A Hitting Set Construction, with Applications to Arithmetic Circuit Lower Bounds

Pascal Koiran

December 7, 2009
LIP*, École Normale Supérieure de Lyon, Université de Lyon
Department of Computer Science, University of Toronto**
Pascal.Koiran@ens-lyon.fr

Abstract. A polynomial identity testing algorithm must determine whether a given input polynomial is identically equal to 0. We give a deterministic black-box identity testing algorithm for univariate polynomials of the form $\sum_{j=0}^{t} c_j X^{\alpha_j}(a+bX)^{\beta_j}$. From our algorithm we derive an exponential lower bound for representations of polynomials such as $\prod_{i=1}^{2^n}(X^i-1)$ under this form.
It has been conjectured that these polynomials are hard to compute by general arithmetic circuits. Our result shows that the "hardness from derandomization" approach to lower bounds is feasible for a restricted class of arithmetic circuits. The proof is based on techniques from algebraic number theory, and more precisely on properties of the height function of algebraic numbers.

1 Introduction

The large body of work on hardness versus randomness tradeoffs shows that the two tasks of proving lower bounds and derandomizing algorithms are roughly equivalent. This equivalence holds both in the boolean and arithmetic world. We focus here on the arithmetic world [10]. The equivalence between lower bounds and derandomization suggests a new approach to lower bounds (see e.g. [10,2]): let us derandomize algorithms first, and much-coveted lower bounds will follow. This "hardness from derandomization" approach is very appealing, but apparently has not yet led to many new lower bound results. There have been some recent advances in derandomization, however, especially for identity testing of small-depth arithmetic circuits, e.g. [24,15] and for the more difficult problem of black-box circuit reconstruction [16]. Also techniques have been developed for obtaining simultaneously lower bounds and identity tests [23], thereby reinforcing the intuition that these two problems are intimately connected.

* UMR 5668 ENS Lyon, CNRS, UCBL, INRIA.
** A part of this work was done during a visit to the Fields Institute.

In this paper we use the "hardness from derandomization" approach to obtain lower bounds for a certain class of arithmetic circuits. More precisely, we prove lower bounds for representations of univariate polynomials under the form

$$\sum_{j=0}^{t} c_j X^{\alpha_j} (a + bX)^{\beta_j}, \tag{1}$$

where the c_j, a and b are rational numbers. Polynomials of this form were first considered in [11] due to their role in the factorization of sparse bivariate polynomials. Indeed, such an expression vanishes identically if and only if $Y - a - bX$ is a linear factor of the bivariate polynomial $\sum_{j=0}^{t} c_j X^{\alpha_j} Y^{\beta_j}$.

Obviously, any univariate polynomial can be expressed under form (1) by expanding it as a sum of monomials (the resulting β_j are all 0). Representation (1) can potentially be much more compact than the "sum of monomials" representation, however, due to the presence of the possibly large exponents α_j and β_j (note that X^{α_j} can be computed in about $\log \alpha_j$ multiplications by repeated squaring; the same trick applies of course to $(a+bX)^{\beta_j}$). The presence of possibly large exponents makes lower bounds and deterministic identity testing nontrivial.

1.1 Lower Bound Statement

A simple version of our lower bound result is as follows.

Theorem 1. *Consider a family of polynomials (P_n) of the form*

$$P_n = \prod_{i=1}^{N_n} (X^i - 1). \tag{2}$$

Assume that P_n can be expressed under form (1) with t polynomially bounded in n and the bit sizes of the c_j, α_j and β_j polynomially bounded in n. Then N_n must be polynomially bounded in n as well.

We define the bit size of c_j as the sum of the bit sizes of its numerator and denominator. Note that there is no restriction on the size of the coefficients a and b in this theorem (they may grow arbitrarily fast as a function of n). Here we have expressed our result as a function of a single parameter n for the sake of clarity. We give in Theorem 4 a more precise (and slightly more general) lower bound where the dependency on each parameter is worked out carefully. In particular, we work with the

projective height $H(c)$ of the tuple $c = (c_j)$. This is a more appropriate notion of "size" of c than the naive bit size used in Theorem 1. The projective height is defined in Section 2.2.

The "obvious" arguments such as degree comparison between (1) and (2) only show that N_n must be *exponentially* bounded in n. Theorem 1 should therefore be viewed as an exponential lower bound. One can also see the exponential nature of our lower bound by considering the polynomials $\prod_{i=1}^{2^n}(X^i - 1)$: it follows from Theorem 4 that for some constant $\epsilon > 0$, these polynomials cannot be expressed under form (1) if t and the bit sizes of the c_j, α_j and β_j are bounded by $2^{\epsilon n}$.

