of Φ in \mathcal{M} .

Proof. The proof amounts to a more or less straightforward formalization of a proof of the ordinary Completeness Theorem by reasoning inside \mathcal{M} instead of “the real world”: Starting with th_Φ construct a suitable semi-theory formula \mathbf{henk}_Φ without parameters for every model of Φ such that \mathbf{henk}_Φ represents a Henkin semi-theory for Φ if $\mathcal{M} \models \mathbf{con}_{\text{th}_\Phi}$ holds. Then a formalization of the Lindenbaum completion leads to \mathbf{true}_Φ , and the remaining formulas result from formalizing the definition of the canonical term model of Φ within \mathcal{M} .⁸ \square

The next proposition establishes an important connection between a model of PA and another model being strongly defined in it:

Proposition 2.2. (embedding in strongly definable models) *Let \mathcal{M} and \mathcal{N} be models of PA. If \mathcal{N} is strongly defined in \mathcal{M} (by some $\rho, [\psi_i]_{i < 5}, \phi$) then there exists an embedding of \mathcal{M} onto an initial segment of \mathcal{N} .*

Proof. Since \mathcal{N} is strongly defined in \mathcal{M} by $\rho, [\psi_i]_{i < 5}, \phi$, in particular the individual $\bar{0}^{\mathcal{N}}$ and the 1-nary function $\mathbf{s}^{\mathcal{M}}$ are defined in \mathcal{M} by ψ_0 and ψ_1 , respectively. Thus, by PA-recursion, there is a formula η defining a function $E: |\mathcal{M}| \rightarrow |\mathcal{N}|$ with $E(\bar{0}^{\mathcal{M}}) = \bar{0}^{\mathcal{N}}$ and $E(\mathbf{s}^{\mathcal{M}}(m)) = \mathbf{s}^{\mathcal{N}}(E(m))$ for all $m \in |\mathcal{M}|$. Because $+\mathcal{N}$ and $\cdot^{\mathcal{N}}$ are defined in \mathcal{M} by ψ_2 and ψ_3 respectively, $E(m \circ^{\mathcal{M}} m') = E(m) \circ^{\mathcal{N}} E(m')$ for $\circ \in \{+, \cdot\}$ and any $m, m' \in |\mathcal{M}|$ is obtained by PA-induction. Finally by PA-induction can be shown that, for all $m, m' \in |\mathcal{M}|$, $m <^{\mathcal{M}} m'$ iff $E(m) <^{\mathcal{N}} E(m')$ holds and $\text{Rng}(E)$ is an

⁸ The detailed proof in [1], pp. 186-191, can be adopted to the terminology used here without difficulties.

initial segment of \mathcal{N} . To do this the fact that $\prec^{\mathcal{M}}$ and $|\mathcal{N}|$ are defined by ϕ_4 and ρ in \mathcal{M} is used. Thus E is an embedding of \mathcal{M} onto an initial segment of \mathcal{N} . \square

This proposition readily leads to the following corollary:

Corollary 2.3. (strongly definable models and Σ_1^{PA} -sentences) *Let \mathcal{M} and \mathcal{N} be models of PA. If \mathcal{N} is strongly defined in \mathcal{M} (by some $\rho, [\psi_i]_{i < 5}, \phi$) then $\mathcal{N} \models \chi$ follows from $\mathcal{M} \models \chi$ for every Σ_1^{PA} -sentence χ .*

Proof. Combine proposition 2.2 and the embedding lemma. \square

Now the Model Chain Lemma can be formulated and proved:

Lemma 2.4. (Model Chain Lemma) *There are not both a sequence $[\mathcal{M}_k]_{k \in \mathbb{N}}$ of models of PA and an $n^* \in \mathbb{N}$ such that the following conditions are both fulfilled:*

- (M 1) \mathcal{M}_{k+1} is strongly defined by $\mathbf{univ}_{\text{PA}}, [\mathbf{int}_{\text{PA}}]_{i < 5}, \mathbf{true}_{\text{PA}}$ in \mathcal{M}_k for every $k \in \mathbb{N}$, and
- (M 2) for every $k \in \mathbb{N}$ there is an $n_k \leq n^*$ such that $\text{Frm}(n_k)$ is a sentence with $\mathcal{M}_k \models \text{Frm}(n_k)$ but $\mathcal{M}_k \not\models \mathbf{true}_{\text{PA}}(\text{Frm}(n_k))$.

