Affine Projections of Symmetric Polynomials

Amir Shpilka
Institute of Computer Science
Hebrew University
Jerusalem, Israel
amirs@cs.huji.ac.il

Abstract

In this paper we introduce a new model for computing polynomials - a depth 2 circuit with a symmetric gate at the top and plus gates at the bottom, i.e the circuit computes a symmetric function in linear functions - $S_m^d(\ell_1, \ell_2, ..., \ell_m)$ *(*S_m^d *is the d'th elementary symmetric polynomial in* m *variables, and the* ℓ_i*'s are linear functions). We refer to this model as the symmetric model. This new model is related to standard models of arithmetic circuits, especially to depth 3 circuits. In particular we show that in order to improve the results of [19], i.e to prove super-quadratic lower bounds for depth 3 circuits, one must first prove a super-linear lower bound for the symmetric model.*

We prove two nontrivial linear lower bounds for our model. The first lower bound is for computing the determinant, and the second is for computing the sum of two monomials. The main technical contribution relates the maximal dimension of linear subspaces on which S_m^d *vanishes, and lower bounds to the symmetric model. In particular we show that an answer of the following problem (which is very natural, and of independent interest) will imply lower bounds on symmetric circuits for many polynomials:*

"What is the maximal dimension of a linear subspace of $\mathcal{C}^m$*, on which* S_m^d *vanishes ?"*

We give two partial solutions to the problem above, each enables us to prove a different lower bound. Using our techniques we also prove quadratic lower bounds for depth 3 circuits computing the elementary symmetric polynomials of degree αn *(where* $0 < \alpha < 1$ *is a constant), thus extending the result of [19]. These are the best lower bounds known for depth 3 circuits over fields of characteristic zero.*

1. Introduction

1.1. Background

Arithmetic circuits and boolean circuits are very natural models for computing polynomials. Similar to most computational models almost no lower bounds are known for these models. The best lower bound is the classical $\Omega(n \log d)$ of [21, 2] for arithmetic circuits computing a polynomial of degree d in n variables over the complex field. No such lower bound is known for circuits over fields of bounded size. In addition there is no lower bound for depth.

Since it is difficult to prove lower bounds for the general model (independent of the characteristic of the field), research focused on restricted models such as monotone circuits, bounded depth circuits and more. Exponential lower bounds were proved for monotone circuits over any field. However, the results for bounded depth circuits depend on the field.

Exponential lower bounds are known for bounded depth Boolean circuits [1, 5, 25, 9]. [20, 16] showed that even if we allow the circuit to use $\bmod p$ gates (for some fixed prime p), then it is still exponentially hard to compute the majority function.

For general bounded depth arithmetic circuits (over characteristic $\neq 2$) there are no exponential lower bounds, in contrast to the boolean case. The best lower bound (beside the classical $\Omega(n \log d)$) is a slightly super-linear lower bound of [13, 15] for polynomials of bounded degree (over any field). Not only is it difficult to prove lower bounds for general bounded depth circuits, it seems that proving lower bounds for depth 3 circuits (the first non-trivial depth) is a difficult task in itself. Over fields of characteristic > 0, [6, 7] proved exponential lower bounds for depth 3 circuits computing a generalized majority function.

In spite of these results for depth 3 circuits over fields of characteristic > 0, it seems that the situation in characteristic 0 is completely different. [11, 12] proved exponential lower bounds for restricted $\Sigma\Pi\Sigma$ circuits, but for general $\Sigma\Pi\Sigma$ circuits, there are no such strong results. It seems that although these circuits look very restricted they are actually quite powerful. For example, [3] showed a quadratic depth 3 formula that computes the elementary symmetric polynomials. This construction shows that the efforts to generalize the result of [6, 7] to circuits over fields of characteristic 0, are in vain. The best lower bound is due to [19], who showed a quadratic lower bound for depth 3 circuits computing some of the elementary symmetric polynomials, thus showing that the construction of [3] is essentially optimal. This is the best lower bound for depth 3 circuits over characteristic 0 so far, no super-quadratic lower bounds are known for this model.

1.2. Results

From now on we will only consider computations over the complex field, $\mathcal{C}$.

