An order-theoretic characterization of the Howard-Bachmann-hierarchy

Jeroen Van der Meeren, Michael Rathjen, Andreas Weiermann

Abstract

In this article we provide an intrinsic characterization of the famous Howard-Bachmann ordinal in terms of a natural well-partial-ordering by showing that this ordinal can be realized as a maximal order type of a class of generalized trees with respect to a homeomorphic embeddability relation. We use our calculations to draw some conclusions about some corresponding subsystems of second order arithmetic. All these subsystems deal with versions of light-face \( \Pi^1_1 \)-comprehension.

1 Introduction

The famous Howard-Bachmann ordinal \( \eta_0 \) (which in the literature is also denoted by \( \psi_{\Omega+1}, \theta_{\Omega+1}, \theta_{\Omega+1}^0, d_{\Omega+1} \)) belongs to the most well-established arsenal of proof-theoretic ordinals of natural theories for developing significant parts of (impredicative) mathematics. Of course \( \eta_0 \) is much bigger than \( \varepsilon_0 \), the proof-theoretic ordinal of first order Peano arithmetic, and it is also bigger than \( \Gamma_0 \), the proof-theoretic ordinal of predicative analysis. The ordinal \( \eta_0 \) is the proof-theoretic ordinal of the first order theory \( ID_1 \), which extends \( PA \) by schemes for smallest fixed points of non-iterated positive inductive definitions. The ordinal \( \eta_0 \) is also the proof-theoretic ordinal of the theory \( KP\omega \) which formalizes an admissible universe containing \( \omega \), and \( \eta_0 \) is also the proof-theoretic ordinal of...
ACA₀ + (Π¹¹ – CA)⁻ which formalizes lightface Π¹¹-comprehension and of the theory RCA₀ + (BI) which extends RCA₀ by a scheme of bar induction.

All these theories are considered to be impredicative. For example the theory ACA₀ + (Π¹¹ – CA)⁻ allows the formation of new sets of natural numbers by using a comprehension formula which may contain a set quantifier ranging over the set which is just defined using the comprehension under consideration. An ordinal analysis for each of these impredicative theories turned out to be difficult. In technical terms this is usually reflected by proof calculi with rules where the complexity of the antecedent is greater than the complexity of the succedent. Typically this is a stumbling block for cut elimination and only very sophisticated methods like Buchholz’ operator-controlled derivations can circumvent it.

It is important to notice that proof-theoretic ordinals are more than just shere set-theoretic ordinals. They usually come equipped with a first order structure with certain built-in functions which generate the ordinal in question in some natural and perspicuous way. On a more combinatorial level such an ordinal is then usually represented by a certain primitive recursive set of terms with a primitive recursive well-ordering relation. A precise description on what the characteristics of a proof-theoretic ordinal are runs under the ”natural well-ordering problem” which is known to be notoriously difficult. Investigations on proof-theoretic ordinals have been undertaken by many people. To name a few: Aczel, Arai, Bachmann, Bridges, Buchholz, Feferman, Gerben, Girard, Gordeev, Isles, Jäger, Kino, Levitz, Okada, Pfeiffer, Pohlers, Probst, Rathjen, Schütte, Setzer, Strahm, Takeuti, Veblen and Weiermann.

A very important aspect of investigations on proof-theoretic ordinals goes back to Diana Schmidt who first recognized the connection between these ordinals and order-theoretic properties of the functions generating them. She characterized completely the order types which could be generated from the ordinal 0 by applying a monotonic increasing function. A monotonic increasing binary function generates out of the singleton set containing the ordinal 0 no order type larger than ε₀. Functions of bigger arities produce easily ordinals bigger than Γ₀ and in fact ordinals of size comparable to the small Veblen ordinal ϑΩ but by no means an ordinal which is of size comparable to η₀.

Diana Schmidt moreover showed that studying bounds on closure ordinals can best be achieved by determining maximal order types of well-partial-orderings which reflect monotonicity properties of the functions in question. With regard to this research program she classified maximal order types for various classes of labelled trees. The ordinals obtained in this way are all around the small Veblen ordinal ϑΩ and way below η₀ and it was for some time not clear whether η₀ can be characterized in terms of closure ordinals.
In Weiermann [24], it was shown that $\eta_0$ could indeed be characterized as a closure ordinal of so-called essentially monotonic increasing functions. Since then it has been open whether a corresponding order-theoretic characterization in terms of maximal order types is possible. Weiermann’s proof made essential use of the linearity of ordinals and did not generalize to partially ordered structures.

Extending Schmidt’s work in [25], the third author provided in a first step, an order-theoretic characterization for the large Veblen ordinal $\vartheta \Omega^\omega$. Quite recently, the authors of this paper were able to provide in [22] much more convincing methods and results which were suitable for being extended to larger ordinals as well. This recent approach is already far reaching but still misses essential ingredients for a order-theoretic characterization of $\eta_0$. This paper will contribute to this problem.

It should also be noticed that H. Friedman defined in 1985 tree-embeddability relations with a so-called gap-condition which generated ordinals of size $\psi \Omega^\omega$, an ordinal which is much bigger than $\eta_0$. So in principal it seemed plausible that it is possible to single out a natural subordering of Friedman’s ordering which would match with $\eta_0$. This paper provides also a positive answer to this challenge: the defined well-partial-order of maximal order type $\eta_0$ can be seen as a natural subordering of Friedman’s tree-ordering.

Moreover, we believe that our analysis will be a starting point for classifying the maximal order types of the full Friedman’s gap-ordering on trees. We expect that analyzing Friedman’s embeddability relations will be rather hard and difficult and we hope that the result provided in this paper yields a roadmap for a more general result. We confine ourselves to $\eta_0$ since this ordinal is somehow the first serious step into impredicativity. To elaborate a little bit more on Friedman’s gap-ordering: we believe that the maximal order type of Friedman’s trees on $n$ labels can be described using the $n^{th}$ regular cardinal number $\Omega_n$. Hence, the uncountable cardinal numbers play a very important role in classifying the strength of Friedman’s well-partial-order.

In section 2, we give some preliminaries that are needed for later sections. In section 4 we yield a well-partial-order of maximal order type $\eta_0$, that can be seen as a natural subordering of Friedman’s tree-ordering. Section 3 is needed to obtain the results in section 4. In the last two sections, section 5 and 6 results from a pure proof-theoretical point of view are studied. We determine bounds on the proof-theoretical ordinals of theories corresponding to our well-partial-orders, from which we obtain unprovability results concerning the well-partial-orderedness of these partial orders. All these theories deal with versions of light-face $\Pi^1_1$-comprehension.
2 Preliminaries

2.1 Well-partial-orderings

Well-partial-orderings are ordered structures that are used in different fields of mathematics. For example in Gröbnerbases [11] and rewrite theory [23]. Moreover, they are well-known objects among logicians. Well-partial-orderings can be seen as generalizations of well-orderings, an important notion used in ordinal analysis [3,14,15]. They are the underlying concepts of the theorem of Higman [8], the theorem of Kruskal [11], Fraïssé's order type conjecture [12] and Friedman's gap-embeddability relation on trees [20]. In [16], the second and third author did a complete proof-theoretical analysis of the theorem of Kruskal.

Definition 2.1. A well-partial-ordering (hereafter wpo) is a partial ordering \((X, \leq_X)\) such that for every infinite sequence \((x_i)_{i=1}^{+\infty}\) of elements in \(X\), there exists two indices \(i\) and \(j\) such that \(i < j\) and \(x_i \leq_X x_j\). We denote the wpo \((X, \leq_X)\) by \(X\) if the ordering is clear from the context.

So a wpo is a well-founded partial ordering that does not admit an infinite antichain. Equivalently, it is a partial order such that every extension is well-founded. In the literature, one is more familiar with the notion of a well-quasi-order. That is a quasi-order (no antisymmetry) with the same distinctive property. However, by canceling out an obvious equivalence relation, one gets a well-partial-order. Therefore, we can restrict ourselves only to wpo's. The interested reader can read more about well-partial-ordering in [6], the most ground-breaking paper on this subject.

Definition 2.2. The maximal order type of the wpo \((X, \leq_X)\) is defined as

\[
\sup\{\alpha: \preceq_X is an extension of \leq_X, \preceq_X is a well-ordering on X and otype(X, \preceq_X) = \alpha\}.
\]

We denote this ordinal as \(o(X, \leq_X)\) or as \(o(X)\) if the ordering on \(X\) is clear from the context.

The maximal order type of a wpo is an important characteristic of that wpo. E.g. one can use it in determining the exact proof-theoretical strength of the wpo under consideration. In [6], it is proved that this supremum is actually a maximum, meaning that every wpo \(X\) has at least one maximal linear extension.

Theorem 2.3 (de Jongh and Parikh [6]). Assume that \((X, \leq_X)\) is a wpo. Then there exists a well-ordering \(\leq\) on \(X\) which is an extension of \(\leq_X\) and \(otype(X, \leq) = o(X, \leq_X)\).
In this paper, we are interested in studying the maximal order type of specific wpos. The technique of reifications is very useful in obtaining upper bounds and lower bounds can be acquired from finding linearizations. Our technique will use the concept of left-sets \( L(x) \) and quasi-embeddings. Of course, this is also interwoven with the previous mentioned techniques.

**Definition 2.4.** Let \((X, \leq_X)\) be a partial order and \(x \in X\). Define the left set \( L_X(x) \) as the set \( \{ y \in X : x \not\leq_X y \} \) and \( l_X(x) := o(L_X(x)) \). We skip the subscript \( X \) if this is clear from the context.

**Theorem 2.5** (de Jongh and Parikh \[6\]). Assume that \( X \) is a partial ordering. If \( L(x) \) is a wpo for every \( x \in X \), then \( X \) is a wpo. (The converse is trivially true.) In this case, \( o(X) = \sup\{l(x) + 1 : x \in X\} \).

Using this result, it can be easily seen that the maximal order type is equal to the height of the root in the tree of finite bad sequences. If one would have that \( l(x) < \alpha \), for every \( x \in X \), then \( o(X) \leq \alpha \). Therefore, we are really interested in characterizing the left-sets of a wpo. This, in combination with the notion of quasi-embeddings, will be the most important building blocks of the proofs in sections \( 3 \) and \( 4 \).

**Definition 2.6.** Let \( X \) and \( Y \) two posets. A map \( e : X \rightarrow Y \) is called a **quasi-embedding** if for all \( x, x' \in X \) with \( e(x) \leq_Y e(x') \) we have \( x \leq_X x' \).

**Lemma 2.7.** If \( X \) and \( Y \) are posets and \( e : X \rightarrow Y \) is a quasi-embedding and \( Y \) is a wpo, then \( X \) is a wpo and \( o(X) \leq o(Y) \).

### 2.2 Constructions on Well-partial-orderings

If we have wpo, one can construct plenty of other well-partial-orderings from it. The two most important examples are disjoint unions and products.

**Definition 2.8.** Let \( X_0 \) and \( X_1 \) be two partial orders. Define the **disjoint union** \( X_0 + X_1 \) as the set \( \{(x, 0) : x \in X_0\} \cup \{(y, 1) : y \in X_1\} \) with the following ordering:

\[(x, i) \leq (y, j) \iff i = j \text{ and } x \leq_X y.\]

We notate the element \( (x, i) \) as \( x \) if it is clear from the context in which set \( x \) lies in. Define the **cartesian product** \( X_0 \times X_1 \) as the set \( \{(x, y) : x \in X_0, y \in X_1\} \) with the following ordering:

\[(x, y) \leq (x', y') \iff x \leq_{X_0} x' \text{ and } y \leq_{X_1} y'.\]

With \( X^n \) we denote the partial ordering \( X \times \cdots \times X \), where \( X \) occurs \( n \) times.
In [8], Higman studied one of the most well-known constructor on well-partial-orderings. It is in some sense a cornerstone in the theory of wpo’s.

**Definition 2.9.** Let \((X, \leq_X)\) be a partial order. Define \((X^*, \leq_X^*)\) as the partial ordering on the set \(X^*\) of finite sequences over \(X\) ordered by

\[
(x_1, \ldots, x_n) \leq_X^* (y_1, \ldots, y_m) \iff (\exists 1 \leq i_1 < \cdots < i_n \leq m)(\forall j \in \{1, \ldots, n\})(x_j \leq_X y_{i_j}).
\]

We notate this partial ordering also as \((X^*, \leq_X)\) or even as \(X^*\).

In the next theorem, we state the connection of the maximal order type of these constructors on well-partial-orderings \(X\) and the original order type \(o(X)\).

**Theorem 2.10** (de Jongh and Parikh[6], Schmidt[17]). If \(X_0, X_1\) and \(X\) are wpo’s, then \(X_0 + X_1, X_0 \times X_1\) and \(X^*\) are still wpo’s, and

\[
o(X_0 + X_1) = o(X_0) \oplus o(X_1),
\]
\[
o(X_0 \times X_1) = o(X_0) \otimes o(X_1),
\]

where \(\oplus\) and \(\otimes\) is the natural sum and product between ordinals, and

\[
o(X^*) = \begin{cases} 
\omega^{o(X)} - 1 & \text{if } o(X) \text{ is finite,} \\
\omega^{o(X)} + 1 & \text{if } o(X) = \varepsilon + n, \text{ with } \varepsilon \text{ an epsilon number and } n < \omega, \\
\omega^{o(X)} & \text{otherwise.}
\end{cases}
\]

In section 4 we will prove that a specific ordering \(T(B(\cdot))\) is a well-partial-ordering with order type equal to the Howard-Bachman number. This ordering can be seen as a subordering of Friedman’s famous ordering on finite trees and it uses the following constructor on wpo’s.

**Definition 2.11.** Let \(X\) be a partial order. Define \(B(X)\) as the partial order where the underlying set is the set of the finite structured binary trees with leaf-labels in \(X\) and with the usual embeddability relation between trees. The internal nodes do not have labels. This means that if a tree \(B \in B(X)\) is embeddable in a tree \(B' \in B(X)\), then either \(B\) and \(B'\) are trees of one node with label \(x \in X\) and \(x' \in X\) respectively with \(x \leq_X x'\) or \(B'\) is a binary tree with a left immediate subtree \(B'_1\) and a right immediate subtree \(B'_2\) and \(B\) is embeddable in \(B'_1 \) or \(B'_2\) or \(B\) has also a left immediate subtree \(B_1\) and a right immediate subtree \(B_2\) and \(B\) is embeddable in \(B'_i\) for \(i = 1, 2\). This implies that if \(B\) is embeddable in \(B'\), then the internal nodes of \(B\) are mapped on the internal nodes of \(B'\) and the leaf-nodes of \(B\) on the leaf-nodes of \(B'\).
One can prove that this construct a well-partial-ordering starting from a \( wpo \) \( X \). We are interested in its maximal order type.

**Definition 2.12.** Define \( \varphi_0 \beta \) as \( \omega^\beta \) and let \( \varphi_\alpha \) be the enumeration function of the common fixed points of all \( \varphi_\gamma \) with \( \gamma < \alpha \). This constructs the so-called Veblen hierarchy.

**Definition 2.13.** Let \( \alpha \) be an ordinal.

\[
\overline{\alpha} := \begin{cases} 
\alpha - 1 & \text{if } \alpha \text{ is finite,} \\
\alpha + 1 & \text{if } \alpha = \varphi_2 \beta + n \text{ with } n \text{ a natural number,} \\
\alpha & \text{otherwise.}
\end{cases}
\]

For a proof of the following theorem, we refer the reader to [17].

**Theorem 2.14.** If \( X \) is a \( wpo \), then \( B(X) \) is a \( wpo \) and \( o(B(X)) = \varepsilon_{o(X)} \).

### 2.3 Generalized tree-structures

In section 4, we present a \( wpo \) of order type \( \eta_0 \). This will be a tree-structure that can be interpreted as a subordering of Friedman's famous \( wpo \) on trees with gap-condition. In section 3 we give some auxiliary \( wpo \)'s with maximal order types strictly below \( \eta_0 \). These \( wpo \)'s can also be seen as tree-representations of ordinals below \( \eta_0 \). In this subsection, we give the definitions of these \( wpo \)'s. The reader can also find this kind of \( wpo \)'s in [22] and [26]. For the actual proofs of well-partial-orderedness and the maximal order types, the reader has to wait until sections 3 and 4. Before we present the definition, we elaborate on the definition of theta-functions.

**Definition 2.15.** Let \( \Omega \) denote the first uncountable ordinal. Every ordinal \( 0 < \alpha < \varepsilon_{\Omega+1} \) can be written as \( \Omega^{\alpha_1} \beta_1 + \cdots + \Omega^{\alpha_n} \beta_n \) with \( \beta_i < \Omega \) and \( \alpha > \alpha_1 > \cdots > \alpha_n \). Define the set of coefficients recursively as \( K(\alpha) = \{\beta_1, \ldots, \beta_n\} \cup K(\alpha_1) \cup \cdots \cup K(\alpha_n) \). Let \( K(0) = \{0\} \). Define then \( k(\alpha) \) as the ordinal \( \max(K(\alpha)) \).

**Definition 2.16.** For an ordinal \( \alpha \), define \( \Omega_0[\alpha] \) as \( \alpha \) and \( \Omega_{n+1}[\alpha] \) as \( \Omega^{\Omega_0[\alpha]} \).

**Definition 2.17.** Let \( P \) denote the set of the additive closed ordinal numbers \( \{\omega^\alpha : \alpha \in ON\} \). For every ordinal \( \alpha < \varepsilon_{\Omega+1} \), define \( \vartheta(\alpha) \) as \( \min\{\zeta \in P : k(\alpha) < \zeta \text{ and } \forall \beta < \alpha( k(\beta) < \zeta \to \vartheta(\beta) < \zeta)\} \). The Howard-Bachmann ordinal number is defined as \( \eta_0 = \vartheta(\varepsilon_{\Omega+1}) = \sup_n (\vartheta(\Omega_n[1])) \).
For more information about the theta-function and its connection with Buchholz’ \( \Psi \)-function, we refer the reader to [16]. There, they introduced the \( \vartheta \)-function in a different way, but one can prove that they coincide with our definition if the argument is above \( \Omega^2 \). It can be shown be an easy cardinality argument that \( \vartheta \alpha < \Omega \).

