Dynamic ordinal analysis

Arnold Beckmann
Institut für Mathematische Logik und Grundlagenforschung
Westfälische Wilhelms-Universität
Einsteinstr. 62, 48149 Münster, Germany*

June 1, 2002

Abstract

Dynamic ordinal analysis is ordinal analysis for weak arithmetics like fragments of bounded arithmetic. In this paper we will define dynamic ordinals – they will be sets of number theoretic functions measuring the amount of $s\Pi^b_1(X)$ order induction available in a theory. We will compare order induction to successor induction over weak theories. We will compute dynamic ordinals of the bounded arithmetic theories $s\Sigma^b_n(X)$-$\mathrm{L}^m\mathrm{IND}$ for $m = n$ and $m = n+1$, $n \geq 0$. Different dynamic ordinals lead to separation. In this way we will obtain several separation results between these relativized theories. We will generalize our results to further languages extending the language of bounded arithmetic.[1]

Keywords: Dynamic ordinal; Bounded arithmetic; Proof-theoretic ordinal; Order induction; Semi-formal system; Cut-elimination.

MSC: Primary 03F30; Secondary 03F05, 03F50.

Contents

*email: Arnold.Beckmann@math.uni-muenster.de

[1]The results for $s\Sigma^b_m(X)$-$\mathrm{L}^m\mathrm{IND}$ are part of the authors dissertation [3]; the results for $s\Sigma^b_m(X)$-$\mathrm{L}^{m+1}\mathrm{IND}$ base on results of Arai [1].

1 Introduction

Bounded arithmetic is designed to characterize low complexity computability, i.e. the polynomial hierarchy. Every primitive recursive function is provable total in $I\Sigma_1$, hence $I\Sigma_1$ is much stronger than bounded arithmetic. By PARIKH's Theorem ([14], or see [5, p.83, Theorem 11]) the provable total functions of $I\Delta_0$ (in the language $\mathcal{L}_{PA}$ of PEANO arithmetic) are bounded by polynomials. Hence $I\Delta_0(\mathcal{L}_{PA})$ is weaker than bounded arithmetic. Furthermore, only a constant number of elements $\leq n$ can be coded in a sequence $s = n^{O(1)}$. What do we mean by this? Assume a faithfull and feasible sequence coding and consider the function $s_n(l)$ obtained by forming the sequence consisting of l copies of n. $s_n(l)$ roughly is of the form

$$s_n(l) = \underbrace{\overbrace{\rule{2em}{0.4pt}}^{\log n \text{ bits}} \ldots \overbrace{\rule{2em}{0.4pt}}^{\log n \text{ bits}}}_{l \text{ times}}$$

I.e., $s_n(l)$ consists of $l \cdot \log n$ bits, hence it has size n^l (roughly). How fast is l, as a function of m, allowed to grow such that $I\Delta_0(\mathcal{L}_{PA})$ can prove the totality of $s_n(l(m))$ as a function of n, m? As remarked above $s_n(l(m))$ has to be bounded by a polynomial in m, n, hence $l(m)$ can only be constant. Thus, metamathematical arguments in general are not formalizable in $I\Delta_0(\mathcal{L}_{PA})$. We now look for a growth rate of $l(m)$ such that $s_n(l(m))$ has a growth rate suitable for bounded arithmetic.

Allowing $l(m) = m$ many elements would result in an exponential growth rate of $s_n(l(m))$, again too strong.

As argued in [14] the right growth rate is obtained by allowing $l(m) = \log m$ many elements $\leq n$ to be coded into one sequence. Then

$$s_n(l(m)) = n^{\log m} \approx 2^{|n|\cdot|m|} =: n\#m$$

where $|m|$ is the number of bits in the binary representation of m. Now bounded arithmetic can be formulated as $I\Delta_0$ in the language $\mathcal{L}_{BA}$ of bounded arithmetic, that is $\mathcal{L}_{PA}$ extended by $|.|, \#$, or, equivalently, as $I\Delta_0 + \Omega_1$ (where Ω_1 is equivalent to the statement $\forall x \exists y(|x|^2 = |y|)$), the latter being the original formulation of bounded arithmetic, see [20].

A stratification of bounded arithmetic, which corresponds to the stratification of the polynomial hierarchy, is obtained by putting restrictions on induction axioms; namely, allowing induction only for certain classes, Σ_i^b, of bounded formulas, and using length induction (LIND) in place of usual successor induction (IND). The most important subtheories of bounded arithmetic are the theories S_2^i, axiomatized by Σ_i^b-LIND, and the theories T_2^i, axiomatized by Σ_i^b-IND. The following is known for these theories:

$$S_2^1 \subseteq T_2^1 \preceq_{\forall\Sigma_2^b} S_2^2 \subseteq T_2^2 \preceq_{\forall\Sigma_3^b} S_2^3 \ldots$$

and their union is the theory $S_2 = T_2 = I\Delta_0(\mathcal{L}_{BA})$ (cf. [5, 6]). Here $T \preceq_{\forall\Sigma_i^b} T'$ means that T' is a $\forall\Sigma_i^b$-conservative extension of T. Furthermore, the class of predicates definable by Σ_i^b (or Π_i^b) formulas is precisely the class of predicates in the ith level Σ_i^p (resp. Π_i^p) of the polynomial hierarchy. In addition, the Σ_i^b-definable functions of S_2^i are precisely the $\square_i^p$-functions, which are the functions which are polynomial time computable with an oracle for Σ_{i-1}^p.

The main open problem for bounded arithmetic is the question if S_2 is finitely axiomatizable. As S_2^i and T_2^i are finitely axiomatizable, this question is equivalent to ask if there exists an i with $S_2^i = T_2^i$. This question is also connected to the open problem whether the polynomial hierarchy collapses, hence also with $P = ?NP$. The connection is that S_2 is finitely axiomatizable if and only if S_2 can prove that the polynomial hierarchy collapses. The precise connection is that $T_2^i = S_2^{i+1}$ if and only if the polynomial hierarchy collapses to its $i+2$nd level, provable in T_2^i (cf. [7, 22]).

The common conjecture is that the answer to all these questions is NO!

Bounded arithmetic still is in lack of a good measure of the proof theoretic strength of its fragments. By this we mean that proof theoretic strength should be assigned to fragments in such a way that different fragments are measured differently, which results in a separation of the fragments.

Good measures for a lot of theories are already known in the literature. An example is the proof theoretic ordinal of a theory. Inspired by GENTZEN's consistency proof for PA one defines

$$\begin{aligned}\mathrm{PO}(T(X)) \;:=\; \sup\{\alpha : \alpha &\text{ is the ordertype of a primitive recursive} \\ &\text{well-ordering} \prec \text{ and } T(X) \vdash \mathrm{Wf}(\prec)\}\end{aligned}$$

where $T(X)$ denotes the theory T in a language extended by adding a free set parameter X and the relation $\in$ (we will call $T(X)$ "the relativization of T"), and where $\mathrm{Wf}(\prec)$ expresses the well-foundedness of $\prec$ by the Π_1^1 sentence

$$(\forall X)\,\mathrm{Found}(\prec, X) \equiv (\forall X)\Big[(\forall x)((\forall y \prec x)(y \in X) \to x \in X) \to (\forall x)(x \in X)\Big].$$

GENTZEN's proof shows that $\mathrm{PO}(PA(X)) = \epsilon_0$. If one applies his methods to sub-theories of $PA(X)$ one obtains

$$\mathrm{PO}(\Delta_0(X)\text{-Ind}) = \omega^2 \quad \mathrm{PO}(\Sigma_1(X)\text{-Ind}) = \omega^\omega \quad \mathrm{PO}(\Sigma_2(X)\text{-Ind}) = \omega^{\omega^\omega} \quad \ldots$$

Now it is an obvious question whether the proof theoretical ordinal is a good measure for bounded arithmetic. The answer is **NO**, because R. SOMMER in [18] has shown that if $T(X)$ is a reasonable theory strictly weaker than $\Sigma_1(X)$-Ind, then $T(X)$'s proof theoretic ordinal is always ω^2. Applied to bounded arithmetic this shows

$$\mathrm{PO}(S_2^1(X)) = \omega^2 \qquad \mathrm{PO}(T_2^1(X)) = \omega^2 \qquad \ldots \qquad \mathrm{PO}(S_2(X)) = \omega^2$$

In this paper we follow another idea to measure proof theoretic strength of certain fragments of arithmetic. Proof theoretic ordinals are sets of static ordinals, "static" in the sense that ordinals are points. We will rather consider sets of ordinal enumerations, which we will call *dynamic ordinals*, "dynamic" in the sense that ordinal enumerations capture a certain dynamic. Ordinal enumerations will be compared by their growth rates. A first approach would look like (here "otyp" denotes "ordertype")

$$\mathrm{DO}(T(X)) \;:=\; \{\lambda n.\, \mathrm{otyp}(\prec_n) : (\prec_n)_n \text{ is a suitable sequence of well-founded orderings and } T(X) \vdash (\forall x)\, \mathrm{Found}(\prec_x, \Pi_1(X))\}$$

where $\mathrm{Found}(\prec, \Phi)$, for Φ is a set of formulas, denotes the schema of all formulas $\mathrm{Found}(\prec, A)$ for $A \in \Phi$, where $\mathrm{Found}(\prec, A)$ is the result of replacing X in $\mathrm{Found}(\prec, X)$ by the formula A.

Dynamic ordinals are arranged with respect to "eventual majorizability". But we immediately see that already very weak theories (which are even weaker than S^0_0) can always define an ordering $\prec$ whose ordertype is the supremum of those of a given sequence $\prec_n$, by defining $\langle n, a\rangle \prec \langle m, b\rangle$ iff $n = m$ and $a \prec_n b$. Also it is provable that the well-foundedness of all $\prec_n$, $n \in \omega$, implies the well-foundedness of $\prec$ (no induction is needed for this). Hence, concerning eventually majorizability, this definition of a dynamic ordinal always yields

$$\mathrm{DO}(T) \;\equiv\; \{\lambda n.\alpha \,:\, \alpha \in \mathrm{PO}(T)\},$$

so nothing is gained.

Another approach could be to fix a suitable well-ordering $\prec$ and then define

$$\mathrm{DO}(T(X)) \;:=\; \{\lambda n.\, \mathrm{otyp}(\prec\restriction f(n)) : f \text{ is a provable total function in } T(X) \text{ and } T(X) \vdash (\forall x)\, \mathrm{Found}(\prec\restriction f(x), \Pi_1(X))\}$$

where $\prec\restriction n$ denotes the restriction of $\prec$ to the domain $\{m : m \prec n\}$, i.e. $\prec\restriction n = \prec \cap \{m : m \prec n\}^2$. The behaviour of this definition of dynamic ordinals for stronger theories than bounded arithmetic is unstudied up to now.

In case of fragments of bounded arithmetic we also restrict $\prec$ to the natural ordering on numbers $<$ and f to terms in the language $\mathcal{L}_{BA}$. Observe that $\mathrm{otyp}(<\restriction n) = n$. Furthermore, the definition of dynamic ordinals now uses the principle $\mathrm{Found}(<\restriction t(x), A)$ which is equivalent to order induction $\mathcal{O}\mathrm{Ind}(t(x), A)$ given by

$$(\forall y \le t(x))\Big((\forall z < y)A(z) \to A(y)\Big) \to (\forall y \le t(x))A(y).$$

For $A \in s\Pi^b_1(X)$ this formula is equivalent to an $s\Sigma^b_2(X)$-formula ($s\Pi^b_1$, $s\Sigma^b_2$ etc. are prenex (or strict) reformulations of Π^b_1 resp. Σ^b_2 etc.).

To sum up, we define the dynamic ordinal of fragments $T(X)$ of relativized bounded arithmetic by

$$\mathrm{DO}(T(X)) \;:=\; \{\lambda x.t \,:\, T(X) \vdash (\forall x)\ \mathcal{O}\mathrm{Ind}(t, s\Pi^b_1(X))\}.$$

(Of course, with t we mean an $\mathcal{L}_{BA}$-term in which at most x occurs as a variable.) For very weak $\mathcal{L}_{BA}(X)$-theories $T(X)$, which have induction restricted to $\Sigma^b_0(X)$-formulas, the set of formulas in the definition of $\mathrm{DO}(T(X))$ has to be restricted to specific subsets of $s\Pi^b_1(X)$ formulas, see Definition 17 on page 14.

Dynamic ordinals are sets of number theoretic functions, i.e. subsets of ${}^\omega\omega$. As said before, we arrange subsets of ${}^\omega\omega$ by eventual majorizability:

$$f \trianglelefteq g \quad :\Leftrightarrow \quad g \text{ eventually majorizes } f \quad \Leftrightarrow \quad (\exists m)(\forall n \geq m) f(n) \leq g(n).$$

For subsets of number theoretic functions $D, E \subset {}^\omega\omega$ we define

$$D \trianglelefteq E \quad :\Leftrightarrow \quad (\forall f \in D)(\exists g \in E) f \trianglelefteq g$$

and from this

$$\begin{aligned} D \equiv E \quad &:\Leftrightarrow \quad D \trianglelefteq E \ \& \ E \trianglelefteq D \\ D \triangleleft E \quad &:\Leftrightarrow \quad D \trianglelefteq E \ \& \ E \not\trianglelefteq D \end{aligned}$$

We immediately see that $\trianglelefteq$ is a partial, transitive, reflexive ordering, $\triangleleft$ is a partial, transitive, irreflexive, not well-founded ordering, and $\equiv$ is an equivalence relation.

Examples

$$\begin{aligned} \{\lambda n.n^k : k \in \omega\} \quad &\equiv \quad \{\lambda n.p(n) : p \text{ a polynomial}\} \\ &\triangleleft \quad \{\lambda n.2^{|n|^k} : k \in \omega\} \quad \triangleleft \quad \{\lambda n.2^n\}. \end{aligned}$$

The next example shows that $\triangleleft$ is not well-founded.

$$\{\lambda n.|n|_{k+1}\} \quad \triangleleft \quad \{\lambda n.|n|_k\}$$

where $|n|_k$ is the k-fold iteration of $|\cdot|$ applied to n. The next example shows that the orderings are not total on ${}^\omega\omega$.

$$\lambda n.\begin{cases} 0 : n \text{ even} \\ 2 : n \text{ odd} \end{cases} \qquad \text{is incomparable to} \qquad (\lambda n.1)\,.$$

We will be able to assign different dynamic ordinals to certain fragments of bounded arithmetic, thus, by the next lemma, these fragments will be separated. Therefore, we can say:

> **Dynamic ordinals are good measures for proof theoretic strengths of fragments of bounded arithmetic.**

Lemma 1. *Let S, T be two theories in the language of bounded arithmetic and assume $\mathrm{DO}(S) \neq \mathrm{DO}(T)$. Then S is separated from T.*

Proof. Assume $f \in \mathrm{DO}(T) \setminus \mathrm{DO}(S)$. By the definition of dynamic ordinals there is a term $t(x)$ and an $s\Pi^b_1(X)$-formula A such that $f(n) = t(n)$ and $T \vdash (\forall x)\, \mathcal{O}\mathrm{Ind}(t(x), A)$. But $f \notin \mathrm{DO}(S)$ implies $S \nvdash (\forall x)\, \mathcal{O}\mathrm{Ind}(t(x), A)$. $\square$

We observe that the separation is given by a $\forall s\Sigma^b_2(X)$-sentence; in case that t is sharply bounded (i.e., $t \equiv |t'|$) this sentence even will be $\forall s\Sigma^b_1(X)$.

