THEORY OF COMPUTING, Volume 7 (2011), pp. 119–129
www.theoryofcomputing.org

Arithmetic Complexity in Ring Extensions

Pavel Hrubeš* Amir Yehudayoff†

Received: May 25, 2009; published: May 9, 2011.

Abstract: Given a polynomial f with coefficients from a field $\mathbb{F}$, is it easier to compute f over an extension ring R than over $\mathbb{F}$? We address this question, and show the following. For every polynomial f, there is a noncommutative extension ring R such that $\mathbb{F}$ is in the center of R and f has a polynomial-size formula over R. On the other hand, if $\mathbb{F}$ is algebraically closed, no commutative extension ring R can reduce formula or circuit complexity of f. To complete the picture, we prove that over any field, there exist hard polynomials with zero-one coefficients. (This is a basic theorem, but we could not find it written explicitly.) Finally, we show that low-dimensional extensions are not very helpful in computing polynomials. As a corollary, we obtain that the elementary symmetric polynomials have formulas of size $n^{O(\log\log n)}$ over any field, and that division gates can be efficiently eliminated from circuits, over any field.

ACM Classification: F.2.1

AMS Classification: 03D15

Key words and phrases: algebraic complexity, algebraic extensions

1 Introduction

We investigate the following general question: *given a field $\mathbb{F}$ and a polynomial f with coefficients from $\mathbb{F}$, is it easier to compute f over a ring extension $R \supseteq \mathbb{F}$ than over $\mathbb{F}$?* As our model of computation we take the standard model for computing polynomials, arithmetic circuits. In principle, the circuit complexity of a polynomial can decrease when working in a larger field or ring. This is related to the fact that in the circuit model, we assume that addition and multiplication of elements of the underlying ring can be

*Partially supported by NSF grant CCF 0832797.
†Partially supported by NSF grant CCF 0832797.

 DOI: 10.4086/toc.2011.v007a008

performed at unit cost, no matter how complicated the ring is. We always assume that the field $\mathbb{F}$ is in the center of the extension ring R we investigate, that is, for all $v \in \mathbb{F}$ and $a \in \mathbb{R}$ we have $va = av$.

One example, where the order of the underlying field is important, is the computation of the elementary symmetric polynomials; the elementary symmetric polynomials have small formulas over large fields, but we do not know whether they have polynomial-size formulas when the field is small. This is essentially due to a powerful interpolation argument, first suggested by Ben-Or, see [9]. In the converse direction, we know how to prove exponential lower bounds on the size of depth-three circuits over finite fields [6, 5], but we do not know how to do that over infinite fields. Nevertheless, it is not known whether circuits over larger fields are more powerful than over smaller ones.

From a different perspective, it was shown in [8] that the permanent is computable by polynomial-size formulas when working over a large Clifford algebra. In this approach, it is assumed that addition and multiplication of elements of the Clifford algebra can be performed at unit cost. In fact, the algebra has an exponential dimension and if we wanted to represent the computations as, say, matrix addition and multiplication, we would require exponentially large matrices. We can view this as computation over a ring extension of the real numbers.

We start with a mildly surprising observation (that generalizes the Clifford-based formulas for the permanent): any polynomial of degree d in n variables has a formula of size $O(dn)$ over some noncommutative ring extension. We then show that the situation is quite different in the case of commutative extensions: if a field $\mathbb{F}$ is algebraically closed and f is a polynomial over $\mathbb{F}$, then for any commutative ring extension R of $\mathbb{F}$, computing f over R is not easier than over $\mathbb{F}$. In this sense, commutative extensions are weaker than noncommutative ones. This, however, does not guarantee that for any field $\mathbb{F}$, there exist polynomials that are hard over any commutative extension of $\mathbb{F}$. To complete the picture, we show that over any field, there exist polynomials with zero-one coefficients which require large circuits over the field. Thus, when the field $\mathbb{F}$ is algebraically closed, the hard polynomials are also hard over any commutative extension of $\mathbb{F}$. This theorem seems to be a "folklore" result; it is a theorem without which proving lower bounds for algebraic circuits would be virtually impossible. However, we could not find an explicit statement of it. We note that a similar argument shows that over any field, there exist zero-one tensors of high tensor rank (or zero-one matrices that are rigid).

Finally, we consider ring extensions of a bounded linear dimension. We show that if $R \supseteq \mathbb{F}$ has, as a vector space, small dimension, then any computation over R can be 'efficiently' simulated by a computation over $\mathbb{F}$ (see Theorem 4.2 for more details). As two applications, we show that elementary symmetric polynomials have formulas of size $n^{O(\log\log n)}$ over any field, and that divisions can be efficiently eliminated from circuits over any field (the latter was known for infinite fields).

2 Preliminaries

We denote the family of sub-multisets of size d of $[n] = \{1,\dots,n\}$ by $[n]^{\{d\}}$ and we denote by $[n]^d$ the family of ordered d-tuples with entries from $[n]$. Logarithms are always to the base two.

