Beweistheorie

Wilfried Buchholz

Skriptum einer 4-std. Vorlesung im Wintersemester 1997/98

Mathematisches Institut der Universität München

Literatur

GENTZEN, G.: *Die Widerspruchsfreiheit der reinen Zahlentheorie.* Math. Ann. **112** (1936), pp. 493-565

GENTZEN, G.: *Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie.* Forschungen zur Logik und zur Grundlegung der exakten Wissenschaften, New Series 4, Leipzig (1938), pp. 19-44

GENTZEN, G.: *Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie.* Math. Ann. **119** (1943), pp. 149-161

GIRARD, J.Y.: *Proof Theory and Logical Complexity.* Studies in Proof Theory, Monographs 1. Napoli: Bibliopolis 1987

POHLERS, W.: *Proof Theory.* Springer LNM 1407

SCHWICHTENBERG, H.: *Proof Theory: some applications of cut-elimination.* In: Barwise, J.(ed.) Handbook of Mathematical Logic. Amsterdam: North-Holland 1977

SCHÜTTE, K.: *Proof Theory.*

TAIT, W.W.: *Normal Derivability in Classical Logic.* In: Springer LNM 72 (1986)

TAKEUTI, G.: *Proof Theory.*

WAINER and WALLEN: *Basic Proof Theory.* In: Aczel, Simmons, Wainer (eds.) Proof Theory. Cambridge University Press 1993.

§1 The Tait calculus for classical 1^{st}-order predicate logic without equality

Formal language

Basic symbols:

1. Variables $v_0, v_1, v_2...$ (denoted by $x, y, z, x_1, ...$)
2. $\neg$, $\wedge$, $\vee$, $\forall$, $\exists$

Let $\mathcal{L}$ be some fixed (countable) language, i.e. set of function and predicate symbols where each symbol $p \in \mathcal{L}$ has a certain *arity* $\#(p) \in \mathbb{N}$.

From now on all syntactic notions such as terms, formulas, sequents,.... are defined with respect to $\mathcal{L}$.

Terms are defined as usual:

1. Every variable is a term.
2. If f is an n-ary function symbol ($n \geq 0$) and $t_1, ..., t_n$ are terms then the string $ft_1...t_n$ is a term.

We use $s, t, s_1, ...$ as syntactic variables for terms.

An *atomic formula* is an expression $pt_1...t_n$ where p is an n-ary predicate symbol and $t_1, ..., t_n$ are terms.

An expression of the form A or $\neg A$, where A is an atomic formula, is called a (positive or negative) *literal.*

Inductive definition of *formulas*

1. Every literal is a formula.
2. If A, B are formulas then also $\wedge AB$ and $\vee AB$ are formulas.
3. If A is a formula then $\forall xA$ and $\exists xA$ are formulas.

As usual we write $A \wedge B$, $A \vee B$ for $\wedge AB$, $\vee AB$.

We use A, B, C, D, F, G as syntactic variables for formulas.

Definition of the negation $\text{neg}(A)$ of a formula A

1. If A is atomic then $\text{neg}(A) := \neg A$ and $\text{neg}(\neg A) := A$.
2. $\text{neg}(A \wedge B) := \text{neg}(A) \vee \text{neg}(B)$, $\text{neg}(A \vee B) := \text{neg}(A) \wedge \text{neg}(B)$.
3. $\text{neg}(\forall xA) := \exists x\,\text{neg}(A)$, $\text{neg}(\exists xA) := \forall x\,\text{neg}(A)$.

Corollary: $\text{neg}(A)$ is a formula, and $\text{neg}(\text{neg}(A)) = A$.

Notation:

(i) From now on we write $\neg A$ for $\text{neg}(A)$.

(ii) $A_1 \to ... \to A_n \to B := \neg A_1 \vee (\neg A_2 \vee (... \vee (\neg A_n \vee B)...))$.

For each term or formula E we define the set $\text{FV}(E)$ [$\text{BV}(E)$] of its free [bound] variables in the usual way. Further let $\text{Var}(E) := \text{FV}(E) \cup \text{BV}(E)$ (the set of all variables occuring in E). If $\mathcal{X}$ is a set of terms and/or formulas then $\text{FV}(\mathcal{X}) := \bigcup\{\text{FV}(E) : E \in \mathcal{X}\}$; analogously $\text{BV}(\mathcal{X})$ and $\text{Var}(\mathcal{X})$ are defined.

Substitution

Definition of $E_x(t)$

1. If E is a term or a literal then $E_x(t)$ is obtained from E by replacing every occurrence of x in E by t.
2. $(A \diamond B)_x(t) := A_x(t) \diamond B_x(t) \qquad (\diamond \in \{\wedge, \vee\})$

3.1. $(QxA)_x(t) := QxA \qquad (Q \in \{\forall, \exists\})$

3.2. $(QyA)_x(t) := Qy\, A_x(t)$, if $x \neq y \qquad (Q \in \{\forall, \exists\})$

Abbreviation: $\text{subst}(E, x, t) \;:\Leftrightarrow\; t$ is free for x in E.

Definition of $\text{rk}(A)$

1. $\text{rk}(A) := 0$, if A is a literal.
2. $\text{rk}(A \diamond B) := \max\{\text{rk}(A), \text{rk}(B)\} + 1$.
3. $\text{rk}(QxA) := \text{rk}(A) + 1$.

Corollary: $\text{rk}(\neg A) = \text{rk}(A) = \text{rk}(A_x(t))$.

Definition

$\mathsf{FV} := \{v_{2i} : i \in \mathbb{N}\}$, $\mathsf{BV} := \{v_{2i+1} : i \in \mathbb{N}\}$.

A formula A is called *decent* if $\text{FV}(A) \subseteq \mathsf{FV}$ and $\text{BV}(A) \subseteq \mathsf{BV}$.

Ter := set of all terms t with $\text{FV}(t) \subseteq \mathsf{FV}$.

Finite sets of decent formulas are called *sequents*.

Syntactic variables for sequents are Γ, Δ.

Proof systems

In the following we mostly write $A_1, ..., A_n$ for $\{A_1, ..., A_n\}$, and A, Γ, Δ for $\{A\} \cup \Gamma \cup \Delta$, etc.

A proof system $\mathfrak{S}$ is given by

- a set of formal expressions called *inference symbols* (syntactic variable $\mathcal{I}$)
- for each inference symbol $\mathcal{I}$ a set $|\mathcal{I}|$ (the *arity* of $\mathcal{I}$), a sequent $\Delta(\mathcal{I})$ and a family of sequents $(\Delta_\iota(\mathcal{I}))_{\iota \in |\mathcal{I}|}$. The elements of $\Delta(\mathcal{I})$ [$\bigcup_{\iota \in |\mathcal{I}|} \Delta_\iota(\mathcal{I})$] are called the *principal formulas* [*minor formulas*] of $\mathcal{I}$.
- for each inference symbol $\mathcal{I}$ a set $\text{Eig}(\mathcal{I})$ which is either empty or a singleton $\{y\}$ with $y \in \mathsf{FV} \setminus \text{FV}(\Delta(\mathcal{I}))$; in the latter case y is called the *eigenvariable* of $\mathcal{I}$.
- for each inference symbol $\mathcal{I}$ an ordinal $\deg(\mathcal{I})$.

NOTATION

By writing

$(\mathcal{I}) \quad \dfrac{\ldots \Delta_\iota \ldots (\iota \in I)}{\Delta}$ [$!y!$]

we express that $\mathcal{I}$ is an inference symbol with $|\mathcal{I}| = I$, $\Delta(\mathcal{I}) = \Delta$, $\Delta_\iota(\mathcal{I}) = \Delta_\iota$, and $\text{Eig}(\mathcal{I}) = \emptyset$ [$\text{Eig}(\mathcal{I}) = \{y\}$, resp.].

If $|\mathcal{I}| = \{0, ..., n\}$ we write $\dfrac{\Delta_0 \;\; \Delta_1 \;\; \ldots \;\; \Delta_n}{\Delta}$, instead of $\dfrac{\ldots \Delta_\iota \ldots (\iota \in I)}{\Delta}$.

Inference symbols $\mathcal{I}$ with $|\mathcal{I}| = \emptyset$ will be called *axioms*.

By writing " $(\mathcal{I})\;\; \Delta$ " we declare $\mathcal{I}$ as an axiom with $\Delta(\mathcal{I}) := \Delta$.

For almost all inference symbols (except axioms) the sequents $\Delta(\mathcal{I}), \Delta_\iota(\mathcal{I})$ are singletons or empty.

Example:

By (Cut_C) $\dfrac{C \qquad \neg C}{\emptyset}$ we express that for each decent formula C, the expression $\mathcal{I} := \mathsf{Cut}_C$ is an inference symbol with $|\mathcal{I}| = \{0,1\}$, $\Delta(\mathcal{I}) = \emptyset$, $\Delta_0(\mathcal{I}) = \{C\}$, $\Delta_1(\mathcal{I}) = \{\neg C\}$.

Inductive definition of $\mathfrak{S}$-derivations

If $\mathcal{I}$ is an inference symbol of $\mathfrak{S}$, and $(d_\iota)_{\iota\in|\mathcal{I}|}$ is a family of $\mathfrak{S}$-*derivations* such that $\mathrm{Eig}(\mathcal{I})\cap\mathrm{FV}(\Gamma) = \emptyset$ where $\Gamma := \Delta(\mathcal{I})\cup\bigcup_{\iota\in|\mathcal{I}|}(\Gamma(d_\iota)\setminus\Delta_\iota(\mathcal{I}))$, then $d := \mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}$ (or $\mathcal{I}d_0...d_{n-1}$ if $|\mathcal{I}| = \{0,...,n{-}1\}$) is an $\mathfrak{S}$-*derivation* with $\Gamma(d) := \Gamma$, $\mathsf{last}(d) := \mathcal{I}$, $\deg(d) := \sup(\{\deg(\mathcal{I})\}\cup\{\deg(d_\iota) : \iota \in |\mathcal{I}|\})$, $lg(d) := \sup\{lg(d_\iota){+}1 : \iota \in |\mathcal{I}|\}$.

$\Gamma(d)$ is called *the endsequent of* d, and $\mathsf{last}(d)$ is called the last inference (symbol) of d.

A proof system $\mathfrak{S}$ also comprises the assignment of an ordinal $\mathrm{o}(d)$ to each $\mathfrak{S}$-derivation d.
If nothing else is said we define $\mathrm{o}(d) := lg(d)$.

Abbreviations

$\mathfrak{S} \ni d \vdash^\alpha_\rho \Gamma \ :\Longleftrightarrow\ d$ is an $\mathfrak{S}$-derivation with $\mathrm{o}(d) \le \alpha$, $\deg(d) \le \rho$, $\Gamma(d) \subseteq \Gamma$,
$\mathfrak{S} \vdash^\alpha_\rho \Gamma \ :\Longleftrightarrow\ \mathfrak{S} \ni d \vdash^\alpha_\rho \Gamma$ for some d,
$\mathfrak{S} \vdash_\rho \Gamma \ :\Longleftrightarrow\ \mathfrak{S} \vdash^\alpha_\rho \Gamma$ for some α.

If $\mathfrak{S}$ is known from the context, we write $d \vdash^\alpha_\rho \Gamma$ instead of $\mathfrak{S} \ni d \vdash^\alpha_\rho \Gamma$, etc.

Remark (for proof systems with $\mathrm{o}(d) = lg(d)$)
Assume that $\mathcal{I} \in \mathfrak{S}$ & $\deg(\mathcal{I}) \le \rho$ & $\mathrm{Eig}(\mathcal{I}) \cap \mathrm{FV}(\Gamma) = \emptyset$ & $\forall\iota \in |\mathcal{I}|(\alpha_\iota < \alpha)$:
$(\forall\iota \in |\mathcal{I}|)\, \mathfrak{S} \ni d_\iota \vdash^{\alpha_\iota}_\rho \Gamma, \Delta_\iota(\mathcal{I}) \implies \mathfrak{S} \ni \mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|} \vdash^\alpha_\rho \Gamma, \Delta(\mathcal{I})$.

The proof system PL1 (Tait's version of Gentzen's calculus of sequents)

(LogAx_A) $A, \neg A$ if $\mathrm{rk}(A) = 0$

$(\bigwedge_{A_0\wedge A_1})$ $\dfrac{A_0 \qquad A_1}{A_0\wedge A_1}$ $\qquad$ $(\bigvee^k_{A_0\vee A_1})$ $\dfrac{A_k}{A_0\vee A_1}$ $(k \in \{0,1\})$

$(\bigwedge^y_{\forall xA})$ $\dfrac{A_x(y)}{\forall xA}$ $!y!$ $\qquad$ $(\bigvee^t_{\exists xA})$ $\dfrac{A_x(t)}{\exists xA}$ $(t \in \mathrm{Ter})$

(Cut_C) $\dfrac{C \qquad \neg C}{\emptyset}$

$\deg(\mathsf{Cut}_C) := \mathrm{rk}(C){+}1$ and $\deg(\mathcal{I}) := 0$ for all other inference symbols.
$\mathrm{o}(d) := lg(d)$.

Displaying derivations:

To increase the readability we often write derivations in tree form, i.e. we write $\dfrac{d_0 \ldots d_n}{\mathcal{I}}$ instead of $\mathcal{I}d_0...d_n$.

Another way of representing derivations (close to the traditional one) is to write them as trees of sequents and to display the respective inference symbols at the right or left end of each inference line. Mostly we will not show the full inference symbol $\mathcal{I}$ but only some kind of abbreviation (e.g. the outermost logical symbol of the principal formula of $\mathcal{I}$) or nothing. The sequents and inference symbols will be arranged in such a way that at each node ν the relation $\dfrac{\Gamma_{\nu*\langle 0\rangle} \ldots \Gamma_{\nu*\langle 1\rangle}}{\Gamma_\nu}\mathcal{I}_\nu$ is satisfied (where Γ_ν [$\mathcal{I}_\nu$, resp.] is the sequent [inference symbol] at node ν). Consequently we have $\Gamma(d|_\nu) \subseteq \Gamma_\nu$, if $d|_\nu$ is the subderivation with root ν.

Two examples: Abbreviation: $\mathsf{Ax}_F := \mathrm{LogAx}_F$

I. $d \;= \bigvee^{t}_{\exists x(\neg Rfx\vee Rx)}\bigvee^{ft}_{\exists x(\neg Rfx\vee Rx)}\bigvee^{0}_{\neg Rft\vee Rt}\bigvee^{1}_{\neg Rfft\vee Rft}\mathsf{Ax}_{Rft} =$

$$= \dfrac{\dfrac{\dfrac{\dfrac{\mathsf{Ax}_{Rft}}{\bigvee^{1}_{\neg Rfft\vee Rft}}}{\bigvee^{0}_{\neg Rft\vee Rt}}}{\bigvee^{ft}_{\exists x(\neg Rfx\vee Rx)}}}{\bigvee^{t}_{\exists x(\neg Rfx\vee Rx)}} \quad = \quad \dfrac{\dfrac{\dfrac{\dfrac{Rft,\neg Rft}{\neg Rfft\vee Rft,\neg Rft}\,(\vee)}{\neg Rfft\vee Rft,\neg Rft\vee Rt}\,(\vee)}{\exists x(\neg Rfx\vee Rx),\neg Rft\vee Rt}\,(\exists)}{\exists x(\neg Rfx\vee Rx)}\,(\exists)$$

II. Let $G := \exists x(F(x)\wedge\neg F(Sx))$,

$d \;= \bigvee^{S0}_{G}\bigwedge_{F(S0)\wedge\neg F(SS0)}\bigvee^{0}_{G}\bigwedge_{F(0)\wedge\neg F(S0)}\mathsf{Ax}_{F(0)}\mathsf{Ax}_{F(S0)}\;\mathsf{Ax}_{F(SS0)} =$

$$= \dfrac{\dfrac{\dfrac{\dfrac{\mathsf{Ax}_{F(0)}\quad \mathsf{Ax}_{F(S0)}}{\bigwedge_{F(0)\wedge\neg F(S0)}}}{\bigvee^{0}_{G}}\qquad\qquad \mathsf{Ax}_{F(SS0)}}{\bigwedge_{F(S0)\wedge\neg F(SS0)}}}{\bigvee^{S0}_{G}} \quad = \quad \dfrac{\dfrac{\dfrac{\dfrac{\neg F(0),G,F(0)\quad \neg F(S0),F(S0)}{\neg F(0),G,F(0)\wedge\neg F(S0),F(S0)}\,(\wedge)}{\neg F(0),G,F(S0)}\,(\exists)\qquad \neg F(SS0),F(SS0)}{\neg F(0),G,F(S0)\wedge\neg F(SS0),F(SS0)}\,(\wedge)}{\neg F(0),G,F(SS0)}\,(\exists)$$

Definition

$\Sigma\models C \iff C$ is a logical consequence from Σ (Σ a set of formulas)

$\Sigma\models\{A_1,...,A_n\} \iff \Sigma\models A_1\vee...\vee A_n$

$\models\Gamma \iff \emptyset\models\Gamma$.

An *axiom system* is a set of *closed* formulas.

Theorem 1.1

$\models\Gamma \implies \mathrm{PL1}\vdash_0\Gamma$.

Corollary

If Σ is an axiom system and C a decent formula then the following holds:

$\Sigma\models C \implies$ There are $A_1,...,A_n\in\Sigma$ such that $\mathrm{PL1}\vdash_0\neg A_1,...,\neg A_n,C$.

Proof of Theorem 1.2:

AX := set of finite sequences $(A_0,...,A_l)$ such that there is a prime formula A with $\{A,\neg A\}\subseteq\{A_1,...,A_l\}$.

Let Π be a finite sequence of decent formulas .

$t_0,t_1,...$: enumeration of Ter.

μ,ν are ranging over finite 0-1-sequences (elements of $\{0,1\}^{<\omega}$).

$\mu\sqsubseteq\nu$:$\Leftrightarrow$ μ is an initial segment of ν (i.e. $\nu=\mu*\tau$ for some $\tau\in\{0,1\}^{<\omega}$)

For each $\nu\in\{0,1\}^{<\omega}$ we define a finite sequence of decent formulas Π_ν.

The definition proceeds by recursion on $lh(\nu)$.

1. $\Pi_{\langle\,\rangle} := \Pi$,

Let $n = lh(\nu)$, and assume that Π_μ is already defined for each $\mu\sqsubseteq\nu$.

2. $\Pi_\nu \in \mathrm{AX}$ or all formulas in Π_ν are literals: $\Pi_{\nu*\langle i\rangle} := \Pi_\nu$,

3. $\Pi_\nu = \Pi', A, \Pi'' \notin \mathrm{AX}$, and $\mathrm{rk}(A) > 0$ while all formulas in Π' are literals:

3.1. $A = A_0 \wedge A_1$: $\Pi_{\nu*\langle i\rangle} := \Pi', A_i, \Pi''$,

3.2. $A = A_0 \vee A_1$: $\Pi_{\nu*\langle i\rangle} := \Pi', A_0, A_1, \Pi''$,

3.3. $A = \forall x B$: $\Pi_{\nu*\langle i\rangle} := \Pi', B_x(u), \Pi''$, where u is the first variable in $\mathsf{FV} \setminus \mathrm{FV}(\Pi_\nu)$,

3.4. $A = \exists x B$: $\Pi_{\nu*\langle i\rangle} := \Pi', B_x(t_k), \Pi'', A$, where k is minimal s.t. $(\, x \in \mathrm{FV}(B) \Rightarrow \forall \mu \sqsubseteq \nu\, B_x(t_k) \notin \Pi_\mu\,)$.

Assumption:

$(i_n)_{n\in\mathbb{N}}$ is a 0-1-sequence such that $\forall n \in \mathbb{N}(\,\Pi_{\langle i_0,\dots,i_{n-1}\rangle} \notin \mathrm{AX}\,)$.

Abbreviation: $\nu(n) := \langle i_0, \dots, i_{n-1}\rangle$, $\Phi := \bigcup_{n\in\mathbb{N}} \Pi_{\nu(n)}$

Definition: If not all formulas in Π_ν are literals let $\mathrm{dp}(\Pi_\nu)$ be the first formula in Π_ν which is not a literal.

Proposition 1

$A \in \Pi_{\nu(n)}$ & $\mathrm{rk}(A) > 0 \implies \exists k \ge n(\,\mathrm{dp}(\Pi_{\nu(k)}) = A\,)$

Proposition 2

a) $\mathrm{rk}(A) = 0 \implies A \notin \Phi \vee \neg A \notin \Phi$,

b) $A_0 \wedge A_1 \in \Phi \implies A_0 \in \Phi$ or $A_1 \in \Phi$,

c) $A_0 \vee A_1 \in \Phi \implies A_0 \in \Phi$ and $A_1 \in \Phi$,

d) $\forall x B \in \Phi \implies \exists u \in \mathsf{FV}(\,B_x(u) \in \Phi\,)$,

e) $\exists x B \in \Phi \implies \forall t \in \mathrm{Ter}(\,B_x(t) \in \Phi\,)$.

Proof:

a) $\mathrm{rk}(A) = 0$ & $A \in \Pi_{\nu(n)}$ & $\neg A \in \Pi_{\nu(m)} \Rightarrow \forall k \ge \max\{m,n\}(A, \neg A \in \Pi_{\nu(k)})$.

Assume that $A \in \Phi$ with $\mathrm{rk}(A) > 0$. Then $A = \mathrm{dp}(\Pi_{\nu(k)})$ for some k.

b) $A = A_0 \wedge A_1$: Then $A_i \in \Pi_{\nu(k)*\langle i\rangle}$ for $i = 0, 1$, and therefore $A_0 \in \Pi_{\nu(k+1)}$ or $A_1 \in \Pi_{\nu(k+1)}$.

c) $A = A_0 \vee A_1$: Then $A_0, A_1 \in \Pi_{\nu(k)*\langle i\rangle}$ for $i = 0, 1$, and therefore $A_0 \in \Pi_{\nu(k+1)}$ and $A_1 \in \Pi_{\nu(k+1)}$.

d) $A = \forall x B$: Then $B_x(u) \in \Pi_{\nu(k)*\langle i\rangle}$ for some $u \in \mathsf{Var}_0$.

e) $A = \exists x B$: Then $\forall n \ge k_0(\,\exists x B \in \Pi_{\nu(n)}\,)$ $(*)$.

By induction on m we prove $B_x(t_m) \in \Phi$:

Assume that $B(t_i) \in \Phi$ for all $i < m$. By Proposition 1 and $(*)$ there is an $n \ge k$ such that $\mathrm{dp}(\Pi_{\nu(n)}) = \exists x B$ and $\forall i < m \exists j \le n(\,B(t_i) \in \Pi_{\nu(j)}\,)$ $(**)$.

By definition of $\Pi_{\nu(n+1)}$ we get $\big(\forall j \le n(\,B(t_m) \notin \Pi_{\nu(j)}\,) \Rightarrow B(t_m) \in \Pi_{\nu(n+1)}\big)$ and thus $B(t_m) \in \Phi$.

$|\mathcal{M}| := \mathrm{Ter}$, $f^{\mathcal{M}}(s_1, \dots, s_n) := f s_1 \dots s_n$, $p^{\mathcal{M}}(s_1, \dots, s_n) :\Leftrightarrow p s_1, \dots, s_n \notin \Phi$

Let $\xi : \mathsf{Var} \to \mathrm{Ter}$ such that $\xi(x) = x$ for $x \in \mathsf{FV}$.

Proposition

a) $t^{\mathcal{M}}[\xi] = t$ for each $t \in \mathrm{Ter}$,

b) $A \in \Phi \Rightarrow \mathcal{M} \not\models A[\xi]$.

Proof of b) by induction on $\mathrm{rk}(A)$:

$\forall xB \in \Phi \;\Rightarrow\; B_x(u) \notin \Phi$ for some $u \in \mathsf{FV} \overset{\text{IH}}{\Rightarrow} \mathcal{M} \not\models B_x(u)[\xi] \;\Rightarrow\; \mathcal{M} \not\models (\forall xB)[\xi]$.

$\exists xB \in \Phi \;\Rightarrow\; B_x(t) \in \Phi$ for all $t \in \mathrm{Ter} \overset{\text{IH}}{\Rightarrow} \mathcal{M} \not\models B_x(t)[\xi]$ for all $t \in \mathrm{Ter} \overset{\text{a)}}{\Rightarrow} \mathcal{M} \not\models B[\xi_x^t]$ for all $t \in \mathrm{Ter} \;\Rightarrow\; \mathcal{M} \not\models (\exists xB)[\xi]$.

$\Pi \subseteq \Phi$ & $\mathcal{M} \not\models A[\xi]$ for each $A \in \Phi \implies \not\models \Pi$.

Hence the assumption was false.

By Königs Lemma (and since $\Pi_\nu \in \mathrm{AX}$ implies $\Pi_{\nu*\langle i\rangle} \in \mathrm{AX}$) the set $\{\nu : \Pi_\nu \notin \mathrm{AX}\}$ is finite.
Let $m := \max\{lh(\nu) : \Pi_\nu \notin \mathrm{AX}\} + 1$.

By recursion on $m \dot{-} lh(\nu)$ we define a PL1-derivation d_ν with $d_\nu \vdash_0 \Pi_\nu$:

1. $\Pi_\nu \in \mathrm{AX}$: Then $d_\nu := \mathrm{LogAx}_A$ with suitable A.

2. $\Pi_\nu \notin \mathrm{AX}$: Then $lh(\nu) < m$ and thus $m \dot{-} lh(\nu*\langle i\rangle) < m \dot{-} lh(\nu)$.
Hence by IH (*) $d_{\nu*\langle i\rangle} \vdash_0 \Pi_{\nu*\langle i\rangle}$ for $i = 0, 1$.

Now not all formulas $A \in \Pi_\nu$ are literals, since otherwise $\Pi_{\nu*\mu} = \Pi_\nu \notin \mathrm{AX}$ for all μ.
Hence one of the following cases holds:

2.1. $\Pi_\nu = \Pi', A_0 \wedge A_1, \Pi''$ and $\Pi_{\nu*\langle i\rangle} = \Pi', A_i, \Pi''$: $d_\nu := \bigwedge_{A_0\wedge A_1} d_{\nu*\langle 0\rangle} d_{\nu*\langle 1\rangle}$.
2.2. $\Pi_\nu = \Pi', A_0 \vee A_1, \Pi''$ and $\Pi_{\nu*\langle 0\rangle} = \Pi', A_0, A_1, \Pi''$: $d_\nu := \bigvee^1_{A_0\vee A_1} \bigvee^0_{A_0\vee A_1} d_{\nu*\langle 0\rangle}$.
2.3. $\Pi_\nu = \Pi', \forall xB, \Pi''$ and $\Pi_{\nu*\langle 0\rangle} = \Pi', B_x(u), \Pi''$ with $u \in \mathsf{FV} \setminus \mathrm{FV}(\Pi_\nu)$: $d_\nu := \bigwedge^u_{\forall xB} d_{\nu*\langle 0\rangle}$.
2.4. $\Pi_\nu = \Pi', \exists xB, \Pi''$ and $\Pi_{\nu*\langle 0\rangle} = \Pi', B_x(t), \Pi''$ with $t \in \mathrm{Ter}$: $d_\nu := \bigvee^t_{\exists xB} d_{\nu*\langle 0\rangle}$.

§2 Cut-Elimination

Substitution

Definition

A proof system $\mathfrak{S}$ is *closed under substitution* iff it satisfies the following conditions:

(I) For each $x \in \mathsf{FV}$, $t \in \mathrm{Ter}$ and $\mathcal{I} \in \mathfrak{S}$ with $\mathrm{Eig}(\mathcal{I}) \cap (\mathrm{FV}(t) \cup \{x\}) = \emptyset$ an inference symbol $\mathcal{I}(x/t) \in \mathfrak{S}$ is defined such that

(a) $|\mathcal{I}(x/t)| = |\mathcal{I}|$ and $\mathrm{Eig}(\mathcal{I}(x/t)) = \mathrm{Eig}(\mathcal{I})$,

(b) $\Delta(\mathcal{I}(x/t)) = \Delta(\mathcal{I})_x(t)$,

(c) $(\forall i \in |\mathcal{I}|)\, \Delta_i(\mathcal{I}(x/t)) = \Delta_i(\mathcal{I})_x(t)$,

(d) $\deg(\mathcal{I}(x/t)) = \deg(\mathcal{I})$.

(II) For $\mathcal{I} \in \mathfrak{S}$ with $\mathrm{Eig}(\mathcal{I}) \neq \emptyset$ and $u \in \mathsf{FV} \setminus \mathrm{FV}(\Delta(\mathcal{I}))$ an inference symbol $\mathcal{I}^u \in \mathfrak{S}$ is defined such that

(a) $|\mathcal{I}^u| = |\mathcal{I}|$ and $\mathrm{Eig}(\mathcal{I}^u) = \{u\}$,

(b) $\Delta(\mathcal{I}^u) = \Delta(\mathcal{I})$,

(c) $(\forall i \in |\mathcal{I}|)\, \Delta_i(\mathcal{I}^u) = \Delta_i(\mathcal{I})_y(u)$, where $\mathrm{Eig}(\mathcal{I}) = \{y\}$,

(d) $\deg(\mathcal{I}^u) = \deg(\mathcal{I})$.

Definition

$\mathrm{LogAx}_A(x/t) := \mathrm{LogAx}_{A_x(t)}$, $\bigwedge_A(x/t) := \bigwedge_{A_x(t)}$, $\bigvee_A^k(x/t) := \bigvee^k_{A_x(t)}$, $\mathsf{Cut}_C(x/t) := \mathsf{Cut}_{C_x(t)}$,

$\bigvee^s_{\exists zA}(x/t) := \bigvee^{s_x(t)}_{(\exists zA)_x(t)}$, $\bigwedge^y_{\forall zA}(x/t) := \bigwedge^y_{(\forall zA)_x(t)}$. $(\bigwedge^y_{\forall zA})^u := \bigwedge^u_{\forall zA}$

Lemma 2.1

With the above definitions PL1 is closed under substitution.

Proof:

ad (II): If $\mathrm{Eig}(\mathcal{I}) = \{y\}$, then $\mathcal{I} = \bigwedge^y_{\forall zA}$ and $\mathcal{I}^u = \bigwedge^u_{\forall zA}$. Hence $|\mathcal{I}^u| = |\mathcal{I}|$, $\mathrm{Eig}(\mathcal{I}^u) = \{u\}$, $\Delta(\mathcal{I}^u) = \Delta(\mathcal{I})$, and $\Delta_0(\mathcal{I}^u) = \{A_z(u)\} \stackrel{(*)}{=} \{A_z(y)_y(u)\} = \Delta_0(\mathcal{I})_y(u)$. (*) $y \notin \mathrm{FV}(\forall zA)$.

ad (I):

1. $\mathcal{I} = \bigwedge^y_{\forall zA}$: Then $y \notin \mathrm{FV}(\forall zA) \cup \{x\}$.

One easily verifies (*) $(\forall zA)_x(t) = \forall zA'$ with $A'_z(y) = A_z(y)_x(t)$.

By (*) we have $\mathcal{I}(x/t) = \bigwedge^y_{\forall zA'}$, and $y \notin \mathrm{FV}(\forall zA')$, since $\forall zA' = (\forall zA)_x(t)$ and $y \notin \mathrm{FV}(\forall zA) \cup \mathrm{FV}(t)$. Hence $\mathcal{I}(x/t)$ is an inference symbol of PL1.

(a),(b),(d) are trivial.

(c) $\Delta_0(\mathcal{I}(x/t)) = \{A'_z(y)\} \stackrel{(*)}{=} \{A_z(y)_x(t)\} = \Delta_0(\mathcal{I})_x(t)$

[Proof of (*): If $x \notin \mathrm{FV}(\forall zA)$ then $(\forall zA)_x(t) = \forall zA$ and $A_z(y) = A_z(y)_x(t)$ (note that $x \neq y$). If $x \in \mathrm{FV}(\forall zA)$ then $(\forall zA)_x(t) = \forall zA_x(t)$ and $A_x(t)_z(y) = A_z(y)_x(t)$, since $z \notin \mathrm{FV}(t)$ and $x \neq y$.]

2. $\mathcal{I} = \bigvee^s_{\exists zA}$: One easily verifies (*) $(\exists zA)_x(t) = \exists zA'$ with $A'_z(s_x(t)) = A_z(s)_x(t)$.

Hence $\mathcal{I}(x/t) = \bigvee^{s_x(t)}_{\exists zA'}$ is an inference symbol of PL1.

(a),(b),(d) are trivial.

(c) $\Delta_0(\mathcal{I}(x/t)) = \{A'_z(s_x(t))\} \stackrel{(*)}{=} \{A_z(s)_x(t)\} = \Delta_0(\mathcal{I})_x(t)$.

[Proof of (*): If $x \notin \mathrm{FV}(\exists zA)$ then $(\exists zA)_x(t) = \exists zA$ and $A_z(s_x(t)) = A_z(s)_x(t)$. If $x \in \mathrm{FV}(\exists zA)$ then $(\exists zA)_x(t) = \exists z A_x(t)$ and $A_x(t)_z(s_x(t)) = A_z(s)_x(t)$, since $z \notin \mathrm{FV}(t) \cup \{x\}$.]

3. Otherwise: trivial.

Now let $\mathfrak{S}$ be a finitary proof system which is closed under substitution and extends PL1.

Let $m_0 \le \omega$ such that $\deg(\mathcal{I}) \le m_0$ and $\mathrm{rk}(A) < m_0$ for all $\mathcal{I} \in \mathfrak{S} \setminus \mathrm{PL1}$ and $A \in \Delta(\mathcal{I})$.

To simplify the writing we assume in addition that $|\mathcal{I}| = \{0\}$ whenever $\mathrm{Eig}(\mathcal{I}) \neq \emptyset$.

For the rest of this section d, d_i denote $\mathfrak{S}$-derivations. $\mathrm{o}(d) := lg(d)$.

Definition of $d(x/t)$ for $x \in \mathsf{FV}$, $t \in \mathrm{Ter}$

For $d = \mathcal{I}d_0...d_{n-1}$ we set $d(x/t) := \begin{cases} d & \text{if } \mathrm{Eig}(\mathcal{I}) = \{x\} \\ \mathcal{I}^u(x/t)\, d_0(y/u)(x/t) & \text{if } \mathrm{Eig}(\mathcal{I}) = \{y\} \text{ with } y \in \mathrm{FV}(t) \setminus \{x\} \\ \mathcal{I}(x/t) d_0(x/t) \ldots d_{n-1}(x/t) & \text{otherwise} \end{cases}$

where u is the first variable in $\mathsf{FV} \setminus (\mathrm{FV}(\Gamma(d)) \cup \mathrm{FV}(t) \cup \{x\})$,

Lemma 2.2 $d \vdash^k_m \Gamma \;\Rightarrow\; d(x/t) \vdash^k_m \Gamma_x(t)$.

Proof:

Let x, t be fixed, and $d = \mathcal{I}d_0...d_{n-1}$. W.l.o.g. $\Gamma(d) = \Gamma$. – We write U' for $U(x/t)$.

1. $\mathrm{Eig}(\mathcal{I}) = \{x\}$: $d' = d$ and $\Gamma = \Gamma_x(t)$, since $x \notin \mathrm{FV}(\Gamma)$.

2. $\mathrm{Eig}(\mathcal{I}) = \{y\}$ with $y \in \mathrm{FV}(t) \setminus \{x\}$: Let $e := \mathcal{I}^u d_0(y/u)$.
We have $d_0 \vdash^{k-1}_m \Gamma, \Delta_0(\mathcal{I})$, and therefore by IH $d_0(y/u) \vdash^{k-1}_m \Gamma, \Delta_0(\mathcal{I})_y(u)$.
Since $\Delta_0(\mathcal{I})_y(u) = \Delta_0(\mathcal{I}^u)$ and $u \notin \mathrm{FV}(\Gamma)$, we obtain $e \vdash^k_m \Gamma$.

Since $\mathrm{Eig}(\mathcal{I}^u) = \{u\}$ and $u \notin \mathrm{FV}(t) \cup \{x\}$, we have $e' = (\mathcal{I}^u)' d_0(y/u)' = d'$, and by case 3. below $e' \vdash^k_m \Gamma'$.

3. Otherwise:

For each $i < n$ we have $d_i \vdash^{k-1}_m \Gamma, \Delta_i(\mathcal{I})$ and therefore by IH $d_i' \vdash^{k-1}_m \Gamma', \Delta_i(\mathcal{I})'$, i.e. $d_i' \vdash^{k-1}_m \Gamma', \Delta_i(\mathcal{I}')$.

Further $\Delta(\mathcal{I}') = \Delta(\mathcal{I})' \subseteq \Gamma'$ and $\mathrm{Eig}(\mathcal{I}') \cap \Gamma' = \emptyset$ (the latter follows from $\mathrm{Eig}(\mathcal{I}) \cap \mathrm{FV}(\Gamma) = \emptyset$ & $\mathrm{Eig}(\mathcal{I}) \cap \mathrm{FV}(t) = \emptyset$ & $\mathrm{Eig}(\mathcal{I}') = \mathrm{Eig}(\mathcal{I})$). Now we get $d' = \mathcal{I}' d_0' ... d_{n-1}' \vdash^{k-1}_m \Gamma'$.

Remark

Given a finite set of variables V and a derivation $d = \mathcal{I}d_0...d_{n-1}$ with $\mathfrak{S} \ni d \vdash^k_m \Gamma$ we may from now on assume without loss of generality (w.l.o.g.) that $\mathrm{Eig}(\mathcal{I}) \cap (\mathrm{FV}(\Gamma) \cup V) = \emptyset$. Because if $\mathrm{Eig}(\mathcal{I}) = \{y\} \subseteq \mathrm{FV}(\Gamma) \cup V$ we may choose some $u \in \mathsf{FV} \setminus (V \cup \mathrm{FV}(\Gamma))$ and consider the modified derivation $d' := \mathcal{I}^u d_0(y/u)$. According to Lemma 2.2 $\mathfrak{S} \ni d' \vdash^k_m \Gamma$.

Lemma 2.3

Assume $\mathfrak{S} \ni d \vdash^k_m \Gamma, C$ with $\mathrm{rk}(C) \ge m_0$.

a) If $C = \forall x A$, then $\mathfrak{S} \vdash^k_m \Gamma, A_x(t)$ for each $t \in \mathrm{Ter}$.

b) If $C = A_0 \wedge A_1$, then $\mathfrak{S} \vdash^k_m \Gamma, A_j$ for $j \in \{0, 1\}$.

c) If $C = A_0 \vee A_1$, then $\mathfrak{S} \vdash^k_m \Gamma, A_0, A_1$.

Proof of a): Let $d = \mathcal{I}d_0...d_{n-1}$, and (w.l.o.g.) $\Gamma(d) = \Gamma, C$.

1. $\forall xA \in \Delta(\mathcal{I})$: Since $\text{rk}(\forall xA) \geq m_0$, $\mathcal{I}$ has to be a PL1-inference. Hence $\mathcal{I} = \bigwedge^y_{\forall xA}$, $y \notin \text{FV}(\Gamma, \forall xA)$ and $d_0 \vdash^{k-1}_m \Gamma, \forall xA, A_x(y)$. By Lemma 2.2 we get $d_0(y/t) \vdash^{k-1}_m \Gamma, \forall xA, A_x(t)$. Now we apply the IH and obtain $\mathfrak{S} \vdash^{k-1}_m \Gamma, A_x(t)$.

2. $\forall xA \notin \Delta(\mathcal{I})$: Then $\Delta(\mathcal{I}) \subseteq \Gamma$ and $d_i \vdash^{k-1}_m \Gamma, \forall xA, \Delta_i(\mathcal{I})$ for $i < n$. By IH we get $d'_i \vdash^{k-1}_m \Gamma, A_x(t), \Delta_i(\mathcal{I})$ for $i < n$. W.l.o.g. $\text{Eig}(\mathcal{I}) \cap \text{FV}(\Gamma, A_x(t)) = \emptyset$. Hence $\mathcal{I}d'_0...d'_{n-1} \vdash^k_m \Gamma, A_x(t)$.