The main achievements of this paper will be the computation of the following dynamic ordinals:

$$\begin{array}{lcll} \mathrm{DO}(T^1_2(X)) & \equiv & \{\lambda n.2_2(c\cdot|n|_2) \ : c\in\omega\} & \equiv \ \mathrm{DO}(S^2_2(X)) \\ \mathrm{DO}(S^1_2(X)) & \equiv & \{\lambda n.2^{c\cdot|n|_2} \ : c\in\omega\} & \\ \mathrm{DO}(sR^2_2(X)) & \equiv & \{\lambda n.2_2(c\cdot|n|_3) \ : c\in\omega\} & \end{array}$$

where $sR^i_2(X)$ is $s\Sigma^b_i(X)$-L^2IND which is induction restricted to double-logarithmic induction on $s\Sigma^b_i(X)$-formulas, and more generally for $m \geq 0$

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-}\mathrm{L}^m\mathrm{Ind}) \quad \equiv \quad \{\lambda n.2_m(c\cdot|n|_{m+1}) \ : \ c\in\omega\}$$

where 2_m is the m-fold iteration of exponentiation. Furthermore, we will use results from Arai [1] to obtain

$$\begin{array}{lcll} \mathrm{DO}(T^0_2(X)) & \equiv & \{\lambda n.2^{c\cdot|n|_2} \ : c\in\omega\} & \equiv \ \mathrm{DO}(S^1_2(X)) \\ \mathrm{DO}(S^0_2(X)) & \equiv & \{\lambda n.c\cdot|n|_2 \ : c\in\omega\} & \\ \mathrm{DO}(sR^1_2(X)) & \equiv & \{\lambda n.2^{c\cdot|n|_3} \ : c\in\omega\} & \end{array}$$

and more generally for $m \geq 0$

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-}\mathrm{L}^{m+1}\mathrm{Ind}) \quad \equiv \quad \{\lambda n.2_m(c\cdot|n|_{m+2}) \ : \ c\in\omega\}$$

Thus by the previous Lemma and remarks these dynamic ordinals lead to relationships of bounded arithmetic theories which we display in Figure 1. Here we mean with $S < T$ that the theories S and T are separated and S is included in the consequences of T; with $S \equiv T$ that S and T have the same dynamic ordinals (this does not imply that S and T prove the same consequences); and with $S \not\subseteq T$ that S is not included in the consequences of T.

2 Bounded arithmetic

Let us recall some definitions. Fragments of bounded arithmetic are first order theories of arithmetic. The language of bounded arithmetic $\mathcal{L}_{BA}$ consists of function symbols 0 (zero), S (successor), + (addition), $\cdot$ (multiplication), $|x|$ (binary length), $\lfloor\frac{1}{2}x\rfloor$ (binary shift right), $x \# y$ (smash), $x \dot{-} y$ (arithmetical subtraction), $\mathrm{MSP}(x,i)$ (Most Significant Part) and $\mathrm{LSP}(x,i)$ (Less Significant Part), and relation symbols = (equality) and $\leq$ (less than or equal).

The meaning of MSP and LSP is given by

$$x = \mathrm{MSP}(x,i)\cdot 2^i + \mathrm{LSP}(x,i) \qquad \text{and} \qquad \mathrm{LSP}(x,i) < 2^i.$$

Restricted exponentiation $2^{\min(x,|y|)}$ can be defined by

$$2^{\min(x,|y|)} \quad = \quad \mathrm{MSP}(y \# 1, |y| \dot{-} x),$$

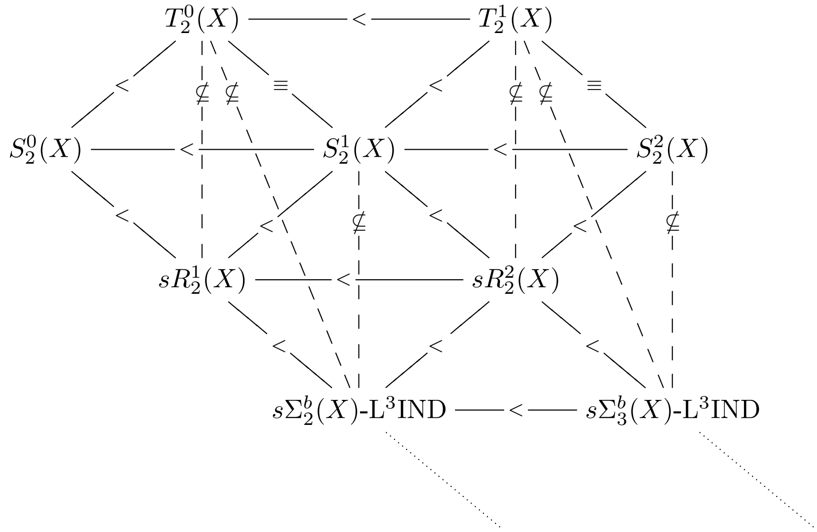

Figure 1: The separations

hence we can assume that restricted exponentiation is also part of our language $\mathcal{L}_{BA}$. We often write 2^t and mean $2^{\min(t,|x|)}$ if $t \leq |x|$ is clear from the context.

Relativized bounded arithmetic is formulated in the language $\mathcal{L}_{BA}(X)$ which is $\mathcal{L}_{BA}$ extended by one set variable X and the element relation $\in$.

$BASIC$ is a finite set of open axioms (cf. [5, 19, 10]) which axiomatizes the non-logical symbols. When dealing with $\mathcal{L}_{BA}(X)$ we assume that $BASIC$ also contains the equality axioms for X.

Bounded quantifiers play an important rôle in bounded arithmetic . We abbreviate

$$(\forall x \leq t)A := (\forall x)(x \leq t \to A) \qquad (\exists x \leq t)A := (\exists x)(x \leq t \wedge A)$$
$$(\forall x < t)A := (\forall x \leq t)(x < t \to A) \qquad (\exists x < t)A := (\exists x \leq t)(x < t \wedge A)$$

and call these *bounded quantifiers*. A quantifier of the form $(Q\,x \leq |t|)A$, $Q \in \{\forall, \exists\}$, is called a *sharply bounded quantifier*. A formula in which all quantifiers are (sharply) bounded is called a *(sharply) bounded formula*. Bounded formulas are stratified into levels:

Definition 2. *1. $\Delta_0^b = s\Sigma_0^b = s\Pi_0^b$ is the set of all sharply bounded formulas.*

2. $s\Sigma_n^b$-formulas are those which have a block of n alternating bounded quantifiers, starting with an existential one, in front of a sharply bounded kernel.

3. $s\Pi_n^b$ is defined dually, i.e. the block of alternating quantifiers starts with an universal one.

In the relativized case $\Delta_0^b(X)$, $s\Sigma_n^b(X)$, $s\Pi_n^b(X)$ are defined analogously.

Induction is also stratified. Let $|x|_0 := x$ and $|x|_{m+1} := |(|x|_m)|$.

Definition 3. Ψ-L^{m}IND *is the schema*

$$\varphi(0) \wedge (\forall x < |t|_m)(\varphi(x) \to \varphi(\mathrm{S}\, x)) \to \varphi(|t|_m)$$

for $\varphi \in \Psi$ *and terms* t*. For* $m = 0$ *this is the usual successor induction schema and will be denoted by* Ψ-IND*. In case* $m = 1$ *we usually write* Ψ-LIND*.*

The theories of bounded arithmetic are given by

$$BASIC + s\Sigma_n^b\text{-L}^m\text{IND}.$$

Usually we do not mention $BASIC$ and simply call this theory $s\Sigma_n^b$-L^{m}IND. Some of the theories have special names:[2]

Definition 4.

$$\begin{aligned} T_2^i &:= s\Sigma_i^b\text{-IND}, \\ S_2^i &:= s\Sigma_i^b\text{-LIND}, \\ sR_2^i &:= s\Sigma_i^b\text{-L}^2\text{IND}. \end{aligned}$$

For theories S, T let $S \subseteq T$ denote that all axioms in S are consequences of T. From the definition of the theories it immediately follows

$$\begin{aligned} s\Sigma_n^b\text{-L}^{m+1}\text{IND} &\subseteq s\Sigma_n^b\text{-L}^m\text{IND}, \\ s\Sigma_n^b\text{-L}^m\text{IND} &\subseteq s\Sigma_{n+1}^b\text{-L}^m\text{IND}. \end{aligned}$$

A little bit more insight is needed for

$$\begin{aligned} s\Sigma_n^b\text{-L}^m\text{IND} &\subseteq s\Sigma_{n+1}^b\text{-L}^{m+1}\text{IND}, \\ S_2^0 = sR_2^0 &= \Delta_0^b\text{-L}^{m+1}\text{IND}, \end{aligned}$$

see [5, 3] for a proof. Figure 2 reflects the just obtained relations. View the diagram as a graph whose nodes are theories. Take any edge in the graph. The way we displayed the graph allows us to destinguish a left and a right end of the edge. Now, the theory on the lefthand side of any edge in the graph is included in the theory on the righthand side. Similar definitions and results can be stated for theories of relativized bounded arithmetic. Furthermore, Krajíček, Pudlák and Takeuti have shown in [13] that $T_2^i(X) \neq S_2^{i+1}(X)$ using an oracle A such that the polynomial hierarchy in A collapses, in particular $P^A \neq NP^A$ (see [2, 21, 9] for a construction of A). Similarly, Krajíček has shown in [11] that $S_2^i(X) \neq T_2^i(X)$. A similar method is used by Pollett in [16] to obtain some refined separations for relativized theories $s\Sigma_n^b(X)$-L^{m}IND (independently from our work). In this article we will prove further refined separations which base on a completely different method called dynamic ordinal analysis. One half of these results are part of the author's dissertation [3], the other half bases on results due to Arai [1].

[2] The definition of T_2^i, resp. S_2^i, in [5, 12] differs from the one used here in that we define theories by induction only on strict formulas, but they all define same theories.

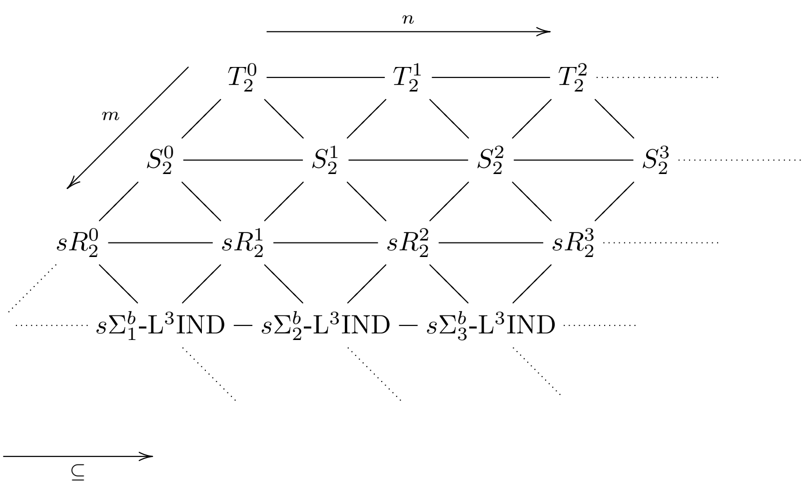

Figure 2: The theories $s\Sigma_n^b$-L^{m}IND

3 Lower bounds on dynamic ordinals

Theories of bounded arithmetic are axiomatized by using successor induction, where dynamic ordinals are based on order induction. In the following we will compare these two kinds of induction. Let us first fix some useful abbreviations. We adopt from set theory the convention of identifying numbers with the set of their predecessors, i.e. $y = \{z : z < y\}$. E.g., we write $y \subseteq X$ instead of $(\forall z < y)(z \in X)$.

$$\begin{aligned}
\mathcal{S}\mathrm{Prog}(x, X) &:= 0 \in X \wedge (\forall y < x)(y \in X \to \mathrm{S}\, y \in X) \\
\mathcal{S}\mathrm{Ind}(x, X) &:= \mathcal{S}\mathrm{Prog}(x, X) \to x \in X \\
\mathcal{O}\mathrm{Prog}(x, X) &:= (\forall y \le x)(y \subseteq X \to y \in X) \\
\mathcal{O}\mathrm{Ind}(x, X) &:= \mathcal{O}\mathrm{Prog}(x, X) \to \mathrm{S}\, x \subseteq X
\end{aligned}$$

Order induction, here called $\mathcal{O}$Ind, is logically equivalent to minimization:

$$(\exists y \le x)A(y) \to (\exists y \le x)(A(y) \wedge (\forall z < y)\neg A(z)).$$

It is well-known (cf. [5, 12]) that over the base theory $BASIC$ the schema $s\Sigma_i^b$-IND is equivalent to minimization for $s\Sigma_i^b$-formulas which is equivalent (by coding one existential quantifier) to minimization for $s\Pi_{i-1}^b$-formulas.

We first examine direct relations between $\mathcal{S}$Ind and $\mathcal{O}$Ind. We will often consider sets $\{y : A(y)\}$ for a formula $A(a)$, and we usually will abbreviate this set by A if the variable a is clear or unimportant. For a set of formulas Φ let $\mathcal{O}\mathrm{Ind}(t, \Phi)$ denote the schema given by all instances $\mathcal{O}\mathrm{Ind}(t, A)$ for $A \in \Phi$; similar for $\mathcal{S}$Ind. When saying "let T be an $\mathcal{L}_{BA}$-theory" we always mean that T includes some weak base theory, e.g. $S_2^0 \subseteq T$.

Lemma 5. *1. $BASIC \vdash \mathcal{O}\mathrm{Ind}(t,A) \rightarrow \mathcal{S}\mathrm{Ind}(t,A)$ for arbitrary formulas A.*

2. Let Φ be a set of $\mathcal{L}_{BA}$-formulas, which is closed under bounded (resp., sharply bounded) universal quantification, T be an $\mathcal{L}_{BA}$-theory, and t (resp., $t = |s|$) be an $\mathcal{L}_{BA}$-term. Then $T \vdash \mathcal{S}\mathrm{Ind}(t,\Phi)$ implies $T \vdash \mathcal{O}\mathrm{Ind}(t,\Phi)$.

Proof. 1. is obvious, because $\mathcal{S}\mathrm{Prog}(t,A)$ immediately implies $\mathcal{O}\mathrm{Prog}(t,A)$ in $BASIC$. Therefore, we obtain $t+1 \subseteq A$ assuming $\mathcal{O}\mathrm{Ind}(t,A)$ and $\mathcal{S}\mathrm{Prog}(t,A)$. In particular $A(t)$.

For 2. assume $T \vdash \mathcal{S}\mathrm{Ind}(t,\Phi)$ and $A \in \Phi$. We argue in T to show $\mathcal{O}\mathrm{Ind}(t,A)$. To this end suppose $\mathcal{O}\mathrm{Prog}(t,A)$, then we have to show $t+1 \subseteq A$. Define $B(y) := y \subseteq A$. Then $B \in \Phi$. We prove $B(t)$ using $\mathcal{S}\mathrm{Ind}(t,B)$. $B(0)$ is obvious. Now assume $y < t$ such that $B(y)$ holds. Then $y \subseteq A$, hence $A(y)$ using $\mathcal{O}\mathrm{Prog}(t,A)$, hence $B(y+1)$. By $\mathcal{S}\mathrm{Ind}(t,B)$, which is available by assumption, we obtain $B(t)$. Thus $t \subseteq A$. One more application of $\mathcal{O}\mathrm{Prog}(t,A)$ yields $A(t)$, hence $t+1 \subseteq A$. □

The tightest connection between $\mathcal{S}\mathrm{Ind}$ and $\mathcal{O}\mathrm{Ind}$ is obtained using the jump set $\mathrm{jp}(x,X)$ defined by:

$$\Big\{ y \leq |x| \ : \ (\forall z \leq 2^{|x|})[(z \in X \wedge z + 2^y \leq 2^{|x|} \rightarrow z + 2^y \in X) \wedge 0 \in X] \Big\}$$

Lemma 6. *Let Φ be a set of $\mathcal{L}_{BA}$-formulas, which is closed under substitution, and T be an $\mathcal{L}_{BA}$-theory. Then $T \vdash \mathcal{S}\mathrm{Ind}(2^{|x|},\Phi)$ if and only if $T \vdash \mathcal{O}\mathrm{Ind}(|x|,\mathrm{jp}(x,\Phi))$.*

Proof. Let $A \in \Phi$ and abbreviate $\mathrm{jp}(x,A)$ by Y. We argue in T.

