Relating ordinals to proofs in a perspicious way

Wilfried Buchholz
Mathematisches Institut der Ludwig-Maximilians-Universität München
Theresienstr. 39, D-80333 München, Germany
email: buchholz@rz.mathematik.uni-muenchen.de

Introduction

Contemporary ordinal-theoretic proof theory (i.e., the part of proof theory concerned with ordinal analysises of strong impredicative theories) suffers from the extreme (and as it seems unavoidable) complexity and opacity of its main tool, the ordinal notation systems. This is not only a technical stumbling block which prevents most proof-theorists from a closer engagement in that field, but it also calls the achieved results into question, at least as long as these results do not have interesting consequences, such as e.g. foundational reductions or intuitively graspable combinatorial independence results. The question is, what have we gained by having shown that the proof-theoretic ordinal of a certain formal theory S equals the ordertype $\| \prec \|$ of some specific term ordering $\prec$ (disregarding for the moment possible applications of the abovementioned kind). I would say, as long as the ordering is simple and natural (as e.g. the standard orderings of order type ε_0 or Γ_0) or occurs already elsewhere in mathematics or theoretical computer science, we have obtained a result of pure mathematics which is interesting and noteworthy by itself. But I'm not sure if the latter can (already) be said about Rathjen's ([Ra95])and Arai's ([Ar96],[Ar97]) most impressive and admirable work on Π^1_2-CA and beyond. Rather I think, that extensive efforts should be made aiming at a substantial simplification and deeper understanding of that work, especially of the used ordinal notation systems (term orderings, resp.). The purpose of the present paper is to present as an example a particularly nice piece of ordinal-theoretic, impredicative proof theory (far below Π^1_2-CA) where such a simplification and deeper understanding has already been achieved. The paper is more or less a condensation and improved presentation of parts of [Bu81], [Bu86], [Bu87] providing an ordinal analysis "from above" for the theories ID_ν of iterated inductive definitions by means of the so-called $\Omega_{\mu+1}$-rule. Special efforts have been made to let the assignment of ordinals to derivations not appear as an ad hoc approach but as a quite natural and perspicuous procedure.

§1 Ordinal functions and ordinal notations

Around the early seventies Feferman suggested the following definition of ordinal functions θ_α for generating large segments of recursive ordinals, which was intended to replace the extremely complex Bachmann-Pfeiffer-Isles definition procedure for hierarchies of normal functions. This definition brought a new impetus into proof theory and was the starting point of a rather successful development.

Definition (Feferman 1970)
Let ν_0 be a fixed countable ordinal.
By transfinite recursion on α one defines sets $C(\alpha,\beta)$ and ordinal functions θ_α:
$C(\alpha,\beta) :=$ least set $X \supseteq \beta \cup \{\Omega_\sigma : 0 < \sigma < \nu_0\}$ such that $\forall \xi,\eta \in X(\xi+\eta \in X)$
and $\forall \xi,\eta \in X(\,\xi < \alpha \implies \theta_\xi(\eta) \in X\,)$,
θ_α enumerates $\{\beta > 0 : C(\alpha,\beta) \cap \beta^+ \subseteq \beta\}$ where $\beta^+ := \min\{\Omega_{\sigma+1} : \beta < \Omega_{\sigma+1}\}$ and $\Omega_\sigma := \aleph_\sigma$.

Dedicated to Solomon Feferman on the occasion of his 70th birthday. I am deeply indebted to Sol for the encouragement and stimulation he gave me over many years through his interest in my work.

A simple cardinal argument shows that $C(\alpha,\beta)$ has essentially the same cardinality as β (namely $card(C(\alpha,\beta)) = \max\{\aleph_0, card(\beta)\}$), no matter how large α is. This yields $\forall\beta < \Omega_{\sigma+1}(\theta_\alpha(\beta) < \Omega_{\sigma+1})$, for any α (see Lemma 1.1 below). Hence the function $\lambda\alpha.\theta_\alpha(\Omega_\sigma)$ maps On into $\Omega_{\sigma+1}$, i.e., it is a so-called *collapsing function*. Important contributions to the further investigation of the θ-functions were made by Weyhrauch, Aczel and (most important) Jane Bridge. Aczel generalized the definition and conjectured relationships between the θ_α's and the Pfeiffer/Isles functions. These conjectures were established by Jane Bridge in her Ph.D. thesis. She also obtained partial results on the recursiveness of the notation system associated with θ_α [Br75]. Starting from Bridge's thesis, recursiveness of the full θ-notation system was established in [Bu75]. There also a variant $\overline{\theta}_\alpha$ of θ_α was introduced which has the advantage that the $<$-relation between $\overline{\theta}$-terms can be characterized in a particularly simple way. This opened the way to a successful use of these ordinal functions in proof-theory. Later it turned out that in proof-theoretic applications actually only the values of $\overline{\theta}_\alpha$ at initial ordinals Ω_σ are used, which led to the idea to define directly functions ψ_σ corresponding to $\lambda\alpha.\overline{\theta}_\alpha(\Omega_\sigma)$

Definition (Buchholz 1981)
Let ν_0 be a fixed countable ordinal.
$\psi_\sigma(\alpha) := \min\{\beta \geq \Omega^-_{\sigma+1} : C(\alpha,\beta) \cap \Omega_{\sigma+1} \subseteq \beta\}$ $(\sigma < \nu_0)$
$C(\alpha,\beta) :=$ closure of β under $+$ and all $\psi_\sigma{\restriction}\alpha$ $(\sigma < \nu_0)$.

$$\Omega^-_{\sigma+1} := \begin{cases} 1 & \text{if } \sigma = 0 \\ \Omega_\sigma & \text{if } \sigma > 0 \end{cases}$$

Remark $C(\alpha,\psi_\sigma(\alpha)) \cap \Omega_{\sigma+1} = \psi_\sigma(\alpha)$

Before we start to prove some basic facts about these ordinal functions ψ_σ, we establish a general Lemma which also applies to stronger systems of ordinal functions.

Definition
Given a countable set $\mathcal{F}$ of ordinal functions and an ordinal β let
$\mathrm{Cl}(\mathcal{F};\beta) :=$ the closure of β under $+$ and all functions in $\mathcal{F}$.
Let $\Re$ be the class of all uncountable regular cardinals.

Lemma 1.1
If $\mathcal{F}$ is a countable set of ordinal functions, then for each $\kappa \in \Re$
the set $\{\beta < \kappa : \mathrm{Cl}(\mathcal{F};\beta) \cap \kappa \subseteq \beta\}$ is closed unbounded in κ.
Proof:
unbounded: Let $\beta_0 < \kappa$ be given. We set $\beta_{n+1} := \min\{\eta : \mathrm{Cl}(\mathcal{F};\beta_n) \cap \kappa \subseteq \eta\}$, $\beta := \sup_{n<\omega}\beta_n$. Since $\forall\eta < \kappa(\, card(\mathrm{Cl}(\mathcal{F};\eta)) < \kappa\,)$, we obtain (by induction on n) $\beta_n < \kappa$, and then $\beta_0 \leq \beta < \kappa$ and $\mathrm{Cl}(\mathcal{F};\beta) \cap \kappa \subseteq \bigcup_{n<\omega}\mathrm{Cl}(\mathcal{F};\beta_n) \cap \kappa \subseteq \bigcup_{n<\omega}\beta_{n+1} \subseteq \beta$.
closed: If $\beta = \sup(X)\ \&\ \forall\eta{\in}X(\mathrm{Cl}(\mathcal{F};\eta) \cap \kappa \subseteq \eta)$ then $\mathrm{Cl}(\mathcal{F};\beta) \cap \kappa \subseteq \bigcup_{\eta\in X}\mathrm{Cl}(\mathcal{F};\eta) \cap \kappa \subseteq \bigcup_{\eta\in X}\eta = \beta$.

Theorem 1.2 (Basic properties of ψ_σ)
(a) $\psi_\sigma(\alpha) < \Omega_{\sigma+1}$ *(collapsing property)*
(b) $\psi_\sigma(0) = \Omega^-_{\sigma+1}$, and $\psi_\sigma(\alpha)$ is closed under $+$
(c) $\alpha_0 < \alpha\ \&\ \alpha_0 \in C(\alpha_0,\psi_\sigma(\alpha_0)) \implies \psi_\sigma(\alpha_0) < \psi_\sigma(\alpha)$
(d) $\psi_\sigma(\alpha_0) = \psi_\sigma(\alpha_1)\ \&\ \underbrace{\alpha_i \in C(\alpha_i,\psi_\sigma(\alpha_i))}_{\textit{normalform condition}}\ (i=0,1) \implies \alpha_0 = \alpha_1$

Proof:
(a) Obviously $C(\alpha,\beta) = \mathrm{Cl}(\{\psi_\sigma{\restriction}\alpha : \sigma < \nu_0\};\beta)$. Hence $\{\beta < \Omega_{\sigma+1} : \beta \geq \Omega^-_{\sigma+1}\ \&\ C(\alpha,\beta) \cap \Omega_{\sigma+1} \subseteq \beta\} \neq \emptyset$ (by Lemma 1.1), and therefore $\psi_\sigma(\alpha) = \min\{\beta \geq \Omega^-_{\sigma+1} : C(\alpha,\beta) \cap \Omega_{\sigma+1} \subseteq \beta\} < \Omega_{\sigma+1}$.

(b) Obviously $C(0,\Omega^-_{\sigma+1}) = \Omega^-_{\sigma+1}$, and therefore $\psi_\sigma(0) = \Omega^-_{\sigma+1}$.
$\psi_\sigma(\alpha)$ is closed under $+$, since $\psi_\sigma(\alpha) = C(\alpha,\psi_\sigma(\alpha)) \cap \Omega_{\sigma+1}$.
(c) $C(\alpha_0,\psi_\sigma(\alpha)) \cap \Omega_{\sigma+1} \subseteq C(\alpha,\psi_\sigma(\alpha)) \cap \Omega_{\sigma+1} \subseteq \psi_\sigma\alpha$ and therefore $\psi_\sigma(\alpha_0) \leq \psi_\sigma(\alpha)$ which implies $C(\alpha_0,\psi_\sigma(\alpha_0)) \subseteq C(\alpha,\psi_\sigma(\alpha))$. From $\alpha_0 < \alpha$ & $\alpha_0 \in C(\alpha,\psi_\sigma(\alpha))$ we conclude $\psi_\sigma(\alpha_0) \in C(\alpha,\psi_\sigma(\alpha))$.
By (a) we have $\psi_\sigma(\alpha_0) < \Omega_{\sigma+1}$ and thus $\psi_\sigma(\alpha_0) \in C(\alpha,\psi_\sigma(\alpha)) \cap \Omega_{\sigma+1} \subseteq \psi_\sigma(\alpha)$.
(d) If $\alpha_0 < \alpha_1$ then the assumption $\alpha_0 \in C(\alpha_0,\psi_\sigma(\alpha_0))$ together with (c) yields $\psi_\sigma(\alpha_0) < \psi_\sigma(\alpha_1)$.

Lemma 1.3
$C(\alpha,\psi_0(\alpha)) = C(\alpha,1)$
Proof by induction on α:
Let us assume that $C(\xi,\psi_0(\xi)) = C(\xi,1)$, for all $\xi < \alpha$ (IH). We have to prove $\psi_0\alpha \subseteq C(\alpha,1)$. Let $\alpha > 0$ (otherwise $\psi_0(\alpha) = 1 \subseteq C(\alpha,1)$). As we will show below the IH implies that $\beta := C(\alpha,1) \cap \Omega_1$ is in fact an ordinal. Obviously $\Omega_1 = \psi_1(0) \in C(\alpha,\beta)$ and $C(\alpha,\beta) \cap \Omega_1 \subseteq C(\alpha,1) \cap \Omega_1 = \beta$ and thus $\psi_0(\alpha) \leq \beta$.
Claim: $\gamma \in C(\alpha,1) \cap \Omega_1 \;\Rightarrow\; \gamma \subseteq C(\alpha,1)$.
Proof by side induction on the definition of $C(\alpha,1)$:
1. $\gamma = \psi_0(\xi)$ with $\xi < \alpha$ & $\xi \in C(\alpha,1)$:
By the above IH we have $C(\xi,\psi_0(\xi)) = C(\xi,1)$. Hence $\gamma = \psi_0(\xi) \subseteq C(\xi,1) \subseteq C(\alpha,1)$.
2. $\gamma = \gamma_0 + \gamma_1$ with $\gamma_0,\gamma_1 \in C(\alpha,1) \cap \Omega_1$:
Then by SIH $\gamma_0,\gamma_1 \subseteq C(\alpha,1)$ which (together with $\gamma_0 \in C(\alpha,1)$) yields $\gamma_0 + \gamma_1 \subseteq C(\alpha,1)$.

Disgression
Of course the above definition of ψ_σ can be generalized in the same way as Aczel [Acz] generalized Feferman's definition, namely by incorporating a countable set $\mathcal{G}$ of cardinal valued ordinal functions into the definition of $C(\alpha,\beta)$. For the simple case $\mathcal{G} = \{\lambda x.\Omega_x\}$ this is equivalent to setting, for any $\sigma,\alpha \in On$, $\psi_\sigma(\alpha) := \min\{\beta \geq \Omega^-_{\sigma+1} : C(\alpha,\beta) \cap \Omega_{\sigma+1} \subseteq \beta\}$ where $C(\alpha,\beta) := \mathrm{Cl}(\psi{\restriction}\alpha\,;\,\beta)$ with $\psi{\restriction}\alpha : On \times \alpha \to On$, $(\rho,\xi) \mapsto \psi_\rho(\xi)$.
A more substantial extension of the θ/ψ-approach was developed by Schütte (unpublished), Pohlers [Po87] and, most important, Jäger [Jä84]. In addition to the ψ_σ's which are collapsing functions for successor regulars, now also collapsing functions for limit regulars, i.e. weakly inaccessible cardinals, are introduced. In order to be able to treat both kinds of collapsing functions uniformly Jäger denoted the ψ-function for a regular cardinal κ, i.e. one with values below κ, by ψ_κ. In this notation the former ψ_σ becomes $\psi_{\Omega_{\sigma+1}}$.
Definition (Jäger 1984):
$\psi_\kappa(\alpha) := \min\{\beta > \kappa^- : C(\alpha,\beta) \cap \kappa \subseteq \beta\}$
$C(\alpha,\beta) := \mathrm{Cl}(\lambda xy.I_x(y), \psi{\restriction}\alpha\,;\,\beta)$ with $\psi{\restriction}\alpha : \Re \times \alpha \to On$, $(\pi,\xi) \mapsto \psi_\pi(\xi)$
$I_\rho :=$ ordering function of $cl(\{\kappa \in \Re : \forall\xi < \rho(I_\xi(\kappa) = \kappa)\})$,
where $cl(A) := \{\sup(X) : X$ nonempty subset of $A\}$ *(topological closure of A)*
$$\kappa^- := \begin{cases} 0 & \text{if } \kappa = I_\rho(0) \\ I_\rho(\sigma) & \text{if } \kappa = I_\rho(\sigma{+}1) \text{ with } \rho,\sigma < \kappa \end{cases}.$$

Later Rathjen developed further extensions up to the use of large large cardinals ([Ra90], [Ra94], [Ra95]). In the first of these extensions Rathjen assumed the existence of a weakly Mahlo cardinal M and utilized the fact that the regular cardinals are stationary in M.
Definition (Rathjen 1990)
$$\psi_\kappa(\alpha) := \begin{cases} \min\{\beta \in \Re : \alpha \in C(\alpha,\beta) \;\&\; C(\alpha,\beta) \cap \kappa \subseteq \beta\} & \text{if } \kappa = M \\ \min\{\beta \in On : \kappa \in C(\alpha,\beta) \;\&\; C(\alpha,\beta) \cap \kappa \subseteq \beta\} & \text{if } \kappa < M \end{cases}$$
$C(\alpha,\beta) := \mathrm{Cl}(\lambda x.\omega^{M+x}, \psi{\restriction}\alpha;\beta)$ with $(\psi{\restriction}\alpha)(\pi,\xi) := \psi_\pi(\xi)$ for $\pi \in \Re \cap (M{+}1)$ and $\xi < \alpha$.