Cut-elimination

Theorem 2.4

$\mathfrak{S} \vdash^k_m \Gamma, C$ & $\mathfrak{S} \vdash^l_m \Gamma, \neg C$ & $m_0 \leq \text{rk}(C) \leq m \implies \mathfrak{S} \vdash^{k+l}_m \Gamma$.

Proof by induction on $k + l$:

Assume $d \vdash^k_m \Gamma, C$ and $e \vdash^l_m \Gamma, \neg C$.

1. C is not a main formula of $\mathcal{I} := \mathsf{last}(d)$:

Then $\Delta(\mathcal{I}) \subseteq \Gamma$ and w.l.o.g. we may assume that $\text{Eig}(\mathcal{I}) \cap \text{FV}(\Gamma) = \emptyset$.

For each $i \in |\mathcal{I}|$ we have $\vdash^{k_i}_m \Gamma, C, \Delta_i(\mathcal{I})$ with $k_i < k$. By IH we get $\vdash^{k_i+l}_m \Gamma, \Delta_i(\mathcal{I})$ for each $i \in |\mathcal{I}|$.

Hence $\vdash^{k+l}_m \Gamma$, since $\Delta(\mathcal{I}) \subseteq \Gamma$ and $\text{Eig}(\mathcal{I}) \cap \text{FV}(\Gamma) = \emptyset$.

1'. $\neg C$ is not main formula of $\mathsf{last}(e)$: symmetric to 1.

2. C is main formula of $\mathsf{last}(d)$, and $\neg C$ is main formula of $\mathsf{last}(e)$:

From $m_0 \leq \text{rk}(C) = \text{rk}(\neg C)$ it follows that $\mathsf{last}(d)$ and $\mathsf{last}(e)$ are inference symbols of PL1.

2.1. C is a literal: Then $\{C, \neg C\} = \Delta(\mathsf{last}(d)) \subseteq \Gamma \cup \{C\}$ and $\{C, \neg C\} = \Delta(\mathsf{last}(e)) \subseteq \Gamma \cup \{\neg C\}$. Hence $\{C, \neg C\} \subseteq \Gamma$, and therefore $\vdash^{k+l}_m \Gamma$.

2.2. $C = \exists xA$: Then $\neg C = \forall xA$, $\mathsf{last}(d) = \bigvee^t_C$, and $\mathsf{last}(e) = \bigwedge^y_{\neg C}$.

Hence $d_0 \vdash^{k_0}_m \Gamma, C, A_x(t)$, $e_0 \vdash^{l_0}_m \Gamma, \neg C, \neg A_x(y)$ with $k_0 < k$, $l_0 < l$.

W.l.o.g. $y \notin \text{FV}(\Gamma, \neg C)$. Therefore by Lemma 2.2 we get $e_0(y/t) \vdash^{l_0}_m \Gamma, \neg C, \neg A_x(t)$.

Now the IH yields $\vdash^{k_0+l}_m \Gamma, A_x(t)$ and $\vdash^{k+l_0}_m \Gamma, \neg A_x(t)$.

Further we have $\text{rk}(A_x(t)) < \text{rk}(C) \leq m$, and therefore $\vdash^{k+l}_m \Gamma$ by $(\mathsf{Cut}_{A_x(t)})$.

2.2'. $C = \forall xA$ or $A_0 \wedge A_1$ or $A_0 \vee A_1$: analogous to 2.2.

Theorem 2.5

$\mathfrak{S} \vdash^k_{m+1} \Gamma$ & $m_0 \leq m \implies \mathfrak{S} \vdash^{2^k}_m \Gamma$.

Proof:

1. $d = \mathsf{Cut}_C d_0 d_1$ with $m_0 \leq \text{rk}(C)$:

Then $\text{rk}(C) \leq m$ and $\vdash^{k_0}_{m+1} \Gamma, C$, $\vdash^{k_1}_{m+1} \Gamma, \neg C$ with $k_0, k_1 < k$.

By IH we get $\vdash^{2^{k_0}}_m \Gamma, C$ and $\vdash^{2^{k_1}}_m \Gamma, \neg C$. Hence by Theorem 2.4 $\vdash^{2^k}_m \Gamma$, since $2^{k_0} + 2^{k_1} \leq 2^k$.

2. otherwise: Then $d = \mathcal{I}d_0...d_{n-1}$ with $\deg(\mathcal{I}) \leq m_0 \leq m$, $\text{Eig}(\mathcal{I}) \cap \text{FV}(\Gamma) = \emptyset$, and $d_i \vdash^{k_i}_{m+1} \Gamma, \Delta_i(\mathcal{I})$ with $k_i < k$ for all $i \in |\mathcal{I}|$. By IH we get $\vdash^{2^{k_i}}_m \Gamma, \Delta_i(\mathcal{I})$ $(\forall \iota \in |\mathcal{I}|)$, and then $\vdash^{2^k}_m \Gamma$.

§3 Ordinalzahlen

In diesem Abschnitt bezeichnen wir mit A, B, C stets Klassen.

Die Struktur $(On, <)$ der Ordinalzahlen ist bis auf Isomorphie eindeutig bestimmt durch folgende Eigenschaften:

(O1) $<$ ist eine lineare Ordnung auf On.

(O2) Jede nichtleere Klasse $C \subseteq On$ besitzt ein kleinstes Element.

(O3) Für jedes $\alpha \in On$ ist $\{\xi \in On : \xi < \alpha\}$ eine Menge.

(O4) Zu jeder Menge $A \subseteq On$ gibt es $\gamma \in On$ mit $\forall \alpha {\in} A (\alpha < \gamma)$.
(On ist also keine Menge, sondern eine echte Klasse).

Üblicherweise definiert man die Ordinalzahlen derart, daß stets $\alpha = \{\xi \in On : \xi \in \alpha\}$. Davon wollen wir hier auch ausgehen, da es an manchen Stellen die Notation vereinfacht.

$\alpha, \beta, \gamma, \delta, \xi, \eta, \zeta$ bezeichnen im folgenden stets Ordinalzahlen.
Die kleinste Ordinalzahl wird mit 0 bezeichnet. Entsprechend obiger Annahme ist $0 = \emptyset$.

V sei die Allklasse (Klasse aller Mengen).

Satz 3.1(Transfinite Induktion)
$\forall \alpha [\, \forall \xi {<} \alpha\, \phi(\xi) \ \to \phi(\alpha)\,] \to \forall \alpha\, \phi(\alpha)$.

Beweis:
Wir beweisen die Kontraposition der Behauptung. Gelte also $\neg \forall \alpha \phi(\alpha)$. Dann ist die Klasse $C := \{\alpha : \neg\phi(\alpha)\}$ nichtleer und besitzt daher ein kleinstes Element α_0. Somit gilt $\forall \xi {<} \alpha_0 \phi(\xi) \wedge \neg\phi(\alpha_0)$.

Satz 3.2 (Transfinite Rekursion)
Zu $G : A \times On \times V \to B$ gibt es genau ein $F : A \times On \to B$ mit
$F(x, \alpha) = G(x, \alpha, F_x | \alpha)$ für alle $(x, \alpha) \in A \times On$. (Dabei sei $F_x := \lambda y.F(x, y)$.)

Definition
$\alpha{+}1 := \min\{\xi : \alpha < \xi\}$.
α heißt Limeszahl, falls $\alpha \neq 0$ und $\neg\exists\xi(\alpha = \xi{+}1)$.
Lim bezeichnet die Klasse aller Limeszahlen und ω die kleinste Limeszahl.
(Die Existenz von Limeszahlen folgt aus dem sog. Unendlichkeitsaxiom der Mengenlehre.)
Die Ordinalzahlen $< \omega$ nennt man *endliche Ordinalzahlen* oder *natürliche Zahlen*.
Die Buchstaben i, j, k, l, m, n bezeichnen im folgenden stets natürliche Zahlen.

Definition
Für jede Menge $A \subseteq On$ sei $\sup(A) := \min\{\gamma : \forall \xi \in A(\xi \leq \gamma)\}$ und
$\sup^+(A) := \min\{\gamma : \forall \xi \in A(\xi < \gamma)\}$ $(= \sup\{\xi{+}1 : \xi \in A\})$.

Lemma 3.3
Für jede nichtleere Menge $A \subseteq On$ gilt: $\sup(A) \notin A \ \Rightarrow\ \sup(A) \in Lim$.

Beweis:

Sei $\alpha := \sup(A) \notin A$. Offenbar ist dann $\alpha \neq 0$. Bleibt zu zeigen $\forall \beta < \alpha(\beta + 1 < \alpha)$.

Sei also $\beta < \alpha$. Dann existiert ein $\xi \in A$ mit $\beta < \xi$, und wegen $\alpha \notin A$ gilt $\beta + 1 \leq \xi < \alpha$.

Definition

Eine binäre Relation $\prec$ heißt *fundiert*, wenn jede nichtleere Menge ein bzgl. $\prec$ minimales Element besitzt.

Lemma 3.4

a) Ist $\prec$ fundiert, so gibt es keine unendliche Folge $(a_n)_{n<\omega}$ mit $a_{n+1} \prec a_n$ für alle n.

(Mit Hilfe des Auswahlaxioms läßt sich auch die Gegenrichtung beweisen.)

b) Ist $\prec \subseteq A \times A$ und $F : A \to On$ mit $\forall x, y(y \prec x \Rightarrow F(y) < F(x))$ so ist $\prec$ fundiert.

Beweis:

a) Sei die Folge $(a_n)_{n<\omega}$ gegeben. Nach Voraussetzung besitzt die Menge $A := \{a_n : n < \omega\}$ ein minimales Element a_m. Dann gilt $\forall x \in A(x \not\prec a_m)$, insbesondere $a_{m+1} \not\prec a_m$.

b) Sei X eine nichtleere Menge. Ist $X \cup A = \emptyset$, so ist jedes Element von X minimal bzgl. $\prec$. Ist $X \cup A \neq \emptyset$, so existiert $\beta := \min\{F(x) : x \in X \cup A\}$. Sei $x_0 \in X \cap A$ mit $F(x_0) = \beta$. Dann ist x_0 minimales Element von X bzgl. $\prec$.

Definition

Sei A eine Menge und $\prec$ eine binäre Relation auf A.

Für jede Menge $X \subseteq A$ sei $\Phi_\prec(X) := \{x \in A : \forall y \prec x(y \in X)\}$.

(Im folgenden wird lediglich die Monotonie von $\mathcal{I}$ bzgl. $\subseteq$ ausgenützt.)

Durch transfinite Rekursion definieren wir: $\Phi_\prec^\alpha := \Phi_\prec(\bigcup_{\xi<\alpha} \Phi_\prec^\xi)$.

Ferner sei $\Phi_\prec^\infty := \bigcup_{\alpha \in On} \Phi_\prec^\alpha$ und $\mathrm{Acc}(A, \prec) := \bigcap\{X \subseteq A : \Phi_\prec(X) \subseteq X\}$.

Lemma 3.5

a) $\alpha < \beta \Rightarrow \Phi_\prec^\alpha \subseteq \Phi_\prec^\beta$.

b) $\Phi_\prec^{\alpha+1} = \Phi_\prec(\Phi_\prec^\alpha)$.

c) $\Phi_\prec(\Phi_\prec^\infty) = \Phi_\prec^\infty$.

d) $\Phi_\prec^\infty = \mathrm{Acc}(A, \prec)$.

Beweis:

a),b) klar.

c) "$\subseteq$": $x \in \Phi_\prec(\Phi_\prec^\infty) \Rightarrow \forall y \prec x \exists \xi(y \in \Phi_\prec^\xi) \Rightarrow \exists \alpha \forall y \prec x \exists \xi < \alpha(y \in \Phi_\prec^\xi) \Rightarrow \exists \alpha(x \in \Phi_\prec(\bigcup_{\xi<\alpha} \Phi_\prec^\xi) = \Phi_\prec^\alpha)$.

"$\supseteq$": $\Phi_\prec^\alpha \subseteq \Phi_\prec^{\alpha+1} = \Phi_\prec(\Phi_\prec^\alpha) \subseteq \Phi_\prec(\Phi_\prec^\infty)$.

d) Aus c) folgt $\mathrm{Acc}(A, \prec) \subseteq \Phi_\prec^\infty$. –

Andererseits erhält man durch transfinite Induktion $\forall \alpha(\Phi_\prec^\alpha \subseteq X)$ für $X \subseteq A$ mit $\Phi_\prec(X) \subseteq X$.

Definition

Sei $\prec$ eine binäre Relation auf der Menge A.

Für $x \in \mathrm{Acc}(A, \prec)$ definieren wir $|x|_\prec := \min\{\alpha : x \in \Phi_\prec^{\alpha+1}\}$.

Offenbar gilt: $(*)$ $x \in \mathrm{Acc}(A, \prec) \Rightarrow |x|_\prec = \sup\{|y|_\prec + 1 : y \prec x\}$.

[Beweis: Sei $x \in \mathrm{Acc}$ und $\alpha := |x|_{\prec}$, also $x \in \Phi_{\prec}^{\alpha+1} = \Phi_{\prec}(\bigcup_{\xi<\alpha} \Phi_{\prec}^{\xi})$). Es folgt $\forall y \prec x(y \in \bigcup_{\xi<\alpha} \Phi_{\prec}^{\xi})$, also $\forall y \prec x(|y|_{\prec} < \alpha)$ und somit $\beta := \sup\{|y|_{\prec} + 1 : y \prec x\} \leq \alpha$. Umgekehrt gilt $\forall y \prec x(|y|_{\prec} < \beta)$, also $\forall y \prec x(y \in \Phi_{\prec}^{\beta})$ und folglich $x \in \Phi_{\prec}^{\beta+1}$, d.h. $\alpha \leq \beta$.]

Satz 3.6

Sei $\prec$ eine binäre Relation auf der Menge A.

a) $\prec$ eingeschränkt auf $\mathrm{Acc}(A, \prec)$ ist fundiert.

b) $\mathrm{Acc}(A, \prec) = A \iff \prec$ fundiert.

c) Ist $\prec$ fundiert, so ist $|x|_{\prec} \in On$ für alle $x \in A$ definiert und es gilt $\forall x, y \in A(y \prec x \Rightarrow |y|_{\prec} < |x|_{\prec})$.

Beweis:

a) folgt aus $(*)$ und Lemma 3.4b.

b) "$\Rightarrow$" folgt aus a).

"$\Rightarrow$": Sei $A^- := A \setminus \mathrm{Acc}(A, \prec)$. Wäre $A^- \neq \emptyset$, so besäße A^- ein minimales Element a; wir hätten also $a \in A \wedge \forall x \prec a(x \in \mathrm{Acc}(A, \prec) \wedge a \notin \mathrm{Acc}(A, \prec)$. Widerspruch.

c) folgt aus b) und $(*)$.

Definition

Eine Funktion $F : On \to On$ heißt *ordnungstreu*, wenn gilt $\forall \alpha, \beta(\beta < \alpha \Rightarrow F(\beta) < F(\alpha))$.

Lemma 3.7

Ist $F : On \to On$ ordnungstreu, so $\alpha \leq F(\alpha)$ für alle α.

Beweis durch transfinite Induktion:

Nach I.V. haben wir $\forall \xi{<}\alpha(\xi \leq F(\xi) < F(\alpha))$, also $(F(\alpha) < \alpha \Rightarrow F(\alpha) < F(\alpha))$ und somit $\alpha < F(\alpha)$.

Definition

$F : On \to On$ heißt Ordnungsfunktion der Klasse $A \subseteq On$, falls F ordnungstreu ist und $\mathrm{ran}(F) = A$ gilt.

Satz 3.8

Jede unbeschränkten Klasse $A \subseteq On$ besitzt genau eine Ordnungsfunktion F,
und zwar gilt $F(\alpha) = \min\{\beta \in A : \forall \xi{<}\alpha(F(\xi) < \beta)\}$.

Beweis:

Existenz:

Durch transfinite Rekursion definieren wir $F : On \to On$ mit $F(\alpha) := \min\{\beta \in A : \forall \xi{<}\alpha(F(\xi) < \beta)\}$.

1. $F(\alpha)$ ist für jedes α definiert, denn $\{F(\xi) : \xi < \alpha\}$ ist Menge und deshalb gibt es $\beta \in A$ mit $\forall \xi{<}\alpha(F(\xi) < \beta)$.

2. F ist ordnungstreu und $\mathrm{ran}(F) \subseteq A$: klar nach Def.

3. $A \subseteq \mathrm{ran}(F)$:

Sei $\gamma \in A$. Da F injektiv ist, ist $\{\xi : F(\xi) < \gamma\}$ eine Menge und es existiert $\alpha := \min\{\xi : \gamma \leq F(\xi)\}$. Nach Definition von α gilt nun $\gamma \leq F(\alpha) \wedge \forall \xi < \alpha(F(\xi) < \gamma)$. Folglich $F(\alpha) = \min\{\beta \in A : \forall \xi < \alpha(F(\xi) < \beta)\} \leq \gamma \leq F(\alpha)$.

Eindeutigkeit: Sei $G : On \to A$ ordnungstreu mit $\mathrm{ran}(G) = A$. Durch transfinite Induktion nach α zeigen wir $G(\alpha) = F(\alpha)$. Gelte schon $G(\xi) = F(\xi)$ für alle $\xi < \alpha$. Wegen $F(\alpha) \in A = \mathrm{ran}(G)$ gibt es ein β mit $G(\beta) = F(\alpha)$. Da F ordnungstreu ist, muß $\alpha \leq \beta$ sein. Da G ordnungstreu ist, folgt sogar $\alpha = \beta$, denn andernfalls wäre $F(\alpha)$ nicht in $\mathrm{ran}(G)$.

Vereinbarung: λ bezeichne im folgenden stets Limeszahlen.

Definition

1. Eine Funktion $F : On \to On$ heißt *stetig*, falls gilt $\forall\lambda(F(\lambda) = \sup_{\xi<\lambda} F(\xi))$.
2. $F : On \to On$ heißt *Normalfunktion*, falls F ordnungstreu und stetig ist.

Lemma 3.9

Ist $F : On \to On$ stetig mit $\forall\alpha(F(\alpha) < F(\alpha + 1))$, so ist F eine Normalfunktion.

Beweis: Durch transfinite Induktion nach α zeigt man $\forall\beta < \alpha(F(\beta) < F(\alpha))$.

Lemma 3.10

Für jede Normalfunktion $F : On \to On$ gilt:

a) $F(\alpha) = \sup\{F(\xi + 1) : \xi \in \alpha\}$, für alle $\alpha > 0$.

b) $\lambda \in Lim \Rightarrow F(\lambda) \in Lim$.

c) $\forall\gamma \geq F(0)\exists!\alpha(F(\alpha) \leq \gamma < F(\alpha + 1))$.

d) G Normalfunktion $\Rightarrow F \circ G$ Normalfunktion.

e) $F(\sup(A)) = \sup(F[A])$ $(= \sup_{\xi\in A} F(\xi))$ für jede nichtleere Menge $A \subseteq On$.

Beweis:

a) Wegen $\forall\xi < \alpha(\xi + 1 \leq \alpha)$ gilt $\gamma := \sup\{F(\xi + 1) : \xi < \alpha\} \leq F(\alpha)$. Ist $\alpha = \beta + 1$, so $F(\alpha) \in \{F(\xi + 1) : \xi < \alpha\}$ und deshalb $F(\alpha) \leq \gamma$. Ist $\alpha \in Lim$, so $F(\alpha) = \sup_{\xi<\alpha} F(\xi) \leq \sup_{\xi<\alpha} F(\xi + 1) = \gamma$.

b) Es ist $0 \leq F(0) < F(\lambda)$. Aus $\gamma < F(\lambda)$ folgt $\exists\xi < \lambda(\gamma < F(\xi))$ und weiter $\exists\xi(\gamma + 1 \leq F(\xi) < F(\lambda))$.

c) Sei $\gamma \geq F(0)$. Wegen $\gamma \leq F(\gamma) < F(\gamma + 1)$ existiert $\alpha := \min\{\xi : \gamma < F(\xi + 1)\}$. Dann $\gamma < F(\alpha + 1)$. Ist $\alpha = 0$, so $F(\alpha) = F(0) \leq \gamma$. Ist $\alpha > 0$, so $F(\alpha) = \sup_{\xi<\alpha} F(\xi+1)$ und $\forall\xi < \alpha(F(\xi+1) \leq \gamma)$, also $F(\alpha) \leq \gamma$.

d) $(F \circ G)(\lambda) = F(G(\lambda)) = F(\sup(\{G(\xi) : \xi < \lambda\})) = \sup_{\xi<\lambda} F(G(\xi))$.

e) Sei $\emptyset \neq A \subseteq On$ und $\alpha := \sup(A)$. Ist $\alpha \in A$, so $F(\alpha) = \sup(F[A])$, da F ordnungstreu. Ist $\alpha \notin A$, so $\alpha \in Lim$ und deshalb $F(\alpha) = \sup_{\xi<\alpha} F(\xi)$. Ferner gilt $\forall\xi \in A(\xi < \alpha) \wedge \forall\xi < \alpha\exists\eta{\in}A(\xi < \eta)$, woraus $\sup_{\xi<\alpha} F(\xi) = \sup(F[A])$ folgt.

Definition von $\alpha + \beta$ durch transfinite Rekursion nach β

$\alpha + 0 := \alpha, \quad \alpha + (\beta + 1) := (\alpha + \beta) + 1, \quad \alpha + \lambda := \sup_{\eta<\lambda}(\alpha + \eta)$.

Abkürzung: $1 := 0 + 1$.

Bemerkung: $\alpha + 1$ hat nun zwei Bedeutungen, die aber übereinstimmen, nämlich

1. $\alpha + 1 = \min\{\xi : \alpha < \xi\}$,
2. $\alpha + 1 =$ Summe von α und 1 gemäß obiger Definition.

Lemma 3.11

a) Für jedes α ist $\beta \mapsto \alpha + \beta$ eine Normalfunktion.

b) $\beta_0 < \beta_1 \;\Rightarrow\; \alpha + \beta_0 < \alpha + \beta_1$

c) $\alpha, \beta \leq \alpha + \beta$

d) $\forall \gamma \geq \alpha \exists! \beta (\alpha + \beta = \gamma)$

e) $\alpha_0 \leq \alpha_1 \;\Rightarrow\; \alpha_0 + \beta \leq \alpha_1 + \beta$

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

g) $\alpha, \beta < \omega \;\Rightarrow\; \alpha + \beta = \beta + \alpha < \omega$

h) $0 < k < \omega \;\Rightarrow\; k + \omega = \omega < \omega + k$

Beweis:

a) folgt aus 3.9. b),d) folgen aus a). c) folgt aus a) und 3.7. e) Induktion nach β. f) Induktion nach γ. g) Induktion nach β. h) $\omega \leq k + \omega = \sup_{n<\omega}(k+n) \leq \omega < \omega + k$.

Definition von $\alpha \cdot \beta$ durch transfinite Rekursion nach β

$\alpha \cdot 0 := 0,\; \alpha \cdot (\beta + 1) := (\alpha \cdot \beta) + \alpha,\; \alpha \cdot \lambda := \sup\{\alpha \cdot \eta : \eta < \lambda\}$.

Lemma 3.12

a) Für jedes $\alpha \geq 1$ ist $\beta \mapsto \alpha \cdot \beta$ eine Normalfunktion.

b) $\alpha_0 \leq \alpha_1 \;\Rightarrow\; \alpha_0 \cdot \beta \leq \alpha_1 \cdot \beta$

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

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

e) $0 \cdot \alpha = \alpha \cdot 0 = 0 \,\&\, 1 \cdot \alpha = \alpha \cdot 1 = \alpha$

f) $\alpha, \beta < \omega \;\Rightarrow\; \alpha \cdot \beta = \beta \cdot \alpha < \omega$

Definition von α^β durch transfinite Rekursion nach β

$\alpha^0 := 1,\; \alpha^{\beta+1} := \alpha^\beta \cdot \alpha,\; \alpha^\lambda := \sup\{\alpha^\eta : \eta < \lambda\}$.

Lemma 3.13

Für $\alpha \geq 2$ gilt:

a) $\beta \mapsto \alpha^\beta$ ist Normalfunktion.

b) $\alpha \leq \gamma \;\Rightarrow\; \alpha^\beta \leq \gamma^\beta$

c) $\alpha^\beta \cdot \alpha^\gamma = \alpha^{\beta+\gamma}$

d) $(\alpha^\beta)^\gamma = \alpha^{\beta \cdot \gamma}$

e) $\beta > \beta_0 > ... > \beta_n \,\&\, \delta_0, ..., \delta_n < \alpha \;\Rightarrow\; \alpha^\beta > \alpha^{\beta_0} \cdot \delta_0 + ... + \alpha^{\beta_n} \cdot \delta_n$.

Beweis:

e) Ind. nach n: I.V. $\Rightarrow\; \alpha^{\beta_0} > \alpha^{\beta_1} \cdot \delta_1 + ... + \alpha^{\beta_n} \cdot \delta_n \;\Rightarrow\; a^\beta \geq \alpha^{\beta_0} \cdot \alpha \geq \alpha^{\beta_0} \cdot \delta_0 + \alpha^{\beta_0} > \alpha^{\beta_0} \cdot \delta_0 + ... + \alpha^{\beta_n} \cdot \delta_n$.

Satz 3.14

a) Zu $\alpha \geq 2$ und $\gamma \geq 1$ existieren eindeutig β, δ, γ_0 mit $0 < \delta < \alpha \,\&\, \gamma_0 < \alpha^\beta \,\&\, \gamma = \alpha^\beta \cdot \delta + \gamma_0$.

b) Zu $\alpha \geq 2$ und $\gamma \geq 1$ existieren eindeutig $\beta_0 > ... > \beta_n$ und $0 < \delta_0, ..., \delta_n < \alpha$ mit $\gamma = \alpha^{\beta_0} \cdot \delta_0 + ... + \alpha^{\beta_n} \cdot \delta_n$. *(Cantorsche Normalform von γ zur Basis α)*

Beweis:

a) *Eindeutigkeit:* Sei $\gamma = \alpha^\beta \cdot \delta + \gamma_0 = \alpha^{\beta_1} \cdot \delta_1 + \gamma_1$ mit $0 < \delta, \delta_1 < \alpha \,\&\, \gamma_0 < \alpha^\beta \,\&\, \gamma_1 < \alpha^{\beta_1}$. Dann $\alpha^\beta \leq \gamma < \alpha^{\beta+1} \,\&\, \alpha^{\beta_1} \leq \gamma < \alpha^{\beta_1+1}$, also $\beta = \beta_1$. Weiter folgt nun $\alpha^\beta \cdot \delta \leq \gamma < \alpha^\beta \cdot (\delta + 1)$ und $\alpha^\beta \cdot \delta_1 \leq \gamma < \alpha^\beta \cdot (\delta_1 + 1)$, also auch $\delta = \delta_1$. Aus $\alpha^\beta \cdot \delta + \gamma_0 = \alpha^\beta \cdot \delta + \gamma_1$ folgt schließlich $\gamma_0 = \gamma_1$.

Existenz: Nach 3.10c existiert ein β mit $\alpha^\beta \leq \gamma < \alpha^{\beta+1}$, d.h. $\alpha^\beta \cdot 1 \leq \gamma < \alpha^\beta \cdot \alpha$. Wiederum mit 3.10 (im wesentlichen) folgt daraus $\alpha^\beta \cdot \delta \leq \gamma < \alpha^\beta \cdot (\delta + 1) = \alpha^\beta \cdot \delta + \alpha^\beta$. Nach 3.11 existiert deshalb ein $\gamma_0 < \alpha^\beta$ mit $\gamma = \alpha^\beta \cdot \delta + \gamma_0$.

b) folgt aus a) und 3.13e mittels Induktion nach γ.

Abkürzung

$\alpha =_{CNF} \omega^{\alpha_0} \cdot k_0 + \ldots + \omega^{\alpha_k} \cdot k_n \;:\Longleftrightarrow\; \alpha = \omega^{\alpha_0} \cdot k_0 + \ldots + \omega^{\alpha_k} \cdot k_n \,\&\, \alpha_0 > \ldots > \alpha_n$

Definition (Additive Hauptzahlen)

$\gamma \in On$ heißt *additive Hauptzahl*, falls $\gamma > 0 \wedge \forall \xi, \eta < \gamma(\xi + \eta < \gamma)$.

$H :=$ Klasse aller additiven Hauptzahlen.

Lemma 3.15

a) $\alpha \mapsto \omega^\alpha$ ist die Ordnungsfunktion der Klasse H aller additiven Hauptzahlen.

b) $\gamma \in H \;\Leftrightarrow\; \gamma > 0 \wedge \forall \xi < \gamma(\xi + \gamma = \gamma)$.

Beweis:

a) 1. Durch Induktion nach α zeigen wir $\omega^\alpha \in H$:

1.1. $\omega^0 \in H$ ist trivial.

1.2. $\xi, \eta < \omega^{\alpha+1} \Rightarrow \xi, \eta < \omega^\alpha \cdot n$ für ein $n < \omega \Rightarrow \xi + \eta < \omega^\alpha \cdot n + \omega^\alpha \cdot n = \omega^\alpha \cdot (n + n) < \omega^{\alpha+1}$.

1.3. $\xi, \eta < \omega^\lambda \Rightarrow \xi, \eta < \omega^\alpha$ für ein $\alpha < \lambda \overset{\text{I.V.}}{\Rightarrow} \xi + \eta < \omega^\alpha < \omega^\lambda$.

2. $\gamma \notin \{\omega^\alpha : \alpha \in On\} \Rightarrow \gamma \notin H$.

Beweis: Sei $1 \leq \gamma \notin \{\omega^\alpha : \alpha \in On\}$. Dann $\gamma = \omega^\beta \cdot n + \gamma_0$ mit $0 < n < \omega \,\&\, \gamma_0 < \omega^\beta$ und $1 < n \vee 0 < \gamma_0$. Für $\eta := \omega^\beta \cdot (n - 1) + \gamma_0$ gilt nun $0 < \eta < \omega^\beta \cdot n \leq \gamma$ und $\omega^\beta < \omega^\beta + \eta = \gamma$, d.h. $\gamma \notin H$.

b) 1. Sei $\gamma \in H$ und $\xi < \gamma$. Dann $\xi + \gamma = \sup\{\xi + \eta + 1 : \eta < \gamma\} \leq \gamma$, da $(\xi, \eta < \beta \Rightarrow \xi + \eta < \gamma \Rightarrow \xi + \eta + 1 \leq \gamma)$. 2. Gelte $\gamma > 0 \,\&\, \forall \xi < \gamma(\xi + \gamma = \gamma)$, und sei $\xi, \eta < \gamma$. Dann $\xi + \eta < \xi + \gamma = \gamma$.

Abkürzung $\alpha =_{NF} \alpha_0 + \ldots + \alpha_n \;:\Longleftrightarrow\; \alpha = \alpha_0 + \ldots + \alpha_n \,\&\, \alpha_0, ..., \alpha_n \in H \,\&\, \alpha_0 \geq \ldots \geq \alpha_n$.

Lemma 3.16

a) Zu jedem $\alpha > 0$ existiert genau ein Tupel $\alpha_0, ..., \alpha_n \in H$ mit $\alpha =_{NF} \alpha_0 + ... + \alpha_n$.

b) Ist $\alpha = \alpha_0 + ... + \alpha_n$ mit additiven Hauptzahlen $\alpha_0 \geq \ldots \geq \alpha_n$, so gilt für jedes $k < n$:
$\alpha_0 + ... + \alpha_k < \alpha$ und $\alpha_{k+1} + ... + \alpha_n < \alpha$.

Beweis:

a) folgt aus dem Satz über die Cantorsche-Normalform (zur Basis ω) unter Berücksichtigung der Gleichung $\omega^\beta \cdot n = \omega^\beta + ... + \omega^\beta$.

b) Die erste Ungleichung ist trivial. Nun zur zweiten Ungleichung.

Fall 1: $\alpha_{k+1} < \alpha_k$. Dann $\alpha_{k+1} + ... + \alpha_n < \alpha_k < \alpha$.

Fall 2: $\alpha_{k+1} = \alpha_k$. Dann $\alpha_{k+1} + ... + \alpha_n \leq \alpha_k + ... + \alpha_{n-1} < \alpha_k + ... + \alpha_n \leq \alpha$.

Definition (Natürliche Summe oder Hessenberg-Summe)

$\alpha\#0 := 0\#\alpha := \alpha$.

Für $\alpha =_{NF} \alpha_0 + ... + \alpha_n$ und $\beta =_{NF} \alpha_{n+1} + ... + \alpha_{n+m}$ sei $\alpha\#\beta := \alpha_{p(0)} + ... + \alpha_{p(m+n)}$,
wobei p eine Permutation von $m + n + 1$ mit $\alpha_{p(0)} \geq ... \geq \alpha_{p(m+n)}$.

Bemerkung

Ist $\alpha =_{CNF} \omega^{\gamma_0} \cdot k_0 + \ldots + \omega^{\gamma_n} \cdot k_n$ und $\beta =_{CNF} \omega^{\gamma_0} \cdot l_0 + \ldots + \omega^{\gamma_n} \cdot l_n$, so gilt
$\alpha\#\beta = \omega^{\gamma_0} \cdot (k_0+l_0) + \ldots + \omega^{\gamma_n} \cdot (k_n+l_n)$.

Lemma 3.17

a) $\alpha\#\beta = \beta\#\alpha$,

b) $(\alpha\#\beta)\#\gamma = \alpha\#(\beta\#\gamma)$,

c) $\alpha_0 \geq ... \geq \alpha_n$ additive Hauptzahlen $\Rightarrow$ $\alpha_0 + ... + \alpha_n = \alpha_0\#...\#\alpha_n$,

d) $\beta < \gamma \;\Rightarrow\; \alpha\#\beta < \alpha\#\gamma$,

e) $\alpha, \beta < \omega^\gamma \;\Rightarrow\; \alpha\#\beta < \omega^\gamma$,

f) $\alpha + \beta \leq \alpha\#\beta$.

Beweis:

a),b),e) klar.

d) Wegen b) und c) reicht es, die Behauptung für $\alpha \in H$ zu beweisen. Sei $\alpha = \omega^{\alpha_0}$, $\beta =_{CNF} \omega^{\delta_0} \cdot k_0 + \ldots + \omega^{\delta_n} \cdot k_n$ und $\gamma =_{CNF} \omega^{\delta_0} \cdot l_0 + \ldots + \omega^{\delta_n} \cdot l_n$ mit $\alpha_0 = \delta_m$ für ein $m \leq n$. Dann $\alpha\#\beta = \omega^{\delta_0} \cdot k'_0 + \ldots + \omega^{\gamma_n} \cdot k'_n$ und $\alpha\#\gamma = \omega^{\delta_0} \cdot l'_0 + \ldots + \omega^{\gamma_n} \cdot l'_n$, wobei $k'_m = k_m{+}1$, $l'_m = l_m{+}1$ und $k'_i = k_i$, $l'_i = l - i$ für $i \neq m$.

f) 1. Ist $\alpha \in H$, so $\alpha + \beta = \beta \leq \alpha\#\beta$ oder $\alpha + \beta = \alpha\#\beta$.

2. Für beliebiges α folgt die Behauptung aus 1. und der Assoziativität von #.

§4 Proof theoretic analysis of $\mathcal{Z}$ via the infinitary system Z^∞

The axiom system $\mathcal{Z}$ of pure number theory

Inductive Definition of sets PR^n of n-ary function symbols

(PR 1) $\mathbf{0}^n \in \mathrm{PR}^n$ $(n \geq 0)$, $\mathsf{S} \in \mathrm{PR}^1$, $\mathrm{I}_i^n \in \mathrm{PR}^n$ $(1 \leq i \leq n)$.

(PR 2) $h \in \mathrm{PR}^m$ & $g_1, ..., g_m \in \mathrm{PR}^n$ & $m, n \geq 1 \implies (\circ h g_1 ... g_m) \in \mathrm{PR}^n$.

(PR 3) $g \in \mathrm{PR}^n$ & $h \in \mathrm{PR}^{n+2} \implies (\mathrm{R} g h) \in \mathrm{PR}^{n+1}$.

Abbreviation: $\mathrm{PR} := \bigcup_{n \in \mathbb{N}} \mathrm{PR}^n$, $\mathbf{0} := \mathbf{0}^0$.

$\mathcal{L}_0 := \mathrm{PR} \cup \{\mathbf{=}\}$, where $\mathbf{=}$ is a binary relation symbol (equality).

The $\mathcal{L}_0$-terms $\mathbf{0}$, $\mathsf{S}\mathbf{0}$, $\mathsf{SS}\mathbf{0}, ...$ are called numerals. For $n \in \mathbb{N}$ let $\underline{n} := \overbrace{\mathsf{S}...\mathsf{S}}^{n}\mathbf{0}$.

$T :=$ set of all closed $\mathcal{L}_0$-terms.

If $t \in T$ then $\mathsf{val}(t)$ denotes its canonical *value*. Hence $\mathsf{val}(\underline{n}) = n$.

$\mathrm{TRUE}_0 :=$ set of all true closed literals of $\mathcal{L}_0$ $[= \{s\mathbf{=}t : s, t \in T \ \&\ \mathsf{val}(s) = \mathsf{val}(t)\} \cup \{\neg(s\mathbf{=}t) : \mathsf{val}(s) \neq \mathsf{val}(t)\}]$

The **language of** $\mathcal{Z}$ is $\mathcal{L}_0(\mathcal{X}) := \mathcal{L}_0 \cup \{X_0, X_1, ...\}$, where $X_0, X_1, ...$ are unary predicate symbols; we call them *set variables*. But note that they are not considered as variables in the proper sense (e.g. $\mathrm{FV}(X_i\mathbf{0}) = \emptyset$). We use X as syntactic variable for $X_0, X_1, ...$

The **axioms of** $\mathcal{Z}$ are the universal closures of the following $\mathcal{L}_0(\mathcal{X})$-formulas:

$x\mathbf{=}x$

$x\mathbf{=}y \to A_z(x) \to A_z(y)$, for each atomic $\mathcal{L}_0(\mathcal{X})$-formula A

$\neg(\mathsf{S}x\mathbf{=}\mathbf{0})$

$\mathsf{S}x\mathbf{=}\mathsf{S}y \to x\mathbf{=}y$

$\mathbf{0}^n x_1 ... x_n \mathbf{=} \mathbf{0}$

$\mathrm{I}_i^n x_1 ... x_n \mathbf{=} x_i$

$(\circ h g_1 ... g_m) x_1 ... x_n \mathbf{=} h g_1 x_1 ... x_n \ldots g_m x_1 ... x_n$

$(\mathrm{R} g h) x_1 ... x_n 0 \mathbf{=} g x_1 ... x_n$

$(\mathrm{R} g h) x_1 ... x_n \mathsf{S} y \mathbf{=} h x_1 ... x_n y (\mathrm{R} g h) x_1 ... x_n y$

$F_x(0) \to \forall x(F \to F_x(\mathsf{S}x)) \to F_x(z)$, for each $\mathcal{L}_0(\mathcal{X})$-formula F

Definition

Let R be an $\mathcal{L}_0$-formula with $\mathrm{FV}(R) = \{x, y\}$ such that the relation $\prec := \{(m,n) \in \mathbb{N}^2 : \mathbb{N} \models R_{y,x}(m,n)\}$ is wellfounded.

By recursion over $\prec$ one defines the *$\prec$-norm* $|n|_\prec$ of $n \in \mathbb{N}$:
$|n|_\prec := \sup\{|m|_\prec + 1 : m \prec n\}$.

