Introduction to Predicative Ordinal Analysis

Summary of a lecture given at the Hilbert–Bernays Summer School on Logic and Computation 2018 in Göttingen.

A preparation for the course "Introduction to Impredicative Ordinal Analysis"

Wolfram Pohlers
WWU Münster

Contents

0.1 Remark The mark ☐ at the end of a theorem, a lemma or an exercise is a link to the proof/solution of the theorem, lemma or exercise. Perhaps not all links will be present (in cases that the solution is planned to be left to the reader).

1 Preface

These are the lecture notes for a course on predicative ordinal analysis given at a previous summer school. Athough I do not adopt all the notions of this course, I will stick to them to a large extent. The present course, however, will go much further. Nevertheless, I will start with a brief summary of predicative ordinal analysis and it could be usefull to have a look at these notes.

2 A brief reminder

2.1 Abstract structures and logical inferences

In a very general setting mathematics is concerned with the study of abstract structures. An abstract structure has the form $\mathfrak{M} = (M, \mathscr{C}, \mathscr{R}, \mathscr{F})$ where M is a non–void set, $\mathscr{C}$ as subset of M, $\mathscr{R}$ a set of relations on M and $\mathscr{F}$ a set of functions on M. Associated to an abstract structure is its abstract language $\mathscr{L}_{\mathfrak{M}} = \mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$ which comprises a set $\mathcal{C}$ of constants for elements of M, a set $\mathcal{R}$ of symbols for the relations in $\mathscr{R}$ and a set $\mathcal{F}$ of symbols for the functions in $\mathscr{F}$.

In general a *signature* for a logical language is a triple $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ where every $R \in \mathcal{R}$ and $f \in \mathcal{F}$ carries its arity $0 < \#R \in \mathbb{N}$ and $0 < \#f \in \mathbb{N}$.

An abstract structure $\mathfrak{M} = (M, \mathscr{C}, \mathscr{R}, \mathscr{F})$ *interprets* a signature $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ if every *constant* $c \in \mathcal{C}$ has an interpretation $c^{\mathfrak{M}} \in M$, every *relation symbol* $R \in \mathcal{R}$ an interpretation $R^{\mathfrak{M}} \subseteq M^{\#R}$ and every *function symbol* $f \in \mathcal{F}$ an interpretation $f^{\mathfrak{M}} \colon M^{\#f} \longrightarrow M$.

We say that a signature *matches* the structure $\mathfrak{M}$ if there is a symbol $c \in \mathcal{C}$, $R \in \mathcal{R}$ and f in $\mathcal{F}$ for every constant in $\mathscr{C}$, every relation in $\mathscr{R}$ and every function in $\mathscr{F}$.

The *closed terms* of a signature $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ are either constants or composed terms of the form $(ft_1, \ldots, t_n)$ where $\#f = n$ and $t_1, \ldots, t_n$ are constants or previously defined composed terms.

Atomic sentences have the form $(Rt_1, \ldots, t_n)$ where R is an n–ary relation symbols and $t_1, \ldots, t_n$ are closed terms. Starting from atomic sentences we can inductively build a logical language $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$ using the familiar boolean operations and quantifications. If quantification is restricted to individuals we talk about a *first order logic* $\mathscr{L}_1(\mathcal{C}, \mathcal{R}.\mathcal{F})$. If we also allow quantifiers ranging over relations we talk about a *second order logic* $\mathscr{L}_2(\mathcal{C}, \mathcal{R}, \mathcal{F})$.

For an abstract structure $\mathfrak{M}$ that interprets the signature of a logical language every closed term t possesses a canonical interpretation $t^{\mathfrak{M}} \in M$.

Defining $\mathfrak{M} \models (Rt_1, \ldots, t_n)$ iff $(t_1^{\mathfrak{M}}, \ldots, t_n^{\mathfrak{M}}) \in R^{\mathfrak{M}}$ and continuing inductivley according to the meaning of the logical operations we obtain a canonical *satisfiability* *relation* $\mathfrak{M} \models F$ for the sentences F in the language $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$. We say that F is *valid in* $\mathfrak{M}$ iff $\mathfrak{M} \models F$.

If $S \cup \{F\}$ is a set of $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentences we call $S \models F$ a *logical inference* iff for every abstract structure $\mathfrak{M}$ that interprets $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ the fact that $\mathfrak{M} \models G$ holds true for all $G \in S$ also implies $\mathfrak{M} \models F$.

An abstract structure $\mathfrak{M}$ *satisfies* a set S of $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentences iff $\mathfrak{M}$ interprets $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ and satisfies all sentences in S.

A set S of $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentences is *consistent* iff there is a structure $\mathfrak{M}$ which satifies S.

A set S of $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentences is *logically valid* iff every structure which interprets $(\mathcal{C}, \mathcal{R}, \mathcal{F})$ satifies S.

2.1 Exercise Give a formal definition of $t^{\mathfrak{M}}$ and $\mathfrak{M} \models F$.

2.2 Exercise Show that $S \models F$ iff $S \cup \{\neg F\}$ is inconsistent.

2.3 Exercise Show that a set S of $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentences is consistent iff there is no $\mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$–sentence F such that $S \models F \wedge \neg F$.

2.2 Formal derivations

We extend a language $\mathscr{L} := \mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$ by adding free individual variables, denoted by $x, y, z, x_1, \ldots$, and free relation variables, denoted by $X, Y, Z, X_1, \ldots$, together with their arities $\#X \in \mathbb{N}$. In forming terms individual variables are treated like constants; in forming formulae relation variables are treated like relation symbols. Terms without occurrences of free individual variables are *closed*, sentences are formulae in which neither individual variables nor relation variables occur freely.

Even if a structure $\mathfrak{M}$ interprets $\mathscr{L}$ there is no canonical interpretation for terms containing free individual variables and no canonical satisfaction relation for formulae containing free variables. Interpretation and satisfaction of terms and formulae need an *assignment* Φ which assigns an element $\Phi(x) \in M$ to every free individual variable x and a set $\Phi(X) \subseteq M^{\#X}$ to every relation variable X. We denote by $\mathfrak{M} \models F[\Phi]$ that $\mathfrak{M}$ satisfies the formula F under the assignment Φ.

We extend the definition of a logical inference to sets of $\mathscr{L}$–formulae.

2.4 Definition Let $S \cup \{F\}$ be a set of $\mathscr{L}$–formulae. Then $S \models F$ is a logical inference iff for every structure $\mathfrak{M}$ that interprets $\mathscr{L}$ and every assignment Φ we have

$$\mathfrak{M} \models G[\Phi] \text{ for all formulae in } G \in S \text{ implies } \mathfrak{M} \models F[\Phi].$$

2.5 Definition Let $\mathscr{L} := \mathscr{L}(\mathcal{C}, \mathcal{R}, \mathcal{F})$ be a logical language. A *formal rule* is a figure

$$P_1, \ldots, P_n \vdash C,$$

where $n \geq 0$ and $\{P_1, \ldots, P_n, C\}$ is a set of $\mathscr{L}$–formulae.

A *formal system* $\mathbb{S}$ is a set of formal rules.

Given a formal system $\mathbb{S}$ we define *formal derivability* $A_1, \ldots, A_m \vdash_{\mathbb{S}} F$ inductively by:

- *If* $A_1, \ldots, A_m \mathrel{\underset{\mathbb{S}}{\vdash}} P_i$ *for* $i = 1, \ldots, n$ *and* $P_1, \ldots, P_n \vdash F$ *is a rule of* $\mathbb{S}$ *then* $A_1, \ldots, A_m \mathrel{\underset{\mathbb{S}}{\vdash}} F$.

A formal system $\mathbb{S}$ is *sound* if for every rule $P_1, \ldots, P_n \vdash C$ in $\mathbb{S}$ we have $P_1, \ldots, P_n \models C$.

2.6 Exercise (Soundness Theorem) Let $\mathbb{S}$ be a sound formal system. Show that $A_1, \ldots, A_n \mathrel{\underset{\mathbb{S}}{\vdash}} F$ entails $A_1, \ldots, A_n \models F$.