For the direction from left to right assume $\mathcal{S}\mathrm{Ind}(2^{|x|},\Phi)$ and $\mathcal{O}\mathrm{Prog}(|x|,Y)$, then we have to show $|x|+1 \subseteq Y$. To this end let $y \leq |x|$ and $z \leq 2^{|x|}$ such that $A(z)$ and $z + 2^y \leq 2^{|x|}$. We have to prove $A(z+2^y)$ and $A(0)$. From $0 \subseteq Y$ we obtain $0 \in Y$ by $\mathcal{O}\mathrm{Prog}(|x|,Y)$, hence $A(0)$. Let $t(a)$ be $(z+2^y) \dot- (z+2^y \dot- a)$ and let $B(a)$ be $A(t(a))$, then $B \in \Phi$. We want to show $B(2^{|x|})$ using the assumption $\mathcal{S}\mathrm{Ind}(2^{|x|},\Phi)$, then this implies the assertion as $t(2^{|x|}) = z + 2^y$.

First we observe $t(a) = \min(a, z+2^y)$. We already have $A(0)$, hence $B(0)$. For the induction step let $a < 2^{|x|}$ such that $B(a)$. In case $a \geq z + 2^y$ we immediately have $B(a+1)$. Otherwise, $B(a)$ yields $A(a)$. As $a+2^0 = a+1 \leq 2^{|x|}$ we can use $0 \in Y$ to obtain $A(a+1)$, hence $B(a+1)$. Now $\mathcal{S}\mathrm{Ind}(2^{|x|},B)$ yields $B(2^{|x|})$ and we are done.

For the other direction from right to left we assume $\mathcal{O}\mathrm{Ind}(|x|,Y)$ and $\mathcal{S}\mathrm{Prog}(2^{|x|},A)$, then we have to show $A(2^{|x|})$. As an intermediate assertion we prove $\mathcal{O}\mathrm{Prog}(|x|,Y)$. So assume $b \leq |x|$ with $b \subseteq Y$, then we have to show $b \in Y$. Going into the definition of Y let $u \leq 2^{|x|}$ such that $A(u)$ and $u + 2^b \leq 2^{|x|}$. We have to conclude $A(u+2^b)$.

In a first case assume $b = 0$, then $u + 2^b = u+1 \leq 2^{|x|}$, hence $u < 2^{|x|}$. Thus the assertion $A(u+1)$ follows using $\mathcal{S}\mathrm{Prog}(2^{|x|},A)$.

In the other case $b = c+1$ for some c, hence $c < b \leq |x|$, hence $c \in Y$ by assumption. Using this and the definition of Y twice we first obtain $A(u+2^c)$

using the assumption $A(u)$, and from this $A((u+2^c)+2^c)$. This shows $A(u+2^b)$ which completes the proof of the intermediate assertion.

Using the assumption $\mathcal{O}\mathrm{Ind}(|x|, Y)$ we obtain $|x|+1 \subseteq Y$, hence $|x| \in Y$. As $A(0)$ we obtain $A(2^{|x|})$, hence the main assertion follows. $\square$

Lemma 7. *Let T be an $\mathcal{L}_{BA}$-theory and Φ a set of $\mathcal{L}_{BA}$-formulas closed under substitution. Suppose $T \vdash \mathcal{S}\mathrm{Ind}(t, \Phi)$. Then $T \vdash \mathcal{S}\mathrm{Ind}(p(t), \Phi)$ for all monotone polynomials p.*

Proof. The monotone polynomials in one variable x can be formally defined as an inductive set $\mathrm{MonPoly}(x)$ by the clauses: $0 \in \mathrm{MonPoly}(x)$, and if $p \in \mathrm{MonPoly}(x)$, then also $p+1 \in \mathrm{MonPoly}(x)$ and $p \cdot x \in \mathrm{MonPoly}(x)$. We can easily find for all monotone polynomials p in one variable x some $q \in \mathrm{MonPoly}(x)$ such that $BASIC \vdash p = q$.

Suppose $T \vdash \mathcal{S}\mathrm{Ind}(t, \Phi)$. We proof the assertion by induction on the generation of $p \in \mathrm{MonPoly}(t)$. The cases $p = 0$ and $p = q + 1$ are simple. We consider the case $p(t) = q \cdot t$. We argue in T. Let $A \in \Phi$ and assume $\mathcal{S}\mathrm{Prog}(q \cdot t, A)$, then we have to show $A(q \cdot t)$.

Define $C(z) :\equiv A(z \cdot t)$. Then $C \in \Phi$, hence we can apply the induction hypothesis obtaining $\mathcal{S}\mathrm{Ind}(q, C)$. Therefore, if we can show $\mathcal{S}\mathrm{Prog}(q, C)$, then this implies $C(q) \equiv A(q \cdot t)$ and we are done.

So we have to show $\mathcal{S}\mathrm{Prog}(q, C)$. $C(0)$ is obvious as $A(0)$ holds. Let $c < q$ such that $C(c)$, show $C(c+1)$. Define $D(u) :\equiv A(c \cdot t + u)$ then $D \in \Phi$. Hence we can apply the assumption and obtain $\mathcal{S}\mathrm{Ind}(t, D)$. In order to show $\mathcal{S}\mathrm{Prog}(t, D)$ first observe that $C(c)$ implies $D(0)$. Now let $v < t$ such that $D(v)$. This can be rewritten as $A(c \cdot t + v)$. An application of $\mathcal{S}\mathrm{Prog}(q \cdot t, A)$ yields $A(c \cdot t + v + 1)$ which is the same as $D(v + 1)$.

Hence we have shown $\mathcal{S}\mathrm{Prog}(t, D)$. Together with $\mathcal{S}\mathrm{Ind}(t, D)$ we obtain $D(t)$ which is the same as $A(c \cdot t + t)$, hence $C(c + 1)$ and we are done. $\square$

Lemma 8. *Let T be an $\mathcal{L}_{BA}$-theory and Φ a set of $\mathcal{L}_{BA}$-formulas closed under bounded universal quantification and substitution. Suppose $T \vdash \mathcal{O}\mathrm{Ind}(t, \Phi)$. Then $T \vdash \mathcal{O}\mathrm{Ind}(p(t), \Phi)$ for all monotone polynomials p.*

Proof. Suppose $T \vdash \mathcal{O}\mathrm{Ind}(t, \Phi)$. Then Lemma 5.1. shows $T \vdash \mathcal{S}\mathrm{Ind}(t, \Phi)$. Hence $T \vdash \mathcal{S}\mathrm{Ind}(p(t), \Phi)$ by Lemma 7. Hence $T \vdash \mathcal{O}\mathrm{Ind}(p(t), \Phi)$ using Lemma 5.2. $\square$

Theorem 9.
1. *$s\Sigma^b_n$-$\mathrm{L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(p(|x|_m), s\Pi^b_n)$ for monotone polynomials p, if $m > 0$ or $n > 0$.*
2. *$s\Sigma^b_0$-$\mathrm{L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(c \cdot |x|_{m+1}, \mathrm{jp}(x, \Delta^b_0))$ for natural numbers c.*

Proof. For 1. let T be $s\Sigma^b_n$-$\mathrm{L}^m\mathrm{IND}$. Using standard techniques (cf. [5]) we can see that T proves the schema $\mathcal{S}\mathrm{Ind}(|x|_m, s\Pi^b_n)$. If $n > 0$, then Lemma 5.2. shows $T \vdash \mathcal{O}\mathrm{Ind}(|x|_m, s\Pi^b_n)$. In case $n = 0$ we have $m > 0$ by assumption. An inspection of the proof of Lemma 5.2. shows that $B(y)$ can be written as $(\forall z \leq |x|_m)(z < y \to A(z))$. Therefore, for sharply bounded A we can choose

B also to be sharply bounded, hence the same proof as in Lemma 5.2. shows $T \vdash \mathcal{O}\mathrm{Ind}(|x|_m, s\Pi^b_0)$.

Thus, we always have $T \vdash \mathcal{O}\mathrm{Ind}(|x|_m, s\Pi^b_n)$, hence the assertion follows from Lemma 8.

In case 2. observe that $s\Sigma^b_0\text{-L}^m\mathrm{IND} \vdash \mathcal{S}\mathrm{Ind}(2^{|x|_{m+1}}, \Delta^b_0)$. Thus, we obtain $s\Sigma^b_0\text{-L}^m\mathrm{IND} \vdash \mathcal{S}\mathrm{Ind}(2^{c\cdot|x|_{m+1}}, \Delta^b_0)$ by applying Lemma 7. Hence the assertion follows from Lemma 6. □

For special theories these results can be rewritten as

$$\begin{array}{rlrl} T^0_2 \vdash & \mathcal{O}\mathrm{Ind}(|t|^c, \mathrm{jp}(x, \Sigma^b_0)) & T^{i+1}_2 \vdash & \mathcal{O}\mathrm{Ind}(2^{|t|^c}, s\Pi^b_{i+1}) \\ S^0_2 \vdash & \mathcal{O}\mathrm{Ind}(c \cdot ||t||, \mathrm{jp}(x, \Sigma^b_0)) & S^{i+1}_2 \vdash & \mathcal{O}\mathrm{Ind}(|t|^c, s\Pi^b_{i+1}) \\ & & sR^{i+1}_2 \vdash & \mathcal{O}\mathrm{Ind}(||t||^c, s\Pi^b_{i+1}) \end{array}$$

where c is a positive integer.

Now we connect order induction for different complexities of formulas using a (big) jump set $\mathrm{Jp}(t, x, X)$ defined by:

$$\Big\{ y \leq t \ : \ t \leq |x| \wedge (\forall z \leq 2^t)[z \subseteq X \wedge z + 2^y \leq 2^t + 1 \rightarrow z + 2^y \subseteq X] \Big\}$$

In the following four statements let $A(a)$ be an arbitrary $\mathcal{L}_{BA}$-formula.

Proposition 10. $BASIC \vdash t \leq |x| \wedge \mathcal{O}\mathrm{Ind}(2^t, A) \rightarrow \mathcal{O}\mathrm{Ind}(t, \mathrm{Jp}(t, x, A))$.

Proof. We argue in $BASIC$.

Suppose $t \leq |x|$, $\mathcal{O}\mathrm{Ind}(2^t, A)$ and $\mathcal{O}\mathrm{Prog}(t, \mathrm{Jp}(t, x, A))$. We have to show $t + 1 \subseteq \mathrm{Jp}(t, x, A))$. To this end we show $\mathcal{O}\mathrm{Prog}(2^t, A)$, because this implies $2^t + 1 \subseteq A$. But then the assertion follows because if $u \leq t$ and z given such that $z + 2^u \leq 2^t + 1$ then $z + 2^u \subseteq 2^t + 1 \subseteq A$, hence $u \in \mathrm{Jp}(t, x, A))$.

We are left to show $\mathcal{O}\mathrm{Prog}(2^t, A)$. Suppose $z \leq 2^t$ with $z \subseteq A$. By assumption $\mathcal{O}\mathrm{Prog}(t, \mathrm{Jp}(t, x, A))$ we have $0 \in \mathrm{Jp}(t, x, A)))$. Thus we obtain $z + 1 \subseteq A$. In particular $A(z)$. □

Lemma 11. $BASIC \vdash t \leq |x| \wedge \mathcal{O}\mathrm{Prog}(2^t, A) \rightarrow \mathcal{O}\mathrm{Prog}(t, \mathrm{Jp}(t, x, A))$.

Proof. We argue in $BASIC$. Assume $t \leq |x|$, $\mathcal{O}\mathrm{Prog}(2^t, A)$ and $u \leq t$ such that $u \subseteq \mathrm{Jp}(t, x, A)$. We have to show $u \in \mathrm{Jp}(t, x, A)$. To this end let $z \leq 2^t$ such that $z \subseteq A$ and $z + 2^u \leq 2^t + 1$. We are left to prove $z + 2^u \subseteq A$.

In a first case assume $u = 0$, then we derive $A(z)$ from the assumptions $z \subseteq A$ and $\mathcal{O}\mathrm{Prog}(2^t, A)$, hence $z + 1 \subseteq A$. Otherwise $u = v + 1$ for some v, hence $v \in \mathrm{Jp}(t, x, A)$ by assumption. Using this twice we obtain from $z \subseteq A$ first $z + 2^v \subseteq A$ and then $(z + 2^v) + 2^v \subseteq A$. But $2^v + 2^v = 2^u$, hence we are done. □

Proposition 12. $BASIC \vdash t \leq |x| \wedge \mathcal{O}\mathrm{Ind}(t, \mathrm{Jp}(t, x, A)) \rightarrow \mathcal{O}\mathrm{Ind}(2^t, A)$.

Proof. We argue in $BASIC$.

Suppose $t \leq |x|$, $\mathcal{O}\mathrm{Ind}(t, \mathrm{Jp}(t,x,A))$ and $\mathcal{O}\mathrm{Prog}(2^t, A)$. By the previous Lemma 11 we derive $\mathcal{O}\mathrm{Prog}(t, \mathrm{Jp}(t,x,A))$, hence $t+1 \subseteq \mathrm{Jp}(t,x,A)$. In particular $t \in \mathrm{Jp}(t,x,A)$. Reading the definition of Jp we obtain from $0 \subseteq A$ already $2^t \subseteq A$. Thus $A(2^t)$ using $\mathcal{O}\mathrm{Prog}(2^t, A)$ again. Altogether we have $2^t + 1 \subseteq A$. □

Similar proofs like the ones of the last three statements show

Proposition 13. $BASIC \vdash \mathcal{S}\mathrm{Ind}(|t|, \mathrm{jp}(t,A)) \leftrightarrow \mathcal{S}\mathrm{Ind}(2^{|t|}, A)$. □

We will show in an example how this can be used to obtain stronger induction in theories. Let $\mathrm{BB}\,\Phi$ be the sharply bounded collection scheme for formulas in Φ. Let $i \geq 1$. By a result of Ressayre we know that the theories S^i_2 and $S^i_2 + \mathrm{BB}\,\Sigma^b_{i+1}$ have the same $\forall\Sigma^b_i$-consequences (see [12] for definitions and results). Let us fix some special collection scheme which is more suitable for our purpose: For a set of formulas Φ let $\overline{\mathrm{BB}}(s\Sigma^b_i, \Phi)$ be the set of all formulas

$$(\forall x \leq |t|)[(\exists y A) \vee B] \rightarrow (\exists w)(\forall x \leq |t|)[A_y((w)_x) \vee B]$$

with $\exists y A \in s\Sigma^b_i$ and $B \in \Phi$. We immediately have that

$$\overline{\mathrm{BB}}(s\Sigma^b_i, s\Pi^b_i) \subseteq \mathrm{BB}\, s\Sigma^b_{i+1}$$

and, by a standard argument using $s\Sigma^b_i$-LMax which is equivalent to $s\Sigma^b_i$-LIND,

$$S^i_2 \vdash \overline{\mathrm{BB}}(\Sigma^b_i, \Sigma^b_\infty).$$

Now $\mathrm{jp}(|x|, s\Pi^b_i) \subseteq s\Delta^b_{i+1}$ using $\overline{\mathrm{BB}}(s\Sigma^b_i, s\Pi^b_i)$. By a result of Pollett [16] we have $sR^i_2 \vdash \mathcal{S}\mathrm{Ind}(||x||, s\Delta^b_{i+1})$, hence

$$sR^i_2 + \overline{\mathrm{BB}}(s\Sigma^b_i, s\Pi^b_i) \vdash \mathcal{S}\mathrm{Ind}(||x||, \mathrm{jp}(|x|, s\Pi^b_i)).$$

Now the last Proposition shows $sR^i_2 + \overline{\mathrm{BB}}(s\Sigma^b_i, s\Pi^b_i) \vdash \mathcal{S}\mathrm{Ind}(2^{||x||}, s\Pi^b_i)$, hence

Corollary 14. $S^i_2 = sR^i_2 + \overline{\mathrm{BB}}(s\Sigma^b_i, s\Pi^b_i)$. □

Similar results hold for other fragments of bounded arithmetic.