**Lemma 2.18.** \( \vartheta \alpha < \vartheta \beta \iff \begin{cases} \alpha < \beta \text{ and } k(\alpha) < \vartheta \beta \\ \beta < \alpha \text{ and } \vartheta \alpha \leq k(\beta). \end{cases} \)

We need the following two additional lemmas. The proofs are rather straightforward.

**Lemma 2.19.** Suppose \( \alpha \) and \( \beta \) are ordinals beneath \( \varepsilon_{\Omega+1} \). Then
\[
\begin{align*}
    k(\alpha \oplus \beta) &\leq k(\alpha) \oplus k(\beta), \\
    k(\alpha \otimes \beta) &\leq \max\{k(\alpha) \oplus k(\beta), k(\alpha) \otimes k(\beta) \otimes \omega\}, \\
    k(\omega^\alpha) &\leq \omega^{k(\alpha)}.
\end{align*}
\]
Furthermore, \( k(\alpha), k(\beta) \leq k(\alpha \oplus \beta) \) and \( k(\alpha) \leq k(\alpha \otimes \beta) \) if \( \beta > 0 \).

**Lemma 2.20.** Suppose \( \alpha_n, \ldots, \alpha_0 \) are countable ordinal numbers with \( \alpha_i < \gamma \) for an epsilon number \( \gamma \). Then \( k(\alpha((\Omega^n \alpha_n + \cdots + \Omega \alpha_1 + \alpha_0)^*)) < \gamma \).

Before we give the definition of the \( wpo \)'s that we use in sections 3 and 4, let us define a specific class of constructors.

**Definition 2.21.** Define \( Map \) as the least set satisfying the following:

1. \( \cdot \in Map \), (\( \cdot \) plays the role of a place holder).
2. If \( X \) is a countable \( wpo \), then \( X \in Map \),
3. If \( W_1, W_2 \in Map \), then \( W_1 + W_2, W_1 \times W_2, W_1^* \) and \( B(W_1) \) are also elements of \( Map \).

Every element \( W \) of \( Map \) can be seen as a mapping from the set of partial orderings to the set of partial orderings: \( W(X) \) is a partial ordering by putting the partial order \( X \) into the \( \cdot \). For example, if \( W = (B(\cdot) \times X)^* \), then \( W(X) \) is the partial ordering \( (B(X) \times X)^* \). Furthermore, if \( X \) is a \( wpo \), then \( W(X) \) is a \( wpo \) and if \( X \) is countable, then so is \( W(X) \). Every element of \( W(X) \) is represented by a term in finitely many elements in \( X \). For example in the case \( W = (B(\cdot) \times X)^* \), the element \( ((B(x_1, x_2), x_3), (B(x_3, x_1), x_2)) \) in \( W(X) \) with \( x_1, x_2, x_3 \in X, x_1, x_2 \in X \) and \( B(a, b) \) the binary tree.
is represented by a term in \(x_1, x_2, x_3\). By deleting all entries of \(X\), we get the naked term \(((B(\cdot, \cdot), x_1), (B(\cdot, \cdot), x_2))\), which we notate as \(w(\cdot, \cdot, \cdot)\), where

\[
w(a, b, c, d) = ((B(a, b), x_1), (B(c, d), x_2)).
\]

Therefore, the element \(((B(x_1, x_2), x_1), (B(x_3, x_1), x_2))\) can be described using this naked term \(w(\cdot, \cdot, \cdot)\) and the elements \(x_1, x_2, x_3 \in X\) as \(w(x_1, x_2, x_3, x_1)\). In general, an element of \(W(X)\) is represented as \(w(x_1, \ldots, x_n)\) using a naked term \(w(\cdot, \ldots, \cdot)\) and elements \(x_1, \ldots, x_n \in X\). We will call this naked term ‘an element of \(W\).

**Definition 2.22.** Take \(W \in \text{Map}\). Define \(\mathcal{T}(W)\) as the least set satisfying the following requirements

1. \(\circ \in \mathcal{T}(W)\),
2. If \(w(\cdot, \cdot, \cdot)\) is an element of \(W\) and \(t_1, \ldots, t_n \in \mathcal{T}(W)\), then \(w(t_1, \ldots, t_n)\) is an element of \(W(\mathcal{T}(W))\) and let \(\circ[w(t_1, \ldots, t_n)] \in \mathcal{T}(W)\). We will say that \(t_1, \ldots, t_n\) have a lower complexity than \(\circ[w(t_1, \ldots, t_n)]\).

Let the underlying ordering \(\leq_{\mathcal{T}(W)}\) be the least binary reflexive and transitive relation on \(\mathcal{T}(W)\) such that

1. \(\circ \leq_{\mathcal{T}(W)} t\) for every \(t\) in \(\mathcal{T}(W)\),
2. if \(s \leq_{\mathcal{T}(W)} t_j\) for a certain \(j\), then \(s \leq_{\mathcal{T}(W)} \circ[w(t_1, \ldots, t_n)]\),
3. if \(w(t_1, \ldots, t_n) \leq_{\mathcal{T}(W)} w'(t'_1, \ldots, t'_{n'})\), then \(\circ[w(t_1, \ldots, t_n)] \leq_{\mathcal{T}(W)} \circ[w'(t'_1, \ldots, t'_{n'})]\).

If it is clear from the context, we also notate \(\leq_{\mathcal{T}(W)}\) as \(\leq\). Sometimes, we notate the elements \(\circ[w(t_1, \ldots, t_n)]\) also as \(\circ w(t_1, \ldots, t_n)\) if \(w(t_1, \ldots, t_n)\) has already enough brackets in its description.

**Notation 2.23.** Suppose \(t = \circ[w(t_1, \ldots, t_n)]\) is an element of \(\mathcal{T}(W)\). We denote the element \(w(t_1, \ldots, t_n)\) of \(W(\mathcal{T}(W))\) also as \(\times t\).

Our general conjecture is that for every \(W \in \text{Map}\), the partial ordering \(\mathcal{T}(W)\) is actually a \(\text{wpo}\) and the maximal order type is equal to \(\vartheta(\circ(W(\Omega)))\) if \(\Omega^3 \leq \circ(W(\Omega)) \leq \varepsilon_{\Omega+1}\). In section 3 we prove for specific \(W \in \text{Map}\) that the ordering...
\( T(W) \) is indeed a \( wpo \) and \( \vartheta(o(W(\Omega))) \) is an upper bound on the maximal order type of \( T(W) \). In section 4 we show that \( T(B(\cdot)) \) is also a \( wpo \) and \( \vartheta(\varepsilon_{\Omega+1}) \) is exactly equal to \( o(T(B(\cdot))) \). This \( wpo \) can be seen as a subordering of Friedman’s famous trees with gap-embeddability relation \[20\]. For cases where \( o(W(\Omega)) > \varepsilon_{\Omega+1} \), some little adaptations of the general formula \( o(T(W)) = \vartheta(o(W(\Omega))) \) are needed because the domain of the theta-function is below \( \varepsilon_{\Omega+1} \). This is however beyond the scope of this article and will be treated in latter work. We believe that generalizations will lead to a full classification of the strength of Friedman’s \( wpo \)’s.

The next lemma is a very important lemma for the rest of the article. We will skip its proof, but one can find some subparts and the general idea of this proof in \[22\].

**Lemma 2.24 (Lifting Lemma).** Assume that \( W \in \text{Map} \) and let \( q \) be a quasi-embedding from the partial ordering \( Y \) to the partial ordering \( Z \). Then for all elements \( y_1, \ldots, y_n, y'_1, \ldots, y'_m \) in \( Y \) and \( w(\cdot, \ldots, \cdot), v(\cdot, \ldots, \cdot) \) in \( W \) the inequality \( w(q(y_1), \ldots, q(y_n)) \leq_{W(Z)} v(q(y'_1), \ldots, q(y'_m)) \) implies \( w(y_1, \ldots, y_n) \leq_{W(Y)} v(y'_1, \ldots, y'_m) \).

Before we go further, we want to show that \( T(B(\cdot)) \) can indeed be seen as a subordering of Friedman’s \( wpo \) with gap-condition. First, we give the definition of his \( wpo \).

**Definition 2.25.** Let \( T_n \) be the set of finite rooted trees with labels in \( \{0, \ldots, n-1\} \). An element of \( T_n \) is of the form \((T, l)\), where \( T \) is a finite rooted tree, which we see as a partial ordering on a set of nodes, and \( l \) is a labeling function, a mapping from \( T \) to the set \( \{0, \ldots, n-1\} \). Define \((T_1, l_1) \preceq_{\text{gap}} (T_2, l_2)\) if there exists an injective order- and infimum-preserving mapping \( f \) from \( T_1 \) to \( T_2 \) such that

1. \( \forall \tau \in T_1, \text{ we have } l_1(\tau) = l_2(f(\tau)) \).

2. \( \forall \tau \in T_1 \text{ and for all immediate successors } \tau' \in T_1 \text{ of } \tau, \text{ we have that if } \tau \in T_2 \text{ and } f(\tau) < \tau < f(\tau'), \text{ then } l_2(\tau) \geq l_2(f(\tau')) = l_1(\tau'). \)

In \[20\], this is the so-called weak gap-embeddability relation.

**Definition 2.26.** Define the partial ordering \( \overline{T}_2 \) as the subset of \((T_2, \preceq_{\text{gap}})\) which consists of all finite rooted trees such that nodes with label 0 has zero or one immediate successor(s) and nodes with label 1 has exactly two immediate successors. Furthermore, every tree in \( \overline{T}_2 \) has a root with label 0.
Note that the trees in $T_n$ are unstructured. The trees in $T(B(\cdot))$ are structured, so to see the resemblance between $T(B(\cdot))$ and $\overline{T}_2$ we have to restrict $\overline{T}_2$ even a little bit more: we say that every tree in $\overline{T}_2$ is structured, meaning that it has a left-hand side and a right-hand side and an embedding between two trees preserves these left-right order.

**Lemma 2.27.** The partial-ordering $T(B(\cdot))$ is order-isomorphic to the partial ordering $\overline{T}_2$.

**Proof.** Define $g : T(B(\cdot)) \rightarrow \overline{T}_2$ as follows. Let $g(\circ)$ be the tree which consists of one node with label 0. Take $t = \circ[B(t_1, \ldots, t_n)]$ with $B(t_1, \ldots, t_n)$ a binary tree with leaf-labels in the set $\{t_1, \ldots, t_n\}$ and assume that $g(t_1), \ldots, g(t_n)$ is already defined. Set $g(t)$ then as the tree consisting of a root with label 0, that root connected with an edge to the root of $B(t_1, \ldots, t_n)$. Give all the internal nodes of $B$ label 1 and plug $g(t_i)$ in the leaves of $B(t_1, \ldots, t_n)$ with label $t_i$ for every $i$. For example, if $t = \circ[B(\circ, \circ[B(\circ, \circ)])]$, with $B(a, b)$ equal to

Then $g(t)$ is

It is easy to see that $g$ is surjective. If we can prove that

$$t \leq_{T(B(\cdot))} t' \iff g(t) \leq_{\overline{T}_2} g(t')$$

we are done. We will prove this by induction on the sum of complexities of $t$ and $t'$. If $t = \circ$ or $t' = \circ$, then this is trivial. Assume both $t$ and $t'$ are different from $\circ$. Let $t = \circ[B(t_1, \ldots, t_n)]$ and $t' = \circ[B'(t'_1, \ldots, t'_m)]$. If $t \leq_{T(B(\cdot))} t'$, then either $t \leq_{T(B(\cdot))} t_i'$ for a certain $i$ or $B(t_1, \ldots, t_n) \leq_{B(T(B(\cdot)))} B'(t'_1, \ldots, t'_m)$. In both cases, the induction hypothesis yields $g(t) \leq_{\overline{T}_2} g(t')$ quite easily. Now assume $g(t) \leq_{\overline{T}_2} g(t')$. We know that the root of $g(t)$, which has label 0, is mapped on a node with
label 0. If it is not mapped onto the root of $g(t')$, then it is mapped onto a node with label 0 in $g(t'_i)$ for a certain $i$. Hence $g(t) \leq_T \tau_2 g(t'_i)$, so $t \leq_T \tau(B(\cdot)) t'_i \leq_T \tau(B(\cdot)) t'$. Now assume that the root of $g(t)$ is mapped onto the root of $g(t')$. Every internal node $a$ of $B$ has to be mapped on an internal node of $B'$, because otherwise the label 0 of the root of the $g(t_i)$ in which the internal node $a$ of $B$ is mapped, gives a contradiction with the gap-condition. Furthermore, every leaf of $B$, in which $g(t_i)$ are plugged in, has label 0 and is mapped on a node in $g(t'_i)$ with label 0. We can conclude that $B(g(t_1), \ldots, g(t_n)) \leq_B \tau(B(\cdot)) B'(g(t'_1), \ldots, g(t'_m))$. The induction hypothesis yields that $g$ is a quasi-embedding from the set $\{t_1, \ldots, t_n, t'_1, \ldots, t'_m\}$ to $\tau_2$. So the Lifting Lemma implies $B(t_1, \ldots, t_n) \leq_B \tau(B(\cdot)) B'(t'_1, \ldots, t'_m)$. Hence, $t \leq_T \tau(B(\cdot)) t'$.

From the previous lemma, one can actually already conclude that $\tau(B(\cdot))$ is a wpo. Therefore, one can think that the well-partial-orderedness proof of $\tau(B(\cdot))$ in Theorem 4.1 is superfluous. However, this well-partial-orderedness proof does not need an extra argument: it follows from the calculation of an upper bound of the maximal order type of $\tau(B(\cdot))$. Therefore, we do not really waste efforts by stating it in Theorem 4.1.

3 Tree-structures below the Howard-Bachmann ordinal

In section 4, we will show that $\tau(B(\cdot))$ is a wpo with maximal order type $\vartheta(\epsilon_{\Omega+1})$. For obtaining these results, we need to approximate this wpo. This is done from ‘below’ and is treated in this section. The next theorems are generalizations of Theorems 9 and 10 in [22]. The proofs follow the same procedures as in that article, but they are more involved.

**Theorem 3.1.** Suppose $Y_{i,j,k}$ and $Z_i$ are countable wpo’s for all indices. If

$$W(X) = \sum_{i=0}^{N} \left( k_{i,1} \sum_{j=0}^{X_{i,j,1} \times X^j} \right) \times \cdots \times \left( \sum_{j=0}^{X_{i,j,n_i} \times X^j} \right) \times X^{m_i} \times Z_i$$

then $\tau(W)$ is a wpo and $o(\tau(W)) \leq o(\vartheta(W(\Omega)))$.

**Proof.** We will prove the theorem by main induction on the ordinal $o(W(\Omega))$. If $W(X)$ is the empty wpo for every $X$, then the theorem follows easily. We can now assume without loss of generality that $Z_i$, $Y_{i,k_i,j}$ are non-empty and $k_{i,j} > 0$ for all $i$ and $l$. If $o(W(\Omega)) < \Omega$, then $\cdot$ does not occur in $W$. Therefore, $n_i = m_i = 0$.
for all \(i\). Hence, \(W(X)\) is equal to a \(wpo\) \(\sum_{i=0}^{N} Z_{i} =: Z\). So \(T(W) \cong Z \cup \{0\}\), where 0 is a new element smaller than every element in \(Z\). This yields that \(T(W)\) is a \(wpo\) and \(o(T(W)) \leq o(Z) + 1 \leq \vartheta(o(Z)) = \vartheta(o(W(\Omega))).\n
Assume from now on that \(o(W(\Omega)) \geq \Omega\). We will prove that \(L(t)\) is a \(wpo\) and \(l(t) = 0 < \vartheta(o(W(\Omega))).\) Assume now \(t = \circ ((t_{1}, \ldots, t_{n_{l}}), (t_{1}, \ldots, t_{m_{l}}), z)\) with \(z \in Z_{l}, t_{p} \in T(W)\) and \(\overline{t_{p}}\) an element in \(\left(\sum_{j=0}^{k_{l,p}} Y_{l,j} \times T(W)\right)^{+}\), meaning

\[
\overline{t_{p}} = (\overline{t_{p}^{1}}, \ldots, \overline{t_{p}^{r_{p}}}),
\]

with \(\overline{t_{p}^{q}} \in \sum_{j=0}^{k_{l,p}} Y_{l,j} \times T(W)\). So

\[
\overline{t_{p}^{q}} = (y_{p,q}, (t_{p,q}^{1}, \ldots, t_{w_{p,q}}^{p,q}))
\]

with \(y_{p,q} \in Y_{l,v_{p,q}}, v_{p,q} \leq k_{l,p}\). Assume \(l(t_{i}), l(t_{k}^{i,j}) < \vartheta(o(W(\Omega)))\) and \(L(t_{i})\) and \(L(t_{k}^{i,j})\) are \(wpo\)’s. We want to prove that \(L(t)\) is a \(wpo\) and \(l(t) < \vartheta(o(W(\Omega))).\)

Suppose \(s\) is an arbitrary element in \(T(W)\), different from \(\circ\). Then

\[
s = \circ ((s_{1}, \ldots, s_{m_{l'}}, (s_{1}, \ldots, s_{m_{l'}}), z'), \overline{s_{p}^{q}} = (s_{1}^{p}, \ldots, s_{r_{p}^{q}}), \overline{s_{q}} = (y_{p,q}, (s_{1}^{p,q}, \ldots, s_{w_{p,q}}^{p,q}))),
\]

with

\[
\begin{align*}
z' &\in Z_{l'}, \\
y_{p,q}' &\in Y_{l',v_{p,q}}, v_{p,q} \leq k_{l,p}.
\end{align*}
\]

We see that \(s \in L(t)\) if \(s_{i} \in L(t), s_{k}^{i,j} \in L(t)\) and one of the following holds:

a. \(l \neq l'\),

b. \(l = l', z' \in L_{Z_{l}}(z)\),

c. \(l = l', z \leq z_{l} z', (t_{1}, \ldots, t_{m_{l}}) \not\leq (s_{1}, \ldots, s_{m_{l'}})\),

d. \(l = l', z \leq z_{l} z', (t_{1}, \ldots, t_{m_{l}}) \leq (s_{1}, \ldots, s_{m_{l'}}), (\overline{t_{1}}, \ldots, \overline{t_{m_{l}}}) \not\leq (\overline{s_{1}}, \ldots, \overline{s_{m_{l'}}})\),
Now if (c.) holds, there must be a minimal index $m(s) \leq m_1$ such that

$$t_1 \leq s_1, \ldots, t_{m(s)-1} \leq s_{m(s)-1}, t_{m(s)} \not\leq s_{m(s)}.$$ 

If (d.) is valid, there must be a minimal index $n(s) \leq n_1$ such that

$$\overline{t_1} \leq \overline{s_1}, \ldots, \overline{t_{n(s)-1}} \leq \overline{s_{n(s)-1}}, \overline{t_{n(s)}} \not\leq \overline{s_{n(s)}}.$$ 

By similar arguments,

$$\overline{t_{n(s)}} = (\overline{t^{n(s)}_1}, \ldots, \overline{t^{n(s)}_{r_{n(s)}}}) \not\leq (\overline{s^{n(s)}_1}, \ldots, \overline{s^{n(s)}_{r_{n(s)}}}) = \overline{s^{n(s)}}$$

holds iff we are in one of the next cases

1. $t^{n(s)}_1 \not\leq s^{n(s)}_1$ for every $i$,

2. there exists an index $i_1$ such that $t^{n(s)}_1 \not\leq s^{n(s)}_i$ for every $i < i_1$, $t^{n(s)}_1 \leq s^{n(s)}_{i_1}$ and $t^{n(s)}_{i_1} \not\leq s^{n(s)}_{i_1}$ for every $i > i_1$,

... 