The following completeness theorem by Kurt Gödel is one of the most important theorems of Mathematical Logic.

2.7 Theorem *(Gödel's completeness theorem) Let* $\mathscr{L}_1$ *be a first order language and* $S \cup \{F\}$ *a set of* $\mathscr{L}_1$*–formulae. Then there is a sound formal system* $\mathbb{S}$ *such that* $S \models F$ *entails* $S_1, \ldots, S_n \mathrel{\underset{\mathbb{S}}{\vdash}} F$ *for a finite subset* $\{S_1, \ldots, S_n\}$ *of* S.

There is a fraternal twin to Gödel's completeness theorem.

2.8 Theorem *(Compactness Theorem) Let* $\mathscr{L}_1$ *be a first order language and* S *a set of* $\mathscr{L}_1$*–sentences. If every finite subset* $S_i \subseteq S$ *is consistent, then* S *is consistent.*

2.9 Exercise Show that Gödel's completeness theorem entails the compactness theorem. (The opposite direction—though true—is much harder to show).

A formal dervation $A_1, \ldots, A_n \mathrel{\underset{\mathbb{S}}{\vdash}} F$ in a formal system $\mathbb{S}$ can be viewed as a finite tree whose root is labelled by F, whose leaves are labelled by the formulae A_i and which is is locally correct with respect to the rules in $\mathbb{S}$. This makes the correctness of a formal proof machine–checkable, i.e., decidable. Admittedly in practice mathematical proofs are not formalized to the point that they become machine–checkable, but they should be formalizable in principle. This fact is responible for the intersubjectibility of mathematical proofs.

For full second order logic there is no compactness theorem, hence also no completeness theorem. So full second order logic is, in principle, useless for mathematical reasoning. Nevertheless there are sound formal systems for second order logic.[1]

2.3 Why ordinal analysis?

Gödel's completeness theorem establishes a tool for the investigation of abstract structures. We can try to characterize a structure $\mathfrak{M}$ by a set of first order sentences which are characteristic for and valid in $\mathfrak{M}$, the axioms for $\mathfrak{M}$. Starting from the axioms for $\mathfrak{M}$ we can argue by logical inferences to ensure that everything we conclude is a theorem of $\mathfrak{M}$. Examples for this approach are group theory, ring theory, field theory and similar algebraic disciplines.

In setting up an axiom system for a structure $\mathfrak{M}$ we are confronted with a couple of problems. First we have to ensure that the set of axioms is consistent. This causes no problems in case of groups, rings, field etc, since there are finite structures which satisfy

[1] These system, however, should rather be viewed as formal systems for a two sorted first order logic.

the finitely many axioms. The second problem is that of categoricity, i.e., the question whether we can characterize the structure by the axioms up to isomorphism. This, however, is in general not possible for a first order axiom system by the compactness theorem. In the case of groups, rings etc., it is not even desirable since we know that there are many non isomorphic groups, rings, ...

The situation is different when we try to axiomatize standard structures which we believe to be familiar with. The first—and probably most important— such structure is the structure $\mathfrak{N}$ of natural numbers. We have (and all our mathematical ancestors had) in some sense a clear intuition of this structure. Here it would be desirable to have an axiomatization up to isomorphism but this is excluded by the compactness theorem.There are categorial second order axiom systems for $\mathfrak{N}$ but, according to the lack of a completeness theorem for second order logic, they are mathematically useless. So we have to resign categoricity.

There are well-established axiom system for $\mathfrak{N}$, e.g. the Peano axioms which we will later introduce in detail. Since we have resigned categoricity it remains the problem of consistency. This is not so easy to solve as in the case of the group—or similar algebraic—axioms since the standard structure $\mathfrak{N}$ is an infinite structure. Therefore any adequate axiomatization of $\mathfrak{N}$ has to incorporate infinity which entails that there exist no finite structures that satisfy these axioms (as e.g. in group theory). But $\mathfrak{N}$ is in some sense the simplest infinite structure. In order to build a structure which satisfies the axioms for $\mathfrak{N}$ we need a structure somehow above $\mathfrak{N}$ which itself needs a consistent axiomatization which then is likely to embrace the axioms for $\mathfrak{N}$.

This exposes a foundational problem. Hilbert in his programme suggested a way to solve this problem (even aiming at solving the consistency problem for all existing mathematics) by formalization. Since a formal proof is a finite figure it should be likely that we can show by finitistic—i.e. purely finite combinatorical—means that there cannot be a proof figure of a contradiction.

This hope was destroyed by Gödel's incompleteness theorems in which he showed that a proof of the consistency of any recursively enumerable axiom system for $\mathfrak{N}$ has to exceed the means of this axiom system. Especially there cannot be a consistency proof for a recursively enumerable axiom system of $\mathfrak{N}$ by finite combinatorics.

However, despite of Gödels incompleteness theorems Gerhard Gentzen in [**?**] gave a consistency proof for the Peano axioms for $\mathfrak{N}$. His proof only used finitistic means except for an application of a transfinite induction along a well–ordering of order–type ε_0. By Gödel's incompleteness theorem it therefore follows that transfinite induction up to ε_0 cannot be provable from the Peano axioms. In a later paper [**?**] he showed that conversely any ordinal less than ε_0 can be represented by a well–ordering whose well–foundedness is provable in Peano 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.

As we will see later the proof–theoretic ordinal of an axiom system in fact incorporates a measure for the *performance* of an axiom with respect to the intended standard structure and the universe of its subsets above it.

The aim of the course is to give an introduction to ordinal analysis on the example of an axiom system for $\mathfrak{N}$ which is equivalent to the Peano axioms.

2.10 Remark (added 2019) *The present course on* impredicative ordinal analysis *will emphasize the fact that only additional axioms for the universe above the standard structure can enhance the performance of an axiom system. This may lead to* impredicative axioms systems *which are the subject of the present course.*

3 Ordinals

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

3.1 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 simply 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).

3.1 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

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

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

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

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

Figure 1: Some set theoretical ordinals

Let WO($\prec$) abbreviate the above sentence. Then $\prec$ is a well–ordering iff $\prec$ is a total, irreflexive and transitive ordering that satisfies WO($\prec$).

3.3 Exercise Prove theorem 3.1.

3.2 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.

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

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

3.4 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

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

Clearly

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

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

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

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

3.6 Lemma *$\beta \in On$, $\mathit{Tran}(a)$ and $a \subsetneq \beta$ imply $a \in \beta$.* □

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

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

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

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

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

3.11 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'.*

3.12 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.

3.13 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 $\xi < \omega$.

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

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

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

and

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

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

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

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

3.3 Basics of ordinal arithmetic

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

3.18 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.

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

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

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

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

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

3.24 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 $\varepsilon_0 = \sup\{\omega^{(n)}(\xi) \,|\, n \in \omega\}$.* □

3.25 Remark By ε–numbers we refer to ordinals that are closed under ω–powers, i.e. for which $\alpha < \varepsilon$ entails also $\omega^\alpha < \varepsilon$.

3.26 Exercise Prove Lemmata 3.21 , 3.22 and 3.24.

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

3.28 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 $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$. □

3.29 Exercise Show that for any ε–number ε we have $2^\varepsilon = \omega^\varepsilon = \varepsilon$.

4 The verification calculus for countable structures

Let $\mathfrak{M} = (M, \mathscr{C}, \mathscr{R}, \mathscr{F})$ be an abstract structures with countable domain M. We will in future tacitly presume that equality $=$ and inequality $\neq$ are among the relations $\mathscr{R}$. Moreover we assume that with each relation in $\mathscr{R}$ also its complement belongs the $\mathscr{R}$.

4.1 Definition We define the closed terms and sentences of $\mathscr{L}_1(\mathfrak{M})$ inductively by the following clauses.