Abbreviations:

$s \prec t := R_{y,x}(s,t)$, $\forall y{\prec}t\, F(y) := \forall y(y{\prec}t \to F(y))$,

$|t|_\prec := |\mathsf{val}(t)|_\prec$ for $t \in T$,

$\mathrm{Prog}_\prec(F) := \forall x(\forall y{\prec}x F(y) \to F(x))$,

$\mathrm{TI}_\prec(F,t) := \mathrm{Prog}_\prec(F) \to \forall x{\prec}t F(x)$.

In this section we will show that *transfinite induction up to* ε_0 is not provable in $\mathcal{Z}$, more precisely we will establish the following

Theorem $\mathcal{Z} \vdash \mathrm{TI}_\prec(X, \underline{n}) \implies |n|_\prec < \varepsilon_0$.

Sketch of the proof:

We define an infinitary proof system Z^∞, which (essentially) results from PL1 by

(i) replacing each inference symbol $\bigwedge^y_{\forall xA}$ by its infinitary version

$$(\textstyle\bigwedge_{\forall xA})\ \frac{\dots A_x(t) \dots (t{\in}T)}{\forall xA} \quad (\omega\text{-rule})$$

(ii) adding the axioms A (for $A \in \mathrm{TRUE}_0$) and $Xs, \neg Xt$ (for $s, t \in T$ with $\mathsf{val}(s) = \mathsf{val}(t)$)

Then we prove:

Lemma 4.5 $\mathrm{PL1} \vdash^k_0 \Gamma \ \&\ \mathrm{FV}(\Gamma) = \emptyset \implies \mathsf{Z}^\infty \vdash^k_0 \Gamma$.

Lemma 4.8 $A \in \mathcal{Z} \implies \mathsf{Z}^\infty \vdash^\alpha_0 A$ for some $\alpha < \omega + \omega$.

Theorem 4.2 $\mathsf{Z}^\infty \vdash^\alpha_{m+1} \Gamma \implies \mathsf{Z}^\infty \vdash^{3^\alpha}_m \Gamma$.

Lemma 4.4 $\mathsf{Z}^\infty \vdash^\beta_0 \mathrm{TI}_\prec(X, \underline{n}) \implies |n|_\prec \le 2^\beta$

From this the announced Theorem is obtained as follows:

$\mathcal{Z} \vdash \mathrm{TI}_\prec(X, \underline{n}) \overset{1.1(\mathrm{Cor.})}{\Longrightarrow} \mathrm{PL1} \vdash^k_0 \neg A_1, ..., \neg A_l, \mathrm{TI}_\prec(X, \underline{n})$ for some $k < \omega$ and $A_1, ..., A_l \in \mathcal{Z} \overset{4.5}{\Longrightarrow}$
$\mathsf{Z}^\infty \vdash^k_0 \neg A_1, ..., \neg A_l, \mathrm{TI}_\prec(X, \underline{n}) \overset{4.8}{\Longrightarrow} \mathsf{Z}^\infty \vdash^\alpha_m \mathrm{TI}_\prec(X, \underline{n})$ with $\alpha < \omega + \omega$ and $m := \max_{1 \le i \le l} \mathrm{rk}(A_i) + 1 \overset{4.2}{\Longrightarrow}$
$\mathsf{Z}^\infty \vdash^\beta_m \mathrm{TI}_\prec(X, \underline{n})$ with $\beta < \varepsilon_0 \overset{4.4}{\Longrightarrow} |n|_\prec \le 2^\beta < \varepsilon_0$.

The infinitary proof system Z^∞

The language of Z^∞ is $\mathcal{L}_0(\mathcal{X})$.

We introduce the following relation $\simeq$ between $\mathcal{L}_0(\mathcal{X})$-formulas and (possibly infinitary) conjunctions or disjunctions of $\mathcal{L}_0(\mathcal{X})$-formulas:

$A_0 \wedge A_1 \simeq \bigwedge_{i\in\{0,1\}} A_i$, $\quad \forall x A \simeq \bigwedge_{t\in T} A_x(t)$, $\quad A_0 \vee A_1 \simeq \bigvee_{i\in\{0,1\}} A_i$, $\quad \exists x A \simeq \bigvee_{t\in T} A_x(t)$

Then we have

– $A \simeq *_{\iota\in J} A_\iota \;\&\; \iota \in J \implies \mathrm{rk}(A_\iota) < \mathrm{rk}(A)$,

– $A \simeq *_{\iota\in J} A_\iota \implies \neg(A) \simeq \overline{*}_{\iota\in J} \neg(A_\iota)$, where $\overline{\bigvee} := \bigwedge$, $\overline{\bigwedge} := \bigvee$,

Definition

$\mathcal{AX}(\mathsf{Z}^\infty)$:= set of all sequents Δ such that

– all elements of Δ are literals,

– $\Delta \cap \mathrm{TRUE}_0 \neq \emptyset$ or Δ contains a subset $\{Xs, \neg Xt\}$ with $\mathsf{val}(s) = \mathsf{val}(t)$.

Remark: $\Delta', \Delta'' \in \mathcal{AX}(\mathsf{Z}^\infty) \implies (\Delta' \setminus \{C\}) \cup (\Delta'' \setminus \{\neg C\}) \in \mathcal{AX}(\mathsf{Z}^\infty)$

Z^∞-inferences

$(\mathsf{Ax}^\infty_\Delta)$ $\quad \Delta \quad$ if $\Delta \in \mathcal{AX}(\mathsf{Z}^\infty)$

$(\bigwedge_A)$ $\quad \dfrac{\ldots A_\iota \ldots (\iota\in J)}{A} \quad$ if $A \simeq \bigwedge_{\iota\in J} A_\iota$

$(\bigvee^\mu_A)$ $\quad \dfrac{A_\mu}{A} \quad$ if $A \simeq \bigvee_{\iota\in J} A_\iota$ and $\mu \in J$

(Cut_C) $\quad \dfrac{C \quad \neg C}{\emptyset}$

(Rep) $\quad \dfrac{\emptyset}{\emptyset}$

$\mathrm{o}(d) := lg(d)$, $\deg(\mathsf{Cut}_C) := \mathrm{rk}(C) + 1$ and $\deg(\mathcal{I}) := 0$ for all other inferences.

Remark: At moment we could do without Rep inferences. They will become important later.

NOTATION

We use $d, d_0, d_1, e, ...$ as syntactic variables for Z^∞-derivations.

$$d \simeq \left\{ \begin{array}{c} d_\iota \\ | \\ \dfrac{\ldots \Gamma_\iota : \alpha_\iota \ldots}{\Gamma : \alpha}\mathcal{I} \end{array} \right. \quad :\Leftrightarrow \quad d = \mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|} \;\&\; \forall\iota \in |\mathcal{I}|(\Gamma(d_\iota) \subseteq \Gamma_\iota \;\&\; \mathrm{o}(d_\iota) \le \alpha_\iota < \alpha) \;\&\; \frac{...\Gamma_\iota...(\iota\in|\mathcal{I}|)}{\Gamma}\mathcal{I},$$

where $\dfrac{...\Gamma_\iota...(\iota\in|\mathcal{I}|)}{\Gamma}\mathcal{I} :\Leftrightarrow \Delta(\mathcal{I}) \subseteq \Gamma \;\&\; \forall\iota \in |\mathcal{I}|(\Gamma_\iota \subseteq \Gamma, \Delta_\iota(\mathcal{I}))$.

Note that $d \simeq \left\{ \dfrac{\cdots\cdots}{\Gamma : \alpha}\mathcal{I} \right.$ implies $\Gamma(d) \subseteq \Gamma$ and $\mathrm{o}(d) \le \alpha$.

Theorem and Definition 4.1

For each formula C we define an operator $\mathcal{R}_C$ such that:

$d \vdash^{\alpha}_{m} \Gamma, C \ \ \& \ \ e \vdash^{\beta}_{m} \Gamma, \neg C \ \ \& \ \ \mathrm{rk}(C) \le m \implies \mathcal{R}_C(d,e) \vdash^{\alpha\#\beta}_{m} \Gamma.$

Proof: $\mathcal{R}_C(d,e)$ is defined by recursion on $\alpha\#\beta$.

1. C is not a main formula of $\mathcal{I} := \mathsf{last}(d_0)$: Then $d \simeq \left\{ \dfrac{\begin{array}{c} d_\iota \\ | \\ \ldots\Gamma, C, \Delta_\iota : \alpha_\iota \ldots (\iota \in I) \end{array}}{\Gamma, C : \alpha}\,\mathcal{I} \right.$.

By IH we get $\mathcal{R}_C(d_\iota, e) \vdash^{\alpha_\iota\#\beta}_{m} \Gamma_\iota$ for all $\iota \in I$. Further we have $\alpha_\iota\#\beta < \alpha\#\beta$ for all $\iota \in I$.

Hence $\mathcal{R}_C(d,e) := \mathcal{I}(\mathcal{R}_C(d_\iota,e))_{\iota\in I} \simeq \left\{ \dfrac{\begin{array}{c} \mathcal{R}_C(d_\iota,e) \\ | \\ \ldots\Gamma, \Delta_\iota : \alpha_\iota\#\beta \ldots (\iota \in I) \end{array}}{\Gamma : \alpha\#\beta}\,\mathcal{I} \right.$ is a derivation as required.

1'. $\neg C$ is not main formula of $\mathsf{last}(e)$: symmetric to 1.

2. C is main formula of $\mathsf{last}(d)$, and $\neg C$ is main formula of $\mathsf{last}(e)$:

2.1. C is a literal: Then $\mathsf{last}(d) = \mathsf{Ax}^{\infty}_{\Delta'}$ and $\mathsf{last}(e) = \mathsf{Ax}^{\infty}_{\Delta''}$ with $\Delta' \in \mathcal{AX}(\mathsf{Z}^\infty)$ and $\Delta'' \in \mathcal{AX}(\mathsf{Z}^\infty)$. Hence $\Delta := (\Delta' \setminus \{C\}) \cup (\Delta'' \setminus \{\neg C\}) \subseteq \Gamma$, and $\Delta \in \mathcal{AX}(\mathsf{Z}^\infty)$ (by ....). We set $\mathcal{R}_C(d,e) := \mathsf{Ax}^{\infty}_{\Delta}$.

2.2. $C \simeq \bigvee_{\iota\in J} C_\iota$:

Then $\neg C \simeq \bigwedge_{\iota \in J} \neg C_\iota$ and $d \simeq \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma, C, C_\mu : \alpha_0 \end{array}}{\Gamma, C : \alpha}\,\bigvee^{\mu}_{C} \right.$, $e \simeq \left\{ \dfrac{\begin{array}{c} e_\iota \\ | \\ \ldots\Gamma, \neg C, \neg C_\iota : \beta_\iota \ldots (\iota \in J) \end{array}}{\Gamma, \neg C : \beta}\,\bigwedge_{\neg C} \right.$.

By IH we get $\mathcal{R}_C(d_0, e) \vdash^{\alpha_0\#\beta}_{m} \Gamma, C_\mu$ and $\mathcal{R}_C(d, e_\mu) \vdash^{\alpha\#\beta_\mu}_{m} \Gamma, \neg C_\mu$.

Further $\mathrm{rk}(C_\mu) < \mathrm{rk}(C) \le m$.

Hence $\mathcal{R}_C(d,e) := \mathsf{Cut}_{C_\mu}\mathcal{R}_C(d_0,e)\mathcal{R}_C(d,e_\mu) \simeq \left\{ \dfrac{\begin{array}{cc} \mathcal{R}_C(d_0,e) & \mathcal{R}_C(d,e_\mu) \\ | & | \\ \Gamma, C_\mu : \alpha_0\#\beta & \Gamma, \neg C_\mu : \alpha\#\beta_\mu \end{array}}{\Gamma : \alpha\#\beta}\,\mathsf{Cut}_{C_\mu} \right.$.

2.2'. $C \simeq \bigwedge_{\iota\in J} C_\iota$: symmetric to (Case 2.2).

Theorem and Definition 4.2

We define an operator $\mathcal{E}$ such that the following holds: $d \vdash^{\alpha}_{m+1} \Gamma \implies \mathcal{E}(d) \vdash^{3^\alpha}_{m} \Gamma$.

Proof:

1. $d \simeq \left\{ \dfrac{\begin{array}{cc} d_0 & d_1 \\ | & | \\ \Gamma, C : \alpha_0 & \Gamma, \neg C : \alpha_1 \end{array}}{\Gamma : \alpha}\,\mathsf{Cut}_C \right.$: Then $\mathrm{rk}(C) \le m$ and, by IH, $\mathcal{E}(d_0) \vdash^{3^{\alpha_0}}_{m} \Gamma, C$ and $\mathcal{E}(d_1) \vdash^{3^{\alpha_1}}_{m} \Gamma, \neg C$.

Hence $\mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1)) \vdash^{3^{\alpha_0}\#3^{\alpha_1}}_{m} \Gamma$ by Theorem 4.1. So we could define $\mathcal{E}(d)$ to be $\mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1))$. But for reasons which become clear later we set $\mathcal{E}(d) := \mathsf{Rep}\,\mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1))$.

2. otherwise: $\mathcal{E}(d) := \mathcal{I}(\mathcal{E}(d_\iota))_{\iota\in I} \simeq \left\{ \dfrac{\begin{array}{c} \mathcal{E}(d_\iota) \\ | \\ \ldots\Gamma, \Delta_\iota : 3^{\alpha_\iota} \ldots (\iota \in I) \end{array}}{\Gamma : 3^\alpha}\,\mathcal{I} \right.$ if $d = \left\{ \dfrac{\begin{array}{c} d_\iota \\ | \\ \ldots\Gamma, \Delta_\iota : \alpha_\iota \ldots (\iota \in I) \end{array}}{\Gamma : \alpha}\,\mathcal{I} \right.$

Theorem 4.3 (Inversion)

a) $A \simeq \bigwedge_{\iota \in J} A_\iota$ & $\mathsf{Z}^\infty \vdash^\alpha_m \Gamma, A \implies \mathsf{Z}^\infty \vdash^\alpha_m \Gamma, A_\iota$ for each $\iota \in J$.

b) $\mathsf{Z}^\infty \vdash^\alpha_m \Gamma, A_0, A_1 \implies \mathsf{Z}^\infty \vdash^\alpha_m \Gamma, A_0, A_1$.

Definition

$\mathcal{L}_0[X] := \mathcal{L}_0 \cup \{X\}$,

A closed $\mathcal{L}_0[X]$-formula (sequent) is called X-positive if it contains no subformula $\neg Xt$.

$\models^\alpha A \;:\Leftrightarrow\; (\mathbb{N}, \{n : |n|_\prec < \alpha\}) \models A$,

$\models^\alpha \{A_1, ..., A_k\} \;:\Leftrightarrow \models^\alpha A_1 \vee ... \vee A_k$

Lemma 4.4 (Boundedness)

Let Γ be X-positive.

$\mathsf{Z}^\infty \vdash^\beta_0 \neg\mathrm{Prog}_\prec(X), \neg X s_1, ..., \neg X s_k, \Gamma$ & $|s_1|_\prec, ..., |s_k|_\prec \leq \alpha \implies \models^{\alpha+2^\beta} \Gamma$.

Corollary $\mathsf{Z}^\infty \vdash^\beta_0 \mathrm{TI}_\prec(X, \underline{n}) \implies |n|_\prec \leq 2^\beta$.

Proof of Corollary:

$\mathsf{Z}^\infty \vdash^\beta_0 \mathrm{TI}_\prec(X, \underline{n}) \stackrel{4.3a}{\Rightarrow} \mathsf{Z}^\infty \vdash^\beta_0 \neg\mathrm{Prog}_\prec(X), \forall y \prec \underline{n} Xy \stackrel{4.4}{\Rightarrow} \models^{2^\beta} \forall y \prec \underline{n} Xy \;\Rightarrow\; |n|_\prec \leq 2^\beta$.

Proof of Lemma by induction on β:

Abbreviations: $\Lambda := \{\neg X s_1, ..., \neg X s_k\}$.

Let $d \vdash^\beta_0 \neg\mathrm{Prog}_\prec(X), \Lambda, \Gamma$.

1.1. $\mathsf{last}(d) = \mathsf{Ax}^\infty_\Delta$ and $\Delta \cap \mathrm{TRUE}_0 \neq \emptyset$: $\Gamma \cap \mathrm{TRUE}_0 \neq \emptyset$ and the claim is trivial.

1.2. $\mathsf{last}(d) = \mathsf{Ax}^\infty_\Delta$ and $Xt, \neg Xs \subseteq \Delta$ with $\mathsf{val}(t) = \mathsf{val}(s)$:

Then $Xt \in \Gamma$ and $\neg Xs \in \Lambda$. Hence $|t|_\prec = |s|_\prec \leq \alpha < \alpha + 2^\beta$ and thus $\models^{\alpha+2^\beta} \Gamma$.

2. $\mathsf{last}(d) = \bigvee^{s_0}_{\neg\mathrm{Prog}_\prec(X)}$: Then $\vdash^{\beta_0}_0 \neg\mathrm{Prog}_\prec(X), \Lambda, \Gamma, \forall y {\prec} s_0\, Xy \wedge \neg X s_0$ with $\beta_0 < \beta$.

By 4.3a (Inversion) we get (1) $\vdash^{\beta_0}_0 \neg\mathrm{Prog}_\prec(X), \Lambda, \Gamma, \forall y {\prec} s_0\, Xy$, and (2) $\vdash^{\beta_0}_0 \neg\mathrm{Prog}_\prec(X), \neg X s_0, \Lambda, \Gamma$.

By IH from (1) we get $\models^{\alpha+2^{\beta_0}} \Gamma, \forall y {\prec} s_0\, Xy$.

(Case 1) $\models^{\alpha+2^{\beta_0}} \Gamma$: Then also $\models^{\alpha+2^\beta} \Gamma$, since $\beta_0 \leq \beta$ and Γ is X-positive.

(Case 2) $\models^{\alpha+2^{\beta_0}} \forall y {\prec} s_0\, Xy$: Then $|m|_\prec < \alpha + 2^{\beta_0}$ for all $m \prec \mathsf{val}(s_0)$, i.e. $|s_0|_\prec \leq \alpha + 2^{\beta_0}$.

From this together with $|s_1|_\prec, ..., |s_k|_\prec \leq \alpha$ and (2) by IH we obtain $\models^{\alpha+2^{\beta_0}+2^{\beta_0}} \Gamma$, and thus $\models^{\alpha+2^\beta} \Gamma$.

3. $\mathsf{last}(d) = \bigwedge_C$ with $C \simeq \bigwedge_{\iota \in I} C_\iota \in \Gamma$: Then, for all $\iota \in I$, $\vdash^{\beta_\iota}_0 \neg\mathrm{Prog}_\prec(X), \Lambda, \Gamma, C_\iota$ and $\beta_\iota < \beta$. Hence, by IH, $\models^{\alpha+2^{\beta_\iota}} \Gamma, C_\iota$ for all ι. Since Γ is X-positive, this implies $\models^{\alpha+2^\beta} \Gamma, C_\iota$ for all ι, and thus $\models^{\alpha+2^\beta} \Gamma, C$.

4. $\bigvee^\mu_C$ with $C \in \Gamma$: as 3.

5. $\mathsf{last}(d) = \mathsf{Rep}$: immediate by IH.

Embedding of $\mathrm{PL1}+\mathcal{Z}$ **into** Z^∞ **; unprovability of** $\mathrm{TI}(\varepsilon_0)$

Lemma 4.5

$\mathrm{PL1} \vdash_0^k \Gamma \;\&\; \mathrm{FV}(\Gamma)=\emptyset \implies \mathsf{Z}^\infty \vdash_0^k \Gamma$

Proof by induction on k:

1. $\{\neg A, A\} \subseteq \Gamma$ where A is atomic: Then either $A \in \mathrm{TRUE}_0$ or $\neg A \in \mathrm{TRUE}_0$ or $A = Xs$. Hence $\mathsf{Z}^\infty \vdash_0^k \Gamma$.

2. $\forall x A \in \Gamma$ und $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(y)$: Then by 2.2 $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(t)$ for all $t \in T$.

Hence by IH $\mathsf{Z}^\infty \vdash_0^{k-1} \Gamma, A_x(t)$ for all $t \in T$ and thus $\mathsf{Z}^\infty \vdash_0^k \Gamma$, since $\forall x A \simeq \bigwedge_{t\in T} A_x(t)$.

3. $\exists x A \in \Gamma$ and $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(t)$: w.l.o.g. $t \in T$ (cf. Lemma 2.2). By IH $\mathsf{Z}^\infty \vdash_0^{k-1} \Gamma, A_x(t)$ and thus $\mathsf{Z}^\infty \vdash_0^k \Gamma$.

4. $(\wedge)$,$(\vee)$: as 2. and 3.

Lemma 4.6

$s,t \in T \;\&\; \mathsf{val}(s)=\mathsf{val}(t) \;\&\; \mathrm{FV}(F) \subseteq \{x\} \implies \mathsf{Z}^\infty \vdash_0^{2\cdot\mathrm{rk}(F)} \neg F_x(s), F_x(t)$.

Proof by induction on $\mathrm{rk}(F)$.

Lemma 4.7 (Derivability of Complete Induction)

Let $\forall x F$ be closed and $r := 2\cdot \mathrm{rk}(F)$. Then for each $t \in T$ we have

$\mathsf{Z}^\infty \vdash_0^{r+2n} \neg F_x(0), \neg\forall x(F \to F_x(\mathsf{S}x)),\ F_x(t)$ where $n := \mathsf{val}(t)$.

Proof by induction on n:

Let $\Gamma := \{\neg F(0), \neg\forall x(F \to F_x(\mathsf{S}x))\}$, i.e. $\Gamma = \{\neg F(0), \exists x(F \wedge \neg F_x(\mathsf{S}x))\}$.

1. $\mathsf{val}(t)=0$: Then $\vdash_0^r \Gamma, F(0)$ by Lemma 4.6.

2. $\mathsf{val}(t)=n{+}1$:

(1) $\vdash_0^{r+2n} \Gamma, F(\underline{n})$	[IH]
(2) $\vdash_0^{r} \neg F(\mathsf{S}\underline{n}), F(t)$	[L.4.6]
(3) $\vdash_0^{r+2n+1} \Gamma, F(\underline{n}) \wedge \neg F(\mathsf{S}\underline{n}), F(t)$	[(1),(2),$(\wedge)$]
(4) $\vdash_0^{r+2n+2} \Gamma, F(t)$	[(3), $\bigvee^{\underline{n}}_{\exists x(F\wedge\neg F_x(\mathsf{S}x))}$]

Lemma 4.8

For each $C \in \mathcal{Z}$ there is an $\alpha < \omega+\omega$ such that $\mathsf{Z}^\infty \vdash_0^\alpha C$.

Proof:

1. $C = \forall\vec{y}\forall z[\,F(0,\vec{y}) \to \forall x(F(x,\vec{y}) \to F(\mathsf{S}x,\vec{y})) \to F(z,\vec{y}))]$:

By Lemma 4.7 we have $\vdash_0^\omega \neg F(0,\vec{s}), \neg\forall x(F(x,\vec{s}) \to F(\mathsf{S}x,\vec{s})), F(t,\vec{s})$ for all $\vec{s}, t \in T$. This yields the claim.

2. $C = \forall\vec{y}\forall x_1\forall x_2(x_1{=}x_2 \to A(x_1,\vec{y}) \to A(x_2,\vec{y}))$ where A is atomic:

Then for any closed $\mathcal{L}_0$-terms $\vec{s}, t_1, t_2$ we have $\vdash_0^0 \neg(t_1{=}t_2), \neg A(t_1,\vec{s}), A(t_2,\vec{s})$. [If $\mathsf{val}(t_1)=\mathsf{val}(t_2)$ then either $\neg A(t_1,\vec{s}) \in \mathrm{TRUE}_0$ or $A(t_2,\vec{s}) \in \mathrm{TRUE}_0$ or $\{\neg A(t_1,\vec{s}), A(t_2,\vec{s})\} = \{\neg Xt, Xt'\}$ with $\mathsf{val}(t)=\mathsf{val}(t')$. Otherwise $\neg(t_1{=}t_2) \in \mathrm{TRUE}_0$.] Hence $\vdash_0^\omega C$.

3. $C = \forall x\forall y(\mathsf{S}x{=}\mathsf{S}y \to x{=}y)$: The claim holds, since $\vdash_0^0 \neg\mathsf{S}s{=}\mathsf{S}t, s{=}t$ for all $s,t \in T$.

4. All other axioms are of the form $\forall\vec{x}A$ with $A_{\vec{x}}(\vec{s}) \in \mathrm{TRUE}_0$ for all $\vec{s} \in T$.

Theorem 4.9
If $\mathcal{Z} \vdash C$ and $\mathrm{FV}(C) = \emptyset$ then $\mathsf{Z}^\infty \vdash_0^\alpha C$ for some $\alpha < \varepsilon_0$.

Proof:

Assume $\mathcal{Z} \vdash C$. Then there are axioms $A_1, ..., A_l \in \mathcal{Z}$ such that $\mathrm{PL1} \vdash \neg A_1, ..., \neg A_l, C$. Now Lemma 4.5 yields $\vdash_0^\omega \neg A_1, ..., \neg A_n, C$, and by Lemma 4.8 we have $\vdash_0^\beta A_i$ $(i = 1, ..., l)$ for some $\beta < \omega + \omega$.

Hence $\vdash_m^{\omega+\omega} C$, with $m := \max_{1 \le i \le l} \mathrm{rk}(A_i){+}1$. Now we apply the Cut-Elimination Theorem 4.2 and obtain $\vdash_0^\alpha C$ with $\alpha := 3^{\cdot^{\cdot^{3^{\omega+\omega}}}} < \varepsilon_0$.

Theorem 4.10 $\mathcal{Z} \vdash \mathrm{TI}_\prec(X, \underline{n}) \implies |n|_\prec < \varepsilon_0$

Proof:

$\mathcal{Z} \vdash \mathrm{TI}_\prec(X, \underline{n}) \overset{4.9}{\Longrightarrow} \mathsf{Z}^\infty \vdash_0^\alpha \mathrm{TI}_\prec(X, \underline{n}) \overset{4.4(\mathrm{Cor.})}{\Longrightarrow} |n|_\prec \le 2^\alpha < \varepsilon_0$.

Corollary. If $\varepsilon_0 \le \|{\prec}\| := \sup\{|n|_\prec + 1 : n \in \mathbb{N}\}$ then $\mathcal{Z} \nvdash \mathrm{Prog}_\prec(X) \to \forall y X y$.

Proof:

$\mathcal{Z} \vdash \mathrm{Prog}_\prec(X) \to \forall y X y \;\Rightarrow\; \mathcal{Z} \vdash \mathrm{TI}_\prec(X, \underline{n})$ (for all $n \in \mathbb{N}$) $\overset{4.4(\mathrm{Cor.})}{\Rightarrow} |n|_\prec \le 2^\alpha$ for all $n \;\Rightarrow\; \|{\prec}\| < \varepsilon_0$.

Provability of transfinite induction in $\mathcal{Z}$

We assume a bijective coding of finite sequences of natural numbers $\mathbb{N}^{<\omega} \to \mathbb{N}$, $(a_0, ..., a_{n-1}) \mapsto \langle a_0, ..., a_{n-1}\rangle$ such that

(i) $\langle\,\rangle = 0$, and $a_i < \langle a_0, ..., a_{n-1}\rangle$ for $i < n$,

(ii) for each fixed n the mapping $(a_0, ..., a_n) \mapsto \langle a_0, ..., a_n\rangle$ is primitive recursive,

(iii) there is a primitive recursive function $(a, i) \mapsto (a)_i$ with $(\langle a_0, ..., a_{n-1}\rangle)_i = a_i$ for $i < n$.

In the following a, b, c, x, y, z denote natural numbers.

Definition of $b \prec' a$

$b \prec' a$ if, and only if, $a = \langle a_0, ..., a_n\rangle$ and one of the following cases holds

(i) $b = \langle a_0, ..., a_{k-1}\rangle$ with $k \le n$,

(ii) $b = \langle a_0, ..., a_{k-1}, b_k, ..., b_m\rangle$ with $k \le \min\{m, n\}$ and $b_k \prec' a_k$.

Inductive Definition of a set OT **of ordinal notations**

1. $0 \in \mathrm{OT}$,
2. $a_0, ..., a_n \in \mathrm{OT}$ & $a_n \preceq' \ldots \preceq' a_0 \implies \langle a_0, ..., a_n\rangle \in \mathrm{OT}$.

Definition $b \prec a :\Leftrightarrow a, b \in \mathrm{OT}$ & $b \prec' a$

Abbreviation: $a \oplus b := a * b$ (concatenation) , $\quad \overline{F}(y) := \forall x(\forall z{\prec}x F(z) \to \forall z{\prec}x{\oplus}\langle y\rangle\, F(z))$

Lemma 4.11 $\mathcal{Z} \vdash \mathrm{Prog}(F) \to \mathrm{Prog}(\overline{F})$

Proof (in $\mathcal{Z}$):

Assume (1) $\mathrm{Prog}(F)$, (2) $\forall y{\prec}b\overline{F}(y)$, (3) $\forall z{\prec}aF(z)$, (4) $c \prec a \oplus \langle b\rangle$.

We have to prove $F(c)$. From (4) it follows that either $c \prec a$ or $c = a \oplus \langle b_1, ..., b_n\rangle$ with $b_n \preceq ... \preceq b_1 \prec b$.

1. $c \prec a$: $F(c)$ follows from (3).
2. $c = a \oplus \langle b_1, ..., b_n\rangle$ with $b_n \preceq ... \preceq b_1 \prec b$:

One easily shows that $\prec'$ is transitive, hence $b_1, ..., b_n \prec b$.

By (2) we get $\overline{F}(b_i)$ for $i = 1, ..., n$, i.e. $\forall x(\forall z{\prec}xF(z) \to \forall z \prec x \oplus \langle b_i\rangle\, F(z))$ for $i = 1, .., n$.

Now by complete induction (on i) we get $\forall z \prec a \oplus \langle b_1, ..., b_i\rangle\, F(z)$ $(=: A(c, i))$.

Hence $\forall z{\prec}cF(z)$ and thus by (1) $F(c)$.

Lemma 4.12 $\mathcal{Z} \vdash \mathrm{TI}_\prec(\overline{F}, y) \to \mathrm{TI}_\prec(F, \langle y\rangle)$.

Proof:

Assume $\mathrm{Prog}(\overline{F}) \to \forall z \prec y\overline{F}(z)$ and $\mathrm{Prog}(F)$. By Lemma 4.11 we get $\mathrm{Prog}(\overline{F}) \land \forall z \prec y\overline{F}(z)$, hence $\overline{F}(y)$, and from this $\forall z \prec \langle y\rangle F(z)$.

Theorem 4.13 $\mathcal{Z} \vdash \mathrm{TI}_\prec(F, \underline{a})$, for each $a \in \mathbb{N}$.

Proof:

If $a = 0$ or $a \notin \mathrm{OT}$ then $\forall z \prec \underline{a}(0 = 1)$ and therefore $\mathrm{TI}_\prec(F, \underline{a})$.

Now let $c_0 := 0$, $c_{m+1} := \langle c_m\rangle$. Then for each $a \in \mathrm{OT}$ there is an m with $a \prec c_m$, and by (meta-)induction on m we obtain $\mathcal{Z} \vdash \mathrm{TI}_\prec(F, c_m)$. [Induction step: IH $\Rightarrow$ $\mathcal{Z} \vdash \mathrm{TI}_\prec(\overline{F}, c_m) \overset{4.12}{\Rightarrow} \mathcal{Z} \vdash \mathrm{TI}_\prec(F, c_{m+1})$]

Definition $o(0) := 0$, $o(\langle a_0, ..., a_n\rangle) := \omega^{o(a_0)} + \ldots + \omega^{o(a_k)}$

Lemma 4.14 o maps $(\mathrm{OT}, \prec)$ isomorphic onto $(\varepsilon_0, <)$.

Proof:

1. From the definition of $\prec'$ we get by induction on a: $a \not\prec' a$ and $\forall b(b \prec' a \vee a = b \vee a \prec' b)$.
2. $\forall a(o(a) < \varepsilon_0)$: trivial.
3. By induction on b we prove: $b \prec a \Rightarrow o(b) < o(a)$.

3.1. If $a = \langle a_0, ..., a_n\rangle$ and $b = \langle a_0, ..., a_{k-1}\rangle$ with $k \leq n$ then $o(b) < o(b) + \omega^{o(a_k)} + \ldots + \omega^{o(a_n)} = o(a)$.

3.2. If $a = \langle a_0, ..., a_n\rangle$, $b = \langle a_0, ..., a_{k-1}, b_k, ..., b_m\rangle$ with $k \leq \min\{m, n\}$, $b_k \prec a_k$, then by IH $o(b_m) \leq ... \leq o(b_k) < o(a_k)$ and thus $o(b) = \omega^{o(a_0)} + \ldots + \omega^{o(a_{k-1})} + \omega^{o(b_k)} + \ldots + \omega^{o(b_m)} < \omega^{o(a_0)} + \ldots + \omega^{o(a_k)} \leq o(a)$.

4. From 1. and 3. it follows that $o|_{\mathrm{OT}}$ is injective, and that $a, b \in \mathrm{OT}$ & $o(b) < o(a)$ implies $b \prec a$.

5. By induction on $\alpha < \varepsilon_0$ we prove $\exists a \in \mathrm{OT}(o(a) = \alpha)$: Let $\alpha \neq 0$. By Lemma 3.16a there are $\alpha_0 \geq ... \geq \alpha_n$ such that $\alpha = \omega^{\alpha_0} + \ldots + \omega^{\alpha_n}$ and $\alpha_0 < \alpha$ (the latter follows from $\alpha_0 \leq \omega^{\alpha_0} \leq \alpha < \varepsilon_0$). Now by IH there are $a_0, ..., a_n \in \mathrm{OT}$ with $\alpha_i = o(a_i)$. From $\alpha_n \leq ... \leq \alpha_0$ we obtain $a_n \preceq ... \preceq a_0$ by 4.
Hence $a := \langle a_0, ..., a_n\rangle \in \mathrm{OT}$ and $o(a) = \alpha$.

Corollary $\prec$ is wellfounded with $|a|_\prec = \begin{cases} o(a) & \text{if } a \in \mathrm{OT} \\ 0 & \text{otherwise} \end{cases}$

Proof: If $a \in \mathrm{OT}$ then $|a|_\prec = \sup\{|b|_\prec{+}1 : b \prec a\} \overset{\mathrm{IH}}{=} \sup\{o(b){+}1 : b \prec a\} \overset{4.14}{=} o(a)$.

The Hydra game

A *hydra* is a finite unlabelled tree. By 0 we denote the hydra consisting only of one node.

Let σ be the rightmost head of a the hydra $h \neq 0$. If Hercules chops off this head the hydra h chooses an arbitrary number n and transforms itself into a new hydra $h[n]$ as follows (where τ is the node immediately below σ, and h^- is h without σ):
Case 1: If τ is the root of h, then $h[n] := h^-$.
Case 2: Otherwise $h[n]$ arises from h^- by sprouting n replicas of $h^-|_\tau$ from the node immediately below τ.

A *hydra game* is a finite or infinite sequence $(h_i)_{i<\alpha}$ of hydras, such that $\forall i{+}1 < \alpha \exists n_i(h_{i+1} = h_i[n_i])$.

Theorem 4.15 Each hydra game terminates, i.e., $\forall h \forall (n_i)_{i<\omega} \exists k(h[n_0][n_1]...[n_k] = 0)$.

Theorem 4.16 $\mathcal{Z} \nvdash \forall h \forall (n_i)_{i<\omega} \exists k(h[n_0][n_1]...[n_k] = 0)$.

Proof of Theorem 4.15: To each hydra h we assign its Gödel number $\ulcorner h \urcorner$ as follows: $\ulcorner h \urcorner := \langle \ulcorner h_0 \urcorner, ..., \ulcorner h_{n-1} \urcorner\rangle$ where $h_0, ..., h_{n-1}$ are the immediate subtrees of h. Obviously the mapping $h \mapsto \ulcorner h \urcorner$ is a bijection from the set of all hydras onto $\mathbb{N}$. Therefore from now on we identify hydras and natural numbers.
The above operation $a \mapsto a[n]$ can be defined by primitive recursion as follows:

1. $\mathsf{tp}(0) := 0$, $0[n] := 0$.
2. If $a = \langle a_0, ..., a_m\rangle$ with $a_m = 0$ then $\mathsf{tp}(a) := 1$ and $a[n] := \langle a_0, ..., a_{m-1}\rangle$.
3. If $a = \langle a_0, ..., a_m\rangle$ with $tp(a_m) = 1$ then $\mathsf{tp}(a) := \omega$ and $a[n] := \langle a_0, ..., a_{m-1}, \underbrace{a_m[0], ..., a_m[0]}_{n+1}\rangle$.
4. If $a = \langle a_0, ..., a_m\rangle$ with $\mathsf{tp}(a_m) = \omega$ then $\mathsf{tp}(a) := \omega$ and $a[n] := \langle a_0, ..., a_{m-1}, a_m[n]\rangle$.

By induction on a we get $o(a[n]) < o(a)$ for each $a \neq 0$. This proves 4.15.

Proof of Theorem 4.16 (Unprovability of termination of the hydra game)

Abbreviation: Let $\lhd \subseteq \mathbb{N}\times\mathbb{N}$ be an arbitrary arithmetical relation.

$\mathrm{WF}_\lhd(G) :\equiv \forall x\exists! y G(x,y) \to \exists x, z_0, z_1 (G(x,z_0) \wedge G(x{+}1,z_1) \wedge \neg z_1 \lhd z_0)$,

$\mathrm{WF}_\lhd(X) :\equiv \mathrm{WF}_\lhd(G)$ with $G(x,y) :\equiv X\langle x,y\rangle$.

$\mathrm{TI}_\lhd(F) :\equiv \mathrm{Prog}_\lhd(F) \to \forall x F$.

$\|\lhd\| := \sup\{|a|_\lhd + 1 : a \in \mathbb{N}\}$ if $\lhd$ is wellfounded.

Lemma 4.17 $\mathcal{Z} \vdash \mathrm{WF}_\lhd(X) \iff \mathcal{Z} \vdash \mathrm{TI}_\lhd(X)$

Proof:

"$\Longrightarrow$": Assume $\mathcal{Z} \vdash \mathrm{WF}_\lhd(X)$; then also $\mathcal{Z} \vdash \mathrm{WF}_\lhd(G)$ for each formula $G(x,y)$.

Now we work "in $\mathcal{Z}$": *Assumption:* $\mathrm{Prog}_\lhd(X) \wedge a \notin X$.

For suitable G we prove $\neg\mathrm{WF}_\lhd(G)$, i.e. $\forall i\exists! b G(i,b) \wedge \forall i, b_0, b_1 (\, G(i,b_0) \wedge G(i{+}1,b_1) \to b_1 \lhd b_0\,)$.

$A(i,s) :\Leftrightarrow \forall j < i((s)_{j+1} \lhd (s)_j \wedge (s)_{j+1} \notin X \wedge \forall x \lhd (s)_j [x < (s)_{j+1} \to x \in X])$,

$G(i,b) :\Leftrightarrow \exists s((s)_0 = a \wedge (s)_i = b \wedge A(i,s))$.

(0) $A(i,s) \wedge A(i,\tilde{s}) \wedge (s)_0 = (\tilde{s})_0 \Rightarrow \forall j \le i((s)_j = (\tilde{s})_j)$,

(1) G is function: cf. (0).

(2) G total: By induction on i we prove $\exists b\, G(i,b)$.
Induction step: $G(i,b_0) \;\Rightarrow\; b_0 \notin X \overset{\mathrm{Prog}_\lhd(X)}{\Longrightarrow} \exists b \lhd b_0 (b \notin X) \Rightarrow \exists b_1\, G(i{+}1,b_1)$.

