Annals of Pure and Applied Logic 60 (1993) 49–88
North-Holland

Proof-theoretic investigations on Kruskal's theorem

Michael Rathjen

Department of Mathematics, Ohio State University, Columbus. OH 43210, USA

Andreas Weiermann

Institut für mathematische Logik und Grundlagenforschung der Westfälischen Wilhelms-Universität Münster, Einsteinstraße 62, W-4400 Münster, Germany

Communicated by D. van Dalen
Received 6 March 1992

Abstract

Rathjen, M. and A. Weiermann, Proof-theoretic investigations on Kruskal's theorem, Annals of Pure and Applied Logic 60 (1993) 49–88.

In this paper we calibrate the exact proof-theoretic strength of Kruskal's theorem, thereby giving, in some sense, the most elementary proof of Kruskal's theorem.

Furthermore, these investigations give rise to ordinal analyses of restricted bar induction.

Introduction

S.G. Simpson in his article [10], "Nonprovability of certain combinatorial properties of finite trees", presents proof-theoretic results, due to H. Friedman, about embeddability properties of finite trees. It is shown there that Kruskal's theorem is not provable in ATR_0. An exact description of the proof-theoretic strength of Kruskal's theorem is not given. On the assumption that there is a bad infinite sequence of trees, the usual proof of Kruskal's theorem utilizes the existence of a minimal bad sequence of trees, thereby employing some form of Π_1^1 comprehension. So the question arises whether a more constructive proof can be given. The need for a more elementary proof of Kruskal's theorem is especially due to the fact that this theorem figures prominently in computer science, because it is the main tool for showing that sets of rewrite rules are terminating (see [3, p. 258], where this challenge is offered).

Correspondence to: A. Weiermann, Institut für mathematische Logik und Grundlagenforschung der Westfälischen Wilhelms-Universität Münster, Einsteinstraße 62, W-4400 Münster, Germany.

Our paper gives a complete proof-theoretic characterization of Kruskal's theorem in terms of ordinal notation systems, subsystems of second-order arithmetic, and subsystems of Kripke–Platek set theory.

The paper is divided into eleven parts.

In Section 1 we introduce an ordinal notation system T_A, which represents the Ackermann-ordinal (see [2] for a definition) in a natural way. It is shown that within ACA_0, Kruskal's theorem implies the well-foundedness of T_A.

The reversal of the latter implication constitutes the content of Section 2. Given a bad sequence of trees, we show how to produce effectively a strictly descending sequence of ordinals in T_A of the same length.

The equivalence of Kruskal's theorem with the well-foundedness of T_A then provides an upper bound for the order-types of simplification orderings, since Kruskal's theorem can be used to prove the well-foundedness of these orderings.

For $\alpha \leqslant \omega$ let $\mathbb{T}(\alpha)$ be the set of finite trees T such that every vertex of T has less than α immediate successors. $\mathbb{T}(\alpha)$ is quasi-ordered by the natural tree-embeddability relation. Let $KT(\alpha)$ be the statement "$\mathbb{T}(\alpha)$ is a well-quasi-ordered" (so $KT(\omega)$ is just Kruskal's theorem). We also show that $\forall n\, KT(n)$ and $KT(\omega)$ are equivalent over ACA_0.

In [10, p. 99], it is stated that Kruskal's theorem is provable in the formal system $T := ACA_0 + \Pi^1_2\text{-}BI$. The investigations of this paper were mainly prompted by this remark. It turns out that this is not quite true. Indeed, $ACA_0 + KT(\omega)$ proves the uniform Π^1_1 reflection principle of the latter theory, $RFN_{\Pi^1_1}(T)$, and is therefore a stronger theory. The remainder of this paper is devoted to proving

$$T := ACA_0 + \Pi^1_2\text{-}BI \nvdash \forall n\, KT(n) \quad \text{and} \quad ACA_0 \vdash KT(\omega) \leftrightarrow RFN_{\Pi^1_1}(T).$$

Sections 4–10 pursue the ordinal analysis of the system $ACA_0 + \Pi^1_2\text{-}BI$, thereby showing that the order type of T_A is the proof-theoretic ordinal of the latter system. The methods used here are perfectly general in that they provide a general framework for analyzing all the theories $ACA_0 + \Pi^1_n\text{-}BI$ for $n \geqslant 2$. Using results from [5], we also establish the proof-theoretic equivalence of $ACA_0 + \Pi^1_n$-BI and $KP\omega^- + \Pi_n$-*Foundation* for all $n \geqslant 2$, where $KP\omega^-$ stands for Kripke–Platek set theory including infinity but with foundation restricted to sets.

An effective version of the ordinal analysis of $ACA_0 + \Pi^1_2\text{-}BI$ is sketched in Section 11, finally establishing $ACA_0 \vdash KT(\omega) \leftrightarrow RFN_{\Pi^1_1}(T)$.

1. An ordinal notation system

Firstly, we need some ordinal-theoretic background. Let On be the class of ordinals. Let $AP := \{\xi \in On : \exists \eta \in On\, [\xi = \omega^\eta]\}$ be the class of additive principal numbers and let $E := \{\xi \in On : \xi = \omega^\xi\}$ be the class of ε-numbers which is enumerated by the function $\lambda\xi.\varepsilon_\xi$.

We write $\alpha =_{\mathrm{NF}} \omega^{\beta} + \delta$ if $\alpha = \omega^{\beta} + \delta$ and either, $\delta = 0$ and $\beta < \alpha$, or $\delta = \omega^{\delta_1} + \cdots + \omega^{\delta_k}$ with $\beta \geqslant \delta_1 \geqslant \cdots \geqslant \delta_k$ and $k \geqslant 1$.

Note that by Cantor's normal form theorem, for every $\alpha \notin E \cup \{0\}$, there are uniquely determined ordinals β and δ such that $\alpha =_{\mathrm{NF}} \omega^{\beta} + \delta$.

Let $\Omega := \aleph_1$. For any $\alpha < \varepsilon_{\Omega+1}$ we define the set $E_{\Omega}(\alpha)$ which consists of the ε-numbers below Ω which are needed for the unique representation of α in Cantor normal form recursively as follows:

1. $E_{\Omega}(0) := E_{\Omega}(\Omega) := \emptyset$,
2. $E_{\Omega}(\alpha) := \{\alpha\}$, if $\alpha \in E \cap \Omega$,
3. $E_{\Omega}(\alpha) := E_{\Omega}(\beta) \cup E_{\Omega}(\delta)$ if $\alpha =_{\mathrm{NF}} \omega^{\beta} + \delta$.

Let $\alpha^* := \max(E_{\Omega}(\alpha) \cup \{0\})$.

We define sets of ordinals $C(\alpha, \beta)$, $C_n(\alpha, \beta)$, and ordinals $\vartheta\alpha$ by main recursion on $\alpha < \varepsilon_{\Omega+1}$ and subsidiary recursion on $n < \omega$ (for $\beta < \Omega$) as follows.

(C1) $\{0, \Omega\} \cup \beta \subseteq C_n(\alpha, \beta)$,

(C2) $\gamma, \delta \in C_n(\alpha, \beta) \wedge \xi =_{\mathrm{NF}} \omega^{\gamma} + \delta \Rightarrow \xi \in C_{n+1}(\alpha, \beta)$,

(C3) $\delta \in C_n(\alpha, \beta) \cap \alpha \Rightarrow \vartheta\delta \in C_{n+1}(\alpha, \beta)$,

(C4) $C(\alpha, \beta) := \bigcup \{C_n(\alpha, \beta) : n < \omega\}$,

(C5) $\vartheta\alpha := \min\{\xi < \Omega : C(\alpha, \xi) \cap \Omega \subseteq \xi \wedge \alpha \in C(\alpha, \xi)\}$,

Lemma 1.1. *$\vartheta\alpha$ is defined for every $\alpha < \varepsilon_{\Omega+1}$.*

Proof. Let $\beta_0 := \alpha^* + 1$. Then $\alpha \in C(\alpha, \beta_0)$ via (C1) and (C2). Since the cardinality of $C(\alpha, \beta)$ is less than Ω there exists a $\beta_1 < \Omega$ such that $C(\alpha, \beta_0) \cap \Omega \subset \beta_1$. Similarly there exists for each $\beta_n < \Omega$ (which is constructed recursively) a $\beta_{n+1} < \Omega$ such that $C(\alpha, \beta_n) \cap \Omega \subseteq \beta_{n+1}$. Let $\beta := \sup\{\beta_n : n < \omega\}$. Then $\alpha \in C(\alpha, \beta)$ and $C(\alpha, \beta) \cap \Omega \subset \beta < \Omega$. Therefore $\vartheta\alpha \leqslant \beta < \Omega$. □

Lemma 1.2. (1) *$\vartheta\alpha \in E$,*

(2) *$\alpha \in C(\alpha, \vartheta\alpha)$,*

(3) *$\vartheta\alpha = C(\alpha, \vartheta\alpha) \cap \Omega$, and $\vartheta\alpha \notin C(\alpha, \vartheta\alpha)$,*

(4) *$\gamma \in C(\alpha, \beta) \Leftrightarrow E_{\Omega}(\gamma) \subseteq C(\alpha, \beta)$,*

(5) *$\alpha^* < \vartheta\alpha$,*

(6) *$\vartheta\alpha = \vartheta\beta \Rightarrow \alpha = \beta$,*

(7) *$\vartheta\alpha < \vartheta\beta \Leftrightarrow (\alpha < \beta \wedge \alpha^* < \vartheta\beta) \vee (\beta < \alpha \wedge \vartheta\alpha \leqslant \beta^*)$,*

(8) *$\beta < \vartheta\alpha \Leftrightarrow \omega^{\beta} < \vartheta\alpha$.*

Proof. (1) and (8) issue from closure $\vartheta\alpha$ under (C2).

(2) and (3) follow from Lemma 1.1 and the definition of $\vartheta\alpha$.

(4) If $E_{\Omega}(\gamma) \subseteq C(\alpha, \beta)$, then $\gamma \in C(\alpha, \beta)$ by (C2). On the other hand, $\gamma \in C_n(\alpha, \beta) \Rightarrow E_{\Omega}(\gamma) \subseteq C_n(\alpha, \beta)$ is easily seen by induction on n.

(5) $\alpha^* \in C(\alpha, \vartheta\alpha)$ holds by (4). As $\alpha^* < \Omega$, this implies $\alpha^* < \vartheta\alpha$ by (3).

(6) Suppose, aiming at a contradiction, that $\vartheta\alpha = \vartheta\beta$ and $\alpha < \beta$. Then

$C(\alpha, \vartheta\alpha) \subseteq C(\beta, \vartheta\beta)$; hence $\alpha \in C(\beta, \vartheta\beta) \cap \beta$ by (2); thence $\vartheta\alpha = \vartheta\beta \in C(\alpha, \vartheta\beta)$, contradicting (3).

(7) Suppose $\alpha < \beta$. Then $\vartheta\alpha < \vartheta\beta$ implies $\alpha^* < \vartheta\beta$ by (5). If $\alpha^* < \vartheta\beta$, then $\alpha \in C(\beta, \vartheta\beta)$; hence $\vartheta\alpha \in C(\beta, \vartheta\beta)$; thus $\vartheta\alpha < \vartheta\beta$. This shows

$$\text{(a)}\ \alpha < \beta \Rightarrow (\vartheta\alpha < \vartheta\beta \Leftrightarrow \alpha^* < \vartheta\beta).$$

By interchanging the roles of α and β, and employing (5), one obtains

$$\text{(b)}\ \beta < \alpha \Rightarrow (\vartheta\alpha < \vartheta\beta \Leftrightarrow \vartheta\alpha \leq \beta^*).$$

(a) and (b) yield (7). □

The Ackermann ordinal is denoted in this context by $\vartheta\Omega^\omega$.

Definition 1.1. Inductive definition of a set $OT(\vartheta)$ of ordinals and a natural number $G_\vartheta\alpha$ for $\alpha \in OT(\vartheta)$.

1. $0, \Omega \in OT(\vartheta), \qquad G_\vartheta 0 := G_\vartheta\Omega := 0$,
2. $\alpha =_{\mathrm{NF}} \omega^\beta + \delta \wedge \beta, \delta \in OT(\vartheta) \Rightarrow \alpha \in OT(\vartheta),\ G_\vartheta\alpha := \max\{G_\vartheta\beta, G_\vartheta\delta\} + 1$,
3. $\alpha = \vartheta\alpha_1 \wedge \alpha_1 \in OT(\vartheta) \Rightarrow \vartheta\alpha_1 \in OT(\vartheta),\ G_\vartheta\alpha := G_\vartheta\alpha_1 + 1$.

Observe that according to Lemma 1.2(1) and 1.2(6), the function G_ϑ is well-defined. Each ordinal $\alpha \in OT(\vartheta)$ has a unique normal form using the symbols $0, \Omega, +, \omega, \vartheta$. Furthermore, if for $\alpha, \beta \in OT(\vartheta)$, represented in their normal form, we were to decide $\alpha < \beta$, we could do this by deciding $\alpha_0 < \beta_0$ for ordinals α_0 and β_0 that appear in these representations and, in addition, satisfy $G_\vartheta\alpha_0 + G_\vartheta\beta_0 < G_\vartheta\alpha + G_\vartheta\beta$. This follows from Lemma 1.2(7) and the recursive procedure for comparing ordinals in Cantor normal form. So we come to see the following fact.

Lemma 1.3. *After a straightforward coding in the natural numbers, we may consider $\langle OT(\vartheta), <\restriction OT(\vartheta)\rangle$ as a primitive recursive ordinal notation system.*

Lemma 1.4. (1) $OT(\vartheta) = \bigcup\{C(\alpha, 0): \alpha < \varepsilon_{\Omega+1}\}$,
(2) $OT(\vartheta) \cap \alpha = \alpha$ *for* $\alpha \in OT(\vartheta) \cap \Omega$.

From now on, we presume an effective coding of $\langle OT(\vartheta), <\restriction OT(\vartheta)\rangle$ in the natural numbers, so that the latter structure can be dealt with in ACA_0 (actually in primitive recursive arithmetic). Of course, the well-foundedness of $\langle OT(\vartheta), <\restriction OT(\vartheta)\rangle$ is not provable in ACA_0.

Next, we recall some basic definitions from [10]. A *finite tree* $\mathscr{T}$ is a finite partial order $\langle T, \leq_T\rangle$ such that,

1. $(\exists r \in T)(\forall t \in T)[t \neq r \rightarrow r \leq_T t \wedge t \nleq_T r]$,
2. $(\forall s \in T)(\forall t \in T)(\forall u \in T)[t \leq_T s \wedge u \leq_T s \rightarrow t \leq_T u \vee u \leq_T t]$.

For a finite tree $\mathcal{T} = \langle T, \leqslant_T \rangle$ and $t, u \in T$ we denote the $\leqslant_T$-infimum of t and u by $t \wedge_T u$. The uniquely determined $\leqslant_T$-minimal element is called the *root* of $\mathcal{T}$.

A finite tree $\mathcal{T} = \langle T, \leqslant_T \rangle$ is *embeddable* into a finite tree $\mathcal{U} = \langle U, \leqslant_U \rangle$ if there exists a one-to-one (embedding-) function $f: T \to U$ such that $f(t \wedge_T u) = f(t) \wedge_U f(u)$ for every $t, u \in T$.

For $a \in T$ let $T^a := \{b \in T: a \leqslant_T b\}$ and $\mathcal{T}^a = \langle T^a, \leqslant_T \restriction T^a \rangle$. An *immediate subtree* $\mathcal{U}$ of $\mathcal{T}$ has the form $\mathcal{T}^a$ where a is an immediate $\leqslant_T$-successor of the root of $\mathcal{T}$. Every immediate subtree $\mathcal{U}$ of $\mathcal{T}$ is embeddable into $\mathcal{T}$.

Theorem 1.1 (Kruskal). *For any ω-sequence $\langle \mathcal{T}_i : i < \omega \rangle$ of finite trees there exist indices i and j such that $i < j < \omega$ and $\mathcal{T}_i$ is embeddable into $\mathcal{T}_j$.*

In [10] it is shown that Kruskal's theorem implies the well-foundedness of Γ_0 within ACA_0. This proof can be extended to also yield the well-foundednes of $\vartheta\Omega^\omega$ from Kruskal's theorem in ACA_0. The proof utilizes a normal form for ordinals $< \vartheta\Omega^\omega$. This normal form can be computed primitive recursively.

We require some notation. Let $\Omega \cdot 0 := 0$ and $\Omega \cdot (n+1) := \Omega \cdot n + \Omega$. Let $\Omega^n \cdot 0 := 0$. If $\beta = \omega^{\beta_1} + \cdots + \omega^{\beta_k}$ and $\beta_1 \geqslant \cdots \geqslant \beta_k$ we set

$$\Omega^n \cdot \beta := \omega^{\Omega \cdot n + \beta_1} + \cdots + \omega^{\Omega \cdot n + \beta_k}.$$

Proposition 1.1. *Let $\alpha \in E \cap \vartheta\Omega^\omega$. Then there exists a unique $n < \omega$ and unique ordinals $\alpha_0, \ldots, \alpha_n < \alpha$ such that $\alpha = \vartheta(\Omega^n \cdot \alpha_n + \cdots + \Omega^0 \cdot \alpha_0)$, and $\alpha_n \neq 0$ if $n \neq 0$.*

Definition 1.2. For any $\alpha \leqslant \omega$ let $\mathbb{T}(\alpha)$ be the set of trees $\mathcal{T}$ such that any vertex in $\mathcal{T}$ has less than α immediate successors. We use $KT(\alpha)$ to abbreviate that $\mathbb{T}(\alpha)$ is well-quasi-ordered.

For $\alpha \in OT(\vartheta)$ let $WF(\alpha)$ stand for "$<$ *restricted to* $\{\beta \in OT(\vartheta): \beta < \alpha\}$ *is well-founded*".

Theorem 1.2. $ACA_0 \vdash (\forall n < \omega)\, KT(n) \to WF(\vartheta\Omega^\omega)$.

Proof. We reason within ACA_0. Let *Tree* be the set of all finite trees. We shall define a primitive recursive mapping

$$o: Tree \to \{\alpha \in OT(\vartheta): \alpha < \vartheta\Omega^\omega\}$$

by recursion on the number of elements of a tree $\mathcal{T}$, $|\mathcal{T}|$. If $\mathcal{T}$ consists only of its root, then let $o(\mathcal{T}) = 0$. Otherwise the root of $\mathcal{T}$ has finitely many immediate successors $a_0, \ldots, a_k$. Let $\mathcal{T}_i$ be the immediate subtrees of $\mathcal{T}$ determined by a_i. Since $|\mathcal{T}_i| < |\mathcal{T}|$, we may assume that $\alpha_i := o(\mathcal{T}_i)$ is already defined. We may further assume that $\alpha_0 \geqslant \cdots \geqslant \alpha_k$. We set $o(\mathcal{T}) = \alpha_0$ if $k = 0$, $o(\mathcal{T}) = \alpha_1 \oplus \alpha_0$ if $k = 1$, where $\oplus$ denotes the commutative natural sum of ordinals, $o(\mathcal{T}) = \omega^{\alpha_0}$ if $k = 2$ and $\omega^{\alpha_0} > \alpha_0$, $o(\mathcal{T}) = \omega^{\alpha_0 + 1}$ if $k = 2$ and $\omega^{\alpha_0} = \alpha_0$ and $o(\mathcal{T}) = \vartheta\alpha_0$ if $k = 3$. To deal with the case $k \geqslant 4$ we introduce some auxiliary notation. For finite sequences of ordinals we define

$$\begin{aligned}&\langle\beta_0,\dots,\beta_m\rangle <_l \langle\gamma_0,\dots,\gamma_n\rangle \Leftrightarrow\\ &\quad\max\{\beta_0,\dots,\beta_m\}\leqslant\max\{\gamma_0,\dots,\gamma_n\}\\ &\quad\wedge[[2\leqslant m<n]\vee[2\leqslant m=n\wedge(\exists j<m)[\beta_j<\gamma_j\wedge(\forall r<j)\,\beta_r=\gamma_r]].\end{aligned}$$

For a finite sequence $\langle\beta_0,\dots,\beta_m\rangle$ $(m\geqslant 1)$ of countable ordinals let $\vartheta(\langle\beta_0,\dots,\beta_m\rangle):=\vartheta(\Omega^m\cdot(1+\beta_0)+\Omega^{m-1}\cdot\beta_1+\cdots+\Omega^0\cdot\beta_m)$. Then by Lemma 1.5 below $\langle\beta_0,\dots,\beta_m\rangle <_l \langle\gamma_0,\dots,\gamma_n\rangle$ implies $\vartheta(\langle\beta_0,\dots,\beta_m\rangle)<\vartheta(\langle\gamma_0,\dots,\gamma_n\rangle)$.