Rings and polynomials We assume that a ring is always *a unital ring*, a ring with multiplicative identity element $1 \in R$. Let $X = \{x_1, x_2, \dots, x_n\}$ be a finite set of variables. A *monomial* is a product $x_{i_1} x_{i_2} \cdots x_{i_k}$ with $i_1 \leq i_2 \cdots \leq i_k$. We shall write it as x_I, where $I = \{i_1,\dots,i_k\}$ is a multiset. A *polynomial over a ring*

R is a finite formal sum

$$f = \sum_I f_I x_I\,,$$

where $f_I \in R$ is the *coefficient* of the monomial x_I in f. Addition of polynomials is defined by $(f+g)_I = f_I + g_I$. A product $f \cdot g$ is defined as $\sum_{I,J} f_I g_J x_{I \cup J}$, where $I \cup J$ is the union of the multisets I and J. We denote by $R[X]$ the ring of polynomials in the set X of variables with coefficients from R. Note that the variables multiplicatively commute with each other, as well as with every element of R, whether the ring R is commutative or not.

If $R^\star$ is a ring extending the ring R, the *dimension* of $R^\star$ over R, $\dim_R(R^\star)$, is defined as the smallest natural number k such that there exist $e_1, \ldots, e_k \in R^\star$ with the following properties:

1. $e_1 = 1$ and every $a \in R$ commutes with every e_i, $i \in \{1, \ldots, k\}$, and

2. every $b \in R^\star$ can be uniquely written as $\sum_{i=1}^k b_i e_i$, with $b_i \in R$.

(If no such k exists, the dimension is infinite.) Observe that if the dimension is finite and R is commutative, then R is in the center of $R^\star$. Specifically, when R is also a field, then $\dim_R(R^\star)$ is the dimension of $R^\star$ as a vector space over R.

Arithmetic circuits and formulas We are interested in the arithmetic complexity of a polynomial f over a field $\mathbb{F}$, and how this complexity can change when computing f over a ring R extending $\mathbb{F}$. The ring R, however, is no longer required to be a field (or even to be commutative).

An *arithmetic circuit* Φ over the ring R and the variables X is a directed acyclic graph with every node of indegree either two or zero, labelled in the following manner: every vertex of indegree 0 is labelled either by a variable in X or by an element of R. Every other node in Φ has indegree two and is labelled either by $\times$ or by $+$. The circuit Φ computes a polynomial $f \in R[X]$ in the obvious manner. An arithmetic circuit is called a *formula* if the outdegree of each node is one (and so the underlying graph is a directed tree). The *size* of a circuit is the number of nodes, and the *depth* of a circuit is the length of the longest directed path.

If f is a polynomial with coefficients from R, we denote by $C_R(f)$ the size of a smallest circuit over R computing f. We denote by $L_R(f)$ the size of a smallest formula over R computing f. Finally, $D_R(f)$ denotes the smallest depth of a circuit computing f in R.

3 General extensions

We start with an elementary property of arithmetic complexity measures. Let $R, R^\star$ be two rings and let $H : R \to R^\star$ be a ring homomorphism. If $f = \sum_I f_I x_I$ is a polynomial in $R[X]$, then $f_H \in R^\star[X]$ denotes the polynomial $\sum_I H(f_I) x_I$.

The following simple proposition tells us that allowing a circuit to use elements of a larger ring cannot make it "weaker."

Proposition 3.1. *Let $f \in R[X]$. Let $H : R \to R^\star$ be a ring homomorphism. Then $C_{R^\star}(f_H) \le C_R(f)$ and $L_{R^\star}(f_H) \le L_R(f)$. In particular, if $R \subseteq R^\star$, then $C_{R^\star}(f) \le C_R(f)$ and $L_{R^\star}(f) \le L_R(f)$.*

Proof. Let Φ be a smallest circuit computing f over R. Let Ψ be the circuit Φ after substituting each constant $a \in R$ by $H(a) \in R^\star$. Thus Ψ computes f_H, and so $C_{R^\star}(f_H) \le C_R(f)$. The proof of $L_{R^\star}(f_H) \le L_R(f)$ is similar. □

The following theorem shows that if one allows *noncommutative* ring extensions of arbitrary dimension, every polynomial can be computed by a polynomial-size formula.

Theorem 3.2. *Let R be a ring. Let $f \in R[X]$ be a polynomial of degree d (recall $|X| = n$). Then there exists $R^\star \supseteq R$ such that $L_{R^\star}(f) = O(dn)$.*

The ring $R^\star$ depends on the polynomial f. Moreover, it is a noncommutative ring of a large finite dimension. (The dimension can be taken roughly as n^d.) As we will see, noncommutativity and a significant increase in dimension are inevitable.