Let us see how the "Mahloness" of M is used for obtaining the crucial property $\forall\alpha < \varepsilon_{M+1}(\,\psi_M(\alpha) < M\,)$. According to Lemma 1.1 the set $\{\beta < M : C(\alpha,\beta)\cap M \subseteq \beta\}$ is closed unbounded in M. If $\alpha < \varepsilon_{M+1}$ then $\alpha \in C(\alpha,\beta)$ for sufficiently large $\beta < M$. Hence also $X := \{\beta < M : \alpha \in C(\alpha,\beta)\ \&\ C(\alpha,\beta)\cap M \subseteq \beta\}$ is club in M (if $\alpha < \varepsilon_{M+1}$). Since $\Re$ is stationary in M, we obtain $X \cap \Re \neq \emptyset$, and thus $\psi_M(\alpha) = \min(X\cap\Re) < M$.

The general strategy behind these extensions is that one tries to produce notations for more and more regular cardinals κ, which in turn gives more and more collapsing functions ψ_κ. In Jäger's definition the regulars are provided by the hierarchy $(I_\rho)_{\rho\in On}$. In defining this hierarchy one already assumes the existence of a weakly Mahlo cardinal M, but one does not exploit this assumption to the same extent as it is done in Rathjen's approach. Actually the limit ordinal $\psi_{\Omega_1}(\sup_n \delta_n)$ ($\delta_0 := 0$, $\delta_{n+1} := I_{\delta_n}(0)$) of [Jä84] is "much" smaller then the limit ordinal $\psi_{\Omega_1}(\varepsilon_{M+1})$ of [Ra90].

Now we return to the simple system of ψ-functions ψ_σ ($\sigma < \nu_0$) defined above. In order to avoid some technical complications we will even assume $\nu_0 = \omega$. But we want to stress that no essential new difficulties would emerge when in all what follows the assumption "$\nu_0 = \omega$" would be replaced by "$\nu_0 < \omega_1^{CK}$".

From now on σ,ρ,μ,ν range over numbers $< \omega$.

Below we will introduce a system of ordinal notations based on the ordinal functions ψ_σ. The canonical way for that is to consider the set T of all terms which are generated from the constant 0 by means of function symbols $\oplus, D_0, D_1, \ldots$ for the ordinal functions $+, \psi_0, \psi_1, \ldots$. Then one looks for a (primitive) recursive characterization of the relation $<_{\mathrm{o}} := \{(a,b) \in \mathrm{T}\times\mathrm{T} : \mathrm{o}(a) < \mathrm{o}(b)\}$, where $\mathrm{o}(a) \in On$ is the canonical interpretation of $a \in \mathrm{T}$. It turns out that the relation $<_{\mathrm{o}}$ has a particularly simple characterization when it is restricted to the subset $\mathrm{OT} \subseteq \mathrm{T}$ of those terms $a \in \mathrm{T}$ which are in "normalform" (i.e. $\mathrm{o}(b) \in C(\mathrm{o}(b), \psi_\sigma(\mathrm{o}(b)))$ for each subterm $D_\sigma b$ of a, and $\mathrm{o}(a_n) \le \ldots \le \mathrm{o}(a_0)$ for each subterm $a_0 \oplus \ldots \oplus a_n$ of a).

Definition $\overline{\mathrm{o}}(a) :=$ ordertype of $(\{x \in \mathrm{OT} : x <_{\mathrm{o}} a\}, <_{\mathrm{o}})$, $\mathrm{OT}_0 := \{a \in \mathrm{OT} : a <_{\mathrm{o}} D_1 0\}$.

It can be proved that every $a \in \mathrm{T}$ has a (unique) normalform $a^* \in \mathrm{OT}$ such that $\mathrm{o}(a^*) = \mathrm{o}(a)$ which yields

$$(\diamondsuit)\quad \{\mathrm{o}(a) : a \in \mathrm{T}\} = \{\mathrm{o}(a) : a \in \mathrm{OT}\}.$$

Obviously $\{\mathrm{o}(a) : a \in \mathrm{T}\} = C(\Omega_\omega, 1)$, and thus by Lemma 1.3 $\{\mathrm{o}(a) : a \in \mathrm{T}\ \&\ a <_{\mathrm{o}} D_1 0\} = \psi_0(\Omega_\omega)$. So, $\{\mathrm{o}(a) : a \in \mathrm{T}\ \&\ a <_{\mathrm{o}} D_1 0\}$ is a segment of On, and the equation $(\diamondsuit)$ implies $\mathrm{o}(a) = \overline{\mathrm{o}}(a)$ for each $a \in \mathrm{OT}_0$. (Of course we will never have $\mathrm{o}(a) = \overline{\mathrm{o}}(a)$ for *all* $a \in \mathrm{OT}$, since T is countable and thus $\overline{\mathrm{o}}(a) < \Omega_1 \le \mathrm{o}(a)$ for any $a \in \mathrm{OT}\setminus\mathrm{OT}_0$.) The proof of $(\diamondsuit)$ itself is rather tedious (cf. [BS88] or [Se98]) and we will not discuss it here. Having a closer look one will realize that actually $(\diamondsuit)$ is not of great importance, since in all (existing) proof theoretic applications of the ψ_σ's one can confine to terms from OT, and the only drawback when one dispenses with $(\diamondsuit)$ is that e.g. instead of saying "the proof theoretic ordinal of ID_ν is $\psi_0(\varepsilon_{\Omega_\nu+1})$" one has to say "the proof theoretic ordinal of ID_ν is the ordertype of $D_0 D_{\nu+1} 0$ in $(\mathrm{OT}_0, <_{\mathrm{o}})$". Usually one avoids the trouble with having to distinguish between $\mathrm{o}(a)$ and $\overline{\mathrm{o}}(a)$ (for $a \in \mathrm{OT}_0$) by following Schütte and incorporating the normalform condition into the definition of $C(\alpha,\beta)$, i.e., by closing $C(\alpha,\beta)$ only under $\widetilde{\psi}_\sigma\restriction\alpha$ (instead of $\psi_\sigma\restriction\alpha$) where $(\widetilde{\psi}_\sigma\restriction\alpha)(\xi) := \begin{cases} \psi_\sigma(\xi) & \text{if } \xi < \alpha \textit{ and } \xi \in C(\xi,\psi_\sigma(\xi)) \\ \text{undefined} & \text{otherwise} \end{cases}$. Then $\{o(a) : a \in \mathrm{OT}\} = C(\Omega_\omega, 1)$ is easy to prove and together with Lemma 1.3 it yields $\mathrm{o}(a) = \overline{\mathrm{o}}(a)$ for all $a \in \mathrm{OT}_0$.

Now we define the set T of terms, a linear ordering $\prec$ on T, for any $a \in \mathrm{T}$ and $\sigma < \omega$ a set $G_\sigma a$ of subterms of a, and the set OT of *ordinal terms* (i.e. terms in normalform) in such a way that, for all $a, c \in \mathrm{OT}$,
(a) $c \prec a \;\Leftrightarrow\; \mathrm{o}(c) < \mathrm{o}(a)$ $[\,\Leftrightarrow\; c <_{\mathrm{o}} a\,]$ and (b) $G_\sigma c \prec a \;\Leftrightarrow\; \mathrm{o}(c) \in C(\mathrm{o}(a), \psi_\sigma \mathrm{o}(a))$.
(Here $G_\sigma a \prec c$ abbreviates $\forall x \in G_\sigma a (x \prec c)$.)

Inductive definition of T

1. $0 \in \mathrm{T}$.
2. If $a \in \mathrm{T}$ and $\sigma < \omega$, then $D_\sigma a \in \mathrm{T}$; we call $D_\sigma a$ a *principal term.*
3. If $a_0, ..., a_n$ are principal terms and $n \geq 1$, then $(a_0 \oplus ... \oplus a_n) \in \mathrm{T}$.

Notation For principal terms $a_0, ..., a_{n-1}$ and $n \geq 0$ we set $a_0 \oplus ... \oplus a_{n-1} := \begin{cases} 0 & \text{if } n = 0 \\ a_0 & \text{if } n = 1 \\ (a_0 \oplus ... \oplus a_{n-1}) & \text{if } n > 1 \end{cases}$

So every $a \in \mathrm{T}$ can be uniquely written as $a = D_{\sigma_0} a_0 \oplus \ldots \oplus D_{\sigma_{n-1}} a_{n-1}$ with $n \geq 0$ and $a_0, ..., a_{n-1} \in \mathrm{T}$.
Further we define: $(a_0 \oplus ... \oplus a_{n-1}) \oplus (b_0 \oplus ... \oplus b_{m-1}) := a_0 \oplus ... \oplus a_{n-1} \oplus b_0 \oplus ... \oplus b_{m-1}$,
and $a \cdot n := \underbrace{a \oplus ... \oplus a}_{n}$ for principal terms a_i, b_i, a.

Definition of $\mathrm{o} : \mathrm{T} \longrightarrow On$
$\mathrm{o}(D_{\sigma_0} a_0 \oplus \ldots \oplus D_{\sigma_{n-1}} a_{n-1}) := \psi_{\sigma_0} \mathrm{o}(a_0) + \ldots + \psi_{\sigma_{n-1}} \mathrm{o}(a_{n-1})$

Definition of $a \prec b$ **for** $a, b \in \mathrm{T}$

1. $0 \prec b \ :\Longleftrightarrow\ b \neq 0$
2. $D_\sigma a \oplus \tilde{a} \prec D_\rho b \oplus \tilde{b} \ :\Longleftrightarrow\ \sigma < \rho$ or ($\sigma = \rho$ & $a \prec b$) or ($\sigma = \rho$ & $a = b$ & $\tilde{a} \prec \tilde{b}$)

Remark. $\prec$ is a linear ordering on T, but it's not a wellordering (e.g. $\ldots \prec D_0 D_0 D_1 0 \prec D_0 D_1 0 \prec D_1 0$).

Definition of $G_\sigma a$

1. $G_\sigma(a_0 \oplus ... \oplus a_{n-1}) := \bigcup_{i<n} G_\sigma a_i$, 2. $G_\sigma D_\mu a := \begin{cases} \{a\} \cup G_\sigma a & \text{if } \sigma \leq \mu \\ \emptyset & \text{if } \mu < \sigma \end{cases}$

Inductive definition of OT

1. $0 \in \mathrm{OT}$.
2. $a \in \mathrm{OT}$ & $G_\sigma a \prec a \Rightarrow D_\sigma a \in \mathrm{OT}$.
3. $a_0, ..., a_n \in \mathrm{OT}$ ($n \geq 1$) principal terms with $a_n \preceq \ldots \preceq a_0 \Longrightarrow (a_0 \oplus ... \oplus a_n) \in \mathrm{OT}$.

The elements of OT are called *ordinal terms*. We identify $n \in \mathbb{N}$ with the ordinal term $\underbrace{D_0 0 \oplus ... \oplus D_0 0}_{n}$.

Abbreviation. $\Omega_0 := \omega := D_0 1$, $\Omega_\sigma := D_\sigma 0$ for $\sigma > 0$.

Theorem 1.4 For $a, c \in \mathrm{OT}$ we have
(a) $c \prec a \Leftrightarrow \mathrm{o}(c) < \mathrm{o}(a)$, [i.e., $c \prec a \Leftrightarrow c <_{\mathrm{o}} a$]
(b) $G_\sigma c \prec a \Leftrightarrow \mathrm{o}(c) \in C(\mathrm{o}(a), \psi_\sigma \mathrm{o}(a))$
Proof by induction on the length of c simultaneous for (a),(b):
We only prove "$\Rightarrow$". The reverse implication of (a) follows from "$\Rightarrow$", since $\prec$ is total. The reverse direction of (b) is more difficult to obtain, but since it is not needed for the proof of (a) nor for any proof in this paper, we omit it here.
(a) Let $c = D_\sigma c_0 \oplus c_1 \oplus ... \oplus c_m$, $a = D_\rho a_0 \oplus a_1 \oplus ... \oplus a_n$ with principal terms $c_1, ..., c_m, a_1, ..., a_n$.
1. $\sigma < \rho$: From $c_m \preceq \ldots \preceq c_1 \preceq D_\sigma c_0$ we get by IH $\mathrm{o}(c_m) \leq \ldots \leq \mathrm{o}(c_1) \leq \mathrm{o}(D_\sigma c_0) = \psi_\sigma \mathrm{o}(c_0) < \Omega_{\sigma+1}$ and thus $\mathrm{o}(c) < \Omega_{\sigma+1} \leq \Omega_\rho \leq \mathrm{o}(D_\rho a_0) \leq \mathrm{o}(a)$.
2. $\sigma = \rho$ and $c_0 \prec a_0$: By IH $\mathrm{o}(c_0) < \mathrm{o}(a_0)$. Since $D_\sigma c_0 \in \mathrm{OT}$, we have $G_\sigma c_0 \prec c_0$ and thus by IH $\mathrm{o}(c_0) \in C(\mathrm{o}(c_0), \psi_\sigma \mathrm{o}(c_0))$. Hence $\psi_\sigma \mathrm{o}(c_0) < \psi_\sigma \mathrm{o}(a_0)$ by Theorem 1.2(c). Now $\mathrm{o}(c) \prec \mathrm{o}(a)$ follows as in 1. (using that $\psi_\sigma \mathrm{o}(a)$ is additively closed).
3. $\sigma = \rho$ & $c_0 = a_0$ & $c_1 \oplus ... \oplus c_m \prec a_1 \oplus ... \oplus a_n$: Immediate by IH.
(b) 1. $c = c_0 \oplus ... \oplus c_{k-1}$ with $k \neq 1$: Then $G_\sigma c_i \prec a$ and thus (by IH) $\mathrm{o}(c_i) \in C := C(\mathrm{o}(a), \psi_\sigma \mathrm{o}(a))$ for $i < k$. This yields $\mathrm{o}(c) = \mathrm{o}(c_0) + \ldots + \mathrm{o}(c_{k-1}) \in C$.
2. $c = D_\mu c_0$ with $\mu < \sigma$: Then $\mathrm{o}(c) \in \Omega_{\mu+1} \subseteq \Omega_\sigma \subseteq C$.
3. $c = D_\mu c_0$ with $\sigma \leq \mu$: Then $\{c_0\} \cup G_\sigma c_0 = G_\sigma c_0 \prec a$ and therefore by IH $\mathrm{o}(c_0) < \mathrm{o}(a)$ & $\mathrm{o}(c_0) \in C$ which yields $\mathrm{o}(c) = \psi_\mu \mathrm{o}(c_0) \in C$.
Corollary $(\mathrm{OT}, \prec)$ is a well ordering.