Now we can proceed in the definition of $o(\mathcal{T})$ for $k\geqslant 4$. Let $S:=\{\langle\alpha_{\pi(0)},\dots,\alpha_{\pi(i)}\rangle\colon\ 1\leqslant i\leqslant k,\ \pi$ a permutation of $\{0,\dots,i\}\}$. Let $\langle\alpha_0^*,\dots,\alpha_{i_0}^*\rangle$ be the $(k-3)$th element of S with respect to $<_l$ and set $o(\mathcal{T})=\vartheta(\langle\alpha_0^*,\dots,\alpha_{i_0}^*\rangle)$.

Note that this assignment of ordinals to trees is weakly increasing, i.e. $o(\mathcal{T}_a)\leqslant o(\mathcal{T})$ if $\mathcal{T}_a$ is an immediate subtree of $\mathcal{T}$. This is due to the fact, that for $\beta_0,\dots,\beta_n<\Omega$ we have $\beta_0,\dots,\beta_n<\vartheta(\Omega^n\cdot\beta_n+\cdots+\Omega^0\cdot\beta_0)$.

Given an embedding of trees $f:\mathcal{T}_1\to\mathcal{T}_2$, we claim that $o(\mathcal{T}_1^a)\leqslant o(\mathcal{T}_2^{f(a)})$ for each $a\in\mathcal{T}_1$. The proof is by induction on $|\mathcal{T}_1^a|$; it just springs from the above observation.

By induction on $G_\vartheta\alpha$ one easily verifies that there is a tree $\mathcal{T}$ such that $o(\mathcal{T})=\alpha$ provided that $\alpha<\vartheta\Omega^\omega$. For $\alpha\in E$ one has to employ Proposition 1.1.

If now $\langle\alpha_k\colon k<\omega\rangle$ were an infinitely descending sequence of ordinals below $\vartheta\Omega^\omega$, then we would get a corresponding sequence of finite trees $\langle\mathcal{T}_k\colon k<\omega\rangle$ with $o(\mathcal{T}_k)=\alpha_k$ for each $k<\omega$. Since $\sup\{\vartheta(\Omega^n+\cdots+\Omega^0)\colon n<\omega\}=\vartheta\Omega^\omega$, there would be an n_0 such that for all $k<\omega$, $\mathcal{T}_k\in\mathbb{T}(n_0)$. Therefore, by $KT(n_0)$, we could pick $i<j<\omega$ such that $\mathcal{T}_i\leqslant\mathcal{T}_j$. Hence $o(\mathcal{T}_i)=\alpha_i\leqslant o(\mathcal{T}_j)=\alpha_j$ by the above claim. Contradiction. □

In the next section we will frequently draw on the following result.

Lemma 1.5. *Suppose* $\alpha_0,\dots,\alpha_m,\beta_0,\dots,\beta_n\in OT(\vartheta)\cap\Omega$. *Assume* $\Omega^n\cdot\beta_n+\cdots+\Omega^0\cdot\beta_0<\Omega^m\cdot\alpha_m+\cdots+\Omega^0\cdot\alpha_0$ *and* $\max(\beta_0^*,\dots,\beta_n^*)\leqslant\max(\alpha_0^*,\dots,\alpha_m^*)$. *Then*

$$\vartheta(\Omega^n\cdot\beta_n+\cdots+\Omega^0\cdot\beta_0)<\vartheta(\Omega^m\cdot\alpha_m+\cdots+\Omega^0\cdot\alpha_0).$$

Proof. This follows from Lemma 1.2(7). □

2. An elementary proof of Kruskal's theorem

From now on we argue in ACA_0 (in fact, RCA_0 would suffice).

Let $T_A:=\{\alpha\in OT(\vartheta)\colon\alpha<\vartheta\Omega^\omega\}$ and $<_{T_A}:=(<\restriction OT(\vartheta))\restriction T_A$. Then $\langle T_A,$

$<_{T_A}\rangle$ is a primitive recursive ordinal notation system for the Ackermann ordinal. We introduce some more notation. A *quasi-order* $\mathbb{X}$ is an ordered pair $\langle X, \leqslant\rangle$ where X is a countable set and $\leqslant$ is a binary, reflexive and transitive relation on X. A (finite or countably infinite) sequence $\boldsymbol{x} = \langle x_0, x_1, \ldots\rangle$ of elements in X is called *bad* if there are no indices i and j such that $i < j < \text{length}(\boldsymbol{x})$ and $x_i \leqslant x_j$. In particular, the empty sequence is bad.

$\mathbb{X}$ is a *well-quasi-order* if all bad sequences in X are finite. Let $Bad(\mathbb{X})$ be the set of all finite bad sequences in X. A *reification* of $\mathbb{X}$ into $\alpha \in OT(\vartheta)$ is a mapping $f: Bad(\mathbb{X}) \to \alpha + 1$ such that $f(\boldsymbol{x}^\frown y) < f(\boldsymbol{x})$ for every $\boldsymbol{x} \in Bad(\mathbb{X})$ and every $y \in X$ such that $\boldsymbol{x}^\frown y \in Bad(\mathbb{X})$.

Lemma 2.1. *The following is provable in ACA_0: If f is a reification of $\mathbb{X}$ into α and α is well-founded, then $\mathbb{X}$ is a well-quasi-order.*

Proof. An infinite bad sequence in $\mathbb{X}$ would give rise to an infinite descending chain below $\alpha + 1$. $\square$

We introduce some more terminology. Let $\mathbb{X} = \langle X, \leqslant\rangle$ be a quasi-order. For $\boldsymbol{x} \in Bad(\mathbb{X})$ let $X_{\boldsymbol{x}} := \{y \in X : \boldsymbol{x}^\frown y \in Bad(\mathbb{X})\}$ and $\mathbb{X}_{\boldsymbol{x}} := \langle X_{\boldsymbol{x}}, \leqslant \upharpoonright X_{\boldsymbol{x}}\rangle$. Let $\mathbb{X}_0 = \langle X_0, \leqslant_0\rangle$ and $\mathbb{X}_1 = \langle X_1, \leqslant_1\rangle$ be quasi-orders.

We define quasi-orders $\mathbb{X}_0 \oplus \mathbb{X}_1$ and $\mathbb{X}_0 \otimes \mathbb{X}_1$ as follows. The domain of $\mathbb{X}_0 \oplus \mathbb{X}_1$ is the disjoint union $X_0 \uplus X_1$ of X_0 and X_1. Therefore $X_0 \uplus X_1$ consists of ordered pairs $\langle 0, x_0\rangle$ and $\langle 1, x_1\rangle$ where $x_0 \in X_0$ and $x_1 \in X_1$. $X_0 \uplus X_1$ is quasi-ordered by a relation $\leqslant_0 \oplus \leqslant_1$ as follows:

$$\langle i, u\rangle \leqslant_0 \oplus \leqslant_1 \langle j, v\rangle \;:\Leftrightarrow\; i = j \wedge u \leqslant_i v.$$

In writing $X_0 \uplus X_1$ we assume without loss of generality that X_0 and X_1 are disjoint and we will identify $x_0 \in X_0$ with $\langle 0, x_0\rangle$ and $x_1 \in X_1$ with $\langle 1, x_1\rangle$.

The domain of $\mathbb{X}_0 \otimes \mathbb{X}_1$ is the cartesian product $X_0 \times X_1$ of X_0 and X_1. $X_0 \times X_1$ is quasi-ordered by the relation $\leqslant_0 \otimes \leqslant_1$ as follows:

$$\langle x_0, x_1\rangle \leqslant_0 \otimes \leqslant_1 \langle x_0', x_1'\rangle \;:\Leftrightarrow\; x_0 \leqslant_0 x_0' \wedge x_1 \leqslant_1 x_1'.$$

For a given quasi-order $\mathbb{X} = \langle X, \leqslant\rangle$ we define the quasi-order $\mathbb{X}^{<\omega}$ as follows: The domain of $\mathbb{X}^{<\omega}$ consists of the set $X^{<\omega}$ of all finite sequences in X. $X^{<\omega}$ is quasi-ordered by a relation $\leqslant^{<\omega}$ as follows:

$$\langle x_0, \ldots, x_m\rangle \leqslant^{<\omega} \langle x_0', \ldots, x_n'\rangle$$

if and only if there exists a sub-sequence $i_0 < \cdots < i_m \leqslant n$ such that $x_l \leqslant x_{i_l}'$ for every $l \leqslant m$. (To be precise, the empty sequence is the bottom element with respect to the ordering $\leqslant^{<\omega}$.)

If $\mathbb{X}_0$ and $\mathbb{X}_1$ are well-quasi-orders then $\mathbb{X}_0 \oplus \mathbb{X}_1$, $\mathbb{X}_0 \otimes \mathbb{X}_1$ and $\mathbb{X}_0^{<\omega}$ are well-quasi-orders, too. According to [8], this can be shown in ACA_0 (especially, Higman's lemma is provable in ACA_0).

A *finite tree* $\mathcal{T}$ *with labels in a quasi-order* $\mathbb{X} = \langle X, \leqslant \rangle$ is an ordered triple $\langle T, \leqslant_T, l_T \rangle$ such that $\langle T, \leqslant_T \rangle$ is a finite tree and $l_T : T \to X$. If $\mathcal{T} = \langle T, \leqslant_T, l_T \rangle$ and $\mathcal{U} = \langle U, \leqslant_U, l_U \rangle$ are finite trees with labels in a quasi-order $\mathbb{X} = \langle X, \leqslant_X \rangle$ we say that $\mathcal{T}$ is *embeddable* into $\mathcal{U}$ if there exists an embedding-function $f : T \to U$ such that $l_T(t) \leqslant_X l_U(f(t))$ for every $t \in T$. Let $T(\mathbb{X})$ be the set of finite trees with labels in $\mathbb{X}$ and let $\leqslant_{T(\mathbb{X})}$ be the corresponding embeddability relation. Let $\mathbb{T}(\mathbb{X}) := \langle T(\mathbb{X}), \leqslant_{T(\mathbb{X})} \rangle$.

Theorem 2.1 (Kruskal). *$\mathbb{T}(\mathbb{X})$ is a well-quasi-order for every well-quasi-order $\mathbb{X}$.*

We want to show that the existence of a reification of $\mathbb{X}$ into $\alpha \in T$ implies the existence of a reification of $\mathbb{T}(\mathbb{X})$ into $\vartheta(\Omega^\omega \cdot \alpha) \in T$. This will imply Kruskal's theorem by taking the domain of $\mathbb{X}$ to be a singleton, i.e., $\alpha = 1$. For technical reasons we introduce the following terminology, which is due to Schmidt [6].

Definition 2.1. Let $\mathbb{X}_i = \langle X_i, \leqslant_i \rangle$ $(i = 0, \ldots, n)$ be pairwise disjoint quasi-orders and let $\alpha_0, \ldots, \alpha_n$ ordinals such that $0 < \alpha_0 < \cdots < \alpha_n \leqslant \omega$.

Let $\mathbb{X} := \mathbb{X}_0 \oplus \cdots \oplus \mathbb{X}_n$. Let

$$T\begin{pmatrix} \mathbb{X}_0 & \cdots & \mathbb{X}_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

be the set of all finite trees $\mathcal{T} = \langle T, \leqslant_T, l_T \rangle$ in $T(\mathbb{X})$ such that for every vertex $t \in T$, if the label $l_T(t)$ of t is in X_i, then t has strictly fewer than α_i immediate successors in $\mathcal{T}$. Let $\leqslant_{T(X,\alpha)}$ be the restriction of the tree-embeddability-relation to

$$T\begin{pmatrix} \mathbb{X}_0 & \cdots & \mathbb{X}_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

and let

$$\mathbb{T}\begin{pmatrix} \mathbb{X}_0 & \cdots & \mathbb{X}_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix} := \left\langle T\begin{pmatrix} \mathbb{X}_0 & \cdots & \mathbb{X}_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}, \leqslant_{T(X,\alpha)} \right\rangle.$$

For ordinals α, β we denote by $\alpha \oplus \beta$ the (commutative) natural sum of α and β and by $\alpha \otimes \beta$ the (commutative) natural product of α and β. These operations are defined as follows: Let $\alpha \oplus 0 := 0 \oplus \alpha := \alpha$ and $\alpha \otimes 0 := 0 \otimes \alpha := 0$.

Now suppose $\alpha = \omega^{\delta_0} + \cdots + \omega^{\delta_n} \geqslant \delta_0 \geqslant \cdots \geqslant \delta_n$, $n \geqslant 0$ and $\beta = \omega^{\delta_{n+1}} + \cdots + \omega^{\delta_m} \geqslant \delta_{n+1} \geqslant \cdots \geqslant \delta_m$, $m \geqslant n+1$. Then

$$\alpha \oplus \beta := \omega^{\delta_{\pi(0)}} + \cdots + \omega^{\delta_{\pi(m)}},$$

where π is a permutation of $\{0, \ldots, m\}$ such that $\delta_{\pi(i)} \geqslant \delta_{\pi(j)}$ if $i < j \leqslant m$. We define $\alpha \otimes \beta$ to be

$$(\omega^{\delta_0 \oplus \delta_{n+1}} \oplus \cdots \oplus \omega^{\delta_0 \oplus \delta_m}) \oplus \cdots \oplus (\omega^{\delta_n \oplus \delta_{n-1}} \oplus \cdots \oplus \omega^{\delta_n \oplus \delta_m}).$$

Every ε-number is closed under these operations.

Theorem 2.2. *The following is provable in ACA_0: Let $\mathbb{X}_i$ $(i = 0, \ldots, n)$ be pairwise disjoint quasi-orders, let f_i be reifications of $\mathbb{X}_i$ into β_i $(i = 0, \ldots, n$; $\beta_i \in OT(\vartheta) \cap \Omega)$ and let $0 < \alpha_0 < \cdots < \alpha_n \leq \omega$. Then there exists a reification of*

$$\mathbb{T}\begin{pmatrix} \mathbb{X}_0 & \cdots & \mathbb{X}_n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

into $\vartheta(\Omega^{\alpha_n} \cdot \beta_n + \cdots + \Omega^{\alpha_0} \cdot \beta_0)$.

The proof requires some further preparation. Let $\mathfrak{X}^0, \ldots, \mathfrak{X}^n$ be names, that means appropriate Gödel-numbers, for pairwise disjoint quasi-orders $\mathbb{X}^0, \ldots, \mathbb{X}^n$ and let $0 < \alpha_0 < \cdots < \alpha_n \leq \omega$. Let

$$Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

be the least class (of names for quasi-orders) which is closed under the following rules.

1. For every $i \leq n$:

$$\mathfrak{X}^i \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

is a name for $\mathbb{X}^i$.

2. For every $i \leq n$ and every $\boldsymbol{x}_i \in Bad(\mathbb{X}^i)$:

$$\mathfrak{X}^i_{\boldsymbol{x}_i} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

is a name for $\mathbb{X}^i_{\boldsymbol{x}_i}$.

3. If

$$\mathfrak{Y}^0, \ldots, \mathfrak{Y}^r \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

are names for $\mathbb{Y}^0, \ldots, \mathbb{Y}^r$ then

$$\bigoplus\{\mathfrak{Y}^l : l \leq r\}, \bigotimes\{\mathfrak{Y}^l : l \leq r\} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

are names for $\bigoplus\{\mathbb{Y}^l : l \leq r\}$ and $\bigotimes\{\mathbb{Y}^l : l \leq r\}$.

4. If

$$\mathfrak{X} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

is a name for $\mathbb{X}$ then

$$\mathfrak{X}^{<\omega} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

is a name for $\mathbb{X}^{<\omega}$.

5. If

$$\mathfrak{Y}^0, \ldots, \mathfrak{Y}^m \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

are names for pairwise disjoint quasi-orders $\mathbb{Y}^0, \ldots, \mathbb{Y}^m$ and $0 < \gamma_0 < \cdots < \gamma_m \leqslant \omega$ and $\Omega^{\gamma_m} + \cdots + \Omega^{\gamma_0} \leqslant \Omega^{\alpha_n} + \cdots + \Omega^{\alpha_0}$ then

$$\mathfrak{T}\begin{pmatrix} \mathfrak{Y}^0 & \cdots & \mathfrak{Y}^m \\ \gamma_0 & \cdots & \gamma_n \end{pmatrix} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

is a name for

$$\mathbb{T}\begin{pmatrix} \mathbb{Y}^0 & \cdots & \mathbb{Y}^m \\ \gamma_0 & \cdots & \gamma_m \end{pmatrix}.$$

Definition 2.2. Let $\mathfrak{X}^0, \ldots, \mathfrak{X}^n$ be names for pairwise disjoint quasi-orders $\mathbb{X}^0, \ldots, \mathbb{X}^n$ and let $0 < \alpha_0 < \cdots < \alpha_n \leqslant \omega$. Let f_i be reifications of X^i into β_i. For each

$$\mathfrak{X} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

let $o(\mathfrak{X})$ be defined recursively as follows:

1. $o(\mathfrak{X}^i) := \vartheta(f_i(\langle\,\rangle))$,
2. $o(\mathfrak{X}^i_{\boldsymbol{x}_i}) := \vartheta(f_i(\boldsymbol{x}_i))$,
3. $o(\bigoplus\{\mathfrak{Y}^l : l \leqslant r\}) := \bigoplus\{o(\mathfrak{Y}^l) : l \leqslant r\}$,
   $o(\bigotimes\{\mathfrak{Y}^l : l \leqslant r\}) := \bigotimes\{o(\mathfrak{Y}^l) : l \leqslant r\}$,
4. $o(\mathfrak{X}^{<\omega}) := \vartheta(o(\mathfrak{X}) + 1)$,
5. $o\left(\mathfrak{T}\begin{pmatrix} \mathfrak{Y}^0 & \cdots & \mathfrak{Y}^m \\ \gamma_0 & \cdots & \gamma_m \end{pmatrix}\right) := \vartheta(\Omega^{\gamma_m} \cdot o(\mathfrak{Y}^m) + \cdots + \Omega^{\gamma_0} \cdot o(\mathfrak{Y}^0))$.