We note that the polynomials P_n were suggested by Agrawal as good candidates for proving lower bounds. As observed by Agrawal [1], if it could be shown that P_n is hard to compute by general arithmetic circuits, it would follow that the permanent is hard to compute by arithmetic circuits. This also follows from a general result (Theorem 5 of [17], see also [7]) which roughly speaking shows the following: if the permanent has polynomial size arithmetic circuits then exponential-size products of easy-to-compute polynomials are themselves easy to compute.

Note also that there is a formal similarity between (2) and the well-known Pochhammer-Wilkinson polynomial $\prod_{i=1}^{n}(X - i)$ where roots of unity are replaced by integers. The Pochhammer-Wilkinson polynomial is widely conjectured to be hard to compute [6,7,20,25]. As explained in Section 6, it is possible to obtain a good lower bound for representations of this polynomial under form (1).

1.2 Main Ideas and Connections to Previous Work

Our lower bound is based on the construction of hitting sets for polynomials of the form (1). Recall that a hitting set $\mathcal{H}$ for a set $\mathcal{F}$ of polynomials is a (finite) set of points such that there exists for any non-identically zero polynomial $f \in \mathcal{F}$ at least one point $a \in \mathcal{H}$ such that $f(a) \neq 0$. Hitting sets are sometimes called "correct test sequences" [8]. By a natural abuse of notation, we will sometimes say that $\mathcal{H}$ is a hitting set for a polynomial f if it a hitting set for the singleton $\{f\}$.

The existence of polynomial size hitting sets for general arithmetic circuits follows from standard probabilistic arguments. A much more difficult problem is to give explicit (deterministic) constructions of small hitting sets. It is easy to see that this problem is equivalent to black-box deterministic identity testing: any hitting set for $\mathcal{H}$ yields an obvious black-box identity testing algorithm (declare that $f \equiv 0$ iff f evaluates to

0 on all the points of $\mathcal{H}$); conversely, assuming that $\mathcal{F}$ contains the identically zero polynomial, the set of points queried by a black box algorithm on the input $f \equiv 0$ must be a hitting set for $\mathcal{F}$.

There is a general connection between lower bounds and derandomization of polynomial identity testing [10]. This connection is especially apparent in the case of black-box derandomization. Namely, let $\mathcal{H}$ be a hitting set for $\mathcal{F}$. The polynomial $P = \prod_{a \in \mathcal{H}}(X - a)$ cannot belong to $\mathcal{F}$ since it is nonzero and vanishes on $\mathcal{H}$. The same remark applies to all nonzero multiples of P. If $\mathcal{F}$ is viewed as some kind of "complexity class", we have therefore obtained a lower bound against $\mathcal{F}$ by exhibiting a polynomial P which does not belong to $\mathcal{F}$. This connection between hitting sets and arithmetic lower bounds has been known for at least 30 years [8], but has led to suprisingly few lower bound results.[1] To the best of our knowledge, only one lower bound of this type is known: Agrawal [2, Corollary 65] has shown that certain multilinear polynomials cannot be computed by circuits with unbounded fanin addition gates of size $n^{2-\epsilon}$ and depth $(2 - \epsilon) \log n$. The lower bound applies to polynomials with coefficients computable in PSPACE (this complexity class was independently defined in [13], where it is called VPSPACE; further results on this class and other space-bounded classes in Valiant's model can be found in [14,21,22]).

We have pointed out in Section 1.1 that a lower bound for P_n against general arithmetic circuits would imply a lower bound for the permanent. For the same reason (Theorem 5 of [17]), a hitting set construction against general arithmetic circuits would imply a lower bound for the permanent.

Our hitting set construction builds on work from [11,12]. In [11] we designed a deterministic identity testing algorithm for expressions of the form (1) as an intermediate step toward an algorithm for the factorization of "supersparse" bivariate polynomials. Our identity testing algorithm was not black-box. Rather, it was based on a structure theorem (a so-called "gap theorem") which makes it possible to recognize easily identically zero expressions. Here we build on this work to construct hitting sets. These sets turn out to be made of roots of unity, explaining why we obtain a lower bound for polynomials of the form (2).

