Lecture Notes in Mathematics

Wolfram Pohlers

Proof Theory

1407

An Introduction

$$\mathbf{PA} \not\vdash \mathbf{TI}(\varepsilon_0, F)$$

$$\mathbf{ID}_1 \not\vdash \mathbf{TI}(\psi\varepsilon_{\Omega+1}, F)$$

Springer

Lecture Notes in Mathematics

Edited by A. Dold, B. Eckmann and F. Takens

1407

Wolfram Pohlers

Proof Theory

An Introduction

Springer-Verlag
Berlin Heidelberg New York London Paris Tokyo Hong Kong

Author

Wolfram Pohlers
Institut für mathematische Logik und Grundlagenforschung
Einsteinstr. 62, 4400 Münster, West Germany

Mathematics Subject Classification (1980): 03 F

ISBN 3-540-51842-8 Springer-Verlag Berlin Heidelberg New York
ISBN 0-387-51842-8 Springer-Verlag New York Berlin Heidelberg

Printed in Germany

Printing and binding: Druckhaus Beltz, Hemsbach/Bergstr.
2146/3140-543210 – Printed on acid-free paper

Preface

This book contains the somewhat extended lecture notes of an introductory course in proof theory I gave during the winter term 1987/88 at the University of Münster, FRG. The decision to publish these notes in the Springer series has grown out of the demand for an introductory text on proof theory. The books by K.Schütte and G.Takeuti are commonly considered to be quite advanced and J.Y.Girard's brilliant book also, is too broad to serve as an introduction.

I tried, therefore, to write a book which needs no previous knowledge of proof theory at all and only little knowledge in logic. This is of course impossible, so the book runs on two levels - a very basic one, at which the book is self-contained, and a more advanced one (chiefly in the exercises) with some cross-references to definability theory. The beginner in logic should neglect these cross-references.

In the presentation I have tried not to use the 'cabal language' of proof theory but a language familiar to students in mathematical logic.

Since proof theory is a very inhomogeneous area of mathematical logic, a choice had to been made about the parts to be presented here. I have decided to opt for what I consider to be the heart of proof theory - the ordinal analysis of axiom systems. Emphasis is given to the ordinal analysis of the axiom system of the impredicative theory of elementary inductive definitions on the natural numbers. A rough sketch of the 'constructive' consequences of ordinal analysis is given in the epilogue.

Many people helped me to write this book. *J.Columbus* suggested and checked nearly all the exercises. *A.Weiermann* made a lot of valuable suggestions especially in the section about alternative interpretations for Ω. *A.Schlüter* did the proof-reading, drew up the subject index and the index of notations and suggested many corrections especially in the part about the autonomous ordinals of $\mathbf{Z}_\infty$.

I am also indebted to the students of the workshop on proof theory in Münster who suggested many more corrections. Last but not least I want to thank all the students attending my course of lectures during the winter term 1987/88. It was their interest in the topic that encouraged me to write this book.

A first version of the typescript was typed by my secretary Mrs. *J.Pröbsting* using the Signum text system. She also wrote the table of contents. Many thanks to all these persons.

July 19, 1989 W. P.

Münster

TABLE OF CONTENTS

INTRODUCTION

The history of proof theory begins with the foundational crisis in the first decades of our century. At the turn of the century, as a reaction to the explosion of mathematical knowledge in the last two centuries, endeavours began to provide the growing body of mathematics with a firm foundation. Some of the notions used then seemed to be quite problematic. This was especially true of those which somehow depended upon that of infinity. On the one hand there was the notion of infinitesimals which embodied 'infinity in the small'. The elimination of infinitesimals by the introduction of limit processes represented a great progress in foundational work (although one may again find a justification for infinitesimals as it is done today in the field of nonstandard analysis). But on the other hand there were also notions which, at least implicitly, depended on 'infinity in the large'. *G.Cantor* in his research about trigonometrical series was repeatedly confronted with such notions. This led him to develop a completely new mathematical theory of infinity, namely set theory. The main feature of set theory is the comprehension principle which allows to form collection of possibly infinitely many objects (of the mathematical universe) as a single object. Cantor called the objects of the mathematical universe 'Mengen' usually translated by 'sets'. Set theory, however, soon turned out to be a source of doubt itself. Since Cantor's comprehension principle allows the collection of all sets x sharing an arbitrary property E(x) into the set $\{x: E(x)\}$ one easily runs into contradictions.[1] For instance if we form the set $M := \{x: x \notin x\}$, then we obtain the well-known Russellian antinomy: $M \in M$ if and only if $M \notin M$. It is easy to construct further antinomies of a

1) Cantor himself was well aware of the distinction between sets and other collections which may lead to contradictions. See his letter to Dedekind from 27.7.1899 [Purkert et al. 1987]

similar sort. Another annoying fact was that the plausible looking axiom of choice

(AC) For any family $(S_k)_{k \in I}$ of non empty sets there is a choice function $f: I \to \bigcup\{S_k : k \in I\}$ such that $f(k) \in S_k$ for all $k \in I$

had as a consequence the apparently paradoxical possibility of wellordering any set. Nobody could imagine what a wellordering of the reals could look like and D.Hilbert, in his famous list of mathematical problems presented in Paris in 1900, stated in his remarks concerning problem one (the Continuum Hypothesis) that it would be extremely desirable to have a direct proof of this mysterious statement.

Today we know that there is no elementary construction of a wellordering of the reals. Any wellordering of the reals has the same degree of constructiveness as the choice function itself. The existence of a choice function, however, is not even provable from the Zermelo Fraenkel axioms for set theory.

All these facts contributed to a feeling of uncertainty among members of the mathematical society about the notion of a set that they were opposed to set theory in general. But it was of course not possible to simply ignore Cantor's discoveries. *Hermann Weyl* in his paper 'Über die neue Grundlagenkrise der Mathematik' [Weyl 1921] tried to convince his contemporaries that the foundational problems arising in set theory were not just exotic phenomena of an isolated branch of mathematics but also concerned analysis, the very heart of mathematics. It was he who introduced the term 'foundational crisis' into the discussion. In his book 'Das Kontinuum' [Weyl 1918] he had already suggested a development of mathematics which avoided the use of unrestricted set constructions. In more modern terms one could say that he proposed a predicative development of mathematics. Others, like L.E.J.Brouwer, already doubted the logical basis of mathematics. Their point of attack was the law of the excluded middle. With the help of the law of the excluded middle it becomes possible to prove the existence of objects without constructing them explicitly. Brouwer suggested developing mathematics on the basis of alternative intuitive principles which excluded the law of the excluded middle. Their formalization – due to *Heyting* – now is known as intuitionistic logic. Both approaches, Weyl's as well as Brouwer's, meant rigid restrictions on mathematics. *D.Hilbert*, then one of the most prominent mathematicians, was not willing to accept any foundation of mathematics which would mutilate existing mathematics. To him the foundational crisis was a nightmare haunting mathematics. In his opinion mathematics was *the* science, the model for all sciences, whose 'truths had been proven on the basis of definitions via infallible inferences' and therefore were 'valid overall in reality'. He felt that this position of mathematics was in danger and therefore wanted to preserve it as it was. He was especially unwilling to give up Cantor's set theory, a paradise from which no one would expel him. In his opinion Cantor's treatment of transfinite ordinals was one of the supreme achievements of human thought. Therefore he planned a program to save mathematics in its existing form. He charted his program in a couple of writings and debated it in several talks (cf. [Hilbert 1932–1935]). Therefore it would be inadequate to try to sketch Hilbert's program in only a few sentences. For a serious evaluation of the status of Hilbert's program today deeper considerations are necessary (cf. JSL 53 (1988)). The part of Hilbert's program, however, which was essential for the development of the kind of proof theory we want to give an introduction to in this lecture may be roughly characterized by the following steps:

I. Axiomatize the whole of mathematics

II. Prove that the axioms obtained in step I are consistent.

Hilbert proposed that step II of his program, the consistency proof, should be carried out within a new mathematical theory which he called '*Beweistheorie*', i.e. *Proof Theory.* According to Hilbert, proof theory should use contentual reasoning in contrast to the formal inferences of mathematics. Hilbert himself was aware of the fact that the reasoning of proof theory must itself not become the subject of criticism. He therefore required proof theory to obtain its results by methods beyond the shadow of a doubt. He suggested using only finitistic methods. By finitistic methods he understood those methods 'without which neither reasoning nor scientific action are possible'. In my personal opinion, finitistic reasoning may be interpreted as combinatorial reasoning over finite domains. Some of Hilbert's students (e.g. Ackermann, J.v.Neumann, P.Bernays) soon obtained concrete results. Following Hilbert's maxim of first developing the mathematical tools necessary for the solution of a general problem by studying special cases of the problem they first tackled subsystems of elementary arithmetic. In fact they succeeded in obtaining consistency proofs for subsystems not containing the scheme of complete induction. It thus seemed to be just a matter of technical refinement to extend these consistency proofs to systems containing the full induction scheme. However, the systems containing complete induction stubbornly resisted all attempts to prove their consistency. That this failure was neither an accident nor was due to the incompetence of the researchers, became clear after the publication of *Kurt Gödel's* paper 'Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme' [Gödel 1931]. In this paper Gödel proved his famous theorems which, roughly speaking, say the following:

I. In any formal system, satisfying certain natural requirements, it is possible to formulate sentences which are true in the intended structure but are also undecidable within the formal system (i.e. neither the sentence nor its negation are provable in the formal system).

II. The consistency proof for any formal system, again satisfying canonical requirements, may not be formalized in the system itself.

One might think that Gödel's theorems meant a sudden end to Hilbert's program. The first theorem shows that step I in Hilbert's program is indeed impossible. This, however, might be remedied by the oberservation that in fact it is not necessary to formalize all possible mathematics. It would suffice just to axiomatize existing mathematics. Today we know that nearly everything in everyday's mathematics (and, except for the Continuum Hypothesis, probably all which Hilbert may have thought of) is formalizable in one single formal system, namely Zermelo Fraenkel set theory with the axiom of choice (ZFC). Most parts are even formalizable in much weaker systems. Gödel II, however, is a lethal blow to Hilbert's program. Since the methods 'without which neither reasoning nor scientific action are possible' (combinatorial reasoning over finite domains, in our interpretation) should itself be available in mathematics, any reasonable axiomatization of mathematics should allow the formalization of Hilbert's finitistic methods. Therefore there is no finitistic consistency proof for an axiomatization of stronger fragments of mathematics (i.e. essentially those containing the scheme of complete induction). Luckily for the development of proof theory, the researchers in the thirties did not interpret these results als having such drastic consequences. It is hard to say why. Gödel's results were known to the Hilbert school. For instance Bernays mentions them in [Bernays 1935a] but although he expresses doubts about the feasibility of finitistic consistency proofs he denies that Gödel's results imply their impossibility. I conjecture that the true reasons were Hilbert's authority as well as the vagueness of his program. Since he gave no precise definition of what he meant by finitistic methods one could hope that these methods comprised a kind of contentual reasoning which cannot be mathematically formalized. As a matter of fact mathematicians did not stop searching for consistency proofs and in 1936 *Gerhard Gentzen* succeeded in proving the consistency of elementary number theory. According to Gödel's second theorem Gentzen's proof had to use nonfinitistic means. Gentzen succeeded in concentrating all nonfinitistic means in one single point – induction along a wellordering of transfinite ordertype. This result confirmed the Hilbert school's opinion that just a slight modification of the finitistic standpoint (i.e. accepting a weak form of transfinite induction) would suffice to make the whole program feasible. In §16 we will discuss the consequences of this 'slight modification' for Hilbert's program. There we will try to argue, in the spirit of Hilbert's program, that Gentzen's proof is of little help. This, however, does not mean that Gentzen's proof and his results are of no importance. Quite on the contrary, in our opinion Gentzen's proof is one of the deepest results in logic. To see why, we propose a reinterpretation of his results.

In point of fact it is very easy to prove the consistency of pure number theory. One simply has to show that there exists a model for it. So what is the advantage of Gentzen's consistency proof? The construction of the model itself needs a certain framework, e.g. set theory. Thus what is obtained by a consistency proof via a model construction in set theory (or some even weaker theory) is that the consistency of set theory also entails the consistency of pure number theory. Gentzen's proof, however, gives much more information. It has already been mentioned that Gentzen's proof is finitistic apart from his use of induction along a wellordering of transfinite ordertype. In our opinion this is the essential contribution of Gentzen's proof. Its consequences are twofold:

1. The induction in Gentzen's proof need only be applied to formulas of a very restricted complexity. In addition the consistency proof never uses the law of the excluded middle. Thus it may be formalized within a system T based on intuitionistic logic with induction along a wellordering of transfinite ordertype where this induction scheme is restricted to formulas of a very low complexity. So the problem of the consistency of pure number theory may be decided within the system T. Although the wellordering is of transfinite order type it can easily be visualized. So it seems to be completely plain that the system T is consistent. By Gödel's second theorem the proof theoretic strength of the system T, as it will be defined later in this lecture, has to exceed that of pure number theory. But the subsystem T_0 of T which is obtained from T by restricting induction to initial segments of the wellordering only can be shown to be equiconsistent with elementary number theory. Thus Gentzen's proof provides a reduction of the consistency problem for elementary number theory to that of a theory T_0, which from a conceptual point of view may be regarded as 'safer' than elementary number theory itself.

This is an example of *reductive proof theory*. In reductive proof theory one generally tries to reduce the consistency problem of a theory T_1 to that of a theory T_2. For a clever choice of T_2 both systems will have the same proof theoretic strength. The principles used in T_2, however, may be easier to visualize and therefore a justification of the system T_2 seems more plausible. This type of proof theory is of great foundational importance (cf. the introduction to [BFPS] by S.Feferman). One important feature of Hilbert's program we did not mention is the 'elimination of ideal elements'. In this sense reductive proof theory contributes to Hilbert's program by eliminating complicated unperspicuous principles. Since both systems (T_1 and T_2 in the above example) are of the same proof theoretical strength reductive proof theory is in full accordance with Gödel's second theorem.

2. The fact that induction along the wellordering is the only nonfinitistic means in Gentzen's proof also suggests using this wellordering as a measure for the transfinite content of pure number theory. Pursuing this idea one had defined the *proof theoretic ordinal* of a formal theory T as the ordertype of the smallest wellordering which is needed for a consistency proof of T. This definition, however, is somehow vague since it says nothing about the means used besides the induction along this wellordering (one tacitly has to assume that these at least have to be formalizable in T). To obtain a more precise definition one calls an ordinal α provable in T if there is a primitive recursively definable wellordering $\prec$ of ordertype α such that the wellordering of $\prec$ is provable in T. It is a consequence of Gödel's second theorem, that the proof theoretic ordinal of T (in the previous sense) cannot be a provable ordinal of T. Therefore one may define the proof theoretic ordinal of T as the least ordinal which is not provable in T. This is the common definition today. The computation of the proof theoretic ordinal of T is called the *ordinal analyis of* T. Gentzen's paper ' Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie' [Gentzen 1943] indicates that he himself already interpreted his result as an ordinal analysis (and not just as a consistency proof).

The intention of this lecture is to give an introduction to the techniques of ordinal analysis. We suppress the aspects of reductive proof theory. Only in the epilogue it will be indicated how the results and methods of ordinal analysis may be used in reductive proof theory. To get acquainted with the basic notions and techniques we reprove Gentzen's result in the first chapter. The second chapter will discuss the limits of Gentzen's methods. There we will reprove S.Feferman's and K.Schütte's results on the limits of predicativity. The emphasis, however, is on the ordinal analysis of impredicative formal systems. To demonstrate this method we will give in chapter III an ordinal analysis for one of the simplest impredicative formal systems, the system ID_1 for noniterated inductive definitions by the *method of local predicativity.* A discussion on the foundational significance of ordinal analysis will be added in the epilogue.

CHAPTER I

ORDINAL ANALYSIS OF PURE NUMBER THEORY

To begin with we follow Hilbert's program and, in a first step, try to axiomatize – if not the whole of mathematics – but the theory of natural numbers. To obtain a feeling how this might be done we start by some heuristic considerations.

The aim of the 'working' mathematician interested in the theory of a certain structure is to discover the 'mathematical facts' which hold in this structure. In order to do this he first has to be able to formulate the 'mathematical facts'. This means that he needs a language in which he may talk about this structure. The mathematical facts which possibly may hold in the structure will then be expressed by sentences in this language. The problem then is to figure out which of the sentences are the true ones. This may be done by pure intuition. But to be really sure about the truth of a sentence it needs a proof. The only way to prove a sentence, however, is to show that it is a logical consequence of some other sentences which already are known to be true in the structure. Tracking back this procedure we finally end up with a set of sentences, the mathematical axioms of the structure, which cannot be proved themselves but either are true by definition or by common agreement. Showing that a sentence is a logical consequence of other sentences usually is done by deriving the sentence from those others through a series of inferences. A set of inference rules will be called a *proof procedure.* Some of the inferences in a proof procedure may have no premises. Those inferences will be called the *logical axioms* of the proof procedure. The choice of the axioms and of the proof procedure is of course not arbitrary. As a first requirement the truth of every mathematical axiom really has to be indubitable and it also must be clear that the truth of the premises of an inference undoubtedly entails the truth of its conclusion (if there is no premise, then the conclusion must be true in every structure, i.e. logically valid.). This will guarantee that all proven sentences really are true. But the 'working' mathematician does not only want to ensure himself about his theorem but he also wants to convince his colleagues about its truth. Therefore there must be a way of checking a proof. Thus the second requirement is, that it must be decidable whether a given sentence is an axiom or not, and it also has to be decidable whether an inference is a correct application of an inference rule or not. Otherwise we had no possibility to check the correctness of a given proof. A proof procedure meeting these requirements will be called *decidable*.

This little heuristic teaches us the following facts about axiomatization:

In order to axiomatize the theory of a structure we

- first need a *formalization of the language of the structure*. The formal language of the structure has to be given in such a way that it becomes decidable whether a symbol string is a wellformed expression or not;
- second need a decidable set of sentences in this language which undoubtedly are true. The sentences in this set are the *axioms of the structure*;
- third need a decidable *proof procedure* which produces logical consequences of the axioms.

A decidable formal language together with a decidable set of mathematical axioms and a decidable proof procedure will be called a *formal system* or sometimes also a *formal theory* for the structure. From this it immediately follows that the set of sentences which are provable in one formal system always is a recursively enumerable set.

By results of mathematical logic there are complete proof procedures for first order languages, i.e. there are proof procedures which produce all logical consequences of a given set of mathematical axioms. This of course must not be mistaken in that way that the proof procedure together with the mathematical axioms produce all true first order sentences of the structure. In general the set of true sentences of a structure is not recursively enumerable but of higher complexity. Thus in general we cannot expect a complete axiomatization even for the first order theory of a structure. Since we have to abandon completeness anyway we may as well regard the second order language of the structure although there is not even a complete proof procedure for second order logic. The only important thing is that there are sound proof procedures. It will then be the task of proof theoretical research to determine the limits of a formal system.

In the present lecture we will not use full second order logic but first order logic with free set variables. We will introduce the notion of a Π^1_1-sentence and then examine the power of formal systems with respect to their provable Π^1_1-sentences.
In the first sections of the following chapter we are going to develop a quite simple formal system for the structure of natural numbers which in the later sections will be analyzed proof theoretically.

§1. The language $\mathscr{L}$ of pure number theory

A structure usually is given by a non void set together with collections of constants, of functions and of relations on that set. In order to obtain a formal language for the structure of natural numbers we first need to specify our picture of this structure. The set of natural numbers is characterized by the facts that every natural number either is zero or the successor of another natural number and that every natural number possesses a uniquely determined successor. Using this characterization we obtain a name (or constant as we are going to call it) for every natural number. We start with $\underline{0}$ as a name for the natural number zero and a symbol $\underline{S}$ for the successor function. Then a constant for every natural number is obtained by successively applying the successor function to the symbol $\underline{0}$. So it should be clear that we at least need a constant for zero and the successor function in our language (and then as well may assume that we already have a constant $\underline{n}$ for every natural number n). The next question to be answered is which functions and relations besides the successor function on the natural numbers we should consider. The most general answer is of course "all possible functions and relations on the set of natural numbers". Since there are uncountably many such functions and relations this already would lead to a language with uncountably many basic symbols. In a formal system only those constants for which there are defining axioms contribute to the power of the formal system. Therefore we would need an uncountable set of axioms which is outside the scope of a formal system since every decidable set already is countable. If we dispense with defining axioms for function or relation constants we may as well treat them as variables. In fact we will introduce a language which has such second order variables. In our framework it will suffice just to introduce set variables. The introduction of bare set variables (or function variables) will in general also not raise the power of a formal system (cf. exercise 3.15.4). But if we add the defining axioms for set variables, i.e. the comprehension scheme, we will obtain a system which is so strong that up to now we have not been able to do its proof theoretic analysis. Therefore we will be more modest and in a first step will restrict ourselves to a system which we are going to call the system of *pure number theory.* The most important functions in number theory are *'plus'* and *'times'*. 'Plus' and 'times' are primitive recursive functions and it is possible to obtain all primitive recursive functions from 'plus' and 'times' (cf. remark 3.12.). Therefore we are going to introduce a seemingly stronger system in which we have a constant for every primitive recursive function and relation. In order to do this we first will introduce names for all primitive recursive functions. In definition 1.1. we will give the syntactical definition of the primitive recursive function terms, while the meaning of those terms becomes clear from definition 1.2. in which we define the evaluation of an n-ary primitive recursive function term f on an n-tuple $t_1,\dots,t_n$ of natural numbers.

1.1. Primitive recursive function terms

(i) S (the symbol for the successor function) is an unary primitive recursive function term.

(ii) P^n_k (the symbol for the k-th projection of an n-tuple) and C^n_k (the symbol for the n-ary constant function with value k) are n-ary primitive recursive function terms, where in the case of P^n_k we require $1 \le k \le n$.

(iii) If $h_1,\dots h_m$ are n-ary primitive recursive function terms and g is an m-ary primitive recursive function term, then $Sub(g,h_1,\dots,h_m)$ is an n-ary primitive recursive function term. (Substitution of functions).

(iv) If g is an n-ary and h an n+2-ary primitive recursive function term, then Rec(g,h) is an n+1-ary primitive recursive function term. (Primitive recursion).

1.2. Inductive definition of $f(t_1,\dots,t_n) = t$ for an n-ary primitive recursive function term f and natural numbers $t_1,\dots,t_n,t$

(i) $S(t_1)=t$ if t is the successor of t_1,

(ii) $C^n_k(t_1,\dots,t_n) = t$ if $t=k$,

(iii) $P^n_k(t_1,\dots,t_n) = t$ if $t=t_k$,

(iv) $Sub(g,h_1,\dots,h_m)(t_1,\dots,t_n)=t$ if there are natural numbers $u_1,\dots,u_m$ such that $h_i(t_1,\dots,t_n) = u_i$ and $g(u_1,\dots,u_m) = t$.

(iv) $\mathrm{Rec}(g,h)(t_1,\dots,t_n,k) = t$ holds if $k=0$ and $g(t_1,\dots,t_n) = t$ or if k is the successor of k_0 and $h(t_1,\dots,t_n,k_0,\mathrm{Rec}(g,h)(t_1,\dots,t_n,k_0)) = t$.

$f(t_1,\dots,t_n) = t$ is to be read as: "The *evaluation* of the n-ary primitive recursive function f on the n-tuple $t_1,\dots,t_n$ of natural numbers yields the value t".

1.3. Definition

The graph of an n-ary primitive recursive function term f is the n+1-ary relation $\{f\}$ given by $\{f\}(t_1,\dots,t_n,t) :\Leftrightarrow f(t_1,\dots,t_n) = t$.

1.4. Definition

An n-ary relation R on $\mathbb{N}$ is *primitive recursive* if its characteristic function χ_R defined by

$$\chi_R(t_1,\dots,t_n) := \begin{cases} 1, & \text{if } R(t_1,\dots,t_n) \\ 0, & \text{otherwise} \end{cases}$$

is primitive recursive.

We do not want to go deeper into the theory of primitive recursive functions. This is the topic of another lecture. The aim of the preceeding definitions was to emphasize that it is possible to name every primitive recursive function by a term. This also means that, via its characteristic function, we have a name for every primitive recursive relation. We now are prepared to introduce the formal language $\mathcal{L}$ for the structure of natural numbers.

1.5 Basic symbols of the language $\mathcal{L}$

1. *Logical symbols*

(i) Countably many number variables denoted by u,v,w,x,y,z,...

(ii) Countably many set variables denoted by U,V,W,X,Y,Z,...

(iii) The *sentential connectives* $\neg, \wedge, \vee$, the *quantifiers* $\forall, \exists$ and the *membership relation* symbol $\in$.

2. *Nonlogical symbols*

(i) A constant $\underline{n}$ for every natural number n.

(ii) An n-ary function constant $\underline{f}$ for every n-ary primitive recursive function term f.

(iii) An n-ary relation symbol $\underline{R}$ for every primitive recursive relation R.

When no confusion is to be feared we often will omit the underlining.

3. Brackets serve as *auxiliary symbols.*

1.6 Inductive definition of the *terms* of the language $\mathscr{L}$

(i) Every number constant $\underline{n}$ is a term and it is $FV_1(\underline{n})=\emptyset$.

(ii) Every number variable x is a term and it is $FV_1(x)=\{x\}$.

(iii) If $t_1,...,t_n$ are terms and $\underline{f}$ is an n-ary function constant, then $(\underline{f}t_1,...,t_n)$ is a term and it is $FV_1(\underline{f}t_1,...,t_n) = FV_1(t_1)\cup...\cup FV_1(t_n)$.

We call $FV_1(t)$ the set of *number variables occuring free in* t.

1.7 Inductive definition of the *formulas* of $\mathscr{L}$

(i) If $t_1,...,t_n$ are terms and $\underline{P}$ is an n-ary relation symbol, then $(\underline{P}t_1...t_n)$ is a formula and it is $FV_1(\underline{P}t_1...t_n) = FV_1(t_1)\cup...\cup FV_1(t_n)$ and $BV_1(\underline{P}t_1...t_n) = FV_2(\underline{P}t_1...t_n) = BV_2(\underline{P}t_1...t_n) = \emptyset$.

(ii) If t is a term and X a set variable, then $t\in X$ is a formula and it is $FV_1(t\in X) = FV_1(t)$, $FV_2(t\in X) = \{X\}$, $BV_1(t\in X) = BV_2(t\in X) = \emptyset$.

(iii) If A and B are formulas, then $(\neg A)$, $(A\wedge B)$ and $(A\vee B)$ are formulas and it is $FV_i(\neg A) = FV_i(A)$, $FV_i(A \overset{\wedge}{\vee} B) = FV_i(A)\cup FV_i(B)$ for i=1,2 and $BV_i(\neg A) = BV_i(A)$, $BV_i(A \overset{\wedge}{\vee} B) = BV_i(A)\cup BV_i(B)$ for i=1,2.

(iv) If A is a formula such that $x\notin BV_1(A)$, then $\forall x A$ and $\exists x A$ are formulas and we define $FV_1(QxA)=FV_1(A)\setminus\{x\}$, $FV_2(QxA)=FV_2(A)$, $BV_1(QxA)=BV_1(A)\cup\{x\}$ and $BV_2(QxA)=BV_2(A)$ for $Q\in\{\forall,\exists\}$.

Formulas which are built according to (i) or (ii) are called *atomic.*

$FV_1(F)$ is the set of *free number variables* occuring in F, $FV_2(F)$ the set of *free set variables* occuring in F. We call $BV_1(F)$ the set of *number variables occurring bound in* F. By FV(F) we denote the set $FV_1(F)\cup FV_2(F)$ of *free variables* of F and by BV(F) the set $BV_1(F)\cup BV_2(F)$ of *bound variables* of F

Sentences are formulas F without free variables, i.e formulas F such that $FV_1(F)\cup FV_2(F) = \emptyset$.

Π^1_1*-sentences* are formulas F such that $FV_1(F) = \emptyset$, i.e. formulas without free number variables.

Up to now we have $BV_2(F)=\emptyset$ for all formulas F. That means that F does not contain bound set variables or bound *second order variables* as they often are synonymized. One therefore calls them *first order formulas.* We usually refer to first order formulas as $\mathscr{L}_1$-formulas. The *second order formulas* or $\mathscr{L}_2$-formulas are obtained by adding the clause

(v) If F is a formula and $X \notin BV_2(F)$, then $(\forall X F)$ and $(\exists X F)$ are formulas such that $FV_1(QXF) = FV_1(F)$, $BV_1(QXF) = BV_1(F)$, $FV_2(QXF) = FV_2(F)\setminus\{X\}$ and $BV_2(QXF) = BV_2(F)\cup\{X\}$ for $Q \in \{\forall,\exists\}$.

1.8 Notational conventions

As syntactical variables for number variables we use the letters u,v,w,x,y,z. Terms are denoted by r,s,t,a,b,c and number constants by $\underline{m},\underline{n},\underline{k},\underline{l}$. U,V,W,X,Y,Z are syntactical variables for set variables. All these symbols will also occur with indices.

By $A[x_1,...,x_n]$ we indicate that the variables $x_1,...,x_n$ really do occur in A, i.e. $FV_1(A) = \{x_1,...,x_n\}$. $A(x_1,...,x_n)$ just means that $x_1,...,x_n$ may occur in A. We use analogous conventions for set variables.

$A_x(s)$ or $t_x(s)$ are obtained from A or t respectively by replacing all occurences of x by s. If there is no danger of confusion we omit the subscript x.

Class terms of the form $\{x:A(x)\}$ do not belong to the language but will be used as defined objects. $s \in \{x:A(x)\}$ then stands for $A_x(s)$. We often write $A_X(B)$ instead of $A_X(\{x:B(x)\})$ and omit the subscript X whenever there is no danger of confusion.

The sentential connectives $\rightarrow$ and $\leftrightarrow$ are defined as usual by $\neg ... \vee \_\_\_$ and $(... \rightarrow \_\_\_) \wedge (\_\_\_ \rightarrow ...)$ respectively.

1.9 Exercises

1. Suppose that t, s are $\mathscr{L}$-terms.
   (i) Give an inductive definition of $t_x(s)$.
   (ii) Show that $t_x(s)$ again is an $\mathscr{L}$-term.
2. Suppose that s is an $\mathscr{L}$-term and F is an $\mathscr{L}$-formula.
   (i) Give an inductive definition of $F_x(s)$.
   (ii) Show that $F_x(s)$ again is an $\mathscr{L}$-formula.
3. Suppose that F and B are $\mathscr{L}$-formulas.
   (i) Give an inductive definition of $F_X(\{x:B(x)\})$.
   (ii) Find formulas F, B such that $F_X(\{x:B(x)\})$ is not an $\mathscr{L}$-formula.
   iii) What prerequisites are needed for F and B in order to obtain $F_X(\{x:B(x)\})$ to be an $\mathscr{L}$-formula?

§ 2. Semantics for $\mathcal{L}$

Hitherto we defined terms and formulas of $\mathcal{L}$ as mere syntactical objects. To give them a mathematical meaning we need an interpretation for the formal language $\mathcal{L}$. The development of such a semantics is the goal of the present section. We will, however, not develop a general theory of semantics for $\mathcal{L}$ but, according to our intention, will restrict ourselves to the so called standard interpretation of $\mathcal{L}$ in the structure $\mathbb{N}$ of natural numbers.

2.1. **Definition**

An *assignment* for $\mathcal{L}$ is a mapping Φ which assigns a number $\Phi(x) \in \mathbb{N}$ to every number variable x and a set $\Phi(X) \subset \mathbb{N}$ to every set variable X.

2.2. **Inductive Definition** of the *value* t^Φ *of an $\mathcal{L}$-term* t with respect to an assignment Φ

(i) $\underline{n}^\Phi = n$

(ii) $x^\Phi = \Phi(x)$

(iii) $(\underline{f}t_1 \dots t_n)^\Phi = n$ if $f(t_1^\Phi, \dots, t_n^\Phi) = n$ according to 1.2.

As an immediate consequence of definition 2.2. we obtain $t^\Phi \in \mathbb{N}$.

2.3. **Inductive definition** of $\mathbb{N} \models A^\Phi$

Suppose that Φ is an assignment for $\mathcal{L}$.

(i) $\mathbb{N} \models (\underline{P}t_1 \dots t_n)^\Phi :\Leftrightarrow \chi_P(t_1^\Phi, \dots, t_n^\Phi) = 1$

i.e., we have $\mathbb{N} \models (\underline{P}t_1 \dots t_n)^\Phi$ iff $P(t_1^\Phi, \dots, t_n^\Phi)$ is true where P is the primitive recursive predicate denoted by $\underline{P}$

(ii) $\mathbb{N} \models (t \in X)^\Phi \Leftrightarrow t^\Phi \in \Phi(X)$

(iii) $\mathbb{N} \models (\neg A)^\Phi \Leftrightarrow \mathbb{N} \not\models A^\Phi$

(iv) $\mathbb{N} \models (A \wedge B)^\Phi \Leftrightarrow \mathbb{N} \models A^\Phi$ and $\mathbb{N} \models B^\Phi$

(v) $\mathbb{N} \models (A \vee B)^\Phi \Leftrightarrow \mathbb{N} \models A^\Phi$ or $\mathbb{N} \models B^\Phi$

(vi) $\mathbb{N} \models \forall x A^\Phi \Leftrightarrow \mathbb{N} \models A_x(\underline{n})^\Phi$ for all $n \in \mathbb{N}$

(vii) $\mathbb{N} \models \exists x A^\Phi \Leftrightarrow \mathbb{N} \models A_x(\underline{n})^\Phi$ for some $n \in \mathbb{N}$

This gives the semantics for $\mathcal{L}_1$-formulas. We obtain the semantics for $\mathcal{L}_2$-formulas by adding the clauses

(viii) $\mathbb{N} \models \forall X A^\Phi \Leftrightarrow \mathbb{N} \models A^\Psi$ for any assignment Ψ which at most differs in the value of $\Psi(X)$ from Φ.

(ix) $\mathbb{N} \models \exists X A^\Phi \Leftrightarrow$ There is an assignment Ψ which at most differs in the value of $\Psi(X)$ from Φ such that $\mathbb{N} \models A^\Psi$.

If $FV_1(t)=\emptyset$ we have $t^\Phi=t^\Psi$ for all assignments Φ and Ψ. For *closed terms* t, i.e. terms t such that $FV_1(t)=\emptyset$, we therefore define $t^{\mathbb{N}}:=t^\Phi$ for an arbitrary assignment Φ. Two closed terms s and t such that $s^{\mathbb{N}}=t^{\mathbb{N}}$ are called *equivalent*. Two formulas F_1 and F_2 are said to be *equivalent* if they only differ in equivalent terms.

The value of t^Φ and the relation $\mathbb{N}\models A^\Phi$ obviously only depend upon $\Phi\restriction FV(t)$ or $\Phi\restriction FV(A)$ respectively. If $FV(t)=\{x_1,\dots,x_n\}$ or $FV(A)=\{x_1,\dots,x_n,X_1,\dots,X_m\}$, we often write $t[k_1,\dots,k_n]$ or $\mathbb{N}\models A[k_1,\dots,k_n,S_1,\dots,S_m]$ repectively instead of "t^Φ or $\mathbb{N}\models A^\Phi$ for an assignment Φ such that $\Phi(x_i)=k_i$ and $\Phi(X_j)=S_j$ hold for $i=1,\dots,n$ and $j=1,\dots,m$".

If F is a sentence we obviously have $\mathbb{N}\models F^\Phi \Leftrightarrow \mathbb{N}\models F^\Psi$ for all assignments Φ and Ψ. In this case we write $\mathbb{N}\models F$ and say that the sentence F *is valid in* $\mathbb{N}$. For Π^1_1-sentences $A[X_1,\dots,X_n]$ we have $\mathbb{N}\models \forall X_1\dots\forall X_n A$ if and only if $\mathbb{N}\models A^\Phi$ holds for any assignment Φ. This is the reason for calling them Π^1_1-sentences although they prima facie are $\mathscr{L}_1$-formulas. For Π^1_1-sentences we always write $\mathbb{N}\models A$ instead of $\mathbb{N}\models \forall X_1\dots\forall X_n A$. This notation sometimes will also be used for arbitrary formulas A. So, for a formula A, $\mathbb{N}\models A$ means '$\mathbb{N}\models A^\Phi$ for all assignments Φ'.

2.4. Exercise

Suppose that L is a first order language which is given by a set $\mathbb{C}$ of individual constants, a set $\mathbb{F}$ of function constants and a set $\mathbb{P}$ of predicate constants. We define L_1 and L_2 analogously to $\mathscr{L}_1$ or $\mathscr{L}_2$ respectively. The semantics for L_1 and L_2 is defined in the following way.

(i) A structure $\mathscr{S}$ for L_1 is a quadruple $(I,\mathscr{C},\mathscr{F},\mathfrak{P})$ which satisfies the following conditions:

(a) $I\neq\emptyset$ is a set.

(b) We have $\mathscr{C}\subset I$ such that for every $c\in\mathbb{C}$ there is a $c^{\mathscr{S}}\in\mathscr{C}$.

(c) $\mathscr{F}$ is a set of function on I, such that for any n-ary function symbol $f\in\mathbb{F}$ there is a function $f^{\mathscr{S}}: I^n\rightarrow I$ in $\mathscr{F}$.

(d) $\mathfrak{P}$ is a set of predicates on I such that there is a $P^{\mathscr{S}}\subset I^n$ in $\mathfrak{P}$ for every n-ary predicate symbol $P\in\mathbb{P}$.

(ii) A structure $\mathscr{S}$ for L_2 is a quintuple $(I,M,\mathscr{C},\mathscr{F},\mathfrak{P})$ such that $(I,\mathscr{C},\mathscr{F},\mathfrak{P})$ is a structure for L_1 and $M\subset Power(I)$ (the power set of I).

(iii) If $\mathscr{S}$ is a structure for L_1 or L_2, then an $\mathscr{S}$-assignment for $\mathscr{L}_i(i=1,2)$ is a mapping Φ which assigns to any x an element $\Phi(x)\in I$ and to any set variable X a set $\Phi(X)\subset I$ or $\Phi(X)\in M$ respectively.

For L-terms t and L-formulas F and an $\mathcal{S}$-assignment Φ we define t^{Φ} and $\mathcal{S} \models F^{\Phi}$ analogously to 2.2 and 2.3 respectively. We write $\mathcal{S} \models F$ for an L_i-formula F if $\mathcal{S} \models F^{\Phi}$ holds for any $\mathcal{S}$-assignment Φ and $\models F$ if $\mathcal{S} \models F$ holds for all L_i-structures $\mathcal{S}$.

Prove the following claims:

(i) $\mathcal{S} \models A \to F \Rightarrow \mathcal{S} \models A \to \forall x F$ if $x \notin FV(A)$

(ii) $\mathcal{S} \models F \to A \Rightarrow \mathcal{S} \models \exists x F \to A$ if $x \notin FV(A)$

(iii) $\models \forall X F \to F_X(Y)$

(iv) $\models F_X(Y) \to \exists X F$

(v) $\mathcal{S} \models A \to F \Rightarrow \mathcal{S} \models A \to \forall X F$ if $X \notin FV(A)$

(vi) $\mathcal{S} \models F \to A \Rightarrow \mathcal{S} \models \exists X F \to A$ if $X \notin FV(A)$

§3. A formal system for pure number theory

Still in the spirit of Hilbert's program we are trying to establish a formal system which derives as much valid sentences of $\mathbb{N}$ as possible. In a first step we are going to deal with those sentences which are valid because of their logical structure. Every formula of $\mathcal{L}$ carries a sentential and a quantifier structure. To clarify the sentential structure of an $\mathcal{L}$-formula which is given by the logical connectives $\neg, \wedge$ and $\vee$ we introduce the sentential subformulas of an $\mathcal{L}$-formula.

3.1 Inductive definition of the set AT(F) of *sentential subformulas* of an $\mathcal{L}$-formula F

(i) If F is atomic or a formula QxA or QXA respectively where $Q \in \{\forall, \exists\}$, then $AT(F) = \{F\}$.

(ii) If F is a formula $\neg A$, then $AT(F) = \{F\} \cup AT(A)$.

(iii) If F is a formula $(A \wedge B)$ or $(A \vee B)$, then $AT(F) = \{F\} \cup AT(A) \cup AT(B)$.

Formulas A such that $AT(A) = \{A\}$ are called *sentential atoms*. By AE we denote the set of all sentential atoms of $\mathcal{L}$. We define $AE(F) := AE \cap AT(F)$.

3.2. **Definition**

(i) A *sentential assignment* is a mapping $\mathbb{B} : AE \to \{t, f\}$.

(ii) The truth value $A^{\mathbb{B}}$ of a formula A under a given sentential assignment $\mathbb{B}$ is given by the usual interpretation of the logical connectives as truth functions (cf. 2.3.(iii)-(v) and 10.12. below).

One should notice that only the values of $\mathbb{B}$ restricted to AE(A) are needed in the computation of $A^{\mathbb{B}}$.

(iii) A formula A is *sententially valid* if $A^{\mathbb{B}} = t$ holds for all sentential assignments $\mathbb{B}$.

3.3. **Lemma**

If $A \in AT(F)$, *then* $FV_i(A) \subset FV_i(F)$ *for* $i = 1,2$.

The proof is an easy induction on the definition of $A \in AT(F)$.

An assignment Φ canonically induces a sentential assignment $\mathbb{B}_\Phi$ by defining $A^{\mathbb{B}\Phi} = t \Leftrightarrow \mathbb{N} \models A^\Phi$ for all sentential atoms A. For these assignments we have the following lemma.

3.4. **Lemma**

$\mathbb{N} \models A^\Phi$ *holds if and only if* $A^{\mathbb{B}\Phi} = t$.

Proof by induction on the length of the formula A

1. If $A \in AE$, then we have $\mathbb{N} \models A^\Phi \Leftrightarrow A^{\mathbb{B}\Phi} = t$ by definition.
2. If A is a formula $\neg B$, then we have $\mathbb{N} \models A^\Phi \Leftrightarrow \mathbb{N} \not\models B^\Phi \Leftrightarrow B^{\mathbb{B}\Phi} = f \Leftrightarrow A^{\mathbb{B}\Phi} = t$.
3. If A is a formula $(B \overset{\wedge}{\vee} C)$, then we have $\mathbb{N} \models A$ if and only if $\mathbb{N} \models B$ and/or $\mathbb{N} \models C$. By the induction hypothesis this holds if and only if $B^{\mathbb{B}\Phi} = t$ and/or $C^{\mathbb{B}\Phi} = t$. But this is equivalent to $(B \overset{\wedge}{\vee} C)^{\mathbb{B}\Phi} = t$.

As a corollary to 3.4. we obtain the following theorem.

3.5. **Theorem**

If F is sententially valid, then $\mathbb{N} \models F$.

Concerning the quantifier structure of an $\mathscr{L}$-formula we just need the following observation.

3.6 **Lemma**

If F *is a formula* $\neg A_x(t) \vee \exists x A$ *or* $\neg \forall x A \vee A_x(t)$, *then* $\mathbb{N} \models F$.

Proof

From $\mathbb{N} \models A_x(t)^\Phi$ we have to conclude $\mathbb{N} \models \exists x A^\Phi$. By induction on the length of A we easily obtain $\mathbb{N} \models A_x(t)^\Phi \Leftrightarrow \mathbb{N} \models A_x(\underline{t^\Phi})^\Phi$. Hence $\mathbb{N} \models \exists x A^\Phi$. In the second case we have to show that $\mathbb{N} \models \forall x A^\Phi$ implies $\mathbb{N} \models A_x(t)^\Phi$. But $\mathbb{N} \models \forall x A^\Phi$ implies $\mathbb{N} \models A_x(\underline{t^\Phi})^\Phi$ and therefore also $\mathbb{N} \models A_x(t)^\Phi$.

In the proof of 3.6. the careful reader will have noticed that the proof needs the additional hypothesis that the term t is substitutable for x in A, i.e. none of the free variables occurring in t must be bound in A. Here and in future we will tacitly assume that this prerequisit always is satisfied. This means no restriction since by renaming the bound variables in A we may always obtain that t is substitutable in A.

Now we are prepared to formulate the axioms and inference rules of the formal system $\mathbf{Z}_1$ of pure number theory. The language of $\mathbf{Z}_1$ is the first order language $\mathscr{L}_1$.

3.7. **Logical axioms** of the formal system $\mathbf{Z}_1$

(i) Every sententially valid formula is a logical axiom of $\mathbf{Z}_1$.

(ii) Every formula of the form $\neg \forall x A \vee A_x(t)$ and $\neg A_x(t) \vee \exists x A$ is a logical axiom of $\mathbf{Z}_1$.

3.8. **Logical inferences** of the formal system $\mathbf{Z}_1$

(mp)	$A, \neg A \vee B \vdash B$		(modus ponens)
($\forall$)	$\neg A \vee B \vdash \neg A \vee \forall x B$	if $x \notin FV_1(A)$	($\forall$-rule)
($\exists$)	$\neg B \vee A \vdash \neg \exists x B \vee A$	if $x \notin FV_1(A)$	($\exists$-rule)

The variable x of a quantifier inference is called its *eigenvariable*.

3.9. **Equality axioms** of the formal system $\mathbf{Z}_1$

Among the constants for the primitive recursive relations we have a constant = for the equality relation. Although we could derive the properties of the equality relation from its defining axioms (contained in 3.10.) we prefer to formulate them explicitly as a separate group of axioms. This is in coincidence with the usual treatment of formal systems where the equality symbol often is regarded as a logical symbol.

(i) $\forall x(x = x)$

(ii) $\forall x \forall y\ (x = y \rightarrow y = x)$

(iii) $\forall x_1 \forall x_2 \forall x_3 (x_1 = x_2 \wedge x_2 = x_3 \rightarrow x_1 = x_3)$

(iv) $\forall x \forall y (x = y \rightarrow t = t_x(y))$

(v) $\forall x \forall y (x = y \rightarrow (F \rightarrow F_x(y)))$.

3.10. Mathematical axioms of the formal system $\mathbf{Z}_1$

(i) The *successor axioms*

$\forall x(\neg \underline{0} = \underline{S}x)$

$\forall x \forall y(\underline{S}x = \underline{S}y \rightarrow x = y)$

$\underline{Sn} = \underline{S}\,\underline{n}$ for all $n \in \mathbb{N}$.

(ii) The *defining axioms for primitive recursive functions* are given by the universal closures of the following formulas

$(\underline{C}^n_k x_1,\dots,x_n) = \underline{k}$

$(\underline{P}^n_k x_1,\dots,x_n) = x_k$

$(\mathrm{Sub}(g,h_1,\dots,h_m)x_1\dots x_n) = (g(h_1 x_1\dots x_n) \dots (h_m x_1\dots x_n))$

$(x = 0 \rightarrow ((Rgh)x_1\dots x_n x) = g x_1\dots x_n) \wedge$

$(x = \underline{S}y \rightarrow ((Rgh)x_1\dots x_n x) = h x_1\dots x_n y((Rgh)x_1\dots x_n y))$

$Rx_1\dots x_n \leftrightarrow \chi_R x_1\dots x_n = \underline{1}$

(iii) The *induction axiom* is given by the scheme

(IND) $A_x(0) \wedge \forall y(A_x(y) \rightarrow A_x(\underline{S}y)) \rightarrow \forall x A$

3.11. Inductive definition of $\mathbf{Z}_1 \vdash F$

We are going to define the formal derivation predicate for $\mathbf{Z}_1$. $\mathbf{Z}_1 \vdash F$ should be read as 'F is formally derivable in $\mathbf{Z}_1$'.

(i) If A is one of the axioms 3.7.,3.9. or 3.10, then $\mathbf{Z}_1 \vdash A$.

(ii) If $\mathbf{Z}_1 \vdash A_i$ ($i = 1$ or $i = 1,2$) holds for the premise(s) of an interference according to 3.8. whose conclusion is A, then we also have $\mathbf{Z}_1 \vdash A$.

3.12. Remark

The system $\mathbf{Z}_1$ is an extension by definitions of the better known system PA of Peano arithmetic. **PA** is formulated in a first order logic with equality. The only nonlogical symbols of **PA** are the binary function symbols '+' for addition and '· ' for multiplication, the unary function symbol $\underline{S}$ for the sucessor function and a constant $\underline{0}$ for the natural number 0. (The equality symbol is counted among the logical symbols). The axioms of the group 3.10.(ii) are then replaced by the defining axioms for $\underline{0}$, $\underline{S}$, '+' and '·', i.e. by the universal closure of the following formulas

$x+\underline{0} = x$ and $x\cdot\underline{0} = \underline{0}$

$x+\underline{S}y = \underline{S}(x+y)$ $x\cdot\underline{S}y = (x\cdot y)+x$

Here as usual we have written $(x+y)$ instead of $(+xy)$ and $(x\cdot y)$ instead of $(\cdot xy)$. Apparently **PA** is a subsystem of $\mathbf{Z}_1$, i.e. we have

(i) $\mathbf{PA} \vdash F \Rightarrow \mathbf{Z}_1 \vdash F$ *for every formula* F *in the language of* **PA**.

We also have the opposite direction of (i) which means that $\mathbf{Z}_1$ is a conservative extension of **PA**. But we may even prove:

(ii) *For every* $\mathcal{L}_1$ *-formula* F *there is a formula* F_P *in the language of* **PA** *such that* $\mathbf{Z}_1 \vdash F \leftrightarrow F_P$.

This means that every symbol of $\mathbf{Z}_1$ can be defined in **PA**. For this reason $\mathbf{Z}_1$ is called an extension of **PA** by definitions. The proofs of (i) and (ii), however, require methods from the theory of recursive functions and will not be given here.

3.13. **Soundness theorem** for $\mathbf{Z}_1$

If $\mathbf{Z}_1 \vdash F$, *then* $\mathbb{N} \models F$.

Proof

By induction on the definition of $\mathbf{Z}_1 \vdash F$ we show that $\mathbf{Z}_1 \vdash F$ implies $\mathbb{N} \models F^\Phi$ for any assignment Φ. If F is a logical axiom then we obtain $\mathbb{N} \models F^\Phi$ by 3.5. or 3.6 respectively. The claim is obvious for the equality axioms and the mathematical axioms. We only should check the induction scheme. Here we have to show that $\mathbb{N} \models A(\underline{0})^\Phi$ and $\mathbb{N} \models \forall y(A(y) \to A(\underline{S}y))^\Phi$ imply $\mathbb{N} \models A(\underline{n})^\Phi$ for all $n \in \mathbb{N}$. But $\mathbb{N} \models \forall y(A(y) \to A(\underline{S}y))^\Phi$ and $\mathbb{N} \models A(\underline{n})^\Phi$ imply $\mathbb{N} \models A(\underline{Sn})^\Phi$. Since we have $\mathbb{N} \models A(\underline{0})^\Phi$ and every natural number is obtained from 0 by finite applications of the successor function we easily obtain by metainduction on n that $\mathbb{N} \models A(\underline{n})^\Phi$ for all $n \in \mathbb{N}$. In a last step we show that the validity in $\mathbb{N}$ is conserved by the inference rules. From $\mathbb{N} \models A^\Phi$ and $\mathbb{N} \models (A \to B)^\Phi$ we immediately obtain $\mathbb{N} \models B^\Phi$. If F is the conclusion of an instance of the $\forall$-rule then F must be a formula $A \to \forall xB$ and we obtain by the induction hypothesis $\mathbb{N} \models (A \to B)^\Phi$ for all assignments Φ. It remains to show that $\mathbb{N} \models A^\Phi$ implies $\mathbb{N} \models B_x(\underline{n})^\Phi$ for all $n \in \mathbb{N}$. For an arbitrary $n \in \mathbb{N}$ we obtain an assignment Ψ by defining $\Psi(x) := n$, $\Psi(y) := \Phi(y)$ for all $y \neq x$ and $\Psi(X) := \Phi(X)$. Since $x \notin FV_1(A)$ we have $\mathbb{N} \models A^\Psi$. $\mathbb{N} \models (A \to B)^\Psi$ therefore also implies $\mathbb{N} \models B^\Psi$. But this implies $\mathbb{N} \models B_x(\underline{\Psi x})^\Phi$ since in $B_x(\underline{\Psi x})$ all occurences of x are replaced by the constant $\underline{\Psi x}$. Hence $\mathbb{N} \models B_x(\underline{n})^\Phi$ for all $n \in \mathbb{N}$.

The case of an $\exists$-inference is treated analogously.

3.14. **Remark**

The soundness theorem assures that $\mathbf{Z}_1$ only derives Π^1_1-*theorems* of $\mathbb{N}$, i.e. Π^1_1-sentences which are valid in the structure $\mathbb{N}$. On the other hand we know that

$\mathbf{Z}_1$ cannot derive all theorems of $\mathbb{N}$ by Gödel's theorem. As already mentioned in the beginning of this chapter, there is a tremendous gap between the set of formulas which are derivable in $\mathbf{Z}_1$ or any other formal system and the formulas which are valid in $\mathbb{N}$. The latter is a Π^1_1-complete set while the set of formally derivable formulas always is Σ^0_1. This naturally arises the question if there is a classification of the formulas which are outside the scope of $\mathbf{Z}_1$ (or of any other formal system). We are going to give such a characterization by defining a norm for the Π^1_1-sentences of $\mathscr{L}$ and then showing that only sentences whose norm is not too large may be derived in $\mathbf{Z}_1$. The definition of this norm function is the aim of the following sections.

3.15. **Exercises**

1. Let F,A be $\mathscr{L}_1$-formulas such that $(BV(A) \cup FV(A)) \cap BV(F) = \emptyset$.
Prove that $\mathbf{Z}_1 \vdash F$ already implies $\mathbf{Z}_1 \vdash F_X(\{x: A(x)\})$.

2. Suppose that L is a language such that $\mathscr{L} \subset L$ and let $\ulcorner\ \urcorner$ be an arithmetization of L. Assume that T is a consistent theory for L such that an L-formula sb(x,y) exists which satisfies $FV(sb(x,y)) = \{x,y\}$ and $T \vdash sb(\ulcorner \underline{F} \urcorner, y) \leftrightarrow y = \underline{\ulcorner F_{x_0}(\ulcorner \underline{F} \urcorner)\urcorner}$. Show that there is no L-formula Tr(x) such that $T \vdash Tr(\ulcorner \underline{F} \urcorner) \leftrightarrow F$ holds for all L-formulas F.

A formal theory T for a language $L_i(T)$ $(i = 1,2)$ is a set of formulas of $L_i(T)$. The relation $T \vdash F$ is defined inductively by:

(i) If $F \in T$ or if F is a logical axiom according to 3.7 or a formula of the shape $\forall X F \rightarrow F_X(Y)$ or $F_X(Y) \rightarrow \exists X F$, then $T \vdash F$.

(ii) If F is the conclusion of a logical inference according to 3.8. whose premises are F_0 or F_0, F_1 respectively and if $T \vdash F_j$ holds for $j = 0$ or $j = 0,1$ respectively, then $T \vdash F$.

(iii) If $i = 2$ and $T \vdash A \rightarrow F$ or $T \vdash F \rightarrow A$, then $T \vdash A \rightarrow \forall X F$ or $T \vdash \exists X F \rightarrow A$ repectively provided that $X \notin FV(A)$.

Suppose that T_1, T_2 are formal theories for $L(T_1)$ or $L(T_2)$. We say that T_2 is a conservative extension of T_1 (written as $T_1 \prec T_2$) if $L(T_1) \subset L(T_2)$ and $T_1 \vdash F \Leftrightarrow T_2 \vdash F$ holds for all $L(T_1)$-formulas F.

3. Suppose that T_1 is a subtheory of the formal theory T_2. Assume that for any $L(T_1)$-structure $\mathfrak{A}$ and any $\mathfrak{A}$-assignment Φ such that $\mathfrak{A} \models T_1^{\Phi}$ there is a $L(T_2)$-structure $\mathfrak{B}$ and a $\mathfrak{B}$-assignment Ψ which satisfies $\mathfrak{B} \models T_2^{\Psi}$ and $\mathfrak{A} \models F^{\Phi} \Leftrightarrow \mathfrak{B} \models F^{\Psi}$ for all $L(T_1)$-formulas F . Show that $T_1 \prec T_2$.
[Hint: Use the completeness theorem for first order logic.]

4. Let the formal theory $\mathbf{Z}_2$ for $\mathscr{L}_2$ be given by the axioms of $\mathbf{Z}_1$ with the scheme (IND) of complete induction expanded to all $\mathscr{L}_2$-formulas. Show that $\mathbf{Z}_2 \vdash F$ implies $\mathbf{Z}_1 \vdash F$ for all $\mathscr{L}_1$-formulas F satisfying $FV_2(F) = \emptyset$.
5. The formal theory ACA_0 for $\mathscr{L}_2$ is given by the axioms of $\mathbf{Z}_1$, with the axiom (Ind) $\forall X(\underline{0} \in X \wedge \forall x(x \in X \rightarrow \underline{S}x \in X) \rightarrow \forall x(x \in X))$ instead of the scheme (IND) of complete induction together with the scheme of arithmetical comprehension, i.e. $\exists X \forall x(x \in X \leftrightarrow F)$ for any $\mathscr{L}_1$-formula F such that $X \notin FV_2(F)$.
Show that $\mathbf{Z}_1 \prec \mathbf{ACA_0}$.
6. The formal theory $\mathbf{\Sigma^0_n}$**-IND** for $\mathscr{L}_1$ is given by the axioms of $\mathbf{Z}_1$ with the scheme (IND) of complete induction restricted to Σ^0_n-formulas.
$\mathbf{\Sigma^0_n}$**-INDR** results from $\mathbf{Z}_1$ by replacing the scheme (IND) by the rule

$$(\Sigma^0_n\text{-INDR}) \qquad \vdash A \rightarrow F_x(\underline{0}),\ \vdash A \rightarrow F \rightarrow F_x(\underline{S}x) \ \Rightarrow\ \vdash A \rightarrow \forall x F$$

for every Σ^0_n-formula F and every $\mathscr{L}_1$-formula A such that $x \notin FV(A)$.
We obtain the formal theory $\mathbf{\Sigma^0_n}$**-INDR'** by restricting the rule (Σ^0_n-INDR) to formulas F of the shape $\exists x_1 \forall x_2 \ldots Q x_n(\langle x, x_1, \ldots, x_n\rangle \in X)$ only.
Show that for every $\mathscr{L}_1$-formula F the following are equivalent:

(i) $\mathbf{\Sigma^0_n}$**-IND** $\vdash F$
(ii) $\mathbf{\Sigma^0_n}$**-INDR** $\vdash F$
(iii) $\mathbf{\Sigma^0_n}$**-INDR'** $\vdash F$.

§ 4. *The infinitary language* $\mathscr{L}_\infty$

We have already mentioned that it is impossible to obtain all Π^1_1-theorems of $\mathbb{N}$ by a finitary formal system. Therefore we are trying to (at least successfully) derive them by an infinitary system. For this purpose we reformulate the Π^1_1-sentences of $\mathbb{N}$ in an infinitary language $\mathscr{L}_\infty$ for which there is a canonical infinite derivation procedure. As far as we know *W. Tait* was the first one who used exactly this approach. The use of infinitary systems in proof theory, however, was already implicitly suggested by *D. Hilbert*. *K. Schütte* was the first one who systematically used infinitary systems in proof theoretic research. The term '*semiformal system*' is due to him.

4.1. **Basic symbols** of the language $\mathscr{L}_\infty$

1. *Logical symbols*
   (i) Countably many set variables
   (ii) The logical symbols $\bigwedge, \bigvee, \in, \notin$.
2. The *nonlogical symbols* of $\mathscr{L}_\infty$ are the same as those of $\mathscr{L}$.

4.2. **Inductive definition** of the *terms* of $\mathcal{L}_\infty$

(i) Every number constant is a term.

(ii) If $t_1,\dots,t_n$ are terms and if $\underline{f}$ is a constant for an n-ary primitive recursive function, then $(\underline{f}t_1,\dots,t_n)$ is a term.

4.3. **Inductive definition** of the *formulas* of $\mathcal{L}_\infty$

(i) If $t_1,\dots,t_n$ are terms and $\underline{R}$ is an n-ary relation constant, then $(\underline{R}t_1\dots t_n)$ is a formula.

(ii) If t is a term and X a set variable, then $(t \in X)$ and $(t \notin X)$ are formulas.

(iii) If I is a nonvoid index set and $(A_i)_{i\in I}$ a sequence of formulas, then $\bigwedge\{A_i : i \in I\}$ and $\bigvee\{A_i : i \in I\}$ are formulas.

We often write $\bigwedge_{i\in I} A_i$ and $\bigvee_{i\in I} A_i$ instead of $\bigwedge\{A_i : i \in I\}$ or $\bigvee\{A_i : i \in I\}$ respectively. As usual we write $A_1 \wedge \dots \wedge A_n$ or $A_1 \vee \dots \vee A_n$ instead of $\bigwedge\{A_1,\dots,A_n\}$ or $\bigvee\{A_1,\dots A_n\}$ respectively.

Formulas built according to one of the clauses (i) or (ii) are called *atomic*.

The *language* $\mathcal{L}_\Omega$ is the sublanguage of $\mathcal{L}_\infty$ which is obtained by restricting the index set I in clause (iii) to countable sets only.

For technical reasons we do not count the negation symbol $\neg$ among the basic symbols of the language. This, however, does not mean any restriction since we may define it in the following way.

4.4. **Inductive definition** of $\neg A$

(i) $\neg\underline{R}t_1 \dots t_n$ is the formula $\overline{R}t_1 \dots t_n$ where $\overline{R}$ means the primive recursive relation complementary to R

(ii) $\neg(t \in X) \equiv (t \notin X)$, $\quad \neg(t \notin X) \equiv (t \in X)$,

(iii) $\neg\bigwedge_{i\in I} A_i \equiv \bigvee_{i\in I} \neg A_i$, $\quad \neg\bigvee_{i\in I} A_i \equiv \bigwedge_{i\in I} \neg A_i$.

4.5. **Lemma**

$\neg\neg A \equiv A$

The proof is an easy induction on the definition of $\neg A$.

4.6. **Remark**

Sometimes we will be forced to extend the language $\mathcal{L}_\infty$ by number variables. We usually will only need finitely many number variables $x_1,\dots,x_n$. We denote this extended language by $\mathcal{L}_\infty(x_1,\dots,x_n)$. The terms of the extended language are then defined by adding the clause

(0) Every number variable is a term

to definition 4.2.

§ 5. Semantics for $\mathscr{L}_\infty$

In $\mathscr{L}_\infty$ we do not have any number variables and therefore do not need an assignment for them. An assignment for $\mathscr{L}_\infty$ is a mapping Φ from the set of variables into the power set of $\mathbb{N}$. We define t^Φ for $\mathscr{L}_\infty$-terms t as in 2.2. Since there are no free number variables in t we always have $t^\Phi = t^{\mathbb{N}}$.

5.1. **Inductive definition** of $\mathbb{N} \models F^\Phi$

(i) $\mathbb{N} \models (\underline{P}t_1...t_n)^\Phi \Leftrightarrow \chi_{\underline{P}}(t_1^{\mathbb{N}},...,t_n^{\mathbb{N}}) = 1$

(ii) $\mathbb{N} \models (t \in X)^\Phi \Leftrightarrow t^{\mathbb{N}} \in \Phi(X)$

$\mathbb{N} \models (t \notin X)^\Phi \Leftrightarrow t^{\mathbb{N}} \notin \Phi(X)$

(iii) $\mathbb{N} \models \bigwedge_{i \in I} A_i^\Phi \Leftrightarrow \mathbb{N} \models A_i^\Phi$ for all $i \in I$

(iv) $\mathbb{N} \models \bigvee_{i \in I} A_i^\Phi \Leftrightarrow \mathbb{N} \models A_i^\Phi$ for some $i \in I$

As in the semantics for $\mathscr{L}$ we denote by $\mathbb{N} \models A$ that $\mathbb{N} \models A^\Phi$ holds for all assignments Φ.

Our first goal is to obtain a more syntactical description of the validity relation for the language $\mathscr{L}_\Omega$. For this reason we are going to introduce a concept of infinitary derivations which completely characterizes the validity of $\mathscr{L}_\Omega$-formulas in $\mathbb{N}$. Again for technical reasons we will not solely derive single formulas but rather finite sets of formulas. These finite formula sets are to be interpreted as the disjunction of their members. As syntactical variables for finite sets of formulas we use capital greek letters such as $\Delta,\Gamma,\Lambda,...$. We always will write Δ,F instead of $\Delta\cup\{F\}$.

5.2. **Inductive definition** of $\models_{\bar{0}} \Delta$

(Ax1) If $\chi_P(t_1^{\mathbb{N}},...,t_n^{\mathbb{N}}) = 1$ and $(\underline{P}t_1...t_n) \in \Delta$, then $\models_{\bar{0}} \Delta$

(Ax2) If $t^{\mathbb{N}} = s^{\mathbb{N}}$ then $\models_{\bar{0}} \Delta, t \in X, s \notin X$

($\bigwedge$) If $\models_{\bar{0}} \Delta, A_i$ for all $i \in I$, then $\models_{\bar{0}} \Delta, \bigwedge\{A_i : i \in I\}$

($\bigvee$) If $\models_{\bar{0}} \Delta, A_i$ for some $i \in I$, then $\models_{\bar{0}} \Delta, \bigvee\{A_i : i \in I\}$

5.3. **Soundness theorem** for $\vdash_{\overline{0}}$
If $\vdash_{\overline{0}} \Delta$, *then* $\mathbb{N} \models \bigvee \{F : F \in \Delta\}$.

Proof
We prove the claim by induction on the definition of $\vdash_{\overline{0}} \Delta$.
(Ax1) Then Δ contains an atomic formula $(\underline{P}t_1 \ldots t_n)$ such that $\chi_P(t_1^{\mathbb{N}},\ldots,t_n^{\mathbb{N}}) = 1$ and by 5.1.(i) we obtain $\mathbb{N} \models (\bigvee \{F : F \in \Delta\})^\Phi$ for every assignment Φ.
(Ax2) For any assignment Φ and term s and t such that $s^{\mathbb{N}} = t^{\mathbb{N}}$ we either have $\mathbb{N} \models (t \in X)^\Phi$ or $\mathbb{N} \models (s \notin X)^\Phi$. Hence $\mathbb{N} \models (\bigvee \Delta \vee (t \in X) \vee (s \notin X))^\Phi$.
$(\bigwedge)$ By the induction hypothesis we have $\mathbb{N} \models (\bigvee \Delta \vee A_i)^\Phi$ for all $i \in I$. If $\mathbb{N} \models \bigvee \Delta^\Phi$, then also $\mathbb{N} \models (\bigvee \Delta \vee \bigwedge_{i \in I} A_i)^\Phi$. If $\mathbb{N} \not\models \bigvee \Delta^\Phi$, then $\mathbb{N} \models A_i^\Phi$ for all $i \in I$. Hence $\mathbb{N} \models \bigwedge_{i \in I} A_i^\Phi$ which implies $\mathbb{N} \models (\bigvee \Delta \vee \bigwedge_{i \in I} A_i)^\Phi$.
$(\bigvee)$ This case is dual to the case of $(\bigwedge)$.

5.4. **Completeness theorem** for $\vdash_{\overline{0}}$
If F is an $\mathscr{L}_\Omega$*-formula such that* $\mathbb{N} \models F$, *then* $\vdash_{\overline{0}} F$.

Due to the presence of free set variables the proof of the completeness theorem is not trivial. We need some preparations for the proof. In this section we briefly write formula instead of $\mathscr{L}_\Omega$-formula.

5.5. **Definition**

(i) We call a finite formula sequence $\Delta = (A_0,\ldots,A_l)$ *reducible* if it contains a formula of the form $\bigwedge_{i \in I} A_i$ or $\bigvee_{i \in I} A_i$. These formulas are the *redexes* of Δ.

(ii) Suppose that $\Delta = (A_0,\ldots,A_l)$ is reducible. A redex $A_k \in \Delta$ is *distinguished* in Δ if there is no redex A_i in Δ such that $0 \le i < k$.

(iii) If Δ is reducible we obtain Δ^r from Δ by cancelling the distinguished redex.

5.6. **Definition**
For a finite sequence Δ of formulas we define a tree B_Δ together with a label-function $\delta : B_\Delta \to \{\Gamma : \Gamma$ is a finite sequence of formulas$\}$. We call B_Δ the *quasi-deductiontree* of Δ,

(i) $\langle\,\rangle \in B_\Delta$ and $\delta(\langle\,\rangle) := \Delta$

(ii) If $\sigma \in B_\Delta$ and $\delta(\sigma)$ is not reducible or an axiom according to (Ax1) or (Ax2), then it is $\sigma * \langle j \rangle \notin B_\Delta$ for all $j < \omega$ (i.e. σ is a top node of the tree).

(iii) If $\sigma \in B_\Delta$ and $\delta(\sigma)$ reducible with distinguished redex $\bigwedge_{i \in I} A_i$, then $\sigma * \langle i \rangle \in B_\Delta$

for all $i \in I$ and $\delta(\sigma * \langle i \rangle) = \delta(\sigma)^r, A_i$

(iv) If $\sigma \in B_\Delta$ and $\delta(\sigma)$ is reducible with distinguished redex $\bigvee_{i \in I} A_i$ and there is a minimal $k_0 \in I$ (in a fixed enumeration of I) such that A_{k_0} does not occur in $B_{\Delta_\sigma} := \bigcup\{\delta(\tau) : \tau \in B_\Delta \wedge \tau \subset \sigma\}$, then $\sigma * \langle k_0 \rangle \in B_\Delta$ and $\delta(\sigma * \langle k_0 \rangle) = \delta(\sigma)^r, A_{k_0}, \bigvee_{i \in I} A_i$. Otherwise we put $\sigma * \langle 0 \rangle \in B_\Delta$ and define $\delta(\sigma * \langle 0 \rangle) = \delta(\sigma)^r, A_0$

A path through the quasideductiontree (B_Δ, δ) is called a *quasideductionpath* of Δ.

We sometimes will not distinguish between Δ a sequence and Δ as a finite set. As in the following lemma, however, the context always makes clear which meaning is to be taken.

5.7. **Syntactical mainlemma**

If every quasideductionpath of Δ contains an axiom, then $\models_{\overline{0}} \Delta$.

Proof

By the hypothesis that every quasideductionpath of Δ contains an axiom we have that every path in B_Δ is finite. Hence B_Δ is wellfounded and we show $\models_0 \delta(\sigma)$ for $\sigma \in B_\Delta$ by induction on B_Δ.

1. If $\delta(\sigma)$ is an axiom we trivially have $\models_{\overline{0}} \delta(\sigma)$.

Otherwise we know that $\delta(\sigma)$ is reducible.

2. If the distinguished redex of $\delta(\sigma)$ is $\bigwedge_{i \in I} A_i$, then we have $\sigma * \langle i \rangle \in B_\Delta$ for all $i \in I$ and obtain $\models_{\overline{0}} \delta(\sigma * \langle i \rangle)$ by the induction hypothesis, i.e. $\models_{\overline{0}} \delta(\sigma)^r, A_i$ for all $i \in I$. By an inference $(\bigwedge)$ this implies $\models_{\overline{0}} \delta(\sigma)^r, \bigwedge_{i \in I} A_i$, i.e. $\models_{\overline{0}} \delta(\sigma)$.

3. If the distinguished redex is $\bigvee_{i \in I} A_i$, then there is a $k_0 \in I$ such that $\sigma * \langle k_0 \rangle \in B_\Delta$. By the induction hypothesis we have $\models_{\overline{0}} \delta(\sigma)^r, A_{k_0}, \bigvee_{i \in I} A_i$ and this implies $\models_{\overline{0}} \delta(\sigma)$ by an $(\bigvee)$-inference.

5.8. **Semantical mainlemma**

Suppose that a finite sequence Δ of formulas has a quasideductionpath which does not contain an axiom. Then there is an assignment Φ such that $\mathbb{N} \not\models F^\Phi$ for all $F \in \Delta$.

Proof

Pick a path f in B_Δ which does not contain an axiom. Then f has the following properties:

(1) If $\sigma \in f$ and $P \in \delta(\sigma)$ is atomic, then $P \in \delta(\tau)$ holds for all $\sigma \subset \tau \in f$.

This is obvious since P is no redex and therefore never will be cancelled.

(2) If $\sigma \in f$ *and* $R \in \delta(\sigma)$ *is a redex, then there is a* $\tau \in f$ *such that* $\sigma \subset \tau$ *and* R *is distinguished in* $\delta(\tau)$.

The proof of (2) is by induction on the number of redexes which have a smaller index than R. If this number is 0, then R is already distinguished in $\delta(\sigma)$. Otherwise let R_0 be distinguished in $\delta(\sigma)$. Then there is a j such that $\sigma * \langle j \rangle \in f$ and R_0 either is cancelled or is the redex with maximal index in $\delta(\sigma * \langle j \rangle)$ (cf. 5.6.(iv)). By the induction hypothesis we then have a $\tau \in f$ such that $\sigma \subset \sigma * \langle j \rangle \subset \tau$ and R is distinguished in $\delta(\tau)$.

(3) If $\sigma \in f$ *and* $(\bigwedge_{i \in I} A_i) \in \delta(\sigma)$, *then there is an* $i \in I$ *and* $\tau \in f$ *such that* $A_i \in \delta(\tau)$.

By (2) we have a $\tau_0 \in f$ such that $(\bigwedge_{i \in I} A_i)$ is the distinguished redex in $\delta(\tau_0)$. By definition 5.6.(iii) we then have $\tau_0 * \langle i \rangle \in B_\Delta$ for all $i \in I$. Since f is a path through B_Δ there is an $i \in I$ such that $\tau_0 * \langle i \rangle \in f$ and we have $A_i \in \delta(\tau_0 * \langle i \rangle)$. We define $\tau := \tau_0 * \langle i \rangle$.

(4) If $\sigma \in f$ *and* $(\bigvee_{i \in I} A_i) \in \delta(\sigma)$, *then for every* $i \in I$ *there is a* $\tau_i \in f$ *such that* $A_i \in \delta(\tau_i)$.

Assume that there is an $i \in I$ such that $A_i \notin \delta(\tau)$ for all $\tau \in f$. Choose i_0 minimal with this property. This means $\forall j < i_0 \exists \tau_j \in f \, A_j \in \delta(\tau_j)$. Let τ_{i_0} be the union of all those τ_j (as finite sequences). Then $\tau_{i_0} \in f$ and by definition 5.6. we have $(\bigvee_{i \in I} A_i) \in \delta(\tau_{i_0})$. By (2) there is a $\sigma_0 \in f$ such that $\tau_{i_0} \subset \sigma_0$ and $(\bigvee_{i \in I} A_i)$ is distinguished in $\delta(\sigma_0)$. By definition 5.6.(iv) we then have $\tau_{i_0} * \langle i_0 \rangle \in f$ and $A_{i_0} \in \delta(\tau_{i_0} * \langle i_0 \rangle)$. Contradiction.

Now we define an assignment Φ by:

$$\Phi(X) = \{n \in \mathbb{N} : \exists t(t^{\mathbb{N}} = n \wedge (\exists \sigma \in f)((t \notin X) \in \delta(\sigma)))\}$$

Then we have

(5) $\mathbb{N} \nvDash F^\Phi$ *for all* $\sigma \in f$ *and* $F \in \delta(\sigma)$.

(5) is proved by induction on the length of F.

1. If $F \equiv (\underline{P} t_1 \ldots t_n)$, then $\mathbb{N} \nvDash (\underline{P} t_1 \ldots t_n)^\Phi$ because otherwise $\delta(\sigma)$ was an axiom.

2. Assume $F \equiv (t \in X)$. If $t^{\mathbb{N}} \in \Phi(X)$, then there is a term s equivalent to t and a $\tau_0 \in f$ such that $(s \notin X) \in \delta(\tau_0)$. But then by (1) $\{t \in X, s \notin X\} \subset \delta(\tau)$ where $\tau = \max(\sigma, \tau_0) \in f$ and we obtain an axiom in f. Hence $t^{\mathbb{N}} \notin \Phi(X)$ which entails $\mathbb{N} \nvDash (t \in X)^\Phi$.

3. $F \equiv (t \notin X)$. Then $t^{\mathbb{N}} \in \Phi(X)$ and $\mathbb{N} \nvDash (t \notin X)^\Phi$.

4. $F \equiv (\bigwedge_{j \in I} A_j)$. By (3) there is a $j \in I$ and $\tau \in f$ such that $A_j \in \delta(\tau)$. By the induction hypothesis we therefore obtain $\mathbb{N} \nvDash A_j^\Phi$ and this entails $\mathbb{N} \nvDash (\bigwedge_{j \in I} A_j)^\Phi$.

5. $F \equiv (\bigvee_{j \in I} A_j)$. By (4) there is a $\tau_j \in f$ such that $A_j \in \delta(\tau_j)$ for every $j \in I$ and we

obtain by the induction hypothesis $\mathbb{N} \not\models A_j^{\Phi}$ for all $j \in I$ and this implies $\mathbb{N} \not\models (\bigvee_{j \in I} A_j)^{\Phi}$.

Proof of 5.4.
If $\not\models_0 F$, then by the syntactical mainlemma there is a quasideductionpath of F which does not contain an axiom. By the semantical mainlemma we then obtain an assignment Φ such that $\mathbb{N} \not\models F^{\Phi}$, i.e. $\mathbb{N} \not\models F$.

5.9. **Exercises**
Show that $\mathbb{N} \models F$ implies $\models_0 F$ for every $\mathcal{L}_{\Omega}$-formula F not containing free set variables without using 5.4.

§ 6. Ordinals

The relation $\models_0 \Delta$ is more syntactically defined than the relation $\mathbb{N} \models F$ in that sense that its definition does not refer to assignments. For finite sets of sentences both definitions essentially coincide. The question is now if there really is more information in the relation $\models_0$ than we already had in the relation $\models$. A derivation $\models_0 \Delta$ may be visualized as an infinite branching wellfounded tree. This was the imagination we had in the definition of the quasideduction-trees. The complexity of this wellfounded tree then is a measure for the complexity of the validity of the formula set Δ. Trees, however, are not easy to compare. Therefore we are looking for a characteristic magnitude of a wellfounded tree in which the essential information of the tree is incorporated. Such a magnitude will be given by the depth of the tree. We then may call a tree more complicated than another if it has a larger depth.
In a wellfounded tree every path is finite. We therefore may define the depth of a tree as the length of its maximal path. Our derivation trees, however, are ω-branching infinite trees. If we look for example at the following tree

```
                .
            4   .
         3  3   .
      2  2  2   .
   1  1  1  1   .
0  0  0  0  0   . . .

        ⟨ ⟩
```

we immediately see that every path is finite so the tree is wellfounded but it obviously does not possess a maximal path. This shows that natural numbers will not suffice for the description of the depth of infinitely branching trees. Therefore we have to improve our concept of numbers.

A natural number n has two different aspects. On the one side the *quantity aspect* which describes that an object of magnitude n has just n elements and on the other hand an *order aspect* which describes that the elements of a set of n elements may be ordered as 0,1,...,n−1. The difference between both aspects of a natural number, however, is a bit hazy since, modulo permutations, there is only one way to order a finite set. The situation changes in the case of an infinite set. As an example we regard the set of all natural numbers. In their usual order they look like 0,1,2,... but we may order them as 1,2,3,...,0 or 0,2,4,6,...,1,3,5,7,9,..., where the order relation is given by the convention that the elements on the left are smaller than those on the right. Since in all orderings we used the same set the quantity aspect will not change while the order aspect did.

If we try to extend counting into the transfinite it is exactly the order aspect of a number we are interested in. We first want to count all natural numbers and, having completed them, go on counting. Such a counting into the transfinite is for instance given by the order 1,2,3,...,0 where we may first complete the counting of all natural numbers and then count one more element.

Of course not every ordering allows counting. If, for example, we regard the ordering of the non negative rational numbers, then we just may count 0 and then do not know how to continue since there is no next element following the element 0. Only those orderings will allow counting which have the property that, after taking away arbitrarily many elements, the remaining set of elements always has a least element provided it is not empty. Such orderings are called *wellorderings* and their formal definition runs as follows.

A binary relation $\prec$ is called an *ordering* of a set A if it satisfies the following conditions

$\forall x \in A(\neg x \prec x)$ *(irreflexivity)*

$\forall x \in A \forall y \in A \forall z \in A(x \prec y \wedge y \prec z \rightarrow x \prec z)$ *(transitivity)*

$\forall x \in A \forall y \in A(x \prec y \vee y \prec x \vee x = y)$ *(linearity)*

We denote an ordering by $(A,\prec)$ and call A the field of $\prec$.

An ordering $(A,\prec)$ is called a *wellordering* of A if it also satisfies the additional condition

$\forall X \subset A(X \neq \emptyset \rightarrow \exists y \in X \forall t \in X(y \neq t \rightarrow y \prec t)$. *(wellfoundedness)*

Two orderings are *equivalent* if there is an order preserving map from the field of the first ordering onto the field of the second. An *ordertype* is an equivalence class of an ordering. As we have seen we need the ordertypes of wellorderings for extending counting into the transfinite. These ordertypes are called *ordinals.* Our hitherto described concept of ordinal is based on a completely naive understanding of the universe of sets. If one tries to make this precise within the framework of an axiomatic set theory one immediately runs into troubles since the equivalence class of an ordering will not be a set. But of course we want the ordinals to be elements of the universe, i.e. to be sets. The problem may be solved by selecting a characteristic representative for an ordinal. The question then is which one to take. In the language of set theory one usually has the symbol $\in$ as the only nonlogical symbol. So it seems to be reasonable to take as representatives those sets which are wellordered by the relation $\in$ itself. That this in fact is a canonical choice becomes even clearer by pursuing the naive theory of ordinals a bit further.

