The polynomial and linear hierarchies in V^0

Leszek Aleksander Kołodziejczyk*and Neil Thapen†

January 11, 2007

Abstract

We show that the bounded arithmetic theory V^0 does not prove that the polynomial time hierarchy collapses to the linear time hierarchy (without parameters). This result follows from a lower bound for bounded depth circuits computing prefix parity, where the circuits are allowed some auxiliary input.

This is a continuation of work in [KT06] where we show that this collapse is not provable in PV under a cryptographic assumption.

Keywords: prefix parity, linear hierarchy, bounded arithmetic, bounded depth circuits

Introduction

One approach to problems of structural complexity is to look at their behaviour in theories of bounded arithmetic. This allows us to consider how complexity classes behave in models not unreasonably different from the real world, and to study what logical resources are necessary to answer complexity-theoretical questions. The most important problem in this area is whether there is a model of full bounded arithmetic in which the polynomial hierarchy does not collapse to a finite level, see e.g. [Kra95].

*Institute of Mathematics, Warsaw University, Banacha 2, 02-097 Warszawa, Poland, lak@mimuw.edu.pl. This work was carried out while the author was visiting the Mathematical Institute of the Academy of Sciences of the Czech Republic in Prague.

†Mathematical Institute, Academy of Sciences of the Czech Republic, Žitná 25, CZ-115 67 Praha 1, Czech Republic, thapen@math.cas.cz. Supported in part by grant AV0Z10190503 and by the Eduard Čech Center grant LC505.

In the present paper we deal with the problem of the relation between the linear and polynomial time hierarchies, and look at it in the rather weak two-sorted theory V^0, which can be thought of as a subtheory of S^1_2. V^0 is strong enough to prove the basic properties of AC^0 circuits, but weak enough that we can use lower bounds on the strength AC^0 circuits to obtain unconditional independence results. We show the existence of a model of V^0 in which a set in the second level Σ^p_2 of the polynomial time hierarchy is not contained in the parameter free linear time hierarchy. Our result uses a simple model-theoretic construction and the following circuit bound: a bounded depth, polynomial size circuit can compute the prefix parities of only an exponentially small fraction of n-bit inputs X, even if it has access to some auxiliary input strings of length $n^{\frac{1}{4}}$ which may depend on X. Here the "prefix parity of X" is the string whose ith bit is the parity of bits $1, \ldots, i$ of X.

In the following sections we explain how complexity classes are defined in nonstandard models of arithmetic and describe the theory V^0 and some of its simple properties; we prove our main result, assuming the lower bound; and we finally prove the lower bound, as a corollary of an old theorem of Ajtai about the fraction of its inputs for which a bounded depth, polynomial size circuit correctly computes the parity bit.

This paper continues work in [KT06] where we show, under a cryptographic assumption, that some important statements about structural complexity theory are not provable in the bounded arithmetic theories S^1_2 and PV. Our assumption is that there is no probabilistic polynomial time algorithm for factoring. This guarantees the existence of a model of S^1_2 in which the injective weak pigeonhole principle fails for a polynomial time function f (that is, f is an injection from n^2 to n for some n) but in which the surjective weak pigeonhole principle holds for all polynomial time functions (that is, for any n and any p-time g, g is not a surjection from n to n^2). We use this to construct a model of PV in which the polynomial time hierarchy does not collapse to the linear time hierarchy; a model of S^1_2 in which an NP set is not in the second level of the linear time hierarchy; and a model of S^1_2 in which the polynomial hierarchy does collapse to the linear hierarchy, but does not collapse to any finite level Σ^p_i. Parameters are allowed in the definition of these hierarchies.

Here we use the same basic technique as [KT06], which is to take a

model of some theory and close an initial segment of it (and possibly a few elements more) under a certain set of functions (there PV, here FAC^0). This gives a new model in which formulas whose quantifiers only range over the common initial segment keep their truth values unchanged, but in which some formulas with bigger quantifiers will change their truth values. The present work has the advantage that it does not use any assumptions. It has two main disadvantages, besides the obvious one that V^0 is a weaker theory than PV.