Fundamental sequences

In order to get a better insight into the structure $(\mathrm{T},\prec)$ and a better understanding of the collapsing functions ψ_σ we now present an assignment of (fundamental) sequences to the elements of T. For each term $a\in\mathrm{T}$ we define its *(cofinality) type* $\mathrm{tp}(a)\in\{0,1,\omega\}\cup\{\Omega_{\mu+1}:\mu<\omega\}$ and a family $(a[x])_{x\in|\mathrm{tp}(a)|}$ of terms, such that the following holds, where $|0|:=\emptyset$, $|1|:=\{0\}$, $|\omega|:=\mathbb{N}$, $|\Omega_{\mu+1}|:=\{D_\mu b:b\in\mathrm{T}\}$:

Theorem 1.5

(a) $x\in|\mathrm{tp}(a)|\implies a[x]\prec a$

(b) $x,x'\in|\mathrm{tp}(a)|\ \&\ x\prec x'\implies a[x]\prec a[x']$

(c) $\mathrm{tp}(a)=1\implies a=a[0]\oplus 1$

(d) $a,c\in\mathrm{OT}\ \&\ c\prec a\ \&\ \mathrm{tp}(a)\neq 1\implies \exists x\in\mathrm{OT}\cap|\mathrm{tp}(a)|(\,c\prec a[x]\,)$

(e) $a,x\in\mathrm{OT}\ \&\ x\in|\mathrm{tp}(a)|\implies a[x]\in\mathrm{OT}$

Note that, according to Theorem 1.5, only for $a\in\mathrm{OT}$ and only relative to $(\mathrm{OT},\prec)$ is the family $(a[x])_{x\in|\mathrm{tp}(a)|}$ a fundamental sequence of a in the proper sense. But in §3 we will give a natural interpretation of the terms $a\in\mathrm{T}$ as wellfounded trees (so-called *tree ordinals*) $\mathfrak{t}(a)$ which harmonizes with the assignment $(a,x)\mapsto a[x]$. For example we will have $\mathfrak{t}(a)=\big(\mathfrak{t}(a[i])\big)_{i\in\mathbb{N}}$ if $\mathrm{tp}(a)=\omega$.

I think, one can well say that all clauses of the following definition are canonical (modulo some minor variations, such as setting $(D_\sigma a)[i]:=(D_\sigma a[0])\cdot i$ instead of $(D_\sigma a[0])\cdot(i{+}1)$ in clause 4.), and therefore it seems reasonable to call $(a[x])_{x\in|\mathrm{tp}(a)|}$ *the canonical* fundamental sequence of $a\in\mathrm{OT}$. Only clause 6. requires some explanation which will be given below in the proof of Theorem 1.5.

Definition of $\mathrm{tp}(a)$ and $a[x]$ for $a\in\mathrm{T}$, $x\in|\mathrm{tp}(a)|$

1. $\mathrm{tp}(0):=0$.
2. $\mathrm{tp}(D_0 0):=1,\ \ (D_0 0)[0]:=0$.
3. $\mathrm{tp}(D_{\mu+1}0):=\Omega_{\mu+1},\ \ (D_{\mu+1}0)[x]:=x$.
4. $\mathrm{tp}(a)=1\ \Rightarrow\ \mathrm{tp}(D_\sigma a):=\omega,\ \ (D_\sigma a)[i]:=(D_\sigma a[0])\cdot(i{+}1)$.
5. $\mathrm{tp}(a)\in\{\omega\}\cup\{\Omega_{\mu+1}:\mu<\sigma\}\ \Rightarrow\ \mathrm{tp}(D_\sigma a):=\mathrm{tp}(a),\ \ (D_\sigma a)[x]:=D_\sigma a[x]$.
6. $\mathrm{tp}(a)=\Omega_{\mu+1}\ \&\ \mu\geq\sigma\ \Rightarrow\ \mathrm{tp}(D_\sigma a):=\omega,\ \ (D_\sigma a)[i]:=D_\sigma a[x_i]$ with $x_0:=\Omega_\mu$, $x_{i+1}:=D_\mu a[x_i]$.
7. $\mathrm{tp}(a_0{\oplus}...{\oplus}a_n):=\mathrm{tp}(a_n),\ \ (a_0{\oplus}...{\oplus}a_n)[x]:=(a_0{\oplus}...{\oplus}a_{n-1})\oplus a_n[x]\qquad(n\geq 1)$.

For technical reasons we also set $a[n]:=a[0]$, if $\mathrm{tp}(a)=1$.

Proof of Theorem 1.5:

(a),(b),(c) are easily verified by induction on $\ell(a)$ (length of the string a).

For a proof of (e) see [BS88, §5] or [Bu86, Lemma 3.3].

(d) is proved by induction on $\ell(a)$. All cases except 6. are straightforward. So let us assume $D_\sigma c\oplus\tilde c\prec D_\sigma a$. Then $c\prec a$, and by IH $c\prec a[x]$ for some $x\in\mathrm{OT}\cap|\mathrm{tp}(a)|$. Hence $D_\sigma c\oplus\tilde c\prec D_\sigma a[x]$. If $\mathrm{tp}(a)\in\{\omega\}\cup\{\Omega_{\mu+1}:\mu<\sigma\}$, it can be proved that $D_\sigma a[x]\in\mathrm{OT}$. Therefore in that case $D_\sigma a[x]$ is the canonical choice for $(D_\sigma a)[x]$.

Now let us assume that $\mathrm{tp}(a)=\Omega_{\mu+1}$ with $\mu\geq\sigma$.

Then "$D_\sigma a[x]\in\mathrm{OT}$" does no longer hold for arbitrary $x\in\mathrm{OT}\cap|\mathrm{tp}(a)|$.

But by induction on $\ell(a)$ one can prove

(1) $\mathrm{tp}(a)=\Omega_{\mu+1}\ \&\ c\in\mathrm{OT}\ \&\ a[\Omega_\mu]\preceq c\prec a\implies\exists x=D_\mu(b{+}1)\in\mathrm{OT}(\,b\in G_\mu c\ \&\ c\prec a[x]\,)$

from which we conclude

(2) $\mathrm{tp}(a)=\Omega_{\mu+1}\ \&\ c\in\mathrm{OT}\ \&\ a[\Omega_\mu]\preceq c\ \&\ \{c\}\cup G_\mu c\prec a\implies$

$$\implies\exists b\in\mathrm{OT}(\,\ell(b)<\ell(c)\ \&\ \{b\}\cup G_\mu b\prec a\ \&\ c\prec a[D_\mu(b{+}1)])$$

Obviously (2) suggests to define $x_0:=\Omega_\mu$, $x_{n+1}:=D_\mu a[x_n]$ in order to obtain by induction on $\ell(c)$

(3) $\mathrm{tp}(a)=\Omega_{\mu+1}\ \&\ c\in\mathrm{OT}\ \&\ \{c\}\cup G_\mu c\prec a\implies\exists n(c\prec a[x_n])$.

(Induction step: $c \prec a[D_\mu(b+1)]$ & $b \prec a[x_n] \Rightarrow c \prec a[D_\mu(b+1)] \preceq a[D_\mu a[x_n]] = a[x_{n+1}]$)

Now $\exists n(D_\sigma c \oplus \tilde{c} \prec (D_\sigma a)[n])$ is obtained as follows:

$\mathrm{OT} \ni D_\sigma c \oplus \tilde{c} \prec D_\sigma a \overset{\sigma\leq\mu}{\Longrightarrow} G_\mu c \subseteq G_\sigma c \prec c \prec a \overset{(3)}{\Rightarrow} \exists n(D_\sigma c \oplus \tilde{c} \prec D_\sigma a[x_n] = (D_\sigma a)[n])$.

Proof of (1):

1. $a = \Omega_{\mu+1}$: Then $c = D_\mu c_0 \oplus \tilde{c} \prec D_\mu(c_0 \oplus 1)$. Let $b := c_0$.
2. $a = a_0 \oplus a_1$ with $\mathrm{tp}(a_1) = \Omega_{\mu+1}$:
Then $c = a_0 \oplus c_1$ with $a_1[\Omega_\mu] \preceq c_1 \prec a_1$, and the claim follows immediately from the IH.
3. $a = D_\rho a_0$ with $\mu < \rho$: Then $\mathrm{tp}(a_0) = \Omega_{\mu+1}$ and $a[x] = D_\sigma a_0[x]$. Further $c = D_\rho c_0 \oplus \tilde{c}$ with $b[\Omega_\mu] \preceq c_0 \prec b$. By IH we get $c_0 \prec a_0[x]$ for some $x = D_\mu(b \oplus 1) \in \mathrm{OT}$ with $b \in G_\mu c_0$. Since $\mu < \rho$, $G_\mu c_0 \subseteq G_\mu c$. From $c_0 \prec a_0[x]$ we get $c \prec D_\rho a_0[x] = a[x]$.

Lemma 1.6

$$a \in \mathrm{OT}_0 \implies \mathrm{tp}(a) \in \{0,1,\omega\} \text{ and } \overline{\mathrm{o}}(a) = \begin{cases} 0 & \text{if } \mathrm{tp}(a) = 0 \\ \overline{\mathrm{o}}(a[0]) + 1 & \text{if } \mathrm{tp}(a) = 1 \\ \sup_{n\in\mathbb{N}}(\overline{\mathrm{o}}(a[n])+1) & \text{if } \mathrm{tp}(a) = \omega \end{cases}$$

Proof of the last claim:
Let $a \in \mathrm{OT}_0$ with $\mathrm{tp}(a) = \omega$. Then $a[n] \prec a$ & $a[n] \in \mathrm{OT}_0$ (by Theorem 1.5a,e) and $\overline{\mathrm{o}}(a[n]) < \overline{\mathrm{o}}(a)$ (by Theorem 1.4a). Now let $\gamma < \overline{\mathrm{o}}(a)$. Then $\gamma = \overline{\mathrm{o}}(c)$ for some $c \in \mathrm{OT}_0$ with $c \prec a$. Theorem 1.5d yields $c \prec a[n]$ for some n. Hence $\gamma < \overline{\mathrm{o}}(a[n])$.

§2 Collapsing of infinitary derivations

In this section we take a fresh start and introduce systems ID_ν^∞ of infinitary derivations together with cut-elimination $\mathcal{E}$ and collapsing operations $\mathcal{D}_\sigma$ by which every ID_ν^∞-derivation of an arithmetical formula A can be transformed into a cut-free derivation of A in ID_0^∞ (i.e. ω-arithmetic). ID_ν^∞ is so to speak an infinitary version of the formal theory ID_ν of ν-times iterated inductive definitions; every ID_ν-derivation h can be translated into an ID_ν^∞-derivation h^∞ of the same formula (or sequent). We will establish an upper bound η_ν for the proof-theoretic ordinal $|\mathrm{ID}_\nu|$ in terms of the operations $h \mapsto h^\infty$, $\mathcal{E}$, $\mathcal{D}_\sigma$, namely we will prove $|\mathrm{ID}_\nu| \leq \eta_\nu := \sup\{\|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\| : m \in \mathbb{N} \text{ and } h \text{ an } \mathrm{ID}_\nu\text{-derivation with endsequent of level } 0\ \}$, where $\|d\|$ denotes the length (or depth) of an infinitary derivation d.

As we will see below, the operations $\mathcal{E}$ and $\mathcal{D}_\sigma$ are closely related to the assignment of fundamental sequences $(a[x])_{x\in|\mathrm{tp}(a)|}$ given in §1. In §3 we will exploit this observation and prove $\|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\| \leq \psi_0\psi_\nu^{m+2}(0)$, which yields $|\mathrm{ID}_\nu| \leq \eta_\nu \leq \psi_0(\varepsilon_{\Omega_\nu+1})$.

Syntax

Let $\mathcal{L}$ be an arbitrary 1st order language (i.e. set of function and predicate symbols). Atomic $\mathcal{L}$-formulas are $Rt_1...t_n$ where R is an n-ary predicate symbol (of $\mathcal{L}$), and $t_1,...,t_n$ are $\mathcal{L}$-terms. Expressions of the shape A or $\neg A$, where A is an atomic $\mathcal{L}$-formula, are called *literals*. $\mathcal{L}$-formulas are built up from literals by means of $\wedge, \vee, \forall x, \exists x$. $\mathrm{FV}(A)$ denotes the set of free variables of A. A formula or term A is called *closed* if $\mathrm{FV}(A) = \emptyset$. The *negation* $\neg A$ of a non-atomic formula A is defined via de Morgan's laws. The *rank* $\mathrm{rk}(A)$ of a formula A is defined by: $\mathrm{rk}(A) := 0$ if A is a literal, $\mathrm{rk}(A \wedge B) := \mathrm{rk}(A \vee B) := \max\{\mathrm{rk}(A), \mathrm{rk}(B)\} + 1$, $\mathrm{rk}(\forall xA) := \mathrm{rk}(\exists xA) := \mathrm{rk}(A) + 1$. By $A(x/t)$ we denote the result of substituting t for (every free occurrence of) x in A (renaming bound variables if necessary). Expressions $\lambda x.F$ are called *predicates* and denoted by $\mathcal{F}$. For $\mathcal{F} = \lambda x.F$ we set $\mathcal{F}[t] := F(x/t)$. If $\mathcal{P}$ is a unary predicate symbol then $B(\mathcal{P}/\mathcal{F})$ denotes the result of substituting $\mathcal{F}$ for $\mathcal{P}$ in B, i.e. the formula resulting from B be replacing every atom $\mathcal{P}t$ by $\mathcal{F}[t]$.

Let X be unary predicate symbol not in $\mathcal{L}$. A *positive operator form in* $\mathcal{L}$ is an $\mathcal{L} \cup \{X\}$-formula $\mathfrak{A}$ in which X occurs only positively (i.e. $\mathfrak{A}$ has no subformula $\neg Xt$) and which has at most one free variable x.
We use the following abbreviations: $\mathfrak{A}(\mathcal{F}, t) := \mathfrak{A}(X/\mathcal{F}, x/t)$, $\mathfrak{A}(\mathcal{F}) \subseteq \mathcal{F} := \forall x(\mathfrak{A}(\mathcal{F}, x) \to \mathcal{F}[x])$.
For each positive operator form $\mathfrak{A}$ we introduce a (new) unary predicate symbol $\mathcal{P}_{\mathfrak{A}}$.

Finite sets of formulas are called *sequents*. They are denoted by Γ, Γ', Δ. We mostly write $A_1, ..., A_n$ for $\{A_1, ..., A_n\}$, and A, Γ, Δ for $\{A\} \cup \Gamma \cup \Delta$, etc.

