Charles University in Prague

Faculty of Mathematics and Physics

DOCTORAL THESIS

Michal Garlík

Model constructions for bounded arithmetic

Department of Algebra

Supervisor of the doctoral thesis:	Prof. RNDr. Jan Krajíček, DrSc.
Study programme:	Mathematics
Specialization:	Algebra, Theory of Numbers and Mathematical Logic

Prague 2015

Acknowledgements

First and foremost I would like to thank my supervisor Jan Krajíček for encouragement and inspiration, for his constant help, patience and careful guidance. I am indebted to my parents for their support during my studies. I wish to thank Samuel Buss, Emil Jeřábek, Leszek Kołodziejczyk, Pavel Pudlák and Neil Thapen for many clarifying discussions and good suggestions, and my colleagues Ján Pich and Sebastian Müller for broadening my view of the world and being good friends. I acknowledge financial support under grants GAUK 5732/20, IAA100190902 of GA AV ČR, and SVV-2014-260107.

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

In .... date .... signature of the author

Název práce: Konstrukce modelů omezené aritmetiky

Autor: Michal Garlík

Katedra: Katedra algebry

Vedoucí disertační práce: Prof. RNDr. Jan Krajíček, DrSc.

Abstrakt: Studujeme konstrukce modelů teorií omezené aritmetiky. Pomocí základních technik teorie modelů podáme nový důkaz Ajtaiovy věty o úplnosti pro nestandardně konečné struktury. Za použití omezené redukované mocniny (zobecnění ultraproduktu) navrhneme dvě nové metody konstrukce modelů omezené aritmetiky. První dá nový důkaz Bussovy dosvědčující věty. Druhou metodou ukážeme, že teorie R^1_2 je silnější než její varianta $strictR^1_2$ za věrohodného výpočetně-složitostního předpokladu (existence dostatečně silné jednosměrné permutace) a že za stejného předpokladu je teorie $PV_1 + \Sigma^b_1(PV) - \mathrm{LLIND}$ silnější než $PV_1 + strict\Sigma^b_1(PV) - \mathrm{LLIND}$. Pro relativizované teorie dokážeme, že $R^1_2(\alpha)$ je silnější $strictR^1_2(\alpha)$ (bez dodatečného předpokladu).

Klíčová slova: formální teorie, nestandardní model, třída složitosti

Title: Model constructions for bounded arithmetic

Author: Michal Garlík

Department: Department of Algebra

Supervisor: Prof. RNDr. Jan Krajíček, DrSc.

Abstract: We study constructions of models of bounded arithmetic theories. Using basic techniques of model theory we give a new proof of Ajtai's completeness theorem for nonstandard finite structures. Working in the framework of restricted reduced powers (a generalization of the ultrapower construction) we devise two methods of constructing models of bounded arithmetic. The first one gives a new proof of Buss's witnessing theorem. Using the second method we show that the theory R^1_2 is stronger than its variant $strictR^1_2$ under a plausible computational assumption (the existence of a strong enough one-way permutation), and that the theory $PV_1 + \Sigma^b_1(PV) - \mathrm{LLIND}$ is stronger than $PV_1 + strict\Sigma^b_1(PV) - \mathrm{LLIND}$ under the same assumption. Considering relativized theories, we show that $R^1_2(\alpha)$ is stronger than $strictR^1_2(\alpha)$ (unconditionally).

Keywords: formal theory, nonstandard model, complexity class

Contents

1. Preface

The general topic of the thesis are nonstandard models of weak theories of arithmetic and their links with problems in computational and proof complexity. The qualification “nonstandard” means that the structure is not isomorphic to the structure of natural numbers (in a particular language), the so called standard model. The qualification “weak” means that the induction axiom scheme (or some related axiom scheme) is restricted to some subclass of bounded formulas. The links with complexity theory stem form the facts that bounded formulas satisfying various additional particular restrictions define sets in various important computational complexity classes and that proofs of bounded formulas in weak theories can be translated into families of efficient propositional proofs. Introductions in Chapters 2 and 3 offer more specific and detailed overview of the topic and of the particular issues we tackle.

The thesis is formed by two papers:

- *A New Proof of Ajtai's Completeness Theorem for Nonstandard Finite Structures*, Archive for Mathematical Logic 54(3-4), (2015), pp. 413-424,

- *Construction of models of bounded arithmetic by restricted reduced powers*, submitted to Archive for Mathematical Logic,

that are included as Chapters 2 and 3, and by Chapter 4 that is a draft of a future paper. In the last chapter we add a few concluding remarks on possible lines for future research. Each chapter has its own bibliography.

2. A New Proof of Ajtai's Completeness Theorem for Nonstandard Finite Structures

This chapter is formed by paper "A New Proof of Ajtai's Completeness Theorem for Nonstandard Finite Structures" published in Archive for Mathematical Logic. It is identical to the original version except the numbering of definitions and statements.

A New Proof of Ajtai's Completeness Theorem for Nonstandard Finite Structures

Michal Garlík [*]

Faculty of Mathematics and Physics

Charles University in Prague

Abstract

Ajtai's completeness theorem roughly states that a countable structure A coded in a model of arithmetic can be end-extended and expanded to a model of a given theory G if and only if a contradiction cannot be derived by a (possibly nonstandard) proof from G plus the diagram of A, provided that the proof is definable in A and contains only formulas of a standard length. The existence of such model extensions is closely related to questions in complexity theory. In this paper we give a new proof of Ajtai's theorem using basic techniques of model theory.

2.1 Introduction

It is well known that various statements of complexity theory can be equivalently expressed in terms of mathematical logic, and model theory in particular. Some of these model-theoretic statements have a similar form asserting the existence of certain extensions of first-order structures. Let us mention informally three specific examples, all discussed already in detail in [3] (we refer the reader there for details of these equivalences).

Call a structure with a finite signature *nonstandard finite* if it is coded in a nonstandard model of true arithmetic and is countably infinite. The statement that parity is not in $\mathbf{AC}^0$ is equivalent to the statement that there are nonstandard finite structures containing a unary predicate whose size is odd in the original nonstandard model coding the structure but the same structure can be encoded in another nonstandard model which thinks that the predicate is of even size. The statement that the pigeonhole principle has no polynomial size propositional proofs in constant depth Frege systems is equivalent to the statement that every nonstandard finite structure has an expansion by a function violating the pigeonhole principle while satisfying induction for all definable sets. Finally, the statement that the class NP is not closed under complementation is equivalent to the existence of a non-3-colorable nonstandard finite graph such that any nonstandard finite structure on its vertices expanding the graph can be extended to a nonstandard finite structure containing a 3-coloring of the graph.

[*] Supported by grant GAUK 5732/2012 and in part by grant IAA100190902 of GA AV ČR. A part of this research has been done while the author was a visiting fellow at the Isaac Newton Institute in Cambridge (programme Semantics and Syntax) in Spring 2012.

Note that Ajtai's original proofs of super-polynomial lower bounds for constant depth circuits for parity (Ajtai [1], independently also Furst, Saxe, Sipser [6]) and for constant depth Frege proofs of the pigeonhole principle (Ajtai [2]) proceeded by establishing first the equivalent model-theoretic statements. It is thus of great interest to understand when such extensions can be constructed. Ajtai [3] and [4] formulated a theorem that can be understood as a completeness theorem for the existence of similar extensions-expansions (see Section 2.4 for the statement). Published version is [3], Ajtai has made available to us [4]. The newer version adds more details to the original version [3] making some assumptions used in [3] explicit now; see also our comments on the A-rule after Theorem 2.9.

Even though for proving lower bounds these models are usually built by some other methods and then only the principle behind the soundness direction of Ajtai's theorem is used, it might also prove useful for future constructions to keep in mind the completeness direction of the theorem, as such constructions do not seem to be trivial to find. In this paper we give a new (and simpler) proof of Ajtai's theorem, which relies on basic techniques of model theory.

2.2 Preliminaries

Definition 2.1. Let $L_0(exp)$ denote the first-order language of arithmetic with symbols $\leq, +, \cdot, 0, 1, 2^x$, where 2^x is a unary function symbol. A bounded quantifier is a quantifier of the form $\exists x \leq t$ or $\forall x \leq t$, where t is an $L_0(exp)$-term that does not include x. A $\Delta_0(exp)$-formula is a formula in the language $L_0(exp)$, in which all quantifiers are bounded. $I\Delta_0(exp)$ will denote the first-order theory in the language $L_0(exp)$ with the following axioms: the axioms of Robinson's arithmetic Q, $2^0 = 1$, $2^{(x+1)} = 2^x + 2^x$ and induction for all $\Delta_0(exp)$-formulas. We define $x \in y$ by the formula

$$\exists u \leq y \; \exists w < 2^x \; y = u \cdot 2^{x+1} + 2^x + w.$$

Let $B \subseteq M \models I\Delta_0(exp)$ and $b \in M$. We will say that b *codes* B in M if for each $x \in M$,

$$x \in B \Leftrightarrow M \models x \in b.$$

Note that bounded $\Delta_0(exp)$-comprehension holds in $I\Delta_0(exp)$, that is, for each $\Delta_0(exp)$-formula $\varphi(x, \bar{z})$, $I\Delta_0(exp)$ proves

$$\forall x \exists y < 2^x \forall u < x (u \in y \leftrightarrow \varphi(u, \bar{z})).$$

Hence a subset of M which is not cofinal in M is $\Delta_0(exp)$-definable in M if and only if it is coded by an element in M. See [7] for details on the theory $I\Delta_0(exp)$.

Definition 2.2. Assume that M is a model of $I\Delta_0(exp)$ and let $a \in M$. Then $\langle a, \leq \rangle$ will denote the structure which has universe $\{m \in M \mid M \models m \leq a\}$ and whose only relation $\leq$ is the restriction of the ordering in M to this universe.

Definition 2.3. Assume $\mathcal{L}$ is a first-order language containing a constant symbol a. Let φ be an $\mathcal{L}$-formula. Then $\varphi^{\leq a}$ is the formula we get by replacing every occurrence of $\forall x, \exists x$ in φ, where x is a variable, by $\forall x \leq a, \exists x \leq a$, respectively. If J is a set of formulas, $J^{\leq a}$ denotes the set $\{\varphi^{\leq a} \mid \varphi \in J\}$.

Definition 2.4. Assume $\mathcal{L}$ is a first-order language and A is an $\mathcal{L}$-structure. $\mathcal{L}(A)$ will denote the language we get from $\mathcal{L}$ by adding new constant symbols $\hat{a}$ for each $a \in A$ to $\mathcal{L}$. $\mathrm{Th}_{\mathcal{L}}(A)$, *the theory of* A, denotes the set of all $\mathcal{L}$-sentences true in A. Let A_A be the structure we get by expansion of A to $\mathcal{L}(A)$ such that for each $a \in A$ the constant symbol $\hat{a}$ is interpreted as a. The *atomic diagram of* A, which will be denoted by $\mathtt{diag}(A)$, is the set of all atomic and negated atomic sentences in the language $\mathcal{L}(A)$ that are true in A_A.

2.3 Proofs Definable in a Structure

Let $\mathbf{H}$ be any logical calculus for predicate logic, e.g. Hilbert-style calculus as defined in Chapter IV of [9]. Proofs in $\mathbf{H}$ can be thought of as finite trees whose nodes are labelled by formulas. The label of a node and the labels of its immediate successors form the conclusion and premises of either the rule modus ponens or the generalization rule. We are going to generalize the notion of an $\mathbf{H}$-proof by allowing possibly infinite proof trees definable in a structure.

Suppose that $\mathcal{L}$ is a first-order language containing a finite number of relation and function symbols and A is a set, $|A| \geq 2$. Let $\mathtt{symb}(\mathcal{L})$ denote the set of symbols of $\mathcal{L}$, that is relation and function symbols, symbols for variables, boolean operations, the existential and universal quantifiers, left and right parentheses and comma. We will want to represent the symbols of $\mathcal{L}(A)$ by the elements of a cartesian product A^i where i is a positive integer. So assume that for some positive integer i, $\mathtt{symb}(\mathcal{L})$ forms a subset of $A^i \setminus \{\langle a, a, \ldots, a\rangle \in A^i \mid a \in A\}$ and each symbol $\hat{a}$ is identified with the constant i-tuple $\langle a, a, \ldots, a\rangle$. Of course if A is finite there are only finitely many variables represented in this way in A^i. It would be easier to assume that A is infinite but later constructions do not need to assume that and we want to maintain maximal generality in this respect. Therefore for every $j > i$ we also consider an extended representation of the symbols of $\mathcal{L}(A)$ that will be denoted by $\mathtt{symb}^{(j,A)}(\mathcal{L}(A))$, such that the symbols of $\mathcal{L}(A)$ in A^i are naturally embedded in A^j in the following way. An element $\langle a_1, \ldots, a_i, a_{i+1}, \ldots, a_j\rangle$ is a non-variable symbol of $\mathtt{symb}^{(j,A)}(\mathcal{L}(A))$ iff $\langle a_1, \ldots, a_i\rangle$ is the corresponding non-variable symbol in A^i and $a_i = a_{i+1} = \ldots = a_j$. An element $\langle a_1, \ldots, a_i, a_{i+1}, \ldots, a_j\rangle$ is a variable symbol of $\mathtt{symb}^{(j,A)}(\mathcal{L}(A))$ iff $\langle a_1, \ldots, a_i\rangle$ is a variable symbol in A^i. Thus we have $|A|^{j-i}$ times more variables in $\mathtt{symb}^{(j,A)}(\mathcal{L}(A))$ than in A^i.

Definition 2.5. Let $\langle P, \leq\rangle$ be a partially ordered set and $a, b \in P$. We say that b is a *successor* of a if $a < b$ and there is no element $c \in P$ with $a < c < b$. We say that b is a *predecessor* of a if a is a successor of b.

Definition 2.6. Suppose that $\mathcal{K}, \mathcal{L}$ are first-order languages, each containing a finite number of relation and function symbols, $\mathcal{K} \subseteq \mathcal{L}$, $\mathcal{K}$ contains a binary relation symbol $\leq$ and a constant symbol a. Assume that A is a $\mathcal{K}$-structure whose universe is linearly ordered by $\leq$ and a is the largest element with respect to $\leq$. Suppose that G is a theory in $\mathcal{L}(A)$. Let q, k, l be positive integers, $T \subseteq A^q$, $\leq_T \subseteq A^{2q}$ and $\Theta \subseteq A^{q+kl}$. We say that $P = \langle T, \leq_T, \Theta\rangle$ is an $\mathbf{H}^{(A)}$*-proof from* G *with formula length* l if the following conditions are satisfied:

(1) The relations $T, \leq_T, \Theta$ are definable in A.

(2) $\langle T, \leq_T \rangle$ is a partially ordered set such that

- there exists an element 0_T in T and $T \models \forall a\, 0_T \leq_T a$,
- $T \models \forall a, b, c\, (a \leq_T c \wedge b \leq_T c \rightarrow a \leq_T b \vee b \leq_T a)$.

(3) $\mathtt{symb}(\mathcal{L}(A)) \subseteq A^r$ for some positive integer r, and $r \leq k$.

(4) The relation Θ is a function from T to A^{kl}, i.e. we can write

$$\bar{\Theta}(a_1, \ldots, a_q) = \langle a_{q+1}, \ldots, a_{q+kl} \rangle \quad \text{iff} \quad \Theta(a_1, \ldots, a_q, a_{q+1}, \ldots, a_{q+kl}).$$

(5) If $\langle a_1, a_2, \ldots a_{q+kl} \rangle \in A^{q+kl}$ and $\Theta(a_1, a_2, \ldots, a_{q+kl})$ then for every integer $i = 0, 1, \ldots, l-1$ we have

$$\langle a_{q+ki+1}, a_{q+ki+2} \ldots, a_{q+ki+k} \rangle \in \mathtt{symb}^{(k,A)}(\mathcal{L}(A))$$

and the sequence $\{\langle a_{q+ki+1}, a_{q+ki+2} \ldots, a_{q+ki+k} \rangle\}_{i=0}^{l-1}$ is an $\mathcal{L}(A)$-formula (padded on the left to length l using the symbol "," of $\mathcal{L}$ to accommodate all $\mathcal{L}(A)$-formulas of length at most l).

(6) If $\bar{c} \in T$ and the set S of its successors is nonempty then one of the two following conditions holds:

(i) $|S| \leq 2$ and the formulas assigned by the function $\bar{\Theta}$ to $\bar{c}$ and its successors are formed according to an inference rule of **H**, i.e. by modus ponens or generalization.

(ii) There exist $\langle a_2, a_3, \ldots, a_q \rangle \in A^{q-1}$ and an $\mathcal{L}(A)$-formula $\varphi(x)$ with one free variable such that $S = \{\langle a_1, a_2, \ldots, a_q \rangle \mid a_1 \in A\}$, for every $a_1 \in A$ we have $\bar{\Theta}(a_1, a_2, \ldots, a_q) = \varphi(a_1)$ and $\bar{\Theta}(\bar{c}) = \forall x \leq a\ \varphi(x)$. In this case we will say that $\forall x \leq a\ \varphi(x)$ was derived from $\{\varphi(a_1) \mid a_1 \in A\}$ by the *A-rule*.

(7) If $\bar{c} \in T$ is a maximal element with respect to $\leq_T$, then $\bar{\Theta}(\bar{c})$ is an instance of an axiom scheme of **H** or a sentence from G.

Remark 2.7. If T is a finite set, then $\langle T, \leq_T \rangle$ satisfying (2) from the previous definition is a finite tree. Without further restrictions on $\langle T, \leq_T \rangle$ considered in this definition it may happen, for example, that it contains a non-maximal element without any successors, preventing the proof from being sound. These problems will be resolved by assuming that the structure A has certain finiteness properties.