We now consider iterations of Jp:

$$\begin{aligned} \mathrm{Jp}_0(t,x,X) &= X, \\ \mathrm{Jp}_{i+1}(t,x,X) &= \mathrm{Jp}(t, |x|_i, \mathrm{Jp}_i(t,x,X)). \end{aligned}$$

By iterated applications of Proposition 10 and Proposition 12 we obtain

Corollary 15.

$$BASIC \vdash t \leq |x|_m \rightarrow [\mathcal{O}\mathrm{Ind}(2_m(t), A) \leftrightarrow \mathcal{O}\mathrm{Ind}(t, \mathrm{Jp}_m(t,x,A))].$$

□

Concerning the complexity of the iterated jump we observe that

$$\mathrm{Jp}_i(t, x, s\Pi_1^b) \subset s\Pi_{i+1}^b$$

hence Theorem 9 and Corollary 15 together show

Theorem 16. *Let* $0 \leq n < m$ *or* $n = m = 1$*, and let* c *be some natural number, then* $s\Sigma_{n+1}^b$-$\mathrm{L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(2_n(|x|_m^c), s\Pi_1^b)$*, hence* $s\Sigma_{n+1}^b$-$\mathrm{L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(2_{n+1}(c \cdot |x|_{m+1}), s\Pi_1^b)$*.*

Proof. Theorem 9 shows $s\Sigma_{n+1}^b$-$\mathrm{L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(|x|_m^c, \mathrm{Jp}_n(|x|_m^c, s(x), s\Pi_1^b))$. To be able to apply Corollary 15 to this we need that $|x|_m^c \leq |s(x)|_n$. But this is granted by the assumptions $n < m$ or $n = m = 1$ (for a suitable choice of s). □

The results obtained so far all relativize to theories including X and $\in$.

Definition 17. *The dynamic ordinal of an* $\mathcal{L}_{BA}(X)$*-theory* $T(X)$ *is defined by*

$$\mathrm{DO}(T(X)) \quad := \quad \{\lambda x.t \,:\, T(X) \vdash (\forall x)\ \mathcal{O}\mathrm{Ind}(t, s\Pi_1^b(X))\}.$$

For very weak $\mathcal{L}_{BA}(X)$*-theories* $T(X)$*, which have their induction restricted to* $\Sigma_0^b(X)$*-formulas, the set of formulas in the definition of* $\mathrm{DO}(T(X))$ *has to be restricted to a specific subset of* $s\Pi_1^b(X)$ *given by* $\mathrm{jp}(t, \Delta_0^b(X))$*. Furthermore, we restrict terms to the form* $|t|$ *because dynamic ordinals of very weak theories are bounded by the dynamic ordinal of* $S_2^1(X)$ *which, as we will see, is eventually majorized by* $\{\lambda n.|n|^c \,:\, c \text{ a numeral}\}$*:*

$$\mathrm{DO}(T(X)) \quad := \quad \{\lambda x.|t| \,:\, T \vdash (\forall x)\ \mathcal{O}\mathrm{Ind}(|t|, \mathrm{jp}(t, \Delta_0^b(X)))\}.$$

(Of course, with t *we always mean an* $\mathcal{L}_{BA}$*-term* t *in which at most* x *occurs as a variable.)*

By definition we have

$$\mathrm{DO}(T(X)) \quad \trianglelefteq \quad \{\lambda n.2^{|n|^c} \,:\, c \text{ number}\} \quad \equiv \quad \{\lambda n.2^{2^{c\cdot||n||}} \,:\, c \text{ number}\}$$

and, for very weak theories T,

$$\mathrm{DO}(T(X)) \quad \trianglelefteq \quad \{\lambda n.|n|^c \,:\, c \text{ number}\} \quad \equiv \quad \{\lambda n.2^{c\cdot||n||} \,:\, c \text{ number}\}$$

where 'number' alsways means a positive integer. These are crude upper bounds simply given by growth rates of functions represented by $\mathcal{L}_{BA}$-terms. Combining them with the lower bounds resulting from Theorem 16, we can estimate the dynamic ordinals for several theories:

$$\begin{array}{lllll}
\mathrm{DO}(T_2^0(X)) & \equiv & \{\lambda n.|n|^c \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{c\cdot||n||} \,:\, c \text{ number}\} \\
\mathrm{DO}(T_2^1(X)) & \equiv & \{\lambda n.2^{|n|^c} \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{2^{c\cdot||n||}} \,:\, c \text{ number}\} \\
\mathrm{DO}(S_2^0(X)) & \trianglerighteq & & & \{\lambda n.(c \cdot ||n||) \,:\, c \text{ number}\} \\
\mathrm{DO}(S_2^1(X)) & \trianglerighteq & \{\lambda n.|n|^c \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{c\cdot||n||} \,:\, c \text{ number}\} \\
\mathrm{DO}(S_2^2(X)) & \equiv & \{\lambda n.2^{|n|^c} \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{2^{c\cdot||n||}} \,:\, c \text{ number}\} \\
\mathrm{DO}(sR_2^1(X)) & \trianglerighteq & \{\lambda n.||n||^c \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{c\cdot|n|_3} \,:\, c \text{ number}\} \\
\mathrm{DO}(sR_2^2(X)) & \trianglerighteq & \{\lambda n.2^{||n||^c} \,:\, c \text{ number}\} & \equiv & \{\lambda n.2^{2^{c\cdot|n|_3}} \,:\, c \text{ number}\}
\end{array}$$

In general

$$\begin{aligned} \mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND}) &\trianglerighteq \{\lambda n.2_m(c\cdot |n|_{m+2}) \;:\; c \text{ number}\} \\ \mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m}\mathrm{IND}) &\trianglerighteq \{\lambda n.2_m(c\cdot |n|_{m+1}) \;:\; c \text{ number}\} \end{aligned}$$

From the lower bound for the dynamic ordinal of a theory $T(X)$ we immediately read how much $s\Pi^b_1(X)$-induction (resp., $\Delta^b_0(X)$-induction for very weak $T(X)$) is available in $T(X)$. E.g., let $\lambda x.t \in \mathrm{DO}(T(X))$. Then $T(X) \vdash \mathcal{O}\mathrm{Ind}(t, s\Pi^b_1(X))$, hence $T(X) \vdash \mathcal{S}\mathrm{Ind}(t, s\Pi^b_1(X))$. For very weak $T(X)$ we have $t = |s|$ and $T(X) \vdash \mathcal{O}\mathrm{Ind}(t, \mathrm{jp}(2^t, \Delta^b_0(X)))$, hence $T(X) \vdash \mathcal{S}\mathrm{Ind}(2^t, \Delta^b_0(X))$.

4 Upper bounds

We will start this section computing upper bounds on dynamic ordinals for theories $s\Sigma^b_m(X)$-$\mathrm{L}^{m+1}\mathrm{IND}$ by applying results from Arai [1], section 2.4.

Assume $f \in \mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND})$. If $m > 0$ then there is by definition some $\mathcal{L}_{BA}$-term $t(x)$ such that $f = \lambda n.t(n)$ and $s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(t, s\Pi^b_1(X))$. Let $\tau := 2^{\min(t,|x|)} - 1$, then $|\tau| = \min(t, |x|) \le t$, hence $s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(|\tau|, \mathrm{jp}(\tau, \Delta^b_0(X)))$ as $\mathrm{jp}(\tau, \Delta^b_0(X)) \subseteq s\Pi^b_1(X)$. In case $m = 0$ there is by definition some $\mathcal{L}_{BA}$-term $t(x)$ such that $t = |t'|$, $f = \lambda n.t(n)$ and $s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(t, \mathrm{jp}(t', \Delta^b_0(X)))$. But then also $s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(|\tau|, \mathrm{jp}(\tau, \Delta^b_0(X)))$. In both cases Lemma 6 shows

$$s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND} \vdash \mathcal{S}\mathrm{Ind}(\tau + 1, X),$$

as $\mathrm{jp}(\tau, X) \in \mathrm{jp}(\tau, \Delta^b_0(X))$ and $2^{|\tau|} = \tau + 1$. By investigations from Arai [1], section 2.4, this implies that there is a polynomial p such that

$$\tau + 1 = 2^{\min(t(n),|n|)} \le 2_m(p(|n|_{m+1})).$$

Thus, there is a natural number c (the degree of p) such that $p(|n|_{m+1}) \trianglelefteq |n|^c_{m+1}$, hence

$$2^{\min(t(n),|n|)} \trianglelefteq 2_{m+1}(c\cdot |n|_{m+2}),$$

hence

$$\min(t(n), |n|) \trianglelefteq 2_m(c\cdot |n|_{m+2}) \triangleleft 2_m(|n|_{m+1}) \equiv |n|$$

(in writing $r(n) \trianglelefteq s(n)$ we always mean that $\lambda n.r(n) \trianglelefteq \lambda n.s(n)$). Hence $\lambda n. \min(t(n), |n|) \equiv f$. Altogether we have

$$f \trianglelefteq \lambda n.2_m(c\cdot |n|_{m+2}),$$

hence

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND}) \quad\equiv\quad \{\lambda n.2_m(c\cdot |n|_{m+2}) \;:\; c \text{ number}\}.$$

We now turn to the computation of upper bounds on dynamic ordinals for theories $s\Sigma^b_m(X)$-$\mathrm{L}^m\mathrm{IND}$.

4.1 Semi-formal systems for bounded arithmetic

We introduce semi-formal systems for bounded arithmetic á là SCHÜTTE [17, 15]. They are a kind of propositional proof systems which are adequate for investigating the provability strength of weak theories.

The language $\mathcal{L}_\omega$ of our semi-formal systems is a finitary version of the infinitary language $\mathcal{L}_\infty$ described in [15]. It is a TAIT-style language consisting of one set variable X, logical connectives $\bigvee, \bigwedge$, numerals $\underline{n}$ for $n \in \omega$, and binary relation symbols $=, \neq, \leq, \not\leq, \in, \notin$. Atomic or prime formulas are $\underline{m} = \underline{n}$, $\underline{m} \neq \underline{n}$, $\underline{m} \leq \underline{n}$, $\underline{m} \not\leq \underline{n}$, $\underline{m} \in X$, $\underline{m} \notin X$ for $m, n \in \omega$, and arbitrary $\mathcal{L}_\omega$-formulas are obtained using the following rule: if A_i for $i < k$ are $\mathcal{L}_\omega$-formulas, then also $\bigwedge_{i<k} A_i$ and $\bigvee_{i<k} A_i$. Typical for a TAIT-language is that negation $\neg$ is not a symbol of the language, instead it is a defined syntactic operation mapping $\mathcal{L}_\omega$-formulas to $\mathcal{L}_\omega$-formulas, which bases on the DE MORGAN laws of negation. E.g., $\neg \underline{m} \notin X$ is the same as $\underline{m} \in X$, $\neg \bigwedge_{i<k} A_i$ the same as $\bigvee_{i<k} \neg A_i$ etc. For a given set $M \subseteq \omega$ we can evaluate an $\mathcal{L}_\omega$-formula F in the standard model $\mathbb{N}$ in the usual way by interpreting X thru M. We write $\mathbb{N} \vDash F_X[M]$ iff this evaluation results in true.

Before we can describe some calculus for $\mathcal{L}_\omega$ we need some measures of certain complexities of $\mathcal{L}_\omega$-formulas. The length or size of an $\mathcal{L}_\omega$-formula F, $\mathrm{lh}(F)$, is the number of atomic formula occurrences in F. We define for $\lambda < \omega$ the λ-rank $\lambda\text{-rk}(F)$ of F over small sub-formulas, i.e. sub-formulas of length bounded by λ, inductively as follows: If F atomic, let $\lambda\text{-rk}(F) := 0$. If $\mathrm{lh}(F) \leq \lambda$ and $\lambda > 1$, let $\lambda\text{-rk}(F) := 0$. Otherwise, F is of the form $\bigwedge_{i<k} A_i$ or $\bigvee_{i<k} A_i$. In this case let

$$\lambda\text{-rk}(F) := \max\{\lambda\text{-rk}(F_i) \ : \ i < k\} + 1.$$

We have definied 1-rk in such a way that it is the same as the usual rank. In particular, $1\text{-rk}(F) = 0$ if and only if F is atomic. Furthermore, $\lambda\text{-rk}(F) < \omega$ for arbitrary $\lambda < \omega$ and $\mathcal{L}_\omega$-formulas F.

Definition 18. *We inductively define the semi-formal system* $\vdash^{\delta}_{\rho,\lambda} \Delta$ *for natural numbers* $\delta, \rho, \lambda \in \omega$ *and a finite set* Δ *of* $\mathcal{L}_\omega$*-formulas by the following clauses:*

- $(Ax1)$ $\vdash^{\delta}_{\rho,\lambda} \Delta, \underline{m} = \underline{m}$ *holds.* $\vdash^{\delta}_{\rho,\lambda} \Delta, \underline{m} \neq \underline{n}$ *holds if* $m \neq n$.
- $(Ax2)$ $\vdash^{\delta}_{\rho,\lambda} \Delta, \underline{m} \notin X, \underline{m} \in X$ *holds.*
- $(\bigwedge)$ $\vdash^{\delta}_{\rho,\lambda} \Delta, \bigwedge_{i<k} F_i$ *holds if for all* $i < k$ *there is some* $\delta_i < \delta$ *such that* $\vdash^{\delta_i}_{\rho,\lambda} \Delta, F_i$.
- $(\bigvee)$ $\vdash^{\delta}_{\rho,\lambda} \Delta, \bigvee_{i<k} F_i$ *holds if there is some* $i_0 < k$ *and* $\delta_0 < \delta$ *such that* $\vdash^{\delta_0}_{\rho,\lambda} \Delta, F_{i_0}$.
- (Cut) $\vdash^{\delta}_{\rho,\lambda} \Delta$ *holds if there is some* $\delta_0 < \delta$ *and some* $\mathcal{L}_\omega$*-formula* F *such that* $\lambda\text{-rk}(F) < \rho$ *and* $\vdash^{\delta_0}_{\rho,\lambda} \Delta, F$ *and* $\vdash^{\delta_0}_{\rho,\lambda} \Delta, \neg F$.

A formula which is derived in an inference is called the main formula of that inference. The formula which disappears thru a (Cut) *is called a cut-formula.*

It is clear from the definition that if $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta$ then there is a derivation-tree $\mathcal{D}_\Delta$ such that each node of $\mathcal{D}_\Delta$ is labeled with a finite set of $\mathcal{L}_\omega$-formulas, the root is labeled with Δ, each leaf with an axiom $(Ax1)$ or $(Ax2)$, and the labels of each node together with those of its sons represent valid inferences which in case of a (Cut) fulfill that the λ-rank of the cut-formula is strictly bounded by ρ, and the depth of $\mathcal{D}_\Delta$ (i.e. the length of the longest path through $\mathcal{D}_\Delta$) is bounded by δ. Furthermore, if $\overset{\delta}{\underset{1,1}{\vdash}} \Delta$, then only atomic cut-formulas are allowed in the derivation, because $1\text{-rk}(F) = 0$ is equivalent to F being atomic. If $\overset{\delta}{\underset{0,\lambda}{\vdash}} \Delta$ then no cut at all occurs in the derivation. In this case we call the derivation cut-free, in the former nearly cut-free.