The first is that our result holds for the parameter-free versions of the hierarchies, while it seems that the natural definition of the hierarchies in nonstandard models would allow parameters – we say more about this below. We deal with parameters in [KT06] essentially by iterating our basic step and using a union-of-chains construction. A similar approach does not appear to be possible here, at least using our circuit lower bound.

The second disadvantage is related to properties of the set which is not provably in the linear time hierarchy. In the present case, this set is rather artificial — for example, it is empty in the standard model, and even in nonstandard models of stronger bounded arithmetic theories. Additionally, it is from the second level Σ^p_2 of the polynomial time hierarchy, whereas the set in [KT06] is in NP. The definition of that NP set depends on having a function f in a model of PV which defines an injection from the set of numbers of length n^2 (in binary notation) into the set of numbers of length n. By the non-provability of the relativized pigeonhole principle in $I\Delta_0$ we know that there are models of V^0 in which there is a definable injection from some $n+1$ to n, giving an injection from strings of length $n+1$ to strings of length n. But this is too small a difference between domain and range, and in the absence of the ability to iterate functions polynomially many times (available in PV, but not in V^0) there seems to be no way of amplifying it. Of course the existence of a model of V^0 with a definable injection from n^2 to n is equivalent to an old open problem, about the provability of the relativized weak pigeonhole principle in $I\Delta_0$. The existence of a definable injection between strings of these lengths, instead of just the numbers, is an interesting question in its own right, and may be an easier version of this open problem.

Definitions

Most of the definitions below are based on [CN06].

We work in a language $\mathcal{L}_A^2$ of two-sorted arithmetic. We will write variables of the number or "first-order" sort as $i, j, k, \ldots$ and variables of the finite set or "second-order" sort (which we will think of as strings) as $X, Y, Z, \ldots$. The language consists of the function and predicate symbols $\{0, 1, +, \cdot, |\ |, \in, \leq, =\}$. Here $+, \cdot, \leq$ only apply to the number sort. $|X|$ is the least upper bound of the set X, or 0 if X is empty; by abuse of notation we will also use it to mean the length of X when we think of X as a string.

A Σ_0^B formula is a formula in this language in which the only quantifiers are bounded number quantifiers, that is, quantifiers of the form $\forall i < t$ or $\exists i < t$ where t is a number term (not containing i). Here a number term is one taking a value of the number sort; it is allowed to contain subterms of the form $|X|$.

A polynomially bounded string quantifier is of the form $\forall X\,(|X| < t \rightarrow \ldots)$ or $\exists X\,(|X| < t \wedge \ldots)$ where t is a number term (not containing the variable X). We will write these as $\forall X < t$ and $\exists X < t$. A linearly bounded string quantifier is defined in the same way, with the important difference that the bounding term t is not allowed to contain multiplication.

For $i \in \mathbb{N}$, the Σ_i^B formulas consist of i alternations of blocks of polynomially bounded string quantifiers, beginning with an existential quantifier, followed by a Σ_0^B formula. The Σ_i^{LIN} or "linear" formulas are defined similarly, but with linearly bounded string quantifiers. Π_i^B and Π_i^{LIN} are defined dually. Σ_∞^B and Σ_∞^{LIN} are the unions of the respective sets of formulas over all $i \in \mathbb{N}$.

It is straightforward to see that for $i \geq 1$ the sets of strings definable in the standard model by Σ_i^B formulas are exactly the sets from the ith level Σ_i^p of the polynomial hierarchy. Similarly the sets of strings definable by Σ_∞^{LIN} formulas are exactly the sets from the linear hierarchy [Lyn82, Imm87] – the linear hierarchy is not defined so robustly, so we do not seem to have the level-by-level correspondence. Hence in a nonstandard model of an arithmetical theory in this language it is natural to identify the polynomial hierarchy with the Σ_∞^B definable sets of strings and the linear hierarchy with the Σ_∞^{LIN} sets of strings.