Suppose that $(A,\prec)$ is a wellordering. $(U,\prec)$ is called a *segment* of $(A,\prec)$ if $U \subset A$ and $a \prec b \in U$ already implies $a \in U$. A segment $(U,\prec)$ is *proper* if $U \neq A$. It is obvious that $(U,\prec)$ is a proper segment if and only if there is a $b \in A$ such that $U = \{a \in A : a \prec b\}$.

On the wellorderings we define an orderrelation $<$ by

$(A,\prec_A) < (B,\prec_B) :\Leftrightarrow$ There is a proper segment of $(B,\prec_B)$ which is equivalent to $(A,\prec_A)$.

It is not very hard to prove that this relation in fact is a wellordering. But since all these considerations only serve as heuristics we will omit the proof. If we look at a representative $(A,\prec)$ of an ordinal β we notice that the ordertype of the set $\{(B,\prec_B) : (B,\prec_B) < (A,\prec)\}$ is exactly the ordinal β, i.e. $\beta = \{\alpha : \alpha < \beta\}$ and this implies that $\alpha < \beta$ holds if and only if $\alpha \in \beta$. Therefore the sets well-ordered by the $\in$-relation are in fact canonical representatives for ordinals.

We are now going to develop the theory of ordinals on the basis of a set theory which needs not to be specified here. The experienced reader may think of **ZFC**. Since there usually is an axiom of foundation in a set theory it suffices to define ordinals as hereditarily transitive sets. We denote the class of ordinals by On. Facts about On which only can be proved on the basis of the set theory will be stated here as *basic properties* without proof. These basic properties may be taken as axioms for On. But we want to emphasize that it is not our aim to give a complete axiomatization of the class On.

6.1. Basic properties of the class On

(O1) On is a transitive class which is wellordered by $\in$. As usual we write $\alpha < \beta$ instead of $\alpha \in \beta$.

(O2) On is unbounded, i.e. $\forall \alpha \in On\, \exists \beta \in On(\alpha < \beta)$.

(O3) If $M \subset On$ and if there is an $\alpha \in On$ and an 1-1 mapping from M onto α, then M is bounded in On (i.e. $\exists \beta \in On\ \forall \xi \in M(\xi \le \beta)$).

(O4) $\exists \alpha (\alpha \in On)$.

6.2. Theorem (Transfinite induction)

If $\forall \xi < \alpha\ A(\xi) \rightarrow A(\alpha)$ *holds for all* $\alpha \in On$, *then* $\forall \xi \in On\, A(\xi)$.

Proof

Assume that $\{\xi \in On : \neg A(\xi)\} \neq \emptyset$ and define $\alpha := \min\{\xi \in On : \neg A(\xi)\}$. Then we have $\forall \xi < \alpha\ A(\xi)$ and by hypothesis we obtain $A(\alpha)$. Contradiction.

6.3. Theorem

(i) There is a least ordinal which will be denoted by 0.

(ii) For every ordinal α *there is a least ordinal* β *such that* $\alpha < \beta$. *We call this ordinal the successor of* α *and denote it by* α'.

Proof

(i) is obvious because of $On \neq \emptyset$.

(ii) is an immediate consequence of (O1) and (O2).

6.4. Lemma

(i) $\alpha < \alpha'$

(ii) $\alpha < \beta \Leftrightarrow \alpha' < \beta'$

(iii) $\alpha < \beta \Rightarrow \alpha' \le \beta$

(iv) $\alpha < \beta' \Rightarrow \alpha \le \beta$

Proof

(i) holds by definition.

(ii) If $\alpha < \beta$, then $\beta \in \{\xi : \alpha < \xi\}$ which implies $\alpha' := \min\{\xi : \alpha < \xi\} \le \beta < \beta'$. From this we obtain by contraposition $\alpha' \le \beta' \Rightarrow \alpha \le \beta$. Since $\alpha = \beta$ entails $\alpha' = \beta'$ we have that $\alpha' < \beta'$ already implies $\alpha < \beta$.

(iii) is obvious by definition.

(iv) If $\alpha < \beta'$, then by (iii) $\alpha' \leq \beta'$ which implies $\alpha \leq \beta$.

6.5. Definition

An ordinal λ which neither is 0 nor the successor of another ordinal is called a *limit ordinal.* The class of limit ordinals is denoted by Lim.

6.6. Lemma

$\lambda \in \mathrm{Lim}$ *and* $\alpha < \lambda$ *imply* $\alpha' < \lambda$.

Proof

$\alpha < \lambda$ implies $\alpha' \leq \lambda$. Since $\lambda \in \mathrm{Lim}$ we have $\lambda \neq \alpha'$ which entails $\alpha' < \lambda$.

As a consequence of lemma 6.6 we obtain $\forall \alpha \in \lambda \, \exists \xi \in \lambda \, (\alpha < \xi)$ for limit ordinals λ. This shows that limit ordinals reflect the basic property (O2) of the class On of ordinals. The existence of limit ordinals cannot be proved. In an axiomatic set theory we therefore need an infinity axiom which requires the existence of at least one limit ordinal. In our notation this axiom would just be $\exists \lambda \, (\lambda \in \mathrm{Lim})$. We will, however, need a stronger form of the axiom of infinity which in the framework of **ZFC** follows from the weaker one. Since set theory is not the topic of this lecture we may as well take the stronger form of the axiom of infinity as basic property.

6.7. Definition

An ordinal $\varkappa$ is *regular* if it satisfies

($\mathbb{R}$1) $\varkappa \in \mathrm{Lim}$

($\mathbb{R}$2) Every $M \subset \varkappa$ for which there is an 1–1 mapping from M onto some $\alpha < \varkappa$ is bounded in $\varkappa$, i.e.. $\exists \eta < \varkappa \, \forall \xi \in M \, (\xi \leq \eta)$.

A regular ordinal reflects the property (O3) of the ordinal universe On. By $\mathbb{R}$ we denote the class of regular ordinals. The axiom of infinity which we need here is the following basic property

(05) $\mathbb{R}$ is unbounded in On, i.e. $\forall \xi \in \mathrm{On} \, \exists \eta \in \mathbb{R} (\xi < \eta)$.

6.8. Theorem

If $M \subset \mathrm{On}$ *is bounded in* On, *then there is a least upper bound for* M. *We denote this bound by* $\sup M$.

Proof

The class $\bar{M} = \{\xi \in On : \forall \eta \in M(\eta \leq \xi)\}$ is not empty and we define $\sup M := \min \bar{M}$.

6.9. **Lemma**

If $\varkappa \in \mathbb{R}$, $M \subset \varkappa$ *and there is an* 1-1 *mapping of* M *onto some* $\alpha < \varkappa$, *then* $\sup M < \varkappa$.

Proof

By ($\mathbb{R}$2) M is bounded in $\varkappa$. Hence $\sup M < \varkappa$.

6.10. **Lemma**

If $\beta < \sup M$, *then there is an* $\eta \in M$ *such that* $\beta < \eta$.

Proof

Using the terminology of 6.8. we have $\beta < \sup M \Rightarrow \beta \notin \bar{M} \Rightarrow \exists \eta \in M(\beta < \eta)$.

6.11. **Theorem**

If $M \neq \emptyset$ *is bounded in* On, *then we either have* $\sup M \in M$ *or* $\sup M \in Lim$.

Proof

If $\sup M = 0$ then $M \neq \emptyset$ entails $M = \{0\}$, i.e. $\sup M = \max M$. If $\sup M = \alpha'$ then by 6.10. there is an $\eta \in M$ such that $\alpha < \eta \leq \alpha'$. Hence $\eta = \alpha'$, i.e. $\alpha' \in M$ and $\sup M = \max M$.

6.12. **Definition**

(i) $\mathbb{N} = \bigcap\{M \subset On : 0 \in M \wedge \forall \xi \in M(\xi' \in M)\}$

(ii) $\omega := \sup \mathbb{N}$

6.13. **Lemma**

(i) $0 \in \mathbb{N} \wedge \forall \xi(\xi \in \mathbb{N} \rightarrow \xi' \in \mathbb{N})$.

(ii) $\omega \notin \mathbb{N}$ *and there is no limit ordinal in* $\mathbb{N}$.

(iii) $\mathbb{N}$ *is a segment of* On.

(iv) ω *is the least limit ordinal and it is* $\mathbb{N} = \omega$.

Proof

Define $\mathfrak{M} = \{M \subset On : 0 \in M \wedge \forall \xi \in M(\xi' \in M)\}$. For any limit ordinal α we have $\alpha \in M$ by 6.6. So $\mathfrak{M}$ is not empty and $\mathbb{N}$ is bounded which implies that ω is defined.

(i) Since $0 \in M$ for all $M \in \mathfrak{M}$ we obtain $0 \in \bigcap\mathfrak{M} = \mathbb{N}$. If $\xi \in \mathbb{N}$ so $\xi \in M$ for all $M \in \mathfrak{M}$ which entails $\xi' \in M$ for all $M \in \mathfrak{M}$. Hence $\xi' \in \bigcap\mathfrak{M} = \mathbb{N}$.

(ii) Assume that $\omega \in \mathbb{N}$. But then by (i) $\omega' \in \mathbb{N}$ and it is $\omega < \omega'$ in contradiction to $\omega = \sup\mathbb{N}$. Assume that there is an $\alpha \in \mathbb{N} \cap \mathrm{Lim}$. Then we have $0 \in \alpha \cap \mathbb{N}$ and by (i) and 6.6. $\xi \in \alpha \cap \mathbb{N}$ implies $\xi' \in \alpha \cap \mathbb{N}$. Hence $\alpha \cap \mathbb{N} \in \mathfrak{M}$ and therefore $\mathbb{N} \subset \alpha \cap \mathbb{N} \subset \mathbb{N}$. This, however, contradicts $\alpha \in \mathbb{N}$.

(iii) In a first step we prove

(1) $\beta' \in \mathbb{N} \Rightarrow \beta \in \mathbb{N}$.

Assume $\beta \notin \mathbb{N}$. We have $0 \in \beta' \cap \mathbb{N}$. If $\xi \in \beta' \cap \mathbb{N}$ we obtain $\xi \leq \beta$ and from $\beta \notin \mathbb{N}$ even $\xi < \beta$. Then $\xi' < \beta'$ and it follows $\xi' \in \beta' \cap \mathbb{N}$. Hence $\beta' \cap \mathbb{N} \in \mathfrak{M}$ and therefore $\mathbb{N} \subset \beta' \cap \mathbb{N} \subset \mathbb{N}$ in contradiction to $\beta' \in \mathbb{N}$.

From (1) we obtain

(2) $\alpha < \beta \in \mathbb{N} \Rightarrow \alpha \in \mathbb{N}$

by induction on β.

For $\beta = 0$ (2) holds trivially. For $\beta \neq 0$ there is a β_0 such that $\beta = \beta_0'$ by (ii). By (1) we have $\beta_0 \in \mathbb{N}$. Now $\alpha < \beta$ implies $\alpha \leq \beta_0$. If $\alpha = \beta_0$, then $\alpha \in \mathbb{N}$ and for $\alpha < \beta_0$ we obtain $\alpha \in \mathbb{N}$ from the induction hypothesis.

(iv) $\mathbb{N}$ is a segment of On which does not contain a limit ordinal. By (O5), however, there are regular ordinals and thus also limit ordinals. So $\mathbb{N}$ is bounded in On and $\sup\mathbb{N}$ exists. By (ii) and 6.11. we have $\sup\mathbb{N} \in \mathrm{Lim}$, i.e. $\omega \in \mathrm{Lim}$. Since $\mathbb{N}$ does not contain limit ordinals ω has to be the least limit ordinal. To show $\mathbb{N} = \omega$ we notice that we already have $\mathbb{N} \subset \omega$. For $\alpha < \omega$ we obtain by 6.10. an $\eta \in \mathbb{N}$ such that $\alpha < \eta$. By (iii) we have $\alpha \in \mathbb{N}$. Hence also $\omega \subset \mathbb{N}$.

6.14. Definition

We define $\aleph_1 := \min\{\varkappa \in \mathbb{R} : \omega < \varkappa\}$. A set M is *countable* if there is an 1–1 mapping from M onto some $\alpha < \aleph_1$. This definition of countability coincides with the notion of countability which we already used intuitively in §4. Instead of $\aleph_1$ we frequently write Ω_1 or even shorter Ω.

6.15 Lemma

(i) *If* M *is a proper segment of* On, *then there is a* $\beta \in \mathrm{On}$ *such that* $M = \beta$.

(ii) *If* M *is a proper segment of a regular ordinal* $\varkappa$, *then there is an ordinal* $\beta < \varkappa$ *such that* $M = \beta$.

(iii) *if* $M \subsetneq \Omega$ *is a segment, then* M *is countable.*

Proof

(i) Define $\bar{M} = \{\xi \in On : \xi \notin M\} \neq \emptyset$ and $\beta := \min \bar{M}$. Then $\beta \subset M$. We show that $\xi \in M$ entails $\xi < \beta$. If not we had $\xi \in M \wedge \beta \leq \xi$ which would imply $\beta \in M$ since M is a segment. This, however, contradicts $\beta \in \bar{M}$.

(ii) Define $\bar{M} := \{\xi < \varkappa : \xi \notin M\}$ and $\beta := \min \bar{M} < \varkappa$. Then β possesses all the required properties.

(iii) is an immediate consequence of (ii).

6.16. **Transfinite induction** (second version)

Suppose we have

(i) $F(0)$,

(ii) $\forall \xi (F(\xi) \rightarrow F(\xi'))$

and

(iii) $\forall \lambda \in Lim (\forall \eta < \lambda\, F(\eta) \rightarrow F(\lambda))$.

Then it follows $\forall \xi \in On\, F(\xi)$.

Proof

Assume that $M = \{\xi \in On : \neg F(\xi)\} \neq \emptyset$. Put $\alpha := \min M$. Then we have $\alpha \neq 0$ by (i). If $\alpha = \beta'$ we had $\beta \notin M$. But this means $F(\beta)$ which by (ii) entails $F(\alpha)$ in contradiction to $\alpha \in M$. If $\alpha \in Lim$ we had $\forall \eta < \alpha\; F(\eta)$ and by (iii) also $F(\alpha)$ producing the same contradiction. Hence $M = \emptyset$ and the theorem is proved.

In the following we will be forced also to deal with partial functions from the ordinals into the ordinals. As in recursion theory partial functions arise in regarding minima of sets which possibly are empty. The reader who does not like partial functions may imagine a set $\infty \notin On$, define $\min \emptyset = \infty$ and expand the $<$-relation by $\alpha < \infty$ for all $\alpha \in On$.

As in recursion theory we define

$f(\alpha) \simeq g(\alpha) :\Leftrightarrow (\alpha \in dom(f) \cap dom(g) \wedge f(\alpha) = g(\alpha)) \vee (\alpha \notin dom(f) \wedge \alpha \notin dom(g))$, where $dom(f) := \{\alpha \in On : f(\alpha) \in On\}$.

6.17. **Definition**

Suppose that $M \subset On$. We recursively define a partial function $OD_M : On \rightarrow M$ by

(i) $OD_M(0) \simeq \min M$

(ii) $OD_M(\alpha') \simeq \min\{\xi \in M : OD_M(\alpha) < \xi\}$

(iii) $OD_M(\lambda) \simeq \min\{\xi \in M : \sup\{OD_M(\eta) : \eta < \lambda\} \leq \xi\}$ for $\lambda \in Lim$.

6.18. Lemma

The function OD_M *is uniquely defined by (i), (ii) and (iii).* $\mathrm{dom}\, OD_M$ *is a segment of* On *and* OD_M *is order preserving.*

Proof

Assume that f_1 and f_2 are partial functions from On to M which satisfy (i), (ii) and (iii). By induction on α we show

(1) $\alpha \in \mathrm{dom}\, f_1 \Rightarrow \alpha \in \mathrm{dom}\, f_2 \wedge f_1(\alpha) = f_2(\alpha)$

and

(2) $\alpha \in \mathrm{dom}\ f_1 \wedge \beta < \alpha \rightarrow \beta \in \mathrm{dom}\, f_1 \wedge f_1(\beta) < f_1(\alpha)$.

For $\alpha = 0$ (2) holds trivially. If $\alpha \in \mathrm{dom}\, f_1$, then $M \neq \emptyset$ and it follows $f_1(\alpha) \simeq \min M \simeq f_2(\alpha)$ which implies $\alpha \in \mathrm{dom}\, f_2$ and $f_1(\alpha) = f_2(\alpha)$.
If $\alpha = \beta_0'$ and $\alpha \in \mathrm{dom}\, f_1$, then $f_1(\alpha) = \min\{\xi \in M : f_1(\beta_0) < \xi\}$. This implies that $\{\xi \in M : f_1(\beta_0) < \xi\} \neq \emptyset$. Hence $\beta_0 \in \mathrm{dom}(f_1)$. $\beta < \alpha$ entails $\beta \le \beta_0$. If $\beta = \beta_0$, then $\beta \in \mathrm{dom}\, f_1$ and $f_1(\beta) < f_1(\alpha)$. If $\beta < \beta_0$, then $\beta \in \mathrm{dom}\, f_1$ and $f_1(\beta) < f_1(\beta_0) < f_1(\alpha)$ by the induction hypothesis for (2). This proves (2). (1) now follows from the induction hypothesis $\beta_0 \in \mathrm{dom}\, f_2$ and from $f_1(\alpha) \simeq \min\{\xi \in M : f_1(\beta_0) < \xi\} \overset{\text{i.h.}}{\simeq} \min\{\xi \in M : f_2(\beta_0) < \xi\} \simeq f_2(\alpha)$.
For $\alpha \in \mathrm{Lim}$ we have $f_1(\alpha) = \min\{\eta \in M : \sup\{f_1(\xi) : \xi < \alpha\} \le \eta\}$. Since $\alpha \in \mathrm{dom} f_1$, $\sup\{f_1(\xi) : \xi < \alpha\}$ has to be defined. Hence $\alpha \subset \mathrm{dom}\, f_1$. If $\beta < \alpha$, then there is a $\beta_0 < \alpha$ such that $\beta < \beta_0$. By the induction hypothesis we then obtain $f_1(\beta) < f_1(\beta_0) \le f_1(\alpha)$. This proves (2). From the induction hypothesis for (1) we obtain $\alpha \subset \mathrm{dom}\, f_2(\xi)$ and $f_1 \restriction \alpha = f_2 \restriction \alpha$. Hence $\sup\{f_1(\xi) : \xi < \alpha\} = \sup\{f_2(\xi) : \xi < \alpha\}$ and it immediately follows $\alpha \in \mathrm{dom}\, f_2$ and $f_2(\alpha) = f_1(\alpha)$.

6.19. Remark

The definition of the function OD_M is a special case of the principle of definition by *transfinite recursion*. This principle is a generalization of the principle of definition by primitive recursion which we already know from 1.2.(iv). Within a framework of axiomatic set theory such as **ZFC** the existence and the uniqueness of the function defined by transfinite recursion becomes provable.

6.20. Definition

For $M \subset$ On the function OD_M and therefore also $\mathrm{dom}\, OD_M$ are uniquely determined. We call $\mathrm{dom}\, OD_M$ the *ordertype* of M and denote it by Otyp(M). According to 6.15. we either have Otyp(M) = On or the existence of an ordinal β such that Otyp(M) = β. We define $\mathrm{ord}_M := OD_M \restriction \mathrm{Otyp}(M)$ and call ord_M the *enumerating function* of M. By 6.18. ord_M is uniquely determined by M.

6.21. **Lemma**

ord_M *is an order preserving function from* Otyp(M) *onto* M.

Proof

In 6.18. we have already shown that ord_M is order preserving. All there remains to show is that ord_M is onto. Therefore we prove

(1) $\eta \in M \rightarrow \exists \alpha \in \mathrm{Otyp}(M)(\mathrm{ord}_M(\alpha) = \eta)$

by induction on η.

Put $\beta := \mathrm{Otyp}(M \cap \eta)$.

1. $\beta = 0$. Then $\eta = \min M$ and therefore $\eta = \mathrm{ord}_M(0)$.
2. $\beta = \beta'_0$. Then $\mathrm{ord}_M(\beta) = \min\{\xi \in M : \mathrm{ord}_M(\beta_0) < \xi\}$. Since $\beta_0 < \beta$ we have $\mathrm{ord}_M(\beta_0) < \eta$ and therefore $\mathrm{ord}_M(\beta) \le \eta$. If we assume $\mathrm{ord}_M(\beta) < \eta$, then we have $\mathrm{ord}_M(\beta) \in M \cap \eta$ and by the induction hypothesis obtain an ordinal $\nu < \beta$ such that $\mathrm{ord}_M(\beta) = \mathrm{ord}_M(\nu)$ in contradiction to the fact that ord_M is order preserving. Hence $\mathrm{ord}_M(\beta) = \eta$.
3. $\beta \in \mathrm{Lim}$. Then $\mathrm{ord}_M(\beta) = \min\{\xi \in M : \sup\{\mathrm{ord}_M(\zeta) : \zeta < \beta\} \le \xi\}$ and by $\sup\{\mathrm{ord}_M(\zeta) : \zeta < \beta\} \le \eta$ we obtain $\mathrm{ord}_M(\beta) \le \eta$. The assumption $\mathrm{ord}_M(\beta) < \eta$ then leads to the same contradiction as above.

6.22. **Theorem**

If $M \subset \mathrm{On}$ *is a segment of* On *and* $f : M \rightarrow \mathrm{On}$ *is an order preserving function, then* $\xi \le f(\xi)$ *holds for all* $\xi \in M$.

Proof

Assume that $S := \{\xi \in M : f(\xi) < \xi\}$ is not empty and define $\xi_0 := \min S$. Then $f(\xi_0) < \xi_0$ and since M is a segment also $f(\xi_0) \in M$. Hence $f(f(\xi_0)) < f(\xi_0)$ in contradiction to the minimality of ξ_0.

6.23. **Theorem**

Suppose that $\varkappa$ *is a regular ordinal.* $M \subset \varkappa$ *is bounded in* $\varkappa$ *if and only if* $\mathrm{Otyp}(M) < \varkappa$. $M \subset \mathrm{On}$ *is bounded if and only if* $\mathrm{Otyp}(M) \in \mathrm{On}$.

Proof

Since $\mathrm{ord}_M : \mathrm{Otyp}(M) \rightarrow M$ is an 1-1 mapping onto M $\mathrm{Otyp}(M) \in \varkappa$ or $\mathrm{Otyp}(M) \in \mathrm{On}$ respectively imply that M is bounded in $\varkappa$ or On respectively. If on the other hand M is bounded or even bounded in $\varkappa$, then we obtain $\sup M \in On$ or $\sup M < \varkappa$. But 6.22. then implies $Otyp(M) \leq \sup M$.

6.24. Definition

Let $M \subset On$ and $\varkappa$ be a regular ordinal $> \omega$.

(i) M is *closed ($\varkappa$-closed)* if $\sup U \in M$ holds for every non empty set $U \subset M$ which is bounded (in $\varkappa$).

(ii) An order preserving mapping $f : M \to On$ is *continuous* ($\varkappa$-*continuous*) if M is ($\varkappa$-)closed and $f(\sup U) = \sup\{f(\xi) : \xi \in U\}$ holds for every non empty set $U \subset M$ which is bounded (in $\varkappa$).

The notions "closed" and "continuous" stem from the fact that the ordinals together with the order topology form a topological space.

6.25. Lemma

The enumerating function of a set $M \subset On$ *is ($\varkappa$-)continuous if and only if* M *is ($\varkappa$-)closed.*

Proof

$\Rightarrow$: Suppose that ord_M is ($\varkappa$-)continuous. Then Otyp(M) is ($\varkappa$-)closed. Assume that $U \neq \emptyset$, $U \subset M$ and U is bounded (in $\varkappa$). Let $B := ord_M^{-1}(U)$. For $\xi \in B$ we have $\xi \leq ord_M(\xi) \in U$. Hence B is also bounded (in $\varkappa$). So $\sup B \in Otyp(M)$ exists and we have $ord_M(\sup B) \in M$. By the continuity of ord_M we obtain $ord_M(\sup B) = \sup\{ord_M(\xi) : \xi \in B\} = \sup U$.

$\Leftarrow$: Let M be ($\varkappa$-)closed. If $M \subset \varkappa$, then also $Otyp(M) \subset \varkappa$. Assume that $U \subset Otyp(M)$, $U \neq \emptyset$ and U is bounded (in $\varkappa$). But then $ord_M(U)$ is bounded (in $\varkappa$) too. Hence $\sup ord_M(U) \in M$ and there is an $\alpha \in Otyp(M)$ such that $ord_M(\alpha) = \sup ord_M(U)$. For $\xi \in U$ we have $ord_M(\xi) \leq ord_M(\alpha)$. Hence $\sup U \leq \alpha$. If we assume $\sup U < \alpha$ we obtain $ord_M(\sup U) < ord_M(\alpha) = \sup(ord_M(U))$. Therefore there exists a $\xi \in U$ such that $ord_M(\sup U) < ord_M(\xi)$, i.e. $\sup U < \xi$ which contradicts the definition of sup U.

6.26. Definition

(i) A continuous order preserving function $f : On \to On$ is called a *normal-function.*

(ii) We call a mapping $f : \varkappa \to \varkappa$ where $\varkappa$ is a regular ordinal $> \omega$ a $\varkappa$-*normal-function* if f is order preserving and continuous.

6.27. Theorem

(i) ord_M *is a normal function if and only if* M *is closed unbounded.*

(ii) ord_M *is a* $\varkappa$*-normal function if and only if* M *is closed unbounded in* $\varkappa$.

Proof

(i) $\Leftarrow$: If M is unbounded we obtain $\mathrm{dom}(\mathrm{ord}_M) = \mathrm{On}$ by 6.23. If M is also closed we obtain the continuity of ord_M by 6.25. Since ord_M is order preserving by 6.21. it is a normal-function.

$\Rightarrow$: Suppose that ord_M is a normal function. Then M is closed by 6.25. Since $\mathrm{dom}(\mathrm{ord}_M) = \mathrm{On}$ it follows by 6.23. that M is unbounded..

(ii) $\Leftarrow$: Otyp(M) cannot be bounded in $\varkappa$ because ord_M maps Otyp(M) 1-1 onto M. Hence $\varkappa \subset \mathrm{Otyp}(M) \le \sup M \le \varkappa$, i.e. $\mathrm{Otyp}(M) = \varkappa$. Since M is $\varkappa$-closed we also obtain the $\varkappa$-continuity of f. So ord_M is a $\varkappa$-normal-function.

$\Rightarrow$: If ord_M is a $\varkappa$-normal-function we obtain by 6.25. that M is $\varkappa$-closed. Since $\mathrm{Otyp}(M) = \varkappa$ M has to be unbounded in $\varkappa$.

6.28. Exercise

1. The open intervals $(\alpha,\beta) := \{\gamma \in \mathrm{On} : \alpha < \gamma < \beta\}$ form a basis of a topology on On. This topology is called the *order topology* on On. This topology also induces a topology on every regular ordinal $\varkappa$.

Prove the following claims:

(i) A set M is closed (in $\varkappa$) in the sense of definition 6.24. if and only M is closed in the order topology on On (on $\varkappa$).

(ii) Let M be closed (in $\varkappa$). An order preserving mapping $f\colon M \to \mathrm{On}$ ($f\colon M \to \varkappa$) is continuous in the sense of definition 6.24. if and only if it is continuous in the order topology on On (on $\varkappa$).

(iii) Let M be closed. Characterize those functions $f\colon M \to \mathrm{On}$ which satisfy $f(\sup U) = \sup\{f(\xi) : \xi \in U\}$ for all nonvoid sets $U \subset M$ which are bounded in M.

2. An ordinal $\varkappa$ is a *cardinal* if $\varkappa$ cannot be mapped by an 1-1 mapping onto an ordinal $\alpha < \varkappa$. Prove:

(i) Every regular ordinal is a cardinal.

(ii) There are cardinals which are not regular.

(iii) The class of cardinals is closed unbounded.

(iv) If $\varkappa, \lambda$ are cardinals such that $\varkappa < \lambda$ and there is no cardinal $\mu \in (\varkappa,\lambda)$, then λ is regular.

§ 7. Ordinal arithmetic

7.1. **Definition** of the *ordinal sum*

$\alpha + 0 := \alpha$

$\alpha + \beta' := (\alpha + \beta)'$

$\alpha + \lambda := \sup\{\alpha + \xi : \xi < \lambda\}$ for limit ordinals λ.

The function $\lambda\xi.\alpha+\xi$ is defined by transfinite recursion on ξ (cf. 6.19.).

7.2 **Lemma**

$\lambda\xi.\alpha+\xi$ is the enumerating function of the class $\{\eta : \alpha \le \eta\}$. Since this class is closed unbounded (in every regular ordinal $\varkappa$) we have that $\lambda\xi.\alpha+\xi$ is a ($\varkappa$-) normal function (for all regular ordinals $\varkappa$).

Proof

Define $M := \{\eta : \alpha \le \eta\}$. We show $\mathrm{ord}_M(\xi) = \alpha + \xi$ by induction on ξ. It is $\mathrm{ord}_M(0) = \min M = \alpha = \alpha + 0$ and $\mathrm{ord}_M(\beta') = \min\{\xi \in M : \mathrm{ord}_M(\beta) < \xi\} \overset{\text{i.h.}}{=} \min\{\xi \in M : \alpha+\beta < \xi\} = (\alpha+\beta)' = \alpha+\beta'$. If $\lambda \in \mathrm{Lim}$ it is $\mathrm{ord}_M(\lambda) = \min\{\xi \in M : \sup\{\mathrm{ord}_M(\eta) : \eta < \lambda\} \le \xi\}$ $\overset{\text{i.h.}}{=}$ $\min\{\xi \in M : \sup\{\alpha+\eta : \eta < \lambda\} \le \xi\} = \min\{\xi \in M : \alpha + \lambda \le \xi\} = \alpha + \lambda$.

By (O2) the class M is unbounded and trivially it is closed.

7.3. **Lemma** (Elementary properties of the ordinal sum)

(i) $0 + \beta = \beta$

(ii) $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$ *(associativity)*

(iii) $\beta < \gamma \Rightarrow \alpha+\beta < \alpha+\gamma$ *(strong monotonicity in the right argument)*

(iv) $\alpha \le \beta \Rightarrow \alpha+\gamma \le \beta+\gamma$ *(weak monotonicity in the left argument)*

Proof

(i) follows from the fact that $\lambda\xi.0+\xi$ is the enumerating function of the class On which obviously is the identity on On.

(ii) is easily proved by induction on γ.

(iii) $\lambda\xi.\alpha+\xi$ is order preserving since it is an enumerating function.

(iv) If $\alpha \le \beta$, then there is an η such that $\alpha+\eta = \beta$. Hence $\alpha+\gamma \le \alpha+(\eta+\gamma) = (\alpha+\eta)+\gamma = \beta+\gamma$.

7.4. **Definition**

An ordinal α is an *additive principal ordinal* or briefly *principal ordinal* if $\alpha \neq 0$

and $\xi,\eta<\alpha$ also imply $\xi+\eta<\alpha$. By $\mathbb{H}$ (for German Hauptzahl) we denote the class of principal ordinals.

7.5. **Lemma**

If $\alpha\notin\mathbb{H}$, *then there are* $\xi,\eta<\alpha$ *such that* $\alpha=\xi+\eta$.

Proof

$\alpha\notin\mathbb{H}$ implies the existence of ordinals $\xi,\eta_0<\alpha$ such that $\alpha\le\xi+\eta_0$. By 7.2. there is an η such that $\xi+\eta=\alpha\le\xi+\eta_0$. But then we have $\eta\le\eta_0<\alpha$.

7.6. **Lemma**

$\mathbb{H}\subset\mathrm{Lim}\cup\{0'\}$

Proof

If $\alpha\notin\mathrm{Lim}\cup\{0'\}$, then either $\alpha=0$ or there is an ordinal α_0 such that $\alpha=\alpha'_0=(\alpha_0+0)'=\alpha_0+0'$. Since $\alpha\neq0'$ we obtain $0'<\alpha$ and $\alpha_0<\alpha$ in the second case. So in both cases α is not principal.

As usual we will denote 0' by the symbol 1, 1' by the symbol 2 etc.

7.7. **Lemma**

$\{1,\omega\}\subset\mathbb{H}$. *There are no further principal ordinals between* 1 *and* ω.

Proof

If $\xi<1=0'$, then $\xi=0$ and $0+0=0<1$. Hence $1\in\mathbb{H}$. By 6.13. there are no limit ordinals below ω. So by 7.6. the only principal ordinal below ω is 1.
If $\xi,\eta<\omega$, then $\xi,\eta\in\mathbb{N}$. We show by induction on η that this implies $\xi+\eta\in\mathbb{N}$, i.e. $\xi+\eta<\omega$. We have $\xi+0=\xi\in\mathbb{N}$. For $\eta=\eta'_0$ it is $\xi+\eta=\xi+\eta'_0=(\xi+\eta_0)'$. By the induction hypothesis we have $\xi+\eta_0\in\mathbb{N}$ and by the definition of $\mathbb{N}$ this implies $(\xi+\eta_0)'\in\mathbb{N}$, i.e. $\xi+\eta\in\mathbb{N}$.

7.8. **Theorem**

The class $\mathbb{H}$ *is closed unbounded in every regular ordinal* $\varkappa>\omega$.

Proof

Let $\alpha<\varkappa$. Define $\alpha_0:=\alpha'$, $\alpha_{n+1}:=\alpha_n+\alpha_n$ and $M:=\{\alpha_n:n<\omega\}$. Then $M\subset\varkappa$ by 7.2. and we have an 1–1 mapping from M onto $\omega<\varkappa$. Hence M is bounded in $\varkappa$ and we obtain $\alpha < \alpha_0 \le \sup M =: \beta < \varkappa$. For $\xi, \eta < \beta$ there is an $n < \omega$ such that $\xi, \eta < \alpha_n$ and we obtain $\xi + \eta \le \alpha_n + \eta < \alpha_n + \alpha_n = \alpha_{n+1} \le \beta$. Hence $\beta \in \mathbb{H} \cap \varkappa$ and $\mathbb{H}$ is unbounded in $\varkappa$.

Now let $U \subset \mathbb{H}$ be bounded in $\varkappa$. Then $\sup U < \varkappa$. For $\xi, \eta < \sup U$ there is a $\rho \in U$ such that $\xi, \eta < \rho$. Hence $\xi + \eta < \rho \le \sup U$ which entails $\sup U \in \mathbb{H}$. This shows that $\mathbb{H}$ is closed in $\varkappa$.

By 7.8. it follows that $\mathbb{H}$ is closed unbounded in On. Hence the enumerating function of $\mathbb{H}$ is a normal function. Its restriction to any regular $\varkappa > \omega$ always is a $\varkappa$-normal function. We define

$$\omega^\xi := \mathrm{ord}_{\mathbb{H}}(\xi).$$

In the exercises we will show that ω^ξ really has the properties of an exponential function.

7.9. **Lemma** (Elementary properties of ω^ξ)

(i) $\lambda\xi.\omega^\xi$ *is a normal function and* $\lambda\xi < \varkappa.\omega^\xi$ *is a* $\varkappa$*-normal function for any regular* $\varkappa > \omega$.

(ii) $0 < \omega^\alpha$

(iii) $\omega^0 = 1$, $\omega^1 = \omega$

(iv) $\alpha < \beta \Rightarrow \omega^\alpha < \omega^\beta$

(v) *If* $\xi < \omega^\alpha$, *then* $\xi + \omega^\alpha = \omega^\alpha$, *i.e.* $\beta \in \mathbb{H}$ *and* $\xi < \beta$ *imply* $\xi + \beta = \beta$.

Proof

(i)–(iv) are obvious.

(v) For $\xi < \omega^\alpha$ and $\alpha = 0$ we have $\xi = 0$. Hence $\xi + \omega^\alpha = \omega^\alpha$. If $\alpha \neq 0$ we have $\omega^\alpha \in \mathrm{Lim}$ and obtain $\xi + \omega^\alpha = \sup_{\eta < \omega^\alpha} (\xi + \eta) \le \omega^\alpha$ since $\omega^\alpha \in \mathbb{H}$. Hence $\omega^\alpha \le \xi + \omega^\alpha \le \omega^\alpha$, i.e. $\xi + \omega^\alpha = \omega^\alpha$.

7.10. **Theorem** (*Additive normal form* for ordinals)

For every $\alpha \in \mathrm{On}$ *which is different from* 0 *there are uniquely determined ordinals* $\alpha_1, \dots, \alpha_n \in \mathbb{H}$ *such that* $\alpha = \alpha_1 + \dots + \alpha_n$ *and* $\alpha_1 \ge \dots \ge \alpha_n$. *This is denoted by* $\alpha =_{NF} \alpha_1 + \dots + \alpha_n$ *and we define* $\mathbb{H}(\alpha) := \{\alpha_1, \dots, \alpha_n\}$.

Proof

a) We prove the existency of $\alpha_1, \dots, \alpha_n$ by induction on α.

We are done if $\alpha \in \mathbb{H}$. Otherwise there are $\beta_1, \beta_2 < \alpha$ such that $\alpha = \beta_1 + \beta_2$. By the induction hypothesis we have $\beta_1 = \alpha_{11} + \dots + \alpha_{1n}$ and $\beta_2 = \alpha_{21} + \dots + \alpha_{2m}$ such that

$\alpha_{11} \geq .. \geq \alpha_{1n}$ and $\alpha_{21} \geq ... \geq \alpha_{2m}$. But then $\alpha = \beta_1 + \beta_2 = (\alpha_{11} + ... + \alpha_{1n}) + (\alpha_{21} + ... + \alpha_{2m}) = \alpha_{11} + ... + \alpha_{1k} + \alpha_{21} + ... + \alpha_{2m}$ where α_{1k} is the last element in the list $\alpha_{11}, ..., \alpha_{1n}$ which is larger or equal to α_{21}. The latter equation holds since by 7.9. $\alpha_{1j} < \alpha_{21}$ always implies $\alpha_{1j} + \alpha_{21} = \alpha_{21}$ such that all ordinals less than α_{21} are swallowed by α_{21}.

b) In a second step we prove the uniqueness.

Assume that $\alpha =_{NF} \alpha_1 + ... + \alpha_n$ and $\alpha =_{NF} \beta_1 + ... + \beta_m$. We prove $n = m$ and $\alpha_i = \beta_i$ by induction on n. Since $\alpha_1, \beta_1 \in \mathbb{H}$ there are ordinals ξ_1 and ξ_2 such that $\alpha_1 = \omega^{\xi_1}$ and $\beta_1 = \omega^{\xi_2}$. Hence $\omega^{\xi_1} \leq \alpha < \omega^{\xi_2'}$ and $\omega^{\xi_2} \leq \alpha < \omega^{\xi_1'}$. This entails $\xi_1 \leq \xi_2$ and $\xi_2 \leq \xi_1$ and we have $\xi_1 = \xi_2$ and therefore also $\alpha_1 = \beta_1$. If $n = 1$ we are done. Otherwise it follows $\alpha_2 + ... + \alpha_n = \beta_2 + ... + \beta_m$ which by the induction hypothesis implies $n = m$ and $\alpha_i = \beta_i$ for $i = 2, ..., n$.

7.11. **Corollary** (*Cantor normal form* to basis ω)

For every ordinal $\alpha \neq 0$ *there are uniquely determined ordinals* $\alpha_1 \geq .. \geq \alpha_n$ *such that* $\alpha = \omega^{\alpha_1} + ... + \omega^{\alpha_n}$.

7.12. **Definition** of the *natural sum* of ordinals

If $\alpha =_{NF} \alpha_1 + ... + \alpha_n$ and $\beta =_{NF} \alpha_{n+1} + ... + \alpha_{n+m}$, we define $\alpha \# \beta := \alpha_{\pi(1)} + ... + \alpha_{\pi(n+m)}$ where $\pi \in S_{n+m}$ is a permutation of the integers $1, ..., n+m$ such that $i < j$ always implies $\alpha_{\pi(i)} \geq \alpha_{\pi(j)}$.

7.13. **Lemma**

(i) $\alpha \# \beta = \beta \# \alpha$.

(ii) $\alpha < \beta$ *implies* $\alpha \# \gamma < \beta \# \gamma$ *and* $\gamma \# \alpha < \gamma \# \beta$.

(iii) If $\gamma \in \mathbb{H}$, $\alpha < \gamma$ *and* $\beta < \gamma$, *then* $\alpha \# \beta < \gamma$.

(iv) $\alpha \# (\beta \# \gamma) = (\alpha \# \beta) \# \gamma$.

7.14. **Definition**

We recursively define the exponentiation to the basis 2 by

(i) $2^0 = 0'$

(ii) $2^{\alpha'} = 2^\alpha \# 2^\alpha$

(iii) $2^\lambda = \sup\{2^\xi : \xi < \lambda\}$ if $\lambda \in Lim$.

This exponentiation has the following properties

7.15. **Lemma**

(i) $\alpha \le 2^\alpha$

(ii) $\alpha < \beta$ *implies* $2^\alpha < 2^\beta$ *and* $2^\alpha + 2^\alpha \le 2^\beta$.

(iii) For all $\alpha \in \mathrm{On}$ *we have* $2^\alpha \le \omega^\alpha$.

All proofs are easy exercises.

7.16. **Exercises**

1. Prove or disprove the following claims

   (i) $(\alpha+\beta)+\gamma = \alpha+(\beta+\gamma)$

   (ii) $\alpha+\beta = \beta+\alpha$

   (iii) If $\alpha \neq 0$ and $\beta+\alpha = \alpha$ for all $\beta < \alpha$, then $\alpha \in \mathbb{H}$.

2. Prove 7.13.

We define the multiplication of ordinals by the following recursion

$\alpha \cdot 0 = 0$

$\alpha \cdot \beta' = \alpha \cdot \beta + \alpha$

$\alpha \cdot \lambda = \sup\{\alpha \cdot \xi : \xi < \lambda\}$ for $\lambda \in \mathrm{Lim}$.

3. Prove or disprove the following claims

   (i) $\alpha < \beta \wedge \gamma > 0 \Leftrightarrow \gamma \cdot \alpha < \gamma \cdot \beta$

   (ii) $\alpha \le \beta \Rightarrow \alpha \cdot \gamma \le \beta \cdot \gamma$

   (iii) $\alpha \cdot (\beta+\gamma) = \alpha \cdot \beta + \alpha \cdot \gamma$

   (iv) $\alpha \cdot (\beta \cdot \gamma) = (\alpha \cdot \beta) \cdot \gamma$

   (v) $(\alpha+\beta) \cdot \gamma = \alpha \cdot \gamma + \beta \cdot \gamma$

4. Define the enumerating function of those ordinals which are not successor ordinals. Prove your claim.

The general exponentiation of ordinals is defined by

$\exp(\alpha,0) = 1$

$\exp(\alpha,\beta') = \exp(\alpha,\beta) \cdot \alpha$

$\exp(\alpha,\lambda) = \sup\{\exp(\alpha,\xi) \mid \xi < \lambda\}$ for $\lambda \in \mathrm{Lim}$

5. Prove the following claims

   (i) $\beta < \gamma \wedge \alpha > 1 \Rightarrow \exp(\alpha,\beta) < \exp(\alpha,\gamma)$

   (ii) $\alpha < \beta \Rightarrow \exp(\alpha,\gamma) \le \exp(\beta,\gamma)$

   (iii) $\exp(\alpha,\beta+\gamma) = \exp(\alpha,\beta) \cdot \exp(\alpha,\gamma)$

   (iv) $\exp(\alpha,\beta \cdot \gamma) = \exp(\exp(\alpha,\beta),\gamma)$

(v) $\forall\alpha\ \exp(\omega,\alpha) \in \mathbb{H}$

(vi) $\forall\alpha\ \exp(\omega,\alpha) = \omega^{\alpha}$

(vii) $\forall\lambda\ (\mathrm{Lim}(\lambda) \Rightarrow \exp(2,\lambda) \in \mathbb{H})$

(viii) $\forall\alpha\ (2^{\alpha} = \exp(2,\alpha))$

(ix) $\forall\alpha > 0\ (\omega^{\alpha} = 2^{\omega\cdot\alpha})$

An ordinal $\alpha>1$ is a *multiplicative principal* ordinal if it is closed under ordinal multiplication i.e. if $\xi,\eta<\alpha$ imply $\xi\cdot\eta<\alpha$.

6. Show the following properties.

(i) If $\alpha>2$ is multiplicative principal, then α is a limit ordinal.

(ii) An ordinal $\alpha>1$ is multiplicative principal if and only if it is $\xi\cdot\alpha = \alpha$ for all $1\le\xi<\alpha$.

(iii) An ordinal $\alpha>1$ is multiplicative principal if and only if $\alpha = 2$ or if there is an ordinal β such that $\alpha = \omega^{(\omega^{\beta})}$.

§ 8. A notation system for a segment of the ordinals

8.1. Definition

(i) $\varepsilon_0 := \min\{\xi : \omega^{\xi} = \xi\}$

(ii) $\omega_0(\beta) := \beta$, $\omega_{n+1}(\beta) := \omega^{\omega_n(\beta)}$, $\omega_\lambda(\beta) := \sup\{\omega_\xi(\beta) : \xi<\lambda\}$ for $\lambda \in \mathrm{Lim}$.

8.2. Lemma

$\varepsilon_0 = \omega_\omega(0)$.

Proof

By induction on n we immediately obtain $\omega_n(0) < \omega_{n+1}(0)$. Therefore the set $\{\omega_n(0) : n<\omega\}$ has no maximum. By 6.11. it follows that $\omega_\omega(0) \in \mathrm{Lim}$. Hence $\omega^{\omega_\omega(0)} = \sup\{\omega^{\xi} : \xi<\omega_\omega(0)\} = \sup\{\omega^{\omega_n(0)} : n<\omega\} = \sup\{\omega_{n+1}(0) : n<\omega\} = \omega_\omega(0)$. This shows that $\varepsilon_0 \le \omega_\omega(0)$. $\varepsilon_0 < \omega_\omega(0)$ is excluded because otherwise we had an $n<\omega$ such that $\omega_n(0) \le \varepsilon_0 < \omega_{n+1}(0)$. But from this we would obtain $\omega_{n+1}(0) \le \omega^{\varepsilon_0} = \varepsilon_0 < \omega_{n+1}(0)$ which is impossible.

8.3 Theorem

$\varepsilon_0 < \Omega$.

Proof

We have $\omega_n(0) < \Omega$ for all n because $\lambda\xi.\omega^\xi$ is an Ω-normal function. Since there is a 1-1 mapping from $\{\omega_n(0) : n < \omega\}$ onto $\omega < \Omega$ this set is bounded in Ω. Hence $\varepsilon_0 = \sup\{\omega_n(0) : n < \omega\} < \Omega$.

8.4. **Inductive definition** of the ordinal set E

(i) $0 \in E$

(ii) If $\alpha, \beta \in E$, then $\alpha+\beta \in E$

(iii) If $\alpha \in E$, then $\omega^\alpha \in E$

8.5. **Theorem**

The ordinal set E *exactly is the segment of ordinals below* ε_0, *i.e.* $E = \varepsilon_0$.

Proof

We show

(1) $\alpha \in E \Rightarrow \alpha < \varepsilon_0$

by induction on the definition of E.

$0 < \varepsilon_0$ is obvious. $\alpha, \beta < \varepsilon_0$ imply $\alpha+\beta < \varepsilon_0$ since $\varepsilon_0 \in \mathbb{H}$. If $\alpha < \varepsilon_0$, then $\omega^\alpha < \omega^{\varepsilon_0} = \varepsilon_0$.

For the opposite inclusion we prove

(2) $\alpha < \varepsilon_0 \Rightarrow \alpha \in E$

by induction on α.

For $\alpha = 0$ this again is obvious.

If $\alpha \in \mathbb{H}$, then there is a $\xi \leq \alpha$ such that $\alpha = \omega^\xi$. $\alpha < \varepsilon_0$ implies $\xi < \alpha$. By the induction hypothesis we then have $\xi \in E$ which by 8.4.(iii) entails $\alpha = \omega^\xi \in E$.

If $\alpha \notin \mathbb{H}$, then there are ordinals $\alpha_1, \alpha_2 < \alpha$ such that $\alpha_1+\alpha_2 = \alpha$. By the induction hypothesis we obtain $\alpha_i \in E$ for $i = 1,2$, and by 8.4.(ii) it follows $\alpha = \alpha_1+\alpha_2 \in E$.

8.5. enables us to denote every ordinal less than ε_0 by an element in E. This notation, however, is not unique. To obtain uniqueness we have to refer to normal forms.

8.6. **Theorem**

For every ordinal $\alpha \in E$ *different from* 0 *there are uniquely determined ordinals* $\alpha_1,\ldots,\alpha_n \in E \cap \alpha$ *such that* $\alpha =_{NF} \omega^{\alpha_1}+\ldots+\omega^{\alpha_n}$.

Proof

By the normal form theorem there are uniquely determined ordinals $\alpha_1,\dots,\alpha_n$ such that $\alpha =_{NF} \omega^{\alpha_1}+\dots+\omega^{\alpha_n}$. $\alpha \in E$ implies $\alpha < \varepsilon_0$ and we obtain $\alpha_1,\dots,\alpha_n < \alpha \subset \varepsilon_0 = E$.

8.7. **Definition** (*arithmetization* of E)

We define a mapping $\ulcorner\;\urcorner : E \to \mathbb{N}$ by:

(i) $\ulcorner 0 \urcorner = 0$

(ii) If $\alpha =_{NF} \omega^{\alpha_1}+\dots+\omega^{\alpha_n}$ then $\ulcorner\alpha\urcorner := \langle 1, \ulcorner\alpha_1\urcorner,\dots,\ulcorner\alpha_n\urcorner\rangle$

$\ulcorner E \urcorner := \{\ulcorner\alpha\urcorner : \alpha \in E\}$

$n \prec m :\Leftrightarrow \exists\alpha \in E\, \exists\beta \in E(\ulcorner\alpha\urcorner = n \wedge \ulcorner\beta\urcorner = m \wedge \alpha < \beta)$

$n \equiv m :\Leftrightarrow \exists\alpha \in E\, \exists\beta \in E(\ulcorner\alpha\urcorner = n \wedge \ulcorner\beta\urcorner = m \wedge \alpha = \beta)$

8.8. **Theorem**

The set $\ulcorner E \urcorner$ *and the relations* $\prec$ *and* $\equiv$ *are primitive recursive.*

Proof

We have

$$n \in \ulcorner E\urcorner \Leftrightarrow n = 0 \vee \mathrm{Seq}(n) \wedge (n)_0 = 1 \wedge \forall x < lh(n)\,(0 < x \Rightarrow (n)_x \in \ulcorner E \urcorner)$$
$$\wedge\; \forall x < lh(n)-1\,(0 < x \Rightarrow (n)_{x+1} \preceq (n)_x).$$

and

$$n_1 \prec n_2 \Leftrightarrow n_1 \in \ulcorner E\urcorner \wedge n_2 \in \ulcorner E \urcorner \wedge [(n_1 = 0 \wedge n_2 \neq 0)$$
$$\vee\; (\exists x < \min\{lh(n_1), lh(n_2)\}((n_1)_x \prec (n_2)_x) \wedge \forall y < x\,(0 < y \Rightarrow (n_1)_y = (n_2)_y))$$
$$\vee\; (lh(n_1) < lh(n_2) \wedge \forall x < lh(n_1)((n_1)_x = (n_2)_x))].$$

Thus the set $\ulcorner E\urcorner$ and the relation $\prec$ are definable by simultaneous course of value recursion. Hence both are primitive recursive.

The following corollary then is an immediate consequence of 8.8.

8.9. **Corollary**

$\varepsilon_0 < \omega_1^{CK}$.

§9. A norm function for Π^1_1-sentences

9.1. Inductive definition of $\models^{\alpha}_{0} \Delta$

(Ax) If $\models_0 \Delta$ holds according to (Ax1) or (Ax2), then $\models^{0}_{0} \Delta$.

($\bigwedge$) If $\models^{\alpha_i}_{0} \Delta, A_i$ holds for all $i \in I$ and $\alpha = \sup\{\alpha_i+1 : i \in I\}$, then $\models^{\alpha}_{0} \Delta, \bigwedge_{i\in I} A_i$.

($\bigvee$) If $\models^{\alpha_0}_{0} \Delta, A_i$ holds for some $i \in I$, then $\models^{\alpha_0'}_{0} \Delta, \bigvee_{i\in I} A_i$.

9.2. Lemma

Let Δ,F be a finite set of $\mathscr{L}_\Omega$-formulas.

(i) It holds $\models_0 \Delta$ if and only if there is an $\alpha<\Omega$ such that $\models^{\alpha}_{0}\Delta$.

(ii) We have $\mathbb{N} \models F$ if and only if there is an $\alpha<\Omega$ such that $\models^{\alpha}_{0}$ F.

Proof

(i) We prove $\models^{\alpha}_{0}\Delta \Rightarrow \alpha<\Omega$ by induction on the definition of $\models^{\alpha}_{0} \Delta$. In the case of (Ax) this is obvious. In the case of an inference ($\bigvee$) we have $\alpha_0<\Omega$ by the induction hypothesis which entails $\alpha_0' < \Omega$ since $\Omega \in \mathrm{Lim}$. In the case of an inference ($\bigwedge$) we have by the induction hypothesis $\alpha_i < \Omega$ for all $i \in I$. Since Δ is a set of $\mathscr{L}_\Omega$ -formulas the index set I and therefore also the set $\{\alpha_i : i \in I\}$ has to be countable. Hence $\alpha = \sup\{\alpha_i : i \in I\} < \Omega$.

The opposite direction is trivial since the definition of $\models^{\alpha}_{0} \Delta$ immediately implies $\models_0 \Delta$.

(ii) By the soundness and the completeness theorem we have $\mathbb{N} \models F$ if and only if $\models_0 F$ and the claim follows from (i).

9.3. Definition

For an $\mathscr{L}_\Omega$-formula F we define

$$|F| := \begin{cases} \min\{\alpha : \models^{\alpha}_{0} F\} & \text{if this is defined} \\ \Omega & \text{otherwise} \end{cases}$$

We call $|F|$ the *norm* of the $\mathscr{L}_\Omega$-formula F.

9.4. Definition of the translation * which maps Π^1_1-sentences of $\mathscr{L}_1$ to formulas of $\mathscr{L}_\infty$

(i) If F is an atomic formula, then $F^* :\equiv F$.

(ii) $(\neg A)^* :\equiv \neg A^*$

(iii) $(A \wedge B)^* :\equiv \bigwedge\{A^*, B^*\}$

(iv) $(A \vee B)^* :\equiv \bigvee\{A^*, B^*\}$

(v) $(\forall x A)^* :\equiv \bigwedge\{A_x(\underline{n})^* : n < \omega\}$

(vi) $(\exists x A)^* :\equiv \bigvee\{A_x(\underline{n})^* : n < \omega\}$.

A formula of $\mathscr{L}_\infty$ which is the $*$-translation of a Π^1_1-formula of $\mathscr{L}_1$ often is called a *Π^1_1-sentence* of $\mathscr{L}_\infty$.
For a Π^1_1-sentence F of $\mathscr{L}$ we define $|F| := |F^*|$.

By an easy induction on the definition of the $\mathscr{L}_1$-formula F we obtain

9.5. **Lemma**
For a Π^1_1-sentence F of $\mathscr{L}$ we have $\mathbb{N} \models F$ if and only if $\mathbb{N} \models F^*$.

9.6. **Theorem**
For an $\mathscr{L}_\Omega$-formula F we have $\mathbb{N} \models F$ if and only if $|F| < \Omega$.

Proof
If $\mathbb{N} \models F$, then by 9.2. there is an $\alpha < \Omega$ such that $\models^{\alpha}_{0} F$. Hence $|F| \le \alpha < \Omega$.
$\mathbb{N} \not\models F$ implies $\not\models_0 F$ by the soundness theorem. But then $\{\alpha : \models^{\alpha}_{0} F\} = \emptyset$ by 9.2. Hence $|F| = \Omega$.

The following corollary is an immediate consequence of 9.5. and 9.6.

9.7. **Corollary**
For a Π^1_1-sentence F in $\mathscr{L}$ we have $\mathbb{N} \models F$ if and only if $|F| < \Omega$.

The claim of 9.7. may be sharpened to

$$\mathbb{N} \models F \Leftrightarrow |F| < \omega_1^{CK}.$$

Here it is even possible to show that

$$\sup\{|F| : \mathbb{N} \models F\} = \omega_1^{CK}.$$

Both proofs use methods of recursion theory and are outside the scope of this lecture.

9.8. **Ecercise**
Show that $|F| < \omega$ holds for true $\mathscr{L}_1$-sentences.

§ 10. The infinitary system $\mathbf{Z}_\infty$

After having introduced a norm for the Π^1_1-sentences of the language of pure number theory it is a natural question to ask which norms are accessed by Π^1_1-sentences provable in $\mathbf{Z}_1$. In order to answer this question we will introduce an infinitary system $\mathbf{Z}_\infty$ whose cut free derivations may be interpreted as $\models_0$ derivations and so obtain an upper bound for the norms of the derivable formulas.

In a first step we define the rank of an $\mathscr{L}_\infty$-formula.

10.1. **Definition** of the *rank* rk(F) of an $\mathscr{L}_\infty$-formula F

(i) If F is atomic, then $\mathrm{rk}(F) := 0$.

(ii) If F is a formula $\bigwedge\{F_i : i \in I\}$ or a formula $\bigvee\{F_i : i \in I\}$ we define $\mathrm{rk}(F) := \sup\{\mathrm{rk}(F_i)+1 : i \in I\}$.

As an immediate consequence we obtain

10.2. **Lemma**

If $F \in \mathscr{L}_\Omega$, *then* $\mathrm{rk}(F) = \mathrm{rk}(\neg F) < \Omega$.

For the following definition we presuppose that M is a subclass of the ordinals and that Δ is a finite set of $\mathscr{L}_\infty$-formulas whose ranks all belong to M.

10.3. **Inductive definition** of $\mathbf{Z}_M \vdash^\alpha_\rho \Delta$

(Ax) If $\models_0 \Delta$ holds by (Ax1) or (Ax2), then we have $\mathbf{Z}_M \vdash^\alpha_\rho \Delta$ for all $\alpha, \rho \in M$.

($\bigwedge$) If we have $\mathbf{Z}_M \vdash^{\alpha_i}_\rho \Delta, A_i$ and $\alpha_i \in \alpha \cap M$ for all $i \in I$, then we also have $\mathbf{Z}_M \vdash^\alpha_\rho \Delta, \underline{\bigwedge\{A_i : i \in I\}}$

($\bigvee$) If $\mathbf{Z}_M \vdash^{\alpha_0}_\rho \Delta, A_i$ and $\alpha_0 \in M \cap \alpha$ holds for some $i \in I$ and $\alpha \in M$, then we also have $\mathbf{Z}_M \vdash^\alpha_\rho \Delta, \underline{\bigvee\{A_i : i \in I\}}$.

(cut) If $\mathbf{Z}_M \vdash^{\alpha_1}_\rho \Delta, A$ and $\mathbf{Z}_M \vdash^{\alpha_2}_\rho \Delta, \neg A$ and $\mathrm{rk}(A) \in M \cap \rho$, then we have $\mathbf{Z}_M \vdash^\alpha_\rho \Delta$ for all $\alpha \in M$ such that $\alpha_1, \alpha_2 < \alpha$. We call rk(A) the *rank of the cut.*

The underlined formulas in the conclusion of the inferences ($\bigwedge$) and ($\bigvee$) are characteristic for the inference. We call it the *mainformula* of the inference. We often interpret an axiom as an inference without premises. The mainformula of an axiom according to (Ax1) (cf.5.2) is $\underline{P}t_1 \dots t_n$. An axiom according to (Ax2) has two mainformulas, $t \in X$ and $s \notin X$.

If M is a recursive set of ordinals, then $\mathbf{Z}_M$ is called a *semiformal system.*

10.4. Lemma

If M is a segment of the ordinals and $\alpha \in M$, *a comparison of the definitions of* $\mid\frac{\alpha}{o}$ *and* $\mathbf{Z}_M \mid\frac{\alpha}{o}$ *shows that there are the following connections.*

$$\mathbf{Z}_M \mid\frac{\alpha}{o} \Delta \Leftrightarrow \alpha \in M \wedge \exists \xi \leq \alpha (\mid\frac{\xi}{o} \Delta).$$

This immediately entails $\mathbf{Z}_M \mid\frac{\alpha}{o} F \Rightarrow |F| \leq \alpha$.

Instead of $\mathbf{Z}_{On}$ we usually write $\mathbf{Z}_\infty$. First of all we will only regard such $M \subset On$ which are segments of Ω. For a segment $M = \beta < \Omega$ we just write $\mathbf{Z}_\beta \mid\frac{\alpha}{\rho} \Delta$ or shortly $\mid\frac{\alpha}{\rho} \Delta$ instead of $\mathbf{Z}_M \mid\frac{\alpha}{\rho} \Delta$ whenever it is clear from the context which set M we are talking about.

10.5 Lemma

If $\mid\frac{\alpha}{\rho} \Delta$, $\alpha \leq \beta$ *and* $\rho \leq \sigma$, *then* $\mid\frac{\beta}{\sigma} \Delta$.

The proof is an easy induction on α.

10.6. Theorem (structural rule)

If $\mid\frac{\alpha}{\rho} \Delta$ *and* $\Delta \subset \Gamma$, *then* $\mid\frac{\alpha}{\rho} \Gamma$.

Proof by induction on α

(Ax) If Δ is an axiom $\mid\frac{\alpha}{\rho} \Delta_0, \underline{P}\, t_1,...,t_n$ or $\mid\frac{\alpha}{\rho} \Delta_0, t \in X, s \notin X$, then Γ is an axiom of the same kind. Hence also $\mid\frac{\alpha}{\rho} \Gamma$.

($\bigwedge$) From the premises $\mid\frac{\alpha_i}{\rho} \Delta_0, A_i$ for all $i \in I$ we have $\mid\frac{\alpha_i}{\rho} \Gamma, A_i$ for all $i \in I$ by the induction hypothesis. Because of $(\bigwedge_{i \in I} A_i) \in \Delta \subset \Gamma$ we obtain $\mid\frac{\alpha}{\rho} \Gamma$ by an $\bigwedge$-inference.

($\bigvee$) From the premise $\mid\frac{\alpha_0}{\rho} \Delta, A_i$ for some $i \in I$ we first obtain $\mid\frac{\alpha_0}{\rho} \Gamma, A_i$ by the induction hypothesis. By an $\bigvee$-inference it then follows $\mid\frac{\alpha}{\rho} \Gamma$.

(cut) From $\mid\frac{\alpha_1}{\rho} \Delta, A$ and $\mid\frac{\alpha_2}{\rho} \Delta, \neg A$ it follows $\mid\frac{\alpha_1}{\rho} \Gamma, A$ and $\mid\frac{\alpha_2}{\rho} \Gamma, \neg A$ by the induction hypothesis. Using a cut we then obtain $\mid\frac{\alpha}{\rho} \Gamma$.

10.7. Theorem ($\bigwedge$-inversion rule)

$\mid\frac{\alpha}{\rho} \Delta, \bigwedge\{A_i : i \in I\}$ *entails* $\mid\frac{\alpha}{\rho} \Delta, A_i$ *for all* $i \in I$.

Proof

1. If $\vdash^{\alpha}_{\rho} \Delta, \bigwedge_{i\in I} A_i$ is an axiom, then $\bigwedge_{i\in I} A_i$ cannot be its main formula. But then $\vdash^{\alpha}_{\rho} \Delta, A_i$ for all $i \in I$ is an axiom of the same kind.

2. If $\bigwedge_{i\in I} A_i$ is not the main formula of the last inference

$$\text{(S)} \vdash^{\alpha_j}_{\rho} \Delta_j, \bigwedge_{i\in I} A_i \Rightarrow \vdash^{\alpha}_{\rho} \Delta, \bigwedge_{i\in I} A_i .$$

then we have $\vdash^{\alpha_j}_{\rho} \Delta_j, A_i$ for all $i \in I$ by the induction hypothesis. By the same inference (S) we thus obtain $\vdash^{\alpha}_{\rho} \Delta, A_i$ for all $i \in I$.

3. If $\bigwedge_{i\in I} A_i$ is the main formula of the last inference, then this inference is a $\bigwedge$-inference whose premises are $\vdash^{\alpha_i}_{\rho} \Delta_0, A_i$ for all $i \in I$. By the structural rule we obtain $\vdash^{\alpha_i}_{\rho} \Delta, A_i, \bigwedge_{i\in I} A_i$ and by the induction hypothesis $\vdash^{\alpha_i}_{\rho} \Delta, A_i$ for all $i \in I$. This implies $\vdash^{\alpha}_{\rho} \Delta, A_i$ for all $i \in I$ by 10.5.

10.8. $\bigvee$**-importation and** $\bigvee$**-exportation**

(i) $\vdash^{\alpha}_{\rho} \Delta, A_1,\dots,A_n$ *implies* $\vdash^{\alpha+n}_{\rho} \Delta, A_1 \vee \dots \vee A_n$.

(ii) $\vdash^{\alpha}_{\rho} \Delta, A_1 \vee \dots \vee A_n$ *implies* $\vdash^{\alpha}_{\rho} \Delta, A_1,\dots,A_n$.

Proof

(i) By iterated application of $\bigvee$-inferences we obtain $\vdash^{\alpha}_{\rho} \Delta, A_1,\dots,A_n \Rightarrow \vdash^{\alpha+1}_{\rho} \Delta, A_2,\dots,A_n, A_1 \vee \dots \vee A_n \Rightarrow \vdash^{\alpha+2}_{\rho} \Delta, A_3,\dots,A_n, A_1 \vee \dots \vee A_n \Rightarrow \dots \Rightarrow \vdash^{\alpha+n}_{\rho} \Delta, A_1 \vee \dots \vee A_n$

(ii) The proof is by induction on α.

1. If $A_1 \vee \dots \vee A_n$ is not the main formula of the last inference, then either Δ is an axiom and so is $\Delta, A_1,\dots,A_n$ or we have the premises $\vdash^{\alpha_j}_{\rho} \Delta_j, A_1 \vee \dots \vee A_n$. But then we have $\vdash^{\alpha_j}_{\rho} \Delta_j, A_1,\dots,A_n$ by the induction hypothesis and obtain $\vdash^{\alpha}_{\rho} \Delta, A_1,\dots,A_n$ by the same inference.

2. If $A_1 \vee \dots \vee A_n$ is the main formula of the last inference, then it is an $\bigvee$-inference whose premise is $\vdash^{\alpha_0}_{\rho} \Delta, A_i, A_1 \vee \dots \vee A_n$. By the induction hypothesis it follows $\vdash^{\alpha_0}_{\rho} \Delta, A_1, \dots A_n$ and by 10.5. $\vdash^{\alpha}_{\rho} \Delta, A_1, \dots A_n$.

10.9. **Tautology lemma**

Suppose that F is an $\mathcal{L}_\infty(x_1,\dots,x_n)$ *formula.* $\mathbf{t}=(t_1,\dots,t_n)$ *and* $\mathbf{s}=(s_1,\dots,s_n)$ *are n-tuples of* $\mathcal{L}_\infty$*-terms such that* s_i *and* t_i *are equivalent for* $i = 1,\dots,n$. *Now if* $F_1 \equiv F_{\mathbf{x}}(\mathbf{t})$, $F_2 \equiv F_{\mathbf{x}}(\mathbf{s})$ *and* $\alpha = rk(F)$, *then we have* $\vdash^{2\alpha}_{0} \Delta, F_1, \neg F_2$ *for all finite formula sets* Δ.

Proof by induction on $rk(F_1)$

1. If F_1 is an atomic formula $\underline{R}t_1\dots t_n$, then we have $F_2 \equiv \underline{R}s_1\dots s_n$. If $\underline{R}t_1\dots t_n$ is valid , we obtain by (Ax1) $\vdash^{0}_{0} \Delta, F_1, \neg F_2$. Otherwise $\neg\underline{R}s_1\dots s_n$ is valid and we

again by (Ax1) obtain $\vdash^0_0 \Delta, F_1, \neg F_2$.

2. If $F_1 \equiv t_i \in X$, then $\neg F_2 \equiv s_i \notin X$. But then $\vdash^0_0 \Delta, F_1, \neg F_2$ holds by (Ax2).

3. If $F_1 \equiv \bigwedge_{i\in I} A^1_i$, then $\neg F_2 \equiv \bigvee_{i\in I} \neg A^2_i$ and by the induction hypothesis we have $\vdash^{2\alpha_i}_0 \Delta, A^1_i, \neg A^2_i$ for all $i \in I$. Using an $\bigvee$-inference we obtain $\vdash^{2\alpha_i+1}_0 \Delta, A^1_i, \neg F_2$ for all $i \in I$. Since $\alpha = \sup\{\alpha_i+1 : i \in I\}$ we have $2\alpha_i+1 < 2(\alpha_i+1) \le 2\alpha$ and it follows $\vdash^{2\alpha}_0 \Delta, F_1, \neg F_2$. by an $\bigwedge$-inference.

10.10. **Inductive definition** of the set AT(F) of *sentential subformulas* of a $\mathcal{L}_\infty$-formula F.

(i) If F is a formula $Pt_1,\ldots,t_n$, $t \in X$, $t \notin X$, $\bigwedge\{A_i : i \in I\}$ or $\bigvee\{A_i : i \in I\}$ with infinite index set I, then $AT(F) = \{F\}$.

(ii) If F is a formula $\bigwedge\{A_i : i<n<\omega\}$ or a formula $\bigvee\{A_i : i<n<\omega\}$, then we define $AT(F) = \{F\} \cup \bigcup\{AT(A_i) : i<n\}$.

As in §3 we call formulas F with $AT(F) = \{F\}$ *sentential atoms* of $\mathcal{L}_\infty$. By AE we denote the set of all sentential atoms. We define $AE(F) := AT(F) \cap AE$.

10.11. **Definition**

(i) Two sentential atoms $A[t_1,\ldots,t_n]$ and $A[s_1,\ldots,s_n]$ are *equivalent*, if we have $t_i^{\mathbb{N}} = s_i^{\mathbb{N}}$ for $i = 1,\ldots,n$.

It is completely obvious that this relation in fact is an equivalence relation.

(ii) Two sentential atoms A_1 and A_2 are *dual*, if there is a sentential atom F such that A_1 is equivalent to F and A_2 equivalent to $\neg F$.

10.12. **Definition**

(i) A *sentential assignment* is a mapping $\mathbb{B} : AE \longrightarrow \{t,f\}$, which assigns different truth values to dual sentential atomic formulas and is compatible with the equivalence of sentential atomic formulas.

(ii) Inductive definition of $A^{\mathbb{B}}$ for $A \in AT(F)$.

1. $A \in AT(F) \cap AE$. Then $A^{\mathbb{B}} = \mathbb{B}(A)$.
2. $(\bigvee\{A_i : i<n<\omega\})^{\mathbb{B}} = t \Leftrightarrow A_i^{\mathbb{B}} = t$ for some $i<n<\omega$.
3. $(\bigwedge\{A_i : i<n<\omega\})^{\mathbb{B}} = t \Leftrightarrow A_i^{\mathbb{B}} = t$ for all $i<n<\omega$.

(iii) A finite formula set $\{F_1,\ldots,F_n\}$ is *sententially valid*, if for every sentential assignment $\mathbb{B}$ there is an $i \in \{1,\ldots,n\}$ such that $F_i^{\mathbb{B}} = t$.

(iv) If Δ is a finite formula set we denote by Δ^a the set of formulas which comes up from Δ if we replace each occurence of a formula of the shape

$\bigvee\{F_i : i<n<\omega\}$ by $\{F_0,\ldots,F_{n-1}\}$ and iterate this process until all finite disjuctions have disappeared. According to 10.12.(ii)2. Δ is sententially valid if and only if Δ^a is.

(v) A finite formula set Δ is *sententially reducible* if Δ^a contains a formula of the form $\bigwedge\{A_i : i<n<\omega\}$. Otherwise Δ is *sententially irreducible.*

10.13. Lemma

A sententially irreducible formula set Δ is sententially valid if and only if Δ^a has the shape Δ_0, F_1, F_2 where F_1, F_2 are dual sentential atoms.

Proof

$\Leftarrow$: If Δ^a has the shape Δ_0, F_1, F_2, then we have $F_1^{\mathbb{B}} = t$ or $F_2^{\mathbb{B}} = t$ for any sentential assignment $\mathbb{B}$. Hence Δ^a and consequently also Δ are sententially valid.

$\Rightarrow$: If Δ^a does not have the shape Δ_0, F_1, F_2, then we assign the truth value f to all sentential atoms in Δ^a. This defines a correct sentential assignment $\mathbb{B}$ because by hypothesis Δ^a does not contain dual sentential atoms. But Δ^a only contains sentential atoms since Δ is irreducible. So we have $F^{\mathbb{B}} = f$ for all $F \in \Delta^a$.

10.14. Lemma

A finite formula set $\Delta, \bigwedge\{F_i : i<n<\omega\}$ is sententially valid if and only if Δ, F_i is sententially valid for all $i<n$.

Proof

$\Rightarrow$: Let $\mathbb{B}$ a sentential assignment such that $F^{\mathbb{B}} = f$ for all $F \in \Delta$ and $F_i^{\mathbb{B}} = f$ for some $i<n$. Then we have $\bigwedge\{F_i : i<n\}^{\mathbb{B}} = f$ and therefore $F^{\mathbb{B}} = f$ for all $F \in \Delta, \bigwedge\{F_i : i<n\}$.

$\Leftarrow$: For a sentential assignment $\mathbb{B}$ we always have $\bigwedge\{F_i : i<n\}^{\mathbb{B}} = F_{i_0}^{\mathbb{B}}$ for some $i_0<n$. Since Δ, F_i is sententially valid for all $i<n$ we obtain $F^{\mathbb{B}} = t$ for all $F \in \Delta, \bigwedge\{F_i : i<n\}$.

10.15 Inductive definition of the *degree* GF *of sentential reducibility* of an $\mathscr{L}_\infty$-formula F

(i) for $F \in AE$ we define $GF = 0$.

(ii) $G(\bigvee\{F_i : i<n<\omega\}) := \max\{GF_i : i<n\}$

(iii) $G(\bigwedge\{F_i : i<n<\omega\}) := \max\{GF_i : i<n\}+1$

For a finite formula set Δ we define $G(\Delta) := \sum_{F\in\Delta} G(F)$. We obviously always have $F(\Delta)<\omega$.

10.16. **Theorem** (sentential completeness)
If Δ is a finite formula set which is sententially valid and $\alpha := \max\{rk(F) : F\in\Delta^a\}$, *then there is an* $m<\omega$ *such that* $\vdash_0^{2\alpha+m} \Delta$.

Proof by induction on $G\Delta^a$.
1. If $G\Delta^a = 0$, then Δ is irreducible. By 10.12. Δ^a is of the shape Δ_0, F_1, F_2 with dual sentential atoms F_1 and F_2. By the tautology lemma we therefore obtain $\vdash_0^{2rkF_1} \Delta^a$. It is $rkF_1 \le \alpha$ and Δ may be obtained from Δ^a by $\bigvee$-importation. Hence there is an $m<\omega$ such that $\vdash_0^{2\alpha+m}\Delta$.
2. $G\Delta^a>0$. Then Δ^a has the form $\Delta_0, \bigwedge\{F_i : i<n\}$ for some $n<\omega$. By 10.13. the set Δ_0,F_i is sententially valid for all $i<n$. For $i<n$ it is $G(\Delta_0,F_i)<G(\Delta)$ and we obtain an $m_i<\omega$ such that $\vdash_0^{2\alpha_i+m_i} \Delta_0,F_i$ by the induction hypothesis. It is $\alpha_i = \max\{rkF : F\in\Delta_0,F_i\} \le \max\{rkF : F\in\Delta\} = \alpha$. For $m := \max\{m_i+1 : i<n\}$ we therefore obtain $\vdash_0^{2\alpha+m}\Delta^a$ by an $\bigwedge$-inference. Using $\bigvee$-importation we obtain the claim.

10.17. **Induction lemma**
For $n<\omega$ *we have* $\vdash_0^{\alpha_n} \neg F_x(\underline{0}), \neg\bigwedge_{k<\omega}(\neg F_x(\underline{k}) \vee F_x(\underline{k}')), F_x(\underline{n})$ *with* $\alpha_n = 2(rkF_x(\underline{0})+n)$.

Proof by induction on n
For n = 0 we have $\vdash_0^{\alpha_0} \neg F_x(\underline{0}), \neg\bigwedge_{k<\omega}(\neg F_x(\underline{k}) \vee F(\underline{k}'), F_x(\underline{0})$ by the tautology lemma 10.8.
For the induction step we have the induction hypothesis

(1) $\vdash_0^{\alpha_n} \neg F(\underline{0}), \neg\bigwedge_{k<\omega}(\neg F(\underline{k}) \vee F(\underline{k}')), F(\underline{n})$.

By a structural inference (1) yields

(2) $\vdash_0^{\alpha_n} \neg F(\underline{0}), \neg\bigwedge_{k<\omega}(\neg F(\underline{k}) \vee F(\underline{k}')), F(\underline{n}), F(\underline{n}')$.

Using the tautology lemma we obtain

(3) $\vdash_0^{\alpha_0} \neg F(\underline{0}), \neg\bigwedge_{k<\omega}(\neg F(\underline{k}) \vee F(\underline{k}')), \neg F(\underline{n}'), F(\underline{n}')$.

From (2) and (3) it follows

(4) $\vdash_0^{\alpha_n+1} \neg F(\underline{0}), \bigvee_{k<\omega}(F(\underline{k}) \wedge \neg F(\underline{k}')), F(\underline{n}) \wedge \neg F(\underline{n}'), F(\underline{n}')$.

using an $\bigwedge$-inference.
By an $\bigvee$-inference we obtain from (4)

(5) $\vdash_0^{\alpha_n+2} \neg F(\underline{0}), \bigvee_{k<\omega}(F(\underline{k}) \wedge \neg F(\underline{k}')), F(\underline{n}')$.

But it is $\alpha_n+2 = 2(rk(F(\underline{0}))+n)+2 = 2(rk\ F(\underline{0})+n') = \alpha_{n+1}$. This completes the induction step.

10.18. Exercises

1. Suppose that Δ,Γ are finite sets of $\mathcal{L}_\Omega$-formulas such that $\vdash_0^\alpha \Delta,\Gamma$ holds. Show the following claim.

There is an $\mathcal{L}_\Omega$-formula A with $rk(A) \leq \alpha$ which is an interpolation formula for Δ and Γ. That means: every predicate constant different from = and every set variable which occurs in A occurs both in Δ and Γ and we have $\vdash_0^{2\cdot\alpha} \Delta, A$ as well as $\vdash_0^{2\cdot\alpha} \Gamma, \neg A$.

2. Show that for all valid sentences we have $\vdash_0^{rk(F)} F$ (cf. exercise 9.8.).

§ 11. *Embedding of* $\mathbf{Z}_1$ *into* $\mathbf{Z}_\Omega$

The *-translation of the $\mathcal{L}_1$- formulas into the $\mathcal{L}_\Omega$-formulas has already been defined in § 9. This translation has the following property.

11.1. Lemma

If $F[x_1,\dots,x_n]$ *is an* $\mathcal{L}_1$*-formula which does not contain further free number variables, then for every* n*-tuple* $k_1,\dots,k_n$ *of natural numbers* $F[\underline{k}_1,\dots,\underline{k}_n]^*$ *is an* $\mathcal{L}_\Omega$*-formula of finite rank. This rank is independent from the choice of the n-tuple.*

Proof by induction on the length of the formula $F[x_1,\dots,x_n]$.

1. The claim holds trivially for atomic formulas.

2. If $F[x_1,\dots,x_n]$ is no sentential atom, then we obtain the claim immediately from the induction hypothesis.

3. Suppose that $F[\mathbf{x}]$ is a formula $\forall y\, G[y,\mathbf{x}]$. For any n+1-tuple $(l,k_1,\dots,k_n)$ the formula $G[l,\underline{\mathbf{k}}]^*$ is an $\mathcal{L}_\infty$-formula of finite rank m, say, by the induction hypothesis. Then $\forall y\ G[y,\underline{\mathbf{k}}]^*$ is the formula $\bigwedge_{l<\omega} G[l,\underline{\mathbf{k}}]^*$ whose rank obviously is m+1.

The case that $F[\mathbf{x}]$ is a formula $\exists y\, G[y,\mathbf{x}]$ is treated analogously.

In fact the *-translations of the $\mathcal{L}_1$-formulas form a fragment of $\mathcal{L}_\Omega$ in the sense as it will be introduced in chapter III.

11.2. Embedding lemma

If $F[x_1,\dots,x_n]$ *is an* $\mathscr{L}_1$*-formula which contains only the indicated number variables such that* $\mathbf{Z}_1 \vdash F[x_1,\dots,x_n]$*, then there is an ordinal* $\alpha < \omega+\omega$ *and an ordinal* $m < \omega$ *such that* $\models^{\alpha}_{m} F[\underline{k}_1,\dots,\underline{k}_n]^*$ *holds for every* n*-tuple* $(k_1,\dots,k_n)$ *of natural numbers.*

Proof

The proof is by induction on the length of the derivation $\mathbf{Z}_1 \vdash F[x_1,\dots,x_n]$.