2.4 Ajtai's Completeness Theorem

Definition 2.8. Assume that $\mathcal{K}, \mathcal{L}$ are first-order languages, $\mathcal{K} \subseteq \mathcal{L}$, $\mathcal{K}$ contains a binary relation symbol $\leq$ and a constant symbol a. Suppose that A is a $\mathcal{K}$-structure such that $\leq$ is a linear order on A and a its largest element with respect to $\leq$. We say that an $\mathcal{L}$-structure B is an *expanded end-extension* of A if it meets the following four requirements:

(1) B is linearly ordered by $\leq$.

(2) The universe of A is a subset of the universe of B and for every $b \in B$, $B \models b \leq a$ iff $b \in A$.

(3) For all $k < \omega$, for all $b_1, \ldots, b_k \in A$ and for every k-ary relation symbol R in $\mathcal{K}$ we have $A \models R(b_1, \ldots, b_k)$ iff $B \models R(b_1, \ldots, b_k)$.

(4) For all $k < \omega$, for all $b_0, b_1, \ldots, b_k \in A$ and for every k-ary function symbol f in $\mathcal{K}$ we have $A \models f(b_1, \ldots, b_k) = b_0$ iff $B \models f(b_1, \ldots, b_k) = b_0$.

Assume further that G is a theory in $\mathcal{L}$. We say that G *has a model over* A if there exists a model B of G such that B is an expanded end-extension of A.

The following theorem is essentially Ajtai's theorem from [4] formulated in our terminology.

Theorem 2.9. (Ajtai [4]). Suppose that

($\star$) $M \models I\Delta_0(exp)$ and a is a nonstandard element of M such that the set $\{b \in M \mid M \models b \leq a\}$ is countable. Assume that A is an expansion of $\langle a, \leq \rangle$ to a first-order language $\mathcal{K}$ containing a finite number of relation and function symbols such that every function and relation of A is coded by an element in M. Also, let a be a constant symbol of $\mathcal{K}$ naming the element a.

Suppose further that $\mathcal{L} \supseteq \mathcal{K}$ is a first-order language containing a finite number of relation and function symbols and G is a theory in $\mathcal{L}$ such that the following conditions are satisfied:

(1) $G \vdash$"$\leq$ is a linear order",

(2) There is a set $\widehat{G}$ coded by an element in M such that $\widehat{G} \cap \mathbb{N}$ is the set of Gödel numbers of the sentences from G.

(3) $G \vdash \forall \bar{u}\ [\exists x \leq a\ \varphi(x, \bar{u}) \rightarrow \exists x \leq a\ [\varphi(x, \bar{u}) \wedge \forall y < x\ \neg\varphi(y, \bar{u})]]$ for every $\mathcal{L}$-formula $\varphi(x, \bar{u})$.

Then the following two statements are equivalent:

(I) There exists a positive integer l and an $\mathbf{H}^{(A)}$-proof of a contradiction from $\mathtt{diag}(A) \cup G$ with formula length l.

(II) G does not have a model over A.

Our setup of the theorem is very close to Ajtai's but there are some slight, cosmetic, differences. For example, we have A ordered by the ordering inherited from M whereas in [4] the ordering is only required to be definable in M, or, in the definition of "having model over" we have the end-extension requirement whereas in [4] there is a new unary predicate $\mathbf{U}$ in the language $\mathcal{L}$ which holds in the extended model exactly of the elements of A.

The only substantial difference in our formulation of the statement is the explicit inclusion of the A-rule. A form of the A-rule appears in Ajtai's proof in [4] as well. It is derived there by a repeated use of the cut rule and the assumption that there is a function which maps an element of the structure to be extended to its predecessor. But this assumption (on the theory G), or some similar one, is not made explicit in the statement of the theorem in [4] (and a theory G postulating the existence of an element in $\mathbf{U}$ strictly between some element of the initial structure and its successor would currently constitute a counterexample to the completeness direction of the theorem). Since the property that no new elements are introduced into the initial structure is essential, we included the A-rule.

We note that condition (3) is only needed for the implication (I) $\Rightarrow$ (II). This condition is usually required in problems of model construction like those mentioned in the introduction.

Let us remark that when G contains only sentences with quantifiers bounded by a one can use the well-known Paris-Wilkie translation of first-order proofs into propositional proofs (see e.g. [10]) and state condition (I) equivalently as follows: There is a (possibly nonstandard) proof of a contradiction from the translated sentences of G such that the proof is in a constant-depth Frege system with the depth being a standard number and the proof is considered as a k-ary relation definable in A for some standard number k. Constant-depth Frege systems are propositional proof systems in which formulas have a bounded alternation of unbounded fan-in conjunctions and disjunctions. In this way one gets the equivalence statement from the example with the pigeonhole principle mentioned in the introduction.

2.5 A New Proof of Ajtai's Completeness Theorem

Ajtai's original proof of the implication (II) $\Rightarrow$ (I) involves a lengthy and explicit construction of a model of G. We simplify this part significantly by utilizing the ideas behind the proof of the theorem due to Barwise and Schlipf [5], and (independently) Ressayre [12], that states that countable recursively saturated structures are resplendent (cf. [8], Theorem 15.7, for a presentation). The proof of the implication (I) $\Rightarrow$ (II) is essentially that of Ajtai.

Lemma 2.10. *Suppose ($\star$) from Theorem 2.9. Assume that $p(x)$ is a type in the language $\mathcal{K}$ in A over $a_0, \ldots, a_{n-1} \in A$, $n \in \omega$, and suppose that there is $d \in M$ such that*

$$\{m \in M \mid M \models m \in d\} \cap \mathbb{N} \\ = \{\ulcorner \varphi(x, x_0, \ldots, x_{n-1}) \urcorner \mid \varphi(x, a_0, \ldots, a_{n-1}) \in p(x)\}.$$

Then $p(x)$ is realized in A.

Proof. There exists a $\Delta_0(exp)$-formula $\mathrm{Tr}(t, u, v, w)$ such that for any $\mathcal{K}$-formula $\psi(\bar{z})$ and any tuple $\bar{c}$ of elements of A of the same length as $\bar{z}$ the following holds:

$$M \models \mathrm{Tr}(a, \langle \bar{e} \rangle, \ulcorner \psi(\bar{z}) \urcorner, \langle \bar{c} \rangle) \quad \text{iff} \quad A \models \psi(\bar{c}),$$

where $\bar{e}$ are the elements of M coding the functions and relations of A. Since all the quantifiers in ψ as well as the values of all the terms in $\psi(\bar{c})$ can be bounded by a, Tr can be constructed as a truth definition for bounded formulas with all quantifiers in it bounded by exponential terms. See e.g. [11] for details of the truth definition.

Now let $\theta(s)$ be the following formula:

$$\exists r \leq a \, \forall y \leq s \, (y \in d \to \mathrm{Tr}(a, \langle \bar{e} \rangle, y, \langle r, a_0, \ldots, a_{n-1} \rangle)) \, .$$

It is a $\Delta_0(exp)$-formula with parameters $a, d, \bar{e}, a_0, \ldots a_{n-1}$ and since $p(x)$ is a type, $M \models \theta(i)$ for every $i \in \mathbb{N}$. Therefore, by overspill, there is a nonstandard $i \in M$ such that $M \models \theta(i)$. It follows that there exists an element in A that satisfies all the formulas from $p(x)$ in A, i.e. $p(x)$ is realized. □

Lemma 2.11. *Suppose* $(\star)$ *from Theorem 2.9. Suppose further that* $\mathcal{L} \supseteq \mathcal{K}$ *is a first-order language containing a finite number of relation and function symbols and* G *is a theory in* $\mathcal{L}$ *such that the following conditions are satisfied:*

(1) $G \vdash$ *"$\leq$ is a linear order",*

(2) *There is a set* $\widehat{G}$ *coded by an element in* M *such that* $\widehat{G} \cap \mathbb{N}$ *is the set of Gödel numbers of the sentences from* G.

Then the following two statements are equivalent:

(I) $\mathrm{Th}_{\mathcal{K}}(A)^{\leq a} \cup G$ *is consistent,*

(II) G *has a model over* A.

Proof. The implication (II)⇒(I) is obvious; let us prove (I)⇒(II)[1]. We will construct a complete theory J in the language

$$\mathcal{L}(A, C) = \mathcal{L} \cup \{b \mid b \in A\} \cup \{c_i \mid i < \omega\},$$

where we denote the constant symbol representing an element b by b itself and where $C = \{c_i \mid i < \omega\}$ are new distinct constant symbols, such that

(i) $G \subseteq J$,

(ii) for every $\mathcal{K}(A)$-sentence σ, $J \vdash \sigma^{\leq a} \Rightarrow A \models \sigma$,

(iii) if $\varphi(x)$ is an $\mathcal{L}(A, C)$-formula with only x free and $J \vdash \exists x \varphi(x)$, then either $J \vdash \varphi(c_i)$ for some $i < \omega$ or $J \vdash \varphi(b)$ for some $b \in A$,

(iv) for all $i < \omega$, $J \vdash a \leq c_i$.

It is clear that the canonical structure for the theory J is the desired model of G over A.

We will construct theories J_i $(i < \omega)$ in the language $\mathcal{L}(A, C)$ such that the following two statements will hold for all $j < \omega$:

[1] As pointed out by the referee, if G is recursively axiomatizable, then (I)⇒(II) can be quickly seen as follows: A is a countable recursively saturated structure (by Lemma 2.10), hence it is resplendent, and having a countable expanded end-extension to a model of G can be characterized by a recursive set of axioms in a language with finitely many new symbols.

(v) If σ is a $\mathcal{K}(A)$-sentence and $J_j \vdash \sigma^{\leq a}$, then $A \models \sigma$.

(vi) There exists $l_j < \omega$ such that all the constants from C occurring in the formulas of J_j are exactly $c_0, c_1, \ldots, c_{l_j-1}$.

Since $\mathrm{Th}_{\mathcal{K}}(A)^{\leq a} \cup G$ is consistent, (v) is true for $j = 0$ and $J_0 := G$. There are no constants from C in J_0, hence (vi) is true as well, with $l_0 = 0$.

Let $\{\varphi_i(x) \mid i < \omega\}$ be an enumeration of all $\mathcal{L}(A, C)$-formulas with only one free variable x such that every such formula occurs in it infinitely many times. (Here we use the assumption that A is countable.) Assume that J_i has been constructed so that (v) and (vi) hold for i. Let $k > i$ be the smallest integer such that the constants from C occurring in $\varphi_k(x)$ are among $c_0, c_1, \ldots, c_{l_i-1}$. We shall construct J_{i+1} by adding to J_i one of the following formulas

- $\forall x \neg\varphi_k(x) \wedge a = c_{l_i}$,
- $\varphi_k(c_{l_i}) \wedge a < c_{l_i}$,
- $\varphi_k(b)$ for some $b \in A$

so that (v) and (vi) hold for $i + 1$.

Before we show that one of these choices can be made, note that if $\varphi_k(x)$ is $a < x$, J_{i+1} has to include the new constant c_{l_i} in its language. As $a < x$ will be dealt with infinitely many times during the construction of all J_j's ($j < \omega$), every constant from C will eventually appear in the language of J_j for some $j < \omega$. Therefore every $\mathcal{L}(A, C)$-formula will be treated at some step of the construction. Hence it is clear that once all J_j's are constructed in the way described above and satisfy (v) and (vi), the theory $J = \bigcup_{j<\omega} J_j$ is complete and has the required properties (i)-(iv).

It remains to show that J_{i+1} can be constructed by adding one of the above formulas to J_i so that (v) is true for $i + 1$. Let $a_0, \ldots, a_{n-1}$ be the elements of A occurring in $J_i \cup \{\varphi_k(x)\}$ and suppose, for a contradiction, that none of the above choices can be made. Then there are $\mathcal{K} \cup \{a_0, \ldots, a_{n-1}\}$-sentences σ, γ, η_r for all $r < n$, and for all $b \in A \setminus \{a_0, \ldots, a_{n-1}\}$ there is a $\mathcal{K} \cup \{a_0, \ldots, a_{n-1}\}$-formula $\xi_b(x)$ such that

$$\begin{aligned} J_i + \forall x \neg\varphi_k(x) \wedge a = c_{l_i} &\vdash \sigma^{\leq a} \\ J_i + \varphi_k(c_{l_i}) \wedge a < c_{l_i} &\vdash \gamma^{\leq a} \\ J_i + \varphi_k(a_r) &\vdash \eta_r^{\leq a} \ \text{ for all } \ r < n \\ J_i + \varphi_k(b) &\vdash \xi_b^{\leq a}(b) \ \text{ for all } \ b \in A \setminus \{a_0, \ldots, a_{n-1}\} \end{aligned}$$

but $A \not\models \sigma$, $A \not\models \gamma$, $A \not\models \eta_r$ for all $r < n$, and $A \not\models \xi_b(b)$ for all elements b in $A \setminus \{a_0, \ldots, a_{n-1}\}$. Using the fact that c_{l_i} does not occur in $J_i + \varphi_k(x)$ it follows that

$$\begin{aligned} J_i + \neg\sigma^{\leq a} &\vdash \exists x \varphi_k(x) \\ J_i + \neg\gamma^{\leq a} &\vdash \forall x(\varphi_k(x) \to x \leq a) \\ J_i + \bigwedge_{r<n} \neg\eta_r^{\leq a} &\vdash \forall x(\varphi_k(x) \to \bigwedge_{r<n} x \neq a_r). \end{aligned}$$

Thus if $\theta(x)$ is a $\mathcal{K} \cup \{a_0, \ldots, a_{n-1}\}$-formula and

$$J_i \vdash \forall x(\varphi_k(x) \to \theta^{\leq a}(x))$$

then

$$J_i \vdash \neg\sigma^{\leq a} \wedge \neg\gamma^{\leq a} \wedge \bigwedge_{r<n} \neg\eta_r^{\leq a} \to (\exists x \leq a)\theta^{\leq a}(x).$$

By the induction hypothesis it follows that $A \models \exists x\theta(x)$. This consideration shows that the set $p(x)$ consisting of all $\mathcal{K} \cup \{a_0, \ldots, a_{n-1}\}$-formulas of the form $\theta(x) \wedge \theta(x) \wedge \ldots \wedge \theta(x)$ (s conjunctions) such that there is a proof from J_i of length s of the sentence

$$\forall x(\varphi_k(x) \to \theta^{\leq a}(x))$$

is a type in A. Moreover, there exists a $\Delta_0(exp)$-formula $\pi(y)$ such that

$$\{m \in M \mid M \models \pi(m)\} \cap \mathbb{N} = \{\ulcorner\delta\urcorner \mid \delta \in p(x)\}.$$

(Here we use the condition (2).) Hence there is d in M such that

$$\{m \in M \mid M \models m \in d\} \cap \mathbb{N} = \{\ulcorner\delta\urcorner \mid \delta \in p(x)\},$$

by $\Delta_0(exp)$-comprehension. By Lemma 2.10, $p(x)$ is realized by some element $b \in A \smallsetminus \{a_0, \ldots, a_{n-1}\}$ (because formulas equivalent to $x = a_r \to \eta_r^{\leq a}$ for $r < n$ are in $p(x)$). But for this b we have $J_i + \varphi_k(b) \vdash \xi_b^{\leq a}(b)$. Since b does not occur in J_i it follows that $J_i \vdash \forall x(\varphi_k(x) \to \xi_b^{\leq a}(x))$ so we have that $\xi_b(x) \wedge \xi_b(x) \wedge \ldots \wedge \xi_b(x)$ is in $p(x)$ for some suitable number of conjuncts. Thus $A \models \xi_b(b)$, a contradiction with our assumption on $\xi_b(x)$. Thus J_{i+1} can be found satisfying (v) for $i+1$ and by its construction it obviously satisfies (vi).

□

Definition 2.12. Suppose that $\mathcal{H}$ is a first-order language, B is an $\mathcal{H}$-structure, k is a positive integer, $X \subseteq B^k$ is a definable set in B and $\leq$ is a definable linear order on X. We say that X is *quasi-finite in B with respect to* $\leq$ if the following requirements are met:

(1) X has a smallest and a largest element,

(2) each definable nonempty subset of X has a smallest element,

(3) each element of X, except for the smallest one, has a predecessor.

Lemma 2.13. *Assume that $\mathcal{H}$ is a first-order language, B is an $\mathcal{H}$-structure, X is a definable set in B which is quasi-finite in B with respect to a definable linear order $\leq$ on X. Suppose that i is a positive integer and $\langle P, \leq_P\rangle$ is a nonempty partial order definable in B so that $P \subseteq X^i$. Then P has a minimal element.*

Proof. Let $Y = X^i$ and $\leq_Y$ be the lexicographic order on Y induced by $\leq$. It is easily checked that $\leq_Y$ is definable in B and that Y is quasi-finite in B with respect to $\leq_Y$. Next we verify that each definable nonempty subset U of Y has a largest element in $\leq_Y$. Indeed, it is either the largest element 1_Y of Y if $1_Y \in U$, or it is the predecessor of the least element of the set $\{x \in Y \mid \neg\exists y\,(y \in U \wedge x \leq_Y y)\}$.