Proof. We can assume that f is homogeneous of degree d. (Otherwise introduce a new variable t and consider polynomial $\sum_j t^{d-j} H_j(f)$ with $H_j(f)$ the j-th homogeneous part of f.) Let us introduce a set Z of dn new variables, $Z = \{z_j^i \,:\, i \in \{1,\dots,d\} \text{ and } j \in \{1,\dots,n\}\}$. Let S be the ring of noncommutative polynomials over the set Z of variables with coefficients from R (defined similarly to a ring of polynomials except that the variables do not commute among themselves). Consider the formula

$$\prod_{i \in \{1,\dots,d\}} (z_1^i x_1 + z_2^i x_2 + \ldots + z_n^i x_n)\,, \tag{3.1}$$

which defines a polynomial $F \in S[X]$. The size of this formula is $O(dn)$. Recall that for $J \in [n]^{\{d\}}$, $F_J \in S$ is the coefficient of x_J in F. Let $\mathcal{I} \subseteq S$ be the ideal generated by the polynomials

$$F_J - f_J, \quad J \in [n]^{\{d\}}\,.$$

It is sufficient to prove that $\mathcal{I} \cap R = \{0\}$. For then $R^\star = S/\mathcal{I}$ extends R as $R/\mathcal{I} \cap R$ is R, and (3.1) taken as a formula over $R^\star$ computes the polynomial f (in this case, $F_J = f_J$ in the ring $R^\star$).

Our goal now is to prove that $\mathcal{I} \cap R = \{0\}$. For $K = (k_1,\dots,k_d) \in [n]^d$, let z_K denote the monomial $z_{k_1}^1 z_{k_2}^2 \cdots z_{k_d}^d$. For the rest of this proof, we call such a monomial an *ordered* monomial. For a multiset $J = \{j_1, j_2, \cdots, j_d\} \in [n]^{\{d\}}$, where $j_1 \le j_2 \le \cdots \le j_d$, let $J' = (j_1, j_2, \dots, j_d) \in [n]^d$, the d-tuple that is the ordering of J. For $K \in [n]^d$, let us define $a_K \in R$ by

$$a_K = \begin{cases} f_J & \text{if } K = J', \\ 0 & \text{if } K \ne J' \text{ for every } J \in [n]^{\{d\}}. \end{cases}$$

(Roughly speaking, this definition counts every K once so as to match f.) Since the variables in Z do not commute, every monomial $h \in S$ can be uniquely written as

$$h = h^{(1)} \cdot z_{K_1} \cdot h^{(2)} \cdot z_{K_2} \cdots h^{(s)} \cdot z_{K_s} \cdot h^{(s+1)},$$

where each z_{K_ℓ} is an ordered monomial, each $h^{(\ell)}$ is a monomial in S, and s is the maximal integer for which such a decomposition is possible. E. g., for $d = 2$ and $n = 2$, the monomial $h = z_1^1 z_2^1 z_1^2$ can be written in such a form with $h^{(1)} = z_1^1$, $z_{K_1} = z_2^1 z_1^2$ and $h^{(2)} = 1$. For h of such a form, define

$$h' = a_{K_1} \cdots a_{K_s} h^{(1)} h^{(2)} \cdots h^{(s)} h^{(s+1)}\,.$$

For $g = \sum_j b_j h_j$, where $b_j \in R$ and the $h_j \in S$ are monomials, set $g' = \sum_j b_j h'_j$. Define the ideal $\bar{\mathcal{I}} = \{g \in S \,:\, g' = 0\}$. Then $\bar{\mathcal{I}} \supseteq \mathcal{I}$, and for every nonzero $a \in R$, we have $a \notin \bar{\mathcal{I}}$ since $a' = a$. □

Corollary 3.3. *Let R be a ring. Then there exists a ring $\overline{R} \supseteq R$ such that for every $f \in \overline{R}[X]$ of degree d, $L_{\overline{R}}(f) = O(dn)$.*

The ring $\overline{R}$ depends neither on f nor on d, and it has infinite dimension over R.

Proof. Let us first show that for every ring R there exists a ring R' such that every polynomial in $f \in R[X]$ of degree d has a formula of size $O(nd)$ in R'. Let $\mathcal{F}_d \subseteq R[X]$ be the set of all polynomials of degree d over R. For every $f \in \mathcal{F}_d$, let $R_f \supseteq R$ be the extension of R with $L_{R_f}(f) = O(dn)$, given by Theorem 3.2. Let R' be direct sum of R_f ,$f \in \mathcal{F}_d$. Each R_f can be canonically embedded into R'. We can also assume that $R \subseteq R'$. This gives $L_{R'}(f) = O(dn)$ for every $f \in \mathcal{F}_d$.

Next, define a sequence of rings $R_0 = R$, and $R_{i+1} = R'_i \supseteq R_i$ (given by the above argument). Define $\overline{R} = \bigcup_{i \geq 0} R_i$. Every $f \in \overline{R}[X]$ has a finite number of coefficients, and hence there exists k such that $f \in R_k[X]$. If f has degree d, then $L_{R_{k+1}}(f) = O(dn)$. Finally, $L_{\overline{R}}(f) = O(dn)$, since $R_k \subseteq \overline{R}$. □

The situation is entirely different in the case of commutative extensions. We now show that if $\mathbb{F}$ is algebraically closed, then circuit size and formula size cannot be reduced by taking a commutative extension of $\mathbb{F}$. A similar statement appears in Chapter 4.3 of [1] and we prove it here for completeness. In other words, given a field $\mathbb{F}$ and a polynomial f over $\mathbb{F}$, the complexity of f in any commutative extension of $\mathbb{F}$ is at least the complexity of f in the algebraic closure of $\mathbb{F}$. Theoretically, this would still allow the alternative that all polynomials from $\mathbb{F}$ would have polynomial-size formulas over the algebraic closure. We shall dispense with this alternative in Theorem 3.7.

