Partially Definable Forcing and Bounded Arithmetic

Albert Atserias [*]   Moritz Müller [†]

July 17, 2014

Abstract

We describe a method of forcing against weak theories of arithmetic and its applications in propositional proof complexity.

1 Introduction

We are interested in the following problem. Given a nonstandard model M of arithmetic we want to expand it by interpreting a binary relation symbol R such that R^M does something prohibitive, e.g. violates the pigeonhole principle in the sense that R^M is a bijection from $n+1$ onto n for some (nonstandard) $n \in M$. The goal is to do so while preserving as much as possible from ordinary arithmetic. More precisely, we want the expansion (M, R^M) to model the least number principle for a class of formulas as large as possible.

Progress concerning this general problem has been very slow. Concerning the pigeonhole principle, the central results are the following. Paris and Wilkie [30] succeeded in getting the least number principle for existential formulas, and Riis [38] pushed their line of argument to handle formulas that may additionally have universal quantifiers bounded by some $b_0 < n^{o(1)}$, i.e. n raised to some infinitesimal power. The most important result in the area is due to Ajtai [1], who could do it for formulas with quantifiers (of both types) bounded by some b_0 that is bigger than any standard power of n. Subsequent work [31, 27] improved Ajtai's bound to $b_0 < 2^{n^{o(1)}}$, which is essentially optimal.

The purpose of this article is to give a common framework for constructing such expansions by forcing.

---

[*]Universitat Politècnica de Catalunya (UPC), Departament de Llenguatges i Sistemes Informàtics (LSI), Jordi Girona Salgado 1-3, Edif. Omega-S122, 08034 Barcelona, Spain, `atserias@lsi.upc.edu`

[†]Kurt Gödel Research Center (KGRC), Universität Wien, Währingerstraße 25, 1090 Wien, Austria, `moritz.mueller@univie.ac.at`

Independence and proof complexity The main motivation for our question is a better understanding of independence from (weak) theories of arithmetic. Pudlák argues that the current understanding is unsatisfactory in that "except for Gödel's theorem which gives only special formulas, no general method is known to prove independence of [arithmetical] Π_1 sentences" [33, section 3]. The question is linked to central open questions in computational complexity theory in general and propositional proof complexity in particular.

In fact, already weak arithmetical theories like those in Buss' hierarchy correspond in a certain precise sense to the complexity classes in the polynomial hierarchy. In this respect, Riis' result implies independence of the pigeonhole principle from Buss' theory $T_2^1(R)$, Ajtai's result implies independence from the theory $I\Delta_0(R)$, and the mentioned improvements imply independence from Buss' $T_2(R)$.[1]

Independence from weak arithmetics is closely related to lower bounds on the size of propositional proofs. The usual textbook systems, i.e. Hilbert style calculi given by finitely many inference rules, are usually called *Frege systems* [14]. For these systems no superpolynomial lower bounds on proof size are known[2] (see [11] for a survey), in particular, there are short (polynomial size) Frege proofs of a sequence of tautologies expressing the pigeonhole principle [9]. Ajtai's result implies a superpolynomial lower bound $n^{\omega(1)}$ for *bounded depth* Frege proofs, i.e. Frege proofs using only formulas of some fixed $\wedge/\vee$-alternation rank. The mentioned improvements of Ajtai's result imply an exponential lower bound $2^{n^{\Omega(1)}}$. Riis' result implies an $n^{\Omega(1)}$ lower bound on a general notion of width, namely *Poizat width*, of refutations of arbitrary infinity axioms in arbitrary proof systems. We refer to section 5 for details.

Paris and Wilkie, Riis, and Ajtai all explicitly refer to their argument as being of the forcing type. Ajtai says his construction of R^M is "done according to the general ideas of Cohen's method of forcing " [1, p.348]. However, the argument is "mostly combinatorial and probabilistic" [1, p.347], relying on specialized and difficult versions of so-called switching lemmas in circuit complexity. As Ben-Sasson and Harsha put it, it is "extremely difficult to understand and explain" [7, p.19:2]. Lots of efforts have been made to simplify and reinterpret Ajtai's argument (e.g. [6, 34, 23, 24, 25, 7]). Conceptually, later improvements [6, 31, 27] of Ajtai's result "eliminate the non-standard model theory" [6, p.367] and the forcing mode of speech. And technically, the mentioned switching lemmas have been improved and simplified (see [4, 48] for surveys).

Still, not much is known on how to apply Ajtai's argument to stronger systems or other principles (cf. [23, Chapter 12]). Perhaps one can say that the abovementioned efforts did not lead to an understanding of Ajtai's argument as instantiating some general method as Pudlák asks for.

---

[1] [23, 13] are monographs, [12, 10] surveys on the subject. The present article does not suppose any familiarity with these theories.

[2] For weaker systems strong lower bounds are known; [36, 34, 49, 5, 41] are surveys with distinct emphases. The present article does not suppose any familiarity with proof complexity.

Comparison with Cohen forcing   This sorry state of affairs clearly contrasts with Cohen forcing in set theory. We recall briefly and informally its set-up. With a model M of ZF and a 'generic' set G external to M one associates a model $M[G]$ containing G. Intuitively, G being 'generic' means being 'random' with respect to possible partial information about it. Forcing is a way to reason about $M[G]$ using partial information about G. A piece of partial information p *forces* φ if any generic G 'satisfying' p leads to a model $M[G]$ satisfying φ. Such pieces can be extended in various, possibly incompatible ways, so we think of them as being partially ordered (the *forcing frame*).

Following Shoenfield [43], reasoning about forcing rests on three central lemmas: the Extension Lemma states that extension preserves forcing, the Truth Lemma asserts that every sentence true in $M[G]$ is already forced by some partial information p about G, and the Definability Lemma states that the forcing relation is in a certain sense definable in M. In turn, these lemmas rest on the Forcing Completeness Theorem, a characterization of the 'semantic' forcing notion above by a handier 'syntactic' notion defined via recursion on logical syntax. This understanding of forcing underlies the "Principal Theorem" [43] stating that $M[G]$ models ZF. This way an independence question is reduced to a combinatorial task of designing an appropriate forcing frame.

In contrast, the mentioned forcing type arguments in bounded arithmetic are not based on some more general background theory of forcing. Ajtai writes "Our terminology will be similar to the terminology of forcing but we actually do not use any result from it" [1, p.348]. Consequently it is not completely clear why one should refer to these arguments as forcing arguments. Technically, the crucial difference is that the Definability Lemma fails. Forcing Completeness is proved neither in the original arguments nor in later presentations [51], [23, section 12.7] that emphasize the forcing mode of speech. In [30, 1] no 'syntactic' notion is defined; in [38] it is, but one for which Forcing Completeness fails.

This work   We propose a general background theory of forcing as a unifying way to address the problem mentioned in the beginning and to understand the arguments of Paris and Wilkie, Riis, and Ajtai [30, 38, 1]. We first sketch the argument of Paris and Wilkie and discuss what such a theory should look like. Section 2 then develops forcing accordingly. The framework naturally accommodates the mentioned forcing arguments [30, 38, 1] as well as Cohen forcing and others.

In our context, a Principal Theorem would state that generic expansions satisfy the least number principle for a certain fragment of formulas. In section 3 we show that this holds true when using a forcing that is in an appropriate sense 'definable' for the fragment in question. A more combinatorial formulation of this property is gained by what we call the *antichain method*. Thereby again, independence questions reduce to a combinatorial task of designing forcing frames.

Section 4 gives the constructions of Paris and Wilkie, Riis and Ajtai within this framework. As a byproduct we obtain the above mentioned 'width' lower bound in section 5.

The framework proposed allows to understand Ajtai's result as instantiating some general method closely following Cohen forcing. In [26] Krajíček also develops such a framework

which is technically and conceptually quite different, and follows Scott and Solovay's forcing with random variables. This and other related work is discussed in Notes at the end of sections 2, 4 and 5. All results are stated and proved in a generally accessible language. In particular, no familiarity with bounded arithmetic or forcing is assumed. This way we hope to bring some open questions in the field to the attention of a wider audience. In particular, techniques from set theoretic forcing may be applicable to these questions.

2 Forcing in general

2.1 Motivation

To start, we sketch the argument of Paris and Wilkie [30], "the first forcing argument in the context of weak arithmetic" [23, p.278], and give some informal discussion.

Theorem 2.1 (Paris, Wilkie 1985)**.** *Let M be a countable model of true arithmetic and $n \in M$ nonstandard. Then M has an expansion (M, R^M) such that R^M is a bijection from $[n+1]$ onto $[n]$ and (M, R^M) satisfies the least number principle for existential formulas with parameters.*

Being a model of true arithmetic means being elementarily equivalent to the standard model $\mathbb{N}$ interpreting, say, the language $L := \{+, \cdot, 0, 1, <\}$. We write $[n]$ for $\{m \in M \mid m <^M n\}$. That (M, R^M) satisfies the least number principle for a formula $\varphi(x)$ means that the set $\varphi(M)$ defined by $\varphi(x)$ in (M, R^M) is empty or has a minimal element (with respect to the linear order $<^M$). That $\varphi(x)$ is a formula *with parameters* means that its language is $L \cup \{R\}$ together with the elements of M as constant symbols which are understood to be interpreted by themselves.

Sketch. That n is nonstandard means that $[n]$ is infinite. Consider the set P of finite partial bijections from $[n+1]$ to $[n]$. We construct an interpretation R^M of R as the union of a chain $\emptyset = p_0 \subseteq p_1 \subseteq \cdots$ in P. Having constructed p_{2i} we choose p_{2i+1} such that it contains the ith element a of $[n+1]$ in its domain and the ith element b of $[n]$ in its image; here, we are refering to fixed, external enumerations of $[n+1]$ and $[n]$. These choices ensure that $R^M = \bigcup_{i \in \mathbb{N}} p_i$ will be a bijection from $[n+1]$ onto $[n]$. In fact, p_{2i+1} 'forces' that the expansion under construction will satisfy $\exists! y Ray$ and $\exists! x Rxb$.

Fix an enumeration of all existential $(L \cup \{R\})$-formulas with parameters. The choice of p_{2i+2} 'forces' that the ith such formula $\varphi(x)$ will not violate the least number principle. This is done as follows: write $\varphi(x) = \exists \bar{y} \psi(x, \bar{y})$ and let $a \in M$. Assume p_{2i+1} is such that $\neg\varphi(a)$ is not 'forced', i.e. *some* continuation of the chain leads to an R^M with $(M, R^M) \models \varphi(a)$. Choose $\bar{b}$ such that $(M, R^M) \models \psi(\bar{b}, a)$. Since $\psi(\bar{b}, a)$ is quantifier free it is propositionally satisfied by the truth values of the atoms Rcd occuring in it – say these are $\ell \in \mathbb{N}$ many. Then there is $q \in P$ containing ℓ pairs such that $p_{2i+1} \cup q$ is in P and 'forces' $\psi(\bar{b}, a)$ and thereby $\varphi(a)$. Now, it is not hard to see that there is an L-formula $\chi(x)$ with parameters that expresses this as a property of a in M. If there is no a satisfying $\chi(x)$ in M, then take

$p_{2i+2} = p_{2i+1}$. Otherwise there is a minimal such a, because M is a model of true arithmetic, and then take $p_{2i+2} = p_{2i+1} \cup q$ for the corresponding q. $\qquad\square$

What is the general line of this argument, that justifies calling it "a simple forcing argument"? [30, p.333] Intuitively, P figures as 'forcing frame' and $p \in P$ 'forces' a sentence φ if φ is true in all those expansions (M, R^M) that result from continuing the construction of the chain from p. The construction is done as it is in order to 'force' certain properties of R^M. For the right notion of genericity, this should be automatic. This way one can define a 'semantic' notion of forcing by stipulating that p 'forces' φ if φ is true in all 'generic' expansions (M, R^M).

Note P can be viewed as a subset of M when identifying the finite partial bijections with their codes in M, but this set is *not* definable in M. We thus ask for some general framework for forcing with undefinable forcing frames P. In fact, we let P be just a second structure. We start with syntactically defined forcing relations in section 2.3. It is a subtle point to come up with the right notion of genericity, a point discussed in section 2.4. Section 2.5 then defines *generic associates $M[G]$* of M. Generic expansions are obtained using a simple type of forcing that we call *conservative* (section 2.6).

The resulting framework includes the Extension Lemma, the Truth Lemma and Forcing Completeness but of course no Definability Lemma. It is sufficiently general to naturally accomodate various forcing type arguments from the literature (cf. section 2.7), including Cohen forcing and, as we shall see in section 4, the mentioned forcing type arguments in bounded arithmetic.

2.2 Basic forcing terminology

A *forcing frame* is a structure $(P, \leq, D_0, D_1, \ldots)$ such that $\leq$ partially orders P and $D_0, D_1, \ldots$ are subsets of P. We use $p, q, r, \ldots$ to range over elements of P, called *conditions*. If $p \leq q$ we say p *extends* q and call p an *extension* of q. If p, q have a common extension, they are *compatible* ($p \| q$) and otherwise *incompatible* ($p \perp q$).

A set of conditions $X \subseteq P$ is *downward-closed* if it contains all extensions of its elements; being *upward-closed* is similarly explained. The set X is *consistent* if it contains a common extension of any two of its elements. If X is both upward-closed and consistent, then it is a *filter*. Further, X is *dense below p* if for every $q \leq p$ there is $r \leq q$ such that $r \in X$. Finally, X is *dense* if it is dense below all conditions, i.e. if every condition has an extension in X.

For the rest of the section we fix

- a countable *forcing frame* $(P, \leq, D_0, D_1, \ldots)$,
- a countable structure M interpreting a countable language L,
- a countable language $L^* \supseteq L$.

The *forcing language* is $L^*(M)$, that is, the language L^* together with the elements of M as new constants. We let $\varphi, \psi, \ldots$ range over $L^*(M)$-sentences.

2.3 Forcing relations

In principle, countless 'syntactic' forcing relations $\Vdash$ may be defined, depending on how $\Vdash$ interacts with the logical symbols. Throughout this paper we assume (first-order) formulas to be written in the logical symbols $\{\forall, \exists, \wedge, \vee, \neg\}$ and we shall restrict attention to two kinds of forcings, namely, *universal* and *existential forcings*.[3] Roughly, the choice depends on whether $\{\forall, \wedge, \neg\}$ or $\{\exists, \vee, \neg\}$ is taken as primitive while the other logical symbols are defined using the usual classical dualities. Existential forcing is widely used, but we shall see it has some disadvantages over universal forcing (cf. Remark 2.22).