(3) $G(i,b_0) \wedge G(i{+}1,b_1) \Rightarrow \exists s[\,(s)_0 = a \wedge (s)_i = b_0 \wedge A(i,s)] \wedge \exists \tilde{s}[\,(\tilde{s})_0 = a \wedge (\tilde{s})_{i+1} = b_1 \wedge A(i{+}1,\tilde{s})] \overset{(0)}{\Rightarrow}$
$b_1 = (\tilde{s})_{i+1} \lhd (\tilde{s})_i = (s)_i = b_0$.

"$\Longleftarrow$": left to the reader.

Definition $b \prec_1 a \;:\Leftrightarrow\; a \neq 0 \;\&\; \exists i(b = a[i])$

Lemma 4.18 $b \prec a \Rightarrow \exists i(b \preceq a[i])$.

Theorem 4.19

a) $\prec_1$ is wellfounded and $\|{\prec_1}\| = \varepsilon_0$.

b) $\mathcal{Z} \nvdash \mathrm{WF}_{\prec_1}(X)$.

Proof:

From $\forall a \neq 0 \forall n(o(a[n]) < o(a))$, it follows that $\prec_1$ is wellfounded, and by induction on $o(a)$ we get $|a|_{\prec_1} \le o(a)$. Hence $\|\prec_1\| \le \varepsilon_0$, since $o(a) < \varepsilon_0$ for all a.

Using Lemma 4.18 we obtain $\forall a \in \mathrm{OT}(o(a) \le |a|_{\prec_1})$ by induction on a:

$0 \neq a \in \mathrm{OT} \Rightarrow\; o(a) = \sup\{o(b){+}1 : b \prec a\} \overset{4.18}{\le} \sup\{o(a[i]){+}1 : i \in \mathbb{N}\} \overset{\mathrm{IH}}{\le} \sup\{|a[i]|_{\prec_1}{+}1 : i \in \mathbb{N}\} = \sup\{|b|_{\prec_1}{+}1 : b \prec_1 a\} = |a|_{\prec_1}$. Hence $\varepsilon_0 = \sup\{o(a){+}1 : a \in \mathrm{OT}\} \le \sup\{|a|_{\prec_1}{+}1 : a \in \mathbb{N}\} = \|{\prec_1}\|$.

b) $\varepsilon_0 \le \|{\prec_1}\| \overset{\mathrm{Cor.4.10}}{\Longrightarrow} \mathcal{Z} \nvdash \mathrm{TI}_{\prec_1}(X) \overset{4.17}{\Rightarrow} \mathcal{Z} \nvdash \mathrm{WF}_{\prec_1}(X)$.

Remark: Theorem 4.16 follows from 4.19 b).

§5 Gentzen's consistency proof for pure number theory

The proof-system Z

The language of Z is $\mathcal{L}_0(\mathcal{X})$ (language of arithmetic with set variables)

Let $\mathcal{AX}(\mathsf{Z})$ be the set of all sequents having one of the following shapes:

(G1) $t{=}t$

(G2) $\neg s{=}t, \neg A_x(s), A_x(t)$ if A is a literal

(S0) $\neg \mathsf{S}t{=}0$

(S1) $\neg \mathsf{S}s{=}\mathsf{S}t, s{=}t$

(PR0) $\mathbf{0}^n t_1...t_n{=}0$

(PR1) $\mathrm{I}^n_i t_1...t_n{=}t_i$

(PR2) $(\circ h g_1...g_m)t_1...t_n{=}hg_1t_1...t_n \dots g_m t_1...t_n$

(PR3.0) $(\mathrm{R}gh)t_1...t_n 0{=}gt_1...t_n$

(PR3.1) $(\mathrm{R}gh)t_1...t_n\mathsf{S}s{=}ht_1...t_n s(\mathrm{R}gh)t_1...t_n s$

The *inference symbols of* Z are those of PL1 plus

(Ax_Δ) Δ for $\Delta \in \mathcal{AX}(\mathsf{Z})$,

$(\mathsf{Ind}^{y,t}_F)$ $\dfrac{\neg F, F_y(\mathsf{S}y)}{\neg F_y(0), F_y(t)}$ $!y!$.

$$\deg(\mathcal{I}) := \begin{cases} \mathrm{rk}(C)+1 & \text{if } \mathcal{I} = \mathsf{Cut}_C \text{ or } \mathsf{Ind}^{y,t}_C \\ 0 & \text{otherwise} \end{cases}$$

In this section $\mathcal{I}$ is always an inference symbol of Z, and $d, d', ...$ denote Z-derivations.

Definition of $\mathcal{I}(x/t)$ and $\mathcal{I}^u$

$\mathsf{Ind}^{y,s}_F(x/t) := \mathsf{Ind}^{y,s_x(t)}_{F_x(t)}$, $(\mathsf{Ind}^{y,s}_F)^u := \mathsf{Ind}^{u,s}_{F_y(u)}$; otherwise $\mathcal{I}(x/t)$ and $\mathcal{I}^u$ are defined as in PL1.

Lemma 5.1 Z is closed under substitution.

Proof: Let $\mathcal{I} = \mathsf{Ind}^{y,s}_F$ and $\mathcal{I}' := \mathcal{I}(x/t)$ with $x \neq y$. Then $|\mathcal{I}'| = \{0\} = |\mathcal{I}|$, $\mathrm{Eig}(\mathcal{I}') = \{y\} = \mathrm{Eig}(\mathcal{I})$, $\deg(\mathcal{I}') = \mathrm{rk}(F_x(t)) = \mathrm{rk}(F) = \deg(\mathcal{I})$, $\Delta(\mathcal{I}') = \{\neg F_x(t)_y(0), F_x(t)_y(s_x(t))\} = \{\neg F_y(0)_x(t), F_y(s)_x(t)\} = \Delta(\mathcal{I})_x(t)$, $\Delta_0(\mathcal{I}') = \{\neg F_x(t), F_x(t)_y(\mathsf{S}y)\} = \{\neg F_x(t), F_y(\mathsf{S}y)_x(t)\} = \Delta_0(\mathcal{I})_x(t)$.
$\Delta(\mathcal{I}^u) = \{\neg F_y(u)_u(0), F_y(u)_u(s)\} = \{\neg F_y(0), F_y(s)\} = \Delta(\mathcal{I})$, $\Delta_0(\mathcal{I}^u) = \{\neg F_y(u), F_y(u)_u(\mathsf{S}u)\} = \{\neg F_y(u), F_y(\mathsf{S}y)_y(u)\} = \Delta_0(\mathcal{I})_y(u)$ (note that $u \notin \mathrm{FV}(F_y(s))$).

Definition

For $m \in \mathbb{N}$, let $\mathcal{Z}_m$ denote the subsystem arising from $\mathcal{Z}$ by deleting all induction axioms $\forall\vec{v}\forall z(F_x(0) \to \forall x(F \to F_x(\mathsf{S}x)) \to F_x(z))$ with $\mathrm{rk}(F) > m$.

Theorem 5.2

If $\mathcal{Z}_m \vdash A$ then $\mathsf{Z} \vdash_{m+1} A$, i.e., there is a Z-derivation d of A such that $\mathrm{rk}(C) \leq m$ for each Cut_C or $\mathsf{Ind}^{y,t}_C$ occurring in d.

Proof:

Let Z_m be the subsystem of Z corresponding to $\mathcal{Z}_m$, i.e., Z_m arises from Z by removing all inferences $\mathsf{Ind}_F^{y,t}$ with $\mathrm{rk}(F) > m$. We first show $\mathsf{Z}_m \vdash_0 A$, for each $A \in \mathcal{Z}_m$:

1. If A is not an induction axiom then $\mathsf{Z}_m \vdash_0 A$ is easy to see.

2. Now let $A \equiv \forall\vec{v}\forall z(F_x(0) \to \forall x(F \to F_x(\mathsf{S}x)) \to F_x(z))$ with $\mathrm{rk}(F) \le m$.

Let $G := \forall x(F \to F_x(\mathsf{S}x))$ and $y \in \mathsf{FV} \setminus \mathrm{FV}(G)$:

$\mathrm{PL1} \vdash_0 F_x(y) \wedge \neg F_x(\mathsf{S}y), \neg F_x(y), F_x(\mathsf{S}y) \;\Rightarrow\; \mathrm{PL1} \vdash_0 \neg G, \neg F_x(y), F_x(\mathsf{S}y) \overset{\mathsf{Ind}}{\Rightarrow} \mathsf{Z}_m \vdash_0 \neg G, \neg F_x(0), F_x(u) \;\Rightarrow$ $\mathsf{Z}_m \vdash_0 \forall z(F_x(0) \to G \to F_x(z)) \;\Rightarrow\; \mathsf{Z}_m \vdash_0 A$.

Now assume $\mathcal{Z}_m \vdash A$. Then there are $A_1, ..., A_n \in \mathcal{Z}_m$ such that $\mathrm{PL1} \vdash_0 \neg A_1, ..., \neg A_n, A$ and (as shown above) $\mathsf{Z}_m \vdash_0 A_i$ for $i = 1, ..., n$. This yields $\mathsf{Z}_m \vdash_r A$ with $r := \max\{m, \mathrm{rk}(A_1), ..., \mathrm{rk}(A_n)\}+1$.

Now we apply Theorem 2.5 (with $\mathfrak{S} := \mathsf{Z}_m$ and $m_0 := m+1$) and obtain $\mathsf{Z}_m \vdash_{m+1} A$, i.e., $\mathsf{Z} \vdash_{m+1} A$.

Definition of $\mathrm{hgt}_d(\nu)$ and $O_d(\nu)$ (Gentzen)

Let d be a Z-derivation and ν a position in d.

$\mathcal{I}_d(\nu) :=$ the inference at position ν in d

$$\mathrm{rk}_d(\nu) := \begin{cases} \mathrm{rk}(C) & \text{if } \mathcal{I}_d(\nu) = \mathsf{Cut}_C \text{ or } \mathsf{Ind}_C^{y,t} \\ 0 & \text{otherwise} \end{cases}$$

$\mathrm{hgt}_d(\nu) := \sup\{\mathrm{rk}_d(\sigma) : \sigma \text{ is strictly below } \nu\}$

$$O_d(\nu) := \begin{cases} \omega_m(O_d(\nu 0) \# O_d(\nu 1)) & \text{if } \mathcal{I}_d(\nu) = \mathsf{Cut}_C \\ \omega_m(O_d(\nu 0) \cdot \omega) & \text{if } \mathcal{I}_d(\nu) = \mathsf{Ind}_C^{y,t} \\ (\sup_{i<k} O_d(\nu i)) + 1 & \text{otherwise} \end{cases}$$

where $m := \mathrm{hgt}_d(\nu 0) - \mathrm{hgt}_d(\nu) = \mathrm{rk}(C) \dot{-} \mathrm{hgt}_d(\nu)$.

$O(d) := O_d(\langle\rangle)$.

Definition (End-pieces)

1. $*$ is an end-piece.
2. If $\mathfrak{a}$ is an end-piece, and d a Z-derivation then $\mathsf{Cut}_C \mathfrak{a} d$ and $\mathsf{Cut}_C d \mathfrak{a}$ are end-pieces.

We use $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$ as syntactic variables for end-pieces. $\mathfrak{a}\{q\} := \begin{smallmatrix} q \\ \cdot\,\vdots\,\cdot \\ \mathfrak{a} \end{smallmatrix}$:= the result of substituting q for $*$ in $\mathfrak{a}$.

Definition of $\Gamma^*(d)$

$$\Gamma^*(\mathcal{I} d_0 ... d_{n-1}) := \begin{cases} (\Gamma^*(d_0) \setminus \{C\}) \cup (\Gamma^*(d_1) \setminus \{\neg C\}) & \text{if } \mathcal{I} = \mathsf{Cut}_C \\ \emptyset & \text{if } \mathcal{I} = \mathsf{Ind}_F^{y,t} \\ \Delta(\mathcal{I}) & \text{otherwise} \end{cases}$$

Lemma 5.3

a) $\Gamma^*(d) \subseteq \Gamma(d)$.

b) If $\Gamma^*(d) = \emptyset$ then either $d = \mathfrak{a}\{\mathsf{Ind}_F^{y,t} d'\}$ or $d = \mathfrak{a}\{\mathsf{Cut}_C d'd''\}$ with $\Gamma^*(d') = \{C\}$ and $\Gamma^*(d'') = \{\neg C\}$.

c) If $A \in \Gamma^*(d)$ then $d = \mathfrak{a}\{\tilde{d}\}$ with $A \in \Delta(\mathsf{last}(\tilde{d}))$.

Proof:

a) trivial.

b) Since $\Gamma^*(d) = \emptyset$, one of the follwing two cases holds.

1. $d = \mathsf{Ind}_F^{y,t} d'$: o.k.

2. $d = \mathsf{Cut}_A d_0 d_1$ with $\emptyset = (\Gamma^*(d_0) \setminus \{A\}) \cup (\Gamma^*(d_1) \setminus \{\neg A\})$:

2.1. $\Gamma^*(d_0) = \{A\}$ and $\Gamma^*(d_1) = \{\neg A\}$: o.k.

2.2. $\Gamma^*(d_0) = \emptyset$: Then by IH either $d_0 = \mathfrak{a}_0\{\mathsf{Ind}_F^{y,t} d'\}$ or $d_0 = \mathfrak{a}_0\{\mathsf{Cut}_C d'd''\}$ with $\Gamma^*(d') = \{C\}$ and $\Gamma^*(d'') = \{\neg C\}$. Hence the claim holds with $\mathfrak{a} := \mathsf{Cut}_A \mathfrak{a}_0 d_1$.

2.3. $\Gamma^*(d_1) = \emptyset$: anlogous to 2.2.

c) 1. $d = \mathsf{Cut}_C d_0 d_1$ and $\Gamma^*(d) = (\Gamma^*(d_0) \setminus \{C\}) \cup (\Gamma^*(d_1) \setminus \{\neg C\})$: Then w.l.o.g. $A \in \Gamma^*(d_0)$ and by IH $d_0 = \mathfrak{a}_0\{\tilde{d}\}$ with $A \in \Delta(\mathsf{last}(\tilde{d}))$. Therefore the claim holds with $\mathfrak{a} := \mathsf{Cut}_C \mathfrak{a}_0 d_1$.

2. Otherwise we have $\Gamma^*(d) = \Delta(\mathsf{last}(d))$ and the claim holds with $\mathfrak{a} := *$

"Definition"

A (Z-)derivation d is called *closed* if it contains no superfluous free variables, i.e., if every variable $x \notin \mathrm{FV}(\Gamma(d))$ occurring free at some position ν in d is the eigenvariable of some inference symbol below ν.
(A precise definition of this notion will be given in §6.)

Theorem 5.4 (Gentzen)

For each Z-derivation d of the empty sequent $\emptyset$ one can construct another Z-derivation d^- of $\emptyset$ such that $O(d^-) < O(d)$.

Proof:

W.l.o.g. we may assume that d is closed. Otherwise we replace all superfluous free variables by the constant 0. This operation does neither change the endsequent nor the ordinal of d.

Then according to Lemma 5.3 one of the following two cases holds.

Case 1: $d = \mathfrak{a}\{\mathsf{Ind}_F^{y,t} d_0\}$ with $\mathrm{FV}(t) = \emptyset$ (since d is closed).

Let us assume that t is a numeral $\underline{n}$ (for the general case cf. §6).

$$
d^- := \quad
\cfrac{\cfrac{\cfrac{\cfrac{d_0(y/0) \quad d_0(y/1)}{\mathsf{Cut}_{F(1)}} \quad d_0(y/2)}{\mathsf{Cut}_{F(2)}} \quad d_0(y/3)}{\vdots} \quad d_0(y/n{-}1)}{\begin{array}{c}\mathsf{Cut}_{F(n-1)} \\ \vdots \\ \mathfrak{a}\end{array}}
$$

Let ν be the position of $*$ in $\mathfrak{a}$. Then $O_d(\nu) = \omega_m(O_d(\nu 0) \cdot \omega) > \omega_m(O_d(\nu 0) \cdot n) = O_{d^-}(\nu)$ (with $m := \mathrm{rk}(F) \dot{-} \mathrm{hgt}_d(\nu)$). This yields $O(d) > O(d^-)$.

Case 2: $d = \dfrac{\begin{matrix}\tilde{d}_0 \\ \vdots \\ \mathfrak{c}_0\end{matrix} \qquad \begin{matrix}\tilde{d}_1 \\ \vdots \\ \mathfrak{c}_1\end{matrix}}{\begin{matrix}\mathsf{Cut}_C \\ \vdots \\ \mathfrak{a}\end{matrix}}$ with $C \in \Delta(\mathsf{last}(\tilde{d}_0))$ and $\neg C \in \Delta(\mathsf{last}(\tilde{d}_1))$.

Here we have four subcases according to the shape of C. Let us consider the case $C = \forall x A$. Then

$$d = \dfrac{\begin{matrix}\dfrac{d_0}{\bigwedge^{y}_{\forall xA}} \\ \vdots \\ \mathfrak{c}_0\end{matrix} \qquad \begin{matrix}\dfrac{d_1}{\bigvee^{t}_{\exists x\neg A}} \\ \vdots \\ \mathfrak{c}_1\end{matrix}}{\begin{matrix}\mathsf{Cut}_{\forall xA} \\ \vdots \\ \mathfrak{b} \\ \vdots \\ \mathfrak{a}_0\end{matrix}} \qquad \text{and} \qquad d^- := \dfrac{\begin{matrix}\dfrac{\begin{matrix}d_0(y/t) \\ \vdots \\ \mathfrak{c}_0\end{matrix} \qquad \begin{matrix}\dfrac{d_1}{\bigvee^{t}_{\exists x\neg A}} \\ \vdots \\ \mathfrak{c}_1\end{matrix}}{\mathsf{Cut}_{\forall xA}} \\ \vdots \\ \mathfrak{b}\end{matrix} \qquad \begin{matrix}\dfrac{\begin{matrix}\dfrac{d_0}{\bigwedge^{y}_{\forall xA}} \\ \vdots \\ \mathfrak{c}_0\end{matrix} \qquad \begin{matrix}d_1 \\ \vdots \\ \mathfrak{c}_1\end{matrix}}{\mathsf{Cut}_{\forall xA}} \\ \vdots \\ \mathfrak{b}\end{matrix}}{\begin{matrix}\mathsf{Cut}_{A(t)} \\ \vdots \\ \mathfrak{a}_0\end{matrix}}$$

where $\mathfrak{a} = \mathfrak{a}_0\{\mathfrak{b}\}$ such that the root of $\mathfrak{b}$ (i.e. the position of $*$ in $\mathfrak{a}_0$) is the uppermost node ν below the root of $\mathfrak{c}_0$ (or $\mathfrak{c}_1$) such that $\mathrm{hgt}_d(\nu) < \mathrm{hgt}_d(\nu 0)$.

Instead of proving $O(d^-) < O(d)$ in general we treat a sufficiently informative example.

Example

By $: m$ we indicate the height $\mathrm{hgt}_d(\nu)$ of the respective position ν.

$$d = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0}{\bigwedge^{y}_{\forall xA}} \quad c_0}{\mathsf{Cut}_{G_0}:m} \quad \dfrac{\dfrac{d_1}{\bigvee^{t}_{\exists x\neg A}} \quad c_1}{\mathsf{Cut}_{G_1}:m}}{\mathsf{Cut}_{\forall xA}:m} \quad b}{\mathsf{Cut}_F:n} \quad a}{\mathsf{Cut}_D:0} \qquad \text{with } \mathrm{rk}(D) = n < \mathrm{rk}(\forall xA) \leq \mathrm{rk}(F) = m.$$

Then $\mathfrak{b} = \dfrac{* \quad b}{\mathsf{Cut}_F}$, $\mathfrak{a}_0 = \dfrac{* \quad a}{\mathsf{Cut}_D}$ and

$$d^- = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0(y/t) \quad c_0}{\mathsf{Cut}_{G_0}:m} \quad \dfrac{\dfrac{d_1}{\bigvee^{k}_{\exists x\neg A}} \quad c_1}{\mathsf{Cut}_{G_1}:m}}{\mathsf{Cut}_{\forall xA}:m} \quad b}{\mathsf{Cut}_F:k} \quad \dfrac{\dfrac{\dfrac{\dfrac{d_0}{\bigwedge^{y}_{\forall xA}} \quad c_0}{\mathsf{Cut}_{G_0}:m} \quad \dfrac{d_1 \quad c_1}{\mathsf{Cut}_{G_1}:m}}{\mathsf{Cut}_{\forall xA}:m} \quad b}{\mathsf{Cut}_F:k}}{\mathsf{Cut}_{A(t)}:n} \quad a}{\mathsf{Cut}_D:0} \qquad \text{with } k := \max\{n, \mathrm{rk}(A(t))\} < m.$$

$O(d^-) = \omega_n\Big(\omega_{k-n}\big(\omega_{m-k}(\gamma_0^- \# \gamma_1 \# \beta) \# \omega_{m-k}(\gamma_0 \# \gamma_1^- \# \beta)\big) \# \alpha\Big) <$
$\omega_n\Big(\omega_{k-n}\big(\omega_{m-k}(\gamma_0 \# \gamma_1 \# \beta)\big) \# \alpha\Big) = \omega_n\big(\omega_{m-n}(\gamma_0 \# \gamma_1 \# \beta) \# \alpha\big) = O(d)$.

Theorem $\mathrm{PRA} + \mathrm{TI}^{\mathrm{qf}}_{\varepsilon_0} \vdash \mathrm{Con}(\mathbf{Z})$.

Proof:

We assume that $\mathbf{Z}$-derivations and ordinals are coded by natural numbers such that

– $O(d) \in \mathrm{OT}$ for each $d \in Z$ ($Z :=$ set of all (codes of) $\mathbf{Z}$-derivations),

– the function $d \mapsto O(d)$ is primitive recursive,

– there is a primitive recursive function g with $\forall d \in Z(g(d) = 0 \leftrightarrow \Gamma(d) = \emptyset)$.

Abbreviation: $F(y,z) :\equiv z \in Z \wedge O(z) = y \;\to\; g(z) \neq 0, \qquad \tilde{F}(y) :\equiv \forall z F(y,z)$

By $\mathrm{TI}_{\prec}(\tilde{F})$ we prove $\forall y \tilde{F}(y)$, i.e. $\forall a \forall d(d \in Z \wedge O(d) = a \to g(d) \neq 0)$:

Assume $d \in Z \wedge O(d) = a$. Then there is a $d^- \in Z$ with $g(d) = 0 \to O(d^-) \prec O(d) = a \wedge g(d^-) = 0$. By IH we have $O(d^-) \prec a \to g(d^-) \neq 0$. Hence $g(d) = 0 \to \bot$, i.e. $g(d) \neq 0$.

It remains to derive $\mathrm{TI}_{\prec}(\tilde{F})$ from $\mathrm{TI}^{\mathrm{qf}}_{\prec}$.

There are primitive recursive functions q, r, p such that

– $a \in \mathrm{OT} \to\; q(a) \in \mathrm{OT} \wedge o(a) = \omega \cdot o(q(a)) + r(a)$,

– $b \in \mathrm{OT} \wedge k \in \mathbb{N} \;\to\; p(b,k) \in \mathrm{OT} \wedge q(p(b,k)) = b \wedge r(p(b,k)) = k$.

Abb.: $G(x) :\equiv\; x \in \mathrm{OT} \to F(q(x), r(x))$

(*) $\mathrm{Prog}_{\prec}(\tilde{F}) \to \mathrm{Prog}_{\prec}(G)$.

Proof:

Assume (1) $\mathrm{Prog}_{\prec}(\tilde{F})$, (2) $\forall y \prec x G(y)$, (3) $x \in \mathrm{OT}$.

To prove: $\forall y \prec q(x) \forall z F(y,z)$ [from this using $\mathrm{Prog}_{\prec}(\tilde{F})$ we get $\tilde{F}(q(x))$ and then $F(q(x), r(x))$].

Let $b \prec q(x)$ and $k \in \mathbb{N}$ be given. We set $a := p(b,k)$. Then $a \in \mathrm{OT} \wedge q(a) = b \prec q(x) \wedge r(a) = k$.

$a, x \in \mathrm{OT} \wedge q(a) \prec q(x) \;\Rightarrow\; a \prec x \;\stackrel{(2)}{\Rightarrow}\; G(a) \;\Rightarrow\; F(q(a), r(a)) \;\Rightarrow\; F(b,k)$.

Using $\mathrm{TI}^{\mathrm{qf}}_{\prec}(G)$ from (*) we get:

$\mathrm{Prog}_{\prec}(\tilde{F}) \Rightarrow \forall x G(x) \Rightarrow \forall x \in \mathrm{OT}\, F(q(x), r(x)) \Rightarrow \forall y \in \mathrm{OT} \forall z F(y,z) \Rightarrow \forall y \in \mathrm{OT} \tilde{F}(y)$.

Trivially holds $\mathrm{Prog}_{\prec}(\tilde{F}) \to \forall y \notin \mathrm{OT}\, \tilde{F}(y)$.

§6 Notations for infinitary derivations; the proof system Z^*

The proof-system Z^*

The system Z^* results from Z by adding the following inference symbols:

– (R_C) $\dfrac{C \quad \neg C}{\emptyset}$ with $\deg(\mathsf{R}_C) := \mathrm{rk}(C)$,

– (E) $\dfrac{\emptyset}{\emptyset}$ with $\deg(\mathsf{E})$ undefined,

and defining $\mathrm{o}(h)$ and $\deg(h)$ for Z^*-derivations $h = \mathcal{I}h_0...h_{n-1}$ as follows

$$\mathrm{o}(h) := \begin{cases} \mathrm{o}(h_0)\#\mathrm{o}(h_1) & \text{if } \mathcal{I} = \mathsf{R}_C \\ \mathrm{o}(h_0)+\omega & \text{if } \mathcal{I} = \mathsf{Ind}_C^{y,t} \\ 3^{\mathrm{o}(h_0)} & \text{if } \mathcal{I} = \mathsf{E} \\ \sup_{i<n}(\mathrm{o}(h_i){+}1) & \text{otherwise} \end{cases}, \quad \deg(h) := \begin{cases} \deg(h_0)\dot{-}1 & \text{if } \mathcal{I} = \mathsf{E} \\ \max(\{\deg(\mathcal{I})\} \cup \{\deg(h_i) : i < n\}) & \text{otherwise} \end{cases}$$

Note that $\mathrm{o}(h)$, $\deg(h)$ for Z^*-derivations h are defined differently from $\mathrm{o}(d)$, $\deg(d)$ for Z-derivations d.

Remark: The definitions of $\mathrm{o}(h)$ and $\deg(h)$ are motivated by the interpretation $h \mapsto h^\infty$ (introduced below) and Theorems 4.1,4.2. For example, since, according to Theorem 4.2, $\mathrm{o}(\mathcal{E}(h_0^\infty)) \le 3^{\mathrm{o}(h_0^\infty)}$ and $\deg(\mathcal{E}(h_0^\infty)) \le \deg(h_0^\infty)\dot{-}1$ holds, we have defined $\mathrm{o}(\mathsf{E}h_0) := 3^{\mathrm{o}(h_0)}$ and $\deg(\mathsf{E}h_0) := \deg(h_0)\dot{-}1$.

Substitution

We set $\mathsf{R}_C(x/t) := \mathsf{R}_{C_x(t)}$, $\mathsf{E}(x/t) := \mathsf{E}$; otherwise we define $\mathcal{I}(x/t)$ (and $\mathcal{I}^u$) as in Z.

Then Z^* is closed under substitution. In the following we only consider ground substitutions, i.e. substitutions (x/t) with $t \in T$. Then $(\mathcal{I}h_0...h_{n-1})(x/t) = \begin{cases} \mathcal{I}h_0...h_{n-1} & \text{if } \mathrm{Eig}(\mathcal{I}) = \{x\} \\ \mathcal{I}(x/t)h_0(x/t)...h_{n-1}(x/t) & \text{otherwise} \end{cases}$

Lemma 6.1

If h is a Z^*-derivation and $t \in T$ then $h(x/t)$ is a Z^*-derivation with $\Gamma(h(x/t)) \subseteq \Gamma(h)_x(t)$, $\deg(h(x/t)) = \deg(h)$, $\mathrm{o}(h(x/t)) = \mathrm{o}(h)$.

Definition: $\mathrm{FV}(\mathcal{I}) := \begin{cases} \mathrm{FV}(\Delta(\mathcal{I})) \cup \mathrm{FV}(t) & \text{if } \mathcal{I} = \bigvee_{\exists yA}^t \\ \mathrm{FV}(\Delta(\mathcal{I})) & \text{otherwise} \end{cases}$

Remark: $\mathrm{Eig}(\mathcal{I}) \cap \mathrm{FV}(\mathcal{I}) = \emptyset$.

Definition of $\mathrm{FV}(h)$

$\mathrm{FV}(\mathcal{I}h_0...h_{n-1}) := \mathrm{FV}(\mathcal{I}) \cup \bigcup_{i<n}(\mathrm{FV}(h_i) \setminus \mathrm{Eig}(\mathcal{I}))$

Lemma 6.2

a) $\mathrm{FV}(\Gamma(h)) \subseteq \mathrm{FV}(h)$,

b) $\mathrm{FV}(h(x/t)) = \mathrm{FV}(h) \setminus \{x\}$, if $t \in T$.

Proof: Let $h = \mathcal{I}h_0...h_{n-1}$.

a) $\Gamma(h) = \Delta(\mathcal{I}) \cup \bigcup_{i<n}(\Gamma(h_i) \setminus \Delta_i(\mathcal{I}))$ and $\mathrm{FV}(\Gamma(h)) \cap \mathrm{Eig}(\mathcal{I}) = \emptyset$ $(*)$.

$\mathrm{FV}(\Delta(\mathcal{I})) \subseteq \mathrm{FV}(\mathcal{I}) \subseteq \mathrm{FV}(h)$.

$\mathrm{FV}(\Gamma(h_i) \setminus \Delta_i(\mathcal{I})) \overset{(*)}{\subseteq} \mathrm{FV}(\Gamma(h_i)) \setminus \mathrm{Eig}(\mathcal{I}) \overset{\mathrm{IH}}{\subseteq} \mathrm{FV}(h_i) \setminus \mathrm{Eig}(\mathcal{I}) \subseteq \mathrm{FV}(h)$.

b) Abb.: $\mathcal{I}' := \mathcal{I}(x/t)$, $h' := h(x/t)$.

1. $\mathrm{Eig}(\mathcal{I}) = \{x\}$: Then $h' = h$ and $x \notin \mathrm{FV}(h)$. Hence $\mathrm{FV}(h') = \mathrm{FV}(h) = \mathrm{FV}(h) \setminus \{x\}$.

2. Otherwise: Then $h' = \mathcal{I}'h'_0...h'_{n-1}$, and by IH $\mathrm{FV}(h'_i) = \mathrm{FV}(h_i) \setminus \{x\}$.
Moreover one easily verifies that $\mathrm{FV}(\mathcal{I}') = \mathrm{FV}(\mathcal{I}) \setminus \{x\}$.
Hence $\mathrm{FV}(h') = \mathrm{FV}(\mathcal{I}') \cup \bigcup_i (\mathrm{FV}(h'_i) \setminus \mathrm{Eig}(\mathcal{I}')) \overset{\text{IH}}{=}$
$(\mathrm{FV}(\mathcal{I}) \setminus \{x\}) \cup \bigcup_i \big((\mathrm{FV}(h_i) \setminus \{x\}) \setminus \mathrm{Eig}(\mathcal{I})\big) =$
$\big((\mathrm{FV}(\mathcal{I}) \cup \bigcup_i (\mathrm{FV}(h_i) \setminus \mathrm{Eig}(\mathcal{I}))\big) \setminus \{x\} = \mathrm{FV}(h) \setminus \{x\}$.

Definition

A Z^*-derivation h is called *closed* iff $\mathrm{FV}(h) = \emptyset$.

Lemma 6.3

a) Every Z^*-derivation h with closed endsequent can be transformed into a closed Z^*-derivation h' with $\Gamma(h') \subseteq \Gamma(h)$, $\deg(h') = \deg(h)$, $\mathrm{o}(h') = \mathrm{o}(h)$.

b) If $h = \mathcal{I}h_0...h_{n-1}$ is closed and $\mathrm{Eig}(\mathcal{I}) = \emptyset$ then $h_0, ..., h_{n-1}$ are closed.

c) If $h = \mathcal{I}h_0$ is closed and $\mathrm{Eig}(\mathcal{I}) = \{x\}$ then $h_0(x/t)$ is closed for each $t \in T$.

Proof:

a) By induction on the cardinality of $\mathrm{FV}(h)$ we prove that each Z^*-derivation h can be transformed into a Z^*-derivation h' with $\Gamma(h') \subseteq \Gamma(h)$, $\deg(h') = \deg(h)$, $o(h') = o(h)$, and $\mathrm{FV}(\Gamma(h')) = \mathrm{FV}(h')$:
If $\mathrm{FV}(h) \setminus \mathrm{FV}(\Gamma(h)) = \emptyset$ then $h' := h$, and the claim follows by L.6.2a.
Now assume that $x \in \mathrm{FV}(h) \setminus \mathrm{FV}(\Gamma(h))$. Then for $h_1 := h(x/0)$ we have $\Gamma(h_1) \subseteq \Gamma(h)_x(0) = \Gamma(h)$, $\deg(h_1) = \deg(h)$, $o(h_1) = o(h)$, and (by L.6.2b) $\mathrm{FV}(h_1) = \mathrm{FV}(h) \setminus \{x\}$. Hence the claim follows by IH.

b) $\mathrm{FV}(h_i) \subseteq \mathrm{FV}(h) \cup \mathrm{Eig}(\mathcal{I})$.

c) $\mathrm{FV}(h_0) \subseteq \mathrm{FV}(h) \cup \{x\} = \{x\} \;\Rightarrow\; \mathrm{FV}(h_0(x/t)) = \mathrm{FV}(h_0) \setminus \{x\} = \emptyset$.

Interpretation of Z^* in Z^∞

For each closed Z^*-derivation h we define its interpretation $h^\infty \in \mathsf{Z}^\infty$ as follows:

0. $(\mathsf{Ax}_\Delta)^\infty := \mathsf{Ax}_\Delta^\infty$,
1. $(\bigwedge^y_{\forall x A} h_0)^\infty := \bigwedge_{\forall x A}\big(h_0(y/t)^\infty\big)_{t \in T}$,
2. $(\mathsf{R}_C h_0 h_1)^\infty := \mathcal{R}_C(h_0^\infty, h_1^\infty)$,
3. $(\mathsf{E}h_0)^\infty := \mathcal{E}(h_0^\infty)$,
4. $(\mathsf{Ind}^{y,t}_F h_0)^\infty := \begin{cases} \mathsf{Rep}\,\mathsf{c}^{y,t}_F & \text{if } n = 0 \\ \mathsf{Rep}\,\mathsf{Cut}_{F(\underline{n})}\,\mathsf{e}_{n-1}\,\mathsf{c}^{y,t}_F & \text{if } n > 0 \end{cases}$ where

   $n := \mathsf{val}(t)$, $\mathsf{c}^{y,t}_F \vdash^{<\omega}_0 \neg F(\underline{n}), F(t)$, $\mathsf{e}_0 := h_0(y/0)^\infty$, $\mathsf{e}_i := \mathsf{Cut}_{F(\underline{i})}\,\mathsf{e}_{i-1}\,h_0(y/\underline{i})^\infty$ for $i > 0$.

$$(\mathsf{Ind}^{y,t}_F h_0)^\infty = \cfrac{\cfrac{\cfrac{\vdots \qquad h_0(y/n{-}1)^\infty}{\mathsf{Cut}_{F(n-1)}} \qquad \mathsf{c}^{y,t}_F}{\mathsf{Cut}_{F(n)}}}{\mathsf{Rep}} \qquad \text{where } \vdots = \cfrac{\cfrac{h_0(y/0)^\infty \quad h_0(y/1)^\infty}{\mathsf{Cut}_{F(1)}} \quad h_0(y/2)^\infty}{\ \ }$$

5. Otherwise: $(\mathcal{I}h_0...h_{n-1})^\infty := \mathcal{I}h_0^\infty \ldots h_{n-1}^\infty$.

Remark

With the help of Theorems 4.1,4.2 one easily verifies that h^∞ is a Z^∞-derivation with $h^\infty \vdash^{\mathrm{o}(h)}_{\deg(h)} \Gamma(h)$.

Let us look at the Ind-case:

$$(\mathsf{Ind}^{y,t}_F h_0)^\infty \simeq \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{array}{c}h_0(y/0)^\infty\\ |\\ \neg F(0),F(1):\gamma\end{array} \quad \begin{array}{c}h_0(y/1)^\infty\\ |\\ \neg F(1),F(2):\gamma\end{array}}{\neg F(0),F(2):\gamma{+}1} \quad \begin{array}{c}h_0(y/2)^\infty\\ |\\ \neg F(2),F(3):\gamma\end{array}}{\vdots} \quad \begin{array}{c}h_0(y/n{-}1)^\infty\\ |\\ \neg F(n{-}1),F(n):\gamma\end{array}}{\neg F(0),F(n):\gamma{+}n{-}1} \quad \begin{array}{c}\mathsf{c}^{y,t}_F\\ |\\ \neg F(n),F(t):k\end{array}}{\dfrac{\neg F(0),F(t):\gamma{+}\tilde{n}}{\neg F(0),F(t):\gamma{+}\omega}}$$

6.4. Definition of $\mathsf{tp}(h)$ and $h[\iota]$ for closed Z^*-derivations h and $\iota \in |\mathsf{tp}(h)|$

By recursion on the build-up of h we define a Z^∞-inference $\mathsf{tp}(h)$ and closed Z^*-derivation(s) $h[\iota]$ in such a way that

$$h^\infty = \mathsf{tp}(h)\Big(h[\iota]^\infty\Big)_{\iota\in|\mathsf{tp}(h)|} = \frac{\ldots h[\iota]^\infty \ldots (\iota\in|\mathsf{tp}(h)|)}{\mathsf{tp}(h)}$$

The definition clauses for $h = \mathsf{R}_C h_0 h_1$ and $h = \mathsf{E} h_0$ can be read off from the corresponding clauses in the definitions of $\mathcal{R}_C$ and $\mathcal{E}$.

1.1. $h = \mathsf{Ax}_\Delta$: $\mathsf{tp}(h) := \mathsf{Ax}^\infty_\Delta$.

1.2. $h = \bigwedge_C h_0 h_1$: $\mathsf{tp}(h) := \bigwedge_C$, $h[i] := h_i$.

1.3. $h = \bigwedge^y_C h_0$: $\mathsf{tp}(h) := \bigwedge_C$, $h[t] := h_0(y/t)$.

1.4. $h = \bigvee^\iota_C h_0$: $\mathsf{tp}(h) := \bigvee^\iota_C$, $h[0] := h_0$.

2. $h = \mathsf{Ind}^{y,t}_F h_0$: $\mathsf{tp}(h) := \mathsf{Rep}$, $h[0] := \begin{cases} c^{y,t}_F & \text{if } n = 0 \\ \mathsf{Cut}_{F(\underline{n})} e_{n-1} c^{y,t}_F & \text{if } n > 0 \end{cases}$ where

$n := \mathsf{val}(t)$, $\mathsf{Z}^* \ni c^{y,t}_F \vdash^{<\omega}_1 \neg F(\underline{n}), F(t)$, $e_0 := h_0(y/0)$, $e_i := \mathsf{Cut}_{F(\underline{i})} e_{i-1} h_0(y/\underline{i})$ for $i > 0$.

3. $h = \mathsf{E}h_0$:

3.1. $\mathsf{tp}(h_0) = \mathsf{Cut}_C$: $\mathsf{tp}(h) := \mathsf{Rep}$, $h[0] := \mathsf{R}_C \mathsf{E}h_0[0] \mathsf{E}h_0[1]$,

3.2. otherwise: $\mathsf{tp}(h) := \mathsf{tp}(h_0)$, $h[\iota] := \mathsf{E}h_0[\iota]$.

