Introduction to Ordinal Analysis

Hilbert–Bernays Summer School on Logic and Computation

July 2015 Göttingen

Wolfram Pohlers
WWU Münster

Contents

Part I
Handouts

1 Motivation

One of the aims of mathematical activities is the study of abstract structures. The perhaps most important structure in mathematics is the "standard model of arithmetic" which we are going to denote by $\mathbb{N}$ in this course.

The only tools that mathematicians have at their disposal to perform this task are axiom systems and (mathematical) logic. By axiom systems they try to characterize structures as accurately as possible and then use logic to derive their theorems from the axioms.

However, it follows from Gödel's incompleteness theorems that no axiom system can be rich enough to logically entail all the theorems of a (sufficiently complex) structure, e.g., of the complexity of $\mathbb{N}$.

A possibility to gauge the performance of an axiom system is ordinal analysis. Here we try to calibrate its performance in terms of infinite magnitudes, i.e., by infinite ordinals.

The origin of this type of *ordinally informative proof theory* goes back to Gerhard Gentzen who—despite of Gödels incompleteness theorem—gave a consistency proof for an axiom system for arithmetic. This proof only used finitistic means except for an application of a transfinite induction along a well–ordering of order–type ε_0 ([1]. By Gödel's incompleteness theorem it therefore follows that transfinite induction up to ε_0 cannot be provable from the axioms of arithmetic. In a later paper [2] he then showed that any ordinal less than ε_0 can be represented by a well–ordering whose well–foundedness is provable in arithmetic. This was the birth of ordinally informative proof theory. Since then we define the proof theoretic ordinal of an axiom system T as the supremum of the order–types of well–orderings which are elementarily definable in the language of T and whose well–foundedness is provable in T.

The aim of the course is an ordinal analysis of an axiom system for $\mathbb{N}$ which is equivalent to the Peano axioms.

2 Ordinals

Our main tool in gauging the range of axiom systems are ordinals. Intuitively an ordinal is the order–type of a well–ordering.

2.1 Ordinals as equivalence classes

Given two finite set M_1 and M_2 there are two methods to compare their size. We can bring the elements of both sets into one–one correspondence and check on which side there are elements left or we simple count the members and compare the numbers. Mathematically speaking the first methods yields the definition

$$\overline{\overline{M_1}} \leq \overline{\overline{M_2}} \quad :\Leftrightarrow \quad \text{there is a } f\colon M_1 \xrightarrow{1-1} M_2$$

and

$$\overline{\overline{M_1}} = \overline{\overline{M_2}} \quad :\Leftrightarrow \quad \overline{\overline{M_1}} \leq \overline{\overline{M_2}} \wedge \overline{\overline{M_2}} \leq \overline{\overline{M_1}} \quad \Leftrightarrow \quad \text{there is a } f\colon M_1 \longleftrightarrow M_2$$

and we call the equivalence class $\{N \mid \overline{\overline{M}} = \overline{\overline{N}}\}$ the *cardinality* of the set M.

Counting the elements of a set M means to order the elements of M. Orders which are suited for counting are order–relations with the property that every non–void subset of the field of the ordering possesses a least element (the candidate for the next element to be counted).

2.1 Definition A binary relation $\prec$ is a well–ordering if it is total, transitive, irreflexive and satisfies $(\forall X)[X \subseteq \mathit{field}(\prec) \wedge X \neq \emptyset \Rightarrow (\exists x \in X)[(\forall y)[y \prec x \rightarrow y \notin X]]]$.

For well–orderings $\prec_1$ and $\prec_2$ we define

$$\prec_1 \leq \prec_2 \;:\Leftrightarrow\; (\exists f)[f\colon \mathit{field}(\prec_1) \longrightarrow \mathit{field}(\prec_2) \text{ order preserving}]$$

and

$$\prec_1 = \prec_2 \;:\Leftrightarrow\; \prec_1 \leq \prec_2 \;\wedge\; \prec_2 \leq \prec_1 .$$

The order–type of a well–ordering $\prec$ is the equivalence class

$$\mathrm{otyp}(\prec) := \{\prec^* \mid \prec = \prec^*\}.$$

2.2 Theorem *Let $\prec$ a well–ordering. Then*

$$(\forall X)\big[(\forall x)[(\forall y)(y \prec x \rightarrow y \in X) \rightarrow x \in X] \rightarrow (\forall x \in \mathit{field}(\prec))(x \in X)\big].$$

Let $WO(\prec)$ abbreviate the above sentence. Then we obtain that $\prec$ is a well–ordering iff $WO(\prec)$.

2.2 Set theoretical ordinals

Since the equivalence classes "cardinality" and "order–type" are proper classes and thus no sets in a set theoretical sense is has become common to represent order–types by set theoretical ordinals.

$0 = \emptyset,\ 1 = \{0\}, 2 = \{\emptyset, \{\emptyset\}\} = \{0, 1\}, \quad \ldots, \quad n + 1 = \{0, \ldots, n\},$
$\omega = \{0, 1, 2, \ldots\} \quad \omega + \omega = \{0, \ldots, \omega, \omega + 1, \ldots\} \quad \ldots$

Figure 1: Some set theoretical ordinals

As a reminder: A set a is transitive if it possesses no $\in$ holes, i.e., if

$$(\forall x \in a)(\forall y \in x)[y \in a].$$

2.3 Definition An ordinal is a transitive set that is well–ordered by the $\in$–relation. Let On be the class of all set theoretical ordinals. We define

$$\alpha < \beta \quad :\Leftrightarrow \quad \alpha \in On \wedge \beta \in On \wedge \alpha \in \beta.$$

Clearly

$$\beta \in On \wedge a \subseteq \beta \wedge Tran(a) \text{ entail } a \in On. \tag{1}$$

Observe that by this definition a set theoretical ordinal coincides with the set of its predecessors, i.e.,

$$\alpha = \{\xi \,|\, \xi < \alpha\}.$$

2.4 Theorem *(Transfinite Induction) If $(\forall \xi < \eta)F(\xi)$ entails $F(\eta)$ for any ordinal η then we already have $(\forall \zeta \in On)F(\zeta)$.*

2.5 Lemma *$\beta \in On$, $Tran(a)$ and $a \subsetneq \beta$ imply $a \in \beta$.*

2.6 Lemma *In presence of the foundation scheme an ordinal is a hereditarily transitive set.*

2.7 Definition An ordinal κ is a *cardinal* iff $(\forall f)(\forall \xi)[f{:}\, \kappa \longleftrightarrow \xi \Rightarrow \kappa \leq \xi]$ where we generally agree that lower case Greek letters are supposed to vary over ordinals.

2.8 Definition An ordinal λ is a limit ordinal iff $(\forall \xi < \lambda)(\exists \eta < \lambda)[\xi < \eta]$.

2.9 Definition For a set M of ordinals let $\sup M = \min\{\xi \,|\, (\forall \eta \in M)[\eta \leq \xi]\} = \bigcup M$.

Recall that $\bigcup M := \{x \,|\, (\exists y \in M)[x \in y]\}$.

2.10 Lemma *If α is an ordinal then $\alpha \cup \{\alpha\}$ is again an ordinal satisfying $(\forall \alpha < \xi)[\alpha \cup \{\alpha\} \leq \xi]$. We call $\alpha \cup \{\alpha\}$ the* successor *of α often denoted by α'.*

2.11 Lemma *An ordinal $\lambda \neq 0$ is a limit ordinal iff $\sup \lambda = \lambda$. If α is not a limit ordinal then $(\sup \alpha)' = \alpha$. Let Lim denote the class of limit ordinals.*

There are three types of ordinals: 0, successor ordinals α' and limit ordinals.

2.12 Definition Let $\omega := \min \mathit{Lim}$. Then $\omega \in \mathit{Lim} \wedge (\forall \eta < \omega)[\eta \notin \mathit{Lim}]$. An ordinal ξ is finite iff $\xi < \omega$.

2.13 Theorem *Every finite ordinal and ω are cardinals.*

2.14 Definition Let $\prec$ be a well–ordering and $x \in \mathit{field}(\prec)$. Then we define

$$\operatorname{otyp}_{\prec}(x) := \sup \{(\operatorname{otyp}_{\prec}(y))' \,|\, y \prec x\}$$

and

$$\operatorname{otyp}(\prec) := \sup \{\operatorname{otyp}_{\prec}(x) \,|\, x \in \mathit{field}(\prec)\}.$$

This definition coincides with the first—informal—definition in so far that $\operatorname{otyp}(\prec)$ is a representative of the equivalence class $\operatorname{otyp}(\prec)$.