Definition of the languages $\mathcal{L}_\sigma$ $(0 \leq \sigma < \omega)$
Let $\mathcal{L}_0$ be the language consisting of the constant 0 (*zero*), the unary function symbol S (*successor*), and predicate symbols R for primitive recursive relations.
$\mathcal{L}_{\sigma+1} := \mathcal{L}_0 \cup \{\mathcal{P}_{\mathfrak{A}} : \mathfrak{A}$ positive operator form in $\mathcal{L}_\sigma\}$
$\mathcal{L}_{<\omega} := \bigcup_{\sigma<\omega} \mathcal{L}_\sigma$

The only closed $\mathcal{L}_0$-terms are the *numerals* $0, S0, SS0, ...$ which we identify with the corresponding natural numbers (elements of $\mathbb{N}$). Arbitrary $\mathcal{L}_0$-terms will be denoted by $t, t_1, ...$, and (number) variables by x, y.
$\mathsf{TRUE}_0 :=$ set of all closed $\mathcal{L}_0$-literals which are true in the standard model $\mathfrak{N}$.

Definition of $\text{lev}(A)$
$\text{lev}(A) := 0$ if A is an $\mathcal{L}_0[X]$-literal
$\text{lev}(\mathcal{P}_{\mathfrak{A}}t) := \text{lev}(\mathfrak{A})$, $\text{lev}(\neg\mathcal{P}_{\mathfrak{A}}t) := \text{lev}(\mathfrak{A}) + 1$
$\text{lev}(A \wedge B) := \text{lev}(A \vee B) := \max\{\text{lev}(A), \text{lev}(B)\}$
$\text{lev}(\forall xA) := \text{lev}(\exists xA) := \text{lev}(A)$
$\text{lev}(\mathcal{P}_{\mathfrak{A}}) := \text{lev}(\mathfrak{A})$, $\text{lev}(\Gamma) := \max\{\text{lev}(A) : A \in \Gamma\}$

Remark
$\text{lev}(\mathcal{P}_{\mathfrak{A}}) < \sigma$ for each predicate symbol $\mathcal{P}_{\mathfrak{A}}$ in $\mathcal{L}_\sigma$,
$\text{lev}(A) \leq \sigma$ for each $\mathcal{L}_\sigma$-formula A.

From now on A, B, C denote $\mathcal{L}_{<\omega}$-formulas.

Proof systems
We will work in a Tait-style calculus with a somewhat unusual notion of derivation which is especially useful for the purposes of this paper.

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 $\{x\}$ with x a variable not in $\text{FV}(\Delta(\mathcal{I}))$; in the latter case x is called the *eigenvariable* of $\mathcal{I}$.

NOTATION
By writing

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

we declare $\mathcal{I}$ as an inference symbol with $|\mathcal{I}| = I$, $\Delta(\mathcal{I}) = \Delta$, $\Delta_\iota(\mathcal{I}) = \Delta_\iota$, and $\text{Eig}(\mathcal{I}) = \emptyset$ [$\text{Eig}(\mathcal{I}) = \{u\}$, resp.].

If $I = \{0, ..., n{-}1\}$ we write $\dfrac{\Delta_0\ \ \Delta_1\ \ \ldots\ \ \Delta_{n-1}}{\Delta}$, instead of $\dfrac{\ldots\Delta_\iota\ldots(\iota\in I)}{\Delta}$.

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 $d:=\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}$ is an $\mathfrak{S}$-*derivation* with $\Gamma(d):=\Gamma$ *(endsequent of d)* and $\mathrm{last}(d):=\mathcal{I}$ *(last inference of d).*

Instead of $\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}$ we also write $\dfrac{\ldots d_\iota\ldots(\iota\in|\mathcal{I}|)}{\mathcal{I}}$ or $\mathcal{I}d_0...d_{n-1}$ or $\dfrac{d_0\ldots d_{n-1}}{\mathcal{I}}$ if $|\mathcal{I}|=\{0,...,n-1\}$.

Abbreviation: $\mathfrak{S}\ni d\vdash\Gamma\;:\Longleftrightarrow\; d$ is an $\mathfrak{S}$-derivation with $\Gamma(d)\subseteq\Gamma$.

Remark
Our notion of derivations differs from the usual one in so far as our derivations have inferences (inference symbols) and not sequents assigned to their nodes. The sequent "belonging" to a node τ of a derivation d is not explicitly displayed, but can be computed by tree recursion from d (similarly for the free assumptions in a natural deduction style derivation).

The Tait-style inference rules in their traditional form $\dfrac{\ldots\Gamma,\Delta_\iota(\mathcal{I})\ldots}{\Gamma,\Delta(\mathcal{I})}$ are reobtained here as follows:

If $\mathcal{I}\in\mathfrak{S}$ and $\mathrm{Eig}(\mathcal{I})\cap\mathrm{FV}(\Gamma)=\emptyset$, then

from $\ldots\mathfrak{S}\ni d_\iota\vdash\Gamma,\Delta_\iota(\mathcal{I})\ldots(\forall\iota\in|\mathcal{I}|)$ we conclude $\mathfrak{S}\ni\mathcal{I}(d_\iota)_{\iota\in|\mathcal{I}|}\vdash\Gamma,\Delta(\mathcal{I})$.

The proof system ID_ν
The language of ID_ν is $\mathcal{L}_\nu$.

The inference symbols of ID_ν are

$(\mathsf{Ax}_\Delta)\quad\dfrac{}{\Delta}\quad$ if $\Delta\in\mathsf{Ax}(\nu)$

All we need to know about $\mathsf{Ax}(\nu)$ is that it is a set of $\mathcal{L}_\nu$-sequents such that
(i) $\Delta\in\mathsf{Ax}(\nu)\Rightarrow\Delta(\vec{x}/\vec{t})\in\mathsf{Ax}(\nu)$
(ii) $\Delta\in\mathsf{Ax}(\nu)$ & $\mathrm{FV}(\Delta)=\emptyset\;\Rightarrow\;\Delta\cap\mathsf{TRUE}_0\neq\emptyset$ or $\{\neg P,P\}\subseteq\Delta$ for some $P=\mathcal{P}_\mathfrak{A}n$.

$(\bigwedge_{A_0\wedge A_1})\quad\dfrac{A_0\qquad A_1}{A_0\wedge A_1}$

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

$(\bigwedge^y_{\forall xA})\quad\dfrac{A(x/y)}{\forall xA}\quad !y!$

$(\bigvee^t_{\exists xA})\quad\dfrac{A(x/t)}{\exists xA}$

$(\mathsf{Cut}_C)\quad\dfrac{C\qquad\neg C}{\emptyset}$

$(\mathsf{Ind}^t_\mathcal{F})\quad\dfrac{}{\neg\mathcal{F}[0],\neg\forall x(\mathcal{F}[x]\to\mathcal{F}[Sx]),\mathcal{F}[t]}$

$(\mathsf{Cl}_{\mathcal{P}_\mathfrak{A}t})\quad\dfrac{\mathfrak{A}(\mathcal{P}_\mathfrak{A},t)}{\mathcal{P}_\mathfrak{A}t}$

$(\mathsf{Ind}^{\mathcal{P}_\mathfrak{A},t}_\mathcal{F})\quad\dfrac{}{\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\neg\mathcal{P}_\mathfrak{A}t,\mathcal{F}[t]}$

The infinitary proof systems $\mathrm{ID}^\infty_\sigma$ $(\sigma < \omega)$

The language of $\mathrm{ID}^\infty_\sigma$ consists of all closed $\mathcal{L}_{<\omega}$-formulas A.

We use P [P_μ, resp.] as syntactic variable for formulas of the form $\mathcal{P}_\mathfrak{A} n$ [with $\mathrm{lev}(\mathfrak{A}) = \mu$, resp.].

The inference symbols of $\mathrm{ID}^\infty_\sigma$ are

(Ax_Δ) $\dfrac{}{\Delta}$, if $\Delta = \{A\}$ with $A \in \mathsf{TRUE}_0$ or $\Delta = \{\neg P, P\}$ with $\mathrm{lev}(P) < \sigma$

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

$(\bigwedge_{\forall x A})$ $\dfrac{\ldots A(x/i) \ldots (i \in \mathbb{N})}{\forall x A}$ $\qquad$ $(\bigvee^k_{\exists x A})$ $\dfrac{A(x/k)}{\exists x A}$ $(k \in \mathbb{N})$

$$(\mathsf{Cut}_C) \quad \frac{C \qquad \neg C}{\emptyset}$$

$(\mathsf{Cl}_{\mathcal{P}_\mathfrak{A} t})$ $\dfrac{\mathfrak{A}(\mathcal{P}_\mathfrak{A}, t)}{\mathcal{P}_\mathfrak{A} t}$ $(\mathrm{lev}(\mathcal{P}_\mathfrak{A}) \le \sigma)$ $\qquad$ $(\widetilde{\Omega}_P)$ $\dfrac{P \qquad \ldots \Delta^P_q \ldots (q \in |P|)}{\emptyset}$ $\qquad (\mathrm{lev}(P) < \sigma)$

$|P| :=$ set of all cutfree ID^∞_μ-derivations, where $\mu := \mathrm{lev}(P)$

$\Delta^P_q := \Gamma(q) \setminus \{P\}$

An ID^∞_μ-derivation d is called *cutfree* if $\text{c-rk}(d) = 0$ where $\text{c-rk}(d)$ is the least number greater than the ranks of all cut-formulas occurring in d, i.e.,

$\text{c-rk}(\mathcal{I}(d_\iota)_{\iota \in I}) := \sup(\{\text{c-rk}(\mathcal{I})\} \cup \{\text{c-rk}(d_\iota) : \iota \in I\})$ with $\text{c-rk}(\mathcal{I}) := \begin{cases} \mathrm{rk}(C) + 1 & \text{if } \mathcal{I} = \mathsf{Cut}_C \\ 0 & \text{otherwise} \end{cases}$

Abbreviation

$\mathrm{ID}^\infty_\sigma \ni d \vdash_m \Gamma \;:\Leftrightarrow\; \mathrm{ID}^\infty_\sigma \ni d \vdash \Gamma \;\&\; \text{c-rk}(d) \le m$

$d \vdash_m \Gamma \;:\Leftrightarrow\; \mathrm{ID}^\infty_\sigma \ni d \vdash_m \Gamma$ for some $\sigma < \omega$.

The set $\mathsf{ID}^\infty_\sigma$ of all $\mathrm{ID}^\infty_\sigma$-derivations is introduced by an inductive definition (as given above for arbitrary proof systems $\mathfrak{S}$) under the assumption that the sets ID^∞_μ for $\mu < \sigma$ are already defined.

We set $\mathsf{ID}^\infty_{<\omega} := \bigcup_{\sigma<\omega} \mathsf{ID}^\infty_\sigma$.

The $(\widetilde{\Omega}_P)$-rule can be motivated as follows (with $\mu := \mathrm{lev}(P) < \sigma$):

Imitating the constructive interpretation of implication we start by saying:

> *"An $\mathrm{ID}^\infty_\sigma$-derivation of $P \to B$ is an operation $q \mapsto d_q$ transforming every cutfree ID^∞_μ-derivation of P into an $\mathrm{ID}^\infty_\sigma$-derivation of B ".*

This may be replaced by the stricter version:

> *"An $\mathrm{ID}^\infty_\sigma$-derivation of $P \to B$ is an operation $q \mapsto d_q$ transforming every cutfree ID^∞_μ-derivation of $A \to P$ into an $\mathrm{ID}^\infty_\sigma$-derivation of $A \to B$ (for any formula A)".*

In terms of the Tait-calculus used here this amounts to the following rule:

(Ω_P) *If for each Δ and each cutfree ID^∞_μ-derivation q of Δ, P , d_q is an $\mathrm{ID}^\infty_\sigma$-derivation of Δ, Γ, then $(d_q)_{q \in |P|}$ is an $\mathrm{ID}^\infty_\sigma$-derivation of $\neg P, \Gamma$ ".*

Now $(\widetilde{\Omega}_P)$ is just a combination of (Ω_P) and (Cut_P).

The following definitions and Theorem 2.1 are needed for the embedding of ID_ν into ID^∞_ν, i.e., for deriving $\neg(\mathfrak{A}(\mathcal{F}) \subseteq \mathcal{F}), \neg \mathcal{P}_\mathfrak{A} n, \mathcal{F}[n]$ by means of $(\widetilde{\Omega}_{\mathcal{P}_\mathfrak{A}})$.

Definitions (Substitution)
For each closed $\mathcal{L}_\sigma$-formula A let $\mathsf{d}_{\neg A,A}$ be the canonical cutfree $\mathrm{ID}^\infty_\sigma$-derivation of $\neg A, A$
(e.g. $\mathsf{d}_{\neg P,P} := \mathsf{Ax}_{\{\neg P,P\}}$, $\mathsf{d}_{\neg A,A} := \mathsf{Ax}_{\{A\}}$ if $A \in \mathsf{TRUE}_0$, $\mathsf{d}_{\neg\forall xA,\forall xA} := \bigwedge_{\forall xA}\Big(\bigvee^i_{\neg\forall xA}\mathsf{d}_{\neg A(x/i),A(x/i)}\Big)_{i\in\mathbb{N}}$).

$$\mathsf{e}^n_{\mathfrak{A},\mathcal{F}} := \bigvee^n_{\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F})}\bigwedge_{\mathfrak{A}(\mathcal{F},n)\wedge\neg\mathcal{F}[n]}\mathsf{d}_{\neg\mathfrak{A}(\mathcal{F},n),\mathfrak{A}(\mathcal{F},n)}\mathsf{d}_{\neg\mathcal{F}[n],\mathcal{F}[n]} \approx \cfrac{\cfrac{\neg\mathfrak{A}(\mathcal{F},n),\mathfrak{A}(\mathcal{F},n) \qquad \neg\mathcal{F}[n],\mathcal{F}[n]}{\mathfrak{A}(\mathcal{F},n)\wedge\neg\mathcal{F}[n],\neg\mathfrak{A}(\mathcal{F},n),\mathcal{F}[n]}}{\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\neg\mathfrak{A}(\mathcal{F},n),\mathcal{F}[n]}$$

Given $\mathcal{P} = \mathcal{P}_\mathfrak{A}$, a predicate $\mathcal{F}$, and a sequent Π we define an operation $\mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}} : \mathsf{ID}^\infty_{\mathrm{lev}(\mathcal{P})} \to \mathsf{ID}^\infty_{<\omega}$ which transforms any derivation $d \in \mathsf{ID}^\infty_{\mathrm{lev}(\mathcal{P})}$ of Γ,Π into a derivation $d^* := \mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}}(d)$ of $\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\Gamma,\Pi(\mathcal{P}/\mathcal{F})$. Roughly speaking d^* results from d by substituting certain occurrences of $\mathcal{P}$ by $\mathcal{F}$. In doing so, some inferences $(\mathsf{Cl}_{\mathcal{P}n})\dfrac{\mathfrak{A}(\mathcal{P},n)}{\mathcal{P}n}$ are turned into $\dfrac{\mathfrak{A}(\mathcal{F},n)}{\mathcal{F}[n]}$ which is not an inference of $\mathrm{ID}^\infty_{<\omega}$.