Proof. Lets start with some simple but useful observations. First, for all formula paths Γ and Δ with $\text{Len}(\Gamma) = \text{Len}(\Delta)$

$$(1) \quad \Gamma \prec_{\neg} \Delta \quad \text{or} \quad \Gamma = \Delta \quad \text{or} \quad \Delta \prec_{\neg} \Gamma$$

directly follows from the definition of \prec_{\neg} . Second, there is a close connection between formula paths ‘in the real world’ and formula path in a model \mathcal{M} of PA for all $n \in \mathbb{N}$: If Γ is a formula path with $\text{Len}(\Gamma) = n$ then $\sharp\Gamma^{\mathcal{M}} \in \mathbf{fpath}^{\mathcal{M}}$ and $\mathbf{len}^{\mathcal{M}}(\sharp\Gamma^{\mathcal{M}}) = \bar{n}^{\mathcal{M}}$ are true where $\sharp\Gamma =_{\text{def}} [\Gamma(0), \Gamma(1), \dots, \Gamma(n-1)]_n$. On the other hand, for any $p \in \mathbf{fpath}^{\mathcal{M}}$ with $\mathbf{len}^{\mathcal{M}}(p) = \bar{n}^{\mathcal{M}}$ we can take $\natural_{\mathcal{M}}p =_{\text{def}}$ the formula path Γ with $\text{Len}(\Gamma) = n$ and $\mathcal{M} \models \mathbf{memb}(\bar{i}, v_0, \underline{\Gamma(i)})[p]$ for every $i < n$. We get for any formula paths Γ, Δ

$$(2) \quad \natural_{\mathcal{M}}\sharp\Gamma^{\mathcal{M}} = \Gamma \quad \text{and} \quad \Gamma \prec_{\neg} \Delta \quad \text{iff} \quad \sharp\Gamma^{\mathcal{M}} \prec_{\neg}^{\mathcal{M}} \sharp\Delta^{\mathcal{M}}$$

because all formulas involved in the definitions of \sharp and $\natural_{\mathcal{M}}$ are Δ_1^{PA} . A further observation using (2) and the definition of “ $\mathbf{henk}_{\text{PA}}$ -leftmost path in \mathcal{M} ” leads to

- (3) If $\sharp\Gamma^{\mathcal{M}}$ is a **henk**_{PA}-leftmost path in \mathcal{M}
 than $\sharp[\Gamma(i)]_{i < m}^{\mathcal{M}}$ is a **henk**_{PA}-leftmost path in \mathcal{M} too.

for all formula paths Γ and $m \leq \text{Len}(\Gamma)$.

To get a contradiction let us now assume that there is a sequence $[\mathcal{M}_k]_{k \in \mathbb{N}}$ of models of PA and an $n^* \in \mathbb{N}$ such that (M1) and (M2) hold. First of all,

- (4) For all $k, m \in \mathbb{N}$ exactly one formula path Γ exists such that
 $\text{Len}(\Gamma) = m$ and $\sharp\Gamma^{\mathcal{M}_k}$ is a **henk**_{PA}-leftmost formula path in \mathcal{M}_k .

To see that there is at most such a formula path take formula paths Γ, Δ such that $\text{Len}(\Gamma^{(k)}) = m = \text{Len}(\Delta)$ and $\sharp\Gamma^{\mathcal{M}_k}, \sharp\Delta^{\mathcal{M}_k}$ are **henk**_{PA}-leftmost formula paths in \mathcal{M}_k . Then, by (2) and the definition of “**henk**_{PA}-leftmost path in \mathcal{M} ” $\Gamma \prec_{\neg} \Delta$ and $\Delta \prec_{\neg} \Gamma$ are impossible, hence $\Gamma = \Delta$ holds by (1). To get a formula path with the desired properties take an $m' \in \mathbb{N}$ with $m \leq m'$ such that $\text{Frm}(m')$ is a sentence true in \mathcal{M}_{k+1} . Then, by (M1) and (D), $\mathcal{M}_k \models \text{true}_{\text{PA}}(\text{Frm}(m'))$ holds, hence there is a **henk**_{PA}-leftmost path p in \mathcal{M}_k with $\text{len}^{\mathcal{M}_k}(p) = \overline{m}^{\mathcal{M}_k}$. Now consider $[\sharp_{\mathcal{M}_k} p(i)]_{i < m}$ and use (3).