2.15 Definition Let M be a class of ordinals. Then $\operatorname{otyp}(M) := \operatorname{otyp}(<\restriction M)$. The inverse function $\mathit{en}_M \colon \operatorname{otyp}(M) \longrightarrow M$ satisfying $\mathit{en}_M(\operatorname{otyp}_{<}(x)) = x$ is the *enumerating function* of M.

Observe that ω is the order–type of the natural numbers in their canonical ordering.

2.3 Basics of ordinal arithmetic

2.16 Definition (Ordinal addition) Let
$\alpha + 0 := \alpha$
$\alpha + \beta' := (\alpha + \beta)'$
and
$\alpha + \lambda := \sup \{\alpha + \xi \,|\, \xi < \lambda\}$.

2.17 Definition (Ordinal multiplication) Let
$\alpha \cdot 0 := 0$
$\alpha \cdot \beta' := (\alpha \cdot \beta) + \alpha$
$\alpha \cdot \lambda := \sup \{\alpha \cdot \xi \,|\, \xi < \lambda\}$ for limit ordinals λ.

2.18 Definition (Ordinal exponentiation) Let
$\alpha^0 := \{0\}\ \ (= 1)$
$\alpha^{\beta'} := (\alpha^{\beta}) \cdot \alpha$
$\alpha^{\lambda} := \sup \{\alpha^{\xi} \,|\, \xi < \lambda\}$.

2.19 Definition An ordinal α is additively indecomposable if $\xi, \eta < \alpha$ entail $\xi + \eta < \alpha$.

2.20 Lemma *The function $\lambda\xi.(\alpha + \xi)$ is the enumerating function of the class $M = \{\xi \,|\, \alpha \le \xi\}$. Hence $\xi < \eta$ iff $\alpha + \xi < \alpha + \eta$.*

2.21 Lemma *The function $\lambda\xi.(\omega^{\xi})$ is the enumerating function of the class of additively indecomposable ordinals. Hence $\xi < \eta$ iff $\omega^{\xi} < \omega^{\eta}$.*

2.22 Definition Let $\varepsilon_0 := \min \{\xi \,|\, \omega^{\xi} = \xi\}$.

2.23 Lemma *Put $\omega^{(0)}(\alpha) := \alpha$ and $\omega^{(n+1)}(\alpha) := \omega^{\omega^{(n)}(\alpha)}$. Then $\varepsilon_0 = \sup \{\omega^{(n)}(0) \,|\, n \in \omega\}$. For any ordinal $\xi < \varepsilon_0$ we get $\varepsilon_0 = \sup \{\omega^{(n)}(\xi) \,|\, n \in \omega\}$.*

2.4 Exercises

2.24 Exercise Prove Lemmata 2.20 , 2.21 and 2.23.

2.25 Exercise Show that $\xi < \alpha$ implies $\xi + \alpha = \alpha$ for additively indecomposable ordinals α.

2.26 Exercise (Cantor normal form) Show that for every ordinal α there are additively indecomposable ordinals $\alpha_1, \ldots, \alpha_n$ such that $\alpha = \alpha_1 + \cdots + \alpha_n$ and $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$.

3 The standard model of arithmetic

In general a structure has the form $\mathfrak{M} = (M, \mathcal{C}, \mathcal{R}, \mathcal{F})$ where

- M is a non–void set, the *domain* of the the structure
- $\mathcal{C}$ is a subset of M, the *constants* of $\mathfrak{M}$.
- Every $R \in \mathcal{R}$ is a subset $M^{\#R}$ where $0 \leq \#R < \omega$ is the arity of the relation R
- Every $f \in \mathcal{F}$ is a function $f \colon M^{\#f} \longrightarrow M$ where $0 < \#f < \omega$ is the arity of f.

3.1 Primitive recursive functions

Let *PRF* be the smallest class of arithmetical functions which

- contains the successor function S.
- contains all n–ary constant functions $C^n_k(z_1, \ldots, z_n) = k$.
- contains all n–ary projection functions $P^n_k(x_1, \ldots, x_n) = x_k$.
- is closed under substitutions, defined by

$$Sub(g, h_1, \ldots, h_m)(x_1, \ldots, x_n) = g(h_1(x_1, \ldots, x_n)) \ldots (h_m(x_1, \ldots, x_n)).$$

- is closed under primitive recursion, defined by

$$\begin{aligned} &Rec(g, h)(0, x_1, \ldots, x_n) = g(x_1, \ldots, x_n) \\ &Rec(g, h)(Sy, x_1, \ldots, x_n) = h(y, Rec(g, h)(y, x_1, \ldots, x_n), x_1, \ldots, x_n) \end{aligned}$$

An n–ary relation $R \subseteq \omega^n$ is primitive recursive iff its characteristic function χ_R, defined by

$$\chi_R(y_1, \ldots, y_n) = \begin{cases} 1 & \text{if } (y_1, \ldots, y_n) \in R \\ 0 & \text{otherwise,} \end{cases}$$

is primitive recursive. Let *PRP* denote the set of primitive recursive relations. Cf. Figure 2 for a list of primitive–recursive functions and relations.

Function/Relation	Name	Definition
$sg(n)$	signum of n (case distinction)	$sg(0) = 0,\ \ sg(S(x)) = 1$
$\overline{sg}(n)$	antisignum of n (case distinction)	$\overline{sg}(0) = 1,\ \ \overline{sg}(S(x)) = 0$
$a + n$	addition	$a + 0 = a,\ \ a + S(n) = S(a + n)$
$a \cdot n$	multiplication	$a \cdot 0 = 0,\ \ a \cdot (Sn) = (a + n) + a$
$a!$	a faculty	$0! = 1,\ \ (Sa)! = a! \cdot a$
a^n	exponentiation	$a^0 = 1,\ \ a^{S(n)} = a^n \cdot a$
$Pd(n)$	predecessor	$Pd(0) = 0,\ \ Pd(S(x)) = x$
$a \dot{-} x$	arithmetical difference	$a \dot{-} 0 = 0,\ \ a \dot{-} S(x) = Pd(a \dot{-} x)$
$\lvert a - x \rvert$	absolute value	$(a \dot{-} x) + (x \dot{-} a)$
$\mathcal{I} := \{(x, x) \mid x \in \omega\}$	identity	$\chi_{\mathcal{I}} = \overline{sg}(\lvert x - y \rvert)$
$x < y, x \leq y$	less than (or equal to)	$(\exists z < y)[x + z = y], x < y \lor x = y$
$\max\{a, b\}$	maximum	$\max\{a, b\} = \begin{cases} a & \text{if } b \leq a \\ b & \text{otherwise} \end{cases}$
x/y	x divides y	$(\exists z < S(y))[x \cdot z = y]$
$Prime(x)$	x is a prime number	$(\forall z < S(x))[z = 1 \lor z = x \lor \neg(z/x)]$
$pn(n)$	Enumeration of the primes	$pn(n) = \begin{cases} 2 & \text{if } n = 0 \\ \mu x < pn(n)! + 2.\ [Prime(x) \land pn(n) < x] & \text{if } n > 0 \end{cases}$
$\langle x_0, \ldots, x_n \rangle$	Coded tuple	$= \begin{cases} 0 & \text{for } n = -1 \text{ (empty sequence)} \\ \prod_{i=0}^{n}[pn(i)^{S(x_i)}] & \text{for } n \geq 0 \end{cases}$
$lh(x)$	length of the tuple coded by x	$lh(\langle x_0, \ldots, x_n \rangle) = n + 1$
$(a)_i$	decoding function	$(\langle x_0, \ldots, x_n \rangle)_i = x_i$ for $0 \leq i \leq n$
$Seq(s)$	s codes a sequence	$s = 0 \lor (\forall i < S(s))[\neg(pn(S(i))/s) \lor pn(i)/s]$
a belongs to a finite set	$a \in \{a_1, \ldots, a_n\}$	$a = a_1 \lor \cdots \lor a = a_n$

Figure 2: Some primitive–recursive functions and relations

3.2 The standard structure $\mathbb{N}$.

3.1 Definition We call the structure $\mathbb{N} = (\omega, 0, \mathit{PRP}, \mathit{PRF})$ the *standard structure* of arithmetic.