It seems natural to allow parameters to be used in defining these sets,

firstly because this captures the idea of limiting the time bounds of our Turing machines to the standard polynomials (while letting the input and the code of the machine range over the whole model); secondly because if there is a model in which the polynomial hierarchy is contained in the linear hierarchy without parameters, then we must have this containment already in the standard model. However we are not currently able to prove our result for V^0 in the version with parameters.

V^0 is a theory of bounded arithmetic in our two-sorted language $\mathcal{L}_A^2$, based on a theory of Zambella [Zam96]. For a complete introduction to the version we use here see [CN06]. V^0 consists of a set 2-BASIC of axioms fixing the basic properties of its language and the following comprehension axiom for each Σ_0^B formula ϕ, possibly with parameters:

$$\exists Z{<}j\, \forall i{<}j\, (i \in Z \leftrightarrow \phi(i)).$$

Notice that together with the properties of the $|\ |$ function, this gives induction for Σ_0^B formulas. In fact, V^0 is conservative over $I\Delta_0$.

Let $\phi(i, \bar{X})$ be any Σ_0^B formula, with a free number variable i, some free string variables $\bar{X}$, and no other free variables. Let $t(\bar{X})$ be any number-valued term. Then ϕ and t naturally give rise to a function $F_{\phi,t}$: the output of $F_{\phi,t}$ on input $\bar{X}$ is the string of length $t(\bar{X})$ whose bits are given by the values of $\phi(i, \bar{X})$ for $i = 1$ to $t(\bar{X})$. We call the functions defined in this way the uniform FAC^0 functions. They correspond to the string functions defined by uniform families of polynomial size, bounded depth circuits.

Now let $\mathbf{M} = (N, M)$ be a model of V^0, where N is the set of number elements and M the set of string elements. For any $S \subseteq M$, let $T \subseteq M$ be the closure in $\mathbf{M}$ of S under all uniform FAC^0 functions and let U be the set of lengths of strings from T. Then (U, T) is a model of V^0; closure under the uniform FAC^0 functions is exactly what is needed to guarantee that comprehension holds.

The main theorem

We first state our lemma about small bounded depth circuits. The proof is postponed until the next section.

Lemma 1. *Let $k \in \mathbb{N}$. Let (C_n) be a family of polynomial-size, bounded depth circuits where each circuit has as input one string X of length n^2 and*

k many auxiliary input strings, each of length $\sqrt{n}$, and each circuit has as output a string Y of length n^2.

Then for all sufficiently large n, for all but a fraction of at most $2^{-\sqrt{n}}$ input strings X, C_n fails to output the prefix parity of X for any choice of auxiliary strings.

Let $\phi(A)$ be the formula "for some X with $|X| = |A|^4$, there is no prefix parity Y of X". This is Σ_2^B, since we can express "Y is the prefix parity of X" in a Σ_0^B way as

$$|Y| = |X| \wedge Y(1) \equiv X(1) \wedge \forall i{<}|X|\,(Y(i+1) \equiv Y(i) \oplus X(i+1)),$$

where we use $X(i)$ to mean the ith bit of the string X.