Consider the set $V = \{x \in P \mid \forall y \in P\,(y \leq_P x \to x \leq_Y y)\}$. V is nonempty because it contains the $\leq_Y$-smallest element of P. Let v be the $\leq_Y$-largest element of V. Either v is a minimal element of $\langle P, \leq_P\rangle$ or the set $W = \{x \in P \mid x <_P v\}$ is nonempty. We will show that the latter case leads to a contradiction. Consider the $\leq_Y$-smallest element w of W. Since $v \in V$ and $w <_P v$ we have $v <_Y w$. Because of the maximality of v in V we get $w \notin V$ and so there must exist $u \in P$ with $u \leq_P w$ and $u <_Y w$. By transitivity of $\leq_P$ we get $u <_P v$ and so $u \in W$ in contradiction to the minimality of w. □

Proof of Ajtai's Completeness Theorem.
First we show the implication (II)⇒(I). If G does not have a model over A, then by Lemma 2.11 there exists a proof of a contradiction from $\mathrm{Th}_{\mathcal{K}}(A)^{\leq a} \cup G$. Since the proof is finite, there exists an $\mathbf{H}^{(A)}$-proof P_0 of a contradiction from $\mathrm{Th}_{\mathcal{K}}(A)^{\leq a} \cup G$ with formula length l for some positive integer l. Thus it remains to find an $\mathbf{H}^{(A)}$-proof from $\mathtt{diag}(A)$ of each of the finitely many sentences of $\mathrm{Th}_{\mathcal{K}}(A)^{\leq a}$ that occur as axioms in P_0 and attach these proofs to P_0. It suffices to show the following claim.

Claim: For every $\mathcal{K}$-formula $\alpha(\bar{x})$, where $\bar{x} = \langle x_1, \ldots, x_n\rangle$ for some positive integer n, there exist $\mathcal{K}(A)$-formulas $\tau^\alpha(\bar{x}, \bar{u})$, $\lambda^\alpha(\bar{x}, \bar{v})$, $\phi^\alpha(\bar{x}, \bar{w})$ (where $\bar{u}, \bar{v}, \bar{w}$ are some tuples of free variables) such that for every $\bar{a} = \langle a_1, \ldots, a_n\rangle \in A^n$, if $A \models \alpha(\bar{a})$ then the triple of relations defined in A by formulas $\tau^\alpha(\bar{a}, \bar{u}), \lambda^\alpha(\bar{a}, \bar{v})$, $\phi^\alpha(\bar{a}, \bar{w})$ is an $\mathbf{H}^{(A)}$-proof of $\alpha^{\leq a}(\bar{a})$ from $\mathtt{diag}(A)$.

To prove the claim, we may assume that in α negation only occurs in front of atomic formulas. We proceed by induction on the logical complexity of α. The claim is obvious if α is an atomic or negated atomic formula. (For example, suppose that $\alpha(x_1, x_2)$ is $x_1 \leq x_2$, we have symbols represented in $\mathtt{symb}^{(2,A)}(\mathcal{L}(A))$ and $b_1 \neq b_2 \in A$ are such that $\langle b_1, b_2\rangle$ represents $\leq$. Then we choose τ^α to define a single element, say a, and for $\phi^\alpha(x_1, x_2, w_0, \ldots, w_6)$ we can take

$$x_1 \leq x_2 \wedge w_0 = a \wedge w_1 = w_2 = x_1 \wedge w_3 = b_1 \wedge w_4 = b_2 \wedge w_5 = w_6 = x_2.)$$

For α of the form $\alpha_1 \wedge \alpha_2$ we just join the $\mathbf{H}^{(A)}$-proofs of $\alpha_1^{\leq a}$, $\alpha_2^{\leq a}$ and of an appropriate axiom by applying modus ponens twice. If α is of the form $\alpha_1 \vee \alpha_2$, the three formulas defining an $\mathbf{H}^{(A)}$-proof of $\alpha^{\leq a}(\bar{a})$ have to distinguish two cases depending on $\bar{a}$ and hence have the form of a disjunction: either $\alpha_1(\bar{a})$ and the $\mathbf{H}^{(A)}$-proof of $\alpha_1^{\leq a}(\bar{a})$ is joined with that of an instance of disjunction introduction (left), or $\neg\alpha_1(\bar{a}) \wedge \alpha_2(\bar{a})$ and the $\mathbf{H}^{(A)}$-proof of $\alpha_2^{\leq a}(\bar{a})$ is joined with that of an instance of disjunction introduction (right). If $\alpha(\bar{x})$ is $\forall x_0\ \beta(x_0, \bar{x})$ we use the induction hypotheses for β to uniformly (in $\bar{a}$'s satisfying $A \models \alpha(\bar{a})$) define a family of $|A|$ disjoint $\mathbf{H}^{(A)}$-proofs of $\beta^{\leq a}(b, \bar{a})$ $(b \in A)$ and join these proofs by an application of the A-rule. Finally, let α be of the form $\exists x_0\ \beta(x_0, \bar{x})$. From the assumptions $(\star)$ it easily follows that the least number principle for $\mathcal{K}$-formulas holds in A. So we apply it to the formula $\beta(x_0, \bar{x})$ with parameters $\bar{x}$ and use the induction hypotheses to uniformly (in $\bar{a}$'s satisfying $A \models \alpha(\bar{a})$) define the $\mathbf{H}^{(A)}$-proof of $\beta^{\leq a}(b, \bar{a})$ with b the least possible such that $A \models \beta(b, \bar{a})$. Then we join this proof by modus ponens with an instance of $\exists$-introduction axiom. This completes the proof of the claim and of the implication (II)$\Rightarrow$(I).

Now we show the implication (I) $\Rightarrow$ (II). Suppose there exists an $\mathbf{H}^{(A)}$-proof P of a contradiction from $\mathtt{diag}(A) \cup G$ with formula length l, but contrary to (II), there exists a model N of G over A. The universe of A is defined in N by the formula $x \leq a$. All the functions and relations of A are definable in N by restricting the functions and relations of the same name in N to elements $\leq a$. Therefore the components $T, \leq_T, \Theta$ of P are defined in N as well. It follows from the way A originated from $M \models I\Delta_0(exp)$ and from the condition (3) of the theorem that the universe of A is quasi-finite in N with respect to $\leq$. We know that $T \subseteq A^q$ for some positive integer q. Therefore Lemma 2.13 implies that each nonempty subset of T which is definable in N has a maximal and a minimal element with respect to $\leq_T$.

The lengths of the formulas of the proof P are bounded by l, the symbols of $\mathcal{L}(A)$ used in these formulas are those of $\mathtt{symb}^{(k,A)}(\mathcal{L}(A))$ for some positive integer k and the language $\mathcal{L}$ contains only finitely many function and relation symbols. Therefore there are only finitely many formula shapes of length l and hence there exists a function $\Gamma : A^{kl} \mapsto \{0, 1\}$ definable in N that assigns truth value in N to each $\mathcal{L}(A)$ formula $\{\langle a_{i+1}, \ldots, a_{i+k}\rangle\}_{i=0}^{l-1} \in (A^k)^l$.

Now let F be the set of those elements t of T that satisfy $\Gamma(\bar{\Theta}(t)) = 0$. For the root 0_T of T we have that $\bar{\Theta}(0_T)$ is a contradiction, so F is nonempty. Since F is a nonempty definable subset of T there exists an element $m \in F$ which is maximal in F with respect to $\leq_T$. If t is a maximal element of $\langle T, \leq_T\rangle$ then $\bar{\Theta}(t)$ is an instance of an axiom scheme of $\mathbf{H}$ or a sentence from $\mathtt{diag}(A) \cup G$, so $\bar{\Theta}(t)$ holds in N. Therefore m cannot be a maximal element of $\langle T, \leq_T\rangle$. Let $Q = \{t \in T \mid m <_T t\}$. Since Q is a nonempty definable subset of T it has a minimal element with respect to $\leq_T$. Let $S \subseteq Q$ be the set of minimal elements of Q with respect to $\leq_T$. It is the set of all $\leq_T$-successors of m and by the definition of $\mathbf{H}^{(A)}$-proof the formulas assigned by $\bar{\Theta}$ to m and its successors must satisfy one of the inference rules. The inference rules have the property that for every element $t \in T$ the formula $\bar{\Theta}(t)$ is true in N if for every successor t' of t the formula $\bar{\Theta}(t')$ is true in N. (The case of the A-rule relies here on the fact that N is a model over A). This contradicts $\Gamma(\Theta(s)) = 1$ for all $s \in S$.

□

Acknowledgement

I wish to thank Jan Krajíček for very helpful discussions and the anonymous referee for valuable comments and suggestions.

Bibliography

[1] M. Ajtai, *Σ_1^1-formulae on finite structures*, Annals of Pure and Applied Logic 24 (1983) 1-48

[2] M. Ajtai, *The Complexity of the Pigeonhole Principle*, Proceedings of the IEEE 29th Annual Symposium on Foundation of Computer Science (1988) 346-355

[3] M. Ajtai, *Generalizations of the Compactness Theorem and Gödel's Completeness Theorem for Nonstandard Finite Structures*, Proceedings of the 4th international conference on Theory and applications of models of computation (2007) 13-33.

[4] M. Ajtai, *A Generalization of Gödel's Completeness Theorem for Nonstandard Finite Structures*, manuscript (2011)

[5] J. Barwise, J. Schlipf, *An Introduction to Recursively Saturated and Resplendent Models* J. Symbolic Logic, 41 (1976) 531-536

[6] M. Furst, J. Saxe, M. Sipser, *Parity, Circuits, and the Polynomial-Time Hierarchy*, Mathematical Systems Theory 17 (1984) 13-27

[7] P. Hájek, P. Pudlák, *Metamathematics of first order arithmetic*, Springer, 1993

[8] R. Kaye, *Models of Peano Arithmetic*, Oxford Logic Guides 15, Oxford University Press, 1991

[9] S. C. Kleene, *Introduction to metamathematics*, D. Van Nostrand Co., Inc., New York, N. Y., 1952.

[10] J. Krajíček, *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, Cambridge University Press, 1995

[11] J. Paris and C. Dimitracopoulos, *Truth definitions for Δ_0 formulae*, Logic and Algorithmic, L'Enseignement Mathematique Monographie no 30, Geneva, (1982) 317-330.

[12] J. P. Ressayre, *Models with Compactness Properties Relative to an Admissible Language*, Annals of Pure and Applied Logic 11 (1977) 31-55.

3. Construction of models of bounded arithmetic by restricted reduced powers

This chapter is formed by paper "Construction of models of bounded arithmetic by restricted reduced powers" submitted to Archive for Mathematical Logic. It is identical to the original version except the numbering of definitions and statements as we refer to some of them in Chapter 4.

Construction of models of bounded arithmetic by restricted reduced powers

Michal Garlík *

Faculty of Mathematics and Physics

Charles University in Prague

Abstract

We present two constructions of models of bounded arithmetic, both in the form of a generalization of the ultrapower construction, that yield nonelementary extensions but do not introduce new lengths. As an application we show, assuming the existence of a one-way permutation g hard against polynomial-size circuits, that $\mathit{strict}R_2^1(g)$ is weaker than $R_2^1(g)$. In particular, if such a permutation can be defined by a term in the language of R_2^1, then $\mathit{strict}R_2^1$ is weaker than R_2^1.

3.1 Introduction

It is well known that some problems in complexity theory can be cast as problems of constructions of expanded extensions of models of bounded arithmetic ([1], [2], [8],[12], [13]). These models are usually required to satisfy some form of bounded induction but at the same time not introduce any new lengths of strings. In [13] Máté encourages the study of possible modifications of known methods of constructing elementary extensions of models of arithmetic so as to obtain nonelementary partial extensions and mentions the restricted ultrapower construction of Kochen and Kripke ([11]) as such an example. Restricted ultrapowers in the context of bounded arithmetic were used in [5]. Modifications of the ultrapower construction can make it easier to meet the requirement of preserving lengths of strings, so we concentrate on them in this paper, which consists of two such constructions. In Section 3.2 we introduce our framework of *restricted reduced powers*. In Section 3.3 we illustrate this framework on Buss's witnessing theorem. Construction A (Theorem 3.1) developed there manages not to introduce new lengths but does not appear to be easily amenable to other situations. Thus, in Section 3.4, we present Construction B (Theorem 3.4), which seems to be a more flexible and promising approach. An example illustrating Construction B is given in Section 3.5. Its statement (Theorem 3.6) could be easily derived from the known witnessing theorem for R_2^1 except for the part that no new lengths are added - but our emphasis here is on developing constructions adding no new lengths. In Section 3.6 we use Construction B together with the assumption of the existence of a one-way permutation g hard against polynomial-size circuits to

*Supported by grant GAUK 5732/2012.

show that $strictR^1_2(g)$ is weaker than $R^1_2(g)$. In particular, if such a permutation is definable by a term in the language of R^1_2, then $strictR^1_2$ is weaker than R^1_2.

3.2 Preliminaries

Let $\mathcal{L}$ be a first-order language, Ω a non-empty set and M an $\mathcal{L}$-structure. For an $\mathcal{L}$-formula $\varphi(x_1, \ldots, x_k)$ and functions $f_1, \ldots, f_k$ from Ω to M denote $[\![\varphi(f_1, \ldots, f_k)]\!]$ the set of $u \in \Omega$ such that

$$M \models \varphi(f_1(u), \ldots, f_k(u)).$$

Let $\mathcal{F}$ be a set of functions from Ω to M and let $\mathcal{G}$ be a filter on the powerset of Ω. We define an equivalence relation $\sim$ on $\mathcal{F}$ by

$$f \sim g \quad \text{iff} \quad [\![f = g]\!] \in \mathcal{G}.$$

We write $[f]$ for the equivalence class of an element $f \in \mathcal{F}$. We define an $\mathcal{L}$-structure $\mathcal{F}/\mathcal{G}$ as follows. The universe is the set of equivalence classes $[f]$ with $f \in \mathcal{F}$. The interpretation of a k-ary relation symbol R and a k-ary function symbol F of $\mathcal{L}$ is given by:

$$\begin{aligned}
&([f_1], \ldots, [f_k]) \in R^{\mathcal{F}/\mathcal{G}} \quad \text{iff} \quad [\![R(f_1, \ldots, f_k)]\!] \in \mathcal{G},\\
&F^{\mathcal{F}/\mathcal{G}}([f_1], \ldots, [f_k]) = [f]\\
&\text{where } f(u) = F^M(f_1(u), \ldots, f_k(u)) \text{ for each } u \in \Omega.
\end{aligned}$$

Here we must have some $f' \in \mathcal{F}$ with $f' \sim f$ for this definition to make sense. Its correctness is then readily verified using that $\mathcal{G}$ is a filter. We call such a structure $\mathcal{F}/\mathcal{G}$, where $\mathcal{F} \subsetneq M^\Omega$, a *restricted reduced power.*

We use restricted reduced powers in the following setting: M is a model of arithmetic, Ω is a definable subset of M, $\mathcal{F}$ consists of some functions definable in M and $\mathcal{G}$ is a filter on the M-definable subsets of Ω. We start by picking a nonstandard number $n \in M$ and we want to construct $\mathcal{F}/\mathcal{G}$ such that it is another model of arithmetic and such that it has, like M, the interval $[0, \ldots, n]$ as an initial segment. That is, $[0, \ldots, n]$ is embedded in $\mathcal{F}/\mathcal{G}$ onto $[[c_0], \ldots, [c_n]]$, where $c_j \in \mathcal{F}$ is the constant function with value j.

We present two specific constructions of restricted reduced powers intended for obtaining models of bounded arithmetic theories S^i_2 and R^i_2. The language L_2 of these theories has non-logical symbols $0, S, +, \cdot, x\#y$, $|x|$, $\lfloor x/2 \rfloor, \leq, \dot{-}, \mathit{MSP}$. Theories S^i_2 and R^i_2 share a finite set BASIC of open axioms fixing the basic properties of this language (see [3] for the axioms; they extend Buss's BASIC axioms from [4] by adding axioms for $\dot{-}$ (subtraction) and for $\mathit{MSP}(x, i) := \lfloor x/2^i \rfloor$ (the ith most significant part of x)). S^i_2 extends BASIC by the axiom scheme Σ^b_i-LIND. This scheme consists of formulas of the form

$$\phi(0) \wedge (\forall y < |x|)(\phi(y) \rightarrow \phi(S(y))) \rightarrow \phi(|x|)$$

where $\phi(x)$ is Σ^b_i and may contain other free variables besides x. Similarly, R^i_2 extends BASIC by the axiom scheme Σ^b_i-LLIND, which consists of formulas of the form

$$\phi(0) \wedge (\forall y < ||x||)(\phi(y) \rightarrow \phi(S(y))) \rightarrow \phi(||x||)$$

where $\phi(x)$ is Σ^b_i and may contain other free variables besides x. We may also accept the last induction scheme only for Σ^b_i formulas in the *strict* form, i.e.

$$(\exists x_1 \leq t_1)(\forall x_2 \leq t_2)\dots(Qx_i \leq t_i)\varphi,$$

where Q is $\exists$ if i is odd and $\forall$ if i is even, $t_1,\dots,t_i$ are L_2-terms and φ is sharply bounded. Then the resulting theory is called $strictR^i_2$.

The following terms, as defined e.g. in [10], will be useful to define pairing and a simple sequence coding in R^1_2:

$$\begin{aligned}
2^{|x|} &:= 1\#x\\
2^{\min(x,|y|)} &:= MSP(2^{|y|}, |y| \dot{-} x)\\
LSP(x,i) &:= x \dot{-} 2^{\min(i,|x|)} \cdot MSP(x,i)\\
\beta_a(w,i) &:= MSP(LSP(w, Si\cdot|a|), i\cdot|a|)\\
bd(a,s) &:= 2(2a\#2s).
\end{aligned}$$

So $LSP(x,i)$ gives the number consisting of the last i bits of x. If w is the number whose binary representation consists of 1 followed by binary representations of numbers $b_1,\dots,b_\ell$, each padded with zeros to be of length $|a|$, then $\beta_a(w,i) = b_i$. The term $bd(a,s)$ gives a bound on the code for a sequence length $|s|$ with each item bounded by a.