According to (4) we can set for every $k \in \mathbb{N}$: $\Gamma^{(k)} =_{\text{def}}$ the formula Γ path of length n^* with $\sharp\Gamma^{\mathcal{M}_k}$ is **henk**_{PA}-leftmost path in \mathcal{M}_k .

For every $k \in \mathbb{N}$ we have:

$$(5) \quad \Gamma^{(k)} \neq \Gamma^{(k+1)}$$

This can be seen in the following way. We have $n_{k+1} \leq n$ and can consider the n_{k+1} -th member of $\Gamma^{(k)}$ and $\Gamma^{(k+1)}$: On one hand, (3) and (4) imply together that p is a **henk**_{PA}-leftmost path in \mathcal{M}_{k+1} with $\text{last}^{\mathcal{M}_{k+1}}(p) = \text{Frm}(n_{k+1})^{\mathcal{M}_{k+1}}$ iff $p = \sharp[\Gamma^{(k+1)}(i)]_{i \leq n_{k+1}}^{\mathcal{M}_{k+1}}$; hence $\mathcal{M}_{k+1} \models \text{true}_{\text{PA}}(\text{Frm}(n_{k+1}))$ iff $\Gamma^{(k+1)}(n_{k+1}) = \text{Frm}(n_{k+1})$ follows by (L). On the other hand we get $\Gamma^{(k)}(n_{k+1}) = \text{Frm}(n_{k+1})$ iff $\mathcal{M}_k \models \text{true}_{\text{PA}}(\text{Frm}(n_{k+1}))$ (by an analogous consideration for k instead of $k+1$) iff $\mathcal{M}_{k+1} \models \text{Frm}(n_{k+1})$ (by (D) and (M1)). Thus (M2) — applied to $k+1$ — leads to $\Gamma^{(k)}(n_{k+1}) = \text{Frm}(n_{k+1}) \neq \Gamma^{(k+1)}(n_{k+1})$.

Moreover, for every $k \in \mathbb{N}$ we have

$$(6) \quad \Gamma^{(k+1)} \not\prec_{\neg} \Gamma^{(k)}.$$

$\Gamma^{(k+1)} \prec_{\neg} \Gamma^{(k)}$ would lead, by (2), to $\sharp\Gamma^{(k+1)\mathcal{M}_k} \prec_{\neg}^{\mathcal{M}_k} \sharp\Gamma^{(k)\mathcal{M}_k}$; hence, because $\sharp\Gamma^{(k)\mathcal{M}_k}$ is a **henk**_{PA}-leftmost path in \mathcal{M}_k , $\text{con}_{\text{henk}_{\text{PA}}^{\mathcal{M}_k}}(\sharp\Gamma^{(k+1)\mathcal{M}_k})$

must be false in \mathcal{M}_k . Thus $\neg \mathbf{con}_{\mathbf{henk}_{\text{PA}}^{\cup}}(\#\Gamma^{(k+1)})$ would be true in \mathcal{M}_k , but this is a Σ_1^{PA} -sentence (because PA is recursive enumerable); hence corollary 2.3 would lead to $\mathcal{M}_{k+1} \models \neg \mathbf{con}_{\mathbf{henk}_{\text{PA}}^{\cup}}(\#\Gamma^{(k+1)})$, what cannot happen because $\#\Gamma^{(k+1)} \mathcal{M}_{k+1}$ is a leftmost $\mathbf{henk}_{\text{PA}}$ -path in \mathcal{M}_{k+1} .

Now we have got the desired contradiction: $\{\Gamma^{(k)} \mid k \in \mathbb{N}\}$ has to be a infinite set of formula paths because of (1), (5) and (6) we have $\Gamma^{(k)} \neq \Gamma^{(l)}$ for all $k, l \in \mathbb{N}$ with $k \neq l$. But this is impossible because $\text{Len}(\Gamma^{(k)}) = n^* + 1$ for all $k \in \mathbb{N}$, and only 2^{n^*+1} formula paths of length $n^* + 1$ exist. \square

3. A proof of the Second G.I.T. following Kreisel

By using the Model Chain Lemma in connection with the Arithmetized Completeness theorem indirect proofs of the Second G.I.T. are easy to get. Only one additional ingredient is necessary — a method to fulfil condition (M2) of the Model Chain Lemma. To this end one can start, for example, with a undecidable sentence from a proof of the First G.I.T. and reconstruct this sentence with $\mathbf{true}_{\text{PA}}$ instead of $\mathbf{prov}_{\text{PA}}$. The mostly direct way to do this starts with the undecidable sentence of the classical proof for the First G.I.T. and, therefore, uses the well-known Diagonalization Lemma. This strategy amounts to the proof of the Second G.I.T. given by Kreisel.