$r_{n(s)}$. there exist indices $i_1 < \cdots < i_{r_{n(s)}-1}$ such that $t^{n(s)}_1 \not\leq s^{n(s)}_{i_1}$ for every $i < i_1$, $t^{n(s)}_1 \leq s^{n(s)}_{i_1}$ and $t^{n(s)}_{i_1} \not\leq s^{n(s)}_{i_1}$ for every $i > i_1$, $t^{n(s)}_{i_1} \leq s^{n(s)}_{i_1}$ and $t^{n(s)}_{i_1} \not\leq s^{n(s)}_{i_1}$ for every $i > i_{r_{n(s)}-1}$.

Now,

$$\overline{t^{n(s)}_i} = (y_{n(s),i}, (t^{n(s),i}_1, \ldots, t^{n(s),i}_{r_{n(s),i}})) \not\leq (y'_{n(s),j}, (s^{n(s),j}_1, \ldots, s^{n(s),j}_{r_{n(s),j}})) = \overline{s^{n(s)}}$$

holds iff one of the following is valid

1. $v_{n(s),i} \not\leq w_{n(s),j}$,

2. $v_{n(s),i} = w_{n(s),j}$ and $y_{n(s),i} \not\leq y'_{l,v_{n(s),i},n(s)} y_{n(s),j}$,

3. $v_{n(s),i} = w_{n(s),j}$, $y_{n(s),i} \leq y'_{l,v_{n(s),i},n(s)} y'_{n(s),j}$ and there exists a minimal index $p_{i,j}(s) \leq v_{n(s),i}$ such that

$$t^{n(s),i}_1 \leq s^{n(s),j}_1, \ldots, t^{n(s),i}_{p_{i,j}(s)-1} \leq s^{n(s),j}_{p_{i,j}(s)-1}, t^{n(s),i}_{p_{i,j}(s)} \not\leq s^{n(s),j}_{p_{i,j}(s)}.$$
We just completely characterized $L(t)$. Using this characterization, we define the following constructor $W'(X)$ in $Map$: let $W'(X)$ be $W'_1(X) + W'_2(X)$ with $W'_1(X)$ equal to

$$
\sum_{i=0, i \neq l}^{N} \left( \left( \sum_{j=0}^{k_{i,1}} Y_{i,j,1} \times X^j \right)^* \times \cdots \times \left( \sum_{j=0}^{k_{i,n}} Y_{i,j,n} \times X^j \right)^* \times X^{m_i} \times Z_i \right)
+ \left( \left( \sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left( \sum_{j=0}^{k_{l,n}} Y_{l,j,n} \times X^j \right)^* \times X^{m_l} \times L_{T_i}(z) \right)
+ \sum_{m=1}^{m_l} \left( \left( \sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left( \sum_{j=0}^{k_{l,n}} Y_{l,j,n} \times X^j \right)^* \times X^{m_{l-1}} \times L_{T(W)}(t_m) \times Z_l \right)
$$

and $W'_2(X)$ equal to

$$
\sum_{n=1}^{n_l} \sum_{q=1}^{r_n} \left[ \left( \sum_{j=0}^{k_{l,1}} Y_{l,j,1} \times X^j \right)^* \times \cdots \times \left( \sum_{j=0}^{k_{l,n}} Y_{l,j,n} \times X^j \right)^* \times X^{m_l} \times Z_l \right.
\left. \times \left( \sum_{j=0}^{k_{l,n+1}} Y_{l,j,n+1} \times X^j \right)^* \times \cdots \times \left( \sum_{j=0}^{k_{l,n}} Y_{l,j,n} \times X^j \right)^* \times X^{m_l} \times Z_l \right]
$$

with

$$W''_{2}(X) := \sum_{j=0, j \neq v_{n,i}}^{k_{l,n}} (Y_{l,j,n} \times X^j)
+ \left( L_{y_{1,v_{1,i},n}}(y_{n,i}) \times X^{v_{n,i}} \right) + \sum_{p=1}^{v_{n,i}} \left( Y_{l,v_{n,i},n} \times X^{v_{n,i}-1} \times L_{T(W)}(t^{n,i}_p) \right).$$

The three cases separated by the sign $+$ in $W'_1(X)$ corresponds to the cases (a.), (b.) and (c.). The index $m$ in the third line of $W'_1(X)$ matches with $m(s)$. $W'_2(X)$ corresponds to case (d). The index $n$ in $W'_2(X)$ matches with $n(s)$ and the index $q$ corresponds to the cases $1, \ldots, r_n(s)$ in which we are for $t^{n,i}_p \leq s^{n,i}_p$. $W''_{2}(X)$ matches with $t^{n,s}_i \leq s^{n,s}_j$, where $p$ in $W''_{2}(X)$ corresponds to $p_{i,j}(s)$. 

Now, following the here-described characterization of $L(t)$ thoroughly step-by-step, one can see that there exists a mapping $f$ from \{w(s_1, \ldots, s_n) \in W(T(W)) : o[w(s_1, \ldots, s_n)] \in L(t)\} into $W'(T(W))$ such that if we have $f(w(s_1, \ldots, s_n)) = w'(s_1', \ldots, s_n')$, then \{s_1', \ldots, s_n'\} $\subseteq$ \{s_1, \ldots, s_n\} $\subseteq$ $L(t)$ and if the inequality $f(w(s_1, \ldots, s_n)) \leq W'(T(W)) f(\overline{w(s_1, \ldots, s_n)})$ holds, then $w(s_1, \ldots, s_n) \leq W(T(W)) \overline{w(s_1, \ldots, s_n)}$. We do not explicitly write out the full details of this argument because in Theorem 4.1 we will do a similar proof (written out in full details) and that proof is less messy.

We pinpoint a mapping $g$ from $L(t)$ into $T(W')$. This mapping will be a quasi-embedding. We do this by induction on the complexity of the terms in $L(t)$. Let $g(o)$ be $o$. Let $s$ be an element of $L(t)$, defined as in (1). By induction, we can assume that $g(s_i)$ and $g(s_{kj})$ are already defined. We know that $s$ is equal to

$$
o w(s_1, \ldots, s_m, s_1^{1,1}, \ldots, s_{w_{m',r',n'}}^{1,1})$$

for a certain element $w(\cdot, \ldots, \cdot)$ in $W$. Denote

$$f(\times s) = f\left(w(s_1, \ldots, s_m, s_1^{1,1}, \ldots, s_{w_{m',r',n'}}^{1,1})\right)$$

as $w'(s_1', \ldots, s_m') \in W'(L(t))$ with \{s_1', \ldots, s_m'\} $\subseteq$ \{s_1, \ldots, s_m, s_1^{1,1}, \ldots, s_{w_{m',r',n'}}^{1,1}\}. Define $g(s)$ as $o[w'(g(s_1'), \ldots, g(s_m'))]$. We want to prove that the mapping $g$ is a quasi-embedding.

We prove by induction on the sum of the complexities of $s$ and $\overline{s}$ that the inequality $g(s) \leq_{T(W')} g(\overline{s})$ implies $s \leq_{T(W)} \overline{s}$. If $s$ or $\overline{s}$ is equal to $o$, then this is trivial. Now, let $s = o[w(s_1, \ldots, s_n)]$ and $\overline{s} = o[\overline{w(s_1, \ldots, s_m)}]$. Assume that $f(w(s_1, \ldots, s_n))$ is equal to $w'(s_1', \ldots, s_m')$ and $f(\overline{w(s_1, \ldots, s_m)})$ is $\overline{w'(s_1', \ldots, s_m')}$. So we have

$$o[w'(g(s_1'), \ldots, g(s_m'))] = g(s) \leq_{T(W')} g(\overline{s}) = o[\overline{w'(g(s_1'), \ldots, g(s_m'))}].$$

Then either $g(s) \leq_{T(W')} g(\overline{s}_q)$ for a certain $q$ or

$$w'(g(s_1'), \ldots, g(s_m')) \leq_{W'(T(W'))} \overline{w'(g(s_1'), \ldots, g(s_m'))}.$$ 

In the first case, $s \leq \overline{s}_q \leq \overline{s}$. In the latter case, the induction hypothesis yields that $g$ is a quasi-embedding from the set \{s_1', \ldots, s_m', \overline{s}_1', \ldots, \overline{s}_m'\} to $T(W')$. Hence, the Lifting Lemma implies $w'(s_1, \ldots, s_m') \leq_{W'(T(W))} \overline{w'(s_1, \ldots, s_m')}$. So $f(\times s) \leq f(\times \overline{s})$, hence $\times s \leq \times \overline{s}$. From this we can conclude that $s \leq \overline{s}$.
So \( g \) is a quasi-embedding from \( L(t) \) in \( T(W') \). If \( o(W''(\Omega)) < o(W(\Omega)) \), the main induction hypothesis yields \( T(W') \) is a \( \omega \)-\( \varepsilon \) and \( o(T'(W')) \leq \vartheta(o(W'(\Omega))) \). Therefore, using Lemma 2.7, \( L(t) \) is a \( \omega \)-\( \varepsilon \) and \( o(L(t)) \leq \vartheta(o(W(\Omega))) \). If additionally \( k(o(W''(\Omega))) < \vartheta(o(W(\Omega))) \), we can conclude that the inequality \( o(L(t)) < \vartheta(o(W(\Omega))) \) holds, the objective that we want to achieve. Thus if we can prove that \( o(W'(\Omega)) < o(W(\Omega)) \) and \( k(o(W''(\Omega))) < \vartheta(o(W(\Omega))) \), we can end the proof of this theorem.

1) \( o(W''(\Omega)) < o(W(\Omega)) \).

For notational convenience, we write \( Y \) instead of \( o(Y) \) for \( \omega \)-\( \varepsilon \)'s \( Y \). Additionally, we write \( \alpha^* \) instead of \( o(\alpha^*) \) for ordinal numbers \( \alpha \).

\[
o(W''_{2i}(\Omega)) < \bigoplus_{j=0}^{k_{l,n}} (Y_{l,j,n} \otimes \Omega^j) \oplus \left( l_{Y_{l,v_n,i,n}}(y_{v_n,i}) \otimes \Omega^{v_n,i} \right) \oplus \Omega^{v_n,i} \
\leq \bigoplus_{j=0}^{k_{l,n}} (Y_{l,j,n} \otimes \Omega^j).
\]

We know that \( \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n} \otimes \Omega^j \right)^* \) is a multiplicative closed ordinal. Therefore, inequality [2] yields

\[
\left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n} \otimes \Omega^j \right) \otimes o(W''_{2i}(\Omega))^* \cdots \otimes o(W''_{2q}(\Omega))^* < \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n} \otimes \Omega^j \right)^*.
\]

From the assumption that \( k_{l,n} > 0 \) and \( Y_{l,k_{l,n},n} \neq \emptyset \), we also have

\[
Z_l \otimes \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n} \otimes \Omega^j \right)^{q-1} \otimes o(W''_{2i}(\Omega))^* \cdots \otimes o(W''_{2q}(\Omega))^* \leq \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n} \otimes \Omega^j \right)^*.
\]

Because

\[
\left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \otimes \Omega^{m_l}
\]

is additive closed, we obtain

\[
o(W''_{2i}(X)) < \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,1} \otimes \Omega^j \right)^* \otimes \cdots \otimes \left( \bigoplus_{j=0}^{k_{l,n}} Y_{l,j,n_l} \otimes \Omega^j \right)^* \otimes \Omega^{m_l}
\]
and

\[
\bigoplus_{m=1}^{m_l} \left( \bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n} \otimes \Omega^j \right) \otimes \Omega^{m_l-1} \otimes L_{\mathcal{T}(W)}(t_m) \otimes Z_l \bigoplus o(W'_2(X))
\]

\[
< \left( \bigoplus_{j=0}^{k_{l,1}} Y_{l,j,1} \otimes \Omega^j \right) \otimes \cdots \otimes \left( \bigoplus_{j=0}^{k_{l,n_l}} Y_{l,j,n} \otimes \Omega^j \right) \otimes \Omega^{m_l}.
\]

Therefore,

\[
o(W'(\Omega)) \leq o(W(\Omega)) < o(W(\Omega)).
\]

2) \(k(o(W'(\Omega))) < o(W(\Omega)).\)

We know that \(o(W(\Omega)) \geq \Omega\), hence \(o(W(\Omega))) is an epsilon number. So from Lemmas 2.19 and 2.20 the claim follows if \(Y_{i,j,k}, Z_i, l_{\mathcal{T}(W)}(t_i)\) and \(l_{\mathcal{T}(W)}(t_i^{i,j})\) are all smaller than \(o(W(\Omega))\). We know that this is true for \(l_{\mathcal{T}(W)}(t_i)\) and \(l_{\mathcal{T}(W)}(t_i^{i,j})\), by the sub-induction hypothesis. Furthermore, we have \(Y_{i,j,k}, Z_k \leq k(o(W(\Omega))) < o(W(\Omega)))\) using Lemma 2.19.

**Theorem 3.2.** If \(W(X) = X^{**}\), then \(\mathcal{T}(W)\) is a \(wpo\) and \(o(\mathcal{T}(W)) \leq o(W(\Omega))) = \vartheta(\Omega^{\Omega^{\Omega^{\omega}}})\).

**Proof.** We show that \(L(t)\) is a \(wpo\) and \(l(t) < \vartheta(\Omega^{\Omega^{\Omega^{\omega}}})\) for every \(t\) in \(\mathcal{T}(W)\) by induction on the complexity of \(t\). The theorem then follows from Theorem 2.5.

If \(t = \circ\), then left-set \(L(t)\) is the empty \(wpo\) and \(l(t) = 0 < \vartheta(\Omega^{\Omega^{\Omega^{\omega}}})\). Assume
now that $t = \circ((t_1^1, \ldots, t_{n_1}^1), \ldots, (t_k^k, \ldots, t_{n_k}^k))$. From the induction hypothesis, we know that $L(t_j^i)$ are wpo's and $l(t_j^i) < \partial(o(W(\Omega)))$. Assume

$$s = \circ((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^l, \ldots, s_{m_l}^l)).$$

Then $t \leq s$ iff $t \leq s_j^i$ for certain $i$ and $j$ or

$$((t_1^1, \ldots, t_{n_1}^1), \ldots, (t_k^k, \ldots, t_{n_k}^k)) \preceq^s ((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^l, \ldots, s_{m_l}^l)).$$

Hence, $s \in L(t)$ iff $s_j^i \in L(t)$ for every $i$ and $j$ and one of the following holds

1. $(t_1^1, \ldots, t_{n_1}^1) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $i$.

2. there exists an index $l_1$ such that $(t_1^1, \ldots, t_{n_1}^1) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $i < l_1$, $(t_1^1, \ldots, t_{n_1}^1) \preceq^s (s_1^1, \ldots, s_{m_1}^1)$ and $(t_1^2, \ldots, t_{n_2}^2) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $i > l_1$.

3. there exist indices $l_1 < l_2$ such that $(t_1^1, \ldots, t_{n_1}^1) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $i < l_1$, $(t_1^1, \ldots, t_{n_1}^1) \preceq^s (s_1^1, \ldots, s_{m_1}^1)$, $(t_1^2, \ldots, t_{n_2}^2) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $l_1 < i < l_2$, $(t_1^2, \ldots, t_{n_2}^2) \preceq^s (s_1^2, \ldots, s_{m_2}^2)$ and $(t_1^3, \ldots, t_{n_3}^3) \not\preceq^s (s_1^3, \ldots, s_{m_3}^3)$ for every $i > l_2$.

... 

k. there exist indices $l_1 < \cdots < l_{k-1}$ such that $(t_1^1, \ldots, t_{n_1}^1) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $i < l_1$, $(t_1^1, \ldots, t_{n_1}^1) \preceq^s (s_1^1, \ldots, s_{m_1}^1)$, $(t_1^2, \ldots, t_{n_2}^2) \preceq^s (s_1^1, \ldots, s_{m_1}^1)$ for every $l_1 < i < l_2$, $(t_1^2, \ldots, t_{n_2}^2) \preceq^s (s_1^2, \ldots, s_{m_2}^2)$, ..., $(t_1^{k-1}, \ldots, t_{n_{k-1}}^{k-1}) \preceq^s (s_1^{k-1}, \ldots, s_{m_{k-1}}^{k-1})$ and $(t_1^k, \ldots, t_{n_k}^k) \not\preceq^s (s_1^k, \ldots, s_{m_k}^k)$ for every $i > l_{k-1}$.