3.2 Definition The first order language $\mathcal{L}_{\mathbb{N}}$ of $\mathbb{N}$ is inductively defined by:

- 0 and every free individual variable x is an $\mathcal{L}_{\mathbb{N}}$–term.
- If $t_1, \ldots, t_n$ are $\mathcal{L}_{\mathbb{N}}$–terms and f is a name for an n–ary function in *PRF* then $(ft_1, \ldots, t_n)$ is an $\mathcal{L}_{\mathbb{N}}$–term.
- If t is an $\mathcal{L}_{\mathbb{N}}$–term and X a set variable then $t \in X$ and $t \notin X$ are atomic $\mathcal{L}_{\mathbb{N}}$–formulas.
- If $t_1, \ldots, t_n$ are $\mathcal{L}_{\mathbb{N}}$–terms and R is a symbol for an n–ary relation in *PRP* then $(Rt_1, \ldots, t_n)$ is an atomic $\mathcal{L}_{\mathbb{N}}$–formula.
- Every atomic $\mathcal{L}_{\mathbb{N}}$–formula is an $\mathcal{L}_{\mathbb{N}}$–formula.
- If F and G are $\mathcal{L}_{\mathbb{N}}$–formulas then $F \wedge G$ and $F \vee G$ are $\mathcal{L}_{\mathbb{N}}$–formula.
- If F is an $\mathcal{L}_{\mathbb{N}}$–formula and x a free individual variable then $(\forall x)F(x)$ and $(\exists x)F(x)$ are $\mathcal{L}_{\mathbb{N}}$–formula in which x is not longer free but bound.

A formula which does not contain free variables is a *sentence*.

We use the language $\mathcal{L}_N$ in *Tait style*. I.e. there is no negation symbol among the logical symbols. We can however define negation using deMorgan's rules.

- For an atomic $(Rt_1, \ldots, t_n)$ we define $\neg(Rt_1, \ldots, t_n) \equiv (\overline{R}t_1, \ldots, t_n)$ where $\overline{R}$ is a symbol for the complement of R.
- For the atomic formulas $t \in X$ and $t \notin X$ we define $\neg(t \in X) :\equiv t \notin X$ and $\neg(t \notin X) :\equiv t \in X$.
- $\neg(F \wedge G) :\equiv \neg F \vee \neg G$ and $\neg(F \vee G) :\equiv \neg F \wedge \neg G$.
- $\neg(\forall x)F(x) :\equiv (\exists x)\neg F(x)$ and $\neg(\exists x)F(x) :\equiv (\forall x)\neg F(x)$.

Observe that although we do not have a constant for every natural number there is a name for every $n \in \omega$. Here we do not mean the set theoretical name but the name in the structure $\mathbb{N}$ which is obtained from 0 by successively applying the successor function S. We call these names *numerals*.

3.2.1 Exercises

3.3 Exercise Show that the primitive–recursive functions are closed under definition by primitive recursive case distinctions.

3.4 Exercise Show that the primitive–recursive relations are closed under all Boolean operations, bounded quantification and substitutions with primitive–recursive functions.

3.5 Exercise Show that for any primitive–recursive function f the bounded search function $\mu \le k.\,(f(x, y_1, \ldots, y_n) = 0)$, defined by

$$\mu x \le k.\,(f(x, y_1, \ldots, y_n) = 0) = \begin{cases} \min\{z \le k \mid f(z, y_1, \ldots, y_n) = 0\} & \text{if this exists} \\ k & \text{otherwise,} \end{cases}$$

is primitive–recursive.

3.6 Exercise Verify the open facts in Figure 2

3.3 The verification calculus

3.7 Definition We divide the sentences of $\mathcal{L}_{\mathbb{N}}$ into two types:
The $\bigwedge$–type comprises:

- All true atomic sentences,
- The sentences $F \wedge G$ and $(\forall x)G(x)$.

The $\bigvee$–type comprises:

- All false atomic sentences,
- The sentences $F \vee G$ and $(\exists x)F(x)$.

3.8 Definition The characteristic sequence $\mathrm{CS}(F)$ of a sentence F is defined by

- $\mathrm{CS}(F) = \emptyset$ if F is an atomic sentence.
- $\mathrm{CS}(F \circ G) = \langle F, G\rangle$ for $\circ \in \{\wedge, \vee\}$.
- $\mathrm{CS}((\mathsf{Q}x)F(x)) = \big\langle F(n) \,\big|\, n \in \omega\big\rangle$ for $\mathsf{Q} \in \{\forall, \exists\}$.

3.9 Lemma *For every sentence $F \in \mathcal{L}_{\mathbb{N}} \cap \bigwedge$–type we have:*

$\quad\mathbb{N} \models F$ *iff* $\mathbb{N} \models G$ *for all* $G \in \mathrm{CS}(F)$.

For every sentence $F \in \mathcal{L}_{\mathbb{N}} \cap \bigvee$–type we have:

$\quad\mathbb{N} \models F$ *iff* $\mathbb{N} \models G$ *for some* $G \in \mathrm{CS}(F)$.

3.10 Definition (The verification calculus) Let Δ be a finite set of $\mathcal{L}_{\mathbb{N}}$–sentences. We define the verification calculus by two rules:

($\bigwedge$) If $F \in \Delta \cap \bigwedge$–type then $\models^{\alpha_G} \Delta, G$ and $\alpha_G < \alpha$ for all $G \in \mathrm{CS}(F)$ implies $\models^{\alpha} \Delta$.

($\bigvee$) If $F \in \Delta \cap \bigvee$–type then $\models^{\alpha_0} \Delta, G$ and $\alpha_0 < \alpha$ for some $G \in \mathrm{CS}(F)$ implies $\models^{\alpha} \Delta$.

3.11 Lemma *For* $\mathrm{rnk}(F) := \sup\{\mathrm{rnk}(G) + 1 \mid G \in \mathrm{CS}(F)\}$ *we obtain*

$\quad\mathbb{N} \models F$ *iff* $\models^{\mathrm{rnk}(F)} F$.

3.4 Π^1_1–completeness

3.12 Definition A *pseudo* Π^1_1*–sentence* is a $\mathcal{L}_{\mathbb{N}}$–formula which must not contain free individual variables but may contain free set variables.

We extend the definitions of $\bigwedge$–type and $\bigvee$–type to pseudo Π^1_1–sentences. Sentences of the form $t \in X$ and $t \notin X$ belong to no type.

3.13 Observation *Let $F(X)$ be a pseudo Π^1_1–sentence. Then*

$$\mathbb{N} \models (\forall X)F(X) \text{ iff } \mathbb{N} \models F(X)[S] \text{ for any set } S \subseteq \omega.$$

3.14 Definition We extend the verification calculus $\models^{\alpha} \Delta$ to a semi–formal system $\vdash^{\alpha}_{\rho} \Delta$ for finite sets Δ of pseudo Π^1_1–sentences. We replace $\models^{\alpha} \Delta$ by $\vdash^{\alpha}_{\rho} \Delta$ in $(\bigwedge)$– and $(\bigvee)$–rules and add the new rules

(X–rule) If $s^{\mathbb{N}} = t^{\mathbb{N}}$ and $\{s \in X, t \notin X\} \subseteq \Delta$ then $\vdash^{\alpha}_{\rho} \Delta$ holds true for all ordinals α

and

(cut) From $\vdash^{\alpha_0}_{\rho} \Delta, F$, $\vdash^{\alpha_0}_{\rho} \Delta, \neg F$ and $\mathrm{rnk}(F) < \rho$ infer $\vdash^{\alpha}_{\rho} \Delta$ for all $\alpha > \alpha_0$.

3.15 Lemma *We have* $\vdash^{2\cdot\mathrm{rnk}(F)}_{0} \Delta, F, \neg F$ *for any pseudo Π^1_1–sentence F.*

3.16 Theorem *(Π^1_1–completeness) For a Π^1_1–sentence we have $\mathbb{N} \models (\forall X)F(X)$ iff $\vdash^{\alpha}_{0} F(X)$ for some ordinal $\alpha < \omega_1^{CK}$ where ω_1^{CK} denotes the least ordinal which is not the order–type of a primitive–recursively definable well–ordering on $\mathbb{N}$.*

3.17 Definition We define the truth complexity of (pseudo) Π^1_1–sentences by

$$\mathrm{tc}\big((\forall X)F(X)\big) := \begin{cases} \min\{\alpha \,|\, \vdash^{\alpha}_{0} F(X)\} & \text{if this exists} \\ \omega_1^{CK} & \text{otherwise.} \end{cases}$$

3.18 Theorem *(Boundedness Theorem) An $\mathcal{L}_{\mathbb{N}}$–definable ordering is a well–ordering iff* $\mathrm{otyp}(\prec) \leq 2^{\mathrm{tc}\left(WO(\prec)\right)} < \omega_1^{CK}$.