Theorem 3.4. *Assume that $\mathbb{F}$ is an algebraically closed field. Let R be a subring of $\mathbb{F}$ and let $R^\star \supseteq R$ be a commutative ring. Then for every $f \in R[X]$, we have $C_{\mathbb{F}}(f) \leq C_{R^\star}(f)$ and $L_{\mathbb{F}}(f) \leq L_{R^\star}(f)$.*

Proof. Let us argue about circuit size; formula size is similar. Let Φ be a circuit over $R^\star$ computing f. Assume that Φ contains $a_1, \ldots, a_k \in R^\star \setminus R$. Let us introduce new variables $z_1, \ldots, z_k$. Let Φ' be the circuit obtained from Φ by replacing every a_i by z_i. Thus Φ' defines a polynomial F with coefficient in $S = R[z_1, \ldots, z_k]$.

Let $\mathcal{I} \subseteq S$ be the ideal generated by the polynomials $F_J - f_J$. (Recall that F_J and f_J are the coefficients of the monomial x_J in F and f respectively.) The ideal $\mathcal{I}$ does not contain 1. This is because the equations $F_J(z_1, \ldots, z_k) - f_J = 0$ have a common solution $a_1, \ldots, a_k$ in $R^\star$, and so every polynomial h in $\mathcal{I}$ satisfies $h(a_1, \ldots, a_k) = 0$. Since $1 \notin \mathcal{I}$, Hilbert's "Weak Nullstellensatz" tells us that there exist $v_1, \ldots, v_k \in \mathbb{F}$ such that for every J, $F_J(v_1, \ldots, v_k) - f_J = 0$.

Let Φ'' be the circuit obtained by replacing every z_i in Φ' by $v_i \in \mathbb{F}$. Thus Φ'' computes the polynomial f, and the size of Φ'' is the same as the size of Φ. □

Our next goal is to show that over every field (and hence over every commutative ring), there exist "hard" polynomials with zero-one coefficients.

Lemma 3.5. *Let $\mathbb{F}$ be a field. Let $F : \mathbb{F}^n \to \mathbb{F}^m$ be a polynomial map of degree $d > 0$, that is, $F = (F_1, \ldots, F_m)$, each F_i is of degree d. Then $|F(\mathbb{F}^n) \cap \{0,1\}^m| \leq (2d)^n$.*

Proof. Without loss of generality assume that $\mathbb{F}$ is algebraically closed. (Otherwise, we can consider the closure $\overline{\mathbb{F}}$ of $\mathbb{F}$ and $|F(\mathbb{F}^n)\cap\{0,1\}^m| \leq |F(\overline{\mathbb{F}}^n)\cap\{0,1\}^m|$.) We use a few notions from algebraic geometry. For formal definitions see, for example, [4]. We start by proving the following stronger claim.

Claim 3.6. *Let $V \subseteq \mathbb{F}^n$ be an irreducible variety of dimension k and degree $r > 0$. Then*

$$|F(V)\cap\{0,1\}^m| \leq r(2d)^k .$$

Proof. The proof of the claim is by induction on k. If $k = 0$, the variety V is a single point and so $|F(V)| \leq 1$. Let $k > 0$, and assume that there exists some $i \in \{1,\dots,m\}$ such that both of the varieties $V_0 = V \cap F_i^{-1}(0)$ and $V_1 = V \cap F_i^{-1}(1)$ are nonempty. (If no such i exists, $|F(V)| \leq 1$.) Since V_0 and V_1 are proper subvarieties of V, the dimension of both V_0 and V_1 is less than k, the dimension of V. Let $\varepsilon \in \{0,1\}$, and consider V_ε. Since F_i has degree at most d, Bezout's Theorem (see Section 2.2 in [3]) tells us that the irreducible components of V_ε, say $V_\varepsilon^1,\dots,V_\varepsilon^{t_\varepsilon}$, satisfy $\sum_{j\in[t_\varepsilon]} r_\varepsilon^j \leq rd$, where $r_\varepsilon^j = \deg(V_\varepsilon^j)$. By the inductive assumption, $|F(V_\varepsilon^j)\cap\{0,1\}^m| \leq r_\varepsilon^j(2d)^{k-1}$ for every $j \in [t_\varepsilon]$. Thus,

$$|F(V)\cap\{0,1\}^m| \leq \sum_{\varepsilon\in\{0,1\}} \sum_{j\in[t_\varepsilon]} |F(V_\varepsilon^j)\cap\{0,1\}^m| \;\leq 2rd(2d)^{k-1} = r(2d)^k.$$

□

Since $\mathbb{F}$ is algebraically closed, it follows that $\mathbb{F}^n$ is an irreducible variety of dimension n and degree 1, and Claim 3.6 implies the lemma. □

The previous lemma shows that the zero-one image of a polynomial map is small. The next theorem uses the lemma to conclude that there exist hard polynomials with zero-one coefficients. The basic idea behind the proof of the theorem is that a polynomial computed by a circuit is an image of a polynomial map.