Pairs are coded as $\langle x,y\rangle := (B+y)\cdot 2B + (B+x)$ where $B = 2^{|\max(x,y)|}$. To project out the first and second coordinate from an ordered pair we use terms $\langle w\rangle_1 := \beta_{\lfloor\frac{1}{2}|w|\rfloor \dot{-} 1}(\beta_{\lfloor\frac{1}{2}|w|\rfloor}(w,0),0)$ and $\langle w\rangle_2 := \beta_{\lfloor\frac{1}{2}|w|\rfloor \dot{-} 1}(\beta_{\lfloor\frac{1}{2}|w|\rfloor}(w,1),0)$.

Let Γ be a class of formulas. *Sharply bounded collection scheme* (also called *replacement scheme*) $BB\Gamma$ is

$$\begin{aligned}
&(\forall x \leq |s|)(\exists y \leq t)A(x,y) \rightarrow\\
&\qquad\qquad (\exists w \leq bd(t,s))(\forall x \leq |s|)\beta_t(w,x) \leq t \wedge A(x,\beta_t(w,x))
\end{aligned}$$

for each $A(x,y) \in \Gamma$ and for all terms s,t, such that $A(x,y), s, t$ may contain other free variables but t and s do not involve x or y.

Sometimes we will use a different coding of a sequence of numbers by a single number such that we have a function $(w)_i$ (having two inputs unlike $\beta_a(w,i)$) to return the ith item of the sequence coded by w and a function $lh(w)$ which gives the length of the sequence coded by w. We can use the coding described in [3] or the one in [6]; important is that the functions $(w)_i$ and $lh(w)$ are Δ^b_1-defined in R^1_2 and R^1_2 proves their basic properties, e.g. that each sequence can be extended by an arbitrary element.

Besides the standard notation Σ^b_i, Π^b_i for classes of bounded formulas we shall use the symbol $\Delta^{<t}_0$, where t is a term, to denote formulas with all quantifiers bounded by t, i.e. each of their quantifiers is of the form $(\exists x < t)$ or $(\forall x < t)$.

If F is a new function symbol then formulas from the *relativized* classes $\Sigma^b_i(F)$ or $\Pi^b_i(F)$ are allowed to use the symbol F freely. The relativized theories $S^i_2(F)$, $R^i_2(F)$ or $strictR^i_2(F)$ have induction axioms for the corresponding relativized class of formulas.

3.3 Construction A

The starting structure M of this construction is a model of the theory S^1_2. We build the set $\mathcal{F}$ using deterministic Turing machines in M which run in polynomial time in M, where the exponents of these polynomials are small nonstandard numbers. For this recall from [9] a formalization of polynomial-time functions in S^1_2. There is a Δ^b_1 formula $\nu_0(e,x,y,z)$ such that if we prove the number $2^{|x|^e}$ exists, $\nu_0(e,x,y,2^{|x|^e})$ says that the Turing machine with code e runs on input x for $|x|^e$ steps and outputs y. In order to simplify notation, we write $\{e\}_1(x)=y$ for $\nu_0(e,x,y,2^{|x|^e})$. Thus, provided a number $b=2^{|x|^v}$ exists in M for some $v\in M$, we can write $\{e\}_1(x)=y$ for $e\le v$ either as a Σ^b_1 formula with parameter b:

$$(\exists z\le b)(z=2^{|x|^e}\wedge \nu_0(e,x,y,z))$$

or equivalently as a Π^b_1 formula with parameter b

$$(\forall z\le b)(z=2^{|x|^e}\to \nu_0(e,x,y,z)),$$

hence it is Δ^b_1.

Theorem 3.1. *Suppose that M is a countable nonstandard model of S^1_2, $n\in M$ is nonstandard, $\psi(x,y)$ is Σ^b_0 in the language L_2 and assume that there is $b\in M$ such that*

$$lh(b)=n^v \tag{3.1}$$
$$(\forall i<lh(b))\,|(b)_i|=n \tag{3.2}$$
$$(\forall i,j<lh(b))\,((b)_i=(b)_j\to i=j) \tag{3.3}$$
$$(\forall i<lh(b))\,(\forall e\le r)\,\neg\psi((b)_i,\{e\}_1((b)_i)), \tag{3.4}$$

where v,r are some nonstandard numbers such that $r\le v$.

Then there is a filter $\mathcal{G}$ on M-definable subsets of

$$\Omega:=\{w\in M\mid(\exists i<lh(b))\;w=(b)_i\}$$

and a set $\mathcal{F}$ of functions $\gamma:\Omega\to M$, $\gamma\in M$, such that $id_\Omega\in\mathcal{F}$ and $\mathcal{F}/\mathcal{G}$ satisfies the following:

$$\mathcal{F}/\mathcal{G}\models S^1_2 \tag{3.5}$$
$$\mathcal{F}/\mathcal{G}\models(\forall y)\neg\psi([id_\Omega],y) \tag{3.6}$$
$$\mathcal{F}/\mathcal{G}\models a\le[c_n]\Rightarrow\mathcal{F}/\mathcal{G}\models a=[c_j]\text{ for some } j\in M, j\le n. \tag{3.7}$$

Proof. We construct the set $\mathcal{F}$ with the help of Turing machines E_k, $k\in\mathbf{N}$, based on one that was used by Wilkie in his model-theoretic proof of Buss's witnessing theorem (unpublished, see [9] for a presentation). First we describe a machine $\hat{E}$ which works on input u and uses a tuple of parameters

$$\langle k,(\exists y\le t(x,\bar{z}))\varphi(x,y,\bar{z}),\bar{d},\bar{d}',\hat{r}\rangle,$$

where $k\in\mathbf{N}$, t is a term and $\varphi(x,y,\bar{z})$ is a Σ^b_0 formula in the language L_2, $\bar{d},\bar{d}'$ are sequences of natural numbers, both of the same length as $lh(\bar{z})$, such that $(\bar{d})_l\le k$ for $l=1,2,\dots,lh(\bar{z})$, and $\hat{r}\in M$. Roughly speaking, the machine uses $\bar{d}, \bar{d}'$ and input u to compute parameters $\bar{a}$ which it assigns to $\bar{z}$ in $(\exists y \leq t(x, \bar{z}))\varphi(x, y, \bar{z})$. It also computes the value of a variable h. Then, using machines with code bounded by $\hat{r}$ as subroutines, it searches for $j < h$ and y such that $y \leq t(j, \bar{a}) \wedge \varphi(j, y, \bar{a})$ is true and such that it cannot find a witness for $y \leq t(j+1, \bar{a}) \wedge \varphi(j+1, y, \bar{a})$. Later we will define machines E_k by specifying for each k the remaining parameters from the above tuple. Then we will be able to clarify our choices of the subroutines to compute $\bar{a}$ and h and to explain its operation in detail. The machine $\hat{E}$ operates as follows:

```
1. $u_0 := u$;
   for $i = 1, 2, \ldots, k$:
     $u_0 := \langle u_0 \rangle_1$ (i.e. project the first coordinate of the pair coded by $u_0$);
2. for $l = 1, 2, \ldots, lh(\bar{z})$:
     $u_l := u$;
     for $i = 1, 2, \ldots, k - (\bar{d})_l$:
       $u_l := \langle u_l \rangle_1$;
3. $j := 0$; $g := u$; $h := |u_0|$; $(\bar{a})_l := \{(\bar{d}')_l\}_1(u_l)$ for $l = 1, 2, \ldots, lh(\bar{z})$;
4. find the first $e \leq \hat{r}$ such that
       $\{e\}_1(g) \leq t(j, \bar{a}) \wedge \varphi(j, \{e\}_1(g), \bar{a})$;
5. if such an $e$ does not exist and $j = 0$, output $\langle u, \langle h+1, g \rangle \rangle$;
6. if such an $e$ does not exist and $j > 0$, output $g$;
7. else put $g := \langle u, \langle j, \{e\}_1(g) \rangle \rangle$;
8. if $j = h$, output $g$;
9. else put $j := j + 1$ and go to 4
```

We shall estimate the code of this Turing machine, $\hat{e}$, considering its parameters as hard-wired in the machine, so $\hat{e}$ will depend on them. Recall the running time of this machine on input u is bounded by $|u|^{\hat{e}}$. The time needed to compute u_0 and $u_1, \ldots, u_{lh(\bar{z})}$ is a standard polynomial in $|u|$. Therefore $|t(j, \bar{a})|$ in step 4 is always bounded by a standard polynomial in $|u|$, say $p(|u|)$. Hence every time the machine succeeds to find e which satisfies the condition in step 4, we have that $|\{e\}_1(g)|$ is bounded by $p(|u|)$. So looking at steps 3 and 7 we see that $|g|$ is always bounded by some standard polynomial $p'(|u|)$. Thanks to this, we can estimate the time we need in step 4 to run some $\{e\}_1$ with $e \leq \hat{r}$ on input g: it is $|g|^e \leq |g|^{\hat{r}} \leq (p'(|u|))^{\hat{r}}$. In step 4 we do this $\hat{r}$ times. The time needed to check the condition in step 4 is a standard polynomial in $|u|$, say $p''(|u|)$. Finally, the cycle containing steps 4 - 9 is repeated at most h times. Thus the total running time is bounded by

$$\left((p'(|u|))^{\hat{r}} + p''(|u|)\right) \cdot \hat{r} \cdot h.$$

Since $h = |u_0|$, it is bounded by a standard polynomial in $|u|$. Thus we can bound the above expression by $|u|^{\hat{r}^2}$. Since the parameter $\hat{r}$ is the only element in the program of the machine which may be nonstandard (otherwise the program is finite), we can bound its code by

$$\hat{e} \leq \hat{r}^2. \tag{3.8}$$

Now we describe the aforementioned machines E_k, $k \in \mathbf{N}$. Let

$$\big\langle (\exists y \leq t_k(x,\bar{z}))\varphi_k(x,y,\bar{z}), \bar{d}'_k, \bar{d}_k \big\rangle,\ k \in \mathbf{N} \tag{3.9}$$

be an enumeration of all tuples where t_k is a term and $\varphi_k(x,y,\bar{z})$ is a Σ^b_0 formula in the language L_2 such that it is logically valid for $x = 0$, and $\bar{d}'_k, \bar{d}_k$ are sequences of natural numbers, both of the same length as $lh(\bar{z})$, such that $(\bar{d}_k)_l \leq k$ for $l = 1, 2, \ldots, lh(\bar{z})$. The machine E_k is defined as $\hat{E}$ with the following parameters: k, the above tuple, and r_k plugged for the parameter $\hat{r}$, where r_k is defined inductively by

$$r_0 := |r|, \quad r_{k+1} := |r_k|.$$

Let e_k be the code of E_k. Put

$$\Omega := \{w \in M \mid (\exists i < lh(b))\ w = (b)_i\}.$$

Note that by (3.2), $\Omega \subseteq \{0,1\}^n$ in M. We will often write $(\exists w \in \Omega)\,\varphi(w)$ for $(\exists i < lh(b))\,\varphi((b)_i)$, where φ is some formula. Define a set of functions $\{\alpha_k : \Omega \to M \mid k \in \mathbf{N}\}$ by

$$\alpha_0(w) := w, \quad \alpha_{k+1}(w) := \{e_k\}_1(\alpha_k(w))$$

and

$$\mathcal{F}' := \{\gamma : \Omega \to M \mid (\exists e, k \in \mathbf{N})(\forall w \in \Omega)\ \gamma(w) = \{e\}_1(\alpha_k(w))\}.$$

With these definitions at hand, we comment on the operation of the machine E_k. Observe that due to steps 5 - 8 its output is always of the form $\langle u, u' \rangle$ where u is its input. But from the definition of α_{k+1} we see that this u is at the same time the output of E_{k-1} (if $k > 0$). Thus after E_k completes step 1, the value of the variable u_0 is equal to the input of the machine E_0, that is, to some $w \in \Omega$. Further, in step 3 we set $h := |u_0|$ and we know that all elements of Ω are of length n. So for every $k \in \mathbf{N}$ the value of the variable h of E_k during its computation of α_{k+1} is n. (The reason why we could not hard-wire n, as we did with the other parameters, is that n is too big.) In steps 2 and 3 the machine E_k computes parameters $\bar{z}$ for the formula $\varphi_k(x,y,\bar{z})$: first it extracts the value of $\alpha_{(\bar{d}_k)_l}$ (with $(\bar{d}_k)_l \leq k$ for $l = 1, 2, \ldots, lh(\bar{z})$) into the variable u_l and then runs $\{(\bar{d}'_k)_l\}_1$ on it. In this way, thanks to our enumeration of parameters (3.9), all tuples of functions from $\mathcal{F}'$ will occur as parameters for all the formulas in question. We will write $\eta_{k,l} : \Omega \to M$ for these parameter functions:

$$\begin{aligned} l_k &:= lh(\bar{d}'_k) = lh(\bar{d}_k) \\ \eta_{k,l}(w) &:= \{(\bar{d}'_k)_l\}_1(\alpha_{(\bar{d}_k)_l}(w)) \text{ for } l = 1, 2, \ldots, l_k \text{ and } k \in \mathbf{N}. \end{aligned} \tag{3.10}$$

In step 4 of E_k running on input $\alpha_k(w)$, the machine searches for a witness of y in $\varphi_k(j, y, \eta_{k,1}(w), \ldots, \eta_{k,l_k}(w))$ repeatedly for $j = 0, 1, 2, \ldots$ until it cannot find it or $j = n$. Since $\varphi_k(0, y, \bar{z})$ is logically valid, E_k never stops at step 5. Observing how its output $\alpha_{k+1}(w)$ is formed, define functions $\iota_k : \Omega \to M$ and $\sigma_k : \Omega \to M$ for $k \in \mathbf{N}$ by

$$\alpha_{k+1}(w) = \langle \alpha_k(w), \langle \iota_k(w), \sigma_k(w) \rangle \rangle.$$

With this notation, we can sum up one of the two E_k's achievements:

$$\begin{aligned} M \models (\forall w \in \Omega)\, \big(&\sigma_k(w) \leq t_k(\iota_k(w), \eta_{k,1}(w), \ldots, \eta_{k,l_k}(w)) \\ &\wedge \varphi_k(\iota_k(w), \sigma_k(w), \eta_{k,1}(w), \ldots, \eta_{k,l_k}(w))\big). \end{aligned} \tag{3.11}$$

The other achievement of E_k, stated in the following claim, explains the choice of r_k in the definition of E_k. First, for all $k \in \mathbf{N}$, set

$$\begin{array}{rcl} \Psi_{-1}(y,w) & := & \neg\psi(w,y) \\ \Psi_k(y,w) & := & \Psi_{k-1}(y,w) \\ & & \wedge \neg\big[\iota_k(w) < n \\ & & \wedge\, y \le t_k(\iota_k(w)+1, \eta_{k,1}(w), \ldots, \eta_{k,l_k}(w)) \\ & & \wedge\, \varphi_k(\iota_k(w)+1, y, \eta_{k,1}(w), \ldots, \eta_{k,l_k}(w))\big]. \end{array}$$

Claim 1. *For $k = -1, 0, 1, 2, \ldots$*

$$M \models (\forall w \in \Omega)(\forall e \le r_k)\, \Psi_k\big(\{e\}_1(\alpha_{k+1}(w)), w\big),$$

where we put $r_{-1} := r$.

The claim is established by induction on k. By assumption (3.4) of the theorem, the claim holds for $k = -1$. Assume it holds for $k-1$ and let $w' \in \Omega$ and $e' \le r_k$. The machine E_k ran on $\alpha_k(w')$ and if $\iota_k(w') < n$, then before it output $\alpha_{k+1}(w')$, it ran in step 4 all $e \le r_k$ on this $\alpha_{k+1}(w')$ and verified that none of them produced a witness for y in

$$\begin{array}{c} y \le t_k(\iota_k(w')+1, \eta_{k,1}(w'), \ldots, \eta_{k,l_k}(w')) \\ \wedge\, \varphi_k(\iota_k(w')+1, y, \eta_{k,1}(w'), \ldots, \eta_{k,l_k}(w')). \end{array}$$

It remains to show that $\{e'\}_1(\alpha_{k+1}(w'))$ satisfies $\Psi_{k-1}(y, w')$. By definition, $\alpha_{k+1}(w') = \{e_k\}_1(\alpha_k(w'))$, and (3.8) gives an estimate $e_k \le r_k^2$. Since we have $e' \le r_k = |r_{k-1}|$, the machine which runs first e_k and then e' has a code which is bounded by r_{k-1}. By the induction hypothesis, $\{e\}_1(\alpha_k(w'))$ satisfies $\Psi_{k-1}(y, w')$ for all $e \le r_{k-1}$. This completes the proof of the claim.

Next, we arrange (3.7) by constructing a suitable filter on the M-definable subsets of Ω. This will be the filter $\mathcal{G}$ from the statement of the theorem. Take some enumeration

$$\gamma_k,\ k \in \mathbf{N} \tag{3.12}$$

of $\mathcal{F}'$. We shall construct elements

$$b = b_0, b_1, b_2, \ldots$$

of M such that for every $k = 1, 2, 3, \ldots$ the following will hold in M:

$$lh(b_k) \ge lh(b_{k-1})/(2n) \tag{3.13}$$

$$(\forall i < lh(b_k))(\exists i' < lh(b_{k-1}))(b_k)_i = (b_{k-1})_{i'} \tag{3.14}$$

$$(\forall i, i' < lh(b_k))((b_k)_i = (b_k)_{i'} \to i = i') \tag{3.15}$$

$$(\forall i < lh(b_k))\, \gamma_{k-1}((b_k)_i) \ge n \vee (\exists j < n)(\forall i < lh(b_k))\, \gamma_{k-1}((b_k)_i) = j. \tag{3.16}$$

Suppose b_k has already been constructed. Working in M, we construct b_{k+1}. First we find a sequence b' all of whose elements are distinct and are exactly those elements $(b_k)_i$ of b_k such that $\gamma_k((b_k)_i) \ge n$. This is done using Σ^b_1−LIND and relies on the fact that the required properties of b' are described by Δ^b_1 formulas. If $lh(b') \ge lh(b_k)/2$, put $b_{k+1} = b'$. If not, we similarly create a sequence b'' with $lh(b'') \ge lh(b_k)/2$ of those elements of b_k on which γ_k is smaller than n, but this time we rearrange them into blocks according to the values of γ_k: each block will be determined by two indices, $i_{min} \leq i_{max}$, such that γ_k gives the same value on the elements $(b'')_i$ with $i_{min} \leq i \leq i_{max}$ and γ_k does not give this value to any of the elements $(b'')_i$ with $i < i_{min}$ or $i_{max} < i$. Then we use Δ^b_1−LIND to show that there exists a block for which $i_{max} - i_{min} + 1 \geq lh(b_{k-1})/(2n)$. We define b_{k+1} to be the sequence consisting of this block. Obviously, b_{k+1} satisfies (3.13) - (3.16).