4. $h = \mathsf{R}_C h_0 h_1$:

4.1. $C \notin \Delta(\mathsf{tp}(h_0))$: $\mathsf{tp}(h) := \mathsf{tp}(h_0)$, $h[\iota] := \mathsf{R}_C h_0[\iota] h_1$.

4.2. $\neg C \notin \Delta(\mathsf{tp}(h_1))$: $\mathsf{tp}(h) := \mathsf{tp}(h_1)$, $h[\iota] := \mathsf{R}_C h_0 h_1[\iota]$.

4.3. $C \in \Delta(\mathsf{tp}(h_0))$ and $\neg C \in \Delta(\mathsf{tp}(h_1))$:

4.3.0. $\mathrm{rk}(C) = 0$: $\mathsf{tp}(h) := \mathsf{Ax}^\infty_\Delta$ with $\Delta := (\Delta(\mathsf{tp}(h_0)) \setminus \{C\}) \cup (\Delta(\mathsf{tp}(h_1)) \setminus \{\neg C\})$.

4.3.1. $C = \forall x A$: Then $\mathsf{tp}(h_1) = \bigvee^t_{\neg C}$ for some $t \in T$.

$\mathsf{tp}(h) := \mathsf{Cut}_{A_x(t)}$, $h[0] := \mathsf{R}_C h_0[t] h_1$, $h[1] := \mathsf{R}_C h_0 h_1[0]$.

4.3.2. $C = \exists x A$ or $A_0 \wedge A_1$ or $A_0 \vee A_1$: analogous to 4.3.1.

Theorem 6.5

If h is a closed Z^*-derivation with $\mathcal{I} = \mathsf{tp}(h)$ and $h \vdash^{\alpha}_{m} \Gamma$, then the following holds:

a) $\Delta(\mathcal{I}) \subseteq \Gamma$,

b) $\deg(\mathcal{I}) \leq m$, i.e. $(\mathcal{I} = \mathsf{Cut}_C \Rightarrow \mathrm{rk}(C) < m)$,

c) For each $\iota \in |\mathcal{I}|$: $h[\iota] \vdash^{\alpha_\iota}_{m} \Gamma, \Delta_\iota(\mathcal{I})$ with $\alpha_\iota < \alpha$.

Proof by straightforward induction on the build-up of h:

W.l.o.g. $\mathrm{FV}(\Gamma) = \emptyset$.

1. $h = \mathsf{Ax}_\Delta$: Then $\mathcal{I} = \mathsf{Ax}^{\infty}_{\Delta}$ and thus $\Delta(\mathcal{I}) = \Delta = \Gamma(h) \subseteq \Gamma$ and $|\mathcal{I}| = \emptyset$.

2.1. $h = \bigwedge^{y}_{\forall xA} h_0$: Then $\mathcal{I} = \bigwedge_{\forall xA}$, $\Delta(\mathcal{I}) = \{\forall xA\} \subseteq \Gamma$ and $h_0 \vdash^{\alpha_0}_{m} \Gamma, A_x(y)$ with $\alpha_0 < \alpha$.

By ... we get $h[t] = h_0(y/t) \vdash^{\gamma}_{m} \Gamma, A_x(t)$ for each $t \in T$.

2.2. $h = \bigvee^{t}_{\exists xA} h_0$: Then $t \in T$, $\mathcal{I} = \bigvee^{t}_{\exists xA}$, $\Delta(\mathcal{I}) = \{\exists xA\} \subseteq \Gamma$ and $h_0 \vdash^{\alpha_0}_{m} \Gamma, A_x(t)$ with $\alpha_0 < \alpha$.

2.3. $h = \bigwedge_{A_0 \wedge A_1} h_0 h_1$ or $h = \bigvee^{k}_{A_0 \vee A_1}$: analogous to 2.1 and 2.2.

3. $h = \mathsf{Cut}_C h_0 h_1$: Then $\mathcal{I} = \mathsf{Cut}_C$, $\mathrm{rk}(C) < m$, and $h[0] = h_0 \vdash^{\alpha_0}_{m} \Gamma, C$, $h[1] = h_1 \vdash^{\alpha_1}_{m} \Gamma, \neg C$ with $\alpha_0, \alpha_1 < \alpha$.

4.1. $h = \mathsf{R}_C h_0 h_1$ with $C = \forall xA$, $\mathsf{tp}(h_0) = \bigwedge_C$, $\mathsf{tp}(h_1) = \bigvee^{t}_{\neg C}$: Then $t \in T$ and $\mathsf{tp}(h) = \mathsf{Cut}_{A(t)}$.

Let $\gamma := \mathrm{o}(h_0)$, $\beta := \mathrm{o}(h_1)$.

Then $h_0 \vdash^{\gamma}_{m} \Gamma, C$, $h_1 \vdash^{\beta}_{m} \Gamma, \neg C$ and $\mathrm{rk}(A(t)) < \mathrm{rk}(C) \leq m$.

By IH we obtain $h_0[t] \vdash^{\gamma_t}_{m} \Gamma, C, A(t)$ with $\gamma_t < \gamma$, and $h_1[0] \vdash^{\beta_0} \Gamma, \neg C, \neg A(t)$ with $\beta_0 < \beta$.

Hence $h[0] = \mathsf{R}_C h_0[t] h_1 \vdash^{\gamma_t \# \beta}_{m} \Gamma, A(t)$ and $h[1] = \mathsf{R}_C h_0 h_1[0] \vdash^{\gamma \# \beta_0}_{m} \Gamma, \neg A(t)$

with $\gamma_t \# \beta$, $\gamma \# \beta_0 < \gamma \# \beta = \mathrm{o}(h) \leq \alpha$.

4.2. $h = \mathsf{R}_C h_0 h_1$ with $\mathsf{tp}(h_0) = \mathcal{I}$ and $C \notin \Delta(\mathcal{I})$: Let $\gamma := \mathrm{o}(h_0)$, $\beta := \mathrm{o}(h_1)$.

Then $h_0 \vdash^{\gamma}_{m} \Gamma, C$, $h_1 \vdash^{\beta}_{m} \Gamma, \neg C$.

By IH we obtain $h_0[\iota] \vdash^{\gamma_\iota}_{m} \Gamma, C, \Delta_\iota(\mathcal{I})$ with $\gamma_\iota < \gamma$, for all $\iota \in |\mathcal{I}|$.

Hence $h[\iota] = \mathsf{R}_C h_0[\iota] h_1 \vdash^{\gamma_\iota \# \beta}_{m} \Gamma, \Delta_\iota(\mathcal{I})$ with $\gamma_\iota \# \beta < \gamma \# \beta = \mathrm{o}(h) \leq \alpha$.

$h_0 \vdash^{\gamma}_{m} \Gamma, C \;\&\; \mathsf{tp}(h_0) = \mathcal{I} \;\&\; C \notin \Delta(\mathcal{I}) \overset{\mathrm{IH}}{\Rightarrow} \Delta(\mathcal{I}) \subseteq \Gamma \;\&\; (\mathcal{I} = \mathsf{Cut}_A \Rightarrow \mathrm{rk}(A) < m)$.

4.3. $h = \mathsf{R}_C h_0 h_1$ with $\mathrm{rk}(C) = 0$ and $C \in \Delta^0 := \Delta(\mathsf{tp}(h_0))$, $\neg C \in \Delta^1 := \Delta(\mathsf{tp}(h_1))$:

Then $\mathcal{I} = \mathsf{Ax}_\Delta$ with $\Delta := (\Delta^0 \setminus \{C\}) \cup (\Delta^0 \setminus \{\neg C\})$, and by IH $\Delta^i \subseteq \Gamma(h_i)$. Hence $\Delta(\mathcal{I}) = \Delta \subseteq \Gamma(h) \subseteq \Gamma$.

5. $h = \mathsf{E} h_0$ with $\mathsf{tp}(h_0) = \mathsf{Cut}_C$: Then $\mathsf{tp}(h) = \mathsf{Rep}$, $h[0] = \mathsf{R}_C \mathsf{E} h_0[0] \mathsf{E} h_0[1]$ and $\deg(h_0) \leq m{+}1$.

Let $\gamma := \mathrm{o}(h_0)$. Then $h_0 \vdash^{\gamma}_{m+1} \Gamma$.

By IH we have $\mathrm{rk}(C) < m{+}1$ and $h_0[0] \vdash^{\gamma_0}_{m+1} \Gamma, C$, $h_0[1] \vdash^{\gamma_1}_{m+1} \Gamma, \neg C$ with $\gamma_0, \gamma_1 < \gamma$.

Hence $\mathsf{E} h_0[0] \vdash^{3^{\gamma_0}}_{m} \Gamma, C$ and $\mathsf{E} h_0[1] \vdash^{3^{\gamma_1}}_{m} \Gamma, \neg C$.

From this together with $\mathrm{rk}(C) \leq m$ we get $h[0] = \mathsf{R}_C \mathsf{E} h_0[0] \mathsf{E} h_0[1] \vdash^{3^{\gamma_1} \# 3^{\gamma_1}}_{m} \Gamma$ and $3^{\gamma_0} \# 3^{\gamma_1} < 3^{\gamma} = \mathrm{o}(h) \leq \alpha$.

6. $h = \mathsf{Ind}^{y,t}_{F} h_0$: Then $\mathcal{I} = \mathsf{Rep}$, $\mathrm{rk}(F) < m$, and $h_0 \vdash^{\gamma}_{m} \Gamma, \neg F(y), F(\mathsf{S}y)$ with $\gamma := \mathrm{o}(h_0)$, $\gamma + \omega \leq \alpha$.

$$h[0] = \cfrac{\cfrac{\cfrac{\cfrac{\cfrac{h_0(y/0) \quad h_0(y/1)}{\mathsf{Cut}_{F(1)}} \quad h_0(y/2)}{\vdots} \quad h_0(y/n{-}1)}{\mathsf{Cut}_{F(n-1)}} \quad c^{y,t}_{F}}{\mathsf{Cut}_{F(n)}} \qquad \text{with } c^{y,t}_{F} \vdash^{<\omega}_{1} \neg F(n), F(t).$$

Lemma 6.6 (Consistency of Z)

Let $\mathsf{Z}^*_\perp$ be the set of all closed Z^*-derivations h with $\Gamma(h) = \emptyset$ & $\deg(h) = 0$.

a) $h \in \mathsf{Z}^*_\perp \;\Rightarrow\; h[0] \in \mathsf{Z}^*_\perp$ & $\mathrm{o}(h[0]) < \mathrm{o}(h)$,

b) There is no Z-derivation d with $\Gamma(d) = \emptyset$.

Proof:

a) $h \in \mathsf{Z}^*_\perp \overset{6.5}{\Rightarrow} \Delta(\mathsf{tp}(h)) \subseteq \Gamma(h) = \emptyset$ & $\deg(\mathsf{tp}(h)) = 0 \;\Rightarrow\; \mathsf{tp}(h) = \mathsf{Rep}$.

$h \in \mathsf{Z}^*_\perp$ & $\mathsf{tp}(d) = \mathsf{Rep} \overset{6.5}{\Rightarrow} h[0] \in \mathsf{Z}^*_\perp$ & $\mathrm{o}(h[0]) < \mathrm{o}(h)$.

b) By transfinite induction up to ε_0 from a) we get $\mathsf{Z}^*_\perp = \emptyset$. Now assume that d is a Z-derivation with $\Gamma(d) = \emptyset$. W.l.o.g. we may assume that d is closed. Let $m := \deg(d)$. Then $\mathsf{E}^m d = \mathsf{E}...\mathsf{E}d \in \mathsf{Z}^*_\perp$. *Contradiction.*

Definition

$\beta_0(\alpha) := \alpha$, $\beta_{m+1}(\alpha) := \beta^{\beta_m(\alpha)}$. $\omega_m := \omega_m(1)$, i.e. $\omega_0 = 1$, $\omega_1 = \omega$, $\omega_2 = \omega^\omega$, $\omega_3 = \omega^{\omega^\omega}$ $(= \omega^{(\omega^\omega)})$, ...

Corollary

a) $\omega \le \alpha \;\Rightarrow\; 3_m(\omega^\alpha) = \omega_m(\omega^\alpha)$,

b) $m > 0 \Rightarrow 3_m(\omega^2) = \omega_{m+1}$.

Proof: a) $3_{m+1}(\omega^\alpha) = 3_m(3^{\omega^\alpha}) = 3_m(3^{\omega\cdot\omega^\alpha}) = 3_m((3^\omega)^{\omega^\alpha}) = 3_m(\omega^{\omega^\alpha}) \overset{\mathrm{IH}}{=} \omega_m(\omega^{\omega^\alpha}) = \omega_{m+1}(\omega^\alpha)$.

b) $3_m(\omega^2) = 3_{m-1}(3^{\omega\cdot\omega}) = 3_{m-1}(\omega^\omega) \overset{\mathrm{a)}}{=} \omega_{m-1}(\omega^\omega) = \omega_{m+1}$.

Let $\mathrm{OT} \subseteq \mathbb{N}$ and $\prec \subseteq \mathbb{N} \times \mathbb{N}$ be the primitive recursive relations defined in §4.

$(\mathrm{OT}, \prec)$ has ordertype ε_0. 0 is the $\prec$-least element of OT.

For $\alpha < \varepsilon_0$ let $\ulcorner\alpha\urcorner$ be the corresponding element of OT.

Theorem 6.7

If $\mathcal{Z}_m \vdash \forall x \exists y A(x,y)$ (A atomic) then there are primitive recursive functions $g : \mathbb{N}^2 \to \mathbb{N}$, $\theta : \mathbb{N}^2 \to \mathbb{N}$ and an $\alpha < \max\{\omega^2, \omega_{m+1}\}$ such that $\forall n(\theta(n,0) = \ulcorner\alpha\urcorner)$ and $\mathbb{N} \models \forall x A(x, f(x))$,

where $f(n) := g(n, \min\{k : \theta(n,k{+}1) \not\prec \theta(n,k)\})$.

Proof:

We assume a canonical arithmetization (coding) $q \mapsto \ulcorner q \urcorner$ of syntax (terms, formulas, sequents, finite derivations, derivation terms etc.). A set M of syntactical objects is called primitive recursive if the set $\{\ulcorner q\urcorner : q \in M\}$ is primitive recursive. An operation (function) Φ on syntactical objects or ordinals $< \varepsilon_0$ is called primitive recursive if there is a primitive recursive function $f : \mathbb{N}^n \to \mathbb{N}$ such that $f(\ulcorner q_1\urcorner, ..., \ulcorner q_n\urcorner) = \ulcorner\Phi(q_1, ..., q_n)\urcorner$ for all $(q_1, ..., q_n)$ in the domain of Φ.

By Theorem 5.2 there is a Z-derivation d of $\exists y A(x,y)$ such that $\mathrm{rk}(C) \le m$ for each Cut_C or $\mathsf{Ind}_C^{y,t}$ occurring in d. So we have $\mathsf{Z}^* \ni d \vdash^{\mathrm{o}(d)}_{m+1} \exists y A(x,y)$ with $\mathrm{o}(d) < \omega^2$.

This d will be fixed for the whole proof.

W.l.o.g. we may assume that $\mathrm{FV}(d) \subseteq \{x\}$. Hence $d(n) := d(x/\underline{n})$ is closed for any $n \in \mathbb{N}$.

Let $\mathcal{L}_0^-$ be a *finite* subset of $\mathcal{L}_0$ (= PR), such that

(i) $0, \mathsf{S} \in \mathcal{L}_0^-$ and each function symbol occurring in d belongs to $\mathcal{L}_0^-$,

(ii) with $p \in \mathcal{L}_0^-$ also each function symbol occurring in (the definition of) p belongs to $\mathcal{L}_0^-$.

Let $\widetilde{\mathsf{Z}}^*$ be the restriction of Z^* to $\mathcal{L}_0^-(\mathcal{X})$.

It is wellknown that the set TRUE_0^- of all true $\mathcal{L}_0^-$-literals is primitive recursive, and that there is a primitive recursive function which for any two $\mathcal{L}_0^-$-terms s, t of equal value computes a $\widetilde{\mathsf{Z}}^*$-derivation c of $s{=}t$ with $\mathrm{o}(c) < \omega$ and $\deg(c) \leq 1$. Obviously the functions $\Gamma(\cdot)$, $\mathrm{o}(\cdot)$, $\deg(\cdot)$, $\mathsf{tp}(\cdot)$ are primitive recursive, and $\cdot[\cdot]$ restricted to $\widetilde{\mathsf{Z}}^*$ is primitive recursive too.

Let $\mathbf{D}$ be the set of all closed $\widetilde{\mathsf{Z}}^*$-derivations.

Definition of the primitive recursive function red : $\mathbf{D} \cup \{0\} \to \mathbf{D} \cup \{0\}$:

$$\mathrm{red}(h) := \begin{cases} 0 & \text{if } h = 0 \text{ or } \mathsf{tp}(h) = \mathsf{Ax}_\Delta^\infty \text{ or } \mathsf{tp}(h) = \bigvee_{\exists yB}^s \text{ with } B_y(s) \in \mathsf{TRUE}_0^- \\ h[1] & \text{if } \mathsf{tp}(h) = \mathsf{Cut}_C \text{ with } C \in \mathsf{TRUE}_0^- \text{ or } C = \neg Xt \\ h[0] & \text{otherwise} \end{cases}$$

Definition: $h(n,k) := \mathrm{red}^{(k)}(\underbrace{\mathsf{E}...\mathsf{E}}_{m}\, d(n))$, $\alpha := 3_m(\mathrm{o}(d))$.

Then $\deg(h(n,0)) = \deg(d) \dot{-} m \leq (m{+}1) \dot{-} m = 1$, and $\mathrm{o}(h(n,0)) = \alpha < 3_m(\omega^2) = \max\{\omega^2, \omega_{m+1}\}$ for all n.

Proposition: If $h(n,k) \neq 0$ then

a) $\mathrm{o}(h(n,k)) \leq \alpha$,

b) $\deg(h(n,k)) \leq 1$,

c) $\Gamma(h(n,k)) \subseteq \{\exists y A(n,y)\} \cup \{B : \neg B \in \mathsf{TRUE}_0^-\} \cup \{Xt : t \in T\}$

Proof by induction on k:

a),b) are obvious, since $\mathrm{o}(h[i]) < \mathrm{o}(h)$, $\deg(h[i]) \leq \deg(h)$, $\mathrm{o}(h(n,0)) = \alpha$, $\deg(h(n,0)) \leq 1$.

c) $k = 0$: $\Gamma(h(n,0)) = \Gamma(\mathsf{d}(n)) = \{\exists y A(n,y)\}$.

$k > 0$: IH $\Rightarrow$ $\Gamma(h(n,k{-}1)) \subseteq \{\exists y A(n,y)\} \cup \{B : \neg B \in \mathsf{TRUE}_0^-\} \cup \{Xt : t \in T\}$

$\stackrel{h(n,k)\neq 0}{\Rightarrow}$ $\mathsf{tp}(h(n,k{-}1)) = \bigvee_{\exists y A(n,y)}^t$ with $\neg A(n,t) \in \mathsf{TRUE}_0^-$ or $= \mathsf{Rep}$ or $= \mathsf{Cut}_C$ with $\mathrm{rk}(C) = 0$

$\Rightarrow$ $\Gamma(h(n,k)) \subseteq \Gamma(h(n,k{-}1)) \cup \{B : \neg B \in \mathsf{TRUE}_0^-\} \cup \{Xt : t \in T\}$ $\stackrel{\mathrm{IH}}{\Rightarrow}$ claim.

Definition:

$\theta(n,k) := \ulcorner \mathrm{o}(h(n,k)) \urcorner$ (where $\mathrm{o}(0) := 0$)

$$g(n,k) := \begin{cases} \mathsf{val}(t) & \text{if } k > 0 \text{ and } \mathsf{tp}(h(n,k{-}1)) = (\bigvee_{\dots}^t) \\ 0 & \text{otherwise} \end{cases}$$

$f(n) := g(n, \min\{k : \theta(n,k{+}1) \not\prec \theta(n,k)\})$.

Now let k be the least number such that $\theta(n,k{+}1) \not\prec \theta(n,k)$.

Assumption: $h(n,k) \neq 0$. Then [by Prop.c)] $\mathsf{tp}(h(n,k)) \neq \mathsf{Ax}$ and thus $\theta(n,k{+}1) = \mathrm{o}(h(n,k{+}1)) \prec \mathrm{o}(h(n,k)) = \theta(n,k)$. Contradiction.

Hence $h(n,k) = 0$ and thus $k > 0$ and [by Definition of red] $\mathsf{tp}(h(n,k{-}1)) = \mathsf{Ax}$ or $\mathsf{tp}(h(n,k{-}1)) = \bigvee_{\exists y B(y)}^l$ with $B(l) \in \mathsf{TRUE}_0^-$. By Proposition c) and Theorem 6.5a from this we get $\mathsf{tp}(h(n,k{-}1)) = \bigvee_{\exists y A(n,y)}^t$ with $A(n,t) \in \mathsf{TRUE}_0^-$. Hence $f(n) = g(n,k) = \mathsf{val}(t)$ and $\mathbb{N} \models A(n, f(n))$. qed

Definition (Fundamental Sequences for ordinals $< \varepsilon_0$)

1. $0[n] := 1[n] := 0$.
2. $\omega^{\alpha+1}[n] := \omega^{\alpha} \cdot (n+1)$.
3. $\omega^{\lambda}[n] := \omega^{\lambda[n]}$, für $\lambda \in Lim$.
4. $\alpha[n] := \alpha_0 + ... + \alpha_{k-1} + \alpha_k[n]$, falls $\alpha =_{NF} \alpha_0 + ... + \alpha_k$.

Proposition: $(\alpha+1)[n] = \alpha$.

Definition $N\alpha := N\alpha_1 + \ldots + N\alpha_k + k$ if $\alpha = \omega^{\alpha_1} + \ldots \omega^{\alpha_k}$ with $k \geq 0$ and $\alpha_k \leq \ldots \leq \alpha_1 < \varepsilon_0$

Lemma 6.8

a) $\alpha \in Lim \;\Rightarrow\; \forall n(\alpha[n] < \alpha[n+1])$ & $\alpha = \sup\{\alpha[n] : n \in \mathbb{N}\}$

b) $\alpha > 0 \;\Rightarrow\; N\alpha[0] < N\alpha$

c) $\alpha[n] < \beta < \alpha \;\Rightarrow\; \alpha[n] \leq \beta[0]$

d) $\alpha[n] < \beta < \alpha \;\Rightarrow\; N\alpha[n] < N\beta$

e) $\beta < \alpha \;\Rightarrow\; \beta \leq \alpha[N\beta]$

Proof:

a),b) obvious.

c) Let $\beta =_{NF} \beta_0 + ... + \beta_k$.

1. Assume $\omega^{\alpha} \cdot (n+1) < \beta < \omega^{\alpha+1}$. Then $k > n$ and $\beta_0 = ... = \beta_n = \omega^{\alpha}$.
From this we get $\omega^{\alpha} \cdot (n+1) \leq \beta_0 + ... + \beta_{k-1} + \beta_k[0] = \beta[0]$.

2. Assume $\omega^{\lambda[n]} < \beta < \omega^{\lambda}$ und $\lambda \in Lim$. Then $\omega^{\lambda[n]} \leq \beta_0 = \omega^{\gamma} < \omega^{\lambda}$.
If $k = 0$ then $\lambda[n] < \gamma < \lambda$ and therefore (by IH) $\lambda[n] \leq \gamma[0]$. Hence $\omega^{\lambda[n]} \leq \omega^{\gamma[0]} = \omega^{\gamma}[0] = \beta[0]$.
If $k > 0$ then $\omega^{\lambda[n]} \leq \beta_0 + ... + \beta_{k-1} + \beta_k[0] = \beta[0]$.

3. Assume $\alpha =_{NF} \alpha_0 + ... + \alpha_m$, $m > 0$ and $\alpha[n] = \alpha_0 + ... + \alpha_{m-1} + \alpha_m[n] < \beta < \alpha$. Then $m \leq k$, $\alpha_m[n] < \beta_m < \alpha_m$ and $\alpha_i = \beta_i$ for $i < m$. By IH we get $\alpha_m[n] \leq \beta_m[0]$ and then $\alpha[n] \leq \beta_0 + ... + \beta_{m-1} + \beta_m[0] \leq \beta_0 + ... + \beta_{k-1} + \beta_k[0] = \beta[0]$.

d) By c) we have $\alpha[n] = \beta[0]...[0]$. Hence $N\alpha[n] \leq N\beta[0] < N\beta$.

e) Let $\alpha \in Lim$. According to a),d) we then have $\forall n(N\alpha[n] < N\alpha[n+1])$, and therefore $N\beta \leq N\alpha[N\beta]$. Together with d) this yields the assertion.

Definition of $H_\alpha : \mathbb{N} \to \mathbb{N}$ for $\alpha < \varepsilon_0$

1. $H_0(n) := n$,
2. $H_\alpha(n) := H_{\alpha[n]}(n+1)$, for $\alpha > 0$.

Lemma 6.9

a) $H_\alpha(n) < H_\alpha(n+1)$,

b) $\beta[m] < \alpha < \beta \;\Rightarrow\; H_{\beta[m]}(n+1) \leq H_\alpha(n)$,

c) $\beta < \alpha$ & $N\beta \leq n \;\Rightarrow\; H_\beta(n) < H_\alpha(n)$,

d) $\alpha > 0 \;\Rightarrow\; H_\alpha(n) = \min\{k \geq n : \alpha[n]...[k-1] = 0\} = n + \min\{l : \alpha[n][n+1]\ldots[n+l-1] = 0\}$.

Proof:

a),b) simultaneous induction on α: Let $\alpha > 0$.

a) 1. $\alpha \in Lim$: $H_\alpha(n) = H_{\alpha[n]}(n{+}1) \overset{\text{IHa}}{<} H_{\alpha[n]}(n{+}3) \overset{\text{IHb}}{\leq} H_{\alpha[n+1]}(n{+}2) = H_\alpha(n{+}1)$.

2. $\alpha = \alpha_0{+}1$: $H_\alpha(n) = H_{\alpha_0}(n{+}1) \overset{\text{IHa}}{<} H_{\alpha_0}(n{+}2) = H_\alpha(n{+}1)$.

b) From $\beta[m] < \alpha < \beta$ we obtain $\beta[m] \leq \alpha[n] < \beta$ by Lemma 6.8.

Hence $H_{\beta[m]}(n) \overset{\text{IHb}}{\leq} H_{\alpha[n]}(n) \overset{\text{IHa}}{<} H_{\alpha[n]}(n{+}1) = H_\alpha(n)$.

c) Induction on α: $\beta < \alpha \overset{\text{L.6.8e}}{\Rightarrow} \beta \leq \alpha[N\beta] \leq \alpha[n] \overset{\text{a)+IH}}{\Rightarrow} H_\beta(n) < H_\beta(n{+}1) \leq H_{\alpha[n]}(n{+}1) = H_\alpha(n)$.

d) Let $k \geq n$ minimal such that $\alpha[n]...[k{-}1] = 0$.

Then $H_\alpha(n) = H_{\alpha[n]}(n{+}1) = ... = H_{\alpha[n]...[k-1]}(k) = H_0(k) = k$.

Abbreviation: $NF(\alpha,\beta)$ $:\Leftrightarrow$ $\alpha = 0$ or $\beta = 0$ or $[\,\alpha = \omega^{\alpha_0} + ... + \omega^{\alpha_n}$ & $\beta = \omega^{\beta_0} + ... + \omega^{\beta_m}$ with $\alpha_0 \geq ... \geq \alpha_n \geq \beta_0 \geq ... \geq \beta_m\,]$.

Proposition: $NF(\alpha,\beta)$ & $\beta > 0 \Rightarrow (\alpha+\beta)[n] = \alpha + \beta[n]$ & $NF(\alpha,\beta[n])$.

Lemma 6.10

a) $NF(\alpha,\beta) \Rightarrow H_{\alpha+\beta} = H_\alpha \circ H_\beta$.

b) $H_{\omega^{\alpha+1}}(n) = H_{\omega^\alpha}^{(n+1)}(n{+}1)$ and $H_{\omega^\lambda}(n) = H_{\omega^{\lambda[n]}}(n{+}1)$ for $\lambda \in Lim$.

c) For each primitive recursive function f there exists a $k \in \mathbb{N}$ such that $\forall\vec{x}(\,f(\vec{x}) < H_{\omega^k}(\max\{\vec{x}\})\,)$.

Proof:

a) Induction on β: 1. $H_{\alpha+0}(n) = H_\alpha(n) = H_\alpha(H_0(n))$.

2. $\beta > 0$: $H_{\alpha+\beta}(n) = H_{(\alpha+\beta)[n]}(n{+}1) = H_{\alpha+\beta[n]}(n+1) \overset{\text{IH}}{=} H_\alpha(H_{\beta[n]}(n+1)) = H_\alpha(H_\beta(n))$.

b) $H_{\omega^{\alpha+1}}(n) = H_{\omega^\alpha\cdot(n+1)}(n{+}1) \overset{\text{a)}}{=} H_{\omega^\alpha}^{(n+1)}(n{+}1)$.

c) From b) it follows that $(k,n) \mapsto H_{\omega^k}(n)$ is (a variant of) the Ackermann function.

Theorem 6.11

Let $\theta : \mathbb{N} \times \mathbb{N} \to \mathrm{OT}$ be primitive recursive, $m \geq 1$, and $\alpha < \omega_{m+1}$ such that $\forall n(\,\theta(n,0) \preceq \ulcorner\alpha\urcorner\,)$.

Then there is an $\tilde{\alpha} < \omega_{m+1}$ such that $\min\{l : \theta(n,l{+}1) \not\prec \theta(n,l)\} < H_{\tilde{\alpha}}(n)$ $(\forall n)$.

Proof:

W.l.o.g. $\forall n(\,\theta(n,0) = \ulcorner\alpha\urcorner\,)$.

Let $w(i,n,l) := N(\omega^i \cdot (\|\theta(n,l{+}1)\|{+}1))$, where $\|\cdot\|$ is the isomorphism from $(\mathrm{OT},\prec)$ onto $(\varepsilon_0,<)$.

One easily sees that w is primitive recursive.

Let $g(i,n,l) := \max\{w(i,n,l),i,n,l\}$.

There exists a $k \geq 1$ such that $g(i,n,l{+}1) < H_{\omega^k}(\max\{i,n,l\})$ and $g(i,n,0) < H_{\omega^k}(\max\{i,n\})$ $(\forall i,n,l)$.

Then we have

(1) $g(k,n,l{+}1) < H_{\omega^k}(g(k,n,l))$ $(\forall n,l)$.

Abbreviation: $\varphi(n,l) := H_{\omega^k\cdot\|\theta(n,l)\|}(g(k,n,l))$.

(2) $\theta(n,l{+}1) \prec \theta(n,l) \Rightarrow \varphi(n,l{+}1) < \varphi(n,l)$.

Proof: $H_{\omega^k\cdot\|\theta(n,l+1)\|}(g(k,n,l+1)) \overset{(1)}{<} H_{\omega^k\cdot\|\theta(n,l+1)\|}H_{\omega^k}(g(k,n,l)) = H_{\omega^k\cdot(\|\theta(n,l+1)\|+1)}(g(k,n,l)) \overset{(*)}{\leq}$
$\leq H_{\omega^k\cdot\|\theta(n,l)\|}(g(k,n,l)) = \varphi(n,l)$.

$(*)$ $\omega^k\cdot(\|\theta(n,l+1)\|+1) \leq \omega^k\cdot\|\theta(n,l)\|$ and $N(\omega^k\cdot(\|\theta(n,l+1)\|+1)) = w(k,n,l) \leq g(k,n,l)$.

(3) $\exists l \leq \varphi(n,0)(\,\varphi(n,l+1) \not< \varphi(n,l)\,)$.

Proof: $[\forall l \leq j(\,\varphi(n,l+1) < \varphi(n,l)\,) \Rightarrow j < \varphi(n,0)]$ and therefore *not* $\forall l \leq \varphi(n,0)(\,\varphi(n,l+1) < \varphi(n,l)\,)$.

(2) & (3) & $\alpha = \|\theta(n,0)\| \;\Rightarrow\; \exists\, l \leq H_{\omega^k\cdot\alpha}(g(k,n,0))[\,\theta(n,l+1) \not\prec \theta(n,l)\,]$.

$H_{\omega^k\cdot\alpha}(g(k,n,0)) < H_{\omega^k\cdot\alpha}H_{\omega^k}(\max\{k,n\}) \leq H_{\omega^k\cdot(\alpha+1)+k}(n)$.

$m \geq 1 \Rightarrow \omega \leq \omega_m$, $\quad \alpha < \omega_{m+1} = \omega^{\omega_m} \;\Rightarrow\; \omega^k\cdot(\alpha+1)+k < \omega^k\cdot(\alpha+2) < \omega^k\cdot\omega^{\omega_m} = \omega^{\omega_m} = \omega_{m+1}$.

Theorem 6.12

If $\mathcal{Z}_m \vdash \forall x\exists y A(x,y)$ (A atomic, $m \geq 1$) then there is an $\alpha < \omega_{m+1}$ such that $\forall n\exists l < H_\alpha(n)\,\mathbb{N} \models A(n,l)$.

Proof:

By Theorem 6.7 there are primitive recursive functions g,θ and an $\alpha_0 < \omega_{m+1}$ such that $\forall n(\,\|\theta(n,0)\| = \alpha_0\,)$ and $\forall n(\,\mathbb{N} \models A(n,f(n))\,)$, where $f(n) := g(n,f^*(n))$, $f^*(n) := \min\{l : \theta(n,l+1) \not\prec \theta(n,l)\}$).

By 6.11 there exists $\beta < \omega_{m+1}$ with $f^*(n) < H_\beta(n)$. Further there exists $k < \omega$ with $\forall n,i(\,g(n,i) < H_{\omega^k}(\max\{n,i\})\,)$, hence $f(n) < H_{\omega^k}(\max\{n,f^*(n)\}) \leq H_{\omega^k}H_\beta(n)$. Since $\omega_{m+1} = \sup_{i\in\mathbb{N}}\omega_m^i$, there is $\gamma := \omega_m^i$ such that $\omega^k,\beta < \gamma$.

It follows that there is an n_0 such that $H_{\omega^k}H_\beta(n) \leq H_\gamma H_\gamma(n) = H_{\gamma+\gamma}(n)$ for all $n \geq n_0$. Hence $f(n) < H_{\omega^k}H_\beta(n_0{+}n) \leq H_{\gamma+\gamma}(n_0{+}n) = H_{\gamma+\gamma+n_0}(n)$ (and $\gamma{+}\gamma{+}n_0 < \omega_{m+1}$).

Corollary: $\mathcal{Z}_m \nvdash$ '$\forall n\exists l(\,\omega_{m+1}[n][n{+}1]...[l] = 0\,)$'.

Proof: Assumption: $\mathcal{Z}_m \vdash$ '$\forall n\exists l(....)$'.

Then there is an $\alpha < \omega_{m+1}$ such that $\forall n\exists l < H_\alpha(n)(\,\omega_{m+1}[n][n{+}1]...[l] = 0\,)$, i.e. $\forall n\exists l < H_\alpha(n)(\,H_{\omega_{m+1}}(n) \leq l{+}1\,)$. This implies $\forall n(\,H_{\omega_{m+1}}(n) \leq H_\alpha(n)\,)$. But by L.6.9c we have $\forall n \geq N(\alpha)(\,H_\alpha(n) < H_{\omega_{m+1}}(n)\,)$. Contradiction.

Disgression

Let $h_\alpha(n) := n$, $h_{\alpha+1}(n) := h_\alpha(n+1)$, and $h_\lambda(n) = h_{\lambda[n]}(n)$ for $\lambda \in \mathit{Lim}$.

Then $h_\alpha(n) \leq H_\alpha(n) < h_\alpha(n+1)$.

Proof by induction on α:

1. $h_{\alpha+1}(n) = h_\alpha(n+1) \overset{\text{IH}}{\leq} H_\alpha(n+1) = H_{\alpha+1}(n)$ and $H_{\alpha+1}(n) = H_\alpha(n+1) \overset{\text{IH}}{<} h_\alpha(n+2) = h_{\alpha+1}(n+1)$.
2. $\alpha \in \mathit{Lim}$: $h_\alpha(n) = h_{\alpha[n]}(n) \overset{\text{IH}}{\leq} H_{\alpha[n]}(n) < H_{\alpha[n]}(n+1) = H_\alpha(n)$ and $H_\alpha(n) = H_{\alpha[n]}(n+1) \overset{\text{L.6.9b}}{\leq} H_{\alpha[n+1]}(n) \overset{\text{IH}}{<} h_{\alpha[n+1]}(n+1) = h_\alpha(n+1)$.

Appendix to §6: Comparison with Gentzen's consistency proof

Modifications of Z^*:

– all Cut_C are deleted,

– in the definition of $(\mathsf{Ind}^{y,t}_F h_0)[0]$ we use $\mathsf{R}_{F(\underline{i})}$ instead of $\mathsf{Cut}_{F(\underline{i})}$,

– $\deg(\mathsf{Ind}^{y,t}_F) := \mathrm{rk}(F)$,

– $\mathrm{o}(h) := \begin{cases} \mathrm{o}(h_0)\#\mathrm{o}(h_1) & \text{if } \mathcal{I} = \mathsf{R}_C \\ \mathrm{o}(h_0)\cdot\omega & \text{if } \mathcal{I} = \mathsf{Ind}^{y,t}_C \\ \omega^{\mathrm{o}(h_0)} & \text{if } \mathcal{I} = \mathsf{E} \\ (\sup_{i<n}\mathrm{o}(h_i))+1 & \text{otherwise} \end{cases}$,

Definition of $\psi : \mathsf{Z} \to \mathsf{Z}^*$

$\psi(d) \in \mathsf{Z}^*$ results from d by replacing every occurrence of a symbol Cut_C [$\mathsf{Ind}^{y,t}_C$, resp.] by $\mathsf{E}^m\mathsf{R}_C$ [$\mathsf{E}^m\mathsf{Ind}^{y,t}_C$, resp.] , with $m := \mathrm{rk}(C)\dot{-}\mathrm{hgt}_d(\nu)$ where ν is the position of Cut_C [Ind_C] in d.

Definition of $\widehat{\psi} : \mathsf{Z}^* \to \mathsf{Z}$

$\widehat{\psi}(h) \in \mathsf{Z}$ results from h by deleting all E's and replacing R_C by Cut_C.

Proposition

$\Gamma(\psi(d)) = \Gamma(d)$, $\deg(\psi(d)) = 0$, $\mathrm{o}(\psi(d)) = O(d)$.

Example

Assume $\mathrm{rk}(D) = 1$ and $\mathrm{rk}(\forall xA), \mathrm{rk}(G_0), \mathrm{rk}(G_1) \le \mathrm{rk}(F) = 3$.

$$d = \cfrac{\cfrac{\cfrac{\cfrac{\cfrac{d_0}{\bigwedge^y_{\forall xA}} \quad c_0}{\mathsf{Cut}_{G_0}} \quad \cfrac{\cfrac{d_1}{\bigvee^k_{\exists x\neg A}} \quad c_1}{\mathsf{Cut}_{G_1}}}{\mathsf{Cut}_{\forall xA} : 3} \quad b}{\mathsf{Cut}_F : 1} \quad a}{\mathsf{Cut}_D : 0} \qquad \psi(d) = \cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{\cfrac{d_0'}{\bigwedge^y_{\forall xA}} \quad c_0'}{\mathsf{R}_{G_0}} \quad \cfrac{\cfrac{d_1'}{\bigvee^k_{\exists x\neg A}} \quad c_1'}{\mathsf{R}_{G_1}}}{\mathsf{R}_{\forall xA}} \quad b'}{\mathsf{R}_F}}{\mathsf{E}}}{\mathsf{E}} \quad a'}{\mathsf{R}_D}}{\mathsf{E}}$$

$O(d) = O_d(\,) = \omega_1\Big(O_d(0) \# O_d(1)\Big) =$
$\omega_1\Big(\omega_2\Big(O_d(00)\#O_d(01)\Big) \# O_d(1)\Big) =$
$\omega_1\Big(\omega_2\Big(O_d(000) \# O_d(001) \# O_d(01)\Big) \# O_d(1)\Big)$.