In this paper we introduce a new model for computing polynomials. The new model is a depth two circuit with unbounded fan-in plus gates at the bottom and a gate computing an elementary symmetric polynomial of the gates of the first level, at the top. Clearly each plus gate computes some linear function ℓ_i, and the top gate computes

$$S_m^d(\ell_1, ..., \ell_m)$$

for some d (m is the number of gates in the first level). We call m the size of the circuit, and d its degree. We call such a circuit a symmetric circuit. We show the following:

- Universality of the model: We prove that every polynomial can be computed in this model. More specifically, we show that for every polynomial f we have
$$s_{\text{sym}}(f) \leq 2^d \cdot d \cdot \text{mon}(f) \ ,$$
where $d = \deg(f)$, $s_{\text{sym}}(f) =$ the size of a smallest symmetric circuit for f, and $\text{mon}(f) =$ the number of monomials of f.

- Relation to $\Sigma\Pi\Sigma$ circuits: We show that this model is weaker than $\Sigma\Pi\Sigma$ circuits (see definition 3.2). In particular we prove the following theorem
$$s_{\text{sym}}(f) > \sqrt{s_3(f)} \ ,$$
where $s_3(f)$ is the size of a smallest $\Sigma\Pi\Sigma$ circuit computing f. As a corollary we get that super-quadratic lower bounds for depth 3 circuits imply super-linear lower bounds for the symmetric model. On the other hand if f has a $\Sigma\Pi\Sigma$ circuit with m multiplication gates, each of degree at most d then
$$s_{\text{sym}}(f) \leq d \cdot 2^d \cdot m \ .$$

- Lower bounds for the symmetric model: We prove the following lower bounds
$$s_{\text{sym}}(DET_{\sqrt{n}}) \geq 2n - 3\sqrt{n} \ ,$$
$$s_{\text{sym}}(\prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i) \geq \frac{3}{2}n - 2 \ .$$
Using our techniques we are able to extend the results of [19] and prove the following lower bound
$$s_3(S_n^d) = \Omega(d(n-d)) \ ,$$
for every $d > 1$. Thus for $d = \alpha n$ for some constant $0 < \alpha < 1$ we have
$$s_3(S_n^{\alpha n}) = \Omega(n^2) \ .$$

Although this model seems a bit obscure, it is interesting to note that similar models are already known. Our model can be viewed as $\mathrm{SYM}(\ell_1, \ldots, \ell_m)$, the similar models $\mathrm{DET}(\ell_{1,1}, \ldots, \ell_{m,m})$ and $\mathrm{PERM}(\ell_{1,1}, \ldots, \ell_{m,m})$ have already been studied. In [22] Valiant shows that the formula size of a polynomial f is at least (up to a small constant) the minimal rank of a matrix A, whose entries are linear forms, such that $det(A) = f$. In [22, 23, 24] Valiant studied p-computable families of polynomials. In this model Valiant shows that the permanent is p-complete (for p-definable polynomials under p-projections), i.e every p-definable polynomial can be represented as the permanent of a matrix with linear functions as entries, such that the rank of the matrix is the p-complexity of the polynomial (actually the linear functions in this model are not general, i.e $\ell \in \{x_i,\ -x_i,\ 0,\ 1\}$). Therefore, our model of symmetric polynomial in linear forms, is an interesting variation of the better known models of permanent and determinant in linear forms.

Another analog of our symmetric model is the model of bounded depth boolean threshold circuits (TC^0 circuits), i.e bounded depth boolean circuits with majority gates. TC^0 circuits received much attention [8, 26, 10], but no exponential lower bounds are known for them. Since the majority function is a symmetric function, we see that over the boolean domain circuits with symmetric gates have also been studied.

1.3. Techniques

The proof of the universality of the model is not very difficult, but it is very interesting to note that there is no such universal property over the real field, $\mathcal{R}$. For example, one can prove that the polynomial $xy + wz$ cannot be represented as an elementary symmetric polynomial in linear forms with real coefficients. The major difference between the real and complex fields which makes the proof work, is that the complex field contains roots of unity of large order.