- For every element $m \in M$ the constant $\underline{m}$ is a closed term.
- If $t_1, \ldots, t_n$ are closed terms of $\mathscr{L}_1(\mathfrak{M})$ and f is a symbol for an n–ary function in $\mathscr{F}$ then $(ft_1, \ldots, t_n)$ is a closed $\mathscr{L}_1(\mathfrak{M})$ term.
- If $t_1, \ldots, t_n$ are closed $\mathscr{L}_1(\mathfrak{M})$–terms and R is a symbol for a relation in $\mathscr{R}$ then $(Rt_1, \ldots, t_n)$ is an (atomic) sentence of $\mathscr{L}_1(\mathfrak{M})$.

- If A and B are sentences of $\mathscr{L}_1(\mathfrak{M})$ then $(A \wedge B)$ and $(A \vee B)$ are sentences of $\mathscr{L}_1(\mathfrak{M})$.
- If $F(\underline{z})$ is an $\mathscr{L}_1(\mathfrak{M})$–sentence then $(\forall x)F(x)$ and $(\exists x)F(x)$ are $\mathscr{L}_1(\mathfrak{M})$–sentences.

Observe the pecularity—which we introduce for technical reasons—that there is no negation symbol in our language. This beomes superfluous by the requiremant that for every relation in $\mathscr{R}$ we also have its complement in $\mathscr{R}$. Therefore we have for every symbol R for a relation in $\mathscr{R}$ also a symbol $\breve{R}$ for its complement. This makes negation definable.

4.2 Definition (Inductive definition of $\neg F$)

- It is $\neg(Rt_1,\dots,t_n) \equiv (\breve{R}t_1,\dots,t_n)$ and $\neg(\breve{R}t_1,\dots,t_n) \equiv (Rt_1,\dots,t_n)$.
- It is $\neg(A \wedge B) \equiv (\neg A \vee \neg B)$ and $\neg(A \vee B) \equiv (\neg A \wedge \neg B)$.
- It is $\neg(\forall x)F(x) \equiv (\exists x)[\neg F(x)]$ and $\neg(\exists x)F(x) \equiv (\forall x)[\neg F(x)]$.

For every $\mathscr{L}_1(\mathfrak{M})$–sentence F we introduce a *characteristic sequence* $\mathrm{CS}(F)$ which should be viewed as the collection of all the sentences that suffice to verify F. If all the sentences in $\mathrm{CS}(F)$ are needed to verify F we say that F is of *conjunction type*, denoted by $F \in \bigwedge\text{–type}$, if already some of the sentences in $\mathrm{CS}(F)$ suffices we say that F is of *disjunction type*, written as $F \in \bigvee\text{–type}$.

4.3 Definition We define

- $\mathrm{CS}(F)$ is the empty sequence if F is an atomic sentence.
- $\mathrm{CS}(A \wedge B) = \mathrm{CS}(A \vee B) = \langle A, B\rangle$.
- $\mathrm{CS}((\forall x)F(x)) = \mathrm{CS}((\exists x)F(x)) = \big\langle F(\underline{z}) \,\big|\, z \in M \big\rangle$ where we assume a fixed enumeration of the countable set M.

In most cases the order in the sequence $\mathrm{CS}(F)$ is unimportant. Therefore we mostly neglect the order and identify $\mathrm{CS}(F)$ with the set of its members. So we write sloppily $\emptyset$ for the empty sequence and briefly $G \in \mathrm{CS}(F)$ to denote that G is a member of the sequence.

4.4 Definition The $\bigwedge$–type of $\mathscr{L}_1(\mathfrak{M})$ comprises

- All the atomic sentences $(Rt_1,\dots,t_n)$ such that $\mathfrak{M} \models (Rt_1,\dots,t_n)$.
- All $\mathscr{L}_1(\mathfrak{M})$–sentences of the form $(A \wedge B)$.
- All $\mathscr{L}_1(\mathfrak{M})$–sentences of the form $(\forall x)F(x)$.

Dually the $\bigvee$–type of $\mathscr{L}(\mathfrak{M})$ comprises

- All the atomic sentences $(Rt_1,\dots,t_n)$ such that $\mathfrak{M} \not\models (Rt_1,\dots,t_n)$.
- All $\mathscr{L}_1(\mathfrak{M})$–sentences of the form $(A \vee B)$.

- All $\mathscr{L}_1(\mathfrak{M})$–sentences of the form $(\exists x)F(x)$.

4.5 Observation *Let F be an $\mathscr{L}_1(\mathfrak{M})$–sentence.*

If $F \in \bigwedge$–type we have $\mathfrak{M} \models F$ iff $\mathfrak{M} \models G$ for all $G \in \mathrm{CS}(F)$.

If $F \in \bigvee$–type we have $\mathfrak{M} \models F$ iff there is a $G \in \mathrm{CS}(F)$ such that $\mathfrak{M} \models G$.

4.6 Definition (The verification calculus for $\mathscr{L}_1(\mathfrak{M})$.) Let Δ be a finite set of $\mathscr{L}_1(\mathfrak{M})$–sentences. We inductively define the verification calculus $\mathfrak{M} \models^{\alpha} \Delta$, where α denotes an ordinal.

($\bigwedge$) If $F \in \bigwedge$–type $\cap \Delta$ and $\mathfrak{M} \models^{\alpha_G} \Delta \cup \{G\}$ and $\alpha_G < \alpha$ holds true for all $G \in \mathrm{CS}(F)$ then we conclude $\mathfrak{M} \models^{\alpha} \Delta$.

($\bigvee$) If $F \in \bigvee$–type $\cap \Delta$ and $\mathfrak{M} \models^{\alpha_0} \Delta \cup \{G\}$ holds true for some $G \in \mathrm{CS}(F)$ then we obtain $\mathfrak{M} \models^{\alpha} \Delta$ for all $\alpha > \alpha_0$.

4.7 Theorem *The verification calculus is sound, i.e.*
If $\mathfrak{M} \models^{\alpha} \{F_1, \dots, F_n\}$ for some ordinal α then $\mathfrak{M} \models F_1 \lor \cdots \lor F_n$.

The verification calculus is complete, i.e.
If $\mathfrak{M} \models F$ for an $\mathscr{L}_1(\mathfrak{M})$–sentence F then there is an ordinal α such that $\mathfrak{M} \models^{\alpha} F$. □

To simplify notations let us from now on assume that the functions in $\mathscr{F}$ contain a sufficently strong coding and decoding functions. We extend the first order language $\mathscr{L}_1(\mathfrak{M})$ to the language $\mathscr{L}(\mathfrak{M})$ by adding a set variable, i.e. a unary relation variable X and extend Definition 4.1 by the additional clause

- If t is a closed $\mathscr{L}_1(\mathfrak{M})$–term then $t \in X$ and $t \notin X$ are $\mathscr{L}(\mathfrak{M})$–formulae.

The satisfaction relation $\mathfrak{M} \models F$ is no longer defined for fomulae F containing the set variable X. To obtain a satisfaction relation also for $\mathscr{L}(\mathfrak{M})$–formulae we define

- $\mathfrak{M} \models F$ iff $\mathfrak{M} \models F[\Phi]$ for all assignments Φ such that $\Phi(X) \subseteq M$.

This means that we treat $\mathscr{L}(\mathfrak{M})$–formulae which contain the set variable semantically as Π^1_1–formulae in full second order semantics. For that reason we refer to them as *pseudo Π^1_1–sentences*.

To define negation also for the language $\mathscr{L}(\mathfrak{M})$ we have to extend Definition 4.2 by the addional rule

- It is $\neg(t \in X) \equiv (t \notin X)$ and $\neg(t \notin X) \equiv (t \in X)$.

We clearly cannot verify atomic formulae of the form $t \in X$ and $t \notin X$. Therefore we extend Definition 4.3 by the clause

- $\mathrm{CS}(t \in X) = \mathrm{CS}(t \notin X) = \emptyset$

We keep Definition 4.4 of $\bigwedge$–type and $\bigvee$–type under the agreement that atomic formulae $(t \in X)$ and $(t \notin X)$ belong neither to $\bigwedge$–type nor to $\bigvee$–type.