In terms of the class of arithmetic circuits studied, the work which seems closest to ours is by Saxena [23]. He gives lower bounds and identity testing algorithms for "diagonal circuits", i.e., sums of powers of (multivariate) linear functions, and more generally for sums of products

[1] As already observed in [8], hitting sets may be difficult to construct precisely because they yield lower bounds.

of a small number of powers of linear functions. Our circuits fall in this category since they compute sums of products of two powers of linear functions. Our results and methods are quite different, however. He uses non-black-box methods, whereas we use black-box methods. Moreover, his lower bounds break down for powers of high degree whereas we can handle high degree powers (indeed, for univariate polynomials the only challenge is to prove lower bounds for polynomials of high degree since any low degree polynomial can be represented efficiently as a sum of monomials, assuming that field constants are given for free).

1.3 Organization of the paper

As in [11,12] we use number-theoretic techniques and in particular properties of the height of algebraic numbers. Some background on the height function is provided in Section 2. Section 3 is technical: we obtain a height lower bound which we use in Section 4 to construct our hitting sets. From there, the lower bound theorem of Section 5 follows easily from the approach outlined in Section 1.2. Finally, we suggest some possible extensions of our results in Section 6.

2 Number Theory Background

In this section we provide some background on the height function, first for algebraic numbers and then more generally for points in projective space.

2.1 Heights of Algebraic Numbers

For any prime number p, the p-adic absolute value on $\mathbb{Q}$ is characterized by the following properties: $|p|_p = 1/p$, and $|q|_p = 1$ if q is a prime number different from p. For any $x \in \mathbb{Q} \setminus \{0\}$, $|x|_p$ can be computed as follows: write $x = p^\alpha y$ where p is relatively prime to the numerator and denominator of y, and $\alpha \in \mathbb{Z}$. Then $|x|_p = 1/p^\alpha$ (and of course $|0|_p = 0$). We denote by $M_\mathbb{Q}$ the union of the set of p-adic absolute values and of the usual (archimedean) absolute value on $\mathbb{Q}$.

Let $d, e \in \mathbb{Z}$ be two non-zero relatively prime integers. By definition, the height of the rational number d/e is $\max(|d|, |e|)$. There is an equivalent definition in terms of absolute values: for $x \in \mathbb{Q}$, $H(x) = \prod_{\nu \in M_\mathbb{Q}} \max(1, |x|_\nu)$. Note in particular that $H(0) = 1$.

More generally, let K be a number field (an extension of $\mathbb{Q}$ of finite degree). The set M_K of *normalized absolute values* is the set of absolute

values on K which extend an absolute value of $M_{\mathbb{Q}}$. For $\nu \in M_K$, we write $\nu|\infty$ if ν extends the usual absolute value, and $\nu|p$ if ν extends the p-adic absolute value. One defines a "relative height" H_K on K by the formula

$$H_K(x) = \prod_{\nu \in M_K} \max(1, |x|_\nu)^{d_\nu}. \tag{3}$$

Here d_ν is the so-called "local degree". For every p (either prime or infinite), $\sum_{\nu|p} d_\nu = [K : \mathbb{Q}]$. The absolute height $H(x)$ of x is $H_K(x)^{1/n}$, where $n = [K : \mathbb{Q}]$. It is independent of the choice of K. The above material is standard in algebraic number theory. More details can be found for instance in [18] or [27]. We will also need a special case of a result due Amoroso and Zannier and already used in [12].

Lemma 1. *Let θ be a root of unity and $a, b \in \mathbb{Q}$ such that $\alpha = a + b\theta$ is not a root of unity. If $\alpha \neq 0$ we have $H(\alpha) \geq C$ where $C > 1$ is an absolute constant.*

Proof. This follows from Theorem 1.1 of [3] since the cyclotomic extension $\mathbb{Q}(\theta)$ is Abelian over $\mathbb{Q}$ (see for instance [26], Section 8.4).□

2.2 Projective Height

One can define a notion of (relative) height for a point $c = (c_0, \ldots, c_t)$ in K^{t+1} by the formula

$$H_K(c) = \prod_{\nu \in M_K} |c|_\nu^{d_\nu},$$

where $|c|_\nu = \max_{0 \leq j \leq t} |c_j|_\nu$. This is the classical notion of height for a point in projective space ([9], section B.2). As a projective notion, $H_K(c)$ should be invariant by scalar multiplication. Indeed, for $\lambda \in K \setminus \{0\}$ we have $H_K(\lambda c) = H_K(c)$. This follows from the product formula:

$$\prod_{\nu \in M_K} |\lambda|_\nu^{d_\nu} = 1$$

for any $\nu \in K \setminus \{0\}$. Note also that the (relative) height of an algebraic number $x \in K$ is equal to the projective height of the point $(1, x) \in K^2$. As in the previous section, we can define an absolute height by the formula $H(c) = H_K(c)^{1/n}$ where $n = [K : \mathbb{Q}]$ and K is chosen so that $c \in K^{t+1}$.