Now, $(t_1^1, \ldots, t_{n_1}^1) \not\preceq^s (s_1^1, \ldots, s_{m_1}^1)$ is valid iff one of the following holds

1. $s_r^j \in L(t_1^1)$ for every $r$,

2. there exists an index $r_{1}^{i,j}$ such that $s_r^j \in L(t_1^1)$ for every $r < r_{1}^{i,j}$, $t_1^i \leq s_{r_{1}^{i,j}}^j$ and $s_r^j \in L(t_2^i)$ for every $r > r_{1}^{i,j}$,

... 

n. there exist indices $r_{1}^{i,j} < \cdots < r_{1}^{i,j-1}$ such that $s_r^j \in L(t_1^1)$ for every $r < r_{1}^{i,j}$, $t_1^i \leq s_{r_{1}^{i,j}}^j$, $s_r^j \in L(t_2^i)$ for every $r_{2}^{i,j} > r > r_{1}^{i,j}$, $t_1^i \leq s_{r_{1}^{i,j-1}}^j$, $s_r^j \in L(t_{n_1}^i)$ and $s_r^j \in L(t_{n_2}^i)$ for every $r > r_{n_1}^{i,j-1}$.
Define
\[ W'(X) := \bigotimes_{i=1}^{k} \left( \bigoplus_{j=1}^{n_i} L(t_j^1)^* \otimes X \otimes \cdots \otimes X \otimes L(t_j^l)^* \right)^* \otimes X^* \otimes \cdots \]

If \( s = \circ((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^p, \ldots, s_{m_p}^p)) \) is an element in \( L(t) \), we can interpret \(((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^p, \ldots, s_{m_p}^p))\) as an element of \( W'(L(t)) \). Denote this interpretation as \( f(\times s) = w(s_1, \ldots, s_n) \), with \( \{s_1, \ldots, s_n\} \subseteq \{s_1^1, \ldots, s_{m_1}^1\} \subseteq L(t) \).

For example assume that case 2. holds, meaning that there exists an index \( l_1 \) such that \((t_1^1, \ldots, t_{n_1}^1) \not\leq^* (s_1^1, \ldots, s_{m_1}^1)\) for every \( i < l_1 \), \((t_1^1, \ldots, t_{n_1}^1) \not\leq^* (s_1^l, \ldots, s_{m_1}^l)\) and \((t_1^2, \ldots, t_{n_2}^2) \not\leq^* (s_1^1, \ldots, s_{m_1}^1)\) for every \( i > l_1 \). Then assume that for \((t_1^1, \ldots, t_{n_1}^1) \not\leq^* (s_1^1, \ldots, s_{m_1}^1)\) subcase 1. holds. Additionally, assume that for every \( i > l_1 \) subcase 2. holds. If we look to \((r_1^1, \ldots, r_{n_2}^1) \not\leq^* (s_1^1, \ldots, s_{m_1}^1)\). In this case, \( f(\times s) = w(s_1, \ldots, s_n) \) is equal to

\[
\left( \left( (s_1^1, \ldots, s_{m_1}^1), \left((s_1^2, \ldots, s_{m_2}^{2_{l_1-1}}), s_1^{2_{l_1-1}}, (s_1^{2_{l_1}}, \ldots, s_{m_2}^{2_{l_1}})\right), \ldots, \left( (s_1^{l_1}, \ldots, s_{m_1}^{l_1}), \left((s_1^{l_1+1}, \ldots, s_{m_1}^{l_1+1}), s_1^{l_1+1}, (s_1^{l_1+1}, \ldots, s_{m_1}^{l_1+1})\right), \ldots, \right) \right),
\]

where \( w(\cdot, \ldots, \cdot) \) is equal to

\[
\left( \left( (s_1^1, \ldots, s_{m_1}^1), \left((s_1^2, \ldots, s_{m_2}^{2_{l_1-1}}), \ldots, (s_1^{2_{l_1}}, \ldots, s_{m_2}^{2_{l_1}})\right), \ldots, \left( (s_1^{l_1}, \ldots, s_{m_1}^{l_1}), \left((s_1^{l_1+1}, \ldots, s_{m_1}^{l_1+1}), \ldots, (s_1^{l_1+1}, \ldots, s_{m_1}^{l_1+1})\right), \ldots, \right) \right),
\]

Using the just-described characterization of the set \( L(t) \), one can see that the inequality \( f(\times s) \leq_{W'(\tau(W))} f(\times s') \) yields \( \times s \leq_{W'(\tau(W))} \times s' \) for every \( s, s' \in
Let \( g \). A similar, but more detailed, argument can be found in the proof of Theorem 4.1.

Define now a mapping \( g \) from \( L(t) \) into \( T(W') \) in the following recursive way. Let \( g(o) = o \). Assume now that \( s = o((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^j, \ldots, s_{m_j}^j)) \in L(t) \) and that \( g(s_i^j) \) is defined for every \( i \) and \( j \). If \( f(xs) = w(s_1, \ldots, s_n) \), let \( g(s) \) be the element \( o[w(g(s_1), \ldots, g(s_n))] \) in \( T(W') \). If \( s \) is the example as before, then \( g(s) \) is defined as

\[
\circ \left( \left( (s_1^1, \ldots, s_{m_1}^1), (s_1^2, \ldots, s_{m_2}^2), \ldots \right) + \left( (s_1^{l_1}, \ldots, s_{m_1}^{l_1}), (s_1^{l_1+1}, \ldots, s_{m_1}^{l_1+1}), \ldots \right) \right).
\]

Is \( g \) a quasi-embedding? We claim that \( g(s) \leq_{T(W')} g(s') \) implies \( s \leq_{T(W)} s' \) by induction on the sum of the complexities of \( s \) and \( s' \). If \( s' = o \) or \( s = o \), then this is trivial. Let

\[
s = o((s_1^1, \ldots, s_{m_1}^1), \ldots, (s_1^j, \ldots, s_{m_j}^j)),
\]

\[
s' = o((s_1^{l_1}, \ldots, s_{m_1}^{l_1}), \ldots, (s_1^{l_j}, \ldots, s_{m_j}^{l_j})),
\]

and assume that \( g(s) \leq g(s') \). If \( f(xs) = w(s_1, \ldots, s_n) \) and \( f(xs') = w'(s_1^{l_1}, \ldots, s_{m_1}^{l_1}) \). Then either \( g(s) \leq g(s_i') \) for a certain \( i \) or \( xg(s) \leq_{w'(T(W'))} \times g(s') \). In the former case, the induction hypothesis yields \( s \leq s_i' \leq s' \). In the latter case, we have \( w(g(s_1), \ldots, g(s_n)) \leq_{w'(T(W'))} w'(g(s_1'), \ldots, g(s_{m_1}')) \). The induction hypothesis yields that \( g \) is a quasi-embedding from the set \( \{s_1, \ldots, s_n, s_1', \ldots, s_{m_1}'\} \) to \( T(W') \). Hence, the Lifting Lemma implies \( f(xs) = w(s_1, \ldots, s_n) \leq_{w'(T(W'))} w'(s_1^{l_1}, \ldots, s_{m_1}^{l_1}) = f(xs') \). From the properties of \( f \), we obtain \( xs \leq_{w'(T(W))} xs' \), hence \( s \leq_{T(W)} s' \).

Now, Lemma 2.7 and Theorem 3.1 yield that \( L(t) \) is a \( w \circ o \) and

\[
o(L(t)) \leq o(T(W')) \leq \vartheta(o(W'(\Omega))).
\]

So we can end this proof if we can show that \( \vartheta(o(W'(\Omega))) < \vartheta \left( \Omega^{o\omega} \right) \). For notational convenience, we write \( \Omega^{\omega} \) instead of \( o(\Omega) \) and \( \Omega^{\omega} \) instead of \( o(\Omega^\omega) \) for
wpo’s Y. From the induction hypothesis, we know that $l(t^i_j) < \vartheta \left( \Omega^{\Omega^{\Omega^\omega}} \right) < \Omega$, hence $l(t^i_1)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t^i_j)^* < \Omega^n$ for a certain finite $n$. Therefore,

$$\left( \bigoplus_{j=1}^{n_i} l(t^i_j)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t^i_j)^* \right) + 1 < \Omega^\omega$$

and

$$\left( \bigoplus_{j=1}^{n_i} l(t^i_j)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t^i_j)^* \right)^* < \Omega^\Omega^\omega.$$

From this it follows that $o(W'(\Omega)) < \Omega^\Omega^\omega$. Now

$$o(W'(\Omega)) = \bigoplus_{i=1}^{k} \left( \left( \bigoplus_{j=1}^{n_i} l(t^i_j)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t^i_j)^* \right)^* \otimes \Omega^* \otimes \cdots \otimes \Omega^* \otimes \cdots \otimes \left( \bigoplus_{j=1}^{n_i} l(t^i_j)^* \otimes \Omega \otimes \cdots \otimes \Omega \otimes l(t^i_j)^* \right)^* \right) \otimes \cdots$$

$$= \bigoplus_{i=1}^{k} \left( \Omega^{\omega(i-1)} \otimes \left( \bigoplus_{j=1}^{n_i} \Omega^{j-1} \otimes \cdots \otimes \Omega^{j-1} \otimes \cdots \otimes l(t^i_j)^* \right)^* \otimes \cdots \otimes \left( \bigoplus_{j=1}^{n_i} \Omega^{j-1} \otimes \cdots \otimes l(t^i_j)^* \right)^* \right).$$

Using Lemmas 2.19 and 2.20, $k(o(W'(\Omega))) < \vartheta \left( \Omega^{\Omega^{\Omega^\omega}} \right)$ is valid if

$$l(t^i_1)^* \otimes \cdots \otimes l(t^i_j)^* < \vartheta \left( \Omega^{\Omega^{\Omega^\omega}} \right).$$

This is true because $l(t^i_j) < \vartheta \left( \Omega^{\Omega^{\Omega^\omega}} \right)$ and $\vartheta \left( \Omega^{\Omega^{\Omega^\omega}} \right)$ is an epsilon number.

These results can be generalized to the following theorem. We do not give the proof as it is very technical and it will not provide new insights because it follows the same procedures as in Theorems 3.1 and 3.2.

**Theorem 3.3.** Suppose that $W \in \text{Map}$ consists only of ·, +, ×, * and countable wpo’s $X_i$. Then $T(W)$ is a wpo and $o(T(W)) \leq \vartheta(o(W(\Omega)))$.

One can also show that $\vartheta(o(W(\Omega)))$ is a lower bound for $o(T(W))$ if $o(W(\Omega)) \geq \Omega^3$. However, the purpose of this section is to proof theorems that are needed in section 4. Therefore, we skip the proof of the lower bound.
4 An order-theoretic approach of the Howard-Bachmann ordinal

The previous section yields

\[ o(T(\star \cdots \star^n)) = \vartheta(\Omega_{2n-1}[\omega]). \]

Therefore, the tree-structures \( T(\star \cdots \star^n) \) give rise to representations of countable ordinals strictly below the Howard-Bachmann ordinal and the 'limit' of these structures will be equal to this famous ordinal. But what do we mean by the 'limit' of these structures? In some sense, the set of binary trees is the limit of an iteration of the \( \star \)-operator. Hence, one can expect that

\[ o(T(B(\cdot))) = \sup_{n<\omega} \vartheta(\Omega_{2n-1}[\omega]) = \vartheta(\varepsilon_{\Omega+1}). \]

In this section, we will prove that this is indeed the case. This result yields that the Howard-Bachmann ordinal can be represented as a tree-structure using binary trees, or more specifically, as the \( \omega \)-theoretic (\( T_{2}, \leq_{gap} \)).

**Theorem 4.1.** \( T(B(\cdot)) \) is a \( \omega \)-theoretic and \( o(T(B(\cdot))) \leq \vartheta(\varepsilon_{\Omega+1}) \).

**Proof.** We prove that \( L(t) \) is a \( \omega \)-theoretic and \( l(t) < \vartheta(\varepsilon_{\Omega+1}) \) for every \( t \) in \( T(B(\cdot)) \) by induction on the complexity of \( t \). The theorem then follows from Theorem 2.5. If \( t = o \), then \( L(t) \) is the empty \( \omega \)-theoretic and \( l(t) = 0 < \vartheta(\varepsilon_{\Omega+1}) \).

Let \( B(t_1, \ldots, t_n) \) be an element of \( B(T(B(\cdot))) \). If we write \( B(t_1, \ldots, t_n) \), we mean that the leaf-labels are elements of \( \{t_1, \ldots, t_n\} \). If it is clear from the context, we sometimes write \( B \) instead of \( B(t_1, \ldots, t_n) \). If \( B(t_1, \ldots, t_n) \) is a tree of height zero with leaf-label \( t_i \), define \( W_B(X) \) as the partial ordering \( B(B(L_B(B(\cdot))))(t_i)) \).

Remark that \( \cdot \) does not occur in \( W_B(X) \). If \( B(t_1, \ldots, t_n) \) is a tree with immediate subtrees \( B_1 \) and \( B_2 \), define \( W_B(X) = W_{B_1}(X) \ast X \ast W_{B_2}(X) \).

We prove by induction on the height of the tree \( B \) that there exists a mapping \( g_B \) from

\[ L_B := \{ D(d_1, \ldots, d_k) \in B(T(B(\cdot))) : B(t_1, \ldots, t_n) \notin B(T(B(\cdot))) \} \]

to the partial ordering \( W_B(T(B(\cdot))) \) such that \( g_B \) is a quasi-embedding and if \( g_B(D) = w(d_1', \ldots, d_k') \), with \( w \) an element in \( W_B \) and \( d_1', \ldots, d_k' \) \( \in T(B(\cdot)) \), then \( \{d_1', \ldots, d_k'\} \subseteq \{d_1, \ldots, d_k\} \).

1) \( \text{height}(B) = 0. \)

Let \( B(t_1, \ldots, t_n) \) be a tree with one node and leaf-label \( t_i \). Then \( D(d_1, \ldots, d_k) \in \)
Note that there exist functions $g_B$ as $w(d'_1, \ldots, d'_m)$, then $m = 0$, hence one can show easily that the desired properties of $g_B$ are valid.

\[ \text{ii) height}(B) > 0. \]

Let $B(t_1, \ldots, t_n)$ be a binary tree with immediate subtrees $B_1$ and $B_2$. By the induction hypothesis, there exist functions $g_{B_1}$ and $g_{B_2}$ with the wanted properties. Now, Pick an arbitrary $D(d_1, \ldots, d_k) \in B(T(B(\cdot)))$. Then $B(t_1, \ldots, t_n) \not\preceq_{B(T(B(\cdot)))} D(d_1, \ldots, d_k)$ is valid iff one of the following holds

1. $D(d_1, \ldots, d_k)$ is a binary tree of height 0 with label $d_i$,

2. $D(d_1, \ldots, d_k)$ is a tree of height strictly larger than 0 with immediate subtrees $D_1$ and $D_2$. $B(t_1, \ldots, t_n) \not\preceq_{B(T(B(\cdot)))} D_i$ for $i = 1, 2$ and one of the following occurs

   \begin{enumerate}
   \item $B_1 \not\preceq_{B(T(B(\cdot)))} D_1$,
   \item $B_1 \preceq_{B(T(B(\cdot)))} D_1$ and $B_2 \not\preceq_{B(T(B(\cdot)))} D_2$.
   \end{enumerate}

Because $B(t_1, \ldots, t_n) \not\preceq_{B(T(B(\cdot)))} D_i$ for $i = 1, 2$, we can also use the above case-study for the trees $D_1$ and $D_2$. This leads us to the following definition. Choose an arbitrary $D(d_1, \ldots, d_k) \in L_B$. Define $E_0$ and $F_0$ as $D(d_1, \ldots, d_k)$. Assume that we have $E_i$ and $F_i$ for a certain $i$ as elements of $B(T(B(\cdot)))$. If $F_i$ is a tree of height strictly larger than 0, we want to define $E_{i+1}$ and $F_{i+1}$. Suppose that $F_i^1$ and $F_i^2$ are the immediate subtrees of $F_i$. Then define $E_{i+1}$ as

\[
\begin{cases}
F_i^1 & \text{in case that 2.(a) holds if we look to the condition } B \not\preceq_{B(T(B(\cdot)))} F_i \\
F_i^2 & \text{in case that 2.(b) holds if we look to the condition } B \not\preceq_{B(T(B(\cdot)))} F_i.
\end{cases}
\]

Let $x_{i+1}$ be the number $j$ such that $E_{i+1} = F_i^j$ and let $F_{i+1}$ be $F_i^{3-x_{i+1}}$, the other immediate subtree of $F_i$. From this definition, we obtain a finite sequence $E_0, E_1, \ldots, E_p, F_p$ with $E_0 = D$ and $F_p$ a tree of height 0. Therefore, $F_p$ consists of only one node with a label, let us say, $s$. Note that $s$ is also a leaf-label of the tree $D$. Define now $g_B(D)$ as follows using the fact that we have $g_{B_1}$ and $g_{B_2}$:

\[ g_B(D) := ((g_{B_{x_1}}(E_1), \ldots, g_{B_{x_p}}(E_p)), s). \]

Note that $B_{x_i} \not\preceq E_i$, which means that $E_i \in L_{B_{x_i}}$. So $g_{B_{x_i}}(E_i)$ is well-defined, hence $g_B(D) \in W_B(T(B(\cdot)))$. Does $g$ satisfies the desired properties?