1. If $F[x_1,\dots,x_n]$ is a sentential axiom, then $F[\underline{k}_1,\dots,\underline{k}_n]^*$ is a sententially valid formula and we obtain $\models^{m}_{0} F[\underline{k}_1,\dots,\underline{k}_n]^*$ for some $m<\omega$ by 10.16. and 11.1. [Here we have to check that the length of the derivation in 10.16. does not depend upon the choice of the terms $\underline{k}_1,\dots,\underline{k}_n$. If we do not want to do that, then we obtain $\models^{\omega}_{0} F[\underline{k}_1,\dots,\underline{k}_n]^*$ which is also sufficient for the proof of the lemma].

2. Suppose that $F[x_1,\dots,x_n]$ is a formula $\neg\forall x\, A[x_1,\dots,x_n] \vee A_x(t)[x_1,\dots,x_n]$. If we choose an n-tuple $\mathbf{\underline{k}} = (\underline{k}_1,\dots,\underline{k}_n)$, then $t_{\mathbf{x}}(\mathbf{\underline{k}})$ is a closed term t_0 such that $t_0^{\mathbb{N}} =: k$ say. By the tautology lemma we have

$$\models^{\alpha}_{0} \neg A_x(\underline{k})[\mathbf{\underline{k}}]^*,\ A_x(t)[\mathbf{\underline{k}}]^* \text{ for } \alpha := 2\,\mathrm{rk}(A_x(\underline{k})[\mathbf{\underline{k}}])^* < \omega.$$

By an $\bigvee$-inference this implies

$$\models^{\alpha+1}_{0} \neg\bigwedge_{k<\omega} A_x(\underline{k})[\mathbf{\underline{k}}]^*, A_x(t)[\mathbf{\underline{k}}]^*$$

which by $\bigvee$-importation entails

$$\models^{\alpha+3}_{0} \neg\bigwedge_{k<\omega} A_x(\underline{k})[\mathbf{\underline{k}}]^* \vee A_x(t)[\mathbf{\underline{k}}]^*.$$

But this is the formula $F[\mathbf{\underline{k}}]^*$.

Completely analogously we obtain $\models^{\alpha}_{0} (A_x(t) \rightarrow \exists x A)^*$.

3. Logical inferences

(mp) By the induction hypothesis we obtain $\alpha_1,\alpha_2 < \omega+\omega$ and $m_1,m_2 < \omega$ such that $\models^{\alpha_1}_{m_1} A^*$ and $\models^{\alpha_2}_{m_2} \neg A^* \vee B^*$. By the structural rule and $\bigvee$-exportation it follows $\models^{\alpha_2}_{m} A^*,B^*$ and $\models^{\alpha_2}_{m} \neg A^*,B^*$ for $m := \max\{m_1,m_2,\mathrm{rk}(B^*)+1\}$. By a cut this implies $\models^{\alpha}_{m} B^*$ for $\alpha := \max\{\alpha_1,\alpha_2\}+1 < \omega+\omega$.

($\forall$) By the induction hypothesis there are ordinals $\alpha_0 < \omega+\omega$ and $m<\omega$ such that $\models^{\alpha_0}_{m} (\neg A \vee B)_x[\underline{k},\mathbf{\underline{k}}]^*$ holds for all $k \in \omega$. By $\bigvee$-exportation and the variable condition $x \notin FV_1(A)$ this entails $\models^{\alpha_0}_{m} \neg A[\mathbf{\underline{k}}]^*, B_x(k)[\mathbf{\underline{k}}]^*$ for all $k \in \omega$. Using an $\bigwedge$-inference we obtain $\models^{\alpha_0+1}_{m} \neg A[\mathbf{\underline{k}}]^*, \bigwedge_{k<\omega} B_x(k)[\mathbf{\underline{k}}]^*$ and by $\bigvee$-importation finally $\models^{\alpha_0+3}_{m} (\neg A \vee \forall x B)[\mathbf{\underline{k}}]^*$.

4. Equality axioms

(i) According to (Ax1) we have $\models^{0}_{0} \underline{n} = \underline{n}$ for all $n \in \omega$. By an $\bigwedge$- inference this implies $\models^{1}_{0} \forall x(x = x)^*$.

(ii) $\models^{0}_{0} \neg\underline{n} = \underline{m}, \underline{m} = \underline{n}$ holds for all $n,m \in \omega$ by (Ax1). Using $\bigvee$-importation and two $\bigwedge$-inferences we obtain $\models^{4}_{0} \forall x \forall y(x = y \rightarrow y = x)^*$.

(iii) The transitivity axiom is proved similarly.

(iv) $\vdash_0^0 \neg \underline{n} = \underline{m}, t_x(\underline{n}) = t_x(\underline{m})$ holds by (Ax1) because we either have $\mathbb{N} \models \underline{n} \neq \underline{m}$ or $\mathbb{N} \models t_x(\underline{n}) = t_x(\underline{m})$. By $\bigvee$-importation and two $\bigwedge$-inferences we again obtain $\vdash_0^4 \forall x \forall y (x = y \to t = t_x(y))^*$.

(v) We have $\vdash_0^\alpha \underline{n} \neq \underline{m}, \neg F_x(\underline{n})^*, F_x(\underline{m})^*$ for $\alpha = 2\operatorname{rk} F_x(\underline{n})^* < \omega$ since we either have $n \neq m$ and therefore an axiom according to (Ax1) or it is $n = m$ and we may derive the formula using the tautology lemma. By $\bigvee$-importation and two $\bigwedge$-inferences it follows $\vdash_0^{\alpha+5} \forall x \forall y (x = y \to F \to F_x(y))^*$.

5. Mathematical axioms

$\vdash_0^0 \underline{0} \neq S\underline{n}$ holds for all $n \in \omega$ according to (Ax1). By an $\bigwedge$-inference this entails $\vdash_0^1 \forall x (\neg 0 = Sx)^*$.

$\vdash_0^0 \underline{S}\underline{n} \neq \underline{S}\underline{m}, \underline{n} = \underline{m}$ holds according to (Ax1). Hence $\vdash_0^4 \forall x \forall y (\underline{S}x = \underline{S}y \to x = y)^*$ as before.

$\vdash_0^0 \underline{Sn} = \underline{S}\underline{n}$ is an axiom according to (Ax1).

We are going to treat the defining equations for primitive recursive functions just in examples.

We have $\vdash_0^0 \underline{C}_k^n \underline{k}_1 \dots \underline{k}_n = \underline{k}$ since this is an axiom according to (Ax1). Hence $\vdash_0^n \forall x_1 \dots x_n (\underline{C}_k^n x_1 \dots x_n = \underline{k})$ by $\bigwedge$-inferences. The case of the constants P_k^n is treated analogously.

$\vdash_0^0 \operatorname{Sub}(g, h_1, \dots, h_m)(\underline{\mathbf{k}}) = g(h_1\underline{\mathbf{k}}) \dots (h_m\underline{\mathbf{k}})$ is an axiom according to (Ax1). By $\bigwedge$-inferences we obtain the translation of the defining axiom for Sub.

We have $\vdash_0^0 \underline{k} \neq \underline{0}, (Rgh)\underline{\mathbf{k}}\underline{k} = g\underline{\mathbf{k}}$ and $\vdash_0^0 \underline{k} \neq \underline{S1}, (\underline{Rgh})\underline{\mathbf{k}}\underline{k} = \underline{h}\underline{\mathbf{k}}\underline{l}(\underline{Rgh}\ \underline{\mathbf{k}}\ \underline{l}))$ according to (Ax1). Using $\bigvee$-importation and some $\bigwedge$-inferences we obtain the translation of the defining axiom for (Rgh).

It holds $\vdash_0^0 \neg R\underline{\mathbf{k}}, \chi_R\underline{\mathbf{k}} = \underline{1}$ and $\vdash_0^0 \chi_R\underline{\mathbf{k}} \neq \underline{1}, R\underline{\mathbf{k}}$ by (Ax1). By $\bigvee$-importation and $\bigwedge$-inferences this implies $\vdash_0^{3+n} \forall x_1 \dots x_n (Rx_1 \dots x_n \leftrightarrow \chi_R x_1 \dots x_n = \underline{1})^*$.

In a last step we have to show that (IND)* is provable in Z_Ω.

By the induction lemma 10.17. we have

$$\vdash_0^{\alpha_k} \neg A_x(\underline{0})^*, (\neg \forall y (A_x(y) \to A_x(\underline{S}y))^*, A_x(\underline{k})^*$$

for all $k < \omega$ and $\alpha_k = 2(\operatorname{rk} F_x(\underline{0}) + n)$. By an $\bigwedge$-inference we obtain

$$\vdash_0^\omega \neg A_x(\underline{0})^*, \neg \forall y (A_x(y) \to A_x(\underline{S}y))^*, (\forall x A)^*$$

and by $\bigvee$-importation this yields $\vdash_0^{\omega+3}$ (IND)*.

This terminates the proof of the embedding lemma. As a last remark we want to emphasize that only the presence of the induction scheme (IND) forced us to regard infinite derivations in Z_Ω. In absence of (IND) the embedding lemma would work with ω instead of $\omega+\omega$ and therefore yield finitary derivations.

11.3. Exercises

1. Let the formal theory $\mathbf{Z}_0$ be $\mathbf{Z}_1$ without IND. Prove that for any $\mathscr{L}_1$-formula $F[x_1,\dots,x_n]$ such that $\mathbf{Z}_0 \vdash F[x_1,\dots,x_n]$ there are natural numbers $i,j<\omega$ such that $\frac{i}{j} F[\underline{k}_1,\dots,\underline{k}_n]$ holds for every n-tuple $(k_1,\dots,k_n)$ of natural numbers.
2. For every $n<\omega$ we define an infinitary system $\mathfrak{Z}_n$ for $\mathscr{L}_\infty$ by the following rules: (Ax), ($\bigwedge$) and ($\bigvee$) are the same as in $\mathbf{Z}_\infty$.

(cut) If we have $\mathfrak{Z}_n \overset{\alpha_0}{\underset{\rho}{\vdash}} \Gamma, A$; $\mathfrak{Z}_n \overset{\alpha_1}{\underset{\rho}{\vdash}} \Delta, \neg A$ and $rk(A)<n+1+\rho$, then we also have $\mathfrak{Z}_n \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma,\Delta$ for all $\alpha>\alpha_0,\alpha_1$.

(IND_n) If we have $\mathfrak{Z}_n \overset{\alpha_0}{\underset{\rho}{\vdash}} \Gamma, F_x(\underline{0})$, $\mathfrak{Z}_n \overset{\alpha_1}{\underset{\rho}{\vdash}} \Gamma, \neg F_x(\underline{k}), F_x(\underline{Sk})$ for all $k\in\mathbb{N}$, $\alpha>\alpha_0,\alpha_1$ and $rk(F_x(\underline{0}))\le n$, then we also have $\mathfrak{Z}_n \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma, F_x(\underline{k})$ for all $k\in\mathbb{N}$.

Prove the following claims:

(i) If $F[x_1,\dots,x_j]$ is an $\mathscr{L}$-formula which contains only the indicated number variables such that Σ^0_n-$\mathbf{IND} \vdash F[x_1,\dots,x_j]$, then there are natural numbers k,m such that $\mathfrak{Z}_n \overset{k}{\underset{m}{\vdash}} F[\underline{i}_1,\dots,\underline{i}_j]^*$ holds for every j-tuple $(i_1,\dots,i_j)$ of natural numbers.

(ii) For every finite set Δ of $\mathscr{L}_\infty$-formulas $\mathfrak{Z}_n \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ implies $\mathbf{Z}_\infty \overset{\omega\cdot\alpha}{\underset{n+1+\rho}{\vdash}} \Delta$.

§12. Cut elimination for $\mathbf{Z}_\Omega$

We start this section by the remark that we also have a soundness theorem for $\mathbf{Z}_\Omega$.

12.1. Soundness theorem for $\mathbf{Z}_\Omega$

If Δ is a finite set of $\mathscr{L}_\Omega$-formulas such that $\overset{\alpha}{\underset{\rho}{\vdash}} \Delta$, then $\mathbb{N} \models \bigvee\{F : F\in\Delta\}$

The proof is essentially the proof of 5.3. In the induction which here may be formulated as induction on α we only have to take into account the additional case of a cut. There we have the induction hypothesises $\mathbb{N} \models \bigvee\{F : F\in\Delta\} \vee A$ and $\mathbb{N} \models \bigvee\{F : F\in\Delta\} \vee \neg A$. But this entails $\mathbb{N} \models \bigvee\{F : F\in\Delta\}$.

From the soundness and the completeness theorem for $\underset{0}{\vdash}$ we can see that the cut rule in fact is superfluous in the system $\mathbf{Z}_\Omega$. From $\mathbf{Z}_\Omega \overset{\alpha}{\underset{\rho}{\vdash}} F$ we obtain $\mathbb{N} \models F$ and thereof $\overset{\beta}{\underset{0}{\vdash}} F$ with $\beta = |F|$. By 10.4,however, $\overset{\beta}{\underset{0}{\vdash}} \Delta$ entails $\overset{\beta}{\underset{0}{\vdash}} \Delta$. So we may infer from $\overset{\alpha}{\underset{\rho}{\vdash}} F$ that $\overset{\beta}{\underset{0}{\vdash}} F$ holds for some $\beta<\Omega$. This shows that the cut rule in $\mathbf{Z}_\Omega$ is in principle eliminable. So we do have a cut elimation theorem for $\mathbf{Z}_\Omega$ but we do not yet have much information about the size of the ordinal β. Of course we know that the norm of the formula F suffices. But this is of little help since it is the aim of our consideration to obtain some information about this norm. Therefore we have to prove the cut elimination theorem for $\mathbf{Z}_\Omega$ in a different way. This proof must be done in such a way that we may keep control over the length of the derivation trees during the elimination procedure. The embedding of $\mathbf{Z}_1$ into $\mathbf{Z}_\Omega$ produced derivation trees of lengths below $\omega+\omega$ and finite cut rank. We are going to show that the resulting cut free derivations will have lengths less than ε_0.

12.2. **Elimination lemma**

If $\mathrm{rk}(F) = \rho$, $\vdash^{\alpha}_{\rho} \Delta, F$ *and* $\vdash^{\beta}_{\rho} \Gamma, \neg F$, *then* $\vdash^{\alpha\#\beta}_{\rho} \Delta, \Gamma$.

Proof by induction on $\alpha\#\beta$

1. Assume that either F or $\neg$F is not the mainformula of the last inference, where again we regard axioms as inferences without premises. Because of the symmetry of the claim without loss of generality we may assume this is the case for F.

1.1. If $\vdash^{\alpha}_{\rho} \Delta,F$ is an axiom so is $\vdash^{\alpha}_{\rho}\Delta$ and we obtain $\vdash^{\alpha\#\beta}_{\rho}\Delta,\Gamma$ by the structural rule 10.5.

1.2. If $\vdash^{\alpha}_{\rho} \Delta,F$ is the conclusion of an inference S whose premises are $\vdash^{\alpha_i}_{\rho} \Delta_i, F$, then we obtain $\vdash^{\alpha_i\#\beta}_{\rho} \Delta_i,\Gamma$ by the induction hypothesis. Because of $\alpha_i\#\beta<\alpha\#\beta$ the same inference S yields $\vdash^{\alpha\#\beta}_{\rho} \Delta,\Gamma$.

2. Now we assume that F as well as $\neg$F are the main formulas of the last inference. We then have to distinguish the following cases.

2.1. $\mathrm{rk}(F) = 0$, i.e. F is an atomic formula. Since axioms are the only inferences whose main formulas are atomic we have that $\vdash^{\alpha}_{\rho} \Delta,F$ as well as $\vdash^{\beta}_{\rho} \Gamma,\neg F$ are axioms with mainformula F or $\neg$F respectively. But then F or $\neg$F must have the form $(t \in X)$. Otherwise F were a formula $\underline{P}t_1...t_n$ such that $\mathbb{N} \models \underline{P}t_1...t_n$ which would contradict the fact that $\neg$F, i.e. $\neg\underline{P}t_1...t_n$ is the main formula of an axiom too, which means $\mathbb{N} \models \neg\underline{P}t_1...t_n$.

Because of the symmetry of the claim we again may assume $F \equiv t \in X$ without loss of generality . But then Δ has to contain a formula $s_1 \notin X$ such that $s_1^{\mathbb{N}} = t^{\mathbb{N}}$ and Γ a formula $s_2 \in X$ such that $s_2^{\mathbb{N}} = t^{\mathbb{N}}$. Then Δ,Γ too, is an axiom according to (Ax2) and it follows $\vdash^{\alpha\#\beta}_{\rho} \Delta,\Gamma$.

2.2. $\mathrm{rk}(F)>0$. By symmetry we again may assume that $F \equiv \bigwedge_{i<\nu} F_i$ for some $\nu \le \omega$. Then we have the following inferences

(1) $\vdash^{\alpha_i}_{\rho}\Delta, F_i, \bigwedge_{i<\nu} F_i$ for all $i<\nu \Rightarrow \vdash^{\alpha}_{\rho} \Delta, \bigwedge_{i<\nu} F_i$

and

(2) $\vdash^{\beta_0}_{\rho} \Gamma, \neg F_{i_0}, \bigvee_{i<\nu} \neg F_i \Rightarrow \vdash^{\beta}_{\rho} \Gamma, \bigvee_{i<\nu}\neg F_i$.

(If the formula $\bigwedge_{i<\nu} F_i$ does not occur in the premise we may add it by an application of the structural rule).

By the induction hypothesis we obtain $\vdash^{\alpha_i \# \beta}_{\rho} \Delta, \Gamma, F_i$ for all $i<\nu$ and $\vdash^{\alpha \# \beta_0}_{\rho} \Gamma, \Delta, \neg F_{i_0}$. Because of $rk(F_{i_0}) < rk(F) = \rho$ and $\alpha_{i_0} \# \beta < \alpha \# \beta$ as well as $\alpha \# \beta_0 < \alpha \# \beta$ we obtain $\vdash^{\alpha \# \beta}_{\rho} \Gamma, \Delta$ by a cut of rank $rk(F_{i_0}) < \rho$.

12.3. First elimination theorem

If $\vdash^{\alpha}_{\rho+1} \Delta$, *then* $\vdash^{2^\alpha}_{\rho} \Delta$.

Proof by induction on α

1. In the case that the last inference is not a cut of rank ρ we either have an axiom $\vdash^{\alpha}_{\rho+1} \Delta$ or the inference has the premises $\vdash^{\alpha_i}_{\rho+1} \Delta_i$ $(i<\nu\le\omega)$. In the case of an axiom we have $\vdash^{2^\alpha}_{\rho} \Delta$ by definition and in the other case we obtain $\vdash^{2^{\alpha_i}}_{\rho} \Delta_i$ for all $i<\nu$ by the induction hypothesis. Since the inference in question is not a cut of rank ρ - if it is a cut it must have cut rank $<\rho$ - we may apply the same inference to the premises $\vdash^{2^{\alpha_i}}_{\rho} \Delta_i$ to obtain $\vdash^{2^\alpha}_{\rho} \Delta$.
2. If the last inference is a cut of rank ρ, then we have the premises $\vdash^{\alpha_1}_{\rho+1} \Delta, A$ and $\vdash^{\alpha_2}_{\rho+1} \Delta, \neg A$ and $rk(A) = \rho$. By the induction hypothesis we obtain $\vdash^{2^{\alpha_1}}_{\rho} \Delta, A$ and $\vdash^{2^{\alpha_2}}_{\rho} \Delta, \neg A$. By an application of the elimination lemma it follows $\vdash^{2^{\alpha_1} \# 2^{\alpha_2}}_{\rho} \Delta$. For $\alpha_0 := \max\{\alpha_1, \alpha_2\}$, we have $\alpha_0 < \alpha$ and $2^{\alpha_1} \# 2^{\alpha_2} \le 2^{\alpha_0} \# 2^{\alpha_0} = 2^{\alpha_0'} \le 2^\alpha$. Hence $\vdash^{2^\alpha}_{\rho} \Delta$.

12.4. **Theorem** (Ordinal analysis of $\mathbf{Z}_1$)

If F *is a* Π^1_1*-sentence such that* $\mathbf{Z}_1 \vdash F$, *then we have* $\vdash^{\beta}_{0} F^*$ *for some* $\beta < \varepsilon_0$.

Proof

If $\mathbf{Z}_1 \vdash F$ holds for a Π^1_1-sentence F, then there is an $\alpha < \omega \cdot 2$ and an $m < \omega$ such that $\vdash^{\alpha}_{m} F^*$ by the embedding lemma. If we define $2_0(\alpha) := \alpha$ and $2_{n+1}(\alpha) := 2^{2_n(\alpha)}$, then m- fold application of the elimination theorem yields $\vdash^{2_m(\alpha)}_{0} F^*$. But we have $\omega+\omega < \omega^\omega = \omega_3(0)$ and this implies $2_m(\alpha) < \omega_{3+m}(0)$ for all $m<\omega$. Hence $2_m(\alpha) < \omega_\omega(0) = \varepsilon_0$.

12.5. **Corollary**

If F *is a* Π^1_1*-sentence such that* $\mathbf{Z}_1 \vdash F$, *then* $|F| < \varepsilon_0$.

The proof is obvious by 12.4. and 10.4.

12.6. Definition

a) $SP_0(\mathbf{Z}_1) := \{|F| : F \text{ is a } \Pi^1_1\text{-formula and } \mathbf{Z}_1 \vdash F\}$.

b) $|\mathbf{Z}_1| := \sup SP_0(\mathbf{Z}_1)$.

We call $SP_0(\mathbf{Z}_1)$ the *(basic-)spectrum* of $\mathbf{Z}_1$, $|\mathbf{Z}_1|$ is the *proof theoretic ordinal* of $\mathbf{Z}_1$.

As a consequence of 12.5. we obtain

12.7. Theorem

We have $SP_0(\mathbf{Z}_1) \subset \varepsilon_0$ *and* $|\mathbf{Z}_1| \le \varepsilon_0$.

The question which now canonically arises is if these bound are the exact ones. It will be answered in the following sections

12.8. Exercises

1. Prove that $|\mathbf{Z}_0| \le \omega$.
2. Show that the bound in the elimination lemma is the best possible one.
3. Prove the elimination lemma and the first elimination theorem for $\mathbf{Z}_n$.
4. Prove the following special case of the elimination lemma:

Assume $n \le \omega$, $k < \omega$ *and* $0 \le i_0 < \ldots < i_k < n$ *and for* $i < n$ *let* P_i *be an atomic formula. If* $\vdash^{\alpha}_{0} \Delta, \bigwedge\{P_i : i < n\}$ *and* $\vdash^{\beta}_{0} \Gamma, \bigvee\{\neg P_i : i < n\}, \neg P_{i_0}, \ldots, \neg P_{i_k}$, *then* $\vdash^{\alpha \# \alpha \# \beta}_{0} \Delta, \Gamma$.

5. Prove the following claims:

 (i) $|\Sigma^0_0\text{-}\mathbf{IND}| \le \omega^{\omega} = \omega_3(0)$.

 (ii) For $n \ge 1$ we have $|\Sigma^0_n\text{-}\mathbf{IND}| \le \omega_{n+2}(0)$.

§ 13. Formalization of transfinite induction

The following considerations hold for any language $\mathcal{L}$ which comprises the language of pure number theory.

13.1. Definition

(i) A relation $\prec \subset \mathbb{N} \times \mathbb{N}$ is *$\mathcal{L}$-definable*, if there is an $\mathcal{L}$-formula A such that $FV(A) = \{x,y\}$ and $n \prec m \Leftrightarrow \mathbb{N} \models A_{x,y}(\underline{n},\underline{m})$.

(ii) The *field of a relation* $\prec$ is defined by $\text{field}(\prec) := \{x : \exists y(x \prec y \vee y \prec x)\}$.

(iii) $\text{Tran}(\prec)$ is the formula $\forall x \forall y \forall z(x \prec y \wedge y \prec z \rightarrow x \prec z)$.

(iv) $\text{LO}(\prec)$ is the formula

$\forall x(\neg x \prec x) \wedge \text{Tran}(\prec) \wedge \forall x \forall y(x \in \text{field}(\prec) \wedge y \in \text{field}(\prec) \rightarrow x \prec y \vee y \prec x \vee x = y)$.

Obviously we have

$\mathbb{N} \models \text{LO}(\prec) \Leftrightarrow \prec$ is a linear order relation on $\mathbb{N}$.

(v) $\text{Prog}(\prec, X)$ is the formula $\forall x(x \in \text{field}(\prec) \wedge \forall y(y \prec x \rightarrow y \in X) \rightarrow x \in X)$.

We have $\mathbb{N} \models \text{Prog}(\prec, S)$ if and only if the class S is *progressive* with respect to the relation $\prec$, i.e., if all $\prec$-predecessors of x belong S, then also x belongs to S.

(vi) By $\text{Fund}(\prec, X)$ we denote the formula

$\text{Tran}(\prec) \wedge (\text{Prog}(\prec, X) \rightarrow \forall x(x \in \text{field}(\prec) \rightarrow x \in X))$.

$\mathbb{N} \models \text{Fund}(\prec, X)$ then means that the relation $\prec$ is transitive and wellfounded.

(vii) $\text{TI}(\prec, X)$ is the formula $\text{LO}(\prec) \wedge \text{Fund}(\prec, X)$.

(viii) $\text{WO}(\prec)$ is the formula $\forall X\, \text{TI}(\prec, X)$.

We have $\mathbb{N} \models \text{TI}(\prec, S)$ if and only if tranfinite induction holds for S.

13.2. Definition

For a wellfounded transitive relation we define the $\prec$-norm for $n \in \mathbb{N}$ by:

$$1)\quad |n|_{\prec} := \begin{cases} \{|m|_{\prec} : m \prec n\} & \text{if } n \in \text{field}(\prec) \\ \Omega & \text{otherwise} \end{cases}$$

$$2)\quad \|\prec\| := \{|n|_{\prec} : n \in \text{field}(\prec)\}$$

13.3. Lemma

If $\prec$ *is a wellfounded transitive relation and* $n \in \text{field}(\prec)$, *then* $|n|_{\prec}$ *and* $\|\prec\|$ *are ordinals.*

Proof

Since any transitive set of ordinals is itself an ordinal, it suffices to show that $|n|_{\prec}$ is a transitive set of ordinals. This will be done by induction along $\prec$. If $\alpha \in |n|_{\prec}$, then $\alpha = |m|_{\prec}$ for some $m \prec n$. Hence $\alpha \in \text{On}$ by the induction hypothesis. So $|n|_{\prec}$ is a set of ordinals. If $\beta \in |m|_{\prec} \in |n|_{\prec}$, then $\beta = |m_0|_{\prec}$ for some $m_0 \prec m \prec n$. Since $\prec$ is transitive we obtain $m_0 \prec n$ and this implies $\beta \in |n|_{\prec}$.

We hitherto have shown that $\|\prec\|$ is a set of ordinals. It remains to show that $\|\prec\|$ is transitive. For $\alpha \in \beta \in \|\prec\|$, however, there is an $m \in \text{field}(\prec)$ such that $\beta = |m|_\prec$. Hence $\alpha = |m_0|$ for some $m_0 \prec m$ which implies $\alpha \in \|\prec\|$. Thus $\|\prec\|$ is transitive.

13.4. Lemma

If $\prec$ is a wellfounded transitive relation, then $\|\prec\| = \sup\{|n|_\prec + 1 : n \in \text{field}(\prec)\}$.

Proof

Define $\gamma := \sup\{|m|_\prec + 1 : m \in \text{field}(\prec)\}$. Then $\|\prec\| \leq \gamma$. For $\alpha < \gamma$ there is an $m \in \text{field}(\prec)$ such that $\alpha \leq |m|_\prec < \|\prec\|$ which also shows $\gamma \leq \|\prec\|$.

13.5. Definition

(i) $\prec_\alpha = \{n : |n|_\prec < \alpha\}$

(ii) $\prec\restriction\alpha = \prec \cap \prec_\alpha^2$, i.e. $\prec\restriction\alpha$ is the relation $\prec$ restricted to elements of $\prec$-norms less than α.

13.6. Lemma

If $\prec$ is wellfounded and $\alpha < \|\prec\|$, then it is $\|\prec\restriction\alpha\| = \alpha$.

Proof

Since $n \in \prec_\alpha$ implies $|n|_\prec < \alpha$ we obviously have $\|\prec\restriction\alpha\| \leq \alpha$. If $\beta < \alpha$ then there is an $m \in \prec_\alpha$ such that $\beta = |m|_\prec$. Hence also $\alpha \leq \|\prec\restriction\alpha\|$.

13.7. Definition

An $\mathcal{L}_\infty$-formula F which has no occurence of X of the form $t \notin X$ is an X-*positive* formula.

13.8. Monotonicity lemma

Suppose that F *is an* X*-positive* $\mathcal{L}_\infty$*-formula and* $S, T \subset \mathbb{N}$ *are classes such that* $S \subset T$. *Then* $\mathbb{N} \models F_X[S]$ *entails* $\mathbb{N} \models F_X[T]$.

Proof by induction on rk(F)

1. The claim is trivial if X does not occur in F.
2. Suppose that $F \equiv t \in X$. Then $\mathbb{N} \models F_X[S] \Rightarrow t^{\mathbb{N}} \in S \Rightarrow t^{\mathbb{N}} \in T \Rightarrow \mathbb{N} \models F_X[T]$.
3. $F \equiv \bigwedge\{A_i : i \in I\}$. Then $\mathbb{N} \models F_X[S] \Rightarrow \mathbb{N} \models A_i[S]$ for all $i \in I \overset{\text{i.h.}}{\Rightarrow} \mathbb{N} \models A_i[T]$ for all $i \in I \Rightarrow \mathbb{N} \models F_X[T]$.
4. $F \equiv \bigvee\{A_i : i \in I\}$. Then $\mathbb{N} \models F_X[S] \Rightarrow \mathbb{N} \models A_i[S]$ for some $i \in I \overset{\text{i.h.}}{\Rightarrow} \mathbb{N} \models A_i[T]$ for some $i \in I \Rightarrow \mathbb{N} \models F[T]$.

13.9. **Boundedness lemma**

Suppose that $\prec$ is a transitive wellfounded $\mathcal{L}_1$-definable relation on $\mathbb{N}$ and Δ is a finite set of X-positive $\mathcal{L}_\infty$-formulas. If $\vdash_0^\alpha \neg\mathrm{Prog}(\prec, X), t_1 \notin X, \ldots, t_n \notin X, \Delta$, then it follows $\mathbb{N} \models \bigvee\{F_X[\prec_\gamma] : F \in \Delta\}$ where $\gamma = \beta + 2^\alpha$ and $\beta = \max\{|t_1^{\mathbb{N}}|_\prec, \ldots, |t_n^{\mathbb{N}}|_\prec\}$.

Proof *by induction on* α

1. In the case of an axiom according to (Ax1) the set Δ contains a true atomic formula. Hence $\mathbb{N} \models \bigvee \Delta_X[\prec_\gamma]$. In the case of an axiom (Ax2) Δ contains a formula $s \in X$ such that $s^{\mathbb{N}} = t_i^{\mathbb{N}}$ holds for some $i \in \{1, \ldots, n\}$. If $\beta_i = |t_i^{\mathbb{N}}|_\prec$, then it is $\beta_i \le \beta < \gamma$ and $\mathbb{N} \models (t_i \in X)[\prec_\gamma]$ since $\beta_i < \gamma$. Hence $\mathbb{N} \models \bigvee \Delta_X[\prec_\gamma]$.
2. Assume that the mainformula of the last inference belongs to Δ. Then we have the premises $\vdash_0^{\alpha_i} \neg\mathrm{Prog}(\prec, X), t_1 \notin X, \ldots, t_n \notin X, \Delta_i$, where Δ_i again only contains X-positive formulas. By the induction hypothesis it follows $\mathbb{N} \models \bigvee \Delta_i[\prec_{\gamma_i}]$ for $\gamma_i = \beta + 2^{\alpha_i}$. Using the monotonicity lemma we first obtain $\mathbb{N} \models \bigvee \Delta_i[\prec_\gamma]$ and therefore also $\mathbb{N} \models \bigvee \Delta[\prec_\gamma]$ since validity is preserved by all inferences.
3. Suppose that the mainformula of the last inference is

$$\neg\mathrm{Prog}(\prec, X), \text{ i.e. } \exists x(x \in \mathrm{field}(\prec) \wedge \forall y(y \prec x \to y \in X) \wedge x \notin X).$$

Then we have the premise

$$\vdash_0^{\alpha_0} \neg\mathrm{Prog}(\prec, X), t \in \mathrm{field}(\prec) \wedge \forall y(\neg y \prec t \vee y \in X) \wedge t \notin X, t_1 \notin X, \ldots, t_n \notin X, \Delta.$$

Thence we obtain by $\bigwedge$-inversion

(1) $\vdash_0^{\alpha_0} \neg\mathrm{Prog}(\prec, X), t \in \mathrm{field}(\prec) \wedge \forall y(\neg y \prec t \vee y \in X), t_1 \notin X, \ldots, t_n \notin X, \Delta$

and

(2) $\vdash_0^{\alpha_0} \neg\mathrm{Prog}(\prec, X), t \notin X, t_1 \notin X, \ldots, t_n \notin X, \Delta$.

Assume that $\mathbb{N} \not\models \bigvee \Delta[\prec_\gamma]$. Applying the induction hypothesis to (1) we obtain

(3) $\mathbb{N} \models \bigvee\{F : F \in \Delta\} \vee (t \in \mathrm{field}(\prec) \wedge \forall y(y \prec t \to y \in X))[\prec_{\gamma_0}]$ for $\gamma_0 = \beta + 2^{\alpha_0}$.

By the monotonicity lemma $\mathbb{N} \not\models \bigvee \Delta[\prec_\gamma]$ entails $\mathbb{N} \not\models \bigvee \Delta[\prec_{\gamma_0}]$ and by (3) we obtain $y \in \prec_{\gamma_0}$ for all $y \prec t^{\mathbb{N}}$, i.e. $|t^{\mathbb{N}}|_\prec \le \gamma_0$. If we define $\beta_0 := \max\{|t^{\mathbb{N}}|, \beta\}$, then we have $\beta_0 \le \gamma_0$. Applying the induction hypothesis (2) we obtain $\mathbb{N} \models \bigvee \Delta[\prec_{\beta_0 + 2^{\alpha_0}}]$.

But it is $\beta_0 \leq \beta + 2^{\alpha_0}$ and $2^{\alpha_0} + 2^{\alpha_0} \leq 2^{\alpha}$. Hence $\beta_0 + 2^{\alpha_0} \leq \beta + 2^{\alpha_0} + 2^{\alpha_0} \leq \beta + 2^{\alpha} = \gamma$ and it follows $\mathbb{N} \models \bigvee \Delta[\prec_\gamma]$ by the monotonicity lemma. This shows that our assumption was wrong.

13.10. **Boundedness theorem**

If $|\mathrm{Fund}(\prec, X)| \leq \alpha$, *then it is* $\|\prec\| \leq 2^{\alpha}$.

Proof

$|\mathrm{Fund}(\prec, X)| \leq \alpha$ by 10.4. implies $\vdash_0^{\alpha} \neg\mathrm{Prog}(\prec, X) \vee \forall x \in \mathrm{field}(\prec)(x \in X)$. Hence $\vdash_0^{\alpha} \neg\mathrm{Prog}(\prec, X), \forall x \in \mathrm{field}(\prec)(x \in X)$ and this implies $\mathbb{N} \models \forall x \in \mathrm{field}(\prec)(|x| \leq 2^{\alpha})$ by the boundedness lemma. Hence $\|\prec\| \leq 2^{\alpha}$.

13.11. **Corollary**

If $\mathbf{Z}_1 \vdash \mathrm{Fund}(\prec, X)$, *then it is* $\|\prec\| < \varepsilon_0$. *That means that all primitive recursive orderings whose wellfoundedness is provable in* $\mathbf{Z}_1$ *are of ordertype less than* ε_0.

13.12. **Remark**

By 13.11. it even follows that any $\mathscr{L}_1$-definable ordering $\prec$ whose wellfoundedness is provable in $\mathbf{Z}_1$ is of ordertype less than ε_0.

13.11. is the bridge to the more common definition of the proof theoretic ordinal of a formal system. Usually one defines:

The proof theoretic ordinal of a formal system T *is the supremum of the ordertypes of all primitive recursive definable order relations whose wellfoundedness is provable in* T.

3.11. shows that ε_0 is also an upper bound for the proof theoretic ordinal of $\mathbf{Z}_1$ defined in the classical way.

13.13. **Exercises**

1. Prove that $\sup\{|F| : \mathbb{N} \models F \wedge F \text{ is a } \Pi_1^1\text{-sentence}\} = \omega_1^{CK}$.

Hint: Show "$\geq$" using the boundedness theorem and prove that $|F| < \omega_1^{CK}$ holds for Π_1^1-sentences F which are valid in $\mathbb{N}$ by showing that the quasideductiontree of F is recursive.

2. Prove the following claims.

   (i) $\mathbf{Z}_0 \vdash \mathrm{Fund}(\prec, X)$ implies $\|\prec\| < \omega$.

   (ii) $|\mathbf{Z}_0| = \omega$.

3. Prove the following stronger versions of the boundedness lemma and the boundedness theorem:

Let $\prec$ be a $\mathcal{L}_1$-definable relation on $\mathbb{N}$.

(i) *Suppose that* Δ *is a finite set of* X-*positive* $\mathcal{L}_\infty$-*formulas. If we have* $\frac{\alpha}{1}\!\!\!\!\models$ $\neg\text{Prog}(\prec,X), t_1 \notin X, \ldots t_n \notin X, \Delta,$ *then* $\mathbb{N} \models \bigvee\{F_X[\prec_\gamma] : F \in \Delta\}$ *where* $\gamma = \beta + 2^\alpha$ *and* $\beta = \max\{|t_1^{\mathbb{N}}|_\prec,\ldots,|t_n^{\mathbb{N}}|_\prec\}$.

(ii) *If* $|\frac{\alpha}{1}$ $\text{Fund}(\prec,X)$, *then* $\|\prec\| \le 2^\alpha$.

4. Show that all primitive recursive orderings whose wellfoundedness is provable in Σ_n^0-**IND** $(n \ge 1)$ are of ordertype less than $\omega_{n+3}(0)$.

5. Prove that $|\frac{4\cdot\alpha}{0}$ $\neg\text{Prog}(\prec,X), t \in X$ where $\alpha := |t^{\mathbb{N}}|_\prec + 1$ holds for all primitive recursive wellorderings $\prec$ and all terms t.

6. Let $\prec$ be a primitive recursive wellordering. Let U be a new predicate constant. The infinitary system $\mathbf{Z}_\infty$ + ProgR($\prec$,U) is the system $\mathbf{Z}_\infty$ with the additional rule

(ProgR($\prec$,U)) If $|\frac{\alpha_s}{\rho}$ $\Delta, s \in U$ and $\alpha_s < \alpha$ for all s such that $s^{\mathbb{N}} \prec t^{\mathbb{N}}$, then $|\frac{\alpha}{\rho}$ $\Delta, t \in U$.

Prove the following claims:

(i) $\mathbf{Z}_\infty$ + ProgR($\prec$,U) $|\frac{6}{0}$ Prog($\prec$,U).

(ii) If $\mathbf{Z}_\infty$ $|\frac{\alpha}{\rho}$ $\neg$Prog($\prec$,U),Δ then $\mathbf{Z}_\infty$ + ProgR($\prec$,U) $|\frac{2\cdot\alpha}{\max\{1,\rho\}}$ Δ.

(iii) Prove the elimination lemma and theorem for $\mathbf{Z}_\infty$ + ProgR($\prec$,U).

(iv) If $\mathbf{Z}_\infty$ + ProgR($\prec$,U) $|\frac{\alpha}{0}$ $t \in U$ then $|t^{\mathbb{N}}| \le \alpha$.

§ 14. On the consistency of formal and semi formal systems

Inspired by Hilbert's program and Gentzen's consistency proof for $\mathbf{Z}_1$ one formerly defined the proof theoretic ordinal of a formal theory T as the ordertype of the least wellordering which is needed for the consistency proof of T. This definition, however, is somewhat problematic since it depends on the means which are allowed besides the induction along the wellordering. Nevertheless we are going to convince ourselves that ε_0 also in the sense of that definiton -properly interpreted- is an upper bound for the proof theoretic ordinal of $\mathbf{Z}_1$. In order to do that we will first sketch that our hitherto considerations also comprehend a consistency proof for $\mathbf{Z}_1$ which besides the transfinite induction along the ordering $\prec$ defined in 8.7. only uses means which are formalizable in $\mathbf{Z}_1$ itself. Since the ordertype of $\prec$ is ε_0 we obtain that ε_0 is an upper bound in the above sense.

14.1. Definition

A (semi-)formal system T is *semantically consistent*, if there is no formula A such that $T \vdash A \wedge \neg A$.

It is easy to see that $\mathbf{Z}_1$ is semantically consistent. We just have to look at the soundness theorem 3.13. in order to conclude that there is no formula A such that $T \vdash A \wedge \neg A$ because otherwise we also had $\mathbb{N} \models A \wedge \neg A$.
An inspection of this proof will show that the only induction we used there is complete induction, i.e. induction along a wellordering (which also may be defined primitve recursively) of ordertype ω. Of course we cannot yet conclude that ω is a candidate for the proof theoretic ordinal of $\mathbf{Z}_1$. The reason for this short induction lies in the fact that this consistency proof is in no way finitistic. We will not enter a discussion about finitistic means. For our purposes it will suffice to call a proof *finitistic* if it may be formalized in $\mathbf{Z}_1$ with the scheme (IND) restricted to Σ_1^0-formulas. Now it is impossible to formalize the notion of validity in $\mathbb{N}$ even in $\mathbf{Z}_1$. To obtain a more finitistic consistency proof it is necessary to describe the consistency of a formal theory in a more syntactical way.

14.2. Definition

A (semi-)formal system T is *syntactically consistent*, if there is a formula A such that $T \nvdash A$.

We are now going to show that for 'reasonable' formal- and semiformal systems the notion of semantical and syntactical consistency coincide.

14.3. Definition

(i) An inference $\vdash A_1, \dots, \vdash A_n \Rightarrow \vdash F$ is a *sentential inference*, if the formula $\neg A_1 \vee \dots \vee \neg A_n \vee F$ is sententially valid.

(ii) An inference $\vdash A_1, \dots, \vdash A_n \Rightarrow \vdash F$ is a *permitted inference* of a formal system T if $T \vdash A_1, \dots, T \vdash A_n$ entails $T \vdash F$.

(iii) A (semi-)formal system T is *sententially closed*, if every sentential inference is a permitted inference of T.

14.4. Theorem

A sententially closed system is semantically consistent if and only if it is syntactically consistent.

Proof

In a semantically consistent formal system no formula $A \wedge \neg A$ is derivable. This shows that it is syntactically consistent.

Now suppose that T is a semantically inconsistent formal system. Then there is a formula A such that $T \vdash A \wedge \neg A$. For every formula F, however, the formula $A \wedge \neg A \rightarrow F$ is sententially valid. Since T is sententially closed we obtain $T \vdash F$ for every formula F which shows that T also is syntactically inconsistent.

By $\mathbf{Z}^{\omega}_{\varepsilon_0}$ we denote the semiformal subsystem of $\mathbf{Z}_{\Omega}$ which contains only formulas of rank below ω and derivations of length below ε_0 and cut rank strictly less than ω.

14.5. **Theorem**

The semiformal system $\mathbf{Z}^{\omega}_{\varepsilon_0}$ *is syntactically consistent.*

Proof

We show that there is a formula F of rank$<\varepsilon_0$ for which we have $\not\models^{\alpha}_{n} F$ for all $\alpha<\varepsilon_0$ and $n<\omega$. Let F be a closed atomic formula such that $\mathbb{N} \not\models F$. If we assume $\models^{\alpha}_{n} F$ for some $\alpha<\varepsilon_0$ and $n<\omega$, then we obtain $\models^{\omega_n\alpha}_{0} F$ and $\omega_n\alpha$ is still less than ε_0. But an easy induction on β shows that $\models^{\beta}_{0} F$ is impossible. F is neither an axiom nor F may be inferred by an inference according to the $\bigwedge$- or $\bigvee$-rule since it then had to contain a logical symbol.

14.6. **Lemma**

The semiformal system $\mathbf{Z}^{\omega}_{\varepsilon_0}$ *is sententially closed.*

Proof

If $\neg A_1 \vee \dots \vee \neg A_n \vee F$ is a sententially valid formula whose rank is less than ε_0, then by 10.16. there is an $\alpha < 2\cdot\max\{rk(A_1),\dots,rk(A_2),rk(F)\}+\omega<\varepsilon_0$ such that $\models^{\alpha}_{0}\ A_1,\dots,\neg A_n,F$. If we assume that there are ordinals $\alpha_1,\dots,\alpha_n<\varepsilon_0$ and $\delta_1,\dots,\delta_n<\omega$ such that $\models^{\alpha_1}_{\delta_1} A_1,\dots,\models^{\alpha_n}_{\delta_n} A_n$ we obtain by cuts $\models^{\beta}_{\delta} F$ for some $\beta<\max\{\alpha_1,\dots,\alpha_n,\alpha\}+\omega<\varepsilon_0$ and $\delta := \max\{\delta_1,\dots,\delta_n,rk(A_1)+1,\dots,rk(A_n)+1\}<\omega$.

From 14.5 and 14.6 we obtain

14.7. **Lemma**

There is no formula F *such that* $rk(F)<\varepsilon_0$ *and* $\models^{\alpha}_{n} F \wedge \neg F$ *holds for some* $\alpha<\varepsilon_0$ *and* $n<\omega$. *That means that the system* $\mathbf{Z}^{\omega}_{\varepsilon_0}$ *is semantically consistent.*

14.8. Theorem

The formal system $\mathbf{Z}_1$ is consistent.

Proof

If $\mathbf{Z}_1 \vdash A \wedge \neg A$, then we obtain by the embedding lemma $\frac{\alpha}{n} A^* \wedge \neg A^*$ with $n<\omega$ and $\alpha<\omega\cdot 2<\varepsilon_0$. Since $rk(A^*)<\omega$ this contradicts 14.7.

Now we have to answer the question what it is gained by this consistency proof in comparison to the consistency proof via the soundness theorem.

We may answer the question in so far that our consistency proof beside transfinite induction up to ε_0 only used finitary means, i.e.means which are at least formalizable in $\mathbf{\Sigma_1^0}$**-IND** ($= \mathbf{Z}_0 + \Sigma_1^0$-IND). Since our real concern will be the impredicative system of chapter III we did not tailor the cut elimination theorem in such a way that we are able to obtain this result in an obvious way. But it is quite easy to sketch how our consistency proof may be formalized in the system $\mathbf{Z}_1$ augmented by the scheme $TI(\prec,X)$. This at least will show that the only means of the consistency proof which really exceeds that of $\mathbf{Z}_1$ is transfinite recursion along a primitive recursive wellordering of ordertype ε_0. Since the exact proof is a bit cumbersome and in fact outside our real concern we just will sketch its strategy.

First one observes that the proof trees resulting from the embedding lemma are in fact recursive trees. Then one has to convince oneself that the cut elimination procedure preserves the recursiveness of the trees. Recursive trees, however, can be formalized in $\mathbf{Z}_1$. To assure the wellfoundedness of these formalized derivation trees one has to assign (codes for) ordinals to the nodes of the tree. As we have seen the ordinals below ε_0 suffice for this purpose and we may represent these ordinals in $\mathbf{Z}_1$ by their codes developed in 8.7. So we obtain a recursive function $\mathfrak{f}$ such that

(A) $\mathbf{Z}_1 + TI(\prec,X) \vdash \mathrm{Proof}_{\mathbf{Z}_1}(x,\ulcorner F \urcorner) \rightarrow (\mathfrak{f}(x) \frac{\alpha}{0} \ulcorner F^* \urcorner)$

for a number variable x. Here $\mathrm{Proof}_{\mathbf{Z}_1}$ is the usual proof predicate for $\mathbf{Z}_1$ and $(x \frac{\alpha}{0} \ulcorner F \urcorner)$ formalizes the sentence

"x is the index of a recursive tree whose nodes are labeled with codes for $\mathscr{L}_\Omega$-formulas and ordinals increasing from top to bottom such that the tree is locally correct with respect to the inference rules of $\mathbf{Z}_\Omega$ and whose bottom node is labeled by the code for the formula F^* and the ordinal α"

(cf. [Pohlers 1981]).

We easily obtain

(B) $\mathbf{Z}_1 + TI(\prec,X) \vdash \forall x(\neg(x \frac{\alpha}{0} \ulcorner \underline{0} = \underline{1} \urcorner))$.

and conclude from (A)

(C) $\mathbf{Z}_1 + TI(\prec, X) \vdash \exists x (Proof_{\mathbf{Z}_1}(x, \ulcorner F \urcorner)) \rightarrow \exists x (x \,|\!\frac{\alpha}{0}\, \ulcorner F^{*} \urcorner)$.

From (B) and (C) we finally obtain

(D) $\mathbf{Z}_1 + TI(\prec, X) \vdash \neg \exists x (Proof_{\mathbf{Z}_1}(x, \ulcorner \underline{0} = \underline{1} \urcorner))$

which gives the formalization of the syntactical consistency of $\mathbf{Z}_1$.

The above sketched formalization has some interesting consequences especially for systems stronger than $\mathbf{Z}_1$. We do not have the time here to go into more details. In the epilogue, however, we try to give a short review of these results. A discussion about how this consistency proof fits into Hilbert's program will be given at the end of this chapter.

§ 15. The wellordering proof in $\mathbf{Z}_1$

In the following section we are going to describe ordinals $<\varepsilon_0$ by their arithmetizations as defined in 8.7. Since equality and the order relation between ordinals are primitive recursive relations we may speak about those ordinals in the language $\mathcal{L}_1$. Nevertheless we are going to keep our familiar notations, i.e. $\alpha, \beta, \gamma, \ldots$ now denote codes for odinals in E and $\alpha = \beta$ as well $\alpha < \beta$ denote the primitive recursive relations $\equiv$ and $\prec$ respectively on the codes of the ordinals as defined in 8.7.

15.1. **Lemma** (provable in $\mathbf{Z}_1$)

For $\mu \neq 0$ *and* $\alpha < \beta + \omega^{\mu}$ *there is a natural number* n *and an ordinal* $\delta < \mu$ *such that* $\alpha < \beta + \omega^{\delta} \cdot n$.

Proof

We informally work in $\mathbf{Z}_1$ (cf. exercise 15.10.1). If $\alpha \leq \beta$, then we are done choosing $\delta = 0$ and $n = 1$. So assume $\beta < \alpha$. Then there is an α_0 such that $\alpha = \beta + \alpha_0 < \beta + \omega^{\mu}$. We develop α_0 in Cantor normal form and obtain $\alpha_0 = \omega^{\alpha_1} + \ldots + \omega^{\alpha_k} < \omega^{\mu}$. Then we have $\alpha_1 < \mu$ and $\alpha_0 < \omega^{\alpha_1} \cdot (k+1)$. Hence $\alpha = \beta + \alpha_0 < \beta + \omega^{\alpha_1} \cdot (k+1) = \beta + \omega^{\delta} \cdot n$ for $\delta := \alpha_1$, $n := k+1$.

15.2. **Lemma**

$\mathbf{Z}_1 \vdash F$ *implies* $\mathbf{Z}_1 \vdash F_X(G)$ *for every class term* $\{x : G(x)\}$ *which is given by a* $\mathcal{L}_1$*-formula* G.

Proof by induction on the definition of $\mathbf{Z}_1 \vdash F$

If F is a logical or mathematical axiom of $\mathbf{Z}_1$, then obviously $F_X(G)$ is an axiom of the same kind.

If F is derived by an inference modus ponens or one of the quantifier inferences, then the claim easily follows from the induction hypothesis. In the case of an quantifier inference we have to ensure that the requirement for the eigenvariable of the inference is not violated. This may easily be obtained by renaming the variable.

By Fund(α,X) we abbreviate the formula $\mathrm{Tran}(\prec) \wedge (\mathrm{Prog}(\prec,X) \rightarrow \forall \xi < \alpha (\xi \in X))$. By TI($\alpha$,X) we denote the formula $\mathrm{LO}(\prec) \wedge \mathrm{Fund}(\alpha,X)$. Essentially Fund($\alpha$,X) says that the relation $\prec \restriction \alpha$ is wellfounded and TI(α,X) that $\prec \restriction \alpha$ is a wellordering. The aim of the current section is to prove that transfinite induction along every proper initial segment of the wellordering $\prec$ is provable in $\mathbf{Z}_1$. It is quite easy to see that $\mathbf{Z}_1$ proves LO($\prec$) (Though easy, the proof in fact is a bit cumbersome. Since not much can be learned by this proof we omit it and take it for granted that $\mathbf{Z}_1$ proves LO($\prec$)). So it remains to show Fund(α,X) for all $\alpha < \varepsilon_0$. Since Fund(0,X) holds trivially we are done if we succeed in proving the following theorem.

15.3. **Theorem**

$\mathbf{Z}_1 \vdash \mathrm{Fund}(\alpha,X)$ *implies* $\mathbf{Z}_1 \vdash \mathrm{Fund}(\omega^\alpha,X)$.

As a consequence of 15.3. we then obtain

15.4. **Theorem**

$\mathbf{Z}_1$ *proves the formula* Fund(α,X) *for all* $\alpha < \varepsilon_0$.

Proof

If $\alpha < \varepsilon_0$, then there is an $n < \omega$ such that $\alpha < \omega_n(0)$. Since $\mathbf{Z}_1 \vdash \forall x (\neg x \prec 0)$ we have $\mathbf{Z}_1 \vdash \mathrm{Fund}(0,X)$. n-fold application of 15.3. leads to $\mathbf{Z}_1 \vdash \mathrm{Fund}(\omega_n(0),X)$, i.e. $\mathbf{Z}_1 \vdash \mathrm{Prog}(\prec,X) \rightarrow \forall x < \omega_n(0)\ (x \in X)$. Because of $\mathbf{Z}_1 \vdash \forall x (x < \alpha \rightarrow x < \omega_n(0))$ this implies $\mathbf{Z}_1 \vdash \mathrm{Prog}(\prec,X) \rightarrow \forall x < \alpha (x \in X)$, i.e. $\mathbf{Z}_1 \vdash \mathrm{Fund}(\alpha,X)$.

In order to prove 15.3. we define a jumpoperator Sp

$$\mathrm{Sp}(X) = \{\alpha : \forall \xi (\xi \subset X \rightarrow \xi + \omega^\alpha \subset X)\} \quad [\xi \subset X \text{ abbreviates } \forall x (x < \xi \rightarrow x \in X)],$$

which enables us to jump from α to ω^α.

15.5. **Lemma**

The following formula is provable in $\mathbf{Z}_1$:

$$\mathrm{Fund}(\alpha,\mathrm{Sp}(X)) \to \mathrm{Fund}(\omega^\alpha,X).$$

Proof

We have the suppositions

(1) $\mathrm{Prog}(\prec,\mathrm{Sp}(X)) \to \alpha \subset \mathrm{Sp}(X)$

and

(2) $\mathrm{Prog}(\prec,X)$

and have to conclude $\omega^\alpha \subset X$. By (2) and lemma 15.6. below it follows

(3) $\mathrm{Prog}(\prec,\mathrm{Sp}(X))$.

From (3) and (1) we at first obtain $\alpha \subset SP(X)$ and this together with (3) implies $\alpha \in \mathrm{Sp}(X)$. If we choose $\xi = 0$ in the definition of SP(X) we obtain $\omega^\alpha \subset X$.

15.3. now is an immediate consequence of 15.5. For if $\mathbf{Z}_1 \vdash \mathrm{Fund}(\alpha,X)$, then we obtain by 15.2.$\mathbf{Z}_1 \vdash \mathrm{Fund}(\alpha,\mathrm{Sp}(X))$. This and 15.5. entail $\mathbf{Z}_1 \vdash \mathrm{Fund}(\omega^\alpha,X)$.

15.6. **Lemma**

$\mathbf{Z}_1 \vdash \mathrm{Prog}(\prec,X) \to \mathrm{Prog}(\prec,\mathrm{Sp}(X))$.

Proof

We have the presupposition

(1) $\mathrm{Prog}(\prec,X)$

and want to show $\mathrm{Prog}(\prec,\mathrm{Sp}(X))$, i.e. $\forall\beta(\beta \subset \mathrm{Sp}(X) \to \beta \in \mathrm{Sp}(X))$. To do that we choose an arbitrary β and assume

(2) $\beta \subset \mathrm{Sp}(X)$.

We then have to prove $\beta \in \mathrm{Sp}(X)$, i.e. $\forall\xi(\xi \subset X \to \xi+\omega^\beta \subset X)$.

Let ξ be a (code for an) ordinal such that

(3) $\xi \subset X$.

The claim is $\xi+\omega^\beta \subset X$. Therefore assume $\eta < \xi+\omega^\beta$.

1. $\beta = 0$. Then we have $\eta \leq \xi$. If $\eta < \xi$, then we have $\eta \in X$ by (3). $\xi \in X$ follows from (1) and (3).

2. $\beta > 0$. Then by 15.1. there is a $\beta_0 < \beta$ and an $n < \omega$ such that $\eta < \xi+\omega^{\beta_0}\cdot n$. We show

(4) $\xi+\omega^{\beta_0}\cdot n \subset X$

by induction on n. For $n = 0$ this is (3). For $n = n_0'$ we have the induction hypothesis

(5) $\xi+\omega^{\beta_0}\cdot n_0 \subset X$.

Since $\beta_0<\beta$ we obtain by (2) $\beta_0 \in Sp(X)$, i.e. $\forall\xi(\xi\subset X \rightarrow \xi+\omega^{\beta_0}\subset X)$. This especially implies $\xi+\omega^{\beta_0}\cdot n_0 \subset X \rightarrow \xi+\omega^{\beta_0}\cdot n_0+\omega^{\beta_0}\subset X$.

Together with (5) we therefore obtain

(6) $\xi+\omega^{\beta_0}\cdot n_0' \subset X$.

This finishes the inductionstep and the lemma is proved.

At this place we should notice that the essential means for the proof of theorem 15.3. is the scheme of complete induction. If T is any formal theory extending $\mathbf{Z}_1$, then 15.3 holds for every wellordering $\prec$ whose defining formula is admitted in the scheme of complete induction on T and for which 15.1. is provable in T.

As a corollary of theorem 15.4. we obtain

15.7. Theorem

ε_0 is the least upper bound of the ordertypes of the primitive recursively definable order relations whose wellfoundedness is provable in $\mathbf{Z}_1$. Moreover we have $\varepsilon_0 = \sup\{\|\prec\| : \prec \text{ is } \mathcal{L}_1\text{-definable and } \mathbf{Z}_1\vdash \mathrm{Fund}(\prec,X)\} = \sup\{\|\prec\| : \prec \text{ is primitive recursive and } \mathbf{Z}_1\vdash \mathrm{Fund}(\prec,X)\}$.

15.8. Theorem

For every $\alpha<\varepsilon_0$ there is a Π^1_1-sentence F such that $\mathbf{Z}_1\vdash F$ and $\alpha<|F|$.

Proof

If $\alpha<\varepsilon_0$, then it is $2^\alpha+1<\varepsilon_0$. By 15.4. we have $\mathbf{Z}_1\vdash \mathrm{Fund}(2^\alpha+1,X)$. But we have $\alpha<|\mathrm{Fund}(2^\alpha+1,X)|$ since by 13.10. $|\mathrm{Fund}(2^\alpha+1,X)|\le\alpha$ implies $2^\alpha+1\le 2^\alpha$.

15.9. Corollary

$SP_0(\mathbf{Z}_1) = |\mathbf{Z}_1| = \varepsilon_0$

15.10. Exercises

1. Suppose that α,β,μ,δ and n are ordinals as in 15.1. such that $\beta+\omega^\mu<\varepsilon_0$. Sketch that $\ulcorner\delta\urcorner$ and n are primitive recursively computable from $\ulcorner\alpha\urcorner$, $\ulcorner\beta\urcorner$ and $\ulcorner\mu\urcorner$.
2. Let $\mathrm{Proof}_{\mathbf{Z}_1}(n,v)$ be an arithmetical proof predicate for $\mathbf{Z}_1$. Define

$u\prec w :\equiv (u<w \wedge \forall x<w\, \neg\mathrm{Proof}_{\mathbf{Z}_1}(x,\ulcorner 0=1\urcorner)) \vee (w<u \wedge \exists x<u\, \mathrm{Proof}_{\mathbf{Z}_1}(x,\ulcorner 0=1\urcorner))$.

and a relation $R\subset \mathbb{N}\times\mathbb{N}$ by:

$R(n,m) :\Leftrightarrow \mathbb{N} \models \underline{n} \prec \underline{m}$.

(i) Show that R is a wellordering on $\mathbb{N}$

(ii) Compute $\|R\|$

(iii) Show $\mathbf{Z}_1 \vdash LO(\prec)$

(iv) Define $F :\equiv \forall z \leq x \neg Proof_{Z_1}(z, \ulcorner 0 = 1 \urcorner)$. Show $\mathbf{Z}_1 \vdash Prog(\prec, F)$

(v) Show $\mathbf{Z}_1 \nvdash Fund(\prec, X)$

§ 16. The use of Gentzen's consistency proof for Hilbert's program.

Our proof theoretic analysis of the formal system $\mathbf{Z}_1$ is essentially the same as Gentzen's original analysis. But contrary to our motivation -ordinal analysis of $\mathbf{Z}_1$- Gentzen originally was guided by the idea to give a consistency proof for $\mathbf{Z}_1$ in the spirit of Hilbert's program. He therefore (in the later version) avoided semiformal systems but showed the syntactic consistency of a formal system which is equivalent to $\mathbf{Z}_1$ by a partial cut elimination. In his proof it becomes immediately plain that its only non finitistic means is transfinite induction along a wellordering of ordertype ε_0. For this reason the Hilbert School believed that only a tiny extension of the finitistic standpoint - here accepting the intuitively plausible fact that the ordering in question really is wellordered - would suffice to carry through Hilbert's program.

Now we want to examine if this extension really is just a tiny one. For this purpose imagine an opponent who doubts the consistency of pure number theory. Let us assume that our opponent is able to understand mathematical reasoning and we are trying to convince him by Gentzen's proof. He really accepts all steps of the proof but in the end begs for an explanation of transfinite induction up to ε_0 since this is beyond his finitistic understanding. We therefore try to substantiate this induction as finitistic as possible. We avoid the notion of an ordinal, only speak about orderings on the natural numbers and so will necessarily end up with a proof which essentially is the same as that we gave in the preceeding section.

Due to his mathematical abilities our opponent very quickly will notice that the crucial point of the proof is lemma 15.5. If we define $Sp_0(X) = X$ and $Sp_{k+1}(X) = Sp(Sp_k(X))$, then we have to start with $Fund(0, Sp_k(X))$ and then by iterated use of lemma 15.5. decrease the number of jumpoperators in order to get to $Fund(\omega_k(0), X)$. But this also means that we need 15.6. in the form $Prog(\prec, Sp_{k-1}(X)) \rightarrow Prog(\prec, Sp_k(X))$. In the proof of 15.6, however, we proved the formula

(4) $\xi+\omega^{\beta_0}\cdot n \subset Sp_{k-1}(X)$

by complete induction. This shows that we need the induction scheme for formulas of the complexity of $Sp_{k-1}(X)$. In our proof we did not pay attention to a parsimonious definition of the jumpoperator. But even under the most careful definition the jumpoperator increases the number of alternations of quantifiers at least by one. That means that Sp(S) in the arithmetical hierachy always is one level higher than S. In order to climb up to ε_0 the constant k has to run through all natural numbers. That means, however, that we cannot restrict the complexity of the formulas in the induction scheme. At this place our opponent will argue that by that we fully use the means of pure number theory and even exceed it. But since he doubts the consistency of pure number theory he cannot accept our proof. We cannot advance a mathematical argument against his argumentation.

Being aware of Gödel's second theorem this situation is not too astonishing. If Gödel's theorem is more than a mere formal triviality but has a genuine meaning, then one cannot expect to bypass it by a tiny extension. Therefore we cannot expect results which incorporate a real progress in the spirit of Hilbert's program.

We want to emphasize, however, that this objections only meet the shortened version of Hilbert's program as we presented it in the introduction. Hilbert originally also spoke of the elimination of ideal elements in mathematics. This subtle part of his program is in fact realisable in many directions. Nevertheless we must be aware that a mathematical proof of the consistency of mathematics is impossible.

Although Gentzen's result is of little help in the spirit of Hilbert's program it has consequences which correspond better to Brouwer's intuitionistic point of view, a standpoint opposed by Hilbert.

By sharpening the considerations in the end of section 14 Gentzen's proof may be interpreted in the following way:

The consistency problem of pure number theory with the unrestricted scheme of complete induction can be reduced to the question, if a system without complete induction without the law of the excluded middle but with transfinite induction along all initial segments of ε_0 for formulas of very restricted complexity is consistent. Since one has a very good picture of an order relation of ordertype ε_0 and the latter system does not allow indirect inferences it intuitively is completely plain that the latter system is consistent although its proof theoretic ordinal is the same as that of pure number theory. This kind of *reductive proof theory* is in full coherence with Gödel's second theorem.

CHAPTER II

The autonomous ordinal of the infinitary system Z_∞ and the limits of predicativity

By the soundness theorem 12.1. for the infinitary system $\mathbf{Z}_\Omega$ and lemma 9.2. we have

$$\mathbf{Z}_\Omega \mid\frac{\alpha}{\rho} F \Rightarrow |F| < \Omega.$$

That means that the ordinal Ω has the following closure property: For $\alpha,\rho \in \Omega$ and $\mathbf{Z}_\Omega \mid\frac{\alpha}{\rho} F$ we also have $|F| \in \Omega$. By Ω we usually denote the first uncountable regular ordinal. Later on we are going to use Ω as a formal symbol, whose standard interpretation is the first regular ordinal $\aleph_1$. But we also will alternatively interpret Ω by other ordinals (cf. chapter III). By recursion theoretic methods (cf. exercise 13.13) it can be shown that Ω keeps the above closure property even in its recursive standard interpretation where Ω is interpreted as ω_1^{CK}, the first recursively regular ordinal. It is now obvious to ask if ω_1^{CK} already is the smallest ordinal above ω having this closure property. By purely recursion theoretic methods this question hardly is to answer. By proof theoretic methods, however, we will establish that there are in fact smaller such ordinals. The smallest one will be ω. But of course our real interest is the question if there are ordinals between ω and ω_1^{CK} having this closure property. If there exist such ordinals, then we already know that they have to be larger than ε_0. This follows from the proof theoretic analysis of $\mathbf{Z}_1$ where we noticed that for every ordinal $\alpha < \varepsilon_0$ there is a Π_1^1-sentence of norm α which is provable in $\mathbf{Z}_1$ and therefore provable with a derivation of length smaller than $\omega \cdot 2$ and finite cut rank. It is also easy to see that there also is a Π_1^1-sentence of norm ε_0 provable with a derivation of length smaller than ε_0 and cut rank ω. In order to tackle the problem we therefore need notations for a segment of the ordinals which is larger than ε_0. The following section will provide us with such a segment.

§ 17. Continuation of the theory of ordinals

17.1. **Definition**

Let $M \subset On$ and $f: M \to On$. We define

(i) $\mathrm{Fix}(f) := \{\xi : f(\xi) = \xi\}$ and $f' = \mathrm{ord}_{\mathrm{Fix}(f)}$

(ii) $M' := \mathrm{Fix}(\mathrm{ord}_M)$.

17.2. **Lemma**

Let $\varkappa > \omega$ be regular. If f is a ($\varkappa$-)normal function, then Fix(f) *is closed unbounded (in $\varkappa$). Hence* f' *again is a ($\varkappa$-)normal function.*

Proof

By 6.27 we only have to show that Fix(f) is closed unbounded (in $\varkappa$). So let $\alpha \in On$ $(\alpha < \varkappa)$. Define $\beta_0 := \alpha$, $\beta_{n+1} := f(\beta_n)$ and $\beta := \sup\{\beta_n : n < \omega\}$. For a $\varkappa$-normal function f we immediately obtain $\beta_n < \varkappa$ by induction on n. Since there is a 1-1 mapping from the set $\{\beta_n : n \in \omega\}$ onto $\omega < \varkappa$ we have $\beta < \varkappa$. If $\beta \in \{\beta_n : n \in \omega\}$, then there is a $k < \omega$ such that $\beta = \beta_k \le f(\beta_k) \le \beta$. Otherwise we have $\beta \in \mathrm{Lim}$ by 6.11. By hypothesis f is a normal function and therefore also continuous. Hence $f(\beta) = \sup\{f(\beta_n) : \varkappa < \omega\} = \sup\{\beta_{n+1} : n < \omega\} = \beta$. Hence $\alpha \le \beta \in \mathrm{Fix}(f)$ and Fix(f) is unbounded (in $\varkappa$). Now suppose that $U \subset \mathrm{Fix}(f)$ is bounded (in $\varkappa$). If $\sup U \in U$ we are done. Otherwise we have $\sup U \in \mathrm{Lim}$ which implies $f(\sup U) = \sup\{f(\xi) : \xi \in U\}$ $= \sup\{\xi : \xi \in U\} = \sup U$. Hence $\sup U \in \mathrm{Fix}(f)$ and Fix(f) is closed.

17.3. **Corollary**

If $M \subset On$ *is closed unbounded (in $\varkappa$), then* M' *is closed unbounded (in $\varkappa$) too.*

Proof

If M is closed unbounded (in $\varkappa$), then ord_M is a ($\varkappa$-)normal function. By 17.2. we then have that $M' = \mathrm{Fix}(\mathrm{ord}_M)$ is closed unbounded (in $\varkappa$).

17.4. **Lemma**

Let $\varkappa > \omega$ be regular. If $\emptyset \neq I \subset On$ *is bounded (in $\varkappa$) and* $\{C_\xi : \xi \in I\}$ *is a family of sets which are closed unbounded (in $\varkappa$), then* $C = \bigcap\{C_\xi : \xi \in I\}$ *is closed unbounded (in $\varkappa$) too.*

Proof

Since all C_ξ are ($\varkappa$-)closed we immediately obtain that C is ($\varkappa$-)closed too. The real problem is to show that C is unbounded (in $\varkappa$). Therefore choose $\alpha \in On(<\varkappa)$. We define a family $(f_n)_{n<\omega}$ of sequences $(f_{n,\xi})_{\xi \in I}$ by

$$f_{0,\xi} := \min\{\gamma \in C_\xi : \alpha \le \gamma \wedge (\forall \eta \in \xi \cap I) f_{0,\eta} < \gamma\}$$

and

$$f_{n+1,\xi} := \min\{\gamma \in C_\xi : (\forall \eta \in I) f_{n,\eta} < \gamma \wedge (\forall \eta \in \xi \cap I) f_{n+1,\eta} < \gamma\}.$$

Since all C_ξ are unbounded (in $\varkappa$) we conclude by induction on ξ that $f_{0,\xi}$ always is defined (and $<\varkappa$). By the induction hypothesis we may assume that $f_{n,\xi}$ is defined (and $<\varkappa$) for all n. Since I is bounded (in $\varkappa$) we have $\sup\{f_{n,\xi} : \xi \in I\} \in On(<\varkappa)$. By the unboundedness of C_ξ we then obtain that $f_{n+1,\xi}$ is defined for all $\xi \in I$ and $n \in \omega$. The sets $\{f_{n,\xi} : n \in \omega\}$ are bounded (in $\varkappa$) and it follows $\beta_\xi := \sup\{f_{n,\xi} : n<\omega\} \in C_\xi$ for all $\xi \in I$. By construction we have $\beta_\xi = \beta_{\xi_0}$ for all $\xi, \xi_0 \in I$. Hence $\alpha \le \beta_{\xi_0} \in C$ for arbitrary ξ_0.

17.5. **Definition**

We define the classes $Cr(\alpha)$ by recursion on α.

(i) $Cr(0) = \mathbb{H}$

(ii) $Cr(\alpha') = (Cr(\alpha))'$

(iii) $Cr(\lambda) = \bigcap\{Cr(\xi) : \xi<\lambda\}$ for $\lambda \in Lim$.

We call $Cr(\alpha)$ the class of α-*critical* ordinals. By φ_α we denote the enumerating function of $Cr(\alpha)$. We usually write $\varphi\alpha\beta$ instead of $\varphi_\alpha(\beta)$.

17.6. **Lemma**

$Cr(\alpha)$ *is closed unbounded in any regular* $\varkappa > \max\{\alpha,\omega\}$. *Hence* $\varphi_\alpha \restriction \varkappa$ *is a* $\varkappa$-*normal function for all* $\varkappa > \max\{\alpha,\omega\}$.

Proof

By 7.8. we have that $Cr(0)$ is closed unbounded in any $\varkappa > \omega$. Using 17.3. and 17.4. we easily obtain by induction on α that all $Cr(\alpha)$ are closed unbounded in any $\varkappa > \max\{\alpha,\omega\}$. Then it is immediate by 6.27 that $\varphi_\alpha \restriction \varkappa$ is a $\varkappa$-normal function.

17.7. **Lemma**

(i) $\varphi 0 \alpha = \omega^\alpha$

(ii) $\varphi 1 0 = \varepsilon_0$

(iii) $\beta<\gamma$ *implies* $\varphi\alpha\beta < \varphi\alpha\gamma$

(iv) $\beta \leq \varphi\alpha\beta$

(v) $\varphi\alpha\lambda = \sup\{\varphi\alpha\xi : \xi < \lambda\}$ *for* $\lambda \in \mathrm{Lim}$

(vi) If $\alpha < \beta$, *then* $\mathrm{Cr}(\beta) \subsetneq \mathrm{Cr}(\alpha)$, $\varphi\alpha(\varphi\beta\gamma) = \varphi\beta\gamma$ *and* $\varphi\alpha\gamma \leq \varphi\beta\gamma$.

Proof

(i) holds by definition.

(ii) We have $\varepsilon_0 = \min\{\xi : \varphi 0\xi = \xi\} = \min \mathrm{Cr}(1) = \varphi 10$.

(iii)–(v) follow from the fact that φ_α is a normal function for all α.

(vi) By induction on β we first obtain $\mathrm{Cr}(\beta) \subset \mathrm{Cr}(\alpha)$. If $\alpha < \beta$, then we have $\varphi\beta\gamma \in \mathrm{Cr}(\beta) \subset \mathrm{Cr}(\alpha+1) = \mathrm{Cr}(\alpha)'$. Hence $\varphi\alpha(\varphi\beta\gamma) = \varphi\beta\gamma$. By $\gamma \leq \varphi\beta\gamma$ we therefore obtain $\varphi\alpha\gamma \leq \varphi\alpha(\varphi\beta\gamma) = \varphi\beta\gamma$. Since $0 \notin \mathrm{Cr}(0) \supset \mathrm{Cr}(\beta)$ it follows $0 < \varphi\beta 0$ which implies $\varphi\alpha 0 < \varphi\alpha(\varphi\beta 0) = \varphi\beta 0$, i.e. $\varphi\alpha 0 \notin \mathrm{Cr}(\beta)$ and therefore $\mathrm{Cr}(\alpha) \neq \mathrm{Cr}(\beta)$.

17.8. **Theorem**

Suppose that $\alpha = \varphi\alpha_1\beta_1$ *and* $\beta = \varphi\alpha_2\beta_2$. *Then we have*

(1) $\alpha = \beta$ *if and only if one of the following conditions is satisfied:*

(i) $\alpha_1 < \alpha_2$ *and* $\beta_1 = \varphi\alpha_2\beta_2$

(ii) $\alpha_1 = \alpha_2$ *and* $\beta_1 = \beta_2$

(iii) $\alpha_2 < \alpha_1$ *and* $\varphi\alpha_1\beta_1 = \beta_2$,

(2) $\alpha < \beta$ *if and only if one of the following conditions is satisfied:*

(i) $\alpha_1 < \alpha_2$ *and* $\beta_1 < \varphi\alpha_2\beta_2$

(ii) $\alpha_1 = \alpha_2$ *and* $\beta_1 < \beta_2$

(iii) $\alpha_2 < \alpha_1$ *and* $\varphi\alpha_1\beta_1 < \beta_2$.

Proof

We simultaneously prove the claims (1) and (2).

(i) If $\alpha_1 < \alpha_2$, then we have $\varphi\alpha_1(\varphi\alpha_2\beta_2) = \varphi\alpha_2\beta_2$. Therefore $\varphi\alpha_1\beta_1 = \varphi\alpha_2\beta_2$ holds if and only if $\beta_1 = \beta$ and $\alpha < \beta$ holds if and only if $\beta_1 < \beta$.

(ii) If $\alpha_1 = \alpha_2$, then (1) and (2) are obvious since φ_α is order preserving.

(iii) is a consequence of (i) and (ii).

Theorem 17.8. is basic for the following parts. Therefore we often will use it without mentioning it explicitly.

17.9. **Lemma**

We have $\varphi\alpha 0 < \varphi\beta 0$ *if and only if* $\alpha < \beta$.

From 17.9. we obtain by 6.22. $\alpha \leq \varphi\alpha 0 \leq \varphi\alpha\beta$ for all α and β. Therefore we have the following lemma.

17.10. Lemma

For all $\alpha, \beta \in \mathrm{On}$ *it is* $\alpha, \beta \leq \varphi\alpha\beta$.

17.11. Theorem

For every ordinal $\alpha \in \mathbb{H}$ *there are uniquely determined ordinals* β *and* γ *such that* $\alpha = \varphi\beta\gamma$ *and* $\gamma < \alpha$.

Proof

The uniqueness of β and γ is an immediate consequence of 17.8. To show their existence we define $\beta := \min\{\xi : \alpha < \varphi\xi\alpha\}$. The ordinal β is defined because of $\alpha \leq \varphi\alpha 0 < \varphi\alpha\alpha$. If $\beta = 0$, then we have $\alpha < \varphi 0 \alpha$. Since $\alpha \in \mathbb{H}$ there is an η such that $\alpha = \varphi 0 \eta < \varphi 0 \alpha$, i.e. $\eta < \alpha$. If $\beta > 0$, then we have $\alpha = \varphi\eta\alpha$ for all $\eta < \beta$. Hence $\alpha \in \mathrm{Cr}(\eta)' = \mathrm{Cr}(\eta')$ for all $\eta < \beta$. This means $\alpha \in \bigcap\{\mathrm{Cr}(\eta') : \eta < \beta\} \subset \mathrm{Cr}(\beta)$. Therefore there is a γ such that $\alpha = \varphi\beta\gamma < \varphi\beta\alpha$ which implies $\gamma < \alpha$.

17.12. Remark

The ordinals in $\mathrm{Cr}(\alpha)$ have the following closure properties. First they are closed under ordinal addition and $\gamma \in \mathrm{Cr}(\alpha)$ and $\xi < \alpha, \eta < \gamma$ imply $\varphi\xi\eta < \gamma$. To prove the latter closure property we observe that if $\gamma \in \mathrm{Cr}(\alpha)$, then there is a γ_1 such that $\gamma = \varphi\alpha\gamma_1$ and we obtain $\varphi\xi\eta < \varphi\alpha\gamma_1 = \gamma$. This means that the ordinals in $\mathrm{Cr}(\alpha)$ are inaccessible for the functions in $\{\varphi_\xi : \xi < \alpha\}$. For this reason we are going to call them α-*critical* ordinals. Ordinals α which themselves are α-critical are even inaccessible for the 2-place function $\lambda\xi\eta.\varphi\xi\eta$. This motivates the following definition.

17.13. Definition

We call an ordinal α *strongly critical* if $\alpha \in \mathrm{Cr}(\alpha)$. By

$$\mathrm{SC} := \{\alpha \in \mathrm{On} : \alpha \in \mathrm{Cr}(\alpha)\}$$

we denote the class of strongly critical ordinals.

We define $\Gamma_\alpha := \mathrm{ord}_{\mathrm{SC}}(\alpha)$.

17.14. Lemma

(i) We have $\alpha \in \mathrm{SC}$ *if and only if it is* $\varphi\alpha 0 = \alpha$.

(ii) If $\alpha \in \mathrm{SC}$ *and* $\beta, \gamma < \alpha$, *then it follows* $\varphi\beta\gamma < \alpha$.

Proof

(i) $\alpha \in SC$ implies $\alpha \in Cr(\alpha)$.Therefore there is an η such that $\alpha = \varphi\alpha\eta$. Now we have $\alpha \le \varphi\alpha 0 \le \varphi\alpha\eta = \alpha$. Hence $\eta = 0$. If $\alpha = \varphi\alpha 0$, then $\alpha \in Cr(\alpha)$ and we obtain $\alpha \in SC$ by definition.

(ii) By (i) $\alpha \in SC$ implies $\alpha = \varphi\alpha 0$. But $\beta < \alpha$ and $\gamma < \alpha = \varphi\alpha 0$ entail $\varphi\beta\gamma < \varphi\alpha 0 = \alpha$ by 17.8.(2) i.

17.15. **Theorem**

The class SC is closed unbounded in every regular $\varkappa > \omega$. Hence Γ is a $\varkappa$-normal function for all regular $\varkappa > \omega$.