In our main lower bound theorem (Theorem 4) we measure the size of the rational tuple $c = (c_j)$ in (1) by its projective height instead of the naive bit size used in Theorem 1. To compute the height of a rational

tuple, we first note that $H(c) = \max_j |c_j|$ if the c_j are relatively prime integers. The general case $c_j \in \mathbb{Q}$ is therefore quite easy: reduce to the same denominator to obtain integer coefficients, divide by their gcd and take the maximum of the absolute values of the resulting integers (so in particular $H(c) \in \mathbb{N}$ for any c in $\mathbb{Q}^{t+1}$).

3 A Height Lower Bound

The goal of this section is to establish the following lower bound.

Proposition 1. *Let (a, b) be a pair of rational numbers different from the five "excluded pairs" $(0, 0)$, $(\pm 1, 0)$ and $(0, \pm 1)$.*

There is a universal constant $C > 1$ such that the inequality

$$H(a + b\theta) \geq C \tag{4}$$

holds for any root of unity θ which is not a 6th root of unity.

The inequality $H(a + b\theta) \geq C$ implies in particular that $a + b\theta$ is not a root of unity, since roots of unity are of height 1.

The main tool in the proof of Proposition 1 is the height lower bound of Lemma 1. In light of this lemma, to complete the proof of Proposition 1 we just need to understand when $a + b\theta$ can be a root of unity.

Lemma 2. *Let θ be a root of unity and (a, b) a pair of rational numbers different from the five excluded pairs $(0, 0)$, $(\pm 1, 0)$ and $(0, \pm 1)$. If θ is not a 6th root of unity then $\alpha = a + b\theta$ is nonzero, and is not a root of unity.*

Proof. We will need some properties of cyclotomic polynomials. Recall that if θ is a root of unity of order n, its minimal polynomial is the cyclotomic polynomials ψ_n. By definition, the conjugates of θ are the other roots of its minimal polynomial. The roots of ψ_n are exactly the roots of unity of order n. There are $\phi(n)$ such roots, where $\phi(n)$ is Euler's totient function. It is known that $\phi(n) \geq \sqrt{n}$ for $n > 6$. We therefore have $\phi(n) \geq 3$ for $n > 6$. From this it follows that $\phi(n) \geq 3$ except for $n = 1, 2, 3, 4$ or 6.

The conclusion of the lemma clearly holds true in the case $b = 0$. We therefore assume in the remainder of the proof that $b \neq 0$.

The only rational roots of unity are $+1$ and -1, which are 6th roots of unity, hence $\alpha \neq 0$. If both θ and $a + b\theta$ happen to be roots of unity then θ lies at the intersection of the unit circle of the complex plane, and of the

circle defined by the condition $|a + bz| = 1$. By excluding the 5 excluded pairs, we have made sure that these two circles are distinct. They have therefore at most 2 intersection points. If θ' is a conjugate of θ, the point $a + b\theta'$ is also a root of unity and must therefore lie at the intersection of the two circles. Since there are at most two intersection points, θ has at most one conjugate. This happens only when θ is a root of a cyclotomic polynomial ψ_n of degree $\phi(n) \leq 2$, and we have seen that there are only 5 possible values for n. The two roots of order 4, $\pm i$, can be ruled out since $a \pm bi$ is a root of unity only when (a, b) is equal to one of the two excluded pairs $(0, \pm 1)$. We are left with the roots of unity of order 1, 2, 3 or 6, that is, with the 6th roots of unity. □

Remark 1. If $\theta^6 = 1$, $a + b\theta$ can be a root of unity for appropriate values of a and b. For instance, if $\theta = e^{i\pi/3}$ then $1 - \theta = e^{-i\pi/3}$. If $\theta = e^{2i\pi/3}$ then $1 + \theta = e^{i\pi/3}$.

Remark 2. In the remainder of this paper we will apply (4) only to p-th roots of unity where p is prime.