Therefore those inferences $(\mathsf{Cl}_{\mathcal{P}n})$ are replaced by

$$\cfrac{\overset{d_0^*}{\mathfrak{A}(\mathcal{F},n)} \qquad \overset{\mathsf{e}^n_{\mathfrak{A},\mathcal{F}}}{\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\neg\mathfrak{A}(\mathcal{F},n),\mathcal{F}[n]}}{\neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\mathcal{F}[n]}\,(Cut)$$

The precise definition of $\mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}}(d)$ runs as follows

$$\mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}}\big(\mathcal{I}(d_\iota)_{\iota\in I}\big) := \begin{cases} \mathsf{Cut}_{\mathfrak{A}(\mathcal{F},n)}\mathcal{S}^{\Pi\cup\Delta_0(\mathcal{I})}_{\mathcal{P},\mathcal{F}}(d_0)\mathsf{e}^n_{\mathfrak{A},\mathcal{F}} & \text{if } \mathcal{I} = \mathsf{Cl}_{\mathcal{P}n} \text{ with } \mathcal{P}n \in \Pi \\ \mathcal{I}^*\big(\mathcal{S}^{\Pi\cup\Delta_\iota(\mathcal{I})}_{\mathcal{P},\mathcal{F}}(d_\iota)\big)_{\iota\in I} & \text{if } \mathcal{I} = \bigwedge_A \text{ or } \bigvee^k_A \text{ with } A \in \Pi \\ \mathcal{I}\big(\mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}}(d_\iota)\big)_{\iota\in I} & \text{otherwise} \end{cases}$$

where $(\bigwedge_A)^* := \bigwedge_{A(\mathcal{P}/\mathcal{F})}$, $(\bigvee^k_A)^* := \bigvee^k_{A(\mathcal{P}/\mathcal{F})}$.

The following theorem is easily verified. Note that the axioms $\mathsf{Ax}_{\{\neg\mathcal{P}n,\mathcal{P}n\}}$ do not belong to $\mathrm{ID}^\infty_{\mathrm{lev}(\mathcal{P})}$!

Theorem 2.1
$\mathrm{ID}^\infty_{\mathrm{lev}(\mathcal{P})} \ni d \vdash_0 \Gamma,\Pi \;\;\&\;\; \mathrm{rk}(\mathfrak{A}(\mathcal{F},x)) < m \implies \mathcal{S}^\Pi_{\mathcal{P},\mathcal{F}}(d) \vdash_m \neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\Gamma,\Pi(\mathcal{P}/\mathcal{F})$.

Embedding of ID_ν into ID^∞_ν
An ID_ν-derivation h is called *closed* if every number variable occurring free in h is the eigenvariable of an inference below that occurrence. Especially $\mathrm{FV}(\Gamma(h)) = \emptyset$ for closed h.
For each closed ID_ν-derivation h we define an ID^∞_ν-derivation h^∞ such that $h^\infty \vdash_m \Gamma(h)$ for some $m \in \mathbb{N}$.
0. $(\mathsf{Ax}_\Delta)^\infty := \mathsf{Ax}_{\Delta'}$ with a suitable $\Delta' \subseteq \Delta$
1. $(\bigwedge^y_{\forall xA}h_0)^\infty := \bigwedge_{\forall xA}\big(h_0(y/i)^\infty\big)_{i\in\mathbb{N}}$, where $h_0(y/i)$ is defined as expected
2. $(\mathsf{Ind}^n_\mathcal{F})^\infty := d_n$ with $d_0 := \mathsf{d}_{\neg\mathcal{F}[0],\mathcal{F}[0]}$, $d_{i+1} := \bigvee^i_{\exists x(\mathcal{F}[x]\wedge\neg\mathcal{F}[Sx])}\bigwedge_{\mathcal{F}[i]\wedge\neg\mathcal{F}[Si]}d_i\,\mathsf{d}_{\neg\mathcal{F}[Si],\mathcal{F}[Si]}$
3. $(\mathsf{Ind}^{\mathcal{P},n}_\mathcal{F})^\infty := \dfrac{\mathsf{Ax}_{\{\neg\mathcal{P}n,\mathcal{P}n\}} \quad \ldots\mathcal{S}^{\{\mathcal{P}n\}}_{\mathcal{P},\mathcal{F}}(q)\ldots(q\in|\mathcal{P}n|)}{\widetilde{\Omega}_{\mathcal{P}n}}$
4. Otherwise: $(\mathcal{I}h_0...h_{n-1})^\infty := \mathcal{I}h_0^\infty\ldots h_{n-1}^\infty$

Theorem 2.2 (Embedding)
$\mathrm{ID}_\nu \ni h \vdash \Gamma \;\;\&\;\; h$ closed $\implies \mathrm{ID}^\infty_\nu \ni h^\infty \vdash_m \Gamma$ for some $m \in \mathbb{N}$.

Proof: straightforward.
Especially $(\mathsf{Ind}^{\mathcal{P},n}_\mathcal{F})^\infty \vdash_m \neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\neg\mathcal{P}n,\mathcal{F}[n]$ (where $\mathcal{P} = \mathcal{P}_\mathfrak{A}$) is obtained from:
$q \in |\mathcal{P}n| \Rightarrow \mathrm{ID}^\infty_{\mathrm{lev}(\mathcal{P})} \ni q \vdash_0 \Delta^{\mathcal{P}n}_q,\mathcal{P}n \overset{\text{Theorem 2.1}}{\Rightarrow} \mathcal{S}^{\{\mathcal{P}n\}}_{\mathcal{P},\mathcal{F}}(q) \vdash_m \neg(\mathfrak{A}(\mathcal{F})\subseteq\mathcal{F}),\Delta^{\mathcal{P}n}_q,\mathcal{F}[n]$.

Abbreviations
$\bigwedge$-For := set of all formulas of the shape $A \wedge B$ or $\forall x A$.
$\bigwedge^{+}$-For := $\mathsf{TRUE}_0 \cup \bigwedge$-For $\cup$ set of all formulas $\mathcal{P}_{\mathfrak{A}} n$.

$$C[k] := \begin{cases} C_k & \text{if } C = C_0 \overset{\vee}{\wedge} C_1 \text{ and } k \in \{0,1\} \\ A(x/k) & \text{if } C = \overset{\exists}{\forall} x A \text{ and } k \in \mathbb{N} \end{cases}$$

Theorem 2.3
By tree recursion one can define operations $\mathcal{J}_C^k$, $\mathcal{R}_C$, $\mathcal{E}$, $\mathcal{D}_\sigma$ on $\mathsf{ID}^\infty_{<\omega}$ with the following properties:
($\bigwedge$-Inversion) $d \vdash_m \Gamma, C$ & $C \in \bigwedge$-For $\implies \mathcal{J}_C^k(d) \vdash_m C[k]$
(Reduction) $e \vdash_m \Gamma, C$ & $d \vdash_m \Gamma, \neg C$ & $C \in \bigwedge^{+}$-For & $\mathrm{rk}(C) \le m \implies \mathcal{R}_C(e,d) \vdash_m \Gamma$.
(Elimination) $d \vdash_{m+1} \Gamma \implies \mathcal{E}(d) \vdash_m \Gamma$.
(Collapsing) $d \vdash_0 \Gamma$ & $\mathrm{lev}(\Gamma) \le \sigma \Rightarrow \mathrm{ID}^\infty_\sigma \ni \mathcal{D}_\sigma(d) \vdash_0 \Gamma$.

Proof:
For $d = \mathcal{I}(d_\iota)_{\iota \in I} \in \mathsf{ID}^\infty_{<\omega}$ and $e \in \mathsf{ID}^\infty_{<\omega}$ we define

$$\mathcal{J}_C^k(d) := \begin{cases} \mathcal{J}_C^k(d_k) & \text{if } \mathcal{I} = \bigwedge_C \\ \mathcal{I}\big(\mathcal{J}_C^k(d_\iota)\big)_{\iota \in I} & \text{otherwise} \end{cases} \qquad (C \in \textstyle\bigwedge\text{-For})$$

$$\mathcal{R}_C(e,d) := \begin{cases} \mathsf{Cut}_{C[k]} \mathcal{J}_C^k(e) \mathcal{R}_C(e,d_0) & \text{if } \mathcal{I} = \bigvee^k_{\neg C} \\ e & \text{if } \mathcal{I} = \mathsf{Ax}_{\{\neg C, C\}} \\ \mathcal{I}\big(\mathcal{R}_C(e,d_\iota)\big)_{\iota \in I} & \text{otherwise (i.e., if } \neg C \notin \Delta(\mathcal{I})) \end{cases} \qquad (C \in \textstyle\bigwedge^{+}\text{-For})$$

$$\mathcal{E}(d) := \begin{cases} \mathcal{R}_C(\mathcal{E}(d_0), \mathcal{E}(d_1)) & \text{if } \mathcal{I} = \mathsf{Cut}_C \text{ with } C \in \bigwedge^{+}\text{-For} \\ \mathcal{R}_{\neg C}(\mathcal{E}(d_1), \mathcal{E}(d_0)) & \text{if } \mathcal{I} = \mathsf{Cut}_C \text{ with } \neg C \in \bigwedge^{+}\text{-For}. \\ \mathcal{I}\big(\mathcal{E}(d_\iota)\big)_{\iota \in I} & \text{otherwise} \end{cases}$$

$$\mathcal{D}_\sigma(d) := \begin{cases} \mathcal{D}_\sigma(d_{\mathcal{D}_\mu(d_0)}) & \text{if } \mathcal{I} = \widetilde{\Omega}_P \text{ with } \mu := \mathrm{lev}(P) \ge \sigma \\ \mathcal{I}\big(\mathcal{D}_\sigma(d_\iota)\big)_{\iota \in I} & \text{otherwise} \end{cases}$$

One easily verifies that the so defined operations have the asserted properties.
Let us look at $\mathcal{D}_\sigma(d)$ for $d = \widetilde{\Omega}_P(d_q)_{q \in \{0\} \cup |P|} \vdash_0 \Gamma$ with $\mathrm{lev}(\Gamma) \le \sigma \le \mu := \mathrm{lev}(P)$.
Then $d_0 \vdash_0 \Gamma, P$ and $d_q \vdash_0 \Gamma, \Delta^P_q$ for all $q \in |P|$ (†).
By IH $\mathrm{ID}^\infty_\mu \ni q_0 := \mathcal{D}_\mu(d_0) \vdash_0 \Gamma, P$. Hence $q_0 \in |P|$ and $\Delta^P_{q_0} \subseteq \Gamma$.
Now (†) yields $d_{q_0} \vdash_0 \Gamma$, and by IH we get $\mathrm{ID}^\infty_\sigma \ni \mathcal{D}_\sigma(d_{q_0}) \vdash_0 \Gamma$.

Remark: The definition of $\mathcal{D}_\sigma(d)$ almost automatically arises if one pursues the goal to eliminate from d all $\widetilde{\Omega}_P$-inferences with $\mathrm{lev}(P) \ge \sigma$.

Definition
For $\mathfrak{A}$ with $\mathrm{lev}(\mathfrak{A}) = 0$ let $\Phi^\alpha_{\mathfrak{A}} := \{n : \mathfrak{A}(\Phi^{<\alpha}_{\mathfrak{A}}, n)\}$, where $\Phi^{<\alpha}_{\mathfrak{A}} := \bigcup_{\xi<\alpha} \Phi^\xi_{\mathfrak{A}}$ $(\alpha \in On)$.
$|n|_{\mathfrak{A}} := \min\{\alpha : n \in \Phi^\alpha_{\mathfrak{A}}\}$ (if $n \in \bigcup_{\alpha \in On} \Phi^\alpha_{\mathfrak{A}}$)
$|\mathrm{ID}_\nu| := \sup\{|n|_{\mathfrak{A}} : \mathrm{lev}(\mathfrak{A}) = 0 \;\&\; \mathrm{ID}_\nu \vdash \mathcal{P}_{\mathfrak{A}} n\}$ *(proof-theoretic ordinal of* ID_ν*)*

By $(\mathfrak{N}, \Phi^{<\alpha})$ we denote the expansion of the standard model $\mathfrak{N}$ where
each predicate constant $\mathcal{P}_{\mathfrak{A}}$ of level 0 is interpreted by $\Phi^{<\alpha}_{\mathfrak{A}}$.

$\|\mathcal{I}(d_\iota)_{\iota \in I}\| := \sup_{\iota \in I}(\|d_\iota\|{+}1)$ (*length* or *depth* of d)

Theorem 2.4 (Boundedness)
$\mathrm{ID}_0^\infty \ni d \vdash_0 \Gamma \;\;\&\;\; \mathrm{lev}(\Gamma)=0 \implies (\mathfrak{N}, \Phi^{<\|d\|}) \models \Gamma$
Proof by induction on $\|d\|$.

Theorem 2.5
If h is a closed ID_ν-derivation of Γ with $\mathrm{lev}(\Gamma)=0$ then
$(\mathfrak{N}, \Phi^{<\alpha}) \models \Gamma$ with $\alpha = \|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\|$ for some $m \in \mathbb{N}$.
Proof:

$$\begin{aligned}
\mathrm{ID}_\nu \ni h \vdash \Gamma &\overset{\text{Embedding}}{\Longrightarrow} \mathrm{ID}_\nu^\infty \ni h^\infty \vdash_m \Gamma \;\text{ for some } m\\
&\overset{\text{Cutelim}}{\Longrightarrow} \mathrm{ID}_\nu^\infty \ni \mathcal{E}^m(h^\infty) \vdash_0 \Gamma\\
&\overset{\text{Collapsing}}{\Longrightarrow} \mathrm{ID}_0^\infty \ni \mathcal{D}_0(\mathcal{E}^m(h^\infty)) \vdash_0 \Gamma\\
&\overset{\text{Boundedness}}{\Longrightarrow} (\mathfrak{N}, \Phi^{<\alpha}) \models \Gamma \text{ with } \alpha := \|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\|
\end{aligned}$$

Definition
$\eta_\nu := \sup\{\|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\| : m \in \mathbb{N}$ and h a closed ID_ν-derivation with endsequent of level 0 $\}$

Then Theorem 2.5 shows that $|\mathrm{ID}_\nu| \le \eta_\nu$. In what follows we will prove $\eta_\nu \le \sup\limits_{m\in\mathbb{N}} \psi_0\psi_\nu^m(0) = \psi_0(\varepsilon_{\Omega_\nu+1})$.

Remark
Note the similarity between
" $\mathcal{D}_\sigma(d) = \mathcal{D}_\sigma(d_{\mathcal{D}_\mu(d_0)})$ if $d = \widetilde{\Omega}_P(d_\iota)_{\iota\in\{0\}\cup|P|}$ with $\mu = \mathrm{lev}(P) \ge \sigma$ "
and
" $(D_\sigma a)[1] = D_\sigma a[D_\mu a[\Omega_\mu]]$ if $a \in \mathrm{T}$ and $\mathrm{tp}(a) = \Omega_{\mu+1}$ with $\mu \ge \sigma$ ".

This observation will be pursued in §3.