Definition 2.2. A *pre-forcing* is a binary relation $\Vdash$ between conditions and $L^*(M)$-sentences. If $p \Vdash \varphi$, we say *p forces φ*. We write

$$[\varphi] := \{p \mid p \Vdash \varphi\}.$$

Definition 2.3. A pre-forcing $\Vdash$ is *universal* or *existential* if it satisfies the conditions of *universal* or *existential forcing recurrence* respectively:

	universal		existential	
$p \Vdash \neg\varphi$	iff	$\forall q \leq p : q \nVdash \varphi$	iff	$\forall q \leq p : q \nVdash \varphi$
$p \Vdash (\varphi \wedge \psi)$	iff	$p \Vdash \varphi$ and $p \Vdash \psi$	iff	$p \Vdash \neg(\neg\varphi \vee \neg\psi)$
$p \Vdash (\varphi \vee \psi)$	iff	$p \Vdash \neg(\neg\varphi \wedge \neg\psi)$	iff	$p \Vdash \varphi$ or $p \Vdash \psi$
$p \Vdash \forall x \chi(x)$	iff	$\forall a \in M : p \Vdash \chi(a)$	iff	$p \Vdash \neg\exists x \neg\chi(x)$
$p \Vdash \exists x \chi(x)$	iff	$p \Vdash \neg\forall x \neg\chi(x)$	iff	$\exists a \in M : p \Vdash \chi(a)$

Observe that a universal or existential pre-forcing is uniquely determined by its restriction to the atomic sentences of the forcing language. Solving the recurrence one sees, for universal pre-forcings, that $p \Vdash \exists x \chi(x)$ if and only if $\bigcup_{a \in M} [\chi(a)]$ is dense below p. For existential pre-forcings, $p \Vdash \forall x \chi(x)$ if and only if $[\chi(a)]$ is dense below p for all $a \in M$. Here are some further direct consequences:

Lemma 2.4. *If $\Vdash$ is a universal or an existential pre-forcing, then*

1. $p \Vdash \neg\neg\varphi$ if and only if $[\varphi]$ is dense below p.

2. (Consistency) $[\varphi] \cap [\neg\varphi] = \emptyset$.

3. $[\varphi] \cup [\neg\varphi]$ is dense.

Definition 2.5. Let $\Vdash$ be a pre-forcing and Φ be a set of $L^*(M)$-formulas.

(a) $\Vdash$ *satisfies Extension for Φ* if for every $\varphi \in \Phi$, the set $[\varphi]$ is downward-closed.

(b) $\Vdash$ *satisfies Stability for Φ* if for every $\varphi \in \Phi$ and $p \in P$, p forces φ whenever $[\varphi]$ is dense below p.

For $\Phi = L^*(M)$ we omit the reference to it.

---

[3]See [19, 3] for examples of forcings that are neither universal nor existential.

(c) $\Vdash$ is a *forcing* if it satisfies Extension and Stability for $L^*(M)$-atoms.

Lemma 2.6.

1. (Extension) *Universal and existential forcings satisfy Extension.*
2. (Stability) *Universal forcings satisfy Stability.*
3. *For a universal forcing $\Vdash$, $p \Vdash \varphi$ if and only if $[\varphi]$ is dense below p.*
4. *For a universal forcing $\Vdash$, $p \nVdash \varphi$ if and only if $q \Vdash \neg\varphi$ for some $q \leq p$.*

Proof. Extension can be shown by a straightforward induction using forcing recurrence. We prove Stability by induction on (the number of logical symbols in) φ. We can assume that φ is written in the logical base $\{\wedge, \neg, \forall\}$.

– For atomic φ, Stability is part of the definition of being a forcing.

– For the $\neg$-step, argue indirectly: if $p \nVdash \neg\varphi$, then by forcing recurrence some $q \leq p$ forces φ, so by Extension and Consistency no extension of q forces $\neg\varphi$. Hence $[\neg\varphi]$ is not dense below p.

– For the $\wedge$-step, note $[(\varphi \wedge \psi)] = [\varphi] \cap [\psi]$ by universal recurrence. If this set is dense below p then so are both $[\varphi]$ and $[\psi]$. By induction p forces both φ and ψ, and hence $p \Vdash (\varphi \wedge \psi)$ by universal recurrence.

– The $\forall$-step is similar.

Statement (3) is immediate by (1) and (2), and statement (4) follows from (3): $p \nVdash \varphi$ if and only if there is $q \leq p$ such that for all $r \leq q$, $r \nVdash \varphi$ (by (3)), if and only if there is $q \leq p$ such that $q \Vdash \neg\varphi$ (by forcing recurrence). $\square$

Example 2.7. Let $\Vdash$ be a universal forcing. A pre-forcing of obvious interest (cf. section 2.1) is:

$$p\|\varphi \text{ if and only if } p \nVdash \neg\varphi, \text{ that is, } q \Vdash \varphi \text{ for some } q \leq p.$$

To explain the notation, observe that $p\|\varphi$ if and only if $p\|q$ for some $q \in [\varphi]$ (by Extension). We have $\Vdash \subseteq \|$ by Consistency. By Stability, $p\|\neg\neg\varphi$ if and only if $p\|\varphi$. Further,

$$
\begin{array}{lll}
p\|\neg\varphi & \text{iff} \quad \exists q \leq p : q \nparallel \varphi & \\
p\|(\varphi \vee \psi) & \text{iff} \quad p\|\varphi \text{ or } p\|\psi & \text{(as existential pre-forcing)} \\
p\|\exists x \chi(x) & \text{iff} \quad \exists a \in M : p\|\chi(a) & \text{(as existential pre-forcing)}
\end{array}
$$

Remark 2.8 (Boolean valued models)**.** The last lemma has a natural topological reading. Namely, $(P, \leq)$ carries the topology whose open sets are the downward-closed sets. Then Extension means that the sets $[\varphi]$ are open and Stability means that they are regularly open (equal to the interior of their closure). The regularly open sets form a complete Boolean algebra in such a way that, for universal forcings, the map $\varphi \mapsto [\varphi]$ is a Boolean valuation.

2.4 Genericity

Let $\Vdash$ be an existential or universal forcing. Ideally, one would like to call a set generic if it intersects every dense set. As in general such sets do not exist, one has to restrict attention to those dense sets coming from a certain 'sufficiently rich' but countable Boolean algebra $\mathcal{B}(\Vdash)$.

In Cohen forcing the forcing frame is a set in M and one simply takes the algebra of its M-definable subsets (cf. Example 2.27). As M models ZF it is not surprising that this algebra is sufficiently rich. For some purposes (cf. Examples 2.29, 2.30, 2.23) already the algebra generated by the $[\varphi]$s is sufficiently rich, but not so in forcing against bounded arithmetic. There, one needs the algebra to contain sets as e.g. $\bigcup_{a \in M} \bigcap_{b \in M} [\varphi(a, b)]$. In [30, 1, 38] algebras are defined ad hoc suitable for their respective situations and there seems to be no canonical choice. That is why we padded the forcing frame by the sets $D_0, D_1, \ldots$: these sets will determine an algebra $\mathcal{B}(\Vdash)$ defined below (Definition 2.11).

Definition 2.9. A set $G \subseteq P$ is *generic* if it is a filter and intersects every dense (in P) set in $\mathcal{B}(\Vdash)$.

Our definition of $\mathcal{B}(\Vdash)$ follows Stern [45]: consider the two-sorted first-order structure (P, M). The first sort carries the forcing frame $(P, \leq, D_0, D_1, \ldots)$ and a second carries the structure M. We let individual variables $\mu, \nu, \xi, \ldots$ range over the first sort and $x, y, z, \ldots$ range over the second sort.

For each L^*-atom $\varphi = \varphi(x_1, \ldots, x_r)$ let R_φ be an $(r + 1)$-ary relation symbol of sort $P \times M^r$. The structure $(P, M)^{\Vdash}$ expands (P, M) by interpreting such a symbol R_φ by $\{p\bar{a} \in P \times M^r \mid p \Vdash \varphi(\bar{a})\}$.

We call the two-sorted first-order language of $(P, M)^{\Vdash}$ the *Stern formalism*. Using forcing recurrence it is straightforward to show:

Lemma 2.10. *For every $L^*(M)$-formula $\varphi(\bar{x})$ there is a formula $\xi \vdash \varphi(\bar{x})$ of the Stern formalism with free variables ξ and $\bar{x}$ and parameters from M that defines $\{p\bar{a} \mid p \Vdash \varphi(\bar{a})\}$ in $(P, M)^{\Vdash}$.*

Definition 2.11. The *forcing algebra* $\mathcal{B}(\Vdash)$ is the set of subsets of P that are definable in $(P, M)^{\Vdash}$.

Here and in the following, *definable* (in a certain structure) always means definable *with parameters* (from the structure).

Clearly, $\mathcal{B}(\Vdash)$ is countable. Hence, by a well-known argument,

Lemma 2.12. *Every condition is contained in some generic set.*

Sketch. Given $p \in P$, let $p_0 := p$, then choose $p_1 \leq p_0$ in the first dense set, then $p_2 \leq p_1$ in the second dense set and so on. The filter generated by the sequence $p_0, p_1, p_2, \ldots$ is generic. $\square$

Lemma 2.13. *If G is generic and $D \in \mathcal{B}(\Vdash)$ is dense below $p \in G$, then there is $q \in G \cap D$ with $q \leq p$.*

8

Proof. If D, p are as stated, then $D(p) := (D \cap \{q \mid q \leq p\}) \cup \{q \mid p \perp q\}$ is dense, and a member of $\mathcal{B}(\Vdash)$: if $D \in \mathcal{B}(\Vdash)$ is defined by $\varphi_D(\xi)$, then $D(p)$ is defined by $(\varphi_D(\xi) \wedge \xi \leq p) \vee \neg \exists \nu (\nu \leq \xi \wedge \nu \leq p)$, a formula (with parameters) of the Stern formalism. By genericity there exists an $r \in G \cap D(p)$. As $p \in G$ and G is consistent, $r \notin \{q \mid p \perp q\}$, so $r \in D \cap \{q \mid q \leq p\}$. $\qquad \square$

2.5   Generic associates

Let $\Vdash$ be a universal or existential forcing. The aim is to define for suitable $G \subseteq P$ (and our fixed structure M) an $L^*(M)$-structure $M[G]$ in such a way, that it models the following theory in the forcing language $L^*(M)$:

$$\mathrm{Th}(G) := \{\varphi \in L^*(M) \mid \exists p \in G : p \Vdash \varphi\}.$$

Obviously this cannot work in general, e.g. $\mathrm{Th}(G)$ may contradict usual first-order equality axioms. But we shall see that this is the only obstacle provided we stick to the idea that the constants from M "name" all the elements of $M[G]$. We first observe that for generic G, the theory $\mathrm{Th}(G)$ is complete and formally consistent in the following sense:

Lemma 2.14. *Let G be generic. For every $L^*(M)$-sentence φ either $\varphi \in \mathrm{Th}(G)$ or $\neg \varphi \in \mathrm{Th}(G)$, but not both.*

Proof. By Lemmas 2.4 (3) and 2.10, G intersects $[\varphi] \cup [\neg \varphi] \in \mathcal{B}(\Vdash)$. Hence $\varphi \in \mathrm{Th}(G)$ or $\neg \varphi \in \mathrm{Th}(G)$ – but not both: assume there would exist $p, q \in G$ forcing φ and $\neg \varphi$ respectively; since G is a filter and filters are consistent, there would exist r extending both p and q; by Extension, then r would force both φ and $\neg \varphi$, contradicting Consistency (of forcing). $\qquad \square$

To define $M[G]$ we rely on some elementary facts about factorizations: for a theory T in a language L containing some constant symbol, the *Herbrand term structure* $\mathfrak{T}(T)$ for T has as universe all closed L-terms, interprets a function symbol $f \in L$ by $\bar{t} \mapsto f(\bar{t})$ and interprets a relation symbol $R \in L$ by $\{\bar{t} \mid R\bar{t} \in T\}$. Note that in $\mathfrak{T}(T)$ every closed term denotes itself. A *congruence* $\sim$ on $\mathfrak{T}(T)$ is an equivalence relation on $\mathfrak{T}(T)$ such that functions in $\mathfrak{T}(T)$ (i.e. interpretations of function symbols of L) map equivalent arguments (i.e. componentwise equivalent argument tuples) to equivalent values and every relation of $\mathfrak{T}(T)$ is a union of equivalence classes of tuples. In this case, let $\mathfrak{T}(T)/\sim$ denote the L-structure induced by $\mathfrak{T}(T)$ on the $\sim$-classes in the natural way. In $\mathfrak{T}(T)/\sim$ every closed term t denotes its $\sim$-class $t/\sim$.

Fact 2.15. If $\sim_T := \{(s, t) \mid s = t \in T\}$ is a congruence on $\mathfrak{T}(T)$, then the atomic sentences true in $\mathfrak{T}(T)/\sim_T$ are precisely those contained in T.

Definition 2.16. Let $G \subseteq P$. If $\sim_{\mathrm{Th}(G)}$ is a congruence on $\mathfrak{T}(\mathrm{Th}(G))$ and every closed term of the forcing language is $\sim_{\mathrm{Th}(G)}$-congruent to a constant $a \in M$, then we say $M[G]$ *is defined* and set
$$M[G] := \mathfrak{T}(\mathrm{Th}(G))/\sim_{\mathrm{Th}(G)} .$$

If G is generic and $M[G]$ defined, then $M[G]$ is a *generic associate of M*.

A generic associate $M[G]$ of M is a *generic extension of M*, if $L = L^*$ and there is an embedding of M into $M[G]$. It is a *generic expansion of M*, if

$$a \mapsto a/\sim_{\mathrm{Th}(G)} : M \cong M[G] \upharpoonright L,$$

that is, if the map that sends each $a \in M$ to its $\sim_{\mathrm{Th}(G)}$-congruence class $a/\sim_{\mathrm{Th}(G)}$ is an isomorphism of M onto the restriction of $M[G]$ to L.

Remark 2.17. Sometimes we assume that $M[G]$ is defined for every generic G. Because this assumption is trivially satisfied in all applications we are aware of, we consider it a mere technicality and make no efforts to avoid it.

Lemma 2.18. *Let G be generic.*

1. *$M[G]$ is defined if for all closed $L^*(M)$-terms t, t', all $L^*(M)$-atoms $\varphi(x)$ and all $p \in P$*

   (a) *if $p \Vdash t = t'$, then $q \Vdash t' = t$ for some $q \leq p$,*
   (b) *if $p \Vdash \varphi(t)$ and $p \Vdash t = t'$, then $q \Vdash \varphi(t')$ for some $q \leq p$,*
   (c) *$q \Vdash t = a$ for some $q \leq p$ and $a \in M$.*

2. *If $M[G]$ is defined, then it has universe $\{a/\sim_{\mathrm{Th}(G)} | a \in M\}$.*

We omit the proof.

Theorem 2.19 (Truth Lemma)**.** *Let G be generic. If $M[G]$ is defined, then $\mathrm{Th}(M[G]) = \mathrm{Th}(G)$.*

Proof. We have to show: $M[G] \models \varphi$ if and only if $p \Vdash \varphi$ for some $p \in G$. We have two cases depending of whether $\Vdash$ is universal or existential. In both cases we proceed by induction on φ.

The case where $\Vdash$ is existential is easy. The base case follows by construction (Fact 2.15). Both the $\vee$-step and the $\exists$-step are trivial. Finally, $\neg\varphi \in \mathrm{Th}(M[G])$, i.e. $\varphi \notin \mathrm{Th}(M[G])$, is equivalent to $\varphi \notin \mathrm{Th}(G)$ by induction and thus to $\neg\varphi \in \mathrm{Th}(G)$ by Lemma 2.14.