4 Hitting Set Construction

It is well known that roots of unity yield hitting sets for sparse polynomials.

Lemma 3. *Let K be a field of characteristic 0 and $f \in K[X]$ a nonzero univariate polynomial of degree at most d with at most m nonzero monomials. Then there are less than $m \log d$ prime numbers p for which $f(X)$ is identically zero modulo $X^p - 1$.*

Here we restrict to fields of characteristic 0 but this lemma is stated in [5] for arbitrary integral domains. A multivariate version can be found in Lemma 5 of [12]. Lemma 3 can be immediately restated in the language of hitting sets:

Lemma 4. *Let K be a field of characteristic 0, $\mathcal{P}$ a set of at least $m \log d$ prime numbers and $\mathcal{H}$ the set of all p-th roots of unity (in the algebraic closure of K) for all $p \in \mathcal{P}$.*

Then $\mathcal{H}$ is a hitting set for the set of all polynomials $f \in K[X]$ of degree at most d with at most m nonzero monomials.

In the next proposition and theorems, the projective height comes into play. Recall that this notion is defined in Section 2; in particular, we explain at the end of that section how to compute $H(c)$ when the c_j are

rational (which is the case in Theorem 2). For a rational tuple, the logarithm of the projective height gives a more appropriate notion of "size" than the naive bit size. In the next proposition, we use the projective height for tuples of algebraic numbers. Namely, following Lenstra [19] we define the height $H(p)$ of a polynomial $p = \sum_{j=0}^{t} c_j X^j \in \overline{\mathbb{Q}}[X]$ as the projective height $H(c)$.

Proposition 2. *Let $p \in \overline{\mathbb{Q}}[X]$ be a polynomial with at most $t+1$ non-zero terms. Assume that p can be written as the sum of two polynomials q and r where each monomial of q has degree at most β and each monomial of r has degree at least γ. Let $x \in \overline{\mathbb{Q}}^*$ be a root of p that is not a root of unity. If $\gamma - \beta > \log(t\,H(p))/\log H(x)$ then x is a common root of q and r.*

The proof of Proposition 2 can be found in [12]. It is essentially the same as the proof of Proposition 2.3 of [19].

Theorem 2 (Gap Theorem for Hitting Sets). *Let $f \in \mathbb{Q}[X]$ be a polynomial of the form (1), with (a, b) different from the five excluded pairs of Proposition 1. Assume without loss of generality that the sequence (β_j) is nondecreasing, and assume also there exists l such that*

$$\beta_{l+1} - \beta_l > \log(t(t+1)H(c))/\log C \tag{5}$$

where C is the constant of Proposition 1, and $H(c)$ is the projective height of the tuple $c = (c_j)$.

Let $\mathcal{H}$ be a set of roots of unity with $\theta^6 \neq 1$ for all $\theta \in \mathcal{H}$.

Let $g = \sum_{j=0}^{l} c_j X^{\alpha_j}(a+bX)^{\beta_j}$ and $h = \sum_{j=l+1}^{t} c_j X^{\alpha_j}(a+bX)^{\beta_j}$. If $\mathcal{H}$ is a hitting set for g and h, $\mathcal{H}$ is also a hitting set for $f = g + h$.

Proof. We need to show that $f(\theta) = 0$ for all $\theta \in \mathcal{H}$ implies $f = 0$. If $\theta \in \mathcal{H}$ is a root of f then $a + b\theta$ is a root of the univariate polynomial $p(X) = \sum_{j=0}^{t} c_j \theta^{\alpha_j} X^{\beta_j}$. The height of p satisfies the inequality $H(p) \leq (t+1)H(c)$. The factor $t+1$ is due to the fact that each monomial of p "comes" from at most $t+1$ terms of (1); see [12], Lemma 3 for a proof. Since $\theta^6 \neq 1$ we have $H(a+b\theta) \geq C > 1$ by Proposition 1. We can therefore apply Proposition 2, and it follows that $x = a + b\theta$ is a common root of the two univariate polynomials $q = \sum_{j=0}^{l} c_j \theta^{\alpha_j} X^{\beta_j}$ and $r = \sum_{j=l+1}^{t} c_j \theta^{\alpha_j} X^{\beta_j}$. This means exactly that $g(\theta) = h(\theta) = 0$.