We now extend the verification calculus $\mathfrak{M} \models^{\alpha} \Delta$ as defined in Definition 4.6 to the *canonical semi–formal system* $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$ for the language $\mathscr{L}(\mathfrak{M})$. For the definition of the semi–formal system we assume that every $\mathscr{L}(\mathfrak{M})$–formula possesses a complexity $\mathrm{rnk}(F)$ which is an ordinal such that $\mathrm{rnk}(G) < \mathrm{rnk}(F)$ holds true for all $G \in \mathrm{CS}(F)$.

4.8 Definition Let Δ be a finite set of $\mathscr{L}(\mathfrak{M})$–formulae. The *canonical semi–formal system* $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$ for the language $\mathscr{L}(\mathfrak{M})$ $\mathfrak{M}$ is given by the following rules.

- (X) If $\{(t \in X), (x \notin X)\} \subseteq \Delta$ and $\mathfrak{M} \models (s = t)$ then $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$ holds true for all ordinals α and ρ.
- $(\bigwedge)$ If $F \in \Delta \cup \bigwedge$–type, $\mathfrak{M} \mid\frac{\alpha_G}{\rho} \Delta \cup \{G\}$ and $\alpha_G < \alpha$ hold true for all $G \in \mathrm{CS}(F)$ then we also have $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$.
- $(\bigvee)$ If $F \in \Delta \cup \bigvee$–type and $\mathfrak{M} \mid\frac{\alpha_0}{\rho} \Delta \cup \{G\}$ for some $G \in \mathrm{CS}(F)$ then we obtain $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$ for all $\alpha > \alpha_0$.
- (Cut) If $\mathfrak{M} \mid\frac{\alpha_0}{\rho} \Delta \cup \{F\}$, $\mathfrak{M} \mid\frac{\alpha_0}{\rho} \Delta \cup \{\neg F\}$ and $\mathrm{rnk}(F) < \rho$ then we have $\mathfrak{M} \mid\frac{\alpha}{\rho} \Delta$ for all ordinals $\alpha > \alpha_0$.

We call the formulae $(s \in X), (t \notin X)$ and F in the above inference rules the *main–formulae* of the corresponding inference. A (cut) has no main–formula but the cut formula F.

4.9 Lemma *For every $\mathscr{L}(\mathfrak{M})$–formula and finite set Δ of $\mathscr{L}(\mathfrak{M})$–formulae we have*

$$\mathfrak{M} \mid\frac{2\cdot \mathrm{rnk}(f)}{0} \Delta \cup \{F, \neg F\}. \qquad \square$$

4.10 Theorem *($\mathscr{L}(\mathfrak{M})$–soundness) Assume $\mathfrak{M} \mid\frac{\alpha}{\rho} \{F_1, \ldots, F_n\}$ for $\mathscr{L}(\mathfrak{M})$–fomulae F_i $i = \{1, \ldots, n\}$. Then $\mathfrak{M} \models (F_1 \vee \cdots \vee F_n)$.* $\square$

4.11 Theorem *($\mathscr{L}(\mathfrak{M})$–completeness) Assume $\mathfrak{M} \models (\forall X)F(X)$ for an $\mathscr{L}(\mathfrak{M})$–fomula $F(X)$. Then there is a countable ordinal α such that $\mathfrak{M} \mid\frac{\alpha}{0} F(X)$.* $\square$

4.12 Definition For a Π^1_1–sentence $(\forall X)F(X)$ we define its *truth–complexity*

$$\mathrm{tc}((\forall X)F(X)) := \begin{cases} \min \{\alpha \mid \mathfrak{M} \mid\frac{\alpha}{0} F(X)\} & \text{if this exists} \\ \omega_1 & \text{otherwise} \end{cases}$$

where ω_1 denotes the first uncountable ordinal. The truth complexity of a pseudo Π^1_1–sentence $F(X)$ is the truth–complexity of the corresponding Π^1_1–sentence $(\forall X)F(X)$. For a countable structure $\mathfrak{M}$ we define its Π^1_1–ordinal

$$\pi^{\mathfrak{M}} := \sup \{\mathrm{tc}(F) \mid F \in \mathscr{L}(\mathfrak{M}) \text{ and } \mathfrak{M} \models F\}.$$

4.13 Remark The Π^1_1–ordinal of a countable structure $\mathfrak{M}$ has many counterparts in abstract recursion theory. E.g. it coincides with the supremum of the order–types of well–orderings that are elementarily definable in M, with the closure ordinal of the structure $\mathfrak{M}$ in the sense of [**?**] and also other characteristic ordinals. [2]

4.14 Observation
a) If X does not occur in F and $\mathfrak{M} \models F$ then $\mathrm{tc}(F) \leq \mathrm{rnk}(F)$.
b) We have $\mathfrak{M} \models (\forall X)F(X)$ for an $\mathscr{L}(\mathfrak{M})$–formula $F(X)$ iff $\mathrm{tc}((\forall X)F(X)) < \omega_1$.

4.15 Definition Let T be an axiom system for a countable structure $\mathfrak{M}$. Then we define the Π^1_1–ordinal of the axiom system T as the ordinal

$$\pi^{\mathfrak{M}}(T) := \sup\{\mathrm{tc}(F) \,|\, F \in \mathscr{L}(\mathfrak{M}) \text{ and } T \models F\}.$$

The distance between $\pi^{\mathfrak{M}}(T)$ and $\pi^{\mathfrak{M}}$ can be viewed as a measure for the performance of the axiom system T.

4.16 Theorem *An axiom system T for a countable structure $\mathfrak{M}$ is consistent iff $\pi^{\mathfrak{M}}(T) \leq \pi^{\mathfrak{M}}$.*

Since Gentzen ([**?**]) we define the *proof theoretic ordinal* $|T|$ of an axiom system T as the supremum of the order–types of well–orderings which are definable in T and whose well–foundedness is provable in T. More precisely let $\mathit{Wf}(\prec, X)$ denote the $\mathscr{L}(\mathfrak{M})$–formula such that $\mathit{WO}(\prec) \equiv (\forall X)\mathit{Wf}(\prec, X)$.

4.17 Definition Let T be an axiom system for a structure $\mathfrak{M}$ then

$$|T| := \{\mathrm{otyp}(\prec) \,|\, \prec \text{ is definable in } \mathscr{L}_1(\mathfrak{M}) \text{ and } T \vdash \mathit{Wf}(\prec, X)\}.$$

4.18 Lemma *Let $\prec$ be a well–ordering which is $\mathscr{L}_1(\mathfrak{M})$–definable in a countable structure $\mathfrak{M}$. Then $\mathfrak{M} \,\big|\!\frac{\alpha}{0}\, \{\neg(\forall x)[(\forall y)[y \prec x \to y \in X] \to x \in X], \underline{z} \in X\}$ entails* $\mathrm{otyp}_{\prec}(z) \leq 2^{\alpha}$. □

4.19 Theorem *Let $\prec$ be an well–orderings which is $\mathscr{L}_1(\mathfrak{M})$–definable in a structure $\mathfrak{M}$. Then* $\mathrm{otyp}(\prec) \leq 2^{\mathrm{tc}(\mathit{Wf}(\prec))}$.

4.20 Corollary *Let T be an axiom system for a structure $\mathfrak{M}$. Then $|T| \leq 2^{\pi^{\mathfrak{M}}(T)}$.*

4.21 Remark Lemma 4.18 and Theorem 4.19 are reformulations of theorems which already occur in [**?**]. They can be sharpened to $\mathrm{otyp}_{\prec}(z) \leq \alpha$ and thus to $\mathrm{otyp}(\prec) \leq \mathrm{tc}(\mathit{Wf}(\prec))$, which has some impact for axiom systems whose Π^1_1–ordinal is not an ε–number.