For $k \in \mathbf{N}$ define

$$\Omega_k := \{w \in M \mid (\exists i < lh(b_k))\, w = (b_k)_i\}$$

(so $\Omega_0 = \Omega$) and let $\mathcal{G}$ be the filter generated by the chain $\Omega_0 \supseteq \Omega_1 \supseteq \ldots$.

Now we have everything we need to form the restricted reduced power $\mathcal{F}'/\mathcal{G}$, as defined in Section 3.2. But first we must verify that $\mathcal{F}'$ is closed under the functions given by function symbols of L_2. Let $\delta_1, \ldots, \delta_m \in \mathcal{F}'$. Recall that for $k' < k \in \mathbf{N}$, $\alpha_{k'}$ can be extracted from α_k by a standard polynomial-time algorithm. From this and from the definition of $\mathcal{F}'$ it follows that for large enough $k \in \mathbf{N}$ each of $\delta_1, \ldots, \delta_m$ can be computed by a standard polynomial-time algorithm from α_k. Each function given by a function symbol F of our language is computable by a standard polynomial-time algorithm. Thus we can compute $F(\delta_1, \ldots, \delta_m)$ by a standard polynomial-time algorithm from α_k, so it is an element of $\mathcal{F}'$.

We derive Łoś's theorem for Σ^b_0 formulas.

Claim 2. *Assume that $\varphi(z_1, \ldots, z_m)$ is a Σ^b_0 formula in the language L_2 and $\delta_1, \ldots, \delta_m \in \mathcal{F}'$. Then*

$$\mathcal{F}'/\mathcal{G} \models \varphi([\delta_1], \ldots, [\delta_m]) \;\textit{ if and only if }\; [\![\varphi(\delta_1, \ldots, \delta_m)]\!] \in \mathcal{G}.$$

Quickly from the definition of $\mathcal{F}'/\mathcal{G}$ follows the claim for the case of φ of the form $t(z_1, \ldots, z_m) = z_0$ (use induction on the complexity of the term t) and then for the case of φ atomic. Then we proceed by induction on the complexity of φ. The induction step for conjunction uses the fact that $\mathcal{G}$ is a filter. This fact also gives the right-to-left implication in the induction step for negation. For the left-to-right implication assume that $\mathcal{F}'/\mathcal{G} \models \neg\varphi([\delta_1], \ldots, [\delta_m])$ and that the claim holds for $\varphi(z_1, \ldots, z_m)$. Let $k \in \mathbf{N}$ be so large that $\delta_1, \ldots, \delta_m$ can be computed from α_k by standard polynomial-time algorithms. Since φ is Σ^b_0, there exists a standard polynomial-time algorithm which on input $\alpha_k(w)$ returns 0 or 1 depending on the truth value of $\varphi(\delta_1(w), \ldots, \delta_m(w))$ in M for $w \in \Omega$. This means that the function $\gamma : \Omega \to \{0, 1\}$ given by α_k composed with this algorithm is an element of $\mathcal{F}'$. So it must appear in the enumeration (3.12) as $\gamma_{k'}$ for some $k' \in \mathbf{N}$. There we constructed $b_{k'+1}$ to satisfy (3.13) - (3.16). Thus, what we know about γ, together with (3.16) and the definition of $\Omega_{k'+1}$, give that either $M \models (\forall w \in \Omega_{k'+1})\, \varphi(\delta_1(w), \ldots, \delta_m(w))$ or $M \models (\forall w \in \Omega_{k'+1})\, \neg\varphi(\delta_1(w), \ldots, \delta_m(w))$. Since the first case leads by our induction hypothesis to a contradiction, the proof of the left-to-right implication in the induction step for negation is finished. It remains to prove the right-to-left implication in the induction step for sharply bounded existential quantifier (the other half of the equivalence is immediate). Assume that for some $k \in \mathbf{N}$, $[\![(\exists y \leq |t|)\varphi(y, \delta_1, \ldots, \delta_m)]\!] \supseteq \Omega_k$. Similarly as before, there is a standard polynomial-time algorithm which runs on $\alpha_{k'}$ for a suitable $k' \in \mathbf{N}$, exhaustively searches for a witness for y in the above formula and outputs this witness if it exists. The composition of $\alpha_{k'}$ with this algorithm is then a function from $\mathcal{F}'$, let us denote it δ. We have $[\![\varphi(\delta, \delta_1, \dots, \delta_m)]\!] \supseteq \Omega_k$ and we can apply the induction hypothesis. This finishes the proof of the claim.

It follows from the claim that $\mathcal{F}'/\mathcal{G}$ is a model of *BASIC*. The next claim asserts that induction holds in $\mathcal{F}'/\mathcal{G}$ up to $[c_n]$ for *strict*Σ^b_1 formulas in the language L_2. Recall that we denoted by c_j, $j \leq n$, the constant function on Ω with value j.

Claim 3. *Suppose that $\delta_1, \dots, \delta_m \in \mathcal{F}'$ and let $\Phi(x)$ be the L_2-formula*

$$(\exists y \leq t(x, [\delta_1], \dots, [\delta_m]))\, \varphi(x, y, [\delta_1], \dots, [\delta_m]),$$

where φ is a Σ^b_0 formula and t is a term. Then

$$\mathcal{F}'/\mathcal{G} \models \neg\Phi([c_0]) \vee (\exists x < [c_n])\big((\Phi(x) \wedge \neg\Phi(S(x))\big) \vee \Phi([c_n]).$$

Without loss of generality, we can assume that $\varphi(0, y, \bar{z})$ is logically valid. Take $k \in \mathbf{N}$ such that the k-th tuple of our enumeration (3.9) contains the formula $(\exists y \leq t(x, \bar{z}))\, \varphi(x, y, \bar{z})$ together with tuples of standard numbers $\bar{d'_k}, \bar{d_k}$ which tell E_k how to compute $\delta_1, \dots, \delta_m$ from α_k, i.e.

$$\delta_l = \eta_{k,l} \text{ for } l = 1, 2, \dots m,$$

where $\eta_{k,l}$ is given by (3.10). Since $\iota_k, \sigma_k \in \mathcal{F}'$, we can apply Łoś's theorem for Σ^b_0 formulas (Claim 2) to (3.11) and get

$$\mathcal{F}'/\mathcal{G} \models \Phi([\iota_k]).$$

Suppose that $[\iota_k] < [c_n]$, for otherwise we are done. Assume for a contradiction that $\mathcal{F}'/\mathcal{G} \models \Phi(S([\iota_k]))$ and let $[\gamma]$ with $\gamma \in \mathcal{F}'$ be a witness of the leading existential quantifier in Φ. There exists some $k' \in \mathbf{N}$, $k' \geq k$, and $e' \in \mathbf{N}$, such that $(\forall w \in \Omega)\, \gamma(w) = \{e'\}_1(\alpha_{k'+1}(w))$. This follows from the definition of $\mathcal{F}'$ and from the existence of a standard polynomial-time algorithm that extracts $\alpha_{k''}$ from $\alpha_{k''+1}$ for every $k'' \in \mathbf{N}$. Claim 1 implies that

$$\begin{aligned} M \models (\forall w \in \Omega)(\forall e \leq r_{k'})\neg\big[\iota_k(w) < n \\ \wedge\ \{e\}_1(\alpha_{k'+1}(w)) \leq t(\iota_k(w) + 1, \delta_1(w), \dots, \delta_m(w)) \\ \wedge\ \varphi(\iota_k(w) + 1, \{e\}_1(\alpha_{k'+1}(w)), \delta_1(w), \dots, \delta_m(w))\big]. \end{aligned}$$

We have $e' < r_{k'}$, since $r_{k'}$ is nonstandard. Since $\mathcal{F}'/\mathcal{G} \models [\iota_k] < [c_n]$, Claim 2 gives

$$\mathcal{F}'/\mathcal{G} \models \neg\big([\gamma] \leq t(S([\iota_k]), [\delta_1], \dots, [\delta_m]) \wedge \varphi(S([\iota_k]), [\gamma], [\delta_1], \dots, [\delta_m])\big).$$

This contradicts our choice of $[\gamma]$ and finishes the proof of the claim.

Note that a consequence of Claim 1 is that for $k = -1, 0, 1, 2, \dots$

$$M \models (\forall w \in \Omega)(\forall e \leq r_k)\, \neg\psi(\alpha_0(w), \{e\}_1(\alpha_{k+1}(w))).$$

With the help of Claim 2, this gives

$$\mathcal{F}'/\mathcal{G} \models (\forall y)\neg\psi([id_\Omega], y). \tag{3.17}$$

Now we define a substructure $\mathcal{F}/\mathcal{G}$ of $\mathcal{F}'/\mathcal{G}$ by setting

$$\mathcal{F} := \{\delta \in \mathcal{F}' \mid \mathcal{F}'/\mathcal{G} \models |[\delta]| \leq [c_n]^k \text{ for some } k \in \mathbf{N}\},$$

i.e., $\mathcal{F}/\mathcal{G}$ is the restriction of $\mathcal{F}'/\mathcal{G}$ to the set of elements with length bounded by a standard power of $[c_n]$. Bounded formulas are absolute between $\mathcal{F}/\mathcal{G}$ and $\mathcal{F}'/\mathcal{G}$. Hence Claim 2 (Łoś's theorem for Σ^b_0 formulas), Claim 3 (induction up to $[c_n]$ for *strict*Σ^b_1 formulas) and (3.17) remain valid with $\mathcal{F}'$ replaced by $\mathcal{F}$.

It remains to show that $\mathcal{F}/\mathcal{G}$ satisfies Σ^b_1-LIND. Since we have induction for *strict*Σ^b_1 formulas up to $[c_n]$, we also have it up to $[c_n]^k$ for all $k \in \mathbf{N}$, i.e., we have

$$\mathcal{F}/\mathcal{G} \models \mathit{strict}\Sigma^b_1 - \text{LIND}.$$

Now *strict*Σ^b_1-LIND is sufficient to prove $BB\Sigma^b_0$, the sharply bounded collection scheme for Σ^b_0 formulas. This scheme allows us to prove that every Σ^b_1 formula is equivalent to a *strict*Σ^b_1 formula. Hence $\mathcal{F}/\mathcal{G} \models S^1_2$. □

As a corollary we get a weaker version of Buss's witnessing theorem:

Corollary 3.2. *Let $\psi(x, y)$ be Σ^b_1 in the language L_2 and suppose that*

$$S^1_2 \vdash (\forall x)(\exists y)\psi(x, y).$$

Then for some $q \in \mathbf{N}$ and $f \in \square^p_1$,

$$\begin{aligned} S^1_2 \vdash (\forall x)(\forall z)\big(lh(x) > z^q \wedge (\forall i < lh(x))|(x)_i| = z \\ \wedge (\forall i, j < lh(x))((x)_i = (x)_j \to i = j) \\ \to (\exists i < lh(x))\psi((x)_i, f((x)_i))\big). \end{aligned}$$

Viewed in the standard model $\mathbf{N}$, the conclusion of the theorem says that for a randomly chosen x of length z the function f witnesses $(\exists y)\psi(x, y)$ with the probability of at least $1 - (z^q - 1)2^{1-z}$. That is, f makes $z^q - 1$ possible errors on elements of length z.

Proof. We can assume that $\psi(x, y)$ is Σ^b_0, since it is equivalent in S^1_2 with its strict form and bounded existential quantifiers can be merged with $(\exists y)$ using the pairing function. Assume that the conclusion of the corollary is false. Let T be the following theory with b, n new constants:

$$\begin{aligned} & S^1_2 + \{lh(b) > n^k \mid k \in \mathbf{N}\} \\ & \quad + (\forall i < lh(b))|(b)_i| = n \\ & \quad + (\forall i, j < lh(b))((b)_i = (b)_j \to i = j) \\ & \quad + \{(\forall i < lh(b))(\forall e < k)\neg\psi((b)_i, \{e\}_1((b)_i)) \mid k \in \mathbf{N}\}. \end{aligned}$$

T is consistent since otherwise in its finite contradictory fragment there would appear some maximal k and we would combine all the polynomial-time machines $e < k$ into one that using ψ as an oracle produces a witness for one of the $(b)_i$'s, contradicting our assumption.

So there exists a countable $M \models T$. In the case that $n = 0$ (hence $lh(b) = 1$ and $(b)_0 = 0$) we have that no standard polynomial-time machine witnesses $\psi(0, y)$. By Parikh's theorem applied to our assumption $S^1_2 \vdash (\forall x)(\exists y)\psi(x, y)$, there is a standard y satisfying $\psi(0, y)$ in M. Therefore this y can be trivially produced by a standard polynomial-time machine, a contradiction. So assume that $n > 0$. Then from the first three lines of T it follows that b and n are nonstandard (in the case $n = 1$ note that there is only one number whose length is 1). By overspill, using the first line of T, we have that $lh(b)$ is at least n^v for some nonstandard $v \in M$, and we may assume $lh(b) = n^v$ with $v \leq n$, since such a shortening of our b still satisfies the axioms of T.

Since $n^v = lh(b) \leq |b|$, 2^{n^v} exists in M. The last line of T says that the following formula holds in M for every standard r:

$$(\forall i < lh(b))(\forall e \leq |v|)\big(e \leq r \to \neg\psi((b)_i, \{e\}_1((b)_i))\big).$$

By overspill, there is a nonstandard $r \in M$ such that the formula holds in M.

Now the model M satisfies the assumptions of Theorem 3.1, which gives a model $\mathcal{F}/\mathcal{G} \models S_2^1 + (\exists x)(\forall y)\neg\psi(x, y)$. This contradicts our assumption. □

It was noted by L. Kołodziejczyk that the corollary gives, in fact, a polynomial-time function which witnesses $(\exists y)\psi(x, y)$ without any errors, as is the case in Buss's witnessing theorem.

Theorem 3.3. *Let $\psi(x, y)$ be Σ_1^b in the language L_2 and suppose that*

$$S_2^1 \vdash (\forall x)(\exists y)\psi(x, y).$$

Then for some $f \in \Box_1^p$,

$$S_2^1 \vdash (\forall x)\psi(x, f(x)).$$

Now the difference between this theorem and Buss's witnessing theorem for S_2^1 is that the witnessing in the latter is provable in the theory PV_1. This difference can be eliminated by checking that relevant arguments in proofs of this section can be carried out in PV_1.

Proof. As before, we can assume that $\psi(x, y)$ is Σ_0^b. Let $\psi'(u, y)$ be the formula

$$\psi(MSP(u, ||u||^2), y).$$

Since $S_2^1 \vdash (\forall u)(\exists y)\psi'(u, y)$, Corollary 3.2 gives $q \in \mathbf{N}$ and $f' \in \Box_1^p$ such that

$$\begin{aligned} S_2^1 \vdash (\forall w)(\forall z)\big(lh(w) > z^q \wedge (\forall i < lh(w))|(w)_i| = z \\ \wedge (\forall i, j < lh(w))((w)_i = (w)_j \to i = j) \\ \to (\exists i < lh(w))\psi'((w)_i, f'((w)_i))\big). \end{aligned}$$

We describe $f \in \Box_1^p$ which uses f' as a subroutine and satisfies the conclusion of the theorem. It works on input x. First it finds z such that $z - |z|^2 = |x|$ by a very simple polynomial-time algorithm (using $|x| < z \leq 2|x|$ for $|x| \geq 64$). Then it produces numbers

$$x \cdot 2^{|z|^2}, x \cdot 2^{|z|^2} + 1, x \cdot 2^{|z|^2} + 2, \ldots, x \cdot 2^{|z|^2} + z^q.$$

This is $z^q + 1$ distinct numbers all of length z and all having as first $|x|$ bits the binary representation of x, provided that $2^{|z|^2} > z^q$. The last inequality is satisfied for every $z \geq 2^q$, therefore we will assume that $|x| \geq 2^q$. Finally, f' is run on each u from these $z^q + 1$ numbers and it is checked whether $\psi'(u, f'(u))$ is true. It has to be true for at least one u by the property of f' above, and $f'(u)$ for the first such u is the output of f. □

We conclude this section with remarks on the disadvantages of Construction A. Firstly, the assumptions (3.1) - (3.4) of Theorem 3.1 are very restrictive. This is connected with the fact that in the proof of the theorem we are using quite a strong computational model (nonstandard deterministic Turing machines running in nonstandard polynomial time) in order to ensure induction. So, in the next construction we will want to work with a weaker computational model and we will have to find a different way to arrange induction. Secondly, despite the property (3.7) of the theorem, the construction cannot guarantee that all elements $[c_j]$, $M \models j \leq n$, are present in $\mathcal{F}/\mathcal{G}$. But this is usually required in the reformulations of statements of complexity theory mentioned in the introduction.