If these two equalities apply to every $\theta \in \mathcal{H}$ we have $g = h = 0$ since $\mathcal{H}$ is supposed to be a hitting set for both g and h. Hence $f = g + h = 0$.□

We are now ready to state our main hitting set theorem. The bound will depend on 3 parameters:

(i) the parameter t in (1).
(ii) d, the maximal value of the α_j.
iii) an upper bound M on the projective height $H(c)$ of the tuple c.

Given t, d and M we define

$$\delta = \log(t(t+1)M)/\log C. \tag{6}$$

Notice that this is essentially the gap bound in (5).

Theorem 3 (Hitting Set Construction). *Let $\mathcal{P}$ be a set of at least $(t+1)(\delta t+1)\log(d+t\delta)$ prime numbers, with δ as in (6) and $p \geq 5$ for all $p \in \mathcal{P}$.*

Let $\mathcal{H}$ be the set of all p-th roots of unity for all $p \in \mathcal{P}$. Then $\mathcal{H}$ is a hitting set for the set of polynomials that can be represented under form (1) with $\alpha_j \leq d$ for all j, the rational tuple c of projective height $H(c) \leq M$, and (a,b) different from the two pairs $(0,\pm 1)$.

Proof. We proceed by reduction to Lemma 4. As in Theorem 2, we will assume without loss of generality that the sequence (β_j) is nondecreasing. We can of course assume that $(a,b) \neq (0,0)$ since the corresponding polynomial in (1) would be identically zero. We will also assume that $(a,b) \neq (\pm 1, 0)$. In that case, f can be written as a sum of $t+1$ monomials of degree at most d and we can apply Lemma 4: $\mathcal{H}$ is a hitting set for f since $|\mathcal{P}| \geq (t+1)\log d$ (the same argument could of course be applied to any pair (a,b) with $b=0$).

The remainder of the proof is divided in two cases. We first consider the case where there is no gap in f in the sense of Theorem 2, that is, $\beta_{l+1} - \beta_l \leq \delta$ for all l. In this case, factoring out the polynomial $(a+bX)^{\beta_0}$ if necessary, we assume without loss of generality that $\beta_0 = 0$. This is legitimate since the nonzero polynomial $(a+bX)^{\beta_0}$ does not vanish at any point of $\mathcal{H}$ (recall that the elements of $\mathcal{H}$ are irrational numbers). From the relations $\beta_0 = 0$ and $\beta_{l+1} - \beta_l \leq \delta$ we find that $\beta_t = \max_l \beta_l \leq \delta t$. Expanding each factor $(a+bX)^{\beta_j}$ in (1) as a sum of monomials, we see that f can be written as a sum of at most $(t+1)(\delta t+1)$ monomials, each of degree at most $d + t\delta$. Lemma 4 therefore implies that $\mathcal{H}$ is a hitting set for f.

We finally consider the case where there are gaps in f. By "breaking f at the gaps", we write $f = \sum_{i=1}^{s} f_i$ where each f_i is a sum of consecutive terms $c_j X^{\alpha_j}(a+bX)^{\beta_j}$ from (1). More precisely, we make sure that there is no gap inside each f_i in the sense that the difference between two consecutive exponents β_j in f_i is bounded by δ, and there is a gap between

f_i and f_{i+1} in the sense that the difference between the smallest exponent β_j in f_{i+1} and the biggest one in f_i is greater than δ.

We have seen that $\mathcal{H}$ is a hitting set for each of the f_i. Applying Theorem 4 repeatedly ($s-1$ times), we see that $\mathcal{H}$ is a hitting set for f as well.□

Remark 3. The pair $(a,b) = (0,\pm 1)$ is excluded from Theorem 3. This case can easily be handled with Lemma 4: f is a sum of $t+1$ monomials of degree at most $d+d'$, where $d' = \max_j \beta_j$. We can therefore replace the set $\mathcal{P}$ in Theorem 3 by a set of prime numbers of cardinality at least $(t+1)\log(d+d')$. By contrast, the bound in Theorem 3 does not depend on d'. Also, we can construct a single hitting set which covers uniformly the two cases $(a,b) \neq (0,\pm 1)$ and $(a,b) = (0,\pm 1)$ by replacing the bound $(t+1)(\delta t+1)\log(d+t\delta)$ in Theorem 3 by the maximum of this bound and $(t+1)\log(d+d')$.

5 Lower Bound Theorem

As explained in Section 1.2, it is straightforward to obtain a lower from our hitting set construction.