Theorem 3.7. *Let $d,n \in \mathbb{N}$ and let $\mathbb{F}$ be a field. Let $m = \binom{n+d-1}{d}$. Then there exists a homogeneous polynomial f of degree d in the variables $x_1,\dots,x_n$ with zero-one coefficients such that*

$$\begin{aligned} C_{\mathbb{F}}(f) &\geq \Omega\left(\sqrt{m}\right) \geq \Omega\left(\binom{n+d-1}{d}^{d/2}\right) \text{ and} \\ L_{\mathbb{F}}(f) &\geq \Omega\left(\frac{m}{\log m}\right) \geq \Omega\left((d\log(n+d))^{-1}\cdot\binom{n+d-1}{d}^{d}\right) . \end{aligned}$$

Proof. Let us start with the first inequality. We first consider a type of circuits we call skeletons. Let $z_1,\dots,z_s$ be new variables. A circuit Γ in the variables $x_1,\dots,x_n,z_1,\dots,z_s$ with no field elements in it is called a *skeleton*. A skeleton Γ computes a polynomial $g = \sum_J g_J x_J \in S[x_1,\dots,x_n]$ with $S = \mathbb{F}[z_1,\dots,z_s]$. We say that a skeleton Γ *defines* a polynomial $f \in \mathbb{F}[x_1,\dots,x_n]$ if there exist $v_1,\dots,v_s \in \mathbb{F}$ such that $f(x_1,\dots,x_n) = g(x_1,\dots,x_n,v_1,\dots,v_s)$. Assume that the size of Γ is at most s. Thus the degree of every g_J, as a polynomial in $z_1,\dots,z_s$, is at most 2^s. Let $F : \mathbb{F}^s \to \mathbb{F}^m$ be the map $F_I(v_1,\dots,v_s) = g_I(v_1,\dots,v_s)$ for every $I \in \{1,\dots,m\}$. (We think of I as determining a monomial of total degree exactly d.) The map F is a polynomial map of degree at most 2^s. Moreover, if Γ defines a homogeneous polynomial f, then the vector of coefficients of f is in the image of F. By Lemma 3.5, the skeleton Γ computes at most $(2^{s+1})^s$ polynomials with zero-one coefficients.

Every skeleton of size at most s contains at most $n+s$ variables, the symbols $+, \times$, and no constants. Thus, there are at most $(n+s+2)^s s^{2s}$ skeletons of size at most s (the indegree of each node is at most two). Hence, skeletons of size at most s define at most $2^{s(s+1)}(n+s+2)^s s^{2s} \leq 2^{cs^2}$ polynomials with zero-one coefficients, where $c > 0$ is a constant (we can assume $s > n$).

Back to considering general circuits. Let Φ be a circuit over $\mathbb{F}$ in the variables $x_1, \ldots, x_n$ of size at most s. The circuit Φ contains at most s elements of $\mathbb{F}$, say $a_1, \ldots, a_s$ (the ordering is arbitrary but fixed). Let Γ be the skeleton obtained from Φ by replacing every a_i by z_i. The size of Γ is at most s. In addition, if Φ computes a polynomial $f \in \mathbb{F}[x_1, \ldots, x_n]$, then Γ defines f. Since there exist 2^m homogeneous polynomials of degree d with zero-one coefficients, in order to compute all such polynomials, we must have $2^{cs^2} \geq 2^m$ and so $s \geq \Omega(m^{1/2})$. For the latter inequality in the statement of the theorem, we use the estimate

$$\binom{n+d-1}{d} \geq \left(\frac{n+d-1}{d}\right)^d .$$

To obtain a lower bound on $L_{\mathbb{F}}(f)$, we use a similar argument. Consider formula skeletons instead of circuit skeletons. The difference is that for formulas, g_J has degree at most s. Hence $F : \mathbb{F}^s \to \mathbb{F}^m$ is a map of degree at most s, and a formula skeleton defines at most $(2s)^s$ polynomials with zero-one coefficients. We note that the number of circuit skeletons above is an upper bound on the number of formula skeletons of size s, and conclude that formula skeletons define at most $(2s)^s(n+s+2)^s s^{2s} \leq 2^{cs\log s}$ polynomials with zero-one coefficients ($c > 0$ is a constant). This gives $2^{cs\log s} \geq 2^m$ and hence $s \geq \Omega(m \log^{-1} m)$. For the latter inequality in the statement of the theorem, we estimate

$$\log \binom{n+d-1}{d} \leq \log \left(\frac{\mathrm{e}(n+d-1)}{d}\right)^d = O(d\log(n+d)) .$$

To match the statement of the theorem, we must show that that the lower bounds on $C_{\mathbb{F}}$ and $L_{\mathbb{F}}$ can be achieved simultaneously. This follows from the fact that the above arguments give that majority of the polynomials with zero-one coefficients have this complexity. $\square$

Corollary 3.8. *The statement of Theorem 3.7 holds for any non-trivial commutative ring R (instead of $\mathbb{F}$).*

