Well-Ordering Principles in Proof Theory and Reverse Mathematics

Michael Rathjen

Department of Pure Mathematics, University of Leeds

Leeds LS2 9JT, UK

E-Mail: M.Rathjen@leeds.ac.uk

Abstract

Several theorems about the equivalence of familiar theories of reverse mathematics with certain well-ordering principles have been proved by recursion-theoretic and combinatorial methods (Friedman, Marcone, Montalbán et al.) and with far-reaching results by proof-theoretic technology (Afshari, Freund, Girard, Rathjen, Thomson, Valencia Vizcaíno, Weiermann et al.), employing deduction search trees and cut elimination theorems in infinitary logics with ordinal bounds in the latter case.

At type level one, the well-ordering principles are of the form

$(*)$ "if X is well-ordered then $f(X)$ is well-ordered"

where f is a standard proof theoretic function from ordinals to ordinals (such f's are always dilators). One aim of the paper is to present a general methodology underlying these results that enables one to construct omega-models of particular theories from $(*)$ and even β-models from the type 2 version of $(*)$.

As $(*)$ is of complexity Π^1_2 such a principle cannot characterize stronger comprehensions at the level of Π^1_1-comprehension. This requires a higher order version of $(*)$ that employs ideas from ordinal representation systems with collapsing functions used in impredicative proof theory. The simplest one is the Bachmann construction. Relativizing the latter construction to any dilator f and asserting that this always yields a well-ordering turns out to be equivalent to Π^1_1-comprehension. This result has been conjectured and a proof strategy adumbrated roughly 10 years ago, but a detailed proof has only been worked out in recent years.

Keywords: infinite proof theory, reverse mathematics, ordinal analysis, ordinal representation systems, cut elimination, dilators, ω-models, β-models
MSC2000: 03F50; 03F25; 03E55; 03B15; 03C70

1 Introduction

The aim to illuminate the role of the infinite in mathematics gave rise to set theory and proof theory alike. Whereas set theory takes much of the infinite for granted (e.g. full separation and powerset), proof theory strives to analyze it from a stance of potential infinity. While the objects that proof theory primarily deals with are rather concrete (e.g. theories, proofs, ordinal representation systems), it is also concerned with abstract higher type[1] properties of concrete objects (e.g. well-foundedness, preservation of well-foundedness). In ordinal analysis the impetus is to associate an ordinal representation system $\mathcal{OR}$ with a theory T in such a way that the former displays the commitments to the

---

[1]Or *ideal* properties in Hilbert's sense.

infinite encapsulated in T. Less poetically, it entails that $\mathbf{PRA} + \mathbf{TI}_{qf}(\mathcal{OR})$ proves all Π_2^0-theorems of T, however, no proper initial segment of $\mathcal{OR}$ suffices for that task (where $\mathbf{PRA}$ refers to primitive recursive arithmetic and $\mathbf{TI}_{qf}(\mathcal{OR})$ stands for quantifier-free transfinite induction along $\mathcal{OR}$).[2] The first ordinal representations were offshoots of ordinal normal forms such as Cantor's and Veblen's from more than a hundred years ago. Gentzen employed the Cantor normal form with base ω to provide an ordinal representation system for ε_0 in his last two papers [22, 23], which mark the beginning of ordinal-theoretic proof theory.

It is often stressed that ordinal representation systems are computable[3] structures, which is true and allows for the treatment of their order-theoretic and algebraic aspects in very weak systems of arithmetic, but it is only one of their distinguishing features. Overstating the computability aspect tends to give the impression that their study is part of the venerable research area of computable orderings. In actual fact, the two subjects have very little in common.

There are many specific ordinal representations systems to be found in proof theory. To make the choice of these systems intelligible one would like to discern general principles involved. A particular challenge is posed by making their algebraic features explicit, i.e. those features not related to effectiveness or well-foundedness. This constitutes a crucial step towards a general theory via an axiomatic approach. As far as I'm aware, it was Feferman [11] who first initiated such a theory by isolating the property of *effective relative categoricity* as the fulcrum of such a characterization. Through the notion of relative categoricity he succeeded in crystallizing the algebraic aspect of ordinal representation systems by way of relativizing them to any set of order-indiscernibles, thereby in effect scaling them up to functors on the category of linear orders with order-preserving maps as morphisms.

Ordinal representation systems understood in this relativized way give rise to well-ordering principles:

<center>"if $\mathfrak{X}$ is a well-ordering then so is $\mathcal{OR}_{\mathfrak{X}}$"</center>

where $\mathcal{OR}_{\mathfrak{X}}$ arises from $\mathcal{OR}$ by letting $\mathfrak{X}$ play the role of the order-indiscernibles. Several theories of reverse mathematics have been shown to be equivalent to such principles involving iconic ordinal representation system from the proof-theoretic literature. From a technical point of view the methods can be roughly divided into recursion-theoretic and combinatorial tools (Friedman, Marcone, Montalban et al.) on the one hand and proof-theoretic tools (Afshari, Freund, Girard, Rathjen, Thomson, Valencia Vizcaíno, Weiermann et al.) on the other hand, where in the latter approach the employment of deduction search trees and cut elimination theorems in infinitary logics with ordinal bounds play a central role.

---

[2]In some books, the proof-theoretic ordinal $|T|$ of a theory T is defined as the supremum of the order-types $|\prec|$ of the T-provable well-orderings $\prec$, i.e.,

$$|T|_{\sup} = \sup\{|\prec| \; : \; \prec \text{ primitive recursive} \; ; T \vdash \mathrm{WO}(\prec)\}$$

where WO expresses that $\prec$ is a well-ordering. This definition, however, conflates extensional aspects, i.e. the order-type of $\prec$ in some background universe of set theory, with intensional aspects, i.e. the niceties of the coding of $\prec$ in the integers. This is liable to lead to results where $|T|_{\sup}$ has nothing to do with the "real" proof-theoretic strength of T. For example, one can define "faithful" theories T_1 and T_2 such that: (a) $|\mathsf{T}_1|_{\sup}$ is the Bachmann-Howard ordinal but T_1 is finitistically reducible to $\mathbf{PA}$, proving the same Π_1^0-theorems as $\mathbf{PA}$; (b) $|\mathsf{T}_2|_{\sup} = \varepsilon_0$ but T_2 proves the same arithmetical theorems as an extension of $\mathbf{ZFC}$ with large cardinal axioms. Thus, paraphrasing the title of Paul Benacerraf's famous 1965 paper, one could say that $|T|_{\sup}$ displays what the proof-theoretic ordinal of T could not be. For more on this, see [48].

[3]They actually possess a very low computational complexity; see [65].

This is a survey paper about these results that tries to illuminate the underlying ideas, emphasizing the proof theory side.

1.1 Reverse mathematics

It is assumed that the reader is roughly familiar the the program of *Reverse Mathematics* (RM) and the language and axioms of the formal system of second order arithmetic, $\mathbf{Z_2}$ as for instance presented in the standard reference [64]. Just as a reminder, RM started out with the observation that for many mathematical theorems τ, there is a weakest natural subsystem $S(\tau)$ of $\mathbf{Z_2}$ such that $S(\tau)$ proves τ. Moreover, it has turned out that $S(\tau)$ often belongs to a small list of specific subsystems of $\mathbf{Z_2}$. In particular, Reverse Mathematics has singled out five subsystems of $\mathbf{Z_2}$, often referred to as the *big five*, that provide (part of) a standard scale against which the strengths of theories can often be measured:

1. $\mathbf{RCA}_0$ Recursive Comprehension

2. $\mathbf{WKL}_0$ Weak König's Lemma

3. $\mathbf{ACA}_0$ Arithmetical Comprehension

4. $\mathbf{ATR}_0$ Arithmetical Transfinite Recursion

5. $\Pi^1_1-\mathbf{CA}_0$ Π^1_1-Comprehension

2 History

The principles we will be concerned with are particular Π^1_2 statements of the form

$$\mathbf{WOP}(f): \qquad \forall X\,[\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(f(\mathfrak{X}))] \qquad\qquad (1)$$

where f is a standard proof-theoretic function from ordinals to ordinals and $\mathrm{WO}(\mathfrak{X})$ stands for '$\mathfrak{X}$ is a well-ordering'.[4] There are by now several examples of functions f familiar from proof theory where the statement $\mathbf{WOP}(f)$ has turned out to be equivalent to one of the theories of reverse mathematics over a weak base theory (usually $\mathbf{RCA}_0$). The first explicit example appears to be due to Girard [26, 5.4.1 theorem] (see also [31]). However, it is also implicit in Schütte's proof of cut elimination for ω-logic [58] and ultimately has its roots in Gentzen's work, namely in his first unpublished consistency proof,[5] where he introduced the notion of a "Reduziervorschrift" [21, p. 102] for a sequent. The latter is a well-founded tree built bottom-up via "Reduktionsschritte", starting with the given sequent and passing up from conclusions to premises until an axiom is reached.

2.1 $2^\mathfrak{X}$ and Arithmetical Comprehension

In our first example, one constructs a linear ordering $2^\mathfrak{X} = (|2^X|, <_{2^X})$ from a given linear ordering $\mathfrak{X} = (X, <_X)$, where $|2^X|$ consists of the formal sums $2^{x_1} + \ldots + 2^{x_n}$ with $x_n <_\mathfrak{X} \ldots <_\mathfrak{X} x_1$, and the ordering between two formal sums $2^{x_0} + \ldots + 2^{x_n}$, $2^{y_0} + \ldots + 2^{y_m}$ is determined as follows:

$$2^{x_0} + \ldots + 2^{x_n} <_{2^X} 2^{y_1} + \ldots + 2^{y_m} \quad \text{iff} \quad \exists i \le \max(m,n)[x_i <_X y_i \wedge \forall j < i\ x_j = y_j]$$
$$\text{or } n < m \wedge \forall i \le n\ x_i = y_i.$$

---

[4]And $\mathbf{WOP}$ is an acronym for the German word "**W**ohl**o**rdnungs**p**rinzip".

[5]The original German version was finally published in 1974 [21]. An earlier English translation appeared in 1969 [19].

Theorem 2.1 (Girard 1987) *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) Arithmetical comprehension.

(ii) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(2^{\mathfrak{X}})]$.

Another characterization from [26], Theorem 6.4.1, shows that arithmetical comprehension is equivalent to Gentzen's Hauptsatz (cut elimination) for ω-logic. Connecting statements of form (1) to cut elimination theorems for infinitary logics will be a major tool in this paper.

2.2 $\mathbf{ACA}_0^+$ and $\varepsilon_{\mathfrak{X}}$

There are several more recent examples of such equivalences that have been proved by recursion-theoretic as well as proof-theoretic methods. The second example is a characterization in the same vein as (1) for the theory $\mathbf{ACA}_0^+$ in terms of the $\xi \mapsto \varepsilon_\xi$ function. $\mathbf{ACA}_0^+$ stands for the theory $\mathbf{ACA}_0$ augmented by an axiom asserting that for any set X the ω jump in X exists. This theory is very interesting as it is currently the weakest subsystem of $\mathbf{Z_2}$ in which one has succeeded to prove Hindman's Ramsey-type combinatorial theorem (asserting that finite colorings have infinite monochromatic sets closed under taking sums) and the Auslander/Ellis theorem of topological dynamics (see [64, X.3]) The source of the pertaining representation system and its relativization is the Cantor normal form.

Theorem 2.2 (Cantor, 1897) *For every ordinal* $\beta > 0$ *there exist unique ordinals* $\beta_0 \geq \beta_1 \geq \cdots \geq \beta_n$ *such that*

$$\beta = \omega^{\beta_0} + \ldots + \omega^{\beta_n}. \tag{2}$$

The representation of β in (2.4) is called the *Cantor normal form*. We shall write $\beta =_{CNF} \omega^{\beta_1} + \cdots + \omega^{\beta_n}$ to convey that $\beta_0 \geq \beta_1 \geq \cdots \geq \beta_k$.

ε_0 denotes the least ordinal $\alpha > 0$ such that $\beta < \alpha \;\Rightarrow\; \omega^\beta < \alpha$, or equivalently, the least ordinal α such that $\omega^\alpha = \alpha$. Every $\beta < \varepsilon_0$ has a Cantor normal form with exponents $\beta_i < \beta$ and these exponents have Cantor normal forms with yet again smaller exponents etc.. As this process must terminate, ordinals $< \varepsilon_0$ can be effectively coded by natural numbers.

We state the result before introducing the functor $\mathfrak{X} \mapsto \varepsilon_{\mathfrak{X}}$.

Theorem 2.3 (Afshari, Rathjen [1]; Marcone, Montalbán [37]) *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) $\mathbf{ACA}_0^+$.

(ii) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\varepsilon_{\mathfrak{X}})]$.

Given a linear ordering $\mathfrak{X} = \langle X, <_X \rangle$, the idea for obtaining the new linear ordering $\varepsilon_{\mathfrak{X}}$ is to create a formal ε-number ε_u for every $u \in X$ such that if $v <_X u$ then $\varepsilon_v <_{\varepsilon_{\mathfrak{X}}} \varepsilon_u$, and in addition fill up the "spaces" between these terms with formal Cantor normal forms.

The ordering $<_{\varepsilon_{\mathfrak{X}}}$

Definition 2.4 Let $\mathfrak{X} = \langle X, <_X \rangle$ be an ordering where $X \subseteq \mathbb{N}$. $<_{\varepsilon_{\mathfrak{X}}}$ and its field $|\varepsilon_{\mathfrak{X}}|$ are inductively defined as follows:

1. $0 \in |\varepsilon_{\mathfrak{X}}|$.

2. $\varepsilon_u \in |\varepsilon_{\mathfrak{X}}|$ for every $u \in X$, where $\varepsilon_u := \langle 0, u \rangle$.

3. If $\alpha_1, \ldots, \alpha_n \in |\varepsilon_{\mathfrak{X}}|$, $n > 1$ and $\alpha_n \leq_{\varepsilon_{\mathfrak{X}}} \ldots \leq_{\varepsilon_{\mathfrak{X}}} \alpha_1$, then

$$\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} \in |\varepsilon_{\mathfrak{X}}|$$

   where $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} := \langle 1, \langle \alpha_1, \ldots, \alpha_n \rangle \rangle$.

4. If $\alpha \in |\varepsilon_{\mathfrak{X}}|$ and α is not of the form ε_u, then $\omega^\alpha \in |\varepsilon_{\mathfrak{X}}|$, where $\omega^\alpha := \langle 2, \alpha \rangle$.

5. $0 <_{\varepsilon_{\mathfrak{X}}} \varepsilon_u$ for all $u \in X$.

6. $0 <_{\varepsilon_{\mathfrak{X}}} \omega^{\alpha_1} + \ldots + \omega^{\alpha_n}$ for all $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} \in |\varepsilon_{\mathfrak{X}}|$.

7. $\varepsilon_u <_{\varepsilon_{\mathfrak{X}}} \varepsilon_v$ if $u, v \in X$ and $u <_X v$.

8. If $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} \in |\varepsilon_{\mathfrak{X}}|$, $u \in X$ and $\alpha_1 <_{\varepsilon_{\mathfrak{X}}} \varepsilon_u$ then $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} <_{\varepsilon_{\mathfrak{X}}} \varepsilon_u$.

9. If $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} \in |\varepsilon_{\mathfrak{X}}|$, $u \in X$, and $\varepsilon_u <_{\varepsilon_{\mathfrak{X}}} \alpha_1$ or $\varepsilon_u = \alpha_1$, then $\varepsilon_u <_{\varepsilon_{\mathfrak{X}}} \omega^{\alpha_1} + \ldots + \omega^{\alpha_n}$.

10. If $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n}$ and $\omega^{\beta_1} + \ldots + \omega^{\beta_m} \in |\varepsilon_{\mathfrak{X}}|$, then

$$\omega^{\alpha_1} + \ldots + \omega^{\alpha_n} <_{\varepsilon_{\mathfrak{X}}} \omega^{\beta_1} + \ldots + \omega^{\beta_m} \text{ iff}$$
$$n < m \wedge \forall i \leq n \ \alpha_i = \beta_i \quad \text{or}$$
$$\exists i \leq \min(n, m)[\alpha_i <_{\varepsilon_{\mathfrak{X}}} \beta_i \wedge \forall j < i \ \alpha_j = \beta_j].$$

Let $\varepsilon_{\mathfrak{X}} = \langle |\varepsilon_{\mathfrak{X}}|, <_{\varepsilon_{\mathfrak{X}}} \rangle$.