Theorem 4 (Main Lower Bound). *Let $\mathcal{P}$ be a set of prime numbers with $p \geq 5$ for all $p \in \mathcal{P}$,*

$$|\mathcal{P}| \geq (t+1)\max(\log(d+d'), (\delta t+1)\log(d+t\delta)) \tag{7}$$

and δ as in (6). The polynomial

$$P = \prod_{i\in\mathcal{P}} (X^{p_i} - 1)$$

cannot be expressed under form (1) if $\alpha_j \leq d$ and $\beta_j \leq d'$ for all j, and if the rational tuple c is of projective height $H(c) \leq M$. The same lower bound applies to all nonzero multiples of P.

Proof. Let f be a polynomial which can be expressed under form (1) with $\alpha_j \leq d$ and $\beta_j \leq d'$ for all j, and $H(c) \leq M$. Let Q be a multiple of P. By Theorem 3 and the remark following it, the set of roots of Q is a hitting set for f. Hence we cannot have $f = Q$, unless $Q = 0$.□

Theorem 1 follows from Theorem 4 since there are $\Omega(N/\log N)$ prime numbers in the interval $[2, N]$.

6 Further Remarks

One can try to extend our results in various ways. One possible direction is prove lower bounds for other polynomials than polynomials of the form $\prod_i (X^i - 1)$. It was recently shown in [4] that nonzero polynomials represented under form (1) have at most $6t - 4$ real roots. As a result, any set of $6t - 3$ real numbers is a hitting set and we have lower bounds for polynomials with many real roots such as $\prod_{i=1}^{2^n}(X - i)$.

Perhaps more importantly, one can look for lower bounds under more general representations than (1). We make two suggestions below.

1. Consider expressions of the form

$$\sum_{j=0}^{t} c_j (a + bX)^{\alpha_j} (c + dX)^{\beta_j}. \tag{8}$$

   Assuming that $b \neq 0$, the change of variable $Y = a+bX$ brings us back to (1) and we can use the black-box algorithm of the present paper or the non-black-box algorithm of [11] to perform deterministic identity testing. Unfortunately, the change of variable $Y = a+bX$ is non-black-box and as a result we do not have a lower bound for polynomials of the form $\prod_i (X^i - 1)$. Nevertheless, the set of real numbers is invariant under this change of variable. As a result, it follows again from [4] that any set of $6t - 3$ real numbers is a hitting set for (8) and we still have exponential lower bounds for polynomials such as $\prod_{i=1}^{2^n}(X - i)$.
   The case $b = 0$ is even simpler: now we have polynomials of the form

$$\sum_{j=0}^{t} c'_j (c + dX)^{\beta_j},$$

   where $c'_j = c_j a^{\alpha_j}$. The change of variable $Y = c + dX$ shows that by Descarte's rule of signs, such a polynomial can have at most $2t+1$ real roots if it is nonzero. We can therefore construct a hitting set (any set of $2t + 2$ real numbers will do) and derive good lower bounds.
2. Consider now expressions of the form

$$\sum_{j=0}^{t} c_j X^{\alpha_j} (a_j + b_j X)^{\beta_j}.$$

   In (1) we have $a_j = a$ and $b_j = b$ for all j. Is deterministic identity testing feasible, either in a black-box or non-black-box way ? Is it possible to derive lower bounds for this form of polynomial representation ?

Acknowledgments

This work was to a large extent triggered by a question of Erich Kaltofen: can the polynomial $(X^n - 1)/(X - 1)$ be represented efficiently under form (1) ?