3.4 Construction B

Assume that

- M is a countable nonstandard model of true arithmetic
- $n \in M \setminus \mathbf{N}$
- $\Omega \subseteq \{0,1\}^n$
- $\mathcal{L}$ is a first-order language whose non-logical symbols are a binary relation $\leq$ and a finite set of function symbols $F_{\mathcal{L}}$ containing a unary symbol S and constant symbols $0, \tilde{n}$. The intended interpretation of each of these function symbols is some function definable in M, with S interpreted by the successor function, $\tilde{n}$ by n, and 0 and $\leq$ interpreted as usual.
- $\psi(x, y, z_1, \ldots, z_{k_0})$ is a $\Delta_0^{<\tilde{n}}$ formula in the language $\mathcal{L}$, i.e., each of its quantifiers is of the form $(\exists w < \tilde{n})$ or $(\forall w < \tilde{n})$.
- $X \in M$ is a set containing the following functions from Ω to M: the identity function id_Ω, for each $v \leq n$ the constant function c_v with value v, and some functions $h_1(x), \ldots, h_{k_0}(x) \in M$.

We define a *straight-line program* (SLP) over $F_{\mathcal{L}}$ of size t with input variables $x_1, \ldots, x_k$ to be a sequence of instructions of the following form: the ith instruction ($i = 1, \ldots, t$) applies a function from $F_{\mathcal{L}}$ to some of the input variables or previously assigned variables $y_1, \ldots, y_{i-1}$ and assigns the outcome of this operation to y_i. Given an assignment of the input variables, the output of the program is the value of y_t. The size of an SLP P will be denoted by $size(P)$.

We will consider SLPs inside M. Let $F_{\mathcal{L}}$ and X be as above and let $P \in M$ be an SLP over $F_{\mathcal{L}}$ of size $t \in M$ with input variables of the form x_g, $g \in X$. We define $Fct_X(P)$ to be the set consisting of all functions $f : \Omega \to M$ where either $f \in X$ or there is some $i = 1, \ldots, t$ such that $f(u)$ is the value at y_i computed by P with the following assignment of the input variables: the value of x_g is $g(u)$.

Fix a parameter $m \in M$, a nonstandard number with $m < n$, and a parameter q, which is a rational number in M and satisfies $0 < q < 1$.

Assume further the following hypothesis:

> *There is a nonstandard $s \in M$ such that for every SLP $P \in M$ of size m^s and every $f \in Fct_X(P)$:*
> $$\Pr_{u \in \Omega}[\psi(u, f(u), h_1(u), \ldots, h_{k_0}(u))] < q. \tag{H}$$

In M define the *master tree* to be the tree of SLPs of size $\leq m^s$. The empty SLP is the root of the master tree and the partial order of the tree is defined by "P is an initial part of Q". For $\ell \geq 0$ put $FCT(\ell) := \bigcup_P Fct_X(P)$, P ranging over SLPs from the master tree of size $\leq \ell$. Denote $FCT := FCT(m^s)$.

For two sets A, B we abbreviate the quantity $\frac{\text{card}(A \cap B)}{\text{card}(B)}$ by $\mu(A/B)$ and call it the *measure of A with respect to B*.

Hence the hypothesis (H) means that

$$\mu([\![\psi(id_\Omega, f, h_1, \ldots, h_{k_0})]\!]/\Omega) < q$$

for all $f \in FCT$.

Our construction also requires the following relation between the parameters n, m and q:

$$q^{1/m^i} < \frac{1}{n} \quad \text{for every } i \in \mathbf{N}. \tag{R}$$

This forces Ω to be large with respect to $\{0,1\}^n$.

Having stated the assumptions of the upcoming construction, we summarize its desired properties in a theorem.

Theorem 3.4. *Let $M, n, \Omega, \mathcal{L}, \psi, X$ be as above. Assume that the parameters m, q as above satisfy the hypothesis* (H) *and the relation (R).*

Then there is $\mathcal{F} \subseteq FCT$ and a filter $\mathcal{G}$ on the M-definable subsets of Ω such that the restricted reduced power $\mathcal{F}/\mathcal{G}$ enjoys the following properties:

(FG 1) *$X \subseteq \mathcal{F}$ and $\mathcal{F}$ is closed under the functions from $F_\mathcal{L}$.*

(FG 2) *$\mathcal{F}/\mathcal{G}$ contains no new lengths $\leq n$, i.e., if $f \in \mathcal{F}$ and $\mathcal{F}/\mathcal{G} \models [f] \leq [c_n]$, then there is $v \leq n$ in M such that $\mathcal{F}/\mathcal{G} \models [f] = [c_v]$.*

(FG 3) *Łoś's theorem holds for $\Delta_0^{<\tilde{n}}$ formulas, i.e., for functions $f_1, \ldots, f_k \in \mathcal{F}$ and for an $\mathcal{L}$-formula $\varphi(x_1, \ldots, x_k)$ which is $\Delta_0^{<\tilde{n}}$, we have*

$$\mathcal{F}/\mathcal{G} \models \varphi([f_1], \ldots, [f_k]) \quad \textit{iff} \quad [\![\varphi(f_1, \ldots, f_k)]\!] \in \mathcal{G}.$$

(FG 4) $\mathcal{F}/\mathcal{G} \models (\forall y) \neg\psi([id_\Omega], y, [h_1], \ldots, [h_{k_0}])$

(FG 5) *Let $\varphi(x)$ be an $\mathcal{L}$-formula of the form*

$$(\exists y_1) \ldots (\exists y_{k'}) \alpha$$

such that $\alpha(x, y_1, \ldots, y_{k'}, [g_1], \ldots, [g_k])$ is $\Delta_0^{<\tilde{n}}$ and has $[g_1], \ldots, [g_k]$ as parameters, where $g_1, \ldots, g_k \in \mathcal{F}$. Then

$$\mathcal{F}/\mathcal{G} \models \neg\varphi([c_0]) \vee \varphi([c_m]) \vee (\exists x < [c_m])\big(\varphi(x) \wedge \neg\varphi(S(x))\big),$$

i.e., in $\mathcal{F}/\mathcal{G}$ induction for φ holds up to m.

Proof. In stages $i = 0, 1, 2, \dots$ we shall construct the following elements of M:

$$P_i, \ell_i, q_i, U_i, Bad_i,$$

where

(PU 1) P_i is an SLP from the master tree, P_0 is the root and P_{i+1} extends P_i.

(PU 2) $\ell_i \geq 0$ is a number such that $\ell_0 := m^s$ and $\ell_0 \geq \ell_1 \geq \ell_2 \geq \dots$ with $\ell_i - size(P_i) \geq m^{s-2i}$.

(PU 3) q_i is a rational number (in M) called the ith *badness coefficient.* $q_0 := q$ and $q_{i+1} := q_i^{1/(2m)}$. (So $0 < q_0 < q_1 < \dots < 1$).

(PU 4) U_i is a subset of Ω, such that $U_0 := \Omega$ and $U_0 \supseteq U_1 \supseteq U_2 \supseteq \dots$.

(PU 5) Bad_i is a collection of subsets of Ω with

$$Bad_0 := \{[\![\psi(id_\Omega, f, h_1, \dots, h_{k_0})]\!] \mid f \in FCT\} \cup \{\emptyset\}$$

and $Bad_0 \subseteq Bad_1 \subseteq Bad_2 \subseteq \dots$, such that for all $i \in \mathbf{N}$ and $V \in Bad_i$, we have $\mu(V/U_i) < q_i$. (Note that this holds for $i = 0$ by (H)).

We shall form the restricted reduced power $\mathcal{F}/\mathcal{G}$ from $\mathcal{F} := \bigcup_{i \in \mathbf{N}} Fct_X(P_i)$ and the filter $\mathcal{G}$ generated by the chain $U_0 \supseteq U_1 \supseteq \dots$.

We now construct P_i, ℓ_i, U_i, Bad_i. Enumerate all tuples of the form

$$(F, Y, \varphi, g_1, \dots, g_k),$$

where F is a symbol from $F_{\mathcal{L}}$, Y is a subset of Ω which is in M, φ is an $\mathcal{L}$-formula, $k \in \mathbf{N}$ and $g_1, \dots, g_k \in FCT$, such that every tuple appears infinitely many times in the enumeration. At each stage $i = 0, 1, 2, \dots$ we apply one of the five steps below to the $\lfloor i/5 \rfloor$-th tuple from the enumeration, such that the number of the applied step, (i) - (v), is congruent to i mod 5. In this way each of these steps deals with every tuple from the enumeration at infinitely many stages.

(i) For $F, g_1, \dots, g_k$ in the tuple under consideration, check whether k is the arity of F and $g_1, \dots, g_k$ are in $Fct_X(P_i)$. If so, let $t = size(P_i)$ and extend P_i by one instruction $y_{t+1} := F(w_1, \dots, w_k)$, where w_j ($j = 1, \dots, k$) is a variable (input or previously assigned) at which g_j is computed. Define P_{i+1} to be the resulting SLP. Otherwise, put $P_{i+1} := P_i$. Leave ℓ_i, U_i, Bad_i unchanged, i.e., set $\ell_{i+1} := \ell_i, U_{i+1} := U_i, Bad_{i+1} := Bad_i$. Obviously, (PU 1) - (PU 5) are satisfied for $i + 1$.

(ii) If the set Y from the tuple under consideration is an element of Bad_i, put $U_{i+1} := U_i \setminus Y$. Otherwise, $U_{i+1} := U_i$. Leave P_i, ℓ_i, Bad_i unchanged. To verify (PU 5) for $i + 1$ we use that $q_i < 1/2 \ \forall i \in \mathbf{N}$, which is a consequence of the assupmtion (R), to estimate

$$\frac{q_i}{1 - q_i} < 2q_i < \frac{1}{q_{i+1}} q_i < q^{\frac{2m-1}{(2m)^{i+1}}} < q^{\frac{1}{(2m)^{i+1}}} = q_{i+1}.$$

So the ratio $\mu(V/U_{i+1})$ from (PU 5) is less than q_{i+1} as required.

(iii) In this step, which is useful for (FG 2) and (FG 3), we make the filter decide whether a $\Delta_0^{<\tilde{n}}$ formula or its negation will hold in $\mathcal{F}/\mathcal{G}$. Let the formula in the considered tuple be a $\Delta_0^{<\tilde{n}}$ formula of the form $\varphi(x_1, \ldots, x_k)$. (Otherwise, leave P_i, ℓ_i, U_i, Bad_i unchanged). Set $W := [\![\varphi(g_1, \ldots, g_k)]\!]$. If

$$\mu(W/U_i) < 1/2,$$

put $U_{i+1} := U_i \setminus W$. Otherwise, $U_{i+1} := U_i \cap W$. Leave P_i, ℓ_i, Bad_i unchanged. To check (PU 5) for $i+1$, we need $2q_i < q_{i+1}$, which was verified in step (ii).

(iv) This step will help ensure (FG 2) and (FG 3) by witnessing a quantifier bounded by $\tilde{n}$. Let the formula φ in the tuple under consideration be of the form $(\exists x < \tilde{n})\varphi'(x, y_1, \ldots, y_k)$, where φ' is $\Delta_0^{<\tilde{n}}$, and assume that $U_i \subseteq [\![\varphi(g_1, \ldots, g_k)]\!]$. From the covering

$$[\![\varphi'(c_0, g_1, \ldots, g_k)]\!], [\![\varphi'(c_1, g_1, \ldots, g_k)]\!], \ldots, [\![\varphi'(c_n, g_1, \ldots, g_k)]\!]$$

of U_i pick a set Z with the largest measure with respect to U_i. Define $U_{i+1} := U_i \cap Z$. Leave P_i, ℓ_i, Bad_i unchanged. Here the new ratio $\mu(V/U_{i+1})$ from (PU 5) is at most $(n+1)\mu(V/U_i) < (n+1)q_i$. Using the assumption (R) we have

$$nq_i + q_i < \frac{q_i}{q_{i+1}} + q_i = q^{\frac{2m-1}{(2m)^{i+1}}} + q^{\frac{1}{(2m)^i}} < 2q^{\frac{2m-1}{(2m)^{i+1}}} < q^{\frac{2m-2}{(2m)^{i+1}}} < q_{i+1}$$

so (PU 5) remains valid for $i+1$.

(v) The goal of this step is to arrange (FG 5) by binary search. Let the formula φ from the tuple under consideration be of the form

$$(\exists y_1) \ldots (\exists y_{k'})\alpha(x, y_1, \ldots, y_{k'}, z_1, \ldots, z_k)$$

where α is $\Delta_0^{<\tilde{n}}$. Moreover, let

$$\begin{aligned} [\![\alpha(c_0, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] &= \Omega \\ \text{and } [\![\alpha(c_m, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] &= \emptyset \end{aligned}$$

for all $f_1, \ldots, f_{k'} \in FCT$.

In steps $j = 0, 1, \ldots, r := \lceil \log m \rceil$ we shall construct inductively in M elements

$$\widetilde{P}_j, \widetilde{\ell}_j, \widetilde{U}_j, \widetilde{Bad}_j, u_j, v_j,$$

such that

- $\widetilde{P}_0 := P_i$, $\widetilde{P}_{j+1}$ extends $\widetilde{P}_j$.
- $\widetilde{\ell}_0 := \ell_i$, $\widetilde{\ell}_j \geq \widetilde{\ell}_{j+1}$ and $\widetilde{\ell}_{j+1} - size(\widetilde{P}_{j+1}) \geq \lfloor (\widetilde{\ell}_j - size(\widetilde{P}_j))/2 \rfloor$.
- $\widetilde{U}_0 := U_i$, $\widetilde{U}_{j+1} \subseteq \widetilde{U}_j$.
- $\widetilde{Bad}_0 := Bad_i$, $\widetilde{Bad}_{j+1} \supseteq \widetilde{Bad}_j$ and for all $j = 0, 1, \ldots, r$ and all $V \in \widetilde{Bad}_j$ we have
$$\mu(V/\widetilde{U}_j) < q_i^{1/2^j}.$$

- $u_0 := 0$, $v_0 := m$ and $v_{j+1} - u_{j+1} \leq \lceil (v_j - u_j)/2 \rceil$.
- For each $j = 0, 1, \ldots, r$ there are $f_1, \ldots, f_{k'} \in Fct_X(\widetilde{P}_j)$ such that
$$[\![\alpha(c_{u_j}, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] \supseteq \widetilde{U}_j.$$
- For each $j = 0, 1, \ldots, r$ and for all $g'_1, \ldots, g'_{k'} \in Fct_X(Q)$, where Q extends $\widetilde{P}_j$ and $size(Q) = \widetilde{\ell}_j$, we have
$$[\![\alpha(c_{v_j}, g'_1, \ldots, g'_{k'}, g_1, \ldots, g_k)]\!] \in \widetilde{Bad}_j.$$

Note that everything is satisfied for $j = 0$ by (PU 5) and by our assumptions. Assume that for some $j < r$ these elements have been constructed. Set $\ell := \lfloor (\widetilde{\ell}_j + size(\widetilde{P}_j))/2 \rfloor$, $w := \lfloor (v_j + u_j)/2 \rfloor$ and

$$\begin{aligned} \mathcal{V} := \{[\![\alpha(c_w, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] \mid & \text{ there is an SLP } P \text{ such that} \\ & f_1, \ldots, f_{k'} \in Fct_X(P), P \text{ extends } \widetilde{P}_j \text{ and } size(P) = \ell\}. \end{aligned}$$

If there is some $V \in \mathcal{V}$ such that

$$\mu(V \,/\, \widetilde{U}_j) \geq q_i^{1/2^{j+1}},$$

then put

$$\begin{aligned} u_{j+1} := w,\ v_{j+1} := v_j,\ \widetilde{\ell}_{j+1} := \widetilde{\ell}_j,\ \widetilde{P}_{j+1} := P, \\ \widetilde{U}_{j+1} := \widetilde{U}_j \cap V,\ \widetilde{Bad}_{j+1} := \widetilde{Bad}_j, \end{aligned}$$

where P is any SLP witnessing that $V \in \mathcal{V}$. Otherwise, put

$$\begin{aligned} u_{j+1} := u_j,\ v_{j+1} := w,\ \widetilde{\ell}_{j+1} := \ell,\ \widetilde{P}_{j+1} := \widetilde{P}_j, \\ \widetilde{U}_{j+1} := \widetilde{U}_j,\ \widetilde{Bad}_{j+1} := \widetilde{Bad}_j \cup \mathcal{V}. \end{aligned}$$

It is easy to check that in both cases the newly constructed elements satisfy all the requirements and that $u_r + 1 = v_r$. Define

$$P_{i+1} := \widetilde{P}_r,\ \ell_{i+1} := \widetilde{\ell}_r,\ U_{i+1} := \widetilde{U}_r,\ Bad_{i+1} := \widetilde{Bad}_r.$$

and note that (PU 1) - (PU 5) are satisfied. In particular, we can estimate

$$size(P_{i+1}) - \ell_{i+1} > \frac{size(P_i) - \ell_i}{2^r} - 1 > \frac{size(P_i) - \ell_i}{2m} \geq m^{s-2i-2}$$

and for any $V \in Bad_{i+1}$ we have $\mu(V/U_{i+1}) < q_i^{1/2^r} < q_i^{1/(2m)} = q_{i+1}$.
Thus, in this step we have found $v := u_r \in \{0, 1, \ldots, m-1\}$, such that

$$[\![\alpha(c_v, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] \supseteq U_{i+1}$$

for some $f_1, \ldots, f_{k'} \in Fct_X(P_{i+1})$, and for all $g'_1, \ldots, g'_{k'} \in Fct_X(Q)$, where Q extends P_{i+1} and $size(Q) = \ell_{i+1}$,

$$[\![\alpha(c_{v+1}, g'_1, \ldots, g'_{k'}, g_1, \ldots, g_k)]\!] \in Bad_{i+1}.$$

Now that U_i and P_i have been defined for all $i \in \mathbf{N}$, we can verify (FG 1) - (FG 5). Recall that we put $\mathcal{F} := \bigcup_{i\in\mathbf{N}} Fct_X(P_i)$ and $\mathcal{G}$ is the filter generated by the chain $U_0 \supseteq U_1 \supseteq \ldots$.