Theorem 2. *There is a model of V^0 in which $\phi(A)$ is not equivalent to any formula $\psi(A)$ in Σ_∞^{LIN} without parameters.*

Proof. It is enough to show that the theory

$$\mathrm{V}^0 + \{\exists A\, \neg(\phi(A) \leftrightarrow \psi(A)) : \psi \in \Sigma_\infty^{LIN}\}$$

is finitely satisfiable. So suppose for a contradiction that we have finitely many linear formulas $\psi_1, \ldots, \psi_m$ and that

$$\mathrm{V}^0 \vdash \bigvee_i \forall A,\, \phi(A) \leftrightarrow \psi_i(A).$$

We define a theory Γ with new constant symbols $U^1, \ldots, U^{m+1}$ and $n_1, \ldots, n_{m+1}$. For each $i = 1, \ldots, m+1$, each $k \in \mathbb{N}$ and each FAC^0 function F, Γ contains the sentence "$|U^i| = n_i^2$ and for all auxiliary strings $Z_1, \ldots, Z_k$, each of length $\sqrt{n_i}$, the output of $F(U^i, \bar{Z})$ is not the prefix parity of U^i". Γ also contains "$n_{i+1} > n_i^4$" for each $i = 1, \ldots, m$.

To see that Γ is finitely satisfiable in $\mathbb{N}$, consider any $k \in \mathbb{N}$ and any finite number of FAC^0 functions. By the bound on any single FAC^0 function given by the lemma, for some large n_1 there is a string U^1 of this length such that none of our finitely many functions can calculate the prefix parity of U^1, for any choice of auxiliary input. We could have chosen n_1 to be arbitrarily large, so there is no problem in finding $n_2 > n_1^4$ with the same property, and so on.

Let $\mathbf{M}$ be a model of Γ together with the theory of true arithmetic in our language $\mathcal{L}_A^2$. For each i, let $\mathbf{M}_i$ be the model of V^0 given by taking the closure in $\mathbf{M}$ of the set $\{\text{strings of length} \leq \sqrt{n_i}\} \cup \{U^i\}$ under all FAC^0 functions in $\mathbf{M}$, as in the previous section.

For each i, by our assumption ϕ must be equivalent in $\mathbf{M}_i$ to some ψ_t. Since we have more models than we have linear formulas ψ, by the pigeonhole principle there must be two models $\mathbf{M}_i$ and $\mathbf{M}_j$, with $i < j$, in both of which ϕ is equivalent to the same ψ_t.

Now let A be the string consisting of $\sqrt{n_i}$ many 1s. $\psi_t(A)$ must have the same truth value in $\mathbf{M}_i$ as in $\mathbf{M}_j$, since it only talks about strings whose lengths are linear in $|A|$ and about numbers polynomial in $|A|$, and these are the same in $\mathbf{M}_i$ and $\mathbf{M}_j$ (since, for $r \in \mathbb{N}$, numbers less than $|A|^r$ can be thought of as r-tuples of numbers less than $|A|$).

However $\phi(A)$ is false in $\mathbf{M}_j$, since $n_j \geq n_i^4$ so $\mathbf{M}_j$ is the same as $\mathbf{M}$ for all strings of length $|A|^4$ and thus contains a prefix parity for every such string. But $\phi(A)$ is true in $\mathbf{M}_i$, since U^i is in $\mathbf{M}_i$ but, by construction, the unique prefix parity (in $\mathbf{M}$) of U^i is not in $\mathbf{M}_i$.

Hence $\phi(A)$ cannot be equivalent to $\psi_t(A)$ in both $\mathbf{M}_i$ and $\mathbf{M}_j$, which gives a contradiction. □

We note that minor changes to the argument show that V^0 does not prove that the polynomial hierarchy collapses to the quadratic time hierarchy, or to any time hierarchy given by polynomials of fixed degree. Also, a similar argument shows that there exists a model in which our formula ϕ is not equivalent to any parameter-free Σ_1^B formula.

The circuit lower bound

It remains to give the proof of Lemma 1, which relies on a result of Ajtai. In the calculations of probabilities below we use the vertical lines $|\delta|$ to mean the absolute value of a real number δ.