References

1. M. Agrawal. A possible pseudorandom generator against arithmetic circuits. Talk at the Daimi Workshop on Algebraic Complexity Theory. Aarhus, September 2008.
2. M. Agrawal. Proving lower bounds via pseudo-random generators. In *Proc. FSTTCS 2005*. Invited survey.
3. F. Amoroso and U. Zannier. A relative Dobrowolski lower bound over Abelian varieties. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. 4*, 29(3):711–727, 2000.
4. Avendano, M. The number of roots of a lacunary bivariate polynomial on a line. *Journal of Symbolic Computation*, 44:1280–1284, 2009.
5. M. Bläser, M. Hardt, R. J. Lipton, and N. K. Vishnoi. Deterministically testing sparse polynomial identities of unbounded degree. *Information Processing Letters*, 109(3):187–192, 2009.
6. P. Bürgisser. On implications between P-NP hypotheses: decision versus computation in algebraic complexity. In *Proc. 26th International Symposium on Mathematical Foundations of Computer Science (MFCS 2001)*, pages 3–17. Springer, 2001. Invited paper.
7. P. Bürgisser. On defining integers in the counting hierarchy and proving lower bounds in algebraic complexity. In *Proc. STACS 2007*, pages 133–144, 2007. Full version: ECCC Report No. 113, August 2006.
8. J. Heintz and C.-P. Schnorr. Testing polynomials which are easy to compute. In *Logic and Algorithmic (an International Symposium held in honour of Ernst Specker)*, pages 237–254. Monographie n^o 30 de L'Enseignement Mathématique, 1982. Preliminary version in *Proc. 12th ACM Symposium on Theory of Computing*, pages 262-272, 1980.
9. M. Hindry and J. H. Silverman. *Diophantine Geometry: an Introduction*, volume 201 of *Graduate Texts in Mathematics*. Springer, 2000.
10. V. Kabanets and R. Impagliazzo. Derandomizing polynomial identity test means proving circuit lower bounds. *Computational Complexity*, 13(1-2):1–46, 2004.
11. E. Kaltofen and P. Koiran. On the complexity of factoring bivariate supersparse (lacunary) polynomials. In *Proc. 2005 International Symposium on Symbolic and Algebraic Computation (ISSAC)*. ACM Press, 2005.
12. E. Kaltofen and P. Koiran. Finding small degree factors of multivariate supersparse (lacunary) polynomials over algebraic number fields. In *Proc. 2006 International Symposium on Symbolic and Algebraic Computation (ISSAC)*. ACM Press, 2006.
13. P. Koiran and S. Perifel. VPSACE and a transfer theorem over the reals. In *Proc. STACS 2007*, volume 4393 of *Lecture Notes in Computer Science*, pages 417–428. Springer-Verlag, 2007. Journal version to appear in *Computational Complexity*.
14. P. Koiran and S. Perifel. VPSPACE and a transfer theorem over the complex field. In *Proc.32nd International Symposium on Mathematical Foundations of Computer Science*, volume 4708 of *Lecture Notes in Computer Science*, pages 359–370. Springer, 2007.
15. Z. S. Karnin and A. Shpilka. Black box polynomial identity testing of depth-3 arithmetic circuits with bounded top fan-in. In *Proc. 23rd IEEE Conference on Computational Complexity (CCC)*, 2008.

16. Z. S. Karnin and A. Shpilka. Reconstruction of generalized depth-3 arithmetic circuits with bounded top fan-in. In *Proc. 24th IEEE Conference on Computational Complexity (CCC)*, 2009.
17. P. Koiran and S. Perifel. Interpolation in Valiant's theory, 2007. http://arxiv.org/abs/0710.0360.
18. S. Lang. *Algebra*. Addison-Wesley, 1993.
19. H. W. Lenstra. Finding small degree factors of lacunary polynomials. In *Number Theory in Progress*, pages 267–276, 1999.
20. R. J. Lipton. Straight-line complexity and integer factorization. In *Proc. First International Symposium on Algorithmic Number Theory*, volume 877 of *Lecture Notes in Computer Science*, pages 71–79. Springer, 1994.
21. M. Mahajan and B. V. R. Rao. Small-space analogues of Valiant's classes. In *Proc. 17th International Symposium on Fundamentals of Computation Theory*, volume 5699 of *Lecture Notes in Computer Science*, pages 250–261. Springer, 2009.
22. B. Poizat. À la recherche de la définition de la complexité d'espace pour le calcul des polynômes à la manière de Valiant. *Journal of Symbolic Logic*, 73(4):1179–1201, 2008.
23. N. Saxena. Diagonal circuit identity testing and lower bounds. In *Proc. 35th International Colloquium on Automata, Languages and Programming (ICALP 2008)*, LNCS 5125, pages 60–71. Springer, 2008.
24. N. Saxena and C. Seshadri. An almost optimal rank bound for depth-3 identities. In *Proc. 24th IEEE Confenrence on Computational Complexity (CCC)*, 2009.
25. M. Shub and S. Smale. On the intractability of Hilbert's Nullstellensatz and an algebraic version of "P=NP". *Duke Mathematical Journal*, 81(1):47–54, 1995.
26. B. L. van der Waerden. *Moderne Algebra*. Springer Verlag, Berlin, 1940. English transl. publ. under the title "Modern algebra" by F. Ungar Publ. Co., New York, 1953.
27. M. Waldschmidt. *Diophantine approximation on linear algebraic groups*. Springer, 2000.