For akzeptable structures with a sufficiently strong coding machinery we get $|T| = \pi^{\mathfrak{M}}(T)$ for axiom systems T which prove the properties of the coding machinery.

4.22 Theorem *Let T be an axiom system for a countable structure $\mathfrak{M}$ for which $\pi^{\mathfrak{M}}(T)$ is an ε–number. Assume that A is an $\mathscr{L}(\mathfrak{M})$–sentence with $\mathfrak{M} \models A$ and put $T' := T \cup \{A\}$. Then $|T'| = |T| = \pi^{\mathfrak{M}}(T)$.* □

[2] Cf. [**?**].

4.23 Exercise Show:

(Str) $\quad \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta\,, \alpha \leq \beta, \rho \leq \sigma, \Delta \subseteq \Gamma \;\Rightarrow\; \mathop{|\!\!\!-}\limits_{\sigma}^{\beta} \Gamma\,.$

($\bigwedge$–Inv) $\quad \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta, F$ and $F \in \bigwedge$–type $\;\Rightarrow\; \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta, G$ for every $G \in \mathrm{CS}(F)$.

($\vee$–Exp) $\quad \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta, F$, $F \in \bigvee$–type and $\mathrm{CS}(F)$ is finite $\;\Rightarrow\; \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta \cup \Gamma$ for $\Gamma = \mathrm{CS}(F)$.

□

4.24 Exercise
a) Show that we have $\mathfrak{M} \mathop{|\!\!\!-}\limits_{o}^{\alpha} F$ for any $\mathscr{L}(\mathfrak{M})$–sentenc F with $\mathrm{rnk}(F) = \alpha$.
b) Show that $\mathfrak{M} \models s = t$ and $\mathfrak{M} \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta(s)$ imply $\mathfrak{M} \mathop{|\!\!\!-}\limits_{\rho}^{\alpha} \Delta(t)$.
c) Show that there is an ordinal α such that $\mathfrak{M} \mathop{|\!\!\!-}\limits_{0}^{\alpha} s \neq t, \neg F(s), F(t)$ for all closed terms s and t. How large is α?

5 The standard model of arithmetic

The standard structure of arithmetic in the widest sense is the structure

$$(\mathbb{N}, \mathrm{Pow}(\mathbb{N}), \mathbb{N}^{\mathbb{N}^{<\omega}}).$$

Since $\mathrm{Pow}(\mathbb{N})$ and $\mathbb{N}^{\mathbb{N}}$ are uncountable this is a very big structure and there is no countable language that might match this strucuture. We will therefore restrict ourselves to a smaller structure be restricting the set of relations and functions and try to find a language that matches this structure.

5.1 Primitive recursive functions

By an arithmetical function we understand a function that maps tuples of natural numbers to a natural number.

5.1 Definition Let $\mathscr{PF}$ 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

$$Rec(g, h)(0, x_1, \ldots, x_n) = g(x_1, \ldots, x_n)$$
$$Rec(g, h)(Sy, x_1, \ldots, x_n) = h(y, Rec(g, h)(y, x_1, \ldots, x_n), x_1, \ldots, x_n)$$

5.2 Definition 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 $\mathscr{PR}$ denote the set of primitive recursive relations.
Cf. Figure 2 for a list of primitive–recursive functions and relations.

5.2 The standard structure $\mathfrak{N}$.

5.3 Definition We call the structure $\mathfrak{N} = (\mathbb{N}, \{0\}, \mathscr{PR}, \mathscr{PF})$ the *standard structure* of elementary arithmetic.

Observe that although we do not have a constant for every natural number there is a name for every $n \in \mathbb{N}$. To obtain a name $\underline{n}$ for $n > 0$ we just apply the successor function n-fold to 0. We call these names *numerals*.

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

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

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

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

is primitive–recursive. □

5.7 Exercise Verify the open facts in Figure 2

5.8 Exercise Let ω_1^{CK} be the first ordinal which is not the order–type of a primitive–recursive order–relation on the natural number.
Sketch a proof of $\pi^{\mathfrak{N}} = \omega_1^{CK}$. □

6 The axiom system NT

6.1 Peano arithmetic

The signature for the language of Peano arithmetic is $P := (\{\underline{0}, \underline{1}\}, \{=, \neq\}, \{+, \cdot\})$. Then $\mathfrak{N}$ clearly interprets $\mathscr{L}(P)$. The non–logical axioms of Peano arithmetic comprise the succesor axioms

$$(\forall x)[x + 1 \neq 0] \text{ and } (\forall x)(\forall y)[x + 1 = y + 1 \to x = y],$$

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	$a \dot{-} 0 = 0, \ \ a \dot{-} S(x) = Pd(a \dot{-} x)$
$\|a - x\|$	absolute value	$(a \dot{-} x) + (x \dot{-} a)$
$\mathcal{I} := \{(x, x) \mid x \in \omega\}$	identity	$\chi_{\mathcal{I}} = \overline{sg}(\|x - y\|)$
$x \leq y, x < y$	less than (or equal to)	$(\exists z < y)[x + z = y], x \leq y \wedge x \neq y$
$\max\{a, b\}$	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	$x \neq 0 \wedge (\forall z < S(x))[z = 1 \vee z = x \vee \neg(z/x)]$
$pn(n)$	Enumeration of the primes	$pn(n) = \begin{cases} 2 & \text{if } n = 0 \\ \mu x < pn(n)! + 2.\ [Prime(x) \wedge pn(n) < x] & \text{if } n > 0 \end{cases}$
$\langle x_0, \ldots, x_n \rangle$	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	$s = 0 \vee (\forall i < S(s))[\neg(pn(S(i))/s) \vee pn(i)/s]$
a belongs to a finite set	$a \in \{a_1, \ldots, a_n\}$	$a = a_1 \vee \cdots \vee a = a_n$

Figure 2: Some primitive–recursive functions and relations

the defining equations

$$(\forall x)[x+\underline{0}=x] \text{ and } (\forall x)(\forall y)[(x+(y+\underline{1})=(x+y)+\underline{1})]$$

for addition and

$$(\forall x)[x\cdot 1=x] \text{ and } (\forall x)(\forall y)[(x\cdot(y+1))=x\cdot y+x]$$

for multiplication and the scheme

$$F(\underline{0})\wedge(\forall x)[F(x)\to F(x+1)]\to(\forall x)F(x)$$

of mathematical induction where F is an arbitrary $\mathscr{L}(P)$–formula.

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.

6.2 Pure logic

To fix the logical framework we introduce a Hilbert style formal system 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*.

6.1 Definition The Boolean atoms of a first order formula F are the subformulae 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.

6.2 Definition The logical axioms of the Hilbert calculus are:

(BOOLE) All Boolean valid formulae

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

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

The identity axioms are

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

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

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

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

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.

6.3 Theorem *a) A formula F is logically valid, i.e., true in any model under any assignment, iff $\vdash F$.*
b) $A_1,\dots,A_n \models F$ iff the formula $A_1 \land \cdots \land A_n \to F$ is logically valid.

6.4 Lemma *a) For any Boolean valid formula $F(x_1,\dots,x_n)$ there is a finite ordinal k such that $\mathfrak{M} \mid\!\frac{k}{0}\, F(\underline{z}_1,\dots,\underline{z}_n)$ holds true for any countable structure $\mathfrak{M}$ for first order predicate logic and any tuple $z_1,\dots,z_n$ of of elements of M.* □

6.5 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 $\mathfrak{M} \mid\!\frac{m}{r}\, F(\underline{z}_1,\dots,\underline{z}_n)$ holds true for any structure $\mathfrak{M}$ that interprets first order predicate laogic and any tuple $z_1,\dots,z_n$ of elements of M..* □

6.3 The axioms of arithmetic

6.6 Definition The non–logical axioms of NT comprise

(MATHAX) All true atomic $\mathscr{L}(\mathfrak{N})$–sentences

(MATHIND) The scheme $F(\underline{0}) \land (\forall x)[F(x) \to F(S(x))] \;\to\; (\forall x)F(x)$ of *mathematical induction*, where $F(\underline{z})$ is any $\mathscr{L}(\mathfrak{N})$–formula.