§3 Majorization of infinitary derivations by tree ordinals

We are now going to relate infinitary derivations $d \in \mathrm{ID}_\nu^\infty$ to ordinals (ordinal notations) $a \in \mathrm{OT}$. Here we heavily utilize the assignment of fundamental sequences from §1, which so to speak turns each $a \in \mathrm{OT}$ into a wellfounded tree, a so-called *tree ordinal* $\mathfrak{o}(a)$, namely $\mathfrak{o}(a) := \big(\mathfrak{o}(a[x])\big)_{x\in|\mathrm{tp}(a)|}$. On the other side, from every derivation $d \in \mathrm{ID}_\nu^\infty$ one obtains a tree ordinal $\mathfrak{o}(d)$ essentially by deleting all inference symbols (and possibly other data) assigned to the nodes of d (namely $\mathfrak{o}(\mathcal{I}(d_\iota)_{\iota\in I}) := \big(\mathfrak{o}(d_\iota)\big)_{\iota\in I}$). Now the first idea which comes into mind is that $\mathfrak{o}(d)$ should equal $\mathfrak{o}(a)$ for suitable $a \in \mathrm{OT}$ (at least if $d = h^\infty$ with $h \in \mathrm{ID}_\nu$). But this doesn't work; instead one can establish a weaker relation between $\mathfrak{o}(d)$ and $\mathfrak{o}(a)$, namely that in a certain sense $\mathfrak{o}(d)$ is "embeddable" into $\mathfrak{o}(a)$. Below we will define a relation $d \lhd \mathfrak{a}$ (d is *majorized* by $\mathfrak{a}$) between infinitary derivations d and tree ordinals $\mathfrak{a}$, which corresponds to this informal notion of embeddability. The main properties of $\lhd$ will be: (i) $d \lhd \mathfrak{a} \;\&\; d \in \mathsf{ID}_0^\infty \;\Rightarrow\; \|d\| \le \|\mathfrak{a}\|$, (ii) $d \lhd \mathfrak{a} \;\&\; d \in \mathsf{ID}_\nu^\infty \;\Rightarrow\; \mathcal{E}(d) \lhd \mathfrak{D}_\nu(\mathfrak{a})$, (iii) $d \lhd \mathfrak{a} \;\Rightarrow\; \mathcal{D}_\sigma(d) \lhd \mathfrak{D}_\sigma(\mathfrak{a})$. Here $\mathfrak{D}_\sigma$ is a collapsing function on tree ordinals defined in close analogy to $\mathcal{D}_\sigma$; at the same time $\mathfrak{D}_\sigma$ is closely related to ψ_σ, it is so to speak the "tree version" of ψ_σ, which becomes clear by comparing the definition of $\mathfrak{D}_\sigma$ (see below) with the corresponding clauses in the definition of the fundamental sequences in §1. Mainly by means of (i)-(iii) we will establish that $\|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\| \le \|\mathfrak{D}_0\mathfrak{D}_\nu^{m+2}(\mathbf{0})\|$ and thus $\eta_\nu \le \sup\limits_{m\in\mathbb{N}} \|\mathfrak{D}_0\mathfrak{D}_\nu^m(\mathbf{0})\|$. Finally we will show that $\|\mathfrak{D}_0(\mathfrak{D}_\nu^m(\mathbf{0}))\| = \psi_0\psi_\nu^m(0)$ which then yields $|\mathrm{ID}_\nu| \le \eta_\nu \le \sup\limits_{m\in\mathbb{N}} \psi_0\psi_\nu^m(0) = \psi_0(\varepsilon_{\Omega_\nu+1})$.

Inductive definition of classes $\mathfrak{T}_\sigma$ of tree ordinals
1. $\mathbf{0} := (\,) \in \mathfrak{T}_\sigma$
2. $\mathfrak{a} \in \mathfrak{T}_\sigma \;\Rightarrow\; \mathfrak{a}{+}\mathbf{1} := (\mathfrak{a}) \in \mathfrak{T}_\sigma$
3. $\forall n \in \mathbb{N}(\mathfrak{a}_n \in \mathfrak{T}_\sigma) \;\Rightarrow\; (\mathfrak{a}_n)_{n\in\mathbb{N}} \in \mathfrak{T}_\sigma$
4. $\mu < \sigma \;\&\; \forall \mathfrak{x} \in \mathfrak{T}_\mu(\mathfrak{a}_\mathfrak{x} \in \mathfrak{T}_\sigma) \;\Rightarrow\; (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu} \in \mathfrak{T}_\sigma$

$\mathfrak{T}_{<\omega} := \bigcup_{\sigma<\omega} \mathfrak{T}_\sigma$. The elements of $\mathfrak{T}_{<\omega}$ are called *tree ordinals* (denoted by $\mathfrak{a}, \mathfrak{b}, \mathfrak{c}$).

Note
Every $\mathfrak{a} \in \mathfrak{T}_\sigma$ is of the form $(\mathfrak{a}_\iota)_{\iota\in I}$ with $I \in \{\emptyset, \{0\}, \mathbb{N}\} \cup \{\mathfrak{T}_\mu : \mu < \sigma\}$.
We define $\|(\mathfrak{a}_\iota)_{\iota\in I}\| := \sup_{\iota\in I}(\|\mathfrak{a}_\iota\| + 1)$.

Abbreviations
$\overline{0} := \mathbf{0}$, $\overline{n{+}1} := \overline{n}{+}\mathbf{1}$, $\mathbf{1} := \overline{1}$, $\mathbf{\Omega}_0 := (\overline{n})_{n\in\mathbb{N}}$, $\mathbf{\Omega}_{\mu+1} := (\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$

Definition of $\mathfrak{a} + \mathfrak{b}$ and $\mathfrak{a} \cdot n$
$\mathfrak{a} + \mathbf{0} := \mathfrak{a}$, $\mathfrak{a} + (\mathfrak{b}_\mathfrak{x})_{\mathfrak{x}\in I} := (\mathfrak{a} + \mathfrak{b}_\mathfrak{x})_{\mathfrak{x}\in I}$ if $I \neq \emptyset$,
$\mathfrak{a} \cdot 0 := \mathbf{0}$, $\mathfrak{a} \cdot (n{+}1) := (\mathfrak{a} \cdot n) + \mathfrak{a}$

Proposition. a) $\mathfrak{a}, \mathfrak{b} \in \mathfrak{T}_\sigma \Rightarrow \mathfrak{a} + \mathfrak{b} \in \mathfrak{T}_\sigma$, b) $\mathfrak{a} + (\mathfrak{b} + \mathfrak{c}) = (\mathfrak{a} + \mathfrak{b}) + \mathfrak{c}$

Definition of $\mathfrak{D}_\sigma : \mathfrak{T}_{<\omega} \longrightarrow \mathfrak{T}_\sigma$
The definition of $\mathfrak{D}_\sigma(\mathfrak{a})$ proceeds by transf. rec. on $\mathfrak{a}$ simultaneously for all $\sigma < \omega$.
$\mathfrak{D}_0(\mathbf{0}) := \mathbf{1}$, $\mathfrak{D}_\sigma(\mathbf{0}) := \mathbf{\Omega}_\sigma$ if $\sigma \neq 0$
$\mathfrak{D}_\sigma(\mathfrak{a}{+}\mathbf{1}) := \big(\mathfrak{D}_\sigma(\mathfrak{a}) \cdot (n{+}1)\big)_{n\in\mathbb{N}}$

$$\mathfrak{D}_\sigma\big((\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in I}\big) := \begin{cases} \big(\mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in I} & \text{if } I \in \{\mathbb{N}\} \cup \{\mathfrak{T}_\mu : \mu{<}\sigma\} \\ \big(\mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_n})\big)_{n\in\mathbb{N}} & \text{if } I = \mathfrak{T}_\mu \text{ with } \mu \geq \sigma \\ \text{with } \mathfrak{x}_0 := \mathbf{\Omega}_\mu,\ \mathfrak{x}_{n+1} := \mathfrak{D}_\mu(\mathfrak{a}_{\mathfrak{x}_n}) & \end{cases}$$

Remark
1. For $\mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in I} \in \mathfrak{T}_\sigma \setminus \{0\}$ we have $\mathfrak{D}_\sigma(\mathfrak{a}) = \begin{cases} \big(\mathfrak{D}_\sigma(\mathfrak{a}_0) \cdot (n{+}1)\big)_{n\in\mathbb{N}} & \text{if } I = \{0\} \\ \big(\mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in I} & \text{otherwise} \end{cases}$.

   This means that on $\mathfrak{T}_\sigma$ the function $\mathfrak{D}_\sigma$ behaves like the ordinal function $\alpha \mapsto \omega^{\Omega_\sigma+\alpha}$ (if $\sigma > 0$) or $\alpha \mapsto \omega^\alpha$ (if $\sigma = 0$).
2. The canonical analogue to the collapsing function $\mathcal{D}_\sigma$ from §2 would be

   $$\mathfrak{D}_\sigma\big((\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in I}\big) := \begin{cases} \mathfrak{D}_\sigma \mathfrak{a}_{\mathfrak{D}_\mu(\mathfrak{a}_0)} & \text{if } I = \mathfrak{T}_\mu \text{ with } \mu \geq \sigma \\ \big(\mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in I} & \text{otherwise} \end{cases}$$

   But we have chosen the above version of $\mathfrak{D}_\sigma$, since this precisely corresponds to ψ_σ and at the same time is not to far removed from $\mathcal{D}_\sigma$.

The following definition and lemma are auxiliary.

Definition of $\mathfrak{a}^\ominus$, $\ll^0$ and $\ll$

$$\mathfrak{a}^\ominus := \begin{cases} \mathfrak{a}_0 & \text{if } \mathfrak{a} = \mathfrak{a}_0{+}\mathbf{1} \text{ or } \mathfrak{a} = (\mathfrak{a}_i)_{i\in\mathbb{N}} \\ \mathfrak{a}_{\mathbf{\Omega}_\mu} & \text{if } \mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu} \end{cases}$$

$\mathfrak{b} \ll^0 \mathfrak{a} \;:\Longleftrightarrow\; (\mathfrak{a} \neq \mathbf{0} \;\&\; \mathfrak{b} = \mathfrak{a}^\ominus)$ or $(\mathfrak{a} = (\mathfrak{a}_i)_{i\in\mathbb{N}} \;\&\; \exists i{\in}\mathbb{N}(\mathfrak{b} = \mathfrak{a}_i))$
$\ll$ ($\underline{\ll}$, resp.) is the transitive (transitive and reflexive, resp.) closure of $\ll^0$.

Lemma 3.1
(a) $\mathfrak{a} \neq \mathbf{0} \;\Rightarrow\; (\mathfrak{c} + \mathfrak{a})^\ominus = \mathfrak{c} + \mathfrak{a}^\ominus \;\;\&\;\; \mathfrak{D}_\sigma(\mathfrak{a})^\ominus = \mathfrak{D}_\sigma(\mathfrak{a}^\ominus)$
(b) $\mathbf{1} \underline{\ll} \mathfrak{a}$ if $\mathfrak{a} \neq \mathbf{0}$

(c) $\mathfrak{b} \ll \mathfrak{a} \;\Rightarrow\; \mathfrak{c}+\mathfrak{b} \ll \mathfrak{c}+\mathfrak{a}$
(d) $\mathfrak{b} \ll \mathfrak{a} \;\Rightarrow\; \mathfrak{D}_\sigma\mathfrak{b} \ll \mathfrak{D}_\sigma\mathfrak{a}$
(e) $\overline{n} \ll \mathbf{\Omega}_\sigma \ll \mathbf{\Omega}_{\sigma+1}$

Definition of $d \lhd \mathfrak{a}$ *(Majorization)*
$d \lhd \mathfrak{a}$ if one of the following clauses holds:
($\lhd$1) $d = \mathcal{I}(d_i)_{i\in|\mathcal{I}|}$ with $\mathcal{I} \neq \widetilde{\Omega}_P$ and $\mathfrak{a} = \mathfrak{b}{+}\mathbf{1}$ with $d_i \lhd \mathfrak{b}$ for all $i \in |\mathcal{I}|$
($\lhd$2) $d = \widetilde{\Omega}_{P_\mu}(d_q)_{q\in\{0\}\cup|P_\mu|}$ & $\mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ & $\forall q \in \{0\}\cup|P_\mu|\,\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow d_q \lhd \mathfrak{a}_\mathfrak{x})$
($\lhd$3) $d \lhd \mathfrak{b}$ & $\mathfrak{b} \ll \mathfrak{a}$
(By convention $0 \lhd \mathfrak{a}$ for any $\mathfrak{a}$.)

Lemma 3.2 $d \lhd \mathfrak{a}$ & $\mathfrak{a} \in \mathfrak{T}_0 \implies \|d\| \le \|\mathfrak{a}\|$.