4.2 Cut-elimination and boundedness

The most important property of a semi-formal derivation $\mathcal{D}$ is that cuts can be eliminated. Beside the pure existence of a cut-free derivation $\mathcal{D}'$ for $\mathcal{D}$ we can control the depth of the derivation during the elimination process. The resulting depth depends on the complexity of cut-formulas that are eliminated. We will prove two cut-elimination theorems, one eliminating arbitrary cut-formulas, and the other eliminating small cut-formulas.

The second important property of cut-free derivations is the boundedness theorem which yields lower bounds on the depth of the derivation. It states that a nearly cut-free derivation of the statement '$<$ restricted to $\{0, \ldots, n-1\}$ is well-founded', i.e. the 'translation' of $\mathcal{O}\mathrm{Ind}(n, X)$ to $\mathcal{L}_\omega$, must have depth at least n.

Let us first state some direct properties, which are easily proven by induction on the depth δ of the derivation.

Proposition 19 (Structural Rule). *Assume $\Delta \subseteq \Delta'$, $\delta \leq \delta'$, $\rho \leq \rho'$, $\lambda \leq \lambda'$ and $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta$, then $\overset{\delta'}{\underset{\rho',\lambda'}{\vdash}} \Delta'$.* □

Proposition 20 (($\bigwedge$)-Inversion). *Suppose $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta, \bigwedge_{i<k} F_i$, then $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta, F_i$ for all $i < k$.* □

Proposition 21 (($\bigvee$)-Exportation). *Suppose $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta, \bigvee_{i<k} F_i$, then $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta, F_0, \ldots, F_{k-1}$* □

Lemma 22 (Elimination Lemma). *Assume F is of the form $\bigwedge_{i<k} F_i$ and $\lambda\text{-rk}(F) \leq \rho$ and $\rho > 0$. If $\overset{\gamma}{\underset{\rho,\lambda}{\vdash}} \Gamma, F$ and $\overset{\delta}{\underset{\rho,\lambda}{\vdash}} \Delta, \neg F$, then $\overset{\gamma+\delta}{\underset{\rho,\lambda}{\vdash}} \Gamma, \Delta$.*

Proof. We prove the assertion by induction on δ. The interesting case is that $\neg F \equiv \bigvee_{i<k} \neg F_i$ is the main formula of the last inference and $\lambda\text{-rk}(\neg F) = \rho$. By eventually applying a Structural Rule we can assume w.l.o.g. that the premise of the last inference is of the form

$$\overset{\delta'}{\underset{\rho,\lambda}{\vdash}} \Delta, \neg F, \neg F_{i_0}$$

for some $\delta' < \delta$ and some $i_0 < k$. Now the induction hypothesis yields

$$\vdash^{\gamma+\delta'}_{\rho,\lambda} \Gamma, \Delta, \neg F_{i_0}$$

Applying $(\bigwedge)$-Inversion to the first derivation yields $\vdash^{\gamma}_{\rho,\lambda} \Gamma, F_{i_0}$, hence

$$\vdash^{\gamma+\delta'}_{\rho,\lambda} \Gamma, \Delta, F_{i_0}$$

by a Structural Rule. Now an application of (Cut) yields the assertion as $\lambda\text{-rk}(F_{i_0}) < \lambda\text{-rk}(F) \leq \rho$ and $\gamma + \delta' < \gamma + \delta$. ☐

Theorem 23 (Elimination Theorem). *If* $\vdash^{\delta}_{\rho+1,\lambda} \Delta$ *and* $\rho > 0$*, then* $\vdash^{2^\delta}_{\rho,\lambda} \Delta$.

Proof. By induction on δ we replace cuts by applications of the Elimination Lemma. ☐

The Elimination Theorem shows that reducing the cut-rank results in an exponential blow-up of the depth of the derivation. The advantage of this reduction is that it is independent from the size of the eliminated cut-formulas. But if the sizes are small another type of cut-elimination produces better, i.e. smaller, growth rates. $\mathfrak{L}_\omega$-Cut-Elimination has been invented in [3]. Here, we present a technically more elegant version.

Lemma 24 ($\mathfrak{L}_\omega$-Cut-Elimination Lemma). *Assume* F *is an* $\mathfrak{L}_\omega$*-formula and* $\rho, \lambda > 0$*. If* $\vdash^{\delta}_{\rho,\lambda} \Delta, F$ *and* $\vdash^{\delta}_{\rho,\lambda} \Delta, \neg F$*, then* $\vdash^{\delta+\text{lh}(F)}_{\rho,\lambda} \Delta$.

Proof. We prove the assertion by induction on the definition of F as an $\mathfrak{L}_\omega$-formula. If F is atomic, we obtain the assertion by a (Cut), because $\text{lh}(F) = 1$ and $\lambda\text{-rk}(F) = 0$.

Otherwise assume F is not atomic. W.l.o.g. we can assume that F is of the form $\bigvee_{i<k} F_i$. $(\bigvee)$-Exportation yields

$$\vdash^{\delta}_{\rho,\lambda} \Delta, F_0, \ldots, F_{k-1}. \tag{1}$$

Using $(\bigwedge)$-Inversion for the second derivation we obtain

$$\vdash^{\delta}_{\rho,\lambda} \Delta, \neg F_0$$

hence the induction hypothesis produces

$$\vdash^{\delta+\text{lh}(F_0)}_{\rho,\lambda} \Delta, F_1, \ldots, F_{k-1}.$$

Next we use $(\bigwedge)$-Inversion and the induction hypothesis to produce

$$\vdash^{\delta+\text{lh}(F_0)+\text{lh}(F_1)}_{\rho,\lambda} \Delta, F_2, \ldots, F_{k-1}.$$

Inductively we obtain

$$\vdash^{\delta+\sum_{i<k}\text{lh}(F_i)}_{\rho,\lambda} \Delta$$

which is our assertion because $\sum_{i<k} \text{lh}(F_i) = \text{lh}(F)$. ☐

Theorem 25 ($\mathcal{L}_\omega$-Cut-Elimination Theorem). *If* $\vdash^{\delta}_{1,\lambda} \Delta$*, then* $\vdash^{\delta\cdot\lambda}_{1,1} \Delta$.

Proof. By induction on δ we replace cuts by applications of the $\mathcal{L}_\omega$-Cut-Elimination Lemma. □

Corollary 26 (Cut-Elimination Corollary). *If* $\vdash^{\delta}_{\rho+1,\lambda} \Delta$*, then* $\vdash^{2_\rho(\delta)\cdot\lambda}_{1,1} \Delta$.

Proof. Applying the Elimination Theorem ρ times, and finally the $\mathcal{L}_\omega$-Cut-Elimination Theorem. □

We have shown how to obtain nearly cut-free derivations, i.e. derivations where only atomic cut-formulas are allowed. Nearly cut-free derivations enjoy a property called 'boundedness' which gives us lower bounds on the depth of the derivation: it states that the depths of derivations of certain well-foundedness principles cannot be small. The boundedness theorem needed here is a $\mathcal{L}_\omega$-version of the one proven in [4]. In order to make this paper self-contained, and because the version needed for $\mathcal{L}_\omega$ is much simpler, we repeat the arguments in an adapted form.

Let us consider the canonical translation of $\mathcal{O}$Prog and $\mathcal{O}$Ind to $\mathcal{L}_\omega$ which we will call again $\mathcal{O}$Prog resp. $\mathcal{O}$Ind:

$$\begin{aligned} \mathcal{O}\mathrm{Prog}(n,X) &:= \bigwedge_{i\le n}\Big(\bigwedge_{j<i}(\underline{j}\in X)\to \underline{i}\in X\Big) \\ \mathcal{O}\mathrm{Ind}(n,X) &:= \mathcal{O}\mathrm{Prog}(n,X)\to\bigwedge_{i\le n}(\underline{i}\in X). \end{aligned}$$

Of course $A \to B$ is an abbreviation of $\bigvee\{\neg A, B\}$. We will prove

$$\vdash^{\delta}_{1,1} \mathcal{O}\mathrm{Ind}(n,X) \quad\Rightarrow\quad \delta\ge n.$$

The main tool for this will be the so called reachability operator $\mathcal{R}^\delta(M)$ which is defined as follows. For a (finite) set $M \subseteq \omega$ let $\overline{\mathrm{en}}_M$ be the enumeration function of $\omega \setminus M$. We define

$$\mathcal{R}^\delta(M) \quad:=\quad \{m\in\omega : m\le \overline{\mathrm{en}}_M(\delta)\}\cup M.$$

The reachability operator gives us the part of X accessed by a derivation using $\mathcal{O}\mathrm{Prog}(n,X)$. Let us state some properties which we will need. We start with some monotonicity properties.

$$\gamma\le\delta \quad\Rightarrow\quad \mathcal{R}^\gamma(M)\subseteq\mathcal{R}^\delta(M). \tag{2}$$

If $M \subseteq N$ then $\overline{\mathrm{en}}_M(\delta) \le \overline{\mathrm{en}}_N(\delta)$, hence

$$M\subseteq N \quad\Rightarrow\quad \mathcal{R}^\delta(M)\subseteq\mathcal{R}^\delta(N). \tag{3}$$

For $M \subset \omega$ and $m \in \omega$ we often write M, m instead of $M \cup \{m\}$. An important calculation is: $\overline{\mathrm{en}}_{M,m}(\delta) \le \overline{\mathrm{en}}_M(\delta+1)$, hence

$$\mathcal{R}^\delta(M,m)\subseteq\mathcal{R}^{\delta+1}(M)\cup\{m\}. \tag{4}$$

The most important property is

$$(\forall i < n) i \in \mathcal{R}^\delta(M) \quad \Rightarrow \quad n \in \mathcal{R}^{\delta+1}(M). \tag{5}$$

To see this assume $n \notin \mathcal{R}^{\delta+1}(M)$, hence $n \notin M$ and $\overline{\mathrm{en}}_M(\delta+1) < n$. Let $m := \overline{\mathrm{en}}_M(\delta+1)$. then $m \notin M$ as m is in the range of $\overline{\mathrm{en}}_M$ which is $\omega \setminus M$, and $\overline{\mathrm{en}}_M(\delta) < m$. Hence $m \notin \mathcal{R}^\delta(M)$ and $m < n$.

An $\mathcal{L}_\omega$-formula F is called X-positive if no $s \notin X$ occurs as a sub-formula in F. The essential property of X-positive formulas is that they are monotone in the following sense: If F is X-positive, $M \subseteq N \subseteq \omega$ and $\mathbb{N} \models F_X[M]$, then $\mathbb{N} \models F_X[N]$. This is easily proven by induction on the definition of F.

Lemma 27 (Boundedness Lemma). *Let Δ be a set of X-positive formulas, and $m_1, \ldots, m_k \in \omega$. Assume*

$$\left|\frac{\delta}{1,1}\right. \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{m}_1 \notin X, \ldots, \underline{m}_k \notin X, \Delta,$$

then $\mathbb{N} \models (\bigvee \Delta)_X[\mathcal{R}^\delta(m_1, \ldots, m_k)]$.

Before proving the Boundedness Lemma let us first draw the desired result.

Theorem 28 (Boundedness Theorem). *If $\left|\frac{\delta}{1,1}\right. \mathcal{O}\mathrm{Ind}(n, X)$, then $\delta \geq n$.*

Proof. From the assumption we obtain $\left|\frac{\delta}{1,1}\right. \neg\, \mathcal{O}\mathrm{Prog}(n, X), \bigwedge_{i \leq n}(\underline{i} \in X)$ by $(\bigvee)$-Exportation. Applying the Boundedness Lemma yields $i \in \mathcal{R}^\delta(\emptyset)$ for all $i \leq n$. In particular $n \in \mathcal{R}^\delta(\emptyset)$, hence $n \leq \overline{\mathrm{en}}_\emptyset(\delta) = \delta$. □

Proof of the Boundedness Lemma. We prove the assertion by induction on δ and destinguish cases according to the last inference in $\left|\frac{\delta}{1,1}\right. \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{\vec{m}} \notin X, \Delta$. If this is an axiom according to $(Ax1)$ then Δ already is an axiom of the same kind and the assertion is obvious. If this is an axiom according to $(Ax2)$ then $\underline{m}_j \in X$ occurs in Δ for some $j \in \{1, \ldots, k\}$. But now the assertion is obvious because $m \in \mathcal{R}^0[\vec{m}]$. If the main formula of the last inference belongs to Δ then the assertion follows from the induction hypothesis, the monotonicity of Δ and its sub-formulas, and the correctness of the inferences.

We now turn to the interesting cases. If the main formula of the last inference is $\neg\, \mathcal{O}\mathrm{Prog}(n, X)$, then we can find, using $(\bigwedge)$-Inversion, some $\gamma < \delta$ and some $m \in \omega$ such that

$$\left|\frac{\gamma}{1,1}\right. \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{\vec{m}} \notin X, \Delta, \bigwedge_{j<m} (\underline{j} \in X) \tag{6}$$

and

$$\left|\frac{\gamma}{1,1}\right. \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{\vec{m}} \notin X, \Delta, \underline{m} \notin X. \tag{7}$$

If there is some $j < m$ such that $j \notin \mathcal{R}^\gamma(\vec{m})$ then the induction hypothesis applied to (6) yields $\mathbb{N} \models \Delta_X[\mathcal{R}^\gamma(\vec{m})]$, and the assertion follows from the monoticity of Δ and (2). Otherwise we have $j \in \mathcal{R}^\gamma(\vec{m})$ for all $j < m$, hence

$$m \in \mathcal{R}^{\gamma+1}(\vec{m})$$

using (5), which together with (4) implies

$$\mathcal{R}^{\gamma}(\vec{m}, m) \subseteq \mathcal{R}^{\gamma+1}(\vec{m}) \cup \{m\} \subseteq \mathcal{R}^{\delta}(\vec{m}). \tag{8}$$

The induction hypothesis applied to (7) together with (8) entails

$$\mathbb{N} \vDash (\bigvee \Delta)_X[\mathcal{R}^{\delta}(\vec{m})]$$

by the monotonicity of Δ.

In case that the last inference is a (Cut) there are $\gamma < \delta$ and an atomic formula F such that

$$\vdash^{\gamma}_{1,1} \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{\vec{m}} \notin X, \Delta, F \tag{9}$$

and

$$\vdash^{\gamma}_{1,1} \neg\, \mathcal{O}\mathrm{Prog}(n, X), \underline{\vec{m}} \notin X, \Delta, \neg F. \tag{10}$$

Assume F is of the form $\underline{m} \in X$ for some $m \in \omega$. In the case $m \notin \mathcal{R}^{\delta}(\vec{m})$ the induction hypothesis applied to (9) yields $\mathbb{N} \vDash (\bigvee(\Delta, \underline{m} \in X))_X[\mathcal{R}^{\delta}(\vec{m})]$ using monotonicity, hence $\mathbb{N} \vDash (\bigvee \Delta)_X[\mathcal{R}^{\delta}(\vec{m})]$. Otherwise $m \in \mathcal{R}^{\delta}(\vec{m})$ and the induction hypothesis applied to (10) leads to $\mathbb{N} \vDash (\bigvee \Delta)_X[\mathcal{R}^{\gamma}(\vec{m}, m)]$. Using (4) we compute

$$\mathcal{R}^{\gamma}(\vec{m}, m) \subseteq \mathcal{R}^{\gamma+1}(\vec{m}) \cup \{m\} \subseteq \mathcal{R}^{\delta}(\vec{m}).$$

Hence $\mathbb{N} \vDash (\bigvee \Delta)_X[\mathcal{R}^{\delta}(\vec{m})]$ using monotonicity.