6.7 Theorem *We have $\mathsf{NT} \vdash F$ for any $\mathscr{L}(\mathfrak{N})$–formula F iff there there are finitely many axioms $A_1,\dots,A_n$ of NT such that $A_1 \land \cdots \land A_n \to F$ is logically valid.*

7 The upper bound

7.1 Embedding of NT

7.1 Lemma *For every natural number n and $\mathscr{L}(\mathfrak{N})$–formula $F(x)$ we have*

$$\mathfrak{N} \mid\!\frac{2(\mathrm{rnk}(F)+n)}{0}\, \neg F(\underline{0}), \neg(\forall x)[F(x) \to F(S(x))], F(\underline{n})\,.$$

□

7.2 Theorem *(Induction Theorem) We have*

$$\mathfrak{N} \mid\!\frac{\omega+3}{0}\; F(\underline{0}) \wedge (\forall x)[F(x) \to F(S(x))] \;\;\to\;\; (\forall x)F(x)\,.$$

7.3 Theorem *(Embedding Theorem) Let $F(x_1, \ldots, x_n)$ be a first order formula whose free individual variables occur all in the list $x_1, \ldots, x_n$. Then $\mathsf{NT} \vdash F(x_1, \ldots, x_n)$ implies that there is a finite ordinal r such that $\mathfrak{N} \mid\!\frac{\omega+\omega}{r}\; F(\underline{z}_1, \ldots, \underline{z}_n)$ holds true for every tuple $z_1, \ldots, z_n$ of natural numbers.* □

7.2 Cut elimination

7.4 Lemma *(Reduction Lemma) Assume $\mathfrak{N} \mid\!\frac{\alpha}{\rho}\; \Delta \cup \{F\}$ and $\mathfrak{N} \mid\!\frac{\beta}{\rho}\; \Gamma \cup \{\neg F\}$ and $\mathrm{rnk}(F) = \rho$ for $F \in \bigwedge\text{–type}$ or $F \notin \bigwedge\text{–type} \cup \bigvee\text{–type}$. Then we obtain $\mathfrak{N} \mid\!\frac{\alpha+\beta}{\rho}\; \Delta \cup \Gamma$.* □

7.5 Theorem *(Elimination Theorem) $\mathfrak{N} \mid\!\frac{\alpha}{\rho+1}\; \Delta$ implies $\mathfrak{N} \mid\!\frac{\omega^\alpha}{\rho}\; \Delta$.* □

7.3 The upper bound

7.6 Theorem *If $\mathsf{NT} \vdash F$ for $F \in \mathscr{L}(\mathfrak{N})$ then $\mathrm{tc}(F) < \varepsilon_0$.* □

7.7 Corollary *We have $\pi^{\mathfrak{N}}(\mathsf{NT}) \leq \varepsilon_0$, hence also $|\mathsf{NT}| \leq 2^{\varepsilon_0} = \varepsilon_0$.*

7.8 Corollary *The theory NT is consistent.*

8 The lower bound

8.1 Ordinal notations

8.1 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 $\alpha > \alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_n$.*

8.2 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.*

8.2 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:

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

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

$$\alpha \subseteq X \quad :\Leftrightarrow \quad (\forall \xi)[\xi < \alpha \rightarrow \xi \in X]$$
$$\alpha \in \mathcal{J}(X) \quad :\Leftrightarrow \quad (\forall \xi)[\xi \subseteq X \rightarrow \xi + \omega^\alpha \subseteq X].$$
$$TI(\alpha) \quad :\Leftrightarrow \quad Prog(X) \rightarrow \alpha \subseteq X.$$

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

8.3 Lemma $\mathsf{NT} \vdash Prog(X) \rightarrow Prog(\mathcal{J}(X))$. □

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

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

So we have $\varepsilon_0 \leq |\mathsf{NT}| \leq 2^{\pi^{\mathfrak{N}}(\mathsf{NT})} \leq 2^{\varepsilon_0} = \varepsilon_0$ which yield the next theorem.

8.6 Theorem $|\mathsf{NT}| = \pi^{\mathfrak{N}}(\mathsf{NT}) = \varepsilon_0$. □

8.7 Theorem *There is am $\mathscr{L}(\mathfrak{N})$–sentence $(\forall x)F(x, X)$ such that $\mathfrak{N} \models (\forall X)(\forall x)F(x, X)$, $\mathsf{NT} \vdash F(\underline{n}, X)$ for any natural number n but $\mathsf{NT} \nvdash (\forall x)F(x, X)$.* □

Part I

Selected proofs and solutions

Section 3

1 *Proof of Lemma 3.21.*

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 $\beta = \alpha + \xi_0'$. 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 $\xi = \sup_{\eta<\beta} \xi_\eta$. □

2 *Proof of Lemma 3.22.*

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 $\eta < \omega^\alpha < \xi$. □

3 *Proof of Lemma 3.24*

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 $\eta \le \varepsilon_0$ and get $\omega^\eta = \sup\{\omega^{\omega^{(n)}(0)} \,|\, n \in \omega\} = \sup\{\omega^{(n+1)}(0) \,|\, n \in \omega\} = \eta$. Hence $\varepsilon_0 \le \eta$. □

4 *Solution to Exercise 3.27*

Since $\alpha \in Lim$ we have $\alpha \le \xi + \alpha = \sup\{\xi + \eta \,|\, \eta < \alpha\} \le \alpha$. □

5 *Solution to Exercise 3.28*

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 3.27 for k the index such that $\xi_k \ge \eta_1$ and $\xi_{k+1} < \eta_1$. □

Section 4

6 *Proof of Theorem 4.10*

By a straightforward induction on α we get that $\mathfrak{M} \,\overset{\alpha}{\underset{0}{|}}\, \Delta$ implies $\mathfrak{M} \models \bigvee \Delta[\Phi]$ for any assignment $\Phi(X) \subseteq M$. So soundness is straightforward. □

7 *Proof of Theorem 4.11* More difficult is completeness. Since there are no free individual variables in $\mathscr{L}(\mathfrak{M})$–formulae every closed term t has a value $t^{\mathfrak{M}} =: z \in M$. Since $\mathfrak{M} \models t^{\mathfrak{M}} = z$ we may w.l.o.g. replace all terms t by the constant $\underline{z}$ for their value.

Let Δ be a finite sequence of $\mathscr{L}(\mathfrak{M})$–sentences. The set Δ is reducible if it contains a sentence in $\bigwedge$–type $\cup \bigvee$–type. The first sentence in Δ which is in $\bigwedge$–type $\cup \bigvee$–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.[3]

[3] Observe that for the structure $\mathfrak{N}$ the search tree S_Δ is primitive–recursively defined.

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

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

(Irr) If $\delta(s)$ is irreducible then $s^\frown\langle 0\rangle \in S_\Delta$ and $\delta(s^\frown\langle 0\rangle) = \delta(s)$.

($\bigwedge$) If $F :\equiv R(\delta(s)) \in \bigwedge$–type and $\mathrm{CS}(F) = \big\langle F_i \,|\, i \in I\big\rangle$ then $s^\frown\langle i\rangle \in S_\Delta$ and $\delta(s^\frown\langle i\rangle) = \delta(s)^r, F_i$.

($\bigvee$) If $F :\equiv R(\delta(s)) \in \bigvee$–type then $s^\frown\langle 0\rangle \in S_\Delta$ and $\delta(s^\frown\langle 0\rangle) = \delta(s)^r, G, F$ 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. Otherwise put $\delta(s^\frown\langle 0\rangle) = \delta(s)^r, F$.