Theorem 3.3
(a) $d \lhd \mathfrak{a} \implies \mathcal{J}^k_C(d) \lhd \mathfrak{a}$
(b) $d \lhd \mathfrak{a} \implies \mathcal{S}^{\Pi}_{\mathcal{P},\mathcal{F}}(d) \lhd \mathbf{\Omega}_\sigma + \mathfrak{a}$ for each σ
(c) $e \lhd \mathfrak{b}$ & $d \lhd \mathfrak{a} \implies \mathcal{R}_C(e,d) \lhd \mathfrak{b}+\mathfrak{a}$
(d) $d \lhd \mathfrak{a} \in \mathfrak{T}_\nu \implies \mathcal{E}(d) \lhd \mathfrak{D}_\nu(\mathfrak{a})$
(e) $d \lhd \mathfrak{a} \implies \mathcal{D}_\sigma(d) \lhd \mathfrak{D}_\sigma(\mathfrak{a})$
Proof by induction on $\mathfrak{a}$:
We only carry out the essential cases of (c),(d),(e).
(c) 1. $d = \mathsf{Ax}_{\{\neg C, C\}}$: $\mathcal{R}(e,d) = e \lhd \mathfrak{b} \underline{\ll} \mathfrak{b}+\mathfrak{a}$.
2. $d = \bigvee^k_{\neg C} d_0$ & $\mathfrak{a} = \mathfrak{a}_0{+}\mathbf{1}$ & $d_0 \lhd \mathfrak{a}_0$:
$\mathcal{R}(e,d_0) \overset{\text{IH}}{\lhd} \mathfrak{b}+\mathfrak{a}_0$ & $\mathcal{J}(e) \overset{\text{(a)}}{\lhd} \mathfrak{b} \underline{\ll} \mathfrak{b}+\mathfrak{a}_0 \implies \mathcal{R}(e,d) = \mathsf{Cut}\,\mathcal{J}(e)\mathcal{R}(e,d_0) \lhd (\mathfrak{b}+\mathfrak{a}_0)+\mathbf{1} = \mathfrak{b}+\mathfrak{a}$.
3. $d = \widetilde{\Omega}_{P_\mu}(d_q)_{q\in I}$ & $\mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ & $\forall q \in I\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow d_q \lhd \mathfrak{a}_\mathfrak{x})$:
IH $\Rightarrow$ $\forall q \in I\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow \mathcal{R}(e,d_q) \lhd \mathfrak{b}+\mathfrak{a}_\mathfrak{x})$ $\Rightarrow$ $\mathcal{R}(e,d) = \widetilde{\Omega}_{P_\mu}\big(\mathcal{R}(e,d_q)\big)_{q\in I} \lhd (\mathfrak{b}+\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu} = \mathfrak{b}+\mathfrak{a}$.
(d) 1. $d = \mathsf{Cut}_C d_0 d_1$ with $C \in \bigwedge^+$-For, and $\mathfrak{a} = \mathfrak{a}_0{+}\mathbf{1}$ & $d_0, d_1 \lhd \mathfrak{a}_0$: IV $\Rightarrow$ $\mathcal{E}(d_i) \lhd \mathfrak{D}_\nu(\mathfrak{a}_0) \overset{\text{(c)}}{\Rightarrow} \mathcal{E}(d) = \mathcal{R}_C(\mathcal{E}(d_0),\mathcal{E}(d_1)) \lhd \mathfrak{D}_\nu(\mathfrak{a}_0)+\mathfrak{D}_\nu(\mathfrak{a}_0) = \mathfrak{D}_\nu(\mathfrak{a}_0)\cdot 2 \;\Rightarrow\; \mathcal{E}(d) \lhd \big(\mathfrak{D}_\nu(\mathfrak{a}_0)\cdot(n{+}1)\big)_{n\in\mathbb{N}} = \mathfrak{D}_\nu(\mathfrak{a})$.
2. $d = \widetilde{\Omega}_{P_\mu}(d_q)_{q\in\{0\}\cup|P_\mu|}$ & $\mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ & $\forall q \in \{0\}\cup|P_\mu|\,\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow d_q \lhd \mathfrak{a}_\mathfrak{x})$:
Since $\mathfrak{a} \in \mathfrak{T}_\nu$, we have $\mu < \nu$ and $\mathfrak{D}_\nu(\mathfrak{a}) = \big(\mathfrak{D}_\nu(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in\mathfrak{T}_\mu}$.
IH $\Rightarrow$ $\forall q \in \{0\}\cup|P_\mu|\,\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow \mathcal{E}(d_q) \lhd \mathfrak{D}_\nu(\mathfrak{a}_\mathfrak{x})) \overset{\text{Def}}{\Rightarrow} \mathcal{E}(d) = \widetilde{\Omega}_{P_\mu}(\mathcal{E}(d_q))_{q\in\{0\}\cup|P_\mu|} \lhd \big(\mathfrak{D}_\nu(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in\mathfrak{T}_\mu}$.
(e) 1. $d = \mathcal{I}(d_i)_{i\in I}$ with $\mathcal{I} \neq \widetilde{\Omega}_P$ and $\mathfrak{a} = \mathfrak{b}{+}\mathbf{1}$ with $d_i \lhd \mathfrak{b}$ for all $i \in I$: IH $\Rightarrow$ $\forall i(\,\mathcal{D}_\sigma(d_i) \lhd \mathfrak{D}_\sigma(\mathfrak{b})\,)$ $\Rightarrow$
$\Rightarrow$ $\mathcal{D}_\sigma(d) = \mathcal{I}(\mathcal{D}_\sigma(d_i))_{i\in I} \lhd \mathfrak{D}_\sigma(\mathfrak{b})+\mathbf{1} \ll \mathfrak{D}_\sigma(\mathfrak{b})+\mathbf{\Omega}_\sigma \underline{\ll} \mathfrak{D}_\sigma(\mathfrak{b})+\mathfrak{D}_\sigma(\mathfrak{b}) \underline{\ll} \mathfrak{D}_\sigma(\mathfrak{b}+\mathbf{1})$.
2. $d = \widetilde{\Omega}_{P_\mu}(d_q)_{q\in I}$ & $\mathfrak{a} = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ & $\forall q \in I\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow d_q \lhd \mathfrak{a}_\mathfrak{x})$:
2.1. $\mu < \sigma$: IH $\Rightarrow$ $\forall q \in I\forall\mathfrak{x}\in\mathfrak{T}_\mu(q \lhd \mathfrak{x} \Rightarrow \mathcal{D}_\sigma(d_q) \lhd \mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x}))$ $\Rightarrow$
$\Rightarrow$ $\mathcal{D}_\sigma(d) = \widetilde{\Omega}_{P_\mu}(\mathcal{D}_\sigma(d_q))_{q\in I} \lhd (\mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x}))_{\mathfrak{x}\in\mathfrak{T}_\mu} = \mathfrak{D}_\sigma(\mathfrak{a})$.
2.2. $\mu \ge \sigma$: Then $\mathcal{D}_\sigma(d) = \mathcal{D}_\sigma(d_{\mathcal{D}_\mu(d_0)})$ and $\mathfrak{D}_\sigma(\mathfrak{a}) = (\mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_n}))_{n\in\mathbb{N}}$ with $\mathfrak{x}_0 = \mathbf{\Omega}_\mu$, $\mathfrak{x}_{n+1} = \mathfrak{D}_\mu(\mathfrak{a}_{\mathfrak{x}_n})$.
$0 \lhd \mathfrak{x}_0 \;\Rightarrow\; d_0 \lhd \mathfrak{a}_{\mathfrak{x}_0} \overset{\text{IH}}{\Longrightarrow} q := \mathcal{D}_\mu(d_0) \lhd \mathfrak{D}_\mu(\mathfrak{a}_{\mathfrak{x}_0}) = \mathfrak{x}_1 \implies d_q \lhd \mathfrak{a}_{\mathfrak{x}_1} \overset{\text{IH}}{\Longrightarrow}$
$\mathcal{D}_\sigma(d) = \mathcal{D}_\sigma(d_q) \lhd \mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_1}) \ll \big(\mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_i})\big)_{i\in\mathbb{N}} = \mathfrak{D}_\sigma(\mathfrak{a})$.

Theorem 3.4 (Embedding)
For each closed ID_ν-derivation h we have $h^\infty \lhd \mathbf{\Omega}_\nu{\cdot}2 + n(h)$,
where $n(\mathcal{I}h_0...h_{m-1}) := \max\{0, n(h_0), ..., n(h_{m-1})\} + 1$
Proof:

By definition $(\mathsf{Ind}_{\mathcal{F}}^{\mathcal{P},n})^\infty = \dfrac{\mathsf{Ax}_{\{\neg\mathcal{P}n,\mathcal{P}n\}} \quad \ldots\mathcal{S}_{\mathcal{P},\mathcal{F}}^{\{\mathcal{P}n\}}(q)\ldots(q\in|\mathcal{P}n|)}{\widetilde{\Omega}_{\mathcal{P}n}}$.

By Theorem 3.3b we have $\forall q\in|P_\mu|\forall\mathfrak{x}\in\mathfrak{T}_\mu(q\lhd\mathfrak{x}\Rightarrow\mathcal{S}_{\mathcal{P},\mathcal{F}}^{\{\mathcal{P}n\}}(q)\lhd\mathbf{\Omega}_\nu+\mathfrak{x})$ which together with $\forall\mathfrak{x}\in\mathfrak{T}_\mu(\mathsf{Ax}_{\{\neg\mathcal{P}n,\mathcal{P}n\}}\lhd\mathbf{\Omega}_\nu+\mathfrak{x})$ yields $(\mathsf{Ind}_{\mathcal{F}}^{\mathcal{P},n})^\infty\lhd\mathbf{\Omega}_\nu+\mathbf{\Omega}_{\mathrm{lev}(\mathcal{P})+1}\lll\mathbf{\Omega}_\nu+\mathbf{\Omega}_\nu=\mathbf{\Omega}_\nu{\cdot}2$.
The other cases are easy.

Theorem 3.5
Let $\nu>0$. If h is a closed ID_ν-derivation of Γ with $\mathrm{lev}(\Gamma)=0$ then
$(\mathfrak{N},\Phi^{<\alpha})\models\Gamma$ with $\alpha=\|\mathfrak{D}_0(\mathfrak{D}_\nu^m(\mathbf{0})\|$ for some $m\in\mathbb{N}$.
Proof:
Theorem 2.5 $\Rightarrow$ $(\mathfrak{N},\Phi^{<\alpha})\models\Gamma$ with $\alpha=\|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\|$ for some $m<\omega$.
$h^\infty \overset{\text{Th.3.4}}{\lhd} \mathbf{\Omega}_\nu{\cdot}2+n \overset{\text{L.3.1c,e}}{\lll} \mathbf{\Omega}_\nu{\cdot}3 \overset{\text{Def}}{\lll} \mathfrak{D}_\nu(\mathbf{1}) \overset{\text{L.3.1b,d}}{\lll} \mathfrak{D}_\nu\mathfrak{D}_\nu(\mathbf{0}) \overset{\text{Th.3.3d,e}}{\Longrightarrow}$
$\mathcal{D}_0(\mathcal{E}^m(h^\infty))\lhd\mathfrak{D}_0\mathfrak{D}_\nu^{m+2}(\mathbf{0}) \overset{\text{L.3.2}}{\Longrightarrow} \|\mathcal{D}_0(\mathcal{E}^m(h^\infty))\|\le\|\mathfrak{D}_0\mathfrak{D}_\nu^{m+2}(\mathbf{0})\|$.

Corollary
$|\mathrm{ID}_\nu|\le\sup_{m\in\mathbb{N}}\|\mathfrak{D}_0(\mathfrak{D}_\nu^m(\mathbf{0})\|$

Now we are going to prove that $\|\mathfrak{D}_0(\mathfrak{D}_\nu^m(\mathbf{0})\|$ equals $\psi_0\psi_\nu^m(0)$. By comparing the definition of $\mathfrak{D}_\sigma$ with the assignment of fundamental sequences in §1 and taking Theorem 1.5 and $(\diamondsuit)$ into consideration this should be more or less clear. In order to obtain a rigorous proof we introduce the canonical interpretation $\mathfrak{t}:\mathrm{T}\longrightarrow\mathfrak{T}_{<\omega}$ and show that this respects the fundamental sequences $(a[x])_{x\in|\mathrm{tp}(a)}$.

Definition of $\mathfrak{t}:\mathrm{T}\longrightarrow\mathfrak{T}_{<\omega}$
$\mathfrak{t}(D_{\sigma_0}a_0\oplus\ldots\oplus D_{\sigma_{n-1}}a_{n-1}):=\mathfrak{D}_{\sigma_0}\mathfrak{t}(a_1)+\ldots+\mathfrak{D}_{\sigma_{n-1}}\mathfrak{t}(a_{n-1})$

Theorem 3.6 For each $a\in\mathrm{T}$ we have
(i) $\mathrm{tp}(a)=1 \Rightarrow \mathfrak{t}(a)=\mathfrak{t}(a[0])+1$,
(ii) $\mathrm{tp}(a)=\omega \Rightarrow \mathfrak{t}(a)=(\mathfrak{t}(a[n]))_{n\in\mathbb{N}}$,
(iii) $\mathrm{tp}(a)=\Omega_{\mu+1} \Rightarrow \mathfrak{t}(a)=(\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\forall x\in|\Omega_{\mu+1}|(\mathfrak{t}(a[x])=\mathfrak{a}_{\mathfrak{t}(x)})$
Proof:
Let $\mathcal{FS}(a)$ abbreviate the claim (i)&(ii)&(iii). Then in a straightforward way one proves
(1) $\mathcal{FS}(a)$ & $\mathcal{FS}(b) \Longrightarrow \mathcal{FS}(b\oplus a)$, (2) $\mathcal{FS}(a) \Longrightarrow \mathcal{FS}(D_\sigma a)$,
from which one obtains $(\forall a\in\mathrm{T})\mathcal{FS}(a)$ by induction on the build up of a.

Theorem 3.7 $a\in\mathrm{OT}_0 \Longrightarrow \overline{\mathrm{o}}(a)=\|\mathfrak{t}(a)\|$
Proof by induction on $\mathfrak{t}(a)$:
By L.1.6 $\mathrm{tp}(a)\in\{0,1,\omega\}$. If $\mathrm{tp}(a)=0$ then $a=0$ and $\mathfrak{t}(a)=\mathbf{0}$. If $\mathrm{tp}(a)=\omega$ then $\mathfrak{t}(a)\overset{3.6}{=}(\mathfrak{t}(a[n]))_{n\in\mathbb{N}}$ and therefore $\|\mathfrak{t}(a)\|=\sup_{n\in\mathbb{N}}(\|\mathfrak{t}(a[n])\|+1)\overset{\text{IH}}{=}\sup_{n\in\mathbb{N}}(\overline{\mathrm{o}}(a[n])+1)\overset{\text{L.1.6}}{=}\overline{\mathrm{o}}(a)$
The case "$\mathrm{tp}(a)=1$" is treated in the same way.

Corollary
(a) $\|\mathfrak{D}_0\mathfrak{D}_\nu^m(\mathbf{0})\|=\psi_0\psi_\nu^m(0)$
(b) $|\mathrm{ID}_\nu|\le\sup_{m\in\mathbb{N}}\psi_0\psi_\nu^m(0)=\psi_0(\varepsilon_{\Omega_\nu+1})$

Proof: We have $|\mathrm{ID}_\nu|\le\sup_{m\in\mathbb{N}}\|\mathfrak{D}_0(\mathfrak{D}_\nu^m(\mathbf{0})\|$ and $\|\mathfrak{D}_0\mathfrak{D}_\nu^m(\mathbf{0})\|\overset{\text{Def}}{=}\|\mathfrak{t}(D_0D_\nu^m0)\|\overset{3.7}{=}\overline{\mathrm{o}}(D_0D_\nu^m0)\overset{\text{Def}}{\le}\mathrm{o}(D_0D_\nu^m0)=\psi_0\psi_\nu^m(0)$. This yields (b), and (a) with $\le$ in place of $=$. To get $=$ in (a) we have to use $(\diamondsuit)$ which implies $\overline{\mathrm{o}}(a)=\mathrm{o}(a)$ for all $a\in\mathrm{OT}$ (cf. §1, pg.4).

§4 Two Applications

Let $\widehat{\mathrm{T}} := \{a \in \mathrm{T} : a \text{ principal term}\}$ and $\widehat{\mathrm{OT}} := \mathrm{OT} \cap \widehat{\mathrm{T}}$.
As one easily sees, the set $\widehat{\mathrm{T}}$ can be inductively generated by

$a_0, ..., a_{n-1} \in \widehat{\mathrm{T}}\ (n \geq 0)\ \&\ \sigma < \omega \implies D_\sigma(a_0 \oplus ... \oplus a_{n-1}) \in \widehat{\mathrm{T}}$.

Hence $\widehat{\mathrm{T}}$ is nothing other than the set of all finite, ordered trees with labels $\sigma < \omega$, and each term $a = a_0 \oplus ... \oplus a_{n-1} \in \mathrm{T}$ can be considered as a tree with immediate subtrees $a_0, ..., a_{n-1} \in \widehat{\mathrm{T}}$ and an unlabeled root. The assignment of (fundamental) sequences $(a[x])_{x \in |\mathrm{tp}(a)|}$ can then be seen as the definition of a reduction procedure (or rewriting relation) $a \hookrightarrow_x a[x]$ on T. In [Bu87] this reduction procedure (restricted to $\mathrm{T}_0 := \{D_0 a_0 \oplus ... \oplus D_0 a_{n-1} : a_0, ..., a_{n-1} \in \mathrm{T}\}$) had been cooked up as a so-called hydra game, where in the i^{th} round of the game (or battle) the hydra a transforms itself into a new hydra $a[n_i]$. Using Theorem 3.6 and Theorem 3.5 one easily concludes that the hydra game terminates (i.e., $\forall a \in \mathrm{T}_0 \forall (n_i)_{i \in \mathbb{N}} \exists k(\, a[n_0][n_1] \ldots [n_k] = 0\,)$), and that this fact is not provable in $\mathrm{ID}_{<\omega}$:
Let W_0 be inductively defined by the rule: $a \in \mathrm{T}_0\ \&\ [\, a \neq 0 \Rightarrow \forall n(a[n] \in W_0)\,] \implies a \in W_0$.