Proof. Let R be a non-trivial commutative ring (i. e., $0 \neq 1$), and let $\mathcal{I}$ be a maximal ideal in R. Thus $\mathbb{F} = R/\mathcal{I}$ is a field. Let f be the polynomial given by Theorem 3.7 with the field $\mathbb{F}$. Let H be the canonical homomorphism $H : R \to \mathbb{F}$. Proposition 3.1 tells us that $C_{\mathbb{F}}(f_H) \leq C_R(f)$. Finally, the polynomial f has zero-one coefficients, and so $f = f_H$. $\square$

The existence of hard polynomials with real coefficients was proved, for example, in [2]. The argument of [2] does not apply to polynomials with zero-one coefficients; the polynomials considered in [2] have algebraically independent coefficients (so-called *generic polynomials*). On the other hand, the circuit lower bound on generic polynomials essentially matches the formula lower bound from Theorem 3.7, whereas the circuit lower bound in Theorem 3.7 is roughly the square root of the "expected" value.

Tensor rank and matrix rigidity Lemma 3.5 can be applied to several other problems; for example, it implies the existence of zero-one tensors of high tensor rank, and zero-one matrices of high rigidity.

Valiant defined the concept of matrix rigidity [12], and proved the existence of matrices of high rigidity over any field. In the case of a finite field, the proof is by a counting argument, and in the infinite case, by a dimension argument. The matrices take as entries all possible field elements. The existence of rigid zero-one matrices was proved in [7], in the case of the field of real numbers. The authors use a real number version of Lemma 3.5 which is due to Warren [13]. Warren's theorem is a stronger version of Lemma 3.5, but it applies exclusively to $\mathbb{R}$.

Here is a sketch of a proof of the existence of tensors of high rank. Consider a three dimensional tensor, say, $T : [n]^3 \to \mathbb{F}$. Recall that a *rank one* tensor is a tensor such that $t(x_1,x_2,x_3) = t_1(x_1)t_2(x_2)t_3(x_3)$ for every $x_1,x_2,x_3 \in [n]$. Also recall that the *tensor rank* of a tensor T is defined as the minimal integer r such that $T = \sum_{i\in[r]} t^{(i)}$ with $t^{(i)}$ of rank one. Let us count the zero-one tensors of rank at most r. Each rank-one tensor $t^{(i)} = t_1^{(i)} t_2^{(i)} t_3^{(i)}$ is defined by $3n$ variables, n variables for each $t_j^{(i)}$. Think of $t^{(i)}$ as a polynomial map from $\mathbb{F}^{3n}$ to $\mathbb{F}^{n^3}$; as such it has degree three. Similarly, a rank-r tensor is a polynomial map from $\mathbb{F}^{3nr}$ to $\mathbb{F}^{n^3}$. Lemma 3.5 tells us that the number of zero-one tensors of rank at most r is at most 6^{3nr}. On the other hand, the number of zero-one tensors is 2^{n^3}. Thus, there exist zero-one tensors of tensor rank $\Omega(n^2)$.

4 Extensions of small dimension

We start by observing that formulas can be assumed to be balanced. (This type of argument is well-known, see, e. g., [10].)

Lemma 4.1. *Let R be a ring. Then for every polynomial f, $c^{-1}D_R(f) \le \log L_R(f) \le D_R(f)$, where $c \ge 1$ is a universal constant.*

Proof. Since the number of nodes in a tree with indegree at most two of depth ℓ is at most 2^ℓ, $\log L_R(f) \le D_R(f)$. The proof of the other inequality is by induction on the size of the formula. Let Φ be a smallest formula for f, that is, $L_R(f) = s$ where s is the size of Φ. Since the indegree is at most two, let u be a node in Φ so that Φ_u, the subformula of Φ rooted at u, is of size between $s/3$ and $2s/3$. Denote by g the polynomial computed by u, and denote by f_a, $a \in R$, the polynomial computed at the output of Φ after deleting the two edges (and hence the two subformulas) going into u and labelling it by a. Thus,

$$f = (f_1 - f_0)g + f_0$$

(this follows by induction on the structure of Φ). All the polynomials g, f_0 and f_1 have formulas of size at most $2s/3$. By induction, every $h \in \{g, f_0, f_1\}$ admits $D_R(h) \le c \log L_R(h)$; let Φ^h be a formula for h of depth $D_R(h) \le c\log(2s/3)$. Set

$$\Psi = (\Phi^{f_1} + (-1) \times \Phi^{f_0}) \times \Phi^g + \Phi^{f_0}.$$

Thus, Ψ computes f and its depth is at most $c\log(2s/3) + 4 \le c \log s$ with $c \ge 1$ a constant. $\square$

The following theorem shows that extensions of low dimensions are not extremely helpful when computing polynomials.