The case where $\Vdash$ is universal is more complicated. The base case and the $\neg$-step are as in the existential case, and the $\wedge$-step is straightforward using the consistency of G. For the $\forall$-step, first assume that some $p \in G$ forces $\forall x \varphi(x)$. Then $p \Vdash \varphi(a)$ for every $a \in M$ by universal recurrence, so $M[G] \models \varphi(a)$ for every $a \in M$ by induction. Hence, $M[G] \models \forall x \varphi(x)$ by Lemma 2.18 (2). Conversely, assume $\forall x \varphi(x) \notin \mathrm{Th}(G)$. We aim to show $\varphi(a) \notin \mathrm{Th}(M[G])$ for some $a \in M$. By Lemma 2.14, $\neg\forall x \varphi(x) \in \mathrm{Th}(G)$, i.e. some $p \in G$ forces $\neg\forall x \varphi(x)$. By universal recurrence this means that for every $q \leq p$ there is $a \in M$ such that $q \not\Vdash \varphi(a)$. By Lemma 2.6 (4) this means: for every $q \leq p$ there is $a \in M$ and there is $r \leq q$ such that $r \Vdash \neg\varphi(a)$. In other words, $D := \bigcup_{a \in M} [\neg\varphi(a)]$ is dense below p. Clearly, $D \in \mathcal{B}(\Vdash)$: it is defined by $\exists x(\xi \vdash \neg\varphi(x))$, a formula (with parameters) of the Stern formalism (cf. Lemma 2.10). As $p \in G$, G intersects D by Lemma 2.13, i.e. there is some $a \in M$ such that $\neg\varphi(a) \in \mathrm{Th}(G)$. Then $\varphi(a) \notin \mathrm{Th}(G)$ by Lemma 2.14, so $\varphi(a) \notin \mathrm{Th}(M[G])$ by induction. $\square$

Corollary 2.20. *Assume $M[G]$ is defined for every generic G and let $p \in P$.*

1. *If $\Vdash$ is existential, then $p \Vdash \varphi$ implies $M[G] \models \varphi$ for every generic G containing p.*

2. *(Forcing Completeness) If $\Vdash$ is universal, then $p \Vdash \varphi$ if and only if $M[G] \models \varphi$ for every generic G containing p.*

3. *If $\Vdash$ is universal, then $\{\varphi \mid p \Vdash \varphi\}$ is closed under logical consequence.*

Proof. By the Truth Lemma $p \Vdash \varphi$ implies $M[G] \models \varphi$ for every generic G containing p. This shows (1) and the forward direction of (2). The backward direction of (2) relies on Lemma 2.6 (4) for universal forcings: if $p \not\Vdash \varphi$, there is $q \leq p$ such that $q \Vdash \neg\varphi$. By Lemma 2.12 there is a generic G containing q. By the Truth Lemma $M[G] \models \neg\varphi$, i.e. $M[G] \not\models \varphi$. Being a filter, G contains p.

To see (3) just note that the set of φ satisfying the right hand side of (2) are closed under logical consequence. $\square$

Remark 2.21. Let $\Vdash$ be a universal forcing and recall Example 2.7 and the discussion in section 2.1. Assume $M[G]$ is defined for every generic G. Then $p\|\varphi$ if and only if $M[G] \models \varphi$ for some generic G containing p. Further, $\{\varphi \mid p\|\varphi\}$ is closed under logical consequence.

Remark 2.22 (Weak forcing). Corollary 2.20 (3) fails for existential pre-forcings $\Vdash$ that are nontrivial in the sense that there exist p_0, φ_0 such that $p_0 \not\Vdash \varphi_0$ and $p_0 \not\Vdash \neg\varphi_0$. Then p_0 does not force $(\varphi_0 \vee \neg\varphi_0)$. Since this is valid $\{\varphi \mid p_0 \Vdash \varphi\}$ is not closed under logical consequence. Assuming $\Vdash$ satisfies Extension for atoms and $M[G]$ is defined for every generic G, one can show that the associated *weak forcing* $\Vdash^*$ is a universal forcing. Here, $p \Vdash^* \varphi$ if and only if $p \Vdash \neg\neg\varphi$. So, in contrast to universal forcing, existential forcing is syntax sensible (if nontrivial) and Forcing Completeness fails. These defects may be repaired when moving to the weak forcing.

Example 2.23 (Keisler forcing). Keisler [21] studies generally existential pre-forcings that satisfy Extension for atoms and the conditions in Lemma 2.18 (1), and proves Forcing Completeness for the associated weak forcing.

We have the following preservation result.

Theorem 2.24. *Let T be a universal L^*-theory. If both*

(i) *for every condition p, the theory T is consistent with*

$$\mathrm{Lit}(p) := \{\varphi \mid p \Vdash \varphi, \varphi \text{ is an } L^*(M)\text{-literal}\},$$

(ii) *and for every closed $L^*(M)$-term t, the set $\bigcup_{a \in M}[t = a]$ is dense,*

then $M[G]$ is defined for every generic G and satisfies T.

Proof. Let G be generic. To show $M[G]$ is defined we verify the three conditions (a), (b), (c) in Lemma 2.18 (1). For (a), if $p \Vdash t = t'$ but $q \nVdash t' = t$ for every $q \leq p$, then $p \Vdash \neg t' = t$ by forcing recurrence. But then $\mathrm{Lit}(p)$ and hence $\mathrm{Lit}(p) \cup T$ is inconsistent, contradicting (i). Condition (b) is similarly verified and (c) is the same as (ii).

To show $M[G] \models T$ is suffices to show that $M[G]$ embeds into a model of T (since T is universal). For this it suffices to show that $T \cup \mathrm{Diag}(M[G])$ is consistent. So let Δ be a finite subset of $\mathrm{Diag}(M[G])$. Then $\Delta \subseteq \mathrm{Th}(G)$ by the Truth Lemma, that is, every literal $\lambda \in \Delta$ is forced by some $p_\lambda \in G$. Since G is consistent it contains a common extension p of all the p_λ's. Then $\Delta \subseteq \mathrm{Lit}(p)$ by Extension and $T \cup \Delta$ is consistent by (i). $\qquad\square$

2.6   Conservative forcing

Let $\Vdash$ be an existential or universal forcing. Which forcings produce generic expansions? We characterize these as follows.

Definition 2.25. The forcing $\Vdash$ is *conservative* if for every condition p and every atomic $L(M)$-sentence φ (i.e. without a symbol from $L^* \setminus L$):

$$p \Vdash \varphi \text{ if and only if } M \models \varphi.$$

Proposition 2.26. *If $\Vdash$ is conservative, then every generic associate is a generic expansion. The converse holds true in case $\Vdash$ is universal and $M[G]$ is defined for every generic G.*

Proof. Let $M[G]$ be a generic associate of M. The map $a \mapsto a/\sim_{\mathrm{Th}(G)}$ is a surjection from M onto $M[G] \upharpoonright L$ (Lemma 2.18). If it is not an isomorphism, then $\mathrm{Th}(M)$ and $\mathrm{Th}(M[G])$ disagree on some atomic $L(M)$-sentence. As $\mathrm{Th}(M[G]) = \mathrm{Th}(G)$ by the Truth Lemma, this contradicts conservativity.

For the second statement, assume $\Vdash$ is not conservative. Using Lemma 2.6 (4) one finds a condition p and an $L(M)$-literal φ such that $p \Vdash \varphi$ and $M \not\models \varphi$. For a generic G containing p (Lemma 2.12) then $M[G] \models \varphi$ by the Truth Lemma. Then $\varphi \in \mathrm{Th}(M[G] \upharpoonright L) \setminus \mathrm{Th}(M)$, so $M[G]$ cannot be an expansion of M. $\qquad\square$

2.7   Some examples

Cohen forcing can be viewed as a natural special case:

Example 2.27 (Cohen forcing)**.** Cohen (set-) forcing starts with a countable transitive standard model M of, say, ZF+GCH and wants $M[G]$ to be an extension of M. In particular $L^* = L = \{\in\}$. Different forcing extensions are obtained by different choices of $(P, \leq)$, a set in M, while the forcing $\Vdash_{\mathrm{Co}}$ is kept fixed. Following e.g. [28] one can define this forcing by universal forcing recurrence stipulating for atoms:

$p \Vdash_{\mathrm{Co}} a \in b$ iff $\{q \mid \exists r \exists c \, ((c, r) \in b \wedge q \leq r \wedge q \Vdash_{\mathrm{Co}} a = c)\}$ is dense below p,

$p \Vdash_{\mathrm{Co}} a = b$ iff $\forall c \in \mathrm{dom}(a \cup b) \, \forall q \leq p \, (q \Vdash_{\mathrm{Co}} c \in a \leftrightarrow q \Vdash_{\mathrm{Co}} c \in b)$.

It is not hard to show that this uniquely determines a universal pre-forcing. The technicality of the definition is to ensure that it is a forcing. Genericity is defined to mean: intersect every dense set that is definable in M. This coincides with our notion for $\emptyset = D_0 = D_1 = \ldots$.

In set theory one defines $M[G]$ as follows: $\in$ is interpreted by itself and the constants $a \in M$ are interpreted by $a_G := \{b_G \mid \exists p \in G : (b, p) \in a\}$. Then $M[G]$ is an extension of M. For this definition of $M[G]$ one can show the Truth Lemma for atoms: $a_G = b_G$ if and only if $a \sim_{\mathrm{Th}(G)} b$ and $a_G \in b_G$ if and only if $\mathrm{Th}(G)$ contains the atom $a \in b$. It follows that $M[G]$ in our sense is defined for every generic G. Second, $M[G]$ in our sense is isomorphic to $M[G]$ in the sense of set theory. Indeed, $\{(a/\sim_{\mathrm{Th}(G)}, a_G) \mid a \in M\}$ is an isomorphism.

Feferman has been the first to explicitly use forcing outside set theory, namely to adress questions in computability theory. But already Cantor's back and forth method can be seen as a forcing argument. Both are conservative forcings:

Example 2.28 (Cantor's Theorem). We give this simple example in some detail, because it reappears in similar form in section 4.

Let $M = (A, A')$ be a countable two-sorted structure where the two sorts A and A' carry dense linear orders without endpoints $\preceq$ and $\preceq'$ respectively (i.e. $L = \{\preceq, \preceq'\}$). Set $L^* := L \cup \{R\}$ for a new binary relation symbol R.

The forcing frame $(P, \leq, D_0, D_1, \ldots)$ is defined as follows: P is the set of all finite partial isomorphisms between A and A'; take $p \leq q$ to mean $p \supseteq q$; finally the sets $D_0, D_1, \ldots$ enumerate the sets $\{p \mid a \in \mathrm{dom}(p)\}, \{p \mid a' \in \mathrm{im}(p)\}$ for $a \in A, a' \in A'$. Each of these sets is dense.

To define a conservative universal pre-forcing $\Vdash_{\mathrm{Ca}}$ it suffices to define $p \Vdash_{\mathrm{Ca}} \varphi$ for φ an atom of the form Rab. Take this to mean $(a, b) \in p$. Then $\Vdash_{\mathrm{Ca}}$ is a forcing: that $\Vdash_{\mathrm{Ca}}$ satisfies Extension for atoms is obvious. Since $\Vdash_{\mathrm{Ca}}$ is conservative we only have to verify Stability for atoms Raa' with $a \in A, a' \in A'$: if $p \nVdash_{\mathrm{Ca}} Raa'$, then $(a, a') \notin p$; choose $b' \neq a'$ such that $q := p \cup \{(a, b')\}$ is a condition ($\mathrm{im}(p)$ is finite); then $q \leq p$ and no extension of q contains (a, a'), so no extension of q forces Raa'; hence $[Raa']$ is not dense below p.

It is easy to see that $M[G]$ is defined for every generic G (e.g. by Lemma 2.18 (1)). By Proposition 2.26 every generic associate $M[G]$ is a generic expansion of M, that is, $a \mapsto a/\sim : M \cong M[G] \restriction L$; here we write $\sim$ for $\sim_{\mathrm{Th}(G)}$. By definition, $M[G]$ interprets R by

$$\{(a/\sim, b/\sim) \mid \exists p \in G : p \Vdash_{\mathrm{Ca}} Rab\} = \{(a/\sim, b/\sim) \mid (a, b) \in \bigcup G\}.$$

Thus $a \mapsto a/\sim : (M, \bigcup G) \cong M[G]$. This and the fact that G intersects all the sets $D_0, D_1, \ldots$, implies $\bigcup G : (A, \preceq) \cong (A', \preceq')$.

Example 2.29 (Feferman forcing). In [17] Feferman considers $M = \mathbb{N}$ interpreting the language L that has relation symbols for the graphs of successor, addition and multiplication. L^* expands L by at most countably many unary predicate symbols. A condition p is a finite consistent set of literals in the new predicates $L^* \setminus L$ and constants from $\mathbb{N}$. A condition p extends another q if $p \supseteq q$. For the sets $D_0, D_1, \ldots$ choose, say, always $\emptyset$. Feferman defines a conservative existential pre-forcing $\Vdash_{\mathrm{Fe}}$ by letting p force an atom involving a new predicate

if and only if the atom belongs to p. It is not hard to see that $\Vdash_{\mathrm{Fe}}$ is a forcing and that $M[G]$ is defined for every generic G. Applications of Feferman forcing in computability theory are surveyed in [29].

Fenner et al. [18] generalize Feferman forcing for the case where $L^* = L \cup \{R\}$ for one new unary predicate R. View a Feferman condition p as the set of functions in $\{0, 1\}^{\mathbb{N}}$ that map $n \in \mathbb{N}$ to 0 or 1 whenever $Rn \in p$ or $\neg Rn \in p$ respectively. Now, instead of using these basic clopen sets as conditions, [18] use perfect sets in $\{0, 1\}^{\mathbb{N}}$. Forcing frames considered in [18] are certain subframes of this forcing frame (cf. [18, Definition 3.3]). Straightforwardly, Fenner et al. let a perfect set p force an atom Rn if and only if every function in p maps n to 1. This determines a conservative existential pre-forcing, that is actually a forcing on the frames considered. For various frames, [18] studies complexity classes relativized by R in generic expansions.

Finally, we mention Robinson forcing in model theory:

Example 2.30 (Finite Robinson forcing)**.** We degrade M to a countably infinite set of constants, i.e. we let $L = \emptyset$. Further, let L^* be a countable language and T be a consistent L^*-theory; $T_\forall$ is the set of universal consequences of T.

The forcing frame $(P, \leq, D_0, D_1, \ldots)$ is defined as follows. A condition p is a finite set of $L^*(M)$-literals that is consistent with $T_\forall$, and again, $p \leq q$ means $p \supseteq q$. The sets $D_0, D_1, \ldots$ enumerate the sets $\bigcup_{a \in M}[t = a]$ for closed $L^*(M)$-terms t. It is easy to see that these sets are dense.