If F is of the form $\underline{m} \notin X$ for some m, then the situation is symmetrical to the previous case. In the remaining case F is an atomic sentence not containing X. The induction hypothesis applied to (9) and (10) combined with monotonicity yields $\mathbb{N} \vDash (\bigvee(\Delta, F))_X[\mathcal{R}^{\delta}(\vec{m})]$ and $\mathbb{N} \vDash (\bigvee(\Delta, \neg F))_X[\mathcal{R}^{\delta}(\vec{m})]$, hence $\mathbb{N} \vDash (\bigvee \Delta)_X[\mathcal{R}^{\delta}(\vec{m})]$. □

4.3 Embedding

We want to embed formal derivations into semi formal systems. Certain measures of the resulting derivations should be controlled. The main process during this embedding is to unfold the induction axioms of the formal derivation through cuts and cut-elimination on the semi-formal side. To obtain optimal bounds we need an intermediate semi-formal system which preprocesses induction. We add the following rule to our previously defined semi-formal system $\vdash^{\delta}_{\rho,\lambda} \Gamma$ obtaining $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Gamma$:

(IND_n) $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Gamma, \neg F_0, F_j$ holds if $0 < j \leq \kappa$ and there are $\delta' < \delta$ and $F_1, \ldots, F_{j-1}$ such that $\lambda\text{-rk}(F_i) \leq n$ for $i \leq j$ and $IND_n \vdash^{\delta',\kappa}_{\rho,\lambda} \Gamma, \neg F_i, F_{i+1}$ for $i < j$.

The extended semi-formal system enjoys similar basic properties as before:

Proposition 29 (Structural Rule). *Assume $\Delta \subseteq \Delta'$, $\delta \leq \delta'$, $\rho \leq \rho'$, $\lambda \leq \lambda'$, $\kappa \leq \kappa'$ and $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta$, then $IND_n \vdash^{\delta',\kappa'}_{\rho',\lambda'} \Delta'$.* □

Proposition 30 (($\bigwedge$)-Inversion). *Suppose $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta, \bigwedge_{i<k} F_i$, then for all $i < k$ we have $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta, F_i$.* □

Proposition 31 (($\bigvee$)-Exportation). *Suppose $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta, \bigvee_{i<k} F_i$, then we have $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta, F_0, \dots, F_{k-1}$.* □

Lemma 32 (Elimination Lemma). *Assume F is of the form $\bigwedge_{i<k} F_i$ and $\lambda\text{-rk}(F) \leq \rho$ and $\rho > 0$.*

If $IND_n \vdash^{\gamma,\kappa}_{\rho,\lambda} \Gamma, F$ and $IND_n \vdash^{\delta,\kappa}_{\rho,\lambda} \Delta, \neg F$, then $IND_n \vdash^{\gamma+\delta,\kappa}_{\rho,\lambda} \Gamma, \Delta$. □

Theorem 33 (Elimination Theorem). *If $IND_n \vdash^{\delta,\kappa}_{\rho+1,\lambda} \Delta$ and $\rho > 0$, then $IND_n \vdash^{2^\delta,\kappa}_{\rho,\lambda} \Delta$.* □

Corollary 34 (Cut-Elimination Corollary). *If $IND_n \vdash^{\delta,\kappa}_{n+\rho+1,\lambda} \Delta$, then $IND_n \vdash^{2_\rho(\delta),\kappa}_{n+1,\lambda} \Delta$.* □

The auxiliary semi-formal system can be embedded into the proper semi-formal system. The cost is an increase of the derivation depth by a factor logarithmically in the additional parameter κ.

Theorem 35. *If $\kappa > 0$ and $IND_n \vdash^{\delta,\kappa}_{n+1,\lambda} \Gamma$, then $\vdash^{\delta\cdot|\kappa|}_{n+1,\lambda} \Gamma$.*

Proof. We use induction on δ. If the last inference is not (IND_n) we obtain the assertion directly (from the induction hypothesis if $\delta > 0$) by the same inference, because the function $\delta \mapsto \delta \cdot |\kappa|$ is strictly monotone. Otherwise, there are some j, δ' with $0 < j \leq \kappa$ and $\delta' < \delta$, and some $F_0, \dots, F_j$ with $\lambda\text{-rk}(F_i) \leq n$ for $i \leq j$ and $\neg F_0, F_j \in \Gamma$ such that

$$\vdash^{\delta'\cdot|\kappa|}_{n+1,\lambda} \Gamma, \neg F_i, F_{i+1} \quad \text{for } i < j$$

using the induction hypothesis.

Now we proceed with cuts using the following strategy, which we picture for $j = 7$:

$$\dfrac{\dfrac{\dfrac{\neg F_0, F_1 \quad \neg F_1, F_2}{\neg F_0, F_2} \quad \dfrac{\neg F_2, F_3 \quad \neg F_3, F_4}{\neg F_2, F_4}}{\neg F_0, F_4} \qquad \dfrac{\dfrac{\neg F_4, F_5 \quad \neg F_5, F_6}{\neg F_4, F_6} \quad \neg F_6, F_7}{\neg F_4, F_7}}{\neg F_0, F_7}$$

Thus, we obtain $\vdash^{\delta'\cdot|\kappa|+|j|}_{n+1,\lambda} \Gamma, \neg F_0, F_j$, hence $\vdash^{\delta\cdot|\kappa|}_{n+1,\lambda} \Gamma$. □

We fix a formal proof system.

Definition 36. *Let $\mathcal{T}$ be $s\Sigma^b_n(X)\text{-L}^m\text{IND}$. We inductively define $\mathcal{T} \vdash \Delta$ for finite sets of $\mathcal{L}_{BA}(X)$-formulas Δ by the following clauses:*

(AxL) *$\mathcal{T} \vdash \Delta$ holds if Δ contains a logical axiom $\neg A, A$ for some atomic formula A.*

(AxE) $\mathcal{T} \vdash \Delta$ *holds if* Δ *contains an equality axiom of the form* $(s = s)$ *or* $(s = t \wedge A(s) \to A(t)$ *for some atomic formula* A *and terms* s, t.

(AxB) $\mathcal{T} \vdash \Delta$ *holds if* Δ *contains an instance of an axiom from* $BASIC$.

(IND) $\mathcal{T} \vdash \Delta$ *holds if* Δ *contains an instance of a formula from* $s\Sigma^b_n(X)$-$\mathrm{L}^m\mathrm{IND}$.

($\wedge$) $\mathcal{T} \vdash \Delta, F_0 \wedge F_1$ *holds if* $\mathcal{T} \vdash \Delta, F_i$ *for all* $i \in \{0,1\}$.

($\vee$) $\mathcal{T} \vdash \Delta, F_0 \vee F_1$ *holds if* $\mathcal{T} \vdash \Delta, F_i$ *for some* $i \in \{0,1\}$.

($\forall$) $\mathcal{T} \vdash \Delta, (\forall x)F$ *holds if there is some free variable* y *not occurring in* $\Delta, (\forall x)F$ *such that* $\mathcal{T} \vdash \Delta, F_x(y)$.

($\exists$) $\mathcal{T} \vdash \Delta, (\exists x)F$ *holds if there is some term* s *such that* $\mathcal{T} \vdash \Delta, F_x(s)$.

(Cut) $\mathcal{T} \vdash \Delta$ *holds if there is some formula* F *such that* $\mathcal{T} \vdash \Delta, F$ *and* $\mathcal{T} \vdash \Delta, \neg F$.

The introduced formal derivation systems are complete and allow partial cut-elimination, i.e., the cuts can be reduced to formulas of the complexity of the axioms which are bounded formulas. Furthermore, we obtain a normal form for derivations. Let $\mathcal{T}$ be a theory formulated in $\mathcal{L}_{BA}(X)$ and Δ a finite set of $\mathcal{L}_{BA}(X)$-formulas. Then we can show that $\mathcal{T} \vdash \Delta$ iff Δ is derivable in the restriction of the previously defined calculus, in which cut-formulas are restricted to bounded formulas and only ($\forall$)-inferences eliminate variables. This eliminated variable has to be the eigenvariable of the inference. We call such a restricted derivation a normal derivation.

This normal form is somehow part of the normal form which Buss et al. call 'a bounded proof which has no free cuts, is in free variable normal form and is restricted by parameter variables' (cf. [5, p. 77, Theorem 9]). In essential, the normal form defined here is that part of the latter normal form which is needed for the forthcoming.

We call an $\mathcal{L}_{BA}(X)$-formula F a sentence if at most the second order variable X occurs free in F. We define a translation * of bounded $\mathcal{L}_{BA}(X)$-sentences into $\mathcal{L}_\omega$.

1. $(Ps_1 \dots s_k)^* :\equiv (P\underline{s_1^{\mathbb{N}}} \dots \underline{s_k^{\mathbb{N}}})$ for predicate symbols P of $\mathcal{L}_{BA}(X)$. Here we think of '$\in X$' and '$\notin X$' as special predicate symbols. I.e., $(s \notin X)^* \equiv (\underline{s^{\mathbb{N}}} \notin X)$.
2. $(F_0 \wedge F_1)^* :\equiv \bigwedge_{i \leq 1} F_i^*$.
3. $(F_0 \vee F_1)^* :\equiv \bigvee_{i \leq 1} F_i^*$.
4. $\Big((\forall x \leq t)F(x)\Big)^* :\equiv \bigwedge_{i \leq\ t^{\mathbb{N}}} F(\underline{i})^*$.
5. $\Big((\exists x \leq t)F(x)\Big)^* :\equiv \bigvee_{i \leq\ t^{\mathbb{N}}} F(\underline{i})^*$.

Observe that the definition of $\mathcal{O}\mathrm{Prog}(n,X)$ and $\mathcal{O}\mathrm{Ind}(n,X)$ on p. 19 is simply the $*$-translation of the former $\mathcal{O}\mathrm{Prog}(\underline{n},X)$ resp. $\mathcal{O}\mathrm{Ind}(\underline{n},X)$ on p. 9.

The next definition needs a technical tool. To each $\mathcal{L}_{BA}$-term t we can associate a monotone $\mathcal{L}_{BA}$-term $\sigma[t]$ which is an upper bound to t, see Buss [5, p. 77] for a definition. To each $\Delta_0^b(X)$-formula F we associate a monotone $\mathcal{L}_{BA}$-term $\Delta(F)$ by recursion on the definition of F.

1. If F is atomic, let $\Delta(F)$ be 1. Observe that $|\Delta(F)| = 1$.
2. If F is of the form $G \wedge H$ or $G \vee H$ let $\Delta(F)$ be $2 \cdot (\Delta(G)+1) \cdot (\Delta(H)+1)$. Observe that $|\Delta(F)| \geq |\Delta(G)| + |\Delta(H)|$.
3. If F is of the form $(\forall x \leq |t|)G$ or $(\exists x \leq |t|)G$ let $\Delta(F)$ be
$$\Delta(G)_x(\sigma[|t|]) \# \sigma[2 \cdot t + 1].$$
Observe that $|\Delta(F)| \geq \sum_{n \leq |t|^{\mathbb{N}}} |\Delta(G)_x(\underline{n})|$.

We easily compute that for $\Delta_0^b(X)$-sentences F
$$\mathrm{lh}(F^*) \leq |\Delta(F)|.$$

We extend the definition of $\Delta(F)$ to $s\Sigma_n^b(X)$-formulas F: If F is of the form
$$(\exists x_1 \leq t_1)(\forall x_2 \leq t_2) \ldots (Qx_n \leq t_n)G$$
with $G \in \Delta_0^b(X)$ and $(Qx_n \leq t_n)G \notin \Delta_0^b(X)$, let $\Delta(F)$ be
$$\Delta(G)_{x_n}(\sigma[t_n])_{x_{n-1}}(\sigma[t_{n-1}]) \ldots_{x_1}(\sigma[t_1]).$$
We easily compute for $s\Sigma_n^b(X)$-sentences F
$$|\Delta(F)|\text{-rk}(F) \leq n.$$

We now come to the main theorem of this subsection, the Embedding Theorem. First some notation. If the free variables of a term t are included in $\{x_0, \ldots, x_p\}$ and $\mathrm{x_0}, \ldots, \mathrm{x_p} \in \omega$ (abbreviated by $\vec{\mathrm{x}} \in \omega$), then we define $t\langle\vec{\mathrm{x}}\rangle :\equiv t_{x_0,\ldots,x_p}(\underline{\mathrm{x_0}}, \ldots, \underline{\mathrm{x_p}})$. Analogously we define $F\langle\vec{\mathrm{x}}\rangle$ for formulas F and $\Gamma\langle\vec{\mathrm{x}}\rangle$ for sets of formulas Γ. To abbreviate we use $F, G\langle\vec{\mathrm{x}}\rangle$ instead of $\{F, G\}\langle\vec{\mathrm{x}}\rangle$ if this does not confuse.

In the following we will often identify a ground term t with its evaluation $t^{\mathbb{N}}$. It will be clear from the context what is meant.

Theorem 37 (Embedding). *Let Γ be a finite set of bounded $\mathcal{L}_{BA}$-formulas. Let the free variables of Γ occur under $\{x_0, \ldots, x_p\}$. Let $\mathcal{T}$ be $s\Sigma_n^b(X)$-$\mathrm{L}^m\mathrm{IND}$ and assume $\mathcal{T} \vdash \Gamma$. Then there are $\delta, \rho < \omega$ and some monotone $\mathcal{L}_{BA}$-term t whose free variables occur among $\{x_0, \ldots, x_p\}$ such that*
$$\forall \vec{\mathrm{x}} \in \omega \quad IND_n \,\Big|\!\frac{\delta, |t|_m\langle\vec{\mathrm{x}}\rangle^{\mathbb{N}}}{\rho, |t|\langle\vec{\mathrm{x}}\rangle^{\mathbb{N}}}\, \Gamma\langle\vec{\mathrm{x}}\rangle^*.$$

Proof. As remarked above we obtain a normal derivation $\mathcal{T} \vdash \Gamma$ in which all cut-formulas are $s\Sigma^b_\infty(X)$-formulas and all formulas in the derivation are bounded. We prove the assertion by induction along this derivation.

In the following we often omit the superscript $*$. Observe that for every bounded formula F containing no variable not in $\{x_0, \ldots, x_p\}$ we can find some $\delta < \omega$ with

$$\forall \vec{\mathrm{x}} \in \omega \quad \left|\frac{\delta}{0,0}\right. \neg F, F\langle\vec{\mathrm{x}}\rangle \tag{11}$$

We distinguish cases according to the last inference:

Cases $(AxL), (AxE), (AxB)$: If Γ is a logical axiom, an equality axiom or an instance of an axiom from $BASIC$ then $(Ax1)$, resp. $(Ax2)$ (and at most four $(\bigvee)$-inferences) yield $\left|\frac{4}{0,0}\right. \Gamma\langle\vec{\mathrm{x}}\rangle$ for any $\vec{\mathrm{x}} \in \omega$.

Case (IND): There is some $s\Sigma^b_n(X)$-formula F and some term t with $\mathcal{S}\mathrm{Ind}(|t|_m, F) \in \Gamma$. If x does not occur as a free variable in F then $F_x(0) \equiv F \equiv F_x(|t|_m)$, thus using (11) we obtain some $\delta < \omega$ with

$$\forall \vec{\mathrm{x}} \in \omega \quad \left|\frac{\delta}{0,0}\right. \neg F_x(0), F_x(|t|_m)\langle\vec{\mathrm{x}}\rangle$$

and three times $(\bigvee)$ yields the assertion.