One then proves (for instance in $\mathbf{RCA}_0$) that $\varepsilon_{\mathfrak{X}}$ is linear ordering. If we denote by $\mathbb{LO}$ the category of linear orderings whose objects are all linear orderings and whose morphisms are the order-preserving maps between linear orderings, then $\mathfrak{X} \mapsto \varepsilon_{\mathfrak{X}}$ gives rise to an endofunctor of $\mathbb{LO}$: For a morphism $f \colon \mathfrak{X} \to \mathfrak{Y}$ define $\varepsilon(f)(t)$ for $t \in |\varepsilon_{\mathfrak{X}}|$ by replacing every expression ε_u occurring in t by $\varepsilon_{f(u)}$. Then, $\varepsilon(f) \colon \varepsilon_{\mathfrak{X}} \to \varepsilon_{\mathfrak{Y}}$. Moreover, the restriction of this functor to wellorderings is a dilator in the sense of Girard [24] as it preserves pullbacks and direct limits (more about this in subsection 4.2).

Ordinal representation systems: 1904-1909 At the beginning of the 20th century, mathematicians were intrigued by Cantor's continuum problem. Hardy in 1904 wanted to "construct" a subset of $\mathbb{R}$ of size $\aleph_1$. In [28] he gave explicit representations for all ordinals $< \omega^2$. Hardy's work then inspired O. Veblen who in his paper from 1908 [69] found new normal forms for ordinals and succeeded in "naming" ordinals of an impressive chunk of the countable ordinals. In doing so, he also furnished proof theorists with the central idea for creating ordinal representation systems that were sufficient for much of their work until the middle 1960s. His φ-function was crucial in the work of S. Feferman and K. Schütte who in the 1960s determined the limits of predicative mathematics with a notion of predicativity based on autonomous progressions of theories (cf. [10, 11, 61, 62]).

Veblen considered *continuous increasing functions* on ordinals. Let ON be the class of ordinals. A (class) function $f : \text{ON} \to \text{ON}$ is said to be *increasing* if $\alpha < \beta$ implies $f(\alpha) < f(\beta)$ and *continuous* (in the order topology on ON) if

$$f(\lim_{\xi < \lambda} \alpha_\xi) = \lim_{\xi < \lambda} f(\alpha_\xi)$$

holds for every limit ordinal λ and increasing sequence $(\alpha_\xi)_{\xi < \lambda}$. f is called *normal* if it is increasing and continuous. By way of contrast, the function $\beta \mapsto \omega + \beta$ is normal whereas $\beta \mapsto \beta + \omega$ is not since for instance $\lim_{\xi < \omega}(\xi + \omega) = \omega$ but $(\lim_{\xi < \omega} \xi) + \omega = \omega + \omega$.

To these normal functions Veblen applied two operations:

(i) *Derivation*

(ii) *Transfinite Iteration.*

The *derivative* f' of a function $f : \mathrm{ON} \to \mathrm{ON}$ is the function which enumerates in increasing order the solutions of the equation

$$f(\alpha) = \alpha,$$

also called the *fixed points* of f. It is a fact of set theory that if f is a normal function,

$$\{\alpha : f(\alpha) = \alpha\}$$

is a proper class and f' will be a normal function, too.

Using the two operations, Veblen defined a hierarchy of normal functions indexed along the ordinals.

Definition 2.5 Given a normal function $f : \mathrm{ON} \to \mathrm{ON}$, define a hierarchy of normal functions as follows:

1. $f_0 = f$

2. $f_{\alpha+1} = f_\alpha{}'$

3.
$$f_\lambda(\xi) = \xi^{th} \text{ element of } \bigcap_{\alpha < \lambda} \{\text{fixed points of } f_\alpha\} \qquad \text{for } \lambda \text{ limit.}$$

Starting with the normal function $f(\xi) = \omega^\xi$, the function f_α is usually denoted by φ_α.

The least ordinal $\gamma > 0$ closed under $\alpha, \beta \mapsto \varphi_\alpha(\beta)$, i.e. the least ordinal > 0 satisfying $(\forall \alpha, \beta < \gamma)\, \varphi_\alpha(\beta) < \gamma$ is the famous ordinal Γ_0 that Feferman and Schütte determined to be the least ordinal 'unreachable' by certain autonomous progressions of theories.

The two-place φ-function gives rise to a normal form theorem.

Theorem 2.6 (φ normal form) *For every additive principal[6] ordinal α there exist uniquely determined ordinals ξ and η such that $\alpha = \varphi_\xi(\eta)$ and $\eta < \alpha$.*

Proof: See [63, Theorem 13.12] or [51, Theorem 5.27]. □

The following comparison theorem encapsulates a procedure for recursively determining the order of φ-expressions, which can then be utilized to develop a representation system for the ordinals below Γ_0.

Theorem 2.7 (φ-comparison) *(i) $\varphi_{\alpha_1}(\beta_1) = \varphi_{\alpha_2}(\beta_2)$ holds iff one of the following conditions is satisfied:*

1. $\alpha_1 < \alpha_2$ and $\beta_1 = \varphi_{\alpha_2}(\beta_2)$

2. $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$

3. $\alpha_2 < \alpha_1$ and $\varphi_{\alpha_1}(\beta_1) = \beta_2$.

(ii) $\varphi_{\alpha_1}(\beta_1) < \varphi_{\alpha_2}(\beta_2)$ holds iff one of the following conditions is satisfied:

1. $\alpha_1 < \alpha_2$ and $\beta_1 < \varphi_{\alpha_2}(\beta_2)$

2. $\alpha_1 = \alpha_2$ and $\beta_1 < \beta_2$

---

[6]This means that $\alpha > 0$ and $\forall \delta, \delta' < \alpha\, \delta + \delta' < \alpha$.

3. $\alpha_2 < \alpha_1$ and $\varphi_{\alpha_1}(\beta_1) < \beta_2$.

Proof: [63, Theorems 13.9, 13.10] or [51, Theorem 5.25]. □

The representation system for Γ_0 can be relativized to any ordering $\mathfrak{X}$ by first introducing formal function symbols φ_u for any $u \in X$ and secondly creating terms out of these using also 0 and formal addition $+$, and finally singling out the normal-forms. The crucial case in defining the ordering is the following:

$$\varphi_u s <_{\varphi_{\mathfrak{X}}} \varphi_v t \quad \text{iff} \quad \begin{aligned} &u <_X v \wedge s <_{\varphi_{\mathfrak{X}}} \varphi_v t \quad \text{or} \\ &u = v \wedge s <_{\varphi_{\mathfrak{X}}} t \quad \text{or} \\ &v <_X u \wedge \varphi_u s <_{\varphi_{\mathfrak{X}}} t, \end{aligned}$$

where all terms are assumed to be in normal form (for details see [37, 55]).

$\varphi_{\mathfrak{X}}$ induces a functor on $\mathbb{LO}$ that characterizes $\mathbf{ATR}_0$.

Theorem 2.8 *(Friedman, unpublished; Marcone and Montalbán; Rathjen and Weiermann) Over $\mathbf{RCA}_0$ the following are equivalent:*

1. $\mathbf{ATR}_0$

2. $\forall \mathfrak{X} [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\varphi_{\mathfrak{X}})]$.

Friedman's proof uses computability theory and also some proof theory. Among other things it uses a result which states that if $P \subseteq \mathcal{P}(\omega) \times \mathcal{P}(\omega)$ is arithmetic, then there is no sequence $\{A_n \mid n \in \omega\}$ such that (a) for every n, A_{n+1} is the unique set such that $P(A_n, A_{n+1})$, and (b) for every n, $A'_{n+1} \leq_T A_n$.

Of the two published proofs of the foregoing theorem, A. Marcone and A. Montalbán employ tools from computability theory as their paper's title, *The Veblen function for computability theorists* [37], clearly indicates. A. Weiermann and the author of this paper use purely proof-theoretic means in [55].

We would like to give the reader some insight into a proof strategy for showing theorems such as 2.3 and 2.8. However, we will be doing this by way of a different example. To state this result, it is convenient to introduce the notion of countable coded ω-model.

Definition 2.9 Let T be a theory in the language of second order arithmetic, $\mathcal{L}_2$. A *countable coded ω-model of T* is a set $W \subseteq \mathbb{N}$, viewed as encoding the $\mathcal{L}_2$-model

$$\mathbb{M} = (\mathbb{N}, \mathcal{S}, \in, +, \cdot, 0, 1, <)$$

with $\mathcal{S} = \{(W)_n \mid n \in \mathbb{N}\}$ such that $\mathbb{M} \models T$ when the second order quantifiers are interpreted as ranging over $\mathcal{S}$ and the first order part is interpreted in the standard way (where $(W)_n = \{m \mid \langle n, m \rangle \in W\}$ with $\langle , \rangle$ being some primitive recursive coding function).

If T has only finitely many axioms, it is obvious how to express $\mathbb{M} \models T$ by just translating the second order quantifiers $QX \ldots X \ldots$ in the axioms by $Qx \ldots (W)_x \ldots$. If T has infinitely many axioms,$+$ one needs to formalize Tarski's truth definition for $\mathbb{M}$. This definition can be made in $\mathbf{RCA}_0$ as is shown in [64], Definition II.8.3 and Definition VII.2.

We write $X \in W$ if $\exists n \; X = (W)_n$.

The notion of countable coded ω-model lends itself to alternative characterizations of Theorems 2.3 and 2.8.[7]

---

[7]The standard systems of reverse mathematics have the induction axiom $\forall X [0 \in X \wedge \forall y (y \in X \to y + 1 \in X) \to \forall y \; y \in X]$ but aren't capable of deducing the induction scheme for all formulas of $\mathcal{L}_2$. Note that an ω-model of such a theory will always satisfy the full induction scheme.

Theorem 2.10 *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\varepsilon_{\mathfrak{X}})]$ *is equivalent to the statement that every set is contained in a countable coded ω-model of* $\mathbf{ACA}_0$.

(ii) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\varphi \mathfrak{X} 0)]$ *is equivalent to the statement that every set is contained in a countable coded ω-model of* $\Delta_1^1\text{-}\mathbf{CA}$ *(or* $\Sigma_1^1\text{-}\mathbf{DC}$*).*

Proof: See [52, Corollary 1.8]. □

Whereas Theorem 2.10 has been established independently by recursion-theoretic and proof-theoretic methods, there is also a result that has a very involved proof and so far has only been shown by proof theory. It connects the Γ-function with the existence of countable coded ω-models of $\mathbf{ATR}_0$. For this we need to introduce the endofunctor $\mathfrak{X} \mapsto \Gamma_{\mathfrak{X}}$ of $\mathbb{LO}$. The linear ordering $\Gamma_{\mathfrak{X}}$ is created out of $\mathfrak{X}$ by adding formal Γ-terms Γ_u for every $u \in X$ with the stipulation that $\Gamma_v <_{\Gamma_{\mathfrak{X}}} \Gamma_u$ if $v <_X u$, and in addition one fills up the "spaces" between these terms with formal sums and Veblen normal forms. The details can be found in [52, Definition 2.5].

Theorem 2.11 ([52, Theorem 1.4]) *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\Gamma_{\mathfrak{X}})]$.

(ii) *Every set is contained in a countable coded ω-model of* $\mathbf{ATR}_0$.

The tools from proof theory employed in the above theorems involve search trees and a plethora of cut elimination techniques for infinitary logic with ordinal bounds. The search tree techniques is a starting point that all proof-theoretic proofs of the theorems of this paper have in common. It consists in producing the search or decomposition tree (in German "Stammbaum") of a given formula. It proceeds by decomposing the formula according to its logical structure and amounts to applying logical rules backwards. This decomposition method has been employed by Schütte [59, 60] to prove the completeness theorem for ω-logic. It is closely related to the method of "semantic tableaux" of Beth [6] and the tableaux of Hintikka [29]. Ultimately, the whole idea derives from Gentzen [20].

2.3 Proof idea of (1)⇒(2) of Theorem 2.11

In it we shall use a simple result, namely that $\mathbf{ATR}_0$ can be axiomatized via a single sentence Π_2^1 sentence $\forall X \, C(X)$ where $C(X)$ is Σ_1^1 (see [64]).

Definition 2.12 Let L_2 be the language of second order arithmetic. We assume that L_2 has relation symbols for primitive recursive relations. Formulas are generated from atomic $R(t_1, \ldots, t_n)$ and negated atomic formulas $\neg R(t_1, \ldots, t_n)$, where R is a symbol an n-ary primitive recursive relation and $t_1, \ldots, t_n$ are numerical terms,[8] via the logical particles $\wedge, \vee, \exists x, \forall x, \exists X, \forall X$; so formulas are assumed to be in negation normal form.

(a) Let $U_0, U_1, U_2, \ldots$ be an enumeration of the free set variables of L_2.

(b) For a closed term t, let $t^{\mathbb{N}}$ be its numerical value. We shall assume that all predicate symbols of the language L_2 are symbols for primitive recursive relations. L_2 contains predicate symbols for the primitive recursive relations of equality and inequality and

---

[8]They are generated from numerical variables $x, y, z, \ldots$, the constants 0, 1, and the function symbols $+, \cdot$ for addition and multiplication, respectively.

possibly more (or all) primitive recursive relations. If R is a predicate symbol in L_2 we denote by $R^{\mathbb{N}}$ the primitive recursive relation it stands for. If $t_1, \ldots, t_n$ are closed terms the formula $R(t_1, \ldots, t_n)$ $(\neg R(t_1, \ldots, t_n))$ is said to be *true* if $R^{\mathbb{N}}(t_1^{\mathbb{N}}, \ldots, t_n^{\mathbb{N}})$ is true (is false).

(c) Henceforth a **sequent** will be a finite list of L_2-formulas *without free number variables*. Sequents will be denoted by upper case Greek letters.

Given sequents Γ and Λ and a formula A, we adopt the convention that Γ, A, Λ denotes the sequent resulting from extending the list Γ by A and then extending it further by appending the list Λ.

(i) A sequent Γ is **axiomatic** if it satisfies at least one of the following conditions:

1. Γ contains a true **literal**, i.e. a true formula of either form $R(t_1, \ldots, t_n)$ or $\neg R(t_1, \ldots, t_n)$, where R is a predicate symbol in L_2 for a primitive recursive relation and $t_1, \ldots, t_n$ are closed terms.

2. Γ contains formulas $s \in U$ and $t \notin U$ for some set variable U and terms s, t with $s^{\mathbb{N}} = t^{\mathbb{N}}$.

(iv) A sequent is **reducible** or a **redex** if it is not axiomatic and contains a formula which is not a literal.

Definition 2.13 (Deduction chains in ω-logic) A **deduction chain** is a finite string

$$\Gamma_0; \Gamma_1; \ldots; \Gamma_k$$

of sequents Γ_i constructed according to the following rules:

(i) $\Gamma_0 = \emptyset$.

(ii) Γ_i is not axiomatic for $i < k$.

(iii) If $i < k$ and Γ_i is not reducible then

$$\Gamma_{i+1} = \Gamma_i, \neg C(U_i).$$

(iv) Every reducible Γ_i with $i < k$ is of the form

$$\Gamma_i', E, \Gamma_i''$$

where E is not a literal and Γ_i' contains only literals.
E is said to be the **redex** of Γ_i.

Let $i < k$ and Γ_i be reducible. Γ_{i+1} is obtained from $\Gamma_i = \Gamma_i', E, \Gamma_i''$ as follows:

1. If $E \equiv E_0 \vee E_1$ then
$$\Gamma_{i+1} = \Gamma_i', E_0, E_1, \Gamma_i'', \neg C(U_i).$$

2. If $E \equiv E_0 \wedge E_1$ then
$$\Gamma_{i+1} = \Gamma_i', E_j, \Gamma_i'', \neg C(U_i)$$
where $j = 0$ or $j = 1$.

3. If $E \equiv \exists x\, F(x)$ then
$$\Gamma_{i+1} = \Gamma_i', F(\bar{m}), \Gamma_i'', \neg C(U_i), E$$
where m is the first number such that $F(\bar{m})$ does not occur in $\Gamma_0; \ldots; \Gamma_i$.

9

4. If $E \equiv \forall x \, F(x)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', F(\bar{m}), \Gamma_i'', \neg C(U_i)$$
for some m.

5. If $E \equiv \exists X \, F(X)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', F(U_m), \Gamma_i'', \neg C(U_i), E$$
where m is the first number such that $F(U_m)$ does not occur in $\Gamma_0; \ldots; \Gamma_i$.

6. If $E \equiv \forall X \, F(X)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', F(U_m), \Gamma_i'', \neg C(U_i)$$
where m is the first number such that $m \neq i+1$ and U_m does not occur in Γ_i.

The set of deduction chains forms a tree $\mathbb{T}$ labeled with strings of sequents. We will now consider two cases.

Case I: $\mathbb{T}$ is not well-founded. Then $\mathbb{T}$ contains an infinite path $\mathbb{P}$. Now define a set M via
$$(M)_i \;=\; \{ t^{\mathbb{N}} \mid t \notin U_i \text{ occurs in } \mathbb{P} \}.$$

Set $\mathbb{M} = (\mathbb{N}; \{ (M)_i \mid i \in \mathbb{N} \}, +, \cdot, 0, 1, <)$.