First, we already noted that $s$ is a leaf-label of $D(d_1, \ldots, d_k)$. Secondly, if $g_{B_{x_i}}(E_i)$ is equal to $w_i(s_{i_1}^i, \ldots, s_{i_n_i}^i)$, then by the induction hypothesis and the fact that $E_i$ is a subtree of $D$, we obtain $\{s_{i_1}^i, \ldots, s_{i_n_i}^i\} \subseteq \{d_1, \ldots, d_k\}$. Hence, if
An order-theoretic characterization of the Howard-Bachmann-hierarchy

\[ g_B(D) = w(d'_1, \ldots, d'_m) \in W_B(T(B(\cdot))) \], then \( \{d'_1, \ldots, d'_m\} \subseteq \{d_1, \ldots, d_k\} \). Now we want to prove that \( g \) is a quasi-embedding. Let \( \vec{E}_0, \vec{E}_1, \ldots, \vec{E}_q, \vec{F}_q \) and \( y_1, \ldots, y_q \) be the finite sequences forthcoming from definition (3), the definitions of \( x_{i+1} \) and \( F_{i+1} \), but now starting with \( \overrightarrow{D}(d'_1, \ldots, d'_l) \in L_B \). Denote the label of the tree \( \vec{F}_q \) of height zero as \( \vec{s} \). Assume further that

\[ g_B(D) = ((g_{B_{i_1}}(E_1), \ldots, g_{B_{i_p}}(E_p)), s) \]

\[ \leq_{W_B(T(B(\cdot)))} g_B(\overrightarrow{D}) = ((g_{B_{i_1}}(E_1), \ldots, g_{B_{i_q}}(E_q)), \vec{s}). \quad (4) \]

We show that inequality \( D(d_1, \ldots, d_k) \leq_{B(T(B(\cdot)))} \overrightarrow{D}(d'_1, \ldots, d'_l) \) holds by induction on \( q \).

From (3), we obtain \( s \leq_{T(B(\cdot))} \vec{s} \). Furthermore, there exist indices \( 1 \leq i_1 < \ldots < i_p \leq q \) such that \( g_{B_{i_j}}(E_j) \leq g_{B_{i_j}}(\vec{E}_j) \). Because the left hand side of this inequality is in \( W_{B_{i_j}}(T(B(\cdot))) \) and the right hand side in \( W_{B_{i_j}}(T(B(\cdot))) \), we obtain \( x_j = y_{i_j} \) for every \( j \). Furthermore, \( E_j \leq_{B(T(B(\cdot)))} \vec{E}_{i_j} \) for every \( j \), because \( g_{B_{i_j}} \) is a quasi-embedding.

If \( q = 0 \), then also \( p = 0 \). Therefore, \( D \) is a tree of height zero with leaf-label \( s \) and \( \overrightarrow{D} \) is a tree of the same height with leaf-label \( \vec{s} \). Hence \( D \leq_{B(T(B(\cdot)))} \overrightarrow{D} \). Now let \( q > 0 \). By construction,

\[ g_B(F_1) = ((g_{B_{i_2}}(E_2), \ldots, g_{B_{i_p}}(E_p)), s), \]

\[ g_B(\overrightarrow{F_{i_1}}) = ((g_{B_{i_1+1}}(E_{i_{1+1}}), \ldots, g_{B_{i_q}}(E_q)), \vec{s}). \]

Because \( g_B(D) \leq_{W_B(T(B(\cdot)))} g_B(\overrightarrow{D}) \) and \( g_{B_{i_1}}(E_1) \) is mapped onto \( g_{B_{i_1}}(\vec{E}_{i_1}) \), we obtain \( g_B(F_1) \leq_{W_B(T(B(\cdot)))} g_B(\overrightarrow{F_{i_1}}) \). Hence, by the sub-induction hypothesis on \( q \), we have the inequality \( F_1 \leq_{B(T(B(\cdot)))} \overrightarrow{F}_{i_1} \). We also know that \( E_1 \leq_{B(T(B(\cdot)))} \overrightarrow{E}_{i_1} \) and \( x_1 = y_{i_1} \). If \( x_1 = 1 \), then \( E_1 \) is the left-immediate subtree of \( F_0 = D \) and \( \overrightarrow{E}_{i_1} \) is the left-immediate subtree of \( \overrightarrow{F}_{i_1-1} \). Furthermore, \( F_1 \) is the right-immediate subtree of \( F_0 = D \) and \( \overrightarrow{F}_{i_1-1} \) is the right-immediate subtree of \( \overrightarrow{F}_{i_1-1} \). We conclude that \( D = F_0 \leq_{B(T(B(\cdot)))} \overrightarrow{F}_{i_1-1} \leq_{B(T(B(\cdot)))} \overrightarrow{F}_0 = \overrightarrow{D} \). The same argument holds for \( x_1 = 2 \). Therefore, \( g_B \) is a quasi-embedding.

Assume \( t = \circ[B(t_1, \ldots, t_n)] \in T(B(\cdot)) \) with \( B(t_1, \ldots, t_n) \) a binary tree in the partial ordering \( B(T(B(\cdot))) \) and assume that \( L(t_i) \) are \( \text{wpo} \)-\( s \) and \( l(t_i) < \partial(\varepsilon_{\Omega+1}) \). We want to prove that \( L(t) \) is a \( \text{wpo} \) and \( l(t) < \partial(\varepsilon_{\Omega+1}) \).

First of all, we define a quasi-embedding \( f \) from \( L(t) \) into \( T(W_B) \). First note that \( d = \circ[D(d_1, \ldots, d_k)] \in L(t) \) iff \( d_i \in L(t) \) and \( D(d_1, \ldots, d_k) \in L_B \). Define \( f(\circ) \) as \( \circ \). Suppose \( d = \circ[D(d_1, \ldots, d_k)] \in L(t) \) and that \( f(d_1), \ldots, f(d_k) \) are already defined. If \( g_B(D) = w(d'_1, \ldots, d'_m) \in W_B(T(B(\cdot))) \) with \( \{d'_1, \ldots, d'_m\} \subseteq \{d_1, \ldots, d_k\} \),
define \( f(d) \) as \( \circ[w(f(d'_1), \ldots, f(d'_m))] \in \mathcal{T}(W_B) \). Now we want to prove that \( f \) is a quasi-embedding.

Assume \( f(d) \leq_{\mathcal{T}(W_B)} f(\overline{d}) \). We prove that this implies \( d \leq_{\mathcal{T}(B^{(i)})} \overline{d} \) by induction on the sum of the complexities of \( d \) and \( \overline{d} \). If \( d \) or \( \overline{d} \) is equal to \( \circ \), then this is trivial. Assume \( g_B(D(d_1, \ldots, d_k)) = w(d'_1, \ldots, d'_m) \), \( g_B(\overline{D(d_1, \ldots, d_k)}) = \overline{w(d'_1, \ldots, d'_m)} \) with \( \{d'_1, \ldots, d'_m\} \subseteq \{d_1, \ldots, d_k\} \) and \( \{\overline{d'_1}, \ldots, \overline{d'_m}\} \subseteq \{\overline{d_1}, \ldots, \overline{d_k}\} \) and

\[
\begin{align*}
f(d) &= f:\circ[D(d_1, \ldots, d_k)] = \circ[w(f(d'_1), \ldots, f(d'_m))] \\
&\leq_{\mathcal{T}(W_B)} f(\overline{d}) = f:\circ[\overline{D(d_1, \ldots, d_k)}] = \circ[\overline{w(f(d'_1), \ldots, f(d'_m))}].
\end{align*}
\]

This implies either \( f(d) \leq_{\mathcal{T}(W_B)} f(\overline{d}_i) \) for some \( i \) or

\[
w(f(d'_1), \ldots, f(d'_m)) \leq_{W_B(\mathcal{T}(W_B))} \overline{w(f(\overline{d'_1}), \ldots, f(\overline{d'_m}))}.
\]

In the former case, the induction hypothesis yields \( d \leq_{\mathcal{T}(B^{(i)})} \overline{d}_i \leq_{\mathcal{T}(B^{(i)})} \overline{d} \). In the latter case, we observe that the induction hypothesis implies that \( f \) is a quasi-embedding from \( \{d_1, \ldots, d_k, \overline{d}_1, \ldots, \overline{d}_i\} \) to \( \mathcal{T}(W_B) \). Therefore, the Lifting Lemma brings the inequality \( g_B(D) \leq_{W_B(\mathcal{T}(W_B))} g_B(\overline{D}) \). Hence we have \( D(d_1, \ldots, d_k) \leq_{\mathcal{B}(\mathcal{T}(B^{(i)}))} \overline{D}(\overline{d}_1, \ldots, \overline{d}_i) \) and \( d \leq_{\mathcal{T}(B^{(i)})} \overline{d} \). We conclude that \( f \) is a quasi-embedding.

By Lemma 2.7 and Theorem 3.3 we obtain that \( L(t) \) is a wpo and

\[
o(L(t)) \leq o(\mathcal{T}(W_B)) \leq \vartheta(W_B(\Omega)).
\]

So if \( \vartheta(W_B(\Omega)) < \vartheta(\varepsilon_{\Omega+1}) \), we can end this proof. We prove simultaneously by induction on the height of the tree \( B \) that

\[
W_B(\Omega) < \Omega_n[1],
\]

\[
k(W_B(\Omega)) < \vartheta(\Omega_n[1]),
\]

for a certain natural number \( n \). Remark that we write \( W_B(\Omega) \) instead of \( o(W_B(\Omega)) \) for notational convenience. From this it follows that \( \vartheta(W_B(\Omega)) < \vartheta(\varepsilon_{\Omega+1}) \).

If the height of the tree \( B \) is zero, we defined \( W_B(X) \) as \( \mathcal{B}(L_T(B^{(i)})(t_i)) \). Because we assumed that \( l(t_i) < \vartheta(\varepsilon_{\Omega+1}) \), we know that \( l(t_i) < \vartheta(\Omega_m[1]) < \Omega \) for a certain natural number \( m \geq 2 \). Hence

\[
k(W_B(\Omega)) = k(o(\mathcal{B}(L(t_i)))) = o(\mathcal{B}(L(t_i))) \leq \varepsilon_{l(t_i)+1}
\]

\[
< \varepsilon_{\vartheta(\Omega_m[1]+1)} = \vartheta(\Omega + \vartheta(\Omega_m[1])) \leq \vartheta(\Omega_{m+1}[1])
\]

and

\[
W_B(\Omega) \leq \varepsilon_{l(t_i)+1} < \Omega < \Omega_{m+1}[1].
\]
Assume that the height of $B$ is strictly larger than zero such that $B_1$ and $B_2$ are immediate subtrees of $B$. Because of the induction hypothesis, we know that there exists a natural number $m$ such that

$$W_{B_1}(\Omega) < \Omega_m[1],$$

$$k(W_{B_1}(\Omega)) < \vartheta(\Omega_m[1]),$$

$$W_{B_2}(\Omega) < \Omega_m[1],$$

$$k(W_{B_2}(\Omega)) < \vartheta(\Omega_m[1]).$$

We defined $W_B(X)$ as $(W_{B_1}(X) + W_{B_2}(X))^* \times X$, hence

$$W_B(\Omega) < \omega^{\Omega_m[1] \oplus \Omega_m[1] \oplus 1} \otimes \Omega < \Omega_{n_1}[1]$$

for a certain $n_1$ large enough. Furthermore, by Lemma 2.19

$$k(W_B(\Omega)) = k\left(\omega^{W_{B_1}(\Omega) \oplus W_{B_2}(\Omega) \oplus 1} \otimes \Omega\right) < \vartheta(\Omega_{n_2}[1])$$

for some $n_2$ large enough. Taking $n = \max\{n_1, n_2\}$, we can end this proof.

Before we prove that the Howard-Bachmann ordinal is a lower bound of the maximal order type of $T(B(\cdot))$, we give a lemma.

Lemma 4.2. Suppose $\alpha < \vartheta(\varepsilon_{\Omega+1})$ and $\alpha \in P$, the set of additively closed ordinals. Then $\alpha$ can be uniquely written as

$$\alpha = \vartheta(\beta),$$

with $\beta = 0$ or

$$\beta = \omega^{\beta_1 \gamma_1 + \gamma_1 + \cdots + \beta_n \gamma_n < \varepsilon_{\Omega+1}},$$

such that $\beta > \beta_1 > \cdots > \beta_n$ and $0 < \gamma_i < \Omega$ and $k(\beta) < \alpha$.

Proof. This follows from standard properties of the $\vartheta$-function. A proof of this fact can be found in an unpublished article of Buchholz.

If you look closer at the proof of the next theorem, one can see how every ordinal number below $\vartheta(\varepsilon_{\Omega+1})$ can be represented as an element of $T(B(\cdot))$. Note that this proof can be carried out in $ACA_0$ if we have a predefined primitive recursive ordinal notation system for $\vartheta(\varepsilon_{\Omega+1})$.

Theorem 4.3. $o(T(B(\cdot))) \geq \vartheta(\varepsilon_{\Omega+1})$. 

An order-theoretic characterization of the Howard-Bachmann-hierarchy
Proof. Define

\[ g : \vartheta(\varepsilon_{\Omega+1}) \to \mathcal{T}(B(\cdot)) \]

in the following recursive way. Let \( g(0) \) be \( \circ \). Pick an arbitrary \( \alpha < \vartheta(\varepsilon_{\Omega+1}) \) and assume that \( g(\beta) \) is already defined for every \( \beta < \alpha \). If \( \alpha = \text{CNF} \omega_{\alpha_1} + \cdots + \omega_{\alpha_n} \) with \( n \geq 2 \), define \( g(\alpha) \) as \( \circ[B] \) with \( B \) the following binary tree:

If \( \alpha < \vartheta(\varepsilon_{\Omega+1}) \) and \( \alpha \in P \), we can write \( \alpha \) as \( \vartheta(\beta) \) as in Lemma 4.2. Because every element in \( K(\beta) \) is strictly smaller than \( \alpha \), we can assume that \( g(\gamma) \) is defined for every \( \gamma \in K(\beta) \). Define \( g(\alpha) \) as \( \circ[f(\beta)] \), where we define the binary tree \( f(\beta) \) in the following recursive way.

Let \( f(0) \) be the binary tree with one node and leaf-label \( \circ \):

Now, let \( \beta = \Omega^{\beta_1} \gamma_1 + \cdots + \Omega^{\beta_n} \gamma_n > \beta_1 > \cdots > \beta_n \geq 0 \) and assume that \( f(\beta_i) \) is already defined for every \( i \). Then define \( f(\beta) \) as
An order-theoretic characterization of the Howard-Bachmann-hierarchy

Note that all labels in the tree $f(\beta)$ are elements of $g(K(\beta) \cup \{0\})$. Additionally, every element in $g(K(\beta) \cup \{0\})$ is a label in the tree $f(\beta)$.

Is $g$ a quasi-embedding? We will show by induction on $\alpha'$ that $g(\alpha) \leq g(\alpha')$ implies $\alpha \leq \alpha'$. If $\alpha$ or $\alpha'$ is equal to zero, then this is trivial, hence we may assume that both $\alpha$ and $\alpha'$ are different from zero. There are now four cases left:

a) $g(\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}) \leq g(\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m})$.

Then either $g(\alpha) \leq g(\alpha'_1)$ or $(g(\alpha_1), \ldots, g(\alpha_m)) \leq^* (g(\alpha'_1), \ldots, g(\alpha'_m))$. In both cases the induction hypothesis yields $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \leq \omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m}$.

b) $g(\vartheta(\beta)) \leq g(\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m})$.

If $\beta = 0$, then it is trivial. Assume $\beta \neq 0$. Then $g(\vartheta(\beta)) \leq g(\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m})$ is only possible if $g(\vartheta(\beta)) \leq g(\alpha'_i)$ for a certain $i$. The induction hypothesis yields $\vartheta(\beta) \leq \omega^{\alpha_1} + \cdots + \omega^{\alpha_n}$.

c) $g(\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m}) \leq g(\vartheta(\beta))$.

It is impossible that $\beta = 0$, because $m \geq 2$. If $0 < \beta < \Omega$, then $g(\vartheta(\beta))$ is equal to

\[
\circ \quad g(\beta) \quad \circ
\]

So $g(\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}) \leq g(\vartheta(\beta))$ can only occur if $g(\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}) \leq g(\beta)$ because $m \geq 2$. By the induction hypothesis, we obtain that $\omega^{\alpha_1} + \cdots + \omega^{\alpha_n} \leq \beta < \vartheta(\beta)$. In the case that $\beta \geq \Omega$, we have that $g(\omega^{\alpha_1} + \cdots + \omega^{\alpha_n}) \leq g(\gamma)$ for a certain $\gamma \in K(\beta) \cup \{0\}$ or for every $\alpha'_i$, there exists a $\gamma_i \in K(\beta) \cup \{0\}$ such that $g(\alpha'_i) \leq g(\gamma_i)$. In the former case, we obtain that $\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m} \leq \gamma \leq k(\beta) < \vartheta(\beta)$. In the latter case, we obtain that $\alpha'_i \leq \gamma_i \leq k(\beta) < \vartheta(\beta)$. Because $\vartheta(\beta)$ is an epsilon number ($\beta \geq \Omega$), we get that $\omega^{\alpha'_1} + \cdots + \omega^{\alpha'_m} < \vartheta(\beta)$.

d) $g(\vartheta(\beta)) \leq g(\vartheta(\beta'))$.