Proof

Pick any $\alpha < \varkappa$. Define $\beta_0 := \alpha + 1$ and $\beta_{k+1} := \varphi\beta_k 0$. Then we have $\beta_0 < \varkappa$ and by induction on k it follows $\beta_k < \varkappa$ for all $k < \omega < \varkappa$. Hence $\beta := \sup\{\beta_k : k < \omega\} < \varkappa$. For $\xi < \beta$ there is a $k < \omega$ such that $\xi < \beta_k \le \beta_m$ for all $m \ge k$. Hence $\varphi\xi\beta_{m+1} = \varphi\xi(\varphi\beta_m 0) = \varphi\beta_m 0 = \beta_{m+1}$ for all $m \ge k$ and we have $\beta \le \varphi\xi\beta = \sup\{\varphi\xi\beta_{m+1} : m \ge k\} = \sup\{\beta_{m+1} : m \ge k\} \le \beta$. This shows that $\beta \in \bigcap\{Cr(\xi)' : \xi < \beta\} \subset Cr(\beta)$, i.e. $\beta \in SC$. Because of $\alpha < \beta < \varkappa$ we have that SC is unbounded in $\varkappa$. To show that SC also is $\varkappa$-closed we assume that $U \subset SC$ is bounded in $\varkappa$. Define $\beta := \sup U$. If $\beta \in U$, then we are done. Otherwise we have $\beta \in Lim$ and for $\xi, \eta < \beta$ there is a $\gamma \in U$ such that $\xi, \eta < \gamma$. Hence $\varphi\xi\eta < \gamma < \beta$ by 17.14. This proves $\beta \le \varphi\xi\beta = \sup\{\varphi\xi\eta : \eta < \beta\} \le \beta$ and we have $\beta \in \bigcap\{Cr(\xi)' : \xi < \beta\} \subset Cr(\beta)$ which implies $\beta \in SC$.

17.16. **Theorem**

For every $\alpha \in \mathbb{H} \setminus SC$ there are uniquely determined ordinals $\beta, \gamma < \alpha$ such that $\alpha = \varphi\beta\gamma$.

Proof

By 17.11. we have $\alpha = \varphi\beta\gamma$ for $\beta \le \alpha$ and $\gamma < \alpha$ and β and γ are uniquely determined. If we assume $\beta = \alpha$, then we obtain $\alpha \in Cr(\alpha)$ which contradicts $\alpha \notin SC$. Hence $\beta < \alpha$.

17.17. **Definition**

We define $\alpha =_{NF} \varphi\beta\gamma :\Leftrightarrow \alpha = \varphi\beta\gamma \wedge \beta < \alpha \wedge \gamma < \alpha$.

17.18. **Definition**

For $\alpha \in On$ we define $\alpha^{\Gamma} := \min\{\gamma \in SC : \alpha < \gamma\}$.

17.19. **Inductive definition** of the set $PC(\alpha)$ (Predicative closure of α).

(i) $\alpha' \subset PC(\alpha)$,

(ii) If $\gamma,\delta \in PC(\alpha)$, then also $\gamma+\delta \in PC(\alpha)$,

(iii) If $\gamma,\delta \in PC(\alpha)$, then also $\varphi\gamma\delta \in PC(\alpha)$.

17.20. **Theorem**

$PC(\alpha) = \alpha^\Gamma$

Proof

We show $PC(\alpha) \subset \alpha^\Gamma$ by induction on the definition of $PC(\alpha)$.

(i) If $\eta < \alpha'$, then we have $\eta \le \alpha < \alpha^\Gamma$.

(ii) Here we obtain $\gamma,\delta < \alpha^\Gamma$ by the induction hypothesis . Since $SC \subset \mathbb{H}$ this implies $\gamma+\delta < \alpha^\Gamma$.

(iii) Again by the induction hypothesis we have $\gamma,\delta < \alpha^\Gamma$. Since $\alpha^\Gamma \in SC$ it follows $\varphi\gamma\delta < \alpha^\Gamma$ by 17.14.(ii).

To prove the opposite direction we show $\xi < \alpha^\Gamma \Rightarrow \xi \in PC(\alpha)$ by induction on ξ. If $\xi \le \alpha$, then we obtain $\xi \in PC(\alpha)$ by 17.19.(i). If $\alpha < \xi < \alpha^\Gamma$, then we have $\xi \notin SC$. If $\xi \notin \mathbb{H}$, then there are $\gamma,\delta < \xi$ such that $\xi = \gamma+\delta$. By the induction hypothesis we have $\gamma,\delta \in PC(\alpha)$, and obtain $\xi = \gamma+\delta \in PC(\alpha)$ by 17.19.(ii). If $\xi \in \mathbb{H}$, then by 17.16. there are $\gamma,\delta < \xi$ such that $\xi = \varphi\gamma\delta$ and we obtain $\xi \in PC(\alpha)$ by 17.19.(iii) and the induction hypothesis.

17.21. **Theorem**

If $\varkappa > \omega$ is regular, then we have $\varkappa \in SC$ and $\alpha^\Gamma < \varkappa$ for all $\alpha < \varkappa$.

Proof

If $\xi,\eta < \varkappa$, then by 17.15. there is a $\sigma \in SC \cap \varkappa$ such that $\xi,\eta < \sigma$. By 17.14 (ii) it follows $\varphi\xi\eta < \sigma < \varkappa$. Hence $\varkappa \le \varphi\xi\varkappa = \sup\{\varphi\xi\eta : \eta < \varkappa\} \le \varkappa$ and we obtain $\varkappa \in \bigcap\{Cr(\xi)' : \xi < \varkappa\} \subset Cr(\varkappa)$. Since by 17.15. $(\alpha,\varkappa) \cap SC \neq \emptyset$ it follows $\alpha^\Gamma < \varkappa$.

17.22. **Theorem**

If we define $\Delta_0(\alpha) = \alpha'$ and $\Delta_{n+1}(\alpha) = \varphi\Delta_n(\alpha)0$, then we have $\alpha^\Gamma = \sup\{\Delta_n(\alpha) : n < \omega\}$.

Proof

By induction on n it easily follows $\Delta_n(\alpha) < \Delta_{n+1}(\alpha)$ for all $n < \omega$. We prove that for all $n < \omega$ there is a $\xi \in PC(\alpha)$ such that $\Delta_n(\alpha) \le \xi$ again by induction on n. If $\alpha = 0$

then $\alpha' \le \varphi 00 \in PC(\alpha)$ by 17.19.(i), (iii). If $\alpha \neq 0$ then $\alpha' \le \alpha+\alpha \in PC(\alpha)$ by 17.19.(i),(ii). For the induction step we have a $\xi \in PC(\alpha)$ such that $\Delta_n(\alpha) \le \xi$. Hence $\Delta_{n+1}(\alpha) = \varphi\Delta_n(\alpha)0 \le \varphi\xi 0 \in PC(\alpha)$ by 17.19.(iii). This proves $\sup\{\Delta_n(\alpha) : n<\omega\} \le \alpha^{\Gamma}$.
To prove the opposite direction we show by induction on the definition of $\beta \in PC(\alpha)$ that there is an $n<\omega$ such that $\beta<\Delta_n(\alpha)$. In the case of (i) this holds for $n = 0$. In the case of (ii) we obtain the claim from the induction hypothesis and the fact that $\Delta_n \in \mathbb{H}$ holds for all $n>0$. In the case of (iii) it is $\beta = \varphi\gamma\delta$ and by the induction hypothesis we have $n_1,n_2<\omega$ such that $\gamma \le \Delta_{n_1}(\alpha), \delta \le \Delta_{n_2}(\alpha)$. We define $n := \max\{n_1,n_2\}$, and obtain $\beta \le \varphi\Delta_n(\alpha)\Delta_n(\alpha) < \varphi\Delta_{n+1}(\alpha)\, 0 = \Delta_{n+2}(\alpha)$.

The following lemma is a special case of 17.20.

17.23. **Lemma**
$\Gamma_0 = PC(0)$.

The ordinals in PC(0) are accessible from from 0 only using the functions + and φ. Therefore all ordinals in PC(0) are represented by terms built up from $\{0,+,\varphi\}$ according to the rules given in 17.19. These terms are called *ordinal terms*. For ordinal terms α in PC(0) we define the degree $\underline{G}\alpha$ as the number of symbols 0,+ and φ which occur in α. So $\underline{G}\alpha$ is the length of the word α in the formal language over the alphabet $\{0,+,\varphi\}$.
According to 17.23. every ordinal $\alpha<\Gamma_0$ is represented by an ordinal term in PC(0). This ordinal term, however, is in general not uniquely determined. In order to obtain an unique representation of an ordinal by an ordinal term we define the set $PC_{NF}(0)$ of terms in *normal form*. $PC_{NF}(0)$ is defined inductively by the clauses

(i) $0 \in PC_{NF}(0)$

(ii) If $\alpha =_{NF} \alpha_1+\ldots+\alpha_n$ and $\{\alpha_1,\ldots,\alpha_n\} \subset PC_{NF}(0)$, then also $\alpha \in PC_{NF}(0)$

(iii) If $\alpha =_{NF} \varphi\alpha_1\alpha_2$ and $\{\alpha_1,\alpha_2\} \subset PC_{NF}(0)$, then also $\alpha \in PC_{NF}(0)$.

It is obvious that $PC_{NF}(0) \subset PC(0)$ and using 7.9, 17.16. and 17.13. we also obtain $PC(0) \subset PC_{NF}(0)$. Therefore every ordinal $\alpha<\Gamma_0$ is represented by an ordinal term $\underline{\alpha}$ in $PC_{NF}(0)$ and it is easy to see that this ordinal term is uniquely determined.
As the *degree* $G\alpha$ *of an ordinal* $\alpha<\Gamma_0$ we define the degree $\underline{G\alpha}$
The set PC_{NF} (0) like the set E in §8 may now be arithmetized. E.g. by defining $\ulcorner 0 \urcorner := \langle 0 \rangle$, $\ulcorner \alpha+\beta \urcorner := \langle 1, \ulcorner\alpha\urcorner, \ulcorner\beta\urcorner \rangle$ and $\ulcorner \varphi\alpha\beta \urcorner := \langle 2, \ulcorner\alpha\urcorner, \ulcorner\beta\urcorner \rangle$. It is not to hard to see that the set $\ulcorner PC(0) \urcorner = \{\ulcorner\alpha\urcorner : \alpha \in PC_{NF}(0)\}$ is again primitive recursive and using 17.8. it follows that the relations defined by

$$\ulcorner\alpha\urcorner \equiv \ulcorner\beta\urcorner :\Leftrightarrow \alpha = \beta \text{ and}$$
$$\ulcorner\alpha\urcorner \prec \ulcorner\beta\urcorner :\Leftrightarrow \alpha < \beta$$

are primitive recursive too. As a corollary to this sketch we obtain the following theorem.

17.24. Theorem

$\Gamma_0 < \omega_1^{CK}$.

17.25. Exercises

1. Show that any cardinal $\varkappa > \omega$ is strongly critical.
2. Prove that for any $\alpha \in PC(0)$ such that $\alpha = \beta$ it holds $G\beta \leq \underline{G}\alpha$. (This shows that another possibility to define $G\beta$ is $G\beta := \min\{\underline{G}\alpha : \alpha \in PC(0) \text{ and } \alpha = \beta\}$).

§18. An upper bound for the autonomous ordinal of $\mathbf{Z}_\infty$

18.1. Definition

The infinitary system $\mathbf{Z}_\infty$ canonically induces the operator

$$\Sigma : \mathfrak{P}(On) \times \mathfrak{P}(On) \to \mathfrak{P}(On)$$

which is given by

$$\Sigma(M,S) = \{|F| : \exists\alpha \in M\, \exists\delta \in S\ (F \text{ is } \Pi_1^1\text{-sentence such that } rk(F) \in S \wedge \vdash_\delta^\alpha F)\}.$$

The norm $|F|$ of a Π_1^1-sentence F, however, depends a little bit on the syntactic definition of the system $\mathbf{Z}_\infty$. An alteration of the inference rules may alter the norm by some finite ordinal. In order to obtain an operator which is not so sensitive for such alterations we redefine the operator Σ by

$$\Sigma(M,S) = \{|F| : \exists\alpha \in M\, \exists\delta \in S\ (F \text{ is } \Pi_1^1\text{-sentence such that } rk(F) \in S \wedge \vdash_\delta^\alpha F)\} \cup \{\alpha+1 : \alpha \in M\}.$$

Since every ordinal already is a subset of On we may turn the operator Σ into an operator $\Delta : On \times On \to On$ by defining

$$\Delta(\alpha,\beta) := \bigcup \Sigma(\alpha,\beta).$$

By diagonalization we obtain an operator

$$\Gamma: On) \to On,$$

defined by

$$\Gamma(\alpha) = \Delta(\alpha,\alpha).$$

We call an ordinal α *closed under* $\mathbf{Z}_\infty$ if $\Gamma(\alpha) \subset \alpha$.

A short review of chapter I is then given by the statement

$$\Delta(\omega\cdot 2,\omega) = \varepsilon_0.$$

In chapter I we moreover have shown

$\Delta(\varepsilon_0,\omega) = \varepsilon_0$.

In the remarks preceeding chapter II we already mentioned that the ordinals $\aleph_1$ and ω_1^{CK} are closed under $\mathbf{Z}_\infty$. In a next step we will see that also the ordinal ω is closed under $\mathbf{Z}_\infty$.

18.2. **Theorem**

ω is closed under $\mathbf{Z}_\infty$.

Proof

From $n,m<\omega$ and $\vdash^n_m F$ we obtain $\vdash^{2_m(n)}_0 F$ by the first elimination theorem 12.3. Hence $|F| \le 2_m(n) < \omega$.

Since Γ is closed under successors, it follows by 18.2. that ω is the least ordinal which is closed under $\mathbf{Z}_\infty$. So the least fixed point of Γ is rather uninteresting (at least if one is interested in clarifying the role of nonfinitistic means in mathematics). However, 18.2. may be interpreted as a mathematical proof for the philosophical statement that is impossible to create an actual infinite domain out of finiteness. The creation of an infinite domain from finite objects was one of the attempts of logicism which tried to establish mathematics on the basis of pure logic. 18.2. shows that this attempt must fail.

The infinitary system $\mathbf{Z}_\infty$ provides a tool to create ordinals autonomously. Starting from a given ordinal μ we successively build the sets

$\Gamma(\mu)$, $\Gamma(\Gamma(\mu))$, $\Gamma^{(3)}(\mu)$, $\Gamma^{(4)}(\mu)$, ...

and finally reach a set

$\mathrm{Aut}(\mu) = \bigcup\{\Gamma^{(n)}(\mu) : n<\omega\}$.

Aut(μ) may be interpreted as the set of ordinals which are autonomously accessible from μ. To see this we have to notice that $\mathbf{Z}_\infty \vdash^{\alpha+n}_0 \mathrm{TI}(\alpha,X)$ holds for some $n<\omega$. Hence $|\mathrm{TI}(\alpha,X)| < \alpha^*$ where α^* denotes the first limit ordinal larger than α. From $\alpha \in \mathrm{Aut}(\mu)$ we obtain $\alpha^* \subset \mathrm{Aut}(\mu)$, since Aut$(\mu)$ is closed under successors. Hence $|\mathrm{TI}(\alpha,X)| \in \mathrm{Aut}(\mu)$ and therefore $|\mathrm{TI}(\alpha,X)| \in \Gamma^{(n)}(\mu)$ for some $n<\omega$. Therefore we obtain an ordinal $\alpha_+ \in \Gamma^{(n-1)}(\mu)$ such that $\vdash^{\alpha_+}_{\alpha_+} \mathrm{TI}(\alpha,X)$. So α is an ordinal whose wellfoundedness is provable from α_+. So we may say that α is accessible from α_+. For $\beta<\mu$ we define $\beta_+ := \beta$. Then we obtain a sequence $\{\alpha_k : k\in\omega\}$ by $\alpha_0 := \alpha$ and $\alpha_{k+1} := (\alpha_k)_+$ which eventually becomes stationary at some $\alpha_m<\mu$. Since accessibility is a transitive process α_0 is accessible from α_m. (This essentially is a sketch of the construction which will be carried out in

§ 19). So we have justified α by assuming α_m as being given. If $\mu = 0$, then we call α *autonomously justified*. By 18.2., however, we have seen that it is impossible to justify infinite ordinals autonomously. If we want to step into the transfinite we already need the existence of an actual infinite set. For the moment it will suffice to require the existence of the ordinal ω. This motivates the following definition.

18.3. Definition

The *transfinite autonomous segment* of the ordinals is the ordinal Aut(ω+1). We usually omit the adjective 'transfinite' and talk of *the* autonomous segment of the ordinals.

For the computation of Aut(ω+1) we have to generalize the first elimination theorem.

18.4. Second elimination theorem

$\vdash^{\alpha}_{\beta+\omega^\rho} \Delta$ *implies* $\vdash^{\varphi\rho\alpha}_{\beta} \Delta$.

Proof by main induction on ρ and side induction on α

If the last inference is not a cut of rank $\geq\beta$, then the claim immediately follows from the induction hypothesis. Therefore we may assume that the last inference is a cut of rank σ such that $\beta\leq\sigma<\beta+\omega^\rho$. For $\rho = 0$ we obtain $\vdash^{\omega^\alpha}_{\beta} \Delta$, i.e. $\vdash^{\varphi 0 \alpha}_{\beta} \Delta$, by the first elimination theorem. If $\rho \neq 0$, then by 15.1. there is a $\rho_0<\rho$ and an $n<\omega$ such that $\beta\leq\sigma<\beta+\omega^{\rho_0}\cdot n$. If the premises of the cut are $\vdash^{\alpha_1}_{\beta+\omega^\rho} \Delta, A$ and $\vdash^{\alpha_2}_{\beta+\omega^\rho} \Delta, \neg A$, then we obtain by the side induction hypothesis $\vdash^{\varphi\rho\alpha_1}_{\beta} \Delta, A$ and $\vdash^{\varphi\rho\alpha_2}_{\beta} \Delta, \neg A$. We have $rk(A) = \sigma<\beta+\omega^{\rho_0}\cdot n$ and obtain by a cut $\vdash^{\delta}_{\beta+\omega^{\rho_0}\cdot n} \Delta$ for $\delta := \max\{\varphi\rho\alpha_1, \varphi\rho\alpha_2\}+1$. Now define $(\varphi\rho_0)^0\delta = \delta$ and $(\varphi\rho_0)^{n+1}\delta = \varphi\rho_0((\varphi\rho_0)^n\delta)$. By n-fold application of the main induction hypothesis we obtain $\vdash^{(\varphi\rho_0)^n\delta}_{\beta} \Delta$ and show by induction on n that it is $(\varphi\rho_0)^n\delta<\varphi\rho\alpha$. For $n = 0$ we have $(\varphi\rho_0)^n\delta = \delta = \max\{\varphi\rho\alpha_1, \varphi\rho\alpha_2\}+1< \varphi\rho\alpha$ since $\alpha_i<\alpha$ and $\varphi\rho\alpha\in\mathbb{H}$. It is $(\varphi\rho_0)^{n+1}\delta = \varphi\rho_0(\varphi\rho_0^n\delta)< \varphi\rho\alpha$, because of $\rho_0<\rho$ and $(\varphi\rho_0^n)\delta<\varphi\rho\alpha$ which holds by the induction hypothesis. So we have $(\varphi\rho_0)^n\delta<\varphi\rho\alpha$ and obtain $\vdash^{\varphi\rho\alpha}_{\beta} \Delta$ from $\vdash^{(\varphi\rho_0)^n\delta}_{\beta} \Delta$.

18.5. Corollary

$\vdash^{\alpha}_{\rho} \Delta$ *implies* $\vdash^{\varphi\rho\alpha}_{0} \Delta$.

Proof

We have $\rho \le \omega^\rho$. So $\vdash^{\alpha}_{\rho} \Delta$ implies $\vdash^{\alpha}_{\omega^\rho} \Delta$ and by 18.4. it follows $\vdash^{\varphi\rho\alpha}_{0} \Delta$.

18.6. **Theorem**

$\mathrm{Aut}(\omega+1) < \Gamma_0$.

Proof

Assume $\alpha,\rho < \Gamma_0$ and $\vdash^{\alpha}_{\rho} F$ for a Π^1_1-sentence F. By 18.5. this implies $\vdash^{\varphi\rho\alpha}_{0} F$ which by 10.3. entails $|F| \le \varphi\rho\alpha$. Since $\Gamma_0 \in SC$ we have $\varphi\rho\alpha < \Gamma_0$ by 17.14. Hence $\Pi(\Gamma_0) < \Gamma_0$. Because of $\omega+1 < \Gamma_0$ we obtain $\mathrm{Aut}(\omega+1) < \Gamma_0$.

18.7. **Exercises**

We are going to examinate two variants $\tilde{\mathbf{Z}}_\infty$ and $\bar{\mathbf{Z}}_\infty$ of the infinitary system $\mathbf{Z}_\infty$. In these variants we are going to derive finite sequences of formulas in which a double occurence of a formula is not automatically cancelled. We identify finite formula-sequences Γ and Δ whenever Γ is a permutation of Δ. We define the following calculi:

(Ax) If Δ is an axiom, then $\bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Delta$ and $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Delta$ holds for all α and ρ.

($\bigwedge$) If $\bar{\mathbf{Z}}_\infty \vdash^{\alpha_i}_{\rho} \Gamma, F_i$ or $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha_i}_{\rho} \Gamma, F_i$ holds for all $i < n \le \omega$, then we have $\bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma, \bigwedge\{F_i : i<n\}$ or $\mathbf{Z}_\infty \vdash^{\alpha}_{\rho} \Gamma, \bigwedge\{F_i : i<n\}$ respectively for all $\alpha \ge \sup\{\alpha_i+1 : i<n\}$.

($\bigvee$) If $\bar{\mathbf{Z}}_\infty \vdash^{\alpha_0}_{\rho} \Gamma, F_i$ or $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha_0}_{\rho} \Gamma, F_i$ holds for some $i < n \le \omega$, then $\bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma, \bigvee\{F_i : i<n\}$ or $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma, \bigvee\{F_i : i<n\}$ respectively holds for all $\alpha > \alpha_0$

(cut) If $\bar{\mathbf{Z}}_\infty \vdash^{\alpha_0}_{\rho} \Gamma, A$ and $\bar{\mathbf{Z}}_\infty \vdash^{\alpha_1}_{\rho} \Gamma, \neg A$ and $rk(A) < \rho$, then we have $\bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma$ for all $\alpha > \max\{\alpha_0, \alpha_1\}$ and if $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha_0}_{\rho} \Gamma, A$ and $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha_1}_{\rho} \Delta, \neg A$ and $rk(A) < \rho$, then we have $\tilde{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma, \Delta$ for all $\alpha > \max\{\alpha_0, \alpha_1\}$

1. Show the following facts:

   (i) The weakening rule is a permitted inference of $\bar{\mathbf{Z}}_\infty$, i.e., $\bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Gamma, A, A$ implies the existence of ordinals β, η such that $\bar{\mathbf{Z}}_\infty \vdash^{\beta}_{\eta} \Gamma, A$.

   (ii) $\exists \alpha, \rho\ \bar{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Delta \Leftrightarrow \exists \alpha, \rho\ \mathbf{Z}_\infty \vdash^{\alpha}_{\rho} \{F : F \in \Delta\}$

   (iii) $\bar{\mathbf{Z}}_\infty$ does not allow cut elimination.

2. (i) Formulate the elimination lemma for $\tilde{\mathbf{Z}}_\infty$ and sketch the proof.

   (ii) Prove the tautology lemma for $\tilde{\mathbf{Z}}_\infty$.

   (iii) Formulate the first elimination theorem for $\tilde{\mathbf{Z}}_\infty$ and sketch its proof.

   (iv) Formulate the second elimination theorem for $\tilde{\mathbf{Z}}_\infty$ and sketch the proof.

   (v) Prove or disprove the sentence: $\exists \alpha, \rho\ \tilde{\mathbf{Z}}_\infty \vdash^{\alpha}_{\rho} \Delta \Leftrightarrow \exists \alpha, \rho\ \mathbf{Z}_\infty \vdash^{\alpha}_{\rho} \{F : F \in \Delta\}$.

   (vi) Is $\tilde{\mathbf{Z}}_\infty$ sententially complete?

Definitions:

1. In the sequel we assume that the $\mathscr{L}_2$-formulas are obtained from atomic and negated atomic formulas by the connectives $\wedge$ and $\vee$ and the quantifiers $\forall$ and $\exists$. Negation for formulas which are not atomic is then defined via the deMorgan laws analogously to the definition for the language $\mathscr{L}_\infty$.

2. We define $rk_1(F)$ for $\mathscr{L}_2$-formulas F inductively by:

   (i) $rk_1(P) := rk_1(\neg P) := 0$ for atomic formulas P

   (ii) $rk_1(A\wedge B) := rk_1(A\vee B) := \max\{rk_1(A), rk_1(B)\}+1$

   (iii) $rk_1(\forall xA) := rk_1(\forall XA) := rk_1(\exists xA) := rk_1(\exists XA) := rk_1(A)+1$

3. $rk_2(F)$ for $\mathscr{L}_2$-formulas F is defined by:

   (i) If F is an $\mathscr{L}_1$-formula, then $rk_2(F) = 0$

   (ii) If $F \equiv A \overset{\wedge}{\vee} B$ is not an $\mathscr{L}_1$-formula, then $rk_2(F) := \max\{rk_2(A), rk_2(B)\}+1$

   (iii) If $F \equiv QxA$ for $Q\in\{\forall,\exists\}$ is not an $\mathscr{L}_1$-formula, then $rk_2(F) := rk_2(A)+1$

   (iv) If $F \equiv QXA$ for $Q\in\{\forall,\exists\}$, then $rk_2(F) := rk_2(A)+1$

4. We extend the interpretation * of $\mathscr{L}_1$-formulas without free occurences of free number variables by adding the following clauses

$$(\forall XF)^* :\equiv \bigwedge\{F_X(\{x: A\})^* : A \text{ is a } \mathscr{L}_\infty(x)\text{-formula and } rk(A)<\omega\}$$

$$(\exists XF)^* :\equiv \bigvee\{F_X(\{x: A\})^* : A \text{ is a } \mathscr{L}_\infty(x)\text{-formula and } rk(A)<\omega\}$$

5. Finally we define the following abbreviations for finite sets of $\mathscr{L}_2$- or $\mathscr{L}_\infty$-formulas respectively:

$\Delta^* := \{F^*: F\in\Delta\}$, $\Delta_X(\{x:A\}) := \{F_X(\{x:A\}): F\in\Delta\}$

6. Inductive definition of $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta$ for finite sets Δ of $\mathscr{L}_2$-formulas without free number variables:

(Ax_1) If Δ^* is an axiom (for $\mathbf{Z}_\infty$), then $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta$ holds for all α and k.

(Ax_2) $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{n}\, \Delta, C, \neg C$ holds for all $\mathscr{L}_1$-formulas C, all formula-sets Δ and all α and $n<\omega$.

$(\bigwedge)$ If $\mathbf{ACA}$ $|\!\frac{\alpha_0}{k}\, \Delta, A$ and $\mathbf{ACA}_\infty \,|\!\frac{\alpha_1}{k}\, \Delta, B$, then $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, A\wedge B$ holds for all $\alpha>\max\{\alpha_0,\alpha_1\}$.

$(\bigvee)$ If $\mathbf{ACA}_\infty \,|\!\frac{\alpha_0}{k}\, \Delta, A$, then $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, A\vee B$ and $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, B\vee A$ hold for all $\alpha>\alpha_0$.

$(\forall_1)$ If $\mathbf{ACA}_\infty \,|\!\frac{\alpha_n}{k}\, \Delta, F_x(\underline{n})$ for all $n\in\mathbb{N}$, then $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, \forall xF$ holds for all $\alpha \geq \sup\{\alpha_n+1: n\in\mathbb{N}\}$.

$(\exists_1)$ If $\mathbf{ACA}_\infty \,|\!\frac{\alpha_0}{k}\, \Delta, F_x(\underline{n})$ for some $n\in\mathbb{N}$, then $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, \exists xF$ holds for all $\alpha>\alpha_0$

$(\forall_2)$ If $\mathbf{ACA}_\infty \,|\!\frac{\alpha_0}{k}\, \Delta, F$ and $X \notin FV_2(\Delta)$, then we have $\mathbf{ACA}_\infty \,|\!\frac{\alpha}{k}\, \Delta, \forall XF$ for all $\alpha>\alpha_0$.

($\exists_2$) If $\mathbf{ACA}_\infty \frac{|\alpha_0}{k} \Delta, F_X(A)$ for some $\mathscr{L}_1$-formula A such that $FV_1(A) \subset \{x\}$, then $\mathbf{ACA}_\infty \frac{|\alpha}{k} \Delta, \exists XF$ holds for all $\alpha > \alpha_0$.

(cut) If $\mathbf{ACA}_\infty \frac{|\alpha_0}{k} \Delta, A$, $\mathbf{ACA}_\infty \frac{|\alpha_1}{k} \Delta, \neg A$ and $rk_2(A) < k$, then $\mathbf{ACA}_\infty \frac{|\alpha}{k} \Delta$ holds for all $\alpha > \max\{\alpha_0, \alpha_1\}$.

7. By **ACA** we denote the theory with language $\mathscr{L}_2$ which is obtained from $\mathbf{ACA}_0$ by adjoining the unrestricted scheme of complete induction.

(i) Let F be an $\mathscr{L}_2$-formula and A an $\mathscr{L}_1$-formula. Show that $rk_2(F) = rk_2(F_X(A))$.

(ii) Prove that if $\mathbf{ACA} \vdash F$ and $FV_1(F) \subset \{x_1, \ldots, x_k\}$, then there are $r, s < \omega$ such that $\mathbf{ACA}_\infty \frac{|\omega+r}{s} F_{x_1, \ldots, x_k}(\underline{n}_1, \ldots, \underline{n}_k)$ for all $(n_1, \ldots, n_k) \in N^k$.

(iii) Show that $\mathbf{ACA}_\infty \frac{|\alpha}{k} \Delta, A$, $\mathbf{ACA}_\infty \frac{|\beta}{k} \Gamma, \neg A$ and $rk_2(A) = k \geq 1$ imply $\mathbf{ACA}_\infty \frac{|\alpha \# \beta}{k} \Delta$.

(iv) Prove that $k > 0$ and $\mathbf{ACA}_\infty \frac{|\alpha}{k+1} \Delta$ imply $\mathbf{ACA}_\infty \frac{|\omega^\alpha}{k} \Delta$.

(v) Show that $\mathbf{ACA}_\infty \frac{|\alpha}{k+1} \Delta$ implies $\mathbf{ACA}_\infty \frac{|\omega_k(\alpha)}{1} \Delta$.

(vi) Let Δ be a set of $\mathscr{L}_\infty$-formulas and A an $\mathscr{L}_\infty(x)$-formula. Show that $\mathbf{Z}_\infty \frac{|\alpha}{\rho} \Delta$ implies $\mathbf{Z}_\infty \frac{|2rk(A)+\alpha}{rk(A)+\rho} \Delta_X(A)$.

(vii) Let Δ be a set of $\mathscr{L}_2$ -formulas and $n := \max\{rk_2(F) : F \in \Delta\}$. Prove that $\mathbf{ACA}_\infty \frac{|\alpha}{1} \Delta$ implies $\mathbf{Z}_\infty \frac{|\omega(n+1)+\alpha}{\omega} \Delta^*$.

(viii) Show that for a Π^1_1-sentence F such that $\mathbf{ACA} \vdash F$ it is $|F| < \varphi 1 \varepsilon_0$.

§19. Autonomous ordinals of $\mathbf{Z}_\infty$

Our next aim is to show that Γ_0 in fact is the least ordinal which comprises ω and is closed under $\mathbf{Z}_\infty$. In order to do this we have to show that every ordinal less than Γ_0 is accessible from ω by an autonomous process similar to that which we have described after 18.2.

As sketched below 17.23. we have primitive recursive codes for all ordinals below Γ_0 such that the equality relation $\equiv$ and order relation $\prec$ between codes for ordinals become primitive recursive. Therefore we have no problems in handling ordinals below Γ_0 in $\mathscr{L}_\infty$. In order to simplify notation we are going to identify ordinals and their codes. As in §15 we denote codes for ordinals by lower case greek letters. The atomic formula $\alpha < \beta$ is only true if $\alpha \in \ulcorner PC(0) \urcorner$, $\beta \in \ulcorner PC(0) \urcorner$ and $\alpha \prec \beta$. $\forall \alpha F(\alpha)$ abbreviates the formula $\bigwedge\{F(\alpha) : \alpha \in \ulcorner PC(0) \urcorner\}$ and $\exists \alpha F(\alpha)$ the formula $\bigvee\{F(\alpha) : \alpha \in \ulcorner PC(0) \urcorner\}$.

It is also comfortable to use class terms S of the form $\{x : F\}$ in the language $\mathscr{L}_\infty$. Since we have no number variables in the language $\mathscr{L}_\infty$ the convention of 1.8. does not make sense. In order to obtain a reasonable definition we have

to extend the language $\mathcal{L}_\infty$ by additional number variables x,y,... to the language $\mathcal{L}_{\infty,\omega}$. Terms of $\mathcal{L}_{\infty,\omega}$ then are defined as the terms of $\mathcal{L}$. Starting with these terms one defines the formulas of $\mathcal{L}_{\infty,\omega}$ in the same way as the formulas of $\mathcal{L}_\infty$. In the definition, however, we have to pay attention to the fact that there always are only finitely many variables free in one formula. The rank rk(S) of a class term $S = \{x: F\}$ then is the rank of the formula F. Now if F is an $\mathcal{L}_{\infty,\omega}$-formula such that $FV_1(F) = \{x\}$ and S is the class term $\{x: F\}$, then for any $\mathcal{L}_\infty$-term t the formula $t \in S$ (which is an abbreviation for the formula $F_x(t)$) again is a wellformed $\mathcal{L}_\infty$-formula.
In detail we agree upon the following abbreviations and notations.

19.1. Definition

(i) For a class term S we denote by $\alpha \subset S$ the formula $\forall\xi(\neg\xi<\alpha \vee \xi \in S)$.

(ii) $\alpha \subset \beta$ denotes the formula $\forall\xi(\neg\xi<\alpha \vee \xi<\beta)$.

(iii) $\mathfrak{K}_\sigma := \{S\colon S$ is a class term such that $\mathrm{rk}(S)<\sigma\}$.

(iv) $\mathrm{Fund}_\sigma(\alpha)$ is the formula $\bigwedge\{\mathrm{Fund}(\alpha,S) : S \in \mathfrak{K}_\sigma\}$

19.2. Lemma (Equality lemma)
If $\vdash^{\alpha}_{\rho} F_x(s)$ *for* $2\cdot\mathrm{rk}\,F \le \alpha$ *and* $\mathrm{rk}\,F<\rho$, *then we obtain* $\vdash^{\alpha+1}_{\rho} \neg s=t, F_x(t)$.

Proof
We have

(1) $\vdash^{2\cdot\mathrm{rk}\,F}_{0} \neg s=t, \neg F_x(s), F_x(t)$,

because either it is $s^{\mathbb{N}} \neq t^{\mathbb{N}}$ and we have an axiom according to (Ax1) or it is $s^{\mathbb{N}} = t^{\mathbb{N}}$ and (1) follows from the tautology lemma by the structural rule. From (1) and the hypothesis $\vdash^{\alpha}_{\rho} F_x(s)$ the claim follows by a cut.

19.3. Lemma (Conjunction lemma)
If F *is a valid sentence such that* $\mathrm{rk}\,F \le \alpha$ *and* $\vdash^{\alpha}_{\rho} \Delta, A$, *then* $\vdash^{\alpha+1}_{\rho} \Delta, A \wedge F$.

Proof
By exercise 10.18 we have $\vdash^{\mathrm{rk}\,F}_{0} F$. This implies $\vdash^{\mathrm{rk}\,F}_{0} \Delta, F$ by the structural rule. Together with the hypothesis $\vdash^{\alpha}_{\rho} \Delta, A$ we obtain $\vdash^{\alpha+1}_{\rho} \Delta, A \wedge F$ by an $\bigwedge$-inference.

19.4. **Lemma** (Detachment lemma)

Suppose that we have $\vdash^{\alpha}_{\rho} \neg F, \Delta$ *and that* F *is a valid sentence such that* $\mathrm{rk}\,F \le \alpha$ *and* $\mathrm{rk}\,F < \rho$. *Then* $\vdash^{\alpha+1}_{\rho} \Delta$. *For* $\mathrm{rk}(F) = 0$ *we even obtain* $\vdash^{\alpha}_{\rho} \Delta$.

Proof

By exercise 10.18. we have $\vdash^{\mathrm{rk}\,F}_{0} F$ and the claim follows by the structural rule and a cut. For $\mathrm{rk}\,F = 0$ we obtain $\vdash^{\alpha}_{\rho} \Delta$ by the elimination lemma.

Hidden applications of the structural rule will no longer be mentioned. We freely will apply the cut rule and $\bigwedge$- rule in the form $\vdash \Delta, A$ and $\vdash \Gamma, \neg A \Rightarrow \vdash \Delta, \Gamma$ as well as $\vdash \Delta, A$ and $\vdash \Gamma, B \Rightarrow \vdash \Delta, \Gamma, A \wedge B$ respectively. Since applications of the structural rule do not increase the length of the derivation this cannot do any harm.

For the rest of the section λ always will denote a limit ordinal. By λ^* we denote the first limit ordinal larger than λ.

We define $SP(S) := \{\eta : \forall\xi(\xi \subset S \to \xi+\eta \subset S)\}$.

19.5. **Lemma**

For all $S \in \mathfrak{K}_\lambda$ *there are* $\alpha < \lambda$ *and* $\rho < \lambda$ *such that* $\vdash^{\alpha}_{\rho} \eta \notin SP(S), \eta \subset S$ *holds for all* η.

Proof

By the tautology lemma we have $\vdash^{\alpha_0}_{0} \eta \notin SP(S), \eta \in SP(S)$ for $\alpha_0 = 2 \cdot \mathrm{rk}(SP(S)) < \lambda$. Hence

(1) $\vdash^{\alpha_0}_{0} \eta \notin SP(S), \neg 0 \subset S, \eta \subset S$

by $\bigwedge$-inversion and $\bigvee$-exportation. On the other side we have $\vdash^{\rho}_{0} \neg\xi < 0, \xi \in S$ according to (Ax1). Using $\bigvee$-importation and an $\bigwedge$-inference this implies

(2) $\vdash^{3}_{0} 0 \subset S$.

From (1) and (2) we obtain the claim by a cut of cut rank $\rho = \mathrm{rk}\,S + 3 < \lambda$.

19.6. **Lemma**

For $S \in \mathfrak{K}_\lambda$ *there are* α *and* ρ *less than* λ *such that* $\vdash^{\alpha}_{\rho} \neg\mathrm{Prog}(SP(S)), \neg\eta \subset SP(S), \eta \subset S$ *holds for all* η.

Proof

By the tautology lemma we have

(1) $\vdash^{\alpha_0}_{0} \neg\mathrm{Prog}(SP(S)), \mathrm{Prog}(SP(S))$ for $\alpha_0 = 2 \cdot \mathrm{rk}(\mathrm{Prog}(SP(S))) < \lambda$.

Hence by $\bigwedge$-inversion and $\bigvee$-exportation

(2) $\vdash_0^{\alpha_0} \neg\text{Prog}(SP(S)), \neg\eta \subset SP(S), \eta \in SP(S)$.

From (2) and lemma 19.5. we obtain by a cut

(3) $\vdash_\rho^{\alpha_1} \neg\text{Prog}(SP(S)), \neg\eta \subset SP(S), \eta \subset S$

for some $\rho < \lambda$ and $\alpha_1 < \lambda$ (we may choose ρ as the rank of SP(S) augmented by some finite ordinal and α_1 as the successor of the maximum of α_0 and the ordinal from lemma 19.5.).

19.7. **Lemma**

For $S \in \mathfrak{K}_\lambda$ *there is an* $\alpha < \lambda$ *such that* $\vdash_0^{\alpha} \neg\xi \subset S, \neg\eta \in SP(S), \xi+\eta \subset S$ *holds for all* η *and* ξ.

Proof

The claim follows by $\bigwedge$-inversion and $\bigvee$-exportation immediately from the tautology lemma.

19.8. **Lemma**

For $S \in \mathfrak{K}_\lambda$ *there are* $\alpha, \rho < \lambda$ *such that* $\vdash_\rho^{\alpha} \neg\text{Prog}(S), \text{Prog}(SP(S))$.

Proof

From the tautology lemma we obtain by $\bigwedge$-inversion and $\bigvee$-exportation

(1) $\vdash_0^{\alpha_0} \neg\xi \subset S, \neg\zeta < \xi, \zeta \in S$ for $\alpha_0 = 2 \cdot rk(S) + 4 < \lambda$,

(2) $\vdash_0^{\alpha_1} \neg\eta \subset SP(S), \neg\eta_0 < \eta, \eta_0 \in SP(S)$ for $\alpha_1 = 2 \cdot rk(SP(S)) + 4 < \lambda$

and

(3) $\vdash_0^{\alpha_2} \neg\eta_0 \in SP(S), \neg\xi \subset S, \xi+\eta_0 \subset S$ for $\alpha_2 < \lambda$.

By a cut it follows from (2) and (3)

(4) $\vdash_{\rho_0}^{\alpha_3} \neg\eta \subset SP(S), \neg\xi \subset S, \neg\eta_0 < \eta, \xi+\eta_0 \subset S$

for $\alpha_3 = \max\{\alpha_1, \alpha_2\} + 1 < \lambda$ and $\rho_0 \in [rk(SP(S))+1, \lambda)$. Again by the tautology lemma $\bigwedge$-inversion and $\bigvee$-exportation we obtain

(5) $\vdash_0^{\alpha_4} \neg\text{Prog}(S), \neg\xi+\eta_0 \subset S, \xi+\eta_0 \in S$

for $\alpha_4 < \lambda$. By a cut it follows from (4) and (5)

(6) $\vdash_{\rho_1}^{\alpha_5} \neg\text{Prog}(S), \neg\eta \subset SP(S), \neg\xi \subset S, \neg\eta_0 < \eta, \xi+\eta_0 \in S$ for $\alpha_5, \rho_1 < \lambda$.

Furthermore we have

(7) $\vdash_0^{2} \neg\zeta < \xi+\eta, \zeta < \xi, \exists\eta_0 (\zeta = \xi+\eta_0 \wedge \eta_0 < \eta)$.

This can be seen in the following way. If $\zeta < \xi+\eta$ and $\xi \leq \zeta$, then there is some η_0 such that $\zeta = \xi+\eta_0$ and $\eta_0 < \eta$ are true sentences. Then we have $\vdash_0^{0} \neg\zeta < \xi+\eta, \zeta < \xi, \zeta = \xi+\eta_0$

and $\vdash^{0}_{0} \neg\zeta<\xi+\eta, \zeta<\xi, \eta_0<\eta$ by (Ax1) and obtain (7) by an $\bigvee$-and an $\bigwedge$-inference. From (6) and the equality lemma we obtain

$$\vdash^{\alpha_6}_{\rho_1} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \neg\xi\subset S, \neg\,\zeta=\xi+\eta_0, \neg\eta_0<\eta, \zeta\in S$$

for all $\zeta\in\mathbb{N}$ and therefore by $\bigvee$-importation and an $\bigwedge$-inference

(8) $\vdash^{\alpha_7}_{\rho_1} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \neg\xi\subset S, \neg\exists\eta_0(\zeta=\xi+\eta_0\wedge\eta_0<\eta), \zeta\in S.$

From (7) and (8) we obtain by a cut

(9) $\vdash^{\alpha_8}_{\rho} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \neg\xi\subset S, \neg\zeta<\xi+\eta, \zeta<\xi, \zeta\in S$

for some $\rho<\lambda$. Using an $\bigwedge$-inference (9) and (1) imply

(10) $\vdash^{\alpha_9}_{\rho} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \neg\xi\subset S, \neg\zeta<\xi+\eta, \zeta<\xi\wedge\neg\zeta<\xi, \zeta\in S.$

From (10) and the detachment lemma we obtain

(11) $\vdash^{\alpha_{10}}_{\rho} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \neg\xi\subset S, \neg\zeta<\xi+\eta, \zeta\in S$ for all $\xi, \zeta\in PC(0)$.

From (11) by $\bigvee$-importation, an $\bigwedge$-inference, again $\bigvee$-importation and an $\bigwedge$-inference it follows

(12) $\vdash^{\alpha_{11}}_{\rho} \neg\mathrm{Prog}(S), \neg\eta\subset SP(S), \eta\in SP(S).$

The claim follows from (12) by $\bigvee$-importation and an $\bigwedge$-inference.Since we have $\lambda\in\mathrm{Lim}$ by hypothesis it is obvious that we may choose $\alpha_{11}<\lambda$ and therefore also $\alpha<\lambda$. Likewise we obtain $\rho<\lambda$ since ρ may be chosen as the rank of S augmented by some finite ordinal.

19.9. **Lemma**

If $\lambda\le\sigma$, then for every $S\in\mathfrak{K}_\lambda$ there is an $\alpha<\lambda$ such that $\vdash^{\alpha}_{0} \neg\mathrm{Fund}_\sigma(\xi), \neg\mathrm{Prog}(S), \xi\subset S.$

Proof

By the tautology lemma and $\bigvee$-exportation we obtain

(1) $\vdash^{\alpha_1}_{0} \neg\mathrm{Tran}(\prec), \neg(\neg\mathrm{Prog}(S)\vee\xi\subset S), \neg\mathrm{Prog}(S), \xi\subset S,$

for $\alpha_1 = 2\cdot\mathrm{rk}(\neg\mathrm{Prog}(S)\vee\xi\subset S)<\lambda$. Since $\lambda\le\sigma$ we have $\mathfrak{K}_\lambda\subset\mathfrak{K}_\sigma$. By $\bigvee$-importation and an $\bigvee$-inference we therefore obtain

(2) $\vdash^{\alpha}_{0} \bigvee\{\neg(\mathrm{Tran}(\prec)\wedge(\neg\mathrm{Prog}(S)\vee\xi\subset S)) : S\in\mathfrak{K}_\sigma\}, \neg\mathrm{Prog}(S), \xi\subset S$ for some $\alpha<\lambda$

which already is the claim.

19.10. **Corollary**

For $\gamma\le\lambda\le\sigma$ we have $\vdash^{\lambda}_{0} \neg\mathrm{Fund}_\sigma(\eta), \mathrm{Fund}_\gamma(\eta).$

Proof

This follows from lemma 19.9 by $\bigvee$-importation, the conjunction lemma and an $\bigwedge$-inference.

19.11. Lemma

We have $\vdash^{\lambda}_{\lambda} \neg\mathrm{Fund}_\lambda(\xi), \neg\mathrm{Fund}_\lambda(\eta), \mathrm{Fund}_\lambda(\xi+\eta)$.

Proof

Assume $S \in \mathfrak{K}_\lambda$. Then we also have $SP(S) \in \mathfrak{K}_\lambda$. By lemma 19.9. it follows

(1) $\vdash^{\alpha_1}_{0} \neg\mathrm{Fund}_\lambda(\xi), \neg\mathrm{Prog}(S), \xi \subset S$

and

(2) $\vdash^{\alpha_2}_{0} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(SP(S)), \eta \subset SP(S)$

for $\alpha_1, \alpha_2 < \lambda$. By the tautology lemma, $\bigwedge$-inversion and $\bigvee$-exportation we obtain $\vdash^{\alpha_3}_{0} \neg\mathrm{Prog}(SP(S)), \neg\eta \subset SP(S), \eta \in SP(S)$ for some $\alpha_3 < \lambda$. This together with (2) implies

(3) $\vdash^{\alpha_4}_{\rho_0} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(SP(S)), \eta \in SP(S)$

for $\alpha_4 < \lambda$ where we may choose $\rho_0 := rk(\xi \subset S \wedge \eta \in SP(S))+1 < \lambda$. By a cut we obtain from (3) and 19.8.

(4) $\vdash^{\alpha_5}_{\rho} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \eta \in SP(S)$

where $\rho < \lambda$ has to be sufficiently large. By (1) and (4) it follows by an $\bigwedge$-inference

(5) $\vdash^{\alpha_6}_{\rho} \neg\mathrm{Fund}_\lambda(\xi), \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \xi \subset S \wedge \eta \in SP(S)$.

By a cut we obtain from (5) and 19.7.

(6) $\vdash^{\alpha_7}_{\rho} \neg\mathrm{Fund}_\lambda(\xi), \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \xi+\eta \subset S$.

Since $\mathrm{Tran}(\prec)$ is a valid sentence such that $rk(\mathrm{Tran}(\prec)) < \omega \le \lambda$ we obtain from (6) by $\bigvee$-importation and the conjunction lemma

(7) $\vdash^{\alpha_8}_{\rho} \neg\mathrm{Fund}_\lambda(\xi), \neg\mathrm{Fund}_\lambda(\eta), \mathrm{Tran}(\prec) \wedge (\neg\mathrm{Prog}(S) \vee \xi+\eta \subset S)$ for all $S \in \mathfrak{K}_\lambda$.

The ordinals α_8 and ρ depend upon the choice of the S. However, they always are less than λ.

The claim follows from (7) by an $\bigwedge$-inference.

19.12. Lemma

We have $\vdash^{\lambda}_{0} \neg\mathrm{Fund}_\lambda(\eta), \neg\xi < \eta, \mathrm{Fund}_\lambda(\xi)$.

Proof

For $S \in \mathfrak{K}_\lambda$ by lemma 19.9. there is an $\alpha_1 < \lambda$ such that

(1) $\vdash^{\alpha_1}_{0} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \eta \subset S$.

Hence by $\bigwedge$-inversion

(2) $\vdash^{\alpha_1}_{0} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \neg\zeta < \eta, \zeta \in S$.

Since $\xi \le \eta \wedge \zeta < \xi \rightarrow \zeta < \eta$ is a valid sentence we have

(3) $\vdash_0^{\alpha_2} \neg\xi<\eta, \neg\zeta<\xi, \zeta<\eta$

for $\alpha_2<\omega$. By (2) and (3) and the elimination lemma we obtain

(4) $\vdash_0^{\alpha_3} \neg\mathrm{Fund}_\lambda(\eta), \neg\mathrm{Prog}(S), \neg\xi<\eta, \neg\zeta<\xi, \zeta\in S$

where $\alpha_3 := \alpha_1 \# \alpha_2 < \lambda$ for all $S\in\mathfrak{K}_\lambda$. By $\bigvee$-importation, an $\bigwedge$-inference, another $\bigvee$-importation and the conjunction lemma it follows

(5) $\vdash_0^{\alpha_4} \neg\mathrm{Fund}_\lambda(\eta), \neg\xi<\eta, \mathrm{Tran}(\prec)\wedge(\neg\mathrm{Prog}(S)\vee \xi\subset S)$ for all $S\in\mathfrak{K}_\lambda$.

Here α_4 depends upon the choice of S but it always may be chosen to be less than λ. Using an $\bigwedge$-inference we obtain the claim.

19.13. **Definition**

For $\eta =_{NF} \eta_1+\ldots+\eta_n$ we define $h(\eta) := \eta_1$

19.14. **Lemma**

There are $\alpha, \rho < \lambda^*$ *such that* $\vdash_\rho^\alpha \neg\mathrm{Fund}_\lambda(h(\eta)), \mathrm{Fund}_\lambda(\eta)$.

Proof

Let $\eta =_{NF} \eta_1+\ldots+\eta_n$. Then for $k = 1,\ldots,n$ the sentence $\eta_k \le h(\eta)$ is valid. By 19.12. and the detachment lemma it follows

(1) $\vdash_0^\lambda \neg\mathrm{Fund}_\lambda(h(\eta)), \mathrm{Fund}_\lambda(\eta_k)$ for $k = 1,\ldots,n$.

From (1) and lemma 19.11. we obtain

(2) $\vdash_\rho^{\alpha_1} \neg\mathrm{Fund}_\lambda(h(\eta)), \mathrm{Fund}_\lambda(\eta_1+\eta_2)\wedge \mathrm{Fund}_\lambda(\eta_3)$.

for α_1 and ρ less than λ^*. By (2) and lemma 19.11. it follows

(3) $\vdash_\rho^{\alpha_3} \neg\mathrm{Fund}_\lambda(h(\eta)), \mathrm{Fund}_\lambda(\eta_1+\eta_2+\eta_3)$,

where α_3 is α_1 augmented by some finite ordinal. By iteration of this procedure we finally obtain the claim.

19.15. **Lemma**

We have $\vdash_0^7 \mathrm{Fund}_\lambda(0)$.

Proof

By (Ax1) we have $\vdash_0^0 \neg\mathrm{Prog}(S), \neg\xi<0, \xi\in S$ for all $S\in\mathfrak{K}_\lambda$. From this we obtain $\vdash_0^3 \neg\mathrm{Prog}(S), 0\subset S$ by $\bigvee$-Importation and an $\bigwedge$-inference. Using $\bigvee$-importation and the conjunction lemma we obtain $\vdash_0^6 \mathrm{Tran}(\prec)\wedge(\neg\mathrm{Prog}(S)\vee 0\subset S)$. The claim follows now immediately by an $\bigwedge$-inference.

To simplify the formulation of the next lemmata we agree upon the following abbreviations.

$A_\lambda(\sigma)$ is the formula $\forall\xi(\neg\xi<\sigma \vee \forall\eta(\neg\mathrm{Fund}_\lambda(\eta)\vee \mathrm{Fund}_\lambda(\varphi\xi\eta)))$

$B_\lambda(\sigma,\tau)$ is the formula $\forall\eta(\neg\eta<\tau\vee \mathrm{Fund}_\lambda(\varphi\sigma\eta))$

$C_\lambda(\mu,\sigma,\tau)$ is the formula $\forall\mu_0(\neg G\mu_0<G\mu\vee \neg\mu_0<\varphi\sigma\tau\vee \mathrm{Fund}_\lambda(\mu_0))$.

19.16. Lemma

For every ordinal μ and limit ordinal λ there are ordinals $\alpha<\lambda^*$ and $\rho<\lambda^*$ such that we have

$$\vdash^{\alpha}_{\rho} \neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg C_\lambda(\mu,\sigma,\tau),\neg \mathrm{Fund}_\lambda(\tau),\neg\mu<\varphi\sigma\tau,\mathrm{Fund}_\lambda(\mu).$$

Proof

If $\varphi\sigma\tau\leq\mu$, then we have

(A) $\vdash^{0}_{0}\neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg C_\lambda(\mu,\sigma,\tau),\neg \mathrm{Fund}_\lambda(\tau),\neg\mu<\varphi\sigma\tau,\mathrm{Fund}_\lambda(\mu)$.

according to (Ax1). We therefore assume $\mu<\varphi\sigma\tau$. If $\mu = 0$, then we obtain by lemma 19.15 and the structural rule

(B) $\vdash^{\omega}_{0}\neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg C_\lambda(\mu,\sigma,\tau),\neg \mathrm{Fund}_\lambda(\tau),\mathrm{Fund}_\lambda(\mu)$..

If $\mu \neq 0$, then we distinguish the following cases.

1. $\mu = \varphi\mu_1\mu_2$ for $\mu_1<\sigma$ and $\mu_2<\varphi\sigma\tau$.

From the tautology lemma we obtain by $\bigwedge$-inversions and $\bigvee$-exportations

(1) $\vdash^{\alpha_1}_{0} \neg A_\lambda(\sigma),\neg\mu_1<\sigma,\neg \mathrm{Fund}_\lambda(\mu_2),\mathrm{Fund}_\lambda(\varphi\mu_1\mu_2)$

for $\alpha_1 = 2\cdot\mathrm{rk}(A_\lambda(\sigma))<\lambda^*$. By (1), the equality- and the detachment lemma it follows

(2) $\vdash^{\alpha_1}_{\rho} \neg A_\lambda(\sigma),\neg \mathrm{Fund}_\lambda(\mu_2),\mathrm{Fund}_\lambda(\mu)$

where we may choose $\rho = \mathrm{rk}(\mathrm{Fund}_\lambda(\mu))+1<\lambda^*$. From the tautology lemma we obtain by $\bigwedge$-inversions and $\bigvee$-exportations

(3) $\vdash^{\alpha_2}_{0} \neg C_\lambda(\mu,\sigma,\tau),\neg G\mu_2<G\mu,\ \neg\mu_2<\varphi\sigma\tau,\mathrm{Fund}_\lambda(\mu_2)$

for $\alpha_2 = 2\cdot\mathrm{rk}(C_\lambda(\mu,\sigma,\tau))<\lambda^*$. Since $G\mu_2<G\mu$ and $\mu_2<\varphi\sigma\tau$ are valid sentences we obtain by the detachment lemma from (3)

(4) $\vdash^{\alpha_2}_{0} \neg C_\lambda(\mu,\sigma,\tau),\mathrm{Fund}_\lambda(\mu_2)$.

By (2) and (4) it follows by cut

(C) $\vdash^{\alpha_3}_{\rho} \neg A_\lambda(\sigma),\neg C_\lambda(\mu,\sigma,\tau),\mathrm{Fund}_\lambda(\mu)$

where again we may choose α_3 and ρ to be less than λ^*.

2. Assume $\mu = \varphi\mu_1\mu_2$ such that $\mu_1 = \sigma$ and $\mu_2<\tau$.

From the tautology lemma we obtain by $\bigwedge$-inversions and $\bigvee$-exportations

(1) $\vdash^{\alpha_1}_{0} \neg B_\lambda(\sigma,\tau),\neg\mu_2<\tau,\mathrm{Fund}_\lambda(\varphi\sigma\mu_2)$

where $\alpha_1 = 2\cdot\text{rk } B_\lambda(\sigma,\tau) < \lambda^*$. By the detachment- and the equality lemma we obtain from (1)

(D) $\frac{|\alpha_1+1}{\rho} \neg B_\lambda(\sigma,\tau), \text{Fund}_\lambda(\mu)$

for $\rho := \text{rk}(\text{Fund}_\lambda(\mu))+1 < \lambda^*$.

3. $\mu = \varphi\mu_1\mu_2$ and $\mu < \tau$.

By lemma 19.12. we have

(1) $\frac{|\lambda}{0} \neg\text{Fund}_\lambda(\tau), \neg\mu < \tau, \text{Fund}_\lambda(\mu)$.

By the detachment lemma we obtain from (1)

(E) $\frac{|\lambda}{0} \neg\text{Fund}_\lambda(\tau), \text{Fund}_\lambda(\mu)$.

4. $\mu = {}_{NF}\mu_1+\ldots+\mu_n$ for $n>1$.

Then we have $G\mu_1 < G\mu$, $\mu_1 < \varphi\sigma\tau$ and $h(\mu) = \mu_1$.

From the tautology lemma we obtain by $\wedge$-inversions and $\vee$-exportations

(1) $\frac{|\alpha_1}{0} \neg C_\lambda(\mu,\sigma,\tau), \neg G\,h(\mu) < G\mu, \neg h(\mu) < \varphi\sigma\tau, \text{Fund}_\lambda(h(\mu))$

for $\alpha_1 = 2\cdot\text{rk}(C_\lambda(\mu,\sigma,\tau)) < \lambda^*$. Using the detachment lemma we obtain from (1)

(2) $\frac{|\alpha_1}{0} \neg C_\lambda(\mu,\sigma,\tau), \text{Fund}_\lambda(h(\mu))$.

By (2) and lemma 19.14. we obtain

(F) $\frac{|\alpha}{\rho} \neg C_\lambda(\mu,\sigma,\tau), \text{Fund}_\lambda(\mu)$,

where α and ρ may be chosen below λ^*.

From (A) - (F), however, we obtain the claim.

19.17. **Lemma**

For every μ there is an $\alpha_\mu < \lambda^*$ and a $\rho < \lambda^*$ such that

$$\frac{|\alpha_\mu}{\rho} \neg A_\lambda(\sigma), \neg B_\lambda(\sigma,\tau), \neg\text{Fund}_\lambda(\tau), \neg\mu < \varphi\sigma\tau, \text{Fund}_\lambda(\mu).$$

Proof

We prove the lemma by metainduction on $G\mu$. If $G\mu = 0$, then we have $\mu = 0$ and the claim is a structural consequence of lemma 19.15. Therefore we assume $G\mu \neq 0$. Then we have

(1) $\frac{|\alpha_{\mu_0}}{\rho} \neg A_\lambda(\sigma), \neg B_\lambda(\sigma,\tau), \neg\text{Fund}_\lambda(\tau), \neg G\mu_0 < G\mu, \neg\mu_0 < \varphi\sigma\tau, \text{Fund}_\lambda(\mu_0)$

for some $\alpha_{\mu_0}, \rho_0 < \lambda^*$, since it either is $G\mu \leq G\mu_0$ and we have (1) according to (Ax1) or it is $G\mu_0 < G\mu$ and (1) holds according to the induction hypothesis. From (1) we obtain by $\vee$-importation and an $\wedge$-inference

(2) $\frac{|\alpha_{\mu_1}+4}{\rho} \neg A_\lambda(\sigma), \neg B_\lambda(\sigma,\tau), \neg\text{Fund}_\lambda(\tau), C_\lambda(\mu,\sigma,\tau)$

where $\alpha_{\mu_1} := \max\{\alpha_{\mu_0} : G\mu_0 < G\mu\}$. From (2) and lemma 19.16. by a cut we get

(3) $\frac{|\alpha_\mu}{\rho} \neg A_\lambda(\sigma), \neg B_\lambda(\sigma,\tau), \neg\text{Fund}_\lambda(\tau), \neg\mu < \varphi\sigma\tau, \text{Fund}_\lambda(\mu)$

for $\alpha_\mu := \max\{\alpha_1, \alpha_{\mu_0}+4\}+1$ and $\rho := \max\{\rho_0, \rho_1, \mathrm{rk}(\mathrm{Fund}_\lambda(\tau))\}+1$ where α_1 and ρ_1 are the ordinals from lemma 19.16.

We define $SP_\lambda(\sigma) := \{\xi\colon \neg\mathrm{Fund}_\lambda(\xi) \vee \mathrm{Fund}_\lambda(\varphi\sigma\xi)\}$. Then we have $SP_\lambda(\sigma) \in \mathfrak{K}_{\lambda^*}$.

19.18. **Lemma**

There are ordinals α *and* ρ *less than* λ^* *such that*

$$\vdash^{\alpha}_{\rho} \neg\mathrm{Fund}_\lambda(\tau), \neg\tau \subset SP_\lambda(\sigma), \forall\eta(\neg\eta<\tau \vee \mathrm{Fund}_\lambda(\varphi\sigma\eta)).$$

Proof

From the tautology lemma we obtain by $\bigvee$-exportation

(1) $\vdash^{\alpha_1}_{0} \neg(\neg\mathrm{Fund}_\lambda(\eta) \vee \mathrm{Fund}_\lambda(\varphi\sigma\eta)), \neg\mathrm{Fund}_\lambda(\eta), \mathrm{Fund}_\lambda(\varphi\sigma\eta)$,

i.e.

(2) $\vdash^{\alpha_1}_{0} \neg\eta \in SP_\lambda(\sigma), \neg\mathrm{Fund}_\lambda(\eta), \mathrm{Fund}_\lambda(\varphi\sigma\eta)$

for $\alpha_1 := 2\cdot\mathrm{rk}(\neg\mathrm{Fund}_\lambda(\eta) \vee \mathrm{Fund}_\lambda(\varphi\sigma\eta)) < \lambda^*$. By 19.12. furthermore we have

(3) $\vdash^{\lambda}_{0} \neg\mathrm{Fund}_\lambda(\tau), \neg\eta<\tau, \mathrm{Fund}_\lambda(\eta)$.

From (2) and (3) we obtain by a cut

(4) $\vdash^{\alpha_2}_{\rho} \neg\eta \in SP_\lambda(\sigma), \neg\mathrm{Fund}_\lambda(\tau), \neg\eta<\tau, \mathrm{Fund}_\lambda(\varphi\sigma\eta)$

for α_2 and ρ less than λ^*. Furthermore we have

(5) $\vdash^{0}_{0} \eta<\tau, \neg\mathrm{Fund}_\lambda(\tau), \neg\eta<\tau, \mathrm{Fund}_\lambda(\varphi\sigma\eta)$

according to (Ax1). By (4) and (5) and an $\bigwedge$-inference we obtain

(6) $\vdash^{\alpha_2+1}_{\rho} \eta<\tau \wedge \neg\eta \in SP_\lambda(\sigma), \neg\mathrm{Fund}_\lambda(\tau), \neg\eta<\tau, \mathrm{Fund}_\lambda(\varphi\sigma\eta)$

which implies by an $\bigvee$-inference, $\bigvee$-importation and an $\bigwedge$-inference

(7) $\vdash^{\alpha_2+5}_{\rho} \neg\tau \subset SP_\lambda(\sigma), \neg\mathrm{Fund}_\lambda(\tau), \forall\eta(\neg\eta<\tau \vee \mathrm{Fund}_\lambda(\varphi\sigma\eta))$.

19.19. **Lemma**

We have $\vdash^{\lambda}_{\lambda} \neg\forall\mu(\neg\mu<\sigma \vee \mathrm{Fund}_\lambda(\mu)), \mathrm{Fund}_\lambda(\sigma)$.

Proof

By 19.9. for every $S \in \mathfrak{K}_\lambda$ there is an $\alpha_1<\lambda$ such that

(1) $\vdash^{\alpha_1}_{0} \neg\mathrm{Fund}_\lambda(\mu), \neg\mathrm{Prog}(S), \mu \subset S$.

Since we have $\vdash^{0}_{0} \mu<\sigma, \neg\mu<\sigma$ and $\vdash^{\alpha_2}_{0} \neg\mathrm{Prog}(S), \neg\mu \subset S, \mu \in S$ for some $\alpha_2<\lambda$ and we obtain from (1) by a cut and an $\bigwedge$-inference

(2) $\vdash^{\alpha_3}_{\rho} \mu<\sigma \wedge \neg\mathrm{Fund}_\lambda(\mu), \neg\mathrm{Prog}(S), \neg\mu<\sigma, \mu \in S$

for α_3 and ρ below λ. By an $\bigvee$-inference it follows

(3) $\frac{\alpha_4}{\rho}\ \neg\forall\mu(\neg\mu<\sigma\vee \mathrm{Fund}_\lambda(\mu)),\ \neg\mathrm{Prog}(S),\neg\mu<\sigma,\mu\in S.$

By an $\bigwedge$-inference, $\bigvee$-importations and the conjunction lemma it finally follows

(4) $\frac{\alpha_5}{\rho}\ \neg\forall\mu(\neg\mu<\sigma\vee \mathrm{Fund}_\lambda(\mu)),\ \mathrm{Tran}(\prec)\wedge(\neg\mathrm{Prog}(S)\vee\sigma\subset S).$

Here α_5 and ρ depend on the choice of S but always may be chosen below λ. Using an $\bigwedge$-inference therefore we obtain

(5) $\frac{\lambda}{\lambda}\ \neg\forall\mu(\neg\mu<\sigma\vee \mathrm{Fund}_\lambda(\mu)),\ \bigwedge\{\mathrm{Tran}(\prec)\wedge(\neg\mathrm{Prog}(S)\vee\sigma\subset S): S\in\mathfrak{K}_\lambda\}$

which is the claim.

19.20. **Lemma**

There are ordinals $\alpha<\lambda^{**}$ *and* $\rho<\lambda^*$ *such that* $\frac{\alpha}{\rho}\ \neg A_\lambda(\sigma),\neg\mathrm{Fund}_{\lambda^*}(\tau),\mathrm{Fund}_\lambda(\varphi\sigma\tau)$.

Proof

By 19.17. there are α_μ and ρ below λ^* such that

$$\frac{\alpha_\mu}{\rho}\ \neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg\mathrm{Fund}_\lambda(\tau),\neg\mu<\varphi\sigma\tau,\mathrm{Fund}_\lambda(\mu)$$

for all μ. By $\bigvee$-importation and an $\bigwedge$-inference this implies

(1) $\frac{\lambda^*}{\rho}\ \neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg\mathrm{Fund}_\lambda(\tau),\forall\mu(\neg\mu<\varphi\sigma\tau\vee \mathrm{Fund}_\lambda(\mu)).$

By (1) and lemma 19.19. we obtain

(2) $\frac{\lambda^*+1}{\rho}\ \neg A_\lambda(\sigma),\neg B_\lambda(\sigma,\tau),\neg\mathrm{Fund}_\lambda(\tau),\mathrm{Fund}_\lambda(\varphi\sigma\tau).$

By lemma 19.18. we have

(3) $\frac{\alpha_1}{\rho}\ \neg\mathrm{Fund}_\lambda(\tau),\neg\tau\subset SP_\lambda(\sigma),B_\lambda(\sigma,\tau),$

for $\alpha_1<\lambda^*$. From (2) and (3) we obtain by cut

(4) $\frac{\lambda^*+2}{\rho}\ \neg A_\lambda(\sigma),\neg\tau\subset SP_\lambda(\sigma),\ \neg\mathrm{Fund}_\lambda(\tau),\mathrm{Fund}_\lambda(\varphi\sigma\tau)$

where we tacitly assume that $\rho<\lambda^*$ is so large that it majorizes the rank of $B_\lambda(\sigma,\tau)$ (and of all cut formulas coming). By (4) we obtain by $\bigvee$-importation

(5) $\frac{\lambda^*+4}{\rho}\ \neg A_\lambda(\sigma),\neg\tau\subset SP_\lambda(\sigma),\tau\in SP_\lambda(\sigma).$

Again by $\bigvee$-importation and an $\bigwedge$-inference we obtain from (5)

(6) $\frac{\lambda^*+7}{\rho}\ \neg A_\lambda(\sigma),\mathrm{Prog}(SP_\lambda(\sigma)).$

By the tautology lemma we have

(7) $\frac{\beta}{0}\ \neg A_\lambda(\sigma),\neg\tau\subset SP_\lambda(\sigma),\tau\subset SP_\lambda(\sigma)$

for $\beta=2\cdot\mathrm{rk}(\tau\subset SP_\lambda(\sigma))<\lambda^*$. From (6) and (7) we obtain by an $\bigwedge$-inference $\frac{\lambda^*+8}{\rho}\ \neg A_\lambda(\sigma),\mathrm{Prog}(SP_\lambda(\sigma))\wedge\neg\tau\subset SP_\lambda(\sigma),\tau\subset SP_\lambda(\sigma)$, from which we obtain

(8) $\frac{\alpha_2}{\rho}\ \neg A_\lambda(\sigma),\neg(\mathrm{Tran}(\prec)\wedge\neg\mathrm{Prog}(SP_\lambda(\sigma))\vee\tau\subset SP_\lambda(\sigma)),\tau\subset SP_\lambda(\sigma)$

by the structural rule and $\bigvee$-importation. Because of $SP_\lambda(\sigma)\in\mathfrak{K}_{\lambda^*}$ we obtain from (8) by an $\bigvee$-inference

(9) $\frac{\alpha_3}{\rho}\ \neg A_\lambda(\sigma),\neg\mathrm{Fund}_{\lambda^*}(\tau),\tau\subset SP_\lambda(\sigma).$

From the tautology lemma we obtain by $\bigwedge$-inversion and $\bigvee$-exportation

(10) $\vdash^{\beta_1}_{0}$ $\neg\mathrm{Prog}(SP_\lambda(\sigma)), \neg\tau \subset SP_\lambda(\sigma), \tau \in SP_\lambda(\sigma)$

for $\beta_1 := 2\cdot\mathrm{rk}\,\mathrm{Prog}(SP_\lambda(\sigma))$. From (6) and (10) it follows by cut

(11) $\vdash^{\alpha_4}_{\rho}$ $\neg A_\lambda(\sigma), \neg\tau \subset SP_\lambda(\sigma), \tau \in SP_\lambda(\sigma)$

and from (9) and (11)

(12) $\vdash^{\alpha_5}_{\rho}$ $\neg A_\lambda(\sigma), \neg\mathrm{Fund}_{\lambda^*}(\tau), \tau \in SP_\lambda(\sigma)$.

This implies by $\bigvee$-exportation

(13) $\vdash^{\alpha_6}_{\rho}$ $\neg A_\lambda(\sigma), \neg\mathrm{Fund}_{\lambda^*}(\tau), \neg\mathrm{Fund}_\lambda(\tau), \mathrm{Fund}_\lambda(\varphi\sigma\tau)$.

From (13) and 19.10. it follows

(14) $\vdash^{\alpha}_{\rho}$ $\neg A_\lambda(\sigma), \neg\mathrm{Fund}_{\lambda^*}(\tau), \mathrm{Fund}_\lambda(\varphi\sigma\tau)$

for $\alpha < \lambda^{**}$ and $\rho < \lambda^*$.

19.21. **Lemma**

Suppose that λ is a limit ordinal such that the set $\{\mu_\sigma : \sigma < \eta\}$ is unbounded in λ and we have $\vdash^{\alpha_\sigma}_{\rho_\sigma}$ $\Delta, \mathrm{Fund}_{\mu_\sigma}(\tau)$ for all $\sigma < \eta$ for $\alpha_\sigma < \beta$ and $\rho_\sigma \le \rho$. Then it follows $\vdash^{\beta}_{\rho}$ $\Delta, \mathrm{Fund}_\lambda(\tau)$.

Proof

Since $\{\mu_\sigma : \sigma < \eta\}$ is unbounded in λ we have for every $S \in \mathfrak{K}_\lambda$ a $\sigma < \eta$ such that $S \in \mathfrak{K}_{\mu_\sigma}$. By $\bigwedge$-inversion we obtain $\vdash^{\alpha_\sigma}_{\rho}$ $\Delta, \mathrm{Tran}(\prec) \wedge (\neg\mathrm{Prog}(S) \vee \tau \subset S)$ from the hypothesis. Since it is $\alpha_\sigma < \beta$ for all $\sigma < \eta$ we obtain the claim by an $\bigwedge$-inference.

19.22. **Lemma**

If $\eta \in \mathrm{Lim}$, then we have $\vdash^{\omega\cdot\eta}_{\omega\cdot\eta}$ $\neg\mathrm{Fund}_{\omega\cdot\eta}(\tau), \mathrm{Fund}_{\omega\cdot\eta}(\varphi 0\tau)$.

Proof

$A_\lambda(0)$ is the formula $\forall\xi(\neg\xi < 0 \vee \forall\eta(\neg\mathrm{Fund}_\lambda(\eta) \vee \mathrm{Fund}_\lambda(\varphi\xi\eta)))$. Therefore we obtain from an axiom (Ax1) by an $\bigvee$-inference and an $\bigwedge$-inference

(1) $\vdash^{2}_{0}$ $A_\lambda(0)$.

For limit ordinals λ we obtain from (1) and lemma 19.20. by a cut

(2) $\vdash^{\alpha_1}_{\rho}$ $\neg\mathrm{Fund}_{\lambda^*}(\tau), \mathrm{Fund}_\lambda(\varphi 0\tau)$,

for $\alpha_1 < \lambda^{**}$ and $\rho < \lambda^*$. By corollary 19.10. it follows

(3) $\vdash^{\lambda^*}_{\lambda^*}$ $\neg\mathrm{Fund}_{\omega\cdot\eta}(\tau), \mathrm{Fund}_{\lambda^*}(\tau)$

for all limit ordinals $\lambda^* \le \omega\cdot\eta$. By (2) and (3) for $\lambda = \omega\cdot\xi$ we obtain by a cut

(4) $\vdash^{\alpha_\xi}_{\omega\cdot(\xi+1)}$ $\neg\mathrm{Fund}_{\omega\cdot\eta}(\tau), \mathrm{Fund}_{\omega\cdot\xi}(\varphi 0\tau)$

for all $\xi < \eta$ such that $\alpha_\xi < \omega\cdot(\xi+2) < \omega\cdot\eta$. Because of $\eta \in \mathrm{Lim}$ the set $\{\omega\cdot\xi : \xi < \eta\}$ is unbounded in $\omega\cdot\eta$ and by lemma 19.21. it follows

(5) $\vdash^{\omega\cdot\eta}_{\omega\cdot\eta} \neg\mathrm{Fund}_{\omega\cdot\eta}(\tau), \mathrm{Fund}_{\omega\cdot\eta}(\varphi 0\tau)$.

19.23. **Lemma** (Euclidian division for ordinals)

For ordinals σ and $\tau \neq 0$ there are uniquely determined ordinals $\rho<\tau$ and η such that $\sigma = \tau\cdot\eta+\rho$.

Proof

Since $1\cdot\sigma\leq\tau\cdot\sigma<\tau\cdot\sigma'$ we have $\{\xi: \sigma<\tau\cdot\xi\} \neq \emptyset$. Define $\eta_0 := \min\{\xi: \sigma<\tau\cdot\xi\}$. Then it is $\eta_0 \neq 0$ and we have by definition that η_0 cannot be a limit ordinal. Hence there is an η such that $\eta_0 = \eta'$. It follows $\tau\cdot\eta\leq\sigma<\tau\cdot\eta' = \tau\cdot\eta+\tau$. Because of $\tau\cdot\eta\leq\sigma$ there is a ρ such that $\sigma = \tau\cdot\eta+\rho$. From $\tau\cdot\eta+\rho = \sigma<\tau\cdot\eta+\tau$ it follows $\rho<\tau$. The uniqueness of η and ρ is obvious.

19.24. **Lemma**

For $\sigma<\omega^{\nu+1}\cdot\eta$ there is an ordinal ξ such that $\sigma<\omega^{\nu}\cdot\xi'<\omega^{(\nu+1)}\cdot\eta$ and for $\mu<\nu$ there is an ordinal ζ such that $\omega^{\nu}\cdot\xi' = \omega^{(\mu+1)}\cdot\zeta$.

Proof

By Euclidian division we obtain $\sigma = \omega^{\nu}\cdot\xi+\rho$ for some $\rho<\omega^{\nu}$. Hence $\sigma<\omega^{\nu}\cdot\xi+\omega^{\nu} = \omega^{\nu}\cdot\xi'$. By $\omega^{\nu}\cdot\xi\leq\sigma<\omega^{(\nu+1)}\cdot\eta$ it follows $\xi<\omega\cdot\eta$ and therefore also $\xi'<\omega\cdot\eta$. Hence $\omega^{\nu}\cdot\xi'< \omega^{(\nu+1)}\cdot\eta$ and we obtain $\sigma< \omega^{\nu}\cdot\xi'< \omega^{(\nu+1)}\cdot\eta$. Now if $\mu<\nu$, then there is a δ such that $\nu = \mu'+\delta$ and it follows $\sigma<\omega^{\nu}\cdot\xi' = \omega^{(\mu+1+\delta)}\cdot\xi' = \omega^{(\mu+1)}\cdot(\omega^{\delta}\cdot\xi')$. The lemma is proved by defining $\zeta := \omega^{\delta}\cdot\xi'$.

19.25. **Lemma**

For all ordinals ν and $\eta \neq 0$ we have

$$\vdash^{\omega(1+\nu+1)\cdot\eta}_{\omega(1+\nu+1)\cdot\eta} \neg\mathrm{Fund}_{\omega(1+\nu+1)\cdot\eta}(\tau), \mathrm{Fund}_{\omega(1+\nu+1)\cdot\eta}(\varphi\nu\tau).$$

Proof by induction on ν

For $\nu = 0$ this follows from lemma 19.22. Therefore let $\nu \neq 0$. By lemma 19.24. we obtain for $\sigma<\omega^{(1+\nu+1)}\cdot\eta$. a $\xi \neq 0$ such that $\sigma<\omega^{1+\nu}\cdot\xi< \omega^{(1+\nu+1)}\cdot\eta$ and for $\mu<\nu$ a ζ such that $\omega^{1+\nu}\cdot\xi = \omega^{1+\mu+1}\cdot\zeta$. Using the induction hypothesis we have

(1) $\vdash^{\omega^{1+\nu}\cdot\xi}_{\omega^{1+\nu}\cdot\xi} \neg\mathrm{Fund}_{\omega^{1+\nu}\cdot\xi}(\upsilon), \mathrm{Fund}_{\omega^{1+\nu}\cdot\xi}(\varphi\mu\upsilon)$.

for all $\mu<\nu$. By $\bigvee$-importation and an $\bigwedge$-inference this implies

(2) $\vdash^{\omega^{1+\nu}\cdot\xi+3}_{\omega^{1+\nu}\cdot\xi} \forall\upsilon(\neg\mathrm{Fund}_{\omega^{1+\nu}\cdot\xi}(\upsilon)\vee \mathrm{Fund}_{\omega^{1+\nu}\cdot\xi}(\varphi\mu\upsilon))$.

Therefore we have

(3) $\frac{\omega^{1+\nu\cdot\xi+4}}{\omega^{1+\nu\cdot\xi}}\ \neg\mu<\nu \vee \forall\upsilon(\neg\mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\upsilon) \vee \mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\varphi\mu\upsilon))$

for all μ because we either have $\mu<\nu$ and (3) follows from (2) by an $\bigvee$-inference or we have $\neg\mu<\nu$ and (3) follows by an $\bigvee$-inference from an axiom according to (Ax1). The formula $A_{\omega^{1+\nu\cdot\xi}}(\nu)$, however, is of the form

$\forall\mu(\neg\mu<\nu \vee \forall\upsilon(\neg\mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\upsilon) \vee \mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\varphi\mu\upsilon)))$

Therefore we obtain from (3) by an $\bigwedge$-inference

(4) $\frac{\omega^{1+\nu\cdot\xi+5}}{\omega^{1+\nu\cdot\xi}}\ A_{\omega^{1+\nu\cdot\xi}}(\nu)$.