For a formula F, let $F \in \mathbb{P}$ mean that F occurs in $\mathbb{P}$, i.e. $F \in \Gamma$ for some $\Gamma \in \mathbb{P}$.

Claim: Under the assignment $U_i \mapsto (M)_i$ we have
$$F \in \mathbb{P} \qquad \Rightarrow \qquad \mathbb{M} \models \neg F.$$

The Claim implies that $\mathbb{M}$ is an ω-model of $\mathbf{ATR_0}$.

Case II: $\mathbb{T}$ is well-founded. We actually want to rule this out. This is the place where the principle
$$\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\Gamma_{\mathfrak{X}})]$$
in the guise of cut elimination for an infinitary proof system enters the stage. Aiming at a contradiction, suppose that $\mathbb{T}$ is a well-founded tree. Let $\mathfrak{X}_0$ be the Kleene-Brouwer ordering on $\mathbb{T}$. Then $\mathfrak{X}_0$ is a well-ordering. In a nutshell, the idea is that a well-founded $\mathbb{T}$ gives rise to a derivation of the empty sequent (contradiction) in the infinitary proof systems $\mathcal{T}_Q^\infty$ from [42]. This is were the really hard work lies and we have to stop here; details are in [52].

3 Toward impredicative theories

The proof-theoretic ordinal functions that figure in the foregoing theorems are all familiar from so-called predicative or meta-predicative proof theory. Thus far a function from genuinely impredicative proof theory is missing. The first such function that comes to mind is of the Bachmann type [4]. We will shortly turn to it.

Veblen in [69] ventured very far into the transfinite in his 1908 paper, way beyond a representation system that incorporates the Γ-function. He extended the two-place φ-function first to an arbitrary finite numbers of arguments, but then also to a transfinite numbers of arguments, with the proviso that in, for example
$$\Phi_f(\alpha_0, \alpha_1, \ldots, \alpha_\eta),$$

only a finite number of the arguments α_ν may be non-zero. In this way Veblen singled out the ordinal $E(0)$, which is often called the *big Veblen number*. $E(0)$ is the least ordinal $\delta > 0$ which cannot be named in terms of representations

$$\Phi_\ell(\alpha_0, \alpha_1, \ldots, \alpha_\eta)$$

with $\eta < \delta$, and each $\alpha_\gamma < \delta$.

3.1 The Bachmann revelation.

In a paper published in 1950, Heinz Bachmann introduced a novel idea that allowed one to "name" much larger ordinals than by the Veblen procedures. He had the amazing idea of using *uncountable ordinals* to keep track of diagonalizations over a hierarchies of functions. In more detail, he used the following steps:

- Define a set of ordinals $\mathfrak{B}$ closed under successor such that with each limit $\lambda \in \mathfrak{B}$ is associated an increasing sequence $\langle \lambda[\xi] : \xi < \tau_\lambda \rangle$ of ordinals $\lambda[\xi] \in \mathfrak{B}$ of length $\tau_\lambda \in \mathfrak{B}$ and $\lim_{\xi < \tau_\lambda} \lambda[\xi] = \lambda$.

- Let Ω be the *first uncountable ordinal.* A hierarchy of functions $(\varphi_\alpha^{\mathfrak{B}})_{\alpha \in \mathfrak{B}}$ is then obtained as follows:

$$\varphi_0^{\mathfrak{B}}(\beta) = \omega^\beta \qquad \varphi_{\alpha+1}^{\mathfrak{B}} = \left(\varphi_\alpha^{\mathfrak{B}}\right)' \tag{3}$$

$$\varphi_\lambda^{\mathfrak{B}} \text{ enumerates } \bigcap_{\xi < \tau_\lambda} (\text{Range of } \varphi_{\lambda[\xi]}^{\mathfrak{B}}) \qquad \lambda \text{ limit, } \tau_\lambda < \Omega$$

$$\varphi_\lambda^{\mathfrak{B}} \text{ enumerates } \{\beta < \Omega : \varphi_{\lambda[\beta]}^{\mathfrak{B}}(0) = \beta\} \qquad \lambda \text{ limit, } \tau_\lambda = \Omega. \tag{4}$$

Distilling Bachmann's idea. What makes Bachmann's approach rather difficult to deal with in proof theory is the requirement to endow every limit ordinal with a fundamental sequence and referring to them when defining φ_λ, notably in the diagonalization procedure enshrined in (4). This layer of complication can be dispensed with though. What underpins the strength of Bachmann's approach can be described without the book-keeping of fundamental sequences. We start by imagining a "big" ordinal Ω. What this means will become clearer as we go along, but definitely Ω should be an ε-number. Using ordinals $< \Omega$ and Ω itself as building blocks, one then constructs further ordinals using Cantor's normal form, i.e., if $\alpha_1 \geq \ldots \geq \alpha_n$ have already been constructed, then we build $\alpha := \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$ provided that $\alpha > \alpha_1$. In this way we can build all ordinals $< \varepsilon_{\Omega+1}$, where the latter ordinal denotes the next ε-number after Ω. Conversely, we can take any $\alpha < \varepsilon_{\Omega+1}$ apart, yielding smaller pieces as long as the exponents in its Cantor normal are smaller ordinals. This leads to the idea of support. More precisely define:

Definition 3.1 (i) $\mathsf{supp}_\Omega(0) = \emptyset$, $\mathsf{supp}_\Omega(\Omega) = \emptyset$.

(ii) $\mathsf{supp}_\Omega(\alpha) = \mathsf{supp}_\Omega(\alpha_1) \cup \cdots \cup \mathsf{supp}_\Omega(\alpha_n)$ if $\alpha =_{CNF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} > \alpha_1$.

(iii) $\mathsf{supp}_\Omega(\alpha) = \{\alpha\}$ if α is an ε-number $< \Omega$.
Note that $\mathsf{supp}_\Omega(\alpha)$ is a finite set.

To define something that is equivalent to what Bachmann achieves in (4), the central idea is to devise an injective function

$$\vartheta \colon \varepsilon_{\Omega+1} \to \Omega$$

such that each $\vartheta(\alpha)$ is an ε-number. Think of ϑ as a *collapsing*, or better *projection* function in the sense of admissible set theory. For obvious reasons, ϑ cannot be order preserving, but the following can be realized:

$$\alpha < \beta \wedge \mathsf{supp}_\Omega(\alpha) < \vartheta(\beta) \;\rightarrow\; \vartheta(\alpha) < \vartheta(\beta).$$

With the help of ϑ-function[9] one obtains an ordinal representation system for the so-called *Bachmann ordinal*, also referred to as the *Bachmann-Howard ordinal*.[10]

Definition 3.2 We inductively define a set $\mathsf{OT}(\vartheta)$.

(i) $0 \in \mathsf{OT}(\vartheta)$ and $\Omega \in \mathsf{OT}(\vartheta)$.

(ii) If $\alpha_1, \ldots, \alpha_n \in \mathsf{OT}(\vartheta)$, $\alpha_1 \geq \cdots \geq \alpha_n$, then $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \in \mathsf{OT}(\vartheta)$.

(iii) If $\alpha \in \mathsf{OT}(\vartheta)$ then so is $\vartheta(\alpha)$.

$(\mathsf{OT}(\vartheta), <)$ gives rise to an elementary ordinal representation system. Here $<$ denotes the restriction to $\mathsf{OT}(\vartheta)$.

The Bachmann-Howard ordinal is the order-type of $\mathsf{OT}(\vartheta) \cap \Omega$.

3.2 Associating a Dilator with Bachmann.

The Bachmann construction can be relativized to an arbitrary linear ordering $\mathfrak{X}$, as was shown in [53], giving rise to a dilator

$$\mathfrak{X} \mapsto \vartheta_{\mathfrak{x}}$$

and the wellordering principle

$$\mathrm{WO}(\mathfrak{X}) \Rightarrow \mathrm{WO}(\vartheta_{\mathfrak{x}}). \tag{5}$$

Definition 3.3 [53, 2.6] Again, let Ω be a "big" ordinal. Let $\mathfrak{X}$ be a well-ordering. With each $x \in X$ we associate a ε-number $\mathfrak{E}_x > \Omega$. The set of ordinals $\mathsf{OT}_{\mathfrak{x}}(\vartheta)$ is inductively defined as follows:

(i) $0 \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$, $\Omega \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$, and $\mathfrak{E}_x \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$ when $x \in X$.

(ii) If $\alpha_1, \ldots, \alpha_n \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$, $\alpha_1 \geq \cdots \geq \alpha_n$, then $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$.

(iii) If $\alpha \in \mathsf{OT}_{\mathfrak{x}}(\vartheta)$ then so is $\vartheta_{\mathfrak{x}}(\alpha)$ and $\vartheta_{\mathfrak{x}}(\alpha) < \Omega$.

To explain the ordering on $\mathsf{OT}_{\mathfrak{x}}(\vartheta)$ one needs to extend supp_Ω: Let $\mathsf{supp}_\Omega^{\mathfrak{x}}(\mathfrak{E}_x) = \emptyset$. One then sets

1. $\mathfrak{E}_x < \mathfrak{E}_y \leftrightarrow x <_{\mathfrak{x}} y$.

2. $\vartheta_{\mathfrak{x}}(\alpha) < \vartheta_{\mathfrak{x}}(\beta) \leftrightarrow ([\alpha < \beta \wedge \mathsf{supp}_\Omega^{\mathfrak{x}}(\alpha) < \vartheta_{\mathfrak{x}}(\beta)] \vee [\exists \gamma \in \mathsf{supp}_\Omega^{\mathfrak{x}}(\beta)\, \vartheta_{\mathfrak{x}}(\alpha) \leq \gamma])$.

$(\mathsf{OT}_{\mathfrak{x}}(\vartheta), <)$ yields an ordinal representation system elementary in $\mathfrak{X}$.

---

[9]This function is a cousin of the θ-function whose definition and properties are owed to Feferman, Aczel, Buchholz, and Bridge (see [63, IX]) . ϑ was first defined in [44] and used in [54].

[10]Bill Howard was looking for a description of the proof-theoretic ordinal of the theory of non-iterated inductive definitions. He was amazed and delighted to find it in Bachmann's paper [4].

The principle (5) turns ot to be equivalent over $\mathbf{RCA}_0$ to asserting that every set is contained in a countable coded ω-model of the theory of *Bar Induction*, $\mathbf{BI}$, whose formalization requires some preparations.

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

$\text{Prog}(\prec, F) := (\forall x)[\forall y(y \prec x \to F(y)) \to F(x)]$ (*progressiveness*)

$\mathbf{TI}(\prec, F) := \text{Prog}(\prec, F) \to \forall x F(x)$ (*transfinite induction*)

$\text{WF}(\prec) := \forall X \mathbf{TI}(\prec, X) :=$
$\forall X(\forall x[\forall y(y \prec x \to y \in X)) \to x \in X] \to \forall x[x \in X])$ (*well-foundedness*).

Let $\mathcal{F}$ be any collection of formulae of $\mathcal{L}_2$. For a 2-place relation $\prec$ we will write $\prec \in \mathcal{F}$, if $\prec$ is defined by a formula $Q(x, y)$ of $\mathcal{F}$ via $x \prec y := Q(x, y)$.

Definition 3.4 BI denotes the bar induction scheme, i.e. all formulae of the form

$$\text{WF}(\prec) \to \mathbf{TI}(\prec, F),$$

where $\prec$ is an arithmetical relation (set parameters allowed) and F is an arbitrary formula of $\mathcal{L}_2$.

By $\mathbf{BI}$ we shall refer to the theory $\mathbf{ACA}_0 + \text{BI}$.

The theorem alluded to above, due to P.F. Valencia Vizcaíno and the author, conjectured in [55, Conjecture 7.2], is the following:

Theorem 3.5 (Rathjen, Valencia Vizcaíno [53]) *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

1. $\forall \mathfrak{X} \, [\text{WO}(\mathfrak{X}) \to \text{WO}(\text{OT}_{\mathfrak{X}}(\vartheta))]$.

2. *Every set is contained in a countable coded ω-model of* $\mathbf{BI}$.

4 Towards a general theory of ordinal representations

We have seen several ordinal representation systems and their relativizations to an arbitrary well-ordering $\mathfrak{X}$. Ordinal representation systems understood in this relativized way give rise to well-ordering principles:

$$(*) \quad \text{"if } \mathfrak{X} \text{ is a well-ordering then so is } \mathcal{OR}_{\mathfrak{X}}\text{"}$$

where $\mathcal{OR}_{\mathfrak{X}}$ arises from $\mathcal{OR}$ by letting $\mathfrak{X}$ play the role of the order-indiscernibles. Any principle of the form $(*)$ is of syntactic complexity Π_2^1; thus cannot characterize stronger comprehensions such as Π_1^1-comprehension.[11] We are therefore led to the idea of higher-order wellordering principles. To this end we require a general theory of ordinal representation systems. A crucial step towards a general theory via an axiomatic approach was taken by Feferman in [11], who singled out the property of *effective relative categoricity* as central to ordinal representation systems. Through the notion of relative categoricity he succeeded in crystallizing the algebraic aspect of ordinal representation systems by way of relativizing them to any set of order-indiscernibles, thereby in effect scaling them up to functors on the category of linear orders with order-preserving maps as morphisms. Below we recall his approach in [11].

---

[11]Π_1^1-comprehension is syntactically of complexity Π_3^1. If it were equivalent to a Π_2^1 statement, say on the basis of $\mathbf{ATR}_0$, then it would follow that Π_1^1-$\mathbf{CA}_0$ proves its own consistency, e.g. by [64, Theorem VIII.5.12].

4.1 Feferman's relative categoricity

Definition 4.1 Let $f_1, \ldots, f_n$ be function symbols with f_i having arity m_i.

The set of terms $Tm(f_1, \ldots, f_n)$ is defined inductively as follows:

(i) $\bar{0} \in Tm(f_1, \ldots, f_n)$;

(ii) each variable x_i is in $Tm(f_1, \ldots, f_n)$;

(iii) if $t_1, \ldots, t_{m_i} \in Tm(f_1, \ldots, f_n)$, then so also

$$\mathsf{f}_i(t_1, \ldots, t_{m_i}).$$

Let Ord denote the class of ordinals. Now suppose given a system of functions $F_i \colon \mathsf{Ord}^{m_i} \to \mathsf{Ord}$ (where $1 \le i \le n$). By an *assignment* we mean any function $\sigma \colon \omega \to \mathsf{Ord}$. With each assignment σ and any $t \in Tm(f_1, \ldots, f_n)$ is associated an ordinal $|t|_\sigma$, determined as follows.

1. $|\bar{0}|_\sigma = 0$;

2. $|x_k| = \sigma(k)$;

3. $|f_i(t_1, \ldots, t_{m_i})|_\sigma = F_i(|t_1|_\sigma, \ldots, |t_{m_i}|_\sigma)$.

The system $\vec{F}$ is said to be **replete** if for every ordinal α the closure of α under $\vec{F}$ is an ordinal.

An ordinal $\kappa > 0$ is said to be $\vec{F}$-*inaccessible* if whenever $\alpha_1, \ldots, \alpha_{m_i} < \rho$ then $F_i(\alpha_1, \ldots, \alpha_{m_i}) < \rho$ holds for all F_i. $\vec{F}$ is **effectively relatively categorical** if, roughly speaking the order relation between any two terms $s(x_1, \ldots, x_k)$, $t(x_1, \ldots, x_k)$ can be effectively determined from the ordering among $x_1, \ldots, x_k$ provided that these all represent $\vec{F}$-inaccessibles. In particular for assignments σ, τ into $\vec{F}$-inaccessibles it means that if

$$\forall i, j \in [1, \ldots, n](\sigma(x_i) < \sigma(x_j) \leftrightarrow \tau(x_i) < \tau(x_j))$$

then

$$
\begin{aligned}
|s(x_1, \ldots, x_k)|_\sigma < |t(x_1, \ldots, x_k)|_\sigma &\leftrightarrow |s(x_1, \ldots, x_k)|_\tau < |t(x_1, \ldots, x_k)|_\tau \\
|t(x_1, \ldots, x_k)|_\sigma < |s(x_1, \ldots, x_k)|_\sigma &\leftrightarrow |t(x_1, \ldots, x_k)|_\tau < |s(x_1, \ldots, x_k)|_\tau \\
|s(x_1, \ldots, x_k)|_\sigma = |t(x_1, \ldots, x_k)|_\sigma &\leftrightarrow |s(x_1, \ldots, x_k)|_\tau = |t(x_1, \ldots, x_k)|_\tau.
\end{aligned}
$$

In his paper [12], with the title *Hereditarily replete functionals over the ordinals*, Feferman extended the foregoing notions also to finite type functionals over ordinals.

4.2 Girard's dilators

A general theory of ordinal representation systems was also initiated by Girard who coined the notion of *dilator* [24].