Step (i) and the definition of $\mathcal{F}$ guarantee that $\mathcal{F}$ is closed under the functions from $F_{\mathcal{L}}$ and that $X \subseteq \mathcal{F}$, i.e., (FG 1) holds.

To check (FG 2), let $f \in \mathcal{F}$ and $\mathcal{F}/\mathcal{G} \models [f] \leq [c_n]$. At some stage $i \in \mathbf{N}$ the formula

$$\nu(y) := (\exists x < \tilde{n})\, x = y$$

will be treated by step (iii) together with f, with the result $U_{i+1} \subseteq [\![\nu(f)]\!]$. (Otherwise $U_{i+1} \subseteq [\![\neg\nu(f)]\!]$, which implies $U_{i+1} \subseteq [\![\neg\, f \leq \tilde{n}]\!]$, contradicting $\mathcal{F}/\mathcal{G} \models [f] \leq [c_n]$). By step (iv) (applied to the same formula at some stage $i' > i$) we have $U_{i'+1} \subseteq [\![c_v = f]\!]$ for some $v \leq n$ in M. Thus, $\mathcal{F}/\mathcal{G} \models [c_v] = [f]$ as required.

We verify (FG 3), Łoś's theorem for $\Delta_0^{<\tilde{n}}$ formulas, by induction on the complexity of the formula φ. By definition, it holds for φ atomic. To prove the inductin step for negation, assume that φ is $\neg\varphi'(x_1, \ldots, x_k)$, $f_1, \ldots, f_k \in \mathcal{F}$ and that the assertion holds for φ'. If $[\![\varphi(f_1, \ldots, f_k)]\!] \in \mathcal{G}$, then $[\![\varphi'(f_1, \ldots, f_k)]\!] \notin \mathcal{G}$, so we get $\mathcal{F}/\mathcal{G} \models \varphi([f_1], \ldots, [f_k])$ by the induction hypothesis. For the other direction, suppose that $\mathcal{F}/\mathcal{G} \models \varphi([f_1], \ldots, [f_k])$. Then $[\![\varphi'(f_1, \ldots, f_k)]\!] \notin \mathcal{G}$ by the induction hypothesis. But $\varphi', f_1, \ldots, f_k$ were treated by step (iii) at some stage i, so we must have $U_{i+1} \subseteq [\![\neg\varphi'(f_1, \ldots, f_k)]\!]$. Hence $[\![\varphi(f_1, \ldots, f_k)]\!] \in \mathcal{G}$ as required. The step for φ of the form $\varphi' \wedge \varphi''$ is easy. Let us consider the case when φ is $(\exists x < \tilde{n})\varphi'(x, x_1, \ldots, x_k)$, $f_1, \ldots, f_k \in \mathcal{F}$ and the assertion holds for φ'. If $[\![\varphi(f_1, \ldots, f_k)]\!] \in \mathcal{G}$, then there is some $i \in \mathbf{N}$ with $U_i \subseteq [\![\varphi(f_1, \ldots, f_k)]\!]$, and there is some $i' > i$ such that $\varphi, f_1, \ldots, f_k$ were considered by step (iv) at stage i'. Hence $U_{i'+1} \subseteq [\![\varphi'(c_v, f_1, \ldots, f_k)]\!]$ for some $v \leq n$ and we get $\mathcal{F}/\mathcal{G} \models \varphi([f_1], \ldots, [f_k])$ by the induction hypothesis. The other direction follows quickly from the induction hypothesis. So, we have verified (FG 3).

Let us check (FG 4). By (PU 5) we have $[\![\psi(id_\Omega, f, h_1, \ldots, h_{k_0})]\!] \in Bad_0$ for every $f \in \mathcal{F}$, and by step (ii) each of these sets was subtracted from U_i at some stage $i \in \mathbf{N}$. Since ψ is $\Delta_0^{<\tilde{n}}$, (FG 3) applies and (FG 4) follows.

Induction (FG 5) is obtained by step (v): Let $g_1, \ldots, g_k \in \mathcal{F}$ and φ be as in (FG 5). We may assume that

$$M \models (\forall z_1, \ldots, z_k)\varphi(0, z_1, \ldots, z_k) \text{ and } M \models (\forall z_1, \ldots, z_k)\neg\varphi(m, z_1, \ldots, z_k).$$

Step (v) treats $\varphi, g_1, \ldots, g_k$ at some stage $i \in \mathbf{N}$ with the result

$$[\![\alpha(c_v, f_1, \ldots, f_{k'}, g_1, \ldots, g_k)]\!] \supseteq U_{i+1}$$

for some $f_1, \ldots, f_{k'} \in \mathcal{F}$ and $v \leq m - 1$, and

$$[\![\alpha(c_{v+1}, g'_1, \ldots, g'_{k'}, g_1, \ldots, g_k)]\!] \in Bad_{i+1}$$

for every $g'_1, \ldots, g'_{k'} \in \mathcal{F}$. The first part together with (FG 3) give immediately $\mathcal{F}/\mathcal{G} \models \varphi([c_v], [g_1], \ldots, [g_k])$. Since by step (ii) each set in Bad_{i+1} is subtracted from $U_{i'}$ at some stage $i' \in \mathbf{N}$, we obtain $\mathcal{F}/\mathcal{G} \models \neg\varphi([c_{v+1}], [g_1], \ldots, [g_k])$ in a similar way. So, (FG 5) holds.

□

3.5 An example

We will show how to use Construction B together with the assumption that one-way permutations exist to get an interesting model of the theory $strictR^1_2$.

Definition 3.5 (ϵ-OWP)**.** Let $\epsilon : \mathbf{N} \to [0,1]$ be a function. A polynomial-time function $g : \{0,1\}^* \to \{0,1\}^*$ is called an *ϵ-OWP (one-way permutation)* if for every n, g is a permutation of $\{0,1\}^n$ and for any polynomial p, for all sufficiently large n and for every boolean circuit C of size at most $p(n)$,

$$\Pr_{x\in\{0,1\}^n}[g(C(x)) = x] < \epsilon(n). \tag{3.18}$$

Let M be a countable nonstandard model of true arithmetic, $n \in M \setminus \mathbf{N}$ and $\delta > 0$ a (standard) rational number. Assume that an ϵ-OWP $g : M \to M$ exists with $\epsilon(x) := 2^{-x^\delta}$. Set

$$\Omega := \{0,1\}^n,\ q := 2^{-n^\delta} \text{ and } m := \log n.$$

Note that the requirement (R) in Theorem 3.4 is satisfied.

Let $\mathcal{L}$ be the language containing symbols $0, \tilde{n}, S, \leq$ (as required in Theorem 3.4) together with the remaining symbols of L_2, i.e. $+, \cdot, |x|, x\#y, \lfloor x/2 \rfloor, \dot{-}, MSP$, and also the symbol $\tilde{g}$ for the one-way permutation. Let $\mathcal{L}$ be interpreted in M in the usual way (with $\tilde{g}$ interpreted as g). Then in M, (3.18) holds for every boolean circuit C of size $\leq n^{s'}$ for some nonstandard $s' \in M$ (by overspill).

Let X be the set consisting of the identity function id_Ω and the constant functions c_v with value v for each $v \leq n$. Let $\psi(x,y)$ be the formula

$$\tilde{g}(y) = x$$

and put $s := |s'|$. Now change the interpretation of each symbol from $F_\mathcal{L}$ (the set of all function symbols of $\mathcal{L}$) so that they give value 0 on arguments of length $\geq n^s$ (our eventual model will not contain elements of length $\geq n^s$). Note that if $f \in Fct_X(P)$, where $P \in M$ is an SLP over $F_\mathcal{L}$ of size m^s with input variables x_h for $h \in X$, then f is computable in M by a boolean circuit of size $\leq n^{O(s)} \cdot m^s < n^{s'}$ (and of depth bounded by $d \cdot m^s$, where d is a bound on the depth of $n^{O(s)}$ size circuits computing functions from $F_\mathcal{L}$). Hence the hypothesis (H) holds for our choice of $\psi(x,y)$ and parameters m, s and q. So we can use Theorem 3.4 to get the model $\mathcal{F}/\mathcal{G}$.

Note that since we have $+, \cdot$ in the language, we can strengthen (FG 5) as follows: In $\mathcal{F}/\mathcal{G}$ induction holds for $(\exists y_1)\ldots(\exists y_{k'})\alpha$ up to m, where α is an $\mathcal{L}$-formula with parameters and with quantifiers bounded by $\tilde{n}^k$ for some $k \in \mathbf{N}$ (instead of bounded just by $\tilde{n}$). Here $\tilde{n}^k$ means $\tilde{n} \cdot \tilde{n} \cdot \ldots \cdot \tilde{n}$ (k-times). This works because each formula

$$x < \tilde{n}^k \to (\exists x_0 < \tilde{n})\ldots(\exists x_{k-1} < \tilde{n})\ x = x_0 + x_1 \cdot \tilde{n} + \ldots + x_{k-1} \cdot \tilde{n}^{k-1}$$

is $\Delta_0^{<\tilde{n}}$, so by Łoś's theorem (FG 3) it holds in $\mathcal{F}/\mathcal{G}$. Hence we can use it to replace each quantifier bounded by $\tilde{n}^k$ with k quantifiers bounded by $\tilde{n}$ and so we get an equivalent (in $\mathcal{F}/\mathcal{G}$) formula with all quantifiers bounded by $\tilde{n}$.

Define a substructure $\mathcal{F}'/\mathcal{G}$ of $\mathcal{F}/\mathcal{G}$ by taking those elements of $\mathcal{F}/\mathcal{G}$ that are of length $\leq \tilde{n}^k$ for some $k \in \mathbf{N}$. The fact that bounded formulas are absolute between $\mathcal{F}/\mathcal{G}$ and $\mathcal{F}'/\mathcal{G}$ helps with verifying that (FG 1) - (FG 4) also hold with $\mathcal{F}$ replaced by $\mathcal{F}'$. Since the theory BASIC consists of open formulas, $\mathcal{F}'/\mathcal{G}$ is a model of BASIC thanks to Łoś's theorem (FG 3). As for induction, compared to just discussed induction in $\mathcal{F}/\mathcal{G}$ we have to bound the existential quantifiers to make sure their witnesses in $\mathcal{F}/\mathcal{G}$ are also in $\mathcal{F}'/\mathcal{G}$. Hence in $\mathcal{F}'/\mathcal{G}$ induction holds up to m for formulas of the form

$$(\exists y_1 \leq t_1) \dots (\exists y_{k'} \leq t_{k'})\alpha$$

where $t_1, \dots, t_{k'}$ are $\mathcal{L}$-terms and α is a sharply bounded formula. Thus we have proved:

Theorem 3.6. *Let $\delta > 0$ be a rational number, $\epsilon(x) := 2^{-x^\delta}$ and assume that an ϵ-OWP exists. Let M be a countable nonstandard model of true arithmetic and let $n \in M$ be a nonstandard number. Suppose that $\tilde{g}$ is a function symbol interpreted in M by the ϵ-OWP. Then there exists $N \models strictR^1_2(\tilde{g})$ such that N restricted to $Log(N)$ coincides with M restricted to $\{x \in M \mid x \leq n^k$ for some $k \in \mathbf{N}\}$ and such that in N,*

$$\begin{aligned}
&|\tilde{g}(x)| = |x| \\
&\tilde{g}(x) = \tilde{g}(y) \to x = y \\
&(\exists y)(\forall x)\neg\tilde{g}(x) = y.
\end{aligned}$$

3.6 Theory $strictR^1_2$ and sharply bounded collection scheme

We separate theories $R^1_2(g)$ and $strictR^1_2(g)$ assuming that g is a one-way function hard against polynomial-size circuits.

We will use Construction B with some slight modifications to its setup. Assume that M is a countable nonstandard model of true arithmetic and $n \in M \setminus \mathbf{N}$. Let $\Omega \subseteq \{0,1\}^{n^2}$ be the set of all numbers w such that

$$(\forall u < n)|\beta_{2^{n-1}}(w, u)| = n.$$

We can view Ω as the set of all $n \times n$ $(0,1)$-matrices in M whose rows represent numbers of length n. Let $\mathcal{L}$ be the language containing symbols $0, \tilde{n}, S, \leq$ as in Theorem 3.4 together with the remaining symbols of L_2, i.e. $+, \cdot, |x|, x\#y, \lfloor x/2 \rfloor$, $\dot{-}$, MSP, and with a unary function symbol $\tilde{g}$ intended for a one-way permutation. Let X be the set consisting of the identity function id_Ω and the constant functions c_v with value v for each $v \leq n$.

There will be one difference in the definition of an SLP over $F_\mathcal{L}$ of size t with input variables x_h, $h \in X$. Namely, we allow it to use an additional instruction to invert one row of the matrix assigned to the input variable x_{id_Ω}. Formally, the additional instruction is of the form

$$y_i := \tilde{g}^{-1}(\beta_{2^{n-1}}(x_{id_\Omega}, v)), \tag{3.19}$$

where $i = 1, \dots, t$, the symbol $\tilde{g}^{-1}$ denotes the inverse of $\tilde{g}$, and $v \leq n$. For an SLP P the set of functions computed at the variables of P, $Fct_X(P)$, is defined as in Section 3.4.

Let $\delta > 0$ be a (standard) rational number and denote

$$q := 2^{-n^\delta}, m := \log n.$$

Let $\psi(x, y)$ be the formula

$$(\forall i < \tilde{n})\, \tilde{g}(\beta_{LSP(x,\tilde{n})}(y, i)) = \beta_{LSP(x,\tilde{n})}(x, i).$$

Here $LSP(x, \tilde{n})$ is just a way to produce a number of length $\tilde{n}$ provided x attains a value from Ω and $\tilde{n}$ is interpreted as n. Then $\psi(x, y)$ says that y is the matrix whose rows are the inverses under $\tilde{g}$ of the rows of x. We need the following hypothesis:

> *There exists a function $g : M \to M$ which for every l permutes the numbers of length l and there exists $s \in M \setminus \mathbf{N}$ such that if $\mathcal{L}$ is interpreted in M in the usual way with $\tilde{g}$ interpreted as g, except that the functions give value 0 on inputs of length $\geq n^s$, then in M,* (H*)
>
> $$\Pr_{u \in \Omega} [\psi(u, f(u))] < q$$
>
> *for all $f \in Fct_X(P)$ for every SLP P over $F_{\mathcal{L}}$ of size m^s with input variables $x_h, h \in X$.*

Theorem 3.7. *Let $M, n, \Omega, \mathcal{L}, X, \delta, q, m, \psi$ be as above and assume the hypothesis* (H*)*. Then there exists a model N of $strictR^1_2(\tilde{g})$ such that N restricted to $Log(N)$ coincides with M restricted to $\{x \in M \mid x \leq n^k$ for some $k \in \mathbf{N}\}$ and the following instance of $BB\Sigma^b_0(\tilde{g})$ does not hold in N:*

$$(\forall x)\big((\forall i < n)(\exists z < 1\#LSP(x, n))\, \tilde{g}(z) = \beta_{LSP(x,n)}(x, i)$$
$$\to (\exists y)(\forall i < n)\, \tilde{g}(\beta_{LSP(x,n)}(y, i)) = \beta_{LSP(x,n)}(x, i)\big).$$

Proof. We can repeat the proof of Theorem 3.4 in the current setup (where we have slightly different SLPs and Ω) with hardly any changes and get a structure $\mathcal{F}/\mathcal{G}$ satisfying the conclusion of that theorem, (FG 1) - (FG 5). As in the previous section, we take the substructure $\mathcal{F}'/\mathcal{G}$ of $\mathcal{F}/\mathcal{G}$ whose universe consists of all elements of $\mathcal{F}/\mathcal{G}$ that are of length $\leq \tilde{n}^k$ for some $k \in \mathbf{N}$. It follows that $\mathcal{F}'/\mathcal{G}$ is a model of $strictR^1_2(\tilde{g})$ such that

$$\mathcal{F}'/\mathcal{G} \models (\forall i < \tilde{n})(\exists z < 1\#LSP([id_\Omega], \tilde{n}))\, \tilde{g}(z) = \beta_{LSP([id_\Omega],\tilde{n})}([id_\Omega], i)$$

but

$$\mathcal{F}'/\mathcal{G} \not\models (\exists y)(\forall i < \tilde{n})\, \tilde{g}(\beta_{LSP([id_\Omega],\tilde{n})}(y, i)) = \beta_{LSP([id_\Omega],\tilde{n})}([id_\Omega], i).$$

□

The proof of the following lemma uses an argument similar to that of S. Cook and N. Thapen [7] described in the footnote on page 7 of [7] and attributed there to R. Impagliazzo.

Lemma 3.8. *Let $M, n, \Omega, \mathcal{L}, X, \delta, q, m, \psi$ be as above and assume that an ϵ-OWP exists with $\epsilon(x) := 2^{-x^\delta}$. Then* (H*) *is true.*

Proof. Let $g : M \to M$ be an ϵ-OWP. Then in M, (3.18) holds for every boolean circuit of size $\leq n^{s'}$ for some nonstandard $s' \in M$ such that $m^{s'} < n$. For the sake of a contradiction, assume that ($\mathrm{H}^\star$) is false. So there is an SLP P of size m^s where $s := |s'|$ such that some $f \in Fct_X(P)$ inverts all rows of a matrix from Ω with probability at least 2^{-n^δ}. P has at most m^s instructions of the form (3.19). These instructions query a fixed set of rows of any matrix from Ω. Assume without loss that the first row is not queried.