Theorem 4.2. *Let R and $R^\star \supseteq R$ be rings such that $\dim_R(R^\star) = k$. Let $f \in R[X]$. Then*

$$C_R(f) \le O(k^3) C_{R^\star}(f) \quad \textit{and} \quad L_R(f) \le (L_{R^\star}(f))^{O(\log k)}.$$

Proof. Let $1 = e_1$ and $e_2, \ldots, e_k$ be elements of $R^\star$ such that every $a \in R^\star$ can be uniquely written as

$$\sum_{i \in \{1,\ldots,k\}} a_i e_i \,.$$

We denote $\overline{a} = \langle a_1, \ldots, a_k \rangle$. Addition and multiplication in $R^\star$ can be performed as $(a+b)_i = a_i + b_i$ and $(a \cdot b)_i = \lambda_i(\overline{a}, \overline{b})$, where λ_i is a bilinear map over R. Every λ_i is computable by a circuit ϕ_i over R of size at most ck^2 and depth at most $c \log k$ with $c > 0$ a constant.

For $g = \sum g_J x_J \in R^\star[X]$ and $i \in \{1, \ldots, k\}$, define $g_i \in R[X]$ as $\sum_J g_{J,i} x_J$, where $g_{J,i} = (g_J)_i$. Denote $\overline{g} = \langle g_1, \ldots, g_k \rangle$. For every $i \in \{1, \ldots, k\}$,

$$(g+h)_i = g_i + h_i \quad \text{and} \quad (g \cdot h)_i = \lambda_i(\overline{g}, \overline{h})\,;$$

for example,

$$(g \cdot h)_i = \sum_{I,J} (g_I h_J)_i x_I x_J = \sum_{I,J} \lambda_i(\overline{g}_I, \overline{h}_J) x_I x_J = \lambda_i\Big(\sum_I \overline{g}_I x_I, \sum_J \overline{h}_J x_J\Big) = \lambda_i(\overline{g}, \overline{h})\,.$$

We start by considering circuit size. Let Φ be an arithmetic circuit over $R^\star$ of size s and depth d computing f. We simulate the computations in $R^\star$ by the computation in R. Let us define a new circuit Ψ over R as follows. For every node u in Φ that computes g^u, the circuit Ψ contains $k+1$ nodes $u_0, \ldots, u_k$ computing $g^u_0, \ldots, g^u_k$. We define Ψ inductively as follows. If u is a leaf, then g^u is either a variable or an element of $R^\star$. In this case, label each u_i by $(g^u)_i$. If $u = v + w$ is a sum node, set $u_i = v_i + w_i$. If $u = v \times w$ is a product node, set $u_i = \lambda_i(\overline{v}, \overline{w})$.

The size of Ψ is at most $O(k^3)s$ and its depth is at most $d \cdot O(\log k)$. If u is the output node of Φ, then u_0 computes the polynomial f_0. Since $f \in R[x_1, \ldots, x_n]$, we have $f = f_0$, and hence u_0 computes f. This shows that $C_R(f) \le O(k^3) C_{R^\star}(f)$.

For formula size, we use Lemma 4.1. Let Φ be a circuit computing f over $R^\star$ of depth $D = D_{R^\star}(f) \le c \log L_{R^\star}(f)$. The above argument tells us that f has a circuit over R of depth at most $D \cdot O(\log k)$ computing f. Lemma 4.1 now tells us that $L_R(f) \le 2^{D \cdot O(\log k)} \le L_{R^\star}(f)^{O(\log k)}$. □

One consequence of Theorem 4.2 is that formulas over $\mathbb{R}$ can polynomially simulate formulas over $\mathbb{C}$ (the same holds for circuits). More exactly, $L_{\mathbb{C}}(f) \le (L_{\mathbb{R}}(f))^{c_1}$ and $C_{\mathbb{C}}(f) \le c_2 C_{\mathbb{R}}(f)$ for appropriate constants $c_1, c_2 > 0$. Here are two more applications of Theorem 4.2.

Elementary symmetric polynomials over finite fields The elementary symmetric polynomial of degree k is the polynomial

$$\sum_{i_1 < i_2 < \cdots < i_k \in [n]} x_{i_1} x_{i_2} \cdots x_{i_k}\,.$$

Ben-Or showed that over large fields the elementary symmetric polynomials have formulas of size $O(n^2)$ (see [9]). This construction is by interpolation, which requires the existence of $n+1$ distinct field elements. Nevertheless, Theorem 4.2 implies that we can find relatively small formulas for the symmetric polynomials over any finite field as well. Indeed, let $\mathbb{F}$ be a finite field and let $\mathbb{E}$ be an extension of $\mathbb{F}$ of dimension $2\log n$, so that the order of $\mathbb{E}$ is at least $n+1$. As mentioned above, over $\mathbb{E}$ we can use interpolation to construct a formula of size $O(n^2)$ for the symmetric polynomial. Now, Theorem 4.2 tells us that we can simulate this formula by a formula of size $n^{O(\log\log n)}$ over $\mathbb{F}$.

Eliminating division nodes over finite fields *A circuit with divisions* is a circuit that has one more type of nodes, divisions gates, that are labelled by $/$. Every node of such a circuit computes a formal rational function, an element of the field $\mathbb{F}(X)$. We require that for any node u/v in the circuit, v computes a nonzero rational function g (though it may happen that $g=0$ on every input from $\mathbb{F}$). In [11], Strassen proved that, over an infinite field, if a polynomial $f \in \mathbb{F}[X]$ of degree d is computable by a circuit Φ of size s with divisions, it is also computable by a circuit Ψ of size $O(d^2 s)$ without divisions (see also [1], Chapter 7.1).