$\mathrm{o}(\psi(d)) = \omega_1\Big(\omega_2\Big(\mathrm{o}(d_0')+1\#\mathrm{o}(c_0') \# \mathrm{o}(d_1')+1\#\mathrm{o}(c_1') \# \mathrm{o}(b')\Big) \# \mathrm{o}(a')\Big)$

Remark

a) $\widehat{\psi}(\psi(d)) = d$,

b) $\forall h(\deg(h) = 0 \;\&\; \widehat{\psi}(h) = d \Rightarrow \mathrm{o}(\psi(d)) \le \mathrm{o}(h))$.

Abbreviation

$d \succ_{\mathcal{G}} d^- \;:\Longleftrightarrow\;$ d^- results from d by a "Gentzen reduction step" in the sense of Theorem 5.4.

Theorem

If d is a closed Z-derivation with $\Gamma(d) = \emptyset$, then $\mathsf{tp}(\psi(d)) = \mathsf{Rep}$ and $d \succ_{\mathcal{G}} \widehat{\psi}(\psi(d)[0])$.

$$\begin{array}{ccc} \mathsf{Z} & \succ_{\mathcal{G}} & \mathsf{Z} \\ \psi \downarrow & & \uparrow \widehat{\psi} \\ \mathsf{Z}^* & \xrightarrow[\cdot\,[0]]{} & \mathsf{Z}^* \end{array}$$

Proof by example: Assume $\mathrm{rk}(D) = 1$ and $\mathrm{rk}(\forall xA), \mathrm{rk}(G_0), \mathrm{rk}(G_1) \leq \mathrm{rk}(F) = 3$.

$$d = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0}{\bigwedge^y_{\forall xA}} \quad c_0}{\mathsf{Cut}_{G_0}} \quad \dfrac{\dfrac{d_1}{\bigvee^k_{\exists x\neg A}} \quad c_1}{\mathsf{Cut}_{G_1}}}{\mathsf{Cut}_{\forall xA}} \quad b}{\mathsf{Cut}_F} \quad a}{\mathsf{Cut}_D}
\qquad
\psi(d) = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0'}{\bigwedge^y_{\forall xA}} \quad c_0'}{\mathsf{R}_{G_0}} \quad \dfrac{\dfrac{d_1'}{\bigvee^k_{\exists x\neg A}} \quad c_1'}{\mathsf{R}_{G_1}}}{\mathsf{R}_{\forall xA}} \quad b'}{\mathsf{R}_F}}{\mathsf{E}}}{\mathsf{E}} \quad a'}{\mathsf{R}_D}}{\mathsf{E}}$$

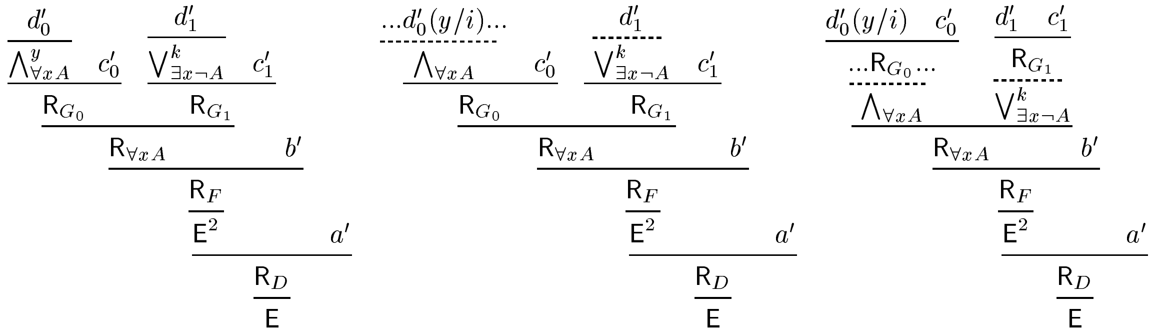
$$\psi(d)[0] = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0'(y/k) \quad c_0'}{\mathsf{R}_{G_0}} \quad \dfrac{\dfrac{d_1'}{\bigvee^k_{\exists x\neg A}} \quad c_1'}{\mathsf{R}_{G_1}}}{\mathsf{R}_{\forall xA}} \quad b'}{\mathsf{R}_F}}{\mathsf{E}} \quad \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0'}{\bigwedge^y_{\forall xA}} \quad c_0'}{\mathsf{R}_{G_0}} \quad \dfrac{d_1' \quad c_1'}{\mathsf{R}_{G_1}}}{\mathsf{R}_{\forall xA}} \quad b'}{\mathsf{R}_F}}{\mathsf{E}}}{\mathsf{R}_{A(k)}}}{\mathsf{E}} \quad a'}{\mathsf{R}_D}}{\mathsf{E}}$$

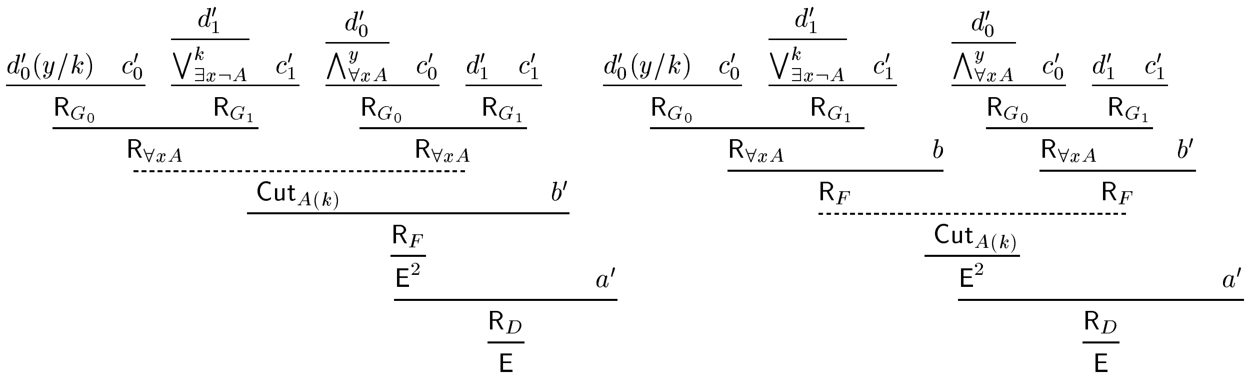
$$\widehat{\psi}(\psi(d)[0]) = \dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{d_0(y/k) \quad c_0}{\mathsf{Cut}_{G_0}} \quad \dfrac{\dfrac{d_1}{\bigvee^k_{\exists x\neg A}} \quad c_1}{\mathsf{Cut}_{G_1}}}{\mathsf{Cut}_{\forall xA}} \quad b}{\mathsf{Cut}_F} \quad \dfrac{\dfrac{\dfrac{\dfrac{d_0}{\bigwedge^y_{\forall xA}} \quad c_0}{\mathsf{Cut}_{G_0}} \quad \dfrac{d_1 \quad c_1}{\mathsf{Cut}_{G_1}}}{\mathsf{Cut}_{\forall xA}} \quad b}{\mathsf{Cut}_F}}{\mathsf{Cut}_{A(k)}} \quad a}{\mathsf{Cut}_D}$$

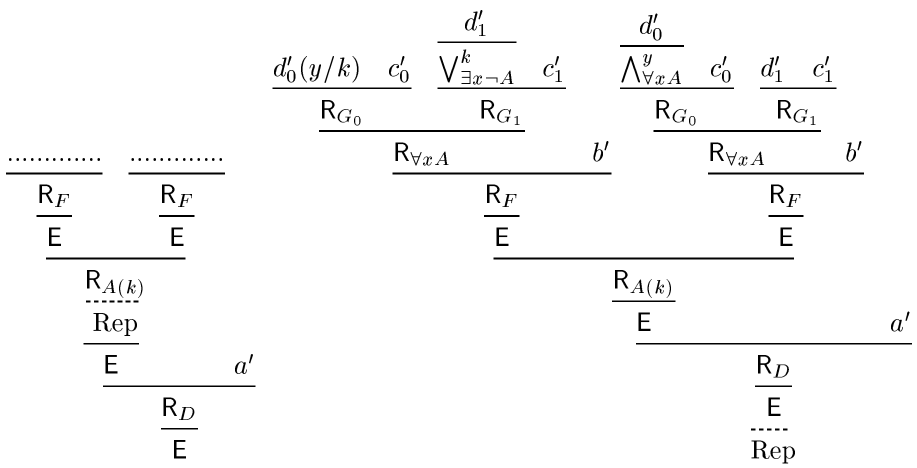
Computation of $\psi(d)[0]$:

By $\frac{\ldots h_i \ldots}{\mathcal{I}}$ (dashed line) we express that $\mathsf{tp}(h) = \mathcal{I}$ and $h[i] = h_i$ holds for the "corresponding" subderivation h of $\psi(d)$.

$\psi(d) =$







§7 Independence Results

Goodstein's Theorem

Definition von $\Phi_b^\alpha : \omega \to \varepsilon_0$ für $2 \le b < \omega$ und $b < \alpha < \varepsilon_0$

$\Phi_b^\alpha(x) := \alpha^{\Phi_b^\alpha(x_1)} \cdot n_1 + \ldots + \alpha^{\Phi_b^\alpha(x_k)} \cdot n_k$, falls $x = \sum_{i=1}^k b^{x_i} \cdot n_i$ mit $x_1 > \ldots > x_k$ und $n_1, ..., n_k \in \{1, ..., b{-}1\}$.

$\mathcal{S}_n(x) := \Phi_{n+1}^{n+2}(x) \quad (n \ge 1,\, x \in \omega)$.

$\mathrm{GS}(a,0) := \mathrm{GS}(a,1) := a, \;\; \mathrm{GS}(a,n+1) := \mathcal{S}_n(\mathrm{GS}(a,n)) \dot{-} 1$ für $n \ge 1$.

Die Folge $(\mathrm{GS}(a,n))_{n \in \mathbb{N}}$ (wobei $a \in \omega$) heißt *Goodstein-Folge* zu a.

Wir werden zeigen, daß jede Goodstein-Folge terminiert (d.h. $\forall a \exists n. \mathrm{GS}(a,n) = 0$), und, daß diese Tatsache nicht in Z beweisbar ist.

Abk.: $\theta_n(x) := \Phi_{n+1}^\omega(x) \quad (n \ge 1,\, x \in \omega)$.

Lemma 7.1 Sei $n \ge 1$.

a) $x < y < \omega \;\Rightarrow\; \mathcal{S}_n(x) < \mathcal{S}_n(y)$ & $\theta_n(x) < \theta_n(y)$,

b) $\theta_n(x) = \theta_{n+1}\mathcal{S}_n(x)$,

c) $x > 0 \;\Rightarrow\; \theta_n(x-1) < \theta_n(x)$.

Beweis durch Induktion nach x:

a) Man zeige: $2 \le b < \alpha$ & $x < y \;\Rightarrow\; \Phi_b^\alpha(x) < \Phi_b^\alpha(y)$. Dabei benutze man folgenden Hilfsatz:

Ist $\xi = \beta^{\xi_0} \cdot \delta_0 + \ldots + \beta^{\xi_k} \cdot \delta_k$ und $\eta = \beta^{\eta_0} \cdot \gamma_0 + \ldots + \beta^{\eta_l} \cdot \gamma_l$ mit $\gamma \ge 2$, $\xi_0 > \ldots > \xi_k$, $\eta_0 > \ldots > \eta_l$ und $0 < \delta_0, ..., \delta_k, \gamma_0, ..., \gamma_l < \beta$, so gilt:

$$\xi < \eta \iff \begin{cases} k < l \;\&\; \forall i \le k(\xi_i = \eta_i \;\&\; \delta_i = \gamma_i) \text{ oder} \\ \exists p \le \min(k,l) \text{ mit } \forall i < p(\xi_i = \eta_i \;\&\; \delta_i = \gamma_i) \;\&\; [\,\xi_p < \eta_p \text{ oder } (\xi_p = \eta_p \;\&\; \delta_p < \eta_p)\,] \end{cases}$$

b) Sei $x = \sum_{i=1}^k (n{+}1)^{x_i} \cdot n_i$ mit $x_1 > \ldots > x_k$ und $n_1, ..., n_k \in \{1, ..., n\}$.

Es gilt $\mathcal{S}_n(x) = \sum_{i=1}^k (n{+}2)^{\mathcal{S}_n(x_i)} \cdot n_i$, und nach a) $\mathcal{S}_n(x_1) > \ldots > \mathcal{S}_n(x_k)$.

Folglich $\theta_{n+1}\mathcal{S}_n(x) = \omega^{\theta_{n+1}\mathcal{S}_n(x_1)} \cdot n_1 + \ldots + \omega^{\theta_{n+1}\mathcal{S}_n(x_k)} \cdot n_k \overset{\mathrm{IH}}{=} \omega^{\theta_n(x_1)} \cdot n_1 + \ldots + \omega^{\theta_n(x_k)} \cdot n_k = \theta_n(x)$.

c) folgt aus 7.2a,d.

Sei nun $a_n := \mathrm{GS}(a,n)$. Dann gilt:

$a_n > 0 \;\Rightarrow\; \mathcal{S}_n(a_n) > 0 \;\Rightarrow\; \theta_{n+1}(a_{n+1}) = \theta_{n+1}(\mathcal{S}_n(a_n) - 1) \overset{7.1\mathrm{c}}{<} \theta_{n+1}(\mathcal{S}_n(a_n)) \overset{7.1\mathrm{b}}{=} \theta_n(a_n)$.

Definition $\mathrm{P}_n(0) := 0$, $\mathrm{P}_n(\alpha+1) := \alpha$, $\;\mathrm{P}_n(\lambda) := \mathrm{P}_n(\lambda[n])$.

Lemma 7.2 Sei $n > 0$.

a) $\alpha > 0 \Rightarrow \mathrm{P}_n(\alpha) < \alpha$,

b) $\mathrm{P}_n(\alpha + \beta) = \alpha + \mathrm{P}_n(\beta)$, falls $NF(\alpha, \beta)$,

c) $\mathrm{P}_n(\omega^\alpha) = \mathrm{P}_n(\omega^{\mathrm{P}_n(\alpha)} \cdot (n{+}1))$, falls $\alpha > 0$,

d) $\mathrm{P}_n(\theta_n(x)) = \theta_n(x \dot{-} 1)$,

e) $n \ge 1 \;\Rightarrow\; \theta_n \mathrm{GS}(a,n) = \mathrm{P}_n ... \mathrm{P}_2 \theta_1(a)$.

Beweis:

a),b) klar.

c) Induktion nach α:

1. $\mathrm{P}_n(\omega^{\alpha+1}) = \mathrm{P}_n(\omega^{\alpha} \cdot (n+1)) = \mathrm{P}_n(\omega^{\mathrm{P}_n(\alpha+1)} \cdot (n+1))$.

2. $\mathrm{P}_n(\omega^{\lambda}) = \mathrm{P}_n(\omega^{\lambda[n]}) \overset{\mathrm{IH}}{=} \mathrm{P}_n(\omega^{\mathrm{P}_n(\lambda[n])} \cdot (n+1)) = \mathrm{P}_n(\omega^{\mathrm{P}_n(\lambda)} \cdot (n+1))$.

d) Induktion nach x: Für $x = 0$ ist die Behauptung trivial.

Sei also $0 < x = \sum_{i=1}^{k}(n+1)^{x_i} \cdot n_i$ mit $x_1 > \ldots > x_k$ und $n_0, ..., n_k \in \{1, ..., n\}$.

Dann $x - 1 = \sum_{i=1}^{k-1}(n+1)^{x_i} \cdot n_i + (n+1)^{x_k} \cdot (n_k - 1) + (n+1)^{x_k - 1} \cdot n + \ldots + (n+1)^0 \cdot n$ und folglich

$\theta_n(x-1) = \omega^{\theta_n(x_1)} \cdot n_1 + \ldots + \omega^{\theta_n(x_k)} \cdot (n_k - 1) + \omega^{\theta_n(x_k - 1)} \cdot n + \ldots + \omega^{\theta_n(0)} \cdot n \overset{\mathrm{HS}}{=}$

$\omega^{\theta_n(x_1)} \cdot n_1 + \ldots + \omega^{\theta_n(x_k)} \cdot (n_k - 1) + \mathrm{P}_n(\omega^{\theta_n(x_k)}) \overset{\mathrm{b}),7.1\mathrm{a}}{=} \mathrm{P}_n(\omega^{\theta_n(x_1)} \cdot n_1 + \ldots + \omega^{\theta_n(x_k)} \cdot n_k) = \mathrm{P}_n(\theta_n(x))$.

HS: $z < x \;\Rightarrow\; \mathrm{P}_n(\omega^{\theta_n(z)}) = \omega^{\theta_n(z-1)} \cdot n + \ldots + \omega^{\theta_n(0)} \cdot n$.

Beweis durch Nebeninduktion nach z:

1. $\mathrm{P}_n(\omega^{\theta_n(0)}) = \mathrm{P}_n(1) = 0$.

2. $z+1 < x$: $\mathrm{P}_n(\omega^{\theta_n(z+1)}) \overset{\mathrm{c})}{=} \mathrm{P}_n(\omega^{\mathrm{P}_n\theta_n(z+1)} \cdot (n+1)) \overset{\mathrm{IH}}{=} \mathrm{P}_n(\omega^{\theta_n(z)} \cdot (n+1)) \overset{\mathrm{b})}{=} \omega^{\theta_n(z)} \cdot n + \mathrm{P}_n(\omega^{\theta_n(z)}) \overset{\mathrm{NIH}}{=}$
$\omega^{\theta_n(z)} \cdot n + \omega^{\theta_n(z-1)} \cdot n + \ldots + \omega^{\theta_n(0)} \cdot n$.

e) Sei $a_i := \mathrm{GS}(a, i)$. Dann $\theta_{n+1}(a_{n+1}) \overset{\mathrm{Def}}{=} \theta_{n+1}(\mathcal{S}_n(a_n) \dot{-} 1) \overset{7.2\mathrm{d}}{=} \mathrm{P}_{n+1}\theta_{n+1}\mathcal{S}_n(a_n) \overset{7.1\mathrm{b}}{=} \mathrm{P}_{n+1}\theta_n(a_n) \overset{\mathrm{IH}}{=}$
$\mathrm{P}_{n+1}\mathrm{P}_n ... \mathrm{P}_2\theta_1(a)$.

Definition $h_0(n) := n, \quad h_{\alpha+1}(n) := h_\alpha(n+1), \quad h_\lambda(n) := h_{\lambda[n]}(n)$.

Lemma 7.3

a) $H_\alpha(n) < h_\alpha(n+1)$,

b) $\alpha > 0 \;\Rightarrow\; h_\alpha(n) = h_{\mathrm{P}_n(\alpha)}(n+1)$.

c) $h_\alpha(n) = \min\{k \geq n : \mathrm{P}_{k-1} ... \mathrm{P}_n(\alpha) = 0\}$.

Proof:

a) siehe §6.

b) Induktion nach α:

1. $\alpha = \beta + 1$: $h_\alpha(n) = h_\beta(n+1) = h_{\mathrm{P}_n(\alpha)}(n+1)$.

2. $\alpha \in Lim$: $h_\alpha(n) = h_{\alpha[n]}(n) = h_{\mathrm{P}_n(\alpha[n])}(n+1) = h_{\mathrm{P}_n(\alpha)}(n+1)$.

c) Für $k = \min\{i \geq n : \mathrm{P}_{i-1} ... \mathrm{P}_n(\alpha) = 0\}$ gilt offenbar

$h_\alpha(n) = h_{\mathrm{P}_n(\alpha)}(n+1) = h_{\mathrm{P}_{n+1}\mathrm{P}_n(\alpha)}(n+2) = \ldots = h_{\mathrm{P}_{k-1} ... \mathrm{P}_n(\alpha)}(k) = h_0(k) = k$.

Satz 7.4

$\mathcal{Z} \nvdash \forall x \exists y[\,\mathrm{GS}(x, y) = 0\,]$.

Proof

Sei $e(0) := 1$, $e(m+1) := 2^{e(m)}$.

Annahme: $\mathcal{Z} \vdash \forall x \exists y[\,\mathrm{GS}(x, y) = 0\,]$.

$\Rightarrow\; \mathcal{Z} \vdash \forall x \exists y[\,\mathrm{GS}(e(x) + x, y) = 0\,] \overset{6.12}{\Rightarrow} \exists \alpha < \varepsilon_0 \forall m \exists n < H_\alpha(m)[\,\mathrm{GS}(e(m) + m, n) = 0\,] \;\Rightarrow$

$\Rightarrow\; \exists p, n[n < H_{\omega_p}(p)\ \&\ \mathrm{GS}(e(p) + p, n) = 0\,]$. ($p \geq N(\alpha)$ mit $\alpha < \omega_p$; vergl. 6.9c)

$\mathrm{GS}(e(p)+p, n) = 0 \;\Rightarrow\; n \geq 2\ \&\ \mathrm{P}_n ... \mathrm{P}_2(\omega_p + p) = \theta_n \mathrm{GS}(e(p)+p, n) = 0 \;\Rightarrow\; H_{\omega_p+p}(1) < h_{\omega_p+p}(2) \leq n+1 \;\Rightarrow$
$H_{\omega_p}(p) = H_{\omega_p+p}(0) < H_{\omega_p+p}(1) \leq n$. *Widerspruch.*

The Paris-Harrington Result

(The proof given here is based on M. Loebl and J. Nešetřil, *An unprovable Ramsey-type Theorem*, Report No 89598-OR, September 1989.)

Abbreviations:

Let $k, m, n, r \in \mathbb{N} (= \omega)$, κ a cardinal, N a set.

$[N]^m := \{X \subseteq N : card(X) = m\}$

Let f be a function with $\mathrm{dom}(f) = [N]^m$:

X is f-homogeneous $\iff$ $\emptyset \neq X \subseteq N$ & $f \restriction [X]^m$ constant.

$N \longrightarrow (\kappa)^m_r \;:\Leftrightarrow\; \forall f : [N]^m \to r \exists X (X\ f\text{-homogeneous} \;\&\; card(X) \geq \kappa)$

$N \overset{*}{\longrightarrow} (\kappa)^m_r \;:\Leftrightarrow\; \forall f : [N]^m \to r \exists X (X\ f\text{-homogeneous} \;\&\; card(X) \geq \max\{\kappa, \min(X)\})$ (for $N \subseteq \mathbb{N}$)

Ramsey Theorem $\forall m, r \in \omega (\omega \longrightarrow (\omega)^m_r)$

Finite Ramsey Theorem $\forall m, r, \kappa \in \omega \exists N \in \omega (N \longrightarrow (\kappa)^m_r)$

PH $\forall m, r, \kappa \in \omega \exists N \in \omega (N \overset{*}{\longrightarrow} (\kappa)^m_r)$

(Proof of the Finite Ramsey Theorem in PA: cf. Hajek, Pudlak Ch.2, Sec.1)

Proof of PH:

Let $m, r, \kappa \in \omega$ be fixed. To prove: $\exists n \in \omega \forall f \neg\Phi(f, n)$

where $\Phi(f, n) \;:\Leftrightarrow\; f : [n]^m \to r \;\&\; \forall X (X\ f\text{-hom} \Rightarrow card(X) < \max\{\kappa, \min(X)\})$.

Assumption: $\forall n \exists f\, \Phi(f, n)$.

By König's Lemma there is a function $f^*[\omega]^m \to r$ such that

$(+)$ $\forall n\, \Phi(f^* \restriction [n]^m, n)$.

By Ramsey's Theorem there exists an infinite f^*-homogeneous set $X \subseteq \omega$. We choose $N < \omega$ such that $card(X \cap N) > \max\{\kappa, \min(X)\}$. For $f := f^* \restriction [N]^m$ we then have $f \restriction [X \cap N]^m = f^* \restriction [X \cap N]^m = constant$. Hence $X \cap N$ is f-homogeneous and $card(X \cap N) > \max\{\kappa, \min(X)\} = \max\{\kappa, \min(X \cap N)\}$, i.e., $\neg\Phi(f, N)$. *Contradiction* to $(+)$.

[[Construction of f^*: Let $\Phi_n := \{f : \Phi(f, n)\}$ and $M(f) := \{i : \exists g \in \Phi_i (f \subseteq g)\}$. — Starting with $f_0 := \emptyset$ we define a sequence $(f_n)_{n \in \omega}$ such that $f_n \in \Phi_n$ & $f_n \subseteq f_{n+1}$ & $card(M(f_n)) \geq \omega$, and we set $f^* := \bigcup_{n \in \omega} f_n$. Definition of f_{n+1}: Let $E := \{f \in \Phi_{n+1} : f_n \subseteq f\}$. Then $M(f_n) = \{n\} \cup \bigcup_{f \in E} M(f)$, and E is finite. This together with $\forall i > n \forall f \in \Phi_i (f \restriction [n{+}1]^m \in \Phi_{n+1})$ implies the existence of an $f_{n+1} \in E$ such that $card(M(f_{n+1})) \geq \omega$.]]

Theorem 7.5 $\forall m \geq 1 \forall k (H_{\omega_m(k)}(k{+}1) < R_m(k))$ with $R_m(k) := \min\{N : N \overset{*}{\longrightarrow} (2m{+}k{+}4)^{m+1}_{k+\Sigma_{i<m} 3^i}\}$

Corollary

a) $\mathcal{Z}_m \nvdash \forall \kappa, r \exists N (N \overset{*}{\longrightarrow} (\kappa)^{m+1}_r)$ $(m \geq 1)$

b) $\mathcal{Z} \nvdash \forall m, \kappa, r \exists N (N \overset{*}{\longrightarrow} (\kappa)^{m+1}_r)$

Proof of the Corollary:

a) Assumption: $\mathcal{Z}_m \vdash \forall \kappa, r \exists N(N \xrightarrow{*} (\kappa)_r^{m+1})$.

Then $\mathcal{Z}_m \vdash \forall \kappa, r \exists N(N \xrightarrow{*} (2m{+}k{+}4)^{m+1}_{k+\Sigma_{i<m} 3^i} \})$.

By 6.12 there is an $\alpha < \omega_{m+1}$ such that $\forall k(R_m(k) < H_\alpha(k))$.

Let $k \in \omega$ such that $\alpha < \omega_m(k)$ and $N(\alpha) \leq k$.

Then $H_{\omega_m(k)}(k{+}1) \overset{7.5}{<} R_m(k) < H_\alpha(k) < H_{\omega_m(k)}(k)$. *Contradiction.*

b) $\mathcal{Z} \vdash \forall n, \kappa, r \exists N(N \xrightarrow{*} (\kappa)_r^{n+1})$

$\Longrightarrow \mathcal{Z}_m \vdash \forall n, \kappa, r \exists N(N \xrightarrow{*} (\kappa)_r^{n+1})$ for suitable m

$\Longrightarrow \mathcal{Z}_m \vdash \forall \kappa, r \exists N(N \xrightarrow{*} (\kappa)_r^{m+1})$.

The rest of this section is spent with the proof of 7.5.

Definition

1. $\omega_0(\alpha) := \alpha$, $\omega_{m+1}(\alpha) := \omega^{\omega_m(\alpha)}$

2. Sei $\alpha = \omega^{\alpha_1} n_1 + ... + \omega^{\alpha_t} n_t$ mit $t \geq 0 \,\&\, n_1, ..., n_t > 0 \,\&\, \alpha > \alpha_1 > ... > \alpha_t$:

$\mathsf{r}(\alpha) := \max\{t, n_1, ..., n_t, \mathsf{r}(\alpha_1), ..., \mathsf{r}(\alpha_t)\}$.

Lemma 7.6

a) $\mathsf{r}(\alpha[k]) \leq \max\{\mathsf{r}(\alpha), k\} + 1$

b) $\mathsf{r}(\alpha) < n \Longrightarrow \mathsf{r}(\alpha[n{-}1]) < n{+}1$

c) $\mathsf{r}(\omega_m(k) + n) \leq n + 1$, falls $1 \leq m$ und $k < n$.

Beweis von c): 1. $k = 0$ und $m = 1$: $\omega_m(k) + n = \omega^0(n+1)$.

2. $k \neq 0$ und $m = 1$: $\mathsf{r}(\omega_m(k) + n) = \max\{2, 1, n, \mathsf{r}(k)\} \leq n{+}1$.

3. $m > 1$: $\mathsf{r}(\omega_m(k) + n) = \max\{2, 1, n, \mathsf{r}(\omega_{m-1}(k))\} = \max\{2, n, k\} \leq n{+}1$.

Definition

Sei $\alpha = \omega^{\alpha_1} n_1 + ... + \omega^{\alpha_t} n_t$ mit $t \geq 0 \,\&\, n_1, ..., n_t > 0 \,\&\, \alpha > \alpha_1 > ... > \alpha_t$:

$\mathsf{S}_i(\alpha) := \omega^{\alpha_i} n_i$, $\mathsf{E}_i(\alpha) := \alpha_i$, $\mathsf{K}_i(\alpha) := n_i$ für $i \in \{1, ..., t\}$

$\mathsf{S}_i(\alpha) := \mathsf{E}_i(\alpha) := \mathsf{K}_i(\alpha) := 0$ für $i \notin \{1, ..., t\}$.

Definition

Für $\alpha > \beta$ sei $\mathsf{d}(\alpha, \beta) := \min\{i : \mathsf{S}_i(\alpha) > \mathsf{S}_i(\beta)\}$, $\mathsf{K}(\alpha, \beta) := \mathsf{K}_{\mathsf{d}(\alpha,\beta)}(\alpha)$, $\mathsf{E}(\alpha, \beta) := \mathsf{E}_{\mathsf{d}(\alpha,\beta)}(\alpha)$.

Lemma 7.7

$\alpha > \beta > \gamma \,\&\, \mathsf{d}(\alpha, \beta) \leq \mathsf{d}(\beta, \gamma) \,\&\, \mathsf{K}(\alpha, \beta) \leq \mathsf{K}(\beta, \gamma) \Longrightarrow \mathsf{E}(\alpha, \beta) > \mathsf{E}(\beta, \gamma)$.

Beweis:

Fall 1: $\mathsf{d}(\alpha, \beta) = \mathsf{d}(\beta, \gamma) = i$. Dann $\mathsf{E}(\alpha, \beta) = \mathsf{E}_i(\alpha)$, $\mathsf{E}(\beta, \gamma) = \mathsf{E}_i(\beta)$ und $\mathsf{K}_i(\alpha) \leq \mathsf{K}_i(\beta)$. Wegen $\alpha < \beta$ und $\mathsf{K}_i(\alpha) \leq \mathsf{K}_i(\beta)$ muß $\mathsf{E}_i(\alpha) > \mathsf{E}_i(\beta)$ sein.

Fall 2: $i = \mathsf{d}(\alpha, \beta) < \mathsf{d}(\beta, \gamma) = j$. Dann $\mathsf{E}(\alpha, \beta) = \mathsf{E}_i(\alpha) \geq \mathsf{E}_i(\beta) > \mathsf{E}_j(\beta) = \mathsf{E}(\beta, \gamma)$.

Definition $(\chi : [\varepsilon_0]^3 \to \{\hat{0}, \hat{1}, \hat{2}\})$

Für $\beta_0 > \beta_1 > \beta_2$ sei

$$\chi(\beta_0, \beta_1, \beta_2) := \begin{cases} \hat{0} & \text{falls } \mathsf{d}(\beta_1, \beta_2) < \mathsf{d}(\beta_0, \beta_1) \\ \hat{1} & \text{falls } \mathsf{d}(\beta_0, \beta_1) \leq \mathsf{d}(\beta_1, \beta_2) \,\&\, \mathsf{K}(\beta_1, \beta_2) < \mathsf{K}(\beta_0, \beta_1) \\ \hat{2} & \text{sonst} \end{cases}$$

Lemma 7.8

Sei $\beta_0 > \ldots > \beta_\ell$ mit $\ell \geq 2$, und sei $c \in \{\hat{0}, \hat{1}, \hat{2}\}$ mit $\{\chi(\beta_i, \beta_{i+1}, \beta_{i+2}) : i \leq \ell - 2\} = \{c\}$.

a) $c \in \{\hat{0}, \hat{1}\} \implies \ell \leq \mathsf{r}(\beta_0)$

b) $c = \hat{2} \implies \mathsf{E}(\beta_0, \beta_1) > \ldots > \mathsf{E}(\beta_{\ell-1}, \beta_\ell)$

Beweis: a) Sei $d_i := \mathsf{d}(\beta_i, \beta_{i+1})$ und $k_i := \mathsf{K}(\beta_i, \beta_{i+1}) \quad (i < \ell)$.

Offenbar ist $d_0, k_0 \leq \mathsf{r}(\beta_0)$. Ist $c = \hat{0}$, so $d_0 > d_1 > \ldots > d_{\ell-1} \geq 1$, woraus $\ell \leq d_0$ folgt.
Ist $c = \hat{1}$, so $k_0 > k_1 > \ldots > k_{\ell-1} \geq 1$, woraus $\ell \leq k_0$ folgt. — b) folgt aus Lemma 7.7.

Definition ($\chi^k_m : [\omega_m(k{+}1)]^{m+1} \to \mathcal{C}^k_m$ für $m \geq 1$, $k \geq 0$)

1. Für $\alpha = \omega^k n_0 + \ldots + \omega^0 n_k > \omega^k n'_0 + \ldots + \omega^0 n'_k = \beta$ setzen wir $\chi^k_1(\alpha, \beta) := \min\{i : n'_i < n_i\}$.

2. Sei $m \geq 1$, $\omega_{m+1}(k{+}1) > \beta_0 > \ldots > \beta_{m+1}$, $\delta_i := \mathsf{E}(\beta_i, \beta_{i+1})$, und $c_i := \chi(\beta_i, \beta_{i+1}, \beta_{i+2})$.

$$\chi^k_{m+1}(\beta_0, ..., \beta_{m+1}) := \begin{cases} \chi^k_m(\delta_0, ..., \delta_m) & \text{falls } c_0 = \ldots = c_{m-1} = \hat{2} \\ (c_0, \ldots, c_{m-1}) & \text{sonst} \end{cases}$$

(Man beachte, daß nach L.7.8 $\omega_m(k{+}1) > \delta_0 > \ldots > \delta_m$, falls $c_0 = \ldots = c_{m-1} = \hat{2}$.

3. $\mathcal{C}^k_1 := \{0, \ldots, k\}$, $\mathcal{C}^k_{m+1} := \mathcal{C}^k_m \cup \{\hat{0}, \hat{1}, \hat{2}\}^m$

Lemma 7.9

Aus $1 \leq m < \ell \,\&\, \omega_m(k + 1) > \beta_0 > \ldots > \beta_\ell \,\&\, c \in \mathcal{C}^k_m$ folgt:
$\forall i \leq \ell{-}m(\, \chi^k_m(\beta_i, ..., \beta_{i+m}) = c\,) \implies \ell < \mathsf{r}(\beta_0) + m$.

Beweis durch Induktion nach m:

I. $m = 1$: Sei $\beta_i = \omega^k n_{i,0} + \ldots + \omega^0 n_{i,k}$. Dann $n_{0,c} > \ldots > n_{\ell,c}$ und somit $\ell \leq n_{0,c} \leq \mathsf{r}(\beta_0) < \mathsf{r}(\beta_0) + m$.

II. $m \to m{+}1$: Sei $1 \leq m \,\&\, m{+}1 < \ell \,\&\, c \in \mathcal{C}^k_{m+1} \,\&\, \omega_{m+1}(k{+}1) > \beta_0 > \ldots > \beta_\ell$ und $\chi^k_{m+1}(\beta_i, ..., \beta_{i+m+1}) = c$ für alle $i \leq \ell{-}m{-}1$.

FALL 1: $c \in \{\hat{0}, \hat{1}, \hat{2}\}^m$.
Sei $c_i := \chi(\beta_i, \beta_{i+1}, \beta_{i+2})$ $(i \leq \ell{-}2)$. Für alle $i \leq \ell{-}m{-}2$ haben wir dann
$(c_i, ..., c_{i+m-1}) = \chi^k_{m+1}(\beta_i, ..., \beta_{i+m+1}) = \chi^k_{m+1}(\beta_{i+1}, ..., \beta_{i+m+2}) = (c_{i+1}, ..., c_{i+m})$. Das impliziert $\forall i \leq \ell{-}m{-}2(\, c_i = c_{i+1} = \ldots = c_{i+m}\,)$ und somit $c_0 = \ldots = c_{\ell-2}$. Zusammen mit $\chi^k_{m+1}(\beta_0, ..., \beta_{m+1}) \in \{\hat{0}, \hat{1}, \hat{2}\}^m$ folgt daraus $c_0 \in \{\hat{0}, \hat{1}\}$ und dann $\ell < \mathsf{r}(\beta_0) + m + 1$ nach Lemma 7.8a.

FALL 2: $c \in \mathcal{C}^k_m$.
Dann $c_0 = \ldots = c_{\ell-2} = \hat{2}$. Nach L.7.8b folgt daraus $\delta_0 > \ldots > \delta_{\ell-1}$. Ferner $c = \chi^k_{m+1}(\beta_i, ..., \beta_{i+m+1}) = \chi^k_m(\delta_i, ..., \delta_{i+m})$ für $i \leq \ell{-}1{-}m$. Nach IV folgt daraus $\ell{-}1 < \mathsf{r}(\delta_0){+}m$. Also $\ell < \mathsf{r}(\delta_0){+}m{+}1 \leq \mathsf{r}(\beta_0){+}m{+}1$.

Proof of Theorem 7.5:

Let $m \geq 1$, $k \geq 0$, $n := m{+}k{+}2$, $\kappa := m{+}n{+}2 = 2m{+}k{+}4$, $r := k + \Sigma_{i<m} 3^i$, $H := H_{\omega_m(k)}(k{+}1)$.

Then the following holds

$H < R_m(k) \;\Leftrightarrow$
$\forall n \leq H \neg(n \xrightarrow{*} (\kappa)^{m+1}_r) \;\Leftrightarrow$

$\forall n \le H \exists f(f : [n]^{m+1} \longrightarrow r \;\&\; \forall X(X \; f\text{-homogeneous} \;\Rightarrow\; card(X) < \max\{\kappa, \min(X)\}) \;\Leftrightarrow$
$\exists N \ge H \exists f(f : [N]^{m+1} \longrightarrow r \;\&\; \forall X(X \; f\text{-homogeneous} \;\Rightarrow\; card(X) < \max\{\kappa, \min(X)\}) \quad (*)$

It remains to prove $(*)$.

Sei $\alpha_0 := \omega_m(k) + n$, $\alpha_{i+1} := \alpha_i[i \dot{-} (m+1)]$, $N := \min\{i : \alpha_i = 0\}$. Dann $\forall i < N(\alpha_i > \alpha_{i+1})$.

HS 1: $\forall i \le n(\mathsf{r}(\alpha_i) + m < \kappa)$ und $\forall i \ge n(\mathsf{r}(\alpha_i) + m < i)$.

Beweis: $i \le n \Rightarrow \mathsf{r}(\alpha_i) + m \le \mathsf{r}(\alpha_0) + m \le n{+}1{+}m < \kappa$.
$\mathsf{r}(\alpha_n) + m \le k{+}1 + m < n$; und nach L.7.6b gilt: $\mathsf{r}(\alpha_i) + m < i \Rightarrow \mathsf{r}(\alpha_{i+1}) = \mathsf{r}(\alpha_i[i{-}m{-}1]) < i{-}m{+}1 \Rightarrow \mathsf{r}(\alpha_{i+1}) + m < i + 1$.