3.19 Definition For a structure $\mathfrak{M}$ we define its Π^1_1*–ordinal*

$$\pi^{\mathfrak{M}} := \sup\{\mathrm{tc}(F) \,|\, \mathfrak{M} \models F\}.$$

3.20 Definition Define $WO(X, \prec)$ as the pseudo Π^1_1–sentence such that $WO(\prec) \equiv (\forall X)WO(X, \prec)$. Let T be an axiom system for $\mathbb{N}$. Then we define its proof–theoretic ordinal

$$|T| := \sup\{\mathrm{otyp}(\prec) \,|\, T \vdash WO(X, \prec)\}$$

where $\prec$ is any elementarily definable ordering, and its Π^1_1–ordinals by

$$\pi^T := \sup\{\mathrm{tc}(F) \,|\, T \vdash F\}.$$

3.21 Remark We have $\pi^{\mathbb{N}} = \omega_1^{CK}$.

3.22 Corollary *Let T be an axiom system for $\mathbb{N}$. Then $|T| \leq 2^{\pi^T}$.*

3.23 Remark Actually we have $|T| = \pi^T$ in general. We will prove this for the axiom system NT.

Also the Boundedness Theorem (Theorem 3.18) can be sharpened to $\mathrm{otyp}(\prec) = \mathrm{tc}\big(\mathit{WO}(\prec)\big)$. The proof is more involved and the sharper bound is not needed in the ordinal analysis of NT.

3.5 Exercises

3.24 Exercise Show:

(Str) $\quad \models^{\alpha}_{\rho} \Delta\,, \alpha \leq \beta, \rho \leq \sigma, \Delta \subseteq \Gamma \;\;\Rightarrow\;\; \models^{\beta}_{\sigma} \Gamma\,.$

($\bigwedge$–Inv) $\models^{\alpha}_{\rho} \Delta, F$ and $F \in \bigwedge$–type $\;\;\Rightarrow\;\; \models^{\alpha}_{\rho} \Delta, G$ for every $G \in \mathrm{CS}(F)$.

($\vee$–Exp) $\models^{\alpha}_{\rho} \Delta, F\,, F \in \bigvee$–type and $\mathrm{CS}(F)$ is finite $\;\;\Rightarrow\;\; \models^{\alpha}_{\rho} \Delta, \Gamma$ for $\Gamma = \mathrm{CS}(F)$.

3.25 Exercise a) Show that $s^{\mathbb{N}} = t^{\mathbb{N}}$ and $\models^{\alpha} F(s)$ imply $\models^{\alpha} F(t)$.
b) Show that $s^{\mathbb{N}} = t^{\mathbb{N}}$ and $\models^{\alpha}_{\rho} \Delta(s)$ imply $\models^{\alpha}_{\rho} \Delta(t)$.
c) Show that there is an α such that $\models^{\alpha}_{0} s \neq t, \neg F(s), F(t)$. How large is α?.

3.26 Exercise Let $\prec$ be a primitive–recursively definable well–ordering with *field*$(\prec) = \mathbb{N}$. Denote by $\mathrm{otyp}_{\prec}(n)$ the order–type of n in $\prec$ and let

$$\mathit{Prog}(X, \prec) \;\;:\Leftrightarrow\;\; (\forall x)[(\forall y)[y \prec x \to y \in X] \to x \in X].$$

Show

$$\models^{4\cdot\mathrm{otyp}_{\prec}(n)}_{0} \neg\mathit{Prog}(X, \prec), n \in X\,.$$

Conclude that $\pi^{\mathbb{N}} = \omega_1^{CK}$.

4 The axiom system NT

4.1 Peano arithmetic

The only non–logical symbols of Peano arithmetic are the constants for 0, 1, and the function symbols $+$ for addition and $\cdot$ for multiplication. The non–logical axioms of Peano arithmetic comprise the defining equations for $+$ and $\cdot$ together with the successor axioms $(\forall x)[x + 1 \neq 0]$ and $(\forall x)(\forall y)[x + 1 = y + 1 \to x = y]$ and the scheme of mathematical induction.

We will, however, give the ordinal analysis for an axiom system NT which comprises symbols for all primitive–recursive functions and –predicates and thus is more expressive than Peano arithmetic. It can, however, be shown that NT is an extension by definitions of Peano arithmetic. This is not completely trivial and rests on the fact that Peano arithmetic proves the existence of a coding machinery. The key to such a machinery is Gödel's β–function whose definition bases on the Chinese remainder theorem.

4.2 Pure logic

To fix the logical framework we introduce a Hilbert style calculus for first order predicate logic. We presuppose familiarity with the language of first order predicate logic with identity where we allow free second order variables in the language. Since we aim at the language of arithmetic we restrict ourselves to unary predicate variables and talk about *set variables*.

4.1 Definition The Boolean atoms of first order formula F are the subformulas of F which are either atomic or the outmost logic symbol of which is a quantifier.

A Boolean valuation for a first order formula F is the assignment of a truth value to every Boolean atom of F.

The truth value of a first order formula under a Boolean valuation is computed according to the familiar rules for the Boolean connectives.

A first order formula is Boolean valid if it is true under any Boolean valuation.

4.2 Definition The logical axioms of the Hilbert calculus are:

(BOOLE) All Boolean valid formulas

($\forall$) All formulas $(\forall x)F(x) \to F(t)$ for any term t

($\exists$) All formulas $F(t) \to (\exists x)F(x)$ for any term t

The identity axioms are

(Ref) $(\forall x)[x = x]$

(Sym) $(\forall x)(\forall y)[x = y \to y = x]$

(Tran) $(\forall x)(\forall y)(\forall z)[x = y \wedge y = z \to x = z]$.

(Com) $(\forall x_1)\dots(\forall x_n)(\forall y_1)\dots(\forall y_n)[\bigwedge_{i=1}^n x_i = y_i \wedge F(x_1,\dots,x_n) \to F(y_1,\dots,y_n)]$

The inference rules are

(mp) $\vdash A$ and $\vdash A \to B \;\Rightarrow\; \vdash B$.

($\forall$) $\vdash A \to F(x) \;\Rightarrow\; \vdash A \to (\forall x)F(x)$ where the *eigenvariable* x must not occur in A.

($\exists$) $\vdash F(x) \to A \;\Rightarrow\; \vdash (\exists x)F(x) \to A$ where the *eigenvariable* x must not occur in A.

4.3 Theorem *A formula F is logically valid, i.e., true in any model under any assignment, iff $\vdash F$.*

4.4 Lemma *For any Boolean valid formula $F(x_1,\dots,x_n)$ there is a finite ordinal k such that $\vdash_0^k F(z_1,\dots,z_n)$ holds true for any tuple $z_1,\dots,z_n$ of numerals.*

4.5 Theorem *For any logically valid formula $F(x_1,\dots,x_n)$ whose free individual variables occur all in the list $x_1,\dots,x_n$ there are finite ordinals m and r such that $\frac{m}{r}\!\!\!\mid\; F(t_1,\dots,t_n)$ for any tuple $t_1,\dots,t_n$ of $\mathcal{L}_{\mathbb{N}}$–terms.*

4.3 The axioms of arithmetic

4.6 Definition The non–logical axioms of NT comprise

(MATHAX) All true atomic $\mathcal{L}_{\mathbb{N}}$–sentences

(MATHIND) The scheme $F(0) \wedge (\forall x)[F(x) \to F(S(x))] \;\; \to \;\; (\forall x)F(x)$ of *mathematical induction*, where $F(x)$ is any $\mathcal{L}_{\mathbb{N}}$–formula.