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

$$\textit{If } S_\Delta \textit{ is well–founded then we have } \overset{|s|}{\underset{0}{|}}\, \delta(s) \textit{ for any node } s \in S_\Delta. \tag{i}$$

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

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

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

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

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

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

If $G \equiv R(z_1, \ldots, z_k)$ then there is a node $s = \langle f(0), \ldots, f(m-1)\rangle =: \overline{f}(m)$ such that $R(\delta(s)) = G$. Then $G \in \bigwedge$–type implies that s is a leave which contradicts the infinity of f. Hence $G \in \bigvee$–type which implies $\mathfrak{M} \not\models G$.

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

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

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

8 *Proof of Lemma 4.18*

We prove by induction on α:

$$\mathfrak{M} \overset{\alpha}{\underset{0}{|}} \neg \mathit{Prog}(X, \prec), \underline{n}_1 \notin X, \ldots, \underline{n}_k \notin X, \Delta \;\Rightarrow \mathfrak{M} \models \bigvee \Delta[\prec\restriction\beta] \tag{i}$$

for a finite set Δ of X–positive formulae 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

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

we have the premise

$$\mathfrak{M} \mathrel{\big|\!\!\frac{\alpha_0}{0}} \neg \mathit{Prog}(X, \prec), (\forall y)[y \prec \underline{z} \rightarrow y \in X] \wedge \underline{z} \notin X, \underline{n}_1 \notin X, \ldots, \underline{n}_k \notin X, \Delta \quad \text{(ii)}$$

for some constant z and some ordinal $\alpha_0 < \alpha$. By $\wedge$–inversion we thus have

$$\mathfrak{M} \mathrel{\big|\!\!\frac{\alpha_0}{0}} \neg \mathit{Prog}(X, \prec), (\forall y)[y \prec \underline{z} \rightarrow y \in X], \underline{n}_1 \notin X, \ldots, \underline{n}_k \notin X, \Delta \quad \text{(iii)}$$

and

$$\mathfrak{M} \mathrel{\big|\!\!\frac{\alpha_0}{0}} \neg \mathit{Prog}(X, \prec), \underline{z} \notin X, \underline{n}_1 \notin X, \ldots, \underline{n}_k \notin X, \Delta\,. \quad \text{(iv)}$$

Towards a contradiction assume

$$\mathfrak{M} \not\models \bigvee \Delta[\prec \restriction \beta].$$

Then we also have $\mathfrak{M} \not\models \bigvee \Delta[\prec \restriction \beta_0]$ 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 $(\forall y)[y \prec z \rightarrow \mathrm{otyp}_\prec(y) < \beta_0]$. i. e., $\mathrm{otyp}_\prec(z) \leq \beta_0$. By induction hypothesis for (iv) we thus get $\mathfrak{M} \models \bigvee \Delta[\prec \restriction \beta_1]$ for

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

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

□

9 *Proof of Theorem 4.22*

From $T' \mathrel{|\!\!-} \mathit{Wf}(\prec, X)$ we get $T \mathrel{|\!\!-} A \rightarrow \mathit{Wf}(\prec, X)$, Since this an $\mathscr{L}(\mathfrak{M})$–sentence there is an ordinal α less than $\pi^{\mathfrak{M}}(T)$ such that $\mathfrak{M} \mathrel{\big|\!\!\frac{\alpha}{0}} \{\neg A \vee \mathit{Wf}(\prec, X)\}$ which entails

$$\mathfrak{M} \mathrel{\big|\!\!\frac{\alpha}{0}} \{\neg A, \neg(\forall x)[(\forall y)[y \prec x \rightarrow y \in X], (\forall x)[x \in \mathit{field}(\prec) \rightarrow x \in X]]\}$$

By Lemma 4.18 we thus get $\mathfrak{M} \models \neg A \vee (\forall x)[x \in \mathit{field}(\prec) \rightarrow x < \alpha]$. Since $\mathfrak{M} \models A$ this entails $\mathrm{otyp}(\prec) \leq 2^{\alpha} < 2^{\pi^{\mathfrak{M}}(T)} = \pi^{\mathfrak{M}}(T)$. Hence $|T| \leq |T'| \leq \pi^{\mathfrak{M}}(T) = |T|$.

□

Section 5

10 *Solution to Exercise 5.4*

If

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

for pairwise disjoint primitive–recursive predicates R_i then put

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

11 *Solution to Exercise 5.5*

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 $\prod_{i=0}^{S(n)} f(i,\vec{x}) = \prod_{i=0}^{n} f(i,\vec{x})\cdot f(S(n),\vec{x})$

12 *Solution to Exercise 5.6*

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

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

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

13 *Sketch of Exercise 5.8*

$\mathfrak{N}\models F$ for an $\mathscr{L}(\mathfrak{N})$–sentence F entails that the search tree for F is well–founded. Hence $\mathfrak{N}\models^{|s|} F$. Since the search tree is primitive–recursively definable it is $|s| < \omega_1^{CK}$. So $\pi^{\mathfrak{N}} \leq \omega_1^{CK}$. For $\alpha < \omega_1^{CK}$ there is primitive–recursive well–ordering $\prec$ such that $\mathrm{otyp}(\prec) = \alpha$. We then have $\mathfrak{N}\models^{\beta} \mathit{Wf}(\prec,X)$ for some ordinal $\beta < \pi^{\mathfrak{N}}$. Hence $\mathfrak{N}\,\vert\!\frac{\beta}{0}\, \{\neg(\forall x)[(\forall y)[y\prec x\to y\in X]\to x\in X], (\forall x\in \mathit{field}(\prec))[x\in X]\}$. By the Boundedness Theorem (Theorem 4.19) we thus get $\alpha\leq 2^{\beta} < \pi^{\mathfrak{N}}$. □

Section 6

14 *Proof of Lemma 6.4*

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

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

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

- A formula F is Boolean valid iff $\bigvee\Delta(F)$ is Boolean valid.
- If $\mathit{Bdeg}(F) = 0$ and F is Boolean valid then $\Delta(F) = \Delta_0, A, \neg A$ for a Boolean atom A.
- If $\mathit{Bdeg}(F) > 0$ then $\Delta(F) = \Delta_0, A \wedge B$ for some formulae 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 4.9 by induction on $\mathit{Bdeg}(\Delta(F))$. □

15 *Proof of Theorem 6.5.*

We prove the theorem by induction on the length of a derivation in the Hilbert calculus. If $F(x_1, \ldots, x_n)$ is Boolean valid so is $F(\underline{z}_1, \ldots, \underline{z}_n)$ and we obtain $\vdash^{\alpha}_{0} F(\underline{z}_1, \ldots, \underline{z}_n)$ by Lemma 6.4.

For a term t let $t_0 := t_{x_1,\ldots,x_n}(\underline{z}_1, \ldots, \underline{z}_n)$ and $z := t_0^{\mathfrak{M}}$. Let $F \equiv (\forall x)G(x_1, \ldots, x_n, x)$ and $G_0 :\equiv G_{x_1,\ldots,x_n}(\underline{z}_1, \ldots, \underline{z}_n, x)$. By Lemma 4.9 we have $\mathfrak{M} \vdash^{\alpha}_{0} \neg G_0(\underline{z}), G_0(\underline{z})$ for $\alpha = 2 \cdot \mathrm{rnk}(F) < \omega$. Since $\mathfrak{M} \models \underline{z} = t_0$. this implies $\mathfrak{M} \vdash^{\alpha}_{0} \neg G_0(\underline{z}), G_0(t_0)$ which by an inference $(\bigvee)$ entails $\mathfrak{M} \vdash^{\alpha+2}_{0} \neg(\forall x)G_0(x) \vee G_0(t_0)$.

Symmetrically we obtain
$\mathfrak{M} \vdash^{\alpha+2}_{0} \neg G_0(t_0) \vee (\exists x)G_0(x)$.

Since $z = z$ is true atomic we obtain $\vdash^{0}_{0} \underline{z} = \underline{z}$ for all $z \in M$ hence $\mathfrak{M} \vdash^{1}_{0} (\forall x)[x = x]$ by $(\bigwedge)$ with empty premise.