HS2: $H \le N$.

Beweis: $n \le i < N \Rightarrow H_{\alpha_i}(i - m - 1) = H_{\alpha_{i+1}}(i + 1 - m - 1)$.

Also $H = H_{\omega_m(k)}(k + 1) = H_{\alpha_n}(n - m - 1) = H_{\alpha_N}(N - m - 1) = N - m - 1 \le N$.

Definition: $f : [N]^{m+1} \to r$, $f(\{i_0, ..., i_m\}_<) := \chi^k_m(\{\alpha_{i_0}, ..., \alpha_{i_m}\}_>)$.

Sei $X = \{i_0, ..., i_\ell\}_<$ f-homogeneous. Dann ist $\{\alpha_{i_0}, ..., \alpha_{i_\ell}\}_>$ χ^k_m-homogen, und wegen Lemma 7.9 gilt $card(X) = \ell + 1 \le \mathsf{r}(\alpha_{i_0}) + m \overset{\text{HS1}}{<} \max\{\kappa, i_0\} = \max\{\kappa, \min(X)\}$.

§8 The collapsing function ψ

Abbreviation: $\Omega := \aleph_1$.

Definition of $C(\alpha)$ and $\psi(\alpha)$ by recursion on α

$C(\alpha) :=$ closure of $\{0\}$ under $+$, $\lambda x.\omega^{\Omega+x}$ and $\xi \mapsto \psi(\xi)$ ($\xi < \alpha \;\&\; \xi \in C(\xi)$).

$\psi(\alpha) := \min\{\xi : \xi \notin C(\alpha)\}$.

Lemma 8.1

a) $\alpha \le \beta \Rightarrow C(\alpha) \subseteq C(\beta) \;\&\; \psi(\alpha) \le \psi(\beta)$,
$\alpha < \beta \;\&\; \alpha \in C(\alpha) \Rightarrow \psi(\alpha) < \psi(\beta)$,

b) $\psi(\alpha) < \Omega$,

c) $\psi(\alpha)$ is an additive principal number,

d) $\psi(\alpha) = C(\alpha) \cap \Omega$,

e) $\alpha \in C(\alpha) \Leftrightarrow \psi(\alpha) < \psi(\alpha{+}1)$.

Proof:

a) The first part is trivial. – Second part: $\alpha < \beta \;\&\; \alpha \in C(\alpha) \Rightarrow \psi(\alpha) \le \psi(\beta) \;\&\; \psi(\alpha) \in C(\beta) \Rightarrow \psi(\alpha) < \psi(\beta)$. The last implication holds, since $\psi(\beta) \notin C(\beta)$.

b) $card(C(\alpha)) < \Omega \Rightarrow \exists \xi < \Omega(\xi \notin C(\alpha)) \Rightarrow \psi(\alpha) < \Omega$.

c) Assume not. Then $\psi(\alpha) = \xi + \eta$ with $\xi, \eta < \psi(\alpha)$, and thus $\psi(\alpha) = \xi + \eta \in C(\alpha)$. Contradiction.

d) "$\subseteq$" follows from b).

Proof of "$\supseteq$" by C-induction, i.e. induction on the definition of $C(\alpha)$:

Let $\gamma \in C(\alpha) \cap \Omega$. We have to prove $\gamma < \psi(\alpha)$.

1. $\gamma = 0$: $0 < \psi(\alpha)$, since $0 \in C(\alpha)$.
2. $\gamma = \xi + \eta$ with $\xi, \eta \in C(\alpha)$: Then $\xi, \eta < \Omega$, and therefore by IH $\xi, \eta < \psi(\alpha)$. By c) we get $\gamma = \xi + \eta < \psi(\alpha)$.
3. $\gamma = \omega^{\Omega+\xi}$: This cannot be.
4. $\gamma = \psi(\xi)$ with $\xi < \alpha$ and $\xi \in C(\xi)$: In this case the claim follows by a).

e) "$\Rightarrow$" follows from a). "$\Leftarrow$": $\alpha \notin C(\alpha) \Rightarrow C(\alpha+1) = C(\alpha) \Rightarrow \psi(\alpha+1) = \psi(\alpha)$.

Remark $\psi(\alpha_0) = \psi(\alpha_1)$ & $\alpha_i \in C(\alpha_i)$ $(i = 0, 1)$ $\Rightarrow$ $\alpha_0 = \alpha_1$.

Lemma 8.2

a) $\omega^{\delta_1} \# ... \# \omega^{\delta_n} \in C(\alpha) \Leftrightarrow \omega^{\delta_1}, ..., \omega^{\delta_n} \in C(\alpha)$,

b) $\xi + \eta \in C(\alpha) \Rightarrow \eta \in C(\alpha)$.

Proof:

a) Proof of "$\Rightarrow$" by C-induction. W.l.o.g. $\delta_0 \geq \ldots \geq \delta_n$.

1. $\omega^{\delta_1} + ... + \omega^{\delta_n} = \xi + \eta$ with $\xi, \eta \in C(\alpha)$:

Then $\xi =_{NF} \omega^{\delta_1} + ... + \omega^{\delta_k} + \omega^{\xi_1} + ... + \omega^{\xi_k}$ and $\eta =_{NF} \omega^{\delta_{k+1}} + ... + \omega^{\delta_n}$. By IH we get $\omega^{\delta_1}, ..., \omega^{\delta_n} \in C(\alpha)$.

2. $\omega^{\delta_1} + ... + \omega^{\delta_n} = \psi(\xi)$ or $\omega^{\Omega+\xi}$: Then $n = 1$ and the claim is trivial.

b) follows from a).

Lemma 8.3

a) $\psi(\alpha) = \min\{\delta \in AP : \forall \xi < \alpha(\xi \in C(\xi) \Rightarrow \psi(\xi) < \delta)\}$.

b) $\alpha \leq \varepsilon_0 \Rightarrow \psi(\alpha) = \omega^\alpha$.

Proof:

a) Let $\delta_0 := \min\{\delta \in AP : \forall \xi < \alpha(\xi \in C(\xi) \Rightarrow \psi(\xi) < \delta)\}$.

By 8.1a,c,d we have $\delta_0 \leq \psi(\alpha) = C(\alpha) \cap \Omega$. By C-induction we get $C(\alpha) \cap \Omega \subseteq \delta_0$.

b) Induction on α: Let $\alpha \leq \varepsilon_0$. Then by IH $\forall \xi < \alpha(\xi < \omega^\xi = \psi(\xi))$.

Hence $\psi(\alpha) = \min\{\delta \in AP : \forall \xi < \alpha(\omega^\xi < \delta)\} = \omega^\alpha$.

Inductive definition of the set T of terms

(T1) $0 \in \mathrm{T}$.

(T2) $a \in \mathrm{T}$ & $\sigma \in \{0, 1\} \Rightarrow D_\sigma a \in \mathrm{T}$ (terms of this kind are called *principal terms*)

(T3) $a_0, ..., a_n \in \mathrm{T}$ $(n \geq 1)$ principal terms $\Rightarrow (a_0, .., a_n) \in \mathrm{T}$

Notation: We use a, b, c as syntactic variables for elements of T. For principal terms a we set $(a) := a$. The empty sequence $()$ is identified with 0.

Definition of $a \prec b$ for $a, b \in \mathrm{T}$

$a \prec b$ if and only if one of the following cases holds

($\prec$ 1) $a = 0$ and $b \neq 0$

($\prec$ 2) $(a = D_0 a_0$ & $b = D_1 b_0)$ or $(a = D_\sigma a_0$ & $b = D_\sigma b_0$ & $a_0 \prec b_0)$

($\prec 3$) $a = (a_0, ..., a_n)$ & $b = (b_0, ..., b_m)$ & $1 \le m{+}n$ &
$[(n < m \;\&\; \forall i{<}n(a_i = b_i)$ or $\exists k \le \min\{m,n\}(a_k \prec b_k \;\&\; \forall i{<}k(a_i = b_i))]$

Lemma 8.4 $a \neq b \;\Rightarrow\; a \prec b$ or $b \prec a$.

Definition of K^*a **for** $a \in \mathrm{T}$

1. $\mathrm{K}^*0 := \emptyset$
2. $\mathrm{K}^*(a_0, ..., a_n) := \bigcup_{i \le n} \mathrm{K}^*a_i$
3. $\mathrm{K}^*D_1a := \mathrm{K}^*a$
4. $\mathrm{K}^*D_0a := \{a\}$

Abbreviation: $\mathrm{K}^*a \prec b \;:\Leftrightarrow\; \forall x{\in}\mathrm{K}^*a(x \prec b)$

Definition of $o(a) \in On$ **for** $a \in \mathrm{T}$

1. $o(0) := 0$
2. $o((a_0, ..., a_n)) := o(a_0) + ... + o(a_1)$
3. $o(D_0a) := \psi(o(a))$
4. $o(D_1a) := \omega^{\Omega + o(a)}$

Remark $\prec$ is not wellfounded: $... \prec D_0D_0\Omega \prec D_0\Omega \prec \Omega$. Hence the mapping $a \mapsto o(a)$ cannot be order preserving.

Inductive definition of the set $\mathrm{OT} \subseteq \mathrm{T}$

(OT1) $0 \in \mathrm{OT}$.

(OT2) $a_0, ..., a_n \in \mathrm{OT}$ $(n \ge 1)$ principal terms & $\alpha_n \preceq ... \preceq a_0 \;\Rightarrow\; (a_0, .., a_n) \in \mathrm{OT}$

(OT3) $a \in \mathrm{OT} \;\Rightarrow\; D_1a \in \mathrm{OT}$

(OT4) $a \in \mathrm{OT}$ & $\mathrm{K}^*a \prec a \;\Rightarrow\; D_0a \in \mathrm{OT}$

Theorem 8.5 For $a, c \in \mathrm{OT}$ the following holds:
a) $c \prec a \;\Leftrightarrow\; o(c) < o(a)$,
b) $\mathrm{K}^*c \prec a \;\Leftrightarrow\; o(c) \in C(o(a))$.

Proof by induction on the length of c:
Due to 8.4 for a) it suffices to prove "$\Rightarrow$".

1. $c = 0$: a) $c \prec a \Rightarrow a \neq 0 \Rightarrow o(a) \neq 0 \Rightarrow o(c) < o(a)$.
b) $\mathrm{K}^*c = \emptyset$ aand $o(c) = 0 \in C(o(a))$.

2. $c = D_\sigma b$:
a) (1) $a = (a_0, ..., a_n)$ with $n \ge 1$, $c = a_0$: trivial.
(2) $a = (D_\tau\tilde{b}, a_1, ..., a_n)$ with $n \ge 0$ and $[(\sigma = 0 \;\&\; \tau = 1)$ or $(\sigma = \tau \;\&\; b \prec \tilde{b})]$:
In the first case we have $o(c) < \Omega \le o(a)$. In the second case by IH $o(b) < o(\tilde{b})$. For $\sigma = \tau = 1$ this yields $o(c) < o(a)$. For $\sigma = \tau = 0$, since $c \in \mathrm{OT}$, we have $\mathrm{K}^*b \prec b$ and therefore by IH $o(b) \in C(o(b))$. Together with $o(b) < o(\tilde{b})$ from this we get $o(c) = \psi(o(b)) < \psi(o(\tilde{b})) = o(D_0\tilde{b}) \le o(a)$.

b) (1) $c = D_1 b$: $\mathrm{K}^* c = \mathrm{K}^* b$ and $o(c) \in C(o(a)) \Leftrightarrow o(b) \in C(o(a))$.
(2) $c = D_0 b$: Then $\mathrm{K}^* c = \{b\}$ and $\mathrm{K}^* b \prec b$. By IH the latter yields $o(b) \in C(o(b))$ and then:
$\mathrm{K}^* c \prec a \Leftrightarrow b \prec a \overset{\text{IH}}{\Leftrightarrow} o(b) < o(a) \Leftrightarrow \psi(o(b)) < \psi(o(a)) \Leftrightarrow o(c) = \psi(o(b)) \in C(o(a))$.

3. $c = (c_0, ..., c_m)$ with $m \geq 1$ and $a = (a_0, ..., a_n)$ with $n \geq 0$:
a) (1) $m < n \;\&\; \forall i \leq m(c_i = a_i)$: trivial.
(2) $k \leq \min(m, n) \;\&\; \forall i < k(c_i = a_i) \;\&\; c_k \prec a_k$: Since $c \in \mathrm{OT}$, we also have $c_m \preceq c_{m-1} \preceq \ldots \preceq c_k$. By IH we get $o(c_m) \leq o(c_{m-1}) \leq \ldots \leq o(c_k) < o(a_k)$ and from this $o(c) < o(a)$, since $o(a_k) \in AP$.

b) As for a) we get $o(c_m) \leq o(c_{m-1}) \leq \ldots \leq o(c_0)$ $(*)$.
Hence: $\mathrm{K}^* c \prec a \Leftrightarrow \mathrm{K}^* c_i \prec a$ $(i = 0, ..., m) \overset{\text{IH}}{\Leftrightarrow} o(c_i) \in C(o(a))$ $(i = 0, ..., m) \Leftrightarrow o(c) = o(c_0) + \ldots + o(c_m) \in C(o(a))$.

Theorem 8.6
The mapping $\mathrm{OT} \to C(\varepsilon_{\Omega+1})$, $a \mapsto o(a)$ is bijective.

Proof:
1. Obviously $C(\varepsilon_{\Omega+1}) \subseteq \varepsilon_{\Omega+1}$ $(+)$.
$a = D_0 b \in \mathrm{OT} \Rightarrow b \in \mathrm{OT} \;\&\; \mathrm{K}^* b \prec b \overset{\text{IH}}{\Rightarrow} o(b) \in C(\varepsilon_{\Omega+1}) \;\&\; o(b) \in C(\mathrm{o}(b)) \overset{(+)}{\Rightarrow} o(a) = \psi(o(b)) \in C(\varepsilon_{\Omega+1})$.

2. From 8.4 and 8.5a it follows that $o|\mathrm{OT}$ is injective.

3. By C-induction we prove: $\gamma \in C(\varepsilon_{\Omega+1}) \Rightarrow \exists a \in \mathrm{OT}(\gamma = o(a))$.
3.1. $\gamma = 0$: trivial.
3.2. $\gamma = \xi + \eta$: ..... $\xi = o((a_0, ..., a_n))$, $\eta = o((b_0, ..., b_m))$ with $a_i, b_i \in \mathrm{OT}$ and $a_n \preceq \ldots \preceq a_0$, $b_m \preceq \ldots \preceq b_0$. Let $b_m \preceq a_k$ and $a_{k+1}, ..., a_n \prec b_m$. Then $a := (a_0, ..., a_k, b_0, ..., b_m) \in \mathrm{OT}$ and $\gamma = o(a)$.
3.3. $\gamma = \omega^{\Omega+\gamma_0}$ with $\gamma_0 \in C(\varepsilon_{\Omega+1})$: By IH there is $a_0 \in \mathrm{OT}$ with $\gamma_0 = o(a_0)$. Hence $\gamma = o(D_1 a_0) \;\&\; D_1 a_0 \in \mathrm{OT}$.
3.4. $\gamma = \psi(\xi)$ with $\xi < \alpha \;\&\; \xi \in C(\varepsilon_{\Omega+1}) \;\&\; \xi \in C(\xi)$:
By IH we have $b \in \mathrm{OT}$ with $\xi = o(b)$. $o(b) \in C(\mathrm{o}(b)) \overset{\text{L.8.5}}{\Rightarrow} \mathrm{K}^* b \prec b \Rightarrow D_0 b \in \mathrm{OT}$ and $\gamma = o(D_0 b)$.

Definition $\mathrm{OT}_0 := \{a \in \mathrm{OT} : a \prec D_1 0\}$

Lemma 8.7
a) $\psi(\varepsilon_{\Omega+1}) = \|(\mathrm{OT}_0, \prec)\|$ (order type of the wellordering $(\mathrm{OT}_0, \prec)$)
b) For each $a \in \mathrm{OT}$ such that $a \prec D_1 a$ we have $o(a) = \sup\{o(b){+}1 : b \in \mathrm{OT} \;\&\; b \prec a\}$.

Proof:
$a \mapsto o(a)$ maps $\{a \in \mathrm{OT} : a \prec D_1 0\}$ order preserving onto $C(\varepsilon_{\Omega+1}) \cap \Omega = \psi(\varepsilon_{\Omega+1})$. Hence $\|(\mathrm{OT}_0, \prec)\| = \psi(\varepsilon_{\Omega+1})$ and further (for $a \in \mathrm{OT}_0$) $o(a) = \sup\{\beta{+}1 : \beta < o(a)\} = \sup\{o(b){+}1 : b \in \mathrm{OT} \;\&\; b \prec D_1 0 \;\&\; o(b) < o(a)\} = \sup\{o(b){+}1 : b \in \mathrm{OT} \;\&\; b \prec a\}$.

§9 Ordinal analysis of ID_1

Inductive Definitions

Let M be a set and $\Phi : \mathcal{P}(M) \to \mathcal{P}(M)$ *monotone*, i.e. $\forall X, Y \in \mathcal{P}(M)(X \subseteq Y \Rightarrow \Phi(X) \subseteq \Phi(Y))$.

$\mathrm{I}_\Phi := \bigcap\{X \in \mathcal{P}(M) : \Phi(X) \subseteq X\}$ is called *the inductively defined set given by* Φ .

Lemma 9.1

(1) $\Phi(\mathrm{I}_\Phi) = \mathrm{I}_\Phi$.

(2) $\Phi(X) \subseteq X \implies \mathrm{I}_\Phi \subseteq X$, for each set $X \subseteq M$.

(I_Φ is the least Φ-closed set and the least fixed point of Φ.)

Proof:

(2) trivial.

(1) (i) $\Phi(X) \subseteq X \overset{(2)}{\Rightarrow} \mathrm{I}_\Phi \subseteq X \overset{\Phi \text{ mon.}}{\Longrightarrow} \Phi(\mathrm{I}_\Phi) \subseteq \Phi(X) \subseteq X$.

Hence $\Phi(\mathrm{I}_\Phi) \subseteq \bigcap\{X \in \mathcal{P}(M) : \Phi(X) \subseteq X\} = \mathrm{I}_\Phi$.

(ii) Let $Y := \Phi(\mathrm{I}_\Phi)$. (i) $\overset{\Phi \text{ mon.}}{\Longrightarrow} \Phi(Y) \subseteq \Phi(\mathrm{I}_\Phi) = Y \overset{(2)}{\Rightarrow} \mathrm{I}_\Phi \subseteq Y = \Phi(\mathrm{I}_\Phi)$.

Definition $\mathrm{I}^\alpha_\Phi := \Phi(\mathrm{I}^{<\alpha}_\Phi)$, where $\mathrm{I}^{<\alpha}_\Phi := \bigcup_{\xi<\alpha} \mathrm{I}^\xi_\Phi \quad (\alpha \in On)$.

Lemma 9.2

a) $\beta < \alpha \Rightarrow \mathrm{I}^\beta_\Phi \subseteq \mathrm{I}^\alpha_\Phi$.

b) $\mathrm{I}^{<\delta}_\Phi = \mathrm{I}^\delta_\Phi \Rightarrow \mathrm{I}^\delta_\Phi = \mathrm{I}_\Phi = \bigcup_{\alpha\in On} \mathrm{I}^\alpha_\Phi$.

c) $\exists\delta < card(M)^+(\mathrm{I}^\delta_\Phi = \mathrm{I}^{<\delta}_\Phi)$.

Beweis:

b) 1. By induction on α we prove $\mathrm{I}^\alpha_\Phi \subseteq \mathrm{I}_\Phi$: IH $\Rightarrow \mathrm{I}^{<\alpha}_\Phi \subseteq \mathrm{I}_\Phi \Rightarrow \mathrm{I}^\alpha_\Phi = \Phi(\mathrm{I}^{<\alpha}_\Phi) \subseteq \Phi(\mathrm{I}_\Phi) = \mathrm{I}_\Phi$.

2. From $\Phi(\mathrm{I}^\delta_\Phi) = \Phi(\mathrm{I}^{<\delta}_\Phi) = \mathrm{I}^\delta_\Phi$ it follows that $\mathrm{I}_\Phi \subseteq \mathrm{I}^\delta_\Phi$. Hence $\mathrm{I}^\delta_\Phi \subseteq \bigcup_{\alpha\in On} \mathrm{I}^\alpha_\Phi \subseteq \mathrm{I}_\Phi \subseteq \mathrm{I}^\delta_\Phi$.

c) By a) we have $\mathrm{I}^{<\alpha}_\Phi \subseteq \mathrm{I}^\alpha_\Phi$. — Assumption: $\mathrm{I}^{<\alpha}_\Phi \subsetneq \mathrm{I}^\alpha_\Phi$ for all $\alpha < card(M)^+$. Then there is a mapping $f : card(M)^+ \to M$ with $f(\alpha) \in \mathrm{I}^\alpha_\Phi \setminus \mathrm{I}^{<\alpha}_\Phi$ for all $\alpha \in card(M)^+$, i.e. f injective. Contradiction.

Definition For $b \in \mathrm{I}_\Phi$ let $|b|_\Phi := \min\{\alpha : b \in \mathrm{I}^\alpha_\Phi\}$.

Corollary. For $b \in \mathrm{I}_\Phi$ the following holds:

a) $b \in \mathrm{I}^\alpha_\Phi \Leftrightarrow |b|_\Phi \leq \alpha$, $\quad b \in \mathrm{I}^{<\alpha}_\Phi \Leftrightarrow |b|_\Phi < \alpha$

b) $|b|_\Phi = \sup\{|c|_\Phi + 1 : c \in \mathrm{I}^{<|b|_\Phi}_\Phi\}$

Proof of b)"$\leq$": Let $\alpha := righthandside$. Then $\mathrm{I}^{<|b|_\Phi}_\Phi \subseteq \mathrm{I}^{<\alpha}_\Phi$ and thus $\mathrm{I}^{|b|_\Phi}_\Phi \subseteq \mathrm{I}^\alpha_\Phi$. Hence $b \in \mathrm{I}^\alpha_\Phi$, i.e. $|b|_\Phi \leq \alpha$.

Definition

If $\prec$ is a binary relation on M, and $\Phi(X) = \{a \in M : \forall b{\prec}a(b \in X)\}$ $(X \in \mathcal{P}(M))$, then

$\mathrm{Acc}(\prec) := \mathrm{Acc}(M, \prec) := \mathrm{I}_\Phi$ $(= \bigcup\{X \subseteq M : \forall a \in M(\forall b{\prec}a(b \in X) \Rightarrow a \in X)\})$.

Lemma 9.3

a) $a \in \mathrm{Acc}(\prec) \iff \forall b{\prec}a(b \in \mathrm{Acc}(\prec)) \quad (a \in M)$

b) $|a|_\Phi = \sup\{|b|_\Phi + 1 : b \prec a\} \quad (a \in \mathrm{Acc}(\prec))$

c) $\prec|_{\mathrm{Acc}(\prec)}$ is wellfounded

d) $S \subseteq M$ & $\forall a{\in}S\forall b{\prec}a(b \in S)$ & $\prec\!|_S$ wellfounded $\Longrightarrow$ $S \subseteq \mathrm{Acc}(\prec)$

e) $\prec$ wellfounded $\Longrightarrow$ $M = \mathrm{Acc}(\prec)$

Proof:

a) follows from 9.1(1).

b) $|a|_\Phi \le \alpha \;\Leftrightarrow\; a \in \Phi(\mathrm{I}_\Phi^{<\alpha}) = \{c \in M : \forall b{\prec}c(b \in \mathrm{I}_\Phi^{<\alpha})\} \;\Leftrightarrow\; \forall b{\prec}a(b \in \mathrm{I}_\Phi^{<\alpha}) \;\Leftrightarrow\; \forall b{\prec}a(|b|_\Phi < \alpha)\}$.

$|a|_\Phi = \min\{\alpha : |a|_\Phi \le \alpha\} = \min\{\alpha : \forall b{\prec}a(|b|_\Phi < \alpha)\} = \sup\{|b|_\Phi + 1 : b \prec a\}$.

c) $b \prec a \in \mathrm{Acc}(\prec) \;\Rightarrow\; |b|_\Phi < |a|_\Phi$.

d) By induction over $\prec\!|S$ we prove: $a \in S \Rightarrow a \in \mathrm{Acc}(\prec)$.

$a \in S \Rightarrow \forall b{\prec}a(b \in S) \overset{\mathrm{IH}}{\Rightarrow} \forall b{\prec}a(b \in \mathrm{Acc}(\prec)) \Rightarrow a \in \mathrm{Acc}(\prec)$.

e) trivial.

Example

$\mathcal{O}(X,x) := x{=}0 \vee (\pi_1 x{=}1 \wedge X\pi_2 x) \vee (\pi_1 x{=}2 \wedge \forall y \exists z(\{\pi_2 x\}(y) \simeq z \wedge Xz))$

Lemma 9.4

$|0|_{\mathcal{O}} = 0$, $|\pi(1,a)|_{\mathcal{O}} = |a|_{\mathcal{O}} + 1$, $|\pi(2,e)|_{\mathcal{O}} = \sup\{|\{e\}(n)|_{\mathcal{O}} + 1 : n \in \mathbb{N}\}$

Lemma 9.5

Let $b \prec a \;:\Leftrightarrow\; a = \pi(1,b) \vee (\pi_1 a = 2 \wedge \exists y(\{\pi_2 a\}(y) \simeq b))$.

Then $\mathcal{O} \subseteq \mathrm{Acc}(\prec)$ and $\forall a \in \mathcal{O}(\,|a|_{\mathcal{O}} = |a|_\prec\,)$.

Proof:

1. $a = 0$: $0 \in \mathrm{Acc}$, since $\{c : c \prec 0\}$. $|0|_\prec = 0 = |0|_{\mathcal{O}}$.

2. $a = \pi(1,a_0)$ and $a_0 \in \mathcal{O}$: IH $\Rightarrow$ $a_0 \in \mathrm{Acc}$ & $|a_0|_{\mathcal{O}} = |a_0|_\prec$ $\Rightarrow$ [since $\{c : c \prec a\} = \{a_0\}$]
$a \in \mathrm{Acc}$ & $|a|_\prec = |a_0|_\prec + 1 = |a_0|_{\mathcal{O}} + 1 = |a|_{\mathcal{O}}$.

3. $a = \pi(2,e)$ & $\forall y(\{e\}(y) \in \mathcal{O})$: IH $\Rightarrow$ $\forall y(\{e\}(y) \in \mathrm{Acc})$ & $\forall y(|\{e\}(y)|_{\mathcal{O}} = |\{e\}(y)|_\prec)$ $\Rightarrow$
[since $\{c : c \prec a\} = \{\{e\}(y) : y \in \mathbb{N}\}$] $|a|_\prec = \sup\{|\{e\}|_\prec + 1 : n \in \mathbb{N}\} = \sup\{|\{e\}|_{\mathcal{O}} + 1 : n \in \mathbb{N}\} = |a|_{\mathcal{O}}$.

The formal theory ID_1

A *positive operator form* is an $\mathcal{L}_0(\mathcal{X})$-formula $\mathcal{A}$, such that

– $\mathrm{FV}(\mathcal{A}) \subseteq \{v_0\}$,

– $\mathcal{A}$ contains no set variable except X_0,

– X_0 occurs only positively in $\mathcal{A}$.

Notation: $\mathcal{A}(F,t)$ denotes the result of replacing in $\mathcal{A}_{v_0}(t)$ every subformula $X_0 s$ by $F(s)$.

$\Phi_{\mathcal{A}}(X) := \{n \in \mathbb{N} : \mathbb{N} \models \mathcal{A}(X,n)\}$, $(X \subseteq \mathbb{N})$. $\Phi_{\mathcal{A}}$ is monotone.

Abbreviation: We write $\mathrm{I}_{\mathcal{A}}$, $\mathrm{I}_{\mathcal{A}}^\alpha$, $|n|_{\mathcal{A}}, \ldots.$ for $\mathrm{I}_{\Phi_{\mathcal{A}}}$, $\mathrm{I}_{\Phi_{\mathcal{A}}}^\alpha$, $|n|_{\Phi_{\mathcal{A}}}, \ldots.$

$\mathcal{L}(\mathrm{ID}_1) := \mathcal{L}_0(\mathcal{X}) \cup \{P_{\mathcal{A}} : \mathcal{A}$ positive operator form $\}$ where the $P_{\mathcal{A}}$'s are new unary predicate symbols.

The axioms of ID_1 are the axioms of $\mathcal{Z}$ in the language $\mathcal{L}(\mathrm{ID}_1)$, and

(ID1) $\forall x(\mathcal{A}(P_{\mathcal{A}},x) \to P_{\mathcal{A}}x)$,

(ID2) for each $\mathcal{L}(\mathrm{ID}_1)$-formula F the universal closure of
$\forall x(\mathcal{A}(F,x) \to F(x)) \to \forall x(P_{\mathcal{A}}x \to F(x))$.

Lemma 9.6

$\mathrm{ID}_1 \vdash \forall x(P_{\mathcal{A}}x \to \mathcal{A}(P_{\mathcal{A}}, x))$.

Proof: Let $F(x) := \mathcal{A}(P_{\mathcal{A}}, x)$.

(ID1) $\Rightarrow \forall x(F(x) \to P_{\mathcal{A}}x) \Rightarrow \forall x(\mathcal{A}(F,x) \to \underbrace{\mathcal{A}(P_{\mathcal{A}}, x)}_{F(x)}) \Rightarrow \forall x(P_{\mathcal{A}}x \to F(x))$.

Question: $\mathrm{ID}_1 \vdash P_{\mathcal{A}}\underline{n} \Longrightarrow |\underline{n}|_{\mathcal{A}} <???$

Remark: Let $\mathcal{A}(X,x) = \forall y{\prec}x Xy$.

Then $\mathrm{Prog}(X) = \forall x(\mathcal{A}(X,x) \to Xx)$. Hence $\mathrm{ID}_1 \vdash \mathrm{Prog}(P_{\mathcal{A}})$.

$\mathrm{ID}_1 \vdash \mathrm{TI}_{\prec}(X, \underline{n}) \Longrightarrow \mathrm{ID}_1 \vdash P_{\mathcal{A}}\underline{n}$.

Proof: $\mathrm{ID}_1 \vdash \mathrm{TI}_{\prec}(X, \underline{n}) \Longrightarrow \mathrm{TI}_{\prec}(P_{\mathcal{A}}, \underline{n}) \Rightarrow \mathrm{ID}_1 \vdash \forall y{\prec}\underline{n} P_{\mathcal{A}} \Rightarrow \mathrm{ID}_1 \vdash P_{\mathcal{A}}\underline{n}$.

The language $\mathcal{L}^*$

For each positive operator form $\mathcal{A}$ and each ordinal $\alpha \le \Omega$ let $P_{\mathcal{A}}^{<\alpha}$ be a new unary predicate symbol.

$\mathcal{L}^* := \mathcal{L}_0(\mathcal{X}) \cup \{P_{\mathcal{A}}^{<\alpha} : \alpha \le \Omega\,,\ \mathcal{A} \text{ positive operator form }\}$

We assume that $P_{\mathcal{A}}^{<\Omega}$ equals $P_{\mathcal{A}}$. Hence $\mathcal{L}(\mathrm{ID}_1) \subseteq \mathcal{L}^*$.

Definition of $\mathrm{rk}(A)$ for each $\mathcal{L}^*$-formula A

1. $\mathrm{rk}(P^{<\alpha}t) := \mathrm{rk}(\neg P^{<\alpha}t) := \omega \cdot \alpha$
2. $\mathrm{rk}(A) := 0$, if A is an $\mathcal{L}_0(\mathcal{X})$-literal
3. $\mathrm{rk}(A \wedge B) := \mathrm{rk}(A \vee B) = \max\{\mathrm{rk}(A), \mathrm{rk}(B)\} + 1$
4. $\mathrm{rk}(\forall x A) := \mathrm{rk}(\exists x A) := \mathrm{rk}(A){+}1$

Definition of $\mathsf{k}_\Pi(A)$ for each $\mathcal{L}^*$-formula A

1. $\mathsf{k}_\Pi(P^{<\alpha}t) := \{0\}$, $\mathsf{k}_\Pi(\neg P^{<\alpha}t) := \{0, \alpha\}$.
2. $\mathsf{k}_\Pi(A) := \{0\}$, if A is an $\mathcal{L}_0(\mathcal{X})$-literal.
3. $\mathsf{k}_\Pi(A \wedge B) := \mathsf{k}_\Pi(A \vee B) := \mathsf{k}_\Pi(A) \cup \mathsf{k}_\Pi(B)$.
4. $\mathsf{k}_\Pi(\forall x A) := \mathsf{k}_\Pi(\exists x A) := \mathsf{k}_\Pi(A)$.

Definition

$\mathsf{k}_\Sigma(A) := \mathsf{k}_\Pi(\neg A)$ and $\mathsf{k}(A) := \mathsf{k}_\Pi(A) \cup \mathsf{k}_\Sigma(A)$. (Hence $\mathsf{k}(A) = \mathsf{k}(\neg A)$.)

$\mathsf{k}(t) := \emptyset$ for $t \in T$, $\quad \mathsf{k}(\xi) := \{\xi\}$ for $\xi \in On$.

Now we 'embed' $\mathcal{L}^*$ into infinitary propositional logic:

Definition

$P_{\mathcal{A}}^{<\alpha}t \simeq \bigvee_{\xi<\alpha}\mathcal{A}(P_{\mathcal{A}}^{<\xi}, t)$, $\quad \neg P_{\mathcal{A}}^{<\alpha}t \simeq \bigwedge_{\xi<\alpha}\neg\mathcal{A}(P_{\mathcal{A}}^{<\xi}, t)$,

$A \wedge B \simeq \bigwedge_{\iota\in\{\underline{0},\underline{1}\}} A_\iota$, $\quad A \vee B \simeq \bigvee_{\iota\in\{\underline{0},\underline{1}\}} A_\iota$, $\quad$ where $A_{\underline{0}} := A$, $A_{\underline{1}} := B$,

$\forall x A \simeq \bigwedge_{t\in T} A_x(t)$, $\quad \exists x A \simeq \bigvee_{t\in T} A_x(t)$

Lemma 9.7

a) $A \simeq *_{\iota\in J} A_\iota \Longrightarrow \mathsf{k}_\Pi(A_\iota) \subseteq \mathsf{k}_\Pi(A) \cup \mathsf{k}(\iota)$,

b) $A \simeq \bigwedge_{\iota\in J} A_\iota \Longrightarrow \exists\sigma \in \mathsf{k}_\Pi(A) \forall\iota \in J[\,\mathsf{k}(\iota) \subseteq \sigma\,]$.

c) $\mathrm{rk}(A) = \omega \cdot \alpha + n$ with $\alpha = \max(\mathsf{k}(A))$ and $n < \omega$.

d) $\mathrm{rk}(A) = \Omega \implies A = P_{\mathcal{A}}t$ or $A = \neg P_{\mathcal{A}}t$.

The infinitary proof system ID_1^∞

The language of ID_1^∞ is $\mathcal{L}^*$.

The inference symbols of ID_1^∞ are

$(\mathsf{Ax}_\Delta^\infty)$	Δ	if $\Delta \in \mathcal{AX}(\mathsf{Z}^\infty)$
$(\bigwedge_A)$	$\dfrac{\ldots A_\iota \ldots (\iota \in J)}{A}$	if $A \simeq \bigwedge_{\iota\in J} A_\iota$
$(\bigvee_A^\mu)$	$\dfrac{A_\mu}{A}$	if $A \simeq \bigvee_{\iota\in J} A_\iota$ and $\mu \in J$
(Cut_C)	$\dfrac{C \quad \neg C}{\emptyset}$	
(Rep)	$\dfrac{\emptyset}{\emptyset}$	
$(\mathsf{Cl}_{P_{\mathcal{A}}t})$	$\dfrac{\mathcal{A}(P_{\mathcal{A}}, t)}{P_{\mathcal{A}}t}$	

Definition of $\mathsf{k}(\mathcal{I})$ for each ID_1^∞-inference $\mathcal{I}$

$\mathsf{k}(\mathsf{Ax}_\Delta^\infty) := \mathsf{k}(\mathsf{Rep}) := \emptyset$, $\mathsf{k}(\bigwedge_A) := \mathsf{k}(\mathsf{Cut}_A) := \mathsf{k}(\mathsf{Cl}_A) := \mathsf{k}(A)$, $\mathsf{k}(\bigvee_A^\mu) := \mathsf{k}(\mu)$.

Inductive Definition of ID_1^∞-derivations

If $\mathcal{I}$ is an inference symbol of ID_1^∞, α is an ordinal, and $(d_\iota)_{\iota\in|\mathcal{I}|}$ is a family of ID_1^∞-*derivations* such that $o(d_\iota) < \alpha$ for all $\iota \in |\mathcal{I}|$, then $d := \alpha\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|} := \dfrac{\ldots d_\iota \ldots}{\alpha : \mathcal{I}}$ is an ID_1^∞-*derivation* with $o(d) := \alpha$.

The cut-elimination operators $\mathcal{R}_C$ and $\mathcal{E}$ are assumed to be adapted to this modified notion of derivation where for $\mathcal{E}$ we use $\omega^{\Omega+\alpha}$ instead of 3^α (cf. proofs of 9.8 and 9.9).

Definition

A set $H \subseteq On$ is called *nice* iff $\{0,1\} \subseteq H$ and H is closed under $\#$, $\lambda\xi.\,\omega\cdot\xi$, $\lambda\xi.\,\omega^{\Omega+\xi}$.

$\tilde{H} := \{\iota : \mathsf{k}(\iota) \subseteq H\}$.

Convention: In the following H is used as syntactical variable for nice sets.

Remark: Every $C(\gamma)$ with $\gamma \geq \omega$ is nice.

Definition

Let $d = \alpha\mathcal{I}(d_\iota)_{\iota\in I}$: $d \triangleleft H$ (d is H-*controlled*) iff $\mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq H$ and $\forall \iota \in I \cap \tilde{H}(d_\iota \triangleleft H)$.

Theorem 9.8

$d_0 \triangleleft H$ & $d_1 \triangleleft H$ & $\mathrm{rk}(C) \neq \Omega \implies \mathcal{R}_C(d_0, d_1) \triangleleft H$.

Proof:

Abbr.: $\alpha := \mathrm{o}(d_0)$, $\beta := \mathrm{o}(d_1)$.

(Case 1) C is not main formula of $\mathcal{I} := \mathsf{last}(d_0)$:

We have $\mathcal{R}_C(d_0, d_1) = (\alpha\#\beta)\mathcal{I}(\mathcal{R}_C(d_{0\iota}, d_1))_{\iota\in I}$ where $d_0 = \alpha\mathcal{I}(d_{0\iota})_{\iota\in I}$.

Since $d_0 \triangleleft H$, we also have $\mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq H$ and $d_{0\iota} \triangleleft H$ for all $\iota \in I \cap \tilde{H}$.

From this (together with $\beta \in H$) we get $\mathsf{k}(\mathcal{I}) \cup \{\alpha\#\beta\} \subseteq H$, and (by IH) $\mathcal{R}_C(d_{0\iota}, d_1) \triangleleft H$ for all $\iota \in I \cap \tilde{H}$, i.e. $\mathcal{R}_C(d_0, d_1) \triangleleft H$.