Definition 2.3. Let $\mathfrak{X}^0, \ldots, \mathfrak{X}^n$ be names for pairwise disjoint quasi-orders and let $0 < \alpha_0 < \cdots < \alpha_n \leqslant \omega$. For every

$$\mathfrak{X} \in Comp\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}$$

we define the complexity number $c(\mathfrak{X})$ recursively as follows:

1. $c(\mathfrak{X}^i) := 0$,
2. $c(\mathfrak{X}^i_{x_i}) := 0$,
3. $c(\bigoplus\{\mathfrak{Y}^l : l \leq r\}) := c(\bigotimes\{\mathfrak{Y}^l : l \leq r\}) := \max\{c(\mathfrak{Y}^l) : l \leq r\} + 1$,
4. $c(\mathfrak{X}^{<\omega}) := c(\mathfrak{X}) + 1$,
5. $c\left(\mathfrak{T}\begin{pmatrix}\mathfrak{Y}^0 & \cdots & \mathfrak{Y}^m \\ \beta_0 & \cdots & \beta^m\end{pmatrix}\right) := \max\{c(\mathfrak{Y}^l) : l \leq m\} + 1$.

Definition 2.4. Let $\mathbb{X}_0 = \langle X_0, \leq_0\rangle$ and $\mathbb{X}_1 := \langle X_1, \leq_1\rangle$ be two quasi-orders. A function $e : X_0 \to X_1$ is called a *quasi-embedding*, if $e(x_0) \leq_1 e(x_0')$ implies $x_0 \leq_0 x_0'$ for every $x_0, x_0' \in X$.

Lemma 2.2. (1) *Let $\mathbb{X}_0 = \langle X_0, \leq_0\rangle$, $\mathbb{X}_1 = \langle X_1, \leq_1\rangle$ and $\mathbb{X}_2 = \langle X_2, \leq_2\rangle$ be quasi-orders and $e_0 : X_0 \to X_1$, $e_1 : X_1 \to X_2$ be quasi-embeddings. Then $e_1 \circ e_0 : X_0 \to X_2$ is a quasi-embedding.*

(2) *Let $\mathbb{X}_0 = \langle X_0, \leq_0\rangle$, $\mathbb{X}_1 = \langle X_1, \leq_1\rangle$ be quasi-orders, $e : X_0 \to X_1$ be a quasi-embedding and $f : X_1 \to \alpha + 1$ be a reification of $\mathbb{X}_1$ into α. Then f and e induce a natural reification of $\mathbb{X}_0$ into α.*

Definition 2.5. Let $\mathbb{X} = \langle X, \leq\rangle$ be a quasi-order and let $x \in X$.

$$L_{\mathbb{X}}(x) := \{y \in X : \neg x \leq y\}.$$

$L_{\mathbb{X}}(x)$ is quasi-ordered by the restriction of $\leq$ to $L_{\mathbb{X}}(x)$.

Let $\mathbb{X}(x) := \langle L_{\mathbb{X}}(x), \leq \restriction L_{\mathbb{X}}(x)\rangle$.

Note that a quasi-order $\mathbb{X}$ is a well-quasi-order if $\mathbb{X}(x)$ is a well-quasi-order for every $x \in X$.

Theorem 2.2 follows from the following lemma. (Our proof of this lemma imitates a corresponding proof of [6].)

Lemma 2.3. *Let $\mathfrak{X}^0, \ldots, \mathfrak{X}^n$ be names for pairwise disjoint quasi-orders $\mathbb{X}^0, \ldots, \mathbb{X}^n$ and let f_i be reifications of $\mathbb{X}^i$ into $\beta_i \in OT(\vartheta) \cap \Omega$. Let $0 < \alpha_0 < \cdots < \alpha_n \leq \omega$. Let*

$$\mathfrak{X} \in Comp\begin{pmatrix}\mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n\end{pmatrix}$$

be a name for a quasi-order $\mathbb{X} = \langle X, \leq\rangle$ and let $x \in X$. Then there exists a name

$$\mathfrak{X}^x \in Comp\begin{pmatrix}\mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n\end{pmatrix}$$

for a quasi-order $\mathbb{X}^x = \langle X^x, \leq^x\rangle$ and a quasi-embedding $e(\mathfrak{X}, x) : L_X(x) \to X^x$ such that $o(\mathfrak{X}^x) < o(\mathfrak{X})$.

We show how the theorem follows from this lemma. Let

$$\langle t_0, \ldots, t_k \rangle \in Bad\left(\mathbb{T}\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}\right)$$

and define recursively (with the use of Lemma 2.3)

$$f(\langle t_0, \ldots, t_k \rangle) := o\left(\cdots\left(\mathfrak{T}\begin{pmatrix} \mathfrak{X}^0 & \cdots & \mathfrak{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}^{\mathfrak{s}_0}\right)\cdots\right)^{\mathfrak{s}_k},$$

where

$$\mathfrak{s}_0 = t_0, \qquad \mathfrak{T}_0 := \mathfrak{T}\begin{pmatrix} \mathfrak{X}^0 & \cdots & \mathfrak{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix}, \qquad e_0 := e(\mathfrak{T}_0, t_0),$$

and, for $i < k$,

$$e_{i+1} := e(\mathfrak{T}_i, t_i) \circ e_i, \qquad \mathfrak{s}_{i+1} = e_{i+1}(t_{i+1}) \quad \text{and} \quad \mathfrak{T}_{i+1} = \mathfrak{T}_i^{\mathfrak{s}_i}.$$

Then this function f is a reification of

$$\mathbb{T}\begin{pmatrix} \mathbb{X}^0 & \cdots & \mathbb{X}^n \\ \alpha_0 & \cdots & \alpha_n \end{pmatrix} \quad \text{into} \quad \vartheta(\Omega^{\alpha_n} \cdot \beta_n + \cdots + \Omega^{\alpha_0} \cdot \beta_0).$$

Proof of Lemma 2.3. By induction on $c(\mathfrak{X})$. Our proof strategy is to define $e_x := e(\mathfrak{X}, x)$ by primitive recursion in previously defined functions.

Case 1: $c(\mathfrak{X}) = 0$. Assume $\mathbb{X} = \mathbb{X}^i_{x_i}$ and $x_i \in Bad(\mathbb{X}^i)$. Then $L_{X^i_{x_i}}(x) = X^i_{\langle x_i, x \rangle}$ since $x \in X^i_{x_i}$. Then $e_x := \mathrm{id} \restriction L_{X^i_{x_i}}(x)$ is quasi-embedding of $\mathbb{X}(x)$ into $\mathbb{X}^x := \mathbb{X}^i_{\langle x_i, x \rangle}$. Furthermore

$$o(\mathfrak{X}^x) = \vartheta(f_i(\langle x_i, x \rangle)) < (f_i(x)) = o(\mathfrak{X}).$$

Case 2: $\mathfrak{X} = \bigoplus\{\mathfrak{Y}^l : l \leq r\}$, $\mathfrak{X} = \bigoplus\{\mathfrak{Y}^l : l \leq r\}$ of $\mathfrak{X} = (\mathfrak{Y})^{<\omega}$. In these cases the assertion follows from [8]. The proof is actually similar to the proof of Case 3 but much simpler.

Case 3:

$$\mathfrak{X} = \mathfrak{T}\begin{pmatrix} \mathfrak{Y}^0 & \cdots & \mathfrak{Y}^m \\ \gamma_0 & \cdots & \gamma_m \end{pmatrix}.$$

Let X be the domain of $\mathbb{X}$ and let $x \in X$. Then x is a finite tree with labels in $\bigoplus\{\mathbb{Y}^l \mid l \leq m\}$. At this stage of the proof we employ a subsidiary induction on the number of vertices (elements) of x.

Subcase 3.1. Assume x is a tree consisting only of a root with a label $y \in Y^i$. By the main induction hypothesis we can pick a quasi-embedding $e(\mathfrak{Y}^i, y) =: e_y : L_{Y^i}(y) \to (Y^i)^y$ such that $o((\mathfrak{Y}^i)^y) < o(\mathfrak{Y}^i)$. The quasi-embedding e_y induces a natural quasi-embedding of

$$\mathbb{T}\begin{pmatrix} \cdots & \mathbb{Y}^i(y) & \cdots \\ \cdots & \gamma_i & \cdots \end{pmatrix}$$

into

$$\mathbb{T}\begin{pmatrix} \cdots & (\mathbb{Y}^i)^y & \cdots \\ \cdots & \gamma_i & \cdots \end{pmatrix} =: \mathbb{X}^x.$$

(Here the "$\cdots$" parts are those which remain unchanged.) Since $\mathbb{X}(x)$ is trivially quasi-embeddable into

$$\mathbb{T}\begin{pmatrix} \cdots & \mathbb{Y}^i(y) & \cdots \\ \cdots & \gamma_i & \cdots \end{pmatrix}$$

we get a quasi-embedding from $\mathbb{X}(x)$ into $\mathbb{X}^x$. The ordinal of the name of $\mathbb{X}^x$ is strictly less than $o(\mathfrak{X})$ by Lemma 1.5.

Subcase 3.2. Assume now that x is a tree with root-label $y \in Y^i$ where the root has exactly $N < \gamma_i$ immediate subtrees $x_0, \ldots, x_{N-1} \in X$.

Assume first that $N = \gamma_k$ for some $k < i$. By the subsidiary induction hypothesis we get quasi-embeddings $e(\mathfrak{X}, x_j) =: e_{x_j} : L_X(x_j) \to X^{x_j}$ such that $o(\mathfrak{X}^{x_j}) < o(\mathfrak{X})$.

By our main induction hypothesis we can pick a quasi-embedding $e(\mathfrak{Y}^i, y) =: e_y : L_{Y^i}(y) \to (Y^i)^y$ such that $o((\mathfrak{Y}^i)^y) < o(\mathfrak{Y}^i)$.

Let $\mathfrak{Z} := \mathfrak{Y}^i \otimes (\bigoplus\{(\mathfrak{X}^{x_0})^{<\omega} \otimes \cdots \otimes (\mathfrak{X}^{x_j})^{<\omega} : j < N\})$ and let $\mathbb{Z} = \langle Z, \leqslant_Z \rangle$ be the corresponding quasi-order. Set

$$\mathfrak{X}^x := \mathfrak{T}\begin{pmatrix} \cdots & \mathfrak{Y}^k \oplus \mathfrak{Y}^i \oplus \mathfrak{Z} & \cdots & (\mathfrak{Y}^i)^y & \cdots \\ \cdots & \gamma_k & \cdots & \gamma_i & \cdots \end{pmatrix}.$$

Then $\mathfrak{X}^z$ is a name for a quasi-order $\mathbb{X}^x$ and $o(\mathfrak{X}^x) < o(\mathfrak{X})$ holds by Lemma 1.5.

If $N \neq \gamma_k$ for all $k < \gamma_i$, put

$$\mathfrak{X}^x := \mathfrak{T}\begin{pmatrix} \cdots & \mathfrak{Y}^i \oplus \mathfrak{Z} & \cdots & (\mathfrak{Y}^i)^y & \cdots \\ \cdots & N & \cdots & \gamma_i & \cdots \end{pmatrix}$$

where the new column has to be inserted at the place which is determined by the ordering of $\{\gamma_0, \ldots, \gamma_i, N\}$.

We shall focus on the case $N = \gamma_k$ since the other case is similar.

We construct a quasi-embedding e_x of $\mathbb{X}(x)$ into $\mathbb{X}^x$ by primitive recursion from $e_{x_0}, \ldots, e_{x_{N-1}}$ and e_y.

Let $z \in L_X(x)$. We define $e_x(z)$ by recursion on the number of vertices of z.

Assume first that z is a tree consisting only of a root which carries a label $v \in Y^j$.

If $j \neq i$ define $e_x(z) := z$. Then $e_x(z) \in X^x$.

If $v \in Y^i$, then $v \in Y^k \uplus Y^i \uplus Z$. So let $e_x(z) := z \in X^x$.

Assume now that z is a tree with root-label $v \in Y^j$, where the root is followed by the immediate subtrees $z_0, \ldots, z_{l-1}$ where $l < \alpha_j$. We can assume inductively that $e_x(z_0), \ldots, e_x(z_l - 1)$ are defined. If $j \neq i$ or $l < N$ let $e_x(z)$ be the tree with root-label v and where the root has the immediate subtrees $e_x(z_1), \ldots, e_x(z_{l-1})$. Then $e_x(z) \in X^x$.

Now assume $j = i$ and $l \geqslant N$.

If $v \in L_{Y^i}(y)$ let $e_x(z)$ be the tree with root label $e_y(v)$ such that the root has the immediate subtrees $e_x(z_1), \ldots, e_x(z_{l-1})$.

Assume also $y \leq_{Y^i} v$, where $\leq_{Y_i}$ is the quasi-ordering on Y^i. Let $\leq_X$ be the quasi-ordering on $\mathbb{X}$. Since $z \in L_X(X)$ and $y \leq_{Y^i} v$,

$$\langle x_0, \ldots, x_{N-1} \rangle \leq_X^{<\omega} \langle z_0, \ldots, z_{l-1} \rangle$$

does not hold. (Otherwise we would have $x \leq_X z$.) So there exists a minimal $s < N$ such that $\langle x_0, \ldots, x_{s-1} \rangle \leq_X^{<\omega} \langle z_0, \ldots, z_{l-1} \rangle$ does not hold. Then $\langle x_0, \ldots x_{s-2} \rangle \leq_X^{<\omega} \langle z_0, \ldots, z_{l-1} \rangle$.

Let $j_0 < l$ be minimal such that $x_0 \leq_X z_{j_0}$. Let $j_1 < l$ minimal such that $j_0 < j_1$ and $x_1 \leq z_{j_1}$. Let finally $j_{s-2} < l$ be minimal such that $j_{s-3} < j_{s-2}$ and $x_{s-2} \leq z_{j_{s-2}}$. Then $z_{j_0}, \ldots, z_{j_{s-2}} \in X$, $\langle z_0, \ldots, z_{j_0 - 1} \rangle \in L_X(x_0)^{<\omega}$, $\langle z_{j_{s-3}+1}, \ldots, z_{j_{s-2}-1} \rangle \in L_X(x_{s-2})^{<\omega}$, and $\langle z_{(j_{s-2})+1}, \ldots, z_{l-1} \rangle \in L_X(x_{s-1})^{<\omega}$. Define $e_x(z)$ to be the tree determined by the immediate subtrees $e_x(x_{j_0}), \ldots, e_x(z_{j_{s-2}})$ and the root labeled with

$$\langle v, \langle \langle e_x(z_0), \ldots, e_x(z_{j_0-1}) \rangle, \ldots, \langle e_x(z_{(j_{s-2})+1}), \ldots, e_x(z_{l-1}) \rangle \rangle \rangle \in \mathbb{Y}^i \oplus \mathbb{Z}.$$

Then $e_x(z) \in X^x$ since $s - 1 < \gamma_k$.

Next, we show that $e_x(z) \leq^x e_x(z')$ implies $z \leq_X z'$ by induction on the sum of vertices of z and z'. If $e_x(z)$ is embeddable into an immediate subtree $\hat{z}$ of $e_x(z')$, then $\hat{z}$ has the form $e_x(z'')$ for some immediate subtree z'' of z'. Then $z \leq^x z'' \leq^x z'$ by the induction hypothesis.

Now we assume that $e_x(z)$ is embeddable into $e_x(z')$ and that $e_x(z)$ is not embeddable into an immediate subtree of $e_x(z')$. Then the root-label r of $e_x(z)$ is less than or equal to the root-label r' of $e_x(z')$ with respect to the appropriate quasi-ordering. Thus, if

$$r = \langle v, \langle \langle e_x(z_0), \ldots, e_x(z_{j_0-1}) \rangle, \ldots, \langle e_x(z_{(j_{s-2})+1}), \ldots, e_x(z_{l-1}) \rangle \rangle \rangle$$

and

$$r' = \langle v', \langle \langle e_x(z'_0), \ldots, e_x(z'_{j'_0-1}) \rangle, \ldots, \langle e_x(z'_{(j'_{s'-2})+1}), \ldots, e_x(z'_{l'-1}) \rangle \rangle,$$

then $s = s'$, $v \leq_{Y^i} v'$,

$$\langle e_x(z_0), \ldots, e_x(z_{j_0-1}) \rangle \, (\leq^x)^{<\omega} \, \langle e_x(z'_0), \ldots, e_x(z'_{j'_0-1}) \rangle, \ldots,$$

and finally

$$\langle e_x(z_{(j_{s-2})+1}), \ldots, e_x(z_{l-1}) \rangle \, (\leq^x)^{<\omega} \, \langle e_x(z'_{(j'_{s'-2})+1}), \ldots, e_x(z'_{l'-1}) \rangle.$$

Furthermore, there exists a permutation π of $\{0, \ldots, s-2\}$ such that $e_x(z_j) \leq^x e_x(z'_{\pi(j)})$ for $j \leq s-2$. Therefore, by the inductive assumption, $z_j \leq_X z'_{\pi(j)}$ for $j \leq s-2$. By combining the above results, we obtain a one-to-one mapping $\rho\colon \{0, \ldots, l-1\} \to \{0, \ldots, l'-1\}$ such that $z_i \leq^x z'_{l(i)}$ holds for $0 \leq i \leq l-1$. Since $v \leq_{Y^i} v'$, we conclude $z \leq_X z'$. □

Corollary 2.1. $ACA_0 \vdash \forall n\, KT(n) \leftrightarrow KT(\omega) \leftrightarrow WF(\vartheta\Omega^\omega)$.

Proof. Theorem 2.2 and Theorem 1.2 □

3. A comparison between two ordinal notation systems for the Howard–Bachmann-ordinal

We have just seen that the ordinal notation system for the Howard–Bachmann-ordinal which is based on the function ϑ is an appropriate tool for the ordinal analysis of Kruskal's theorem. For a perspicious proof-theoretic analysis of Π^1_2 bar induction we need another concept for representing the Howard–Bachmann-ordinal namely the ψ-function which is due to Buchholz. See, for example, [1] for a definition.

Let $\Omega(1, \omega) := \Omega^{\omega}$ and $\Omega(n+1, \omega) := \Omega^{\Omega(n,\omega)}$. We are going to show that $\vartheta(\Omega(n, \omega)) = \psi(\Omega(n+1, \omega))$ is true for every natural number n. This technical result will be needed for comparing the proof-theoretic strength of Kruskal's theorem and Π^1_2-BI over ACA_0. (This section is very technical and may be skipped at first reading.)

Definition 3.1. Inductive definition of sets of ordinals $C_n(\alpha)$, $C(\alpha)$, and of ordinals $\psi\alpha$ by main recursion on α and side recursion on $n < \omega$.

1. $\{0, \Omega\} \subseteq C_n(\alpha)$,
2. $\beta, \gamma \in C_n(\alpha) \Rightarrow \omega^{\beta} + \gamma \in C_{n+1}(\alpha)$.
3. $\beta \in C(\beta) \cap \alpha \cap C_n(\alpha) \Rightarrow \psi\beta \in C_{n+1}(\alpha)$,
4. $C(\alpha) := \bigcup \{C_n(\alpha): n < \omega\}$.

Let $\psi\alpha := \min\{\xi: \xi \notin C(\alpha)\}$.

The following Lemmata can be gathered from [1] or [4].

Lemma 3.1. $\psi\alpha < \Omega$.

Lemma 3.2. (1) $\psi\alpha = C(\alpha) \cap \Omega$,
(2) $\alpha \in C(\alpha) \cap \beta \Rightarrow \psi\alpha < \psi\beta$,
(3) $\alpha, \beta < \psi\gamma \Rightarrow \omega^{\alpha} + \beta < \psi\gamma$.