By lemma 19.20. there is an $\alpha<\omega^{1+\nu\cdot\xi}+\omega\cdot 2$ and a $\rho<\omega^{1+\nu\cdot\xi}+\omega$ such that

(5) $\frac{\alpha}{\rho}\ \neg A_{\omega^{1+\nu\cdot\xi}}(\nu), \neg\mathrm{Fund}_{\omega^{1+\nu\cdot\xi+\omega}}(\tau), \mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\varphi\nu\tau)$.

From (4) and (5) we obtain by a cut

(6) $\frac{\alpha_1}{\omega^{1+\nu\cdot\xi+\omega}}\ \neg\mathrm{Fund}_{\omega^{1+\nu\cdot\xi+\omega}}(\tau), \mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\varphi\nu\tau)$

for some $\alpha_1<\omega^{1+\nu\cdot\xi}+\omega\cdot 2$. Since $\omega^{1+\nu\cdot\xi}+\omega\leq\omega^{(1+\nu+1)\cdot\eta}$ it follows from (6) and lemma 19.10. by cut

(7) $\frac{\alpha_1+1}{\omega^{1+\nu\cdot\xi+\omega}}\ \neg\mathrm{Fund}_{\omega^{(1+\nu+1)\cdot\eta}}(\tau), \mathrm{Fund}_{\omega^{1+\nu\cdot\xi}}(\varphi\nu\tau)$.

Because of $\sigma<\omega^{1+\nu\cdot\xi}$ we obtain from (7) and 19.10.

(8) $\frac{\alpha_1+2}{\omega^{1+\nu\cdot\xi+\omega}}\ \neg\mathrm{Fund}_{\omega^{(1+\nu+1)\cdot\eta}}(\tau), \mathrm{Fund}_{\sigma}(\varphi\nu\tau)$.

Now, we have $\alpha_1<\omega^{1+\nu\cdot\xi}+\omega\cdot 2<\omega^{(1+\nu+1)\cdot\eta}$, as it is $\omega^{1+\nu\cdot\xi}<\omega^{1+\nu+1\cdot\eta}$ and $\omega\cdot 2<\omega^2$ [cf. lemma 24.7.] and therefore also $\alpha_1+2<\omega^{(1+\nu+1)\cdot\eta}$. By lemma 19.21. we finally obtain

(9) $\frac{\omega^{(1+\nu+1)\cdot\eta}}{\omega^{(1+\nu+1)\cdot\eta}}\ \neg\mathrm{Fund}_{\omega^{(1+\nu+1)\cdot\eta}}(\tau), \mathrm{Fund}_{\omega^{(1+\nu+1)\cdot\eta}}(\varphi\nu\tau)$.

19.26. **Definition**

We define $\zeta_0 = \varphi 10$ and $\zeta_{n+1} = \varphi\zeta_n 0$.

19.27. **Lemma**

We have $\sup\{\zeta_n : n<\omega\} = \Gamma_0$.

Proof

This is an immediate consequence of 17.22.

In a next step we are going to convince ourselves that all ordinals ζ_n are autonomously accessible in $\mathbf{Z}_\infty$. In order to do that we prove the following lemmata.

19.28. **Lemma**

We have $\frac{\omega^{3+1}}{\omega^{3+1}}\ \mathrm{Fund}_{\omega^3}(\zeta_0)$.

Proof

By lemma 19.25. we have

(1) $\vdash^{\omega^3}_{\omega^3} \neg\mathrm{Fund}_{\omega^3}(0), \mathrm{Fund}_{\omega^3}(\varphi 10)$.

By lemma 19.15. and a cut this implies

(2) $\vdash^{\omega^3+1}_{\omega^3+1} \mathrm{Fund}_{\omega^3}(\zeta_0)$.

19.29. **Lemma**

For all $n<\omega$ *it is* $\vdash^{\zeta_n\cdot\omega+1}_{\zeta_n\cdot\omega+1} \mathrm{Fund}_{\zeta_n\cdot\omega}(\zeta_{n+1})$.

Proof

We have $\omega^{1+\zeta_n} = \zeta_n$ for all $n<\omega$. By lemma 19.25. it holds

(1) $\vdash^{\zeta_n\cdot\omega}_{\zeta_n\cdot\omega} \neg\mathrm{Fund}_{\zeta_n\cdot\omega}(0), \mathrm{Fund}_{\zeta_n\cdot\omega}(\varphi\zeta_n 0)$.

From (1) and lemma 19.15. we obtain the claim by a cut.

19.30. **Lemma**

We have $\vdash^{\omega^2+\omega}_{\omega^2+\omega} \neg\mathrm{Fund}_{\omega^2}(\zeta_n), \mathrm{Fund}_{\omega^2}(\zeta_n\cdot\omega+1)$.

Proof

By lemma 19.14. it is

(1) $\vdash^{\alpha}_{\rho} \neg\mathrm{Fund}_{\omega^2}(\zeta_n), \mathrm{Fund}_{\omega^2}(\zeta_n+1)$

for $\alpha, \rho<\omega^2+\omega$. By lemma 19.22. we have

(2) $\vdash^{\omega^2}_{\omega^2} \neg\mathrm{Fund}_{\omega^2}(\zeta_n+1), \mathrm{Fund}_{\omega^2}(\varphi 0(\zeta_n+1))$.

It is $\varphi 0(\zeta_n+1) = \omega^{(\zeta_n+1)} = \zeta_n\cdot\omega$. Hence from (1) and (2) by cut

(3) $\vdash^{\alpha_1}_{\rho} \neg\mathrm{Fund}_{\omega^2}(\zeta_n), \mathrm{Fund}_{\omega^2}(\zeta_n\cdot\omega)$

for α_1 and ρ less than $\omega^2+\omega$. Because of $h(\zeta_n\cdot\omega+1) = \zeta_n\cdot\omega$ we obtain by 19.14.

(4) $\vdash^{\alpha_2}_{\rho} \neg\mathrm{Fund}_{\omega^2}(\zeta_n\cdot\omega), \mathrm{Fund}_{\omega^2}(\zeta_n\cdot\omega+1)$

for α_2 and ρ below $\omega^2+\omega$. From (3) and (4) and a cut we obtain the claim.

19.31. **Lemma**

For $S\in\mathfrak{K}_\sigma$ we have $\vdash^{5\cdot(\sigma+\alpha)}_{0} \neg\mathrm{Prog}(<,S), \forall\xi(\xi<\alpha \rightarrow \xi\in S)$.

Proof

The proof is by induction on α. From the induction hypothesis we have

(1) $\vdash^{5\cdot(\sigma+\beta)}_{0} \neg\mathrm{Prog}(<,S), \forall\xi(\xi<\beta \rightarrow \xi\in S)$ for all $\beta<\alpha$.

By the tautology lemma it holds

(2) $\vdash^{2\cdot\sigma}_{0} \beta\notin S, \beta\in S$.

From (1) and (2) we obtain by an $\bigwedge$-inference

(3) $\vdash_0^{5\cdot(\sigma+\beta)+1} \neg\mathrm{Prog}(<,S), \forall\xi(\xi<\beta \rightarrow \xi\in S)\wedge\beta\notin S, \beta\in S$

for all $\beta<\alpha$. This implies

(4) $\vdash_0^{5\cdot(\sigma+\beta)+2} \neg\mathrm{Prog}(<,S), \beta\in S$

for all $\beta<\alpha$ by an $\bigvee$-inference. From (4) we obtain

(5) $\vdash_0^{\delta} \neg\mathrm{Prog}(<,S), \neg\beta<\alpha, \beta\in S$

for all β where $\delta = 5\cdot(\sigma+\beta)+2$ for $\beta<\alpha$ and $\delta = 0$ for $\alpha\le\beta$. By $\bigvee$-importation and an $\bigwedge$-inference it finally follows

(6) $\vdash_0^{5\cdot(\sigma+\alpha)} \neg\mathrm{Prog}(<,S), \forall\xi(\xi<\alpha \rightarrow \xi\in S)$.

19.32. **Corollary**

$|\mathrm{Fund}(\alpha,X)|<\alpha^*$.

Proof

From lemma 19.31. it follows

(1) $\vdash_0^{5\cdot\alpha} \neg\mathrm{Prog}(<,X), \forall\xi(\xi<\alpha \rightarrow \xi\in X)$.

Hence

(2) $\vdash_0^{5\cdot\alpha+m} \mathrm{Fund}(\alpha,X)$

for some ordinal $m<\omega$. Since $5\cdot\alpha+m<\alpha^*$ this implies the claim.

19.33. **Lemma**

$\alpha\le|\mathrm{Fund}(\omega^\alpha,X)|$.

Proof

We have $2^\alpha\le\omega^\alpha\le 2^{|\mathrm{Fund}(\omega^\alpha,X)|}$ by the boundedness theorem. Hence $\alpha\le|\mathrm{Fund}(\omega^\alpha,X)|$.

19.34. **Lemma**

$\zeta_0 \subset \mathrm{Aut}(\omega+1)$.

Proof

By theorem 15.7 we have $\mathbf{Z}\vdash \mathrm{Fund}(\alpha,X)$ for all $\alpha<\varepsilon_0$. Hence $\vdash_n^{\omega+m} \mathrm{Fund}(\alpha,X)$ for all $\alpha<\varepsilon_0$ by the embedding lemma 11.2. Since $\omega\in\mathrm{Aut}(\omega+1)$ and $\mathrm{Aut}(\omega+1)$ is closed under successors we obtain by 19.32. that $\alpha^*\subset\mathrm{Aut}(\omega+1)$.

19.35. **Lemma**

$\mathrm{Aut}(\omega+1)$ *is closed under* $\lambda\xi.\,\omega^\xi$.

Proof

We show

$\alpha \in Aut(\omega+1) \Rightarrow \omega^\alpha \in Aut(\omega+1)$.

For $\alpha < \varepsilon_0$ this follows immediately from 19.34. Therefore assume $\varepsilon_0 \leq \alpha$. By 19.31. we have $\vdash_0^{5\cdot(\omega^2+\alpha)} \neg Prog(<,S), \forall\xi(\xi<\alpha \rightarrow \xi \in S)$ for all $S \in \mathfrak{K}_{\omega^2}$. Since $\varepsilon_0 \leq \alpha$ it is $5\cdot(\omega^2+\alpha) = \alpha + n$ for some $n<\omega$. Therefore we obtain $\vdash_0^\beta Fund_{\omega^2}(\alpha)$ for some $\beta<\alpha^*$. Hence $\beta \in Aut(\omega+1)$. By 19.22. we have $\vdash_{\omega^3}^{\omega^3} \neg Fund_{\omega^2}(\alpha), Fund_{\omega^2}(\omega^\alpha)$ as well as $\vdash_{\omega^3}^{\omega^3} \neg Fund(\omega^\alpha), Fund(\omega^{\omega^\alpha})$. By two cuts it follows $\vdash_{\omega^3}^{\beta+2} Fund_{\omega^2}(\omega^{\omega^\alpha})$ which by $\bigwedge$-inversion yields $\vdash_{\omega^3}^{\beta+2} Fund(\omega^{\omega^\alpha},X)$. Since $\beta+2<\alpha$ and $\omega^3<\varepsilon_0\leq\alpha$ we have $\beta+2 \in Aut(\omega+1)$ as well as $\omega^3 \in Aut(\omega+1)$. Hence $|Fund(\omega^{\omega^\alpha},X)| \in Aut(\omega+1)$ which by 19.33. also implies $\omega^\alpha \in Aut(\omega+1)$.

19.36. **Lemma**

$\zeta_n \in Aut(\omega+1)$ *implies* $\zeta_n\cdot\omega+1 \in Aut(\omega+1)$.

Proof

By 19.31. it holds $\vdash_0^{\zeta_n} \neg Prog(<,S), \forall\xi(\xi<\zeta_n \rightarrow \xi \in S)$ for all $S \in \mathfrak{K}_{\omega^2}$. Therefore there is an $\alpha<\zeta_n^*$ such that $\vdash_0^\alpha Fund_{\omega^2}(\zeta_n)$. This together with 19.30. implies $\vdash_{\omega^2+\omega}^{\alpha+1} Fund_{\omega^2}(\zeta_n\cdot\omega+1)$. By 19.22. we have $\vdash_{\omega^3}^{\omega^3} \neg Fund_{\omega^2}(\zeta_n\cdot\omega+1), Fund_{\omega^2}(\omega^{(\zeta_n\cdot\omega+1)})$. By cut and $\bigwedge$-inversion it follows $\vdash_{\omega^2+\omega}^{\alpha+2} Fund(\omega^{(\zeta_n\cdot\omega+1)},X)$. Since $Aut(\omega+1)$ is closed under successors $\zeta_n \in Aut(\omega+1)$ implies $\alpha+2<\zeta_n^* \subset Aut(\omega+1)$ and by 19.33. we obtain $\zeta_n\cdot\omega+1 \in Aut(\omega+1)$.

19.37. **Lemma**

$\zeta_0 \in Aut(\omega+1)$

Proof

By lemma 19.28. we have

(1) $\vdash_{\omega^3+1}^{\omega^3+1} Fund_{\omega^3}(\zeta_0)$.

By $\bigwedge$-inversion this implies

(2) $\vdash_{\omega^3+1}^{\omega^3+1} Fund(\zeta_0,X)$.

Since $\omega^{\zeta_0} = \zeta_0$ we obtain $\zeta_0 \in Aut(\omega+1)$ by 19.33. and 19.34.

19.38. **Lemma**

$\zeta_n \in Aut(\omega+1)$ *implies* $\zeta_{n+1} \in Aut(\omega+1)$.

Proof

$\zeta_n \in Aut(\omega+1)$ implies $\zeta_n \cdot \omega+1 \in Aut(\omega+1)$ by lemma 19.36. By lemma 19.29 we have $\vdash^{\zeta_n \cdot \omega+1}_{\zeta_n \cdot \omega+1} Fund_{\zeta_n;\omega}(\zeta_{n+1})$. By $\bigwedge$-inversion this implies $\vdash^{\zeta_n \cdot \omega+1}_{\zeta_n \cdot \omega+1} Fund(\zeta_{n+1},X)$. Since $\omega^{\zeta_n+1} = \zeta_{n+1}$ we obtain $\zeta_{n+1} \in Aut(\omega+1)$ by lemma 19.33.

19.39. **Corollary**

For all $n<\omega$ *we have* $\zeta_n \in Aut(\omega+1)$.

Proof

This follows immediately from 19.37. and 19.38. by induction on n.

19.40. **Theorem** (K.Schütte, S. Feferman)

It is $Aut(\omega+1) = \Gamma_0$.

Theorem 19.33. is the reason why Γ_0 is known as *the* boundary ordinal of predicativity. We are going to explain this in some more detail. We call a conception P *impredicative* or *circular* if the definition of P refers to a totality to which P itself belongs to. An example for an impredicative concept is the Russell set $M = \{x : x \notin x\}$ which we already mentioned in the introduction. There we defined M referring to a set universe V. More exactly the definition of M should be $M := \{x \in V : x \notin x\}$. Since we wanted M to be a set itself we have $M \in V$ and immediately obtain the contradiction from the fact that $M \in M \Leftrightarrow M \in V \wedge M \notin M$ and $M \in V$.

That impredicative definitions need not necessarily lead to immediate contradictions becomes clear by another example for an impredicative concept. Here we look at a monotone operator $\Gamma: P\mathbb{N} \rightarrow P\mathbb{N}$. Monotonicity for Γ here means that $S_1 \subset S_2$ also implies $\Gamma(S_1) \subset \Gamma(S_2)$. We obtain the least fixed point I_Γ of Γ by the definition $I_\Gamma := \bigcap\{S : \Gamma(S) \subset S\}$. However, it is easy to see that we also have $\Gamma(I_\Gamma) \subset I_\Gamma$, i.e. we defined I_Γ by referring to the set $\mathfrak{M} = \{S \subset \mathbb{N} : \Gamma(S) \subset S\}$ of Γ-closed sets to which I_Γ itself belongs. Operators of this kind are known as (monotone) inductive definitions. Inductive definitions are ubiquitous in mathematics in general and in mathematical logic especially and will therefore be the research object of the following chapter. Alarmed by the Russellian antinomy some mathematicians heavily doubted in impredicative definitions (and therefore also in inductive definitions). Thenceforth there have been and still are suggestions to construct mathematics by predicative means solely.

The most general approach to obtain a predicatively guaranteed segment of the ordinals is given by an autonomous creation process as we described it in the definition of Aut(ω+1). There we started by a semiformal system $\mathbf{Z}_{\omega+1}$ and therefore have all (codes for) ordinals which are provable in $\mathbf{Z}_{\omega+1}$. This guarantees the segment $\Gamma(\omega+1)$ of ordinals. But then we may argue in the system $\mathbf{Z}_{\Gamma(\omega+1)}$ and therefore will obtain the segment $\Gamma^2(\omega+1)$. Now we argue in $\mathbf{Z}_{\Gamma^2(\omega+1)}$, obtain $\Gamma^3(\omega+1)$ and so on. This kind of justification becomes even more striking if one regards infinitary systems for set theory as we will do it in a continuation of this lecture. There S_α denotes the subsystem which only allows formulas of ranks $\le\alpha$ and derivations of length $\le\alpha$. An ordinal β is acceptable to S_α if (a code for) β and the proof of the fact that it is (a code for) an ordinal both belong to S_α. One may then prove that the autonomous closure of S_ω again exactly is the ordinal Γ_0).

By theorem 18.2. we see that we in fact have to start with an infinite ordinal in order to obtain more than just finite ordinals. From theorem 18.6. it then follows that this autonomous creation process starting with an ordinal below Γ_0 never will access the ordinal Γ_0. On the other hand lemma 19.27. together with theorem 19.31. show that every ordinal below Γ_0 is accessible from the simplest infinite ordinal ω. In this sense Γ_0 is the least ordinal which is not predicatively definable.

Any formal system T whose proof theoretic ordinal is less or equal than Γ_0 may be embedded into the system $\mathbf{Z}_{\Gamma_0}$ in that sense that the Π^1_1-sentences provable in T are also provable in $\mathbf{Z}_{\Gamma_0}$. It therefore allows a predicative interpretation even if it looks impredicative at first glance. Therefore one usually calls a formal system predicative whenever its proof theoretic ordinal is less or equal than Γ_0.

A simple example for a predicative system stronger than $\mathbf{Z}_1$ is the system **ACA** in the exercises. There are a lot of formal systems between pure number theory an impredicative systems. The reader interested in predicativity should consult Feferman's various papers on predicativity (cf. the bibliography). The concern of this lecture, however, is to demonstrate on the example of one of the simplest impredicative systems by which means the boundary of predicativity may be overcome. This will be done in the following chapter.

CHAPTER III

Ordinal analysis of the formal theory for noniterated inductive definitions

§20. A summary of the theory of monotone inductive defintions over the natural numbers

We already frequently used inductive definitions during this lecture. We used it to define terms, formulas, derivations in formal and infinitary systems and also other concepts. Inductive definitions, however, are not only used in mathematical logic but are ubiquitous in mathematics. Whenever we define a set as the least set which comprehends a given set and is closed under certain operations, we use an inductive definition . One of the simplest examples is the subspace $\langle A\rangle$ of a vector space V generated by a set $A \subset V$. $\langle A\rangle$ is defined as the smallest set which comprehends A and itself is a vector space, i.e. is closed under addition and scalar multiplication. This does not look like an inductive definition as we are used to. But we also may put it into the more familiar form of a definition by clauses:

(1.i) $a \in A \Rightarrow a \in \langle A\rangle$

(1.ii) $a_1,\ldots,a_n \in \langle A\rangle \wedge \alpha_1,\ldots,\alpha_n \in K \Rightarrow \alpha_1 a_1 + \ldots + \alpha_n a_n \in \langle A\rangle$

where K denotes the ground field.

A more complex example for an inductive definition is the σ-algebra M_σ induced by a given set M over a domain Ω. The inductive definition of M_σ by clauses is:

(2.i) $M \subset M_\sigma$ and $\{\emptyset,\Omega\} \subset M_\sigma$.

(2.ii) $\forall i \in \omega(a_i \in M_\sigma) \Rightarrow \bigcup_{i<\omega} a_i \in M_\sigma$ und $\bigcap_{i<\omega} a_i \in M_\sigma$

(2.iii) $a \in M_\sigma \Rightarrow \bar{a} \in M_\sigma$, where $\bar{a}$ denotes the complement of a in Ω.

But also here M_σ may be defined as the smallest set which comprehends M and is closed under complements, countable unions and countable intersections. In general any set inductively defined by clauses may be regarded as the least fixed point of a certain operator Γ. In the case of the inductive definition of $\langle A\rangle$ the operator Γ would be given by

$\Gamma(S) := A \cup \{\Sigma\alpha_i a_i : \alpha_i \in K \wedge a_i \in S\}$.

In the case of the σ-algebra the definition of Γ would look like

$\Gamma(S) := \{\emptyset,\Omega\} \cup M \cup \{a : \exists f((f \text{ is a function and } \operatorname{dom} f = \omega)$
$\wedge \forall n<\omega(f(n) \in S \wedge (a = \bigcup\{f(n) : n<\omega\} \vee a = \bigcap\{f(n) : n \in \omega\}))) \vee \exists b \in S(a = \overline{b})\}.$

In both cases the operators are monotone, i.e. whenever we have $S \subset T$ it follows $\Gamma(S) \subset \Gamma(T)$. The least fixed point of a monotone operator Γ is the intersection of all sets which are Γ-closed. In §19 we already mentioned that fixed points of monotone operators are in general impredicatively defined.

The present section is supposed to provide a condensed (and therefore very rough) introduction to the theory of monotone inductive definitions on the natural numbers. The reader who is interested in more details is adviced to consult [Moschovakis 1974] and [Barwise 1975].

As a generalization of the above examples we obtain the following definition.

20.1. **Definition**

An *inductive definition* over the natural numbers is a monotone operator
$\Gamma : P\mathbb{N} \longrightarrow P\mathbb{N}$.

A set $A \subset \mathbb{N}$ is *Γ-closed* if $\Gamma(A) \subset A$. We define
$I_\Gamma := \bigcap\{A : \Gamma(A) \subset A\}$.

and call I_Γ *the fixed point* of the inductive definition Γ.

It is easy to see that I_Γ is the least fixed point of Γ. For a Γ-closed set A we always have $I_\Gamma \subset A$ and by the monotonicity of Γ also $\Gamma(I_\Gamma) \subset \Gamma(A) \subset A$. This shows $\Gamma(I_\Gamma) \subset \bigcap\{A : \Gamma(A) \subset A\} = I_\Gamma$ which means that I_Γ itself is Γ-closed. By monotonicity this implies $\Gamma(\Gamma(I_\Gamma)) \subset \Gamma(I_\Gamma)$ which proves that $\Gamma(I_\Gamma)$ is Γ-closed too. Hence $I_\Gamma \subset \Gamma(I_\Gamma)$ and we have $\Gamma(I_\Gamma) = I_\Gamma$. So I_Γ is a fixed point of Γ which by definition must be minimal.

So we have proven the following theorem:

20.2. **Theorem**

We have $\Gamma(I_\Gamma) = I_\Gamma$ *and* $\Gamma(S) \subset S$ *implies* $I_\Gamma \subset S$.

We already did emphazise that the fixed point of a monotone operator is impredicatively defined. In the case of an inductive definition, however, this impredicativity is not as fatal as it was in the case of the Russellian antinomy. By describing the least fixed point of an inductive definition Γ as the intersection of all Γ-closed sets we did not yet pay attention to the fact that an inductive definition is intended to build up its fixed point by successive application of the induction clauses. Interpreting an inductive definition as a monotone operator this means that we construct its fixed point in stages. We start with $\Gamma(\emptyset)$, then build $\Gamma(\Gamma(\emptyset))$, $\Gamma^3(\emptyset)$, $\Gamma^4(\emptyset)$,... etc. Hitherto we mostly regarded operators whose fixed point already is completed after ω-fold application of Γ which means that we obtained every element of the fixed point by finitely many applications of Γ. Examples for such inductive definitions are the definition of terms and $\mathcal{L}$-formulas. But there are also inductive definitions whose fixed points are more complicated. Examples for such inductive definitions are the σ-algebra M_σ, the definition of the $\mathcal{L}_\infty$-formulas, the definition of the relation $\models_0\Delta$ etc. In general we will have to apply the operator Γ a transfinite number of times in order to construct the fixed point of Γ from below. That means that we will need ordinals in the construction of the fixed point.

We define

$$I_\Gamma^\sigma = \Gamma(I_\Gamma^{<\sigma}) \text{ with the abbreviation } I_\Gamma^{<\sigma} := \bigcup\{I_\Gamma^\eta : \eta<\sigma\}.$$

From this definition we obtain

$$I_\Gamma^0 = \Gamma(\emptyset)\,,\ I_\Gamma^1 = \Gamma(\Gamma\emptyset))\,,\ I_\Gamma^n = \Gamma^n(\emptyset),\ I_\Gamma^\omega = \Gamma(\bigcup\{\Gamma^n(\emptyset) : n<\omega\}) \text{ etc.}$$

We call I_Γ^σ the σ-th *stage* in the inductive definition Γ. For $\sigma<\tau$ we have by definition $I_\Gamma^{<\sigma} \subset I_\Gamma^{<\tau}$ which implies $I_\Gamma^\sigma \subset I_\Gamma^\tau$ by the monotonicity of Γ. We defined I_Γ^ξ for all $\xi\in On$. But there is a $\sigma<\Omega$ such that $I_\Gamma^\sigma = I_\Gamma^{<\sigma}$ since otherwise we obtain for every $\sigma<\Omega$ an $n<\omega$ such that $n\in I_\Gamma^\sigma\setminus I_\Gamma^{<\sigma}$ which gives us a 1-1 mapping from Ω onto ω in contradiction to the regularity of Ω. Therefore there is a least ordinal σ_Γ which has this property. It is called the *closure ordinal of* Γ. By induction on ξ we obtain $I_\Gamma^\xi \subset I_\Gamma$. By the induction hypothesis we have $I_\Gamma^{<\xi} \subset I_\Gamma$ which implies $I_\Gamma^\xi = \Gamma(I_\Gamma^{<\xi}) \subset \Gamma(I_\Gamma) = I_\Gamma$ by the monotonicity of Γ. Hence $I_\Gamma^{\sigma_\Gamma} \subset I_\Gamma$ and we also have $I_\Gamma \subset I_\Gamma^{\sigma_\Gamma}$ because $I_\Gamma^{\sigma_\Gamma}$ is Γ-closed by definition. So we have $I_\Gamma = I_\Gamma^{\sigma_\Gamma}$ which shows that I_Γ can be obtained in stages from below.

If we assume that the operator Γ is definable by an $\mathcal{L}$-formula A, i.e. $\Gamma(S) = \{n\in\mathbb{N} : \mathbb{N}\models A_{X,x}[S,n]\}$ for $S\subset\mathbb{N}$, then we obtain $I_\Gamma^\sigma = \{n\in\mathbb{N} : \mathbb{N}\models A_{X,x}[I_\Gamma^{<\sigma},n]\}$. This shows that the definition of of I_Γ^σ is predicative relative to $I_\Gamma^{<\sigma}$. The definition of I_Γ^σ only refers to the previously constructed sets I_Γ^ξ for $\xi<\sigma$. Therefore the definition of the fixed point I_Γ is at least locally predicative. Of course it will only also be globally predicative if the closure ordinal of Γ is an ordinal below Γ_0 (which in general is not the case). Later we will see that it is exactly the local predicativity of the definition of I_Γ which makes the proof theoretical analysis of the theory for noniterated inductive definitions feasible. But before we may tackle the proof theoretical analysis of this theory we have to work out a suitable formal system for it.

20.3. **Definition**

An operator $\Gamma : P\mathbb{N} \to P\mathbb{N}$ is called *$\mathcal{L}_1$-definable*, if there is an $\mathcal{L}_1$-formula A such that $FV(A) = \{X, x\}$ and $\Gamma(S) = \{n \in \mathbb{N} : \mathbb{N} \models A_{X,x}[S,n]\}$. In this case we synonymously call Γ *arithmetical* or *arithmetically definable.*

Of course we may not expect that every arithmetical operator already is monotone. To ensure that we will only obtain monotone operators we are going to restrict ourselves to positive arithmetical operators. An operator is a *positive arithmetical operator* if it is definable by an X-positive $\mathcal{L}_1$-formula and an $\mathcal{L}_1$-formula A is an X-*positive formula* if its $\mathcal{L}_\infty$ translation A^* is an X-positive formula in the sense of definition 13.7.

20.4. **Lemma**

If Γ is a positive arithmetical operator, then Γ is monotone.

Proof

Let A be the defining formula of Γ and assume $S \subset T$. Then we obtain $\Gamma(S) = \{n \in \mathbb{N} : \mathbb{N} \models A[S,n]\} = \{n \in \mathbb{N} : \mathbb{N} \models A^*[S,n]\} \subset \{n \in \mathbb{N} : \mathbb{N} \models A^*[T,n]\} = \{n \in \mathbb{N} : \mathbb{N} \models A[T,n]\} = \Gamma(T)$ by lemma 9.5. and the monotonicity lemma 13.8.

20.5. **Remark**

By the positive arithmetical inductive definitions we, in fact, grasped the essential part of all monotone arithmetical inductive definitions. They comprehend all those monotone inductive definitions whose monotonicity is logically provable, i.e. is provable without using mathematical axioms. This is a consequence of the interpolation theorem by Lyndon and Craig.

Let A be the defining formula of Γ and assume that $X \subset Y \wedge A \to A_X(Y)$ is provable in first order predicate logic. Then by the Lyndon – Craig interpolation theorem we obtain an interpolation formula B such that $X \subset Y \wedge A \to B$ and $B \to A_X(Y)$. Since Y occurs positively in $X \subset Y \wedge A$ we know that Y has to occur positively in B too and for the special case $Y = X$ we obtain $A \leftrightarrow B$. Hence A is logically equivalent to an X-positive formula which shows that Γ is a positive operator. It in fact is not easy to find canonical examples for monotone operators which are not positive. We just marginally mention that the addition of monotone operators instead of just positive operators would not increase the proof theoretical strength of the formal system. The formulation and therefore also the analysis of the system, however, would become somewhat more complicated.

We are going to fix the following definitions.

20.6. Definitions

Let Γ be a positive arithmetical operator. We define:

(i) $I_\Gamma := \bigcap\{S \subset \mathbb{N} : \Gamma(S) \subset S\}$

(ii) $I_\Gamma^\sigma := \Gamma(I_\Gamma^{<\sigma})$, $I_\Gamma^{<\sigma} := \bigcup\{I_\Gamma^\xi : \xi < \sigma\}$

(iii) $|\Gamma| := \min\{\sigma < \Omega : I_\Gamma^\sigma = I_\Gamma^{<\sigma}\}$

(iv) $|n|_\Gamma := \begin{cases} \min\{\sigma : n \in I_\Gamma^\sigma\}, & \text{if this exists} \\ \Omega & \text{, otherwise} \end{cases}$

20.7. Lemma

$|\Gamma| = \sup\{|n|_\Gamma + 1 : n \in I_\Gamma\}$.

Proof

Define $\sigma := \sup\{|n|_\Gamma + 1 : n \in I_\Gamma\}$. For $\xi < \sigma$ then there is an $n \in I_\Gamma$ such that $\xi < |n|_\Gamma + 1$. By definition we have $I_\Gamma^{<|n|_\Gamma} \subsetneq I_\Gamma^{|n|_\Gamma}$. Because of $\xi \leq |n|_\Gamma$ we obtain $I_\Gamma^{<\xi} \subsetneq I_\Gamma^\xi$. Hence $\xi < |\Gamma|$ which entails $\sigma \leq |\Gamma|$. On the other hand if $\eta < |\Gamma|$, then we have $I_\Gamma^{<\eta} \subsetneq I_\Gamma^\eta$ and there is an $n \in I_\Gamma^\eta \setminus I_\Gamma^{<\eta}$ which means $|n|_\Gamma = \eta$. Hence $\eta < \sigma$ and we also have $|\Gamma| \leq \sigma$.

20.8. Remark

Without proving it we shall mention that for arithmetically definable monotone inductive defintions Γ it always holds $|\Gamma| \leq \omega_1^{CK}$. On the other hand there are positive arithmetical inductive definitions whose closure ordinal exactly is ω_1^{CK}. The proof of these statements needs methods of generalized recursion theory [e.g. cf. Moschovakis 1974, Barwise 1975, Hinman 1978].

20.9. Exercises

1. Let A be an arbitrary set, $M \subset P(A)$ and $\Gamma : P(A) \to P(A)$ the monotone operator which inductively defines the σ-algebra induced by M. Show $|\Gamma| \leq \aleph_1$.

2. Closure properties of inductively defined sets.

We call a set $M \subset P\mathbb{N}$ inductively defined if there is a positive arithmetical operator and a tuple $\mathbf{c} \in \mathbb{N}^m$ such that $x \in M \Leftrightarrow (x, \mathbf{c}) \in I_\Gamma$.

(i) Simultaneous inductive definitions

Suppose that F and G are X,Y-positive formulas. We define Δ_F^α and Δ_G^α by:

$$\Delta_F^\alpha := \{n \in \mathbb{N} : \mathbb{N} \models F_{X,Y,x}[\Delta_F^{<\alpha}, \Delta_G^{<\alpha}, n]\},\ \Delta_F^{<\alpha} := \bigcup_{\beta<\alpha} \Delta_F^\beta$$

and

$$\Delta_G^\alpha := \{n \in \mathbb{N} : \mathbb{N} \models G_{X,Y,x}[\Delta_F^{<\alpha}, \Delta_G^{<\alpha}, n]\},\ \Delta_G^{<\alpha} := \bigcup_{\beta<\alpha} \Delta_G^\beta$$

Prove that $\Delta_F := \bigcup_{\alpha \in On} \Delta_F^\alpha$ as well as $\Delta_G := \bigcup_{\alpha \in On} \Delta_G^\alpha$ are inductively defined.

(ii) Transitivity theorem

Let F be an X,Y-positive an G an X-positive formula. The operator $\Gamma : P\mathbb{N} \to P\mathbb{N}$ is defined by $\Gamma(S) := \{n \in \mathbb{N} | \mathbb{N} \models F_{X,Y,x}[I_G, S, n]\}$.

Show that I_Γ is inductively defined.

(iii) Stage comparison theorem

Let F and G be X-positive formulas. Define

$$n \leq^*_{F,G} m :\Leftrightarrow n \in I_F \wedge |n|_F \leq |m|_G$$

$$n <^*_{F,G} m :\Leftrightarrow |n|_F < |m|_G$$

Prove that there are positive arithmetical operators whose least fixed points are $\leq^*_{F,G}$ and $<^*_{F,G}$ repectively.

3. Prove the following claims.

(i) If Γ is a positive arithmetical operator then for all $\alpha \in On$ Γ^α is inductively defined.

(ii) If Γ is a positive arithmetical operator and $\alpha < \omega_1^{CK}$, then $\mathbb{N} - \Gamma^\alpha$ is an inductively defined set. (Use $\omega_1^{CK} = \sup\{|\Gamma| : \Gamma$ is a positive arithmetical operator$\}$.)

(iii) The sets definable by a Π_1^1-formula are exactly the inductively definable sets.

§ 21. *The formal system* $\mathbf{ID}_1$ *for noniterated inductive definitions*

In a next step we introduce a formal theory which formalizes the existence of inductively definable sets. Our intention is to enlarge the formal system $\mathbf{Z}_1$ such that we may not only talk about arithmetically definable classes (as we did in $\mathbf{Z}_1$) but also about classes which are fixed points of positive arithmetical operators. Since there is no canonical way to talk about ordinals in $\mathbf{Z}_1$ we are going to introduce fixed points of arithmetically definable operators Γ according to 20.6. as the intersection of all sets which are Γ-closed.

21.1. **Basic symbols** of the language $\mathscr{L}_I$ (cf. 1.5.)

(i) The logical symbols of $\mathscr{L}_I$ are those of the language $\mathscr{L}$.

(ii) All nonlogical symbols of the language $\mathscr{L}$ are also nonlogical symbols of $\mathscr{L}_I$.

(iii) If A is an X-positive $\mathscr{L}_1$-formula such that $FV(A) = \{x,X\}$, then $\underline{I}_A$ is a set constant.

21.2. **Definition** of the *terms* and *formulas* of $\mathscr{L}_I$

We define the terms $\mathscr{L}_I$ completely analogously to the $\mathscr{L}_1$-terms (cf. 1.6.).
The inductive definition of the $\mathscr{L}_I$-formulas is the same as the inductive definiton of the $\mathscr{L}_1$-formulas but the clause (ii) in the definition 1.7. is extended by

(ii.1) If $\underline{I}_A$ is a set constant and t is a term, then $t \in \underline{I}_A$ is a formula such that $BV(t \in \underline{I}_A) = FV_2(t \in \underline{I}_A) = \emptyset$ and $FV_1(t \in \underline{I}_A) = FV_1(t)$. $t \in \underline{I}_A$ is an atomic formula.

21.3. **Semantics** for $\mathscr{L}_I$

We obtain the semantics for the language $\mathscr{L}_I$ from the semantics for $\mathscr{L}_1$ by interpreting the set constant $\underline{I}_A$ by the fixed point I_A of the operator Γ_A given by $\Gamma_A(S) = \{n \in \mathbb{N} : \mathbb{N} \models A_{X,x}[S,n]\}$.

21.4. **The formal system $\mathbf{ID}_1$**

We are going to formulate the formal system $\mathbf{ID}_1$ in the language $\mathscr{L}_I$. $\mathbf{ID}_1$ extends the system $\mathbf{Z}_1$ which means that all axioms and inference rules of $\mathbf{Z}_1$ (cf. §3) are also axioms and rules of $\mathbf{ID}_1$. One should notice, however, that the extension of the language also means a strengthening of all schemes. So for instance arbitrary $\mathscr{L}_I$-formulas are allowed in the scheme (IND).
The group of mathematical axioms of $\mathbf{Z}_1$ (cf. 3.10.) is extended by:

(iv) the fixed point axioms

consisting of the schemes

($\mathbf{ID}^1_A$) $\forall x(A_X(\underline{I}_A, x) \rightarrow x \in \underline{I}_A)$

($\mathbf{ID}^2_A$) $\forall x(A_X(S,x) \rightarrow x \in S) \rightarrow \forall x(x \in \underline{I}_A \rightarrow x \in S)$,

where S denotes an arbitrary class term of the language $\mathscr{L}_I$. We formalize by the scheme $\mathbf{ID}^1$ that $\underline{I}_A$ is closed under the operator Γ_A. By $\mathbf{ID}^2_A$ it is formalized that $\underline{I}_A$ is contained in every Γ_A-closed class.
The definition of the relation $\mathbf{ID}_1 \vdash F$ is completely analogously to 3.11.

21.5. **Soundness theorem** for $\mathbf{ID}_1$
If $\mathbf{ID}_1 \vdash F$, *then* $\mathbb{N} \models F$.

The proof of the soundness theorem for $\mathbf{ID}_1$ is exactly the same as the proof of the soundness theorem for $\mathbf{Z}_1$. We use induction on the definition of $\mathbf{ID}_1 \vdash F$. The only additional work is to check the axioms $\mathbf{ID}^1_A$ and $\mathbf{ID}^2_A$. The validity of these axioms, however, is obvious by 20.2.

Now once again we are in the situation to have a formal system which produces theorems of $\mathbb{N}$. Similar as in the case of pure number theory we want to obtain an estimate for the norm of the provable Π^1_1-sentences of $\mathbf{ID}_1$. Therefore in a next step we will have to examine how inductive definitions may be represented in the language $\mathcal{L}_\infty$.

21.6. **Exercise**
Let A be an X-positive $\mathcal{L}_1$-formula such that $FV(A) \subset \{X,x\}$.
We define $Cl_A(Y) :\equiv \forall x(A_X(Y) \rightarrow x \in Y)$. Show the following statements:

(i) $n \in I_A \Leftrightarrow \mathbb{N} \models Cl_A(X) \rightarrow \underline{n} \in X$

(ii) $\vdash \forall x(x \in X \rightarrow x \in Y) \rightarrow \forall x(A \rightarrow A_X(Y))$

(iii) $\mathbf{ID}_1 \vdash t \in \underline{I}_A \rightarrow Cl_A(X) \rightarrow t \in X$

(iv) $\mathbf{ID}_1 \vdash t \in \underline{I}_A \leftrightarrow A_{X,x}(\underline{I}_A, t)$

(v) $\mathbf{ID}_1 \vdash t \in \underline{I}_A \Leftrightarrow \mathbf{ID}_1 \vdash Cl_A(X) \rightarrow t \in X$

(vi) If $\prec$ is an $\mathcal{L}_1$-definable order relation and $A(X,x)$ is the X-positive formula $\forall y(y \prec x \rightarrow y \in X)$, then we have $\mathbf{ID}_1 \vdash \mathrm{Fund}(\prec, X) \Leftrightarrow \mathbf{ID}_1 \vdash \forall x(x \in \mathrm{Feld}(\prec) \rightarrow x \in \underline{I}_A)$.

21.7. **Note**
The first one (as far as we know) who introduced a formal system for generalized inductive definitions was *G.Kreisel* in [Kreisel 1963a]. Comprehensive studies on formal systems for (also iterated) inductive definitions can be found in [Feferman 1970a] and in the articles by *Feferman* and *Feferman and Sieg* in [BFPS].

§ 22. Inductive definitions in $\mathcal{L}_\infty$

Infinitary languages owe their power to the possibility of reflecting the properties of the ordinals of the real world by transfinitely long formulas. Here we are going to use this fact to express the stages and the fixed point of an inductive definition in the infinitary language $\mathcal{L}_\infty$. The language $\mathcal{L}_\infty$ is too complex for an immediate proof theoretical analysis. For the interpretation of $\mathbf{ID}_1$, however, we will only need a fragment $\mathcal{L}_\infty^I$ of $\mathcal{L}_\infty$.

22.1. Definition

An $\mathcal{L}_\infty$-formula is called *arithmetical* if it is the *-translation of an $\mathcal{L}_1$-formula. A Π_1^1-sentence of $\mathcal{L}_\infty$ is the *-translation of a Π_1^1-sentence of $\mathcal{L}_1$. In the sequel we are going to identify $\mathcal{L}_1$-formulas and their *-translations. So we regard $\mathcal{L}_1$ as a sublanguage of $\mathcal{L}_\infty$. Albeit there are no number variables in $\mathcal{L}_\infty$ we will speak of arithmetical $\mathcal{L}_\infty$-formulas A such that $FV_1(A) = \{x_1,...,x_n\}$. In this context we mean the $\mathcal{L}_1$-formula A. Of course A will not become a wellformed $\mathcal{L}_\infty$-formula unless all number variables are replaced by closed terms.

22.2. Recursive definition the formula $(t \in \underline{I}_A^{<\alpha})$

Suppose that $\alpha \le \Omega$ and let A be an X-positive arithmetical formula such that $FV(A) = \{X, x\}$. Then we define:

(i) $(t \in \underline{I}_A^{<0})$ is the formula $(\underline{0} = \underline{1})$

(Since $t \in \underline{I}_A^{<0}$ is the false formula $\underline{I}_A^{<0}$ will represent the empty set.)

(ii) $(t \in \underline{I}_A^{<\alpha})$ is the formula $\bigvee\{A_{X,x}(\underline{I}_A^{<\xi}, t) : \xi < \alpha\}$.

By $t \in \underline{I}_A^{\alpha}$ we denote the formula $A_{X,x}(\underline{I}_A^{<\alpha}, t)$. This notation coincides with the intended interpretation (cf.20.6.ii)

Instead of $\neg(t \in \underline{I}_A^{<\alpha})$ we usually write $(t \notin \underline{I}_A^{<\alpha})$. However, we want to stress that $(t \notin \underline{I}_A^{<\alpha})$ is *not* an atomic formula but the formula $\bigwedge\{\neg A_{X,x}(\underline{I}_A^{<\xi}, t) : \xi < \alpha\}$. We also use the notation $t \notin \underline{I}_A^{\alpha}$ instead of $\neg A_{X,x}(\underline{I}_A^{<\alpha}, t)$.

22.3. Definition of the formulas of $\mathcal{L}_\infty^I$

The formulas of $\mathcal{L}_\infty^I$ are all of the shape $F_{X_1,...,X_n}(S_1,...,S_n)$ where F is a Π_1^1-sentence and for $k = 1,...,n$ S_k is a set term $\{x : x \in \underline{I}_A^{<\alpha}\}$ for some X-positive arithmetical formula A and some ordinal α.

The conventions how to replace a set variable X by a class term S had been defined in 1.8. In the sequel we will write $\underline{I}_A^{<\alpha}$ instead of $\{x : x \in \underline{I}_A^{<\alpha}\}$ and $\neg\underline{I}_A^{<\alpha}$ instead of $\{x : x \notin \underline{I}_A^{<\alpha}\}$.

22.4. Definition

For an $\mathcal{L}_\infty$-formula F we define the X-*rank* $rk_X(F)$ inductively by:

(i) $rk_X(F) = 0$ if F is atomic

(ii) $rk_X(\bigwedge\{F_\iota : \iota \in I\}) = rk_X(\bigvee\{F_\iota : \iota \in I\}) = \sup\{rk_X(F_\iota)+1 : \iota \in I \wedge X \in FV(F_\iota)\}$.

We obviously have $rk_X(F) \le rk(F)$. For a Π_1^1-formula F it even holds that rk(F) =

$rk_X(F)+n$ for some finite ordinal n.

22.5. Lemma

For $\alpha \geq \omega$ *we have* $rk(t \in \underline{I}_A^{<\alpha}) = (rk_X A+1)\cdot\alpha$

Proof

Before showing the lemma we notice that in general it is $rk(F_X(S)) \leq rk(F)+rk(S)$. With the additional hypothesis $X \in FV(F)$ and $rk(S) \geq rk(F)$, however, we even obtain $rk(F_X(S)) = rk(S)+rk_X(F)$ by induction on $rk(F)$. For atomic F this is obvious. For $F \equiv \bigwedge\{F_\iota : \iota \in I\}$ or $F \equiv \bigvee\{F_\iota : \iota \in I\}$ we have $rk(F_X(S)) = \sup\{rk(F_{\iota X}(S))+1 : \iota \in I\} = \sup\{rk(F_{\iota X}(S))+1 : \iota \in I \wedge X \in FV(F_\iota)\}$, since for $X \notin FV(F_\iota)$ and $X \in FV(F_\varkappa)$ we have $rk(F_{\iota X}(S)) = rk(F_\iota) < rk(F) \leq rk(S) \leq rk(F_{\varkappa X}(S))$. By the induction hypothesis it follows $rk(F_X(S)) = \sup\{rk(S)+rk_X(F_\iota) : \iota \in I \wedge X \in FV(F_\iota)\} = rk(S)+\sup\{rk_X(F_\iota) : \iota \in I \wedge X \in FV(F_\iota)\} = rk(S)+rk_X(F)$.

We show the lemma by induction on α.

For $\alpha = \omega$ we have $rk(t \in \underline{I}_A^{<\alpha}) = \sup\{rk(t \in \underline{I}_A^n)+1 : n<\omega\} = \omega$ since all formulas $(t \in \underline{I}_A^n)$ have a finite rank which at least is n. Let us assume that $\alpha > \omega$. Then we have $rk(t \in \underline{I}_A^{<\alpha}) \geq \omega > rk(A)$ and may therefore use the introductory remark. Hence $rk(t \in \underline{I}_A^{<\alpha}) = \sup\{rk(t \in \underline{I}_A^\xi)+1 : \xi<\alpha\} = \sup\{rk(t \in \underline{I}_A^{<\xi})+rk_X(A)+1 : \xi<\alpha\} \stackrel{i.h.}{=}$ $\sup\{(rk_X(A)+1)\cdot\xi + rk_X(A)+1 : \xi<\alpha\} = \sup\{(rk_X(A)+1)\cdot(\xi+1) : \xi<\alpha\} =$ $\sup\{rk_X(A)+1)\cdot\xi : \xi\leq\alpha\} = (rk_X(A)+1)\cdot\alpha$.

22.6. Corollary

We have $rk(t \in \underline{I}_A^{<\alpha}) < \alpha+\omega$ *and* $rk(t \in \underline{I}_A^\alpha) < \alpha+\omega$.

Proof

The claim is obvious for $\alpha<\omega$. For $\omega\leq\alpha$ we have $rk(t \in \underline{I}_A^{<\alpha}) = (rk_X(A)+1)\cdot\alpha$. Since $rk_X(A)<\omega$ we obtain $rk_X(A)+1 = n<\omega$. Let $\alpha =_{NF} \omega^{\alpha_1}+\ldots+\omega^{\alpha_k}$. For $\alpha_i \neq 0$ it is $n\cdot\omega^{\alpha_i} = \omega^{\alpha_i}$. Hence $n\cdot\alpha < \alpha+\omega$. Furthermore we obtain $rk(t \in \underline{I}_A^\alpha) = rk(t \in \underline{I}_A^{<\alpha}) + rk_X(A) < \alpha+\omega$, since $rk(t \in \underline{I}_A^{<\alpha}) = \alpha+n_1$ for some $n_1<\omega$ and $rk_X(A)<\omega$.

22.7. Lemma

$I_A^{<\alpha} = \{n \in \mathbb{N} : \mathbb{N} \models \underline{n} \in \underline{I}_A^{<\alpha}\}$.

Proof by induction on α

For $\alpha = 0$ we have $I_A^{<\alpha} = \emptyset$ and $\{n \in \mathbb{N} : \mathbb{N} \models \underline{n} \in \underline{I}_A^{<\alpha}\} = \{n \in \mathbb{N} : \mathbb{N} \models \underline{0} = \underline{1}\} = \emptyset$.

For $\alpha \neq 0$ we have $n \in I_A^{<\alpha}$ if and only if there is a $\xi<\alpha$ such that $n \in I_A^\xi$, i.e.

$\mathbb{N}\models A[I_A^{<\xi},\underline{n}]$. By induction hypothesis this is equivalent to $\mathbb{N}\models A(\underline{I}_A^{<\xi},\underline{n})$ which implies $\mathbb{N}\models\bigvee\{A(\underline{I}_A^{<\xi},\underline{n}):\xi<\alpha\}$, i.e. $\mathbb{N}\models\underline{n}\in\underline{I}_A^{<\alpha}$. On the other hand, if we have $\mathbb{N}\models\underline{n}\in\underline{I}_A^{<\alpha}$, then there is a $\xi<\alpha$ such that $\mathbb{N}\models\underline{n}\in\underline{I}_A^{\xi}$. By the induction hypothesis this implies $n\in I_A^{\xi}$ which entails $n\in I_A^{<\alpha}$.

22.8. **Corollary**

$I_{\Gamma_A}=\{n\in\mathbb{N}:\mathbb{N}\models\underline{n}\in\underline{I}_A^{<\Omega}\}$.

Proof

We have $I_{\Gamma_A}=I_{\Gamma_A}^{<\sigma_0}$ for some $\sigma_0<\Omega$. But then $I_{\Gamma_A}=I_{\Gamma_A}^{<\Omega}$, and the claim follows from 22.7.

22.9. **Definition**

If we augment the definition 9.4. by the additional clause

(vii) $(t\in\underline{I}_A)^*:=t\in\underline{I}_{A^*}^{<\Omega}$,

then we obtain a translation F^* for every $\mathcal{L}_I$-formula F.

22.10 **Lemma**

(i) If F is an $\mathcal{L}_I$-formula, then F^* is an $\mathcal{L}_\infty^I$-formula such that $rk(F^*)\le\Omega+n$ *for some* $n<\omega$.

(ii) We have $\mathbb{N}\models F$ *if and only if* $\mathbb{N}\models F^*$.

Proof

(i) An $\mathcal{L}_I$-formula F has the form $F\equiv A_{X_1\ldots X_n}(\underline{I}_{B_1},\ldots,\underline{I}_{B_n})$ where A is a Π_1^1-sentence Then we obtain $F^*\equiv A^*(\underline{I}_{B_1^*}^{<\Omega},\ldots,\underline{I}_{B_n^*}^{<\Omega})$ which obviously is an $\mathcal{L}_\infty^I$-formula. If none of the set-variables X_k ($k=1,\ldots,n$) occurs in A we have $rk(F^*)<\omega<\Omega+n$. Otherwise it follows from 22.6 that there is an $n<\omega$ such that $rk(F^*)=\Omega+n$.

(ii) follows from definition 22.9. by 22.8.

We already mentioned that $\mathcal{L}_I$ is a fragment of $\mathcal{L}_\infty$ without defining the fragments of $\mathcal{L}_\infty$. This will remedied by the following definition.

22.11. **Definition**

(i) A collection $\mathfrak{F}$ of $\mathcal{L}_\infty$-formulas is *closed under first order operations* if $\mathfrak{F}$ is closed under the sentential conjunctives $\neg,\vee$ and $\wedge$ and if $F(\underline{n})\in\mathfrak{F}$ for some $n\in\mathbb{N}$ implies $\bigwedge\{F(\underline{n}):n<\omega\}\in\mathfrak{F}$ as well as $\bigvee\{F(\underline{n}):n<\omega\}\in\mathfrak{F}$. $\mathfrak{F}$ is *closed* *under substitutions* if $F(s) \in \mathfrak{F}$ for some term s implies $F(t) \in \mathfrak{F}$ for all terms t.

(ii) The set SF(F) of subformulas of an $\mathscr{L}_\infty$-formula F is inductively defined by

(a) If F is atomic, then $SF(F) := \{F\}$

(b) If F is a formula $\bigwedge\{A_k : k \in I\}$ or a formula $\bigvee\{A_k : k \in I\}$, then we define $SF(F) := \{F\} \cup \bigcup\{SF(A_k) : k \in I\}$.

(iii) A collection $\mathfrak{F}$ of $\mathscr{L}_\infty$- formulas is *closed under subformulas* if $F \in \mathfrak{F}$ implies $SF(F) \subset \mathfrak{F}$.

(iv) A collection $\mathfrak{F}$ of $\mathscr{L}_\infty$- formulas is called a *fragment* of $\mathscr{L}_\infty$ if it satisfies the following conditions:

(a) $\mathscr{L}_1 \subset \mathfrak{F}$

(b) $\mathfrak{F}$ is closed under first order operations and substitutions

(c) $\mathfrak{F}$ is closed under subformulas.

The following lemma is an easy exercise.

22.12. **Lemma**

The language $\mathscr{L}^{\mathrm{I}}_\infty$ is a fragment of $\mathscr{L}_\infty$.

In fact we could define $\mathscr{L}^{\mathrm{I}}_\infty$ as the fragment of $\mathscr{L}_\infty$ which is induced by the *-translation defined in 22.9.

For a fragment $\mathfrak{F}$ of $\mathscr{L}_\infty$ and a finite set Δ of $\mathfrak{F}$-formulas we define the relation $\mathfrak{F} \models_0 \Delta$ in analogy to 5.2.

22.13. **Inductive definition** of $\mathfrak{F} \models_0 \Delta$

(Ax1) If $\chi_P(t_1^{\mathbb{N}},\dots,t_n^{\mathbb{N}}) = 1$ and $(\underline{P}t_1 \dots t_n) \in \Delta \subset \mathfrak{F}$, then $\mathfrak{F} \models_0 \Delta$

(Ax2) We have $\mathfrak{F} \models_0 \Delta, t \in X, s \notin X$ if $\Delta, t \in X, s \notin X \subset \mathfrak{F}$ and $s^{\mathbb{N}} = t^{\mathbb{N}}$

($\bigwedge$) If $\mathfrak{F} \models_0 \Delta, A_i$ for all $i \in I$ and $\bigwedge\{A_i : i \in I\} \in \mathfrak{F}$, then $\mathfrak{F} \models_0 \Delta, \bigwedge\{A_i : i \in I\}$

($\bigvee$) If $\mathfrak{F} \models_0 \Delta, A_i$ for some $i \in I$ and $\bigvee\{A_i : i \in I\} \in \mathfrak{F}$, then $\mathfrak{F} \models_0 \Delta, \bigvee\{A_i : i \in I\}$

We will not have to redefine $\mathfrak{F} \models^{\alpha}_0 \Delta$. Its definition is as in 9.1. The relation $\mathfrak{F} \models^{\alpha}_0 \Delta$ informally says that there is an $\mathfrak{F}$-deduction tree of Δ whose length is exatcly α.

If $\mathfrak{F}$ is a fragment of $\mathscr{L}_\infty$ and Δ is a finite set of formulas in $\mathfrak{F}$ for which we have $\mathbf{Z}_\infty \models^{\alpha}_0 \Delta$, then an easy induction on α shows that all formulas of the derivation tree belong to $\mathfrak{F}$. This is essentially due to the facts that fragments are closed under subformulas and substitutions and cut free derivations do have the subformula property, i.e. all formulas in the premise of an inference are subformulas of some formula in the conclusion. In analogy to 10.4. we therefore obtain

22.14. **Lemma**

Let $\mathfrak{F}$ be a fragment of $\mathscr{L}_\infty$, Δ a finite set of formulas in $\mathfrak{F}$ and M a segment of the ordinals. Then we have

$$\mathbf{Z_M} \overset{\alpha}{\underset{0}{\models}} \Delta \Leftrightarrow \alpha \in \mathrm{M} \wedge \exists \xi \leq \alpha\, (\mathfrak{F} \overset{\xi}{\underset{0}{\models}} \Delta).$$

This entails

$$\alpha \in \mathrm{M} \wedge \mathfrak{F} \overset{\alpha}{\underset{0}{\models}} \Delta \Rightarrow \mathbf{Z_M} \overset{\alpha}{\underset{0}{\models}} \Delta.$$

No we are going to show that for countable fragments of $\mathscr{L}_\infty$ we have a completeness theorem similar to 5.4.

22.15. **Soundness and completeness theorem** for $\mathfrak{F} \underset{0}{\models}$.

If $\mathfrak{F}$ is a countable fragment and Δ is a finite set of $\mathfrak{F}$-formulas, then we have $\mathbb{N} \models \bigvee\{\mathrm{F} : \mathrm{F} \in \Delta\}$ if and only if we have $\mathfrak{F} \underset{0}{\models} \Delta$.

The soundness of $\mathfrak{F} \underset{0}{\models}$ follows as in 5.3. by induction on the definition of $\mathfrak{F} \underset{0}{\models} \Delta$. In order to show the opposite direction we are going to copy the proof in 5.4. All we have to do is to convince ourselves that the label function δ in definition 5.6. only can take finite sets of $\mathfrak{F}$-formulas as values. But this follows from the fact that $\mathfrak{F}$ is closed under subformulas.

The completeness theorem, however, which we really are looking for is more complicated. We desire the following theorem.

22.16. **Soundness and completeness theorem** for $\mathscr{L}^{\mathrm{I}}_\infty \underset{0}{\models}$

For any $\mathscr{L}_{\mathrm{I}}$-formula F *we have $\mathbb{N} \models \mathrm{F}$ if and only if $\mathscr{L}^{\mathrm{I}}_\infty \underset{0}{\models} \mathrm{F}^*$.*

Proof

By 22.10. we have $\mathbb{N} \models \mathrm{F} \Leftrightarrow \mathbb{N} \models \mathrm{F}^*$. Therefore all we need is a theorem of the form $\mathbb{N} \models \mathrm{F}^* \Leftrightarrow \mathscr{L}^{\mathrm{I}}_\infty \underset{0}{\models} \mathrm{F}^*$. This theorem, however, is not just a special case of 22.15., since $\mathscr{L}^{\mathrm{I}}_\infty$ - as we defined it yet- is not a countable fragment of $\mathscr{L}_\infty$. The soundness of the infinitary calculus $\mathscr{L}^{\mathrm{I}}_\infty \underset{0}{\models}$ follows easily by induction on the definition of $\mathscr{L}^{\mathrm{I}}_\infty \underset{0}{\models} \mathrm{F}$. The problem is to show completeness. We will do this in two different ways which, however, will yield two different theorems.

The first way uses remark 20.8. in which - without proof- we stated that the stages of an inductive definition already become stationary at ω_1^{CK} the first recursively regular ordinal above ω. Instead of interpreting Ω by $\aleph_1$ -the first uncountable regular ordinal - we may therefore as well interpret Ω by ω_1^{CK} without spoiling the soundness of the system $\mathscr{L}_\infty^I \models_0$. (From §25 it will follow that this alternative interpretation is also a sound interpretation for the ordinal notations developed in §§ 23 and 24.) Since ω_1^{CK} is a countable ordinal we obtain $\mathscr{L}_\infty^I$ as a countable fragment of $\mathscr{L}_\infty$ and theorem 22.16. is a special case of 22.15. In the second approach we leave the interpretation of Ω as the first uncountable regular ordinal but extend the calculus $\mathscr{L}_\infty^I \models_0$ to a calculus $\mathscr{L}_\infty^{I*} \models_0$ by adding a new rule

(Cl_Ω) $\mathscr{L}_\infty^{I*} \models_0 \Delta, A(\underline{I}_A^{<\Omega}, \underline{n}) \Rightarrow \mathscr{L}_\infty^{I*} \models_0 \Delta, \underline{n} \in \underline{I}_A^{<\Omega}$.

Because of $\mathbb{N} \models A(\underline{I}_A^{<\Omega}, \underline{n}) \Rightarrow \mathbb{N} \models \underline{n} \in \underline{I}_A^{<\Omega}$ the addition of this rule will not spoil the soundness of the calculus $\mathscr{L}_\infty^{I*} \models_0$. To show its completeness we, in analogy to 5.6., define quasideduction trees for the extended calculus $\mathscr{L}_\infty^{I*} \models_0$.

We adopt the clauses (i)-(iii) in the definition 5.6. Clause (iv), however, only makes sense if the distinguished redex is a countable disjunction, i.e. if it is not of the form $\underline{n} \in \underline{I}_A^{<\Omega}$. Otherwise it could happen that the quasideduction path of $\underline{n} \in \underline{I}_A^{<\Omega}$ will not become finite although we have $\mathbb{N} \models \underline{n} \in \underline{I}_A^{<\Omega}$. Therefore we only adopt clause (iv) for those cases in which the distinguished redex is a disjunction of countable length, i.e. is not of the form $\underline{n} \in \underline{I}_A^{<\Omega}$. For this case we introduce an additional clause

(v) If $\sigma \in B_\Delta$ and $\delta(\sigma)$ reducible with distinguished redex $\underline{n} \in \underline{I}_A^{<\Omega}$, then $\sigma * \langle 0 \rangle \in B_\Delta$ and $\delta(\sigma * \langle 0 \rangle) := \delta(\sigma)^r, A(\underline{I}_A^{<\Omega}, \underline{n})$.

We call the resulting tree B_Δ the *extended quasideduction tree* of Δ. Now we have to check that the syntactical and the semantical main lemma also holds for extended quasideduction trees.

22.17. Extended syntactical main lemma for $\mathscr{L}_\infty^{I*} \models_0$.

Suppose that every path in the extended quasideduction tree of a finite set Δ of $\mathscr{L}_\infty^I$-formulas contains an axiom. Then $\mathscr{L}_\infty^{I*} \models_0 \Delta$.

The proof follows the proof of 5.7. We have only to consider one additional case:

If the distinguished redex is of the form $\underline{n} \in \underline{I}_A^{<\Omega}$, then we have $\sigma * \langle 0 \rangle \in B_\Delta$ and $\delta(\sigma * \langle 0 \rangle) = \delta(\sigma)^r$, $A(\underline{I}_A^{<\Omega}, \underline{n})$. By the induction hypothesis we obtain

$\mathscr{L}_\infty^{I*} \models_0 \delta(\sigma)^r, A(\underline{I}_A^{<\Omega}, \underline{n})$

which by the Cl_Ω-rule implies

$\mathscr{L}_\infty^{I*} \models_0 \delta(\sigma)^r, \ \underline{n} \in I_A^{<\Omega}$.

22.18. **Extended semantical main lemma** for $\mathscr{L}_\infty^{I*} \models_0$

Let $\Delta \subset \mathscr{L}_\infty^I$ be a finite set of formulas such that there is a path in the quasi-deduction tree of Δ which does not contain an axiom. Then there is an assignment Φ such that $\mathbb{N} \not\models F^\Phi$ holds for all $F \in \Delta$.

The proof is mainly the same as that of 5.8. Again we choose a path f in B_Δ which does not contain an axiom. The properties (1)–(4) of 5.8. for f are conserved with the restriction, that 5.8.(4) only holds for countable disjunctions, i.e. for disjunctions which do not have the form $\underline{n} \in I_A^{<\Omega}$. As an additional property for f we obtain

(4') If $\sigma \in f$ and $\underline{n} \in I_A^{<\Omega} \in \delta(\sigma)$, then there is a $\tau \in f$ such that $A(I_A^{<\Omega}, \underline{n}) \in \delta(\tau)$.

The proof of (4') follows from 5.8.(1) and the definition clause (v).

Now we are going to define an assignment Φ by:

$\Phi(X) := \{n \in \mathbb{N} : t^{\mathbb{N}} = n \wedge (\exists \sigma \in f)((t \notin X) \in \delta(\sigma))\}$

For an ordinal $\xi < \Omega$ we denote by F^ξ the $\mathscr{L}_\infty^I$-formula obtained from the formula F by replacing all positive occurences of $\underline{n} \in I_A^{<\Omega}$ by $\underline{n} \in I_A^{<\xi}$.

Then we obtain

(5') $\mathbb{N} \not\models F^{\xi,\Phi}$ for all $\xi < \Omega$, $\sigma \in f$ and $F \in \delta(\sigma)$.

by induction on $rk(F^\xi)$. The cases in which F is not of the form $\underline{n} \in I_A^{<\Omega}$ are treated as in 5.8.(5) 1.–4.

If $F \equiv \underline{n} \in I_A^{<\Omega}$, then F^ξ is the formula $\underline{n} \in I_A^{<\xi}$ and by (4') there is a $\tau \in f$ such that $A(I_A^{<\Omega}, \underline{n}) \in \delta(\tau)$, i.e. $\underline{n} \in I_A^{\Omega} \in \delta(\tau)$. We have $rk(\underline{n} \in I_A^{\eta}) < rk(\underline{n} \in I_A^{<\xi})$ for all $\eta < \xi$ and therefore obtain $\mathbb{N} \not\models \underline{n} \in I_A^{\eta}$ for all $\eta < \xi$ by the induction hypothesis. This, however, implies $\mathbb{N} \not\models \underline{n} \in I_A^{<\xi}$.

Since $\mathbb{N} \not\models F^\xi$ for all $\xi < \Omega$ already implies $\mathbb{N} \not\models F$ the extended semantical main lemma is an immediate consequence of (5').

22.19. **Soundness and completeness theorem for** $\mathscr{L}_\infty^{I*} \models_0$.

For any $\mathscr{L}_I$-formula F *we have* $\mathbb{N} \models F$ *if and only if* $\mathscr{L}_\infty^{I*} \models_0$.

22.16. and 22.19. are of course different theorems because in 22.16. we have a countable fragment whereas in 22.19. we are talking about an uncountable fragment. A reinspection of the proof of 22.19., however, shows that we used the interpretation of Ω as the first uncountable regular ordinal only to assure the soundness of the Cl_Ω-rule. Therefore 22.19. will hold for any interpretation of Ω satisfying the Cl_Ω-rule (especially for ω_1^{CK}). So, in the calculus $\mathcal{L}_\infty^{I*} \models_0$ the symbol Ω may be (and should be) viewed as a yet indetermined ordinal constant whose defining axiom is the Cl_Ω-rule. 22.19. then establishes the soundness and correctness of this calculus. By interpreting Ω as ω_1^{CK} in $\mathcal{L}_\infty^{I*} \models_0$ we obtain from 22.16. and 22.19.

22.20. **Corollary**
For any $\mathcal{L}_1$-formula F *we have* $\mathcal{L}_\infty^{I*} \models_0 F$ *if and only if* $\mathcal{L}_\infty^{I} \models_0 F$.

The next step in the ordinal analysis of the formal system $\mathbf{ID}_1$ is the introduction of a semiformal system for the fragment $\mathcal{L}_\infty^I$ of $\mathcal{L}_\infty$. It will be necessary to introduce the Cl_Ω-rule into a semiformal system for $\mathcal{L}_\infty^I$ as it was necessary to introduce the cut into $\mathbf{Z}_\Omega$. (More arguments for the necessity of the Cl_Ω-rule in a semi formal system for $\mathcal{L}_\infty^I$ will be given at the beginning of §26). 22.20 is a semantical proof for the fact that the additional rule Cl_Ω-rule in $\mathcal{L}_\infty^{I*} \models_0$ is eliminable. This resembles the beginning of §12 where we showed the eliminability of the cut rule in the calculus $\mathbf{Z}_\Omega$ by a similar argument. But, as in the situation of $\mathbf{Z}_\Omega$, the semantical proof will not be sufficient for an ordinal analysis of $\mathbf{ID}_1$. In §26 we will therefore give a syntactical proof of the eliminability of the Cl_Ω-rule in a semi formal system.
Before introducing a semiformal system for $\mathcal{L}_\infty^I$ we have to assure that there is a sufficently strong ordinal notation system. It is quite easy to see that the wellordering of the orderrelation of ordertype Γ_0 introduced in §17 may be proved in the formal system $\mathbf{ID}_1$. We will not give the proof now since it will follow as a corollary of a later theorem (29.8). The predicative segment of the ordinals therefore cannot be sufficient for the ordinal analysis of $\mathbf{ID}_1$. In order to obtain a semiformal system in which the provable Π_1^1-sentences of $\mathbf{ID}_1$ may be interpreted we need a larger recursive segment of the ordinals. The development of such a segment will be the aim of the following sections.

22.21. **Exercise**
Let A[X,x] be an X-positive $\mathcal{L}_1$-formula which only contains the indicated free variables and let Δ be a finite set of $\mathcal{L}_\infty^I$-formulas with only positive occurences of Y. Prove the following statements:

(i) $\mathcal{L}_\infty^I \models_0^\alpha \Delta_Y(I_A^{<\eta})$ and $\eta \le \mu \le \Omega \Rightarrow \mathcal{L}_\infty^I \models_0^\alpha \Delta_Y(I_A^{<\mu})$

(ii) $\mathcal{L}_\infty^I \models_0^\alpha \Delta_Y(I_A^{<\eta}) \Rightarrow \mathcal{L}_\infty^I \models_0^\alpha \Delta_Y(I_A^{<\alpha})$

(iii) $\mathcal{L}_\infty^I \models_0^\alpha \underline{n} \in I_A^{<\Omega} \Rightarrow |n|_A < \alpha$

§23. More about ordinals

In §17. we introduced the enumerating function Γ of the strongly critical ordinals and showed that Γ indeed is a normal function. Now it would be an obvious idea to introduce Γ as a new basic function for a notation system. This idea (which as far as I know was performed by Veblen) in fact leads to a notation system which extends Γ_0 but by no means is large enough for an ordinal analysis of $\mathbf{ID}_1$. Any notation system which is sufficient for the ordinal analysis of $\mathbf{ID}_1$ must have some essential impredicative feature. Of course it is impossible to give a precise definition of what we mean by an essential impredicative feature. Roughly speaking one could say that an ordinal notation system has an impredicative feature if it cannot be defined autonomously but its definition needs external points. The external point of the system presented below will be the ordinal Ω. The history of the development of this notation system is quite involved and we are not going into the details of this history. The only facts we want to mention are that the first system of a comparable strength has been introduced by *H.Bachmann* in 1950. The system presented here is an initial segment of a much stronger system which has been developed by *W.Buchholz.* This system on its part is a simplification of the Θ-systems which go back to ideas of *S.Feferman* and have been worked out by *P.Aczel, J.Bridge (Kister)* and *W.Buchholz.*

23.1. **Inductive definition** of the sets $B(\alpha)$ and the function ψ

(B1) $\{0,\Omega\}\subset B(\alpha)$.

(B2) If $\xi,\eta\in B(\alpha)$, then also $\xi+\eta\in B(\alpha)$ and $\varphi\xi\eta\in B(\alpha)$.

(B3) If $\xi\in B(\alpha)\cap\alpha$, then $\psi\xi\in B(\alpha)$.

(ψ1) $\psi\alpha := \min\{\xi : \xi\notin B(\alpha)\}$.

The sets $B(\alpha)$ and therefore also the function ψ are defined by recursion on α. For fixed α the set $B(\alpha)$ is defined inductively. It is easy to see that in this inductive definition every $\xi\in B(\alpha)$ has a finite norm. This shows that there is a 1-1 mapping from $B(\alpha)$ onto ω.

The reader should notice that the definition of the sets $B(\alpha)$ is very simple. In the first step, i.e. in the definition of $B(0)$, we just form a kind of Skolem hull of the ordinals 0 and Ω (as points) and the functions $+$ and φ. Then we denote the first ordinal which does not belong to the segment contained in this Skolem hull by $\psi 0$ and form the Skolem hull of $0,\Omega,\psi 0$ and $+$ and φ. Iterating this process α times leads to the set $B(\alpha)$. Therefore we are going to call $B(\alpha)$ the α*-th iterated Skolem hull* of the ordinals 0 and Ω as points.

23.2. **Lemma**

$\psi\alpha$ is defined for all $\alpha \in \mathrm{On}$ and we always have $\psi\alpha \in (0,\Omega)$.

Proof

By (B1) and (ψ1) we have $0<\psi\alpha$ and $\Omega \neq \psi\alpha$. Since there is a 1-1 mapping from $B(\alpha)$ onto ω and Ω is a regular ordinal we obtain that the set $\{\xi<\Omega : \xi \in B(\alpha)\}$ is bounded in Ω. Therefore there is a $\xi<\Omega$ such that $\xi \notin B(\alpha)$. Hence $\psi\alpha<\Omega$.

23.3. **Lemma**

(i) If $\alpha \le \beta$, then $B(\alpha) \subset B(\beta)$ and $\psi\alpha \le \psi\beta$.

(ii) $\alpha \in B(\beta) \cap \beta$ implies $\psi\alpha < \psi\beta$.

(iii) If $\alpha \le \beta$ and $[\alpha,\beta) \cap B(\alpha) = \emptyset$, then $B(\alpha) = B(\beta)$.

(iv) For $\lambda \in \mathrm{Lim}$ we have $B(\lambda) = \bigcup\{B(\xi) : \xi < \lambda\}$.

Proof

(i) By induction on the definition of $\xi \in B(\alpha)$ we easily obtain that $\alpha \le \beta$ and $\xi \in B(\alpha)$ imply $\xi \in B(\beta)$. Hence $B(\alpha) \subset B(\beta)$ and $\psi\alpha = \min\{\xi : \xi \notin B(\alpha)\} \le \min\{\xi : \xi \notin B(\beta)\} = \psi\beta$.
(ii) By (i) we already have $\psi\alpha \le \psi\beta$. Because of $\alpha \in B(\beta) \cap \beta$ we have $\psi\alpha \in B(\beta)$ by (B3). Hence $\psi\alpha \neq \psi\beta$.
(iii) $B(\alpha) \subset B(\beta)$ is obvious by (i). To prove the opposite direction we show $\xi \in B(\beta) \Rightarrow \xi \in B(\alpha)$ by induction on the definition of $\xi \in B(\beta)$. The cases (B1) and (B2) again are either trivial or immediate consequences of the induction hypothesis. In the case of (B3) there is a $\xi_0 \in B(\beta) \cap \beta$ such that $\xi = \psi\xi_0$. By the induction hypothesis we have $\xi_0 \in B(\alpha)$. Because of $[\alpha,\beta) \cap B(\alpha) = \emptyset$ it is $\xi_0 < \alpha$ and by (B3) we obtain $\xi = \psi\xi_0 \in B(\alpha)$.
(iv) Define $C := \bigcup\{B(\xi) : \xi < \lambda\}$. Then we have $C \subset B(\lambda)$ by (i). For the opposite direction we again show $\xi \in B(\lambda) \Rightarrow \xi \in C$ by induction on the definition of $\xi \in B(\lambda)$. The cases (B1) and (B2) are either trivial or immediate consequences of the induction hypothesis. In the case of (B3) there is a $\xi_0 \in B(\lambda) \cap \lambda$ such that $\xi = \psi\xi_0$. By the induction hypothesis it is $\xi_0 \in C$. Therefore there is a $\rho_0 < \lambda$ such that $\xi_0 \in B(\rho_0)$. Defining $\rho := \max\{\xi_0', \rho_0\}$ we obtain $\xi_0 \in B(\rho) \cap \rho$ and $\rho < \lambda$. By (B3) it follows $\xi \in B(\rho) \subset C$.

It follows from 23.2. and 23.3. that ψ is a monotone function from On into Ω. By a cardinality argument ψ cannot be strictly monotone. We are going to examine the segments on which ψ is strictly monotone.

23.4. **Lemma**

$\beta<\psi\alpha$ *implies* $\beta^\Gamma\subset B(\alpha)$.

Proof

Assume that there is a $\xi<\beta^\Gamma$ such that $\xi\notin B(\alpha)$ and let ξ be minimal with this property. Since $\beta<\psi\alpha$ we then have $\beta<\xi<\beta^\Gamma$. Hence $\xi\notin SC$. But then there are $\xi_1,\xi_2<\xi$ such that $\xi=\xi_1+\xi_2$ or $\xi=\varphi\xi_1\xi_2$. By the minimality of ξ we have $\xi_1,\xi_2\in B(\alpha)$ and obtain $\xi\in B(\alpha)$ by (B2). A contradiction.

23.5. **Theorem**

For all $\alpha\in On$ *we have* $\psi\alpha\in SC$.

Proof

Assume that $\psi\alpha\notin SC$ for some $\alpha\in On$. Then there is a $\beta\in SC$ such that $\beta<\psi\alpha<\beta^\Gamma$. By 23.4, however, $\beta<\psi\alpha$ implies $\beta^\Gamma\subset B(\alpha)$, i.e. $\beta^\Gamma\leq\psi\alpha$. A contradiction.

23.6. **Theorem**

For all $\alpha\in On$ *it is* $B(\alpha)\cap\Omega=\psi\alpha$.

Proof

$\psi\alpha\subset B(\alpha)\cap\Omega$ follows from (ψ1) and 23.2.

For the opposite direction we show $\xi\in B(\alpha)\cap\Omega\Rightarrow\xi<\psi\alpha$ by induction on the definition of $\xi\in B(\alpha)$. In the case of (B1) this follows from 23.2. In the case of (B2) we obtain the claim from the induction hypothesis by 23.5. and $SC\subset\mathbb{H}$. In the case of (B3) we obtain $\xi<\psi\alpha$ by 23.3.(ii).

Since $\psi\alpha$ is not strictly increasing it cannot be a normal function. We shall see, however, that it is at least continuous.

23.7. **Theorem**

For $\lambda\in Lim$ *we have* $\psi\lambda=\sup\{\psi\xi:\xi<\lambda\}$.

Proof

Define $\rho:=\sup\{\psi\xi:\xi<\lambda\}$. Since ψ is monotone we obtain $\rho\leq\psi\lambda$. By 23.3.(iv) we have $B(\lambda)=\bigcup\{B(\xi):\xi<\lambda\}$. For $\eta<\psi\lambda$ we have $\eta\in B(\lambda)\cap\Omega$ by 23.6. Therefore there is a $\xi<\lambda$ such that $\eta\in B(\xi)\cap\Omega$ which implies $\eta<\psi\xi\leq\rho$ by 23.6. Hence $\psi\lambda\leq\rho$.

23.8. **Lemma**

$\psi(\alpha+1) \le \psi(\alpha)^\Gamma$

Proof

It is $\psi(\alpha+1) = B(\alpha+1)\cap\Omega$ and we show $\xi\in B(\alpha+1)\cap\Omega \Rightarrow \xi<\psi(\alpha)^\Gamma$ by induction on the definition of $\xi\in B(\alpha+1)$.

In the case of (B1) we have $\xi = 0$ and the claim is obvious.

If $\xi\in B(\alpha+1)$ holds according to (B2) we have $\xi = \xi_1+\xi_2$ or $\xi = \varphi\xi_1\xi_2$ and $\xi_i\in B(\alpha+1)$ for i = 1,2. But then $\xi_i\le\xi$, and $\xi_i\in B(\alpha+1)\cap\Omega$ and we obtain from the induction hypothesis $\xi_i<(\psi\alpha)^\Gamma$. Since $(\psi\alpha)^\Gamma\in SC$ this also implies $\xi<(\psi\alpha)^\Gamma$.

If $\xi\in B(\alpha+1)$ holds by (B3) then there is a $\xi_0\in B(\alpha+1)\cap(\alpha+1)$ such that $\xi = \psi\xi_0$. By 23.3.(i), however, it follows $\psi\xi_0\le\psi\alpha<(\psi\alpha)^\Gamma$.

23.9. **Lemma**

(i) $\alpha\in B(\alpha+1)$ *implies* $\psi(\alpha+1) = (\psi\alpha)^\Gamma$.

(ii) $\alpha\notin B(\alpha)$ *implies* $B(\alpha+1) = B(\alpha)$ *and therefore also* $\psi(\alpha+1) = \psi\alpha$.

Proof

(i) $\alpha\in B(\alpha+1)$ by 23.3. implies $\psi\alpha<\psi(\alpha+1)$. According to 23.5. it is $\psi(\alpha+1)\in SC$ which entails $(\psi\alpha)^\Gamma\le\psi(\alpha+1)$. Together with 23.8. this implies $(\psi\alpha)^\Gamma = \psi(\alpha+1)$.