Let $\mathbb{ORD}$ be the category whose objects are the ordinals and whose arrows are the strictly increasing functions between ordinals. $\mathbb{ORD}$ enjoys two basic properties.

Lemma 4.2 *(i)* $\mathbb{ORD}$ *has pullbacks.*

(ii) Every ordinal is a direct limit $\varinjlim(n_p, f_{pq})_{p \subseteq q \in I}$ of a system of finite ordinals n_p with I being a set of finite sets of ordinals.

Proof: (i): Let $f\colon \alpha \to \gamma$ and $g\colon \beta \to \gamma$ be in $\mathbb{ORD}$. One easily checks that any $h\colon \delta \to \gamma$ such that

$$\mathrm{ran}(h) \quad = \quad \mathrm{ran}(f) \cap \mathrm{ran}(g) \tag{6}$$

is a pullback of $f\colon \alpha \to \gamma$ and $g\colon \beta \to \gamma$. Note that $h\colon \delta \to \gamma$ is uniquely determined by (6) and that such an h always exists: Let δ be the Mostowski collapse of $\mathrm{ran}(f) \cap \mathrm{ran}(g)$, $\mathsf{clpse}_{\mathrm{ran}(f)\cap\mathrm{ran}(g)}$ be the pertaining collapsing function, and h be the inverse of the collapsing function.

(ii): Fix an ordinal α. Let I be the collection of finite subsets of α. I is ordered by the inclusion relation, which is also directed, i.e. for $s, t \in I$ there exists $q \in I$ such that $s \subseteq q$ and $t \subseteq q$. Let $s \in I$. Then $s = \{\alpha_0, \ldots, \alpha_{n-1}\}$ for uniquely determined ordinals $\alpha_0 < \ldots < \alpha_{n-1}$. Define $n_s := n$ and $f_s\colon n_s \to s$ by $f_s(i) = \alpha_i$. By design, f_s is an order-preserving bijection between n_s and s. For $p \subseteq q \in I$ define $f_{pq}\colon n_p \to n_q$ by $f_{pq} := (f_q)^{-1} \circ f_p$. Thus, f_{pp} is the identity on n_p, and if $p \subseteq q \subseteq r$, one has $f_{pr} = f_{qr} \circ f_{pq}$. As a result,

$$\alpha = \varinjlim(n_p, f_{pq})_{p \subseteq q \in I}.$$

$\square$

Of special interest are the endofunctors of this category that can be characterized by the two foregoing notions.

Definition 4.3 A *dilator* is an endofunctor of the category $\mathbb{ORD}$ preserving direct limits and pullbacks.

Dilators can be characterized in straightforward way.

Lemma 4.4 *For a functor $F\colon \mathbb{ORD} \to \mathbb{ORD}$, the following are equivalent:*

(i) F is a dilator.

(ii) F has the following two properties:

    1. Whenever $f\colon \alpha \to \gamma$, $g\colon \beta \to \gamma$, $h\colon \delta \to \gamma$, then

$$\mathrm{ran}(h) = \mathrm{ran}(f) \cap \mathrm{ran}(g) \quad \Rightarrow \quad \mathrm{ran}(F(h)) = \mathrm{ran}(F(f)) \cap \mathrm{ran}(F(g)) \tag{7}$$

    2. For every ordinal α and $\eta < F(\alpha)$ there exists a finite ordinal n and strictly increasing function $f\colon n \to \alpha$ such that $\eta \in \mathrm{ran}(F(f))$.

Proof: It follows from the proof of Lemma 4.2(i) that preservation of pullbacks and condition (ii)(1) are equivalent. It remains to show that in $\mathbb{ORD}$ preservation of direct limits and condition (ii)(2) amount to the same. This can be gleaned from the proof of Lemma 4.2(ii), however, as the details are not relevant for this paper, we just refer to [24] for the details. $\square$

There is an easy but important consequence of the above, i.e. of Lemma 4.2(ii).

Lemma 4.5 *A dilator is completely determined by its behavior on the subcategory of finite ordinals and morphisms between them, $\mathbb{ORD}_\omega$.*

At this point, it is perhaps not immediately gleanable what the connection between ordinal representation systems and dilators might be. This is what we turn to next.

Denotation systems and dilators   Girard (cf. [24, 25, 27]) provided an alternative account of dilators that assimilates them more closely to ordinal representation systems and in particular to Feferman's notion of relative categoricity.

Definition 4.6 Let ON be the class of ordinals and $F : \mathrm{ON} \to \mathrm{ON}$. A *denotation-system* for F is a class $\mathcal{D}$ of ordinal *denotations* of the form

$$(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$$

together with an assignment $D : \mathcal{D} \to \mathrm{ON}$ such that the following hold:

1. If $(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$ is in $\mathcal{D}$, then $\alpha_0 < \ldots < \alpha_{n-1} < \alpha$ and $D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha) < F(\alpha)$.

2. Every $\beta < F(\alpha)$ has a unique denotation $(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$ in $\mathcal{D}$, i.e. $\beta = D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$.

3. If $(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$ is a denotation and $\gamma_0 < \ldots < \gamma_{n-1} < \gamma$, then $(c; \gamma_0, \ldots, \gamma_{n-1}; \gamma)$ is a denotation.

4. If $D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha) \leq D(d; \alpha'_0, \ldots, \alpha'_{m-1}; \alpha)$, $\gamma_0 < \ldots < \gamma_{n-1} < \gamma$, $\gamma'_0 < \ldots < \gamma'_{m-1} < \gamma$, and for all $i < n$ and $j < m$,

   $$\alpha_i \leq \alpha'_j \Leftrightarrow \gamma_i \leq \gamma'_j \quad \text{and} \quad \alpha_i \geq \alpha'_j \Leftrightarrow \gamma_i \geq \gamma'_j$$

   then

   $$D(c; \gamma_0, \ldots, \gamma_{n-1}; \gamma) \leq D(d; \gamma'_0, \ldots, \gamma'_{m-1}; \gamma).$$

In a denotation $(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$, c is called the *index*, α is the *parameter* and $\alpha_0, \ldots, \alpha_{n-1}$ are the **coefficients** of the denotation. If $\beta = D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$, the index c represents some 'algebraic' way of describing β in terms of the ordinals $\alpha_0, \ldots, \alpha_{n-1}, \alpha$.

Lemma 4.7 *Every denotation system D induces a dilator F_D by letting $F_D(\alpha)$ be the least ordinal η that does not have a denotation of the form $D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$, and for any arrow $f : \alpha \to \delta$ of the category $\mathbb{ORD}$ letting*

$$F_D(f) : F_D(\alpha) \to F_D(\delta)$$

be defined by

$$F_D(f)(D(c; \alpha_0, \ldots, \alpha_{n-1}; \alpha)) := D(c; f(\alpha_0), \ldots, f(\alpha_{n-1}); \delta).$$

The converse is also true.

Lemma 4.8 *To every dilator F one can assign a denotation system D_F such that $\beta < F(\alpha)$ is denoted by*

$$(\gamma; \alpha_0, \ldots, \alpha_{n-1}; \alpha)$$

where n is the least finite ordinal such that there exists a morphism $f : n \to \alpha$ with $\beta \in \mathrm{ran}(F(f))$. Moreover, $\gamma < F(n)$ is uniquely determined by $F(f)(\gamma) = \beta$, and $\alpha_0 = f(0), \ldots, \alpha_{n-1} = f(n-1)$.

Proof: *F* being a dilator, we know that such an *f* exists. *f* is also uniquely determined since n is chosen minimal: for if $g : n \to \alpha$ also satisfied $\beta \in \mathrm{ran}(F(g))$, then their joint pullback $h : m \to \alpha$ would satisfy $\beta \in \mathrm{ran}(h)$ as well, so that $n = m$ and $\mathrm{ran}(h) = \mathrm{ran}(f) = \mathrm{ran}(g)$, meaning $f = g$. $\qquad\square$

5 Higher order wellordering principles

We already noted that a statement of the form $\mathbf{WOP}(f)$ is Π^1_2 and therefore cannot be equivalent to a theory whose axioms have an essentially higher complexity, like for instance Π^1_1-comprehension. The idea thus is that to get equivalences one has to climb up in the type structure. Given a functor

$$F : (\mathbb{LO} \to \mathbb{LO}) \to (\mathbb{LO} \to \mathbb{LO}),$$

where $\mathbb{LO}$ is the class of linear orderings, we consider the statement:[12]

$$\mathbb{HWOP}(F): \qquad \forall f \in (\mathbb{LO} \to \mathbb{LO}) \, [\mathbf{WOP}(f) \to \mathbf{WOP}(F(f))].$$

There is also a variant of $\mathbb{HWOP}(F)$ which should basically encapsulate the same "power". Given a functor

$$G : (\mathbb{LO} \to \mathbb{LO}) \to \mathbb{LO}$$

consider the statement:

$$\mathbb{HWOP}_1(G): \qquad \forall f \in (\mathbb{LO} \to \mathbb{LO}) \, [\mathbf{WOP}(f) \to \mathrm{WO}(G(f))].$$

It was conjectured in [38] that a principle of the form $\mathbb{HWOP}(F)$ might be equivalent to Π^1_1-comprehension. In [56] is was claimed that for a specific functor $\mathcal{B} : (\mathbb{LO} \to \mathbb{LO}) \to \mathbb{LO}$, $\mathbb{HWOP}_1(\mathcal{B})$ is equivalent to Π^1_1-comprehension. $\mathcal{B}$ is a functor that takes an arbitrary dilator F as input and returns an ordinal representation system; this amounts to combining the Bachmann procedure with the closure under F. [56] also adumbrated the steps and a proof strategy for this result.

5.1 Bachmann meets a dilator

The idea to combine the Bachmann construction with an arbitrary dilator one finds in [38] and [56]. The details were worked out in Anton Freund's thesis [13].[13]

Definition 5.1 Again, let Ω be a "big" ordinal. Let $\mathcal{D}$ be a system of denotations.

(i) $0 \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$, $\Omega \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$.

(ii) If $\alpha_1, \ldots, \alpha_n \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$, $\alpha_1 \geq \cdots \geq \alpha_n$, then $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$.

(iii) If $\alpha \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$ then so is $\vartheta_{\mathcal{D}}(\alpha)$ and $\vartheta_{\mathcal{D}}(\alpha) < \Omega$.

(iv) If $\alpha_0, \ldots, \alpha_{n-1} \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$, $\alpha_0 < \cdots < \alpha_{n-1} < \Omega$, and $\sigma = (c; 0, \ldots, n-1; n) \in \mathcal{D}$, then
$$\mathfrak{E}^{\sigma}_{\alpha_0, \ldots, \alpha_{n-1}} \in \mathsf{OT}_{\mathcal{D}}(\vartheta) \text{ and } \Omega < \mathfrak{E}^{\sigma}_{\alpha_0, \ldots, \alpha_{n-1}}.$$

To explain the ordering on $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ one needs to extend supp_{Ω}: Let

$$\mathsf{supp}^{\mathcal{D}}_{\Omega}(\mathfrak{E}^{\sigma}_{\alpha_0, \ldots, \alpha_{n-1}}) = \{\alpha_0, \ldots, \alpha_{n-1}\}.$$

The ordering between $\vartheta_{\mathcal{D}}$-terms is defined as before:

$$\vartheta_{\mathcal{D}}(\alpha) < \vartheta_{\mathcal{D}}(\beta) \leftrightarrow ([\alpha < \beta \wedge \mathsf{supp}^{\mathcal{D}}_{\Omega}(\alpha) < \vartheta_{\mathcal{D}}(\beta)] \vee [\exists \gamma \in \mathsf{supp}^{\mathcal{D}}_{\Omega}(\beta) \, \vartheta_{\mathcal{D}}(\alpha) \leq \gamma]).$$

---

[12]$\mathbb{HWOP}$ being an acronym for the German "**H**öherstufiges **W**ohl**o**rdnungs**p**rinzip".

[13]Of course, Definition 3.3, which harks back to [53], set an important precedent.

How do we compare terms of the form $\mathfrak{E}^{\sigma}_{\alpha_0,\ldots,\alpha_{n-1}}$ and $\mathfrak{E}^{\tau}_{\beta_0,\ldots,\beta_{m-1}}$? To this end let $\sigma = (c; 0, \ldots, n-1; n)$ and $\tau = (e; 0, \ldots, m-1; m)$; also let k be the number of elements of $\{\alpha_0, \ldots, \alpha_{n-1}, \beta_0, \ldots, \beta_{m-1}\}$. In a first step, define $f\colon n \to k$ and $g\colon m \to k$ such that

$$f(i) < g(j) \quad \leftrightarrow \quad \alpha_i < \beta_j$$
$$f(i) = g(j) \quad \leftrightarrow \quad \alpha_i = \beta_j.$$

Then

$$\mathfrak{E}^{\sigma}_{\alpha_0,\ldots,\alpha_{n-1}} \quad < \quad \mathfrak{E}^{\tau}_{\beta_0,\ldots,\beta_{m-1}}$$

if and only if

$$D(c; f(0), \ldots, f(n-1); k) \quad < \quad D(e; g(0), \ldots, g(m-1); k).$$

It is rather rather instructive to see why $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ is a wellordering. We shall prove that with the help of strong comprehension.

Theorem 5.2 (Π^1_1-**CA**$_0$) *For every dilator $\mathcal{D}$, $(\mathsf{OT}_{\mathcal{D}}(\vartheta), <)$ is a wellordering.*

Proof: Let I be the well-founded part of $\mathsf{OT}_{\mathcal{D}}(\vartheta) \cap \Omega := \{\alpha \in \mathsf{OT}_{\mathcal{D}}(\vartheta) \mid \alpha < \Omega\}$. Π^1_1-comprehension ensures that I is a set. Then $\mathfrak{I} := (I, < \restriction I \times I)$ is a wellordering. Now let

$$M := \{\beta \in \mathsf{OT}_{\mathcal{D}}(\vartheta) \mid \mathsf{supp}^{\mathcal{D}}_{\Omega}(\beta) \subseteq I\} \ \text{ and } \ \mathfrak{M} := (M, < \restriction M \times M).$$

We claim that $\mathfrak{M}$ is a wellordering, too. To see this, note that the set of terms of the form $\mathfrak{E}^{\sigma}_{\alpha_0,\ldots,\alpha_{n-1}}$ with $\alpha_0, \ldots, \alpha_{n-1} \in I$ is wellordered as they are in one-one and order preserving correspondence with the denotations $D(c; \alpha_0, \ldots, \alpha_{n-1}; \mathfrak{I})$ where $\sigma = (c; 0, \ldots, n-1; n)$. Now, M is obtained from I by adding the latter terms as well as Ω, and then closing off under $+$ and $\xi \mapsto \omega^{\xi}$ (or more precisely, Cantor normal forms). It is known from Gentzen's [23] that this last step preserves wellorderedness.[14]

Thus, if we can show that $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ is the same as $\mathfrak{M}$ we are done. For this it is enough to establish closure of M under $\vartheta_{\mathcal{D}}$.

$$\mathsf{Claim}: \quad \forall \alpha \in M \ \vartheta_{\mathcal{D}}(\alpha) \in M.$$

The Claim is proved by induction on α. So assume $\alpha \in M$ and

$$(*) \quad \forall \beta \in M \ [\beta < \alpha \to \vartheta_{\mathcal{D}}(\beta) \in I].$$

To be able to conclude that $\vartheta_{\mathcal{D}}(\alpha) \in I$, it suffices to show that $\delta \in I$ holds for all $\delta < \vartheta_{\mathcal{D}}(\beta)$. We establish this via a subsidiary induction on the syntactic complexity of δ (so this is in effect an induction on naturals). If $\delta = 0$ this is immediate, and if $\delta = \omega^{\delta_1} + \ldots + \omega^{\delta_n}$ with $\delta > \delta_1 \geq \ldots \geq \delta_n$ this follows from the subsidiary induction hypothesis and the fact that I is closed under $+$ and $\xi \mapsto \omega^{\xi}$.

This leaves the case that $\delta = \vartheta_{\mathcal{D}}(\rho)$ for some ρ. Since $\vartheta_{\mathcal{D}}(\rho) < \vartheta_{\mathcal{D}}(\alpha)$ there are two subcases to consider.

Case 1: $\rho < \alpha$ and $\mathsf{supp}^{\mathcal{D}}_{\Omega}(\rho) < \vartheta_{\mathcal{D}}(\alpha)$. By the subsidiary induction hypothesis, $\mathsf{supp}^{\mathcal{D}}_{\Omega}(\rho) \subseteq I$ as the syntactic complexity of terms in $\mathsf{supp}^{\mathcal{D}}_{\Omega}(\rho)$ is not bigger than that of ρ and thus smaller than that of δ. Therefore, $\rho \in M$, and since $\rho < \alpha$ it follows from $(*)$ that $\delta = \vartheta_{\mathcal{D}}(\rho) \in I$.