Similarly we obtain $\vdash^{0}_{0} \underline{s} \neq \underline{t}, \underline{t} = \underline{s}$ for all elements $s, t \in M$ by an inference $(\bigwedge)$ with empty premises. Hence

$$\mathfrak{M} \vdash^{4}_{0} (\forall x)(\forall y)[x \neq y \vee y = x]$$

by two ($\bigvee$) and two inferences ($\bigwedge$).

Similarly we have $\left|\frac{0}{0}\right. r \neq s, r \neq t, s = t$ by ($\bigwedge$) and obtain $\left|\frac{7}{0}\right. (\forall x)(\forall y)(\forall z)[x \neq y \vee y \neq z \vee x = z]$ by four ($\bigvee$)– and three ($\bigwedge$)–inferences.

An easy induction on $\text{rnk}(F(x))$ shows $\left|\frac{2\cdot\text{rnk}(F)}{0}\right. s \neq t, \neg F(s), F(t)$. 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. □

16 *Proof of Lemma 7.1 by induction on n.*

Let $\beta := 2 \cdot \text{rnk}(F)$. By Lemma 4.9 we have

$$\mathfrak{N} \left|\frac{\beta}{0}\right. \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(0)\,. \tag{i}$$

By induction hypothesis we have

$$\mathfrak{N} \left|\frac{\beta+2n}{0}\right. \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(n) \tag{ii}$$

and by Lemma 4.9

$$\mathfrak{N} \left|\frac{\beta}{0}\right. \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], \neg F(S(n)), F(S(n))\,. \tag{iii}$$

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

$$\mathfrak{N} \left|\frac{\beta+2n+1}{0}\right. \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(n) \wedge \neg F(S(n)), F(S(n)) \tag{iv}$$

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

$$\mathfrak{N} \left|\frac{\beta+2(n+1)}{0}\right. \neg F(0), \neg(\forall x)[F(x) \to F(S(x))], F(S(n))\,. \qquad \square$$

17 *Proof of Theorem 7.3.*

If $\mathsf{NT} \vdash F(x_1, \ldots, x_n)$ there are finitely many axioms $A_1, \ldots, A_m$ in NT such that $\vdash A_1 \wedge \cdots \wedge A_m \to F(x_1, \ldots, x_n)$. By Theorem 6.5 we therefore find finite ordinals α and ρ such that $\mathfrak{N} \left|\frac{\alpha}{\rho}\right. \{\neg A_1, \ldots, \neg A_m, F(\underline{z}_1, \ldots, \underline{z}_n)\}$ for any tuple $z_1, \ldots, z_n$ of natural numbers. By Exercise 4.24 a) we obtain $\mathfrak{N} \left|\frac{\alpha_i}{0}\right. \{A_i\}$ for $\alpha_i = \text{rnk}(A_i) < \omega$ for every axiom in NT which is not an instance of the scheme of Mathematical Induction. For every instance A_j of Mathematical Induction we obtain $\mathfrak{N} \left|\frac{\omega+3}{0}\right. \{A_j\}$ by Theorem 7.2. Since all the formulae in $\mathscr{L}(\mathfrak{N})$ have finite rank we obtain the claim by a series of cut. □

Section 7

18 *Proof of Lemma 7.4*

We induct on β.

If $\neg F$ is not the main–formula of the last inference leading to $\mathfrak{N} \mid\frac{\beta}{\rho} \Gamma \cup \{\neg F\}$ then we either have $\mathfrak{N} \mid\frac{\alpha}{\rho} \Gamma$ —and obtain the claim trivially—or we have premise(s) $\mathfrak{N} \mid\frac{\beta_\iota}{\rho} \Gamma_\iota \cup \{\neg F\}$ for $\beta_\iota < \beta$ and obtain by induction hypothesis

$$\mathfrak{N} \mid\frac{\alpha+\beta_\iota}{\rho} \Delta \cup \Gamma_\iota \,.$$

This entails $\mathfrak{N} \mid\frac{\alpha+\beta}{\rho} \Delta \cup \Gamma$ by the same inference.

So assume that $\neg F$ is the main–formula of the last inference. If $\rho = \mathrm{rnk}(F) = 0$ then $F \notin \mathscr{L}_1(\mathfrak{N})$ since $\neg F \in \bigvee$–type and only atomic $\mathscr{L}_1(\mathfrak{N})$–sentences in $\bigwedge$–type are allowed as main–formulae of inferences. Therefore $\neg F$ is a formula $t \in X$ or $t \notin X$. W.l.o.g. we assume the former. Since F is the main–formula there is a formula $s \notin X$ in Γ such that $\mathfrak{N} \models s = t$. We therefore have $\mathfrak{N} \mid\frac{0}{0} \Gamma \cup \{t \in X\}$. We now show by induction on α

$$\mathfrak{N} \mid\frac{\alpha}{0} \Delta \cup \Gamma \,. \tag{i}$$

If $\mathfrak{N} \mid\frac{\alpha}{0} \Delta \cup \{t \notin X\}$ holds according to the (X)–rule, then there is formula $r \in X$ in Δ such that $\mathfrak{N} \models r = t$. Hence $\mathfrak{N} \models r = s$ and we obtain $\mathfrak{N} \mid\frac{0}{0} \Delta \cup \Gamma$ by an (X)–rule. Otherwise we have the premises $\mathfrak{N} \mid\frac{\alpha_\iota}{0} \Delta_\iota \cup \{t \notin X\}$ and obtain $\mathfrak{N} \mid\frac{\alpha_\iota}{0} \Delta_\iota \cup \Gamma$ by induction hypothesis and finally $\mathfrak{N} \mid\frac{\alpha}{0} \Delta \cup \Gamma$ by the same inference.

So assume $\rho > 0$. Since $\neg F \in \bigvee$–type the last inference has the form

$$\mid\frac{\beta_0}{\rho} \Gamma, \neg F, \neg G \;\Rightarrow\; \mid\frac{\beta}{\rho} \Gamma, \neg F \tag{ii}$$

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

19 *Proof of Theorem 7.5.*

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

$$\mathfrak{N} \mid\frac{\alpha_0}{\rho+1} \Delta \cup \{F\}\,, \;\; \mathfrak{N} \mid\frac{\alpha_0}{\rho+1} \Delta \cup \{\neg F\} \;\Rightarrow \mathfrak{N} \mid\frac{\alpha}{\rho+1} \Delta$$

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

$$\mathfrak{N} \mid\frac{\omega^{\alpha_0}}{\rho} \Delta \cup \{F\}\,, \;\; \mathfrak{N} \mid\frac{\omega^{\alpha_0}}{\rho} \Delta \cup \{\neg F\}$$

and, since $\omega^{\alpha_0} + \omega^{\alpha_0} < \omega^{\alpha}$, obtain $\mathfrak{N} \mathrel{\big|\!\frac{\omega^\alpha}{\rho}} \Delta$ by the Reduction Lemma . $\square$

20 *Proof of Theorem 7.6*

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

21 *Proof of Corollary 8.2*

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

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

The set *On*:

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

The $\prec$–relation on *On*.

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

22 *Proof of Lemma 8.3*

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 $\nu \in X$. □

23 *Proof of Lemma 8.4*

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 8.3 we thus get $Prog(X) \to \omega^{\alpha} \subseteq X$. □

24 *Proof of Theorem 8.5.*

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 8.4. (This time it is an induction from outside). □

25 *Proof of Theorem 8.7*

We have $\mathfrak{N} \models (\forall X)(\forall x)[Prog(X) \wedge x \in On \to x \in X]$ and NT $\vdash Prog(X) \wedge \underline{n} \in On \to \underline{n} \in X$ by (the proof of) Theorem 8.5. But NT $\vdash (\forall x)[Prog(X) \wedge x \in On \to x \in X]$ would imply $\mathrm{otyp}(\prec) < \varepsilon_0$ by Theorem 7.6 while $\mathrm{otyp}(\prec) = \varepsilon_0$ holds true by the construction of the relation $\prec$. Contradiction! □

These handouts will be continued during the Sommer–school.