Otherwise, let $s' :\equiv \sigma[\Delta(F)]_x(|t|_m) + t$ and $s :\equiv |s'|_m$, then the free variables of s' and s are among $\{x_0, \ldots, x_p\}$. Using (11) we obtain some δ such that for any $\vec{\mathrm{x}} \in \omega$ and $i < |t|_m\langle\vec{\mathrm{x}}\rangle^{\mathbb{N}}$

$$\left|\frac{\delta}{0,0}\right. \neg F_x(\underline{i})\langle\vec{\mathrm{x}}\rangle^*, F_x(\underline{i})\langle\vec{\mathrm{x}}\rangle^* \quad \text{and} \quad \left|\frac{\delta}{0,0}\right. \neg F_x(\mathrm{S}\,\underline{i})\langle\vec{\mathrm{x}}\rangle, F_x(\underline{i+1})\langle\vec{\mathrm{x}}\rangle,$$

hence

$$\left|\frac{\delta+2}{0,0}\right. (\exists x \leq |t|_m)(F \wedge \neg F_x(\mathrm{S}\,x))\langle\vec{\mathrm{x}}\rangle^*, \neg F_x(\underline{i})\langle\vec{\mathrm{x}}\rangle^*, F_x(\underline{i+1})\langle\vec{\mathrm{x}}\rangle^*$$

by $(\bigwedge)$ and $(\bigvee)$. Observe $t\langle\vec{\mathrm{x}}\rangle \leq s'\langle\vec{\mathrm{x}}\rangle$, hence $|t|_m\langle\vec{\mathrm{x}}\rangle \leq s\langle\vec{\mathrm{x}}\rangle$. For $i \leq |t|_m\langle\vec{\mathrm{x}}\rangle$ we have $\Delta(F_x(\underline{i}))\langle\vec{\mathrm{x}}\rangle \leq s'\langle\vec{\mathrm{x}}\rangle$, hence $|s'\langle\vec{\mathrm{x}}\rangle|\text{-rk}(F_x(\underline{i})\langle\vec{\mathrm{x}}\rangle) \leq n$. Therefore, we can apply (IND_n) to produce

$$IND_n \left|\frac{\delta+3, s\langle\vec{\mathrm{x}}\rangle}{0, |s'|\langle\vec{\mathrm{x}}\rangle}\right. \neg(\forall x \leq |t|_m)(F \rightarrow F_x(\mathrm{S}\,x))\langle\vec{\mathrm{x}}\rangle^*, \neg F_x(0)\langle\vec{\mathrm{x}}\rangle^*, F_x(|t|_m)\langle\vec{\mathrm{x}}\rangle^*.$$

Four times $(\bigvee)$ yields

$$\forall \vec{\mathrm{x}} \in \omega \quad IND_n \left|\frac{\delta+7, s\langle\vec{\mathrm{x}}\rangle}{0, |s'|\langle\vec{\mathrm{x}}\rangle}\right. \mathcal{S}\mathrm{Ind}(|t|_m, F)\langle\vec{\mathrm{x}}\rangle^*.$$

Case $(\vee)$: The assertion follows directly from the induction hypothesis.

Case $(\wedge)$: The assertion follows from the induction hypothesis after replacing the upper bounds by some common bounds (using Structural Rule). We may always take the sum of the inductively given terms.
Therefore, in the other cases we will assume common upper bounds.

Case (Cut): There is some bounded formula F such that the free variables of F are among $\{x_0, \dots, x_p\}$, $\mathcal{T} \vdash \Gamma, F$ and $\mathcal{T} \vdash \Gamma, \neg F$. Thus, the induction hypothesis yields some $\delta, \rho < \omega$, and some monotone $\mathcal{L}_{BA}$-term t with

$$IND_n \left|\frac{\delta,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma, F\langle\vec{\mathrm{x}}\rangle \quad \text{and} \quad IND_n \left|\frac{\delta,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma, \neg F\langle\vec{\mathrm{x}}\rangle$$

for all $\vec{\mathrm{x}} \in \omega$. Let ρ' be the maximum of ρ and $0\text{-rk}(F\langle\vec{0}\rangle) + 1$, then

$$|t|\langle\vec{\mathrm{x}}\rangle\text{-rk}(F\langle\vec{\mathrm{x}}\rangle) < \rho'$$

for all $\vec{\mathrm{x}} \in \omega$. Applying Structural Rules and a (Cut) in the semi-formal system yields $IND_n \left|\frac{\delta+1,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho',|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma$.

Case $(\exists)$: There are some term s, some variable x and some formula F such that (w.l.o.g. by eventually using a Structural Rule) $(\exists x)F \in \Gamma$ and $\mathcal{T} \vdash \Gamma, F_x(s)$. By assumption $(\exists x)F$ is bounded, thus there is some bounded formula B and some term u such that $(\exists x)F \equiv (\exists x \leq u)B$ and $F_x(s) \equiv s \leq u \wedge B_x(s)$. The induction hypothesis and $(\bigwedge)$-Inversion produce some $\delta, \rho < \omega$ with $\rho > 0$ and some t whose free variables occur under $\{x_0, \dots, x_p\}$ such that

$$IND_n \left|\frac{\delta,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma, s \leq u\langle\vec{\mathrm{x}}\rangle \tag{12}$$

and

$$IND_n \left|\frac{\delta,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma, B_x(s)\langle\vec{\mathrm{x}}\rangle \tag{13}$$

for all $\vec{\mathrm{x}} \in \omega$.

Fix $\vec{\mathrm{x}} \in \omega$. If $s\langle\vec{\mathrm{x}}\rangle \not\leq u\langle\vec{\mathrm{x}}\rangle$ then we obtain from $(Ax1)$ by a (Cut) with (12) $IND_n \left|\frac{\delta+1,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma\langle\vec{\mathrm{x}}\rangle$. Otherwise $s\langle\vec{\mathrm{x}}\rangle \leq u\langle\vec{\mathrm{x}}\rangle$. Hence $(\bigvee)$ applied to (13) yields

$$IND_n \left|\frac{\delta+1,|t|_m\langle\vec{\mathrm{x}}\rangle}{\rho,|t|\langle\vec{\mathrm{x}}\rangle}\right. \Gamma,$$

because

$$\bigvee_{n \leq u\langle\vec{\mathrm{x}}\rangle^{\mathbb{N}}} B_x(\underline{n})\langle\vec{\mathrm{x}}\rangle^* \equiv ((\exists x \leq u)B\langle\vec{\mathrm{x}}\rangle)^* \equiv ((\exists x)F\langle\vec{\mathrm{x}}\rangle)^* \in \Gamma^*.$$

Case $(\forall)$: There are some formula F and some variables x, y satisfying (w.l.o.g., by eventually using a Structural Rule) $(\forall x)F \in \Gamma$, y does not occur under the free variables of $\Gamma, (\forall x)F$ and $\mathcal{T} \vdash \Gamma, F_x(y)$. By assumption $(\forall x)F$ is a bounded formula. Hence there are some bounded formula G and some term u with $(\forall x)F \equiv (\forall x \leq u)G$ and $F_x(y) \equiv y \leq u \rightarrow G_x(y)$. The induction hypothesis and $(\bigvee)$-Exportation yield some $\delta, \rho < \omega$, $\rho > 0$ and some monotone term t whose free variables are among $\{x_0, \dots, x_p, y\}$, such that

$$IND_n \left|\frac{\delta,|t|_m\langle\vec{\mathrm{x}},\mathrm{y}\rangle}{\rho,|t|\langle\vec{\mathrm{x}},\mathrm{y}\rangle}\right. \Gamma, y \not\leq u, G_x(y)\langle\vec{\mathrm{x}}, \mathrm{y}\rangle \tag{14}$$

for all $\vec{\mathrm{x}}, \mathrm{y} \in \omega$. Fix $\vec{\mathrm{x}} \in \omega$.

Let $t' :\equiv \sigma[t]_y(u)$. The free variables of t' are among $\{x_0, \ldots, x_p\}$ and $t\langle \vec{\mathrm{x}}, \mathrm{y}\rangle \leq t'\langle \vec{\mathrm{x}}\rangle$ for $\mathrm{y} \leq u\langle \vec{\mathrm{x}}\rangle$.

Let $\mathrm{y} \leq u\langle \vec{\mathrm{x}}\rangle$. Then (14) leads to $IND_n \vdash^{\delta, |t'|_m\langle \vec{\mathrm{x}}\rangle}_{\rho, |t'|\langle \vec{\mathrm{x}}\rangle} \Gamma, G_\varphi(\underline{\mathrm{y}}), \underline{\mathrm{y}} \not\leq u\langle \vec{\mathrm{x}}\rangle$. By $(Ax1)$ and a (Cut) we obtain

$$IND_n \vdash^{\delta+1, |t'|_m\langle \vec{\mathrm{x}}\rangle}_{\rho, |t'|\langle \vec{\mathrm{x}}\rangle} \Gamma, G_\varphi(\underline{\mathrm{y}})\langle \vec{\mathrm{x}}\rangle.$$

Applying $(\bigwedge)$ produces $IND_n \vdash^{\delta+2, |t'|_m\langle \vec{\mathrm{x}}\rangle}_{\rho, |t'|\langle \vec{\mathrm{x}}\rangle} \Gamma\langle \vec{\mathrm{x}}\rangle$, because

$$\bigwedge_{\mathrm{y} \leq u\langle \vec{\mathrm{x}}\rangle^{\mathbb{N}}} G_x(\underline{\mathrm{y}})\langle \vec{\mathrm{x}}\rangle^* \equiv ((\forall x \leq u) G\langle \vec{\mathrm{x}}\rangle)^* \equiv ((\forall x) F\langle \vec{\mathrm{x}}\rangle)^* \in \Gamma\langle \vec{\mathrm{x}}\rangle^*.$$

□

4.4 Computing upper bounds

Assume $f \in \mathrm{DO}(s\Sigma^b_m(X)\text{-L}^m\mathrm{IND})$ and $m > 0$. By definition there is some $\mathcal{L}_{BA}$-term $t(x)$ such that $f = \lambda n.t(n)$ and $s\Sigma^b_m(X)\text{-L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(t, \Pi^b_1(X))$, hence

$$s\Sigma^b_m(X)\text{-L}^m\mathrm{IND} \vdash \mathcal{O}\mathrm{Ind}(t, X).$$

By the Embedding Theorem 37 there are some $\delta, \rho < \omega$ and some $\mathcal{L}_{BA}$-term $s(x)$ in which at most x occurs as a free variable such that

$$\forall \mathrm{x} \in \omega \quad IND_m \vdash^{\delta, |s|_m\langle \mathrm{x}\rangle^{\mathbb{N}}}_{\rho, |s|\langle \mathrm{x}\rangle^{\mathbb{N}}} \mathcal{O}\mathrm{Ind}(t\langle \mathrm{x}\rangle^{\mathbb{N}}, X).$$

Fix $n \in \omega$. The Cut-Elimination Corollary 34 for the auxiliary semi-formal system yields

$$IND_m \vdash^{2_\rho(\delta), |s(n)|_m}_{m+1, |s(n)|} \mathcal{O}\mathrm{Ind}(t(n), X).$$

Hence, by Theorem 35

$$\vdash^{2_\rho(\delta) \cdot |s(n)|_{m+1}}_{m+1, |s(n)|} \mathcal{O}\mathrm{Ind}(t(n), X).$$

Now the Cut-Elimination Corollary 26 shows

$$\vdash^{2_m(2_\rho(\delta) \cdot |s(n)|_{m+1}) \cdot |s(n)|}_{1,1} \mathcal{O}\mathrm{Ind}(t(n), X).$$

The Boundedness Theorem 28 applied to this yields

$$f(n) \quad = \quad t(n) \quad \leq \quad 2_m(2_\rho(\delta) \cdot |s(n)|_{m+1}) \cdot |s(n)|. \tag{15}$$

Let $c := 2_\rho(\delta)$. As s is an $\mathcal{L}_{BA}$-term it represents a function of polynomial growth rate, hence there is some $d < \omega$ such that $|s(n)| \leq |n|^d$ for n big enough, hence we compute for n big enough and some c' (remember $m > 0$)

$$\begin{aligned} f(n) &\leq 2_m(c' \cdot |n|_{m+1}) \cdot |n|^d \\ &\leq 2_m((c' + d) \cdot |n|_{m+1}). \end{aligned}$$

Altogether, we have

$$f \trianglelefteq \lambda n.2_m((c' + d) \cdot |n|_{m+1}),$$

hence

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^m\mathrm{IND}) \quad \equiv \quad \{\lambda n.2_m(c \cdot |n|_{m+1}) \;:\; c \text{ number}\}$$

for $m > 0$.

5 Proofs revisited

If we examine the proofs of our results we discover that they are nearly independent from the language $\mathcal{L}_{BA}$. We only needed that $+$ and $\cdot$ are included in $\mathcal{L}_{BA}$ for the lower bounds (e.g. Lemma 7). A language-dependent reformulation of our results reads as follows: Suppose the underlying language is $\mathcal{L}$ and $+$ and $\cdot$ are included in $\mathcal{L}$.

Assume $n + 1 > m \geq 0$ and $f \in \mathrm{DO}(s\Sigma^b_n(X)\text{-L}^m\mathrm{IND}(\mathcal{L}))$. Similar to the last section we obtain some $\mathcal{L}$-term t and some number c such that

$$f(x) \quad \leq \quad 2_n(c \cdot |t(x)|_{m+1}) \cdot |t(x)|,$$

c.f. (15). If $n > 0$ then we can compute

$$\begin{aligned} f(x) \quad &\leq \quad 2_n(c \cdot |t(x)|_{m+1} + |t(x)|_{n+1}) \\ &\leq \quad 2_n((c+1) \cdot |t(x)|_{m+1}) \end{aligned}$$

as $n \geq m$. Together with the results from the beginning of Section 4 and $\mathrm{DO}(T) \trianglelefteq \{\lambda x.|t(x)| \;:\; t \text{ an } \mathcal{L}\text{-term}\}$ for very weak $\mathcal{L}$-theories T, we obtain for $n + 1 \geq m \geq 0$

$$\begin{aligned} &\mathrm{DO}(s\Sigma^b_n(X)\text{-L}^m\mathrm{IND}(\mathcal{L})) \\ &\quad \trianglelefteq \quad \{\lambda x.2_n(c \cdot |t(x)|_{m+1}) \;:\; c \text{ a number and } t \text{ an } \mathcal{L}\text{-term}\}. \end{aligned}$$

On the other hand, 2. of Theorem 9 and Theorem 16 yield

$$\begin{aligned} &\mathrm{DO}(s\Sigma^b_n(X)\text{-L}^m\mathrm{IND}(\mathcal{L})) \\ &\quad \trianglerighteq \quad \{\lambda x.2_n(c \cdot |t(x)|_{m+1}) \;:\; c \text{ a number and } t \text{ an } \mathcal{L}\text{-term}\} \end{aligned}$$

for $0 \leq n \leq m$. Hence

$$\begin{aligned} &\mathrm{DO}(s\Sigma^b_n(X)\text{-L}^m\mathrm{IND}(\mathcal{L})) \\ &\quad \equiv \quad \{\lambda x.2_n(c \cdot |t(x)|_{m+1}) \;:\; c \text{ a number and } t \text{ an } \mathcal{L}\text{-term}\} \end{aligned}$$

for $m = n$ or $m = n + 1$ (and $n \geq 0$).