If $\beta' = 0$, then $\beta$ must also be zero, hence $\vartheta(\beta) \leq \vartheta(\beta')$. Now assume that $\beta' > 0$. $g(\vartheta(\beta)) \leq g(\vartheta(\beta'))$ is possible if $g(\vartheta(\beta)) \leq g(\gamma')$ for a certain $\gamma' \in K(\beta') \cup \{0\}$ or if $f(\beta) \leq f(\beta')$. In the former case, we obtain by the induction hypothesis that $\vartheta(\beta) \leq \gamma' \leq k(\beta') < \vartheta(\beta')$. If the latter case occurs, then for every $\gamma \in K(\beta)$, there exists a $\gamma' \in K(\beta') \cup \{0\}$ such that $g(\gamma) \leq g(\gamma')$. Hence, $k(\beta) \leq k(\beta') < \vartheta(\beta')$. For ending the proof of $\vartheta(\beta) \leq \vartheta(\beta')$, we need to show that
\[ \beta \leq \beta'. \]

We proof by induction on \( \delta' \) that \( f(\delta) \leq f(\delta') \) implies \( \delta \leq \delta' \) for every \( \delta \) with \( K(\delta) \subseteq K(\beta) \) and every \( \delta' \) with \( K(\delta') \subseteq K(\beta') \). If this is true can conclude that \( \beta \leq \beta' \). If \( \delta' = 0 \) or \( \delta = 0 \), then this is trivial. Hence we may assume that both \( \delta \) and \( \delta' \) are different from zero. Assume \( \delta' = NF \Omega^\delta \gamma_1 + \cdots + \Omega^\delta \gamma_m > 0 \). Then \( f(\delta') \) is equal to

\[
\begin{array}{c}
\text{b} \\
\text{g}(\gamma_1) \\
\vdots \\
\text{g}(\gamma_2) \\
\text{g}(\gamma_3) \\
\text{f}(\delta_1) \\
\text{f}(\delta_2) \\
\text{f}(\delta_3) \\
\text{f}(\delta_m)
\end{array}
\]

There are four different subcases:

i) The root of \( f(\delta) \) is mapped into \( f(\delta_1') \)
Then \( f(\delta) \leq f(\delta_1') \), hence, by the induction hypothesis \( \delta \leq \delta_1' \leq \delta' \).

ii) The root of \( f(\delta) \) is mapped on \( b \)
This is only possible if \( \delta = \Omega^\delta \gamma_1 \). In this case we have that \( f(\delta_1) < f(\delta_1') \), hence \( \delta_1 < \delta_1' \). This implies \( \delta < \delta' \).

iii) The root of \( f(\delta) \) is mapped into the right immediate subtree of \( f(\delta') \)
If \( m = 1 \), then \( \delta \) has to be 0, hence \( \delta \leq \delta' \). If \( m > 1 \), then \( f(\delta) \leq f(\delta') \) yields \( f(\delta) \leq f(\Omega^\delta \gamma_2 + \cdots + \Omega^\delta \gamma_m) \), hence \( \delta < \Omega^\delta \gamma_1 + \cdots + \Omega^\delta \gamma_m = \delta' \).

iv) The root of \( f(\delta) \) is mapped on the root of \( f(\delta') \)
Let \( \delta = \Omega^\delta \gamma_1 + \cdots + \Omega^\delta \gamma_n \). If the immediate left subtree of \( f(\delta) \) is mapped into \( f(\delta_1') \), then like in case ii), we obtain that \( \delta < \delta' \). So suppose not, then \( f(\delta_1') \leq f(\delta) \) and \( g(\gamma_1) \leq g(\gamma_1') \). Hence \( \delta_1 \leq \delta_1' \) and \( \gamma_1 \leq \gamma_1' \). If \( \delta_1 < \delta_1' \), then \( \delta < \delta' \). Assume \( \delta_1 = \delta_1' \). If \( \gamma_1 < \gamma_1' \), then \( \delta < \delta' \). Hence assume \( \gamma_1 = \gamma_1' \). If \( m = 1 \),
then $\delta = \delta'$. So assume $m > 1$. There are two cases:

- $n = 1$
  We obtain easily that $\delta < \delta'$.

- $n \geq 2$
  Then $f(\Omega^{\delta_2} \gamma_2 + \cdots + \Omega^{\delta_n} \gamma_n) \leq f(\Omega^{\delta'_2} \gamma_2 + \cdots + \Omega^{\delta'_n} \gamma_n)$. Hence the induction hypothesis yields $\Omega^{\delta_2} \gamma_2 + \cdots + \Omega^{\delta_n} \gamma_n \leq \Omega^{\delta'_2} \gamma_2 + \cdots + \Omega^{\delta'_n} \gamma_n$, hence $\delta \leq \delta'$.

\section{Bounds on the proof-theoretical ordinals of $RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^- \text{ and } RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-$}

In the last section, we show some independence and provability results about the already studied $\omega$-po's. To obtain such results, we need bounds on the proof-theoretical ordinals of theories with light-face $\Pi_1^1$-comprehension, which are obtained in this intermediate section. More specifically, we calculate the proof-theoretical ordinal of $RCA_0 + (\Pi_1^1 - CA_0)^-$. As a side question, we were wondering what would happen with the ordinal if we replace $RCA_0$ by $RCA_0^*$. We could not pinpoint down the exact strength of these theories, but we could do it for restricted versions. These restricted theories are fortunately strong enough to obtain the independence and provability results that we want.

**Definition 5.1.** Let $n \leq \omega$. A $\Pi_1^1(\Pi_0^0)$-formula is a formula of the form $\forall X B(X)$, where $B(X)$ is $\Pi_n^0$. $B$ can contain set and numerical parameters. Note that a $\Pi_0^0$-formula is the same as an arithmetical formula, hence a $\Pi_1^1(\Pi_0^0)$ is a standard $\Pi_1^1$-formula. A $(\Pi_1^1(\Pi_0^0))^-$-formula is a formula of the form $\forall X B(X)$, where $B(X)$ is $\Pi_n^0$ and $\forall X B(X)$ contains no free set parameters. It is allowed that $B$ contains numerical parameters.

**Definition 5.2.** Let $(\mathcal{F} - CA_0)$ be the following well-known comprehension scheme

$$\exists Z \forall n(n \in Z \leftrightarrow A(n)),$$

where $A(n)$ is a formula in the class $\mathcal{F}$. If $\mathcal{F} = (\Pi_1^1(\Pi_0^0))^-$, we denote $(\mathcal{F} - CA_0)$ by $(\Pi_1^1(\Pi_0^0)^- - CA_0)^-$

We show the following results.

**Theorem 5.3.**

1. $|RCA_0 + (\Pi_1^1(\Pi_3^0) - CA_0)^-| = \vartheta(\Omega^\omega)$,

2. $|RCA_0^* + (\Pi_1^1(\Pi_3^0) - CA_0)^-| \geq \varphi \omega 0 = \vartheta(\Omega \cdot \omega)$. 

For clarity, we give the definition of $ACA_0$, $RCA_0$ and $RCA_0^*$

**Definition 5.4.** Both $ACA_0$ and $RCA_0$ are theories in the language of second order arithmetic $L_2$. The axioms of $ACA_0$ are the basic axioms in second order arithmetic and the induction axiom for sets together with comprehension for arithmetical formulas. $RCA_0$ consists of the basic axioms, $\Sigma^0_1$-induction and $\Delta^0_1$-comprehension. Let $L_2(exp)$ be the language $L_2$ augmented by a binary operation symbol $exp$ that denotes the exponential function. Define $RCA_0^*$ as the $L_2(exp)$-theory consisting of the basic axioms, the exponentiation axioms, $\Sigma^0_1$-induction and $\Delta^0_1$-comprehension.

From more information on reverse mathematics and theories in second order arithmetic, we refer the reader to [21].

### 5.1 Lower bounds

These proofs and definitions follow the procedure as in [16], but they need some refinements. Firstly, we give a primitive recursive ordinal notation system for $\vartheta(\varepsilon_{\Omega+1})$ which is suitable for using in $ACA_0 + (F–CA_0)$. Then, we introduce an notation system without $\omega$-exponentiation which is suitable to use in $RCA_0 + (F–CA_0)$ and $RCA_0^* + (F–CA_0)$.

**Definition 5.5.** Define inductively a set $OT(\vartheta)$ of ordinals and a natural number $G_\vartheta \alpha$ for $\alpha \in OT(\vartheta)$ as follows:

1. $0 \in OT(\vartheta)$ and $G_\vartheta(0) := 0$,
2. if $\alpha = \Omega^{\alpha_1} \beta_1 + \cdots + \Omega^{\alpha_n} \beta_n$ with $n \geq 1$, $\alpha_1 > \cdots > \alpha_n$ and $\Omega > \beta_1, \ldots, \beta_n > 0$, then
   (a) if $(n > 1$ or $\alpha_1 > 0$) and $\alpha_1,\ldots,\alpha_n, \beta_1,\ldots, \beta_n \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta \alpha := \max\{G_\vartheta(\alpha_1),\ldots, G_\vartheta(\alpha_n), G_\vartheta(\beta_1),\ldots, G_\vartheta(\beta_n)\} + 1$,
   (b) if $n = 1$, $\alpha_1 = 0$ and $\alpha = \omega^{\delta_1} + \cdots + \omega^{\delta_m} > \delta_1 \geq \cdots \geq \delta_m$ with $m \geq 2$ and $\delta_1,\ldots, \delta_m \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta \alpha := \max\{G_\vartheta(\delta_1),\ldots, G_\vartheta(\delta_m)\} + 1$,
3. if $\alpha = \vartheta \beta$ and $\beta \in OT(\vartheta)$, then $\alpha \in OT(\vartheta)$ and $G_\vartheta \alpha := G_\vartheta \beta + 1$.

Because $\vartheta \beta$ is always additively closed and $\vartheta$ is injective, the function $G_\vartheta$ is well-defined. Remark that $\Omega \in OT(\vartheta)$, because $\Omega = \Omega^1 \cdot 1$ and $1 = \vartheta(0)$.

**Lemma 5.6.** If $\xi \in OT(\vartheta)$, then $K(\xi) \subseteq OT(\vartheta)$. Furthermore, $G_\vartheta(k(\xi)) \leq G_\vartheta(\xi)$ for all $\xi$ in $OT(\vartheta)$.
An order-theoretic characterization of the Howard-Bachmann-hierarchy

\textbf{Proof.} We proof this by induction on $G_{\vartheta}(\xi)$. If $\xi = 0$, then this trivially holds. If $\xi = \vartheta(\xi')$, then $K(\xi) = \{\xi\}$, hence this also trivially holds. Assume $\xi = \Omega^k\beta_1 + \cdots + \Omega^{\xi_n}\beta_n$ with $n \geq 1$, $\xi_1 > \cdots > \xi_n$ and $\Omega > \xi_1, \ldots, \xi_n > 0$. If $n = 1$ and $\xi_1 = 0$, then also $K(\xi) = \{\xi\}$, hence the proof is valid. Let now $n > 1$ or $\xi_1 > 0$. Then $K(\xi) = \{\beta_1, \ldots, \beta_n\} \cup K(\xi_1) \cup \cdots \cup K(\xi_n)$. The induction hypothesis yields $K(\xi) \subseteq OT(\vartheta)$. Furthermore, $G_{\vartheta}(k(\xi_i)) \leq G_{\vartheta}(\xi_i) < G_{\vartheta}(\xi)$. Therefore, the strict inequality $G_{\vartheta}(k(\xi)) = G_{\vartheta}(\max_i\{k(\xi_i), \beta_i\}) < G_{\vartheta}(\xi)$ holds.

Each ordinal $\alpha \in OT(\vartheta)$ has a unique normal form using the symbols $0, \omega, \Omega, +, \vartheta$. The relation $\alpha < \beta$ can expressed using the ordinals appearing in their normal form (by Lemma 2.18), which have strictly smaller $G_{\vartheta}$-values by the previous Lemma 5.6. Hence, the following lemma holds.

\textbf{Lemma 5.7.} If we use a specific coding of $(OT(\vartheta), <_{OT(\vartheta)})$ in the natural numbers, then $(OT(\vartheta) \cap \Omega, <_{OT(\vartheta)})$ can be interpreted as a primitive recursive ordinal notation system for the ordinal $\vartheta(\varepsilon_{\Omega+1})$. Furthermore, one can choose this coding in such a way that $\forall \xi \in K(\alpha)(\xi \leq_N \alpha)$.

This is the ordinal notation system that we will use if we work in the theory $ACA_0 + (F-CA_0)$. However, in the theory $RCA_0 + (F-CA_0)$ and $RCA^*_0 + (F-CA_0)$, we use a different ordinal notation system $OT'(\vartheta)$ without $\omega$-exponentiation.

\textbf{Definition 5.8.} Define inductively a set $OT'(\vartheta)$ of ordinals and a natural number $G'_{\vartheta}\alpha$ for $\alpha \in OT'(\vartheta)$ as follows:

1. $0 \in OT'(\vartheta)$ and $G'_{\vartheta}(0) := 0$,
2. if $\alpha = \Omega^n\beta_1 + \cdots + \Omega^0\beta_0$, with $\Omega > \beta_n > 0$ and $\Omega > \beta_0, \ldots, \beta_{n-1}$, then
   \begin{enumerate}
   \item if $n > 0$ and $\beta_0, \ldots, \beta_n \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_{\vartheta}\alpha := \max\{G'_{\vartheta}(\beta_0), \ldots, G'_{\vartheta}(\beta_n)\} + 1$,
   \item if $n = 0$ and $\alpha = NF \delta_1 + \cdots + \delta_m$ with $m \geq 2$ and $\delta_1, \ldots, \delta_m \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_{\vartheta}\alpha := \max\{G'_{\vartheta}(\delta_1), \ldots, G'_{\vartheta}(\delta_m)\} + 1$,
   \end{enumerate}
3. if $\alpha = \vartheta\beta$ and $\beta \in OT'(\vartheta)$, then $\alpha \in OT'(\vartheta)$ and $G'_{\vartheta}\alpha := G'_{\vartheta}\beta + 1$.

NF stands for normal form.

Also in this ordinal notation system, $G'_{\vartheta}$ is well-defined.

\textbf{Lemma 5.9.} If $\xi \in OT'(\vartheta)$, then $K(\xi) \subseteq OT'(\vartheta)$. Furthermore, $G'_{\vartheta}(k(\xi)) \leq G'_{\vartheta}(\xi)$ for all $\xi$ in $OT'(\vartheta)$. 

Proof. If \( \xi = 0 \), then this trivially holds. If \( \xi = \vartheta(\xi') \), then \( K(\xi) = \{ \xi \} \), hence this also trivially holds. Assume \( \xi = \Omega^n \xi_n + \cdots + \Omega^0 \xi_0 \) with \( \Omega > \xi_n > 0 \) and \( \Omega > \xi_1, \ldots, \xi_{-1} \). If \( n = 0 \), then also \( K(\xi) = \{ \xi \} \), hence the proof is valid. Let now \( n > 0 \). Then \( K(\xi) \subseteq \{ n, \ldots, 0, \xi_0, \ldots, \xi_n \} \). We know \( \xi_0, \ldots, \xi_n \in OT'(\vartheta) \). Additionally, it is trivial to show \( \{ n, \ldots, 0 \} \subseteq OT'(\vartheta) \). Hence, \( K(\xi) \subseteq OT'(\vartheta) \). It is also trivial to show that \( G'(n) \leq 2 \) for all natural numbers \( m \). Therefore, \( G'(n) \leq G'(\xi) \) because \( G'(\xi_n) \geq 1 \). Also, \( G'(\xi_i) < G'(\xi) \). Hence \( G'(k(\xi)) = G'(\max(\{ n, \xi_i \}) < G'(\xi) \).

Like for the first ordinal notation system, we have to following lemma.

**Lemma 5.10.** If we use a specific coding of \((OT'(\vartheta), \prec_{OT'(\vartheta)})\) in the natural numbers, \((OT'(\vartheta) \cap \Omega, \prec_{OT'(\vartheta)})\) can be interpreted as a primitive recursive ordinal notation system for the ordinal \( \vartheta(\Omega^\omega) \). Furthermore, one can choose this coding in such a way that \( \forall \xi \in K(\vartheta)(\xi \leq_\Omega \alpha) \).

From now on, we fix primitive recursive ordinal notation systems \( OT(\vartheta) \cap \Omega \) and \( OT'(\vartheta) \cap \Omega \). If we mention \( ACA_0 \) in the beginning of a theorem, then we assume that we work in \( OT(\vartheta) \). Similarly, if we mention \( RCA_0 \) or \( RCA'_0 \) in the beginning of a theorem, we assume that we work in \( OT'(\vartheta) \).

**Definition 5.11.**

\[
\begin{align*}
Prog(<, F) &:= \forall x (\forall y (y < x \rightarrow F(y)) \rightarrow F(x)), \\
TI(<, F) &:= Prog(<, F) \rightarrow \forall x F(x), \\
WF(<) &:= \forall X TI(<, X),
\end{align*}
\]

where the formula \( Prog(<, F) \) stands for progressiveness, \( TI(<, F) \) for transfinite induction and \( WF(<) \) for well-foundedness. \( F(x) \) is an arbitrary \( L_2(\exp) \)-formula if we work in \( RCA'_0 \) or an arbitrary \( L_2 \)-formula if we work in \( RCA_0 \) or \( ACA_0 \). For an element \( \alpha \in OT(\vartheta) \), the formula \( WF(\alpha) \) stands for ‘\( \prec_{OT(\vartheta)} \) restricted to \( \{ \beta \in OT(\vartheta) : \beta < \alpha \} \) is well-founded’. We sometimes also denote this as \( WF(<\alpha) \). Similarly for \( OT'(\vartheta) \).

**Definition 5.12.**

1. \( \text{Acc} := \{ \alpha < \Omega : WF(<\alpha) \} \),

2. \( M := \{ \alpha : K(\alpha) \subseteq \text{Acc} \} \),

3. \( \alpha \prec_\Omega \beta \Leftrightarrow \alpha, \beta \in M \wedge \alpha < \beta \).

The next lemma shows that \( \text{Acc} \), \( M \) and \( \prec_\Omega \) can be expressed by a \((\Pi_1(\Pi_0^b))^*\)-formula.
Lemma 5.13. $\text{Acc}, M$ and $<_\Omega$ are expressible by a $(\Pi^1_1(\Pi^0_3))^{-}$-formula.

Proof. The proof is the same for the ordinal notation system $OT(\psi)$ and $OT'(\psi)$.

$$WF(\alpha) = \forall X (\forall x (\forall y (y < x \to y \in X) \to x \in X) \to \forall x (x \in X)),$$

where $<$ is $<_\alpha$. It is easy to see that the prenex normal form of the formula $WF(\alpha)$ is $(\Pi^1_1(\Pi^0_3))^{-}$, hence $\text{Acc}$ can be expressed by such a formula. $M$ can be represented by the formula $\forall \xi \leq N \alpha (\xi \in \text{K}(\alpha) \to \xi \in \text{Acc})$. Because $\xi \in \text{K}(\alpha)$ is elementary recursive, both $M$ and $<_{\Omega}$ are also expressible by a $(\Pi^1_1(\Pi^0_3))^{-}$-formula.

\[\blacksquare\]

Lemma 5.14. $(\text{RCA}_0^*) \alpha, \beta \in \text{Acc} \implies \alpha + \beta \in \text{Acc}$.

Proof. Obvious. \[\blacksquare\]

Lemma 5.15. $(\text{ACA}_0) \alpha \in \text{Acc} \implies \omega^\alpha \in \text{Acc}$.