For atomic φ we let $p \Vdash_{\mathrm{Ro}} \varphi$ if and only if $T_\forall \cup p \vdash \varphi$ (slighty deviating from [19]). This determines an existential pre-forcing $\Vdash_{\mathrm{Ro}}$ that is easily seen to be a forcing. By Theorem 2.24, $M[G]$ is defined for every generic G and satisfies $T_\forall$. Note $\bigcup G$ is roughly the same as $\mathrm{Diag}(M[G])$. Hence the Truth Lemma essentially says, that generic associates are *finitely generic for T*, so in particular such structures exist ([19, Theorem 5.11]). Their theory can be seen as a generalized model-companion for T. We refer to [19, 21] for more information and applications.

2.8   Notes

Forcing has been developed in many different settings ([20, 3] survey some), and the development here follows these known lines. We refer to the examples in section 2.7 for a comparison with some of them.

Forcing against bounded arithmetic has been developed by Takeuti and Yasumoto [46, 47] following not Cohen's original method but its reformulation by Scott and Solovay [40] as a method to construct Boolean valued models (cf. Remark 2.8). Scott [40] describes such a model for a 3rd order theory of the reals by interpreting the language over real valued random variables.

Krajíček [26] develops such *forcing with random variables* in full detail as a method to study bounded arithmetics by using algorithmically restricted random variables.

3 Principal theorems

In set theory (Example 2.27) independence results are based on the "Principal Theorem" [43] stating that every generic extension $M[G]$ of a countable model M of ZF again models ZF. In bounded arithmetic one is often interested in constructing generic expansions of a countable nonstandard model M of true arithmetic. One then needs the generic expansions to model some bounded arithmetic, i.e. certain least number principles.

In this section we fix

- a countable forcing frame $(P, \leq, D_0, D_1, \ldots)$
- a conservative universal forcing $\Vdash$,
- an ordered countable L-structure M satisfying the least number principle (defined below).
- a countable language $L^* \supset L$.

A model is *ordered* if it interprets the symbol $<$ by some linear order on its universe. Given an ordered model N and $b_0 \in N$, the quantifiers $\forall x < b_0$ and $\exists x < b_0$ are called b_0-*bounded*.

Remark 3.1. Due to conservativity, forcing recurrence works for bounded quantifiers as it does for unbounded quantifiers:

$$
\begin{aligned}
p \Vdash \forall x < b_0 \chi(x) &\quad \text{iff} \quad \forall a <^M b_0 : p \Vdash \chi(a), \\
p \Vdash \exists x < b_0 \chi(x) &\quad \text{iff} \quad p \Vdash \neg \forall x < b_0 \neg \chi(x).
\end{aligned}
$$

Note, $p \Vdash \exists x < b_0 \chi(x)$ if and only if $\bigcup_{a <^M b_0} [\chi(a)]$ is dense below p.

Definition 3.2. Let N be an ordered model, $b_0 \in N$ and Φ be a set of formulas in the language of N with parameters from N.

(a) N satisfies the *least number principle for* Φ if every nonempty subset of its universe that is definable by a formula in Φ has a $<^N$-least element.

(b) N satisfies the least number principle for Φ *up to* b_0 if it satisfies the least number principle for $\{(\varphi(x) \wedge x < b_0) \mid \varphi(x) \in \Phi\}$.

We omit reference to Φ, if it is the set of all formulas in the language of N with parameters from N.

3.1 Partial definability

Recall the notation $p \| \varphi$ from Example 2.7 and Remark 2.21.

Definition 3.3. Let $b_0 \in M$ and $\varphi = \varphi(\bar{x})$ be an $L^*(M)$-formula.

(a) $\Vdash$ is *definable for* φ if for every $p \in P$ the set $\{\bar{a} \mid p \| \varphi(\bar{a})\}$ is definable in M.

(b) $\Vdash$ is *densely definable for φ up to b_0* if for every $p \in P$ there is $q \leq p$ such that $\{\bar{c} <^M b_0 \mid q \| \varphi(\bar{c})\}$ is definable in M.

We say $\Vdash$ is (densely) definable (up to b_0) for a set Φ of $L^*(M)$-formulas if $\Vdash$ is (densely) definable (up to b_0) for every $\varphi \in \Phi$.

Here, for $\bar{c} = c_1 \cdots c_k$ by $\bar{c} <^M b_0$ we mean $c_i <^M b_0$ for every $1 \leq i \leq k$.

Although we are not going to use it, we include the following simple observation to illustrate the definition.

Proposition 3.4. *Let $b_0 \in M$ and Φ be a set of $L^*(M)$-formulas that is closed under negations. Then*

1. $\Vdash$ *is definable for Φ if and only if for every $\varphi(x) \in \Phi$ and $p \in P$ the set $\{\bar{c} \mid p \Vdash \varphi(\bar{c})\}$ is definable in M.*

2. $\Vdash$ *is densely definable for Φ up to b_0 if and only if for every $\varphi(x) \in \Phi$ and $p \in P$ there is $q \leq p$ such that $\{\bar{c} <^M b_0 \mid q \Vdash \varphi(\bar{c})\}$ is definable in M.*

Proof. Forward: note $p \Vdash \varphi$ if and only if $p \Vdash \neg\neg\varphi$ (by Stability), if and only if $p \nVdash \neg\varphi$. Backward: note $p \| \varphi$ if and only if $p \nVdash \neg\varphi$. $\qquad\qquad\square$

Recall that generic associates are expansions (Proposition 2.26).

Theorem 3.5 (Principal). *Let $b_0 \in M$ and Φ be a set of $L^*(M)$-formulas. If $\Vdash$ is densely definable for Φ up to b_0, then every generic expansion of M satisfies the least number principle for Φ up to b_0.*

In particular, if $\Vdash$ is definable for Φ, then every generic expansion of M satisfies the least number principle for Φ.

Proof. The second statement follows from the first noting that definability implies dense definability up to any $b_0 \in M$. To prove the first, let $M[G]$ be a generic expansion of M and $\varphi(x) \in \Phi$ be such that $M[G] \models \exists x < b_0 \varphi(x)$. We look for a least element in the set defined by $\varphi(x)$ in $M[G]$. It suffices to find $a <^M b_0$ such that $M[G] \models \varphi(a)$ and $M[G] \not\models \varphi(b)$ for every $b <^M a$. Define
$$D_\varphi := \bigcup\nolimits_{a <^M b_0} \bigcap\nolimits_{b <^M a} [(\varphi(a) \wedge \neg\varphi(b))].$$

Claim D_φ is dense below every condition forcing $\exists x < b_0 \varphi(x)$.

Proof of Claim Given p forcing $\exists x < b_0 \varphi(x)$ we are looking for some $q \leq p$ in D_φ. By universal recurrence $\bigcup_{a \in M} [a < b_0 \wedge \varphi(a)]$ is dense below p. By conservativity each set $[a < b_0 \wedge \varphi(a)]$ equals $[\varphi(a)]$ or $\emptyset$ depending on whether $a <^M b_0$ or not. Hence $\bigcup_{a <^M b_0} [\varphi(a)]$ is dense below p, so for some $b <^M b_0$ there is an extension $q_b \leq p$ forcing $\varphi(b)$.

Dense definability applied to $\varphi \in \Phi$ and $q_b \in P$ gives some $\tilde{q} \leq q_b$ such that
$$C := \{c <^M b_0 \mid \tilde{q} \nVdash \neg\varphi(c)\}$$

16

is definable in M. By Extension $\tilde{q} \Vdash \varphi(b)$, so $\tilde{q} \nVdash \neg\varphi(b)$ by Consistency. Hence $b \in C$, so $C \neq \emptyset$. Because M satisfies the least number principle, C has a least element $a \leq^M b <^M b_0$. As $a \in C$ we have $\tilde{q} \nVdash \neg\varphi(a)$, so by forcing recurrence we find $q_a \leq \tilde{q}$ forcing $\varphi(a)$. Then $q_a \leq \tilde{q} \leq q_b \leq p$. To show $q_a \in D_\varphi$, it suffices to show $q_a \Vdash \neg\varphi(b')$ for every $b' <^M a$. But any $b' <^M a \leq^M b <^M b_0$ is not in C by minimality of a, so $\tilde{q} \Vdash \neg\varphi(b')$ and hence also $q_a \Vdash \neg\varphi(b')$ by Extension. $\quad \dashv$

Choose $p_0 \in G$ forcing $\exists x < b_0 \varphi(x)$ by the Truth Lemma. First note $D_\varphi \in \mathcal{B}(\Vdash)$ as it is defined by the following formula (with parameters) of the Stern formalism (cf. Lemma 2.10):

$$\exists x(x < b_0 \wedge \forall y(y < x \to (\xi \vdash (\varphi(x) \wedge \neg\varphi(y))))).$$

The claim and Lemma 2.13 imply that there is a condition $p \in G \cap D_\varphi$. Hence there is $a <^M b_0$ such that for every $b <^M a$ we have $p \Vdash (\varphi(a) \wedge \neg\varphi(b))$. By the Truth Lemma $M[G] \models \varphi(a)$ and $M[G] \models \neg\varphi(b)$ for every $b <^M a$. Thus a is a least element as we are looking for. $\quad \square$

Here is a dual formulation of the Principal Theorem:

Corollary 3.6. *Let $b_0 \in M$ and Φ be a set of $L^*(M)$-formulas. If for every $\varphi(\bar{x}) \in \Phi$ and $p \in P$ there is $q \leq p$ such that*
$$\{\bar{c} <^M b_0 \mid q \Vdash \varphi(\bar{c})\}$$
is definable in M, then every generic expansion of M satisfies transfinite induction *for Φ up to b_0, that is, for every $\varphi(x) \in \Phi$ the sentence*

$$\forall y < b_0(\forall z < y \varphi(z) \to \varphi(y)) \to \forall x < b_0 \varphi(x).$$

Proof. The assumption implies that $\Vdash$ is densely definable for $\neg\Phi$ up to b_0 (see the proof of Proposition 3.4). Now observe that the least number principle for $\neg\Phi$ up to b_0 is equivalent to transfinite induction for Φ up to b_0. $\quad \square$

Remark 3.7. Assume P is definable in the sense that there is a first-order interpretation of $(P, \leq)$ in M. If $\Vdash$ is definable for $L^*(M)$-atoms, then an easy induction shows that $\Vdash$ is definable for all $L^*(M)$-formulas; by the Principal Theorem then every generic expansion of M satisfies the least number principle.

Example 3.8. In set theory Cohen forcing (Example 2.27) or Easton forcing use definable forcing frames. In arithmetic, Feferman forcing (Example 2.29) uses definable forcing frames. This is due to the fact that it starts with the standard model. Simpson [44] gives an example of a definable forcing frame starting with a nonstandard model of arithmetic. In [22] Knight pads M with some additional sorts such that $(P, \leq)$ becomes definable.

Lemma 3.9.

1. *Let Ψ be the set of $L^*(M)$-formulas φ such that $\Vdash$ is definable for φ. Then Ψ is closed under disjunctions and existential quantification.*

2. *Let $b_0 \in M$ and Ψ be the set of $L^*(M)$-formulas φ such that $\Vdash$ is densely definable for φ up to b_0. Then Ψ is closed under disjunctions and b_0-bounded existential quantification.*

Proof. (1) and closure under disjunction in (2) follow easily from the recurrence in Example 2.7. We show closure under b_0-bounded existential quantification in (2): let $\varphi(y\bar{x}) \in \Psi$ and $p \in P$. We are looking for $q \leq p$ such that $\{\bar{a} <^M b_0 \mid q \Vdash \exists y < b_0 \varphi(y\bar{a})\}$ is definable in M. Because $\varphi \in \Psi$ we find $q \leq p$ such that $\{a\bar{a} <^M b_0 \mid q \Vdash \varphi(a\bar{a})\}$ is definable in M. Then also

$$\{\bar{a} <^M b_0 \mid \exists a <^M b_0 : q \Vdash \varphi(a\bar{a})\}$$

is definable in M. By conservativity $a <^M b_0$ is equivalent with $s \Vdash a < b_0$ for any condition s. Hence the above set equals

$$\{\bar{a} <^M b_0 \mid \exists a \in M : q \Vdash (a < b_0 \wedge \varphi(a\bar{a}))\},$$

and this is the set we want (see the recurrence in Example 2.7). $\qquad\square$

3.2 Definable antichains

We sketch a method to establish dense definability. We are going to apply it in the next section. The method is intended for the typical situation where P is an (in general undefinable) subset of M and there are $L(M)$-formulas $\varphi(x, y), \psi(x, y)$ such that for all $p, q \in P$

$$(p \leq q \iff M \models \varphi(p, q)) \quad \text{and} \quad (p \| q \iff M \models \psi(p, q)).$$

In this case, the following two lemmas reduce dense definability of forcing to the definability of predense antichains refining given definable antichains.

We recall some standard forcing terminology: an *antichain* is a set of pairwise incompatible conditions. An antichain A is *maximal in* $X \subseteq P$ if $A \subseteq X$ and every $p \in X$ is compatible with some element of A. A set $X \subseteq P$ is *predense* (*below p*) if every condition (extending p) is compatible with some condition in X. E.g. an antichain is predense if and only if it is maximal in P. We write

$$X \downarrow q := \{p \in X \mid p \leq q\} \quad \text{and} \quad X \downarrow Y := \bigcup_{q \in Y} X \downarrow q.$$

The method is based on the simple observation that in order to define the forcing for some φ it suffices to define a maximal antichain in $[\varphi]$:

Lemma 3.10. *If $p \leq q$ and X is a maximal antichain in $[\varphi] \downarrow q$, then $p \| \varphi$ if and only if p is compatible with some condition in X.*

Proof. If $p \| \varphi$, then there is $r \in [\varphi]$ extending p. Then $r \in [\varphi] \downarrow q$ since $r \leq p \leq q$. By maximality of X, r is compatible with some condition in X, and hence, as $r \leq p$, so is p. The converse is immediate by Extension. $\qquad\square$

To find maximal antichains we intend to proceed by induction on φ. How to get, say, a maximal antichain in $[\neg\varphi]$ from a maximal antichain X in $[\varphi]$? This is easy if one finds a predense antichain A extending X – simply take $A \setminus X$. More generally,

Lemma 3.11. *Let A be a predense antichain and let Φ_A be the set of $L^*(M)$-sentences φ such that $A \subseteq [\varphi] \cup [\neg\varphi]$.*

1. *If $\varphi \in \Phi_A$, then $A \cap [\varphi]$ is a maximal antichain in $[\varphi]$.*

2. *If $\varphi, \psi \in \Phi_A$, then $\neg\varphi, (\varphi \wedge \psi) \in \Phi_A$. In this case $A \cap [\neg\varphi] = A \setminus (A \cap [\varphi])$ and $A \cap [\varphi \wedge \psi] = (A \cap [\varphi]) \cap (A \cap [\psi])$.*

3. *If $b_0 \in M$ and $\chi(a) \in \Phi_A$ for all $a <^M b_0$, then $\forall x < b_0\chi(x) \in \Phi_A$. In this case $A \cap [\forall x < b_0\chi(x)] = \bigcap_{a<^M b_0}(A \cap [\chi(a)])$.*

Proof. We only show (1). Obviously $A \cap [\varphi]$ is an antichain in $[\varphi]$. To see maximality, note that any $p \in [\varphi]$ is compatible with some condition in A by predensity, and since such a condition cannot be in $[\neg\varphi]$ (by Extension and Consistency), it must be in $[\varphi]$. $\square$

Given some fixed antichain A, the previous lemma describes how to define the forcing for $\varphi \in \Phi_A$ by induction on φ. The following, more versatile lemma describes how one can do when constructing suitable antichains on the way. What one needs in every step is a predense antichain that *refines* the current antichain in the following sense:

Definition 3.12. For $X, Y \subseteq P$ we say X *refines* Y if every condition in X that is compatible with some condition in Y already extends some condition in Y.