Case 2: There exists $\xi \in \mathsf{supp}^{\mathcal{D}}_{\Omega}(\alpha)$ such that $\delta \leq \xi$. $\alpha \in M$ entails that $\xi \in I$, and thus $\delta \in I$. $\qquad\qquad\square$

---

[14]For details, consult [63, VIII.2], [40, 9.6.2] or [50, Section 4].

Theorem 5.3 (Freund) *The following are equivalent over* **RCA**$_0$:

(i) Π_1^1-*comprehension.*

(ii) *For every denotation system* $\mathcal{D}$, $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ *is a wellordering.*

A detailed proof was given in Anton Freund's thesis [13] and in his papers [14, 15, 16]. Below, however, we will be presenting another proof, partly going back to [56], since the purposes here is to convey the intuitions behind the result to the reader and present them without employing category-theoretic terminology.

The direction (i)$\Rightarrow$(ii) of Theorem 5.3 is taken care of by Theorem 5.2. Next, we will be concerned with giving a proof sketch for the direction ($*$) (ii)$\Rightarrow$(i). For ($*$) it is enough to show that every set $\bar{Q} \subseteq \mathbb{N}$ is contained in a countable coded model $\mathfrak{M} = (M, E)$ of Kripke-Platek set theory (with Infinity), where the interpretation E of the elementhood relation is well-founded. This is so because a Π_1^1-definable class of naturals, $\{n \in \mathbb{N} \mid F(n, \bar{Q})\}$, with second order parameter $\bar{Q}$ is Σ_1-definable over $\mathfrak{M}$ (see [5, IV.3.1]), and therefore a set in the background theory. Important techniques include Schütte's method of search trees [59] for ω-logic but adjusted to α-logic for any ordinal (or rather well-ordering) α and proof calculi similar to the one used for the ordinal analysis of Kripke-Platek set theory (originally due to Jäger [32, 33]). It is not necessary, though, to use the machinery for the ordinal analysis of **KP**. Instead, one can also directly employ the older techniques from ordinal analyses of the formal theory of non-iterated arithmetic inductive definitions as presented in [39] and [40]. Technically, though, the easiest approach appears to be to work with an extension of Peano arithmetic with inductive definitions and an extra sort of ordinals to express the stages of an inductive definition as in Jäger's **PA**$_\Omega$ from [34].

Definition 5.4 In what follows, we fix a set $\bar{Q} \subseteq \mathbb{N}$ and a well-ordering $\mathfrak{X} = (X, <_X)$ with $X \subseteq \mathbb{N}$. Let L_1 be the first-order language of arithmetic with number variables $x, y, z, \ldots$ (possibly with subscripts), the constant 0, as well as function and relation symbols for all primitive recursive functions and relations, and a unary predicate symbol Q. Let $L_1(P)$ be the extension of L_1 by a further "fresh" unary relation symbol P. The *atomic formulas* of $L_1(P)$ are of the form $Q(t)$, $P(t)$ and $R(s_1, \ldots, s_k)$, where $t, s_1, \ldots, s_k$ are terms and R is a relation symbol for a primitive recursive k-ary relation. The *literals* of $L_1(P)$ are the atomic formulas and their negations $\neg Q(t)$, $\neg P(t)$ and $\neg R(s_1, \ldots, s_k)$. The *formulas* of $L_1(P)$ are then generated from the literals via the logical connectives $\wedge, \vee$ and the quantifiers $\forall x$ and $\exists x$. Note that the formulas are in negation normal form. As per usual, negation is a defined operation, using deletion of double negations and de Morgan's laws. A formula is said to be *P-positive* if it contains no occurrences of the form $\neg P(t)$. A P-positive formula in which at most the variable x occurs free will be called an *inductive operator form*; expressions $A(P, x)$ are meant to range over such forms.

L_1 will be further extended to a new language $L_{\mathfrak{x}}$ by adding a new sort of ordinal variables $\alpha, \beta, \gamma, \ldots$ (possibly with subscripts), a new binary relation symbol $<$ for the less relation on the ordinals[15] and a binary relation symbol P_A for each inductive operator form $A(P, x)$.[16] For each element u of the wellordering $\mathfrak{X}$ we also introduce an ordinal constant α_u into the language $L_{\mathfrak{x}}$. Ordinal constants inherit an ordering from $\mathfrak{X}$, namely α_u is considered to be smaller than α_v if $u <_X v$. In future, when we talk about the least ordinal constant satisfying a certain condition we are referring to that ordering.

Since subscripts are often a nuisance to humans, we will use overlined ordinal variables, i.e. $\bar{\alpha}, \bar{\beta}, \ldots$, as metavariables for these constants.

---

[15]The usual ordering on the naturals will be notated by $<_N$.

[16]Observe that the place holder predicate P of $L_1(P)$ is not part of the language $L_{\mathfrak{x}}$.

The number terms $s, t, \ldots$ of $L_{\mathfrak{x}}$ are the number terms of L_1; the ordinal terms of $L_{\mathfrak{x}}$ are the ordinal variables and constants. The formulas of $L_{\mathfrak{x}}$ are inductively generated as follows:

1. If R is a k-ary relation symbol of L_1 and $s_1, \ldots, s_k$ are number terms, then $R(s_1, \ldots, s_k)$ is an atomic formula of $L_{\mathfrak{x}}$.

2. $\mathfrak{a} < \mathfrak{b}$, $\mathfrak{a} = \mathfrak{b}$ and $P_A(s, \mathfrak{a})$ are atomic formulas of $L_{\mathfrak{x}}$ whenever $\mathfrak{a}$ and $\mathfrak{b}$ are ordinal terms and s is a number term.

3. Negations of atomic formulas of $L_{\mathfrak{x}}$ are formulas of $L_{\mathfrak{x}}$. The latter together with the atomic formulas make up the *literals* of $L_{\mathfrak{x}}$.

4. If B, C are formulas of $L_{\mathfrak{x}}$, then so are $B \wedge C$ and $B \vee C$.

5. If $F(x)$ is a formula of $L_{\mathfrak{x}}$, then so are $\exists x F(x)$ and $\forall x F(x)$.

6. If $G(\alpha)$ is a formula of $L_{\mathfrak{x}}$, then so are $\exists \alpha G(\alpha)$ and $\forall \alpha F(\alpha)$.

7. If $G(\alpha)$ is a formula of $L_{\mathfrak{x}}$ and $\mathfrak{b}$ is an ordinal term, then so are $(\exists \alpha < \mathfrak{b}) G(\alpha)$ and $(\forall \alpha < \mathfrak{b}) F(\alpha)$.

As per usual, $B \to C$ is an abbreviation for $\neg B \vee C$ (with $\neg$ being the obvious defined operation of negation), and $B \leftrightarrow C$ abbreviates $(B \to C) \wedge (C \to B)$.

In future, we write $P_A^{\mathfrak{a}}(s)$ for $P_A(s, \mathfrak{a})$. We also write $P_A^{<\mathfrak{a}}(s)$ for $(\exists \beta < \mathfrak{a}) P_A^{\beta}(s)$ and $P_A(s)$ for $\exists \alpha\, P_A^{\alpha}(s)$.

Definition 5.5 The axioms of $\mathbf{PA}_{\mathfrak{x}}$ fall into several groups:

1. True literals: These are of three forms.

   (a) Let R be a symbol for a k-ary primitive recursive relation and $s_1, \ldots, s_k$ are closed number terms. $R(s_1, \ldots, s_k)$ is a true literal if $\mathcal{R}(s_1^{\mathbb{N}}, \ldots, s_1^{\mathbb{N}})$ is true in the naturals, where $\mathcal{R}$ denotes the predicate pertaining to R and $s_i^{\mathbb{N}}$ denotes the value of the closed number term s_i in the standard interpretation; if $R(s_1, \ldots, s_k)$ is false under the standard interpretation, then $\neg R(s_1, \ldots, s_k)$ is a true literal.

   (b) $Q(s)$ is a true literal if $s_i^{\mathbb{N}} \in \bar{Q}$, and $\neg Q(t)$ is a true literal if $t_i^{\mathbb{N}} \notin \bar{Q}$, where s, t are closed numeral terms.

   (c) Let α_u, α_v are ordinal constants with $u, v \in \mathfrak{X}$. $\alpha_u < \alpha_v$ is a true literal if $u <_X v$, and otherwise $\neg \alpha_u < \alpha_v$ is a true literal. $\alpha_u = \alpha_u$ is a true literal whereas $\neg \alpha_u = \alpha_v$ is a true literal if $u \neq v$.

2. Stage axioms: These are all sentences of the form $\forall \alpha\, \forall x\, [A(P_A^{<\alpha}, x) \to P_A^{\alpha}(x)]$.

3. Reflection: These are all sentences of the form $\forall x\, [A(P_A, x) \to \exists \alpha\, P_A^{\alpha}(s)]$.

4. Let $A_0, A_1, \ldots$ be a fixed enumeration of all stage and reflection axioms. Note that all of these are closed formulas (sentences) and that none of them contains any ordinal constants.

5.2 Deduction chains in $\mathbf{PA}_x$

Sequents are finite lists of sentences of L_x; they will be notated by upper case Greek letters.[17]

Definition 5.6   (i) A sequent Γ is *axiomatic* if it contains a true *literal*.[18]

(ii) A sequent is *reducible* or a *redex* if it is not axiomatic and contains a formula which is not a literal.

Definition 5.7 A $\mathbf{PA}_x$-*deduction chain* is a finite string

$$\Gamma_0; \Gamma_1; \ldots; \Gamma_k$$

of sequents Γ_i constructed according to the following rules:

(i) $\Gamma_0 = \emptyset$.

(ii) Γ_i is not axiomatic for $i < k$.

(iii) If $i < k$ and Γ_i is not reducible then

$$\Gamma_{i+1} = \Gamma_i, \neg A_i$$

where A_i is the i-th formula in the list from Definition 5.5(4).

(iv) Every reducible Γ_i with $i < k$ is of the form

$$\Gamma_i', E, \Gamma_i''$$

where E is not a literal and Γ_i' contains only literals. E is said to be the *redex* of Γ_i.

Let $i < k$ and Γ_i be reducible. Γ_{i+1} is obtained from $\Gamma_i = \Gamma_i', E, \Gamma_i''$ as follows:

1. If $E \equiv E_0 \vee E_1$ then
$$\Gamma_{i+1} = \Gamma_i', E_0, E_1, \Gamma_i'', \neg A_i.$$

2. If $E \equiv E_0 \wedge E_1$ then
$$\Gamma_{i+1} = \Gamma_i', E_j, \Gamma_i'', \neg A_i$$
   where $j = 0$ or $j = 1$.

3. If $E \equiv \exists x \in \mathbb{N}\, F(x)$ then
$$\Gamma_{i+1} = \Gamma_i', F(\bar{m}), \Gamma_i'', \neg A_i, E$$
   where m is the first number such that $F(\bar{m})$ does not occur in $\Gamma_0; \ldots; \Gamma_i$.[19]

4. If $E \equiv \forall x \in \mathbb{N}\, F(x)$ then
$$\Gamma_{i+1} = \Gamma_i', F(\bar{m}), \Gamma_i'', \neg A_i$$
   for some m.

---

[17]As before, Γ, A denotes the sequent obtained from Γ via appending the formula A to the list Γ. Similarly, $\Gamma, A_1, \ldots, A_r, \Lambda$ stands for the list $C_1, \ldots, C_n, A_1, \ldots, A_r, D_1, \ldots, D_q$ if Γ is the list $C_1, \ldots, C_n$ and Λ is the list $D_1, \ldots, D_q$.

[18]Note that if Γ contains formulas $R(\vec{s})$ and $\neg R(\vec{t})$, where $s_i^{\mathbb{N}} = t_i^{\mathbb{N}}$ and R is a symbol for a primitive recursive relation or equals Q, then Γ is axiomatic.

[19]There is the slightly irksome possibility that x does not occur free in $F(x)$. Then m can be any number for it doesn't matter which number one choses.

5. If $E \equiv \exists \xi\, F(\xi)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', F(\bar{\delta}), \Gamma_i'', \neg A_i, E$$

where the ordinal constant $\bar{\delta}$ is picked as follows. If there occurs an ordinal constant $\bar{\gamma}$ in $\Gamma_0; \ldots; \Gamma_i$ such that $F(\bar{\gamma})$ does not occur in $\Gamma_0; \ldots; \Gamma_i$ then let $\bar{\delta}$ be the least such (with leastness understood in the sense of $\mathfrak{X}$). If, however, for all ordinal constants $\bar{\gamma}$ in $\Gamma_0; \ldots; \Gamma_i$ the formula $F(\bar{\gamma})$ already occurs in $\Gamma_0; \ldots; \Gamma_i$, then $\bar{\delta}$ can be any ordinal constant.[20]

6. If $E \equiv \forall \xi\, F(\xi)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', F(\bar{\delta}), \Gamma_i'', \neg A_i$$

where $\bar{\delta}$ is any ordinal constant.

7. If $E \equiv (\exists \xi < \bar{\alpha})\, F(\xi)$. then
$$\Gamma_{i+1} \;=\; \Gamma_i', \bar{\delta} < \bar{\alpha} \,\wedge\, F(\bar{\delta}), \Gamma_i'', \neg A_i, E$$

where the ordinal constant $\bar{\delta}$ is picked as follows: If there occurs an ordinal constant $\bar{\gamma}$ in $\Gamma_0; \ldots; \Gamma_i$ such that $\bar{\gamma} < \bar{\alpha}$ is true and $F(\bar{\gamma})$ does not occur in $\Gamma_0; \ldots; \Gamma_i$, let $\bar{\delta}$ be the least such. If, however, for all ordinal constants $\bar{\gamma}$ in $\Gamma_0; \ldots; \Gamma_i$ with $\bar{\gamma} < \bar{\alpha}$ true the formula $F(\bar{\gamma})$ already occurs in $\Gamma_0; \ldots; \Gamma_i$, then $\bar{\delta}$ can be any ordinal constant.

8. If $E \equiv (\forall \xi < \bar{\alpha})\, F(\xi)$ then
$$\Gamma_{i+1} \;=\; \Gamma_i', \bar{\delta} < \bar{\alpha} \to F(\bar{\delta}), \Gamma_i'', \neg A_i$$

where $\bar{\delta}$ is any ordinal constant.

9. If $E \equiv P_A^{\bar{\alpha}}(s)$. then
$$\Gamma_{i+1} \;=\; \Gamma_i', A(P_A^{<\bar{\alpha}}, s), \Gamma_i'', \neg A_i.$$

10. Let $E \equiv \neg P_A^{\bar{\alpha}}(s)$. Then
$$\Gamma_{i+1} \;=\; \Gamma_i', \neg A(P_A^{<\bar{\alpha}}, s), \Gamma_i'', \neg A_i.$$

The set of $\mathbf{PA}_{\mathfrak{X}}$-deduction chains forms a tree $\mathbb{T}_{\mathfrak{X}}^{\bar{Q}}$ labeled with strings of sequents. We will now consider two cases.

- **Case I:** There is a wellordering $\mathfrak{X}$ such that $\mathbb{T}_{\mathfrak{X}}^{\bar{Q}}$ is **not** well-founded.

  Then $\mathbb{T}_{\mathfrak{X}}^{\bar{Q}}$ contains an infinite path $\mathbb{P}$.

- **Case II**: All $\mathbb{T}_{\mathfrak{X}}^{\bar{Q}}$ are well-founded.

---

[20] One might be tempted to decree that $\bar{\delta}$ be the first ordinal constant (in the sense of $\mathfrak{X}$) that does not occur in $\Gamma_0; \ldots; \Gamma_i$. However, this is problematic for two reasons. Firstly, $\mathfrak{X}$ might be finite and thus such a constant might not exists. Secondly, and more seriously, such a choice would heavily depend on $\mathfrak{X}$ and would not be natural from a category-theoretic standpoint that demands are purely finitistic and syntactic treatment.

Case I: Let $\mathbb{P}$ be an infinite path through $\mathbb{T}_{\mathfrak{X}}^{\bar{Q}}$. Let $\mathsf{Ord}_{\mathbb{P}}$ be the set of ordinal constants that occur in sentences of $\mathbb{P}$. The language $L_{\mathbb{P}}$ is the restriction of $L_{\mathfrak{X}}$ that has only the ordinal constants in $\mathsf{Ord}_{\mathbb{P}}$. We now define a structure $\mathfrak{M}_{\mathbb{P}}$ for this language. The number-theoretic part of $\mathfrak{M}_{\mathbb{P}}$ is just the standard model of the naturals plus the interpretation of Q as $\bar{Q}$ while the ordinal part has as its universe $\mathsf{Ord}_{\mathbb{P}}$ with the ordering:

$$\bar{\alpha} <_{\mathfrak{M}_{\mathbb{P}}} \bar{\beta} \quad \text{iff} \quad \bar{\alpha} < \bar{\beta} \text{ is a true literal.}$$

Note that $<_{\mathfrak{M}_{\mathbb{P}}}$ is a wellordering because $\mathfrak{X}$ is.