Proof. A proof of this lemma goes back to Gentzen. See [14][18]. \[\blacksquare\]

Definition 5.16. Let $\text{Prog}_\Omega(X)$ be the formula

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

Let $\text{Acc}_\Omega$ be the set $\{\alpha \in M : \psi(\alpha) \in \text{Acc}\}$.

Lemma 5.17. $(\text{RCA}_0^*) \text{Prog}_\Omega(\text{Acc}_\Omega)$.

Proof. We work in $OT'(\psi)$. Assume $\alpha \in M$ and $(\forall \beta <_\Omega \alpha) (\psi(\beta) \in \text{Acc})$. We want to prove that $\psi(\alpha) \in \text{Acc}$. We show that $(\forall \xi < \psi(\alpha))(\xi \in \text{Acc})$ by induction on $G_\psi(\xi)$, from which the theorem follows. If $\xi = 0$, then this trivially holds. So assume $\xi > 0$.

a) Assume $\xi \notin P$

Because $\xi < \Omega$, we have $\xi = N\xi_1 + \cdots + \xi_n > \xi_1 \geq \cdots \geq \xi_n$ ($n \geq 2$). The induction hypothesis yields $\xi_1 \in \text{Acc}$. Hence from Lemma 5.14, we obtain $\xi \in \text{Acc}$.

b) Assume $\xi = \psi(\xi')$

From $\psi(\xi') < \psi(\alpha)$, we obtain either $\xi' < \alpha$ and $k(\xi') < \psi(\alpha)$ or $\psi(\xi') \leq k(\alpha)$. In the former case, $G_\psi(k(\xi')) \leq G_\psi(\xi') < G_\psi(\xi)$ and the induction hypothesis implies $k(\xi') \in \text{Acc}$, hence $K(\xi'') \subseteq \text{Acc}$. So $\xi' \in M$, from which it follows $\xi' <_{\Omega} \alpha$, hence $\psi(\xi') \in \text{Acc}$. In the latter case, we know that $k(\alpha) \in \text{Acc}$, because $\alpha \in M$. Therefore, $\psi(\xi') \in \text{Acc}$. \[\blacksquare\]
We actually do not need the following lemma, because we already know the proof-theoretical ordinal of $ACA_0 + (\Pi^1_1 C A_0)^\sim$. However, just for the completeness, we mention it here.

**Lemma 5.18.** $(ACA_0)$ $Prog_\Omega(Acc_\Omega)$.

**Proof.** The proof uses $OT(\theta)$ and follows the same procedure as Lemma 5.17. The only difference is the usage of Lemma 5.15.

**Lemma 5.19.** Let $A(a)$ be a $(\Pi^1_1(\Pi^0_3))\sim$-formula. Define $A_k$ as

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

Then $RCA_0 + (\Pi^1_1(\Pi^0_3) - C A_0)^\sim$ proves $Prog_\Omega(\{\xi : A(\xi)\}) \rightarrow A_k$.

**Proof.** We will prove that by outer induction on $k$. First note that $A(a)$ can be expressed by a set in $RCA_0 + (\Pi^1_1(\Pi^0_3) - C A_0)^\sim$. Assume that $Prog_\Omega(\{\xi : A(\xi)\})$ and pick an arbitrary $\alpha$ such that $(\forall \beta <_\Omega \alpha) A(\beta)$. If $\alpha \notin M$, the assertion trivially follows. Assume $\alpha \in M$. If $k = 0$, the proofs follows easily from $Prog_\Omega(\{\xi : A(\xi)\})$. Assume $k = l + 1$. Let $B(\zeta)$ be

$$(\forall \beta <_\Omega \alpha + \Omega^k \zeta) A(\beta).$$

$B(\zeta)$ is a $\Pi^0_3$-formula in $A$. We prove by induction on $\zeta$ that $(\forall \zeta \in Acc) B(\zeta)$ is true.

The assumption yields $B(0)$. Assume that $\zeta \in Acc$ and $\zeta$ is a limit. If $\beta <_\Omega \alpha + \Omega^l \zeta$, then there exists a $\zeta' < \zeta$ such that $\beta <_\Omega \alpha + \Omega^l \zeta'$. Because $B(\zeta')$ is valid, one obtains $A(\beta)$. Assume that $\zeta = \zeta' + 1 \in Acc$. By the induction hypothesis, one obtains $(\forall \beta <_\Omega \alpha + \Omega^l \zeta') A(\beta)$. Because $A_1$ is valid, we get $(\forall \beta <_\Omega \alpha + \Omega^l \zeta' + \Omega^l) A(\beta)$, hence $(\forall \beta <_\Omega \alpha + \Omega^l (\zeta' + 1)) A(\beta) = B(\zeta)$.

From $\Pi^0_3$-induction one can conclude $(\forall \zeta \in Acc) B(\zeta)$. Now, if $\beta <_\Omega \alpha + \Omega^{l+1}$, there exists a certain $\zeta \in Acc$, such that $\beta <_\Omega \alpha + \Omega^{l+1} \zeta$. Therefore, $(\forall \beta <_\Omega \alpha + \Omega^{l+1}) A(\beta)$ follows from $(\forall \zeta \in Acc) B(\zeta)$.

**Theorem 5.20.** $|RCA_0 + (\Pi^1_1(\Pi^0_3) - C A_0)^\sim| \geq \theta(\Omega^\omega)$.

**Proof.** This follows from Lemmas 5.17, 5.19 and the fact that $Acc_\Omega$ is expressible by a $(\Pi^1_1(\Pi^0_3))\sim$-formula.

**Lemma 5.21.** Let $A(a)$ be a $(\Pi^1_1(\Pi^0_3))\sim$-formula. Define $A_k$ as

$$(\forall \beta <_\Omega \Omega \cdot k) A(\beta).$$

Then $RCA_0^* + (\Pi^1_1(\Pi^0_3) - C A_0)^\sim$ proves $Prog_\Omega(\{\xi : A(\xi)\}) \rightarrow A_k$.
Proof. We will prove that by outer induction on $k$. First note that $A(a)$ can be expressed by a set in $RCA_0^* + (\Pi^1_1(\Pi^0_3) - CA_0)^-$. It is easy to see that the case $k = 0$ holds. Assume $k = l + 1$ and $\text{Prog}_\Omega(\{\xi : A(\xi)\})$. Then we know $(\forall \beta < \Omega \cdot l) A(\beta)$. Pick an arbitrary $\beta < \Omega \cdot (l + 1)$. If $\beta < \Omega \cdot l$, we obtain $A(\beta)$. Hence assume that $\beta \geq \Omega \cdot l$. There exists a $\xi < \Omega$ such that $\beta = \Omega \cdot l + \xi$. Let $B(\xi)$ be $A(\Omega \cdot l + \xi)$. $B(\xi)$ is a $\Pi^0_3$-formula in $A$. We prove by induction on $\xi$ that $(\forall \xi \in \text{Acc}) B(\xi)$ is true. From this, the theorem follows.

If $\xi = 0$, then $B(\xi)$ is true because $(\forall \beta < \Omega \cdot l) A(\beta)$ and $\text{Prog}_\Omega(\{\xi : A(\xi)\})$ imply $A(\Omega \cdot l)$. Assume $\xi \in \text{Acc}$ is a limit and assume $(\forall \xi' < \Omega) B(\xi)$. From $\text{Prog}_\Omega(\{\xi : A(\xi)\})$, we obtain $B(\xi)$. Let $\xi = \xi' + 1 \in \text{Acc}$. Then $B(\xi)$ follows from $B(\xi')$ and $\text{Prog}_\Omega(\{\xi : A(\xi)\})$.

Theorem 5.22. $|RCA_0^* + (\Pi^1_1(\Pi^0_3) - CA_0)^-| \geq \vartheta(\omega \cdot \omega) = \varphi_\omega 0$.

Proof. This follows from Lemmas 5.17, 5.21 and the fact that $\text{Acc}_\Omega$ is expressible by a $(\Pi^1_1(\Pi^0_3))$-formula.

5.2 Upper bounds

In this subsection, we give an upper bound for $|RCA_0 + (\Pi^1_1(\Pi^0_3) - CA_0)^-|$. For this, we use the fact that $|\Pi^1_2 - BI_0| = \vartheta(\omega^3)$ (see [21]).

Lemma 5.23. For every arithmetical formula $B(X)$ with all free set variables indicated, there is a $\Delta_0$-formula $R(x, X, f)$ such that

1. $ACA_0 \vdash B(X) \rightarrow \exists f \forall x R(x, X, f)$
2. If $T$ is a theory with $RCA_0 \subseteq T$ and $F(x, y)$ is an arbitrary formula,
   
   $T \vdash \forall x \exists y F(x, y) \land \forall x \exists z (lh(z) = x \land \forall i < x F(i, z))$,

   then $T \vdash \forall x R(x, X, F) \rightarrow B(X)$, where $R(x, X, F)$ results from $R(x, X, f)$ by replacing subformulae of the form $f(t) = s$ by $F(t, s)$.

Proof. This proof is a little adaptation of the normal form theorem V.1.4 in [21]. We can assume that $B(X)$ is in prenex normal form

$B(X) \equiv \forall x_1 \exists y_1 \ldots \forall x_r \exists y_r S(x_1, y_1, \ldots, x_r, y_r, X),$

with $S$ quantifier-free. Then over $ACA_0$, we have

$B(X) \leftrightarrow \exists f_1 \forall x_1 x_2 \exists y_2 \ldots S(x_1, f_1(x_1), x_2, y_2, \ldots, x_r, y_r, X)$

$\leftrightarrow \exists f_1 \exists f_2 \forall x_1 x_2 x_3 \exists y_3 \ldots S(x_1, f_1(x_1), x_2, f_2(x_2, x_2), x_3, y_3, \ldots, x_r, y_r, X)$

$\vdots$

$\leftrightarrow \exists f_1 \ldots f_r \forall x_1 \ldots x_r S(x_1, f_1(x_1), x_2, f_2(x_2, x_2), \ldots, x_r, f_r(x_1, \ldots, x_r), X)$

$\leftrightarrow \exists f \forall x S((x_1, (f_1)((x_1)_1)), \ldots, (x_r, (f_r)((x_1)_1, \ldots, (x_r)_1), X),$
where \((f)_k((x)_1, \ldots, (x)_k) = f((k, (x)_1, \ldots, (x)_k))\) and \((\ldots)\) is some primitive recursive tupel coding. Now let \(R(x, X, f) := S((x)_1, (f)_1((x)_1), \ldots, X)\). Note that the right-to-left directions are provable in \(RCA_0\), so that part 2. follows from reading the equivalences from bottom to top.

By adapting the previous lemma, we get the following corollary.

**Corollary 5.24.** If in addition to the conditions of Lemma 5.23(2.), \(T\) also satisfies
\[
T \vdash \forall x \exists y \mathcal{G}(x, y) \land \forall x \exists z(lh(z) = x \land \forall i < x \mathcal{G}(i, (z)_i)),
\]
then \(T \vdash \forall x R(x, \mathcal{G}, \mathcal{F}) \rightarrow B(\mathcal{G})\), where \(R(x, \mathcal{G}, \mathcal{F})\) results from \(R(x, X, \mathcal{F})\) by replacing \(t \in X\) by \(\mathcal{G}(t, 0)\).

**Lemma 5.25.** Assuming that the conditions for \(\mathcal{F}\) and \(\mathcal{G}\) from Lemma 5.23 and Corollary 5.24 are satisfied, then there exists a \(\Delta^0_1\)-formula \(P\) such that
\[
T \vdash \forall x R(x, \mathcal{G}, \mathcal{F}) \iff \forall x P(\mathcal{G}[x], \mathcal{F}[x])
\]

**Proof.** By induction on the build-up of \(R\).

**Theorem 5.26.** Assume that \(B(X, n)\) is a \(\Pi^0_3\)-formula, then
\[
\Pi^1_2 - BI_0 \vdash \forall n (\exists A \in Rec((\Pi^1_2)^-)) B(A(\cdot), n) \iff (\forall X \subseteq \omega) B(X, n),
\]
where \(B(A(\cdot), n)\) results from \(B(X, n)\) by replacing \(t \in X\) by \(A(t)\).

**Proof.** Pick an arbitrary natural number \(n_0\). Assume that we have
\[
(\exists A \in Rec((\Pi^1_2)^-)) \neg B(A(\cdot), n_0).
\]
In [10] it is proven that \(\Sigma^1_1\)-\(\omega\)-model reflection holds in \(\Pi^1_2 - BI_0\). Therefore, there exists an \(\omega\)-model \(M\) such that \(n_0 \in M\) and \(M \models ACA_0 + \neg B(A, n_0)\). Define \(X\) as the set \(\{m \in \omega : M \models A(m)\}\). This set exists by arithmetical comprehension. Hence, \(\neg B(X, n_0)\) is valid.

For the reverse implication, assume \((\exists X \subseteq \omega) \neg B(X, n_0)\). Using Lemmas 5.23 and 5.25 on the formula \(\neg B(X, n_0)\), we have \(\exists X \forall x P(X[x], f[x])\), thus \(\exists h \forall x P(h_0[x], h_1[x])\). Define \(I(\cdot)\) as \(\exists h(\sigma \subseteq h \land \forall x P(h_0[x], h_1[x]))\). Define a class function \(\mathcal{H}\) by \(\mathcal{H}(0) = \emptyset\) and \(\mathcal{H}(n + 1) = \mathcal{H}(n) \cup \langle \mu k : I(\mathcal{H}(n) \cap k) \rangle\). Note that \(\mathcal{H}\) has a \(\Sigma^1_2\)-graph. As \(\Pi^1_2 - BI_0\) proves \(\Sigma^1_2\)-induction on the naturals, one can prove that \(\mathcal{H}(n)\) is always defined by \(\exists h \forall x P(h_0[x], h_1[x])\). Also note that \(\mathcal{H}\) is recursive in \(I\). We can show that \(I\) is recursive in a \((\Pi^1_2)^-\)-formula. This because \(I\) is recursive in \(W\), the set of all recursive well-founded trees, which is \((\Pi^1_2)^-\)-complete. Therefore, we obtain that \(\mathcal{H} \in Rec((\Pi^1_2)^-)\). We also have \(\forall x P(\mathcal{H}_0[x], \mathcal{H}_1[x])\) by this construction. Letting \(\mathcal{G} = \mathcal{H}_0\) and \(\mathcal{F} = \mathcal{H}_1\) in Lemma 5.25 we get \(\forall x R(x, \mathcal{G}, \mathcal{F})\). By Corollary 5.24 we get \(\neg B(\mathcal{G}, n_0)\).
Corollary 5.27. \(|RCA_0 + (\Pi^1_1(\Pi^0_0) - CA_0)^-| = \vartheta(\Omega^\omega).\)

**Proof.** Follows from Theorems 5.20, 5.26 and the fact that \(|\Pi^1_2 - BI_0| = \vartheta(\Omega^\omega).\)

In a similar way, one could also show the following lemma.

**Lemma 5.28.** Assume that \(B(X, n)\) is a \(\Pi^0_2\)-formula, then

\[\Pi^1_1 - BI_0 \vdash \forall n[\forall A \in \text{Rec}((\Pi^1_1)^-)B(A(\cdot), n) \leftrightarrow (\forall X \subseteq \omega)B(X, n)].\]

Therefore, \(|\Pi^1_1 - BI_0| = \varphi_\omega 0\) yields that the ordinal of \(RCA_0^* + (\Pi^1_1(\Pi^0_0) - CA_0)^-\) is bounded above by \(\varphi_\omega 0\). Actually one can prove that the ordinal of the theory \(RCA_0^* + (\Pi^1_1(\Pi^0_0) - CA_0)^-\) is even much lower. Using \(WKL_0\) and the normal form theorem II.2.7 in [21], one can prove that every \((\Pi^1_1(\Pi^0_0))^\ominus\)-formula is equivalent with a \((\Pi^0_2)^\ominus\)-formula (and if the original formula has extra parameters, then the equivalent one will also have those parameters). The intuitive idea behind this is that \(WKL_0\) proves that the projection of closed set is a closed set, meaning that \(\exists X \forall x \ldots\) can be reduced to \(\exists z \ldots\) (a closed set can be seen in some sense as a \(\Pi^0_2\)-formula). Furthermore, one can proof the following lemmas.

**Lemma 5.29.** Let \(F\) be \((\Pi^1_1(\Pi^0_0))^\ominus\) or \((\Pi^0_0)^\ominus\). Then the first order part of the theory \(WKL_0 + F - CA_0\) is the same as that of \(RCA_0 + F - CA_0\).

**Proof.** Can be proved easily by little adaptations of paragraph IX.2 in [21].

In a similar way,

**Lemma 5.30.** The first order part of \(RCA_0 + (\Pi^0_2)^\ominus - \text{IND}\) is \(I\Sigma_3\).

**Proof.** Can be proved by adaptations of paragraph IX.1 in [21].

Hence, \(|RCA_0 + (\Pi^1_1(\Pi^0_0) - CA_0)^-| = |I\Sigma_3| = \omega^{\omega^\omega}\). In a similar but more technical way, one could prove that \(RCA_0^* + (\Pi^1_1(\Pi^0_0))^\ominus\) is \(\Pi_4\)-conservative over \(I\Sigma_2\). Hence, \(|RCA_0^* + (\Pi^1_1(\Pi^0_0))^\ominus| = \omega^{\omega^\omega}\). We want to thank Leszek Kołodziejczyk, who reminded us of these facts.

There are still some open questions left, for example what is the ordinal of \(RCA_0 + (\Pi^1_1(\Pi^0_1) - CA_0)^-\) or \(RCA_0 + (\Pi^1_1 - CA_0)^-\) etc. We state the following conjectures.

**Conjecture 5.31.**
6 Independence results

6.1 Independence results for $ACA_0 + (\Pi_1 - CA_0)^-$

It is well-known that $|ACA_0 + (\Pi_1 - CA_0)^-| = \varphi(\omega)$. Now,

**Theorem 6.1.** $ACA_0 \vdash \forall X (X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$.

**Proof.** Theorem 4.3 can be carried out in $ACA_0$ if we use the primitive recursive ordinal notation system $OT(\varphi) \cap \Omega$ for $\varphi(\varepsilon + 1)$. Therefore, one has the following theorem,

**Theorem 6.2.** $ACA_0 + (\Pi_1 - CA_0)^- \not\vdash \forall X (X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$.