Lemma 3.13. *Let φ, ψ be $L^*(M)$-sentences, $\chi(x)$ an $L^*(M)$-formula, $b_0 \in M$ and $p \in P$.*

1. *If X is a maximal antichain in $[\varphi] \downarrow p$, and $A \subseteq P \downarrow p$ is an antichain that is predense below p and refines X, then $A \setminus (A \downarrow X)$ is a maximal antichain in $[\neg\varphi] \downarrow p$.*

2. *If X and Y are maximal antichains in $[\neg\varphi] \downarrow p$ and $[\neg\psi] \downarrow p$ respectively, and $A \subseteq P \downarrow p$ is an antichain that is predense below p and refines $X \cup Y$, then $A \setminus (A \downarrow (X \cup Y))$ is a maximal antichain in $[\varphi \wedge \psi] \downarrow p$.*

3. *If for every $a <^M b_0$, the set X_a is a maximal antichain in $[\neg\chi(a)] \downarrow p$, and $A \subseteq P \downarrow p$ is an antichain that is predense below p and refines $\bigcup_{a<^M b_0} X_a$, then $A \setminus (A \downarrow \bigcup_{a<^M b_0} X_a)$ is a maximal antichain in $[\forall x < b_0\chi(x)] \downarrow p$.*

Proof. We only show (3). Obviously, $A' := A \setminus (A \downarrow \bigcup_{a<^M b_0} X_a)$ is an antichain in $P \downarrow p$. To show $A' \subseteq [\forall x < b_0\chi(x)]$, we assume $q \in A \setminus [\forall x < b_0\chi(x)]$ and show $q \in A \downarrow \bigcup_{a<^M b_0} X_a$. Since $q \not\Vdash \forall x < b_0\chi(x)$ there are $a_0 <^M b_0$ and $r \leq q$ such that $r \Vdash \neg\chi(a_0)$ (Remark 3.1 and Lemma 2.6 (4)). By maximality of X_{a_0}, the condition r, and hence also q, is compatible with some condition in $X_{a_0} \subseteq \bigcup_{a<^M b_0} X_a$. Since $q \in A$ and A refines $\bigcup_{a<^M b_0} X_a$, we get $q \in A \downarrow \bigcup_{a<^M b_0} X_a$.

To see that A' is maximal, let $q \leq p$ force $\forall x < b_0\chi(x)$. Then q is compatible with some $r \in A$ since A is predense below p. We claim $r \in A'$, i.e. $r \notin A \downarrow \bigcup_{a<^M b_0} X_a$. But otherwise $r \Vdash \neg\chi(a_0)$ for some $a_0 <^M b_0$ (by Extension) while $q \Vdash \chi(a_0)$ (by Remark 3.1), so r and q would not be compatible (by Extension and Consistency). $\square$

19

4 Forcing against bounded arithmetic

In this section we fix

- a countable language L containing $\{+, \cdot\, , 0, 1, <\}$.
- a countable L-structure M that is a proper elementary extension of an L-expansion of $(\mathbb{N}, +, \cdot\, , 0, 1, <)$.
- $L^* := L \cup \{R\}$ for a new binary relation symbol $R \notin L$.

We shall use the following notation. For $n \in M$ we write

$$[n] := \{a \in M \mid a <^M n\}.$$

A relation R over M is *bounded (in M)* if there is $b \in M$ such that any component of any tuple in R is $<^M b$. As $\mathbb{N}$ codes every bounded (in $\mathbb{N}$) relation by an element, M codes every definable bounded (in M) relation by an element. If $m \in \mathbb{N}$ is such a code we let $\|m\|$ denote the cardinality of the coded relation. This is not to be confused with $|m|$ where $|m| = \lceil \log(m+1) \rceil$ for $m > 0$ and $|0| = 0$. Using the definitions of these functions in the standard model $(\mathbb{N}, +, \cdot\, , 0, 1, <)$, we get corresponding functions $\| \cdot \|^M$ and $| \cdot |^M$ in M and we shall omit the superscripts.

For arbitrary $n, m \in M$

$$n <^M m^{o(1)}$$

means that $n^\ell <^M m$ for every $\ell \in \mathbb{N}$.

4.1 Paris-Wilkie forcing

Let $n \in M$ be nonstandard, i.e. such that $[n]$ is infinite. We define a forcing frame $(P, \leq, D_0, D_1, \ldots)$ as follows. Note that every (standard-)finite bijection from a subset of $[n+1]$ onto a subset of $[n]$ is coded by an element in M. We let P be the set of all these codes. Note that P is not definable in M. As partial order we use $p \leq q$ if and only if $p \supseteq q$. Here, and below, we blur the distinction between p and the bijection coded. The family $D_0, D_1, \ldots$ enumerates the sets $\{p \mid b \in \mathrm{dom}(p)\}, \{p \mid c \in \mathrm{im}(p)\}$ for $b \in [n+1], c \in [n]$. It is easy to see that these sets are dense.

To determine a universal pre-forcing $\Vdash_{\mathrm{PW}}$ it suffices to define $p \Vdash_{\mathrm{PW}} \varphi$ for atoms φ. We want a conservative forcing, so it suffices to define $p \Vdash_{\mathrm{PW}} \varphi$ for φ an $L^*(M)$-atom that is not an $L(M)$-atom. Such an atom has the form Rst for closed $L(M)$-terms s, t. We set

$$p \Vdash_{\mathrm{PW}} Rst \iff (s^M, t^M) \in p.$$

It is straightforward to check that $\Vdash_{\mathrm{PW}}$ is a forcing and that $M[G]$ is defined for every generic G (cf. Example 2.28).

Lemma 4.1. $\Vdash_{\mathrm{PW}}$ *is definable for quantifier free $L^*(M)$-formulas.*

We prove this exemplifying the antichain method from section 3.2. However, a direct proof would be equally easy. We are in the "typical situation" that we have $L(M)$-formulas $\varphi(x, y), \psi(x, y)$ such that for all $p, q \in P$

$$(p \leq q \Longleftrightarrow M \models \varphi(p, q)) \quad \text{and} \quad (p \| q \Longleftrightarrow M \models \psi(p, q)).$$

E.g. $\psi(x, y)$ is a formula expressing that both x and y code partial bijections that agree on arguments on which they are both defined.

Lemma 4.1. Let $\varphi = \varphi(\bar{x})$ be a quantifier free $L^*(M)$-formula. For $\bar{c}$ from M let $T(\bar{c})$ be the set of those $a \in M$ that are denoted by some closed term in $\varphi(\bar{c})$. Further, let $A_{\bar{c}}$ be the set of all minimal partial bijections p such that $\text{dom}(p)$ contains $T(\bar{c}) \cap [n + 1]$ and $\text{im}(p)$ contains $T(\bar{c}) \cap [n]$. As $T(\bar{c})$ is finite, $A_{\bar{c}} \subseteq P$, and it is easily seen to be a predense antichain. We use Lemma 3.11: observe the set $\Phi_{A_{\bar{c}}}$ defined there contains all atomic subsentences of $\varphi(\bar{c})$ and is closed under Boolean combinations.

First write an $L(M)$-formula $\alpha(z, \bar{x})$ such that $\alpha(z, \bar{c})$ defines $A_{\bar{c}}$ (in M). Then, for every atomic subformula $\psi(\bar{x})$ of $\varphi(\bar{x})$, write a formula $\tilde{\psi}(z, \bar{x})$ such that $\tilde{\psi}(z, \bar{c})$ defines $A_{\bar{c}} \cap [\psi(\bar{c})]$. This is a maximal antichain in $[\psi(\bar{c})]$ by Lemma 3.11. Using the recurrence of this lemma, one can find such a formula for every Boolean combination of such $\psi(\bar{x})$s, and in particular for $\varphi(\bar{x})$. Lemma 3.10 (for $q = \emptyset$) then implies that $\Vdash_{\text{PW}}$ is definable for $\varphi(\bar{x})$. $\square$

Theorem 4.2 (Paris, Wilkie 1985). *Let $n \in M$ be nonstandard. Then M has an L^*-expansion (M, R^M) such that R^M is a bijection from $[n + 1]$ onto $[n]$ and (M, R^M) satisfies the least number principle for existential $L^*(M)$-formulas.*

Proof. Choose a generic G (Lemma 2.12). Up to isomorphism, then $M[G]$ expands M by a bijection R^M from $[n + 1]$ onto $[n]$ (cf. Example 2.28). By Lemmas 4.1 and 3.9, $\Vdash_{\text{PW}}$ is definable for existential $L^*(M)$-formulas. Now the result follows from the Principal Theorem. $\square$

4.2 Riis forcing

In this section, we assume that L also contains a function symbol $-$ for subtraction (cut off at 0) and some function symbols $lh(x)$ and $(x)_y$ for sequence coding such that in the "standard" model every finite sequence is of the form $((n)_0, \ldots, (n)_{lh(n)-1})$ for some $n \in \mathbb{N}$.

Let $\Delta_0^{b_0}(R)$ denote the closure of the set of quantifier-free $L^*(M)$-formulas by b_0-bounded quantification, i.e. $\exists x < b_0$ and $\forall x < b_0$ (cf. page 15). If we additionally allow (unrestricted) existential quantification we get the set $\Sigma_1^{b_0}(R)$.

Theorem 4.3 (Riis 1993). *Let $b_0, n \in M$ be such that $1 <^M b_0 <^M n^{o(1)}$. Then M has an L^*-expansion (M, R^M) such that R^M is a bijection from M onto $[n]$ and (M, R^M) satisfies the least number principle for $\Sigma_1^{b_0}(R)$.*

Remark 4.4. The reader familiar with bounded arithmetic will notice the following. Use Buss' language for L and choose n and b_0 such that both $b_0 <^M n^{o(1)}$ and $|n| <^M b_0^{o(1)}$. By

the second inequality $M \models |t(n)| < b_0$ for every (parameter free) L-term $t(x)$ and hence $\Sigma_1^{b_0}(R)$ includes all $\Sigma_1^b(R)$ formulas with parameters bounded by some L-term in n. Thus the corresponding cut in the expansion above carries a model of $T_2^1(R)$.

We refer to section 5 for an application of Theorem 4.3 in proof complexity.

Let n, b_0 satisfy the assumption of the theorem.

Definition 4.5. A relation X over M is *small* if it is empty or there are $\ell \in \mathbb{N}$ and an $L(M)$-definable surjection from $[b_0]^\ell$ onto X.

We define the forcing. Note that every small relation is bounded in M because this holds in the standard model for any choice of standard b_0. In particular, every small bijection from a subset of M onto a subset of $[n]$ is $L(M)$-definable and bounded in M, and hence coded by an element of M. Let $P \subseteq M$ be the set of all these codes. Again we set $p \leq q$ if $p \supseteq q$, and let $D_0, D_1, \ldots$ enumerate the sets $\{p \mid a \in \mathrm{dom}(p)\}, \{p \mid c \in \mathrm{im}(p)\}$ for $a \in M, c \in [n]$. These sets are dense: note that both the domain and range of $p \in P$ are small; but neither M nor $[n]$ are small (by the assumption that $b_0 <^M n^{o(1)}$), so both $(M \setminus \mathrm{dom}(p))$ and $([n] \setminus \mathrm{im}(p))$ are infinite.

The forcing relation is defined as in the previous section: $p \Vdash_{\mathrm{Ri}} Rst$ if and only if $(s^M, t^M) \in p$. This determines a conservative universal pre-forcing, and in fact a forcing (cf. Example 2.28).

Lemma 4.6. $\Vdash_{\mathrm{Ri}}$ *is definable for $\Delta_0^{b_0}(R)$-formulas.*

Proof. This follows by a slight modification of the proof of Lemma 4.1: define $T(\bar{c})$ to be the set of $a \in M$ denoted by some closed term $t(\bar{d})$ obtained from a term $t(\bar{y})$ in $\varphi(\bar{c})$ by substituting some values $\bar{d} <^M b_0$ for its free variables $\bar{y}$. Observe that $T(\bar{c})$ is small. One can then proceed as in Lemma 4.1. $\square$

In fact, Lemma 4.6 holds for $\Sigma_1^{b_0}(R)$-formulas, but the version stated is sufficient to derive the theorem:

Theorem 4.3. Clearly, $M[G]$ is defined for every generic G, and up to isomorphism expands M by a bijection R^M from M onto $[n]$ (cf. Example 2.28). By Lemmas 4.6 and 3.9 (1), $\Vdash_{\mathrm{Ri}}$ is definable for $\exists \Delta_0^{b_0}(R)$-formulas, i.e. formulas otained from $\Delta_0^{b_0}(R)$-formulas by existential quantification. By the Principal Theorem, $M[G]$ satisfies the least number principle for $\exists \Delta_0^{b_0}(R)$-formulas. So it suffices to show that every $\Sigma_1^{b_0}(R)$-formula is equivalent in $M[G]$ to such a formula. This in turn follows from $M[G]$ satisfying $\Delta_0^{b_0}(R)$-collection, i.e. for every $\Delta_0^{b_0}(R)$-formula $\varphi(x, y)$

$$M[G] \models (\forall x < b_0 \exists y \; \varphi(x, y) \rightarrow \exists z \forall x < b_0 \; \varphi(x, (z)_x)).$$

To see this, assume $M[G] \models \forall x < b_0 \exists y \; \varphi(x, y)$ and consider the formula

$$\psi(u) := \exists z (lh(z) = b_0 - u \wedge \forall x < b_0 (x < lh(z) \rightarrow \varphi(x, (z)_x))).$$

Trivially, $M[G] \models \psi(b_0)$, so $\psi(u)$ defines a nonempty set in $M[G]$. As $\psi(u)$ is $\exists \Delta_0^{b_0}(R)$, this set contains a least element $b \leq^M b_0$. It is easy to verify $M[G] \models \forall u(\psi(u) \rightarrow \psi(u-1))$, so $b = 0$ follows. $\square$

4.3 Ajtai forcing

We prove Ajtai's result [1] including its improvements from [27, 31]. Compared to Theorem 4.3 it embodies an exponential improvement concerning the bound b_0, namely, it assumes only $|b_0| <^M n^{o(1)}$ instead of $b_0 <^M n^{o(1)}$. On the other hand, it only concerns b_0-bounded formulas, i.e $\Delta_0^{b_0}(R)$ (cf. section 4.2).