Definition 3.2. Inductive definition of a set $OT(\psi)$ of ordinals and a natural number $G_{\psi}\alpha$ for $\alpha \in OT(\psi)$.

1. $\{0, \Omega\} \subseteq OT(\psi)$, $\quad G_{\psi}0 := G_{\psi}\Omega := 0$,
2. $\alpha = \alpha_1 + \cdots + \alpha_n$, $\alpha > \alpha_1 \geq \cdots \geq \alpha_n$, $\alpha_1, \ldots, \alpha_n \in OT(\psi) \cap AP \Rightarrow$ $\alpha \in OT(\psi)$, $G_{\psi}\alpha := \max\{G_{\psi}\alpha_1, \ldots, G_{\psi}\alpha_n\} + 1$,
3. $\alpha = \omega^{\alpha_1}$, $\alpha > \alpha_1$, $\alpha_1 \in OT(\psi) \Rightarrow \alpha \in OT(\psi)$, $G_{\psi}\alpha := G_{\psi}\alpha_1 + 1$,
4. $\alpha = \psi\alpha_1$, $\alpha_1 \in C(\alpha_1)$, $\alpha_1 \in OT(\psi) \Rightarrow \alpha \in OT(\psi)$, $G_{\psi}\alpha := G_{\psi}\alpha_1 + 1$.

Lemma 3.3. (1) $OT(\psi) = C(\varepsilon_{\Omega+1})$,
(2) $OT(\psi) \cap \Omega \subset \psi\varepsilon_{\Omega+1}$,
(3) $OT(\psi) \cap \alpha = \alpha$ *for* $\alpha \leq \psi\varepsilon_{\Omega+1}$.

To see that $OT(\psi)$ may be considered as a primitive recursive ordinal notation system we must be able to decide the relation $\alpha \in C(\beta)$ for $\alpha, \beta \in OT(\psi)$. For this purpose we introduce the auxiliary concept of coefficient sets.

Definition 3.3. Inductive definition of a set of ordinals $K\alpha$ for $\alpha \in OT(\psi)$.

1. $K0 := K\Omega := \emptyset$,
2. $K(\alpha_1 + \cdots + \alpha_n) := K\alpha_1 \cup \cdots \cup K\alpha_n$,
3. $K(\omega^{\alpha_1}) = K\alpha_1$,
4. $K\psi\alpha_1 := K\alpha_1 \cup \{\alpha_1\}$.

Let $k\alpha := \max(K\alpha \cup \{0\})$ and $h\alpha := k\alpha + \omega^{\alpha}$(cf. [4]).

Lemma 3.4. $\alpha \in OT(\psi) \Rightarrow (K\alpha < \beta \Leftrightarrow \alpha \in C(\beta))$.

Lemma 3.5. (1) $\alpha \in OT(\psi) \Rightarrow \alpha^*,\ k\alpha \in OT(\psi),\ G_\psi\alpha^* \leq G_\psi\alpha,\ G_\psi k\alpha \leq G_\psi\alpha,$ $k\alpha = k\alpha^*$,

(2) $\alpha \in OT(\psi) \wedge k\alpha < \alpha \Rightarrow \psi\alpha \in OT(\psi)$.

(3) $\alpha \in OT(\psi) \Rightarrow \psi h\alpha \in OT(\psi)$,

Lemma 3.6. $\alpha, \beta \in OT(\psi) \Rightarrow (\alpha^* < \psi\beta \Leftrightarrow k\alpha < \beta)$.

Definition 3.4. Recursive definition of $\hat{\alpha}$ for $\alpha \in OT(\vartheta)$

1. $\hat{0} := 0, \quad \hat{\Omega} := \Omega$,
2. $(\alpha_1 + \cdots + \alpha_n)\hat{} := \hat{\alpha}_1 + \cdots + \hat{\alpha}_n$,
3. $(\omega^{\alpha_1})\hat{} := \omega^{\hat{\alpha}_1}$,
4. $(\vartheta\alpha_1)\hat{} := \psi h\hat{\alpha}_1$.

Lemma 3.7. (1) $\alpha \in OT(\vartheta) \Rightarrow \hat{\alpha} \in OT(\psi)$,

(2) $\alpha, \beta \in OT(\vartheta) \Rightarrow (\alpha < \beta \Leftrightarrow \hat{\alpha} < \hat{\beta})$,

(3) $\alpha \in OT(\vartheta) \Rightarrow (\alpha^*)\hat{} = \hat{\alpha}^*$.

Proof. By a simultaneous induction on $G_\vartheta\alpha + G_\vartheta\beta$. We consider only the non-trivial case for (2). Let $\alpha = \vartheta\alpha_1$ and $\beta = \vartheta\beta_1$. Assume first that $\alpha_1 < \beta_1$ and $\alpha_1^* < \vartheta\beta_1$. Then we see $\hat{\alpha}_1 < \hat{\beta}_1$ and $\hat{\alpha}_1^* < \psi h\hat{\beta}_1$ by the induction hypothesis. Thus $k\hat{\alpha}_1{}^* < h\hat{\beta}_1$ and therefore $h\hat{\alpha}_1 = k\hat{\alpha}_1 + \omega^{\hat{\alpha}_1} = k\hat{\alpha}_1{}^* + \omega^{\hat{\alpha}_1} < k\hat{\beta}_1 + \omega^{\hat{\beta}_1} = h\hat{\beta}_1$. This yields $(\vartheta\alpha_1)\hat{} < (\vartheta\beta_1)\hat{}$. Assume now $\alpha_1 > \beta_1$ and $\vartheta\alpha_1 \leq \beta_1^*$. Then $(\vartheta\alpha_1)\hat{} \leq \hat{\beta}_1^*$. Since $\hat{\beta}_1{}^* \leq (h\beta_1)\hat{}{}^*$ we see $\hat{\beta}_1{}^* \leq \psi h\hat{\beta}_1$ by Lemma 3.6. Thus $(\vartheta\alpha_1)\hat{} < (\vartheta\beta_1)\hat{}$.

Corollary 3.1. (1) $\vartheta\alpha \leq \psi h\hat{\alpha}$,

(2) $\vartheta(\Omega(n, \omega)) \leq \psi(\Omega(n+1, \omega))$.

Proof. This follows from Lemma 3.7 and the well known fact that $f\beta \geq \beta$ holds for any strictly monotonic ordinal function f. $\square$

We now define the reverse embedding $\alpha \mapsto \bar{\alpha}$ of $OT(\psi)$ into $OT(\vartheta)$. Of course the trivial setup $\psi\alpha \mapsto \vartheta\bar{\alpha}$ will induce an order-preserving mapping of $OT(\psi)$ into $OT(\vartheta)$. But this yields only $\psi\alpha \leqslant \vartheta\bar{\alpha} \leqslant \psi h\hat{\bar{\alpha}}$ and this is not an equality in the interesting cases.

Lemma 3.8. *For each $\alpha \in OT(\psi)\setminus\{0\}$ there exist uniquely determined $\alpha_1, \ldots, \alpha_n \in OT(\psi)$ and $\beta_1, \ldots, \beta_n \in OT(\psi)$ such that $\alpha = \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n)$, $\alpha_1 > \cdots > \alpha_n$ and $\beta_1, \ldots, \beta_n < \Omega$.*

Definition 3.5. $\alpha =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots \Omega^{\alpha_n}(1+\beta_n)$
$:\Leftrightarrow\ \alpha = \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n) \wedge \alpha_1 > \cdots > \alpha_n \wedge \beta_1, \ldots, \beta_n < \Omega.$

Lemma 3.9. $\alpha =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n)$
$\Rightarrow\ \alpha^* = \max\{\alpha_1^*, \ldots, \alpha_n^*, \beta_1^*, \ldots, \beta_n^*\}.$

Definition 3.6. Induction definition of $\bar{\alpha}$ for $\alpha \in OT(\psi)$.

1. $\bar{0} := 0, \quad \bar{\Omega} := \Omega,$
2. $\overline{\alpha_1 + \cdots + \alpha_n} := \overline{\alpha_1} + \cdots + \overline{\alpha_n},$
3. $\overline{\omega^{\alpha_1}} := \omega^{\overline{\alpha_1}},$
4. $\overline{\psi 0} := \vartheta 0,$
5. $\overline{\psi\alpha} := \vartheta(\Omega\overline{\alpha_n} + \cdots + \vartheta(\Omega\overline{\alpha_2} + \vartheta(\Omega\overline{\alpha_1} + \overline{\beta_1}) + \overline{\beta_2}) \cdots + \overline{\beta_n}),$
   if $\alpha =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n)$.

Lemma 3.10. $\alpha \in OT(\psi)\ \Rightarrow\ \bar{\alpha} \in OT(\vartheta)$.

The following lemma is of crucial importance.

Lemma 3.11. (1) *If $\alpha, \beta \in OT(\psi)$, then $(\alpha < \beta \Leftrightarrow \bar{\alpha} < \bar{\beta})$.*

(2) *If $\alpha =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n) + \Omega^{\alpha_{n+1}}(1+\beta_{n+1}) + \cdots + \Omega^{\alpha_m}(1+\beta_m) \in OT(\psi)$, if $\beta =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n) + \Omega^{\gamma_{n+1}}(1+\delta_{n+1}) \in OT(\psi)$, and if $\alpha \cup K\alpha < \beta$, then*

$$\vartheta(\Omega\overline{\alpha_m} + \cdots + \vartheta(\Omega\overline{\alpha_{n+1}} + \vartheta(\Omega\overline{\alpha_n} + (\cdots) + \overline{\beta_n}) + \overline{\beta_{n+1}}) + \cdots + \overline{\beta_m})$$
$$< \vartheta(\Omega\overline{\gamma_{n+1}} + \vartheta(\Omega\overline{\alpha_n} + (\cdots) + \overline{\beta_n}) + \delta_{n+1}),$$

(3) *If $\beta =_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n) + \Omega^{\gamma_{n+1}}(1+\delta_{n+1}) \in OT(\psi)$, $\alpha < \beta$ and $\alpha \in OT(\psi)$, then*

$$\bar{\alpha}^* < \vartheta(\Omega\overline{\gamma_{n+1}} + \vartheta(\Omega\overline{\alpha_n} + (\cdots) + \overline{\beta_n}) + \delta_{n+1}).$$

Proof. By main induction on $2^{G_\psi\alpha} + 2^{G_\psi\beta}$ and side induction on $m - n$.

(1) The critical case is $\alpha = \psi\alpha_0$, $\beta = \psi\beta_0$. $K\alpha_0 < \alpha_0$, $K\beta_0 < \beta_0$ and $\alpha < \beta$. Let

$\alpha_0 :=_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1) + \cdots + \Omega^{\alpha_n}(1+\beta_n) + \Omega^{\alpha_{n+1}}(1+\beta_{n+1}) + \cdots + \Omega^{\alpha_m}(1+\beta_m)$

and

$$\beta_0 :=_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1)+\cdots+\Omega^{\alpha_n}(1+\beta_n)+\Omega^{\gamma_{n+1}}(1+\delta_{n+1})+\cdots+\Omega^{\gamma_r}(1+\delta_r),$$

where $\Omega^{\alpha_{n+1}}(1+\beta_{n+1})<\Omega^{\gamma_{n+1}}(1+\delta_{n+1})$. Let

$$\gamma_0 :=_{\Omega\text{-NF}} \Omega^{\alpha_1}(1+\beta_1)+\cdots+\Omega^{\alpha_n}(1+\beta_n)+\Omega^{\gamma_{n+1}}(1+\delta_{n+1}).$$

Then $\alpha_0<\gamma_0$ and $K\alpha_0<\gamma_0$. By the main induction hypothesis applied to (2) we see that

$$\begin{aligned}\overline{\psi\alpha_0}&<\vartheta(\Omega\overline{\gamma_{n+1}}+\vartheta(\Omega\overline{\alpha_n}+(\cdots)+\overline{\beta_n})+\overline{\delta_{n+1}})\\&<\vartheta(\Omega\overline{\gamma_{n+2}}+\vartheta(\Omega\overline{\gamma_{n+1}}+(\cdots)+\overline{\delta_{n+1}})+\overline{\delta_{n+2}})\\&\leqslant\cdots\leqslant\overline{\psi\beta_0}.\end{aligned}$$

(2) The assertion holds for $m-n=0$. We assume first that $\gamma_{n+1}>\alpha_{n+1}$ is true. Let

$$\begin{aligned}\alpha' :=_{\Omega\text{-NF}}& \Omega^{\alpha_1}(1+\beta_1)+\cdots+\Omega^{\alpha_n}(1+\beta_n)+\Omega^{\alpha_{n+1}}(1+\beta_{n+1})\\&+\cdots+\Omega^{\alpha_{m-1}}(1+\beta_{m-1}).\end{aligned}$$

Then $\max\{\alpha', k\alpha'\}\leqslant\max\{\alpha, k\alpha\}<\beta$. We conclude by the side induction hypothesis

$$\begin{aligned}&\vartheta(\Omega\overline{\alpha_{m-1}}+\cdots+\vartheta(\Omega\overline{\alpha_{n+1}}+\vartheta(\Omega\overline{\alpha_n}+(\cdots)+\overline{\beta_n})+\overline{\beta_{n+1}})+\cdots+\overline{\beta_{m-1}})\\&\quad<\vartheta(\Omega\overline{\gamma_{n+1}}+\vartheta(\Omega\overline{\alpha_n}+(\cdots)+\overline{\beta_n})+\delta_{n+1})=:\gamma'.\end{aligned}$$

The main induction hypothesis for (1) yields $\overline{\gamma_{n+1}}>\overline{\alpha_{n+1}}>\cdots>\overline{\alpha_m}$. The main induction hypothesis for (3) yields $\overline{\beta_m}^*<\gamma'$ and $\overline{\alpha_m}^*<\gamma'$. Therefore

$$\vartheta(\Omega\overline{\alpha_m}+\cdots+\vartheta(\Omega\overline{\alpha_{n+1}}+\vartheta(\Omega\overline{\alpha_n}+(\cdots)+\overline{\beta_n})+\overline{\beta_{n+1}})+\cdots+\overline{\beta_m})<\gamma'.$$

Assume now $\gamma_{n+1}=\alpha_{n+1}$ and $\beta_{n+1}<\delta_{n+1}$. If $\alpha_m<\alpha_{n+1}=\gamma_{n+1}$ the assertion follows as before. Assume now $\underline{\alpha_m=\alpha_{n+1}}=\gamma_{n+1}$ and $\beta_{n+1}<\delta_{n+1}$. The main induction hypothesis for (a) yields $\overline{\beta_{n+1}}<\overline{\delta_{n+1}}$. Thus

$$\vartheta(\Omega\overline{\alpha_m}+(\cdots)+\overline{\beta_{n+1}})<\vartheta(\Omega\overline{\gamma_{n+1}}+(\cdots)+\overline{\delta_{n+1}}).$$

(3) The claim is true for $\alpha\in\{0,\Omega\}$. If α is not of the form $\psi\alpha_0$ with $K\alpha_0<\alpha_0$ then the assertion follows immediately from the main induction hypothesis for (3). Let $\alpha=\psi\alpha_0$ and $K\alpha_0\cup\{\alpha_0\}=K\alpha<\beta$. The main induction hypothesis for (2) yields the assertion. □

Corollary 3.2. $\psi(\Omega(n+1,\omega))=\vartheta(\Omega(n,\omega))$.

Remark. Let $\bar\theta$ be the collapsing function defined in [7]. Then $\bar\theta(\Omega(n,\omega))0=\vartheta(\Omega(n,\omega))$ for every $n\in\mathbb{N}$.

4. Majorization relations and fundamental functions

In this section we restate for convenience the concepts of majorization relations and fundamental functions which are developed in [1]. These concepts are needed for carrying through the ordinal analysis of the restricted bar induction schemata. All missing proofs can be found in [1].

Definition 4.1. 1. $\alpha \lhd_\mu \beta$ means $\alpha < \beta$ and for all δ, η:

$$\alpha \leq \delta \leq \min\{\beta, \eta\},\ \delta, \mu \in C(\eta) \Rightarrow \alpha \in C(\eta).$$

2. $\alpha \lhd \beta \ :\Leftrightarrow\ \alpha \lhd_0 \beta$,
3. $\alpha \unlhd \beta \ :\Leftrightarrow\ (\alpha \lhd \beta \vee \alpha = \beta)$.

Lemma 4.1. (1) $\alpha \lhd \beta \Rightarrow \alpha \lhd_\mu \beta$,
(2) $\alpha < \beta \Rightarrow \alpha \lhd_\alpha \beta$,
(3) $\alpha < \beta < \gamma \;\&\; \alpha \lhd_\mu \gamma \Rightarrow \alpha \lhd_\mu \beta$,
(4) $0 < \beta < \varepsilon_0 \Rightarrow \alpha \lhd \alpha + \beta$,
(5) $\alpha < \beta < \Omega \Rightarrow \alpha \lhd \beta$,
(6) $\alpha \lhd \beta \Rightarrow \alpha + 1 \unlhd \beta$.

Lemma 4.2. $\alpha \lhd_\mu \beta,\ \beta \lhd_\mu \gamma \Rightarrow \alpha \lhd_\mu \gamma$.

Lemma 4.3. $\alpha \lhd_\mu \beta,\ \beta < \omega^{\gamma+1} \Rightarrow \omega^\gamma + \alpha \lhd_\mu \omega^\gamma + \beta$.

Corollary 4.1. $\omega^\alpha \cdot n \lhd \omega^\alpha \cdot (n+1)$.

Lemma 4.4. $\alpha \lhd_\mu \beta \Rightarrow \omega^\alpha \cdot n \lhd_\mu \omega^\beta$.

Lemma 4.5. $\alpha \lhd_\mu \beta$, $\mu \in C(\alpha)$, $\beta \in C(\beta)$ *implies*:
(1) $\alpha \in C(\alpha)$,
(2) $\psi\alpha \lhd_\mu \psi\beta$.

Corollary 4.2. $\alpha = \alpha_0 + 1 \in C(\alpha) \Rightarrow \alpha_0 \in C(\alpha) \;\&\; \psi\alpha_0 \lhd \psi\alpha$.

Definition 4.2. A function f with the *domain* $dom(f) \subseteq OT(\psi)$ is said to be a *fundamental function* if the following holds.

F1. If $\beta \in dom(f)$ and $\alpha < \beta$, then $\alpha \in dom(f)$ and $f(\alpha) \lhd_\alpha f(\beta)$.

F2. If $\beta \in dom(f)$ and $f(0) \leq \delta < f(\beta)$, then there is an $\alpha < \beta$ such that $f(\alpha) \leq \delta < f(\alpha+1)$ and $f(\alpha) \lhd f(\alpha+1)$.

F3. If $\alpha \in dom(f)$ and $f(\alpha) \in C(\eta)$, then $\alpha \in C(\eta)$.

Lemma 4.6. *If f is a fundamental function and $\alpha \in dom(f)$, then $\alpha \leq f(\alpha)$.*

Definition 4.3. Let Id_β be the function with domain $dom(Id_\beta) := \{\alpha \in OT(\psi) \colon \alpha \leqslant \beta\}$ and $Id_\beta(\alpha) := \alpha$ for all $\alpha \in dom(Id_\beta)$.

Lemma 4.7. *Id_β is a fundamental function.*

Definition 4.4. Let f be a fundamental function.