(Case 2) $C \simeq \bigvee_{\iota\in J} C_\iota$ is main formula of $\mathsf{last}(d_0)$, and $\neg C \simeq \bigwedge_{\iota\in J} \neg C_\iota$ is main formula of $\mathsf{last}(d_1)$:
Then $\mathsf{last}(d_1) = \bigwedge_{\neg C}$, and, since $\mathrm{rk}(C) \neq \Omega$, $\mathsf{last}(d_0) = \bigvee_C^\mu$.
Hence $\mathcal{R}_C(d_0, d_1) = (\alpha\#\beta)\mathsf{Cut}_{C_\mu}\mathcal{R}_C(d_{00}, d_1)\mathcal{R}_C(d_0, d_{1\mu})$, where $d_0 = \alpha\bigvee_C^\mu d_{00}$ and $d_1 = \beta\bigwedge_{\neg C}(d_{1\iota})_{\iota\in J}$.

We have $\mathsf{k}(C_\mu) \cup \{\alpha, \beta\} \subseteq \mathsf{k}(\mu) \cup \mathsf{k}(C) \cup \{\alpha, \beta\} \subseteq H$, $d_{00} \triangleleft H$, and $d_{1\mu} \triangleleft H$ (since $\mathsf{k}(\mu) \subseteq H$).

From $\alpha, \beta \in H$ we get $\alpha\#\beta \in H$. By IH we get $\mathcal{R}_C(d_{00}, d_1) \triangleleft H$ and $\mathcal{R}_C(d_0, d_{1\mu}) \triangleleft H$. Together with $\mathsf{k}(C_\mu) \cup \{\alpha\#\beta\} \subseteq H$ this yields $\mathcal{R}_C(d_0, d_1) \triangleleft H$.

Theorem 9.9
a) $d \vdash^\alpha_{\rho+1} \Gamma$ & $\rho \neq \Omega \implies \mathcal{E}(d) \vdash^{\omega^{\Omega+\alpha}}_\rho \Gamma$,
b) $d \triangleleft H \implies \mathcal{E}(d) \triangleleft H$.

Proof of b):

(Case 1) $d = \alpha\mathsf{Cut}_C d_0 d_1$ with $\mathrm{rk}(C) \neq \Omega$.

Then $\alpha \in H$, $d_0, d_1 \triangleleft H$ and $\mathcal{E}(d) = \omega^{\Omega+\alpha}\mathsf{Rep}\mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1))$.

By IH $\mathcal{E}(d_0), \mathcal{E}(d_1) \triangleleft H$, and therefore by 9.8 $\mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1)) \triangleleft H$.

From $\alpha \in H$ we get $\omega^{\Omega+\alpha} \in H$ and thus $\mathcal{E}(d) \triangleleft H$.

(Case 2) otherwise: Then $d = \alpha\mathcal{I}(d_\iota)_{\iota\in I}$ and $\mathcal{E}(d) = \omega^{\Omega+\alpha}\mathcal{I}(\mathcal{E}(d_\iota))_{\iota\in I}$.

$d \triangleleft H \;\Rightarrow\; \mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq H$ & $d_\iota \triangleleft H$ for $\iota \in I \cap \tilde{H} \overset{\text{IH}}{\Rightarrow}$
$\mathsf{k}(\mathcal{I}) \cup \{\omega^{\Omega+\alpha}\} \subseteq H$ & $\mathcal{E}(d_\iota) \triangleleft H$ for $\iota \in I \cap \tilde{H} \;\Rightarrow\; \mathcal{E}(d) \triangleleft H$.

Definition
For each $\mathcal{L}(\mathrm{ID}_1)$-formula B let $B[\alpha, \beta]$ denote the result of replacing in B every negative occurrence of $P_\mathcal{A}$ by $P_\mathcal{A}^{<\alpha}$, and every positive occurrence of $P_\mathcal{A}$ by $P_\mathcal{A}^{<\beta}$.

$\Gamma[\alpha, \beta] := \{B[\alpha, \beta] : B \in \Gamma\}$

$\models \{A_1, ..., A_n\} \;:\Longleftrightarrow\;$ $A_1 \vee ... \vee A_n$ is true in the standard interpretation where the set variables may be interpreted arbitrarily.

Definition Let $\mathcal{H}$ be a collection of nice sets:
$d \lhd^* \mathcal{H} \;:\Leftrightarrow\; \forall H \in \mathcal{H}(d \lhd H)$, $\quad \mathcal{H}[\iota] := \{H \in \mathcal{H} : \mathsf{k}(\iota) \subseteq H\}$.

Lemma 9.10
Let $d = \alpha\mathcal{I}(d_\iota)_{\iota \in I}$: $\; d \lhd^* \mathcal{H} \;\Leftrightarrow\; \mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq \bigcap \mathcal{H} \;\&\; \forall \iota \in I(d_\iota \lhd^* \mathcal{H}[\iota])$.

Proof:
$d \lhd^* \mathcal{H} \;\Leftrightarrow$
$\forall H \in \mathcal{H}(d \lhd H) \;\Leftrightarrow$
$\forall H \in \mathcal{H}(\mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq H \;\&\; \forall \iota \in I \cap \tilde{H}(d_\iota \lhd H)) \;\Leftrightarrow$
$\mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq \bigcap \mathcal{H} \;\&\; \forall \iota \in I \forall H \in \mathcal{H}(\mathsf{k}(\iota) \subseteq H \Rightarrow d_\iota \lhd H) \;\Leftrightarrow$
$\mathsf{k}(\mathcal{I}) \cup \{\alpha\} \subseteq \bigcap \mathcal{H} \;\&\; \forall \iota \in I(d_\iota \lhd^* \{H \in \mathcal{H} : \mathsf{k}(\iota) \subseteq H\})$.

Abbreviation: $\mathcal{H}_\gamma := \{C(\xi) : \xi \geq \gamma\}$. — Note that $\bigcap \mathcal{H}_\gamma = C(\gamma)$, and $\mathcal{H}_\gamma[\iota] = \mathcal{H}_\gamma$ if $\mathsf{k}(\iota) \subseteq C(\gamma)$.

Theorem 9.11
$d \vdash^\alpha_{\Omega+1} \Gamma \;\&\; \omega \leq \gamma \in C(\gamma) \;\&\; d \lhd^* \mathcal{H}_\gamma \implies \models \Gamma[\psi\gamma, \psi(\gamma + \omega^{\Omega+\alpha})]$

Proof:

1. $d = \left\{ \dfrac{\begin{array}{cc} d_0 & d_1 \\ | & | \\ \Gamma, Pt : \alpha_0 & \Gamma, \neg Pt : \alpha_1 \end{array}}{\Gamma : \alpha} (\mathsf{Cut}_{Pt}) \right.$: $\quad$ Let $\gamma_0 := \gamma + \omega^{\Omega+\alpha_0}$, $\gamma_1 := \gamma_0 + \omega^{\Omega+\alpha_1}$, $\tilde{\gamma} := \gamma + \omega^{\Omega+\alpha}$.

$d \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; d_0, d_1 \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; d_0 \lhd^* \mathcal{H}_\gamma \;\&\; d_1 \lhd^* \mathcal{H}_{\gamma_0}$. $\;d_0 \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; \alpha_0 \in C(\gamma) \;\Rightarrow\; \gamma_0 \in C(\gamma_0)$.

IH $\implies \models \Gamma, Pt\,[\psi\gamma, \psi\gamma_0]$ and $\models \Gamma, \neg Pt\,[\psi\gamma_0, \psi\gamma_1]$ (and $\gamma < \gamma_0 < \gamma_1 < \tilde{\gamma}$)

$\qquad \implies \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], P^{<\psi\gamma_0}t$ and $\models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \neg P^{<\psi\gamma_0}t \implies \models \Gamma[\psi\gamma, \psi\tilde{\gamma}]$

For the rest of this proof we set $\gamma_\iota := \gamma + \omega^{\Omega+\alpha_\iota}$ and $\tilde{\gamma} := \gamma + \omega^{\Omega+\alpha}$.

2. $d = \left\{ \dfrac{\begin{array}{cc} d_0 & d_1 \\ | & | \\ \Gamma, A : \alpha_0 & \Gamma, \neg A : \alpha_1 \end{array}}{\Gamma : \alpha} (\mathsf{Cut}_{A}) \right.$ with $\mathrm{rk}(A) < \Omega$:

IH $\implies \models \Gamma[\psi\gamma, \psi\gamma_0], A$ and $\models \Gamma[\psi\gamma, \psi\gamma_1], \neg A \implies \models \Gamma[\psi\gamma, \psi\tilde{\gamma}]$.

3. $d = \left\{ \dfrac{\begin{array}{c} d_0 \\ | \\ \Gamma, \mathcal{A}(P,t) : \alpha_0 \end{array}}{\Gamma : \alpha} (\mathsf{Cl}_{Pt}) \right.$:

$d_0 \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; \alpha_0 \in C(\gamma) \;\Rightarrow\; \gamma_0 \in C(\gamma_0) \;\Rightarrow\; \psi\gamma_0 < \psi\tilde{\gamma}\; (*)$.

IH $\implies \models \Gamma[\psi\gamma, \psi\gamma_0], \mathcal{A}(P^{<\psi\gamma_0}, t) \overset{(*)}{\implies} \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], P^{<\psi\tilde{\gamma}}t \quad (= \Gamma[\psi\gamma, \psi\tilde{\gamma}]$, since $Pt \in \Gamma)$.

4. $d = \left\{ \dfrac{\begin{array}{c} d_\xi \\ | \\ ...\Gamma, \neg\mathcal{A}(P^{<\xi}, t) : \alpha_\xi ...(\xi < \Omega) \end{array}}{\Gamma : \alpha} (\bigwedge_{\neg Pt}) \right.$: Note that $\{\xi \in \Omega : \mathsf{k}(\xi) \subseteq C(\gamma)\} = \psi(\gamma)$.

$d \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; \forall \xi < \psi\gamma(d_\xi \lhd^* \mathcal{H}_\gamma[\xi] = \mathcal{H}_\gamma) \overset{\text{IH}}{\Rightarrow} \forall \xi < \psi\gamma \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \neg\mathcal{A}(P^{<\xi}, t) \;\Rightarrow$
$\models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \neg P^{<\psi\gamma} \quad (= \Gamma[\psi\gamma, \psi\tilde{\gamma}]$, since $\neg Pt \in \Gamma)$.

5. $d = \begin{cases} \dfrac{\begin{matrix} d_\xi \\ | \\ ...\Gamma, \neg\mathcal{A}(P^{<\xi}, t) : \alpha_\xi ...(\xi < \delta) \end{matrix}}{\Gamma : \alpha} (\bigwedge_{\neg P^{<\delta} t}) \end{cases}$ with $\delta < \Omega$:

$d \lhd^* \mathcal{H}_\gamma \Rightarrow \forall \xi \in \delta \cap \psi(\gamma)(d_\xi \lhd^* \mathcal{H}_\gamma[\xi] = \mathcal{H}_\gamma)$ & $\{\delta\} = \mathsf{k}(\bigwedge_{\neg P^{<\delta} t}) \subseteq C(\gamma) \cap \Omega = \psi(\gamma) \;\Rightarrow\; (\forall \xi < \delta) d_\xi \lhd^* \mathcal{H}_\gamma \overset{\mathrm{IH}}{\Rightarrow}$ $(\forall \xi < \delta) \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \neg\mathcal{A}(P^{<\xi}, t) \;\Rightarrow \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \neg P^{<\delta} t$.

6. $d = \begin{cases} \dfrac{\begin{matrix} d_0 \\ | \\ \Gamma, \mathcal{A}(P^{<\mu}, t) : \alpha_0 \end{matrix}}{\Gamma : \alpha} (\bigvee^{\mu}_{P^{<\delta} t}) \end{cases}$ with $\mu < \delta \le \Omega$:

$d \lhd^* \mathcal{H}_\gamma \;\Rightarrow\; d_0 \lhd^* \mathcal{H}_\gamma \overset{\mathrm{IH}}{\Rightarrow} \models \Gamma[\psi\gamma, \psi\tilde{\gamma}], \mathcal{A}(P^{<\mu}, t)$.

It remains to prove $(\delta = \Omega \Rightarrow \mu < \psi\tilde{\gamma})$: $d \lhd^* \mathcal{H}_\gamma \Rightarrow \{\mu\} \subseteq C(\gamma) \;\Rightarrow\; \mu \in C(\gamma) \cap \Omega = \psi(\gamma) \le \psi\tilde{\gamma}$.

Embedding of ID_1 into ID_1^∞

Definition

$\mathcal{H}_\Gamma := \{C(\gamma) : \mathsf{k}(\Gamma) \subseteq C(\gamma) \;\&\; \gamma \ge \omega\}$,

$d \Vdash^\alpha_\rho \Gamma \;:\Longleftrightarrow\; d \vdash^\alpha_\rho \Gamma \;\&\; d \lhd^* \mathcal{H}_\Gamma$,

$\mathrm{ID}_1^\infty \Vdash^\alpha_\rho \Gamma \;:\Longleftrightarrow\; d \Vdash^\alpha_\rho \Gamma$ for some ID_1^∞-derivation d

Remark: $A \in \Gamma \;\Rightarrow\; \mathrm{rk}(A) \in \bigcap \mathcal{H}_\Gamma$

Lemma 9.12 For $\alpha \in \bigcap \mathcal{H}_\Gamma$:

a) $A \simeq \bigwedge_{\iota \in J} A_\iota \in \Gamma \;\&\; \forall \iota \in J(\mathrm{ID}_1^\infty \Vdash^{\alpha_\iota}_\rho \Gamma, A_\iota \;\&\; \alpha_\iota < \alpha) \;\Rightarrow\; \mathrm{ID}_1^\infty \Vdash^\alpha_\rho \Gamma$.

b) $A \simeq \bigvee_{\iota \in J} A_\iota \in \Gamma \;\&\; \mathrm{ID}_1^\infty \Vdash^{\alpha_0}_\rho \Gamma, A_\mu \;\&\; \alpha_0 < \alpha \;\&\; \mu \in J \;\&\; \mathsf{k}(\mu) \subseteq \bigcap \mathcal{H}_\Gamma \;\Rightarrow\; \mathrm{ID}_1^\infty \Vdash^\alpha_\rho \Gamma$.

c) $\mathrm{ID}_1^\infty \Vdash^{\alpha_0}_\rho \Gamma, A \;\&\; \mathrm{ID}_1^\infty \Vdash^{\alpha_1}_\rho \Gamma, \neg A \;\&\; \alpha_0, \alpha_1 < \alpha \;\&\; \mathsf{k}(A) \subseteq \bigcap \mathcal{H}_\Gamma \;\Rightarrow\; \mathrm{ID}_1^\infty \Vdash^\alpha_\rho \Gamma$.

Proof:

We have $\mathsf{k}(\Gamma) \cup \{\alpha\} \subseteq \bigcap \mathcal{H}_\Gamma$.

a) Assume $\forall \iota \in J(d_\iota \Vdash^{\alpha_\iota}_\rho \Gamma, A_\iota)$, and let $d := \alpha \bigwedge_A (d_\iota)_{\iota \in J}$.
Then $\mathsf{k}(\bigwedge_A) = \mathsf{k}(A) \subseteq \mathsf{k}(\Gamma)$, and it remains to prove $d_\iota \lhd^* \mathcal{H}_\Gamma[\iota]$:
$d_\iota \Vdash^{\alpha_\iota}_\rho \Gamma, A_\iota \;\Rightarrow\; d_\iota \lhd^* \mathcal{H}_{\Gamma, A_\iota} \overset{(*)}{\Rightarrow} d_\iota \lhd^* \mathcal{H}_\Gamma[\iota]$. $\quad (*)$: $\mathsf{k}(\Gamma, A_\iota) \subseteq \mathsf{k}(\Gamma) \cup \mathsf{k}(\iota)$.

b) Assume $d_0 \Vdash^{\alpha_0}_\rho \Gamma, A_\mu$, and let $d := \alpha \bigvee^\mu_A d_0$. Then $\mathsf{k}(\bigvee^\mu_A) = \mathsf{k}(\mu) \subseteq \bigcap \mathcal{H}_\Gamma$, and it remains to prove $d_0 \lhd^* \mathcal{H}_\Gamma$:
$\mathsf{k}(A_\mu) \subseteq \mathsf{k}(\Gamma) \cup \mathsf{k}(\mu) \subseteq \bigcap \mathcal{H}_\Gamma \;\Rightarrow\; \mathcal{H}_\Gamma = \mathcal{H}_{\Gamma, A_\mu}$. $\quad d_0 \Vdash^{\alpha_0}_\rho \Gamma, A_\mu \;\Rightarrow\; d_0 \lhd^* \mathcal{H}_{\Gamma, A_\mu}$.

c) Assume $d_0 \Vdash^{\alpha_0}_\rho \Gamma, A$, $d_1 \Vdash^{\alpha_1}_\rho \Gamma, \neg A$, and let $d := \alpha \mathsf{Cut}_A d_0 d_1$. It remains to prove $d_i \lhd^* \mathcal{H}_\Gamma$:
$\mathsf{k}(A) \subseteq \bigcap \mathcal{H}_\Gamma \;\Rightarrow\; \mathcal{H}_\Gamma = \mathcal{H}_{\Gamma, (\neg)A}$. $\quad d_i \Vdash^{\alpha_i}_\rho \Gamma, (\neg)A \;\Rightarrow\; d_i \lhd^* \mathcal{H}_{\Gamma, (\neg)A}$.

Lemma 9.13

$t_0, t_1 \in T \;\&\; \mathsf{val}(t_0) = \mathsf{val}(t_1) \;\&\; \mathrm{FV}(A) \subseteq \{x\} \;\Rightarrow\; \mathrm{ID}_1^\infty \Vdash^{2 \cdot \mathrm{rk}(A)}_0 \neg A(x/t_0), A(x/t_1)$.

Proof by induction on $\mathrm{rk}(A)$:

1. $A = (\neg)Xt$ or $(\neg)(r{=}s)$: trivial.

2. $A \simeq \bigvee_{\iota\in J} A_\iota$: Then $A(x/t_i) \simeq \bigvee_{\iota\in J} A_\iota(x/t_i)$.

IH $\;\Rightarrow\;$ $\mathrm{ID}_1^\infty \Vdash_0^{2\cdot\mathrm{rk}(A_\iota)} \neg A_\iota(t_0), A_\iota(t_1) \quad (\forall \iota \in J)$

$\overset{9.12,(*_1)}{\Rightarrow}$ $\mathrm{ID}_1^\infty \Vdash_0^{2\cdot\mathrm{rk}(A_\iota)+1} \neg A_\iota(t_0), A(t_1) \quad (\forall \iota \in J)$

$\overset{9.12,(*_2)}{\Rightarrow}$ $\mathrm{ID}_1^\infty \Vdash_0^{2\cdot\mathrm{rk}(A)} \neg A(t_0), A(t_1)$

$(*_1)$ $\mathsf{k}(\iota) \cup \{2\cdot\mathrm{rk}(A_\iota){+}1\} \subseteq \bigcap \mathcal{H}_{\neg A_\iota(t_0),A(t_1)}$, $(*_2)$ $2\cdot\mathrm{rk}(A) \in \bigcap \mathcal{H}_{\neg A(t_0),A(t_1)}$.

Lemma 9.14

$\mathrm{PL1} \vdash_0^k \Gamma \;\&\; \mathrm{FV}(\Gamma) = \emptyset \implies \mathrm{ID}_1^\infty \Vdash_0^{\Omega+k} \Gamma$

Proof:

1.1. $\{\neg A, A\} \subseteq \Gamma$ with $A = Xt$ or $A = (s{=}t)$: Then $\Gamma \in \mathcal{AX}(\mathsf{Z}^\infty)$.

1.2. $\{\neg Pt, Pt\} \subseteq \Gamma$: The claim follows from 9.13.

2. $\forall x A \in \Gamma$, $k > 0$ and $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(y)$:

Then $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(t)$ for all $t \in T$, and by IH $\mathrm{ID}_1^\infty \Vdash_0^{\Omega+k-1} \Gamma, A_x(t)$ for all $t \in T$.

By 9.12 this implies $\mathrm{ID}_1^\infty \Vdash_0^{\Omega+k} \Gamma$, since $\forall x A \simeq \bigwedge_{t\in T} A_x(t)$.

3. $\exists x A \in \Gamma$, $k > 0$ and $\mathrm{PL1} \vdash_0^{k-1} \Gamma, A_x(t)$: w.l.o.g. $t \in T$.

IH$\Rightarrow$ $\mathrm{ID}_1^\infty \Vdash_0^{\Omega+k-1} \Gamma, A_x(t) \overset{9.12,\,\mathsf{k}(t)=\emptyset}{\Longrightarrow} \mathrm{ID}_1^\infty \Vdash_0^{\Omega+k} \Gamma$.

Lemma 9.15

If $\mathrm{FV}(\forall x F) = \emptyset$, $t \in T$, and $\delta = 2\cdot\mathrm{rk}(F)$ then

$\mathrm{ID}_1^\infty \Vdash_0^{\delta+2\mathsf{val}(t)} \neg F(0), \neg\forall x(F(x) \to F(\mathsf{S}x)), F(t)$.

Proof by induction on n:

1. $\mathsf{val}(t) = 0$: By 9.13 $\mathrm{ID}_1^\infty \Vdash_0^\delta \neg F(0), F(t)$.

2. $\mathsf{val}(t) = n{+}1$: Let $G := \neg\forall x(F(x) \to F(\mathsf{S}x))$.

$$\left.\begin{array}{l}\text{IH} \Rightarrow \Vdash_0^{\delta+2n} \neg F(0), G, F(\underline{n}) \\ 9.13 \Rightarrow \Vdash_0^{\delta} \neg F(\mathsf{S}\underline{n}), F(t)\end{array}\right\} \Rightarrow \Vdash_0^{\delta+2n+1} \neg F(0), G, F(\underline{n}) \wedge \neg F(\mathsf{S}\underline{n}), F(t) \overset{\mathsf{k}(\underline{n})=\emptyset}{\Longrightarrow} \Vdash_0^{\delta+2(n+1)} \neg F(0), G, F(t).$$

Lemma 9.16

If $B(X)$ is an X-positive $\mathcal{L}_1$-formula then for all $\alpha \le \Omega$:

$\mathrm{ID}_1^\infty \Vdash_0^{\delta+\alpha'} \neg\forall x(\mathcal{A}(F,x) \to F(x)), \neg B(P^{<\alpha}), B(F)$, where $\delta := 2\cdot\mathrm{rk}(F)$ and $\alpha' := 2\cdot\mathrm{rk}(B(P^{<\alpha})) + 2$.

Proof by induction on $\mathrm{rk}(B(P^{<\alpha}))$: Let $G := \neg\forall x(\mathcal{A}(F,x) \to F(x))$ $[= \exists x(\mathcal{A}(F,x) \wedge \neg F(x))\,]$.

1. $B(X) \equiv Xt$: Then $\alpha' = 2\omega\alpha + 2 = \omega\alpha + 2$.

IH $\Rightarrow \Vdash_0^{\delta+\omega\xi+m} G, \neg\mathcal{A}(P^{<\xi}, t), \mathcal{A}(F,t)$ for all $\xi < \alpha$

$\Rightarrow \Vdash_0^{\delta+\omega\alpha} G, \neg P^{<\alpha}t, \mathcal{A}(F,t)$

$\Rightarrow \Vdash_0^{\delta+\omega\alpha+1} G, \neg P^{<\alpha}t, \mathcal{A}(F,t) \wedge \neg F(t), F(t)$ $\quad$ [since $\Vdash_0^\delta \neg F(t), F(t)$]

$\Rightarrow \Vdash_0^{\delta+\omega\alpha+2} G, \neg P^{<\alpha}t, F(t)$.

2. B $\mathcal{L}_0$-literal: 9.13.

3. $B \equiv \forall y B_0(X,y)$: Then $\mathrm{rk}(B(P^{<\alpha})) = \beta_0 + 1$ mit $\beta_0 := \mathrm{rk}(B_0(P^{<\alpha}, y))$.
IH $\Rightarrow \Vdash_0^{\delta+2\beta_0+2} G, \neg B_0(P^{<\alpha}, t), B_0(F,t)$ for all $t \in T$
$\Longrightarrow \Vdash_0^{\delta+2\beta_0+3} G, \exists y \neg B_0(P^{<\alpha}, y), B_0(F,t)$ for all $t \in T$
$\Longrightarrow \Vdash_0^{\delta+2(\beta_0+1)+2} G, \neg\forall y B_0(P^{<\alpha}, y), \forall y B_0(F,y)$.

Lemma 9.17
a) $\mathrm{ID}_1^\infty \Vdash_0^{\Omega+\omega} \forall x(\mathcal{A}(P_\mathcal{A}, x) \to P_\mathcal{A} x)$,
b) $\mathrm{ID}_1^\infty \Vdash_0^{\Omega\cdot 2+\omega} \forall \vec{y}[\forall x(\mathcal{A}(F(\cdot,\vec{y}), x) \to F(x,\vec{y})) \to \forall x(P_\mathcal{A} x \to F(x,\vec{y}))]$.

Proof:
a) 9.13 $\Rightarrow \Vdash_0^{\Omega+k_0} \neg\mathcal{A}(P,t), \mathcal{A}(P,t)$ [with $\Omega + k_0 = 2\mathrm{rk}(\mathcal{A}(P,x))$] $\stackrel{(\mathsf{Cl}_{Pt})}{\Longrightarrow} \Vdash_0^{\Omega+k_0+1} \neg\mathcal{A}(P,t), Pt$ $(\forall t \in T)$ $\Rightarrow$ $\Vdash_0^{\Omega+k_0+4} \forall x(\mathcal{A}(P,x) \to Px)$.
b) 9.16 $\Rightarrow \Vdash_0^{\delta+\Omega+2} \neg\forall x(\mathcal{A}(F(\cdot,\vec{s}), x) \to F(x,\vec{s})), \neg Pt, F(t,\vec{s}))$ for all $t, \vec{s} \in T$ [$\delta := 2\mathrm{rk}(F) < \Omega + \omega$]
$\Rightarrow \Vdash_0^{\Omega\cdot 2+5} \neg\forall x(\mathcal{A}(F(\cdot,\vec{s}), x) \to F(x,\vec{s})), \forall x(Px \to F(x,\vec{s}))$ for all $\vec{s} \in T$.

Lemma 9.18 If A is an axiom of ID_1 then $\mathrm{ID}_1^\infty \Vdash_0^{\Omega\cdot 2+\omega} A$.

Proof by means of 9.13, 9.15, 9.17.

Theorem 9.19 $\mathrm{ID}_1 \vdash P_\mathcal{A}\underline{n} \;\Rightarrow\; |n|_\mathcal{A} < \psi(\varepsilon_{\Omega+1})$.

Proof:
Assume $\mathrm{ID}_1 \vdash P\underline{n}$. Then there are ID_1-axioms $A_1, ..., A_l$ such that $\mathrm{PL1} \vdash_0 \neg A_1, ..., \neg A_l, P\underline{n}$.
$\Vdash_0^{\Omega\cdot 2} \neg A_1, ..., \neg A_l, P\underline{n}$ [by 9.14] and $\Vdash_0^{\Omega\cdot 2+\omega} A_i$ $(i = 1, ..., l)$ [by 9.18]
$\stackrel{9.12c}{\Longrightarrow} \Vdash_{\Omega+m+1}^{\Omega\cdot 3} P\underline{n}$ with $\Omega{+}m := \max\{\Omega, \mathrm{rk}(A_1), ..., \mathrm{rk}(A_l)\}$ [note that $\mathsf{k}(A_i) \subseteq \{0, \Omega\}$]
$\stackrel{9.9}{\Longrightarrow} \Vdash_{\Omega+1}^{\alpha} P\underline{n}$ with $\alpha \in C(\varepsilon_{\Omega+1}) \cap \varepsilon_{\Omega+1}$
$\stackrel{9.11}{\Longrightarrow} \models P^{<\beta}\underline{n}$ with $\beta := \psi(\omega + \omega^{\Omega+\alpha}) < \psi(\varepsilon_{\Omega+1})$.

APPENDIX to §4

Remark: Theorem 4.16 follows from 4.19 b).

Zu zeigen: $\mathcal{Z} \vdash \forall h \forall (n_i)_{i<\omega} \exists k (h[n_0][n_1]...[n_k] = 0) \implies$
$\mathcal{Z} \vdash \forall x \exists! y X\langle x, y\rangle \to \exists x, z_0, z_1 (X\langle x, z_0\rangle \wedge X\langle x{+}1, z_1\rangle \wedge \neg z_1 \prec_1 z_0)$

$\forall h \forall (n_i)_{i<\omega} \exists k (h[n_0][n_1]...[n_k] = 0) \implies$
$\forall (h_i)_{i<\omega} [\forall i (h_{i+1} \prec_1 h_i \vee h_{i+1} = h_i = 0) \to \exists k (h_k = 0)] \implies$
$\forall (h_i)_{i<\omega} \exists k \neg (h_{k+1} \prec_1 h_k)$.

Ann.: $\forall k (h_{k+1} \prec_1 h_k)$. Dann existiert ein k mit $h_k = 0$ und folglich $h_{k+1} \not\prec_1 h_k$).

ad Lemma 4.17

"$\Longrightarrow$":
$\vdash \forall x \exists! y G_a(x, y)$,
$\vdash \mathrm{WF}_\triangleleft(G_a) \to \exists x, z_0, z_1 (G_a(x, z_0) \wedge G_a(x{+}1, z_1) \wedge \neg(z_1 \triangleleft z_0))$
$\vdash \mathrm{Prog}_\triangleleft(X)$ & $a \notin X \to \forall x, z_0, z_1 (G_a(x, z_0) \wedge G_a(x{+}1, z_1) \to z_1 \triangleleft z_0)$
$\vdash \mathrm{WF}_\triangleleft(G_a) \wedge \mathrm{Prog}_\triangleleft(X) \wedge a \notin X \ \to \ \bot$.

$A \wedge \neg B\,,\ A \to B \vdash B \wedge \neg B$

"$\Longleftarrow$":
Assume $\forall x \exists y X\langle x, y\rangle$. By $\triangleleft$-induction on z we prove $\exists x X\langle x, z\rangle \to \exists x, z_0, z_1 (X\langle x, z_0\rangle \wedge X\langle x{+}1, z_1\rangle \wedge \neg(z_1 \triangleleft z_0))$. So let $X\langle x, z\rangle$. Then there exists z_1 such that $X\langle x_0{+}1, z_1\rangle$. If $\neg(z_1 \triangleleft z)$ we are done. If $z_1 \triangleleft z$ the claim follows from the IH.

APPENDIX to §8

Def.: $\lambda x.\tilde{\varepsilon}_x :=$ ordering function of $\{0\} \cup \{\beta : \omega^\beta = \beta\}$

Hence $\tilde{\varepsilon}_0 = 0$ and $\tilde{\varepsilon}_{1+\alpha} = \varepsilon_\alpha$.

Def.: $\Gamma := \{\alpha : \alpha \in C(\alpha)\}$

Def.: $\sup^+ X := \min\{\delta \in AP : \forall \xi \in X (\xi < \delta)\}$

Lemma A.1 $\alpha < \varphi_2(0)$ & $\beta < \tilde{\varepsilon}_{\alpha+1} \ \Rightarrow\ \Omega{\cdot}\alpha{+}\beta \in \Gamma$ & $\psi(\Omega{\cdot}\alpha{+}\beta) = \omega^{\tilde{\varepsilon}_\alpha + \beta}$.

Proof by induction on $\Omega{\cdot}\alpha{+}\beta$:
1. $\alpha = \beta = 0$: $\psi(0) = 1 = \omega^0 = \omega^{\tilde{\varepsilon}_\alpha + \beta}$.
2. $0 < \beta$:
By IH $\forall \eta < \beta (\Omega{\cdot}\alpha{+}\eta \in \Gamma$ & $\psi(\Omega{\cdot}\alpha{+}\eta) = \omega^{\tilde{\varepsilon}_\alpha + \eta})$. Hence $\psi(\Omega{\cdot}\alpha{+}\beta) =$
$\sup^+\{\psi(\zeta) : \zeta \in \Gamma$ & $\zeta < \Omega{\cdot}\alpha{+}\beta\} = \sup^+\{\psi(\Omega{\cdot}\alpha{+}\eta) : \eta < \beta\} = \sup^+\{\omega^{\tilde{\varepsilon}_\alpha + \eta} : \eta < \beta\} = \omega^{\tilde{\varepsilon}_\alpha + \beta}$.
3. $\beta = 0$ & $\alpha \in Lim$:
By IH $\forall \xi < \alpha (\Omega{\cdot}\xi \in \Gamma$ & $\psi(\Omega{\cdot}\xi) = \omega^{\tilde{\varepsilon}_\xi})$. Hence
$\psi(\Omega{\cdot}\alpha) = \sup^+\{\psi(\zeta) : \zeta \in \Gamma$ & $\zeta < \Omega{\cdot}\alpha\} = \sup^+\{\psi(\Omega{\cdot}\xi) : \xi < \alpha\} = \sup^+\{\omega^{\tilde{\varepsilon}_\xi} : \xi < \alpha\} = \omega^{\tilde{\varepsilon}_\alpha}$.

4. $\beta = 0$ & $\alpha = \alpha_0{+}1$: Let $\gamma := \omega^{\tilde{\varepsilon}_\alpha}$, i.e., $\gamma = \tilde{\varepsilon}_{\alpha_0+1}$.

By IH $\forall \eta < \gamma(\Omega{\cdot}\alpha_0{+}\eta \in \Gamma \;\&\; \psi(\Omega{\cdot}\alpha_0{+}\eta) = \omega^{\tilde{\varepsilon}_{\alpha_0}+\eta})$ and thus $\psi(\Omega{\cdot}\alpha_0{+}\gamma) = \sup^+\{\psi(\Omega{\cdot}\alpha_0{+}\eta) : \eta < \gamma\} = \sup^+\{\omega^{\tilde{\varepsilon}_{\alpha_0}+\eta} : \eta < \gamma\} = \omega^{\tilde{\varepsilon}_{\alpha_0}+\gamma} = \gamma$.

HS: $\gamma \le \eta < \Omega \;\Rightarrow\; \Omega{\cdot}\alpha_0{+}\eta \notin C(\Omega{\cdot}\alpha_0{+}\gamma)$.

Proof: $\Omega{\cdot}\alpha_0{+}\eta \in C(\Omega{\cdot}\alpha_0{+}\gamma) \;\&\; \eta < \Omega \;\Rightarrow\; \eta \in C(\Omega{\cdot}\alpha_0{+}\gamma) \cap \Omega = \psi(\Omega{\cdot}\alpha_0{+}\gamma) = \gamma$.

$\psi(\Omega{\cdot}\alpha) = \psi(\Omega{\cdot}\alpha_0{+}\Omega) \overset{\text{HS}}{=} \psi(\Omega{\cdot}\alpha_0{+}\gamma) = \gamma = \omega^{\tilde{\varepsilon}_\alpha}$.

Theorem A.2

$\alpha < \varphi_{n+2}(0) \;\&\; \beta < \varphi_{n+1}(\alpha{+}1) \;\Rightarrow\; \psi(\Omega^{n+1}{\cdot}(1{+}\alpha) + \Omega^n{\cdot}\beta) = \varphi_n(\varphi_{n+1}(\alpha) + \beta)$.

Corollary $\psi(\Omega^{n+1}) = \varphi_{n+1}(0)$.

Let $\nu \in On$ and $\Omega := \varepsilon_\nu$.

Definition of $C(\alpha)$ and $\psi(\alpha)$ by recursion on α

$C(\alpha) :=$ closure of $\{0\}$ under $+$, $\lambda x.\omega^{\Omega+x}$ and $\xi \mapsto \psi(\xi)$ $(\xi < \alpha \;\&\; \xi \in C(\xi))$.

$\psi(\alpha) := \min\{\xi : \xi \notin C(\alpha)\}$.

Lemma A.3

a) $\alpha \le \beta \Rightarrow C(\alpha) \subseteq C(\beta) \;\&\; \psi(\alpha) \le \psi(\beta)$,
   $\alpha < \beta \;\&\; \alpha \in C(\alpha) \Rightarrow \psi(\alpha) < \psi(\beta)$,

b) $\psi(\alpha)$ is an additive principal number,

c) $\alpha_0 < \alpha \;\&\; \forall \xi(\alpha_0 \le \xi < \alpha \Rightarrow \xi \notin C(\alpha_0)) \Longrightarrow C(\alpha_0) = C(\alpha) \;\&\; \psi(\alpha_0) = \psi(\alpha)$.

Proof:

a) The first part is trivial. – Second part: $\alpha < \beta \;\&\; \alpha \in C(\alpha) \Rightarrow \psi(\alpha) \le \psi(\beta) \;\&\; \psi(\alpha) \in C(\beta) \Rightarrow$
$\psi(\alpha) < \psi(\beta)$. The last implication holds, since $\psi(\beta) \notin C(\beta)$.

b) Assume not. Then $\psi(\alpha) = \xi + \eta$ with $\xi, \eta < \psi(\alpha)$, and thus $\psi(\alpha) = \xi + \eta \in C(\alpha)$. Contradiction.

c) By definition $C(\alpha_0)$ is closed under $\psi|\alpha_0$. From the premise it follows, that $C(\alpha_0)$ is even closed under $\psi|\alpha$. Hence $C(\alpha) \subseteq C(\alpha_0)$, since $C(\alpha)$ is the least set closed under $\psi|\alpha$ (and 0, $+$, $\lambda x.\omega^{\Omega+x}$).

Lemma A.4

a) $\alpha \le \varepsilon_0 \;\&\; \alpha < \Omega \;\Rightarrow\; \psi(\alpha) \le \omega^\alpha$.

b) $C(\varepsilon_0) \cap \Omega \subseteq \varepsilon_0$.

c) $\varepsilon_0 \le \alpha \le \Omega \;\Rightarrow\; C(\alpha) = C(\varepsilon_0)$.

d) $\psi(\alpha) \le \varepsilon_\alpha$.

e) $\Omega < \varphi_2(\nu_0) \Rightarrow C(\alpha) \subseteq \varphi_2(\nu_0)$, for each $\alpha \in On$.

Proof:

a) Induction on α: Let $\alpha \le \varepsilon_0 \;\&\; \alpha < \Omega$. By IH $\forall \xi < \alpha(\psi(\xi) \le \omega^\xi < \omega^\alpha)$. Hence $C(\alpha) \cap \Omega \subseteq \omega^\alpha$.
Since $\omega^\alpha < \Omega$, this implies $\omega^\alpha \notin C(\alpha)$, and therefore $\psi(\alpha) \le \omega^\alpha$.

b) By a) $\forall\xi < \varepsilon_0(\psi(\xi) < \varepsilon_0)$. This yields the claim.

c) Let $\varepsilon_0 < \alpha \leq \Omega$: By b) $\forall\xi(\varepsilon_0 \leq \xi < \alpha \Rightarrow \xi \notin C(\varepsilon_0))$. Hence $C(\alpha) = C(\varepsilon_0)$ by A.3c.

d) By IH $\forall\xi < \alpha(\psi(\xi) \leq \varepsilon_\xi < \varepsilon_\alpha)$ $(*)$.

Case 1: $\alpha < \varepsilon_0$: By a) $\psi(\alpha) \leq \omega^\alpha < \varepsilon_0$.

Case 2: $\Omega < \varepsilon_\alpha$: Then $C(\alpha) \subseteq \varepsilon_\alpha$ and thus $\psi(\alpha) \leq \varepsilon_\alpha$.

Case 3: $\varepsilon_0 \leq \alpha$ & $\varepsilon_\alpha \leq \Omega$: Then $\varepsilon_0 \leq \alpha \leq \Omega$ and $\varepsilon_0 < \Omega$. Hence $\psi(\alpha) \overset{\text{c)}}{=} \psi(\varepsilon_0) \overset{\text{a)}}{\leq} \omega^{\varepsilon_0} < \varepsilon_\alpha$.

e) $\varphi_2(\nu_0) > \Omega$ is closed under $+$, $\lambda x.\omega^{\Omega+x}$, and [by d)] also under ψ. This yields the claim.