The proofs of the lower bounds are more involved, they are based on a lemma relating the complexity of computing a polynomial, f, in the symmetric model, and the maximal dimension of a subspace $A \subset \mathcal{C}^n$ such that $f|_A = 0$. In order to use this lemma (lemma 4.1 of the paper) it is necessary to find an answer to the following problem

Problem 1.1 *What is the maximal dimension of a linear subspace $A \subset \mathcal{C}^m$ such that $S_m^d|_A = 0$?*

Solving this problem is an open problem of independent interest. We are able to give two partial answers, each is tight for a different range of degrees.

Theorem 1.1 *For every affine subspace, A, such that $\deg(S_m^d|_A) < d$ we have that*

$$dim(A) \leq \max(m - d, d - 1)\ ,$$

$$dim(A) < \frac{m+d}{2}\ .$$

The proof of this theorem uses some algebraic tools such as partial derivatives, symmetrization of polynomials and the algebraic independence of the elementary symmetric polynomials. This theorem also improves a theorem of a similar nature from [19].

1.4. Organization of the paper

In section 2 we define the model and prove the theorems that show that it is a universal model (i.e. that it can compute all the polynomials). We also prove an upper bound for every polynomial.

In section 3 we show the connection to depth 2 and depth 3 arithmetic circuits, and prove that super-quadratic lower bounds for depth 3 circuits imply super-linear lower bounds to the sym model.

In section 4 we state our main theorems, and deduce lower bounds from them. In section 5 we prove the theorems, and discuss their tightness.

2 Definition of the model

Definition 2.1 *We denote by $S_n^d(x_1, ..., x_n)$ the d'th elementary symmetric polynomial on n variables,*

$$S_n^d(x_1, ..., x_n) = \sum_{\substack{T \subset [n] \\ |T| = d}} \prod_{i \in T} x_i\ .$$

We also denote by $T_n^d(x_1,...,x_n)$ the d'th power sum,

$$T_n^d(x_1,...,x_n) = \sum_{i=1}^{n} {x_i}^d \ .$$

We are now ready to define our model.

Definition 2.2 *A depth 2 symmetric circuit is a layered graph with two levels. Each node at the first level computes a linear function in the set of variables X. Denote these linear functions with $L_1,...,L_m$. At the second level there is only one node, which computes $S_m^d(L_1,...,L_m)$. In other words it computes the d'th elementary symmetric polynomial in the linear functions computed in the first level of the circuit. The polynomial computed at the node of the second level is the output of the circuit. The size of this circuit is m, i.e. it is the number of nodes in the first level. The degree of the circuit is d.*

We say that a function f is computed by a symmetric circuit of size m and degree d if there is such a circuit that computes f as its output. For a function f that can be computed by a symmetric circuit we denote by $s_{\rm sym}(f)$ the size of a smallest symmetric circuit, of degree $\deg(f)$, that computes f.

Comment 1 *Although our definition allows the circuit to be of arbitrarily high degree, we will only be interested in circuits which have the same degree as the degree of the polynomial they compute. So from now on we will only be interested in computations of the form*

$$f = S_m^{\deg(f)}(\ell_1,...,\ell_m) \ .$$

It is easy to see that a symmetric circuit computes a polynomial of degree at most d in X. It is not clear however, that it can compute all the polynomials of degree d. The following theorem assures us that it can, and gives a generic upper bound for every polynomial.

Theorem 2.1 *Every polynomial $f \in \mathcal{C}[x_1,...,x_n]$ of degree d can be computed by a symmetric circuit. Moreover, $s_{\rm sym}(f) \leq d \cdot 2^d \cdot {\rm mon}(f)$, where ${\rm mon}(f) =$ the number of monomials of f.*

Remark 1 *The proof is a bit technical and it may be skipped without harm if the interest is in later sections.*

The proof of Theorem 2.1 is in two stages. First we show that every polynomial can be represented as a sum of powers of linear functions, then we show that every sum of powers of linear functions can be computed by a symmetric circuit.

In order to show that every polynomial can be represented as a sum of powers of linear functions, we first show how to represent one monomial as a sum of powers of linear functions.