Theorem 3 ([Ajt83]). *Let (C_n) be a polynomial size family of bounded depth circuits, where each C_n has n input bits. Let P_n be the fraction of input strings X of length n for which the output bit of C_n is the parity of X. Then for any $\epsilon > 0$, for all sufficiently large n,*

$$|P_n - \tfrac{1}{2}| < 2^{-n^{1-\epsilon}}.$$

Note that by the nonuniformity of this result the bound n can be chosen so as to depend only on ϵ and the depth d and size exponent r of the circuit family. Otherwise for arbitrarily large n there would exist some circuit D_n of depth d and size n^r with distance from $\frac{1}{2}$ greater than $2^{-n^{1-\epsilon}}$. These circuits would thus define a family (D_n) violating the theorem.

It is also worth noting that our argument appears to need Ajtai's strong bound on the advantage away from $\frac{1}{2}$ here. The bounds that can be obtained from Håstad's method of switching lemmas do not seem to be strong enough. See Chapter 8 of [Hås87].

Now consider a polynomial size family (C_n) of bounded depth circuits, where the nth circuit takes n^2 input bits and has n output bits. We think of the input as a $n \times n$ binary matrix $\bar{X}$ with rows $X_1, \ldots, X_n$. We imagine the circuit as attempting to output a "parity vector", the ith entry of which is the parity of the vector X_i. We will write C_n^i for the subcircuit which, on input $\bar{X}$, calculates bit i of the circuit's output.

We will show that the circuit outputs the correct parity vector with an appropriately small probability.

Lemma 4. *Take any $\epsilon > 0$. Fix n sufficiently large. We omit the subscript n in what follows.*

For $k \leq n$ let P^k be the probability, over input matrices $\bar{X}$, that $C^i(\bar{X}) =$ parity(X_i) for every $i \leq k$. Then

$$|P^k - \tfrac{1}{2^k}| < 2 \cdot 2^{-n^{1-\epsilon}}.$$

Proof. Let d be the depth and r the size exponent of the circuit family (C_n). We take the n given by Theorem 3 with parameters ϵ, $d+4$ and $r+1$; all the circuits in the proof will be of this size or smaller. Let $\delta = 2^{-n^{1-\epsilon}}$.

The proof is by induction. The base case, $k = 1$, follows from Theorem 3 and an averaging argument. Suppose that, for a random matrix $\bar{X}$, $C^1(\bar{X}) =$ parity(X_1) with probability more than $\frac{1}{2}+\delta$. Then there must be some fixed vectors $Z_2, \ldots, Z_n$ such that if we take a random n-bit vector X_1 and give C^1 the matrix with rows $X_1, Z_2, \ldots, Z_n$ as input, then C^1 outputs the correct parity of X_1 with probability more than $\frac{1}{2} + \delta$, which is impossible. The same argument works if the probability over $\bar{X}$ is less than $\frac{1}{2} - \delta$.

Suppose the lemma is true for k. Say that $P^k = \frac{1}{2^k} + \alpha$, where $|\alpha| < 2\delta$. We will calculate P^{k+1}.

First let $\Pr(C^{k+1}(\bar{X}) = \text{parity}(X_{k+1})) = \frac{1}{2} + \beta$. By averaging, $|\beta| < \delta$.

Now consider the following function f, which takes as input a matrix $\bar{X}$ and tries to output the parity of X_{k+1}. If $C^1(\bar{X}), \ldots, C^k(\bar{X})$ correctly output the parities of $X_1, \ldots, X_k$, then f outputs $C^{k+1}(\bar{X})$. Otherwise f outputs $\neg C^{k+1}(\bar{X})$. Let $\Pr(f(\bar{X}) = \text{parity}(X_{k+1})) = \frac{1}{2} + \gamma$.