4.7 Theorem *A pseudo Π^1_1–sentence F is a* theorem *of* NT *iff it is a logical consequence of the axioms, i.e., iff there there are finitely many axioms $A_1,\dots,A_n$ of* NT *such that $A_1 \wedge \cdots \wedge A_n \to F$ is logically valid.*

5 The upper bound

5.1 Embedding of NT

5.1 Lemma *For every natural number n and $\mathcal{L}_{\mathbb{N}}$–formula $F(x)$ we have*

$$\Big|\!\frac{2(\mathrm{rnk}(F)+n)}{0}\; \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(n)\,.$$

5.2 Theorem *(Induction Theorem) We have*

$$\Big|\!\frac{\omega+3}{0}\; F(0) \wedge (\forall x)[F(x) \to F(S(x))] \;\; \to \;\; (\forall x)F(x)\,.$$

5.3 Theorem *(Embedding Theorem) Let $F(x_1,\dots,x_n)$ be a $\mathcal{L}_{\mathbb{N}}$–formula whose free individual variables occur all in the list $x_1,\dots,x_n$. Then* $\mathsf{NT} \vdash F(x_1,\dots,x_n)$ *implies that there is a finite ordinal r such that $\Big|\!\frac{\omega+\omega}{r}\; F(z_1,\dots,z_n)$ holds true for every tuple $z_1,\dots,z_n$ of numerals.*

5.2 Cut elimination

5.4 Lemma *(Reduction Lemma) From $\Big|\!\frac{\alpha}{\rho}\; \Delta, F$ and $\Big|\!\frac{\beta}{\rho}\; \Gamma, \neg F$ for $F \in \bigwedge$–*type *and* $\mathrm{rnk}(F) = \rho$ *we obtain $\Big|\!\frac{\alpha+\beta}{\rho}\; \Delta, \Gamma$.*

5.5 Theorem *(Elimination Theorem) $\Big|\!\frac{\alpha}{\rho+1}\; \Delta$ implies $\Big|\!\frac{\omega^\alpha}{\rho}\; \Delta$.*

5.3 The upper bound

5.6 Theorem *If* $\mathsf{NT} \vdash F$ *for a pseudo* Π^1_1*–sentence* F *then* $\mathrm{tc}(F) < \varepsilon_0$.

5.7 Corollary *We have* $\pi^{\mathsf{NT}} \leq \varepsilon_0$*, hence also* $|\mathsf{NT}| \leq \varepsilon_0$.

5.8 Corollary *The theory* NT *is consistent.*

6 The lower bound

6.1 Ordinal notations

6.1 Theorem *For every ordinal* α *less than* ε_0 *there are ordinals* $\alpha_1, \ldots, \alpha_n$ *such that* $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ *and* $\alpha > \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$.

6.2 Corollary *There is is notation system* $\ulcorner\alpha\urcorner$ *such that for every* $\alpha < \varepsilon_0$ *the numeral* $\ulcorner\alpha\urcorner$ *denotes the ordinal* α*. The set* $On = \{\ulcorner\alpha\urcorner \mid \alpha < \varepsilon_0\}$ *and the relation* $\ulcorner\alpha\urcorner \prec \ulcorner\beta\urcorner \;:\Leftrightarrow\; \alpha < \beta$ *are primitive–recursive.*

6.2 The well–ordering proof

In the sequel we identify ordinals α and their notations. We denote members of On by lower case Greek letters and write $\alpha < \beta$ instead of $\alpha \prec \beta$.

We use the following abbreviations:

$$\begin{aligned} \alpha \subseteq \beta \;&:\Leftrightarrow\; (\forall \xi)[\xi < \alpha \to \xi < \beta] \text{ (which unabbreviated is)} \\ &\Leftrightarrow\; \alpha \in On \wedge \beta \in On \wedge (\forall \xi \in On)[\xi \prec \alpha \to \xi \prec \beta] \end{aligned}$$

$$\begin{aligned} \mathit{Prog}(X) \;&:\Leftrightarrow\; \mathit{Prog}(X, \prec) \\ &\Leftrightarrow\; (\forall \xi)[\xi \subseteq X \to \xi \in X] \end{aligned}$$

$$\alpha \subseteq X \;:\Leftrightarrow\; (\forall \xi)[\xi < \alpha \to \xi \in X]$$

$$\alpha \in \mathcal{J}(X) \;:\Leftrightarrow\; (\forall \xi)[\xi \subseteq X \to \xi + \omega^\alpha \subseteq X].$$

$$\mathit{TI}(\alpha) \;:\Leftrightarrow\; \mathit{Prog}(X) \to \alpha \subseteq X.$$

The formula $\mathit{TI}(\alpha)$ then expresses transfinite induction up to α.

6.3 Lemma $\mathsf{NT} \vdash \mathit{Prog}(X) \to \mathit{Prog}(\mathcal{J}(X))$.

6.4 Lemma $\mathsf{NT} \vdash \mathit{TI}(\alpha)$ *entails* $\mathsf{NT} \vdash \mathit{TI}(\omega^\alpha)$.

6.5 Theorem *For any ordinal* $\alpha < \varepsilon_0$ *there is a primitive–recursively definable well–ordering* $\prec$ *of order–type* α *such that* $\mathsf{NT} \vdash \mathit{WO}(\prec)$.

6.6 Theorem $|\mathsf{NT}| = \pi^{\mathsf{NT}} = \varepsilon_0$.

6.7 Theorem *There is a pseudo* Π^1_1*–sentence* $(\forall x)F(x, X)$ *such that* $\mathbb{N} \models (\forall X)(\forall x)F(x, X)$*,* $\mathsf{NT} \vdash F(n, X)$ *for any numeral* n *but* $\mathsf{NT} \nvdash (\forall x)F(x, X)$.

Part II
Selected proofs and solutions

Section 2

Proof of Lemma 2.5.

Let $\xi := \min(\beta \setminus a)$. Then $\xi \subseteq a$ by minimality of ξ. For $\eta \in a$ we get $\eta \neq \xi$ and $\xi \notin \eta$ by transitivity of a, Hence $\eta \in \xi$. So we have $a = \xi \in \beta$. $\square$

Proof of Lemma 2.6.

Clearly every ordinal is hereditarily transitive. If conversely α is hereditarily transitive, then $\in$ is transitive on α and well–founded. Again by well–foundedness of $\in$ we have $\alpha \notin \alpha$. It remains to prove linearity. We show that any hereditarily transitive α is linearly ordered by $\in$ by $\in$–induction.
For $\xi, \eta \in \alpha$ let $\gamma := \xi \cap \eta$. If $\gamma = \xi \subsetneq \eta$ we get $\xi \in \eta$ by Lemma 2.5. If $\gamma \neq \xi$ we get $\gamma \in \xi$ by Lemma 2.5. If $\gamma = \eta$ we have $\eta \in \xi$ and $\gamma \neq \eta$ entails $\gamma \in \eta \cap \xi = \gamma$ which contradicts $\in$–foundation. $\square$

Proof of Lemma 2.20.

We show by induction on β that for $\alpha < \beta$ there is an ordinal ξ such that $\alpha + \xi = \beta$. If β is a successor γ' then $\alpha \leq \gamma$. If $\alpha = \gamma$ we choose $\xi := 0$. If $\alpha < \gamma$ there is by induction hypothesis a ξ_0 such that $\gamma = \alpha + \xi_0$, hence $\gamma = \alpha + \xi_0'$. If β is a limit ordinal then we get for every $\eta < \beta$ an ordinal ξ_η such that $\eta = \alpha + \xi_\eta$. Hence $\beta = \sup_{\eta<\beta}(\alpha + \xi_\eta) = \alpha + \xi$ for $\xi = \sup_{\eta<\beta} \xi_\eta$. $\square$

Proof of Lemma 2.21.

It follows by induction on α that $\xi, \eta < \omega^\alpha$ imply $\xi + \eta < \omega^\alpha$. This is obvious for $\alpha = 0$. For $\alpha = \beta'$ we obtain that $\xi, \eta < \omega^\alpha = \omega^\beta \cdot \omega$ implies $\xi < \omega^\beta \cdot n$ and $\eta < \omega^\beta \cdot m$, hence $\xi + \eta < \omega^\beta \cdot (n + m) < \omega^\beta \cdot \omega = \omega^\alpha$. For $\alpha \in$ *Lim* the claim follows from the induction hypothesis.