Also the predicates P_A have to be furnished with an interpretation:

$$P_A^{\mathfrak{M}_{\mathbb{P}}}(n, \bar{\beta}) \quad \text{iff} \quad \neg P_A(\bar{\beta}, \bar{n}) \text{ is a formula of a sequent occurring in } \mathbb{P}.$$

The aim is to show that $\mathfrak{M}_{\mathbb{P}}$ models $\mathbf{PA}_{\mathsf{Ord}_{\mathbb{P}}}$.

Definition 5.8 For a sentence F of $L_{\mathbb{P}}$ we write $F \in \mathbb{P}$ if F occurs in a sequent that belongs to $\mathbb{P}$.

Lemma 5.9   *1. $\mathbb{P}$ does not contain any true literals.*

   2. If $\mathbb{P}$ contains $E_0 \vee E_1$ then $\mathbb{P}$ contains E_0 and E_1.

   3. If $\mathbb{P}$ contains $E_0 \wedge E_1$ then $\mathbb{P}$ contains E_0 or E_1.

   4. If $\mathbb{P}$ contains $\exists x F(x)$ then $\mathbb{P}$ contains $F(\bar{n})$ for all n.

   5. If $\mathbb{P}$ contains $\forall x F(x)$ then $\mathbb{P}$ contains $F(\bar{n})$ for some n.

   6. If $\mathbb{P}$ contains $\exists \xi F(\xi)$ then $\mathbb{P}$ contains $F(\bar{\beta})$ for all $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$.

   7. If $\mathbb{P}$ contains $\forall \xi F(\xi)$ then $\mathbb{P}$ contains $F(\bar{\beta})$ for some $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$.

   8. If $\mathbb{P}$ contains $(\exists \xi < \bar{\alpha}) F(\xi)$ then $\mathbb{P}$ contains $F(\bar{\beta})$ for all $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$ such that $\bar{\beta} < \bar{\alpha}$ holds true.

   9. If $\mathbb{P}$ contains $(\forall \xi < \bar{\alpha}) F(\xi)$ then $\mathbb{P}$ contains $F(\bar{\beta})$ for some $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$.

   10. If $\mathbb{P}$ contains $\neg P_A^{\bar{\alpha}}(s)$ then $\mathbb{P}$ contains $\neg A(P^{<\bar{\alpha}}, s)$.

   11. If $\mathbb{P}$ contains $P_A^{\bar{\alpha}}(s)$ then $\mathbb{P}$ contains $A(P^{<\bar{\alpha}}, s)$.

Proof: The proof is routine. For an existential formula note that such a formula will become the redex of a formula infinitely many times as it never vanishes. Thus, for every formula $\exists x F(x)$ in $\mathbb{P}$, the substitution instance $F(\bar{n})$ will be in $\mathbb{P}$ for every n. Likewise, for an existential formula $\exists \xi F(\xi)$ in $\mathbb{P}$, the substitution instance $F(\bar{\zeta})$ will be in $\mathbb{P}$ for every $\bar{\zeta}$ in $\mathsf{Ord}_{\mathbb{P}}$; although not necessarily for all ordinal constants $\bar{\eta}$. Similar consideration apply to bounded existential ordinal quantifiers. $\square$

Definition 5.10 For the next Lemma (or rather its proof) it is convenient to introduce a measure for the complexity of a $L_{\mathbb{P}}$-sentences.

An ordinal constant $\bar{\alpha}$ is of the form α_u for a unique $u \in X$. Define its rank, $rk(\bar{\alpha})$, to be the order-type of the initial segment of $<_X$ determined by u.

Let u_0 be the least element of X with regard to $<_X$.

   1. $|F| = 0$ if F is a literal not containing ordinal constants.

2. $|\bar{\alpha} < \bar{\beta}| = |\neg \bar{\alpha} < \bar{\beta}| = \max(rk(\bar{\alpha}), rk(\bar{\beta})) \cdot \omega$.

3. $|\bar{\alpha} = \bar{\beta}| = |\neg \bar{\alpha} = \bar{\beta}| = \max(rk(\bar{\alpha}), rk(\bar{\beta})) \cdot \omega$.

4. $|P_A(s, \bar{\alpha})| = |\neg P_A(s, \bar{\alpha})| = rk(\bar{\alpha}) \cdot \omega + \omega$.

5. $|B \wedge C| = |B \vee C| = \max(|B|, |C|) + 1$.

6. $|\exists x F(x)| = |\forall x F(x)| = |F(\bar{0})| + 1$.

7. $|\exists \xi F(\xi)| = |\forall \xi F(\xi)| = \max(\tau \cdot \omega, |F(\alpha_{u_0})| + 1)$ where τ is the order-type of $\mathfrak{X}$.

8. $|(\exists \xi < \bar{\alpha})F(\xi)| = |(\forall \xi < \bar{\alpha})F(\xi)| = \max(rk(\bar{\alpha}) \cdot \omega, |F(\alpha_{u_0})| + 1)$.

Lemma 5.11 *If $F \in \mathbb{P}$ then $\mathfrak{M}_{\mathbb{P}} \models \neg F$.*

Proof: The proof proceeds by induction on $|F|$. The claim is obvious for literals as $\mathbb{P}$ does not contain true literals.

If F is of the form $\neg P_A(s, \bar{\alpha})$ then, by definition of $\mathfrak{M}_{\mathbb{P}}$, we have $\mathfrak{M}_{\mathbb{P}} \models P_A(s, \bar{\alpha})$, whence $\mathfrak{M}_{\mathbb{P}} \models \neg F$.

Let be F is of the form $P_A(s, \bar{\alpha})$. Then, by Lemma 5.9, $A(P^{<\bar{\alpha}}, s) \in \mathbb{P}$. One also has that $|A(P^{<\bar{\alpha}}, s)| < |P_A(s, \bar{\alpha})|$, and hence, by the induction hypothesis, $(*)$ $\mathfrak{M}_{\mathbb{P}} \models \neg A(P^{<\bar{\alpha}}, s)$. If $\mathfrak{M}_{\mathbb{P}} \models P_A(s, \bar{\alpha})$ were to hold, this would mean that $\neg P_A(s, \bar{\alpha}) \in \mathbb{P}$ and thus $\neg A(P^{<\bar{\alpha}}, s) \in \mathbb{P}$ by Lemma 5.9, yielding $\mathfrak{M}_{\mathbb{P}} \models A(P^{<\bar{\alpha}}, s)$, again by invoking the induction hypothesis, which contradicts $(*)$. As a result, $\mathfrak{M}_{\mathbb{P}} \models P_A(s, \bar{\alpha})$.

If F is of the form $\exists \xi G(\xi)$ then $G(\bar{\beta}) \in \mathbb{P}$ for all $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$, and hence, by the induction hypothesis, $\mathfrak{M}_{\mathbb{P}} \models \neg G(\bar{\beta})$ holds for all $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$, thence $\mathfrak{M}_{\mathbb{P}} \models \neg F$.

If F is of the form $\forall \xi G(\xi)$ then $G(\bar{\beta}) \in \mathbb{P}$ for some $\bar{\beta} \in \mathsf{Ord}_{\mathbb{P}}$, and hence, by the induction hypothesis, $\mathfrak{M}_{\mathbb{P}} \models \neg G(\bar{\beta})$, thence $\mathfrak{M}_{\mathbb{P}} \models \neg F$.

The remaining cases are similar. $\square$

Corollary 5.12   *(i)* $\mathfrak{M}_{\mathbb{P}} \models \forall \alpha \forall x \, [A(P_A^{<\alpha}, x) \leftrightarrow P_A^{\alpha}(x)]$.

(ii) $\mathfrak{M}_{\mathbb{P}} \models \forall x \, [A(P_A, x) \leftrightarrow P_A(x)]$.

(iii) *If $U \subseteq \mathbb{N}$ is a set such that $\forall x \, [A(U, x) \to U(x)]$, then $\{n \in \mathbb{N} \mid \mathfrak{M}_{\mathbb{P}} \models P_A(\bar{n})\} \subseteq U$.*

Proof: (i): The formula $\forall \alpha \forall x \, [A(P_A^{<\alpha}, x) \to P_A^{\alpha}(x)]$ is a stage axiom, and hence its negation $\neg \forall \alpha \forall x \, [A(P_A^{<\alpha}, x) \to P_A^{\alpha}(x)]$ is in $\mathbb{P}$, thus $\mathfrak{M}_{\mathbb{P}} \models \forall \alpha \forall x \, [A(P_A^{<\alpha}, x) \to P_A^{\alpha}(x)]$.

Conversely, if $\mathfrak{M}_{\mathbb{P}} \models P_A^{\bar{\beta}}(s)$ then $\neg P_A^{\bar{\beta}}(s) \in \mathbb{P}$, thus $\neg A(P^{\bar{\alpha}}, s) \in \mathbb{P}$, yielding $\mathfrak{M}_{\mathbb{P}} \models A(P^{\bar{\alpha}}, s)$ by Lemma 5.11.

(ii): Note that the forward part of the equivalence is a reflection axiom. So the formula $\neg \forall \alpha \forall x \, [A(P_A, x) \to P_A(x)]$ is in $\mathbb{P}$, and therefore $\mathfrak{M}_{\mathbb{P}} \models \forall \alpha \forall x \, [A(P_A, x) \to P_A(x)]$ by Lemma 5.11. If $\mathfrak{M}_{\mathbb{P}} \models P_A(s)$, then $\mathfrak{M}_{\mathbb{P}} \models P_A^{\bar{\beta}}(s)$ for some $\bar{\beta}$. The latter entails $\mathfrak{M}_{\mathbb{P}} \models A(P_A^{<\bar{\beta}}, s)$ by (i), which on account of positivity implies $\mathfrak{M}_{\mathbb{P}} \models A(P_A, s)$.

(iii): Assume $\forall x \, [A(U, x) \to U(x)]$. By induction over $\mathsf{Ord}_{\mathbb{P}}$ we prove

$$\{n \in \mathbb{N} \mid \mathfrak{M}_{\mathbb{P}} \models P_A^{\bar{\alpha}}(\bar{n})\} \subseteq U,$$

from which (iii) follows. By induction hypothesis,

$$\{n \in \mathbb{N} \mid \mathfrak{M}_{\mathbb{P}} \models P_A^{\bar{\beta}}(\bar{n})\} \subseteq U$$

holds for all $\bar{\beta}$ such that $\mathfrak{M}_\mathbb{P} \models \bar{\beta} < \bar{\alpha}$. Thus,

$$\{n \in \mathbb{N} \mid \mathfrak{M}_\mathbb{P} \models A(P_A^{<\bar{\alpha}}, \bar{n})\} \subseteq \{n \in \mathbb{N} \mid A(U, n)\}$$

by positivity, and hence, by (i), $\{n \in \mathbb{N} \mid \mathfrak{M}_\mathbb{P} \models P_A^{\bar{\alpha}}(\bar{n})\} \subseteq U$. $\qquad\square$

Corollary 5.13 (**ACA**$_0$) *The class $\{n \in \mathbb{N} \mid H(n, \bar{Q})\}$ with $H(x, U)$ being a Π_1^1-formula and all free variables exhibited, is first-order definable over $\mathfrak{M}_\mathbb{P}$. Thus it is a set.*

Proof: It is well-known[21] (and provable in **ACA**$_0$) that every class $\{n \in \mathbb{N} \mid H(n, \bar{Q})\}$ is many-one reducible to one of the form

$$\{n \in \mathbb{N} \mid \forall X (\forall u [A(X, u) \to u \in X] \to n \in X)\}. \tag{8}$$

The latter class, however, is the same as $\{n \in \mathbb{N} \mid \mathfrak{M}_\mathbb{P} \models P_A(\bar{n})\}$ by Corollary 5.12. $\qquad\square$

From the foregoing we conclude that Π_1^1-comprehension obtains (on the basis of **ACA**$_0$) if for all sets of naturals Q there exists a wellordering $\mathfrak{X}$ such that the the search tree $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ is ill-founded. Thus we want to exclude the possibility that for some $\bar{Q}$ all of the search trees $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ for wellorderings $\mathfrak{X}$ are well-founded. To this we turn next.

Case II: All $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ are well-founded. The strings of sequents that make up the nodes of $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ can be linearly ordered by the Kleene-Brouwer tree ordering, which is a wellordering on account of the well-foundedness of the tree. Let $\mathsf{KB}_\mathfrak{X}^{\bar{Q}}$ be the Kleene-Brouwer ordering of $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ (see e.g. [64, V.1.2] for a definition). Owing to the careful design of $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$, the map $\mathfrak{X} \mapsto (\mathbb{T}_\mathfrak{X}^{\bar{Q}}, \mathsf{KB}_\mathfrak{X}^{\bar{Q}})$ gives rise to an endofunctor $F_{\bar{Q}}$ of the category of wellorderings. To see this, let $\mathfrak{Y}$ be another wellordering and $f : \mathfrak{X} \to \mathfrak{Y}$ be a an order preserving function. Each element of $\mathbb{T}_\mathfrak{X}^{\bar{Q}}$ can be viewed as a notation $S(\alpha_{u_0}, \ldots, \alpha_{u_{n-1}})$ where $\alpha_{u_0}, \ldots, \alpha_{u_{n-1}}$ are the ordinal constants occurring in it with $u_0, \ldots, u_{n-1} \in X$ and $u_0 <_X \ldots <_X u_{n-1}$. Now define $F_{\bar{Q}}(f)$ by

$$F_{\bar{Q}}(f)(S(\alpha_{u_0}, \ldots, \alpha_{u_{n-1}})) := S(\alpha_{f(u_0)}, \ldots, \alpha_{f(u_{n-1})}).$$

Then

$$F_{\bar{Q}}(f) : (\mathbb{T}_\mathfrak{X}^{\bar{Q}}, \mathsf{KB}_\mathfrak{X}^{\bar{Q}}) \to (\mathbb{T}_\mathfrak{Y}^{\bar{Q}}, \mathsf{KB}_\mathfrak{Y}^{\bar{Q}}).$$

We shall denote the denotation system associated with this dilator by $\mathcal{D}_{\bar{Q}}$. More concretely, $\mathcal{D}_{\bar{Q}}$ can be visualized as consisting of the denotations

$$(S(x_0, \ldots, x_{n-1}); u_0, \ldots, u_{n-1}; \mathfrak{X})$$

where $S(\alpha_{u_0}, \ldots, \alpha_{u_{n-1}}) \in \mathbb{T}_\mathfrak{X}^{\bar{Q}}$.

For notational reasons, let's drop the subscript $\bar{Q}$, writing $\mathcal{D} := \mathcal{D}_{\bar{Q}}$. Next, $\mathcal{D}$ is wedded to the Bachmann construction as in Definition 5.1, giving rise to the ordinal representation system $\mathsf{OT}_\mathcal{D}(\vartheta)$. The aim is now to show that the well-foundedness of $\mathsf{OT}_\mathcal{D}(\vartheta)$ leads to a contradiction. Let $\mathfrak{J}$ be the suborder of $\mathsf{OT}_\mathcal{D}(\vartheta) \cap \Omega$ of $\mathsf{OT}_\mathcal{D}(\vartheta)$ obtained by allowing only elements α with $\alpha < \Omega$. By assumption, the search tree $\mathbb{T}_\mathfrak{J}^{\bar{Q}}$ is well-founded. A first step consists in realizing that $\mathbb{T}_\mathfrak{J}^{\bar{Q}}$ is actually contained in $\mathsf{OT}_\mathcal{D}(\vartheta)$. Moreover, this order-morphism sends elements of $\mathbb{T}_\mathfrak{J}^{\bar{Q}}$ to ε-numbers above Ω in $\mathsf{OT}_\mathcal{D}(\vartheta)$.

---

[21]See for instance [30, III.3.2] or [64, V.1.8].