The $\text{wpo } T(B(\cdot))$ can be seen as the limit of $T(\cdot \cdots \cdot \cdots)$, because a binary tree can be seen as an iteration of the $\ast$-operator. For example, one can interpret an element of $\{a\}^{\ast\ast}$ as a binary tree which goes only one time to the left and every node has label $a$. So, it would be interesting if $ACA_0 + (\Pi_1 - CA_0)^- \vdash \forall X (X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$ is a wpo for every natural number $n$ because this theory does not prove the ‘limit’. The maximal order type of $T(\cdot \cdots \cdot \cdots)$ is strictly below the proof-theoretical ordinal of $ACA_0 + (\Pi_1 - CA_0)^-$. Therefore, one can really expect this provability assertion. However, in the proof of the upper bound of the maximal order type of $T(\cdot \cdots \cdot \cdots)$ one uses several induction schemes and it is not immediately clear that the proof goes through in $ACA_0 + (\Pi_1 - CA_0)^-$. Theorem 6.4 shows that this is indeed possible and uses the so-called minimal bad sequence argument. For more information about the minimal bad sequence argument and its reverse mathematical strength, we refer the reader to [13].

**Lemma 6.3.** Over $RCA_0$, the following are equivalent

1. $ACA_0$,
2. Higman’s theorem, i.e. $\forall X (X \text{ is a wpo} \rightarrow X^* \text{ is a wpo})$.
3. $\forall X (X \text{ is a well-quasi-order} \rightarrow X^* \text{ is a well-quasi-order})$.
Theorem 6.4. For all $n$, $ACA_0 + (\Pi_1^1 - CA_0)^- \vdash \mathcal{T}(\cdots \ast \cdots \ast)$ is a w.p.o.'. 

Proof. Fix a natural number $n$. We reason in $ACA_0 + (\Pi_1^1 - CA_0)^-$. Code the elements of $\mathcal{T}(\cdots \ast \cdots \ast)$ as natural numbers such that $\circ$ has code 0 and the code of $t_i$ is strictly smaller than the code of $t = \circ[w(t_1, \ldots, t_i, \ldots, t_n)]$. This coding can be done primitive recursively. We say that the leaves of $\circ[w(t_1, \ldots, t_i, \ldots, t_n)]$ are $\{t_1, \ldots, t_n\}$ and we assume that there is a primitive recursive relation ‘...is a leaf of ...’. If $\sigma$ decodes a finite sequence, then this sequence is equal to $((\sigma)_0, \ldots, (\sigma)_{lh(\sigma) - 1})$, where $lh(\sigma)$ is the length and $(\sigma)_i$ is the $i$-th element of the sequence. An infinite sequence $(\sigma_i)_{i < \omega}$ is decoded as a set $\{(i, \sigma_i) : i < \omega\}$. $\{(i, \sigma_i) : i < \omega\}$ stands for the element $\sigma_i$. Note that one can recursively construct the set $\{(\sigma_0, \ldots, \sigma_i) : i < \omega\}$ from the original set $\{(i, \sigma_i) : i < \omega\}$, where $(\sigma_0, \ldots, \sigma_i)$ is decoded as a natural number. Furthermore, if one has the set $\{(\sigma_0, \ldots, \sigma_i) : i < \omega\}$, one can reconstruct the original set $\{(i, \sigma_i) : i < \omega\}$ from it in a recursive way.

Now, assume that $\mathcal{T}(\cdots \ast \cdots \ast)$ is not a well-partial-order. Then there exists an infinite sequence $(t_i)_{i < \omega}$ in $\mathcal{T}(\cdots \ast \cdots \ast)$ such that $\forall i, j(i < j \rightarrow t_i \not\leq t_j)$. Define $\chi(\sigma)$ as

- $\sigma$ is a finite sequence of elements in $\mathcal{T}(\cdots \ast \cdots \ast)$
- $\exists Z(Z$ is an infinite bad sequence in $\mathcal{T}(\ast \cdots \ast)$ and $\forall i < lh(\sigma)((\sigma)_i = (Z)_i)$

and $\psi(\sigma)$ as

- $\sigma$ is a finite sequence of elements in $\mathcal{T}(\ast \cdots \ast)$
- $\forall Y[(Y$ is an infinite bad sequence in $\mathcal{T}(\ast \cdots \ast)$$ \rightarrow i < lh(\sigma) \forall k < i ((\sigma)_k = (Y)_k) \rightarrow (\sigma)_i \not\leq (Y)_i)]$.

Note that $(\sigma)_i \leq (Y)_i$ is interpreted as the inequality relation between natural numbers and not between elements of $\mathcal{T}(\ast \cdots \ast)$. Using $(\Pi_1^1 - CA_0)^-$, there exists a set $S$ such that $\sigma \in S \leftrightarrow \chi(\sigma) \wedge \psi(\sigma)$. Choose now two arbitrary elements $s, s'$ in $S$. We want to prove that either $s$ is an initial segment of $s'$ or $s'$ an initial segment of $s$. Assume there is an index $i < \min\{lh(s), lh(s')\}$ such that
\((s)_i < (s')_i \) and \(\forall k < i, (s)_k = (s')_k\). The case \((s)_i > (s')_i\) can be handled in a similar way. Note that \((s)_i < (s')_i\) is seen as an inequality between natural numbers and not between elements of \(T^{(\star \cdots \star)}\). Because \(s\) is in \(S\), \(s\) can be extended to an infinite bad sequence \(Y\) in \(T^{(\star \cdots \star)}\). This, however, contradicts \(\psi(s')\) because 
\((Y)_k = (s)_k = (s')_k\) for all \(k < i\), but \((Y)_i = (s)_i < (s')_i\).

If \(s \in S\), one can show by minimization in \(RCA_0\), that there is a \(z \in T^{(\star \cdots \star)}\) such that \(s\preceq(z) \in S\). Therefore, there exists an infinite sequence \((s_i)_{i<\omega}\) in \(T^{(\star \cdots \star)}\) such that \(S = \{(s_0, \ldots, s_i) : i < \omega\}\).

Now, define \(subS\) as the set of all pairs \((i, t)\) such that \(t\) is a leaf of \(s_i\). Remark that \(subS\) is definable in \(RCA_0\), because

\[(i, t) \in subS \iff \exists \sigma(\sigma \in S \text{ and } lh(\sigma) = i + 1 \text{ and } t \text{ is a leaf of } \sigma_i) \]
\[\iff \forall \sigma((\sigma \in S \text{ and } lh(\sigma) = i + 1) \rightarrow t \text{ is a leaf of } \sigma_i)\]

On \(subS\), we define the following ordering: \((i, t) \leq (j, t') \iff t \leq T^{(\star \cdots \star)} t'\). With this ordering \(subS\) is a quasi-order. We want to prove that it is a well-quasi-order.

Assume that this is not true. This implies the existence of an infinite bad sequence \((n_i, s'_i)\) in \(subS\). This implies \(s'_i \leq s_{n_i}\) for all \(i\). Construct now an infinite subsequence \(H\) such that \((n_i)_{i<\omega}\) is strictly increasing and the first element of \(H\) is \((n_0, s'_0)\). So \(H = \{(i, (n_j, s'_j)) : i < \omega\}\) with \(j_0 = 0\). This is possible in \(RCA_0\), because the number of leaves of an element in \(T^{(\star \cdots \star)}\) is finite.

Construct now recursively a set \(S'\) such that all elements that lie in \(S'\) have the form \((\sigma_0, \ldots, \sigma_i)\) such that if \(i < n_0\), the element \(\sigma_i\) is equal to \(s_i\) and if \(i \geq n_0\), then \(\sigma_i\) is equal to \(s'_{j_i-n_0}\). The existence of \(S'\) is however in contradiction with the definition of \(S\). This is because \(\psi((s_0, \ldots, s_{n_0}))\), \(s'_{j_0} = s'_0 < s_{n_0}\) as a natural number and \((\sigma_i)_{i<\omega}\) is an infinite bad sequence. Take for example \(k < l < \omega\). If \(k < l < n_0\), then \(\sigma_k \not\leq \sigma_l\) follows from \(s_k \not\leq s_l\). If \(n_0 < k < l\), then \(\sigma_k \not\leq \sigma_l\) follows from \(s'_{j_k-n_0} \not\leq s'_{j_l-n_0}\). Assume \(k < n_0 \leq l\). Then \(\sigma_k \not\leq \sigma_l\) follows from the fact that otherwise \(s_k \leq s'_{j_l-n_0} \leq s_{n_l-n_0}\) with \(k < n_0 = n_{j_0} \leq n_{j_l-n_0}\), a contradiction. So we conclude that \(subS\) is indeed a well-quasi-order.

By Lemma [6.3], we obtain that \((subS)^{\star \cdots \star}\) is a well-quasi-ordering. Now, look at the infinite sequence \((s_i)_{i<\omega}\) in \(T^{(\star \cdots \star)}\). Rewrite every element \(\times s_i\) to an element in \((subS)^{\star \cdots \star}\) and call it \(\overline{s_i}\). For example, if \(s_i = o((s_1^1, \ldots, s_{n_1}^1), \ldots, (s_1^k, \ldots, s_{n_k}^k))\), then \(\overline{s_i}\) is equal to \(((i, s_1^1), \ldots, (i, s_{n_1}^1)), \ldots, ((i, s_1^k), \ldots, (i, s_{n_k}^k))\). Because we know that \((subS)^{\star \cdots \star}\) is a well-quasi-order, there exist two indices \(i < j\) such that \(\overline{s_i} \leq (subS)^{\star \cdots \star} \overline{s_j}\). Therefore, \(s_i \leq T^{(\star \cdots \star)} s_j\), a contradiction. □
Remark that the previous proof can actually be simplified. We wrote it however in this way, so that the reader can see what is needed to prove the well-
partial-orderedness of $\mathcal{T}(W)$: the proof can be carried out in $RCA_0 + (\Pi^1_1(\Pi^0_3)-CA_0)^- + \forall X(X$ is a wpo $\rightarrow X^*$ is a wpo$)$. This theory is the same as $ACA_0 + (\Pi^1_1-CA_0)^-$. A general guideline is that for a natural theory $T \supseteq RCA_0$ with $RCA_0 \vdash \forall X('X$ is a wpo $\rightarrow W(X)$ is a wpo$)$, the following holds:

$$T + (\Pi^1_1(\Pi^0_3)-CA_0)^- \vdash \mathcal{T}(W)$$

and

$$T + (\Pi^1_1(\Pi^0_3)-CA_0)^- \nvdash \mathcal{T}(W')$$

where $W'$ is the ‘limit’ of $W$. The notion ‘limit’ is of course very vague. For example a binary tree in $B(X)$ can be seen in some sense as an element of $X^{**}$ and a finite sequence in $X^*$ can be seen as an element of $X^n$. Therefore, $B(\cdot)$ is the limit of $\cdot^{**}$ and $\cdot$ is the limit of $\cdot^n$. We do not give a specific definition of the notion ‘limit’ as this is only a guideline and not a real theorem.

6.2 Independence results for $RCA_0 + (\Pi^1_1(\Pi^0_3)-CA_0)^-$

We know that $|RCA_0 + (\Pi^1_1(\Pi^0_3)-CA_0)^-| = \vartheta(\Omega^\omega)$. Furthermore, one can easily show as in Theorem 4.3 that $RCA_0 \vdash \forall X('X$ is a wpo $\rightarrow WF(\vartheta(\Omega^\omega))$. Therefore, $RCA_0 + (\Pi^1_1(\Pi^0_3)-CA_0)^- \nvdash \forall X('X$ is a wpo $\rightarrow \mathcal{T}(X^*)$ is a wpo$)$.

Again, as in the $ACA_0$-case, one can search for provability results: $\mathcal{T}(X^*)$ can be seen as the limit of $\mathcal{T}(X^n)$. Therefore, it would be interesting if $RCA_0 + (\Pi^1_1(\Pi^0_3)-CA_0)^- \vdash \forall X('X$ is a wpo $\rightarrow 'X^n$ is a wpo$)$ (see [4] for more information). The theorem $\forall X('X$ is a wpo $\rightarrow 'X^n$ is a wpo$)$ is however provable in $RCA_0 + CAC$, where $CAC$ is the principle saying that every infinite sequence in a partial order has a subsequence that is either a chain or an antichain. This implies that $RCA_0 + CAC + (\Pi^1_1(\Pi^0_3)-CA_0)^- \vdash \forall X('X$ is a wpo$)$ is a wpo$).

Acknowledgements

The first author wants to thank his funding organization Fellowship of the Research Foundation - Flanders (FWO). The second author’s research was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. The authors also wish to thank Leszek Kołodziejczyk for his helpful comments and Ryota Akiyoshi for his fruitful discussions with the first author.
References

[1] Thomas Becker and Volker Weispfenning, *Gröbner bases*, Graduate Texts in Mathematics, vol. 141, Springer-Verlag, New York, 1993. A computational approach to commutative algebra, In cooperation with Heinz Kredel. MR1213453 (95c:13018)

[2] Wilfried Buchholz, Solomon Feferman, Wolfram Pohlers, and Wilfried Sieg, *Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies*, Lecture Notes in Mathematics, vol. 897, Springer-Verlag, Berlin, 1981. MR655036 (84d:03074)

[3] Wilfried Buchholz and Kurt Schütte, *Proof theory of impredicative subsystems of analysis*, Studies in Proof Theory, vol. 2, Bibliopolis, Naples, 1988. MR969597 (90e:03075)

[4] Peter Cholak, Alberto Marcone, and Reed Solomon, *Reverse mathematics and the equivalence of definitions for well and better quasi-orders*, J. Symbolic Logic 69 (2004), no. 3, 683–712. MR2078917 (2005e:03020)

[5] P. Clote, *The metamathematics of Fraïssé’s order type conjecture*, Recursion theory week (Oberwolfach, 1989), 1990, pp. 41–56. MR1071513 (91i:03109)

[6] D. H. J. de Jongh and Rohit Parikh, *Well-partial orderings and hierarchies*, Nederl. Akad. Wetensch. Proc. Ser. A 80 = Indag. Math. 39 (1977), no. 3, 195–207. MR0447056 (56 #5371)

[7] Jean-Yves Girard, *Proof theory and logical complexity*, Studies in Proof Theory. Monographs, vol. 1, Bibliopolis, Naples, 1987. MR903244 (89a:03113)

[8] Graham Higman, *Ordering by divisibility in abstract algebras*, Proc. London Math. Soc. (3) 2 (1952), 326–336. MR0049867 (14,238e)

[9] Jeffry L. Hirst, *Reverse mathematics and ordinal exponentiation*, Ann. Pure Appl. Logic 66 (1994), no. 1, 1–18. MR1263322 (95a:03078)

[10] Gerhard Jäger and Thomas Strahm, *Bar induction and ω model reflection*, Ann. Pure Appl. Logic 97 (1999), no. 1-3, 221–230. MR1682073 (2000f:03178)

[11] Joseph B. Kruskal, *The theory of well-quasi-ordering: A frequently discovered concept*, J. Combinatorial Theory Ser. A 13 (1972), 297–305. MR0306057 (46 #5184)

[12] Richard Laver, *On Fraïssé’s order type conjecture*, Ann. of Math. (2) 93 (1971), 89–111. MR0279005 (43 #4731)

[13] Alberto Marcone, *On the logical strength of Nash-Williams’ theorem on transfinite sequences*, Logic: from foundations to applications (Staffordshire, 1993), 1996, pp. 327–351. MR1428011 (98a:03092)

[14] Wolfram Pohlers, *Proof theory*, Lecture Notes in Mathematics, vol. 1407, Springer-Verlag, Berlin, 1989. An introduction. MR1026933 (91h:03078)

[15] , *Proof theory*, Universitext, Springer-Verlag, Berlin, 2009. The first step into impredicativity. MR2457679 (2010a:03002)

[16] Michael Rathjen and Andreas Weiermann, *Proof-theoretic investigations on Kruskal’s theorem*, Ann. Pure Appl. Logic 60 (1993), no. 1, 49–88. MR1212407 (94j:03120)
[17] Diana Schmidt, *Well-partial orderings and their maximal order types*, Habilitationsschrift, 1979.

[18] Kurt Schütte, *Proof theory*, Springer-Verlag, Berlin-New York, 1977. Translated from the revised German edition by J. N. Crossley, Grundlehren der Mathematischen Wissenschaften, Band 225. MR0505313 (58 #21497)

[19] Kurt Schütte and Stephen G. Simpson, *Ein in der reinen Zahlentheorie unbeweisbarer Satz über endliche Folgen von natürlichen Zahlen*, Arch. Math. Logik Grundlag. 25 (1985), no. 1-2, 75–89. MR822617 (88a:03147)

[20] Stephen G. Simpson, *Nonprovability of certain combinatorial properties of finite trees*, Harvey Friedman’s research on the foundations of mathematics, 1985, pp. 87–117. MR835255 (87i:03122a)

[21] , *Subsystems of second order arithmetic*, Second, Perspectives in Logic, Cambridge University Press, Cambridge, 2009. MR2517689 (2010e:03073)

[22] Jeroen Van der Meeren, Michael Rathjen, and Andreas Weiermann, *Well-partial-orderings and the big Veblen number*, Archive for Mathematical Logic (2014), 1–38 (English).

[23] Andreas Weiermann, *Proving termination for term rewriting systems*, Computer science logic (Berne, 1991), 1992, pp. 419–428. MR1232906 (94c:68107)

[24] , *Bounds for the closure ordinals of essentially monotonic increasing functions*, J. Symbolic Logic 58 (1993), no. 2, 664–671. MR1233931 (94f:03064)

[25] , *An order-theoretic characterization of the Schütte-Veblen-hierarchy*, Math. Logic Quart. 39 (1993), no. 3, 367–383. MR1270384 (95e:03159)

[26] , *A computation of the maximal order type of the term ordering on finite multisets*, Mathematical theory and computational practice, 2009, pp. 488–498. MR2545923 (2011h:06004)