Strassen's argument works over any field which is large enough: Without loss of generality, we can assume that Φ contains exactly one division node u/v that computes g_1/g_2 with polynomials $g_1, g_2 \in \mathbb{F}[X]$. (This is true as we can efficiently simulate the denominator and numerator separately without using divisions.) For the argument to work, it is sufficient to find $a_1,\ldots,a_n \in \mathbb{F}$ such that $g_2(a_1,\ldots,a_n) \neq 0$. Since the degree of g_2 is at most 2^s, it is enough to have $|\mathbb{F}| \geq 2^s+1$ (as the Schwartz-Zippel Lemma tells us).

Here is how we can make it work over small fields as well. If $\mathbb{F}$ is not large enough, we can take a field extension $\mathbb{E}$ of degree greater than s and compute f over $\mathbb{E}$ efficiently. Now, Theorem 4.2 implies that we can, in fact, compute f over $\mathbb{F}$ by a circuit of size $O(s^3 d^2 s) = \mathrm{poly}(s,d)$.

Acknowledgement We wish to thank Avi Wigderson for helpful discussions.

References

[1] PETER BÜRGISSER, MICHAEL CLAUSEN, AND MOHAMMAD A. SHOKROLLAHI: *Algebraic complexity theory*. Volume 315 of *Grundlehren der mathematischen Wissenschaften*. Springer, 1997. 123, 128

[2] PETER BÜRGISSER, THOMAS LICKTEIG, AND MICHAEL SHUB: Test complexity of generic polynomials. *J. Complexity*, 8(3):203–215, 1992. [doi:10.1016/0885-064X(92)90022-4] 125

[3] V. I. DANILOV: Algebraic varieties and schemes. In I. R. SHAFAREVICH, editor, *Algebraic Geometry: Algebraic Curves, Algebraic Manifolds and Schemes*, volume 23 of *Algebraic Geometry I, Encyclopedia of Mathematical Sciences*, pp. 167–297. Springer, 1994. 124

[4] ZEEV DVIR: Extractors for varieties. In *Proc. 24th IEEE Conf. Comput. Complexity (CCC'09)*, pp. 102–113. IEEE Comp. Soc. Press, 2009. [doi:10.1109/CCC.2009.7] 124

[5] D. GRIGORIEV AND A. RAZBOROV: Exponential lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. *Appl. Algebra Engrg. Comm. Comput.*, 10(6):465–487, 2000. [doi:10.1007/s002009900021] 120

[6] DIMA GRIGORIEV, MAREK KARPINSKI, AND ANDREW CHI-CHIH YAO: An exponential lower bound on the size of algebraic decision trees for Max. *Comput. Complexity*, 7(3):193–203, 1998. [doi:10.1007/s000370050010] 120

[7] P. PUDLÁK AND V. RÖDL: Some combinatorial-algebraic problems from complexity theory. *Discrete Math.*, 136(1–3):253–279, 1994. [doi:10.1016/0012-365X(94)00115-Y] 126

[8] RENÉ SCHOTT AND G. STACEY STAPLES: Reductions in computational complexity using Clifford algebras. *Adv. in Appl. Clifford Algebr.*, 20:121–140, 2010. [doi:10.1007/s00006-008-0143-2] 120

[9] AMIR SHPILKA AND AVI WIGDERSON: Depth-3 arithmetic formulae over fields of characteristic zero. In *Proc. IEEE 14th Conf. Comput. Complexity (CCC'99)*, pp. 87–96. IEEE Comp. Soc. Press, 1999. [doi:10.1109/CCC.1999.766267] 120, 127

[10] P. M. SPIRA: On time-hardware complexity tradeoffs for Boolean functions. In *Proc. 4th. Hawaii Intern. Symp. Syst. Sciences.*, pp. 525–527, 1971. 126

[11] VOLKER STRASSEN: Vermeidung von Divisionen. *J. Reine Angew. Math.*, 264:184–202, 1973. [doi:10.1515/crll.1973.264.184] 128

[12] LESLIE G. VALIANT: Graph-theoretic properties in computational complexity. *J. Comput. System Sci.*, 13(3):278–285, 1976. [doi:10.1016/S0022-0000(76)80041-4] 126

[13] HUGH E. WARREN: Lower bounds for approximations by nonlinear manifolds. *Trans. Amer. Math. Soc.*, 133:167–178, 1968. [doi:10.2307/1994937] 126

AUTHORS

Pavel Hrubeš
Postdoctoral fellow
Princeton University, NJ
pahrubes@gmail.com

Amir Yehudayoff
Senior lecturer
Technion, Haifa, Israel
amir.yehudayoff@gmail.com

ABOUT THE AUTHORS

PAVEL HRUBEŠ graduated from Charles University at Prague in 2004; his advisor was Pavel Pudlák. His thesis focused on proof complexity.

AMIR YEHUDAYOFF graduated from Weizmann Institute in 2008; his advisor was Ran Raz. His thesis focused on algebraic complexity with an emphasis on multilinear computation.