Lemma 5.14 *There is an order-morphism that maps $\mathbb{T}_{\mathfrak{J}}^{\bar{Q}}$, equipped with the Kleene-Brouwer ordering, into $\mathsf{OT}_{\mathcal{D}}(\vartheta)$. Moreover, this order-morphism sends elements of $\mathbb{T}_{\mathfrak{J}}^{\bar{Q}}$ to ε-numbers above Ω in $\mathsf{OT}_{\mathcal{D}}(\vartheta)$.*

Proof: An element of $\mathbb{T}_{\mathfrak{J}}^{\bar{Q}}$ is a string of sequents, $S(\alpha_0, \ldots, \alpha_{n-1})$, whose ordinal constants $\alpha_0 < \ldots < \alpha_{n-1}$ are from $\mathfrak{J}$. Letting $c := S(x_0, \ldots, x_{n-1})$ and $\sigma := D(c; 0, \ldots, n-1; n)$, define the desired map by

$$S(\alpha_0, \ldots, \alpha_{n-1}) \mapsto \mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma}.$$

By definition of $\mathcal{D}$ and the ordering of $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ it is (quite) clear that this mapping preserves the order. $\qquad\square$

An important step consists in viewing $\mathbb{T}_{\mathfrak{J}}^{\bar{Q}}$ as the skeleton of a proof in an infinite proof system that enjoys partial cut elimination, for this will mean that one obtains a cut-free proof of the empty sequent, which is impossible. The infinitary system we have in mind is a variant of $\mathbf{PA}_{\mathfrak{J}}$. To be precise, Definition 5.5 only lists the axioms and subsection 5.2 only defines deductions chains in $\mathbf{PA}_{\mathfrak{J}}$. To render it a proper proof system, we have to add inference rules.

Definition 5.15 $\mathbf{PA}_{\mathfrak{J}}^{*}$ has the same axioms as $\mathbf{PA}_{\mathfrak{J}}$ except for the stage and reflection axioms, which will be turned into rules.

The inference rules of $\mathbf{PA}^*_{\mathfrak{I}}$ are:[22]

$$(\wedge) \qquad \frac{\Gamma, E_0 \quad \Gamma, E_1}{\Gamma, E_0 \wedge E_1}$$

$$(\vee) \qquad \frac{\Gamma, E_i}{\Gamma, E_0 \vee E_1} \quad \text{if } i = 0 \text{ or } i = 1$$

$$(\forall_{\mathbb{N}}) \qquad \frac{\Gamma, F(\bar{n}) \quad \text{for all } n \in \mathbb{N}}{\Gamma, (\forall u \in \mathbb{N})\, F(u)}$$

$$(\exists_{\mathbb{N}}) \qquad \frac{\Gamma, F(s)}{\Gamma, (\exists u \in \mathbb{N})\, F(u)} \quad s \text{ number term}$$

$$(b\forall) \qquad \frac{\Gamma, F(\bar{\beta}) \quad \text{for all } \bar{\beta} < \bar{\alpha}}{\Gamma, (\forall \xi < \bar{\alpha}) F(\xi)}$$

$$(b\exists) \qquad \frac{\Gamma, F(\bar{\beta})}{\Gamma, (\exists \xi < \bar{\alpha}) F(\xi)} \quad \text{if } \bar{\beta} < \bar{\alpha}$$

$$(\forall) \qquad \frac{\Gamma, F(\bar{\alpha}) \quad \text{for all } \bar{\alpha}}{\Gamma, \forall \beta\, F(\beta)}$$

$$(\exists) \qquad \frac{\Gamma, F(\bar{\alpha})}{\Gamma, \exists \beta\, F(\beta)}$$

$$(P_A) \qquad \frac{\Gamma, A(P_A^{<\bar{\alpha}}, s)}{\Gamma, P_A(\bar{\alpha}, s)}$$

$$(\neg P_A) \qquad \frac{\Gamma, \neg A(P_A^{<\bar{\alpha}}, s)}{\Gamma, \neg P_A(\bar{\alpha}, s)}$$

$$(\text{Ref}) \qquad \frac{\Gamma, A(P_A, s)}{\Gamma, \exists \xi\, P_A(\xi, s)}$$

$$(\text{Cut}) \qquad \frac{\Gamma, A \quad \Gamma, \neg A}{\Gamma}$$

where in the above rules s is always a number term.

The first observation worthy of note is that any sequent provable in $\mathbb{T}_{\mathfrak{I}}^{\bar{Q}}$ is provable in $\mathbf{PA}^*_{\mathfrak{I}}$ in a very uniform way. $\mathbb{T}_{\mathfrak{I}}^{\bar{Q}}$ consists of deduction chains $\Gamma_0; \ldots; \Gamma_{r-1}$, and we say that Γ_{r-1} is *provable in* $\mathbb{T}_{\mathfrak{I}}^{\bar{Q}}$. More precisely, we write

$$\mathbb{T}_{\mathfrak{I}}^{\bar{Q}} \;\Big|\!\!\frac{S(\alpha_0, \ldots, \alpha_{n-1})}{}\; \Gamma$$

if there exists a deduction chain $\Gamma_0; \ldots; \Gamma_{r-1}$ such that $S(\alpha_0, \ldots, \alpha_{n-1})$ is the notation for this chain and Γ is Γ_{r-1}.

It has always been a challenge of impredicative ordinal-theoretic proof theory to control infinitary derivations. They have to be of a very uniform kind so as to be able to prove

---

[22]There is a slightly annoying aspect to this proof system. $\mathbf{PA}^*_{\mathfrak{I}}$ is supposed to be a proper deduction calculus, which entails that we have to pay attention to structural rules, too. These boring rules can be moved under the carpet by conceiving of sequents as finite sets of formulas. Thus, in what follows, sequents in the sense of $\mathbf{PA}^*_{\mathfrak{I}}$ will be understood as finite sets of formulas. Notations like Γ, A and Γ, Λ will consequently have to be deciphered as $\Gamma \cup \{A\}$ and $\Gamma \cup \Lambda$, respectively.

(partial) cut elimination for such derivations. The problem arises for first time when analyzing systems such as $\mathbf{ID}_1$, $\mathbf{KP}$ and bar induction whose proof-theoretic ordinal is the Bachmann-Howard ordinal. There are several technical approaches from the 1980s, e.g. [32, 39, 33, 9, 40], that could be enlisted to analyze $\mathbf{PA}^*_{\mathfrak{I}}$, too. These days, though, Buchholz' technique of operator controlled from [8] is most widely used and we will also refer to it when defining the uniform derivations of $\mathbf{PA}^*_{\mathfrak{I}}$. However, we will not go into much detail as these techniques are standard by now.[23] For our purposes, an *operator* $\mathcal{H}$ is just a monotone, inclusive and idempotent function from subsets of $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ to subsets of $\mathsf{OT}_{\mathcal{D}}(\vartheta)$ such that for all X, $\mathcal{H}(X)$ contains 0 and Ω and the following holds:

- If α has Cantor normal form $\omega^{\alpha_1} + \ldots + \omega^{\alpha_n}$ then $\alpha \in \mathcal{H}(X) \Leftrightarrow \alpha_1, \ldots, \alpha_n \in \mathcal{H}(X)$;

- If $\alpha_0, \ldots, \alpha_{n-1} \in X$, $\alpha_0 < \cdots < \alpha_{n-1} < \Omega$, and $\sigma = (c; 0, \ldots, n-1; n) \in \mathcal{D}$, then $\mathfrak{E}^\sigma_{\alpha_0,\ldots,\alpha_{n-1}} \in \mathcal{H}(X)$.

We shall write

$$\mathcal{H} \,\big|\!\frac{\delta}{\rho}\, \Gamma$$

to convey that the sequent Γ is deducible in $\mathbf{PA}^*_{\mathfrak{I}}$ controlled by the operator $\mathcal{H}$ with length δ and cut rank ρ.

For an operator $\mathcal{H}$ and finite subset U of $\mathsf{OT}_{\mathcal{D}}(\vartheta)$, $\mathcal{H}[U]$ stands for the operator with $\mathcal{H}[U](V) := \mathcal{H}(U \cup V)$. One then obtains the following interpretation result with the length of the derivation determined by the order-morphism of Lemma 5.14.

Theorem 5.16 *For every operator $\mathcal{H}$,*

$$\mathbb{T}^{\bar{Q}}_{\mathfrak{I}} \,\big|\!\frac{S(\alpha_0,\ldots,\alpha_{n-1})}{}\, \Gamma \quad \Rightarrow \quad \mathcal{H}[\{\alpha_0, \ldots, \alpha_{n-1}\}] \,\big|\!\frac{\mathfrak{E}^\sigma_{\alpha_0,\ldots,\alpha_{n-1}}}{\Omega+k}\, \Gamma$$

for some $k < \omega$, where $c := S(x_0, \ldots, x_{n-1})$ and $\sigma := D(c; 0, \ldots, n-1; n)$.

With $\mathcal{H}$ controlled derivations one obtains partial cut elimination in that all provable sequents consisting entirely of Σ-formulas, i.e. formulas without unbounded universal ordinal quantifiers, have a cut-free proof. The bottom node of the tree $\mathbb{T}^{\bar{Q}}_{\mathfrak{I}}$ is its maximum element in the Kleene-Brouwer ordering. Denoting it by S_0, we have $\mathbb{T}^{\bar{Q}}_{\mathfrak{I}} \,\big|\!\frac{S_0}{}\, \emptyset$ with $\emptyset$ standing for the empty sequent (without any formulas). Owing to the interpretation theorem 5.16, we then have

$$\mathcal{H} \,\big|\!\frac{\mathfrak{E}^{\sigma_0}}{\Omega+3}\, \emptyset \tag{9}$$

where $c := S_0$ and $\sigma_0 := D(c)$. Now, it is possible to turn the derivation of (9) into a cut-free one. Technically, however, the cut elimination procedure requires a hierarchy of ever stronger operators $\mathcal{H}^*_\xi$ with $\xi \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$ that, with growing ordinal index, accommodate more and more ordinals engendered by the collapsing function $\vartheta_{\mathcal{D}}$, which is to say that the sets $\mathcal{H}^*_\xi(X)$ enjoy ever stronger closure properties with regard to $\vartheta_{\mathcal{D}}$ as ξ increases. As this is a standard technique in ordinal analysis (see [8, 41, 51]) at the lowest level of impredicativity, we will not go into details. The upshot is that by applying cut elimination to the derivation in (9) we arrive at a cut-free $\mathbf{PA}^*_{\mathfrak{I}}$-derivation of the empty sequent

$$\mathcal{H}^*_\rho \,\big|\!\frac{\tau}{0}\, \emptyset$$

for specific $\rho, \tau \in \mathsf{OT}_{\mathcal{D}}(\vartheta)$ with $\rho < \Omega$. But this is absurd for a cut-free proof as (Cut) is the only inference that could yield the empty sequent as a consequence; hence case II has been ruled out, thus it's always case I, concluding this direction of the main theorem's proof.

---

[23]E.g. in the textbook [41].

5.3 A glimpse of Anton Freund's work

In his thesis and a series of papers [13, 14, 15, 16], Anton Freund presents a much broader perspective on the connection between Π^1_1-Comprehension and various Bachmann-Howard principles. In particular, he developed a nice categorical approach to the support operation as a natural transformation, thereby introducing the notion of a prae-dilator.

Definition 5.17 The functor
$$[\,\cdot\,]^{<\omega} \colon \mathbb{LO} \to \mathbf{Set}$$
is defined by letting

$$
\begin{aligned}
[\mathfrak{X}]^{<\omega} &= \quad \text{the set of finite subsets of } X \\
[f]^{<\omega} \colon [\mathfrak{X}]^{<\omega} &\to [\mathfrak{Y}]^{<\omega} \\
[f]^{<\omega}(\{x_1, \ldots, x_n\}) &= \{f(x_1), \ldots, f(x_n)\}
\end{aligned}
$$

whenever $f \colon \mathfrak{X} \to \mathfrak{Y}$ and $x_1, \ldots, x_n \in X$.

An endofunctor $T \colon \mathbb{LO} \to \mathbb{LO}$ is called a *prae-dilator* if there exists a natural transformation
$$\mathsf{supp} \colon T \Rightarrow [\,\cdot\,]^{<\omega}$$
i.e., the square

$$
\begin{array}{ccc}
T(\mathfrak{X}) & \xrightarrow{\mathsf{supp}_{\mathfrak{X}}} & [\mathfrak{X}]^{<\omega} \\
{\scriptstyle T(f)}\Big\downarrow & & \Big\downarrow{\scriptstyle [f]^{<\omega}} \\
T(\mathfrak{Y}) & \xrightarrow[\mathsf{supp}_{\mathfrak{Y}}]{} & [\mathfrak{Y}]^{<\omega}
\end{array}
$$

commutes for every $f \colon \mathfrak{X} \to \mathfrak{Y}$ in $\mathbb{LO}$, and for $\sigma \in T(\mathfrak{X})$,

$$\sigma \in \mathrm{ran}(T(\iota_\sigma)),$$

where $\iota_\sigma \colon \mathsf{supp}_{\mathfrak{X}}(\sigma) \hookrightarrow \mathfrak{X}$ denotes the obvious inclusion map.

It's not the purpose of this paper to present more details of Freund's work, however, we strongly recommend looking at the original sources [13, 14, 15, 16]. To finish this subsection, we quote the following characterization theorem from [13, 15, 16].

Theorem 5.18 *The following are equivalent over* $\mathbf{RCA}_0$:

(i) *The principle of* Π^1_1-*comprehension.*

(ii) *The abstract Bachmann-Howard principle:* Every dilator has a well-founded Bachmann-Howard fixed point.

(iii) *The computable Bachmann-Howard principle:* For every dilator T the linear order $\vartheta(T)$ is well-founded.

6 There are much stronger constructions than Bachmann's

In the list of strong theories for which an ordinal analysis has been achieved, the theory of non-iterated inductive arithmetic definitions (embodied in $\mathbf{PA}_\Omega$) is just the starting point. Likewise, Bachmann's ordinal representation system is just the first in a long line of ever bigger ones. It is thus natural to ponder what happens when one marries the stronger representation systems with a dilator in the same vein as 5.1. As some $\mathbf{WOP}(f)$ principles turned out to be equivalent to the existence of countable coded ω-models, one is led to think that there is a stronger notion of model that pertains to principles such as $\mathbb{HWOP}(F)$. After ω-models come β-models and we take interest in statements of the form *"every set belongs to a countably coded β-model of T"* for various theories T.

Definition 6.1 Every ω-model $\mathfrak{M}$ of a theory T in the language of second order arithmetic is isomorphic to a structure
$$\mathfrak{A} = \langle \omega; \mathfrak{X}; 0, +, \times, \ldots \rangle$$
where $\mathfrak{X} \subseteq \mathcal{P}(\omega)$.

$\mathfrak{A}$ is a β-*model* if the concept of well ordering is absolute with respect to $\mathfrak{A}$, i.e. for all $X \in \mathfrak{X}$,

$$\mathfrak{A} \models \mathrm{WO}(<_X) \quad \text{iff} \quad <_X \text{ is a well ordering,}$$

where $n <_X m :\Rightarrow 2^n 3^m \in X$.

A natural test case is to consider β-models of Π^1_1-comprehension. There already exists a characterization of the statement that every set is contained in an ω-model of Π^1_1-comprehension in [57] due to Ian Alexander Thomson and the author of this paper, as well as a characterization of the statement that every set is contained in an ω-model of Π^1_1-comprehension plus bar induction due to Ian Alexander Thomson [67]. This requires relativizing a stronger ordinal representation system than Bachmann's. The construction gives rise to a dilator, too.

Definition 6.2 We do not want to elaborate greatly on the details of the definition of the representation system $\mathsf{OT}(\Omega_\omega \cdot \mathfrak{X})$ for a wellordering $\mathfrak{X}$, but like to give the main ideas. The crucial extension here is that there is not just one strong ordinal Ω as in Definition 3.3 but infinitely many $\Omega_1, \Omega_2, \ldots, \Omega_n, \ldots$ with Ω_ω denoting their limit. $\mathfrak{X}$ enters the representation system $\mathsf{OT}(\Omega_\omega \cdot \mathfrak{X})$ as follows: For $u \in \mathfrak{X}$ one has $\Omega_\omega \cdot u \in \mathsf{OT}(\Omega_\omega \cdot \mathfrak{X})$, and if $u <_\mathfrak{X} v$ and $\alpha < \Omega_\omega$, then $\Omega_\omega + \alpha < \Omega_\omega \cdot u$ and $\Omega_\omega \cdot u + \alpha < \Omega_\omega \cdot v$.

Moreover, each Ω_n comes equipped with its own (partial) collapsing function $\vartheta_n^\mathfrak{X}$, mapping the elements of $\mathsf{OT}(\Omega_\omega \cdot \mathfrak{X})$ below Ω_n, more precisely

$$\vartheta_n^\mathfrak{X} : \mathsf{OT}(\Omega_\omega \cdot \mathfrak{X}) \longrightarrow (\Omega_{n-1}, \Omega_n)$$

where (Ω_{n-1}, Ω_n) denotes the interval of terms strictly inbetween Ω_{n-1} and Ω_n with $\Omega_0 := 0$. Furthermore, each Ω_n has its own support function supp_{Ω_n} and $\vartheta_n^\mathfrak{X}(\alpha)$ is only defined if the ordinals in $\mathsf{supp}_{\Omega_n}(\alpha)$ are strictly less than α, for which we write $\mathsf{supp}_{\Omega_n}(\alpha) < \alpha$. $\mathsf{supp}_{\Omega_n}(\alpha)$ is defined as follows:

(i) $\mathsf{supp}_{\Omega_n}(0) = \emptyset$, $\mathsf{supp}_{\Omega_n}(\Omega_\tau) = \emptyset$ for $\tau \in \omega + 1$.