Lemma 2.2 *There exist 2^d linear function $L_1,...,L_{2^d}$ such that*

$$\prod_{i=1}^{d} x_i = \sum_{i=1}^{2^d} (L_i)^d \ .$$

The proof of the lemma is just an arithmetic analog of the well known inclusion-exclusion principle.

Proof: We first prove the following claim:

Claim 2.3

$$\prod_{i=1}^{d} x_i = (-1)^d \frac{1}{d!} \sum_{T \subset [d]} (-1)^{|T|} (\sum_{i \in T} x_i)^d \ .$$

Proof of claim: The coefficient of $\prod_{i=1}^d x_i$ in the right hand-side is 1, therefore it is sufficient to prove that the coefficient of any other monomial is 0 there. Since the polynomial we compute is homogeneous, any monomial other then $\prod_{i=1}^d x_i$ does not contain all the variables, so let's look at the monomial $\prod_{i=1}^k {x_i}^{\alpha_i}$ (where $k < d$). This monomial is computed by the sub-sum

$$(-1)^d \frac{1}{d!} \sum_{[k] \subset T \subset [d]} (-1)^{|T|} (\sum_{i \in T} x_i)^d.$$

The coefficient of $\prod_{i=1}^{k} {x_i}^{\alpha_i}$ in $(\sum_{i \in T} x_i)^d$ is $\binom{d}{\alpha_1 \alpha_2 \ldots \alpha_k}$ (for every $[k] \subset T \subset [d]$). Therefore the coefficient of $\prod_{i=1}^{k} {x_i}^{\alpha_i}$ in the right hand-side is

$$(-1)^d \frac{1}{d!} \sum_{[k] \subset T \subset [d]} (-1)^{|T|} \binom{d}{\alpha_1 \alpha_2 \ldots \alpha_k} =$$

$$(-1)^{d+k} \frac{1}{d!} \binom{d}{\alpha_1 \alpha_2 \ldots \alpha_k} \sum_{r=0}^{d-k} \binom{d-k}{r} (-1)^r = 0.$$

♣

Now we can write

$$\left((-1)^d \frac{1}{d!} (-1)^{|T|} \right)^{\frac{1}{d}} = c_T$$

for some $c_T \in \mathcal{C}$. So we get:

$$\prod_{i=1}^{d} x_i = \sum_{T \subset [d]} (c_T (\sum_{i \in T} x_i))^d = \sum_{i=1}^{2^d} {L_i}^d.$$

for some linear forms $L_1, ..., L_{2^d}$. ♣

Since every polynomial is a sum of monomials we get that every polynomial can be represented as a sum of powers of linear functions.

Lemma 2.4 *Every polynomial $f \in \mathcal{C}[x_1, ..., x_n]$ of degree d can be written as sum of at most $2^d \cdot \mathrm{mon}(f)$ powers of linear functions.*

Proof: Let $f \in \mathcal{C}[x_1, ..., x_n]$ be a polynomial of degree d. Let z be a new variable. Multiply every monomial of f of degree k, such that $k < d$, with z^{d-k}. After doing so we get a new function $\tilde{f} \in \mathcal{C}[x_1, ..., x_n, z]$ such that

$$\tilde{f}(x_1, ..., x_n, 1) = f(x_1, ..., x_n).$$

According to Lemma 2.2 for every monomial M of $\tilde{f}$ there exists 2^d linear functions $\tilde{L}_{M,1}, .., \tilde{L}_{M,2^d}$ such that $M = \sum_{i=1}^{2^d} (\tilde{L}_{M,i})^d$. Since

$$\tilde{f} = \sum_{monomial} M$$

we get that

$$\tilde{f} = \sum_{monomial} \sum_{i=1}^{2^d} (\tilde{L}_{M,i})^d \ .$$

Hence there exists $r = 2^d \cdot \mathrm{mon}(\tilde{f})$ linear functions, $\{\tilde{L}_i\}_{i=1}^r$, satisfying $\tilde{f} = \sum_{i=1}^{r} (\tilde{L}_i)^d$. Let $L_i = \tilde{L}_i|_{z=1}$, we get $f = \sum_{i=1}^{r} (L_i)^d$.