1. Let $\omega^\gamma + f$ be the function with domain $dom(\omega^\gamma + f) := \{\alpha \in dom(f) \colon f(\alpha) < \omega^{\gamma+1}\}$ and $(\omega^\gamma + f)(\alpha) := \omega^\gamma + f(\alpha)$ for all $\alpha \in dom(\omega^\gamma + f)$.
2. Let ω^f be the function with domain $dom(\omega^f) := dom(f)$ and $(\omega^f)(\alpha) := \omega^{f(\alpha)}$ for all $\alpha \in dom(\omega^f)$.
3. Let ψf be the function with domain $dom(\psi f) := \{\alpha \in dom(f) \colon \alpha < \Omega, f(\alpha) \in C(f(\alpha))\}$ and $(\psi f)(\alpha) := \psi(f(\alpha))$ for all $\alpha \in dom(\psi f)$.

Lemma 4.8. *If f is a fundamental function, then also $\omega^\gamma + f$, ω^f and ψf are fundamental functions.*

Lemma 4.9. *If f is a fundamental function with $\alpha, \Omega \in dom(f)$, $\alpha < \beta = \psi f(\alpha)$ and $f(\alpha) \lhd f(\Omega)$, then also $f(\beta) \lhd f(\Omega)$.*

Corollary 4.3. *If f is a fundamental function with $\Omega \in dom(f)$, then $f(\psi(f(0))) \lhd f(\Omega)$.*

5. Ordinal analysis of restricted bar induction

In this section we determine the proof-theoretic strength of the subsystems of second-order arithmetic $ACA_0 + \Pi^1_2\text{-}BI$ which is based on arithmetical comprehension and Π^1_2 bar induction. From this result we shall gather the unprovability of Kruskal's Theorem in $ACA_0 + \Pi^1_2\text{-}BI$ as well as the proof-theoretic equivalence of $ACA_0 + \Pi^1_2\text{-}BI$ and Kripke–Platek set theory plus infinity axiom but with foundation restricted to set-theoretic Π_2 formulas. Our device for dealing with Π^1_2 bar induction will be Buchholz' Ω-rule (cf. [1]).

To set the context, we fix some notations. The language of second-order arithmetic, $\mathcal{L}_2$, consists of free numerical variables $a, b, c, d, \ldots$, bound numerical variables $x, y, z, \ldots$, free set variables $U, V, W, \ldots$, bound set variables $X, Y, Z, \ldots$, the constant 0, a symbol for each primitive recursive function, and the symbols $=$ and $\in$ for equality in the first sort and the elementhood relation, respectively. The *numerical terms of* $\mathcal{L}_2$ are build up in the usual way; $r, s, t, \ldots$ are syntactic variables for them. *Formulas* are obtained from atomic formulas $(s = t)$, $(s \in U)$ and negated atomic formulas $\neg(s = t)$, $\neg(s \in U)$ by closing under $\wedge$, $\vee$ and quantification $\forall x$, $\exists x$, $\forall X$, $\exists X$ over both sorts; so we stipulate that formulas are in negation normal form.

The classes of Π^1_n- and Σ^1_n-*formulas* are defined as usual (with $\Pi^1_0 = \Sigma^1_0 = \bigcup\{\Pi^0_n \colon n \in \mathbb{N}\}$. $\neg A$ is defined by de Morgan's laws; $A \rightarrow B$ stands for $\neg A \vee B$. All theories in $\mathscr{L}_2$ will be assumed to contain the axioms and rules of classical two sorted predicate calculus, with equality in the first sort. In addition, it will be assumed that they comprise the system ACA_0, ACA_0 contains all axioms of elementary number theory, i.e., the usual axioms for 0, $'$ (successor), the defining equations for the primitive recursive functions, the *induction axiom*

$$\forall X\,[0 \in X \wedge \forall x\,(x \in X \rightarrow x' \in X) \rightarrow \forall x\,(x \in X)],$$

and all instances of *arithmetical comprehension*

$$\exists Z\,\forall x\,[x \in Z \leftrightarrow F(x)],$$

where $F(a)$ is an *arithmetic formula*, i.e., a formula without set quantifiers.

For a 2-place relation $<$ and an arbitrary formula $F(a)$ of $\mathscr{L}_2$ we define

$$\begin{aligned} &Prog(<, F) := \forall x\,[\forall y\,(y < x \rightarrow F(y)) \rightarrow F(x)] \qquad (\textit{progressiveness})\\ &TI(<, F) := Prog(<, F) \rightarrow \forall x\,F(x) \qquad (\textit{transfinite induction})\\ &WF(<) := \forall X\,TI(<, X) := \forall X\,(\forall x[\forall y\,(y < x \rightarrow y \in X) \rightarrow x \in X]\\ &\qquad\qquad\qquad \rightarrow \forall x\,[x \in X]) \qquad (\textit{well-foundedness}). \end{aligned}$$

Let $\mathscr{F}$ be any collection of formulas of $\mathscr{L}_2$. For a 2-place relation $<$ we will write $< \in \mathscr{F}$, if $<$ is defined by a formula $Q(x, y)$ of $\mathscr{F}$ via $x < y := Q(x, y)$. In addition, we will use the notation $< \in \mathscr{F}^-$ to express that $Q(x, y)$ contains no set parameters.

Definition 5.1. 1. $(\Pi^1_n\text{-}BI)_0$ denotes ACA_0 extended by the Π^1_n bar induction scheme, i.e., all formulas of the form

$$WF(<) \rightarrow TI(<, F),$$

where $< \in \Pi^1_0$ and $F \in \Pi^1_n$.

2. $(\Pi^1_n\text{-}BI)^-_0$ denotes the modification, where $<$ is required to contain no set variables.

3. $(\Pi^1_n\text{-}BI)^-$ and $(\Pi^1_n\text{-}BI)$ denote the corresponding theories augmented by the scheme

$$F(0) \wedge \forall x\,(F(x) \rightarrow F(x')) \rightarrow \forall x\,F(x)$$

for every $\mathscr{L}_2$-formula $F(a)$.

For technical purposes it is convenient to reformulate $(\Pi^1_2\text{-}BI)_0$ in a sequent calculus.

Definition 5.2. The sequent calculus version of $(\Pi^1_2\text{-}BI)_0$ derives finite sets of formulas denoted by $\Gamma, \Theta, \Xi, \Lambda, \ldots$. The intended meaning of Γ is the disjunction of all formulas of Γ. We use the notation Γ, A for $\Gamma \cup \{A\}$ and Γ, Ξ for $\Gamma \cup \Xi$. Let Γ be an arbitrary finite set of formulas.

The *axioms of* $(\Pi^1_2\text{-}BI)_0$ are:

$(Ax1)$ $\Gamma, A, \neg A$ for every prime formula A.

$(Ax2)$ Γ, Δ if Δ is a set of prime and negated prime formulas such that $\bigvee \Delta$ (i.e., the disjunction of all formulas of Δ) is a tautological consequence of the equality axioms and the defining equalities of the primitive recursive functions.

(IA) $\Gamma, \forall X\,[0 \in X \wedge \forall x\,(x \in X \rightarrow x' \in X) \rightarrow \forall x\,(x \in X)]$.

$(\Pi^1_0\text{-}CA)$ $\Gamma, \exists Z\,\forall x\,[x \in Z \leftrightarrow F(x)]$ where F is arithmetic.

The logical *rules of inferences* are:

$(\wedge)$ $\vdash \Gamma, A$ and $\vdash \Gamma, B \Rightarrow \vdash \Gamma, A \wedge B$,

$(\vee)$ $\vdash \Gamma, A_i \Rightarrow \vdash \Gamma, A_0 \vee A_1$ if $i \in \{0, 1\}$,

$(\forall_1)$ $\vdash \Gamma, F(a) \Rightarrow \vdash \Gamma, \forall x\, F(x)$,

$(\forall_2)$ $\vdash \Gamma, F(U) \Rightarrow \vdash \Gamma, \forall X\, F(X)$,

$(\exists_1)$ $\vdash \Gamma, F(t) \Rightarrow \vdash \Gamma, \exists x\, F(x)$,

$(\exists_2)$ $\vdash \Gamma, F(U) \Rightarrow \vdash \Gamma, \exists X\, F(X)$,

(Cut) $\vdash \Gamma, A$ and $\vdash \Gamma, \neg A \Rightarrow \vdash \Gamma$,

where in $(\forall_1)$ and $(\forall_2)$ the free variable a, respectively U is not to occur in the conclusion.

The only *non-logical rule of inference of* $(\Pi^1_2\text{-}BI)_0$ is the Π^1_2 bar induction rule:

$$\vdash \Gamma, WF(<) \quad \text{and} \quad \vdash \Gamma, \exists y\,\exists Y\,\forall Z\,[y < a \wedge \neg B(y, Y, Z)], \exists Z\, B(a, U, Z) \\ \Rightarrow \vdash \Gamma, \exists Z\, B(b, V, Z),$$

where $<$ and B are arithmetic; and a and U are not to occur in $\Gamma, \forall x\,\forall Y\,\exists X\, B(x, Y, Z)$.

We shall conceive axioms as inferences with an empty set of premises. The *minor formulas* (m.f.) of an inference are those formulas which are rendered prominently in its premises. The *principal formulas* (p.f.) of an inference are the formulas rendered prominently in its conclusion. (Cut) has no principal formula. So any inference has the form

$(*)$ for all $i < k$ $\vdash \Gamma, \Xi_i \Rightarrow \vdash \Gamma, \Xi$

$(0 \leqslant k \leqslant 2)$, where Ξ consists of p.f. and X_i is the set of m.f. in the ith premise. The formulas in Γ are called *side formulas* (s.f.) of $(*)$.

Derivations of $(\Pi^1_2\text{-}BI)_0$ are defined inductively, as usual. $\mathscr{D}, \mathscr{D}', \mathscr{D}_0, \ldots$ range as syntactic variables over $(\Pi^1_2\text{-}BI)_0$ derivations. All this is completely standard, and we refer to [9] for notions like 'length of a derivation of $\mathscr{D}$' (abbreviated by $|\mathscr{D}|$), 'last inference of $\mathscr{D}$', 'direct subderivation of $\mathscr{D}$'. We write $\mathscr{D} \vdash \Gamma$ to mean that $\mathscr{D}$ is a derivation of Γ.

The most important feature of sequent calculi is cut-elimination. Our sequent calculus $(\Pi^1_2\text{-}BI)_0$ admits cut-elimination concerning cuts whose cut formula is neither a principal formula of a non-logical rule of inference nor a principal formula of an axiom. This is a general phenomenon which will be exploited next. To state this fact concisely, let us introduce a measure of complexity, $gr(A)$, the *grade of a formula A*:

1. $gr(A) = 0$, if A is a prime formula or negated prime formula.
2. $gr(\forall X F(X)) = gr(\exists X F(X)) = \omega$, if $F(U)$ is arithmetic.
3. $gr(A \wedge B) = gr(A \vee B) = \max\{gr(A), gr(B)\} + 1$.
4. $gr(\forall x\, H(x)) = gr(\exists x\, H(x)) = gr(H(0)) + 1$.
5. $gr(\forall X\, G(X)) = gr(\exists X\, G(X)) = gr(G(U)) + 1$, if G is not arithmetic.

The *cut-rank*, $\rho(\mathscr{D})$, of a derivation $\mathscr{D}$ is also defined by induction: Let $\mathscr{D}_i$, $i < k$, be the direct subderivations of $\mathscr{D}$. If the last inference of $\mathscr{D}$ is (Cut) with m.f. A and $\neg A$ let $\rho(\mathscr{D}) := \sup\{gr(A) + 1, \sup\{\rho(\mathscr{D}_i): i < k\}\}$. Otherwise, let $\rho(\mathscr{D}) := \sup\{\rho(\mathscr{D}_i): i < k\}$. By $(\Pi^1_2\text{-}BI)_0 \vdash^k_\rho \Gamma$ we mean that there is a derivation $\mathscr{D} \vdash \Gamma$ in $(\Pi^1_2\text{-BI})_0$ such that $|\mathscr{D}| \leq k$ and $\rho(\mathscr{D}) \leq \rho$.

Theorem 5.1 (Cut-elimination). *Let* $2^k_0 := k$ *and* $2^k_{m+1} := 2^l$ *where* $l := 2^k_m$. *If* $(\Pi^1_2\text{-}BI)_0 \vdash^k_{\omega+n+1} \Gamma$ *then* $(\Pi^1_2\text{-}BI)_0 \vdash^p_{\omega+1} \Gamma$ *where* $p = 2^k_n$.

Proof. Observe that $gr(A) < \omega + 1$ holds for every p.f. A of an axiom or non-logical rule of $(\Pi^1_2\text{-}BI)_0$. So the result follows by the standard cut-elimination procedure for sequent calculi (cf. [9]). □

Proposition 5.1. $(\Pi^1_2\text{-}BI)_0$ *proves* $WF(<) \rightarrow TI(<, F)$ *for any* $< \in \Pi^1_0$ *and* $F \in \Pi^1_2$.

Proof. Let $F(x)$ be $\forall Y \exists Z\, B(x, Y, Z)$ with B arithmetic. Then

$$(\Pi^1_2\text{-}BI)_0 \vdash \forall y\, [y < a \rightarrow \forall Y \exists Z\, B(y, Y, Z)],$$
$$\exists y \exists Y \forall Z\, [y < a \wedge \neg B(y, Y, Z)]$$

and

$$(\Pi^1_2\text{-}BI)_0 \vdash \exists Y \forall Z \neg B(a, Y, Z), \exists Z\, B(a, U, Z)$$

yield

$$(\Pi^1_2\text{-}BI)_0 \vdash C(a), \exists y \exists Y \forall Z\, [y < a \wedge \neg B(y, Y, Z)], \exists Z\, B(a, U, Z)$$

with

$$C(a) \equiv \forall y\, [y < a \rightarrow \forall Y \exists Z\, B(y, Y, Z)] \wedge \exists Y \forall Z \neg B(a, Y, Z).$$

Here we assume that a and U are 'fresh' variables. By $(\exists_1)$, we get

$$(\Pi^1_2\text{-}BI)_0 \vdash \exists x\, C(x), \exists y \exists Y \forall Z\, [y<a \wedge \neg B(y, Y, Z)], \exists Z\, B(a, U, Z). \quad (1)$$

Using the Π^1_2 bar induction rule on (1) and

$$(\Pi^1_2\text{-}BI)_0 \vdash \neg WF(<),\ WF(<), \tag{2}$$

we obtain

$$(\Pi^1_2\text{-}BI)_0 \vdash \neg WF(<),\ \exists x\, C(x),\ \exists Z\, B(a, U, Z);$$

thence, by $(\forall_2)$, $(\forall_1)$, and $(\vee)$, this becomes

$$(\Pi^1_2\text{-}BI)_0 \vdash \neg WF(<) \vee (\exists x\, C(x) \vee \forall x\, \forall Y\, \exists Z\, B(x, Y, Z)).$$

Therefore $(\Pi^1_2\text{-}BI)_0 \vdash WF(<) \rightarrow TI(<, F)$. (Note that $A \rightarrow B \equiv \neg A \vee B$.) $\square$

In the next section we shall embed $(\Pi^1_2\text{-}BI)_0$ into an infinitary calculus T^*. To handle this with optimal bounds, we have to resort to very well behaved derivations.

Definition 5.3. Let $\exists\Sigma^1_2$ be the collection of formulas of the form $\exists y\, \exists X\, \forall Y A(y, X, Y)$, where $A(0, U, V)$ is arithmetic.

A $(\Pi^1_2\text{-}BI)_0$ derivation $\mathscr{D} \vdash \Gamma$ is said to be *nice* if $\rho(\mathscr{D}) \leq \omega + 1$, and, for every Σ^1_2-formula A that serves as the minor formula of an inference $(\exists_1)$ in this derivation, either A is an element of Γ or A is never a side formula of an inference in $\mathscr{D}$.

Note that if none of the formulas occuring in a dervation $\mathscr{D}_0$ has a $\exists\Sigma^1_2$ subformula, then $\mathscr{D}_0$ is automatically nice. We say that a formula C is *essentially* Σ^1_2 (written *ess*-Σ^1_2) if

$$C \in \exists\Sigma^1_2 \cup \bigcup \{\Sigma^1_i : i \leq 2\} \cup \bigcup \{\Pi^1_j : j < 2\}.$$

Lemma 5.1. *Let Γ be a set of ess-Σ^1_2 formulas, $\Xi = \{\exists Z_1\, B_1(t_1, Z_1), \ldots, \exists Z_r\, B_r(t_r, Z_r)\} \subseteq \Sigma^1_2$, and $\Theta = \{\exists y_1\, \exists Z_1\, B_1(y_1, Z_1), \ldots, \exists y_r\, \exists Z_r\, B_r(y_r, Z_r)\}$.*

(1) *If $\mathscr{D} \vdash \Gamma, \Xi$, then we can find a nice $\mathscr{D}^+ \vdash \Gamma, \Theta$.*

(2) *If $\mathscr{D}_0 \vdash \Gamma$, then there is a nice $\mathscr{D}_0^+ \vdash \Gamma$.*

Proof. (1) By Theorem 5.1, we may assume $\rho(\mathscr{D}) \leq \omega + 1$. We proceed by induction on $|\mathscr{D}|$. If Γ, Ξ is an axiom, then so is Γ; hence Γ, Θ is an axiom. The derivation consisting merely of this axiom is of course nice.

Now suppose $0 < |\mathscr{D}|$. If neither a m.f. nor a p.f. of the last inference (l.i.) of $\mathscr{D}$ is Σ^1_2, then the assertion follows by using the induction hypothesis on the premises and reapplying the same inference, since this does not change the stock of Σ^1_2-formulas. Note that $\rho(\mathscr{D}) \leq \omega + 1$.

Next assume that a formula $\exists Z\, B(t, Z) \in \Sigma^1_2$ is a m.f. of the last inference of $\mathscr{D}$. Then this must be an instance of $(\exists_1)$ because of $\Gamma, \Xi \subset ess\text{-}\Sigma^1_2$. So the p.f. is of the form $\exists y\, \exists Z\, B(y, Z) \in \Gamma$, and the direct subderivation of $\mathscr{D}$ takes the form

$\mathcal{D}_0 \vdash \Lambda, \Xi'$ with $\Lambda \subseteq \Gamma$ and $\Xi' = \Xi, \exists Z\, B(y, Z)$. Applying the induction hypothesis we get a nice

$$\mathcal{D}^+ \vdash \Lambda, \Theta, \exists y\, \exists Z\, B(y, Z),$$

thence $\mathcal{D}^+ \vdash \Gamma, \Theta$.

Finally, suppose that $\exists Y\, C(Y) \in \Sigma^1_2$ is the p.f. of the last inference of $\mathcal{D}$. This then must be an instance of $(\exists_2)$. So there is a nice derivation $\mathcal{D}_0 \vdash \Gamma, \Xi, C(U)$ such that $\rho(\mathcal{D}_0) \leq \omega + 1$ and $|\mathcal{D}_0| < |\mathcal{D}|$. Inductively we find a nice derivation

$$\mathcal{D}_0^+ \vdash \Gamma, \Theta, C(U).$$

By use of $(\exists_2)$, we can continue $\mathcal{D}_0^+$ to a nice derivation of $\Gamma, \Theta, \exists Y C(Y)$. If $\exists Y\, C(Y) \in \Gamma$, then we are done. Otherwise, $\exists Y\, C(Y) \in \Xi$; thus an application of $(\exists_1)$ gives us a nice derivation $\mathcal{D}^+ \vdash \Gamma, \Theta$, since in this case $\exists Y\, C(Y)$ does not appear as a s.f. in $\mathcal{D}_0^+$.