We construct a probabilistic boolean circuit C' of size smaller than $n^{s'}$ that satisfies

$$\Pr[g(C'(x)) = x] \geq 2^{-n^\delta}, \tag{3.20}$$

where the probability is taken uniformly over all numbers x of length n and over the random bits used by C'.

The input of the circuit is (the binary representation of) a random number of length n, which we want to find a preimage of. C' uses $(n-1)^2$ random bits, which represent $(n-1)$ random numbers of length n. First, C' forms an $n \times n$ (0,1)-matrix A whose first row is the input and the remaining rows are obtained from the $(n-1)$ random numbers by applying to each of them a circuit computing g. After that, C' simulates the computation of f on A by replacing the instructions of P with polynomial-size circuits computing the functions of $\mathcal{L}$ and by easily providing correct answers to the queries of the form (3.19). The first row of the resulting matrix is the output of C'.

The size of C' is at most $n^{O(s)} \cdot m^s < n^{s'}$. Since g is a permutation, A is uniformly distributed on Ω and hence C' indeed satisfies (3.20). It follows that there exists a setting of the $(n-1)^2$ random bits used by C' such that for the resulting circuit C we have

$$\Pr_{x \in \{0,1\}^n}[g(C(x)) = x] \geq 2^{-n^\delta},$$

a contradiction. □

The next corollary follows from Theorem 3.7, Lemma 3.8 and the result of B. Allen [3] that R^i_2 proves $BB\Sigma^b_i$ for $i \geq 1$.

Corollary 3.9. *Let $\epsilon(x) := 2^{-x^\delta}$. If an ϵ-OWP exists then strict$R^1_2(\tilde{g})$ is weaker than $R^1_2(\tilde{g})$. If an ϵ-OWP is definable by a term in the language of R^1_2, then strictR^1_2 is weaker than R^1_2.*

3.7 Acknowledgements

I wish to thank Samuel Buss, Jan Krajíček and Neil Thapen for very helpful suggestions and discussions, Leszek Kołodziejczyk for comments that helped to improve this paper and for making me aware of the *strict*R^1_2 versus R^1_2 problem and Chris Pollett for clarifying the current status of this problem.

Bibliography

[1] M. Ajtai, *Generalizations of the Compactness Theorem and Gödel's Completeness Theorem for Nonstandard Finite Structures*, Proceedings of the 4th international conference on Theory and applications of models of computation (2007) 13-33.

[2] M. Ajtai, *A Generalization of Gödel's Completeness Theorem for Nonstandard Finite Structures*, manuscript (2011)

[3] B. Allen, *Arithmetizing Uniform NC*, Annals of Pure and Applied Logic 53 (1991) l-50

[4] S. Buss, *Bounded Arithmetic*, Bibliopolis, 1986

[5] S. Buss, *Weak End Extensions of Models of Bounded Arithmetic*, unpublished manuscript (1986)

[6] P. Clote, G. Takeuti *Bounded arithmetic for NC, ALogTIME, L and NL*, Annals of Pure and Applied Logic 56 (1992) 73-177

[7] S. Cook, N. Thapen, *The strength of replacement in weak arithmetic*, ACM Transactions on Computational Logic 7 (2006) 749-764

[8] M. Garlík, *A New Proof of Ajtai's Completeness Theorem for Nonstandard Finite Structures*, Archive for Mathematical Logic 54(3-4), (2015), pp. 413-424.

[9] P. Hájek, P. Pudlák, *Metamathematics of first order arithmetic*, Springer, 1993

[10] J. Johannsen, C. Pollett, *On the Δ_1^b-Bit-Comprehension Rule*, Logic Colloquium '98, Lecture Notes in Logic, ASL 13 (2000) 262-279

[11] S. Kochen, S. Kripke, *Non-standard models of Peano Arithmetic*, Enseign. Math. 28(2) (1982) 211-231.

[12] J. Krajíček, *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, Cambridge University Press, 1995

[13] A. Máté, *Nondeterministic polynomial-time computations and models of arithmetic*, Journal of the Association for Computing Machinery, 37(1) (1990) 175-193

4. Further Results

In this chapter we further develop Construction B from Chapter 3 and use it to improve upon two results from Chapter 3, allowing a richer langauge of theories and removing a hypothesis from the separation of relativized theories.

In [4] Cook an Thapen investigated various theories weaker than S_2^1 and showed (sometimes under a computational assumption) that some of these theories do not prove either $BB\Sigma_1^b$ or $BB\Sigma_0^b$. Theorem 4.1 below can be viewed as an extension of this type of results to a theory which has more induction than theories considered there, namely $strict\Sigma_1^b(PV) - \mathrm{LLIND}$. This theorem strengthens Theorem 3.4 by allowing the language of PV.

Theorem 4.2 below improves upon the separation of the relativized theories in Corollary 3.9. C. Pollett [8] asks whether the theories R_2^1 and $strictR_2^1$ are distinct. We can summarize an answer as follows: They are, assuming a plausible computational assumption (Corollary 3.9), and the relativized versions of these theories, $R_2^1(\alpha)$ and $strictR_2^1(\alpha)$, are distinct (unconditionally) by Theorem 4.2.

This chapter uses the same notation as defined in Preliminaries 3.2 of Chapter 3.

4.1 A separation of two theories in the language of PV

Assuming the existence of a one-way permutation hard against polynomial-size circuits we will separate two theories extending the theory PV_1. PV_1 (defined in [6], called QPV in [3]) is a first-order variant of the theory PV.

PV is an equational theory defined by Cook [2]. Its language includes a function symbol for every polynomial time algorithm and these function symbols are introduced inductively based on Cobham's characterization of the class of polynomial time functions as the least class containing several basic functions and closed under composition and limited recursion on notation. PV contains defining equations for its function symbols and it uses induction on notation as a derivation rule.

The theory PV_1 can be axiomatized by universal axioms defining the function symbols of PV and the scheme of induction for open formulas. We would like to use the symbols of L_2 as well, so we need to identify them with some functions of PV. We will do so by including L_2 in the basic language of PV and by including BASIC in the axioms of PV_1.

We will use the fact that PV_1 is a universal theory. This is because each open induction axiom can be rewritten as a universal formula using a function of PV simulating the binary search (see [5]).

Let $\Sigma_i^b(PV)$ and $strict\Sigma_i^b(PV)$ denote classes of formulas defined like Σ_i^b and $strict\Sigma_i^b$, respectively, but in the language of PV.

Recall the definition of an ϵ-OWP (Definition 3.5).

Theorem 4.1. *Let $\delta > 0$ be a rational number, $\epsilon(x) := 2^{-x^\delta}$ and suppose that an ϵ-OWP exists. Then $PV_1 + strict\Sigma_1^b(PV) - \mathrm{LLIND}$ is weaker than $PV_1 + \Sigma_1^b(PV) - \mathrm{LLIND}$.*

Proof. We will proceed as in Section 3.6. The setup of the proof is the same as described by the paragraphs above Theorem 3.7 (including the hypothesis (H*)) with one exception: the language $\mathcal{L}$ will have, in addition to symbols listed there, finitely many new function symbols. By a theorem of Muchnik [7] the class of polynomial time functions has a finite basis, i.e., there is a finite set of functions (called a basis) such that the class of polynomial time functions is the smallest class of functions containing the functions from the basis and closed under permutation and renaming of variables and under composition of functions. The new function symbols added to $\mathcal{L}$ are to be interpreted as such a basis.

Then, exactly as we did in the proof of Theorem 3.7, we repeat the proof of Theorem 3.4 in the current setup and we get $\mathcal{F}/\mathcal{G}$ satisfying (FG 1) - (FG 5). We take the substructure $\mathcal{F}'/\mathcal{G}$ of $\mathcal{F}/\mathcal{G}$ whose universe consists of all elements of length $\leq \tilde{n}^k$ for some $k \in \mathbf{N}$. We expand $\mathcal{F}'/\mathcal{G}$ into the language of PV by interpreting each symbol of PV by an appropriate composition of functions from the basis. Using Łoś's theorem (FG 3) and the fact that PV_1 is a universal theory we have $\mathcal{F}'/\mathcal{G} \models PV_1$. From (FG 5) we get $\mathcal{F}'/\mathcal{G} \models \mathit{strict}\Sigma_1^b(PV) - \mathrm{LLIND}$. Since SLPs are allowed to query a row of a matrix from Ω by the instruction (3.19), it follows that

$$\mathcal{F}'/\mathcal{G} \models (\forall i < \tilde{n})(\exists z < 1\#LSP([id_\Omega], \tilde{n}))\, \tilde{g}(z) = \beta_{LSP([id_\Omega],\tilde{n})}([id_\Omega], i).$$

By (FG 4) and our choice of $\psi(x, y)$ we get

$$\mathcal{F}'/\mathcal{G} \not\models (\exists y)(\forall i < \tilde{n})\, \tilde{g}(\beta_{LSP([id_\Omega],\tilde{n})}(y, i)) = \beta_{LSP([id_\Omega],\tilde{n})}([id_\Omega], i).$$

Thus $BB\psi$, which is provable in $PV_1 + \Sigma_1^b(PV) - \mathrm{LLIND}$, does not hold in $\mathcal{F}'/\mathcal{G}$. The theorem hence follows from Lemma 3.8, which holds for our richer language $\mathcal{L}$ as well (with the same proof). □

4.2 A relativized separation

We consider the relativized theories $R_2^1(\alpha)$ and $\mathit{strict}R_2^1(\alpha)$ for a new unary relation symbol α. We prove unconditionally that these theories are distinct. Recall that in Chapter 3 (Corollary 3.9) we proved their separation under the hypothesis of the existence of a one-way permutation hard against polynomial-size circuits.

Define

$$Bit(x, i) := MSP(x, i) \dot{-} 2 \cdot MSP(x, i + 1).$$

The comprehension axiom for a formula $\varphi(x)$, denoted $\mathrm{COMP}_\varphi(a)$, is the formula

$$(\exists y < 2^{|a|})(\forall x < |a|)(Bit(y, x) = 1 \leftrightarrow \varphi(x)).$$

We will choose $\varphi(x)$ to be $\alpha(x)$ and we will show that $\mathrm{COMP}_\alpha(a)$ separates the theories above.

Theorem 4.2. *For a new unary relation symbol α, $\mathit{strict}R_2^1(\alpha)$ is weaker than $R_2^1(\alpha)$.*

Proof. $R_2^1(\alpha) \vdash \mathrm{COMP}_\alpha(a)$ is proved by the same recursive doubling trick that B. Allen [1] used to show that R_2^i proves $BB\Sigma_i^b$.

To build a model of $\mathit{strictR}_2^1(\alpha)$ where $\mathrm{COMP}_\alpha(a)$ does not hold we employ Theorem 3.4 with the simplification that Ω is a singleton, and hence the filter and all notions used to construct it become trivial.

Assume that M is a countable nonstandard model of true arithmetic, $n \in M \setminus \mathbf{N}$, and $\Omega := \{0\}$. Let $\mathcal{L}$ consist of L_2 together with the constant $\tilde{n}$ and the relation symbol α. Let X, the set of inputs for SLPs, be $\{0, 1, 2, \ldots, n\}$. We define $\psi(y)$ to be

$$(\forall x < \tilde{n})(\mathit{Bit}(y, x) = 1 \leftrightarrow \alpha(x)).$$

(We stop using the symbol id_Ω occuring in Theorem 3.4 since it is now 0.) Instead of the probabilistic assumption (H) we now have that there is an interpretation which is in M of α on the set $\{0, 1, 2, \ldots, n-1\}$ and there exists a nonstandard $s \in M$ such that for every SLP $P \in M$ of size $(\log n)^s$ and every $e \in \mathit{Fct}_X(P)$,

$$M \models \neg\psi(e).$$

This is true because in M there are 2^n strings of length n, whereas there is a constant $c \in \mathbf{N}$ such that there are less than $2^{c(\log n)^{s+1}}$ possible outputs by SLPs of size at most $(\log n)^s$. Hence we can pick a string v of length n which is not computed by the SLPs and we interpret α by

$$\alpha(i) \leftrightarrow \mathit{Bit}(v, i) = 1.$$

Now we can allow $q = 0$ and continue as in the proof of Theorem 3.4. The only thing that does not become completely trivial is ensuring induction for $\exists\Delta_0^{\leq n}$ formulas up to $\log n$. This is done in step (v) of the proof, which under current circumstances is reduced to the pruning of the tree of SLPs in search for witnesses of the unbounded existential quantifiers in a formula for which we want to arrange induction.

The result is a substructure N of M and we take its substructure N' whose universe consists of all elements of N of length $\leq \tilde{n}^k$ for some $k \in \mathbf{N}$. It is readily verified that $N' \models \mathit{strictR}_2^1(\alpha)$, that N' contains an element e with $|e| = \tilde{n}$, and that $N' \models \neg\mathrm{COMP}_\alpha(e)$.

□

Bibliography

[1] B. Allen, *Arithmetizing Uniform NC*, Annals of Pure and Applied Logic 53 (1991) l–50

[2] S. A. Cook, *Feasibly constructive proofs and the propositional calculus*, Proceedings of the 7th Annual ACM Symposium on Theory of Computing (1975) 83–97.

[3] S. A. Cook, *Relating the provable collapse of P to NC^1 and the power of logical theories*, Proof Complexity and Feasible Arithmetics (P. Beame and S. R. Buss, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science 39, American Mathematical Society (1998) 73–92.

[4] S. Cook, N. Thapen, *The strength of replacement in weak arithmetic*, ACM Transactions on Computational Logic 7 (2006) 749–764

[5] J. Krajíček, *Bounded Arithmetic, Propositional Logic, and Complexity Theory*, Cambridge University Press, 1995

[6] J. Krajíček, P. Pudlák, and G. Takeuti, *Bounded arithmetic and the polynomial hierarchy*, Annals of Pure and Applied Logic 52 (1991) 143–153.

[7] A. A. Muchnik, *On two approaches to the classification of recursive functions*, Problems in Mathematical Logic, Complexity of Algorithms and Classes of Computable Functions (A.A. Muchnik, V.A. Kozmidiadi eds.), MIR, Moscow (1970) 123–138.

[8] C. Pollett, *Structure and Definability in General Bounded Arithmetic Theories*, Annals of Pure and Applied Logic 100 (1999) 189–245.

5. Concluding Remarks

Our results show that it is possible to come up with a new method for constructing nonstandard models of bounded arithmetic theories and use it to solve problems open for some time. We think that further development of the method described as Construction B in Chapter 3 is possible; results reported in Chapter 4 certify to this. It seems to be crucial to have suitable open problems as aims, to guide the development of the method. Below we list two which we think are quite stimulating.

Problem 1. ($I\Delta_0$ and the weak pigeonhole principle.) Given a nonstandard model M of true arithmetic and a nonstandard $a \in M$, can we expand M by adding a relation R on $[0, a]$ such that the resulting structure satisfies induction on $[0, a]$ for all formulas in the language of arithmetic extended by R and such that R is an injective map from $[0, 2a]$ to $[0, a]$, i.e. R violates the weak pigeonhole principle? Or, is it at least true that the weak pigeonhole principle for R is unprovable in $I\Delta_0(R)$, a theory of arithmetic with induction for bounded formulas in the language extended by R?

Problem 2. (A lower bound for $F_d(\oplus)$.) Does for every $d \geq 3$ exist $\epsilon > 0$ such that proofs of the tautologies PHP_k for $k \geq 1$ in $F_d(\oplus)$ must have size at least 2^{k^ϵ}? Here the pigeonhole principle PHP_k is a propositional formulation of the statement that no map from $[k+1]$ into $[k]$ is injective and $F_d(\oplus)$ is a Frege proof system operating with formulas of depth at most d over the DeMorgan language $(0, 1, \neg, \vee, \wedge)$ with the parity connective $\oplus$.

In particular, is $(\forall x)\mathrm{PHP}(x, R)$ unprovable in $Q_2V_1^0$? $\mathrm{PHP}(x, R)$ is a first-order formulation of the pigeonhole principle for a binary relation symbol R and $Q_2V_1^0$ is a bounded arithmetic theory with a parity quantifier.

Problem 1 originated with A. MacIntyre in 1980s. M. Ajtai's beautiful method [1] was successful in solving Problem 1 for an injective map R from $[0, a]$ to $[0, a-1]$, i.e., for the (usual) pigeonhole principle instead of the weak one. But neither this method nor its later improvements by Krajíček et al. [3] and Pitassi et al. [4] were able to solve the above problems. In case of Problem 1, the main obstacle is that we do not have an analogue of the switching lemma, which is a combinatorial argument that together with a particular forcing construction forms the method of Ajtai.

We would like to investigate various model-theoretic constructions to tackle these problems. The aim is to find a construction so that when some combinatorial task enters the picture it will be easier to deal with than obstacles like the switching lemma.

Some recent progress on Problem 2 has been achieved by Buss, Kołodziejczyk and Zdanowski [2] by reducing the question to depth $d = 3$ only.

Bibliography

[1] M. Ajtai, *The Complexity of the Pigeonhole Principle*, Proceedings of the IEEE 29th Annual Symposium on Foundation of Computer Science (1988), 346–355

[2] S. R. Buss, L. A. Kołodziejczyk, K. Zdanowski, *Collapsing modular counting in bounded arithmetic and constant depth propositional proofs*, to appear in Transactions of the AMS

[3] J. Krajíček, P. Pudlák and A. Woods, *An exponential lower bound to the size of bounded depth frege proofs of the pigeonhole principle*, Random Structures and Algorithms, 7:1 (1995), 15–39

[4] T. Pitassi, P. Beame and R. Impagliazzo, *Exponential lower bounds for the pigeonhole principle*, Computational Complexity, 3 (1993), 97–308