We claim that $|\gamma| < \delta$. Otherwise, again by our averaging argument, there are some fixed values of $Z_1, \ldots, Z_k, Z_{k+2}, \ldots, Z_n$ for the rows other than $k+1$ over which, for a random row X_{k+1}, f correctly calculates $\text{parity}(X_{k+1})$ with too high (or too low) a probability. This allows us to violate Theorem 3 by defining a bounded depth circuit for $\text{parity}(X_{k+1})$ as follows. Take $C^1, \ldots, C^n$ and hardwire in $Z_1, \ldots, Z_k, Z_{k+2}, \ldots, Z_n$ as the appropriate rows of the input. At the bottom of the circuit, check whether $C^1, \ldots, C^k$ compute the parities of $Z_1, \ldots, Z_k$ correctly (since these strings are fixed, we can hardwire in their parities); if so, output the output of C^k; otherwise output the inverse of C^k. This construction adds no more than four levels to depth of the circuit C_n and no more than $2n$ nodes to the size. This completes the proof of our claim, since we chose n large enough to work for circuits of this depth and size.

Now for a random matrix $\bar{X}$, f is correct in precisely two cases:

1. $C^1, \ldots, C^k$ are all correct and C^{k+1} is correct. The probability of this is P^{k+1}.

2. $C^1, \ldots, C^k$ are not all correct and C^{k+1} is not correct. The probability of this is

$$\begin{aligned} 1 - \Pr(C^1, \ldots, C^k \text{ all correct}) - \Pr(C^{k+1} \text{ correct}) + P^{k+1} \\ = 1 - (\tfrac{1}{2^k} + \alpha) - (\tfrac{1}{2} + \beta) + P^{k+1}. \end{aligned}$$

Now we can equate our two expressions for $\Pr(f \text{ is correct})$ to get

$$\tfrac{1}{2} + \gamma = P^{k+1} + 1 - (\tfrac{1}{2^k} + \alpha) - (\tfrac{1}{2} + \beta) + P^{k+1}$$

and hence

$$P^{k+1} = \tfrac{1}{2}(\tfrac{1}{2^k} + \alpha + \beta + \gamma).$$

But $|\alpha| < 2\delta$ and $|\beta|, |\gamma| < \delta$. So the advantage of P^{k+1} away from $\frac{1}{2^{k+1}}$ is smaller than 2δ, as required. □

Proof of Lemma 1. First observe that from a small bounded depth circuit computing prefix parities of strings of length n^2, we can easily produce a small bounded depth circuit computing parity vectors of $n \times n$ matrices. So, using Lemma 4, for large n any circuit C_n with fixed auxiliary inputs will successfully calculate the prefix parity for at most a fraction $2^{-n^{\frac{2}{3}}}$ of inputs X; but there are only $2^{k\sqrt{n}}$ possible auxiliary strings. Hence there are no more than a fraction $2^{-n^{\frac{2}{3}}} \cdot 2^{k\sqrt{n}} \leq 2^{-\sqrt{n}}$ of inputs X for which there is at least one auxiliary string which helps to compute the prefix parity of X. $\square$

References

[Ajt83] M. Ajtai, Σ_1^1 *formulae on finite structures*, Annals of Pure and Applied Logic **24** (1983), 1–48.

[CN06] S. Cook and P. Nguyen, *Foundations of proof complexity: Bounded arithmetic and propositional translations*, 2006, book in preparation available online at `http://www.cs.toronto.edu/~sacook/`.

[Hås87] J. T. Håstad, *Computational limitations for small depth circuits*, MIT Press, 1987.

[Imm87] N. Immerman, *Languages that capture complexity classes*, SIAM Journal on Computing **16** (1987), 760–778.

[Kra95] J. Krajíček, *Bounded arithmetic, propositional logic, and complexity theory*, Cambridge University Press, 1995.

[KT06] L. A. Kołodziejczyk and N. Thapen, *The polynomial and linear hierarchies in models where the weak pigeonhole principle fails*, preprint, 2006.

[Lyn82] James F. Lynch, *Complexity classes and theories of finite models.*, Mathematical Systems Theory **15** (1982), no. 2, 127–144.

[Zam96] D. Zambella, *Notes on polynomially bounded arithmetic*, Journal of Symbolic Logic **61** (1996), 942–966.