(ii) $\mathsf{supp}_{\Omega_n}(\alpha) = \mathsf{supp}_{\Omega_n}(\alpha_1) \cup \cdots \cup \mathsf{supp}_{\Omega_n}(\alpha_n)$ if $\alpha =_{CNF} \omega^{\alpha_1} + \cdots + \omega^{\alpha_n} > \alpha_1$.

(iii) $\mathsf{supp}_{\Omega_n}(\Omega_\omega \cdot u + \alpha) = \mathsf{supp}_{\Omega_n}(\alpha)$ for $\alpha < \Omega_\omega$ and $u \in \mathfrak{X}$.

(iv) For $\beta = \vartheta_m^{\mathfrak{X}}(\gamma)$, $\mathsf{supp}_{\Omega_n}(\beta) = \emptyset$ if $m < n$ and

$$\mathsf{supp}_{\Omega_n}(\beta) = \{\gamma\} \cup \mathsf{supp}_{\Omega_n}(\gamma)$$

if $m \geq n$.

Finally, the ordering between terms $\vartheta_n^{\mathfrak{X}}(\alpha)$ and $\vartheta_m^{\mathfrak{X}}(\beta)$ is determined as follows:

$$\vartheta_n^{\mathfrak{X}}(\alpha) < \vartheta_m^{\mathfrak{X}}(\beta) \text{ iff } n < m \vee [n = m \wedge \alpha < \beta].$$

Theorem 6.3 (Rathjen, Thomson [57]) *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) *Every set is contained in an ω-model of Π_1^1-comprehension.*

(ii) $\forall \mathfrak{X} \, [\mathrm{WO}(\mathfrak{X}) \to \mathrm{WO}(\mathsf{OT}(\Omega_\omega \cdot \mathfrak{X}))].$

Proof: [57]. □

The construction of Definition 6.2 lends itself to a combination with any dilator $\mathcal{D}$, following the pattern of Definition 5.1, giving rise to an ordinal representation system notated by $\mathsf{OT}_{\mathcal{D}}(\Omega_\omega)$.

Definition 6.4 In $\mathsf{OT}_{\mathcal{D}}(\Omega_\omega)$, in place of the objects $\Omega_\omega \cdot u$ for $u \in \mathfrak{X}$ one has the following closure property of Definition 6.2:

If $\alpha_0, \ldots, \alpha_{n-1} \in \mathsf{OT}_{\mathcal{D}}(\Omega_\omega)$, $\alpha_0 < \cdots < \alpha_{n-1} < \Omega_\omega$, and $\sigma = (c; 0, \ldots, n-1; n) \in \mathcal{D}$, then

$$\mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma} \in \mathsf{OT}_{\mathcal{D}}(\Omega_\omega) \text{ and } \Omega_\omega < \mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma}.$$

Objects $\mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma}$ act as ε-numbers above Ω_ω. The ordering between terms $\mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma}$ and $\mathfrak{E}_{\beta_0, \ldots, \beta_{m-1}}^{\tau}$ is defined as in 6.2. Furthermore, the support functions need to be extended to these terms:

$$\mathsf{supp}_{\Omega_n}(\mathfrak{E}_{\alpha_0, \ldots, \alpha_{n-1}}^{\sigma}) = \mathsf{supp}_{\Omega_n}(\alpha_0) \cup \ldots \cup \mathsf{supp}_{\Omega_n}(\alpha_{n-1}).$$

Terms $\vartheta_n^{\mathcal{D}}(\alpha)$ are generated by the following clause:

If $\mathsf{supp}_{\Omega_n}(\alpha) < \alpha$ then $\vartheta_n^{\mathcal{D}}(\alpha) \in \mathsf{OT}_{\mathcal{D}}(\Omega_\omega)$ and $\vartheta_n^{\mathcal{D}}(\alpha) < \Omega_n$.

Finally, the ordering between terms $\vartheta_n^{\mathcal{D}}(\alpha)$ and $\vartheta_m^{\mathcal{D}}(\beta)$ is determined as in 6.2:

$$\vartheta_n^{\mathcal{D}}(\alpha) < \vartheta_m^{\mathcal{D}}(\beta) \text{ iff } n < m \vee [n = m \wedge \alpha < \beta].$$

By combining the techniques of [57, 67] and section 5, one arrives at a characterization of β-models of Π_1^1-comprehension.

Theorem 6.5 *Over* $\mathbf{RCA}_0$ *the following are equivalent:*

(i) *Every set is contained in a countably coded β-model of Π_1^1-comprehension.*

(ii) *For every dilator $\mathcal{D}$, $\mathsf{OT}_{\mathcal{D}}(\Omega_\omega)$ is a well-ordering.*

Proof: See [17]. □

In all likelihood, Theorem 6.5 is not a singularity. For instance, using a representation system with an inaccessible as in the ordinal analysis of Δ_2^1-comprehension plus bar induction (see [35]), one should arrive at a dilatorial characterization of β-models of Δ_2^1-comprehension. Similar characterizations should be obtainable for β-models of theories axiomatizing recursively Mahlo universes (cf. [45, 47, 2]) or Π_3-reflecting universes (cf. [46, 3]) and much more, the general conjecture, stated somewhat loosely, reads as follows:

Conjecture 6.6 *Ordinal representation systems* OT *used in ordinal analyses of strong systems T give rise to functors*

$$F_{\mathsf{OT}} : \mathsf{DIL} \longrightarrow \mathsf{WO}$$

sending dilators to ordinal representation systems. The assertion that F_{OT} sends dilators to wellorderings is equivalent to the statement that every set is contained in a countably coded β-model of T.

References

[1] B. Afshari and M. Rathjen: *Reverse Mathematics and Well-ordering Principles: A pilot study*, Annals of Pure and Applied Logic 160 (2009) 231-237.

[2] T. Arai : *Proof theory for theories of ordinals I: recursively Mahlo ordinals*, Annals of Pure and Applied Logic, 122 (2003) 1–85.

[3] T. Arai: 2004, *Proof theory for theories of ordinals II: Π_3-Reflection*, Annals of Pure and Applied Logic 129 (2004) 39–92.

[4] H. Bachmann: *Die Normalfunktionen und das Problem der ausgezeichneten Folgen von Ordinalzahlen*, Vierteljahresschrift Naturforsch. Ges. Zürich 95 (1950) 115–147.

[5] J. Barwise: *Admissible Sets and Structures* (Springer, Berlin 1975).

[6] E.W. Beth: *The Foundations of Mathematics*, (North Holland, Amsterdam, 1959)

[7] W. Buchholz: *Induktive Definitionen und Dilatoren*, Archiv f́ur Mathematische Logik und Grundlagenforschung 27 (1988) 51–60.

[8] W. Buchholz: *A simplified version of local predicativity,* in: Aczel, Simmons, Wainer (eds.): *Leeds Proof Theory 90* (Cambridge University Press, Cambridge, 1993) 115–147.

[9] W. Buchholz and K. Schütte: *Proof theory of impredicative subsystems of analysis.* (Bibliopolis, Napoli, 1988).

[10] S. Feferman: *Systems of predicative analysis*, Journal of Symbolic Logic 29 (1964) 1–30.

[11] S. Feferman: *Systems of predicative analysis II. Representations of ordinals*, Journal of Symbolic Logic 33 (1968) 193–220.

[12] S. Feferman, *Hereditarily replete functionals over the ordinals*, 1970, in [36], 1970, 289–301.

[13] A. Freund: *Type-Two Well-Ordering Principles, Admissible Sets, and Π_1^1-Comprehension.* Ph.D. thesis, 2018, University of Leeds.

[14] A. Freund, *A categorical construction of Bachmann-Howard fixed points*, Bull. Lond. Math. Soc. 51 (5) (2019) 801–814.

[15] A. Freund: Π_1^1-*Comprehension as a Well-Ordering Principle*, Advances in Mathematics 355 (2019), article 106767.

[16] A. Freund: *Computable Aspects of the Bachmann-Howard Principle*, Journal of Mathematical Logic (2020) 2050006.

[17] A. Freund, M. Rathjen: *A proof-theoretic characterization of β-models of Π_1^1-comprehension.* In preparation.

[18] H. Friedman, A. Montalbán, A. Weiermann: *Phi function*, draft, 2007.

[19] G. Gentzen: *The Collected Papers of Gerhard Gentzen.* Translated and edited by M.E. Szabo (North-Holland, Amsterdam, 1969).

[20] G. Gentzen: *Untersuchungen über das logische Schließen*, Mathematische Zeitschrift 39 (1935) 176–210, 405–431.

[21] G. Gentzen: *Der erste Widerspruchsfreiheitsbeweis für die klassische Zahlentheorie.* Archiv für Mathematische Logik und Grundlagenforschung 16 (1974) 97–118.

[22] G. Gentzen: *Neue Fassung des Widerspruchsfreiheitsbeweises für die reine Zahlentheorie, Forschungen zur Logik und zur Grundlegung der exacten Wissenschaften*, Neue Folge 4, Leipzig, Hirzel (1938) 19–44.

[23] G. Gentzen: *Beweisbarkeit und Unbeweisbarkeit von Anfangsfällen der transfiniten Induktion in der reinen Zahlentheorie, Mathematische Annalen*, 119 (1943) 140–161.

[24] J.-.Y. Girard: *A survey of Π_2^1-logic. Part I: Dilators*, Annals of Mathematical Logic 21 (1981) 75–219.

[25] J.-Y. Girard: *Introduction to Π_2^1-logic*, Synthese 62 (1985) 191–216.

[26] J.-Y. Girard: *Proof Theory and Logical Complexity*, vol. 1 (Bibliopolis, Napoli, 1987).

[27] J.-Y. Girard, D. Normann: *Set recursion and Π_2^1-logic*, Annals of Pure and Applied Logic 28 (1985) 255–286.

[28] G.H. Hardy: *A theorem concerning the infinite cardinal numbers*, Quarterly Journal of Mathematics 35 (1904) 87–94.

[29] K.J.J. Hintikka: *Form and content in quantification theory*, Acta Philosophica Fennica 8 (1955) 7–55.

[30] P.G. Hinman: *Recursion-theoretic hierarchies* (Springer, Berlin, 1978).

[31] J. L. Hirst: *Reverse mathematics and ordinal exponentiation*, Annals of Pure and Applied Logic 66 (1994) 1–18.

[32] G. Jäger: *Beweistheorie von KPN*, Archiv f. Math. Logik 2 (1980) 53–64.

[33] G. Jäger: *Zur Beweistheorie der Kripke–Platek Mengenlehre über den natürlichen Zahlen*, Archiv f. Math. Logik 22 (1982) 121–139.

[34] G. Jäger: *Fixed points in Peano arithmetic with ordinals.* Annals of Pure and Applied Logic 60 (1993) 119–132.

[35] G. Jäger and W. Pohlers: *Eine beweistheoretische Untersuchung von* Δ_2^1-CA *und verwandter Systeme*, Sitzungsberichte der Bayerischen Akademie der Wissenschaften, Mathematisch–Naturwissenschaftliche Klasse (1982).

[36] Kino, A., Myhill, J., Vesley, R., 1970, *Intuitionism and Proof Theory*, Amsterdam: North-Holland.

[37] A. Marcone, A. Montalbán: *The Veblen functions for computability theorists*, Journal of Symbolic Logic 76 (2011) 575–602.

[38] A. Montalbán: *Open Questions in Reverse Mathematics*, Bulletin of Symbolic Logic 2011.

[39] W. Pohlers: *Cut elimination for impredicative infinitary systems, part I: Ordinal analysis of* ID_1, Archiv f́ur Mathematische Logik und Grundlagenforschung 21 (1981) 69–87.

[40] W. Pohlers: *Proof Theory: An Introduction.* Lecture Notes in Mathematics Vol. 1407 (Springer-Verlag, Berlin, 1989).

[41] W. Pohlers: *Proof Theory: The first step into impredicativity.* Universitext (Springer-Verlag, Berlin, 2009).

[42] M. Rathjen: *The strength of Martin-Löf type theory with a superuniverse. Part II.* Archive for Mathematical Logic 40 (2001) 207-233.

[43] M. Rathjen: *Untersuchungen zu Teilsystemen der Zahlentheorie zweiter Stufe und der Mengenlehre mit einer zwischen* $\Delta_2^1 - CA$ *und* $\Delta_2^1 - CA + BI$ *liegenden Beweisstärke.* Publication of the Institute for Mathematical Logic and Foundational Research of the University of Münster (1988).

[44] M. Rathjen: *Lecture notes: Selected topics in proof theory* (Münster University, 1989).

[45] M. Rathjen: *Proof-Theoretic Analysis of KPM*, Arch. Math. Logic 30 (1991) 377–403.

[46] M. Rathjen: *Proof theory of reflection.* Annals of Pure and Applied Logic 68 (1994) 181–224.

[47] M. Rathjen: *The recursively Mahlo property in second order arithmetic.* Mathematical Logic Quarterly 42, 59–66 (1996).

[48] M. Rathjen: *The Realm of Ordinal Analysis.* In: Set and Proofs, London Math. Soc. Lecture Note Ser. 258, edited by S. B. Cooper and J. K. Truss, Cambridge University Press, Cambridge (1999), 219–279.

[49] M. Rathjen: *Investigations of subsystems of second order arithmetic and set theory in strength between* Π_1^1-**CA** *and* Δ_2^1-**CA** + **BI***: Part I* In: R. Schindler (ed.): *Ways of Proof Theory* (Ontos Verlag, Frankfurt, 2010) 363–439.

[50] M. Rathjen: *Replacement versus collection in constructive Zermelo-Fraenkel set theory.* Annals of Pure and Applied Logic 136 (2005) 156–174.

[51] M. Rathjen: *Proof Theory*, Lecture Notes, Leeds, 2012.

[52] M. Rathjen: ω-*models and well-ordering principles.* In: Neil Tennant (ed.): *Foundational Adventures: Essays in Honor of Harvey M. Friedman.* (College Publications, London, 2014) 179–212.

[53] M. Rathjen, P.F. Valencia Vizcaíno: *Well ordering principles and bar induction.* In: R. Kahle, M. Rathjen (eds.): *Gentzen's centenary: The quest for consistency* (Springer, Berlin, 2015) 533–561.

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

[55] M. Rathjen, A. Weiermann: *Reverse Mathematics and Well-ordering Principles.* In: S. Cooper, A. Sorbi (eds.): *Computability in Context: Computation and Logic in the Real World* (Imperial College Press, 2011) 351–370.

[56] M. Rathjen: *Omega Models and Well-Ordering Principles: β-models and functors,* Slides from a talk at the *Reverse Mathematics Workshop* at the University of Chicago, 17 September 2011.

[57] M. Rathjen, I.A. Thomson: *Well-ordering principles, ω-models and Π_1^1-comprehension.* In: Kahle, Rathjen (eds.): *The Legacy of Kurt Schütte.* (Springer, Heidelberg, 2020) 169–213.

[58] K. Schütte: *Beweistheoretische Erfassung der unendlichen Induktion in der Zahlentheorie,* Mathematische Annalen 122 (1951) 369–389.

[59] K. Schütte: *Ein System des verknüpfenden Schließens,* Archiv für mathematische Logik und Grundlagenforschung 2 (1956) 55–67.

[60] K. Schütte: *Beweistheorie,* (Springer, Berlin, 1960).

[61] K. Schütte: *Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik,* Archiv für Mathematische Logik und Grundlagenforschung 67 (1964) 45–60.

[62] K. Schütte: *Predicative well-orderings,* in: Crossley, Dummett (eds.), Formal systems and recursive functions (North Holland, 1965) 176–184.

[63] K. Schütte: *Proof Theory* (Springer, Berlin, 1977).

[64] S.G. Simpson: *Subsystems of Second Order Arithmetic,* second edition, (Cambridge University Press, 2009).

[65] R. Sommer: *Ordinal arithmetic in $I\Delta_0$.* In: P. Clote and J. Krajicek (eds.): *Arithmetic, proof theory, and computational complexity* (Clarendon Press, Oxford, 1993) 320–363.

[66] G. Takeuti: *Proof Theory,* (North-Holland, Amsterdam, 1975).

[67] I.A. Thomson: *Well-Ordering Principles and Π_1^1-Comprehension + Bar Induction.* PhD thesis, University of Leeds, 2017. Available online at http://etheses.whiterose.ac.uk/22206/

[68] P.F. Valencia Vizcaíno: *Some Uses of Cut Elimination.* PhD Thesis (University of Leeds, 2013).

[69] Veblen, O., 1908, *Continuous increasing functions of finite and transfinite ordinals, Trans. Amer. Math. Soc.,* 9: 280–292.