Conversely we observe that between ω^α and $\omega^{\alpha+1}$ all ordinals are additively decomposable. For if $\omega^\alpha < \xi < \omega^\alpha \cdot \omega$ there is an $n < \omega$ such that $\omega^\alpha \cdot n \leq \xi < \omega^\alpha \cdot (n + 1)$. Hence $\xi = \omega^\alpha \cdot n + \eta$ for $\eta < \omega^\alpha < \xi$. $\square$

Proof of Lemma 2.23

Since $\alpha < \varepsilon_0$ implies $\omega^\alpha < \varepsilon_0$ we get $\omega^{(n)}(\xi) < \varepsilon_0$ for $\xi < \varepsilon_0$ by induction on n. For $\eta := \sup\{\omega^{(n)}(0) \,|\, n \in \omega\}$ we thus have $\eta \leq \varepsilon_0$ and get $\omega^\eta = \sup\{\omega^{\omega^{(n)}(0)} \,|\, n \in \omega\} = \sup\{\omega^{(n+1)}(0) \,|\, n \in \omega\} = \eta$. Hence $\varepsilon_0 \leq \eta$. $\square$

Solution to Exercise 2.25

Since $\alpha \in$ *Lim* we have $\alpha \leq \xi + \alpha = \sup\{\xi + \eta \,|\, \eta < \alpha\} \leq \alpha$. $\square$

Solution to Exercise 2.26

Induction on α. The claim is clear for additively indecomposable ordinals α. If α is additively decomposable then $\alpha = \xi + \eta$ for $\xi, \eta < \alpha$. By induction hypothesis we get $\xi =_{NF} \xi_1 + \cdots + \xi_m$ and $\eta =_{NF} \eta_1 + \cdots + \eta_m$. Hence $\alpha = \xi_1 + \cdots + \xi_m + \eta_1 + \cdots + \eta_n =_{NF} \xi_1 + \cdots + \xi_k + \eta_1 + \cdots + \eta_n$ by Exercise 2.25 for k the index such that $\xi_k \geq \eta_1$ and $\xi_{k+1} < \eta_1$. □

Section 3

Solution to Exercise 3.3

If

$$f(x_1, \ldots, x_n) = \begin{cases} g_1(x_1, \ldots, x_n) & \text{if } R_1(x_1, \ldots, x_n) \\ \vdots & \\ g_n(x_1, \ldots, x_n) & \text{if } R_n(x_1, \ldots, x_n) \\ h(x_1, \ldots, x_n) & \text{otherwise} \end{cases}$$

for pairwise disjoint primitive–recursive predicates R_i then put

$$\begin{aligned} f(x_1, \ldots, x_n) = (&\sum_{i=1}^{n} g_i(x_1, \ldots, x_n) \cdot \chi_{R_i}(x_1, \ldots, x_n)) \\ &+ h(x_1, \ldots, x_n) \cdot \overline{sg}(\sum_{i=1}^{n} \chi_{R_i}(x_1, \ldots, x_n)). \end{aligned}$$

Solution to Exercise 3.4

It is $\chi_{\neg A} := \overline{sg}(\chi_A)$, $\chi_{(A \,\wedge\, B)} = \chi_a \cdot \chi_B$ and for $P(z, \vec{x}) \;\Leftrightarrow\; (\forall x \leq z) R(x, \vec{x})$ we have $\chi_P(z, \vec{x}) = \prod_{i=0}^{z} \chi_R(i, \vec{x})$, where $\prod_{i=0}^{0} f(i, \vec{x}) = f(0, \vec{x})$ and $\prod_{i=0}^{S(n)} f(i, \vec{x}) = \prod_{i=0}^{n} f(i, \vec{x}) \cdot f(S(n), \vec{x})$

Solution to Exercise 3.5

Define $F(0, y_1, \ldots, y_n) = 0$ and

$$F(S(k), y_1, \ldots, y_n) = \begin{cases} F(k, n) & \text{if } (\exists z \leq k)[f(z, y_1, \ldots, y_n) = 0] \\ S(k) & \text{otherwise} \end{cases}$$

and check that $F(k, y_1, \ldots, y_n) = \mu z \leq k.(f(z, y_1, \ldots, y_n) = 0)$.

Proof of Lemma 3.15 by induction on rnk(F).

If F is an atomic sentence then either F or $\neg F$ is true, hence in $\bigwedge$–type. By an $(\bigwedge)$ rule we thus obtain $\vdash_0^{\alpha} \Delta, F, \neg F$ for any ordinal α.

If F is a formula $t \in X$ then we obtain $\vdash_0^{\alpha} \Delta, t \in X, t \notin X$ for any α by an (X)–rule.

If F is in $\bigwedge$–type then we obtain $\vdash_0^{2 \cdot \mathrm{rnk}(G)} \Delta, G, \neg G$ for all $G \in \mathrm{CS}(F)$ by induction hypothesis. Hence $\vdash_0^{2 \cdot \mathrm{rnk}(G)+1} \Delta, G, \neg F$ by a $(\bigvee)$–rule and finally $\vdash_0^{2 \cdot \mathrm{rnk}(F)} \Delta, F, \neg F$

by a $(\bigwedge)$–rule. $\square$

Proof of Theorem 3.16

By a straightforward induction on α we get that $\vdash_0^\alpha \Delta$ implies $\mathbb{N} \models \bigvee\Delta[\Phi]$ for any assignment $\Phi(X) \subseteq \omega$. So soundness is straightforward.

More difficult is completeness, Since there are no free individual variables in pseudo Π^1_1–sentences every term has a computable value. W.l.o.g. we may therefore assume that all terms are replaced by the numerals denoting their value.

Let Δ be a finite sequence of pseudo Π^1_1–sentences (from now on just called sentences for short). Δ is reducible if it contains a sentence in $\bigwedge\text{–type} \cup \bigvee\text{–type}$. The first sentence in Δ which is in $\bigwedge\text{–type} \cup \bigvee\text{–type}$ is its *redex* $R(\Delta)$. The reduct Δ^r of a reducible Δ is obtained by cancelling its redex. We define the search tree S_Δ together with a label–function δ that assigns a finite sequence of sentences to the nodes of S_Δ.

- $\langle\,\rangle \in S_\Delta$ and $\delta(\langle\,\rangle) = \Delta$.

Now let $s \in S_\Delta$ and assume that $\delta(s)$ is not an instance of an X–rule.

- If $\delta(s)$ is irreducible then $s^\frown\langle 0\rangle \in S_\Delta$ and $\delta(s^\frown\langle 0\rangle) = \delta(s)$.
- If $F :\equiv R(\delta(s)) \in \bigwedge\text{–type}$ and $\mathrm{CS}(F) = \langle F_i \,|\, i \in I\rangle$ then $s^\frown\langle i\rangle \in S_\Delta$ and $\delta(s^\frown\langle i\rangle) = \delta(s)^r, F_i$.
- If $F :\equiv R(\delta(s)) \in \bigvee\text{–type}$ then $s^\frown\langle 0\rangle \in S_\Delta$ and $\delta(s^\frown\langle i\rangle) = \delta(s)^r, G, F$ where G is the first sentence in $\mathrm{CS}(F)$ which is not $\bigcup_{s_0 \subseteq s} \delta(s_0)$ if such a sentence exists.

Observe that S_Δ is primitive–recursively definable.

By an easy induction on the order–type $|s|$ in a well–founded tree S_Δ we immediately get:

$$\text{\textit{If } } S_\Delta \text{ \textit{is well–founded then we have} } \vdash_0^{|s|} \delta(s) \text{ \textit{for any node} } s \in S_\Delta. \tag{i}$$

If S_Δ is not well–founded it contains an infinite path f. Let

$$\delta(f) := \bigcup_{n\in\omega} \delta(\langle f(0), \dots, f(n)\rangle).$$

We define an assignment $\Phi(X) := \{n \,|\, (n \notin X) \in \delta(f)\}$ and prove

$$\mathbb{N} \not\models G[\Phi] \text{ \textit{for all} } G \in \delta(f). \tag{ii}$$

by induction on $\mathrm{rnk}(G)$.

If $G \equiv (n \in X)$ then $n \notin X$ cannot belong to $\delta(f)$ since f is infinite. Hence $n \notin \Phi(X)$. If $G \equiv (n \notin X)$ then $n \in \Phi(X)$.

If $G \equiv R(n_1, \dots, n_k)$ then there is a node $s = \overline{f}(m)$ such that $R(\delta(s)) = G$. Then $G \in \bigwedge\text{–type}$ implies that s is a leave which contradicts the infinity of f. Hence $G \in \bigvee\text{–type}$ which implies $\mathbb{N} \not\models G$.

For non–atomic $G \in \bigwedge\text{–type}$ we have $\mathrm{CS}(G) \neq \emptyset$. Therefore there is a $H \in \mathrm{CS}(F) \cap \delta(f)$ and $\mathbb{N} \not\models H[\Phi]$ follows by induction hypothesis. Hence $\mathbb{N} \not\models G[\Phi]$.

If $G \in \bigvee$–type then, since f is infinite, we get $H \in \delta(f)$ for all $H \in \mathrm{CS}(G)$. Hence $\mathbb{N} \not\models H[\Phi]$ for all $H \in \mathrm{CS}(G)$ which entails $\mathbb{N} \not\models G[\Phi]$.

If we assume $\not\vdash^{\alpha}_{0} F(X)$ for all $\alpha < \omega_1^{CK}$ we obtain by (i) that $S_{F(x)}$ cannot be well–founded. So there is by (ii) an assignment Φ such that $\mathbb{N} \not\models F(X)[\Phi]$ which implies $\mathbb{N} \not\models (\forall X)F(X)$.