(ii) is an immediate consequence of 23.3.(iii).

23.10. **Theorem**

Define $\sigma := \min\{\xi : \Gamma_\xi = \xi\}$. *Then we have*

(i) $\forall\xi\le\sigma(\psi\xi = \Gamma_\xi)$

and

(ii) $\forall\xi\le\Omega(\sigma\le\xi \Rightarrow \psi\xi = \sigma)$.

Proof

We show (i) by induction on $\xi\le\sigma$. If $\xi = 0$, then a comparison of the definitions 17.19. and 23.1. shows that $\psi 0 = B(0)\cap\Omega = PC(0) = \Gamma_0$.

For $\xi = \xi_0+1$ we obtain by the induction hypothesis $\xi_0<\Gamma_{\xi_0} = \psi\xi_0$. Hence $\xi_0\in B(\xi_0) \subset B(\xi)$ and we obtain $\psi(\xi) = (\psi\xi_0)^\Gamma = \Gamma_\xi$ by 23.9.

For $\xi\in Lim$ we have $\psi\xi = \sup\{\psi\eta : \eta<\xi\} \overset{i.h.}{=} \sup\{\Gamma_\eta : \eta<\xi\} = \Gamma_\xi$ since according to 17.15. Γ is a normal function.

To prove (ii) we show $\xi\in[\sigma,\Omega] \Rightarrow B(\xi) = B(\sigma)$ by induction on ξ. The claim is trivial for $\xi = \sigma$. For a limit ordinal ξ we obtain $B(\xi) = \bigcup\{B(\eta) : \eta<\xi\} = B(\sigma)$ by the induction hypothesis. For $\sigma<\xi_0+1<\Omega$ we have $B(\xi_0) = B(\sigma)$ by the induction hypothesis. According to 23.9.(ii) it suffices to show $\xi_0 \notin B(\xi_0)$. From the induction hypothesis and (i) it follows $\psi\xi_0 = B(\xi_0)\cap\Omega = B(\sigma)\cap\Omega = \psi\sigma = \Gamma_\sigma = \sigma$. Hence $\psi\xi_0 = \sigma\leq\xi_0<\Omega$, i.e. $\xi_0 \notin B(\xi_0)$ by 23.6.

Theorem 23.10. shows the role of the ordinal Ω in the definition of the sets $B(\alpha)$. As already mentioned in the introduction Ω is our external point. Without the ordinal Ω in the definition clause (B1) the function ψ would become stationary at σ. Then the effect of definition clause (B3) would be equivalent to augmenting the ordinal notation system of §17 by the function Γ. But because of $\Omega\in B(\Omega+1)$ we now have $\sigma = \psi\Omega\in B(\Omega+1)$ and therefore also $\sigma^\Gamma = \psi(\Omega+1)$. This shows that the segment of the ordinals contained in $B(\alpha)$ is larger than just σ. We are going to examine the size of this segment.

23.11. **Lemma**

For $\alpha\in On$ *we have* $B(\alpha)\subset\Omega^\Gamma$.

Proof

$\xi\in B(\alpha) \Rightarrow \xi<\Omega^\Gamma$ follows by induction on the definition von $\xi\in B(\alpha)$. The claim is obvious in the case of (B1). In the case of (B2) it follows from induction hypothesis since $\Omega^\Gamma\in SC$ and in the case of (B3) the claim holds trivially because of $\psi\xi_0<\Omega$.

23.12. **Theorem**

For $\alpha\in On$ *we have* $B(\alpha)\subset B(\Omega^\Gamma)$ *and* $\psi\alpha\leq\psi(\Omega^\Gamma)$.

Proof

We show $\xi\in B(\alpha) \Rightarrow \xi\in B(\Omega^\Gamma)$ by induction on the definition of $\xi\in B(\alpha)$. The cases (B1) and (B2) do not cause any problem. In the case of (B3) we have $\xi = \psi\xi_0$ for some $\xi_0\in B(\alpha)\cap\alpha$. By the induction hypothesis and 23.11. it follows $\xi_0\in B(\Omega^\Gamma)\cap\Omega^\Gamma$ which implies $\xi = \psi\xi_0\in B(\Omega^\Gamma)$ by (B3).

From theorems 23.12. and 23.6. it follows that $\psi(\Omega^\Gamma)$ is the largest segment of the ordinals accessible by definition 23.1. All ordinals in $B(\Omega^\Gamma)$ are represented by terms built up from the constants $0,\Omega$ by the functions $+$, φ and ψ. In order to see that this really induces a recursive or even primitive recursive notation system we, however, have to work a bit harder.

Up to now we have the following normal forms for ordinals:

1. The Cantor normal form

   $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n :\Leftrightarrow \alpha = \alpha_1 + \ldots + \alpha_n \wedge \{\alpha_1, \ldots, \alpha_n\} \subset \mathbb{H} \wedge \alpha > \alpha_1 \geq \ldots \geq \alpha_n$ for $\alpha \notin \mathbb{H}$.

2. The normal form for principal ordinals which are not strongly critical.

   $\alpha =_{NF} \varphi\alpha_1\alpha_2 \Leftrightarrow \alpha = \varphi\alpha_1\alpha_2 \wedge \alpha > \alpha_1, \alpha_2$ for $\alpha \in \mathbb{H} \setminus SC$.

This was completely sufficient for the ordinals below Γ_0 since there are no strongly critical ordinals. In $B(\Omega^\Gamma)$, however, there are strongly critical ordinals of the form $\psi\xi$. For those ordinals we define a normal form in the following way.

23.13. Definition

$\alpha =_{NF} \psi\xi :\Leftrightarrow \alpha = \psi\xi \wedge \xi \in B(\xi)$

23.14. Lemma

Suppose that $\alpha =_{NF} \psi\alpha_0$ *and* $\beta =_{NF} \psi\beta_0$. *Then we have* $\alpha = \beta \Leftrightarrow \alpha_0 = \beta_0$ *and* $\alpha < \beta \Leftrightarrow \alpha_0 < \beta_0$.

Proof

From $\alpha_0 < \beta_0$ and $\alpha_0 \in B(\alpha_0) \subset B(\beta_0)$ we obtain $\psi\alpha_0 < \psi\beta_0$ by 23.3. On the other hand $\beta_0 \leq \alpha_0$ and $\beta_0 \in B(\beta_0) \subset B(\alpha_0)$ also imply $\psi\beta_0 \leq \psi\alpha_0$.

23.15. Lemma

For every ordinal $\alpha < \Omega^\Gamma$ *there is a uniquely determined ordinal* $\alpha_0 \in B(\alpha_0) \cap \Omega^\Gamma$ *such that* $\psi\alpha =_{NF} \psi\alpha_0$. *We have* $\alpha_0 = \min\{\xi : \alpha \leq \xi \in B(\alpha)\}$.

Proof

By 23.14. the uniqueness of α_0 is obvious. So we just have to prove the existence. In a first step we convince ourselves that the set $\{\xi : \alpha \leq \xi \in B(\alpha)\}$ is not empty. If we define $\Delta_0 = \Omega + 1$ and $\Delta_{n+1} = \varphi\Delta_n 0$, then we have

(1) $\sup\{\Delta_n : n < \omega\} = \Omega^\Gamma$

and

(2) $\Delta_n \in B(\alpha)$ for all $n < \omega$ and $\alpha \in On$.

(1) follows from 17.22. and (2) from the definition of $B(\alpha)$.

By (1) and (2) we obtain the existence of $\alpha_0 := \min\{\xi : \alpha \leq \xi \in B(\alpha)\}$. For α_0 we have $\alpha \leq \alpha_0$ and $[\alpha, \alpha_0) \cap B(\alpha) = \emptyset$. Hence $B(\alpha) = B(\alpha_0)$, and we obtain $\psi\alpha = \psi\alpha_0$ as well as $\alpha_0 \in B(\alpha_0)$.

23.16. Theorem

For every $\alpha \in SC \cap \psi(\Omega^\Gamma)$ *there is a uniquely determined ordinal* $\alpha_0 \in B(\Omega^\Gamma)$ *such that* $\alpha =_{NF} \psi\alpha_0$.

Proof

$\alpha \in SC \cap \psi(\Omega^\Gamma)$ implies $\alpha \in SC \cap B(\Omega^\Gamma) \cap \Omega$. The only possibility for a strongly critical ordinal α to get into the set $B(\Omega^\Gamma) \cap \Omega$ is clause (B3). Therefore there is an $\alpha_1 \in B(\Omega^\Gamma)$ such that $\alpha = \psi\alpha_1$ and by 23.15. there exists an ordinal α_0 such that $\alpha =_{NF} \psi\alpha_0$.

For technical reasons we will define certain subterm sets $P(\alpha)$ and $N(\alpha)$ for ordinals $\alpha \in B(\Omega^\Gamma)$. There are five different types of ordinals in the set $B(\Omega^\Gamma)$:

1. *The ordinal* 0. We define $H(0) = P(0) = N(0) = \emptyset$
2. *Additively decomposable ordinals* $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$ such that $n > 1$. We define $H(\alpha) = \{\alpha_1, \ldots, \alpha_n\}$ and $P(\alpha) = N(\alpha) = \emptyset$
3. *Predicatively decomposable principal ordinals* $\alpha =_{NF} \varphi\alpha_1\alpha_2$. Here we define $P(\alpha) = \{\alpha_1, \alpha_2\}$, $H(\alpha) = \{\alpha\}$ and $N(\alpha) = \emptyset$
4. *Strongly critical ordinals* $\alpha < \Omega$ of the form $\alpha =_{NF} \psi\alpha_0$. Here we define $N(\alpha) = \{\alpha_0\}$ and $H(\alpha) = P(\alpha) = \{\alpha\}$.
5. *The regular ordinal* Ω. For Ω we define $N(\Omega) = P(\Omega) = H(\alpha) = \{\Omega\}$.

We already mentioned that the ordinals in $B(\Omega^\Gamma)$ may be represented by terms built up from 0 and Ω by the functions $+, \varphi$ and ψ. Of course there are different terms which represent the same ordinal. The terms $\psi(\psi\Omega)$ and $\psi\Omega$ for instance are both representations for the ordinal σ. To obtain a unique representation we have to restrict ourselves to ordinal terms in normal form. The set T of terms in normal form is inductively defined by the following definition.

23.17. Inductive definition of the set T of ordinal terms in normal form.

(T1) $\{0, \Omega\} \subset T$.

(T2) If $\alpha_1, \ldots, \alpha_n \in T$ and $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$, then $\alpha \in T$.

(T3) If $\alpha_1, \alpha_2 \in T$ and $\alpha =_{NF} \varphi\alpha_1\alpha_2$, then $\alpha \in T$.

(T4) If $\alpha_0 \in T$ and $\alpha =_{NF} \psi\alpha_0$, then $\alpha \in T$.

The next lemma follows by an easy induction on the definition of T.

23.18. **Lemma**

$T \subset B(\Omega^\Gamma)$

The opposite direction, however, is much harder to prove. The idea is to show

(*) $\xi \in B(\Omega^\Gamma) \Rightarrow \xi \in T$

by induction on the definition of $\xi \in B(\Omega^\Gamma)$.

This induction, however, by no means is straightforward. If, for instance, we have $\xi \in B(\Omega^\Gamma)$ according to clause (B2), then we have $\xi = \xi_1 + \xi_2$ and obtain $\xi_i \in T$ by the induction hypothesis. But we are not allowed to apply clause (T2) since we do not know if we also have $\xi =_{NF} \xi_1 + \xi_2$. To overcome this difficulty we prove (*) by induction on the inductive norm of ξ. Since the closure ordinal of the inductive definition of $B(\alpha)$ obviously is ω we only have to deal with finite norms. If we denote by $B^n(\alpha)$ the n-th stage in the inductive definition of $B(\alpha)$, then we have to show that $\xi =_{NF} \xi_1 + \ldots + \xi_m \in B^n(\alpha)$ already implies $\xi_k \in B^{n-1}(\alpha)$ for $k = 1, \ldots, m$. Then we can use the induction hypothesis and apply (T2). The remaining cases may be treated in the same way. The original inductive definition of the sets $B(\alpha)$, however, is not well suited for this rather technical strategy. Therefore we are going to redefine the sets $B(\alpha)$ in a more technical way.

23.19. **Definition**

The set $B'(\alpha)$ and the function ψ' are inductively defined by the clauses

(B'1) $\{0,1,\Omega\} \subset B'(\alpha)$

(B'2) If $H(\xi) \neq \{\xi\}$ and $H(\xi) \subset B'(\alpha)$, then $\xi \in B'(\alpha)$,

(B'3) If $P(\xi) \neq \{\xi\}$ and $P(\xi) \subset B'(\alpha)$, then $\xi \in B'(\alpha)$,

(B'4) If $\xi \in B'(\alpha) \cap \alpha$, then $\psi'\xi \in B'(\alpha)$.

(ψ'1) $\psi'\alpha := \min\{\xi : \xi \notin B'(\alpha)\}$.

By $B'^n(\alpha)$ we denote the n-th stage in the inductive definition of $B'(\alpha)$. We need some more properties of the stages of $B'(\alpha)$.

23.20. **Lemma**

If $\alpha \in B'^n(\beta)$ *then* $H(\alpha) \cup P(\alpha) \subset B'(\beta)$.

Proof by induction on n.

For $n = 0$ we have $H(\alpha) \cup P(\alpha) \subset \{0,1,\Omega\} \subset B'(\beta)$.

If $\alpha \in B'^n(\beta)$ holds according to (B'2), then we have $H(\alpha) \subset B'^{n-1}(\alpha) \subset B'(\alpha)$ and $P(\alpha) = \emptyset$.
If $\alpha \in B'^n(\beta)$ holds according to (B'3), then we have $H(\alpha) = \{\alpha\} \in B'(\alpha)$ and $P(\alpha) \subset B'^{n-1}(\alpha) \subset B'(\alpha)$.
If $\alpha \in B'^n(\beta)$ holds according to (B'3) then we have $H(\alpha) = P(\alpha) = \{\alpha\}$ and the claim is trivial.

23.21. **Lemma**
For all ordinals α *we have* $B(\alpha) = B'(\alpha)$ as well as $\psi\alpha = \psi'\alpha$.

Proof
The proof is by induction on α.
First we show $\xi \in B(\alpha) \Rightarrow \xi \in B'(\alpha)$ by side induction on the definition of $\xi \in B(\alpha)$.
In the case of a clause (B1) we obtain $\xi \in B'(\alpha)$ by a clause (B'1).
If $\xi \in B(\alpha)$ according to (B2) then we have $\xi = \xi_1 + \xi_2$ or $\xi = \varphi\xi_1\xi_2$ and $\{\xi_1, \xi_2\} \subset B(\alpha)$. By the induction hypothesis it follows $\{\xi_1, \xi_2\} \subset B'(\alpha)$. If $\xi = \varphi\xi_1\xi_2$ this already implies $\xi \in B'(\alpha)$ by (B'3). If $\xi = \xi_1 + \xi_2$, then we have $H(\xi) \subset H(\xi_1) \cup H(\xi_2)$. By 23.20. it follows $H(\xi_1) \cup H(\xi_2) \subset B'(\alpha)$ and this implies $\xi \in B'(\alpha)$ either trivially or by a clause (B'2).
If $\xi \in B(\alpha)$ according to (B2) then we have $\xi = \psi\xi_0$ and $\xi_0 \in B(\alpha) \cap \alpha$. We have $\xi_0 \in B'(\alpha) \cap \alpha$ by the side induction hypothesis and $\psi\xi_0 = \psi'\xi_0$ by the main induction hypothesis. Hence $\xi \in B'(\alpha)$ by a clause (B'3).
For the opposite direction we show $\xi \in B'^n(\alpha) \Rightarrow \xi \in B(\alpha)$ by side induction on n. For $n = 0$ we have $\xi \in \{0, \Omega\}$ and are done by (B1) or $\xi = 1 = \varphi 00$. But $1 \in B(\alpha)$ follows from (B1) by an application of (B2).
If $\xi \in B'^n(\alpha)$ by (B'2) then we have $H(\alpha) \subset B(\alpha)$ by the induction hypothesis and obtain $\xi \in B(\alpha)$ by iterated application of (B2).
If $\xi \in B'^n(\alpha)$ by (B'3) then we have $P(\alpha) \subset B(\alpha)$ by the induction hypothesis and obtain $\xi \in B(\alpha)$ by iterated application of (B2).
If $\xi \in B'^n(\alpha)$ by (B'4) then $\xi = \psi'\xi_0$ and $\xi_0 \in B'^{n-1}(\alpha) \cap \alpha$. By the side induction hypothesis it follows $\xi_0 \in B(\alpha) \cap \alpha$ and by the main induction hypothesis $\psi'\xi_0 = \psi\xi_0$. Hence $\xi \in B(\alpha)$ by a clause (B3).
Now, since we have proven $B(\alpha) = B'(\alpha)$, it also follows $\psi\alpha = \psi'\alpha$.

Due to lemma 23.21. we may identify $B'(\alpha)$ and $B(\alpha)$ as well as ψ and ψ'. We will therefore omit the superscript '. From now on we denote by $B^n(\alpha)$ the n-th stage in the definition of $B'(\alpha)$.
We may sharpen lemma 23.20 in the following way.

23.22. **Lemma**

If $\alpha \in B^n(\beta)$ *and* $\alpha =_{NF} \alpha_1 + \ldots + \alpha_m$ *for* $m>1$ *or* $1<\alpha =_{NF} \varphi\alpha_1\alpha_2$, *then we have* $\alpha_i \in B^{n-1}(\beta)$ *for* $i \in \{1,\ldots,m\}$ *or* $i \in \{1,2\}$ *respectively.*

Proof

Since $\alpha>1$ and $\alpha \notin SC$ the only possibility for $\alpha \in B^n(\beta)$ is by a clause (B2'). Then we have $H(\alpha) \cup P(\alpha) \subset B^{n-1}(\beta)$. For $\alpha =_{NF} \alpha_1 + \ldots + \alpha_m$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2$ we have $H(\alpha) \cup P(\alpha) = \{\alpha_1,\ldots,\alpha_m\}$ or $H(\alpha) \cup P(\alpha) = \{\alpha_1,\alpha_2\}$ respectively which entails the claim.

The proof of the fact that $\alpha =_{NF} \psi\alpha_0 \in B^n(\beta)$ implies $\alpha_0 \in B^{n-1}(\beta)$ is essentially harder although it is easy to see that $\alpha =_{NF} \psi\alpha_0 \in B^n(\beta)$ implies $\alpha_0 \in B(\beta)$. We know that $\psi\alpha_0 < \Omega$, which together with $\psi\alpha_0 \in B(\beta)$ implies $\psi\alpha_0 < \psi\beta$. Hence $\alpha_0 < \beta$ and we obtain $\alpha_0 \in B(\beta)$ since $\alpha =_{NF} \psi\alpha_0$ implies $\alpha_0 \in B(\alpha_0)$. But we do not yet know that α_0 already entered $B(\beta)$ before α. On the other hand we know that there is an $\alpha_1 \in B^{n-1}(\beta)$ such that $\alpha = \psi\alpha_1$ since this is the only way by which α can get into $B^n(\beta)$. Then by 23.15. we have $\alpha_0 = \min\{\xi : \alpha \le \xi \in B(\alpha)\}$.

23.23. **Lemma**

For $\delta \in On$ *define* $\delta(\alpha) := \min\{\xi : \alpha \le \xi \in B(\delta)\}$. *Then* $\alpha \in B^n(\beta)$ *implies* $\delta(\alpha) \in B^n(\beta)$.

Proof

The proof is by induction on n. As a preliminary remark we prove that $\alpha \in \mathbb{H}$ implies $\delta(\alpha) \in \mathbb{H}$ as well as $\alpha \in SC$ implies $\delta(\alpha) \in SC$. If we assume $\alpha \in \mathbb{H}$ and $\delta(\alpha) \notin \mathbb{H}$ then there is an $\eta \in [\alpha,\delta(\alpha)) \cap H(\delta(\alpha))$. Then by 23.22. it follows $\eta \in B(\delta)$ in contradiction to the minimality of $\delta(\alpha)$. The proof for SC runs completely analogously. Trivially $\alpha_1 < \alpha_2$ always implies $\delta(\alpha_1) \le \delta(\alpha_2)$.

If $\alpha \in B(\delta)$, then it is $\delta(\alpha) = \alpha$ and the claim is obvious. If $\alpha \notin B(\delta)$ and $\alpha < \Omega$, then we have by 23.6. $B(\delta) \cap \Omega \subset \alpha$ which implies $\delta(\alpha) = \Omega \in B^n(\beta)$ for all n. So it remains the case $\Omega < \alpha \notin B(\delta)$. We distinguish the following subcases:

1. $\alpha =_{NF} \alpha_1 + \ldots + \alpha_m$ for some $m>1$. Then we have $\alpha_i \in B^{n-1}(\beta)$ by 23.22. and obtain by the induction hypothesis $\delta(\alpha_i) \in B^{n-1}(\beta) \cap B(\delta)$. Since $\alpha_i \in \mathbb{H}$ and $\alpha_1 \ge \ldots \ge \alpha_m$ imply $\delta(\alpha_i) \in \mathbb{H}$ and $\delta(\alpha_1) \ge \ldots \ge \delta(\alpha_m)$, it is $\alpha \le \delta(\alpha_1) + \ldots + \delta(\alpha_m)$ and because of $\delta(\alpha_1) + \ldots + \delta(\alpha_m) \in B(\delta)$ even $\alpha < \delta(\alpha_1) + \ldots + \delta(\alpha_m)$. Therefore there is an $i<m$ such that $\gamma := \alpha_1 + \ldots + \alpha_i = \delta(\alpha_1) + \ldots + \delta(\alpha_i)$ and $\alpha_{i+1} < \delta(\alpha_{i+1})$. This implies $\alpha < \gamma + \delta(\alpha_{i+1})$. Hence $\delta(\alpha) \le \gamma + \delta(\alpha_{i+1})$. We claim $\delta(\alpha) = \gamma + \delta(\alpha_{i+1})$. If we assume $\delta(\alpha) < \gamma + \delta(\alpha_{i+1})$, then we obtain an $\varepsilon \in B(\delta)$ such that $\alpha < \varepsilon < \gamma + \delta(\alpha_{i+1})$. But then $\gamma + \alpha_{i+1} \le \alpha < \varepsilon < \gamma + \delta(\alpha_{i+1})$ and we obtain an ε_1 such that $\varepsilon = \gamma + \varepsilon_1$ and $\alpha_{i+1} < \varepsilon_1 < \delta(\alpha_{i+1})$. Since $\varepsilon \in B(\delta)$, 23.20.

and (B'2) imply $\varepsilon_1 \in B(\delta)$ and this contradicts the definition of $\delta(\alpha_{i+1})$.

2. $\alpha =_{NF} \varphi\alpha_1\alpha_2$. By 23.22. we have $\alpha_i \in B^{n-1}(\beta)$ and obtain by the induction hypothesis $\delta(\alpha_i) \in B^{n-1}(\beta)$. There are the following cases:

2.1. $\alpha_1 = \delta(\alpha_1)$. Then we obviously have $\delta(\alpha) = \varphi\alpha_1\delta(\alpha_2)$ and it follows $\delta(\alpha) \in B^n(\beta)$.

2.2. $\alpha_1 < \delta(\alpha_1)$. If $\alpha \le \delta(\alpha_2)$, then $\delta(\alpha) \le \delta(\alpha_2)$ which together with $\delta(\alpha_2) \le \delta(\alpha)$ implies $\delta(\alpha) = \delta(\alpha_2)$ and we are done because of $\delta(\alpha_2) \in B^{n-1}(\beta) \subset B^n(\beta)$.

Therefore we assume $\delta(\alpha_2) < \alpha$. We define $\alpha_3 := \min\{\xi : \alpha \le \varphi\delta(\alpha_1)\xi\}$ and claim

(*) $\alpha_3 \in B^{n-1}(\beta) \cap B(\delta)$

From (*) it follows $\delta(\alpha) \le \varphi\delta(\alpha_1)\alpha_3$. If we assume $\varphi\alpha_1\alpha_2 < \delta(\alpha) < \varphi\delta(\alpha_1)\alpha_3$, then there are ordinals ξ_1 and ξ_2 such that $\delta(\alpha) =_{NF} \varphi\xi_1\xi_2$. But then we have $\xi_1 < \delta(\alpha_1)$, since $\xi_1 = \delta(\alpha_1)$ implies $\alpha < \delta(\alpha) = \varphi\delta(\alpha_1)\,\xi_2$ and $\xi_2 < \alpha_3$, and $\xi_1 > \delta(\alpha_1)$ implies $\delta(\alpha) < \alpha_3$ and $\varphi\,\delta(\alpha_1)\delta(\alpha) = \delta(\alpha) > \alpha$. Both consequences contradict the definition of α_3. From $\xi_1 < \delta(\alpha_1)$ and $\xi_1 \in B(\delta)$, however, we obtain $\xi_1 < \alpha_1$ which implies $\alpha < \xi_2 \in B(\delta)$ in contradiction to the definition of $\delta(\alpha)$. Hence $\delta(\alpha) = \varphi\delta(\alpha_1)\alpha_3$ which implies $\delta(\alpha) \in B^n(\beta)$ by (B3').

It remains to show (*). If we assume $\alpha_3 \in \mathrm{Lim}$, then we obtain $\varphi\alpha_1\alpha_2 = \alpha = \varphi\delta(\alpha_1)\alpha_3$ because $\varphi\delta(\alpha_1)$ is continuous. Since $\alpha_1 < \delta(\alpha_1)$ this implies $\alpha_2 = \varphi\delta(\alpha_1)\alpha_3$ in contradiction to $\alpha_2 \le \delta(\alpha_2) < \alpha$. So α_3 cannot be a limit ordinal. If $\alpha_3 = 0$, then we are done. Therefore assume $\alpha_3 = \eta'$. Then we have $\alpha_2 \le \delta(\alpha_2) < \alpha = \varphi\alpha_1\alpha_2 \le \varphi\alpha_1\eta'$ and $\varphi\delta(\alpha_1)\eta < \varphi\alpha_1\alpha_2 \le \varphi\delta(\alpha_1)\eta'$. Because of $\alpha_1 < \delta(\alpha_1)$ it follows $\varphi\delta(\alpha_1)\eta < \alpha_2 < \varphi\,\alpha_1\alpha_2$, i.e. $\varphi\delta(\alpha_1)\eta < \alpha_2 \le \delta(\alpha_2) < \varphi\alpha_1\alpha_2 \le \varphi\delta(\alpha_1)\eta'$. Since $\delta(\alpha_2) \in B(\delta) \cap B^{n-1}(\beta)$ we have $B^{n-1}(\beta) \cap (\varphi\delta(\alpha_1)\eta, \varphi\delta(\alpha_1)(\eta+1)) \ne \emptyset$ as well as $B(\delta) \cap (\varphi\delta(\alpha_1)\eta, \varphi\delta(\alpha_1)(\eta+1)) \ne \emptyset$. Now we prove the following auxiliary lemma.

23.24. **Lemma**

If $B^n(\beta) \cap (\varphi\alpha\eta, \varphi\alpha\eta') \ne \emptyset$, *then we have* $\eta+1 \in B^n(\beta)$.

Proof by induction on n.

Define $M^n := B^n(\beta) \cap (\varphi\alpha\eta, \varphi\alpha\eta')$ and assume $\sigma \in M^n$. Since $(\varphi\alpha\eta, \varphi\alpha(\eta+1)) \cap SC = \emptyset$ $\sigma \notin SC$ the only possibility for σ to come into $B^n(\beta)$ therefore is clause (B2') or (B'3).

If $\sigma =_{NF} \sigma_1 + \ldots + \sigma_m$ such that $\sigma_i \in B^{n-1}(\beta)$, then we have $\sigma_1 \in [\varphi\alpha\eta, \varphi\alpha\eta')$ because otherwise we had $\sigma_1 + \ldots + \sigma_m < \varphi\alpha\eta$. If $\varphi\alpha\eta < \sigma_1$, then we obtain $\eta+1 \in B^{n-1}(\beta) \subset B^n(\beta)$ by the induction hypothesis. If $\varphi\alpha\eta = \sigma_1$, then we have $\eta \in B^{n-1}(\beta)$. Since $1 \in B^m(\beta)$ for all $m < \omega$ we obtain $\eta' \in B^n(\beta)$ by (B'2).

If we assume $\sigma =_{NF} \varphi\sigma_1\sigma_2$ and $\sigma_i \in B^{n-1}(\beta)$, then we have $\varphi\alpha\eta < \varphi\sigma_1\sigma_2 < \varphi\alpha\eta'$. But then it is $\sigma_1 < \alpha$ since $\alpha = \sigma_1$ implies $\eta < \sigma_2 < \eta'$ and $\alpha < \sigma_1$ already $\eta < \sigma < \eta'$. Hence

$\varphi\alpha\eta < \sigma_2 < \eta' \leq \varphi\alpha\eta'$ and we obtain by the induction hypothesis $\eta' \in B^{n-1}(\beta) \subset B^n(\beta)$. By the proof of lemma 23.24. the proof of lemma 23.23. is completed too.

23.25. **Lemma**

If $\alpha =_{NF} \psi\alpha_0 \in B^n(\beta)$, *then we have* $n>0$ *and* $\alpha_0 \in B^{n-1}(\beta) \cap \beta$.

Proof

In the remark foregoing 23.23. we already mentioned that if $\alpha =_{NF} \psi\alpha_0 \in B^n(\beta)$, then it is $\alpha_0 < \beta$ and there is an $\alpha_1 \in B^{n-1}(\beta)$ such that $\alpha = \psi\alpha_1$. Then we have $\alpha_0 = \min\{\xi : \alpha_1 \leq \xi \in B(\alpha_1)\}$ and it follows $\alpha_0 \in B^{n-1}(\beta)$ by 23.23.

23.26. **Theorem**

$T = B(\Omega^\Gamma)$

Proof

$T \subset B(\Omega^\Gamma)$ holds by 23.18. For the opposite direction we show $B^n(\Omega^\Gamma) \subset T$ by induction on n. $B^0(\Omega^\Gamma)$ only contains the ordinals $0, 1, \Omega$. $\{0, \Omega\} \subset T$ holds by (T1) $1 \in T$ follows from (T1) by (T3). If $\alpha =_{NF} \alpha_1 + \ldots + \alpha_m \in B^{n+1}(\Omega^\Gamma)$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2 \in B^{n+1}(\Omega^\Gamma)$, then by 23.22. we obtain $\alpha_i \in B^n(\Omega^\Gamma)$ for all $i \in \{1, \ldots, m\}$ or $i = 1, 2$, respectively.
By the induction hypothesis we have $\alpha_i \in T$ for all i and obtain $\alpha \in T$ by (T2) or (T3). If $\alpha =_{NF} \psi\alpha_0$, then we have $\alpha_0 \in B^{n-1}(\beta)$ by 23.25. and obtain $\alpha \in T$ by the induction hypothesis and (T4).

It is of course easy to define codes for the elements of T. Therefore we may regard T as a set of natural numbers. Our aim is to obtain T as primitive recursive set. This, however, in not an immediate consequence of definition 23.17. The stumblingblocks are the normal form requirements in the premisses of clauses (T2)-(T4). This is not so harmful in the case of the clauses (T2) and (T3) since the normal form conditions there only need checking the $<$-relation between formerly defined terms and we may define T and the $<$-relation in the approved manner by simultaneous course of value recursion.
More irritating is the case of a clause (T4) since there $\alpha =_{NF} \psi\alpha_0$ means that $\alpha = \psi\alpha_0 \wedge \alpha_0 \in B(\alpha_0)$. We could manage this case in the same way if we succeeded in defining a primitive recursive function K, say, such that $\alpha_0 \in B(\alpha_0)$ holds if and only if $K\alpha_0 < \alpha_0$. Then again we could define T and $<$ (and possibly also K) by simultaneous course of values recursion. We will not be able to define

$K\alpha_0$ as a function that takes a single ordinal term as value but as a finite set of subterms of α_0 which will be primitive recursively computable from T. The *degree* $G\alpha$ of an ordinal term $\alpha \in T$ is the norm of α in the inductive definition of T. In order to define $K\alpha$ we recapitulate the conditions which are necessary and sufficient for $\alpha \in B(\beta)$.

23.27. **Lemma**

$\alpha \in B(\beta)$ holds if and only if one of the following conditions is satisfied:

(1) $\alpha \in \{0, \Omega\}$

(2) $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$ *and* $\{\alpha_1, \ldots, \alpha_n\} \subset B(\beta)$,

(3) $\alpha =_{NF} \varphi\alpha_1\alpha_2$ *and* $\{\alpha_1, \alpha_2\} \subset B(\beta)$,

(4) $\alpha = \psi\alpha_0$, $\alpha_0 < \beta$ *and* $\alpha_0 \in B(\beta)$.

According to 23.27. we define the finite set $K\alpha$ of subterms of an ordinal term α by:

23.28. **Inductive definition** of $K\alpha$

(K1) $K0 = K\Omega = \emptyset$

(K2) $K\alpha = K\alpha_1 \cup \ldots \cup K\alpha_n$, if $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$

(K3) $K\alpha = K\alpha_1 \cup K\alpha_2$, if $\alpha =_{NF} \varphi\alpha_1\alpha_2$

(K4) $K\alpha = \{\alpha_0\} \cup K\alpha_0$, if $\alpha = \psi\alpha_0$

We call $K\alpha$ the *set of components* of α and write $K\alpha < \beta$ instead of $\forall\xi \in K\alpha(\xi < \beta)$.

23.29. **Lemma**

It is $\alpha \in B(\beta)$ *if and only if* $K\alpha < \beta$.

Proof

This follows immediately from 23.27. by induction on $G\alpha$.

23.30. **Lemma**

For $\alpha, \beta \in T$ *we have* $\alpha < \beta$ *if and only if one of the following conditions is satisfied:*

(1) $\alpha = 0$ *and* $\beta \neq 0$,

(2) $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$, $\beta =_{NF} \beta_1 + \ldots + \beta_m$, n, m > 1 *and* $\exists i \leq n\ \forall j \leq i(\alpha_j = \beta_j \wedge \alpha_{i+1} < \beta_{i+1})$, *where we define* $\alpha_{n+1} = 0$ and $\beta_{m+1} = \ldots = \beta_{n+1} = 0$ if $m < n$.

(3) $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$, $\beta \in \mathbb{H}$ *and* $\alpha_1 < \beta$,

(4) $\alpha \in \mathbb{H}$, $\beta =_{NF} \beta_1 + \ldots + \beta_n$, $n > 1$ *and* $\alpha \leq \beta_1$,

(5) $\alpha =_{NF} \varphi\alpha_1\alpha_2$, $\beta =_{NF} \varphi\beta_1\beta_2$ *and* $\alpha < \beta$ *according to* 17.8.(2),

(6) $\alpha =_{NF} \varphi\alpha_1\alpha_2$, $\beta \in SC$ *and* $\alpha_1, \alpha_2 < \beta$,

(7) $\alpha \in SC$, $\beta =_{NF} \varphi\beta_1\beta_2$ *and* $\alpha \leq \beta_1$ *or* $\alpha \leq \beta_2$,

(8) $\alpha =_{NF} \psi\alpha_0, \beta =_{NF} \psi\beta_0$ *and* $\alpha_0 < \beta_0$,

(9) $\alpha =_{NF} \psi\alpha_0$ *and* $\beta = \Omega$.

It is obvious that $\alpha \in \mathbb{H}$ or $\alpha \in SC$, respectively, are primitive recursively decidable since this can be read off the syntactical form of α. All terms in $\mathbb{H}$ are of the shape $\varphi\alpha_1\alpha_2$, while all terms in SC have the form $\psi\alpha_0$ or Ω.

Now we are going to replace the clause (T4) in 23.17. by

(T4') $\alpha_0 \in T \wedge K\alpha_0 < \alpha_0 \Rightarrow \psi\alpha_0 \in T$,

and define simultaneously the set T, the 'function' $K\alpha$ according to 23.28. and for $\alpha, \beta \in T$ the relation $\alpha < \beta$ according to 23.30. by course of values recursion. Then we obtain:

23.31. **Theorem**

The set T *is primitive recursive and* $<$ *is a primitive recursive order relation on* T.

23.32. **Corollary**

$\psi(\Omega^\Gamma) < \omega_1^{CK}$

23.33. **Exercises**

1. $\forall\alpha,\beta\,(\beta \in B(\alpha) \Rightarrow \Omega\cdot\beta \in B(\alpha))$

2. Define $\sigma := \min\{\rho : \rho > \psi(\Omega\cdot\xi) \wedge \Gamma_\rho = \rho\}$. Prove that under the assumptions $\Omega\cdot\xi \in B(\Omega\cdot\xi)$ and $\psi(\Omega\cdot\xi) = \Gamma_{\psi(\Omega\cdot\xi)}$ we have:
   (i) $\forall\delta \leq \sigma\ \psi(\Omega\cdot\xi+\delta) = \Gamma_{\psi(\Omega\cdot\xi)+\delta}$
   (ii) $\forall\delta(\sigma \leq \delta \leq \Omega \Rightarrow \psi(\Omega\cdot\xi+\delta) = \sigma)$

3. Show the following statements:
   (i) $\forall\xi < \Gamma''0\ (\Gamma'\xi = \psi(\Omega\cdot(1+\xi)) \wedge \Omega\cdot\xi \in B(\Omega\cdot\xi))$
   (ii) $\psi\,\Omega^2 = \Gamma''0$
   (iii) $\psi(\Omega^2+1) = \Gamma''0^\Gamma$

4. Prove that for $\alpha \in B(\Omega^\Gamma)$ we have:
   (i) $\alpha < \Gamma_0 \Leftrightarrow \alpha < \Omega \wedge K\alpha = \emptyset$
   (ii) $\alpha < \psi\Omega \Leftrightarrow K\alpha \cup \{\alpha\} < \Omega$

§24. Collapsing functions

Our goal is to use the ordinal terms of T for the introduction of a semiformal system for the language $\mathscr{L}^1_\Omega$.
The derivation trees of $\mathscr{L}^1_\Omega$ are in general Ω-branching trees. Therefore we will have to measure the length of the derivation trees of a semiformal system for $\mathscr{L}^1_\Omega$ by ordinals above Ω. In order to obtain bounds for the norms of Π^1_1-formulas, which as we know are ordinals below Ω, we need a collapsing function, which allows us to collapse the ordinals in T which are larger than Ω into the ordinals of T below Ω. The development of this collapsing function is the topic of the current section. Since by 23.2. it is $\psi\alpha<\Omega$ we in principle already have a collapsing function. The set T of ordinal terms in normal form, however, is not closed under the function ψ. In order to obtain $\psi\alpha\in T$ we need $\alpha\in B(\alpha)$ which is wrong in general. Therefore we will carefully enlarge α to an ordinal term $h\alpha$ such that $h\alpha\in B(h\alpha)$ is always true and, vaguely speaking, $h\alpha$ essentially carries the same information as α. After having succeeded in doing this we may collapse the ordinals terms greater or equal than Ω by the collapsing function $D\alpha := \psi(h\alpha)$.
We start with a rather technical lemma.

24.1. Lemma

(i) $\alpha\in K\beta$ *implies* $K\alpha\subset K\beta$,

(ii) If $\alpha\in K\beta$, *then it is* $G\alpha<G\beta$,

(iii) $\alpha\notin K\alpha$.

Proof

The statements in (i) and (ii) are proved by induction on $G\beta$. If $\beta=0$ or $\beta=\Omega$, then we have $K\beta=\emptyset$ and both claims are trivial.

If $\beta =_{NF} \beta_1+\ldots+\beta_n$ and $\alpha\in K\beta$, then it is $\alpha\in K\beta_i$ for some $i\in\{1,\ldots,n\}$. By the induction hypothesis for (i) we obtain $K\alpha\subset K\beta_i\subset K\beta$ which proves (i). By the induction hypothesis for (ii) it follows $G\alpha<G\beta_i<G\beta$ which also proves (ii).

If $\beta =_{NF} \varphi\beta_1\beta_2$ and $\alpha\in K\beta$, so then we obtain $\alpha\in K\beta_i$ for some $i\in\{1,2\}$ and we have the same proof as above.

If $\beta =_{NF} \psi\beta_0$ and $\alpha\in K\beta$, then we have $\alpha=\beta_0$ or $\alpha\in K\beta_0$. In the first case it follows $K\alpha = K\beta_0\subset K\beta$ and $G\alpha = G\beta_0<G\beta$ and in the second case with the corresponding induction hypothesis $K\alpha\subset K\beta_0\subset K\beta$ or $G\alpha<G\beta_0<G\beta$.

(iii) is an immediate consequence of (ii).

24.2. **Definition**

$k\alpha := \max K\alpha \cup \{0\}$

24.3. **Lemma**

$Kk\alpha < k\alpha$

Proof

We either have $K\alpha = \emptyset$, and therefore $Kk\alpha = K0 = \emptyset$, or $k\alpha \in K\alpha$. In the latter case we obtain $Kk\alpha \subset K\alpha$, i.e. $Kk\alpha \leq k\alpha$ by 24.1.(i). Since we have $k\alpha \notin Kk\alpha$ by 24.1.(iii) it follows $Kk\alpha < k\alpha$.

In the formation of $k\alpha$, however, we loose too much information about α. So we cannot define $h\alpha$ to be just $k\alpha$. To keep all the information about α we define:

24.4. **Definition**

$h\alpha := k\alpha + \omega^\alpha$

24.5. **Lemma**

$K\alpha = Kh\alpha < h\alpha$

Proof

It is $Kh\alpha = Kk\alpha \cup K\alpha = K\alpha \leq k\alpha < k\alpha + \omega^\alpha = h\alpha$.

24.6. **Definition**

If $\alpha =_{NF} \alpha_1 + ... + \alpha_n$, then we define $r(\alpha) := \alpha_n$. We call $r(\alpha)$ the *(additive) remainder* of α.

24.7. **Lemma**

If $\alpha < \gamma$ *and* $\beta < r(\gamma)$, *then it follows* $\alpha + \beta < \gamma$.

Proof

If $\alpha =_{NF} \alpha_1 + ... + \alpha_n$ and $\gamma =_{NF} \gamma_1 + ... + \gamma_m$, then there is an $i \leq n$ such that $\alpha_j = \gamma_j$ for all $j \leq i$ and $\alpha_{i+1} < \gamma_{i+1}$. If $\beta < r(\gamma) \leq \gamma_{i+1}$, then we obtain $\alpha_{i+1} + ... + \alpha_n < \gamma_{i+1}$ which implies $\alpha + \beta = \alpha_1 + ... + \alpha_n + \beta < \alpha_1 + ... + \alpha_n + \beta + \gamma_{i+1} + ... + \gamma_m = \alpha_1 + ... + \alpha_i + \gamma_{i+1} + ... + \gamma_m = \gamma$.

24.8. Lemma

$\alpha<\beta$ *and* $K\alpha<h\beta$ *imply* $h\alpha<h\beta$.

Proof

$\alpha<\beta$ implies $\omega^{\alpha}<\omega^{\beta}$. $K\alpha<h\beta$ implies $k\alpha<h\beta=k\beta+\omega^{\beta}$. Hence $\omega^{\alpha}<r(h\beta)$. By 24.7. this entails $k\alpha+\omega^{\alpha}<k\beta+\omega^{\beta}$.

24.9. Inductive definition of the set $SC(\alpha)$ of the *strongly critical subterms* of an ordinal term α.

(i) $SC(0):=\emptyset$ and $SC(\Omega):=\{\Omega\}$

(ii) If $\alpha\notin\mathbb{H}$, then we define $SC(\alpha):=\bigcup\{SC(\gamma)\colon\gamma\in H(\alpha)\}$

(iii) If $\alpha\notin SC$, then we define $SC(\alpha):=\bigcup\{SC(\gamma)\colon\gamma\in P(\alpha)\}$

(iv) If $\alpha\in SC$, then $SC(\alpha):=\{\alpha\}$

By $SC_{\Omega}(\alpha)$ we denote the set $SC(\alpha)\cap\Omega$.

As an immediate consequence we obtain by induction on $G\alpha$:

24.10. Lemma

We have $\alpha\in B(\beta)\Leftrightarrow SC(\alpha)\subset B(\beta)$. *Since* $SC(\alpha)\subset SC_{\Omega}(\alpha)\cup\{\Omega\}$ *and we always have* $\Omega\in B(\alpha)$ *this may be sharpened to* $\alpha\in B(\beta)\Leftrightarrow SC_{\Omega}(\alpha)\subset B(\beta)$

24.11. Lemma

We have $K\alpha<h\beta$ *if and only if* $SC_{\Omega}(\alpha)<\psi h\beta$.

Proof

By 23.29. we have $K\alpha<h\beta$ if and only if $\alpha\in B(h\beta)$. By 24.10. this is equivalent to $SC_{\Omega}(\alpha)\subset B(h\beta)\cap\Omega=\psi h\beta$.

24.12. Definition

(i) $$D\alpha=\begin{cases}\alpha, & \text{if}\quad \alpha<\Omega\\ \psi h\alpha & \text{if}\quad \Omega\le\alpha,\end{cases}$$

(ii) $\alpha\ll\beta\;:\Leftrightarrow\;\alpha<\beta\wedge D\alpha<D\beta$.

We call the function $D\colon B(\Omega^{\Gamma})\longrightarrow\Omega$ the *collapsing function* for $B(\Omega^{\Gamma})$.

We read the relation $\alpha\ll\beta$ as 'α *is essentially less than* β'.

24.13. **Lemma**

(i) For $\alpha \in T$ *it is* $D\alpha \in T$ *and* $D\alpha < \Omega$.

(ii) For $\beta \in T \cap \Omega$ *we have* $\alpha \ll \beta$ *if and only if* $\alpha < \beta$.

(iii) $\alpha \ll \beta$ *implies* $D\alpha \ll D\beta$.

24.14. **Lemma**

(i) $SC_\Omega(\alpha) \leq D\alpha$.

(ii) $\alpha \notin SC$ *or* $\Omega \leq \alpha$ *imply* $SC_\Omega(\alpha) < D\alpha$.

Proof

If $\alpha < \Omega$, then we have $SC_\Omega(\alpha) \leq \alpha = D\alpha$. If $\alpha \notin SC$, then we even obtain $SC_\Omega(\alpha) < \alpha = D\alpha$. Therefore assume $\Omega \leq \alpha$. Because of $k\alpha < h\alpha$ we always have $\alpha \in B(h\alpha)$. By 24.10. this implies $SC_\Omega(\alpha) \subset B(h\alpha) \cap \Omega = \psi h\alpha$, i.e. $SC_\Omega(\alpha) < D(\alpha)$.

24.15. **Theorem** (Characterization of the $\ll$-relation)

It holds $\alpha \ll \beta$ *if and only if we have* $\alpha < \beta$ *and* $SC_\Omega(\alpha) < D\beta$.

Proof

We start with the direction from left to right. Assume $\alpha \ll \beta$. This implies $\alpha < \beta$. If $\beta < \Omega$, then we immediately obtain $SC_\Omega(\alpha) \leq \alpha < \beta = D\beta$. If $\Omega \leq \beta$ and $\alpha < \Omega$, then it follows $SC_\Omega(\alpha) \leq \alpha = D\alpha < D\beta$. If finally $\Omega \leq \alpha < \beta$, then $\psi h\alpha = D\alpha < D\beta = \psi h\beta$ already implies $h\alpha < h\beta$. Because of $K\alpha < h\alpha < h\beta$ it follows by 24.11. $SC_\Omega(\alpha) < \psi h\beta = D\beta$.

For the opposite direction we assume $\alpha < \beta$ and $SC_\Omega(\alpha) < D\beta$. If $\beta < \Omega$, then $\alpha \ll \beta$ already follows from $\alpha < \beta$. So assume $\Omega \leq \beta$. Then we have $D\beta = \psi h\beta \in SC$. If $\alpha < \Omega$, then $SC_\Omega(\alpha) < D\beta \in SC$ immediately implies $\alpha < D\beta$. Therefore assume $\Omega \leq \alpha < \beta$. Since $SC_\Omega(\alpha) < \psi h\beta$ we have $SC_\Omega(\alpha) \subset B(h\beta)$ which by 24.10. and 23.29. implies $K\alpha < h\beta$. By 24.8. we then obtain $h\alpha < h\beta$. Hence $D\alpha = \psi h\alpha < \psi h\beta = D\beta$.

24.16. **Lemma**

(i) $\alpha \ll \beta \ll \gamma$ *imply* $\alpha \ll \gamma$.

(ii) It is $\alpha \lesssim \Omega + \alpha$ *and* $\beta \neq 0$ *implies* $\alpha \ll \alpha \# \beta$.

(iii) If $\alpha_i < \varphi\alpha_1\alpha_2$, *then we have* $\alpha_i \ll \varphi\alpha_1\alpha_2$ *for* $i = 1,2$.

(iv) If $\alpha \ll \beta$, *then* $\alpha \# \gamma \ll \beta \# \gamma$ *and* $\Omega + \alpha \ll \Omega + \beta$.

(v) If $\alpha \ll \beta$ *and* $\rho \notin SC$, *then* $\varphi\rho\alpha \ll \varphi\rho\beta$.

(vi) $\alpha_1 \ll \alpha_2$ *and* $\beta_1 \ll \varphi\alpha_2\beta_2$ *imply* $\varphi\alpha_1\beta_1 \ll \varphi\alpha_2\beta_2$.

(vii) If $\alpha_1, \alpha_2 \ll \beta$ *and* $\beta \in \mathbb{H}$, *then* $\alpha_1 \# \alpha_2 \ll \beta$.

(viii) $D\alpha \ll \alpha \# \Omega$.

Proof

(i) is obvious.

(ii) $\alpha \le \Omega + \alpha$ is again obvious. By 24.14. it is $SC_\Omega(\alpha) = SC_\Omega(\Omega+\alpha) \le D(\Omega+\alpha)$. Hence $\alpha \ll \Omega + \alpha$.

$\beta \ne 0$ implies $\alpha < \alpha \# \beta$. If $\alpha = 0$, then we obtain the claim because $SC_\Omega(0) = \emptyset$. Otherwise we have $\alpha \# \beta \notin SC$ and obtain $SC_\Omega(\alpha) \subset SC_\Omega(\alpha \# \beta) < D(\alpha \# \beta)$ by 24.14. Hence $\alpha \ll \alpha \# \beta$ by 24.15.

(iii) It suffices to show $SC_\Omega(\alpha_i) < D\varphi\alpha_1\alpha_2$. If $\varphi\alpha_1\alpha_2 \notin SC$, then we obtain $SC_\Omega(\alpha_i) \subset SC_\Omega(\varphi\alpha_1\alpha_2) < D\varphi\alpha_1\alpha_2$. If $\varphi\alpha_1\alpha_2 \in SC$, then it follows $SC(\alpha_i) \le \alpha_i < \{\varphi\alpha_1\alpha_2\} = SC_\Omega(\varphi\alpha_1\alpha_2) \le D\varphi\alpha_1\alpha_2$.

(iv) $\alpha \# \gamma < \beta \# \gamma$ follows from $\alpha < \beta$. We have $SC_\Omega(\alpha \# \gamma) = SC_\Omega(\alpha) \cup SC_\Omega(\gamma)$. $\alpha \ll \beta$ implies $SC_\Omega(\alpha) < D\beta < D(\beta \# \gamma)$ by (ii) and, since $\beta \ne 0$, also $SC_\Omega(\gamma) \le D\gamma < D(\beta \# \gamma)$. Hence $SC_\Omega(\alpha \# \gamma) < D(\beta \# \gamma)$ and it follows $\alpha \# \gamma \ll \beta \# \gamma$. This proves the first part of (iv). For the second part we have $\Omega + \alpha < \Omega + \beta$ since $\alpha < \beta$ and obtain by (ii) $SC_\Omega(\Omega+\alpha) = SC_\Omega(\alpha) < D\beta \le D(\Omega+\beta)$. Hence $\Omega + \alpha \ll \Omega + \beta$.

(v) From $\alpha < \beta$ we obtain $\varphi\rho\alpha < \varphi\rho\beta$. It is $SC_\Omega(\varphi\rho\alpha) = SC_\Omega(\rho) \cup SC_\Omega(\alpha)$. Since $\alpha \ll \beta$ we have $SC_\Omega(\alpha) < D\beta \le D(\varphi\rho\beta)$ and $\rho \notin SC$ implies $SC_\Omega(\rho) < D\rho \le D(\varphi\rho\beta)$.

(vi) $\varphi\alpha_1\beta_1 < \varphi\alpha_2\beta_2$ follows from 17.8. It is $SC_\Omega(\varphi\alpha_1\beta_1) = SC_\Omega(\alpha_1) \cup SC_\Omega(\beta_1)$. By $\alpha_1 \ll \alpha_2$ we have $SC_\Omega(\alpha_1) < D\alpha_2 \le D(\varphi\alpha_2\beta_2)$ and by $\beta_1 \ll \varphi\alpha_2\beta_2$ it follows $SC_\Omega(\beta_1) < D(\varphi\alpha_2\beta_2)$. Pulling these results together we obtain $SC_\Omega(\varphi\alpha_1\beta_1) < D(\varphi\alpha_2\beta_2)$ and it follows $\varphi\alpha_1\beta_1 \ll \varphi\alpha_2\beta_2$.

It should be noted that for the proof of (vi) it would suffice to have the assumption $\alpha_1 < \alpha_2$ and $SC_\Omega(\alpha_1) < D(\varphi\alpha_2\beta_2)$ instead of $\alpha_1 \ll \alpha_2$.

(vii) $\alpha_1 \# \alpha_2 < \beta$ follows from $\alpha_1, \alpha_2 < \beta \in \mathbb{H}$. Since $SC_\Omega(\alpha_1 \# \alpha_2) = SC_\Omega(\alpha_1) \cup SC_\Omega(\alpha_2) < D\beta$ we also obtain $\alpha_1 \# \alpha_2 \ll \beta$.

(viii) is obvious since $D\alpha < \alpha \# \Omega$ and $SC_\Omega(\alpha) \subset SC_\Omega(\alpha \# \Omega)$.

The $\ll$-relation will be crucial for the definition of the semiformal system $\mathbf{ID}_\infty$. As already mentioned the derivation trees of $\mathbf{ID}_\infty$ will in general be Ω-branching trees whose nodes are labeled by ordinals in T. The derivation of a Π^1_1-formula, however, will just be a $\mathbf{Z}_\infty$-derivation, i.e. an ω-branching tree. To obtain an ordinal analysis of $\mathbf{ID}_1$ we therefore will have to collapse the Ω-branching derivation trees for Π^1_1-sentences in $\mathbf{ID}_\infty$ into ω-branching derivation trees of $\mathbf{Z}_\infty$. The effect of the collapsing procedure on the derivation trees and the assigned ordinals will be controlled by the collapsing function D on the ordinals of T.

So in order to obtain an ordinal assignment to the nodes of the derivations in $\mathbf{ID}_\infty$ which still is a correct assignment after the collapsing procedure (i.e. an assignment which is increasing in the direction from the top nodes to the bottom node) it will not suffice just to assign ordinals which are increasing but we will have to assign ordinals which are essentially increasing i.e. increasing in the sense of $\ll$. This of course causes problems in the case of an inference with Ω-many premises since for no ordinal $\alpha \in T$ there are Ω-many different ordinals $\alpha_\xi \ll \alpha$. To meet this difficulty we imagine a partial function f: On $\rightarrow$ On which enumerates the ordinals of an inference with infinitely many premises. The domain of f then corresponds to the 'number' of the premises of the inference. For such functions f and an ordinal α we are going to define a relation $f \ll \alpha$ which will be sufficient for the collapsing property.

24.17. Definition

Let f: On $\rightarrow$ On be a partial function such that dom f is a segment of On and α be an ordinal. We say that f *is essentially less than* α, in symbols $f \ll \alpha$, if the following conditions are satisfied:

(1) $\forall \xi \in \operatorname{dom} f\ (f\xi < \alpha)$

(2) $\forall \beta \forall \xi \in \operatorname{dom} f\ (\alpha \lesssim \beta \wedge \xi \ll \beta \Rightarrow f\xi \ll \beta)$

24.18. Lemma

(i) $f \ll \alpha$ *and* $\alpha \lesssim \beta$ *imply* $f \ll \beta$.

(ii) If $f\xi \ll \alpha$ *for all* $\xi \in \operatorname{dom} f$, *then we have* $f \ll \alpha$.

Proof

(i) We have $f\xi < \alpha \leq \beta$ for all $\xi \in \operatorname{dom} f$. If $\beta \lesssim \gamma$ and $\xi \ll \gamma$, then by $\alpha \lesssim \beta$ we also have $\alpha \lesssim \gamma$ and obtain $f\xi \ll \gamma$ by 24.17.(2).

(ii) $f\xi \ll \alpha$ immediately implies $f\xi < \alpha$ for all $\xi \in \operatorname{dom} f$. Now if $\alpha \lesssim \beta$, then we obtain $f\xi \ll \alpha \lesssim \beta$ which shows that 24.17.(2) is satisfied.

It follows from 24.18.(ii) that the relation $f \ll \alpha$ in fact is a generalization of the relation $f(\xi) \ll \alpha$ for all $\xi \in \operatorname{dom} f$. It will therefore suffice to secure $\alpha_\xi \ll \alpha$ for the premises of an infinitary inference in order to obtain a correct ordinal assignment. The details of the definition will be given in §26.

24.19. Lemma

Suppose $f \ll \alpha$. *Then we have* *(i)* $\lambda\xi.\Omega + f\xi \ll \Omega + \alpha$,

(ii) $\lambda\xi.(f\xi \# \gamma) \ll \alpha \# \gamma$

and

(iii) $\gamma \notin SC \Rightarrow \lambda\xi.\varphi\gamma(f\xi) \ll \varphi\gamma\alpha$.

Proof

(i) It is $\mathrm{dom}\,\lambda\xi.\Omega + f\xi = \mathrm{dom} f$. For $\xi \in \mathrm{dom} f$ it is $\Omega + f\xi < \Omega + \alpha$. If $\Omega + \alpha \lesssim \beta$ and $\xi \in \mathrm{dom} f$ such that $\xi \ll \beta$, then we obtain $\alpha \lesssim \beta$ by 24.16.(ii). By 24.17.(2) it follows $f\xi \ll \beta$ and we obtain $SC_\Omega(\Omega + f\xi) = SC_\Omega(f\xi) < D\beta$. Hence $\Omega + f\xi \ll \beta$.

We prove (ii) and (iii) simultaneously. To do so we define $g := \lambda\xi.(f\xi \# \gamma)$ or $g := \lambda\xi.\varphi\gamma(f\xi)$ and $\alpha_g := \alpha \# \gamma$ or $\alpha_g := \varphi\gamma\alpha$ respectively. Then in any case we have $\mathrm{dom}\, g = \mathrm{dom} f$, $\alpha \lesssim \alpha_g$ and $\gamma \ll \alpha_g$. For $\xi \in \mathrm{dom}\, g$ we obviously always have $g\xi < \alpha_g$. Now if $\alpha_g \lesssim \beta$ and $\xi \ll \beta$ for some $\xi \in \mathrm{dom}\, g$, then it follows $\alpha \lesssim \beta$ which first proves $f\xi \ll \beta$. We have $SC_\Omega(g\xi) = SC_\Omega(f\xi) \cup SC_\Omega(\gamma)$ and obtain $SC_\Omega(f\xi) < D\beta$ from $f\xi \ll \beta$. Because of $\gamma \ll \alpha_g \lesssim \beta$ we also obtain $SC_\Omega(\gamma) < D\beta$. Altogether we have $g\xi \ll \beta$. Hence $g \ll \beta$.

24.20. **Remark**

The reasons we gave for the definition of the relation $f \ll \alpha$ were purely technical. Of course it were the technical necessities of the proof of the cut elimination theorem which led us to the above definition. There is, however, another aspect under which this relation seems to be interesting. The ordinal Ω relativized to the notation system T (i.e. the term interpretation of Ω as we will call it in the following section) looses its regularity. We have $T \cap \Omega = \psi(\Omega^\Gamma) = \sup\{\psi\Delta_n : n < \omega\}$ which shows that Ω relativized to T has cofinality ω. But $\Omega \cap T$ should in some relativized sense reflect the regularity of Ω. We shall see in what sense. It is easy to see that a regular ordinal may be characterized in the following way:

$\varkappa$ is regular $\Leftrightarrow \forall f(\mathrm{Fun}(f) \Rightarrow \forall \eta < \varkappa\, (\sup\{f\xi : \xi < \eta\} < \varkappa))$

It is now obvious that $\Omega \cap T$ cannot be a model of the sentence 'Ω is regular' if f still ranges over all functions. But perhaps it should be possible to restrict the range of the quantifier $\forall f$. If for instance we restrict the quantifier to functions which are $\varkappa$-recursive, then we obtain the notion of a recursively regular ordinal. The question is if there also is a class $\mathcal{F}$ of functions such that $\Omega \cap T$ becomes regular relative to that class of functions, i.e. such that

$$(T, <\restriction T, \mathcal{F}) \models \text{'}\Omega \text{ is regular'}.$$

To obtain such a class we call a function f an *admissible function for* T if $\mathrm{rg} f \subseteq \Omega$ and there is some $\alpha \in T$ such that $f \ll \alpha$. We then have the following lemma.

24.21. Lemma

Let M *be a subset of* T *which, relative to* T, *is bounded in* Ω *and* f *is a function which is admissible for* T, *then* $\sup\{f\xi : \xi \in M\} \in T \cap \Omega$.

Proof

There is an $\alpha \in T \cap \Omega$ such that $\xi < \alpha$ holds for all $\xi \in M$. On the other hand there is a $\beta \in T$ such that $f \ll \beta$. Now $\xi < \alpha < \Omega$ implies $\xi \ll \alpha \lessapprox \alpha \# \beta \in T$. Since also $\beta \lessapprox \alpha \# \beta$ we obtain from $f \ll \beta$ already $f\xi \ll \alpha \# \beta$. Hence $f\xi < D(\alpha \# \beta)$ for all $\xi \in M$ and the proof of the lemma is completed.

As a consequence of lemma 24.21. we obtain that $(T, <\restriction T, \mathcal{F}) \models$ 'Ω is regular' holds if $\mathcal{F}$ is the class of functions which are admissible for the notation system T. One easily checks that $Id_\Omega \ll \Omega$. Hence Id_Ω is a function which is admissibe for T. Together with 24.19. this provides us with a wide class of functions which are admissible for T. It is exactly this class which will be relevant for the cut elimination procedure in §27.

24.21 Exercises

1. Prove the following statements:
   - (i) $\beta \in SC(\alpha) \wedge \beta \geq \Omega \Rightarrow \beta = \Omega$
   - (ii) $SC_\Omega(\alpha) \subset B(\beta) \Leftrightarrow \alpha \in B(\beta)$
   - (iii) $\alpha < \Omega \Rightarrow SC_\Omega(\alpha) \leq \alpha$
2. Compute:
   - (i) $D\Omega$
   - (ii) $D(\varepsilon_{\Omega+1})$
   - (iii) $D(\Omega+\varepsilon_0)$
   - (iv) $D(\Omega+\psi\Omega)$
3. Compute the following sets:
   - (i) $\{\alpha : \alpha \ll \Omega\}$
   - (ii) $\{\alpha : \alpha \ll \Omega \cdot \omega\}$
4. Let $f, g : On \to On$ be functions such that $\operatorname{dom} f = \operatorname{dom} g \in On$. Show:
   - (i) $id\restriction\eta \ll \eta$ for all $\eta \in T$
   - (ii) $f \ll \alpha, g \ll \alpha \wedge \alpha \in \mathbb{H} \Rightarrow f+g \ll \alpha$

§25. Alternative interpretations for Ω

Hitherto we always assumed that Ω denotes the first regular ordinal above ω. But we already indicated that there might be alternative interpretations for Ω. In the first version of the proof of 22.16.,the completeness theorem for $\mathscr{L}^{I}_{\infty}$, we already did interpret Ω by ω_1^{CK}, the first recursive regular ordinal above ω. For the treatment of the theory of arithmetically definable inductive definitions this interpretation is in fact the natural one. In the present section we will show that this and further interpretations of Ω are consistent with the development of the ordinal notation system in §§ 23 and 24.

In a first remark we notice that in the definition of ordinal addition and of the φ-function the ordinal Ω is of no importance. We therefore may presume that + and φ with all their properties developed in §§ 7 and 17 are independent from the interpretation of Ω.

In the case of the ψ-function the situation is completely different. Since we defined the function ψ and the sets $B(\alpha)$ simultaneously the ordinal Ω enters via clause (B1). Theorem 23.10. shows the importance of the ordinal Ω in the development of the ordinal notation system. It is now an obvious question to ask to what extend the notation system will be changed, when in clause (B1) we replace the first uncountable regular, which from now on will be denoted by $\aleph_1$, by some other ordinal.

To tackle this question we are going to take Ω as a symbol or variable for an ordinal without further information. When we try to develop the theory of §§ 23 and 24 with a free variable Ω, then already the proof of lemma 23.2. becomes impossible. The regularity of Ω played a crucial role for the proof. A further inspection of §§ 23 and 24, however, shows that this in fact was the only place where we used the regularity of Ω. In later applications we always used lemma 23.2. Therefore it is an obvious idea to use lemma 23.2. as a defining axiom for Ω. We introduce the axiom

(Ax_Ω) $\forall\alpha(\psi\alpha<\Omega)$..

By assuming (Ax_Ω) we may develop the theory of §§ 23 and 24 for ordinal terms built up from the functions $+,\varphi,\psi$ and the variable Ω without serious problems. Some places, however, need some caution. So for instance lemma 23.22., where we need $\alpha \neq \Omega$. These and more silly difficulties are easily avoided by the additional requirement $\Omega \in SC$. This requirement, however, is not essential but just for convenience (cf. exercise). We leave it to the reader to convince himself that (Ax_Ω) and $\Omega \in SC$ are in fact sufficient to obtain §§ 23 and 24. In order to distinguish ordinals from ordinal-terms we will, for the moment,

denote ordinal-terms by lower case latin letters. After the identification of ordinal terms and their standard interpretations we will be able to drop again this distinction.
An assignment $V(\Omega)\in SC$ for the variable Ω will be called an *interpretation* V for Ω. If V is an interpretation for Ω, then we define sets $B^V(\alpha)$ and a function ψ^V in the following way:

25.1. **Inductive definition** of the sets $B^V(\alpha)$ and the function ψ^V

(B^V1) $\{0, V\Omega\}\subset B^V(\alpha)$.

(B^V2) If $\xi =_{NF} \xi_1+..+\xi_n$ and $\{\xi_1,...,\xi_n\}\subset B^V(\alpha)$, then also $\xi\in B^V(\alpha)$,

(B^V3) If $\xi =_{NF} \varphi\xi_1\xi_2$ and $\{\xi_1,\xi_2\}\subset B^V(\alpha)$, then also $\xi\in B^V(\alpha)$,

(B^V4) If $\xi\in B^V(\alpha)\cap\alpha$ and $\xi\in B^V(\xi)$, then $\psi^V\xi\in B(\alpha)$.

(ψ^V1) $\psi^V\alpha := \min\{\xi : \xi\notin B^V(\alpha)\}$.

If we define $\eta =_{NF} \psi^V\xi :\Leftrightarrow \eta = \psi^V\xi \wedge \xi\in B^V(\xi)$ then clause (B^V4) takes the form

(B^V4') If $\eta =_{NF} \psi^V\xi \wedge \xi\in B^V(\alpha)\cap\alpha$, then $\eta\in B^V(\alpha)$.

The only essential difference to definition 23.1. or rather to the variation of 23.1. given in 23.19. is the fact that in clause (B^V3) we have built in the normal form condition for $\psi^V\xi$.

Since according to theorem 23.26. all ordinal-terms in the set $B(\Omega^\Gamma)$ are uniquely represented by terms of the set T, we may extend the interpretation V to the terms in $B(\Omega^\Gamma)$ as in definition 25.2. below. The value of an ordinal-term a in the extended interpretation V will be denoted by a^V.

25.2. **Definition** of a^V for $a\in B(\Omega^\Gamma)$

(i) $0^V := 0$, $\Omega^V := V(\Omega)$.

(ii) If $a =_{NF} a_1+...+a_n$, then $a^V := a_1^V+...+a_n^V$.

(iii) If $a =_{NF} \varphi a_1 a_2$, then $a^V := \varphi a_1^V a_2^V$.

(iv) If $a =_{NF} \psi a_0$, then $a^V := \psi^V a_0^V$.

The *degree* Ga of an ordinal term a is defined analogously to the degree of an ordinal $\alpha\in T$, i.e. Ga is the stage of the term a in the inductive definition of the set T of ordinal terms.

We call the interpretation $St(\Omega) := \aleph_1$ the *standard interpretation* for Ω. In §§ 23 and 24 we developed the theory of the standard interpretation for Ω. It follows from lemma 23.25. that the sets $B(\alpha)$ and $B^{St}(\alpha)$ and therefore also the functions ψ and ψ^{St} coincide. Theorem 23.26. may be interpreted as the statement that *every ordinal $\alpha \in B(\aleph_1^\Gamma)$ is the standard interpretation of an uniquely determined ordinal term $a_\alpha \in B(\Omega^\Gamma)$.* Now let V be an interpretation for Ω. By defining $\alpha^V := a_\alpha^V$ we obtain by 25.2. a mapping ${}^V : B(\aleph_1^\Gamma) \to On$. For $M \subset B(\aleph_1^\Gamma)$ we denote by M^V the image of M under ^{V}. The mapping St then obviously is the identity on $B(\aleph_1^\Gamma)$. Therefore we are going to drop again the distinction between ordinals and ordinal terms and identify ordinal terms and their standard interpretations.

25.3 Definition

An interpretation V for Ω is *good relative to an ordinal* β if $\psi^V\alpha^V < V(\Omega)$ holds for all $\alpha \in B(\beta) \cap (\beta+1)$. An interpretation which is good relative to $\aleph_1^\Gamma$ is a *good interpretation.*

The result of lemma 23.2. may now be reformulated as:

25.4 Theorem

The standard interpretation for Ω is a good interpretation.

25.5. Lemma

Let V be an interpretation. Then we have for all $\alpha \in B(\aleph_1^\Gamma)$

(i) $\alpha \in \mathbb{H} \Rightarrow \alpha^V \in \mathbb{H}$,

(ii) $\alpha \in SC \Rightarrow \alpha^V \in SC$.

Proof

If $\alpha = \aleph_1$, then we have $\alpha^V = V\Omega$ and obtain $V\Omega \in SC \subset \mathbb{H}$ since V is an interpretation.
If $\alpha =_{NF} \varphi\alpha_1\alpha_2$, then we have $\alpha^V = \varphi\alpha_1^V\alpha_2^V$ which implies $\alpha^V \in \mathbb{H}$.
If $\alpha =_{NF} \psi\alpha_0$, then it is $\alpha^V = \psi^V\alpha_0^V$ and by 23.5. (whose proof does not need (Ax_Ω)) we obtain $\alpha^V \in SC$.

It is $\Gamma'(0) = \min\{\xi : \Gamma_\xi = \xi\}$. By theorem 23.10. it follows that the function ψ restricted to $\Gamma'(0)$ coincides with the function $\lambda\xi.\Gamma_\xi$. Since the definition of the function $\lambda\xi.\Gamma_\xi$ does not depend upon the value of $V\Omega$ we have that ψ^V and ψ coincide below $\Gamma'(0)$ for every interpretation for which we have $\Gamma'(0) \le V\Omega$. (Otherwise ψ^V could be shifted a litte bit at the place where $V\Omega$ comes into

$B^V(\beta))$. But we have to be careful. In order to use the above argument we have to know that theorem 23.10 holds for all interpretations, i.e. we have to convince ourselves that we do not need (Ax_Ω) in the proof of 23.10. The facts we need in the proof of 23.10. are the following:

(1) $B^V(0) \cap V\Omega = P(0)$.

(2) $\xi < \eta \Rightarrow B^V(\xi) \subset B^V(\eta)$. Hence $\xi < \eta \Rightarrow \psi^V\xi \le \psi^V\eta$.

(3) If $\xi < \Gamma'(0)$, then $\xi \in B^V(\xi+1) \wedge \Gamma_\xi = \psi^V\xi \Rightarrow \psi^V(\xi+1) = (\psi^V\xi)^\Gamma$.

(4) $\psi^V \restriction \Gamma'(0)$ is continuous.

Since $\Gamma_0 < \Gamma'(0) \le V\Omega$ (1) again follows from a comparison of the definitions of $P(0)$ and $B^V(0)$. (2) holds trivially. What we really need to prove is (3) and (4). (3) is lemma 23.9. (i) and (4) is lemma 23.7. and in both proofs we used (Ax_Ω). We therefore have to reprove (3) and (4) without using (Ax_Ω). To prove (3) we observe that $\xi \in B^V(\xi+1)$ implies $\psi^V\xi < \psi^V(\xi+1) \in SC$. So it remains to show

(5) $\psi^V(\xi+1) \le (\psi^V\xi)^\Gamma$.

Assume that $(\psi^V\xi)^\Gamma < \psi^V(\xi+1)$. Then $(\psi^V\xi)^\Gamma \in B^V(\xi+1)$. Since $(\psi^V\xi)^\Gamma \in SC$ and $B^V(\xi+1) \cap SC$ only contains ordinals of the form $\psi^V\eta \le \psi^V\xi$ or $V\Omega$ we obtain $(\psi^V\xi)^\Gamma = V\Omega$. But then we have $V\Omega = (\psi^V\xi)^\Gamma = \Gamma_{\xi+1}$. $\xi < \Gamma'(0)$, however, also implies $\Gamma_{\xi+1} < \Gamma'(0) \le V\Omega$. A contradiction.

To show (4) we simultaneously prove $\sup\{\psi^V\eta : \eta < \xi\} = \psi^V\xi$ for $\xi \in \Gamma'(0) \cap Lim$ and

(6) $\xi < \Gamma'(0) \Rightarrow \psi^V\xi = \Gamma_\xi$

by induction on ξ.

Define $\rho := \sup\{\psi^V\eta : \eta < \xi\}$. Then $\rho \le \psi^V\xi$ by (2). Assume $\rho < \psi^V\xi$. Then $\rho \in B^V(\xi)$. We obviously have $\rho \in SC$. All strongly critical ordinals in $B^V(\xi)$ different from $V\Omega$ are of the form $\psi^V\eta$ for some $\eta < \xi$. Hence $\rho = V\Omega$ and using the induction hypothesis for (6) we obtain $V\Omega = \rho = \sup\{\Gamma_\eta : \eta < \xi\} = \Gamma_\xi < \Gamma'(0) \le V\Omega$. Contradiction.

This proves (4). For the proof of (6) we distinguish the following cases:

$\xi = 0$. Then $\psi^V 0 = \Gamma_0$ by (1).

$\xi = \xi_0+1$. Then $\xi_0 < \Gamma_{\xi_0} = \psi^V\xi_0 \in B^V(\xi_0) \subset B^V(\xi_0+1)$ and by (3) and the induction hypothesis it follows $\psi^V\xi = (\psi^V\xi_0)^\Gamma = (\Gamma_{\xi_0})^\Gamma = \Gamma_\xi$.

If $\xi \in Lim$ we immediately obtain the claim from (4) and the induction hypothesis.

25.6. **Lemma**

Let V be an interpretation such that $\Gamma'(0) \le V(\Omega)$. Then we have $\psi^V\xi = \Gamma_\xi$ and $\xi^V = \xi$ for all $\xi < \Gamma'(0)$.

Proof

We already have shown $\psi^V\xi = \Gamma_\xi$ for all $\xi < \Gamma'(0)$. We prove $\xi^V = \xi$ by induction on $G\xi$. If ξ is not of the form $\psi^V\eta$ we obtain the claim immediately from the induction hypothesis. If $\Gamma'(0) > \xi =_{NF} \psi\eta$ then we have $\eta < \Gamma'(0)$ and, using the induction hypothesis, obtain $\xi^V = \psi^V\eta^V \stackrel{i.h.}{=} \psi^V\eta = \Gamma_\eta = \psi\eta = \xi$.

Another property which is provable without using (Ax_Ω) is the following:

(7) $\xi \in B^V(\xi) \cap \eta \Rightarrow \psi^V\xi < \psi^V\eta$.

To prove (7) we assume $\xi \in B^V(\xi) \cap \eta$ and observe that by (2) we then have $\xi \in B^V(\eta)$ and $\psi^V\xi \leq \psi^V\eta$. By (B^V4) it follows $\psi^V\xi \in B^V(\eta)$ which implies $\psi^V\xi \neq \psi^V\eta$.

In order to have lemma 25.6. we from now on tacitly assume that for all interpretations we have $\Gamma'(0) \leq V\Omega$.

Now let Θ be an ordinal and suppose that V is an interpretation which is good relative to Θ.

25.7. **Lemma**

(i) For all ordinals α,β *in* $B(\Theta+1)$ *we have* $\alpha < \beta$ *if and only if* $\alpha^V < \beta^V$,

(ii) For all ordinals β *in* $B(\Theta) \cap (\Theta+1)$ *we have that* $\alpha \in B(\beta)$ *implies* $\alpha^V \in B^V(\beta^V)$.

(iii) *For all ordinals* α,β *in* $B(\Theta+1)$ *such that* $\alpha \in Cr(\beta)$ *we also have* $\alpha^V \in Cr(\beta^V)$.

Proof

We prove claims (i) and (ii) simultaneously by induction on $2^{G\alpha} + 2^{G\beta}$.

(i) In the proof of claim (i) we follow the distinction by cases of 23.30. It suffices to show $\alpha < \beta \Rightarrow \alpha^V < \beta^V$. The opposite direction then is an immediate consequence.

If $\alpha = 0$ and $\beta \neq 0$, then we have also $\alpha^V = 0$ and $\beta^V \neq 0$ which imply $\alpha^V < \beta^V$.

If $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$, then it is $2^{G\alpha_i} + 2^{G\alpha_{i+1}} \leq 2^{G\alpha} < 2^{G\alpha} + 2^{G\beta}$ for $i = 1,\ldots,n-1$. By 23.22. we have $\{\alpha_1,\ldots,\alpha_n\} \subset B(\Theta+1)$ and obtain by the induction hypothesis first $\alpha_1^V \geq \ldots \geq \alpha_n^V$ and by 25.5. also $\alpha_k^V \in \mathbb{H}$. Hence $\alpha^V =_{NF} \alpha_1^V + \ldots + \alpha_n^V$. For $\beta =_{NF} \beta_1 + \ldots + \beta_m$ we analogously obtain $\beta^V = \beta_1^V + \ldots + \beta_m^V$. Now if $\alpha_j = \beta_j$ for all $j \leq i$ and $\alpha_{i+1} < \beta_{i+1}$, then we obtain by the induction hypothesis $\alpha_j^V = \beta_j^V$ for all $j \leq i$ and $\alpha_{i+1}^V < \beta_{i+1}^V$. By 23.30.(2) (which does not depend upon (Ax_Ω) but only on the results of §7) it then follows $\alpha^V < \beta^V$.

If $\beta \in \mathbb{H}$ and $\alpha_1 < \beta$, then we also have $\beta^V \in \mathbb{H}$ and obtain $\alpha_1^V < \beta^V$ by the induction hypothesis. Hence $\alpha^V < \beta^V$ by 23.30.(3).

If $\alpha \in \mathbb{H}$, $\beta =_{NF} \beta_1 + \ldots + \beta_n$ and $\alpha \leq \beta_1$, then as before we obtain $\beta^V =_{NF} \beta_1^V + \ldots + \beta_n^V$ and $\alpha^V \leq \beta_1^V$ by the induction hypothesis . By 23.30.(4) it follows $\alpha^V < \beta^V$.

If $\alpha =_{NF} \varphi\alpha_1\alpha_2$ and $\beta =_{NF} \varphi\beta_1\beta_2$, then we have $\{\alpha_1,\alpha_2,\beta_1,\beta_2\} \subset B(\Theta+1)$ by 23.22. and obtain $\alpha^V < \beta^V$ by 17.8.(2) and the induction hypothesis.

If $\alpha =_{NF} \varphi\alpha_1\alpha_2$ and $\beta \in SC$ such that $\alpha_1,\alpha_2 < \beta$, then we obtain $\alpha^V = \varphi\alpha_1^V\alpha_2^V$ and $\alpha_1^V,\alpha_2^V < \beta^V$ by 23.22. and the induction hypothesis. By 25.5. it follows $\beta^V \in SC$ and by 17.14. we obtain $\alpha^V < \beta^V$.

If $\alpha \in SC$ and $\beta =_{NF} \varphi\beta_1\beta_2$, then we have $\alpha < \beta_1$ or $\alpha = \beta_1$ and $\beta_2 \neq 0$ or $\alpha < \beta_2$. By 23.22. and the induction hypothesis it follows $\alpha^V < \beta_1^V$, $\alpha^V = \beta_1^V$ and $\beta_2^V \neq 0$ or $\alpha^V < \beta_2^V$. By 25.5. we also have $\alpha^V \in SC$ and obtain $\alpha^V = \varphi\alpha^V 0 < \varphi\beta_1^V 0 \leq \varphi\beta_1^V\beta_2^V = \beta^V$ in the first case. In the second case it follows $\alpha^V = \varphi\alpha^V 0 = \varphi\beta_1^V 0 < \varphi\beta_1^V\beta_2^V = \beta^V$ and in the third case $\alpha^V < \beta_2^V \leq \varphi\beta_1^V\beta_2^V = \beta^V$.