Notice that if f is homogeneous then $f = \tilde{f}$ and the L_i's are linear forms. ♣

Next we show that sum of powers of linear functions can be represented as an elementary symmetric function in linear functions.

Lemma 2.5 *Let w be a primitive root of unity of order d. We have:*

$$\sum_{i=1}^{k} {y_i}^d = -S_{kd}^d(-y_1, -wy_1, -w^2 y_1, \ldots$$

$$\ldots, -w^{d-1} y_1, -y_2, -wy_2, \ldots, -w^{d-1} y_k) \ .$$

Proof: $S_m^r(z_1,...,z_m)$ is the coefficient of t^{m-r} in the polynomial $\prod_{i=1}^m (t+z_i)$. So the coefficient of t^{kd-d} in

$$\prod_{\substack{i=1,...,k \\ j=0,...,d-1}} (t - w^j y_i)$$

is

$$S_{kd}^d(-y_1, -wy_1, -w^2y_1, \ldots$$
$$\ldots, -w^{d-1}y_1, -y_2, -wy_2, ..., -w^{d-1}y_k) \ .$$

Now

$$\prod_{\substack{i=1,...,k \\ j=0,...,d-1}} (t - w^j y_i) = \prod_{i=1}^k (t^d - {y_i}^d).$$

Therefore the coefficient of t^{kd-d} is $-\sum_{i=1}^k {y_i}^d$. And the result follows. ♣

Let ρ be a $2d$ primitive root of unity. We have

$$\sum_{i=1}^k {y_i}^d = \rho^d S_{kd}^d(-y_1, -\rho^2 y_1, -\rho^4 y_1, \ldots$$
$$\ldots, -\rho^{2d-2}y_1, -y_2, -\rho^2 y_2, ..., -\rho^{2d-2}y_k) =$$
$$S_{kd}^d(-\rho y_1, -\rho^3 y_1, -\rho^5 y_1, \ldots$$
$$\ldots, -\rho^{2d-1}y_1, -\rho y_2, -\rho^3 y_2, ..., -\rho^{2d-1}y_k).$$

Now we can deduce our theorem.

Proof of Theorem 2.1: Put $r = 2^d \cdot \mathrm{mon}(f)$, and write $f = \sum_{i=1}^r (L_i)^d$ (Lemma 2.4). We have

$$f = \sum_{i=1}^r (L_i)^d = S_{rd}^d(-wL_1, -w^3L_1, -w^5L_1, \ldots$$
$$\ldots, -w^{2d-1}L_1, -wL_2, -w^3L_2, ..., -w^{2d-1}L_r)$$

($w^{2d} = 1$ a primitive root of unity). ♣

NOTE: The proof of theorem 2.1 is valid over any fields that has a $2d$-primitive root of unity, an element c such that $c^d = d!$, and the existence of $\frac{1}{d!}$.

3. Symmetric circuits vs. depth-3 circuits

In this section we show the relation between the symmetric model and depth 2 and 3 circuits. This relation actually motivated the study of symmetric circuits. In [19] quadratic lower bounds were proved for depth 3 circuits computing some explicit polynomials. However, proving super-quadratic lower bounds for $\Sigma\Pi\Sigma$ circuits, over fields of characteristic 0, remains an open problem. We prove that super-quadratic lower bounds for $\Sigma\Pi\Sigma$ circuits imply super-linear lower bounds for symmetric circuits. The difficulty to prove super-quadratic lower bounds for depth 3 circuits motivates the attempts to prove super-linear lower bounds for symmetric circuits, which we hope will be an easier task.

Definition 3.1 *A $\Sigma\Pi$ circuit is a layered depth 2 circuit with a plus gate at the top and product gates at the bottom. The size of a $\Sigma\Pi$ circuit is the number of edges in it. For a polynomial f we denote by $s_2(f)$ the size of a smallest $\Sigma\Pi$ circuit that computes f.*

This model of computation is not very interesting since the number of gates at the bottom level is clearly the number of monomials of f. Thus, for a homogeneous polynomial f we have

$$s_2(f) = \deg(f) \cdot \mathrm{mon}(f) \ .$$

Together with theorem 2.1 we get the following corollary.