(2) follows from (1) with $\Xi = \emptyset$. $\square$

6. The infinitary calculus T^*

The *formulas* of T^* arise from $\mathcal{L}^2$-formulas by replacing free numerical variables by numerals, i.e., terms of the form $0, 0', 0'', \ldots$. Especially, every formula A of T^* is an $\mathcal{L}_2$-formula, and thus $gr(A)$ is understood. A formula without second-order variables will be called constant. We are going to measure the length of derivations by ordinals. For technical reasons we are compelled to diverge from the ordinal notation system $OT(\vartheta)$ of Section 1. Instead we are going to use the set of ordinals $OT(\psi)$ of Section 3.

Definition 6.1. 1. A formula B is said to be *weak* if it belongs to $\Pi^1_0 \cup \Pi^1_1$.

2. Two closed terms s and t are said to be equivalent if they yield the same value when computed.

3. A formula is called constant if it contains no set variables. The truth or falsity of such a formula is understood with respect to the standard structure of the integers.

4. $\bar{0} := 0$, $\overline{m+1} := \bar{m}'$.

Definition 6.2. Inductive definition of $T^* \vdash^\alpha_\rho \Gamma$ for $\alpha \in OT(\psi)$ and $\rho < \omega + \omega$.

1. If A is a true constant prime formula or negated prime formula and $A \in \Gamma$, then $T^* \vdash^\alpha_\rho \Gamma$.

2. If Γ contains formulas $A(s_1, \ldots, s_n)$ and $\neg A(t_1, \ldots, t_n)$ of grade 0 or ω, where s_i and t_i $(1 \leq i \leq n)$ are equivalent terms, then $T^* \vdash^\alpha_\rho \Gamma$.

3. If $T^* \vdash^\beta_\rho \Gamma_i$ and $\beta \lhd \alpha$ hold for every premiss Γ_i of an inference $(\wedge)$, $(\vee)$,

$(\exists_1)$, $(\forall_2)$ or (Cut) with a cut formula having grade $<\rho$, and conclusion Γ, then $T^* \vdash_\rho^\alpha \Gamma$.

4. If $T^* \vdash_\rho^{\alpha_0} \Gamma,\ F(U)$ holds for some $\alpha_0 \lhd \alpha$ and a non-arithmetic formula $F(U)$ (i.e., $gr(F(U)) \geqslant \omega$), then $T^* \vdash_\rho^\alpha \Gamma,\ \exists X\, F(X)$.

5. (ω-rule). If $T^* \vdash_\rho^\beta \Gamma,\ A(\bar{m})$ is true for every $m < \omega$. $\forall x\, A(x) \in \Gamma$, and $\beta \lhd \alpha$, then $T^* \vdash_\rho^\alpha \Gamma$.

6. (Ω-rule). Let f be a fundamental function satisfying
   a. $\Omega \in \mathrm{dom}(f)$ and $f(\Omega) \unlhd \alpha$,
   b. $T^* \vdash_\rho^{f(0)} \Gamma,\ \forall X\, F(X)$, where $\forall X\, F(X) \in \Pi_1^1$, and
   c. $T^* \vdash_0^\beta \Xi,\ \forall X\, F(X)$ implies $T^* \vdash_\rho^{f(\beta)} \Xi, \Gamma$ for every set of weak formulas Ξ and $\beta < \Omega$.

Then $T^* \vdash_\rho^\alpha \Gamma$ holds.

Remark 6.1. The derivability relation $T^* \vdash_\rho^\alpha \Gamma$ is modelled upon the relation $PB^* \vdash_n^\alpha F$ of [1], the main difference being the sequent calculus setting instead of P- and N-forms and a different assignment of cut-degrees. The allowance for transfinite cut-degrees will enable us to deal with arithmetical comprehension.

Lemma 6.1. (1)$T^* \vdash_\delta^\alpha \Gamma \;\&\; \Gamma \subseteq \Delta \;\&\; \alpha \unlhd \beta \;\&\; \delta \leqslant \rho \;\Rightarrow\; T^* \vdash_\rho^\beta \Delta$.
(2) $T^* \vdash_\rho^\alpha \Gamma, A \wedge B \;\Rightarrow\; T^* \vdash_\rho^\alpha \Gamma, A \;\&\; T^* \vdash_\rho^\alpha \Gamma, B$.
(3) $T^* \vdash_\rho^\alpha \Gamma, A \vee B \;\Rightarrow\; T^* \vdash_\rho^\alpha \Gamma, A, B$.
(4) $T^* \vdash_\rho^\alpha \Gamma,\ F(t) \;\Rightarrow\; T^* \vdash_\rho^\alpha \Gamma,\ F(s)$ *if t and s are equivalent.*
(5) $T^* \vdash_\rho^\alpha \Gamma,\ \forall x\, F(x) \;\Rightarrow\; T^* \vdash_\rho^\alpha \Gamma,\ F(s)$ *for every term s.*
(6) *If* $T^* \vdash_\rho^\alpha \Gamma,\ \forall X\, G(X)$ *and* $gr(G(U)) \geqslant \omega$, *then* $T^* \vdash_\rho^\alpha \Gamma,\ G(U)$.

Proof. By induction on α. The inductions can be carried out straightforwardly. (5) requires (4). As to (6), observe that $\forall X\, G(X)$ cannot be the main formula of an axiom. □

Lemma 6.2. $T^* \vdash_0^{2\cdot\alpha} \Gamma,\ A(s_1, \ldots, s_k),\ \neg A(t_1, \ldots, t_k)$ *if* $\alpha \geqslant gr(A(s_1, \ldots, s_k))$ *and s_i and t_i are equivalent terms.*

Proof. By induction on $gr(A)$. □

Lemma 6.3. (1) $T^* \vdash_0^{2m} \neg(0 \in U),\ (\exists x)[x \in U \wedge \neg(x' \in U)],\ \bar{m} \in U$,
(2) $T^* \vdash_0^{\omega+5} \forall X\, [0 \in X \wedge \forall x\, (x \in X \rightarrow x' \in X) \rightarrow \forall x\, (x \in X)]$.

Proof. For (1) use induction on m. For a detailed proof see [4, 10.17].

(2) is an immediate consequence of (1) using Lemma 6.1(1), the ω-rule, $(\vee)$, and $(\forall_2)$.

Definition 6.3. For formulas $F(U)$ and $A(a)$, $F(A)$ denotes the result of replacing each occurrence of the form $(e \in U)$ in $F(U)$ by $A(e)$. The expression $F(A)$ is a formula if the bound variables in $A(a)$ are chosen in an appropriate way, in particular, if $F(U)$ and $A(a)$ have no bound variables in common.

Lemma 6.4. *Let $\Delta(U) = \{F_1(U), \ldots, F_k(U)\}$ be a set of weak formulas such that U doesn't occur in $\forall X F_i(X)$ $(1 \leq i \leq k)$. For an arbitrary formula $A(a)$ we then have:*

$$T^* \vdash_0^\alpha \Delta(U) \Rightarrow T^* \vdash_0^{\Omega+\alpha} \Delta(A).$$

Proof. By induction on α. Suppose $\Delta(U)$ is an axiom. Then either $\Delta(A)$ is an axiom too, or $T^* \vdash_0^{\omega+\omega} \Delta(A)$ can be obtained through use of Lemma 6.2. Therefore $T^* \vdash_0^{\Omega+\alpha} \Delta(A)$ by Lemma 6.1(1). If $T^* \vdash_0^\alpha \Delta(U)$ is the result of an inference, then this inference must be different from $(\exists_2)$, *(Cut)*, and *(Ω-rule)*. Therefore the assertion follows easily from the induction hypothesis. □

Lemma 6.5. *Let Γ, $\forall X\, (FX)$ be a set of weak formulas. If $T^* \vdash_0^\alpha \Gamma, \forall X F(X)$ and $\alpha < \Omega$, then $T^* \vdash_0^\alpha \Gamma, F(U)$.*

Proof. By induction on α. Note that $\forall X F(X)$ cannot be a principal formula of an axiom, since $\exists X \neg F(X)$ does not surface in such a derivation. Also, due to $\alpha < \Omega$, the derivation doesn't involve instances of the Ω-rule. Therefore the proof is straightforward. □

The role of the *(Ω-rule)* in our calculus T^* is enshrined in the next lemma.

Lemma 6.6. *$T^* \vdash_0^{\Omega \cdot 2} \exists X F(X), \neg F(A)$ for every arithmetic formula $F(U)$ and arbitrary formula $A(a)$.*

Proof. Let $f(\alpha) := \Omega + \alpha$ with $dom(f) := \{\alpha \in OT(\psi): \alpha \leq \Omega\}$. Then

$$T^* \vdash_0^{f(0)} \forall X \neg F(X), \exists X F(X), \neg F(A) \tag{1}$$

according to Lemma 6.2. For $\alpha < \Omega$ and every set of weak formulas Θ, we have by Lemmata 6.4 and 6.5,

$$T^* \vdash_0^\alpha \Theta, \forall X \neg F(X) \Rightarrow T^* \vdash_0^{f(\alpha)} \Theta, \neg F(A).$$

Therefore, by Lemma 6.1(1),

$$T^* \vdash_0^\alpha \Theta, \forall X \neg F(X) \Rightarrow T^* \vdash_0^{f(\alpha)} \Theta, \exists X F(X), \neg F(A). \tag{2}$$

The assertion now follows from (1) and (2) by the Ω-rule. □

7. The reduction procedure for T^*

Lemma 7.1. *Let C be a formula of grade ρ. Suppose C is a prime formula or of either form $\exists X H(X)$, $\exists x G(x)$ or $A \vee B$. Let $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_k}$ with $\delta \leq \omega^{\alpha_k} \leq$*

$\cdots \leqslant \omega^{\alpha_1}$. *Then we have*

$$T^* \vdash_\rho^\alpha \Delta, \neg C \;\&\; T^* \vdash_\rho^\delta \Gamma, C \Rightarrow T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma.$$

Proof. By induction on δ.

1. Let Γ, C be an axiom. Then there are three cases to consider.

1.1. Γ is an axiom. Then so is Δ, Γ. Hence $T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma$.

1.2. C is a true constant prime formula or negated prime formula. A straight-forward induction on α then yields $T^* \vdash_\rho^\alpha \Delta$, and thus $T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma$ by 6.1(1).

1.3. $C \equiv A(s_1, \ldots, s_n)$ and Γ contains a formula $\neg A(t_1, \ldots, t_n)$ where s_i and t_i are equivalent terms. From $T^* \vdash_\rho^\alpha \Delta, \neg A(s_1, \ldots, s_n)$ one receives $T^* \vdash_\rho^\alpha \Delta, \neg A(t_1, \ldots, t_n)$ by use of Lemma 6.1(4). Thence $T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma$ follows by use of Lemma 6.1(1), since $\neg A(t_1, \ldots, t_n) \in \Gamma$.

2. Suppose $C \equiv A \vee B$ and $T^* \vdash_\rho^{\delta_0} \Gamma, C, A_0$ with $A_0 \in \{A, B\}$ and $\delta_0 \lhd \delta$. Inductively we get

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, A_0. \tag{1}$$

Next use Lemma 6.1(2) on $T^* \vdash_\rho^\alpha \Delta, \neg A \wedge \neg B$ to obtain

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, \neg A_0. \tag{2}$$

Whence use a cut on (1) and (2) to get the assertion.

3. Suppose $C \equiv \exists x\, G(x)$ and $T^* \vdash_\rho^{\delta_0} \Gamma, C, G(t)$ with $\delta_0 \lhd \delta$. Inductively we get

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, G(t). \tag{3}$$

By Lemma 6.1(1),(5), we also get

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, \neg G(t); \tag{4}$$

thus (3) and (4) yield $T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma$ by (Cut).

4. Suppose the last inference was $(\exists_2)$ with p.f. C. Then $C \equiv \exists X\, H(X)$ and $T^* \vdash_\rho^{\delta_0} \Gamma, C, H(U)$ for some $\delta_0 \lhd \delta$ and $gr(H(U)) \geqslant \omega$. Inductively we get

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, H(U). \tag{5}$$

By Lemma 6.1(1),(6) we also get

$$T^* \vdash_\rho^{\alpha+\delta_0} \Delta, \Gamma, \neg H(U). \tag{6}$$

From (5) and (6) we obtain

$$T^* \vdash_\rho^{\alpha+\delta} \Delta, \Gamma.$$

5. Let $T^* \vdash_\rho^\delta \Gamma, C$ be derived by the Ω-rule with fundamental function f. Then the assertion follows from the I.H. by the Ω-rule using the fundamental function $\alpha + f$.

6. In the remaining cases the assertion follows from the I.H. used on the premises and by reapplying the same inference. $\square$

Lemma 7.2. $T^* \vdash^{\alpha}_{\eta+1} \Gamma \Rightarrow T^* \vdash^{\omega^\alpha}_{\eta} \Gamma$.

Proof. By induction on α. We only treat the crucial case when $T^* \vdash^{\alpha_0}_{\eta+1} \Gamma, D$ and $T^* \vdash^{\alpha_0}_{\eta+1} \Gamma, \neg D$, where $\alpha_0 \lhd \alpha$, and $gr(D) = \eta$. Inductively this becomes $T^* \vdash^{\omega^{\alpha_0}}_{\eta} \Gamma, D$ and $T^* \vdash^{\omega^{\alpha_0}}_{\eta} \Gamma, \neg D$. Since D or $\neg D$ must be one of the forms exhibited in Lemma 7.1, we obtain $T^* \vdash^{\omega^{\alpha_0}+\omega^{\alpha_0}}_{\eta} \Gamma$ by Lemma 7.1. As $\omega^{\alpha_0} + \omega^{\alpha_0} \lhd \omega^\alpha$, we can use Lemma 6.1(1) to get the assertion. □

Theorem 7.1 (Collapsing Theorem). *If Γ is a set of weak formulas and $\alpha \in C(\alpha)$, then we have*

$$T^* \vdash^{\alpha}_{\omega} \Gamma \Rightarrow T^* \vdash^{\psi\alpha}_{0} \Gamma.$$

Proof. By induction on α. Observe that for $\beta < \delta < \Omega$, we always have $\beta \lhd \delta$.

1. If Γ is an axiom, then the assertion is trivial.

2. Let $T^* \vdash^{\alpha}_{\omega} \Gamma$ be the result of an inference other than (*Cut*) and Ω-rule. Then we have $T^* \vdash^{\alpha_0}_{\omega} \Gamma_i$ with $\alpha_0 \lhd \alpha$ and Γ_i being the ith premiss of that inference. $\alpha_0 \lhd \alpha$ and $\alpha \in C(\alpha)$ imply $\alpha_0 \in C(\alpha_0)$ and $\psi\alpha_0 \lhd \psi\alpha$ by Lemma 4.5. Therefore $T^* \vdash^{\psi\alpha_0}_{0} \Gamma_0$ by the I.H., hence $T^* \vdash^{\psi\alpha}_{0} \Gamma$ by reapplying the same inference.

3. Suppose $T^* \vdash^{\alpha}_{\omega} \Gamma$ results by the Ω-rule with respect to Π^1_1-formula $\forall X\, F(X)$ and a fundamental function f. Then $\Omega \in dom(f)$ and $f(\Omega) \unlhd \alpha$. Also

$$T^* \vdash^{f(0)}_{\omega} \Gamma, \forall X\, F(X), \tag{1}$$

and, for every set of weak formulas Ξ and $\beta < \Omega$,

$$T^* \vdash^{\beta}_{0} \Xi, \forall X\, F(X) \Rightarrow T^* \vdash^{f(\beta)}_{\omega} \Xi, \Gamma. \tag{2}$$

From $\Omega \in dom(f)$ we get $f(0) \lhd f(\Omega)$ by (F1); thus $f(\Omega) \unlhd \alpha$ yields $f(0) \in C(f(0))$ and $f(\Omega) \in C(f(\Omega))$ using Lemma 4.5. Therefore the I.H. used on (1) supplies us with $T^* \vdash^{\psi(f(0))}_{0} \Gamma, \forall X\, F(X)$. Hence with $\Xi = \Gamma$ we get

$$T^* \vdash^{f(\psi(f(0)))}_{\omega} \Gamma \tag{3}$$

from (2). Now Corollary 4.3 ensures that $f(\beta) \lhd f(\Omega)$, where $\beta = \psi(f(0))$; hence $f(\beta) \in C(f(\beta))$ by Lemma 4.5 since $f(\Omega) \in C(f(\Omega))$. So using the I.H. on (3), we obtain

$$T^* \vdash^{\psi(f(\beta))}_{0} \Gamma, \tag{4}$$

thus $T^* \vdash^{\psi\alpha}_{0} \Gamma$ as $f(\beta) \lhd \alpha$.

4. Suppose $T^* \vdash^{\alpha_0}_{\omega} \Gamma, A$ and $T^* \vdash^{\alpha_0}_{\omega} \Gamma, \neg A$, where $\alpha_0 \lhd \alpha$ and $gr(A) < \omega$. Inductively we then get $T^* \vdash^{\psi\alpha_0}_{0} \Gamma, A$ and $T^* \vdash^{\psi\alpha_0}_{0} \Gamma, \neg A$. Let $gr(A) = n - 1$. Then (*Cut*) yields

$$T^* \vdash^{\beta_1}_{n} \Gamma \tag{5}$$

with $\beta_1 = (\psi\alpha_0) + 1$. Applying Lemma 7.2, we get $T^* \vdash^{\omega^{\beta_1}}_{n-1} \Gamma$, and by repeating this process we arrive at

$$T^* \vdash_0^{\beta_n} \Gamma,$$

where $\beta_{k+1} := \omega^{\beta_k}$ $(1 \leq k < n)$. Since $\psi\alpha_0 < \psi\alpha$, we have $\beta_n < \psi\alpha$; thus $T^* \vdash_0^{\psi\alpha} \Gamma$. $\square$

8. The refined reduction lemma

We would like to define nice derivations also in the context of T^*. A T^*-derivation is a tuple consisting of its direct subderivations (d.s.), the set of minor formulas (m.f.), the set of principal formulas (p.f.) (possibly empty), the ordinal bounds for the length and cut-rank, and a symbol indicating the last inference (l.i.). If there were not the Ω-rule, we needn't to be specific about this. We will write $\mathscr{D} \vdash_\rho^\alpha \Gamma$ if $\mathscr{D}$ is a derivation witnessing $\vdash_\rho^\alpha \Gamma$. The tuple

$$\mathscr{D} = \langle \mathscr{D}_0, \mathscr{T}, f, \forall X\, H(X), \alpha, \rho, \Omega \rangle$$

is said to be a nice T^*-derivation if: $\mathscr{D}_0 \vdash_\rho^\alpha \Gamma$, $\forall X\, H(X)$ is nice; f is a fundamental function with $\Omega \in dom(f)$ and $f(\Omega) \trianglelefteq \alpha$; $\forall X\, H(X) \in \Pi_1^1$; and for every set of weak formulas Ξ and $\beta < \Omega$ the following is valid:

If $\mathscr{D}'$ is a T^*-derivation satisfying $\mathscr{D}' \vdash_0^\beta \Xi, \forall X\, H(X)$, then $\mathscr{T}(\mathscr{D}')$ is a nice T^*-derivation satisfying $\mathscr{T}(\mathscr{D}') \vdash_\rho^{f(\beta)} \Xi, \Gamma$. (So $\mathscr{T}$ is a function on proofs.) Nice derivations allow us to improve on the Reduction Lemma 7.1.