Theorem 4.7 (Ajtai 1988). *Assume $b_0, n \in M$ are such that $|b_0| <^M n^{o(1)}$. Then M has an L^*-expansion (M, R^M) such that R^M is a bijection from $[n+1]$ onto $[n]$ and (M, R^M) satisfies the least number principle for $\Delta_0^{b_0}(R)$ up to b_0.*

Remark 4.8. The reader familiar with bounded arithmetic will notice the following. Use Buss' language for L and choose $b_0, n \in M$ such that both $|b_0| < n^{o(1)}$ and $|n| < |b_0|^{o(1)}$. By the second inequality b_0 bounds $t^M(n)$ for every L-term $t(x)$. Thus the corresponding cut carries a model of $T_2(R)$.

We refer to section 5 for an application of Theorem 4.7 in proof complexity.

For $m \in \mathbb{N}$ consider the following finite forcing frame $(P(m), \leq)$ (without a family $D_0, D_1, \dots$): the conditions are the partial bijections from $[m+1]$ to $[m]$ and $p \leq q$ means $p \supseteq q$. Again, we blur the distinction of the bijection and its code in $\mathbb{N}$. The *size* $\|p\|$ of a condition p is its cardinality, i.e. the number of pigeons mapped. The *rank* of a set $X \subseteq P(m)$ is the maximal size of a condition in X (and, say, 0 if X is empty).

Now fix M and $n, b_0 \in M$ satisfying the assumptions of Theorem 4.7. Observe that there are uniform definitions of $P(m)$ in $(\mathbb{N}, +, \cdot, 0, 1, <)$ in the sense that there is a $\{+, \cdot, 0, 1, <\}$-formula $\varphi(x, y)$ such that $P(m) = \{p \in \mathbb{N} \mid \mathbb{N} \models \varphi(m, p)\}$ for every $m \in \mathbb{N}$. Applied in M, these definitions give forcing frames $(P(m), \leq)$ with size function $\|\cdot\|$ also for (nonstandard) $m \in M$. Further note that M defines the function $m \mapsto m^\epsilon$ (rounded up) for any (standard) rational $0 < \epsilon < 1$.

We now define the forcing frame P. It is going to be an undefinable subframe of the definable frame $P(n)$. The set $\{p \in P(n) \mid \|p\| <^M n - n^\epsilon\}$ is definable in M for every standard rational $0 < \epsilon < 1$. We let P be the union of all these sets. As usual $p \leq q$ means $p \supseteq q$, and the family $D_0, D_1, \dots$ enumerates the sets $\{p \in P \mid b \in \text{dom}(p)\}$ and $\{p \in P \mid c \in \text{im}(p)\}$ for $b \in [n+1], c \in [n]$. It is easy to see that these sets are dense (in P).

We define the forcing as in the previous two sections: we let $p \in P$ force an atom Rst if $(s^M, t^M) \in p$ and denote by $\Vdash_{\text{Aj}}$ the resulting conservative universal pre-forcing. It is easy to see that $\Vdash_{\text{Aj}}$ is a forcing and that $M[G]$ is defined for every generic G (cf. Example 2.28).

Lemma 4.9. $\Vdash_{\text{Aj}}$ *is densely definable for $\Delta_0^{b_0}(R)$ up to b_0.*

Proof. We use the antichain method from section 3.2. Let $p \in P$ and $\varphi(\bar{x}) \in \Delta_0^{b_0}(R)$. We have to find $r \leq p$ such that the set $\{\bar{a} <^M b_0 \mid r\|\varphi(\bar{a})\}$ is definable in M. By Lemma 3.10 this set equals $\{\bar{a} <^M b_0 \mid \exists q \in X_{\bar{a}} : r\|q\}$ if for every $\bar{a} <^M b_0$, $X_{\bar{a}}$ is a maximal antichain in $[\varphi(\bar{a})] \downarrow r$. Definabiliy of this set follows from Lemma 4.10 below. $\square$

Theorem 4.7. Use the above lemma and argue as in the proofs of Theorems 4.2 and 4.3. $\square$

Lemma 4.10. *Let $p \in P$. For every $\varphi(\bar{x}) \in \Delta_0^{b_0}(R)$ there is $r \in P, r \leq p$ and a sequence of sets $(X_{\bar{a}})_{\bar{a} <^M b_0}$ in M such that for every $\bar{a} <^M b_0$, the set $X_{\bar{a}}$ is a maximal antichain in $[\varphi(\bar{a})] \downarrow r$.*

Here, by saying that a sequence $(X_{\bar{a}})_{\bar{a}<^M b_0}$ of subsets of M is *in M* we mean that the set $\{(\bar{a}, c) \mid \bar{a} <^M b_0, c \in X_{\bar{a}}\}$ is coded in M.

To prove this we intend to use Lemma 3.13. Therefore we need to define predense antichains refining given sets and it is here where the finite combinatorics enter the argument (cf. Introduction). The idea is to show that suitable antichains exist in $P(m)$ for $m \in \mathbb{N}$ sufficiently large. Then M codes these antichains for the infinite $P(n)$. As a first problem, predensity does not make much sense in finite frames nor in $P(n)$. Therefore we shall calibrate the notion in the definition below. Second, suitable antichains need not to exist, but they do exist after restricting attention to conditions that extend a suitably chosen condition r. This choice is done according to the Switching Lemma 4.13 below, the combinatorial core of the argument. Details follow.

Definition 4.11. Let $m, k \in \mathbb{N}$, $q \in P(m)$ and $X \subseteq P(m)$. Then X is *k-predense (in $P(m)$)* if every condition of size at most $m - k$ is compatible with a condition in X.

Clearly, if $k \leq k'$, then k-predensity implies k'-predensity. For $m \in M$, $p \in P(m)$ and $X \subseteq P(m)$ write

$$X^p := \{q \setminus p \mid q \in X, p \| q\} \quad \text{and} \quad X \cup p := \{q \cup p \mid q \in X, p \| q\}.$$

Note that $P(m)^p \cong P(m - \|p\|)$ via a size preserving isomorphism. Of course, being a bounded set, such an isomorphism is coded in M for every $m \in M$ and $p \in P$. By saying that an antichain is k-predense in $P(m)^p$ we mean that its image under this isomorphism is k-predense in $P(m - \|p\|)$. In the same way k-predensity is explained in $P(n)^p$.

Lemma 4.12. *Let $X \subseteq P$, $p, q \in P, q \leq p$ and let φ be an $L^*(M)$-sentence and $\tilde{b} \leq^M |b_0|$. If X is a maximal antichain in $[\varphi] \downarrow p$ and has rank at most $\|p\| + \tilde{b}$, then $X \cup q$ is a maximal antichain in $[\varphi] \downarrow q$ of rank at most $\|q\| + \tilde{b}$.*

Proof. As $X \subseteq P \downarrow p$, X^p has rank at most $\tilde{b}$. Then $X \cup q = X^p \cup q$ has rank at most $\|q\| + \tilde{b} \leq^M \|q\| + |b_0|$ and, in particular, $X \cup q \subseteq P$. Clearly, $X \cup q$ is an antichain. To show containment in $[\varphi] \downarrow q$, let $r \in X \cup q$ and choose $s \in X, q \| s$ such that $r = s \cup q$. Since $X \subseteq [\varphi]$, we have $s \cup q \in [\varphi] \downarrow q$ by Extension. To show maximality, let $r \in [\varphi] \downarrow q$. By maximality of X, r is compatible with some $s \in X$. As $r \leq q$, q is compatible with s. Then $s \cup q \in X \cup q$ is compatible with r. $\square$

Lemma 4.13 (Switching). *Let $X_0, \dots, X_{N-1}$ be subsets of $P(m)$ of rank at most k. Let $\ell \in \mathbb{N}$ and assume that*

$$\frac{(m - \ell)^k}{(\ell + 1)^{4k} \cdot k^{3k}} > N.$$

Then there is $q \in P(m)$ of size at most $m - \ell$ such that for every $i < N$ there is an antichain $A_i \subseteq P(m)^q$ refining X_i^q that is $2k$-predense in $P(m)^q$ and has rank at most $2k$.

This lemma can be proved by the probabilistic method or a direct (involved) counting argument. Details can be found in [23, Lemma 12.3.10].

Lemma 4.14. *Let $p \in P$ and $X, Y \subseteq P(n)^p$ have rank at most $|b_0|$.*

1. *If X is an antichain in $P(n)^p$, then $X \cup p$ is an antichain in P.*

2. *If X is $|b_0|$-predense in $P(n)^p$, then $X \cup p$ is predense in P below p.*

3. *If X refines Y in $P(n)^p$, then $X \cup p$ refines $Y \cup p$ in P.*

Proof. We only show (2). Note $X \cup p \subseteq P$ because it has rank at most $\|p\| + |b_0|$. Let $q \in P, q \leq p$ and choose $0 < \epsilon < 1$ such that $\|q\| <^M n - n^\epsilon$. Then $\|q \setminus p\| = \|q\| - \|p\| <^M n - n^\epsilon - \|p\| <^M (n - \|p\|) - |b_0|$. Since $(q \setminus p) \in P(n)^p$ and X is $|b_0|$-predense in $P(n)^p$, there is $r \in X$ such that $q \setminus p$ is compatible with r in $P(n)^p$. Then $q \cup r = q \cup (r \cup p)$ extends both q and $r \cup p \in X \cup p$. As $q \cup r$ has size $<^M \|q\| + |b_0|$ it is in P, so q and $r \cup p$ are compatible in P. $\square$

The rest of the argument is straightforward. We give the details:

Lemma 4.10. Let $p \in P$. Call an $L^*(M)$-formula $\varphi(\bar{x})$ *good* if for every positive $c \in \mathbb{N}$ and every $p \in P$ there are $r \in P, r \leq p$, and $(X_{\bar{a}})_{\bar{a} <^M b_0}$ in M such that every $X_{\bar{a}}$ is a maximal antichain in $[\varphi(\bar{a})] \downarrow r$ and additionally has rank at most $\|r\| + |b_0|/c$; here and in the following, $|b_0|/c$ stands for $\lfloor |b_0|/c \rfloor$ computed in M. Clearly, the set of good formulas is closed under logical equivalence. It thus suffices to show that it contains all atomic formulas and is closed under conjunctions, negations and b_0-bounded universal quantification.

We show that atomic formulas are good. Let $c \in \mathbb{N}$ be positive, $p \in P$ and $\varphi(\bar{x})$ have the form Rst for $L(M)$-terms $s = s(\bar{x}), t = t(\bar{x})$; take $r := p$ and define $X_{\bar{a}} := \{r \cup \{(s^M(\bar{a}), t^M(\bar{a}))\}\}$ or $X_{\bar{a}} := \emptyset$ depending on whether $r \cup \{(s^M(\bar{a}), t^M(\bar{a}))\}$ is a partial bijection from $[n+1]$ to $[n]$ or not. Then $X_{\bar{a}}$ has rank at most $\|r\| + 1 \leq^M \|r\| + |b_0|/c$. Similarly, for an $L(M)$-atom $\varphi(\bar{x})$ set $r := p$ and $X_{\bar{a}} := \{r\}$ or $X_{\bar{a}} := \emptyset$ depending on whether $M \models \varphi(\bar{a})$ or not.

We show that negations of good formulas are good. Let $\varphi(\bar{x})$ be good, $c \in \mathbb{N}$ positive and $p \in P$. Choose $r \in P, r \leq p$, and $(X_{\bar{a}})_{\bar{a} <^M b_0}$ in M such that for every $\bar{a} <^M b_0$, $X_{\bar{a}}$ is a maximal antichain in $[\varphi(\bar{a})] \downarrow r$ of rank at most $\|r\| + |b_0|/(2c)$. Choose $0 < \epsilon < 1$ such that $\|r\| <^M n - n^\epsilon$. Then $n^\epsilon <^M n - \|r\| =: m$. Recall that as partial orders $P(n)^r \cong P(m)$, via an isomorphism that is coded in M and preserves the size $\| \cdot \|$.

We intend to apply the Switching Lemma in M to get refining antichains for the sequence $(X_{\bar{a}}^r)_{\bar{a} <^M b_0}$. We check its assumptions: every $X_{\bar{a}}^r$ has rank at most $k := |b_0|/(2c)$; the sequence has length $N := b_0^{\ell_0}$ for $\ell_0 \in \mathbb{N}$ the length of $\bar{x}$; there is a (standard) rational $0 < \epsilon' < 1$ (e.g. $\epsilon' := 1/5$) such that the inequality of the Switching Lemma is satisfied for $\ell := m^{\epsilon'}$. Thus the lemma applies, that is, we find $r' \in P(n)^r$ of size at most $m - m^{\epsilon'}$ such that, writing $s := (r \cup r')$, the following holds: for every $\bar{a} <^M b_0$ there is $A_{\bar{a}} \subseteq (P(n)^r)^{r'} = P(n)^s$ coded in M such that in $P(n)^s$

(a) $A_{\bar{a}}$ is an antichain that is $|b_0|/c$-predense (hence $|b_0|$-predense),

(b) $A_{\bar{a}}$ refines $(X_{\bar{a}}^r)^{r'} = X_{\bar{a}}^s \subseteq P(n)^s$,

(c) $A_{\bar{a}}$ has rank at most $|b_0|/c$.

Note the sequence $(A_{\bar{a}})_{\bar{a} <^M b_0}$ is in M. Further, s has size $\|r\| + \|r'\| \leq^M n - m + m - m^{\epsilon'} <^M n - n^{\epsilon \epsilon'}$, so $s \in P$. Then, in P

(d) $(A_{\bar{a}} \cup s) \subseteq P \downarrow s$ is an antichain, predense below s (Lemma 4.14 (1),(2)),

(e) $(A_{\bar{a}} \cup s)$ refines $X_{\bar{a}}^s \cup s = X_{\bar{a}} \cup s$ (Lemma 4.14 (3)),

(f) $(A_{\bar{a}} \cup s)$ has rank at most $\|s\| + |b_0|/c$.

By Lemma 4.12, $(X_{\bar{a}} \cup s)$ is a maximal antichain in $[\varphi(\bar{a})] \downarrow s$. By (d) and (e) the assumptions of Lemma 3.13 (1) are satisfied. We thus get a maximal antichain in $[\neg\varphi(\bar{a})] \downarrow s$ setting

$$B_{\bar{a}} := (A_{\bar{a}} \cup s) \setminus ((A_{\bar{a}} \cup s) \downarrow (X_{\bar{a}} \cup s)).$$

Then $(B_{\bar{a}})_{\bar{a} <^M b_0}$ is in M and has rank at most $\|s\| + |b_0|/c$ by (f).

We show that conjunctions of good formulas are good. Let $\varphi(\bar{x})$ and $\psi(\bar{y})$ be good, $c \in \mathbb{N}$ positive and $p \in P$. We can assume that $\bar{x}$ equals $\bar{y}$. As seen, $\neg\varphi(\bar{x})$ is good, so there are $\tilde{r} \in P, \tilde{r} \leq p$, and maximal antichains $\tilde{X}_{\bar{a}}$ in $[\neg\varphi(\bar{a})] \downarrow \tilde{r}$ of rank at most $\|\tilde{r}\| + |b_0|/(2c)$. Since also $\neg\psi(\bar{x})$ is good, there are $r \in P, r \leq \tilde{r}$, and maximal antichains $Y_{\bar{a}}$ in $[\neg\psi(\bar{a})] \downarrow r$ of rank at most $\|r\| + |b_0|/(2c)$. By Lemma 4.12, $X_{\bar{a}} := \tilde{X}_{\bar{a}} \cup r$ is a maximal antichain in $[\neg\varphi(\bar{a})] \downarrow r$ of rank at most $\|r\| + |b_0|/(2c)$. Now, we are looking for refining antichains for the sequence $(X_{\bar{a}}^r \cup Y_{\bar{a}}^r)_{\bar{a} <^M b_0}$. This sequence is in M and each set in it has rank at most $|b_0|/(2c)$. As above, the Switching Lemma gives $s \leq r$ and $A_{\bar{a}} \subseteq P(n)^s$ such that (d), (f) and

(e′) $(A_{\bar{a}} \cup s)$ refines $(X_{\bar{a}} \cup s) \cup (Y_{\bar{a}} \cup s)$.