Now if $\alpha =_{NF} \psi\alpha_0$ and $\beta =_{NF} \psi\beta_0$, then we have $2^{G\alpha_0} + 2^{G\alpha_0} \leq 2^{G\alpha} < 2^{G\alpha} + 2^{G\beta}$ and analogously $2^{G\beta_0} + 2^{G\beta_0} < 2^{G\alpha} + 2^{G\beta}$. We have $\alpha_0 \in B(\alpha_0)$ and by 23.25. $\alpha_0 \leq \Theta$. Hence $\alpha_0 \in B(\Theta)$. Similarly we also obtain $\beta_0 \in B(\beta_0) \cap B(\Theta)$. Therefore we may apply the induction hypothesis for claim (ii) to α_0 and β_0 and obtain $\alpha_0^V \in B^V(\alpha_0^V)$ and $\beta_0^V \in B^V(\beta_0^V)$. This implies $\alpha^V =_{NF} \psi^V\alpha_0^V$ and $\beta^V =_{NF} \psi^V\beta_0^V$. By the induction hypothesis for claim (i) we have $\alpha_0^V < \beta_0^V$ and obtain $\alpha^V < \beta^V$ by (7).

If finally $\alpha =_{NF} \psi\alpha_0$ and $\beta = \aleph_1$, then we obtain as above $\alpha_0 \in B(\Theta) \cap \Theta+1$. Since V is a good interpretation relative to Θ it follows $\alpha^V = \psi^V\alpha_0^V < V\Omega = \beta^V$.

(ii) If $\alpha = 0$ or $\alpha = \aleph_1$, then we have $\alpha^V \in B^V(\beta^V)$ according to (B1).

If $\alpha =_{NF} \alpha_1 + \ldots + \alpha_n$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2$, then we either obtain $\alpha^V =_{NF} \alpha_1{}^V + \ldots + \alpha_n{}^V$ or $\alpha^V =_{NF} \varphi\alpha_1{}^V\alpha_2{}^V$ by the induction hypothesis for (i) and 25.5. By the induction hypothesis for (ii) and (B^V2) or (B^V3) respectively we obtain $\alpha^V \in B^V(\beta^V)$.

If $\alpha =_{NF} \psi\alpha_0$, then by 23.25. we have $\alpha_0 \in B(\beta) \cap \beta \subset B(\Theta) \cap (\Theta+1)$. By the induction hypothesis for claim (ii) and (i) we obtain $\alpha_0^V \in B^V(\alpha_0{}^V) \cap B^V(\beta^V) \cap \beta^V$. By a clause ($B^V4$), however, this implies $\alpha^V = \psi^V\alpha_0^V \in B^V(\beta^V)$.

We prove (iii). If $\alpha \in Cr(\beta)$, then we either have $\alpha \in SC$ and $\beta \leq \alpha$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2$ for some $\alpha_1 \geq \beta$. In the first case we obtain $\alpha^V \in SC$ and $\beta^V \leq \alpha^V$ by (i). Hence $\alpha^V \in Cr(\beta^V)$. In the second case it follows that $\alpha^V = \varphi\alpha_1^V\alpha_2^V$. By (i) it is $\alpha_1^V \geq \beta^V$ which implies $\alpha^V \in Cr(\beta^V)$.

25.8. **Lemma**

If $\beta \in B(\Theta) \cap \Theta+1$, *then for every* $\alpha \in B^{V,n}(\beta^V)$ *there is an* $\eta \in B(\beta)$ *such that* $\alpha = \eta^V$. *It is* $\eta \in \mathbb{H}$ *whenever* $\alpha \in \mathbb{H}$.

Proof

We prove the claim by main induction on β with side induction on n.

If $\alpha = 0$ or $\alpha = V\Omega$, then we define $\eta := 0$ or $\eta := \aleph_1$. In both cases we have $\eta \in B(\beta)$.

If $\alpha =_{NF} \alpha_1+...+\alpha_n$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2$, then by induction hypothesis there are ordinals $\eta_1,...,\eta_n$ or η_1,η_2, respectively, which are in $B(\beta)\cap B(\Theta)$ such that $\alpha_k = \eta_k^V$ holds for $k = 1,...,n$ or $k = 1,2$ respectively. Now we define $\eta := \eta_1+...+\eta_n$ or $\eta := \varphi\eta_1\eta_2$ respectively. Then it is $\eta \in B(\beta)$. In the first case we have $\alpha_k \in \mathbb{H}$ and obtain $\eta_k \in \mathbb{H}$ by the induction hypothesis. $\eta_1^V \geq ... \geq \eta_n^V$ implies $\eta_1 \geq ... \geq \eta_n$ by 25.7. In the second case we have $\eta_1,\eta_2 \leq \eta$. $\eta_1 = \eta$ implies $\eta_1 \in SC$ and $\eta_2 = 0$. Hence $\alpha_1 = \eta_1^V \in SC$ by 25.5 and $\alpha_2 = 0$. This, however, contradicts $\alpha =_{NF} \varphi\alpha_1\alpha_2$. If $\eta_2 = \eta$, then $\eta_2 \in Cr(\eta_1+1)$. By 25.7. this implies $\eta_2^V \in Cr(\eta_1^V+1)$ which entails $\varphi\alpha_1\alpha_2 = \varphi\eta_1^V\eta_2^V = \eta_2^V = \alpha$ in contradiction to $\alpha =_{NF} \varphi\alpha_1\alpha_2$. Hence $\eta_1,\eta_2 < \eta$ and we have $\eta =_{NF} \eta_1+...+\eta_n$ as well as $\eta =_{NF} \varphi\eta_1\eta_2$ and in both cases it follows $\eta^V = \alpha$. In the second case we obviously also have $\eta \in \mathbb{H}$.
If $\alpha =_{NF} \psi^V\alpha_o$, then by the side induction hypothesis there is an $\eta_o \in B(\beta) \subset B(\Theta)$ such that $\eta_o^V = \alpha_o$. From the normal form condition we obtain $\eta_o^V \in B^V(\eta_o^V)$. Since $\eta_o^V = \alpha_o < \beta^V$ we obtain $\eta_o < \beta$ by 25.7. Hence $\eta_o \in B(\Theta)\cap(\Theta+1)$ which implies $\eta_o \in B(\eta_o)$ by the main induction hypothesis and 25.7. If we define $\eta := \psi\eta_o$, then we have $\eta =_{NF} \psi\eta_o$. Hence $\eta \in SC\cap B(\beta)$ and $\eta^V = \psi^V\eta_o^V = \alpha$.

25.9. **Corollary**
Assume $\beta \in B(\Theta)$. *Then we have* $B(\beta)^V = B^V(\beta^V)$.

Proof
$\alpha \in B(\beta)$ implies $\alpha^V \in B^V(\beta^V)$ by 25.7.(ii). If conversely $\alpha \in B^V(\beta^V)$, then by 25.8. there is an $\eta \in B(\beta)$ such that $\alpha = \eta^V$. Hence $\alpha \in B(\beta)^V$.

As an immediate consequence of 25.9. we also obtain

25.10. **Corollary**
For every good interpretation V *we have* $B(\aleph_1^\Gamma)^V = B^V((V\Omega)^\Gamma)$.

Proof
Define the sequence Δ_n as in the proof of 23.14. Then we have $\Delta_n \in B(\Delta_n)$ for all $n<\omega$ and $B(\aleph_1^\Gamma) = \bigcup\{B(\Delta_n): n<\omega\}$. V is good relative to all Δ_n. Hence $B(\aleph_1^\Gamma)^V = \bigcup\{B(\Delta_n)^V: n<\omega\} = \bigcup\{B^V(\Delta_n^V): n<\omega\} = B^V(V\Omega^\Gamma)$ since $\sup\{\Delta_n^V: n<\omega\} = V\Omega^\Gamma$.

25.11 **Theorem**
(i) If V *is good relative to* Θ *and* $\beta \in B(\Theta)$, *then* ^{V} *is an isomorphism from* $B(\beta)$ *onto* $B(\beta)^V$.

(ii) Assume that V *is a good interpretation. Then* ^{V} *is an isomorphism from* $B(\aleph_1^\Gamma)$ *onto* $B^V((V\Omega)^\Gamma)$

Proof

(i) follows from 25.7. and 25.9. while (ii) follows from 25.7. and 25.10.

The main concern of the present section is to show that the segment of the notation system does not depend upon the interpretation of Ω. For this purpose we are going to show that the mapping ^{V} is the identity on the ordinal segment in $B(\Theta)$. This is a consequence of the following lemma.

25.12. **Lemma**

If V *is good relative to* Θ*, then we have* $B^V(\alpha^V)\cap V\Omega = \psi^V\alpha^V$ *for all* $\alpha \in B(\Theta)\cap(\Theta+1)$.

Proof

Since V is good relative to Θ we have $\psi^V\alpha^V < V\Omega$ for all $\alpha \in B(\Theta)\cap(\Theta+1)$ and prove $B^V(\alpha^V)\cap V\Omega = \psi^V\alpha^V$ as in 23.6.

25.13. **Theorem**

If V *is a good interpretation relative to* Θ*, then we have*

(i) $B^V(\alpha^V)\cap V\Omega = B(\alpha)\cap\aleph_1$ *for all* $\alpha \in B(\Theta)\cap(\Theta+1)$

and

(ii) $\alpha^V = \alpha$ *for all* $\alpha < \psi(\Theta)$.

Proof

If V is a good interpretation relative to Θ, then for all $\alpha \in B(\Theta)\cap(\Theta+1)$ ^{V} is an order isomorphism from $B(\alpha)$ onto $B^V(\alpha^V)$ mapping $\aleph_1$ to $V\Omega$. According to 25.12 $B(\alpha)\cap\aleph_1$ and $B^V(\alpha^V)\cap V\Omega$ are segments. Thus ^{V} has to be the identity map on $B(\alpha)\cap\aleph_1$.

We prove the second part of the theorem by induction on $G\alpha$. For $\alpha = 0$ it is $\alpha^V = 0$.

If $\alpha =_{NF} \alpha_1+\ldots+\alpha_n$ or $\alpha =_{NF} \varphi\alpha_1\alpha_2$, then we have $\alpha_k^V = \alpha_k$ for $k = 1,\ldots,n$ or $k = 1,2$ respectively by the induction hypothesis. Hence $\alpha^V = \alpha_1^V+\ldots+\alpha_n^V = \alpha_1+\ldots+\alpha_n = \alpha$ or $\alpha^V = \varphi\alpha_1^V\alpha_2^V = \varphi\alpha_1\alpha_2 = \alpha$ respectively.

If $\alpha =_{NF} \psi\alpha_0$, then we have $\alpha_0 \in B(\Theta)\cap\Theta$ and obtain $B^V(\alpha_0^V)\cap V\Omega = B(\alpha_0)\cap\aleph_1$ by the first part. Hence $\alpha^V = \psi^V\alpha_0^V = \psi\alpha_0 = \alpha$.

25.14. **Corollary**

For a good interpretation V *we have* $\alpha^V = \alpha$ *for all* $\alpha \in B(\aleph_1^\Gamma) \cap \aleph_1$.

Proof

Let $\{\Delta_n : n < \omega\}$ be as above. For every $\alpha \in B(\aleph_1^\Gamma) \cap \aleph_1$ there is an $n < \omega$ such that $\alpha \in B(\Delta_n) \cap \aleph_1$. Since V is good relative to Δ_n we obtain $\alpha^V = \alpha$ by 25.13.

Corollary 25.14. shows that the segment of the notation system below $\aleph_1$ is invariant under reinterpretations of Ω, provided they are good interpretations. This of course becomes wrong for the ordinals above $\aleph_1$. These ordinals will be moved by a reinterpretation of Ω. The moving of these ordinals is characterized by the function $\alpha \mapsto \alpha^V$. As we saw in 25.7. this function is order preserving.

25.15. **Theorem**

Assume $\Theta \in B(\Theta)$. *If* V *is a good interpretation relative to* Θ, *then we have* $\psi\Theta < V\Omega$.

Proof

Assume that V is a good interpretation relative to Θ such that $V\Omega \leq \psi\Theta$. By 25.13. it follows that $V\Omega \leq \psi\Theta = \psi^V\Theta^V$ in contradiction to the hypothesis that V is good relative to Θ.

25.16. **Corollary**

For any good interpretation V *we have* $\psi(\aleph_1^\Gamma) \leq V\Omega$.

Proof

Again we denote by $\{\Delta_n : n < \omega\}$ the fundamental sequence for $\aleph_1^\Gamma$ (cf. 25.10 and 23.14.). Then V is good relative to all Δ_n. Since $\Delta_n \in B(\Delta_n)$ holds for all $n < \omega$ we obtain by 25.15. $\psi(\Delta_n) < V\Omega$ for all $n < \omega$. Hence $\psi(\aleph_1^\Gamma) \leq V\Omega$.

Our hitherto only example of a good interpretation for Ω is the standard interpretation. We are now going to show that there are much more good interpreations. We will even be able to give a precise characterization of the good interpretations.

25.17. **Lemma**

Let V be an interpretation and Θ an ordinal such that $\psi(\Theta) \le V\Omega \le \aleph_1$. Then we have:

(i) $\beta^V \le \beta$ holds for all $\beta \in B(\Theta)$

(ii) V is a good interpretation relative to all $\beta < \Theta$.

Proof

We prove (i) and (ii) simultaneously by main induction on $G\beta$ and side induction on β. We first show (i).

If $\beta = 0$, then $\beta^V = \beta$ and if $\beta = \aleph_1$, then $\beta^V = V\Omega \le \aleph_1 = \beta$.

If $\beta =_{NF} \beta_1 + ... + \beta_n$ or $\beta =_{NF} \varphi\beta_1\beta_2$, then we obtain the claim immediately from the main induction hypothesis.

Let $\beta =_{NF} \psi\beta_0$. Then we have $\beta_0 \in B(\Theta) \cap \Theta$ and V is good relative to β_0 according to the main induction hypothesis for (ii). But we also have $\beta_0 \in B(\beta_0) \cap \beta_0 + 1$ by the normal form condition and obtain $\beta^V = \psi^V \beta_0^V = \psi\beta_0 = \beta$ by 25.13.

To prove (ii) we have to show that $\xi \in B(\beta) \cap \beta + 1$ implies $\psi^V \xi^V < V\Omega$. So assume $\xi \in B(\beta) \cap \beta + 1$. Then we have $\psi\xi \le \psi\beta$.

For $\beta = 0$ it is $\xi = 0$. Hence $\psi^V \xi^V = \Gamma_0 = \psi 0 \le \psi\Theta < V\Omega$ by 25.6.

If $\beta \in Lim$ we distinguish the following cases.

1. $\beta = \xi$. Then we have $\xi \in B(\xi)$ which implies that $\psi\xi$ is in normal form. If $\beta \le \psi\beta$ we have $\xi = \beta < \psi\beta \le \psi\Omega = \Gamma'(0)$ which by 25.6. implies $\psi^V \xi^V = \psi\xi = \Gamma_\xi < \Gamma'(0) \le V\Omega$. Now assume $\psi\beta < \beta$. We have $\psi\beta \in B(\Theta) \cap \Theta$ since $\beta \in B(\Theta) \cap \Theta$. We may now apply the side induction hypothesis for (i) and obtain $\psi^V \beta^V = (\psi\beta)^V \le \psi\beta < \psi\Theta < V\Omega$.

2. $\xi < \beta$. Since $B(\beta) = \bigcup\{B(\eta) : \eta < \beta\}$ there is an $\eta < \beta$ such that $\xi \in B(\eta) \cap \eta + 1$. By side induction hypothesis for (ii) we have that V is good relative to η which entails $\psi^V \xi^V < V\Omega$.

We finally assume $\beta = \beta_0 + 1$. If $\beta \le \psi\beta$, then we again obtain $\xi \le \beta < \psi\beta \le \psi\Omega = \Gamma'(0)$. Hence $\psi^V \xi^V = \Gamma\xi < \Gamma'(0) \le V\Omega$ by 25.6. Now assume $\psi\beta < \beta$. There is an $\alpha \in B(\aleph_1^\Gamma)$ such that $\psi\xi =_{NF} \psi\alpha$ and in the terminology of lemma 23.23. it is $\alpha = \xi(\xi) \ge \xi$. If $\xi = \beta$, then $\alpha = \xi$ as $\xi \in B(\beta)$ and therefore $\xi(\xi) = \beta(\beta) = \min\{\eta : \beta \le \eta \in B(\beta)\} = \beta$ and we obtain $\psi\beta \in B(\Theta) \cap \beta$ from $\beta \in B(\Theta) \cap \Theta$ and $\psi\beta < \beta$. By the side induction hypothesis for (i) it then follow $\psi^V \xi^V = \psi^V \beta^V \le \psi\beta < \psi\Theta \le V\Omega$. If $\xi < \beta$, then $\psi\xi =_{NF} \psi\alpha \in B(\beta)$ which by 23.25. implie $\alpha \in B(\beta) \cap \beta \subset B(\Theta)$. By the side induction hypothesis for (ii) we know that V is good relative to β_0. So $\xi \le \alpha$ by 25.7. implies $\xi^V \le \alpha^V$ and as $\psi\alpha < \psi\beta < \beta$ we finally obtain $\psi^V \xi^V \le \psi^V \alpha^V \le \psi\alpha < \psi\beta < \psi\Theta \le V\Omega$ by the side induction hypothesis for (i).

25.18. **Lemma**

Suppose that Θ is a limit ordinal such that $\psi\Theta \le V\Omega$. Then we have $\psi^V \xi^V = (\psi\xi)^V = \psi\xi < \psi\Theta$ for all $\xi \in B(\Theta) \cap \Theta$.

Proof

If $\aleph_1 \leq V\Omega$, then we prove as in 23.2. that V is a good interpretation.
If $V\Omega < \aleph_1$ and $\xi \in B(\Theta) \cap \Theta$, then we have $\xi \in B(\eta) \cap \eta+1$ for some $\eta < \Theta$. By the hypothesis $\psi\Theta \leq V\Omega$ and 25.17. it follows that V is a good relative to η. Hence $\psi^V\xi^V = (\psi\xi)^V = \psi\xi$ by 25.13.

25.19. **Theorem**
An interpretation V *is good if and only if* $\psi(\aleph_1^\Gamma) \leq V\Omega$.

Proof

If V is good, then $\psi(\aleph_1^\Gamma) \leq V\Omega$. by 25.16. For the opposite direction assume $\psi(\aleph_1^\Gamma) \leq V\Omega$ and $\xi \in B(\aleph_1^\Gamma) \cap (\aleph_1^\Gamma+1) = B(\aleph_1^\Gamma) \cap \aleph_1^\Gamma$. Then we obtain $\psi^V\xi^V = \psi\xi < \psi(\aleph_1^\Gamma) \leq V\Omega$ by 25.18. So V is a good interpretation.

25.20. **Theorem**
The following interpretations are good interpretations:

(i) $V\Omega := \aleph_1$. This is the *standard interpretation* denoted by St

(ii) $V\Omega := \omega_1^{CK}$. We call this interpretation the *recursive standard interpretation* and denote it by Rec.

(iii) $V\Omega := \psi(\aleph_1^\Gamma)$. We call this interpretation the *term interpretation* of Ω.

Proof

(i) and (iii) are already proved. (ii) follows from 25.19. and corollary 23.32.

In a last remark we return to (Ax_Ω).

25.21. **Definition**
We call an interpretation V a *global model of* (Ax_Ω) if we have $\forall\xi\, (\psi^V\xi < V\Omega)$. We call it a *local model of* (Ax_Ω) if it holds $\forall\xi \in B((V\Omega)^\Gamma)\, (\psi^V\xi < V\Omega)$.

If V is a global model of (Ax_Ω), then the theory of §§23 and 24 holds for the sets $B^V(\alpha)$ and the function ψ^V in the same way as it did for the sets $B(\alpha)$ and the functions ψ. This shows that we may replace $\aleph_1$ by any ordinal Θ which satisfies (Ax_Ω) without changing the theory. Therefore it is worthwhile to obtain a characterization of the interpretations which are global models of (Ax_Ω).

25.22. **Theorem**

The following statements are equivalent:

(i) V *is a global model for* (Ax_Ω)

(ii) V *is a good interpretation such that* $\psi(\aleph_1^\Gamma) < V\Omega$

(iii) V *is a local model for* (Ax_Ω) *and it is* $\psi^V(V\Omega^\Gamma) < V\Omega$

Proof

From (i) we trivially obtain that V is a good interpretation. But then we also have $\psi(\aleph_1^\Gamma) = \sup\{\psi(\Delta_n(\aleph_1)) : n<\omega\} = \sup\{\psi^V(\Delta_n(\aleph_1))^V : n<\omega\} = \psi^V(\sup\{\Delta_n(\aleph_1)^V : n<\omega\}) = \psi^V(\sup\{\Delta_n(V\Omega) : n<\omega\}) = \psi^V(V\Omega^\Gamma) < V\Omega$ by 25.13. and the continuity of ψ^V (cf. 17.22 and 23.7.).

Assume (ii) and choose any $\xi \in B^V((V\Omega)^\Gamma)$ then by 25.10. there is an $\eta \in B(\aleph_1^\Gamma)$ such that $\xi = \eta^V$. Hence $\psi^V\xi = \psi^V\eta^V < V\Omega$. So V is a local model for (Ax_Ω). We recursively define a sequence $\underline{\Delta}_n^V$ by $\underline{\Delta}_0^V := V\Omega+1$ and $\underline{\Delta}_{n+1}^V := \varphi(\underline{\Delta}_n^V)0$ and obtain $\sup\{\underline{\Delta}_n^V : n<\omega\} = (V\Omega)^\Gamma$ as well as $\underline{\Delta}_n^V \in B^V(\xi) \cap (V\Omega)^\Gamma$ for all $n<\omega$ and all ordinals ξ. It is obvious that $\underline{\Delta}_n^V = (\Delta_n(\aleph_1))^V$ (cf. 17.22.) and we will also obtain

$$(*) \quad \psi^V(V\Omega)^\Gamma = \sup\{\psi^V\underline{\Delta}_n^V : n<\omega\}.$$

From (*), 25.14. and 23.7. we obtain $\psi^V(V\Omega)^\Gamma = \sup\{\psi\Delta_n(\aleph_1) : n<\omega\} = \psi(\aleph_1^\Gamma) < V\Omega$. We prove (*). $\sigma := \sup\{\psi^V\underline{\Delta}_n^V : n<\omega\} \le \psi^V(V\Omega)^\Gamma$ is obvious. By 25.13. it follows $\sigma = \sup\{\psi\Delta_n(\aleph_1) : n<\omega\} = \psi(\aleph_1^\Gamma) < V\Omega$. If we assume $\sigma < \psi^V(V\Omega)^\Gamma$, then we obtain an $m \in \omega$ such that $\sigma \in B^V(\underline{\Delta}_m^V) \cap V\Omega$. Hence $\psi(\aleph_1^\Gamma) = \sigma < \psi^V\underline{\Delta}_m^V = \psi\Delta_m(\aleph_1)$, a contradiction.

Now assume (iii). We first observe that $\alpha \le \beta$ and $[\alpha,\beta) \cap B^V(\beta) = \emptyset$ imply $B^V(\alpha) = B^V(\beta)$ and therefore also $\psi^V\alpha = \psi^V\beta$. This is the relativized form of 23.3.(iii) whose proof does not use (Ax_Ω). We easily obtain $B^{V,n}(\beta) \subset B^V(V\Omega^\Gamma) \cap (V\Omega)^\Gamma$ for all $n<\omega$ and all ordinals β by induction on n. Hence $B^V(\beta) \subset B^V(V\Omega^\Gamma) \cap (V\Omega)^\Gamma$. Now let ξ be an arbitrary ordinal. If $\xi < (V\Omega)^\Gamma$ then $\Delta(\xi) := \min\{\eta : \xi \le \eta \in B^V(\xi)\}$ is defined because there is an $n<\omega$ such that $\xi < \Delta_n^V \in B^V(\xi)$. But then $\psi^V\xi = \psi^V(\Delta(\xi)) < V\Omega$ since V locally satisfies (Ax_Ω). If $(V\Omega)^\Gamma \le \xi$, then we obtain $\psi^V\xi = \psi^V(V\Omega)^\Gamma < V\Omega$. So V is a global model of (Ax_Ω).

25.23. **Corollary**

If $\psi(\aleph_1^\Gamma) < V\Omega$, *then* V *is a global model of* (Ax_Ω).

Proof

This follows from 25.19. and 25.22.

As an immediate consequence of 25.23. we obtain

25.24. **Theorem**

The recursive standard interpretation is a global model of (Ax_Ω).

25.25. **Remark**

The term interpretation, however, is not a global model of (Ax_Ω). It is easy to visualize why. If we interpret Ω by $\psi(\aleph_1^\Gamma)$, then $\psi(\aleph^\Gamma)$ comes into the set $B^V(V\Omega^\Gamma)$ which enlarges the segment belonging to $B^V(V\Omega^\Gamma)$. Hence $V\Omega = \psi(\aleph_1^\Gamma) < \psi^V(V\Omega^\Gamma) = \mathrm{Otyp}(B(\aleph_1^\Gamma))$.

25.26. **Exercises**

1. Prove the following statement.
If $V\Omega \in \Gamma'(0)\backslash SC$, then we have $\Gamma_\alpha = \psi^V\alpha$.

2. Describe the behaviour of ψ^V for $V\Omega \in SC \cap \Gamma'(0)$.

3. In this exercise we are going to drop the assumption $V\Omega \in SC$ in the definition of an interpretation. Such a generalized interpretation V for Ω is called good relative to β, if $\psi^V\alpha^V < \sup\{\lambda \in SC: \lambda \leq V\Omega\}$ holds for all $\alpha \in B(\beta) \cap (\beta+1)$.

   (i) Show that the theorems 25.19. and 25.22. still hold for this definition. [Hint: Define $\mathbb{H}^V := \mathbb{H} \cup \{V\Omega\}$, $SC^V := SC \cup \{V\Omega\}$ and modify the normal form conditions by replacing $\mathbb{H}$ by $\mathbb{H}^V$ and SC by SC^V respectively. Then check all lemmas and theorems of the preceeding section.]

   (ii) Show that lemma 25.7.(i) does not hold if we weaken the definition of an interpretation being good relative to β as follows: V is good relative to β, if $\psi^V\alpha^V < V\Omega$ holds for all $\alpha \in B(\beta) \cap (\beta+1)$.

4. Assume $V\Omega =_{NF} \psi\alpha$, then we have $\psi\alpha < \psi^V\alpha^V$.

5. Assume again $V\Omega \in SC$. Prove the equivalence of the following statements.
   (i) V is a global model for (Ax_Ω).
   (ii) It holds $\psi^V((V\Omega)^\Gamma) = \psi(\aleph_1^\Gamma)$ and $\psi^V\xi^V = \psi\xi$ for all $\xi \in B(\aleph_1^\Gamma)$.
   (iii) It holds $\psi^V((V\Omega)^\Gamma) = \psi(\aleph_1^\Gamma)$ and $(\psi\xi)^V = \psi\xi$ for all $\xi \in B(\aleph_1^\Gamma)$.
   (iv) $\psi^V((V\Omega)^\Gamma) = B^V((V\Omega)^\Gamma) \cap V\Omega$.
   (v) $\psi^V((V\Omega)^\Gamma) < V\Omega$.
   (vi) ψ^V is continuous and we have $\psi^V(\xi+1) = (\psi^V\xi)^\Gamma$.

6. Let again be $V\Omega \in SC$. Show that the following statements are equivalent.
   (i) V is a good interpretation.
   (ii) For all $\xi \in B(\aleph_1{}^\Gamma)$ we have $\psi^V \xi^V = \psi\xi$.
   (iii) For all $\xi \in B(\aleph_1{}^\Gamma)$ we have $(\psi\xi)^V = \psi\xi$.
   (iv) For all ordinals ξ it is $B^V(\xi) \subset (V\Omega)^\Gamma$.
   (v) For all ordinals ξ it holds $\psi^V \xi \leq (V\Omega)^\Gamma$.
   (vi) For all $\xi \in B(\aleph_1{}^\Gamma)$ it is $B^V(\xi^V) \subset (V\Omega)^\Gamma$.
   (vii) For all $\xi \in B(\aleph_1{}^\Gamma)$ it holds $\psi^V \xi^V \leq (V\Omega)^\Gamma$.
   (viii) For all $\xi \in B(\aleph_1{}^\Gamma)$ it holds $(\psi\xi)^V \leq (V\Omega)^\Gamma$.

7. Show that the following statements are *false*
   (i) If ψ^V is continuous, then V is a global model for (Ax_Ω).
   (ii) If $\psi^V(\xi+1) \leq (\psi^V\xi)^\Gamma$ holds for all ordinals ξ, then V is a global model for (Ax_Ω).

8. Let V be the term interpretation. Show $\psi^V((V\Omega)^\Gamma) = \mathrm{Otyp}(B(\aleph_1{}^\Gamma)) = (\psi(\aleph_1{}^\Gamma))^\Gamma$.

§26. The semiformal system $\mathbf{ID}_\infty$

In chapter I we obtained an infinitary system $\mathbf{Z}_\Omega$ by adding the cut rule to the validity relation $\models^{\mathfrak{A}}_{\overline{0}}$ for $\mathcal{L}_\Omega$. The ordinal analysis for the system $\mathbf{Z}_1$ of pure number theory then had been obtained by embedding $\mathbf{Z}_1$ into a semiformal subsystem of $\mathbf{Z}_\Omega$. In a similar manner we will now construct an infinitary system $\mathbf{ID}_\infty$ from the validity relation $\mathcal{L}^I_\infty \models_{\overline{0}}$ for $\mathcal{L}^I_\infty$. It should be clear that, in analogy to the situation in the case of $\mathbf{Z}_\Omega$, we will need the cut rule. But this alone will not suffice. Also the (Cl_Ω)-rule will be necessary. Since the cut rule as well as the (Cl_Ω)-rule preserve validity the addition of both rules to the validity relation $\mathcal{L}^I_\infty \models_{\overline{0}}$ will not disturb the soundness of the resulting infinitary system $\mathbf{ID}_\infty$. The necessity of the (Cl_Ω)-rule can be motivated in the following ways.

Our aim is an ordinal analysis of $\mathbf{ID}_1$. To obtain an ordinal analysis it will not suffice just to embed $\mathbf{ID}_1$ into the infinitary system $\mathbf{ID}_\infty$. What we really need is a semiformal subsystem of $\mathbf{ID}_\infty$. A semiformal system is obtained from an infinitary system by restricting the ordinals to a recursive ordinal notation system. Now assume that the infinitary system $\mathbf{ID}_\infty$ is obtained from the validity relation $\mathcal{L}^I_\infty \models_{\overline{0}}$ just by adding the cut rule and try to obtain a semiformal subsystem by restricting the ordinals in $\mathbf{ID}_\infty$ to any recursive ordinal notation system. In order to interpret the fixed points $\underline{I}_A$ of the language $\mathscr{L}_I$ we need class terms of the form $\{x: x \in \underline{I}_A^{<\Omega}\}$. So the notation system has to contain a symbol Ω. In order to obtain a sound interpretation of the derivations of $\mathbf{ID}_\infty$ we in general will have to interpret Ω by the ordinal ω_1^{CK}. But there is no recursive notation system which contains a segment of length ω_1^{CK}. The ordinal Ω viewed from the notation system, however, merely represents the ordertype of the ordinals below Ω of the notation system and this must be an ordinal less than ω_1^{CK}. This has the consequence that it will not longer be possible to import the property $I_A^{<\Omega} = I_A^{\Omega}$ from the real world into the semiformal system as we did it in the completeness proof for $\mathscr{L}_\infty^I \models_{\bar{0}}$ (where we interpreted Ω by ω_1^{CK}). In order to prove $I_A^{<\Omega} = I_A^{\Omega}$ we therefore need the Cl_Ω-rule in the semiformal sytem. Another consequence is that a semiformal subsystem of $\mathbf{ID}_\infty$ necessarily has to be unsound (cf. theorem 26.16).

There is also a more proof theoretical reason for the necessity of something like the Cl_Ω-rule. If we anticipate the result of §29 that the wellordering of the ordinal Γ_0 is provable in the formal theory $\mathbf{ID}_1$ and take into account the results of chapter II the necessity of a new 'infinity axiom' for any semiformal system which allows an ordinal analysis for $\mathbf{ID}_1$ is not surprising. Due to chapter II the autonomous ordinal of any semiformal system whose only infinity axiom is the existence of ω is Γ_0[3]. Since $\mathbf{ID}_1$ proves the wellordering of Γ_0 there cannot be an ordinal analysis of $\mathbf{ID}$ in a semiformal system without an additional infinity axiom. In the case of $\mathbf{ID}_\infty$ the new infinity axiom will be given by the Cl_Ω-rule which may be taken as a defining rule for the ordinal symbol Ω.

Already in the case of the infinitary system $\mathbf{Z}_\Omega$ and its semiformal subsystems there are important differences. We have proved that the infinitary system is complete. Since all valid Π_1^1-sentences have norms below ω_1^{CK} the system $\mathbf{Z}_\Omega$ remains complete even when we interpret Ω by ω_1^{CK}. A semiformal system, however, will always be incomplete. The reason for this incompleteness is the fact that the segment covered by any recursive notation system for the ordinals will always be an ordinal $\beta < \omega_1^{CK}$ while on the other side there is a Π_1^1-sentence F whose norm is an ordinal between β and ω_1^{CK} .This implies that F cannot be provable in the semiformal system $\mathbf{Z}_\beta$ since there are not enough ordinals in

3) Since every infinite recursive ordinal notation system contains the ordinal ω as a segment we can import all properties of ω from the real world into every semiformal subsystem of $\mathbf{Z}_\Omega$. Therefore we do not need an explicit infinitiy axiom for ω in $\mathbf{Z}_\Omega$.

the notation system. The difference between the complete infinitary system and its semiformal system in the case of $\mathbf{ID}_\infty$ will become even more drastic. We already mentioned that a semiformal system containing an infinity axiom for Ω necessarily is unsound. The reason for this unsoundness again is the different meaning of the ordinal Ω in the real world and in the notation system (cf. 26.17).

26.1. Definition

Recall that be $\underline{I}_A^{<\alpha}$ we always denote the class term $\{x: x \in \underline{I}_A^{<\alpha}\}$ and by $\neg\underline{I}_A^{<\alpha}$ the class term $\{x: x \notin \underline{I}_A^{<\alpha}\}$. We introduce the convention that $\Delta(\underline{I}_A^{<\alpha})$ always denotes a positive occurence of $\underline{I}_A^{<\alpha}$ in one of the formulas of Δ while always negative occurences are denoted by by $\Delta(\neg\underline{I}_A^{<\alpha})$.

According to definition 23.3. every $\mathscr{L}_\infty^I$-formula F has the form

$$F \equiv A_{X_1,\dots,X_n,Y_1,\dots,Y_m}(\underline{I}_{A_1}^{<\alpha_1},\dots,\underline{I}_{A_n}^{<\alpha_n},\neg\underline{I}_{B_1}^{<\beta_1},\dots,\neg\underline{I}_{B_m}^{<\beta_m}).$$

where A is a Π_1^1-formula whose only free variables are $X_1,\dots,X_n,Y_1,\dots,Y_m$. So there are only finitely many occurences of class terms $\neg\underline{I}_B^{<\beta}$ in F. For a formula F we define

(i) $\mathrm{stg}_\bigwedge F = \{\beta: \omega < \beta$ and $\neg\underline{I}_B^{<\beta}$ occurs in $F\}$.

By $\mathrm{stg}_\bigwedge F$ we measure the length of all essentially infinite conjunctions occuring in F. We call $\mathrm{stg}_\bigwedge F$ the set of $\bigwedge$-stages of F.

(ii) For a finite formula set Δ we define $\mathrm{stg}_\bigwedge \Delta := \bigcup\{\mathrm{stg}_\bigwedge F: F \in \Delta\}$.

(iii) We shortly write $\Delta \lessapprox \alpha$ instead of $\forall\xi(\xi \in \mathrm{stg}_\bigwedge \Delta \Rightarrow \xi \lessapprox \alpha)$.

26.2. Inductive definition of $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for a finite set Δ of $\mathscr{L}_\infty^I$-formulas.

(Ax) If we have $\vdash_{\overline{0}} \Delta$ according to (Ax1) or (Ax2), then it holds $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for all $\alpha,\rho \in B(\Omega^\Gamma)$, such that $\Delta \lessapprox \alpha$.

($\bigwedge$) If $\Gamma, \bigwedge\{A_\xi: \xi < \lambda\} \subset \Delta$ and $\mathbf{ID}_\infty \overset{f\xi}{\underset{\rho}{\vdash}} \Gamma, A_\xi$ for all $\xi < \lambda = \mathrm{dom} f \le \Omega$, then it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for all $\alpha \in B(\Omega^\Gamma)$ such that $f \ll \alpha$ and $\Delta \lessapprox \alpha$.

($\bigvee$) If $\Gamma, \bigvee\{A_\xi: \xi < \lambda\} \subset \Delta$ and $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Gamma, A_{\xi_0}$ for some $\xi_0 < \lambda$, then we also have $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for all $\alpha \in B(\Omega^\Gamma)$ such that $\alpha_0 \ll \alpha$ and $\Delta \lessapprox \alpha$.

(Cl_Ω) If $\Gamma, t \in \underline{I}_A^{<\Omega} \subset \Delta$ and $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Gamma, t \in \underline{I}_A^{\Omega}$, then it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for all $\alpha \in B(\Omega^\Gamma)$ such that $\alpha_0 \ll \alpha$ and $\Delta \lessapprox \alpha$.

(cut) If $\Gamma \subset \Delta$ and we have $\mathbf{ID}_\infty \overset{\alpha_1}{\underset{\rho}{\vdash}} \Gamma, A$ as well as $\mathbf{ID}_\infty \overset{\alpha_2}{\underset{\rho}{\vdash}} \Gamma, \neg A$ for $\mathrm{rk}(A) < \rho$, then it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for all $\alpha \in B(\Omega^\Gamma)$ such that $\alpha_1, \alpha_2 \ll \alpha$ and $\Delta \lessapprox \alpha$.

In the definition of the relation $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}}$ we restricted the ordinals to ordinals in the set $B(\Omega^\Gamma)$. Since $B(\Omega^\Gamma)$ is a primitive recursive set of ordinals we have introduced a semiformal system.

The following lemmata are immediate consequences of the definition of $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$.

26.3. **Lemma**

If $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$, *then we have* $\Delta \ll \alpha$.

26.4. **Lemma**

If $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$, $\alpha \ll \beta$ *and* $\rho \le \sigma$, *then* $\mathbf{ID}_\infty \frac{|\beta}{\sigma} \Delta$.

The proof of 26.4., which is by induction on α, uses 24.16.(i) and 24.19.(ii).

26.5. **Lemma** (structural rule)

(i) If $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$ *and* $\Delta \subset \Gamma \ll \alpha$, *then it follows* $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Gamma$.

(ii) $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$ *and* $\Gamma \ll \beta$ *imply* $\mathbf{ID}_\infty \frac{|\alpha\#\beta}{\rho} \Delta,\Gamma$.

Proof

(ii) is a consequence of (i) because $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$ by 24.16.(ii) and 26.4. implies $\mathbf{ID}_\infty \frac{|\alpha\#\beta}{\rho} \Delta$. Since $\Gamma \ll \beta$ we have $\Delta,\Gamma \ll \alpha\#\beta$ and obtain the claim by (i).

We prove (i) by induction on α.

If we have $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$ according to (Ax), then, by the hypothesis $\Gamma \ll \alpha$, we also have $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta,\Gamma$ according to (Ax).

In the case of an $\bigwedge$-inference we have the premises $\mathbf{ID}_\infty \frac{|\alpha_\xi}{\rho} \Delta_0, A_\xi$ for all $\xi < \lambda$ together with $\Delta_0, \bigwedge_{\xi<\lambda} A_\xi \subset \Delta \subset \Gamma$ and $\Delta \ll \alpha$. But since $\Gamma \ll \alpha$ we obtain the claim by an $\bigwedge$-inference.

In the case of an $\bigvee$- or (Cl_Ω)-inference we have the premise $\mathbf{ID}_\infty \frac{|\alpha_0}{\rho} \Delta_0, \xi_0$ for some $\xi_0 \le \lambda$ and $\Delta_0, \bigvee_{\xi<\lambda} A_\xi \subset \Delta \subset \Gamma$ as well as $\Delta \ll \alpha$. Because of $\Gamma \ll \alpha$, $\Delta_0, \bigvee_{\xi<\lambda} A_\xi \subset \Delta \subset \Gamma$ and $\alpha_0 \ll \alpha$ we obtain $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Gamma$ by an inference according to the $\bigvee$-or (Cl_Ω)-rule.

In the case of a cut we have the premises $\mathbf{ID}_\infty \frac{|\alpha_1}{\rho} \Delta_1, A$ and $\mathbf{ID}_\infty \frac{|\alpha_2}{\rho} \Delta_1, \neg A$ where $\Delta_1 \subset \Delta \subset \Gamma$. Since we have $\alpha_i \ll \alpha$ for $i = 1,2$ and $\Gamma \ll \alpha$ we obtain $\mathbf{ID}_\infty \frac{|\alpha}{\rho} \Delta$ by a cut.

26.6. **Definition**

Let F be a $\mathscr{L}^1_\Omega$-formula. We define

$$SF := \begin{cases} 0, & \text{if } \mathrm{stg}_\wedge F < \Omega \\ \Omega, & \text{if } \Omega \le \mathrm{stg}_\wedge F, \end{cases}$$

and call SF the level of the formula F. For a finite set Δ of formulas we define

$$S\Delta := \max\{SF \colon F \in \Delta\}.$$

Formulas of level 0 do not contain Ω-branching conjunctions. $\mathcal{L}^{\mathrm{I}}_{\infty}$-formulas of level 0 play the role of Π^1_1-sentences (cf. the remark following lemma 26.15.).

26.7. **Lemma** (collapsing lemma)
If $S\Delta = 0$ *and* $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta$ *for some* $\rho \le \Omega$, *then we already have* $\mathbf{ID}_\infty \vdash^{D\alpha}_{\rho} \Delta$.

Proof by induction on α
Introductory remark: The claim is obvious for $\alpha < \Omega$. We therefore may assume $\Omega \le \alpha$. From $\Delta \lll \alpha$ we obtain $\xi \lll \alpha$ and therefore also $D\xi \lll D\alpha$ for all $\xi \in \mathrm{stg}_\bigwedge \Delta$. If moreover $S\Delta = 0$, then we have $\xi < \Omega$ for all $\xi \in \mathrm{stg}_\bigwedge \Delta$ and obtain $\xi = D\xi \lll D\alpha$. Hence $\Delta \lll D\alpha$.
An immediate consequence of the introductory remark is that for an axiom $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta$ with $S\Delta = 0$ we also have $\mathbf{ID}_\infty \vdash^{D\alpha}_{\rho} \Delta$ as an axiom.
Now suppose that the last inference in the derivation of Δ is an inference

(S) $\mathbf{ID}_\infty \vdash^{\alpha_\eta}_{\rho} \Delta_\eta \Rightarrow \mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta$.

If (S) is an $\bigvee$- or Cl_Ω-inference, then we have $\mathrm{stg}_\bigwedge \Delta_\eta \subset \mathrm{stg}_\bigwedge \Delta$, $\Delta \lll \alpha$ and $\alpha_\eta \ll \alpha$. But then we also obtain $S\Delta_\eta = 0$, $\Delta \lll D\alpha$ and $D\alpha_\eta \ll D\alpha$. By the induction hypothesis it then follows $\mathbf{ID}_\infty \vdash^{D\alpha_\eta}_{\rho} \Delta_\eta$ from which we obtain $\mathbf{ID}_\infty \vdash^{D\alpha}_{\rho} \Delta$ by an inference (S).
If (S) is an $\bigwedge$-inference, then we have $\mathrm{stg}_\bigwedge \Delta_\eta \subset \mathrm{stg}_\bigwedge \Delta$, $\Delta \lll \alpha$ and $\lambda\eta.\alpha_\eta \ll \alpha$. From $S\Delta = 0$ it follows $\lambda := \mathrm{dom}\, \lambda\eta.\alpha_\eta < \Omega$ since we either have $\lambda \le \omega$ or $\lambda \in \mathrm{stg}_\bigwedge \Delta$. In both cases we have $\lambda \lll \alpha$. Together with $\xi < \lambda < \Omega$ it follows $\xi \ll \lambda \lll \alpha$. Since $\lambda\xi.\alpha_\xi \ll \alpha$ this implies $\alpha_\xi \ll \alpha$. Hence $D\alpha_\xi \ll D\alpha$ for all $\xi < \lambda$, i.e. $\lambda\xi.D\alpha_\xi \ll D\alpha$ by 24.19. By the induction hypothesis we have $\mathbf{ID}_\infty \vdash^{D\alpha_\eta}_{\rho} \Delta_\eta$ for all $\eta < \lambda$, and obtain $\mathbf{ID}_\infty \vdash^{D\alpha}_{\rho} \Delta$ by an $\bigwedge$-inference.
If (S) is a cut, then we have the premises $\mathbf{ID}_\infty \vdash^{\alpha_1}_{\rho} \Gamma, F$ and $\mathbf{ID}_\infty \vdash^{\alpha_2}_{\rho} \Gamma, \neg F$ such that $\mathrm{rk}(F) = \mathrm{rk}(\neg F) < \rho \le \Omega$ and $\Gamma \subset \Delta$. Since $\rho \le \Omega$ F neither contains a term of the form $\underline{I}^{<\Omega}_{A}$ nor $\neg \underline{I}^{<\Omega}_{A}$. Hence $S(F) = S(\neg F) = 0$. Now we may apply the induction hypothesis and obtain $\mathbf{ID}_\infty \vdash^{D\alpha_1}_{\rho} \Gamma, F$ and $\mathbf{ID}_\infty \vdash^{D\alpha_2}_{\rho} \Gamma, \neg F$. Since $D\alpha_i \ll D\alpha$ for $i = 1,2$ we obtain $\mathbf{ID}_\infty \vdash^{D\alpha}_{\rho} \Delta$ by a cut.

26.8. **Lemma** (persistency lemma)
If $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta_X(\underline{I}^{<\lambda}_{A})$ *and* $\lambda \le \mu \le \Omega$, *then* $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta_X(\underline{I}^{<\mu}_{A})$.

Proof by induction on α
We first observe that we always have $\mathrm{stg}_\bigwedge \Delta(\underline{I}^{<\lambda}_{A}) = \mathrm{stg}_\bigwedge \Delta(\underline{I}^{<\mu}_{A})$. If we have $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta(\underline{I}^{<\lambda}_{A})$ according to (Ax) we therefore also have $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta(\underline{I}^{<\lambda}_{A})$ as an

axiom. If the main formula of the last inference is different from $t \in \underline{I}_A^{<\lambda}$, then we obtain the claim immediately from the induction hypothesis. Let $t \in \underline{I}_A^{<\lambda}$ be the main formula of the last inference. Since we may assume that $\lambda < \Omega$, this inference must be an $\bigvee$-inference with premise $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\lambda}), t \in \underline{I}_A^{\xi}$ for some $\xi < \lambda \le \mu$. By the induction hypothesis we obtain $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\mu}), t \in \underline{I}_A^{\xi}$ and it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\mu}), t \in \underline{I}_A^{<\mu}$ by an $\bigvee$-inference.

26.9. **Lemma** (boundedness lemma)

If $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta_X(\underline{I}_A^{<\lambda})$ *and* $\alpha \le \Omega$ *, then we have* $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta_X(\underline{I}_A^{<\alpha})$.

Proof

We already mentioned that we always have $stg_\bigwedge(\Delta(\underline{I}_A^{<\lambda})) = stg_\bigwedge(\Delta(\underline{I}_A^{<\alpha}))$. If $t \in \underline{I}_A^{<\lambda}$ is not the main formula of the last inference, then the claim follows immediately from the induction hypothesis and the persistency lemma. If $t \in \underline{I}_A^{<\lambda}$ is the main formula of the last inference, then it is an $\bigvee$- or Cl_Ω-inference. Then we have the premise $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\lambda}), t \in \underline{I}_A^{\xi}$ for some $\xi \le \lambda$. Recall that $t \in \underline{I}_A^{\xi}$ is an abbreviation for $A(\underline{I}_A^{<\xi}, t)$. By the induction hypothesis we therefore obtain $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\alpha_0}), A(\underline{I}_A^{<\alpha_0}, t)$, i.e. $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0(\underline{I}_A^{<\alpha_0}), t \in \underline{I}_A^{\alpha_0}$, for some $\alpha_0 < \alpha$. By persistency and an $\bigvee$-inference it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta(\underline{I}_A^{<\alpha}), t \in \underline{I}_A^{<\alpha}$.

As Corollary we then obtain

26.10. **Lemma**

If $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ *holds for some* $\alpha < \Omega$, *then we may eliminate all* Cl_Ω*-inferences in this derivation.*

Proof by induction on α.

If the last inference is not an inference according to the Cl_Ω-rule, then the claim is immediate from the induction hypothesis. Therefore suppose that it is an Cl_Ω-inference

$$\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0, t \in \underline{I}_A^{\Omega} \text{ and } \Delta_0, t \in \underline{I}_A^{<\Omega} \subset \Gamma \Rightarrow \mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma.$$

By the boundedness lemma we obtain $\mathbf{ID}_\infty \overset{\alpha_0}{\underset{\rho}{\vdash}} \Delta_0, t \in \underline{I}_A^{\alpha_0}$ for $\alpha_0 < \alpha < \Omega$. By an $\bigvee$-inference it then follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Gamma$.

If Δ is a set of the form $\Delta[\underline{I}_{A_1}^{<\Omega}, \dots, \underline{I}_{A_n}^{<\Omega}]$ without further occurences of class terms of the form $\underline{I}_A^{<\Omega}$, then we denote by Δ^α the set $\Delta[\underline{I}_{A_1}^{<\alpha}, \dots, \underline{I}_{A_n}^{<\alpha}]$.

26.11. **Theorem**

(i) If $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for $\rho \le \Omega$ and $S\Delta = 0$, then it follows $\mathbf{Z}_\Omega \overset{D\alpha}{\underset{\rho}{\vdash}} \Delta^{D\alpha}$.

(ii) From $\mathbf{Z}_\Omega \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ for α and ρ in $B(\Omega^\Gamma)$, we obtain also $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$.

Proof

(i): first we obtain $\mathrm{ID}_\infty \overset{D\alpha}{\underset{\rho}{\vdash}} \Delta$ by the collapsing lemma and then $\overset{D\alpha}{\underset{\rho}{\vdash}} \Delta^{D\alpha}$ by the boundedness lemma. Because of $S\Delta = 0$ the set Δ does not contain class terms of the shape $\neg \underline{I}_A^{<\Omega}$ which implies that all formulas in $\Delta^{D\alpha}$ are $\mathscr{L}_\Omega$-formulas. By 26.10. we may eliminate all Cl_Ω-inferences in this derivation. Since $\rho \le \Omega$ an easy induction on $D\alpha$ now shows that this derivation only contains formulas of $\mathscr{L}_\Omega$.

(ii): This claim is trivial since $\mathbf{Z}_\Omega$ is a subsystem of $\mathbf{ID}_\infty$.

26.12. **Corollary** (Soundness for $\mathbf{ID}_\Omega$)

If we denote by $\mathbf{ID}_\Omega$ the subsystem of $\mathbf{ID}_\infty$ which only contains derivations of the form $\mathbf{ID}_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta$ where $\alpha, \rho \le \Omega$ and $S\Delta = 0$, then $\mathbf{ID}_\Omega \overset{\alpha}{\underset{\rho}{\vdash}} F$ implies $\mathbb{N} \models F$.

Proof

The proof is immediate from 26.11. and 12.1.

26.13. **Corollary** (Boundedness theorem)

If $\mathbf{ID}_\infty \overset{\alpha}{\underset{\Omega}{\vdash}} \underline{n} \in \underline{I}_A^{<\Omega}$, then we have $|n|_A < D\alpha$.

Proof

It is $S(\underline{n} \in \underline{I}_A^{<\Omega}) = 0$. By 26.11. and 26.12. it therefore follows $\mathbb{N} \models \underline{n} \in \underline{I}_A^{<D\alpha}$. By 22.7. this implies $n \in I_A^{<D\alpha}$, i.e. $|n|_A < D\alpha$.

26.14. **Lemma**

If F is an $\mathscr{L}_\infty^I$-sentence F such that $rk(F) < \psi(\Omega^\Gamma)$ and $\mathbb{N} \models F$, then we have $\mathbf{ID}_\infty \overset{rk(F)}{\underset{0}{\vdash}} F$.

Proof

The proof is by induction on $rk(F)$.

If F is atomic, then F is of the shape $\underline{P}t_1 \ldots t_n$ since F is an $\mathscr{L}_\infty^I$-sentence. From $\mathbb{N} \models F$ we then obtain $\mathbf{ID}_\infty \overset{0}{\underset{0}{\vdash}} F$ by (Ax1).

If F has the shape $\bigwedge_{\xi<\lambda} A_\xi$ and we have $\mathbb{N} \models F$, then we obtain $\mathbb{N} \models A_\xi$ for all $\xi < \lambda$.

By the induction hypothesis this implies $\mathbf{ID}_\infty \vdash_0^{rk(A_\xi)} A_\xi$ for all $\xi<\lambda$. Because of $rk(A_\xi)<rk(F)<\Psi(\Omega^\Gamma)$ it follows $rk(A_\xi)\ll rk(F)$ for all $\xi<\lambda$ which implies $\lambda\xi.rk(A_\xi)\ll rk(F)$. By an $\bigwedge$-inference it follows $\mathbf{ID}_\infty \vdash_0^{rk(F)} F$.

If F is a sentence $\bigvee_{\xi<\lambda} A_\xi$ with $\mathbb{N}\models F$, then there is a $\xi<\lambda$ such that $\mathbb{N}\models A_\xi$. By the induction hypothesis it follows $\mathbf{ID}_\infty \vdash_0^{rk(A_\xi)} A_\xi$. Since $rk(A_\xi)\ll rk(F)$ this implies $\mathbf{ID}_\infty \vdash_0^{rk(F)} F$ by an $\bigvee$-inference.

An inspection of the proof of 26.14. shows that the proof does not depend upon the fact that the ordinals in the derivation belong to a recursive notation system. It makes the proof even clumsier. We easily may simplify it in order to obtain the result that for true $\mathcal{L}^I_\Omega$-sentences F we have $\mathcal{L}^I_\infty \vdash_0^{rk(F)} F$. So we obtain as a corollary the following theorem.

26.15. **Theorem**

For any true $\mathcal{L}^I_\infty$ -sentence F of level 0 we have $|F|\le rk(F)$.

Theorem 26.15. has an interesting consequence. As shown in exercise 26.18.2 for any Π^1_1-sentence F there is an $\mathcal{L}^I_\infty$ -sentence F_∞ of level 0 such that $\mathbb{N}\models F \leftrightarrow F_\infty$. The problem to find the shortest sentence F_∞ which has this property is by 26.15. closely connected to the problem of ordinal analysis. For Π^1_1-sentences F whose validity is provable in $\mathbf{ID}_1$ it can be shown that the coresponding $\mathcal{L}^I_\infty$ -formula F_∞ always has a rank below $\Psi(\varepsilon_{\Omega+1})$ (cf. exercises).

As a further consequence of 26.15. we obtain the unsoundness of the system $\mathbf{ID}_\infty$.

26.16. **Theorem**

The semi formal system $\mathbf{ID}_\infty$ is unsound.

Proof

The proof needs a result of recursion theory. There one may prove the existence of an arithmetically definable inductive definition Γ such that $|\Gamma| = \omega_1^{CK}$. Since $\psi(\Omega^\Gamma)<\omega_1^{CK}$ there is an $n\in I_\Gamma$ such that $|n| = \psi(\Omega_\Gamma)$, i.e. $n\notin I_\Gamma^{<\psi(\Omega^\Gamma)}$. By lemma 26.14. we have $\mathbf{ID}_\infty \vdash_0^{\alpha_\xi} \underline{n}\notin \underline{I}_\Gamma^\xi$ for all $\xi\in B(\Omega^\Gamma)\cap\Omega$, where we defined $\alpha_\xi := rk(n\notin I_\Gamma^\xi) \le (m+1)\cdot\xi+m$ and $m := rk(\Gamma)<\omega$. Since $\lambda\xi.(m+1)\cdot\xi+m\ll\Omega$ we obtain $\mathbf{ID}_\infty \vdash_0^\Omega \underline{n}\notin \underline{I}_\Gamma^{<\Omega}$. On the other hand, however, it holds $\mathbb{N}\models \underline{n}\in \underline{I}_\Gamma^{<\Omega}$.

26.17. Remark

The proof of lemma 26.16. shows that the reason for the unsoundness of ID_∞ is the fact that the segment of the notation system is bounded below ω_1^{CK}. But then the unsoundness of ID_∞ only can be remedied by allowing the whole segment of ordinals up to ω_1^{CK}. Since this is impossible for any recursive notation system all semiformal systems which are strong enough to allow the embedding of the formal system ID_1 necessarily have to be unsound.

26.18. Exercises

1. Assume that $F[X_1,\dots,X_n]$ is a Π^1_1-formula without further occurences of set variables. Show that $X_1,\dots,X_n$ occur positively in $F[X_1,\dots,X_n]$ if and only if $F_{X_1,\dots,X_n}[I_A^{<\Omega},\dots,I_A^{<\Omega}]$ is a formula of level 0.

2. Show that for every Π^1_1-sentence F in the language $\mathscr{L}$ there is a sentence F_∞ *(not containing set parameters)* of level 0 in the language $\mathscr{L}^I_\Omega$ such that $\mathbb{N} \models F \leftrightarrow F_\infty$ and vice versa. [Hint: Use exercises 20.9.].

§ 27. Cut elimination for $ID^{\Omega+\omega}_{\Omega\Gamma}$

27.1. Lemma

If $ID_\infty \overset{\alpha}{\underset{\Omega}{\vdash}} \Delta$ *and* $S\Delta = 0$, *then it follows* $ID_\infty \overset{\Omega+\alpha}{\underset{0}{\vdash}} \Delta$.

Proof by induction on α

If the last inference is not a cut, then we obtain the claim immediately from the lemmata 24.16. (iv) and 24.19 (i) and the induction hypothesis. We therefore assume that the last inference is a cut

$$ID_\infty \overset{\alpha_1}{\underset{\Omega}{\vdash}} \Gamma, A \text{ and } ID_\infty \overset{\alpha_2}{\underset{\Omega}{\vdash}} \Gamma, \neg A \Rightarrow ID_\infty \overset{\alpha}{\underset{\Omega}{\vdash}} \Delta$$

where $\Gamma \subset \Delta$ and $\alpha_1,\alpha_2 \ll \alpha$. By the induction hypothesis we obtain $ID_\infty \overset{\Omega+\alpha_1}{\underset{0}{\vdash}} \Gamma, A$ and $ID_\infty \overset{\Omega+\alpha_2}{\underset{0}{\vdash}} \Gamma, \neg A$. Because of $rk(A) < \Omega$ we also have $S(A) = S(\neg A) = 0$. By 26.11. and the persistency lemma it follows $Z_\Omega \overset{D(\Omega+\alpha_1)}{\underset{0}{\vdash}} \Gamma^\beta, A$ and $Z_\Omega \overset{D(\Omega+\alpha_2)}{\underset{0}{\vdash}} \Gamma^\beta, \neg A$, for some $\beta < \Omega$ such that $D(\Omega+\alpha_i) \le \beta$ holds for $i = 1,2$.

For $\sigma := rk(A)$ we obtain $Z_\Omega \overset{D(\Omega+\alpha_1)\#D(\Omega+\alpha_2)}{\underset{\sigma}{\vdash}} \Gamma^\beta$ by the elimination lemma and then $Z_\Omega \overset{\varphi\sigma(D(\Omega+\alpha_1)\#D(\Omega+\alpha_2))}{\underset{0}{\vdash}} \Gamma^\beta$ by 18.5., the corollary to the second elimination theorem. By 26.11. and the structural rule it follows $ID_\infty \overset{\varphi\sigma(D(\Omega+\alpha_1)\#D(\Omega+\alpha_2))}{\underset{0}{\vdash}} \Delta^\beta$. Now it is $SC_\Omega(\varphi\sigma(D(\Omega+\alpha_1)\#D(\Omega+\alpha_2))) = SC_\Omega(\sigma) \cup \{D(\Omega+\alpha_1), D(\Omega+\alpha_2)\}$. Since $\alpha_i \ll \alpha$ we have $D(\Omega+\alpha_i) < D(\Omega+\alpha)$ for $i = 1,2$. From $stg_\wedge A \lesssim \alpha_1 \ll \alpha$ and $stg_\wedge \neg A \lesssim \alpha_2 \ll \alpha$

and $\sigma = \lambda+n$ for some $\lambda \in stg_{\wedge}(A) \cup stg_{\wedge}(\neg A) \cup \{0\}$ we obtain $SC_{\Omega}(\sigma) < D(\Omega+\alpha)$. Hence $\varphi\sigma(D(\Omega+\alpha_1)\#D(\Omega+\alpha_2)) \ll \Omega+\alpha$ and it follows $\mathbf{ID}_\infty \vdash^{\Omega+\alpha}_{0} \Delta$ by the persistency lemma.

We learn from lemma 27.1. that our next aim must be to decrease the cut rank of $\mathbf{ID}_\infty$-derivations to Ω. The first step is the following lemma.

27.2.**Lemma** (Predicative elimination lemma for $\mathbf{ID}_\infty$)
If $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, A$, $\mathbf{ID}_\infty \vdash^{\beta}_{\rho} \Gamma, \neg A$ *and* $rk(A) = \rho \neq \Omega$, *then it follows* $\mathbf{ID}_\infty \vdash^{\alpha\#\beta}_{\rho} \Delta, \Gamma$.

Using the lemmata 24.16. and 24.19. we may proof 27.2. literally as 12.2. Crucial for the proof is the hypothesis $rk(A) \neq \Omega$ since this assures that A cannot be the main formula of a (Cl_{Ω})-inference.
As a consequence of 27.2. we obtain

27.3. **Lemma**
If $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho+1} \Delta$ *holds for some* $\rho \neq \Omega$, *then it follows* $\mathbf{ID}_\infty \vdash^{\omega^\alpha}_{\rho} \Delta$.

Using 27.2. and 24.16. the proof of 27.3. is literally the proof of 12.3.

By 27.3. we already may decrease the cut rank of $\mathbf{ID}_\infty$- derivations to $\Omega+1$. The last and crucial step is the elimination of a cut of rank Ω. This step will be achieved by the following impredicative elimination lemma. To prepare this lemma we first need an inversion lemma of the following kind.

27.4. **Lemma** (Inversion lemma)
From $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \bigwedge_{\xi<\lambda} A_\xi$ *it follows* $\mathbf{ID}_\infty \vdash^{\alpha\#\xi}_{\rho} \Delta, A_\xi$ *for all* $\xi < \lambda$ *and* $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, A_\xi$ *for all* $\xi \ll \lambda$.

Proof
We show both claims by induction on α. If $\bigwedge_{\xi<\lambda} A_\xi$ is not the main formula of the last inference, then the claim either – in the case of an axiom – is trivial or an immediate consequence of the induction hypothesis. We therefore assume that $\bigwedge_{\xi<\lambda} A_\xi$ is the main formula of the last inference. Then we have the premises $\mathbf{ID}_\infty \vdash^{f\xi}_{\rho} \Gamma, A_\xi$ for all $\xi < \lambda$ with $f \ll \alpha$ such that $\Gamma, \bigwedge_{\xi<\lambda} A_\xi \subset \Delta$ and obtain $\mathbf{ID}_\infty \vdash^{f\xi\#\xi}_{\rho} \Gamma, A_\xi$ or $\mathbf{ID}_\infty \vdash^{f\xi}_{\rho} \Gamma, A_\xi$ respectively from the induction hypothesis. Since $\alpha \neq 0$ we have $\xi \ll \alpha\#\xi$ for all $\xi < \lambda$ and obtain $SC_{\Omega}(f\xi) < D(\alpha\#\xi)$. Hence $SC_{\Omega}(f\xi\#\xi) =$

$SC_\Omega(f\xi)\cup SC_\Omega(\xi) < D(\alpha \# \xi)$. Since also $f\xi \# \xi < \alpha \# \xi$ this implies $f\xi \# \xi \ll \alpha \# \xi$. Hence $ID_\infty \vdash^{\alpha\#\xi}_{\rho} \Delta, A_\xi$ by 26.4. and the structural rule. For the second claim we assume $\xi \ll \lambda$. Either we have $\lambda \in stg_\wedge(\Delta, {}_{\xi<\lambda}\bigwedge A_\xi) \lesssim \alpha$ or $\lambda \le \omega$. In the first case we obtain $\xi \ll \alpha$ and since $f \ll \alpha$ also $f\xi \ll \alpha$. In the second case we either have $\alpha < \omega$ or again $\xi \ll \alpha$. In both cases it follows $f\xi \ll \alpha$. From the induction hypothesis $ID_\infty \vdash^{f\xi}_{\rho} \Gamma, A_\xi$ we therefore obtain $ID_\infty \vdash^{\alpha}_{\rho} \Delta, A_\xi$ by 26.4. and the structural rule.

27.5. **Lemma** (Impredicative elimination lemma)

Let Δ *be a finite set of formulas of level* 0. *Then* $ID_\infty \vdash^{\alpha}_{\Omega} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}, t \in I^{<\Omega}_{A}$ *and* $ID_\infty \vdash^{\beta}_{\Omega} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}, t \notin I^{<\Omega}_{A}$ *imply* $\vdash^{\alpha\#\beta\#\Omega\cdot(n+1)}_{\Omega} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}$.

Proof

We introduce the following abbreviations. For $1 \le k \le n$ let Δ_k be the set $\Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_k \notin I^{<\Omega}_{A_k}$. If $\boldsymbol{\xi} = (\xi_1, \ldots, \xi_k)$ is any k-tuple, then we denote by $\Delta^{\boldsymbol{\xi}}_k$ the set $\Delta, t_1 \notin I^{<\xi_1}_{A_1}, \ldots, t_k \notin I^{<\xi_k}_{A_k}$. $\Sigma\boldsymbol{\xi}$ is a shorthand for $\xi_1 \# \ldots \# \xi_k$.

Now let $\boldsymbol{\xi}$ be an arbitrary n-tuple of ordinals $< \Omega$. From the hypothesis

(1) $ID_\infty \vdash^{\alpha}_{\Omega} \Delta_n,\ t \in I^{<\Omega}_{A}$

we obtain by the inversion lemma

(2) $ID_\infty \vdash^{\alpha\#\Sigma\boldsymbol{\xi}}_{\Omega} \Delta^{\boldsymbol{\xi}}_n, t \in I^{<\Omega}_{A}$.

Since $S(\Delta^{\boldsymbol{\xi}}_n, t \in I^{<\Omega}_{A}) = 0$ we obtain by (2), the collapsing and the boundedness lemma

(3) $ID_\infty \vdash^{D(\alpha\#\Sigma\boldsymbol{\xi})}_{\Omega} \Delta^{\boldsymbol{\xi}}_n,\ t \in I^{<D(\alpha\#\Sigma\boldsymbol{\xi})}_{A}$.

From the hypothesis

(4) $ID_\infty \vdash^{\beta}_{\Omega} \Delta_n, t \notin I^{<\Omega}_{A}$

we obtain by the inversion lemma

(5) $ID_\infty \vdash^{\beta\#\Sigma\boldsymbol{\xi}\#\eta}_{\Omega} \Delta^{\boldsymbol{\xi}}_n, t \notin I^{\eta}_{A}$

for all $\eta < D(\alpha\#\Sigma\boldsymbol{\xi})$. Since $\lambda\eta.\beta\#\Sigma\boldsymbol{\xi}\#\eta \ll \beta\#\Sigma\boldsymbol{\xi}\#D(\alpha\#\Sigma\boldsymbol{\xi})$ we obtain by (5) and an $\bigwedge$-inference

(6) $ID_\infty \vdash^{\beta\#\Sigma\boldsymbol{\xi}\#D(\alpha\#\Sigma\boldsymbol{\xi})}_{\Omega} \Delta^{\boldsymbol{\xi}}_n, t \notin I^{<D(\alpha\#\Sigma\boldsymbol{\xi})}_{A}$.

Now we have $D(\alpha\#\Sigma\boldsymbol{\xi}) \lesssim \beta\#\Sigma\boldsymbol{\xi}\#D(\alpha\#\Sigma\boldsymbol{\xi}) \ll \alpha\#\beta\#\Omega\#\Sigma\boldsymbol{\xi}$ and obtain

(7) $ID_\infty \vdash^{\alpha\#\beta\#\Omega\#\Sigma\boldsymbol{\xi}}_{\Omega} \Delta^{\boldsymbol{\xi}}_n$

from (3) and (6) by a cut.

If we define $\boldsymbol{\xi}_i := (\xi_1, \ldots, \xi_{n-i})$, then we have $\lambda\xi.\ \alpha\#\beta\#\Omega\cdot i\#\Sigma\boldsymbol{\xi}_i\#\xi \ll \alpha\#\beta\#\Omega\cdot(i+1)\#\Sigma\boldsymbol{\xi}_i$. From (7) we therefore obtain

(8) $ID_\infty \vdash^{\alpha\#\beta\#\Omega\cdot 2\#\Sigma\boldsymbol{\xi}_1}_{\rho} \Delta^{\boldsymbol{\xi}_1}_{n-1}, t_n \notin I^{<\Omega}_{A_n}$

by an $\bigwedge$-inference. By iteration we then have

(9) $ID_\infty \vdash^{\alpha \# \beta \# \Omega\cdot(i+1) \# \Sigma\xi_i}_{\Omega} \Delta^{\xi_i}_{n-i}, t_{n-i+1} \notin I^{<\Omega}_{A_{n-i+1}}, \ldots, t_n \notin I^{<\Omega}_{A_n}$

for all $i \in \{1,\ldots,n\}$. For $i = n$ this is

(10) $ID_\infty \vdash^{\alpha \# \beta \# \Omega\cdot(n+1)}_{\Omega} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}$.

27.6. **Theorem** (Impredicative elimination theorem)
$ID_\infty \vdash^{\alpha}_{\Omega+1} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}$ implies $ID_\infty \vdash^{\omega^{\Omega+\alpha}}_{\Omega} \Delta, t_1 \notin I^{<\Omega}_{A_1}, \ldots, t_n \notin I^{<\Omega}_{A_n}$ for all finite sets Δ of $\mathcal{L}^I_\Omega$ -formulas of level 0.

Proof by induction on α.
If the last inference is not a cut of rank Ω, then the claim follows immediately from the induction hypothesis, 24.16. and 24.19. Now assume that the last inference is a cut

$$ID_\infty \vdash^{\alpha_1}_{\Omega+1} \Gamma, t \in I^{<\Omega}_A \text{ and } ID_\infty \vdash^{\alpha_2}_{\Omega+1} \Gamma, t \notin I^{<\Omega}_A \Rightarrow ID_\infty \vdash^{\alpha}_{\Omega+1} \Delta_n$$

of rank Ω, where Δ_n is the abbreviation defined in the proof of 27.5. By the induction hypothesis it follows

(1) $ID_\infty \vdash^{\omega^{\Omega+\alpha_1}}_{\Omega} \Gamma, t \in I^{<\Omega}_A$ and $ID_\infty \vdash^{\omega^{\Omega+\alpha_2}}_{\Omega} \Gamma, t \notin I^{<\Omega}_A$.

From (1) we obtain by the impredicative elimination lemma

(2) $ID_\infty \vdash^{\omega^{\Omega+\alpha_1} \# \omega^{\Omega+\alpha_2} \# \Omega\cdot(n+1)}_{\Omega} \Gamma$.

Since $\omega^{\Omega+\alpha_1} \# \omega^{\Omega+\alpha_2} \# \Omega\cdot(n+1) \ll \omega^{\Omega+\alpha}$ and $\Delta_n \subset \Gamma$ we obtain from (2) by 26.4. and the structural rule

(3) $ID_\infty \vdash^{\omega^{\Omega+\alpha}}_{\Omega} \Delta_n$.

27.7. **Lemma**
If $S\Delta = 0$ and $ID_\infty \vdash^{\alpha}_{\Omega+1} \Delta$, then it follows $ID_\infty \vdash^{\omega^{\Omega+\alpha}}_{0} \Delta$.

Proof
From $ID_\infty \vdash^{\alpha}_{\Omega+1} \Delta$ and $S\Delta = 0$ we obtain $ID_\infty \vdash^{\omega^{\Omega+\alpha}}_{\Omega} \Delta$ by the impredicative elimination theorem. By lemma 27.1. this implies $ID_\infty \vdash^{\omega^{\Omega+\alpha}}_{0} \Delta$.

By ID^β_α we will denote the semiformal system whose language only contains formulas of ranks $\ll \alpha$ and in which only derivations of lengths $\ll \alpha$ and cut ranks $< \beta$ are allowed. $ID^\beta_\alpha \vdash \Delta$ means that there are ordinals $\xi \ll \alpha$ and $\rho < \beta$ such that $ID_\infty \vdash^{\xi}_{\rho} \Delta$

27.8. **Theorem** (Cut elimination for $ID^{\Omega+\omega}_{\Omega\Gamma}$)
If $S\Delta = 0$ and $ID_\infty \vdash^{\alpha}_{\Omega+n} \Delta$, then we already have $ID_\infty \vdash^{D(\omega_n(\Omega+\alpha))}_{0} \Delta$.

Proof by induction on n.

For n = 0 this is obvious from 27.1. by the collapsing lemma. Now we assume $n = n_0'$. If $n_0 = 0$, then we obtain fom $ID_\infty \overset{\alpha}{\underset{\Omega+1}{\vdash}} \Delta$ by 27.7. $ID_\infty \overset{\omega^{\Omega+\alpha}}{\underset{\Omega}{\vdash}} \Delta$ and by 27.1. and the collapsing lemma $ID_\infty \overset{D(\omega^{\Omega+\alpha})}{\underset{0}{\vdash}} \Delta$. If $n_0 \neq 0$, then $ID_\infty \overset{\alpha}{\underset{\Omega+n_0+1}{\vdash}} \Delta$ implies $ID_\infty \overset{\omega^\alpha}{\underset{\Omega+n_0}{\vdash}} \Delta$ by 27.3. By the induction hypothesis it then follows

(*) $ID_\infty \overset{D(\omega_{n_0}(\Omega+\omega^\alpha))}{\underset{0}{\vdash}} \Delta$.

Since $\omega_{n_0}(\Omega+\omega^\alpha) \lll \omega_n(\Omega+\alpha)$ we obtain $D(\omega_{n_0}(\Omega+\omega^\alpha)) \lll D(\omega_n(\Omega+\alpha))$ which together with (*) implies the claim.

27.9. **Corollary** (Elimination theorem for $ID^{\Omega+\omega}_{\varepsilon_{\Omega+1}}$)

If $ID^{\Omega+\omega}_{\varepsilon_{\Omega+1}} \vdash F$ *holds for a formula* F *of level* 0, *then there is some* $\alpha < \psi\varepsilon_{\Omega+1}$ *such that* $ID_\infty \overset{\alpha}{\underset{0}{\vdash}} F$.

Proof

$ID^{\Omega+\omega}_{\varepsilon_{\Omega+1}} \vdash F$ implies the existence of some ordinal $\beta \ll \varepsilon_{\Omega+1}$ and $n < \omega$ such that $ID_\infty \overset{\beta}{\underset{\Omega+n}{\vdash}} F$. By the cut elimination theorem 27.8. it then follows $ID_\infty \overset{D(\omega_n(\Omega+\beta))}{\underset{0}{\vdash}} F$. $\beta \ll \varepsilon_{\Omega+1}$ implies $\omega_n(\Omega+\beta) \ll \varepsilon_{\Omega+1}$ and this entails $D(\omega_n(\Omega+\beta)) < D(\varepsilon_{\Omega+1}) = \psi\varepsilon_{\Omega+1}$.

27.10. **Exercise**

Show the following soundness theorem.

$S(F) = 0$ and $ID_\infty \overset{\alpha}{\underset{\Omega+n}{\vdash}} F$ imply $\mathbb{N} \models F$.

§28. Embedding of ID_1 *into* $ID^{\Omega+\omega}_{\varepsilon_{\Omega+1}}$

28.1 **Lemma** ($\bigvee$-importation and $\bigvee$-exportation for ID_∞)

(i) $ID_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta, A_1, \ldots, A_n$ *implies* $ID_\infty \overset{\alpha+n}{\underset{\rho}{\vdash}} \Delta, A_1 \vee \ldots \vee A_n$.

(ii) $ID_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta, A_1 \vee \ldots \vee A_n$ *implies* $ID_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta, A_1, \ldots, A_n$.

Proof.

The proofs are literally the same as those of 10.8.

28.2. **Lemma**

If $ID_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta_x(s)$ *and* t *is a term which is equivalent to* s, *then we also have* $ID_\infty \overset{\alpha}{\underset{\rho}{\vdash}} \Delta_x(t)$.

Proof

The proof is straightforward by induction on α.

28.3. Definition

For a finite set $M := \{\xi_1,\ldots,\xi_n\}$ of ordinals we define $\sum^{\#} M := \xi_1 \# \ldots \# \xi_n$.

(i) For an $\mathscr{L}^I_\infty$-formula F it is $no(F) := 1 + \sum^{\#}\{\xi : \xi \in stg_\wedge F \cup stg_\wedge \neg F\} \# 2\,rk\,F$.

(ii) For a finite set Δ of formulas we put $no(\Delta) := \sum^{\#}\{no(F) : F \in \Delta\}$.

28.4. **Lemma** (Monotonicity lemma for $\mathbf{ID}_\infty$)

Suppose that $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \neg F_x(\underline{n}), G_x(\underline{n})$ *holds for all* $n \in \mathbb{N}$. *Then we also obtain* $\mathbf{ID}_\infty \vdash^{\alpha \# no(A)}_{\rho} \Delta, \neg A_X(F), B_X(G)$ *for all equivalent* X*-positive formulas* A *and* B.

Proof

The proof is by induction on rk A. First we recall that two formulas A and B are equivalent if there are a formula F and terms $s_1,\ldots,s_n$, $t_1,\ldots,t_n$ such that s_k and t_k are equivalent for $k \in \{1,\ldots,n\}$ and it is $A \equiv F_{x_1,\ldots x_n}[s_1,\ldots,s_n]$ as well as $B \equiv F_{x_1,\ldots x_n}[t_1,\ldots,t_n]$.

If A is an atomic formula without occurrences of X, then we have $\neg A_X(F) \equiv \neg A$ and $B_X(G) \equiv B$. It holds $\vdash_0^0 \Delta, \neg A, B$ either according to (Ax1) or to (Ax2). Because of $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \neg F_x(\underline{n}), G_x(\underline{n})$ we have $\Delta \ll \alpha$ and therefore also $\Delta, \neg A, B \ll \alpha$. Hence $\mathbf{ID}_\infty \vdash^{\alpha}_{0} \Delta, \neg A, B$ by (Ax).

If A is the atomic formula $t \in X$, then we have $A_X(F) \equiv F_x(t)$ and $B_X(G) \equiv G_x(s)$ where t and s are equivalent. From the hypothesis $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \neg F_x(\underline{n}), G_x(\underline{n})$ we obtain by lemma 28.2. $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \neg F_x(t), G_x(s)$, i.e. $\mathbf{ID}_\infty \vdash^{\alpha}_{\rho} \Delta, \neg A_X(F), B_X(G)$.

If A is not atomic, then without loss of generality we may assume that A is of the shape $\bigvee_{\xi<\lambda} A_\xi$. The other case is symmetric. But then each of the formulas $\neg A_\xi$ again is X-positive. By the induction hypothesis it now follows $\mathbf{ID}_\infty \vdash^{\alpha \# no(A_\xi)}_{\rho} \Delta, \neg A_\xi(F), B_\xi(G)$ for all $\xi < \lambda$ and by an $\bigvee$-inference we obtain $\mathbf{ID}_\infty \vdash^{\alpha \# no(A_\xi)+1}_{\rho} \Delta, \neg A(F), B_\xi(G)$ for all $\xi < \lambda$. We now claim that $\lambda\xi.\alpha \# no(A_\xi)+1 \ll \alpha \# no(A)$ and infer $\mathbf{ID}_\infty \vdash^{\alpha \# no(A)}_{\rho} \Delta, \neg A(F), B(G)$ by an $\bigwedge$-inference.

It remains to show that $\lambda\xi.\alpha \# no(A_\xi)+1 \ll \alpha \# no(A)$. Since A is an $\mathscr{L}^I_\infty$-formula we only have to consider the following two cases.