Lemma 8.1 (The Refined Reduction Lemma). *Let $B \equiv \exists y\, \exists Z\, A(y, Z)$ be $\exists\Sigma_2^1$. Let $\alpha = \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \geq \alpha_1 \geq \cdots \geq \alpha_n$, $\beta = \omega^{\beta_1} + \cdots + \omega^{\beta_k} \geq \beta_1 \geq \cdots \geq \beta_k$ such that $\beta_1 \leq \alpha_n$. Suppose that $\mathscr{D}' \vdash_{\omega+1}^\alpha \Gamma, \neg B$ and $\mathscr{D} \vdash_{\omega+1}^\beta \Lambda, B$ are nice T^*-derivations, where $\Gamma, \Lambda \subseteq ess\text{-}\Sigma_2^1$. There there is a nice T^*-derivation $\mathscr{D}^* \vdash_{\omega+1}^{\alpha+\beta} \Gamma, \Lambda$.*

Proof. By induction on β. If the last inference (l.i.) does not have a Σ_2^1 or $\exists\Sigma_2^1$ principal formula (p.f.), the assertion follows by using the I.H. on the direct subderivations (d.s.) and reapplying the same inference. (If this is the Ω-rule with fundamental function f, then the new fundamental function will be $\alpha + f$.) Note that such an inference does not change the stock of 'critical' formulas.

If the last inference is $(\exists_2)$ with Σ_2^1 p.f. A, then the d.s. $\mathscr{D}_0$ might not be nice. But this problem can be overcome by adding A as a side formula throughout the derivation $\mathscr{D}_0$, which gives rise to a nice derivation $\mathscr{D}_1$. So we can apply the I.H. on $\mathscr{D}_1$ and subsequently $(\exists_2)$ to get the assertion.

Now suppose that the p.f. of the l.i. is $E \equiv \exists x\, \exists Z\, A(x, Z)$ and $E \in \exists\Sigma_2^1$. Then the d.s. of $\mathscr{D}$ has the form $\mathscr{D}_0 \vdash_{\omega+1}^{\beta_0} \Lambda_0$, C, where $\beta_0 \lhd \beta$, $C \equiv \exists Z\, A(t, Z)$, and $\Lambda_0 \subseteq \Lambda, B$. Note that $\mathscr{D}_0$ is also nice. Suppose $C \in \Lambda$. The I.H. provides us with a nice derivation

$$\mathscr{D}^+ \vdash_{\omega+1}^{\alpha+\beta_0} \Lambda_0 \backslash \{B\}, \Gamma, C.$$

So we get a nice derivation $\mathscr{D}^* \vdash_{\omega+1}^{\alpha+\beta} \Gamma, \Lambda$ by weakening. If $C \notin \Lambda$, then the niceness of $\mathscr{D}$ implies that the l.i. of $\mathscr{D}_0$ is $(\exists_2)$ with p.f. C. So the d.s. $\mathscr{D}_1$ of $\mathscr{D}_0$ then has the form $\mathscr{D}_1 \vdash_{\omega+1}^{\beta_1} \Lambda_1, A(t, U)$ with $\Lambda_1 \subseteq \Lambda_0$ and $\beta_1 \lhd \beta_0$. Using the I.H. on $\mathscr{D}_1$, we find a nice derivation

$$\mathscr{D}_2 \vdash_{\omega+1}^{\alpha+\beta_1} \Lambda_1 \backslash \{B\}, \Gamma, A(t, U).$$

Suppose that $E \equiv B$. Then $\neg B \equiv \forall y\, \forall Z\, \neg A(y, Z)$. Using inversion (Lemma 6.1(2) and 6.1(6)) on $\mathscr{D}'$, we get a derivation

$$\mathscr{D}_3 \vdash_{\omega+1}^{\alpha} \Gamma, \neg A(t, U),$$

which is also nice, since it has the same stock of Σ_2^1 minor and side formulas as $\mathscr{D}'$. Using (*Cut*) on $\mathscr{D}_2$ and $\mathscr{D}_3$ (and weakening) we get a nice derivation

$$\mathscr{D}^* \vdash_{\omega+1}^{\alpha+\beta} \Gamma, \Lambda.$$

Finally, suppose E is different from B. Then $E \in \Gamma$. Then apply $(\exists_2)$ followed by $(\exists_1)$ to $\mathscr{D}_2$, to get a nice derivation $\mathscr{D}^* \vdash_{\omega+1}^{\alpha+\beta} \Lambda_1 \backslash \{B\}, \Gamma, E$ (note that $\alpha + \beta_1 \lhd \alpha + \beta_0 \lhd \alpha + \beta$). As $\Lambda_1 \backslash \{B\} \subseteq \Lambda$ and $E \in \Gamma$, the result follows by weakining. $\square$

9. Embedding $(\Pi_2^1\text{-}BI)_0$ into T^*

The objective of this section is to embed $(\Pi_2^1\text{-}BI)_0$ into T^*, so as to obtain an upper bound for the proof-theoretic ordinal of $(\Pi_2^1\text{-}BI)_0$ by putting to use results of the previous section. The treatment of the Π_2^1 bar induction rule requires the following lemma.

Lemma 9.1. *Let $E(a)$ and $F(a, U, V)$ be arithmetic formulas. If Λ is an arbitrary set of T^* formulas and $T^* \vdash_{\omega+1}^{\delta} \Lambda, \exists y\, [E(y) \wedge \forall Y\, \exists Z\, F(y, Y, Z)]$, then*

$$T^* \vdash_{\omega+1}^{\Omega \cdot 2+\delta \cdot 5} \Lambda, \exists y\, \forall Y\, \exists Z\, [E(y) \wedge F(y, Y, Z)].$$

Proof. By induction on δ. Let $G_0 \equiv \exists y\, [E(y) \wedge \forall Y\, \exists Z\, F(y, Y, Z)]$. If G_0 is not the p.f. of the last inference, then we get the desired derivation by employing the I.H. on the premises and reapplying the same inference.

Now let G_0 be the p.f. of the last inference, which must be $(\exists_1)$. Hence $T^* \vdash_{\omega+1}^{\delta_0} \Lambda, G_0, E(t) \wedge \forall Y\, \exists Z\, F(y, Y, Z)$ for some $\delta_0 \lhd \delta$. Let G be $\exists y\, \forall Y\, \exists Z\, [E(y) \wedge F(y, Y, Z)]$. The I.H. yields $T^* \vdash_{\omega+1}^{\alpha} \Lambda, G, E(t) \wedge \forall Y\, \exists Z\, F(t, Y, Z)$, where $\alpha = \Omega \cdot 2 + \delta_0 \cdot 5$. Hence, by Lemma 6.1(2), we get

$$T^* \vdash_{\omega+1}^{\alpha} \Lambda, G, E(t) \tag{1}$$

and

$$T^* \vdash_{\omega+1}^{\alpha} \Lambda, G, \forall Y\, \exists Z\, F(t, Y, Z). \tag{2}$$

Using Lemma 6.1(6) on (2), we get

$$T^* \vdash^{\alpha}_{\omega+1} \Lambda, G, \exists Z\, F(t, V, Z). \tag{3}$$

From Lemma 6.6, followed by Lemma 6.1(3), we obtain

$$T^* \vdash^{\Omega \cdot 2}_{\omega+1} \exists Z\, [E(t) \wedge F(t, V, Z)], \neg E(t), \neg F(t, V, U).$$

By $(\forall_2)$, this becomes

$$T^* \vdash^{\Omega \cdot 2+1}_{\omega+1} \exists Z\, [E(t) \wedge F(t, U, Z)], \neg E(t), \forall Z \neg F(t, V, Z). \tag{4}$$

Applying (Cut) to (3) and (4) gives

$$T^* \vdash^{\alpha+2}_{\omega+1} \Lambda, G, \exists Z\, [E(t) \wedge F(t, V, Z)], \neg E(t);$$

hence

$$T^* \vdash^{\alpha+4}_{\omega+1} \Lambda, G, \neg E(t) \tag{5}$$

using $(\forall_2)$ and $\exists_1$). Finally employ (Cut) on (1) and (2) to obtain $T^* \vdash^{\alpha+5}_{\omega+1} \Lambda, G$. As $\alpha + 5 \trianglelefteq \Omega \cdot 2 + \delta \cdot 5$, this finishes the proof. $\square$

Theorem 9.1. *Let Γ be a set of ess-Σ^1_2 formulas. Suppose that Γ^* results from Γ by replacing all the free variables in formulas of Γ with numerals. Then*

$$(\Pi^1_2 - BI)_0 \vdash^{k}_{\omega+1} \Gamma \Rightarrow T^* \vdash^{\Omega^{f(k)}}_{\omega+1} \Gamma^*,$$

where $f(k) := (k+1) \cdot 3$.

Proof. By induction on k.

Using Lemma 5.1(2), we can assume that we have a nice derivation of Γ. During this proof we shall also ensure that the corresponding T^*-derivation will be nice.

1. If Γ is an axiom $(Ax1)$ or $(Ax2)$, then Γ^* is also an axiom of T^*.

2. If Γ is an axiom (IA), then this follows from Lemma 6.3(2).

3. Suppose Γ is an axiom $(\Pi^1_0\text{-}CA)$, i.e., Γ contains a formula $(\exists Z)(\forall x)[x \in Z \leftrightarrow A(x)]$, with A arithmetic. Setting $F(U) := (\forall x)[x \in U \leftrightarrow A^*(x)]$, we obtain

$$T^* \vdash^{\omega+\omega}_{0} F(A^*) \tag{1}$$

using Lemma 6.2. By Lemma 6.6 we have

$$T^* \vdash^{\Omega \cdot 2}_{0} \exists Z\, F(Z), \neg F(A^*). \tag{2}$$

Thus (1) and (2) yield (since $gr(F(A^*)) < \omega$) $T^* \vdash^{\Omega^{f(0)}}_{\omega} \exists Z\, F(Z)$, thence $T^* \vdash^{\Omega^{f(0)}}_{\omega+1} \Gamma^*$.

Observe that in the cases, we have been considering so far, the infinitary derivations didn't contain inferences with Σ^1_2 m.f. or p.f.; so they are automatically nice.

4. Suppose $(\Pi^1_2\text{-}BI)_0 \vdash^{k}_{\omega+n-1} \Gamma$ is the result of an inference $(\exists_2)$ with principal formula $\exists X\, H(X) \in \Gamma$ having grade ω. Then $(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Gamma_0, H(U)$ for some

$k_0 < k$ and $\Gamma_0 \subseteq \Gamma$. By I.H., we then have a nice T^*-derivation $\mathscr{D}_0 \vdash^{\Omega^{f(0)}}_{\omega+1} \Gamma^*_0$, $H^*(U)$. The $(\exists_2)$-rule of T^* is not available in this situation since $gr(H^*(U)) < \omega$. We have to resort to Lemma 6.6 to get a nice

$$\mathscr{D}_1 \vdash^{\Omega \cdot 2}_0 \exists X\, H^*(X), \neg H^*(U).$$

Applying (Cut) gives us a nice

$$\mathscr{D}_2 \vdash^{\Omega^{f(k)}}_{\omega+1} \Gamma^*.$$

5. Suppose $(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Gamma, A$ and $(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Gamma, \neg A$ for some $k_0 < k$ and $gr(A) < \omega + 1$. Inductively we then get

$$T^* \vdash^{\Omega^{f(k_0)}}_{\omega+1} \Gamma^*, A^* \quad \text{and} \quad T^* \vdash^{\Omega^{f(k_0)}}_{\omega+1} \Gamma^*, \neg A^*.$$

So, by (Cut), we obtain $T^* \vdash^{\Omega^{f(k)}}_{\omega+1} \Gamma^*$ niceness preserving.

6. Let $(\Pi^1_2\text{-}BI)_0 \vdash^{k}_{\omega+1} \Gamma$ be the result of $(\forall_1)$. Then $(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Gamma_0, F(a)$ for some $k_0 < k$ and $\Gamma_0, \forall x\, F(x) = \Gamma$. Inductively we get $T^* \vdash^{\Omega^{f(k_0)}}_{\omega+1} \Gamma^*_0, F^*(\bar{n})$ for every n. Thus the assertion follows by the ω-rule. Since $(\forall_1)$ does not violate niceness in the original derivation, the ω-rule won't do this in the infinitary derivation.

7. Let $(\Pi^1_2\text{-}BI)_0 \vdash^{k}_{\omega+1} \Gamma$ be the result of a logical inference other than the ones already treated. Then the assertion follows using the I.H. on the premises and subsequently reapplying the same inference in T^*. This of course preserves niceness.

8. Finally, we have to deal with the Π^1_2 bar induction rule. So suppose $\Gamma = \Lambda, \exists Z\, B(b, V, Z)$,

$$(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Lambda, WF(<), \tag{3}$$

and

$$(\Pi^1_2\text{-}BI)_0 \vdash^{k_0}_{\omega+1} \Lambda, \exists y\, \exists Y\, \forall Z\, [y < a \wedge \neg B(y, Y, Z], \exists Z\, B(a, U, Z), \tag{4}$$

where $k_0 < k$ and a, U are 'fresh' variables. Put

$$C(\bar{m}) :\equiv \exists y\, \exists Y\, \forall Z[y < \bar{m} \wedge \neg B(y, Y, Z)]^*,$$

and

$$A(\bar{m}) :\equiv \forall Y\, \exists Z\, B(\bar{m}, Y, Z)^*.$$

By I.H. we then have, for $\beta = \Omega^{f(k_0)}$ and for each $m < \omega$, a nice T^*-derivation

$$\mathscr{D}_m \vdash^{\beta}_{\omega+1} \Lambda^*, C(\bar{m}), \exists Z\, B(\bar{m}, U, Z)^*;$$

thus, using $(\forall_2)$, we obtain a nice derivation

$$\mathscr{D}^1 \vdash^{\alpha_0}_{\omega+1} \Lambda^*, C(\bar{m}), A(\bar{m}) \tag{5}$$

with $\alpha_0 = \Omega^{f(k_0)+1}$ (note that $\exists Z\, B(\bar{m}, U, Z)^* \notin \Sigma^1_2$), and also, by the I.H. used on (3), we get a nice

$$\mathscr{D}^2 \vdash^{\alpha_0}_{\omega+1} \Lambda^*, WO(<)^*. \tag{6}$$

Moreover, Lemma 6.6 provides us with a nice

$$\mathscr{D}' \vdash_0^{\Omega\cdot 2} \neg\forall X\, TI(<^*, X),\ TI(<^*, A),$$

simply because the formulas occuring in this derivation do not have $\exists\Sigma_2^1$ sub-formulas. By Lemma 6.1(3) we get a nice

$$\mathscr{D}^3 \vdash_0^{\Omega\cdot 2} \neg\forall X\, TI(<^*, X),\ \exists z\,[\neg C_0(z) \wedge \neg A(z)],\ \forall z\, A(z), \tag{7}$$

where

$$C_0(b) \equiv \forall y\,[y < b \rightarrow \exists Y\, \forall Z\, B(y, Y, Z)]^*.$$

Claim. *If $\beta \leqslant \Omega \cdot 2$, Ξ a set of formulas which does not have Σ_2^1 subformulas, and $\mathscr{D}^*$ is a nice derivation such that*

$$\mathscr{D}^* \vdash_0^{\beta} \Xi,\ \exists z\,[\neg C_0(z) \wedge \neg A(z)],\ \forall z\, A(z),$$

then, for all $m < \omega$, there is a nice

$$\mathscr{D}_m^* \vdash_{\omega+1}^{g(\beta)} \Xi,\ \Gamma^*,\ A(\bar{m}),$$

where

$$g(\beta) = \alpha_0 \cdot \omega^{\beta} = \Omega^{f(k_0)+1} \cdot \omega^{\beta}.$$

We prove the Claim by induction on β. Put

$$D := \exists x\,[\neg C(z) \wedge \neg A(z)].$$

If neither D nor $\forall z\, A(z)$ is a principal formula of the last inference, then the assertion is easily obtained by the induction hypothesis. Suppose now that $\forall z\, A(z)$ is the principal formula. Then there is some $\beta_0 \lhd \beta$ such that for all $m < \omega$ there is a nice $\mathscr{D}_m \vdash_0^{\beta_0} \Xi,\ D,\ \forall z\, A(z),\ A(\bar{m})$. Inductively we then can pick a nice $\mathscr{D}'_m \vdash_{\omega+1}^{g(\beta_0)} \Xi,\ \Gamma^*,\ A(\bar{m})$, thus $\mathscr{D}_m^* \vdash_{\omega+1}^{g(\beta)} \Xi,\ \Gamma^*,\ A^*(\bar{m})$ for some nice $\mathscr{D}_m^*$ (note that $g(\beta_0) \lhd g(\beta)$).