Lemma 3.1. (Diagonalization Lemma) *For every formula ϕ containing exactly v_0 free there is a sentence γ such that $\mathcal{M} \models \gamma \leftrightarrow \phi(\underline{\theta})$ for every model \mathcal{M} of PA.*

Proof. Because the literature is full of proofs of the Diagonalization lemma I will omit the proof here.⁹ \square

Now a reformulation of Kreisel's proof of the Second G.I.T. can be given:

Theorem 3.2. (Second G.I.T.) *If PA is consistent then $\not\vdash_{\text{PA}} \mathbf{con}_{\text{th}_{\text{PA}}}$.*

Proof. Assume $\vdash_{\text{PA}} \mathbf{con}_{\text{th}_{\text{PA}}}$. Because PA is supposed to be consistent there exists a model \mathcal{M}^* of PA (by the 'ordinary' Completeness Theorem). We can recursively define a sequence $[\mathcal{M}_k]_{k \in \mathbb{N}}$ of models of PA by setting $\mathcal{M}_0 =_{\text{def}} \mathcal{M}^*$ and $\mathcal{M}_{k+1} =_{\text{def}}$ the \mathcal{L}_{PA} -structure \mathcal{M} strongly defined by $\mathbf{univ}_{\text{PA}}$,

⁹ For instance, [1], p. 176., [4], p. 158, and [5], p. 37f. contain proofs of the Diagonalization Lemma.

$[\mathbf{int}_{\text{PA}_i}]_{i < 5}$, $\mathbf{true}_{\text{PA}}$ in \mathcal{M}_k by using the Arithmetized Completeness Theorem because $\mathcal{M} \models \mathbf{con}_{\text{th}_{\text{PA}}}$ holds for every model \mathcal{M} of PA by assumption.

Now use the Diagonalization Lemma to find a sentence γ such that $\mathcal{M}_k \models \gamma \leftrightarrow \neg \mathbf{true}_{\text{PA}}(\underline{\gamma})$ holds for every $k \in \mathbb{N}$, and set $n^* = \text{gn}(\gamma)$.

Thus we have found a sequence $[\mathcal{M}_k]_{k \in \mathbb{N}}$ of models of PA and an $n^* \in \mathbb{N}$ fulfilling both conditions of the Model Chain Lemma: (M1) directly follows from the definition of $[\mathcal{M}_k]_{k \in \mathbb{N}}$, (M2) is clear by the choice of n^* and by setting $n_k =_{\text{def}} n^*$ for all $k \in \mathbb{N}$. But this is, according to this very Lemma, impossible. \square

References

- [1] Adamowicz, Z., and P. Zbierski, *Logic of Mathematics. A Modern Course of Classical Logic*, John Wiley & Sons, New York 1997.
- [2] Boolos, G., *The Logic of Provability*, Cambridge University Press, Cambridge, 1993.
- [3] Feferman, S., “Arithmetization of metamathematics in a general setting”, *Fundamenta Mathematicae* XLIX (1960), p. 35–92.
- [4] Hajék, P., and P. Pudlák, *Metamathematics of First-Order Arithmetic*, Springer Verlag, Berlin 1993.
- [5] Kaye, R., *Models of Peano Arithmetic*, Clarendon Press, Oxford, 1991.
- [6] Kikuchi, M., and K. Tanaka, “On formalization of model-theoretic proofs of Gödel’s theorems”, *Notre Dame Journal of Formal Logic* 35 (1994), p. 403–412.
- [7] Smoryński, C., “The incompleteness theorems”, p. 821–865 in *Handbook of Mathematical Logic*, edited by J. Barwise, North Holland, Amsterdam, 1977.
- [8] Ullrich, D., “A purely model-theoretical proof of the Second Incompleteness Theorem of Gödel using Berry’s Paradox”, submitted to *Acta Mathematica Universitatis Comenianae*.

DIRK ULLRICH
Fachbereich 3
Mathematik TU Berlin
Straße des 17. Juni 136, Sekr. MA 1-1
D-10623 Berlin, Germany
e-mail: ullrich@fb3-s7.math.tu-berlin.de