This has several consequences:

1. We can add arbitrary functions of polynomial growth rate to $\mathcal{L}_{BA}$, still obtaining the 'same' results – the only difference is that the theories are formulated in the new language $\mathcal{L}$. E.g., $S^1_2(\mathcal{L}; X) < T^1_2(\mathcal{L}, X)$, $S^1_2(\mathcal{L}; X) < sR^2_2(\mathcal{L}, X)$, and $T^1_2(\mathcal{L}; X) \not\subseteq sR^2_2(\mathcal{L}, X)$.

2. We can also consider languages including functions of stronger growth rates. E.g., for $k > 0$ let $\mathcal{L}_k$ be $\mathcal{L}_{BA}$ without $\#$ but extended by $\#_2, \dots, \#_k$, where $\#_1 = \cdot$ and $x \#_{k+1} y = 2^{|x| \#_k |y|}$. Hence $x \#_{k+1} y = 2_k(|x|_k \cdot |y|_k)$ and $\#_2 = \#$. Therefore, $\mathcal{L}_1$ is $\mathcal{L}_{BA}$ without $\#$, $\mathcal{L}_2$ is $\mathcal{L}_{BA}$, and $\mathcal{L}_3$ is $\mathcal{L}_{BA}$ extended by $\#_3$ etc. Observe that for an $\mathcal{L}_k$-term $t(x)$ in which at most x occurs as a free variable there exits a $d \in \omega$ such that $\lambda x.t(x) \trianglelefteq \lambda x.2_k(d \cdot |x|_k)$. The theories T^i_k, S^i_k, sR^i_k are T^i_2, S^i_2, resp. sR^i_2 reformulated in the language $\mathcal{L}_k$.

We compute for $m + 1 \le k$

$$\begin{aligned} \lambda x.2_n(c \cdot |t(x)|_{m+1}) &\trianglelefteq \lambda x.2_n(c \cdot 2_{k-(m+1)}(d \cdot |x|_k)) \\ &\trianglelefteq \lambda x.2_{k-1+n-m}(d' \cdot |x|_k) \end{aligned}$$

for certain $d, d' \in \omega$. And for $m + 1 > k$ we have

$$\begin{aligned} \lambda x.2_n(c \cdot |t(x)|_{m+1}) &\trianglelefteq \lambda x.2_n(c \cdot |d \cdot |x|_k|_{m+1-k}) \\ &\trianglelefteq \lambda x.2_n(c' \cdot |x|_{m+1}) \end{aligned}$$

for certain $d, c' \in \omega$.

Using these estimations we obtain

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND}(\mathcal{L}_k)) \equiv \{\lambda x.2_{k-2}(c \cdot |x|_k) : c \text{ number}\}$$

for $m + 2 \le k$, and

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m+1}\mathrm{IND}(\mathcal{L}_k)) \equiv \{\lambda x.2_m(c \cdot |x|_{m+2}) : c \text{ number}\}$$

for $m + 2 > k$. We also have

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m}\mathrm{IND}(\mathcal{L}_k)) \equiv \{\lambda x.2_{k-1}(c \cdot |x|_k) : c \text{ number}\}$$

for $m + 1 \le k$, and

$$\mathrm{DO}(s\Sigma^b_m(X)\text{-L}^{m}\mathrm{IND}(\mathcal{L}_k)) \equiv \{\lambda x.2_m(c \cdot |x|_{m+1}) : c \text{ number}\}$$

for $m + 1 > k$.

Furthermore, we obtain

$$\mathrm{DO}(T^1_k(X)) \trianglerighteq \{\lambda x.2_k(c \cdot |x|_k) : c \text{ number}\}$$

and, for $m < k$,

$$\begin{aligned} &\mathrm{DO}(s\Sigma^b_{m+1}(X)\text{-L}^{m}\mathrm{IND}(\mathcal{L}_k)) \\ &\quad \trianglelefteq \{\lambda x.2_{m+1}(c \cdot |t(x)|_{m+1}) : c \text{ a number and } t \text{ an } \mathcal{L}_k\text{-term}\} \\ &\quad \equiv \{\lambda x.2_k(c \cdot |x|_k) : c \text{ number}\}. \end{aligned}$$

Hence

$$\begin{array}{ccccccc} \mathrm{DO}(S^0_k(X)) & \equiv & \mathrm{DO}(sR^1_k(X)) & \equiv & \dots & \equiv & \mathrm{DO}(s\Sigma^b_{k-2}(X)\text{-L}^{k-1}\mathrm{IND}(\mathcal{L}_k)) \\ \mathrm{DO}(T^0_k(X)) & \equiv & \mathrm{DO}(S^1_k(X)) & \equiv & \dots & \equiv & \mathrm{DO}(s\Sigma^b_{k-1}(X)\text{-L}^{k-1}\mathrm{IND}(\mathcal{L}_k)) \\ \mathrm{DO}(T^1_k(X)) & \equiv & \mathrm{DO}(S^2_k(X)) & \equiv & \dots & \equiv & \mathrm{DO}(s\Sigma^b_{k}(X)\text{-L}^{k-1}\mathrm{IND}(\mathcal{L}_k)). \end{array}$$

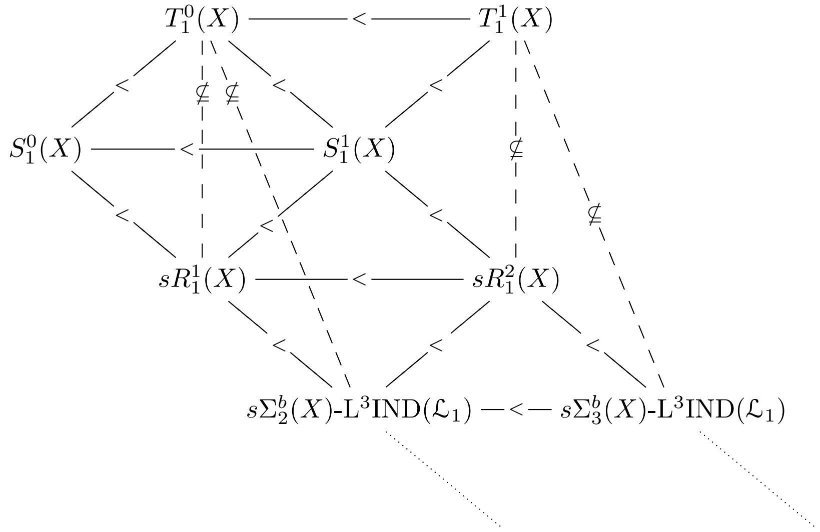

Figure 3: The $\mathfrak{L}_1$-theories

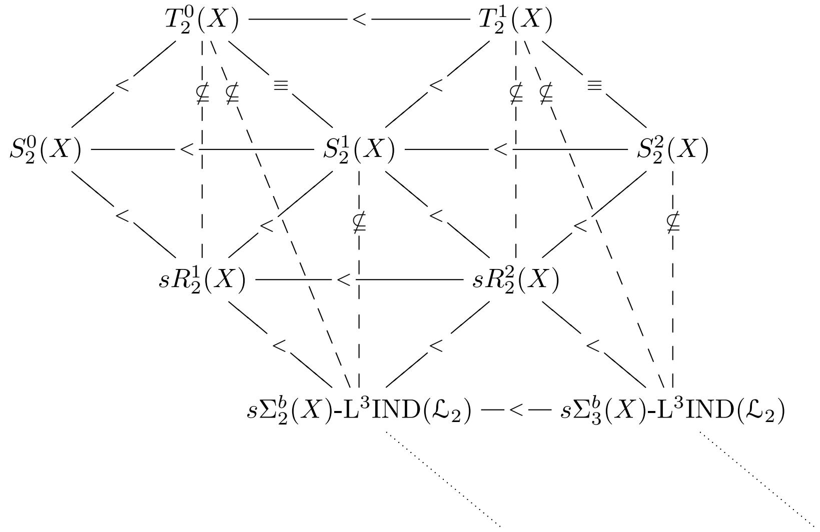

Figure 4: The $\mathfrak{L}_2$-theories

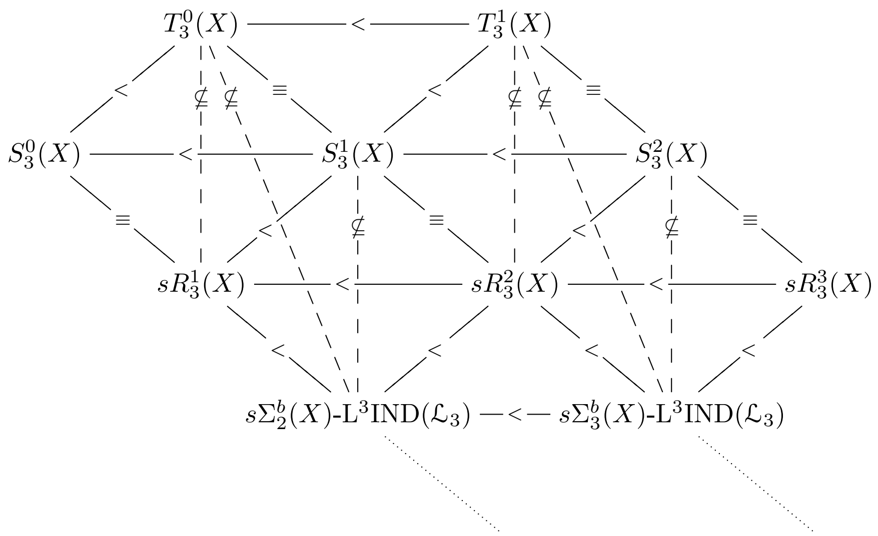

Figure 5: The $\mathcal{L}_3$-theories

In Figure 3-5 we have displayed the relationships of $\mathcal{L}_1$, $\mathcal{L}_2$, resp. $\mathcal{L}_3$-theories. Here we mean with $S < T$ that the theories S and T are separated and S is included in the consequences of T; with $S \equiv T$ that S and T have the same dynamic ordinals (this does not imply that S and T prove the same consequences); and with $S \not\subseteq T$ that S is not included in the consequences of T.

$T^0_2(X)$ and $S^1_2(X)$, resp. $T^1_2(X)$ and $S^2_2(X)$, must have the same dynamic ordinal because it is well-known that S^1_2 is $\forall\Sigma^b_1$-conservative over T^0_2 (if we identify T^0_2 with PV_1) (i.e. $T^0_2 \preceq_{\forall\Sigma^b_1} S^1_2$), resp. S^2_2 is $\forall\Sigma^b_2$-conservative over T^1_2 (i.e. $T^1_2 \preceq_{\forall\Sigma^b_2} S^2_2$). Different dynamic ordinals would lead to separations contradicting these conservativity results.

In thinking the other way round we can postulate the following (weak) **conjectures** by reading them out of our results. E.g., from Figure 5 we draw the following conjectures:

$$\begin{array}{ccccc} & & S^1_3 & \preceq_{\forall\Sigma^b_1(\mathcal{L}_3)} & sR^2_3, \\ T^1_3 & \preceq_{\forall\Sigma^b_2(\mathcal{L}_3)} & S^2_3 & \preceq_{\forall\Sigma^b_2(\mathcal{L}_3)} & sR^3_3; \end{array}$$

more general we can postulate

$$\begin{array}{ccccccc} & & S^1_k & \preceq_{\forall\Sigma^b_1(\mathcal{L}_k)} & \cdots & \preceq_{\forall\Sigma^b_1(\mathcal{L}_k)} & s\Sigma^b_{k-1}\text{-L}^{k-1}\text{IND}(\mathcal{L}_k), \\ T^1_k & \preceq_{\forall\Sigma^b_2(\mathcal{L}_k)} & S^2_k & \preceq_{\forall\Sigma^b_2(\mathcal{L}_k)} & \cdots & \preceq_{\forall\Sigma^b_2(\mathcal{L}_k)} & s\Sigma^b_k\text{-L}^{k-1}\text{IND}(\mathcal{L}_k). \end{array}$$

References

[1] Toshiyasu Arai. Some results on cut-elimination, provable well-orderings, induction and reflection. *Ann. Pure Appl. Logic*, 95:93–184, 1998.

[2] Theodore Baker, John Gill, and Robert Solovay. Relativizations of the $\mathcal{P} =? \mathcal{NP}$ question. *SIAM J. Comput.*, 4:431–442, 1975.

[3] Arnold Beckmann. *Seperating fragments of bounded predicative arithmetic.* PhD thesis, Westf. Wilhelms-Univ., Münster, 1996.

[4] Arnold Beckmann and Wolfram Pohlers. Applications of cut-free infinitary derivations to generalized recursion theory. *Ann. Pure Appl. Logic*, 94:7–19, 1998.

[5] Samuel R. Buss. *Bounded arithmetic*, volume 3 of *Stud. Proof Theory, Lect. Notes.* Bibliopolis, Naples, 1986.

[6] Samuel R. Buss. Axiomatizations and conservation results for fragments of bounded arithmetic. In Wilfried Sieg, editor, *Logic and Computation*, number 106 in Contemp. Math., pages 57–84, Providence, July 1990. Amer. Math. Soc.

[7] Samuel R. Buss. Relating the bounded arithmetic and the polynomial time hierarchies. *Ann. Pure Appl. Logic*, 75:67–77, 1995.

[8] Peter Clote and Jan Krajíček, editors. *Arithmetic, proof theory, and computational complexity.* Oxford Logic Guides. Oxford University Press, New York, 1993. Papers from the conference held in Prague, July 2–5, 1991, Oxford Science Publications.

[9] Johan Håstad. Computational limitations of small-depth circuits. *MIT Press, Cambridge, MA*, 1987.

[10] Jan Johannsen. A note on sharply bounded arithmetic. *Arch. Math. Logik Grundlag.*, 33:159–165, 1994.

[11] Jan Krajíček. Fragments of bounded arithmetic and bounded query classes. *Trans. Amer. Math. Soc.*, 338:587–98, 1993.

[12] Jan Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory.* Cambridge University Press, Heidelberg/New York, 1995.

[13] Jan Krajíček, Pavel Pudlák, and Gaisi Takeuti. Bounded arithmetic and the polynomial hierarchy. *Ann. Pure Appl. Logic*, 52:143–153, 1991.

[14] Rohit J. Parikh. Existence and feasibility in arithmetic. *J. Symbolic Logic*, 36:494–508, 1971.

[15] Wolfram Pohlers. *Proof Theory. An Introduction.* Number 1407 in Lect. Notes Math. Springer, Berlin/Heidelberg/New York, 1989.

[16] Chris Pollett. Structure and definability in general bounded arithmetic theories. *Ann. Pure Appl. Logic*, 100:189–245, 1999.

[17] Kurt Schütte. *Proof Theory.* Number 225 in Grundl. math. Wiss. Springer, Heidelberg/New York, 1977.

[18] Richard Sommer. Ordinal arithmetic in $I\Delta_0$. In Clote and Krajíček [8], pages 320–363. Papers from the conference held in Prague, July 2–5, 1991, Oxford Science Publications.

[19] Gaisi Takeuti. RSUV isomorphism. In Clote and Krajíček [8], pages 364–86. Papers from the conference held in Prague, July 2–5, 1991, Oxford Science Publications.

[20] Alex J. Wilkie and Jeff B. Paris. On the scheme of induction for bounded arithmetic formulas. *Ann. Pure Appl. Logic*, pages 261–302, 1987.

[21] Andrew C. Yao. Separating the polynomial-time hierarchy by oracles. *Proc. 26th Ann. IEEE Symp. on Foundations of Computer Science*, pages 1–10, 1985.

[22] Domenico Zambella. Notes on polynomially bounded arithmetic. *J. Symbolic Logic*, 61:942–966, 1996.