Proof of Theorem 3.18

If $\mathrm{tc}\big(\mathit{WO}(\prec)\big) < \omega_1^{CK}$ we have $\vdash^{\alpha}_{0} \mathit{WO}(\prec)$ hence $\mathbb{N} \models WO(\prec)$ and $\prec$ is well–founded.

For the opposite direction we prove by induction on α:

$$\vdash^{\alpha}_{0} \neg \mathit{Prog}(X,\prec), n_1 \notin X, \ldots, n_k \notin X, \Delta \;\Rightarrow \mathbb{N} \models \bigvee \Delta[\prec\restriction\beta] \tag{i}$$

for a finite set Δ of X–positive formulas where $\prec\restriction\beta = \{n \,|\, \mathrm{otyp}_{\prec}(x) < \beta\}$ for $\beta = \max\{\mathrm{otyp}_{\prec}(n_1), \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha}$.

If the last inference in (i) affects Δ. we get the claim from the induction hypothesis, the semantical correctness of the inference rules and the monotonicity of X–positive sentences.

If the last inference affects

$$\neg \mathit{Prog}(X,\prec) \equiv (\exists x)[(\forall y)[y \prec x \to y \in X] \wedge x \notin X]$$

we have the premise

$$\vdash^{\alpha_0}_{0} \neg \mathit{Prog}(X,\prec), (\forall y)[y \prec n \to y \in X] \wedge n \notin X, n_1 \notin X, \ldots, n_k \notin X, \Delta \tag{ii}$$

for some numeral n. By $\wedge$–inversion we thus have

$$\vdash^{\alpha_0}_{0} \neg \mathit{Prog}(X,\prec), (\forall y)[y \prec n \to y \in X], n_1 \notin X, \ldots, n_k \notin X, \Delta \tag{iii}$$

and

$$\vdash^{\alpha_0}_{0} \neg \mathit{Prog}(X,\prec), n \notin X, n_1 \notin X, \ldots, n_k \notin X, \Delta\,. \tag{iv}$$

Towards a contradiction assume

$$\mathbb{N} \not\models \bigvee \Delta[\prec\restriction\beta].$$

Then we also have $\mathbb{N} \not\models \bigvee \Delta[\prec\restriction\beta_0]$ for $\beta_0 := \max\{\mathrm{otyp}_{\prec}(n_1), \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha_0}$. The induction hypothesis for (iii) then yields $(\forall y)[y \prec n \to \mathrm{otyp}_{\prec}(y) < \beta_0]$. i. e., $\mathrm{otyp}_{\prec}(n) \le \beta_0$. By induction hypothesis for (iv) we thus get $\mathbb{N} \models \bigvee \Delta[\prec\restriction\beta_1]$ for

$$\begin{aligned}
\beta_1 &= \max\{\mathrm{otyp}_{\prec}(n), \mathrm{otyp}_{\prec}(n_1), \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha_0} \\
&\le \max\{\mathrm{otyp}_{\prec}(n_1), \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha_0} + 2^{\alpha_0} \\
&= \max\{\mathrm{otyp}_{\prec}(n_1). \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha_0+1} \\
&\le \max\{\mathrm{otyp}_{\prec}(n_1), \ldots, \mathrm{otyp}_{\prec}(n_k)\} + 2^{\alpha}.
\end{aligned}$$

Contradiction. Setting $k = 0$ and $\Delta = \{(\forall x)[x \in X]\}$ in (i) we obtain the theorem.

□

Solution to Exercise 3.26

We prove the claim by induction on $\prec$. Let $m \prec n$ and $\alpha = \mathrm{otyp}_{\prec}(m)$. Then we have

$$\left|\frac{4\cdot\alpha}{0}\right. \neg \textit{Prog}(X,\prec), \neg(m \prec n), m \in X$$

for all m either by induction hypothesis or by an $\bigwedge$–rule with empty premise. Hence

$$\left|\frac{4\cdot\alpha+2}{0}\right. \neg \textit{Prog}(X,\prec), (\forall y)[y \prec n \to y \in X]$$

From Lemma 3.15 we have

$$\left|\frac{0}{0}\right. n \notin X, n \in X$$

and obtain

$$\left|\frac{4\cdot\alpha+3}{0}\right. \neg \textit{Prog}(X,\prec), (\forall y)[y \prec n \to y \in X] \wedge n \notin X, n \in X,$$

hence

$$\left|\frac{4\cdot\alpha+4}{0}\right. \neg \textit{Prog}(X,\prec), n \in X$$

and $4 \cdot \alpha \leq \mathrm{otyp}_{\prec}(n)$.

For every $\alpha < \omega_1^{CK}$ there is a primitive recursive well–ordering $\prec$ such that $\mathrm{otyp}(\prec) = \alpha$. W. l. o. g we may assume $\alpha \in \textit{Lim}$. Hence

$$\left|\frac{\mathrm{otyp}(\prec)}{0}\right. \neg \textit{Prog}(X,\prec), (\forall x)[x \in X]$$

which implies $\alpha \leq \pi^{\mathbb{N}}$. Since we already have $\pi^{\mathbb{N}} \leq \omega_1^{CK}$ it follows $\pi^{N} = \omega_1^{CK}$. $\square$

Section 4

Proof of Lemma 4.4

W. l. o. g. we assume that the language of first order logic is in Tait style. For a formula F we define its Boolean decompositions

$$\Delta(F) := \begin{cases} \Delta(A) \cup \Delta(B) & \text{if } F \equiv A \vee B \\ \{F\} & \text{otherwise} \end{cases}$$

and its Boolean degree

$$\textit{Bdeg}(F) := \begin{cases} \max\{\textit{Bdeg}(A), \textit{Bdeg}(B)\} + 1 & \text{if } F \equiv A \wedge B \\ 0 & \text{otherwise.} \end{cases}$$

For a finite set Δ of formulas we define $\textit{Bdeg}(\Delta)$ as the sum of the Boolean degrees of the formulas in Δ.
We observe:

- A formula F is Boolean valid iff $\bigvee\Delta(F)$ is Boolean valid.
- If $\textit{Bdeg}(F) = 0$ and F is Boolean valid then $\Delta(F) = \Delta_0, A, \neg A$ for a Boolean atom A.
- If $\textit{Bdeg}(F) > 0$ then $\Delta(F) = \Delta_0, A \wedge B$ for some formulas A and B. F is Boolean valid iff $\bigvee(\Delta_0, A)$ and $\bigvee(\Delta_0, B)$ are Boolean valid.

The claim now follows from Lemma 3.15 by induction on $\textit{Bdeg}(\Delta(F))$. $\square$

Proof of Theorem 4.5.

We prove the theorem by induction on the length of a derivation in the Hilbert calculus. If $F(x_1, \dots, x_n)$ is Boolean valid so is $F(t_1, \dots, t_n)$ and we obtain $\mid\!\frac{\alpha}{0}\, F(z_1, \dots, z_n)$ by Lemma 4.4.

By Lemma 3.15 we have $\mid\!\frac{\alpha}{0}\, \neg F(t), F(t)$ for $\alpha = 2 \cdot \mathrm{rnk}(F) < \omega$ and obtain $\mid\!\frac{\alpha+2}{0}\, \neg(\forall x)F(x) \vee F(t)$ by inferences $(\bigvee)$. Symmetrically we obtain $\mid\!\frac{\alpha+2}{0}\, \neg F(t) \vee (\exists x)F(x)$.

Since $t = t$ is true atomic we obtain $\mid\!\frac{0}{0}\, t = t$ for all terms t by $(\bigwedge)$ with empty premise. Hence $\mid\!\frac{1}{0}\, (\forall x)[x = x]$.

Similarly we obtain $\mid\!\frac{0}{0}\, s \neq t, t = s$ for all terms s and t by and inference $(\bigwedge)$. Hence $\mid\!\frac{3}{0}\, (\forall x)(\forall y)[x \neq y \vee y = x]$ by $(\bigvee)$ and two inferences $(\bigwedge)$.

Similarly we have $\mid\!\frac{0}{0}\, r \neq s, r \neq t, s = t$ by $(\bigwedge)$ and obtain $\mid\!\frac{4}{0}\, (\forall x)(\forall y)(\forall z)[x \neq y \vee y \neq z \vee x = z]$ by two $(\bigvee)$– and three $(\bigwedge)$–inferences.

An easy induction on $\mathrm{rnk}(F(x))$ shows $\mid\!\frac{2\cdot\mathrm{rnk}(F)}{0}\, s \neq t, \neg F(s), F(t)$. By iteration we obtain the translation of (Com).