First assume that $A \equiv t \in I_F^{\leq \lambda}$ for $\omega \leq \lambda$. Then we have $stg_\wedge A = \emptyset$ and $stg_\wedge \neg A = \{\lambda\}$ and for $\xi < \lambda$ it is $stg_\wedge A_\xi = \emptyset$ and $stg_\wedge \neg A_\xi = \{\xi\}$. Hence $\alpha \# no(A_\xi)+1 = \alpha \# \xi \# 2rk(A_\xi) + 1 \ll \alpha \# \xi \# 2\cdot(rk(F)+1)\cdot\xi + rk(F)+1 < \alpha \# \lambda \# 2\cdot(rk(F)+1)\cdot(\xi+1) \leq \alpha \# \lambda \# 2(rk(F)+1)\cdot\lambda = \alpha \# no(A)$. We have $SC_\Omega(\alpha \# no(A_\xi)) = SC_\Omega(\alpha) \cup SC_\Omega(\xi)$ and $SC_\Omega(\alpha \# no(A)) = SC_\Omega(\alpha) \cup SC_\Omega(\lambda)$. Now if $\alpha \# no(A) \ll \beta$ and $\xi \ll \beta$, then we obtain $SC_\Omega(\alpha) < D\beta$ from the first and $SC_\Omega(\xi) < D\beta$ from the second hypothesis. Hence $SC_\Omega(\alpha) \cup SC_\Omega(\xi) < D\beta$ and it follows $\lambda\xi.\alpha \# no(A_\xi)+1 \ll \alpha \# no(A)$

If A is not of the shape $t \in I_A^{<\alpha}$ for some $\alpha \geq \omega$, then it is $A \equiv \bigvee_{\xi<\lambda} A_\xi$ for some $\lambda \leq \omega$. For $\xi < \lambda$ and $\sigma \in \mathrm{stg}_\wedge A_\xi \cup \mathrm{stg}_\wedge \neg A_\xi$ we then either have $\sigma \in \mathrm{stg}_\wedge A \cup \mathrm{stg}_\wedge \neg A$ or $\sigma < \omega$. But this implies $\mathrm{no}(A_\xi)+1 \ll \mathrm{no}(A)$ for all $\xi < \lambda$. According to 24.16. and 24.18. it then follows $\lambda\xi.\alpha \# \mathrm{no}(A_\xi)+1 \ll \alpha \# \mathrm{no}(A)$.

28.5. **Lemma** (Tautology lemma)

If F_1 *and* F_2 *are equivalent formulas and* Δ *is a finite set of* $\mathscr{L}_\infty^1$ *- formulas. then we have* $\mathbf{ID}_\infty \vdash_0^{\mathrm{no}(\Delta,F_1)} \Delta, \neg F_1, F_2$.

Proof

Assume that X is a set variable which does not occur in F_1 (and therefore also not in F_2). Without loss of generality we may assume that neither $\neg F_1$ nor F_2 are elements of Δ. According to (Ax2) we have $\mathrm{ID}_\infty \vdash_0^{\mathrm{no}(\Delta)} \Delta, \underline{n} \notin X, \underline{n} \in X$. By the monotonicity lemma this implies $\mathrm{ID}_\infty \vdash_\rho^{\mathrm{no}(\Delta) \# \mathrm{no}(F_1)} \Delta, \neg F_1, F_2$ and because of $\{\neg F_1, F_2\} \cap \Delta = \emptyset$ we have $\mathrm{no}(\Delta) \# \mathrm{no}(F_1) = \mathrm{no}(\Delta, F_1)$.

28.6. **Lemma**

If Δ *is a sententially valid finite set of formulas, then there is an* $m < \omega$ *such that* $\mathbf{ID}_\infty \vdash_\rho^{\mathrm{no}(\Delta)+m} \Delta$.

Proof

The lemma follows from the tautology lemma by induction on the degree of sentential reduciblity of the set Δ. It is literally the same proof as for 10.16.

28.7. **Lemma**

Assume that a finite set Δ *of formulas contains a quantifier axiom. Then there is an* $m < \omega$ *such that* $\mathbf{ID}_\infty \vdash_0^{\mathrm{no}(\Delta)+m} \Delta$.

Proof

This lemma too is an immediate consequence of the tautology lemma. Its proof runs as case 2 in the proof of the embedding lemma 11.2.

For an X-positive arithmetical formula A we denote by $\mathrm{Cl}_A(F)$ the formula $\bigwedge_{m<\omega}(A_{X,x}(F,\underline{m}) \rightarrow F_x(\underline{m}))$. Then we have $\mathrm{stg}_\wedge \mathrm{Cl}_A(F) \subset \mathrm{stg}_\wedge F \cup \mathrm{stg}_\wedge \neg F$ which implies $\mathrm{stg}_\wedge \mathrm{Cl}_A(F) \lessdot \mathrm{no}(F)$. Obviously the formula $\mathrm{Cl}_A(F)$ expresses that the class $\{x : F\}$ is closed under the monotone operator induced by A.

28.8. **Lemma** (Closure lemma)

$ID_\infty \mid\frac{\Omega\cdot 2+2\cdot rk(A)+4}{0}\ Cl_A(I_A^{<\Omega})$.

Proof

It is $no(I_A^\Omega) = \Omega\cdot 2+2\cdot rk_X A$. Since $rk_X A \leq rk A$ we have by the tautology lemma

(1) $ID_\infty \mid\frac{\Omega\cdot 2+2\cdot rk A}{0}\ \underline{n}\notin I_A^\Omega, \underline{n}\in I_A^\Omega$.

From (1) we obtain by an Cl_Ω-inference

(2) $ID_\infty \mid\frac{\Omega\cdot 2+2\cdot rk A+1}{0}\ \underline{n}\notin I_A^\Omega, \underline{n}\in I_A^{<\Omega}$.

By $\bigvee$-importation and an $\bigwedge$-inference this implies

(3) $ID_\infty \mid\frac{\Omega\cdot 2+2\cdot rk A+4}{0}\ Cl_A(I_A^{<\Omega})$.

28.9. **Lemma** (Generalized induction lemma)

Let A be an X-positive arithmetical formula. Then we have

$ID_\infty \mid\frac{\omega^{no(F)}\# \xi+1}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^\xi, F_x(\underline{k})$.

Proof by transfinite inducution on ξ

For $\xi = 0$ it holds

(0) $ID_\infty \mid\frac{\omega^{no(F)+1}}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^{<0}, F_x(\underline{k})$ by (Ax1).

For $\xi \neq 0$ we obtain from the induction hypothesis

(1) $ID_\infty \mid\frac{\omega^{no(F)}\# \eta+1}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^\eta, F_x(\underline{k})$ for all $\eta<\xi$.

By an $\bigwedge$-inference this implies

(2) $ID_\infty \mid\frac{\omega^{no(F)}\# \xi+1}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^{<\xi}, F_x(\underline{k})$.

From (0) or (2) respectively, the monotonicity lemma and the structural rule we obtain

(3) $ID_\infty \mid\frac{\omega^{no(F)}\# \xi\# 2\cdot rk(A)+1}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^\xi, A_{X,x}(F,\underline{k}), F_x(\underline{k})$,

since for arithmetical A we have $no(A) = 1+2rk(A)<\omega$.

From the tautology lemma it follows

(4) $ID_\infty \mid\frac{\omega^{no(F)}\# \xi}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^\xi, \neg F_x(\underline{k}), F_x(\underline{k})$.

From (3) and (4) we obtain

(5) $ID_\infty \mid\frac{\omega^{no(F)}\# \xi\# 2rk(A)+2}{0}\ \neg Cl_A(F), A_{X,x}(F,\underline{k})\wedge\neg F_x(\underline{k}), \underline{k}\notin I_A^\xi, F_x(\underline{k})$

by an $\bigwedge$-inference. Using an $\bigvee$-inference we therefrom infer

(6) $ID_\infty \mid\frac{\omega^{no(F)}\# \xi+1}{0}\ \neg Cl_A(F), \underline{k}\notin I_A^\xi, F_x(\underline{k})$.

This, however, is the claim.

The translation of the axiom ID^2_A is an immediate consequence of the lemma. Nevertheless we formulate this as a theorem.

28.10. **Theorem** (Generalized induction theorem)
For any X-positive arithmetical formula A *we have*

$$\vdash_0^{\omega^{no(F)\#\Omega+5}} (Cl_A(F) \rightarrow \forall x(x \in \underline{I}_A \rightarrow F))^*$$

Proof

By the generalized induction lemma we have

$$ID_\infty \vdash_0^{\omega^{no(F)\#\xi+1}} \neg Cl_A(F), \underline{k} \notin \underline{I}_A^\xi, F_x(\underline{k})$$

for all $\xi < \Omega$. By an $\bigwedge$-inference this implies

$$ID_\infty \vdash_0^{\omega^{no(F)\#\Omega}} \neg Cl_A(F), \underline{k} \notin \underline{I}_A^{<\Omega}, F_x(\underline{k}).$$

By $\bigvee$-Importation an $\bigwedge$- inference and again $\bigvee$-importation it finally follows

$$ID_\infty \vdash_0^{\omega^{no(F)\#\Omega+5}} (Cl_A(F) \rightarrow \forall x(x \in \underline{I}_A \rightarrow F))^*$$

As a side remark we will justify the name generalized induction lemma for lemma 28.9. by showing that transfinite induction along a wellordering in fact is a consequence of the lemma.

We start with an arithmetically definable order relation $\prec$ and regard the formula $x \in \text{field}(\prec) \wedge \forall y(y \prec x \rightarrow y \in X)$. Let us abbreviate this formula by $A_\prec$. $A_\prec$ then obviously is an X-positive arithmetical formula. We denote its fixed-point $\underline{I}_{A_\prec}$ by $Acc(\prec)$. $Acc(\prec)$ represents the accessible part of the relation $\prec$. If we treat a fixed relation $\prec$ we just write Acc instead of $Acc(\prec)$. The formula $Cl_{A_\prec}(F)$ then coincides with the formula $Prog(\prec, F)$ defined in §13. For $k \in \text{field}(\prec)$ we obtain $k \in Acc^{|k|}$ by induction on $|k|$, the order type of k in the ordering $\prec$ as defined in 13.2. Namely for $m \prec k$ we have $|m| < |k|$ and obtain $m \in Acc^{|m|} \subset Acc^{<|k|}$ from the induction hypothesis. Hence $A_\prec(Acc^{<|k|}, k)$ which means $k \in Acc^{|k|}$. We define $n := rk(A_\prec)+1$. Then it is $n < \omega$ and $rk(\underline{k} \notin \text{field}(\prec) \vee \underline{k} \in Acc(\prec)^\xi) \leq n\cdot(\xi+1)+1$. By 26.14. we now obtain the following lemma.

28.11. **Lemma**
Let $|k|$ *be the ordertype of* k *in the wellordering* $\prec$ *of ordertype* $<\psi\Omega^\Gamma$. *Then we have* $ID_\infty \vdash_0^{n\cdot(|k|+1)+1} \underline{k} \notin \text{field}(\prec), \underline{k} \in Acc(\prec)^{|k|}$.

28.12. **Lemma** (Transfinite induction)
If $\|\prec\| = \xi < \psi(\Omega^\Gamma)$, *then we have* $ID_\infty \vdash_{n\cdot\xi}^{\omega^{no(F)\#\xi+1}} \neg Prog(\prec, F), \underline{k} \notin \text{field}(\prec), F_x(\underline{k})$ *for all* $k \in \mathbb{N}$.

Proof

Using the cut rule we obtain from 28.11. and the generalized induction lemma $\mathbf{ID}_\infty \mid\frac{\omega^{no(F)} \# |k|+1 \# n\cdot(|k|+1)+1}{n\cdot(|k|+1)} \neg Prog(\prec, F), \underline{k} \notin field(\prec), F_x(\underline{k})$ for all $k \in \mathbb{N}$. Since $|k| \ll \xi$ this immediately implies the claim.

The (translation of) the formula which expresses transfinite induction along the wellordering $\prec$ is now easily obtained from the transfinite induction lemma by $\bigvee$-importation, an $\bigwedge$-inference and repeated $\bigvee$-importation. But of course one can do better than 28.12. There is of course an canonical cut free derivation of transfinite induction. The canonical infinite proof for transfinite induction is a generalization of the proof given in 10.17. for complete induction. By translating the proof of 10.17. into the system $\mathbf{ID}_\infty$ we also obtain the following theorem.

28.13. **Theorem** (Complete induction)
$\mathbf{ID}_\infty \mid\frac{\omega^{no(F)}+\omega}{0} (\forall x(\forall y(y < x \to F_x(y)) \to F) \to \forall x F)^*$

28.14. **Theorem** (Embedding of $\mathbf{ID}_1$)
Suppose that F *is an* $\mathscr{L}_I$*-formula such that* $FV_1(F) = \{x_1, \ldots, x_n\}$ *and* $\mathbf{ID}_1 \vdash F$. *Then there are ordinals* $\alpha \ll \Omega^3 \cdot \omega^\omega$ *and* $m < \omega$ *such that* $\mathbf{ID}_\infty \mid\frac{\alpha}{\Omega+m} F_{\mathbf{x}}(\mathbf{k})^*$ *holds for any n-tuple* $\mathbf{k} = (k_1, \ldots, k_n)$.

Proof

The proof runs similar as the proof of the embedding theorem 11.2. for the formal system $\mathbf{Z}_1$. It is by induction on the length of the derivation in $\mathbf{ID}_1$. The embedding of the logical axioms is obtained from lemmata 28.6. and 28.7. The derivability of the equality axioms and those mathematical axioms which also are mathematical axioms of $\mathbf{Z}_1$ follows as in the proof of 11.2. The provability of the scheme of complete induction in the semiformal system follows from 28.13., that of $\mathbf{ID}^1_A$ follows from 28.8. and that of $\mathbf{ID}^2_A$ from 28.10. Therefore all axioms of $\mathbf{ID}_1$ are provable in the semiformal system. The induction step is then proved as in 11.2.

28.15. **Theorem** (Ordinal analysis of $\mathbf{ID}_1$)
For any formula F *of level* 0 *which is provable in* $\mathbf{ID}_1$ *and does not contain free numbervariables we have* $|F| < \psi\varepsilon_{\Omega+1}$.

Proof

From $\mathbf{ID}_1 \vdash F$ and $SF^* = 0$ we obtain by the embedding theorem $\mathbf{ID}^{\Omega+\omega}_{\varepsilon_{\Omega+1}} \vdash F^*$. Using the elimination theorem for $\mathbf{ID}^{\Omega+\omega}_{\varepsilon_{\Omega+1}}$ it follows $\mathbf{ID}_\infty \overset{\alpha}{\underset{0}{\vdash}} F^*$ for some $\alpha < \psi\varepsilon_{\Omega+1}$. By 26.11., however, this implies $\mathbf{Z}_\Omega \overset{\alpha}{\underset{0}{\vdash}} F^*$ which means $|F| \le \alpha < \psi\varepsilon_{\Omega+1}$.

The ordinal analysis of $\mathbf{ID}_1$ immediately gives us the following corollaries.

28.16. **Corollary**

(i) $SP_0(\mathbf{ID}_1) \subset \psi\varepsilon_{\Omega+1}$.

(ii) $|\mathbf{ID}_1| \le \psi\varepsilon_{\Omega+1}$.

28.17. **Corollary**

If $\prec$ *is an order relation for which we have* $\mathbf{ID}_1 \vdash \mathrm{Fund}(\prec, X)$, *then it is* $\|\prec\| < \psi\varepsilon_{\Omega+1}$.

As a further corollary we obtain

28.18. **Theorem**

If $\mathbf{ID}_1 \vdash \underline{n} \in I_A$, *then it is* $|n|_A < \psi\varepsilon_{\Omega+1}$.

Proof

From $\mathbf{ID}_1 \vdash \underline{n} \in I_A$ we obtain by the embedding theorem 28.14. and the cut elimination theorem 27.9. $\mathbf{ID}_\infty \overset{\alpha}{\underset{0}{\vdash}} \underline{n} \in I_A^{<\Omega}$ for some $\alpha < \psi\varepsilon_{\Omega+1}$. By the boundedness theorem 26.13. it then follows $|n|_A \le \alpha < \psi\varepsilon_{\Omega+1}$.

22.19. **Exercise**

1. Assume that Δ is a finite set of $\mathcal{L}^1_\infty$ – formulas possibly containing the set variable X. Show that $\mathbf{ID}_\infty \overset{\alpha}{\underset{0}{\vdash}} \Delta$ implies $\mathbf{ID}_\infty \overset{\alpha\#\Omega}{\underset{0}{\vdash}} \Delta_X[I_A^{<\Omega}]$.

2. Assume that $F[X]$ is an X-positive Π^1_1-formula of $\mathcal{L}^1_\infty$ without further occurences of set variables. Prove the following claims.

(i) $\mathbf{ID}_\infty \overset{\alpha}{\underset{\Omega+n}{\vdash}} F_X(I_A^{<\Omega})$ implies $\mathbf{ID}_\infty \overset{\Omega\#\omega\cdot 2\#\alpha}{\underset{\Omega+k}{\vdash}} Cl_A(X) \rightarrow F$ for some $k < \Omega$.

(ii) $\mathbf{ID}_\infty \overset{\alpha}{\underset{0}{\vdash}} Cl_A(X) \rightarrow F$ implies $\mathbf{ID}_\infty \overset{\alpha\#\Omega\#\omega}{\underset{\Omega+k}{\vdash}} F_X(I_A^{<\Omega})$ for some $k < \Omega$.

§ 29. The wellordering proof in $\mathbf{ID}_1$

We showed in §23. that the ordinals in $B(\Omega^\Gamma)$ may be represented by a primitive recursive wellordering on the natural numbers. Therefore we may talk about the ordinals in $B(\Omega^\Gamma)$ – via codes – in $\mathbf{ID}_1$. Again we will identify ordinals and their codes. In this section we will denote by $<$ the order relation on (the codes of) the ordinals. The goal of the section is the proof that the wellordering of every proper segment of the segment contained in $B(\varepsilon_{\Omega+1})$ is provable in $\mathbf{ID}_1$. Before we start the proof we will sketch its strategy.

We are going to define two order relations $<_\Omega$ and $<_0$ on the ordinals of $B(\varepsilon_{\Omega+1})$. $<_0$ is the usual order relation restricted to ordinals below Ω whereas $<_\Omega$ is a order relation which no longer is arithmetically definable but only by a formula of level Ω (i.e. by a Π^1_1-formula). The definition of $<_\Omega$ is done in such a way that $TI(<_\Omega\restriction\Omega,X)$ holds trivially. Since we have the full scheme of complete induction available in $\mathbf{ID}_1$ we may now copy the wellordering proof of §16 and obtain $TI(<_\Omega\restriction\omega_n(\Omega+1),X)$ for all $n<\omega$. Then we will show that in $\mathbf{ID}_1$ every transfinite induction along $<_\Omega$ up to an ordinal $\alpha \geq \Omega$ may be condensed into a transfinite induction along $<_0$ up to $\psi\alpha$. Since the segment contained in $B(\varepsilon_{\Omega+1})$ is $\psi\varepsilon_{\Omega+1} = \sup\{\psi(\omega_n(\Omega+1)) : n\in\omega\}$ we obtain the transfinite induction along all its proper segments as a theorem of $\mathbf{ID}_1$.

29.1. Definitions

(i) $\alpha<_0\beta :\Leftrightarrow \alpha<\beta<\Omega$.

(ii) $\alpha\subset_0 X$ denotes the formula $\alpha\in \mathrm{field}(<_0)\wedge\forall\eta(\eta<_0\alpha\to\eta\in X)$.

$\alpha\subset_0 X$ then is an X-positive arithmetical formula. We denote its fixed point by Acc. Acc then represents the accessible part of $<_0$.

(iii) $\mathfrak{M} := \{\alpha :\ SC_\Omega(\alpha)\subset \mathrm{Acc}\}$.

(iv) $\alpha<_\Omega\beta :\Leftrightarrow \alpha\in\mathfrak{M}\wedge\alpha<\beta$.

If $A(X,x)$ is an X-positive arithmetical formula and Γ_A the operator induced by A, then the axiom $\mathbf{ID}^1_A$ says

(1) $\Gamma_A(I_A)\subset I_A$

an the axiom $\mathbf{ID}^2_A$ may be read as

(2) $\Gamma_A(\{x:F\})\subset\{x:F\}\to I_A\subset\{x:F\}$.

The monotonicity of a positive operator is already provable in pure logic and therefore also in $\mathbf{ID}_1$. So

(3) $\Gamma_A(\Gamma_A(I_A))\subset\Gamma_A(I_A)$

is by (1) a theorem of $\mathbf{ID}_1$. By (3) and (2) we obtain

(4) $\underline{I}_A \subset \Gamma_A(\underline{I}_A)$

and by (1) and (4) finally

(5) $\mathbf{ID}_1 \vdash \Gamma_A(\underline{I}_A) = \underline{I}_A$.

For the special case that $\underline{I}_A = \mathrm{Acc}$ we therefore have:

29.2. Lemma

$\forall\alpha(\alpha \in \mathrm{Acc} \leftrightarrow \alpha < \Omega \wedge \forall\eta < \alpha(\eta \in \mathrm{Acc}))$ *is a theorem of* $\mathbf{ID}_1$.

By $\mathrm{Prog}_i(F)$ we are going to abbreviate the formula

$$\forall\xi \in \mathrm{field}(<_i)((\forall\eta <_i \xi\, F(\eta)) \rightarrow F(\xi))$$

for $i \in \{0, \Omega\}$.

$\mathrm{TI}_i(\alpha, F)$ then denotes the formula

$$\alpha \in \mathrm{field}(<_i) \wedge \mathrm{Prog}_i(F) \rightarrow \forall\xi(\xi <_i \alpha \rightarrow F(\xi)).$$

Using these abbreviations the axioms for Acc can be formulated in the following way:

($\mathbf{ID}^1_{\mathrm{Acc}}$) $\mathrm{Prog}_0(\mathrm{Acc})$

and

($\mathbf{ID}^2_{\mathrm{Acc}}$) $\mathrm{Prog}_0(F) \rightarrow \mathrm{Acc} \subset \{x : F\}$

29.3. Lemma

$\mathrm{Acc} \subset \Omega$ *is a theorem of* $\mathbf{ID}_1$.

Proof

We have $\mathrm{Field}(<_0) = \{\alpha : \alpha < \Omega\}$. Since $\mathrm{Prog}_0(\mathrm{Field}(<_0))$ holds trivially we obtain $\mathrm{Acc} \subset \mathrm{Field}(<_0)$ by ($\mathbf{ID}^2_{\mathrm{Acc}}$).

29.4. Lemma

$\mathrm{Prog}(<, F) \rightarrow \forall\xi \in \mathrm{Acc}\, F(\xi)$ *is a theorem of* $\mathbf{ID}_1$.

Proof

For $\alpha < \Omega$ we have $\xi < \alpha \leftrightarrow \xi <_0 \alpha$. Therefore $\forall\xi <_0 \alpha(F(\xi))$ also implies $\forall\xi < \alpha\, F(\xi)$. Together with $\mathrm{Prog}(<, F)$ this yields $F(\alpha)$. Hence $\mathrm{Prog}_0(F)$, and we obtain the claim by $\mathbf{ID}^2_{\mathrm{Acc}}$.

29.5. **Lemma** (provable in $\mathbf{ID}_1$)

The class Acc *is closed under ordinal addition.*

Proof

Define $\mathrm{Acc}_+ := \{\alpha : \forall \rho \in \mathrm{Acc}(\rho+\alpha \in \mathrm{Acc})\}$. We claim

(1) $\mathrm{Prog}_0(\mathrm{Acc}_+)$.

To prove this we may assume the hypothesis $\alpha<\Omega$ and $\forall\eta<\alpha(\eta\in\mathrm{Acc}_+)$ and have to show $\alpha\in\mathrm{Acc}_+$, i.e. $\forall\rho\in\mathrm{Acc}(\rho+\alpha\in\mathrm{Acc})$. By 29.2. it suffices to show $\xi\in\mathrm{Acc}$ for all $\xi<\rho+\alpha$. For $\xi\leq\rho$ we obtain $\xi\in\mathrm{Acc}$ from $\rho\in\mathrm{Acc}$ and 29.2. If $\rho<\xi<\rho+\alpha$, then there is an $\eta<\alpha$ such that $\xi=\rho+\eta$. By hypothesis we then have $\eta\in\mathrm{Acc}_+$ which implies $\rho+\eta\in\mathrm{Acc}$. This proves (1). By (1) and $\mathbf{ID}^2_{\mathrm{Acc}}$ we obtain $\mathrm{Acc}\subset\mathrm{Acc}_+$ which means that for $\alpha,\beta\in\mathrm{Acc}$ we also have $\beta\in\mathrm{Acc}_+$ and therefore $\alpha+\beta\in\mathrm{Acc}$.

29.6. **Lemma**

The formula $\mathrm{Prog}_\Omega(F)\rightarrow\mathrm{Prog}_0(F)$ *is provable in* $\mathbf{ID}_1$.

Proof

We have the hypothesis

(1) $\mathrm{Prog}_\Omega(F)$

(2) $\alpha<\Omega$

and

(3) $\forall\eta<_0\alpha\ F(\eta)$

and have to show $F(\alpha)$.

From $\eta<_\Omega\alpha$ we obtain by (2) $\eta<_0\alpha$ and by (3) $F(\eta)$. Hence $F(\alpha)$ by (1).

29.7. **Lemma** (provable in $\mathbf{ID}_1$)

The class Acc *is closed under the φ-function.*

Proof

We define $\mathrm{Acc}_\varphi := \{\alpha : \forall\xi\in\mathrm{Acc}(\varphi\alpha\xi\in\mathrm{Acc})\vee\alpha\notin\mathfrak{M}\vee\Omega\leq\alpha\}$

and show

(1) $\mathrm{Prog}_\Omega(\mathrm{Acc}_\varphi)$.

From

(2) $\forall\eta<_\Omega\alpha\,(\eta\in\mathrm{Acc}_\varphi)$

we have to conclude $\alpha\in\mathrm{Acc}_\varphi$. This is trivial for $\alpha\notin\mathfrak{M}$ or $\Omega\leq\alpha$. We therefore may assume $\alpha\in\mathfrak{M}\wedge\alpha<\Omega$. Then it remains to prove

(3) $\forall\xi\in Acc(\varphi\alpha\xi\in Acc)$.

We show (3) according to 29.4. by induction on ξ. If we choose any $\xi\in Acc$ we then have the induction hypothesis

(4) $\forall\eta<\xi(\varphi\alpha\eta\in Acc)$.

In order to obtain $\varphi\alpha\xi\in Acc$ it by 29.2. suffices to prove

(5) $\rho<\varphi\alpha\xi\rightarrow\rho\in Acc$.

This will be done by side induction on $G\rho$. If $\rho\notin \mathbb{H}$, then we have $\mathbb{H}(\rho)\subset Acc$ by the side induction hypothesis . By 29.5. this implies $\rho\in Acc$.

If $\rho\in SC$, then we have $\rho\leq\alpha$ or $\rho\leq\xi$. If $\rho\leq\xi$ we obtain $\rho\in Acc$ immediately from $\xi\in Acc$. If $\rho\leq\alpha$, then there is a $\nu\in SC_\Omega(\alpha)$ such that $\rho\leq\nu$. Because of $\alpha\in\mathfrak{M}$ we have $\nu\in Acc$ and it again follows $\rho\in Acc$.

Now assume $\rho\in\mathbb{H}\setminus SC$. Then there are ρ_1,ρ_2 such that $\rho=_{NF}\varphi\rho_1\rho_2$. We now have to distinguish the following cases:

1. $\rho_1=\alpha$ and $\rho_2<\xi$. Then we obtain $\varphi\rho_1\rho_2\in Acc$ from (4).
2. $\alpha<\rho_1$ and $\rho\leq\xi$. Then $\rho\in Acc$ follows from $\xi\in Acc$ by 29.2.
3. $\rho_1<\alpha$ and $\rho_2<\varphi\alpha\xi$. Then for every $\nu\in SC_\Omega(\rho_1)$ there is a $\mu\in SC_\Omega(\alpha)\subset Acc$ such that $\nu\leq\mu$. Hence $SC_\Omega(\rho_1)\subset Acc$ which implies $\rho_1<_\Omega\alpha$. By (2) it therefore follows $\rho_1\in Acc_\varphi$. By the side induction hypothesis we have $\rho_2\in Acc$. Since $\rho_1\in\mathfrak{M}\cap\Omega$ this implies $\varphi\rho_1\rho_2\in Acc$. This finishes the proof of (1).

We now have to show that $\alpha,\beta\in Acc$ also imply $\varphi\alpha\beta\in Acc$. $\alpha,\beta\in Acc$ imply $\alpha,\beta<\Omega$ since $Acc\subset\Omega$. Since $SC_\Omega(\alpha)\leq\alpha$, we have $SC_\Omega(\alpha)\subset Acc$, i.e. $\alpha\in\mathfrak{M}$. From (1) we obtain $Prog_0(Acc_\varphi)$ by 29.6. Using $\mathbf{ID}^2_{Acc}$ this implies $Acc\subset Acc_\varphi$ Hence $\alpha\in Acc_\varphi\cap\mathfrak{M}\cap\Omega$. Together with $\beta\in Acc$ this implies $\varphi\alpha\beta\in Acc$.

29.8. Theorem

$\mathbf{ID}_1\vdash TI(\Gamma_0,X)$.

Proof

If we define $\Lambda_0=\varphi00$ and $\Lambda_{n+1}=\varphi\Lambda_n0$, then by 17.21. we have $\Gamma_0=\sup\{\Lambda_n: n<\omega\}$. From $0\in Acc$ and 29.7 we obtain by induction on n $\mathbf{ID}_1\vdash\forall n(\Lambda_n\in Acc)$. For every $\xi<\Gamma_0$ there is an $n<\omega$ such that $\mathbf{ID}_1\vdash\xi<\Lambda_n$. Hence $\mathbf{ID}_1\vdash\Gamma_0\subset Acc$. Together with 29.4. this implies $\mathbf{ID}_1\vdash Prog(<,X)\rightarrow\forall\xi<\Gamma_0(\xi\in X)$, i.e. $\mathbf{ID}_1\vdash TI(\Gamma_0,X)$.

Already in §22. we referred to theorem 29.8. It shows that $\mathbf{ID}_1$ is not predicatively interpretable. We soon will see, however, that the wellordering proved in 29.8 by no means exhausts the power of $\mathbf{ID}_1$.

29.9. Definition

$\mathrm{Acc}_\Omega := \{\alpha : \alpha \notin \mathfrak{M} \vee \alpha \leq K\alpha \vee \psi\alpha \in \mathrm{Acc}\}$

29.10. Lemma (provable in $\mathbf{ID}_1$)

$\mathrm{Prog}_\Omega(\mathrm{Acc}_\Omega)$

Proof

Assume $\alpha \in \mathrm{Field}(<_\Omega)$ and $\forall \eta <_\Omega \alpha(\eta \in \mathrm{Acc}_\Omega)$. We have to show $\alpha \in \mathrm{Acc}_\Omega$. If $\alpha \notin \mathfrak{M}$ or $\alpha \leq K\alpha$, then we are done. We therefore assume $\alpha \in \mathfrak{M}$ and $K\alpha < \alpha$ and have to show $\psi\alpha \in \mathrm{Acc}$. In order to do that it suffices to prove

(1) $\rho < \psi\alpha \rightarrow \rho \in \mathrm{Acc}$.

We show (1) by induction on $G\rho$. If $\rho \notin SC$, then we obtain the claim from 29.5. or 29.7. and the induction hypothesis. We therefore assume $\rho \in SC$. Then there is an $\rho_0 < \alpha$ such that $\rho =_{NF} \psi\rho_0$. Hence $K\rho_0 < \rho_0 < \alpha$. If $\xi \in SC_\Omega(\rho_0)$, then there is an η such that $\xi =_{NF} \psi\eta$. Thus $\eta \in K\xi \subset K\rho_0 < \alpha$ which implies $\xi = \psi\eta < \psi\alpha$. Hence $SC_\Omega(\rho_0) < \psi\alpha$. Since we always have $G\xi \leq G\rho_0 < G\rho$ for $\xi \in SC_\Omega(\rho_0)$ we obtain $SC_\Omega(\rho_0) \subset \mathrm{Acc}$ by the induction hypothesis. This implies $\rho_0 <_\Omega \alpha$ and therefore also $\rho_0 \in \mathrm{Acc}_\Omega$. But this implies $\rho \in \mathrm{Acc}$.

29.11. Lemma (Condensation lemma)

From $\mathbf{ID}_1 \vdash TI_\Omega(\alpha, X) \wedge \alpha \in \mathfrak{M} \wedge K\alpha < \alpha$ *it follows* $\mathbf{ID}_1 \vdash \psi\alpha \in \mathrm{Acc}$ *and therefore also* $\mathbf{ID}_1 \vdash TI(\psi\alpha, X)$.

Proof

From $\mathbf{ID}_1 \vdash TI_\Omega(\alpha, X)$ we especially obtain

(1) $\mathbf{ID}_1 \vdash TI_\Omega(\alpha, \mathrm{Acc}_\Omega)$.

By 29.10. we have

(2) $\mathbf{ID}_1 \vdash \mathrm{Prog}_\Omega(\mathrm{Acc}_\Omega)$.

From (1) and (2) it follows

(3) $\mathbf{ID}_1 \vdash \forall \xi(\xi <_\Omega \alpha \rightarrow \xi \in \mathrm{Acc}_\Omega)$

and from (2) and (3)

(4) $\mathbf{ID}_1 \vdash \alpha \in \mathrm{Acc}_\Omega$.

Because of $\alpha \in \mathfrak{M}$ and $K\alpha < \alpha$ we obtain from (4)

(5) $\mathbf{ID}_1 \vdash \psi\alpha \in \mathrm{Acc}$.

From (5) we obtain by 29.2. and 29.4.

(6) $\mathbf{ID}_1 \vdash \mathrm{Prog}(<, X) \rightarrow \forall \xi < \psi\alpha (\xi \in X)$

and this means $\mathbf{ID}_1 \vdash \mathrm{TI}(\psi\alpha, X)$.

29.12. **Lemma**

$\mathbf{ID}_1 \vdash \mathrm{TI}_\Omega(\Omega+1) \wedge \Omega+1 \in \mathfrak{M} \wedge K(\Omega+1) < \Omega+1$

Proof

Because of $SC_\Omega(\Omega+1) = \emptyset$ and $K(\Omega+1) = \emptyset$ we trivially have $\Omega+1 \in \mathfrak{M} \wedge K(\Omega+1) < \Omega+1$. Now we assume $\mathrm{Prog}_\Omega(X)$ and have to show $\forall \xi \leq_\Omega \Omega (\xi \in X)$. From $\xi <_\Omega \Omega$, however, we obtain $\xi < \Omega$ and $SC_\Omega(\xi) \subset \mathrm{Acc}$. Since Acc is closed under + and φ this already implies $\xi \in \mathrm{Acc}$. By 29.6. $\mathrm{Prog}_\Omega(X)$ also implies $\mathrm{Prog}_0(X)$ and we obtain $\xi \in X$ by $\mathbf{ID}^2_{\mathrm{Acc}}$. Hence $\forall \xi <_\Omega \Omega (\xi \in X)$. This together with the hypothesis $\mathrm{Prog}_\Omega(X)$ also implies $\Omega \in X$.

29.13. **Lemma**

$\mathbf{ID}_1 \vdash \alpha \in \mathfrak{M} \wedge K(\alpha) < A \wedge \mathrm{TI}_\Omega(\alpha, X)$ *also implies* $\mathbf{ID}_1 \vdash \omega^\alpha \in \mathfrak{M} \wedge K(\alpha) < \alpha \wedge \mathrm{TI}_\Omega(\omega^\alpha, X)$.

Proof

Because of $SC_\Omega(\omega^\alpha) = SC_\Omega(\alpha)$ and $K(\omega^\alpha) = K(\alpha)$ we obtain $\omega^\alpha \in \mathfrak{M} \wedge K(\alpha) < \alpha$ immediately from $\alpha \in \mathfrak{M}$. Since $\mathbf{Z}_1$ is a subsystem of $\mathbf{ID}_1$ the second part of the claim follows as in 15.3.

We now define a sequence by

$\zeta_0 = \Omega+1$, $\zeta_{n+1} = \omega^{\zeta_n}$.

Then we have

(1) $\forall n (\zeta_n \in B(0) \subset B(\zeta_n))$, i.e. $\forall n (K\zeta_n < \zeta_n)$

by an easy induction on n and

(2) $\psi\varepsilon_{\Omega+1} = \sup(\psi\zeta_n)$.

To prove (2) we observe that $\zeta_n < \varepsilon_{\Omega+1}$ holds for all $n < \omega$. By (1) this implies $\psi\zeta_n < \psi\varepsilon_{\Omega+1}$ for all $n < \omega$. On the other hand if $\eta < \psi\varepsilon_{\Omega+1}$, then we show by induction on $G\eta$, that there is an $n < \omega$ such that $\eta < \psi\zeta_n$. For $\eta \notin SC$ this is immediate from the induction hypothesis. If $\eta \in SC$, then there is an η_0 such that $\eta =_{NF} \psi\eta_0 < \psi\varepsilon_{\Omega+1} < \Omega$. But then $\eta_0 < \varepsilon_{\Omega+1}$ and we obtain an $n < \omega$ such that $\eta_0 < \zeta_n$. Hence $\eta < \psi\zeta_n$.

29.14. **Theorem**

For all $\xi < \psi\varepsilon_{\Omega+1}$ *we have* $\mathbf{ID}_1 \vdash \mathrm{TI}(\xi, X)$.

Proof

For $\xi<\psi\varepsilon_{\Omega+1}$ there is an $n<\omega$ auch that $\xi<\psi\zeta_n$. By 29.12. and 29.13. we obtain $\mathbf{ID}_1\vdash\zeta_n\in\mathfrak{M}\wedge K\zeta_n<\zeta_n\wedge TI_\Omega(\zeta_n,X)$ for all $n<\omega$. By the condensation lemma this implies $\mathbf{ID}_1\vdash TI(\psi\zeta_n,X)$. Hence also $\mathbf{ID}_1\vdash TI(\xi,X)$.

29.15. **Theorem**

For every $\alpha<\psi\varepsilon_{\Omega+1}$ *there is a* Π^1_1*-sentence* F *such that* $\alpha\leq|F|$ *and* $\mathbf{ID}_1\vdash F$.

Proof

Literally as the proof of 15.8.

29.16. **Theorem**

$Sp_0(\mathbf{ID}_1)=|\mathbf{ID}_1|=\psi\varepsilon_{\Omega+1}$

Proof

This is obvious from 29.15. and 28.16.

In a last step we are going to convince ourselves that the bound given in 28.18. is in fact an exact one. That means that we have to show that for every $\alpha<\psi\varepsilon_{\Omega+1}$ there are an X-positive formula A and an $n\in\omega$ such that $\mathbf{ID}_1\vdash n\in\underline{I}_A$ and $\alpha\leq|n|_A$ hold.

To show this we choose an arithmetically definable wellordering $\prec$ and regard the fixed point of the operator $\Gamma_\prec$ which is given by

$$\Gamma_\prec(S):=\{n:\forall m(m\prec n\rightarrow m\in S)\}.$$

By $|n|_\Gamma$ we now denote the norm of the element n in the inductive definition Γ and by $|n|_\prec$ the norm of n in the wellordering $\prec$ as defined in 13.2. Then we obtain:

29.17. **Lemma**

For all $n\in\mathrm{Field}(\prec)$ *it is* $|n|_\Gamma=|n|_\prec$.

Proof

We first show $n\in Acc^{|n|_\prec}$ by $\prec$-induction. By the induction hypothesis we have $m\in Acc^{|m|_\prec}$ for all $m\prec n$. Since $|m|_\prec<|n|_\prec$ this implies $\forall m\prec n(m\in Acc^{<|n|_\prec})$. Hence $n\in Acc^{|n|_\prec}$ which proves $|n|_\Gamma\leq|n|_\prec$. For the other direction we show $|n|_\prec\leq|n|_\Gamma$ by induction on $|n|_\prec$. If $\xi<|n|_\prec=\{|m|_\prec: m\prec n\}$, then there is an $m\prec n$ such that $\xi=|m|_\prec\underset{IV}{=}|m|_\Gamma$. But since $m\in Acc^{<|n|_\Gamma}$ it follows $|m|_\Gamma<|n|_\Gamma$. Hence $|n|_\prec\leq|n|_\Gamma$.

29.18. **Theorem**

For every $\alpha < \psi\varepsilon_{\Omega+1}$ *there is an* X*-positive arithmetical formula* A *and an* $n \in \omega$ *such that* $\alpha \leq |n|_A$ *and* $\mathbf{ID}_1 \vdash \underline{n} \in \underline{I}_A$.

Proof

For every $\alpha < \psi\varepsilon_{\Omega+1}$ there is an $n < \omega$ such that $\alpha < \psi\Delta_n$. In the proof of 29.14 we have shown that $\mathbf{ID}_1 \vdash TI_\Omega(\Delta_n, X)$ holds for all $n < \omega$. By the condensation lemma this implies $\mathbf{ID}_1 \vdash \psi\Delta_n \in Acc$ for all $n < \omega$. By lemma 29.17., however, it is $|\psi\Delta_n|_{Acc} = \psi\Delta_n$.

Epilogue

The main goal of this lecture was to introduce to the techniques of impredicative ordinal analysis. The axiom system for noniterated inductive definitions served as an simple example for an impredicative theory. Of course, this is just a first step into the world of impredicativity. The most straightforward way to obtain more complicated axiom systems is to consider iterated inductive definitions. These theories are treated in [BFPS]. There it is also shown how these theories are connected to subsystems of classical analysis, i.e. second order number theory with comprehension. The real fascination of impredicative systems, however, becomes not visible till one considers subsystems of set theory. Pitily there are no text books in this area. The best references here are the papers [1979] and [1986] of *G.Jäger*. An impressing variety of subsystems of set theory is presentend by *M.Rathjen* [1989]. This paper is a good example for the interplay between recursion theoretical, set theoretical, model theoretical and proof theoretical methods in the ordinal analysis of subsystems of set theory. A survey of these methods is given in Pohlers [1990]. A text book with the title "Admissible Proof Theory" is in preparation and will appear in the Springer series "Ergebnisse der Mathematik und ihrer Grenzgebiete".

We will close this book by giving some comments on the 'constructive' meaning of ordinal analysis. In §14 we already indicated that it is sufficient for an ordinal analysis to regard only recursive proof trees of the semiformal system. This can be used to show

$$(1)\quad T \vdash F \quad \Leftrightarrow \quad Z_1 + TI(<|T|, X) \vdash F$$

for all Π^1_1-sentences F. Here $TI(<|T|,X)$ means that we allow induction along all initial segments of the primitive recursive wellordering $\prec_T$ of ordertype $|T|$ which has been obtained from the notation system used in the ordinal analysis of T. A detailed proof is in Pohlers [BFPS]. The axiom system **PRA** for 'primitive recursive analysis' is essentially the system $\boldsymbol{\Sigma^0_1}$ – **INDR** of exercise 3.15.6. (often considered as second order theory but without strong comprehensions). By the (formal) reflection principle (REF(T)) for an axiom system T one denotes the principle

$$(REF(T))\quad Bew_T(\ulcorner F \urcorner) \to F$$

where Bew_T is a provability predicate for T. $(\Pi^0_2 - REF(T))$ is the scheme (REF(T)) with F restricted to Π^0_2 – sentences, i.e. sentences of the form $\forall \vec{x}\, \exists y\, G(\vec{x},y)$ with G quantifier free. For a primitive recursive order relation $\prec$ on the natural numbers we denote by PRWO($\prec$) that there are no primitive recursive infinitely $\prec$-descending sequences. Then we have

(2) $$\mathbf{PRA} \vdash \mathrm{PRWO}(\prec) \leftrightarrow (\Pi^0_2 - \mathrm{REF}(\mathrm{Z}_1 + \mathrm{TI}(\prec, \mathrm{X}))).$$

This is considered to be a folklore result of proof theory. Its proof needs a principle known as 'continuous cut elimination' originally developed by G.E.Mints. The most beautiful proof has been given by W.Buchholz in [1988a]. From (1) and (2) it already follows that **PRA** + PRWO($\prec_T$) proves the consistency of T. Moreover the theory **PRA** has a beautiful computational aspect. It has primitive recursive Π^0_2 – Skolem functions, i.e. if $\mathbf{PRA} \vdash \forall \vec{x}\, \exists y\, G(\vec{x}, y)$ for a quantifier free formula $G(\vec{x},y)$, then there is a primitive recursive function f such that $\mathbb{N} \models G(\vec{x}, f(\vec{x}))$. This result can be extended to **PRA** + PRWO($\prec$) in so far that this theory has Skolem functions which can be obtained from the basic functions C^n_k, P^n_k and S by substitution, primitive recursion (cf. 1.1.) and the $\prec$-descending μ-operator which for a given n+1-ary function f searches for the value

$$(\mu_\prec f)(\vec{x}) := \min\{y : (\neg f(\vec{x}, y+1) \prec f(\vec{x},y))\}$$

i.e. $\mu_\prec f(\vec{x})$ computes the length of a $\prec$-descendent sequence $f(\vec{x}, 0) \succ f(\vec{x},1) \succ f(\vec{x}, 2) \succ \ldots$. The class of these functions, the $\prec$-descendent functions, can also be obtained by $\prec$-recursion, i. e. using the scheme

$$f(\vec{x}, y) = \begin{cases} h(\vec{x}, f(\vec{x}, g(\vec{x}, y))) & \text{if } g(\vec{x},y) \prec y \\ k(\vec{x},y) & \text{otherwise} \end{cases}$$

in addition to substitution and primitive recursion. The functions can also be characterized using the Hardy hierarchy of computable functions which is given by

$$H_0(x) = x$$
$$H_{\alpha+1}(x) = H_\alpha(x+1)$$
$$H_\lambda(x) = H_{\lambda[x]}(x) \text{ for limit ordinals } \lambda$$

where $\{\lambda[n] : n < \omega\}$ is a fundamental sequence for λ, i.e. $\sup\{\lambda[n] : n < \omega\} = \lambda$ and $\lambda[n] < \lambda[n+1]$ for all $n < \omega$. It can be shown that the $\prec$-descending functions, where $\prec$ is an initial segment of $\prec_T$, are all majorizable by the function $H_{|T|}$. From (1) and (2) we obtain a characterization of the Π^0_2 – Skolem functions of the theory T. If $T \vdash \forall \vec{x}\, \exists y\, G(\vec{x},y)$, then we obtain $\mathbf{Z}_1 + \mathrm{TI}(\prec, \mathrm{X}) \vdash \forall \vec{x}\, \exists y\, G(\vec{x},y)$ for an initial segment $\prec$ of $\prec_T$ by (1). This entails

$$\mathbf{PRA} \vdash \mathrm{Bew}_{(\mathbf{Z}_1 + \mathrm{TI}(\prec, \mathrm{X}))}(\ulcorner \forall \vec{x}\, \exists y\, G(\vec{x},y) \urcorner)$$

which by (2) implies $\mathbf{PRA} + \mathrm{PRWO}(\prec) \vdash \forall \vec{x}\, \exists y\, G(\vec{x},y)$. Hence T has Π^0_2-Skolem functions which are $\prec$- descendent for initial segments of $\prec_T$. A recursive

function f with index e is provably recursive in T, if $T \vdash \forall\vec{x}\,\exists y\, T(e,\vec{x},y)$, where T denotes the Kleene predicate), i.e. if T proves f to be total. The provably recursive functions of T are thus Π^0_2 – Skolem functions and therefore majorizable by $H_{|T|}$.
Since (1) is a side result of the method of local predicativity (cf. [BFPS]) we obtain as as a corollary of the (proof of the) ordinal analysis for T a characterization of the Π^0_2 – Skolem functions, and thus also of the provable recursive functions of T. This characterization may be considered as a very constructive one since the wellorderings $\prec_T$ obtained from the ordinal analysis are so simple that it causes no problems to implement them on a computer. (For the system obtained in chapter III this has been done by K.Stroetmann in Münster). Therefore there is a program, implementable on a real computer, computing the provably recursive functions of T. As a matter of fact, however, these functions increase so incredibly fast that they only are computable for very small arguments.
The above stated facts are scattered in the literature. The best reference here is Takeuti's book [1987 CH.2 §12] where he proves similar results for the case of pure number theory.
A textbook treating this material systematically is still a challenge.

Bibliography

Abbreviations:

AMLG	Archiv für Mathematische Logik und Grundlagenforschung
APAL	Annals of Pure and Applied Logic (vorher:Annals of Mathematical Logic)
BFPS	Buchholz, Feferman, Pohlers, Sieg: Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies, LNM 897, 1981
HB	Handbook of Mathematical Logic
HF	L. Harrington, M.Morley, A. Scedrov and S.G. Simpson (eds.) Harvey Friedman's research on the foundations of mathematics, NH
	J. Barwise (ed.), North Holland, Amsterdam 1977
IPT	Intuitionism and Proof Theory, Proceedings of the summer conference at Buffalo, N.Y. 1968 A. Kino, J. Myhill, R.E. Vesley (eds.), Amsterdam-London 1970
JSL	Journal of Symbolic Logic
LMPS III	Logic, Methodology and Philosophy of Science III, B. van Rootselaar, J.F. Staal (eds.), Amsterdam 1968
LMPS VI	Logic, Methodology and Philosophy of Science VI L.J.Cohen, J.Los, H. Pfeiffer, K.-P. Podewski, Amsterdam 1982
LNM	Lecture Notes in Mathematics
MA	Mathematische Annalen
NH	North-Holland, Amsterdam
SDBA	Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch-Naturwissenschaftliche Klasse
SR	Stanford Report
ZML	Zeitschrift für mathematische Logik und Grundlagenforschung der Mathematik

W.Ackermann:

1925	Begründung des 'Tertium non datur' mittels der Hilbert'schen Theorie der Widerspruchsfreiheit, MA 93, pp.1-36
1928	Zum Hilbert'schen Aufbau der reellen Zahlen, MA 99, pp.118-133

1937 Die Widerspruchfreiheit der allgemeinen Mengenlehre, MA 114, pp.305-315

1940 Zur Widerspruchsfreiheit der Zahlentheorie, MA 117, pp.162-194

1950 Widerspruchsfreier Aufbau der Logik I. Typenfreies System ohne tertium non datur, JSL 15, pp.33-57

1951 Konstruktiver Aufbau eines Abschnitts der zweiten Cantor'schen Zahlenklasse, Mathematische Zeitschrift 53, pp.403-413

1952 Widerspruchsfreier Aufbau einer typenfreien Logik I (erweitertes System), Mathematische Zeitschrift 55, pp.364-384

1953 Widerspruchsfreier Aufbau einer typenfreien Logik II, Mathematische Zeitschrift 57, pp. 155-166

P. Aczel:

1977 An introduction to inductive definition, in HB pp. 739-782

H. Bachmann:

1950 Die Normalfunktion und das Problem der ausgezeichneten Folgen von Ordnungszahlen, Vierteljahresschr. Nat. Ges., Zürich 95, pp. 5-37

1955 Transfinite Zahlen, Springer, Berlin, Göttingen, Heidelberg

J. Barwise:

1969 Infinitary logic and admissible sets, JSL 34, pp. 226-252

1975 Admissible sets and structures, Springer, Berlin, Heidelberg, New York

P. Bernays:

1928 Zusatz zu Hilberts Vortrag über 'Die Grundlagen der Mathematik', Abhandlungen des mathematischen Seminars der Universität Hamburg 6, pp. 89-92 Transl. in 'From Frege to Gödel'(ed. J.van Heijenoort) Harvard University Press, Cambridge [1967], pp. 486-489

1930 Die Philosophie der Mathematik und die Hilbertsche Beweistheorie, Blätter für deutsche Philosophie 4, pp.326-367

1932 Methoden des Nachweises von Widerspruchsfreiheit und ihrer Grenzen, in W.Sax (ed.) Verhandlungen der internationalen Mathematiker Kongresses Zürich Bd.2, pp.342-343

1935 Quelques points essentiels de la metamathematique, L'Enseignement Mathematique 34, pp. 70-95

1935a Hilberts Untersuchungen über die Grundlagen der Arithmetik, in [Hilbert 1935]

1941 Sur les questions methodologiques actuelles de la theorie hilbertienne de la demonstration, in F.Gonseth (ed.) Les Entretiens de Zuerich sur les Fondements et la methode des Sciences Mathematiques, Leemen, Zürich, pp.144-152, 153-161

1950 Mathematische Existenz und Widerspruchsfreiheit, in Etudes de Philosophie des Sciences, en Hommage a F.Gonseth a l'occasion de son 60eme Anniversaire, Griffon Neuchatel, pp. 11-25

1951 Über das Induktionschema in der rekursiven Zahlentheorie. in A.Menne et al. (eds.) Kontrolliertes Denken, Alber, Freiburg, pp. 10-17

1953 Existence et non-contradiction en mathematiques. Avec une note de G.Bouligand, Revue Philosophique de la France e de l'Etranger 143, pp. 85-87

1954 Über den Zusammenhang des Herbrand'schen Satzes mit den neueren Ergebnissen von Schütte und Stenius, in J.C.H.Gerretsen et al. (eds.) Proceedings of the international Congress of Mathematicians 1954, 2 NH, p. 397

1954a Zur Beurteilung der Situation in der beweistheoretischen Forschung, Theoria 2, pp. 153-154

1970 The original Gentzen consistency proof for number theory, in IPT, pp. 409-417

1971 Zum Symposium über die Grundlagen der Mathematik, Dialectica 25, pp. 171-195

1976 Abhandlungen zur Philosophie der Mathematik, Wissenschaftliche Buchgesellschaft, Darmstadt

J. (Bridge) Kister:

1975 A simplification of the Bachmann method for generating large countable ordinals, JSL 40, pp. 171-185

J. (Bridge) Kister and J,N. Crossley:

1986/87 Natural well-orderings, AMLG 26, pp. 57-76

J. (Bridge) Kister, D. van Dalen and A.S. Troelstra:

1987 Ω-Bibliography of mathematical logic, vol VI, Proof Theory, constructive mathematics, Springer, Berlin, Heidelberg, NY

D.K. Brown and S.G. Simpson:

1986 Which set existence axioms are needed to prove the separable Hahn-Banach theorem?, APAL 31, pp. 123-144

W. Buchholz:

1974 Rekursive Bezeichnungssysteme für Ordinalzahlen auf der Grundlage der Feferman-Aczel'schen Normalfunktionen, Dissertation München

1975 Normalfunktionen und konstruktive Systeme von Ordinalzahlen, LNM 500, pp. 4-25

1976 Über Teilsysteme von $\overline{\Theta}(\{g\})$, AMLG 18, pp. 85-98

1977 Eine Erweiterung der Schnitteliminationsmethode, Habilitationsschrift, München

1981 The $\Omega_{\mu+1}$-rule, in BFPS, pp. 188-233

1981a Ordinal analysis of ID_ν, in BFPS, pp. 234-260

1986 A new system of proof-theoretic ordinal functions, APAL 32, pp. 195-207

1986 An independence result for $(\Pi^1_1\text{-CA})+(BI)$,APAL ?

1988 Induktive Definitionen und Dilatoren, AMLG 27, pp. 51-60

1988a Notation systems for infinitary derivations, Preprint München

W. Buchholz and W. Pohlers:

1978 Provable wellorderings of formal theories for transfinitely iterated inductive definitions, JSL 43, pp. 118-125

W. Buchholz and K. Schütte:

1976 Die Beziehungen zwischen den Ordinalzahlsystemen Σ und $\Theta(\omega)$, AMLG 17, pp. 179-190

1980 Syntaktische Abgrenzung von formalen Systemen von Π^1_1-Analysis und Δ^1_2-Analysis, SDBA

1983 Ein Ordinalzahlensystem für die beweistheoretische Abgrenzung der Π^1_2-Separation und Bar-Induktion, SDBA

198? Proof Theory of Impredicative Subsystems of Analysis. Bibliopolis, To appear

A. Cantini:

198? A note on the theory of admissible sets with ϵ-induction restricted to formulas with one quantifier and related systems, Preprint, München

1983 A note on a predicatively reducible theory of iterated elementary induction, Preprint, München

1985 Majorizing provably recursive functions in fragments of PA, AMLG 25, pp. 21-31

1985 a On weak theories of sets and classes which are based on strict Π_1^1-reflection, ZML 31, pp. 321–332

S. Feferman:

1962 Transfinite recursive progressions of axiomatic theories, JSL 27, 259–316

1964 Systems of predicative analysis, JSL 29, pp. 1–30

1966 Predicative provability in set theory, Bull. AMS 72, pp. 486–489

1967 Autonomous transfinite progressions and the extent of predicative mathematics, LMPS III, pp. 121–135

1968 Lectures on proof theory, LNM 70, pp.1–107

1968 a Systems of predicative analysis II. Representations of ordinals, JSL 33, pp. 193–220

1970 Hereditarily replete functionals over the ordinals, in IPT, pp. 289–301

1970 a Formal theories for transfinite iterations of generalized inductive definitions and some subsystems of analysis, in IPT, pp. 303–326

1971 Ordinals and functionals in proof theory, in Proc.Int. Cong. Maths. (Nice, 1970) 1, pp. 229–233

1972 Infinitary properties, local functors and systems of ordinal functions, LNM 255, pp. 63–97

1974 Predicatively reducible systems of set theory, in Axiomatic Set Theory, Part II, AMS Proc. Symp. Pure Math. 13, pp. 11–32

1975 A language and axioms for explicit mathematics, LNM 450, pp. 87–139

1977 Theory of finite type related to mathematical practice , in HB, pp. 913–971

1978 Recursion theory and set theory; a marriage of convenience, in J.E. Fenstad et al. (eds.), Generalized Recursion Theory II, NH, pp. 55–98

1979 A more perspicious formal system for predicativity, in K. Lorenz (ed.), Konstruktionen versus Positionen I, de Gruyter, Berlin, pp. 87–139

1979 a Constructive theory of functions and classes, in M. Boffa et al. (eds.), Logic Colloquium 1978, NH, pp. 159–224

1979 b Generalizing set-theoretical model theory and an analogue theory on admissible sets, in J. Hintikka et al. (eds.), Essays on mathematical and philosophical logic, Reidel Dordrecht, pp. 171–195

1982 Iterated inductive fixed-point theories, in G. Metakides (ed.) Patras Logic Symposion, NH, pp. 171–196

1982a Monotone inductive definitions, in A.S. Troelstra et al. (eds.), The L.E.J. Brouwer Centenary Symposium, NH, pp. 77–89

1982b Inductively presented systems and the formalization of meta-mathematics, in D. van Dalen et al. (eds.), Logic Colloquium 1980, NH, pp. 95–128

1984 Toward useful type-free theories, JSL 49, pp. 75–111

1984a Foundational ways, in Perspectives in Mathematics, Birkhäuser, Basel, pp. 147–158

1985 Working foundations, Synthese 62, pp. 229–254

198? Reflecting on incompleteness

1987 Proof theory: A personal report, in G. Takeuti 1987, pp. 447–485

1988 Hilbert's program relativized: Proof-theoretical and foundational reductions, JSL 53 pp.364–384

S. Feferman and G. Jäger:

1983 Choice principles, the bar rule and autonomously iterated comprehension schemes in analysis, JSL 48, pp. 63–70

S. Feferman and G. Kreisel:

1966 Persistent and invariant formulas relative to theories of higher order, Bull. AMS 72, pp. 480–485

S. Feferman and W. Sieg:

1981 Iterated inductive definitions and subsystems of analysis, in BFPS, pp. 16–77

1981a Proof theoretic equivalences between classical and constructive theories for analysis, in BFPS, pp. 78–142

H. Friedman:

1967 Subsystems of set theory and analysis, Ph.D. thesis, M.I.T.

1969 Bar induction and Π^1_1-CA, JSL 34, pp. 353–362

1970 Iterated inductive definitions and Σ^1_2-AC, in IPT, pp. 435–442

1975 Some systems of second order arithmetic and their use, in Proc. of the International of Mathematicians, Vancouver 1974, vol. 1, pp. 235–242

1977 Set theoretic foundations for constructive analysis, Ann. Math., pp. 1–28

1978 Classically and intuitionistically provably recursive functions, in LNM 669, pp. 21–27

1980 A strong conservative extension of Peano arithmetic, in J. Barwise et al. (eds.), The Kleene Symposium, pp. 113–122

H. Friedman and R. Jensen:

1968 Note on admissible ordinals, in LNM 72, pp. 77–79

H. Friedman, K. McAloon and S. Simpson

1982 A finite combinatorial principle which is equivalent to the 1-consistency of predicative analysis, in G. Metakides (ed.), Patras Logic Symposion, NH, pp. 197–230

H. Friedman, S.G. Simpson and R.L. Smith:

1983 Countable algebra and set existence axioms, APAL 25, pp. 141–181

H. Friedman and A. Scedrov:

1986 Intuitionistically provable recursive well-orderings, APAL 30, pp. 165–171

R.O. Gandy:

1974 Inductive definitions, in J.E.Fenstad, P.G. Hinman (eds.), Generalized Recursion Theory, NH, pp. 265–299

G. Gentzen:

1934/5 Untersuchungen über das Logische Schließen I, II, Mathematische Zeitschrift 39, pp. 176–210, 405–431

1936 Die Widerspruchsfreiheit der reinen Zahlentheorie, MA 112, pp. 493–565

1943 Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie, MA 119, pp. 140–161

1969 On the relation between intuitionistic and classical arithmetic, in Szabo (ed.), The Collected Papers of Gerhard Gentzen, NH, pp. 53–67

H. Gerber:

1970 Brouwer's bar theorem and a system of ordinal notations, in IPT, pp. 327–338

J.Y. Girard:

1981 Π^1_2-logic, part 1: dilators, APAL 21, pp. 75–219

1982 A survey of Π^1_2-logic, in LMPS VI, pp. 89–107

1982 a Proof Theoretic investigations of inductive definitions, in E. Engels, H. Läuchli, V. Strassen (eds.), Logic and Algorithmic L'Enseignement Math., Genf, pp. 207–236

1982 b Herbrand's theorem and proof theory, in J. Stern (ed.), Proc.of the Herbrand Symposium Logic Coll. 1981, pp. 29–38

1987 Proof theory and logical complexity, Studies in proof theory, Monographs 1, Bibliopolis, Neapel

J.Y. Girard and J. Vauzeilles:

1984 Functors and ordinal notations I: A functorial construction of the Veblen hierarchy, JSL 49, pp. 713–729

1984 a Functors and ordinal notations II: A functorial construction of the Bachmann hierarchy, JSL 49, pp. 1079–1114

K. Gödel:

1931 Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I, Monatshefte Math.Phys. 38,pp. 173–198

1931/2 Diskussion zur Grundlegung der Mathematik, Erkenntnis 2, pp. 147–151

1932/3 Zur intuitionistischen Arithmetik und Zahlentheorie; Ergebnisse eines math. Koll. 4, pp. 34–38

1958 Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, Dialectica 12, pp. 280–287

L. Henkin:

1954 A generalization of the concept of ω-consistency, JSL 19, pp. 183–196

D. Hilbert:

1932 Gesammelte Abhandlungen, Bd. I, Springer, Berlin, Reprint New York 1965

1933 Gesammelte Abhandlungen, Bd. II, Springer, Berlin, Reprint New York 1965

1935 Gesammelte Abhandlungen, Bd. III, Springer, Berlin, Reprint New York 1965

D. Hilbert and P. Bernays:

1934 Grundlagen der Mathematik I, Springer, Berlin

1970 Grundlagen der Mathematik II, Springer, Berlin, Heidelberg, NY

P.G. Hinman:

1978 Recursion-theoretic hierarchies, Springer, Berlin, Heidelberg, NY

W.A. Howard:

1963 The Axiom of choice (Σ^1_1-AC_{01}), bar induction and bar recursion, SR, pp. 71–114

1963 a Transfinite induction and transfinite recursion, SR, pp. 207–262

1972 A system of abstract constructive ordinals, JSL 37, pp. 355–374

D. Isles:

1970 Regular ordinals and normal forms, in IPT, pp. 339–361

1971 Natural well-orderings, JSL 36, pp. 288–300

G. Jäger:

1979 Die konstruktible Hierarchie als Hilfsmittel zur beweistheoretischen Untersuchung von Teilsystemen der Analysis, Dissertation, München

1980 Beweistheorie von KPN, AMLG 20, pp. 53–64

1982 Zur Beweistheorie der Kripke-Platek-Mengenlehre über den natürlichen Zahlen, AMLG 22, pp. 121–139

1982 a Iterating admissibility in proof theory, in J. Stern (ed.), Proc. of the Herbrand Logic Colloquium 1981, NH, pp. 137–146

1983 A well-ordering proof for Feferman's theory T_0, AMLG 23, pp. 65–77

1984 The strength of admissibility without foundation, JSL 49, pp. 867–879

1984 a ρ-inaccessible ordinals, collapsing functions and a recursive notation system, AMLG 24, pp. 49–62

1984 b A version of Kripke-Platek set theory which is conservative over Peano arithmetic, ZML 30, pp. 3–9

1986 Theories for admissible sets: A unifying approach to proof theory, Habilitationsschrift, Bibliopolis Neapel

G. Jäger and W. Pohlers:

1982 Eine beweistheoretische Untersuchung von $(\Delta^1_2\text{-CA})+(\text{BI})$ und artverwandter Systeme, SDBA

198? Admissible proof theory, to appear

G. Jäger and K. Schütte:

1979 Eine syntaktische Abgrenzung der $(\Delta^1_1\text{-CA})$-Analysis, SDBA

R.G. Jeroslow:

1973 Redundancies in the Hilbert-Bernays derivability conditions for Gödel's second incompleteness theorem, JSL 38, pp. 359–367

G. Kreisel:

1951 On the interpretation of non-finitist proofs I, JSL 16, pp. 241–267

1952 On the interpretation of non-finitist proofs II, JSL 17, pp. 43–58

1958 Mathematical significance of consistency proofs, JSL 23, pp. 155–182

1958 a Hilbert's programme, dialectica 12, pp. 346–372

1959 Interpretation of classical analysis by means of constructive functionals of finite type, in A. Heyting (ed.), Constructivity in Mathematics, NH, pp. 101–128

1960 Ordinal logics and the characterisation of informal concepts of proof, in Proc. of the International Congress of Mathematicians at Edinburgh 1958, Cambridge University Press, Cambridge, pp. 289–299

1960 a La predicativite, Bull. Snc. Math. France 88, pp. 371–391

1963 Axiomatic results on second order arithmetic, SR, pp. 23–70

1963 a Generalized inductive definitions, SR, pp. 115–139

1968 A survey of proof theory, JSL 33, pp. 321–388

1970 Principles of proof and ordinals implicit in given concepts, in IPT, pp. 489–516

1971 A survey of proof theory II, in J.E.Fenstad (ed.), Proc. Second Scandinavian Logic Symposium, NH, pp. 109–170

1977 Wie die Beweistheorie zu ihren Ordinalzahlen kam und kommt, Jahresberichte DMV 78, pp. 177–223

1982 Finiteness theorems in arithmetic: An application of Herbrand's theorem for Σ_2-formulas, in J. Stern (ed.), Proc. Herbrand Symposium, NH, pp. 39–55

1987 Proof theory: Some personal recollections, in G. Takeuti 1987, pp. 395–405

G. Kreisel, G.E. Mints and S.G. Simpson:

1975 The use of abstract language in elementary metamathematics: Some pedagogic examples, LNM 453, pp. 38–131

G. Kreisel and A. Levy:

1968 Reflection principles and their use for establishing the complexity of axiomatic systems, ZML 14, pp. 97–142

G. Kreisel, J. Shoenfield and H. Wang:

1959 Number theoretic concepts and recursive well-orderings, AMLG, pp. 42–64

G. Kreisel and G. Takeuti:

1974 Formally self-referential propositions for cut-free classical analysis and related systems, Diss. Math. 118, pp. 1–50

S. Kripke:

1964 Transfinite recursion on admissible ordinals I, II, JSL 29, pp. 161-162

H. Levitz:

1970 On the relationship between Takeuti's ordinal diagrams O(n) and Schütte's system of ordinal notations Σ(n), in IPT, pp. 377-405

M.H. Löb:

1955 Solution of a problem of Leon Henkin, JSL 20, pp. 118-155

H. Luckhardt:

1973 Extensional Gödel functional interpretation: A consistency proof of classical analysis, LNM 306

G.E. Mints:

1971 Exact estimates of the provability of transfinite induction in the initial segments of arithmetic, Soviet Math. 1, pp. 85-91

1971 Quanitfier-free and one-quantifier systems, Zap.Nauch.Sem.,LOM Stek. Akad. Nank SSSR 20, pp. 115-133

1976 What can be done in PRA, Zap. Nanch. Sem., LOMI, vol. 60, pp. 93-102

Y.N. Moschovakis:

1974 Elementary induction on abstract structures, NH

1974a On non-monotone inductive definability, in Fund. Math. 82, pp. 39-83

J.v.Neumann:

1927 Zur Hilbert'schen Beweistheorie, Mathematische Zeitschrift 26, pp.1-46

1931 Bemerkungen zu den Ausführungen von Herrn S.Lesniewski über meine Arbeit 'Zur Hilbert'schen Beweistheorie', Fundamenta Mathematica 17,pp.331-334

H. Ono:

1987 Reflection principles in fragments of Peano arithmetic, ZML, pp. 317-333

S. Orey:

1956 On ω-consistency and related properties, JSL 21, pp. 246-256

P. Päppinghaus:

1985 Ptykes in Gödels T und verallgemeinerte Rekursion über Mengen und Ordinalzahlen, Habil.-Schrift, Hannover

C. Parsons:

1966 On a number-theoretic choice scheme and its relation to induction, Notices AMS 13, pp. 740

1971 Proof-theoretic analysis of restricted induction schemata, JSL 36, pp. 361

1971a On a number-theoretic choice scheme II, JSL 36, pp. 587

1972 On n-quantifier-induction, JSL 37, pp. 466–482

J. Paris and L. Harrington:

1977 A mathematical incompleteness in Peano-arithmetic, in HB, pp. 1133–1142

H. Pfeiffer:

1964 Ausgezeichnete Folgen für gewisse Abschnitte der zweiten und weiteren Zahlklassen, Dissertation, Hannover

1970 Ein Bezeichnungssystem für Ordinalzahlen, AMLG 13, pp. 74–90

R. Platek:

1966 Foundations of recursion theory, Ph. D. Thesis, Stanford

W. Pohlers:

1973 Eine Grenze für die Herleitbarkeit der transfiniten Induktion in einem schwachen Π^1_1-Fragment der klassischen Analysis, Dissertation, München

1975 An upper bound for the provability of transfinite induction in systems with N-times iterated definitions, in LNM 500, pp. 271–289

1977 Beweistheorie der iterierten induktiven Definitionen, Habil.-Schrift, München

1978 Ordinals connected with formal theories for transfinitely iterated inductive definitions, JSL 43

1981 Cut elimination for impredicative infinitary systems, part I: Ordinal analysis of ID_1, AMLG 21, pp. 69–87

1981a Proof-theoretical analysis of ID_ν by the method of local predicativity, in BFPS pp. 261–357

1982 Cut elimination for impredicative infinitary systems, part II: Ordinal analysis for iterated inductive definitions, AMLG 22, pp. 113–129

1982a Admissibility in proof theory, in LMPS VI, pp. 123–139

1986 Beweistheorie, in: Jahrbuch Überblicke Mathematik, Bibliographigraphisches Institut

1987 Ordinal notations based on a hierarchy of inaccessible cardinals, APAL 33, pp. 157–179

1987 a Contributions of the Schütte school in Munich to proof theory, in Takeuti 1987, pp. 406–431

1990 Proof theory and ordinal analysis. To appear.

W. Purkert and H.J.Ilgauds

1987 Georg Cantor 1845–1918, Birkhäuser, Basel, Boston, Stuttgart

M.Rathjen: Untersuchungen zu Teilsystemen der Zahlenthoerie zweiter Stufe und der Mengenlehre mit einer zwischen Δ_2^1 – CA und Δ_2^1 – CA + BI liegenden Beweisstärke. Dissertation, Münster 1989

W. Richter:

1965 Extensions of the constructive ordinals, JSL 30, pp. 193–211

H. Rogers:

1967 Theory of recursive functions and effective computability, McGraw-Hill, New York

J.B. Rosser:

1936 Extensions of some theorems of Gödel and Church, JSL 1, pp. 87–91

B. Scarpellini:

1969 Some applications of Gentzen's second consistency proof, MA 181, pp. 325–344

1970 On cut elimination in intuitionistic systems of analysis, in IPT, pp. 271–285

1971 A model for bar recursion of higher types, Compo. Math. 23

1971 a Proof theory and intuitionistic systems, Springer, Berlin, Heidelberg, NY

1972 Formally constructive model for bar recursion of higher types, ZML 18, pp. 321–383

1972 a Induction and transfinite induction in intuitionistic systems, APAL 4, pp. 173–227

1973 On bar induction of higher types for decidable predicates, APAL 5, pp. 77–164

U. Schmerl:

1979 A fine structure generated by reflection formulas over primitve recursive arithmetic, in M. Boffa, D. van Dalen, K. McAloon (eds.), Logic Colloquium 1978, NH

1982 A proof theoretical fine structure in systems of ramified analysis, AMLG 22, pp. 167–186

1982 a Iterated reflection principles and the ω-rule, JSL 47, pp. 721–733

1982 b Number theory and the Bachmann/Howard ordinal, in: J. Stern (ed.), Proc. of the Herbrand Symposium Log. Coll. 1981, pp. 287–298

K. Schütte:

1951 Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie, MA 122, pp. 369–389

1952 Beweistheoretische Untersuchung der verzweigten Analysis, MA 124, pp. 123–147

1954 Kennzeichnung von Ordnungszahlen durch rekursiv erklärte Funktionen MA 127, pp. 16–32

1960 Beweistheorie, Springer, Berlin

1964 Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik, AMLG 67, pp. 45–60

1965 Predicative well-orderings, in Crossley, Dummett (eds.), Formal systems and recursive functions, NH, pp. 176–184

1969 Ein konstruktives System von Ordinalzahlen, AMLG 11, pp. 126–137, AMLG 12, pp. 3–11

1975 Primitiv rekursive Ordinalzahlfunktionen, SDBA

1976 Einführung der Normalfunktionen Θ_x ohne Auswahlaxiom und ohne Regularitätsbedingung, AMLG 17, pp. 171–178

1977 Proof Theory, Springer, Berlin

1986/7 Majorisierungsrelationen und Fundamentalfolgen eines Ordinalzahlsystems von G. Jäger, AMLG 26, pp. 29–55

1987 Eine beweistheoretische Abgrenzung des Teilsystems der Analysis mit Π^1_2-Separation und Bar-Induktion, SDBA

1988 Ein Wohlordnungsbeweis für das Ordinalzahlensystem T(J), AMLG 27, pp.5–20

K. Schütte and S.G. Simpson:

1985 Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen, ALMG 25, pp. 75–89

H. Schwichtenberg:

1977 Proof theory: Some applications of cut-elimination, in HB, pp. 867–895

1987 Ein einfaches Verfahren zur Normalisierung unendlicher Herleitungen in E. Börger (ed.), Computation Theory and Logic, Lecture in Computer Science 270, pp. 334-384

J.R. Shoenfield:

1967 Mathematical logic, Edison-Verlag, Reading Mass.

W. Sieg:

1977 Trees in metamathematics (Theories of inductive definitions and subsystems of analysis), Ph.D. thesis, Stanford

1981 Inductive definitions, constructive ordinals and normal derivations, in BFPS, pp. 143-187

1984 Foundations for analysis and proof theory, Synthese 60, pp. 159-200

1985 Fragments of arithmetic, in APAL 28, pp. 33-71

1085 a Reductions of theories for analysis, in G. Dorn and P. Weingartner (eds.), Foundations of Logic and Linguistics, NY, London, pp. 199-230

1987 Relative Konsistenz, in E. Börger (ed.), Computation Theory and Logic, Lecture Notes in Computer Science 270, pp. 360-381

1988 Hilbert's program sixty years later, JSL 53, pp. 338-384

S.G. Simpson:

1980 Set theoretic aspects of ATR_0, in D. van Dalen, D. Lascar and J. Smiley (eds.), Logic colloquium 1980, NH, pp. 255-271

1980 a Σ^1_1 and Π^1_1 transfinite induction, ibid., pp. 239-253

1984 Which set existence axioms are needed to prove the Cauchy/Peano theorem for ordinary differential equations?, JSL 49, pp. 783-802

1985 Reverse mathematics, in A. Nerode and R. Shore (eds.), Proc. Symp. Pure math., AMS 42, pp. 461-471

1985 a Friedman's research on subsystems of second order arithmetic, in HF, pp. 137-159

1985 b Nichtbeweisbarkeit von gewissen kombinatorischen Eigenschaften endlicher Bäume, AMLG 25, pp. 45-65

1985 c Nonprovability of certain combinatorial properties of finite trees, in AF, pp. 87-118

1987 Subsystems of Z_2 and reverse mathematics, in G.Takeuti 1978, pp. 432-446

1988 Partial realizations of Hilbert's program, JSL 53, pp. 349-363

S.G. Simpson and R.L.Smith:

1986 Factorization of polynomials and Σ_1^0 induction, APAL 31, pp. 289-306

R.L. Smith:

1985 The consistency strengths of some finite forms of the Higman and Kruskal theorems, in HF, pp. 119-136

G. Smorynski:

1977 The incompleteness theorems, in HB, pp. 821-865

1985 Self reference and modal logic, Springer, Berlin, Heidelberg, NY

C. Spector:

1961 Inductively defined sets of numbers, in: Infinitistic methods, Proc. Warszaw Symp., Pergamon Press, Oxford, pp. 97-102

1962 Provably recursive functions of analysis, in Recursive Function Theory, AMS, Proc. Symp. Pure Math. 5, pp. 1-27

W.W. Tait:

1965 Functionals defined by transfinite recursion, JSL 30, pp. 155-174

1965 a Infinitely long terms of transfinite type, in Crossley and Dummett (eds.), Formal Systems and Recursive Functions, Proc. 8th Logic Colloquium, Oxford 1963, NH, pp. 176-185

1968 Normal derivability in classical logic, in J. Barwise (ed.), LNM 72, pp. 204-236

1970 Applications of the cut elimination theorem to some subsystems of classical analysis, in IPT, pp. 475-488

S. Takahashi:

1986 Monotone inductive definitions in T_0, Dissertation, Stanford

G. Takeuti:

1953 On a generalized logic calculus, Japan J. Math. 23, pp. 39-96; errata/addenda ibid 24, pp. 149-156

1957 On the theory of ordinal numbers, J. Math. Soc. Japan 9, pp. 93-113; errata/addenda ibid 12, p. 127

1957 a Ordinal diagrams, J. Math. Soc. Japan 9, pp. 386-394

1958 On the formal theory of the ordinal diagrams, Ann. Jap. Ass. Phil. Sci. 1, pp. 151-170

1960 Ordinal diagrams II, J. math. Soc. Japan 12, pp. 385-391

1961 On the inductive definition with quantifiers of second order, J.

Math. Soc. Japan 13, pp. 333–341

1963 A remark on Gentzen's paper "Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie" I–II, Proc. Japan Acad. 39, pp. 263–269

1975 Consistency proofs and ordinals, in J. Diller and G. Müller (eds.), Proof theory Symposium, Kiel 1974, LNM 500, pp. 365–369

1985 Proof theory and set theory, Synthese 62, pp. 255–263

1967 Consistency proofs of subsystems of classical analysis, Ann. Math. 86, pp. 299–348

1975 Proof theory, NH

1978 Two applications of logic to mathematics, Pubs. of the math. Soc. of Japan, Iwanami Shoten and Princeton University Press.

1987 Proof Theory, Second edition, NH

G. Takeuti and M. Yasugi:

1973 The ordinals of the systems of second order arithmetic with the provably Δ^1_2–comprehension axiom and with the Δ^1_2–comprehension axiom respectively, Japan J. Math. 41, pp. 1–67

1968 Reflection principles of subsystems of analysis, in H.A. Schmidt, K. Schütte and H.-J. Thiele (eds.), Contributions to Mathematical Logic, Proc. of the Logic Colloquium, NH

1976 Fundamental sequences and ordinal diagrams, Comm. Math. Univ. St. Pauli (Tokyo) 25, pp. 1–80

1981 An accessibility proof of ordinal diagrams, J. Math. Soc. Japan 33, pp. 1–21

A.S. Troelstra:

1973 (ed.) Metamathematical investigations of intuitionistic arithmetic and analysis, LNM 344, with contributions by A.S. Troelstra, C.A. Smorynsky, J.I. Zucker and W.A. Howard

1977 Aspects of constructive mathematics, in HB, pp. 973–1052

A.M. Turing:

1939 Systems of logic based on ordinals, Proc. London Math. Soc. Ser. II 45, pp. 161–228

O. Veblen:

1908 Continuous increasing functions of finite and transfinite ordinals, Trans, AMS 9, pp. 280–292

H. Wang:

1953 Certain predicates defined by induction schemata, JSL 18, pp. 49–59

R. Weyhrauch:

1975 Relations between some hierarchies of ordinal functions and functionals, Dissertation, Stanford

H. Weyl:

1918 Das Kontinuum, Veit, Leipzig

1921 Über die neue Grundlagenkrise der Mathematik. Math. Zeitschrift 10 pp. 39–79

1925 Die heutige Erkenntnislage der Mathematik. Symposion 1, pp.1–32

1929 Consistency in mathematics, Rice Institut Pamphlet 16, pp. 245–265

1949 Philosophy of mathematics and natural sciences, Princeton

J.E. Zucker:

1971 Proof-theoretic studies of systems of iterated inductive definition and subsystems of analysis, Dissertation, Stanford

SUBJECT INDEX

INDEX OF NOTATION