By Lemma 4.12, $(X_{\bar{a}} \cup s)$ and $(Y_{\bar{a}} \cup s)$ are maximal antichains in $[\neg\varphi(\bar{a})] \downarrow s$ and $[\neg\psi(\bar{a})] \downarrow s$ respectively. By Lemma 3.13 (2)

$$B_{\bar{a}} := (A_{\bar{a}} \cup s) \setminus ((A_{\bar{a}} \cup s) \downarrow ((X_{\bar{a}} \cup s) \cup (Y_{\bar{a}} \cup s)))$$

is a maximal antichain in $[(\varphi \wedge \psi)(\bar{a})] \downarrow s$; it has rank $\leq^M \|s\| + |b_0|/c$ by (f).

Finally, we show that the set of good formulas is closed under b_0-bounded universal quantification. Let $\varphi(x\bar{x})$ be good, $c \in \mathbb{N}$ positive and $p \in P$. Then $\neg\varphi(x\bar{x})$ is good, so there are $r \in P, r \leq p$, and maximal antichains $X_{a\bar{a}}$ in $[\neg\varphi(a\bar{a})] \downarrow r$ of rank at most $\|r\| + |b_0|/(2c)$. Applying the Switching Lemma on the sequence $(\bigcup_{a <^M b_0} X_{a\bar{a}}^r)_{\bar{a} <^M b_0}$ gives $s \leq r$ and $A_{\bar{a}} \subseteq P(n)^s$ such that (d), (f) and

(e″) $(A_{\bar{a}} \cup s)$ refines $\bigcup_{a <^M b_0}(X_{a\bar{a}} \cup s)$.

Similarly as before, (f) and Lemmas 4.12 and 3.13 (3) imply that

$$B_{\bar{a}} := (A_{\bar{a}} \cup s) \setminus ((A_{\bar{a}} \cup s) \downarrow \bigcup_{a <^M b_0}(X_{a\bar{a}} \cup s))$$

is a rank $\leq^M \|s\| + |b_0|/c$ maximal antichain in $[\forall x < b_0 \varphi(x\bar{a})] \downarrow s$.

$\square$

4.4 Notes

Compared to Riis' original argument [38] our proof of Theorem 4.3 relies on properties of universal forcing while Riis uses an existential forcing (recall Remark 2.22). Our argument is simpler in that it sidesteps the technically more involved proof in [38] of a definability lemma for $\Sigma_1^{b_0}(R)$.

Compared with other proofs of Theorem 4.7, roughly, the predense antichains in our argument correspond to the complete systems in [27] and [51], to branches in shallow decision trees in [50, 26] or to the small covers in [1]. Our exposition is close to [23, Section 12.3].

Forcing type arguments for Ajtai's result have been given in [1, 51] and [23, section 12.7] and more recently in [26]. In [23, section 12.7] Krajíček presents the method of k-evaluations of propositional formulas [27] as a forcing type argument. Our proof constructs for certain φ a predense antichain together with its maximal part in $[\varphi]$. These pairs of sets give a modified notion of $|b_0|$-evaluation. As in Zambella [51], our argument sidesteps a detour through propositional logic like in [1, 27, 50, 23]. It is simpler than Zambella's in that it avoids the restriction to "internal" generics [51].

All known arguments for Ajtai's result, including Krajíček's more recent proof in [26], use the Switching Lemma in one or another form. The main obstacle to generalize Ajtai's argument to other principles is the difficulty to find analogues of this lemma. According to the proof presented here, its role in the argument is to provide the existence of refining antichains, a combinatorial property of the forcing frame. The rest of the argument is taken over by the general machinery, i.e. the Principal Theorem and the antichain method.

5 Forcing and propositional logic

The *spectrum* of a sentence φ_0 is the set of naturals $m \geq 1$ such that φ_0 has a model of size m. Given $m \geq 1$, there is a propositional formula $\langle \varphi_0 \rangle_m$ whose satisfying assignments describe size m models of φ_0. If φ_0 has empty spectrum, all of these formulas are unsatisfiable, and one may ask for short or otherwise simple propositional refutations. If such a sentence also does not have infinite models, short refutations exist even in treelike resolution [39], but in the presence of an infinite model things are hard to understand [39, 25, 15]. We refer to [16] for a survey on the complexity of first-order spectra.

5.1 Propositional translation

Let L_0 be a finite language and φ_0 be an L_0-sentence of the form

$$\forall \bar{x} \bigwedge_{j \in J} C_j(\bar{x}),$$

where J is a finite nonempty set, $\bar{x} = x_0, \ldots, x_{k-1}$ and the C_j's are clauses. We assume further that φ_0 is *unnested*, that is, its atoms have the form $y = z$ or $E\bar{y}$ or $f\bar{y} = z$ for $E, f \in L_0$.

Remark 5.1. Skolemization allows to compute from a given sentence one of the specified form with the same spectrum. Moreover, the computed sentence has an infinite model if and only if the given sentence has one.

We write propositional formulas in de Morgan language $\{\wedge, \vee, \neg\}$ using propositional constants $\top$ and $\bot$ for "true" and "false" respectively and propositional atoms of the form $E\bar{b}, f\bar{b}c$ where $\bar{b} \in [m]^r, c \in [m]$ and E and f are r-ary relation and function symbols from L_0 respectively. The *translation* $\langle \varphi_0 \rangle_m$ for $m \geq 1$ is the following set of clauses. For every $\bar{a} \in [m]^k$ and $j \in J$ consider the first-order clause $C_j(\bar{a})$; replace first-order atoms $E\bar{b}, f\bar{b} = c$ in $C_j(\bar{a})$ by the propositional atoms $E\bar{b}, f\bar{b}c$, and replace $b = c$ by the propositional constant $\top$ or $\bot$ according to the truth value of $b = c$. Further, $\langle \varphi_0 \rangle_m$ contains the *functionality clauses* $\bigvee_{c \in [m]} f\bar{b}c$ and $(\neg f\bar{b}c \vee \neg f\bar{b}d)$ where $f \in L_0$ is an r-ary function symbol, $\bar{b} \in [m]^r$ and $c, d \in [m], c \neq d$.

Every *functional* assignment, i.e. one that satisfies the functionality clauses, describes an L_0-model on $[m]$ in the obvious way, and it satisfies $\langle \varphi_0 \rangle_m$ if and only if the described model satisfies φ_0.

Example 5.2. Let $L_0 := \{f, g\}$ for a unary function symbol f and a constant (0-ary function) symbol g and consider:

$$\varphi_0 \quad := \quad \forall xyz \, ((\neg fx = z \vee \neg fy = z \vee x = y) \wedge (\neg fx = y \vee \neg g = y)).$$

For $m \geq 1$ the translation $\mathrm{PHP}_m := \langle \varphi_0 \rangle_m$ contains the functionality clauses plus $(\neg fac \vee \neg fbc), (\neg fac \vee \neg gc)$ for $a, b, c \in [m], a \neq b$; here, we omit $\bot$ from clauses and clauses containing $\top$. This is a version of the functional, injective, m to $m - 1$ pigeonhole principle (cf. [37] for a survey).

5.2 An application of Riis' Theorem

Fix an unnested L_0-sentence $\varphi_0 = \forall \bar{x} \bigwedge_{j \in J} C_j(\bar{x})$ with empty spectrum and an infinite model. We aim to show that every semantic refutation of $\langle \varphi_0 \rangle_m$ contains a "complex" formula.

A *semantic refutation* of $\langle \varphi_0 \rangle_m$ is a sequence of propositional formulas in the atoms of $\langle \varphi_0 \rangle_m$ that ends in $\bot$ and every formula is either a clause in $\langle \varphi_0 \rangle_m$ or logically implied by two previous formulas.

Our complexity measure for propositional formulas is the height of a *Poizat tree* computing the formula. This is a full m-branching ordered tree whose inner nodes are labeled with *queries* $(E, \bar{b})$ or $(f, \bar{b})$ for $\bar{b} \in [m]^r$, where E and f are r-ary relation and function symbols from L_0 respectively; its leafs are labeled with truth values $0, 1$. Moreover, no label occurs twice on some path.

Every path in such a tree corresponds in a natural way to a partial functional assignment: if it contains a node labeled $(E, \bar{a})$ and its $(i + 1)$th successor, the propositional atom $E\bar{a}$ evaluates to 0 or 1 depending on whether $i = 0$ or not; if it contains a node labeled $(f, \bar{a})$ and its $(i+1)$th successor, every propositional atom $f\bar{a}b$ is evaluated to 1 or 0 depending on whether $b = i$ or not. Two paths (in possibly different trees) are *compatible* if there exists

28

a functional assignment containing the two partial assignments corresponding to the paths; two leafs (in possibly different trees) are compatible if so are the paths leading to them.

Conversely, every functional assignment determines a branch of a given Poizat tree. The tree *computes* a given propositional formula if for every functional assignment A the truth value of the formula under A coincides with the label of the leaf on the branch determined by A.

Definition 5.3. The *Poizat width* of a formula is the minimal height of a Poizat tree computing it.

The Poizat width of a formula is well-defined: a formula with ℓ atoms has Poizat width at most ℓ, since it can be computed by the trivial tree that queries all appearing atoms one after another. In particular, the clauses in $\langle \varphi_0 \rangle_m$ have constant Poizat width (i.e. independent of m). Note that a functionality clause is computed by the one-node tree with label 1, so has Poizat width 0.

Theorem 5.4. *There is $\epsilon > 0$ such that for every sufficiently large m every semantic refutation of $\langle \varphi_0 \rangle_m$ contains a formula of Poizat width at least m^ϵ.*

Proof. Choose for every $m \geq 1$ a semantic refutation of $\langle \varphi_0 \rangle_m$ of minimal Poizat width $h(m)$ (the Poizat width of a sequence of formulas is understood as the maximal Poizat width of a formula in the sequence). To extend this sequence of refutations to the pseudo-finite we need to code it in arithmetic. We can keep this technically simple using a suitably rich language L.

View the nodes of an m-branching tree as a set of (codes of) sequences over $[m]$ that is closed under initial segments. To talk about them we use function symbols $lh(x), (x)_z$ and $(x)_{<z}$ for the length, the $(z+1)$th component and the length z initial segment of the sequence x; we further use a symbol $x - y$ for subtraction (cf. section 4.2). Using a symbol (x, y) for pairing and identifying $E, f, \ldots \in L_0$ with numbers $\dot{E}, \dot{f}, \ldots \in \mathbb{N}$ queries $(\dot{E}, \bar{a}), (\dot{f}, \bar{a})$ become naturals; we assume these to be bigger than 2. To code the refutations we use a ternary function symbol π such that the function $\pi(m, i, \cdot)$ coincides with the labeling of the Poizat tree of the $(i+1)$th formula in the refutation of $\langle \varphi_0 \rangle_m$; on arguments not coding nodes of the tree, the function has value 2. Note that the formula $\pi(m, i, x) < 2$ defines the set of leaves in the tree for the $(i+1)$th formula of the mth refutation. Additionally, we add unary function symbols *lines* and h such that *lines*(m) and $h(m)$ are the length and the Poizat width of the mth refutation.

Further, we let L contain the language L_0 and interpret it over $\mathbb{N}$ such that it models φ_0. Let M be a countable L-model that is a proper elementary extension of the 'standard' L-model $\mathbb{N}$. In particular, $M \models \varphi_0$.

Assume the theorem fails. Then there is, for every ℓ, an $m \geq \ell$ such that $h(m)+1 < m^{1/\ell}$. Apply this to a nonstandard $\ell \in M$, and find nonstandard $n, b_0 \in M$ such that $h^M(n) <^M b_0 <^M n^{o(1)}$.

Theorem 4.3 gives a bijection R^M from M onto $[n]$ such that (M, R^M) satisfies the least number principle for $\Sigma_1^{b_0}(R)$. Note that R copies M's interpretation of L_0 onto $[n]$, so

(M, R^M) defines a structure on $[n]$ that satisfies φ_0. But as M is a model of true arithmetic, (M, R^M) does not code the 'fictitious' assignment that describes this structure. We show, however, that every tree $\pi^M(n, i, \cdot)$ contains a leaf corresponding to this 'fictitious' assignment. Such a leaf is obtained from a sequence c of length $h^M(n) + 1 \leq^M b_0$ that answers all queries according to the 'fictitious' assignment and, say, stays constant upon reaching a point with a Boolean label.

Formally, this means that $\forall z < lh(c) \; \psi(n, i, c, z)$ holds in (M, R^M) for a suitable formula ψ. E.g. if $L_0 = \{E, f\}$ for a unary relation symbol E and a unary function symbol f, the formula $\psi(n, i, c, z)$ reads

$$\exists y y' \; ((\pi(n, i, (c)_{<z}) = (\dot{E}, y') \wedge Ryy' \wedge ((c)_z > 0 \leftrightarrow Ey))$$
$$\vee (\pi(n, i, (c)_{<z}) = (\dot{f}, y') \wedge Ryy' \wedge Rf(y)(c)_z))$$
$$\vee (\pi(n, i, (c)_{<z}) < 3 \wedge (c)_z = (c)_{z-1}).$$

If c codes such a sequence of length at least $h^M(n) + 1$, then $(c)_{<\ell}$ is a leaf for the minimal ℓ such that $\pi^M(n, i, (c)_{<\ell}) <^M 2$.

Claim For every $i <^M lines^M(n)$, (M, R^M) satisfies

$$\exists x \; (\pi(n, i, x) < 2 \wedge \forall z < lh(x) \; \psi(n, i, x, z)).$$

Proof of Claim As argued above, it suffices to show the existence of a suitable sequence, i.e. it suffices to show the claim with $\pi(n, i, x) < 2$ replaced by $lh(x) = h(n) + 1$. To this end consider the formula $\chi(u)$ obtained replacing $\pi(n, i, x) < 2$ by $lh(x) = h(n) + 1 - u$. Since $h^M(n) + 1 \leq^M b_0$ the formula $\chi(u)$ is $\Sigma_1^{b_0}(R)$ and holds on $h^M(n) + 1$ in (M, R^M). Hence the set it defines has a minimum a_0. But $\chi(u)$ implies $\chi(u - 1)$ in (M, R^M), so $a_0 = 0$. $\quad\dashv$

Consider the standard model $\mathbb{N}$ for a moment and let $m \geq 1$. For a clause $\langle C_j(\bar{a}) \rangle_m$ of $\langle \varphi_0 \rangle_m$ (where $j \in J, \bar{a} \in [m]^k$) there is a trivial tree $T_j^{\bar{a}}$ querying all appearing atoms one after another; if $h_j \in \mathbb{N}$ denotes the number of these atoms, then any branch in $T_j^{\bar{a}}$ makes h_j many queries each having the form $(E, \bar{a}'), (f, \bar{a}')$ with the components of $\bar{a}'$ appearing among those of $\bar{a}$.