Then "$a \in W_0$" says that each $\hookrightarrow$-reduction sequence starting with a terminates. Hence "$\forall a \in \mathrm{T}_0 (a \in W_0)$" expresses termination of the hydra game. Now using Theorem 3.6 by induction on $\mathfrak{t}(a)$ we get
$\forall a \in \mathrm{T}_0(\, a \in W_0\ \&\ |a|_{W_0} = \|\mathfrak{t}(a)\|\,)$.
The unprovability result is obtained as follows
$\mathrm{ID}_\nu \vdash \forall x (D_0 D_\nu^x 0 \in W_0) \stackrel{\text{Th.3.5}}{\Longrightarrow} \exists m \forall n (\, |D_0 D_\nu^n 0|_{W_0} < \|\mathfrak{D}_0 \mathfrak{D}_\nu^m(0)\|\,) \implies$
$\implies \exists m(\, |D_0 D_\nu^m 0|_{W_0} < \|\mathfrak{t}(D_0 D_\nu^m 0)\| = |D_0 D_\nu^m 0|_{W_0}\,)$. *Contradiction.*

Another interesting observation about the system $(\mathrm{OT}, \prec)$ is due to Okada [Ok88] and provides a rather short proof of H. Friedman's result that the extended Kruskal Theorem on finite labeled trees implies the wellfoundedness of $(\mathrm{OT}, \prec)$ (provably in ACA_0). This runs as follows.
First we define a binary relation $\sqsubseteq$ on $\widehat{\mathrm{T}}$ such that $a \sqsubseteq b$ is equivalent to "there exists a homeomorphic embedding $f : a \to b$ satisfying Friedman's *gap condition* (including the gap condition for the root)".

Definition of $a \sqsubseteq b$ for $a, b \in \widehat{\mathrm{T}}$
Let $a = D_\rho(a_0 \oplus ... \oplus a_{m-1})$ and $b = D_\sigma(b_0 \oplus ... \oplus b_{n-1})$.
$a \sqsubseteq b$ iff one of the following two clauses holds
(i) $\rho = \sigma$ and $\exists$ injective $q : \{0, ..., m-1\} \to \{0, ..., n-1\}$ such that $a_i \sqsubseteq b_{q(i)}$ for $i < m$,
(ii) $\rho \leq \sigma$ and $\exists j < n (\, a \sqsubseteq b_j\,)$.

Then we define a relation $\prec^* \subseteq \prec$ and prove $\forall a, b \in \widehat{\mathrm{OT}}(\, a \sqsubseteq b \Rightarrow a \preceq^* b\,)$.

Definition
$a \prec^* b :\Leftrightarrow a \prec b\ \&\ \forall \rho (G_\rho a \preceq G_\rho b)$ (with $X \preceq Y :\Leftrightarrow \forall x \in X \exists y \in Y (x \preceq y)$)

Lemma 4.1
(a) $a \prec^* b \implies D_\sigma a \prec^* D_\sigma b$
(b) $D_\rho a \preceq^* b\ \&\ \rho \leq \sigma\ \&\ G_\sigma b \prec b \implies D_\rho a \prec^* D_\sigma b$
Proof:
(a) $a \prec^* b\ \&\ G_\rho D_\sigma a \neq \emptyset \Rightarrow G_\rho D_\sigma a = \{a\} \cup G_\rho a \preceq \{b\} \cup G_\rho b = G_\rho D_\sigma b$.
(b) 1. $D_\rho a \preceq^* b\ \&\ \mu \leq \rho \leq \sigma \Rightarrow G_\mu D_\rho a \preceq G_\mu b \subseteq G_\mu D_\sigma b$. Hence $\forall \mu (G_\mu D_\rho a \preceq G_\mu D_\sigma b)$.
2. Proof of $D_\rho a \prec D_\sigma b$: Let $\rho = \sigma$ (otherwise the claim is trivial). Then $a \in G_\sigma(D_\rho a) \preceq G_\sigma b \prec b$.

Theorem 4.2
$a, b \in \widehat{\mathrm{OT}}\ \&\ a \sqsubseteq b \implies a \preceq b$

Proof: By induction on $\ell(b)$ we prove the stronger statement $a \preceq^* b$.
Let $a = D_\rho(a_0 \oplus ... \oplus a_{m-1})$ and $b = D_\sigma(b_0 \oplus ... \oplus b_{n-1})$.
(i) $\rho = \sigma$ & $\forall i < m(a_i \sqsubseteq b_{q(i)})$ & $\forall i,j < m(i \neq j \Rightarrow q(i) \neq q(j))$: By IH we have $a_i \preceq^* b_{q(i)}$ for $i < m$. From this we get $(a_0 \oplus ... \oplus a_{m-1}) \preceq^* (b_0 \oplus ... \oplus b_{n-1})$ and then by L.4.1a $a = D_\sigma(a_0 \oplus ... \oplus a_{m-1}) \preceq^* D_\sigma(b_0 \oplus ... \oplus b_{n-1}) = b$.
(ii) $\rho \leq \sigma$ and $\exists j < n(a \sqsubseteq b_j)$: By IH we have $a \preceq^* b_j \preceq^* (b_0 \oplus ... \oplus b_{n-1}) =: c$. Since $b = D_\sigma c \in \mathrm{OT}$, we also have $G_\sigma c \prec c$. By L.4.1b this yields $a = D_\rho(a_0 \oplus ... \oplus a_{m-1}) \prec^* D_\sigma c = b$.

References

[Acz] Aczel, P.: *A new approach to the Bachmann method for describing large countable ordinals.* (Unpublished)

[Ar96] Arai, T.: *Systems of ordinal diagrams.* Preprint (1996)

[Ar97] Arai, T.: *Proof Theory for Theories of Ordinals II: Σ_1-stability.* Preprint (1997)

[Br75] Bridge, J.: *A simplification of the Bachmann method for generating large countable ordinals.* JSL 40 (1975), pp. 171-185.

[Bu75] Buchholz, W.: *Normalfunktionen und konstruktive Systeme von Ordinalzahlen.* In: Proof Theory Symposion, Kiel 1974, Springer LNM 500, pp. 4-25 (1975)

[Bu81] Buchholz, W.: *The $\Omega_{\mu+1}$-rule* and *Ordinal analysis of* ID_ν. In: Buchholz, W., Feferman, S., Pohlers, W., Sieg, W., Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies. Springer LNM 897 (1981)

[Bu86] Buchholz, W.: *A new system of proof-theoretic ordinal functions.* APAL 32, pp. 195-207 (1986)

[Bu87] Buchholz, W.: *An independence result for (Π^1_1-CA)+BI.* APAL 33 (1987), pp. 131-155

[BS88] Buchholz, W. and Schütte, K.: *Proof Theory of Impredicative Subsystems of Analysis.* Studies in Proof Theory, Monographs 2. Napoli: Bibliopolis 1988

[Fe87] Feferman, S.: *Proof Theory: A Personal Report.* In: Takeuti, G.: Proof Theory (Second edition). Elsevier (1987), 447-485.

[Jä84] Jäger, G.: *ρ-inaccessible ordinals, collapsing functions and a recursive notation system.* Arch. math. Logik 24 (1984), pp. 49-62.

[Ok88] Okada, M.: *Note on H. Friedman's independence result on the extended Kruskal Theorem.* Discrete Mathematics (1988)

[Po87] Pohlers, W.: *Ordinal notations based on a hierarchy of inaccessible cardinals.* APAL 33 (1987), pp. 157-179.

[Ra90] Rathjen, M.: *Ordinal Notations Based on a Weakly Mahlo cardinal.* Arch. Math. Logic 29 (1990), pp. 249-263.

[Ra94] Rathjen, M.: *Proof Theory of Reflection.* APAL 68 (1994) pp. 181-224

[Ra95] Rathjen, M.: *Recent Advances in Ordinal Analysis: Π^1_2-CA and related systems.* The Bulletin of Symbolic Logic 1/4 (1995) pp. 468-485

[Se98] Setzer, A.: *Well-ordering proofs for Martin-Löf type theory.* APAL 92 (1998) pp. 113-159

Appendix (Proof of (1), (2) in the proof of Theorem 3.6)

(1) $\mathcal{FS}(a)$ & $\mathcal{FS}(b) \implies \mathcal{FS}(b \oplus a)$,
(2) $\mathcal{FS}(a) \implies \mathcal{FS}(D_\sigma a)$.
Proof:
(1) 0. $a = 0$: $b \oplus a = b$.

1. $\mathrm{tp}(a) = \omega$: Then $\mathrm{tp}(b \oplus a) = \omega$ and $\mathfrak{t}(b \oplus a) = \mathfrak{t}(b) + \mathfrak{t}(a) \overset{\mathcal{FS}(a)}{=} \mathfrak{t}(b) + (\mathfrak{t}(a[n]))_{n\in\mathbb{N}} = (\mathfrak{t}(b) + \mathfrak{t}(a[n]))_{n\in\mathbb{N}} = (\mathfrak{t}(b \oplus a[n]))_{n\in\mathbb{N}} = (\mathfrak{t}((b \oplus a)[n]))_{n\in\mathbb{N}}$.
2. $\mathrm{tp}(a) = \Omega_{\mu+1}$: By assumption $\mathfrak{t}(a) = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\forall x \in |\Omega_{\mu+1}|(\mathfrak{t}(a[x]) = \mathfrak{a}_{\mathfrak{t}(x)})$.
Hence $\mathfrak{t}(b \oplus a) = \mathfrak{t}(b) + (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu} = (\mathfrak{t}(b) + \mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\mathfrak{t}((b \oplus a)[x]) = \mathfrak{t}(b \oplus a[x]) = \mathfrak{t}(b) + \mathfrak{t}(a[x]) = \mathfrak{t}(b) + \mathfrak{a}_{\mathfrak{t}(x)}$.

(2) 0. $a = 0$ & $\sigma = 0$: $\mathrm{tp}(D_\sigma a) = 1$ & $(D_\sigma a)[0] = 0$ and hence $\mathfrak{t}(D_\sigma a) = \mathfrak{D}_0(\mathbf{0}) = \mathfrak{t}((D_\sigma a)[0]) + \mathbf{1}$.
1. $a = 0$ & $\sigma = \mu + 1$: $\mathrm{tp}(D_\sigma a) = \Omega_{\mu+1}$ and $\mathfrak{t}(D_\sigma a) = (\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\mathfrak{t}((D_\sigma a)[x]) = \mathfrak{t}(x)$.
2. $\mathrm{tp}(a) = 1$: $\mathrm{tp}(D_\sigma a) = \omega$ & $(D_\sigma a)[n] = (D_\sigma a[0]) \cdot (n{+}1)$.
By assumption $\mathfrak{t}(a) = \mathfrak{t}(a[0]) + \mathbf{1}$. Hence
$\mathfrak{t}(D_\sigma a) = \mathfrak{D}_\sigma \mathfrak{t}(a) = \big(\mathfrak{D}_\sigma(\mathfrak{t}(a[0])) \cdot (n{+}1)\big)_{n\in\mathbb{N}} = \big(\mathfrak{t}(D_\sigma a[0]) \cdot (n{+}1)\big)_{n\in\mathbb{N}} = \big(\mathfrak{t}((D_\sigma a)[n])\big)_{n\in\mathbb{N}}$.
3. $\mathrm{tp}(a) = \omega$: $\mathrm{tp}(D_\sigma a) = \omega$ & $(D_\sigma a)[n] = D_\sigma a[n]$.
By assumption $\mathfrak{t}(a) = \big(\mathfrak{t}(a[n])\big)_{n\in\mathbb{N}}$. Hence
$\mathfrak{t}(D_\sigma a) = \mathfrak{D}_\sigma \mathfrak{t}(a) = \big(\mathfrak{D}_\sigma(\mathfrak{t}(a[n]))\big)_{n\in\mathbb{N}} = \big(\mathfrak{t}(D_\sigma a[n])\big)_{n\in\mathbb{N}} = \big(\mathfrak{t}((D_\sigma a)[n])\big)_{n\in\mathbb{N}}$.
4. $\mathrm{tp}(a) = \Omega_{\mu+1}$ with $\mu < \sigma$: $\mathrm{tp}(D_\sigma a) = \Omega_{\mu+1}$ & $(D_\sigma a)[x] = D_\sigma a[x]$.
By assumption $\mathfrak{t}(a) = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\mathfrak{t}(a[x]) = \mathfrak{a}_{\mathfrak{t}(x)}$.
Hence $\mathfrak{t}(D_\sigma a) = \big(\mathfrak{D}_\sigma(\mathfrak{a}_\mathfrak{x})\big)_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\mathfrak{t}((D_\sigma a)[x]) = \mathfrak{t}(D_\sigma a[x]) = \mathfrak{D}_\sigma(\mathfrak{t}(a[x])) = \mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{t}(x)})$.
5. $\mathrm{tp}(a) = \Omega_{\mu+1}$ with $\mu \geq \sigma$: $\mathrm{tp}(D_\sigma a) = \omega$ & $(D_\sigma a)[n] = D_\sigma a[x_n]$ with $x_0 := \Omega_\mu$ and $x_{n+1} := D_\mu a[x_n]$.
By assumption $\mathfrak{t}(a) = (\mathfrak{a}_\mathfrak{x})_{\mathfrak{x}\in\mathfrak{T}_\mu}$ with $\forall x \in |\Omega_{\mu+1}|(\mathfrak{t}(a[x]) = \mathfrak{a}_{\mathfrak{t}(x)})$ (**).
Hence $\mathfrak{t}(D_\sigma a) = \mathfrak{D}_\sigma(\mathfrak{t}(a)) = \big(\mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_n})\big)_{n\in\mathbb{N}}$ with $\mathfrak{x}_0 := \mathbf{\Omega}_\mu$, $\mathfrak{x}_{n+1} := \mathfrak{D}_\mu(\mathfrak{a}_{\mathfrak{x}_n})$.
It remains to prove: $\mathfrak{t}((D_\sigma a)[n]) = \mathfrak{D}_\sigma(\mathfrak{a}_{\mathfrak{x}_n})$.
Since $\mathfrak{t}((D_\sigma a)[n]) = \mathfrak{D}_\sigma(\mathfrak{t}(a[x_n]))$, this amounts to $\mathfrak{t}(a[x_n]) = \mathfrak{a}_{\mathfrak{x}_n}$.
Due to (**) it remains to prove $\mathfrak{t}(x_n) = \mathfrak{x}_n$.

$\mathfrak{t}(x_0) = \mathfrak{t}(\Omega_\mu) = \mathbf{\Omega}_\mu = \mathfrak{x}_0$,
$\mathfrak{t}(x_n) = \mathfrak{x}_n \overset{(**)}{\Rightarrow} \mathfrak{t}(a[x_n]) = \mathfrak{a}_{\mathfrak{x}_n} \Rightarrow \mathfrak{t}(x_{n+1}) = \mathfrak{D}_\mu(\mathfrak{t}(a[x_n])) = \mathfrak{D}_\mu(\mathfrak{a}_{\mathfrak{x}_n}) = \mathfrak{x}_{n+1}$.