The embedding of the inference rules follows directly from the induction hypotheses and the variable conditions in the $(\forall)$– and $(\exists)$–rules. □

Proof of Lemma 5.1 by induction on n.

Let $\beta := 2 \cdot \mathrm{rnk}(F)$. By Lemma 3.15 we have

$$\mid\!\frac{\beta}{0}\, \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(0) . \tag{i}$$

By induction hypothesis we have

$$\mid\!\frac{\beta+2n}{0}\, \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(n) \tag{ii}$$

and by Lemma 3.15

$$\mid\!\frac{\beta}{0}\, \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], \neg F(S(n)), F(S(n)) . \tag{iii}$$

From (ii) and (iii) we obtain by an inference $(\bigwedge)$

$$\mid\!\frac{\beta+2n+1}{0}\, \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(n) \wedge \neg F(S(n)), F(S(n)) \tag{iv}$$

and finally with an inference $(\bigvee)$.

$$\mid\!\frac{\beta+2(n+1)}{0}\, \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(S(n)) .$$ □

Section 5

Proof of Lemma 5.4

We induct on β. If the last inference is

$$\left|\frac{\beta_\iota}{\rho}\right. \Gamma_\iota, \neg F \quad \text{for} \quad \iota \in I \Rightarrow \left|\frac{\beta}{\rho}\right. \Gamma, \neg F$$

we obtain

$$\left|\frac{\alpha+\beta_\iota}{\rho}\right. \Delta, \Gamma_\iota$$

for all $\iota \in I$ by induction hypothesis and thence the claim by the same inference.

If the last inference is

$$\left|\frac{\beta_\iota}{\rho}\right. \Gamma, \neg F, \neg G \;\Rightarrow \left|\frac{\beta}{\rho}\right. \Gamma, \neg F \tag{i}$$

for $G \in \mathrm{CS}(F)$ we obtain by the induction hypothesis $\left|\frac{\alpha+\beta_0}{\rho}\right. \Delta, \Gamma, \neg G$. By $(\bigwedge)$–inversion and the structural rule (Str) we get $\left|\frac{\alpha}{\rho}\right. \Delta, \Gamma, G$ from the first premise and, since $\mathrm{rnk}(G) < \mathrm{rnk}(F) = \rho$, obtain $\left|\frac{\alpha}{\rho}\right. \Delta, \Gamma$ by cut. □

Proof of Theorem 5.5.

We induct on α. If the last inference is not a cut of rank ρ the claim follows directly from the induction hypothesis. If it is a cut

$$\left|\frac{\alpha_0}{\rho+1}\right. \Delta, F\,,\ \left|\frac{\alpha_0}{\rho+1}\right. \Delta, \neg F \;\Rightarrow \left|\frac{\alpha}{\rho+1}\right. \Delta$$

with $\mathrm{rnk}(F) = \rho$ we obtain by induction hypothesis

$$\left|\frac{\omega^{\alpha_0}}{\rho}\right. \Delta, F\,,\ \left|\frac{\omega^{\alpha_0}}{\rho}\right. \Delta, \neg F$$

and, since $\omega^{\alpha_0} + \omega^{\alpha_0} < \omega^\alpha$, obtain $\left|\frac{\omega^\alpha}{\rho}\right. \Delta$ by the Reduction Lemma . □

Proof of Theorem 5.6

If $\mathsf{NT} \vdash F$ we obtain $\left|\frac{\omega+\omega}{r}\right. F$ for $r < \omega$. By r–fold application of the Elimination Theorem we thus obtain $\left|\frac{\alpha}{0}\right. F$ for $\alpha < \varepsilon_0$. Hence $\mathrm{tc}\big(F\big) < \varepsilon_0$. □

Proof of Corollary 5.8

$\mathsf{NT} \vdash 0 = 1$ entails $\left|\frac{\alpha}{0}\right. 0 = 1$,i.e., $\models^{\alpha} 0 = 1$ which is impossible since $0 = 1$ is false in $\mathbb{N}$. □

Section 6

Proof of Corollary 6.2

For $\alpha < \varepsilon_0$ we define by simultaneous course–of–values recursion.
The codes:

$$\ulcorner 0 \urcorner = \langle 0, 0\rangle, \quad \ulcorner \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \urcorner := \langle 1, \ulcorner \alpha_1 \urcorner, \ldots, \ulcorner \alpha_n \urcorner \rangle,$$

The set On:

$$\ulcorner 0 \urcorner \in On$$
$$\ulcorner \alpha_1 \urcorner \succeq \ulcorner \alpha_2 \urcorner \succeq \cdots \succeq \ulcorner \alpha_n \urcorner \Rightarrow \ulcorner \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \urcorner \in On$$

The $\prec$–relation on On.

$$\ulcorner\alpha\urcorner \neq \ulcorner 0\urcorner \Rightarrow \ulcorner 0\urcorner \prec \ulcorner\alpha\urcorner$$

$$\begin{aligned}\ulcorner 0\urcorner \neq \ulcorner\alpha\urcorner \prec \ulcorner\beta\urcorner \;\Leftrightarrow\; & (\exists z < \max\{lh(\ulcorner\alpha\urcorner), lh(\ulcorner\beta\urcorner)\})[0 < z \\ & \wedge (\forall i < z)[(\ulcorner\alpha\urcorner)_i = (\ulcorner\beta\urcorner)_i \wedge (\ulcorner\alpha\urcorner)_z \prec (\ulcorner\beta\urcorner)_z]].\end{aligned} \qquad \square$$

Proof of Lemma 6.3

We work in NT. Under the hypotheses

$$Prog(X) \tag{i}$$

$$\xi \subseteq \mathcal{J}(X) \tag{ii}$$

we want to prove $\xi \in \mathcal{J}(X)$, i.e.

$$(\forall\eta)[\eta \subseteq X \to \eta + \omega^\xi \subseteq X]. \tag{iii}$$

So let

$$\eta \subseteq X \tag{iv}$$

$$\nu < \eta + \omega^\xi. \tag{v}$$

If $\nu \leq \eta$ we get $\nu \in X$ by (iv) and (i). So assume

$$\eta < \nu < \eta + \omega^\xi.$$

Then $\xi \neq 0$ and we get

$$\eta < \nu = \eta + \omega^{\nu_1} + \cdots + \omega^{\nu_n} < \eta + \omega^\xi \tag{vi}$$

with $\nu_i < \xi$. Since $\eta \in X$ by (iv) and (i) and $\nu_1 < \xi \subseteq \mathcal{J}(X)$ we get $\eta + \omega^{\nu_1} \in X$ by (ii). By induction on n (which is a formal induction on $lh(\ulcorner\nu\urcorner)$ in NT!) we finally obtain $\nu \in X$. $\square$

Proof of Lemma 6.4

We work in NT. Assume $Prog(X) \to \alpha \subseteq X$. Substituting $\mathcal{J}(X)$ for X entails $Prog(\mathcal{J}(X)) \to \alpha \subseteq \mathcal{J}(X)$, hence $Prog(\mathcal{J}(X)) \to \omega^\alpha \subseteq X$. Together with Lemma 6.3 we thus get $Prog(X) \to \omega^\alpha \subseteq X$. $\square$

Proof of Theorem 6.5.

For $\alpha < \varepsilon_0$ there is a finite n such that $\alpha < \omega^{(n)}(0)$. We trivially have $TI(0)$ and obtain $TI(\omega^{(n)}(0))$, hence also $TI(\alpha)$ by n–fold application of Lemma 6.4. (This time it is an induction from outside). $\square$

Proof of Theorem 6.7

We have $\mathbb{N} \models (\forall X)(\forall x)[Prog(X) \wedge x \in On \to x \in X]$ and NT $\vdash Prog(X) \wedge n \in$

$On \to n \in X$ by (the proof of) Theorem 6.5. But $\mathsf{NT} \vdash (\forall x)[\mathit{Prog}(X) \land x \in On \to x \in X]$ would imply $\mathrm{otyp}(\prec) < \varepsilon_0$ by Theorem 5.6 while $\mathrm{otyp}(\prec) = \varepsilon_0$ holds true by the construction of the relation $\prec$. $\square$

References

[1] G. GENTZEN, *Die Widerspruchsfreiheit der reinen Zahlentheorie*, **Mathematische Annalen**, vol. 112 (1936), pp. 493–565.

[2] ——, *Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie*, **Mathematische Annalen**, vol. 119 (1943), pp. 140–161.

[3] W. POHLERS, **Proof theory. The first step into impredicativity**, Universitext, Springer-Verlag, Berlin/Heidelberg/New York, 2009.