We can assume that $\pi^{\mathbb{N}}(m, i, \cdot)$ equals $T_j^{\bar{a}}$ when the $(i+1)$th formula is $\langle C_j(\bar{a}) \rangle_m$, and that $\pi^{\mathbb{N}}(m, i, \cdot)$ equals the one-node tree with label 1 when the $(i+1)$th formula is a functionality clause. Otherwise there are $i', i'' < i$ such that the following *soundness condition* holds:

if ℓ is a leaf of $\pi^{\mathbb{N}}(m, i, \cdot)$ labeled 0 and ℓ', ℓ'' are leafs of $\pi^{\mathbb{N}}(m, i', \cdot)$ and $\pi^{\mathbb{N}}(m, i'', \cdot)$ such that ℓ, ℓ', ℓ'' are pairwise compatible, then at least one of ℓ', ℓ'' is labeled 0.

Observe that compatibility can be expressed by an L-formula. By elementary equivalence then, the soundness condition holds in M for the (nonstandard) trees $\pi^M(n, i, \cdot)$. Now consider the formula "the line y is false":

$$\exists x \; (\pi(n, y, x) = 0 \wedge \forall z < lh(x) \; \psi(n, y, x, z)).$$

30

We can assume that for every $m \geq 1$ the tree $\pi^{\mathbb{N}}(m, lines^{\mathbb{N}}(m) - 1, \cdot)$ is the one-node tree with label 0. Then trivially $lines^M(n) - 1$ satisfies "line y is false" in (M, R^M). The universal quantifier in this formula can be bounded by b_0, so it is $\Sigma_1^{b_0}(R)$. Hence there exists a minimal $i_0 <^M lines^M(n)$ satisfying it.

Let $\ell_0 \in M$ be a leaf witnessing the quantifier $\exists x$ in "line i_0 is false". Then $\pi^M(n, i_0, \cdot)$ cannot be one of the $T_j^{\bar{a}}$ for $j \in J, \bar{a} \in [n]^k$. Otherwise the h_j many queries below ℓ_0 are answered in a way falsifying $C_j(\bar{a})$. Then the tuple $\bar{b}$ that is mapped to $\bar{a}$ by R^M falsifies $C_j(\bar{x})$ in M and this contradicts $M \models \varphi_0$. Of course, $\pi^M(n, i_0, \cdot)$ also cannot be a one-node tree labeled 1. Hence there are $i', i'' <^M i_0$ satisfying the soundness condition (in M). Choose leafs ℓ' and ℓ'' in $\pi^M(n, i', \cdot)$ and $\pi^M(n, i'', \cdot)$ according the claim. These are such that ℓ, ℓ', ℓ'' are pairwise compatible (in M). Then the soundness condition (in M) implies that ℓ' or ℓ'' is labeled 0. But then i' or i'' satisfies "line y is false" and this contradicts the minimality of i_0. $\qquad\square$

5.3   An application of Ajtai's Theorem

Theorem 4.7 has the following well-known consequence mentioned in the Introduction. We include a brief sketch for completeness.

Recall PHP_m from Example 5.2. Fix a *Frege system* F, that is, a finite set of inference rules. We only assume that F is sound in the sense that every inference rule $\frac{\alpha_0, \ldots, \alpha_{r-1}}{\alpha_r}$ is such that α_r is logically implied by the $\alpha_i, i < r$. An *F-refutation* of PHP_m is a finite sequence of propositional formulas in the atoms of PHP_m that ends in $\perp$ and every formula is either a clause in PHP_m or derived by an inference rule $\frac{\alpha_0, \ldots, \alpha_{r-1}}{\alpha_r}$ from F, that is, equal to a substitution instance of α_r with the corresponding substitution instances of the $\alpha_i, i < r$, appearing earlier in the sequence. The *size* of an F-refutation is its length when it is coded as a binary string; its *depth* is the maximal depth of a formula in it, where the depth of a formula is the maximal number of times the connective $\wedge, \vee, \neg$ changes along a branch in the formula tree.

Theorem 5.5. *For every $d \in \mathbb{N}$ there is $\epsilon > 0$ such that for every sufficiently large m every F-refutation of PHP_m of depth at most d has size at least 2^{m^ϵ}.*

Sketch. Assume the theorem fails at depth $d \in \mathbb{N}$. As in the proof of Theorem 5.4, we find a structure M and $b_0, n \in M$ satisfying the assumptions of Theorem 4.7 and a (code of an) F-refutation $\pi \in M$ of PHP_{n+1} of depth d and size $<^M b_0$. By Theorem 4.7 there is a bijection R^M from $n + 1$ onto n such that (M, R^M) satisfies the least number principle for $\Delta_0^{b_0}(R)$ up to b_0.

Let $True(x)$ be a $\Delta_0^{b_0}(R)$-"truth predicate" for (codes of) depth $\leq d$ propositional formulas of length $<^M b_0$: an atom fab (resp. gc) should satisfy $True(x)$ if and only if $(a, b) \in R^M$ (resp. $c = n$); moreover, $True(x)$ should satisfy Tarski's conditions – e.g. a conjunction of two formulas α, β in the sense of M should satisfy $True(x)$ if and only if so do both α and β. Then $True(x)$ is satisfied by every clause from PHP_{n+1}. Further, if $\beta \in M$ is a formula obtained from a standard formula $\alpha \in \mathbb{N}$ by substituting for atoms X certain formulas

$\gamma_X \in M$, then β satisfies $True(x)$ if and only if α is satisfied by the assignment mapping X to the truth value of $True(\gamma_X)$.

Thus, if $\frac{\alpha_0,\dots,\alpha_{r-1}}{\alpha_r}$ is a rule in F and $\beta_i \in M$ are obtained from α_i via some substitution and β_r satisfies $\neg True(x)$, then there is $i < r$ such that β_i satisfies $\neg True(x)$ (in particular, $r \neq 0$). Then $\neg True((\pi)_y)$ is a $\Delta_0^{b_0}(R)$-formula that defines a set containing $lh(\pi) - 1 <^M b_0$ but no minimum, a contradiction. $\qquad\square$

5.4 Notes

Width lower bounds on resolution refutations follow from certain 'expansion' properties of sets of clauses [8], and are characterized by the existence of winning strategies for the adversary in a prover-adversary game [35, 2]. For a different notion of width, Dantchev and Riis [15] established a general width lower bound on treelike R(k) refutations of principles of the form $\langle \varphi_0 \rangle_m$. Segerlind et al. [42] transfered width lower bounds for resolution in the sense of [8], to width lower bounds for daglike R(k) – this time in the sense of having small height decision trees. This argument uses special properties of R(k). Poizat trees are called as they are because queries correspond to basic operations of machines in the sense of [32].

References

[1] M. Ajtai. The complexity of the pigeonhole principle. *Proceedings of the 29th Annual Symposion on the Foundations of Computer Science*, pages 346–355, 1988.

[2] A. Atserias and V. Dalmau. A combinatorial characterization of resolution width. *Journal of Computer and System Sciences*, 74(3):323–334, 2008.

[3] J. Avigad. Forcing in proof theory. *The Bulletin of Symbolic Logic*, 10(3):305–333, 2004.

[4] P. Beame. A switching lemma primer. *Technical Report UW-CSE-95-07-01, University of Washington*, 1994.

[5] P. Beame and T. Pitassi. Propositional proof complexity: Past, present, and future. *Bulletin of the European Association for Theoretical Computer Science, The Computational Complexity Column (ed. E. Allender)*, 65:66–89, 1998.

[6] S. Bellantoni, T. Pitassi, and A. Urquhart. Approximation and small-depth Frege proofs. *Journal SIAM Journal on Computing*, 21(6):1161–1179, 1992.

[7] E. Ben-Sasson and P. Harsha. Lower bounds for bounded depth Frege proofs via Buss-Pudlák games. *ACM Transactions on Computational Logic*, 11(3):19:1–19:17, 2010.

[8] E. Ben-Sasson and A. Wigderson. Short proofs are narrow- resolution made simple. *Journal of the ACM*, 48(2):149–169, 2001.

[9] S. Buss. Polynomial size proofs of the propositional pigeonhole principle. *Journal of Symbolic Logic*, 52:916–927, 1987.

[10] S. R. Buss. Bounded arithmetic and propositional proof complexity. *In H. Schwichtenberg (ed.), Logic of Computation, Springer*, pages 67–122, 1995.

[11] S. R. Buss. Some remarks on the lengths of propositional proofs. *Archive for Mathematical Logic*, 34:377–394, 1995.

[12] S. R. Buss. First-order proof theory of arithmetic. *Chapter II in S. R. Buss (ed.), Handbook of Proof Theory*, pages 79–147, 1998.

[13] S. A. Cook and P. Nguyen. *Logical Foundations of Proof Complexity*. Cambridge University Press, 2010.

[14] S. A. Cook and A. R. Reckhow. The relative efficiency of propositional proof systems. *The Journal of Symbolic Logic*, 44(1):36–50, 1979.

[15] S. Dantchev and S. Riis. On relativization and complexity gap for resolution-based proof systems. *In M. Baaz, A. Makowsky (eds.), Proceedings of the 17th International Workshop Computer Science Logic, Springer Lecture Notes in Computer Science 2803*, pages 142–154, 2003.

[16] A. Durand, N. D. Jones, J. A. Makowsky, and M. More. Fifty years of the spectrum problem: survey and new results. *The Bulletin of Symbolic Logic*, 18(4):481–644, 2012.

[17] S. Feferman. Some applications of forcing and generic sets. *Fundamentae Mathematicae*, 56:325–345, 1965.

[18] S. Fenner, L. Fortnow, S. A. Kurtz, and L. Li. An oracle builder's toolkit. *Information and Computation*, 182:95–136, 2003.

[19] J. Hirschfeld and W. H. Wheeler. *Forcing, Arithmetic, Division Rings*, volume 454. Lecture Notes in Mathematics, 1975.

[20] W. Hodges. *Building Models by Games*. Cambridge University Press, 1985.

[21] H. J. Keisler. Forcing and the omitting types theorem. *in Morley (ed.) Studies in Model Theory, Studies in Mathematics, The Mathematical Association of America*, 8:96–133, 1973.

[22] J. F. Knight. Generic expansions of structures. *The Journal of Symbolic Logic*, 38(4):561–570, 1973.

[23] J. Krajíček. *Bounded Arithmetic, Propositional Logic, and Complexity Theory*. Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1995.

[24] J. Krajíček. On Frege and extended Frege proof systems. *P. Clote and J. Remmel (eds.), Feasible Mathematics II, Birkhäuser*, pages 284–319, 1995.

[25] J. Krajíček. Combinatorics of first order structures and porpositional proof systems. *Archive of Mathematical Logic*, 43:427–441, 2004.

[26] J. Krajíček. *Forcing with random variables and proof complexity*, volume 382. London Mathematical Society Lecture Note Series, Cambridge University Press, 2011.

[27] 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):15–39, 1995.

[28] K. Kunen. *Set Theory*. Studies in Logic 34, College Publications, London, revised edition, 2013.

[29] P. Odifreddi. Forcing and reducibilities. *The Journal of Symbolic Logic*, 48(2):288–310, 1983.

[30] J. Paris and A. J. Wilkie. Counting problems in bounded arithmetic. *in: Methods in Mathematical Logic*, 1130:317–340, 1985.

[31] T. Pitassi, P. Beame, and R. Impagliazzo. Exponential lower bounds for the pigeonhole principle. *Computational Complexity*, 3:97–108, 1993.

[32] B. Poizat. *Les petits cailloux. Une approche modèle-théorique de l'algorithmie*, volume 3. Nur al Matiq wal-Ma'rifah, Aléas, Lyon, 1995.

[33] P. Pudlák. A bottom-up approach to foundations of mathematics. *Proceedings Gödel'96, Logical Foundations of Mathematics, Computer Science and Physics – Kurt Gödel's Legacy, Springer Lecture Notes in Logic*, 6:81–97, 1996.

[34] P. Pudlák. The lengths of proofs. *Chapter VIII in S. R. Buss (ed.): Handbook of Proof Theory*, pages 547–637, 1998.

[35] P. Pudlák. Proofs as games. *American Mathematical Monthly*, pages 541–550, 2000.

[36] A. A. Razborov. Lower bounds for propositional proofs and independence results in bounded arithmetic. *In F. Meyer auf der Heide, B. Monien (eds.), Proceedings of the 23rd International Colloquium Automata, Languages and Programming Springer Lecture Notes in Computer Science*, 1099:48–62, 1996.

[37] A. A. Razborov. Proof complexity of pigeonhole principles. In *Proceedings of the 5th International Conference on Developments in Language Theory*, pages 100–116. Springer, 2002.

[38] S. Riis. Finitisation in bounded arithmetic. *BRICS Report Series RS-94-23*, 1994.

[39] S. Riis. A complexity gap for tree-resolution. *Computational Complexity*, 10:179–209, 2001.

[40] D. Scott. A proof of the independence of the continuum hypothesis. *Mathematical Systems Theory*, 1(2):89–111, 1967.

[41] N. Segerlind. The complexity of propositional proofs. *The Bulletin of Symbolic Logic*, 13(4):417–481, 2007.

[42] N. Segerlind, S. Buss, and R. Impagliazzo. A switching lemma for small restrictions and lower bounds for k-DNF resolution. *SIAM Journal on Computing*, 33(5):1171–1200, 2004.

[43] J. R. Shoenfield. Unramified forcing. *Axiomatic Set Theory, Proceedings of Symposia in Pure mathematics, American Mathematical Society*, VIII:357–381, 1971.

[44] S. G. Simpson. Forcing and models of arithmetic. *Proceedings of the American Mathematical Society*, 43(1):193–194, 1974.

[45] J. Stern. A new look at the interpolation theorem. *The Journal of Symbolic Logic*, 40(1):1–13, 1975.

[46] G. Takeuti and M. Yasumoto. Forcing on bounded arithmetic. *in P. Hájek (ed.), Gödel'96, Lecture Notes in Logic*, 6:120–138, 1996.

[47] G. Takeuti and M. Yasumoto. Forcing on bounded arithmetic 2. *The Journal of Symbolic Logic*, 63(3):The Journal of Symbolic Logic, 1998.

[48] N. Thapen. Notes on switching lemmas, unpublished manuscript. 2009.

[49] A. Urquhart. The complexity of propositional proofs. *The Bulletin of Symbolic Logic*, 1(4):425–467, 1995.

[50] A. Urquhart and X. Fu. Simplified lower bounds for propositional proofs. *Notre Dame Journal of Formal Logic*, 73(4):523–544, 1996.

[51] D. Zambella. Forcing in finite structures. *Mathematical Logic Quarterly*, 43(3):401–412, 1997.