Now let D be the principal formula. Then there are $\beta_0 \lhd \beta$, a term t, and a nice

$$\mathscr{D}_0^* \vdash_0^{\beta_0} \Xi,\ \forall z\, A(z),\ D,\ \neg C_0(t) \wedge \neg A(t).$$

Let n be the value of t. Using the I.H. and Lemma 6.1(4), we get a nice

$$\mathscr{D}_1^* \vdash_0^{g(\beta_0)} \Xi,\ \Gamma^*,\ A(\bar{m}),\ \neg C_0(\bar{n}) \wedge \neg A(\bar{n}).$$

By Lemma 6.1(2) we therefore find nice $\mathscr{D}_2^*$ and $\mathscr{D}_3^*$ such that

$$\mathscr{D}_2^* \vdash_{\omega+1}^{g(\beta_0)} \Xi,\ \Gamma^*,\ A(\bar{m}),\ \neg C_0(\bar{n}) \tag{8}$$

and

$$\mathscr{D}_3^* \vdash_{\omega+1}^{g(\beta_0)} \Xi,\ \Gamma^*,\ A(\bar{m}),\ \neg A(\bar{n}) \tag{9}$$

We would like to replace

$$\neg C_0(\bar{n}) \equiv \exists y\,[y <^* \bar{n} \wedge \forall Y\, \exists Z\, \neg B(y, Y, Z)^*]$$

with

$$\neg C(\bar{n}) \equiv \exists y\, \forall Y\, \exists Z\, [y <^* \bar{n} \wedge \neg B(y, Y, Z)^*].$$

This can be done using Lemma 9.1. By inspection of the proof of 9.1, one readily verifies that this transformation leads from nice derivations to nice derivations. So from (8) we obtain a nice $\mathscr{D}_4^*$ satisfying

$$\mathscr{D}_4^* \vdash_{\omega+1}^{\eta} \Xi, \Gamma^*, A(\bar{m}), \neg C(\bar{n}), \tag{10}$$

where $\eta := g(\beta_0) \cdot 4$. If we now use the Refined Reduction Lemma 8.1 on (5) and (10), we obtain a nice

$$\mathscr{D}_5^* \vdash_{\omega+1}^{\eta+\alpha_0} \Xi, \Gamma^*, A(\bar{m}), A(\bar{n}). \tag{11}$$

Note that $A(\bar{n})$ is not the minor formula of $(\exists_1)$ in $\mathscr{D}_3^*$ or $\mathscr{D}_5^*$ ((9) and (11)). Thence, by using the Reduction Lemma 7.1 on (9) and (11), we arrive at a nice

$$\mathscr{D}_6^* \vdash_{\omega+1}^{\eta+\alpha_0+\alpha_0} \Xi, \Gamma^*, A(\bar{m}).$$

Since $\eta + \alpha_0 + \alpha_0 = \Omega^{f(k_0)} \cdot (\omega^{\beta_0} \cdot 4 + 2) \lhd \Omega^{f(k_0)} \cdot \omega^{\beta} = g(\beta)$, this furnishes proof of the Claim. Letting $\Xi = \{\neg WF(<^*)\}$ and $\beta = \Omega \cdot 2$, the Claim, used on (7), yields a nice

$$\mathscr{D}_7^* \vdash_{\omega+1}^{\eta} \neg WF(<^*), \Gamma^*, A(\bar{m}), \tag{12}$$

where $\eta = \Omega^{f(k_0)+2}$. Finally, applying (Cut) to (6) and (12), will provide us with a

$$\mathscr{D}_m^* \vdash_{\omega+1}^{\Omega^{f(k)}} \Gamma^*, A(\bar{m}) \tag{13}$$

for all $m < \omega$. Since $A(\bar{m}) \equiv \forall Y\, \exists Z\, B^*(\bar{m}, Y, Z)$ and $\exists Z\, B^*(\bar{m}, V, Z) \in \Gamma^*$ for some $m < \omega$, the assertion follows from (13) and Lemma 6.1(6). □

Corollary 9.1. *Let A be a Π_1^1 sentence such that $(\Pi_2^1 - BI)_0 \vdash A$, then there is some $\alpha < \Omega^{\Omega^\omega}$ such that $T^* \vdash_0^{\psi\alpha} \Gamma$.*

Proof. Employing Theorem 9.1, we find an $n < \omega$ such that $T^* \vdash_{\omega+1}^{\Omega^n} A$; thus $T^* \vdash_{\omega}^{\omega^{\Omega^n}} A$ by Lemma 7.2. Theorem 7.1 then yields $T^* \vdash_0^{\psi(\omega^{\Omega^n})} A$. □

Corollary 9.2. *Let $<$ be an arithmetic well-ordering such that$(\Pi_2^1$-$BI)_0 \vdash WF(<)$. Then the order-type of $<$ is less than $\psi\Omega^{\Omega^\omega}$.*

Proof. From a cut-free proof of $WF(<)$ in T^* of length less than α one can extract that the order-type of $<$ is majorized by 2^α (cf. [4, 13.10] or [8]). □

Corollary 9.3. *$(\Pi_2^1$-$BI)_0$ does not prove $\forall n\, KT(n)$.*

Proof. Immediate by Theorem 1.1, Corollary 9.2, and Corollary 3.2. □

10. A well-ordering proof in $(\Pi^1_2\text{-}BI)_0^-$

By Corollary 9.2 and $\psi\Omega^{\Omega^\omega} \leq \vartheta\Omega^\omega$ (see Corollary 3.2), we know that the well-foundedness of the primitive recursive well-ordering $<\restriction\vartheta\Omega^\omega$ cannot be proved in $(\Pi^1_2\text{-}BI)_0$. This section is aimed at showing $(\Pi^1_2\text{-}BI)_0^- \vdash WF(<\restriction\vartheta\Omega^n)$ for any (meta) n. In the sequel $\alpha, \beta, \gamma, \delta, \ldots$ are supposed to range over $OT(\vartheta)$. $<$ will be used to denote the ordering on $OT(\vartheta)$. We are going to work informally in second-order arithmetic. Especially, variables $X, Y, Z, \ldots$ are ranging over subsets of $\mathbb{N}$. Note also that we consider $OT(\vartheta)$ to be a subset of $\mathbb{N}$. The proofs require some terminology.

Definition 10.1. 1. $Acc := \{\alpha < \Omega : WF(<\restriction\alpha)\}$,
2. $M := \{\alpha : E_\Omega(\alpha) \subseteq Acc\}$,
3. $\alpha <_\Omega \beta \;:\Leftrightarrow\; \alpha, \beta \in M \wedge \alpha < \beta$.

Remark. Acc is a Π^1_1-definable class. Therefore M and $<_\Omega$ are also Π^1_1.

Lemma 10.1. *$\alpha, \beta \in Acc \Rightarrow \alpha + \omega^\beta \in Acc$.*

Proof. Familiar from Gentzen's proof in PA. The proof just requires ACA_0 (cf. [4, Section 15]). □

Lemma 10.2. *$Acc = M \cap \Omega$ $(:= \{\alpha \in M : \alpha < \Omega\}.)$*

Proof. If $\alpha \in Acc$, then $E_\Omega(\alpha) \subseteq Acc$ as well; hence $\alpha \in M \cap \Omega$. If $\alpha \in M \cap \Omega$, then $E_\Omega(\alpha) \subseteq M \cap \Omega$, so $\alpha \in Acc$ follows from Lemma 10.1 □

Definition 10.2. Let $Prog_\Omega(X)$ stand for

$$(\forall\alpha \in M)[(\forall\beta <_\Omega \alpha)(\beta \in X) \rightarrow \alpha \in X].$$

Let $Acc_\Omega := \{\alpha \in M : \vartheta\alpha \in Acc\}$.

Lemma 10.3. *$Prog_\Omega(Acc_\Omega)$.*

Proof. Assume $\alpha \in M$ and $(\forall\beta <_\Omega \alpha)(\beta \in Acc_\Omega)$. We have to show $\vartheta\alpha \in Acc$. It suffices to show

$$\beta < \vartheta\alpha \Rightarrow \beta \in Acc. \tag{1}$$

We shall employ induction on $G_\vartheta(\beta)$, i.e., the length of (the term that represents) β. If $\beta \notin E$, then (1) follows easily by the inductive assumption and Lemma 10.1. Now suppose $\beta = \vartheta\beta_0$. Then we have to distinguish two cases according to Lemma 1.2.

Case 1: $\beta \leq \alpha^*$. Since $\alpha \in M$, we have $\alpha^* \in E_\Omega(\alpha) \subseteq Acc$; therefore $\beta \in Acc$.

Case 2: *$\beta_0 < \alpha$ and $\beta_0^* < \vartheta\alpha$.* As the length of β_0^* is less than the length of β, we get $\beta_0^* \in Acc$; thus $E_\Omega(\beta_0) \subseteq Acc$, therefore $\beta_0 \in M$. By the assumption at the beginning of the proof, we then get $\beta_0 \in Acc_\Omega$; hence $\beta = \vartheta\beta_0 \in Acc$. □

It should be noted that the proof of Lemma 10.3 only requires complete induction for Π_1^1-classes. The next lemma is where we really need $(\Pi_2^1\text{-}BI)_0^-$.

Lemma 10.4. *Let $A(a)$ be a Π_1^1 formula. Letting A_k be the formula*

$$\forall\alpha[(\forall\beta <_\Omega \alpha)\, A(\beta) \rightarrow (\forall\beta <_\Omega \alpha + \Omega^k)\, A(\beta)],$$

$(\Pi_2^1\text{-}BI)_0^-$ proves $Prog_\Omega(\{\xi: A(\xi)\}) \rightarrow A_k$.

Proof. We proceed by outer induction on k. Assume $Prog_\Omega(\{\alpha: A(\alpha)\})$ and $(\forall\beta <_\Omega \delta)A(\beta)$. Then also $(\forall\beta \leqslant_\Omega \delta)A(\beta)$. For $k = 0$ this gives the assertion, since $\gamma < \delta + \Omega^0$ implies $\gamma \leqslant \delta$. Now let $k = m + 1$. So we get A_m by the inductive assumption. Let $B(\eta)$ be the formula $(\forall\beta <_\Omega \delta + \Omega^m \cdot \eta)\, A(\beta)$. $B(\eta)$ is (in ACA_0) provably equivalent to a Π_2^1 formula. Suppose that $\eta \in Acc$ and $(\forall\rho < \eta)\, B(\eta)$. Clearly, $B(0)$ holds. If η is a limit, then for $\beta <_\Omega \delta + \Omega^m \cdot \eta$ there exists $\rho < \eta$ such that $\beta <_\Omega \delta + \Omega^m \cdot \rho$, hence $B(\eta)$ holds. Now let η be a successor $\gamma + 1$. Then $(\forall\beta <_\Omega \delta + \Omega^m \cdot \gamma)\, A(\beta)$. Using A_m, this implies $(\forall\beta < \delta + \Omega^m \cdot \gamma + \Omega^m)\, A(\beta)$, thus $(\forall\beta <_\Omega \delta + \Omega^m \cdot (\gamma + 1))\, A(\beta)$, hence $B(\eta)$. By the above considerations, we have

$$(\forall\eta \in Acc)[(\forall\rho < \eta)\, B(\rho) \rightarrow B(\eta)],$$

hence, using Π_2^1 bar induction along $<\restriction(\eta + 1)$, we obtain

$$(\forall\eta \in Acc)\, B(\eta).$$

Now for every $\beta <_\Omega \delta + \Omega^k$, there is some $\eta \in Acc$ such that $\beta <_\Omega \delta + \Omega^m \cdot \eta$. Therefore

$$(\forall\beta <_\Omega \delta + \Omega^k)\, A(\beta)$$

follows from $(\forall\eta \in Acc)\, B(\eta)$. □

Theorem 10.1. *For any n, $(\Pi_2^1\text{-}BI)_0^- \vdash WF(<\restriction\vartheta\Omega^n)$.*

Proof. Obviously $(\forall\beta <_\Omega 0)\, A(\beta)$ holds for any formula. Therefore we obtain

$$(\Pi_2^1\text{-}TI)_0^- \vdash Prog_\Omega(\{\xi: A(\xi)\}) \rightarrow (\forall\beta <_\Omega \Omega^k)\, A(\beta) \tag{1}$$

for any Π_1^1 formula $A(\alpha)$ and $k \in \mathbb{N}$. Since the formula '$\xi \in Acc_\Omega$' can be written Π_1^1, Lemma 10.3 along with (1) implies, for any $k \in \mathbb{N}$,

$$(\Pi_2^1\text{-}BI)_0^- \vdash (\forall\beta <_\Omega \Omega^k)[\beta \in Acc_\Omega].$$

Therefore, for any n, $(\Pi_2^1\text{-}BI)_0^- \vdash WF(<\restriction\vartheta\Omega^n)$. □

Corollary 10.1. *The proof-theoretic ordinal of $(\Pi_2^1\text{-}BI)_0$ and $(\Pi_2^1\text{-}BI)_0^-$ is $\vartheta\Omega^\omega$.*

Proof. Theorem 10.1 and Corollary 9.2. □

Corollary 10.2. *For every* $n < \omega$, $(\Pi^1_2\text{-}BI)^-_0 \vdash KT(\bar{n})$.

Proof. For fixed n the proof of $KT(n)$ requires only $WF(\vartheta\Omega^m)$ for some $m < n + 3$. This follows from the proof of Theorem 2.2. □

The methods of the last six sections can also be employed to determine the proof-theoretic ordinals of the systems $(\Pi^1_n\text{-}BI)_0$, $(\Pi^1_n\text{-}BI)^-_0$, $(\Pi^1_n\text{-}BI)$, and $(\Pi^1_n\text{-}BI)^-$ for $n > 2$.

For T a theory, let $|T|$ denote the proof-theoretic ordinal of T.

Letting $\Omega(1, \alpha) = \Omega^\alpha$, and $\Omega(n + 1, \alpha) := \Omega^{\Omega(n,\alpha)}$ for $n > 1$, the following is true.

Theorem 10.2. *Let* $n \geqslant 2$.

(1) $|(\Pi^1_n\text{-}BI)_0| = |(\Pi^1_n - BI)^-_0| = |KP\omega^- + \Pi_n\text{-}Foundation| = \vartheta\Omega(n - 1, \omega)$.

(2) $|(\Pi^1_n\text{-}BI)| = |(\Pi^1_n - BI)^-| = \vartheta\Omega(n - 1, \varepsilon_0)$.

Proof. The results about the set theories are from [5]. □

11. $ACA_0 \vdash KT \leftrightarrow RFN_{\Pi^1_1}((\Pi^1_2\text{-}BI)_0)$

In view of Theorem 10.1 one is naturally led to search for a natural strengthening of $(\Pi^1_2\text{-}BI)_0$ that proves $WF(\vartheta\Omega^\omega)$ and is not stronger than $ACA_0 + WF(\vartheta\Omega^\omega)$. Of course, $(\Pi^1_2\text{-}BI)$ proves $WF(\vartheta\Omega^\omega)$, because the outer induction on k in the proof of Lemma 20.4 can be carried out as a formal induction within $(\Pi^1_2\text{-}BI)$. Hence $(\Pi^1_2\text{-}BI) \vdash \forall n\, WF(\vartheta\Omega^n)$, thus $(\Pi^1_2\text{-}BI) \vdash WF(\vartheta\Omega^\omega)$. However, $(\Pi^1_2\text{-}BI)$ has proof-theoretic ordinal $\vartheta\Omega^{\varepsilon_0}$; so this is not the right candidate. It turns out that the Uniform Reflection Principle for $(\Pi^1_2\text{-}BI)_0$, $RFN_{\Pi^1_1}((\Pi^1_2\text{-}BI)_0)$, provides the appropriate strengthening. This scheme asserts the Π^1_1 soundness (with parameters) of $(\Pi^1_2\text{-}BI)_0$.

Definition 11.1. $RFN_{\Pi^1_2}((\Pi^1_2\text{-}BI)_0)$ is the scheme

$$\forall n\, [Pr_{(\Pi^1_2\text{-}BI)_0}(\ulcorner F(\bar{n})\urcorner) \rightarrow F(n)]$$

for Π^1_1 formulas $F(a)$ with at most one free variable a. Here $Pr_{(\Pi^1_2-BI)_0}$ denotes the Σ^0_1 provability predicate for $(\Pi^1_2\text{-}BI)_0$ and $\ulcorner F(\bar{n})\urcorner$ signifies the Gödel number of $F(a)$ when a is replaced by the nth numeral (for details see [11]).

The proof of Theorem 10.1 can be formalized in ACA_0. Therefore $ACA_0 \vdash \forall x\, Pr_{(\Pi^1_2\text{-}BI)_0}(WF(\vartheta\Omega^{\bar{x}}))$. So we come to see the following.

Theorem 11.1. $ACA_0 + RFN_{\Pi^1_1}((\Pi^1_2\text{-}BI)_0) \vdash WF(\vartheta\Omega^\omega)$.

The remainder of the section will be devoted to outlining the proof of the next theorem.

Theorem 11.2. $ACA_0 + WF(\vartheta\Omega^\omega) \vdash RFN_{\Pi^1_1}((\Pi^1_2\text{-}BI)_0)$.

We first give the rough idea of that proof. So let $\mathcal{D} \vdash F(\bar{n})$ be a $(\Pi^1_2\text{-}BI)_0$ proof of a Π^1_1 formula $F(\bar{n}) \equiv \forall X A(x, \bar{n})$. Using Theorem 9.1, we can pick a T^* derivation $\mathcal{D}^*$ such that $\mathcal{D}^* \vdash^{\Omega^m}_{\omega+1} A(U, \bar{n})$ for some $m < \Omega$ which is determined by $\mathcal{D}$. Furthermore, we can transform $\mathcal{D}^*$ into a cut-free derivation $\mathcal{D}^{**} \vdash^{\alpha}_0 A(U, \bar{n})$ with $\alpha = \psi\Omega^{\Omega^m}$. Since $\alpha < \Omega$, the derivation $\mathcal{D}^{**}$ does not contain instances of the Ω-rule. Especially, all the formulas occurring in $\mathcal{D}^{**}$ have to be subformulas of $A(U, \bar{n})$. Thence, by induction on α one verifies that $A(X, n)$ is true for any set X of natural numbers. Since the formula $A(U, \bar{n})$ is Σ^0_k for some k, this very last step requires only a truth predicate for Σ^0_k formulas (with parameters) which is available in ACA_0.

A minor problem with the above scetch is that we didn't exactly specify the amount of transfinite induction that it requires. Closer inspection reveals that we can restrict ourselves to ordinals from $C(\Omega^{\Omega^\omega})$. Since the order-type of $<\restriction C(\Omega^{\Omega^\omega})$ is not much bigger than $\sigma := \psi\Omega^{\Omega^\omega}$, namely $\omega^{\omega^{\sigma\cdot\omega}}$, the well-foundedness of $<\restriction C(\Omega^{\Omega^\omega})$ is still provable in $ACA_0 + WF(\psi\Omega^{\Omega^\omega})$.

A considerably greater challenge is offered by the question: How can we formalize infinitary derivations in $ACA_0 + WF(\psi\Omega^{\Omega^\omega})$? Obviously we have to use an effective counterpart of the above construction, where we work with codes for T^*-derivations instead of using the T^*-derivations themselves. The main idea here is that we can do everything with recursive proof-trees instead of arbitrary derivations. A *proof-tree* is a tree, with each node labeled by: A sequent, a rule of inference or the designation 'Axiom', two sets of formulas specifying the set of principal and minor formulas, respectively, of that inference, and two ordinals (length and cut-rank) such that the sequent is obtained from those immediately above it through application of the specified rule of inference. The well-foundedness of a proof-tree is then witnessed by the (first) ordinal 'tags' which are in reverse order of the tree order.

If the inference is an instance of (Ω*-rule*), the label should also provide an index for the fundamental function f and an index for the functional $\mathcal{T}$ that gives the transformations on proof-trees (cf. Section 8); hence both are required to be recursive.

We then have to show that none of our manipulations on T^*-derivations leads us beyond this class of recursive proof-trees. The latter is guaranteed by the fact that for the embedding of $(\Pi^1_2\text{-}BI)_0$ into T^* we only need instances of the (Ω*-rule*), where the transformation on proof-trees is given by a recursive functional, and by the fact that all the operations in the cut-elimination procedure are of a local nature, i.e., they give rise to recursive functionals.

To carry out all the details of this constructivization would mean to produce another lengthy paper. But it is high time that we finished this paper; so we simply quit at this point.

References

[1] W. Buchholz and K. Schütte, Proof Theory of Impredicative Subsystems of Analysis (Bibliopolis, Napels, 1988).
[2] N. Dershowitz and M. Okada, Proof-theoretic techniques for term-rewriting theory, Proceedings of the Third Annual Symposium on Logic in Computer Science, Edinburgh, Scotland (1988) 104–111.
[3] J.H. Gallier, What's so special about Kruskal's theorem and the ordinal Γ_0? A survey of some results in proof theory, Ann. Pure Appl. Logic 53 (1991) 199–260.
[4] W. Pohlers, Proof Theory: An introduction, Lecture Notes in Math. 1407 (Springer, Berlin, 1990).
[5] M. Rathjen, Fragments of Kripke–Platek set theory with infinity, Leeds Proof Theory 90 (Cambridge Univ. Press, Cambridge, 1990).
[6] D. Schmidt, Well-Partial Orderings and Their Maximal Order Types, Habilitationsschrift, Heidelberg, 1979.
[7] K. Schütte, Proof Theory (Springer, Berlin, 1977).
[8] K. Schütte and S.G. Simpson, Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen, Archiv Math. Logik Grundlag. 25 (1985) 75–89.
[9] H. Schwichtenberg, Proof theory: some applications of cut-elimination, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 867–895.
[10] S.G. Simpson, Nonprovability of certain combinatorial properties of finite trees, in: L.A. Harrignton et al., eds., H. Friedman's Research on the Foundations of Mathematics (North-Holland, Amsterdam, 1985) 87–117.
[11] G. Smorynski, The incompleteness theorems, in: J. Barwise, ed., Handbook of Mathematical Logic (North-Holland, Amsterdam, 1977) 821–865.