Corollary 3.1 *For a homogeneous polynomial f we have: $s_{\mathrm{sym}}(f) \leq 2^{\deg(f)} \cdot s_2(f)$.*

Definition 3.2 *A $\Sigma\Pi\Sigma$ circuit is a layered depth 3 circuit, with a plus gate at the top, product gates at the middle level and plus gates at the bottom. The size of a $\Sigma\Pi\Sigma$ circuit is the number of edges between the bottom level and the middle level. For a polynomial f we denote by $s_3(f)$ the size of a smallest $\Sigma\Pi\Sigma$ circuit computing f.*

From the definition it is clear that if f is computed by a $\Sigma\Pi\Sigma$ circuit with s multiplication gates, then f has a representation of the form

$$f = \sum_{j=1}^{s} M_j$$

where

$$M_j = \prod_{i=1}^{\deg(M_j)} \ell_{i,j} \ ,$$

each $\ell_{i,j}$ is a linear function in the input variables, and $\deg(M_j)$ is the fan-in of the j'th multiplication gate. We stress that the $\ell_{i,j}$'s are linear functions and may involve a constant term (and indeed without this ability the model is homogeneous, and very restricted as shown in [11, 12]), and that different multiplication gates may use the same linear function.

A slight modification of a theorem of [3] (see also [19]) relates the sizes of a depth 3 formulae for f and a symmetric circuit for it (since it is a straight forward generalization we omit the proof).

Theorem 3.1 ([3]) *For every d, m and any m linear functions $L_1, ..., L_m$, there is a $\Sigma\Pi\Sigma$ formulae of size $m(m+1)$ computing $S_m^d(L_1, \ldots, L_m)$.*

As a corollary we get that if f can be computed by a symmetric circuit of size m then there is a depth 3 formulae of size $m(m+1)$ computing f. Combining this with corollary 3.1 we get:

Theorem 3.2 *For a homogeneous polynomial f we have:*

$$\sqrt{s_3(f)} - 1 \leq s_{\mathrm{sym}}(f) \leq 2^{\deg(f)} \cdot s_2(f) \ .$$

As a corollary we see that every super-quadratic lower bound for depth 3 circuits will imply super-linear lower bound for the symmetric model. Hence it seems an easier task to prove super-linear lower bounds for the symmetric model. We prove two nontrivial linear lower bounds, but although our lower bounds are linear it is important to notice that they do not follow from the results on $\Sigma\Pi\Sigma$ circuits, and their proofs require new techniques.

Despite the fact that we cannot prove any strong separation between the symmetric model and depth 3 circuits we conjecture the following.

Conjecture 3.1 *There is a polynomial, f, in n variables such that $s_3(f)$ is polynomial in n and $s_{\mathrm{sym}}(f)$ is exponential in n. In particular, we suspect that*

$$f(x_1, ..., x_n) = \prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i$$

is such a polynomial.

4 Results

In this section we prove lower bounds for symmetric circuits and depth 3 circuits. Our bounds for the symmetric model uses the following lemma that shows that good upper bounds on the dimension of every affine subspace $A \in \mathcal{C}^m$ such that $S^d_m|_A = 0$ imply lower bounds on many functions. Before stating the lemma we need the following definition.

Definition 4.1 *We say that a polynomial f in n variables is non-trivial, if it cannot be represented as a polynomial of less than n independent linear functions.*

Lemma 4.1 *Assume that there is a function g such that whenever $S^d_m|_A = 0$, where A is an affine subspace, the dimension of A is at most $g(m,d)$. Then for every non-trivial polynomial f in n variables, that has a subspace $B \subset \mathcal{C}^n$ of dimension D satisfying $f|_B = 0$, we must have*

$$g(s_{\text{sym}}(f), \deg(f)) \geq D \ .$$

Proof: Assume that there is a symmetric circuit for f of size m and degree $d = \deg(f)$. Then we can write

$$f(x_1, ..., x_n) = S^d_m(L_1, ..., L_m)$$

for some linear functions $L_1, ..., L_m$. Let $B \subset \mathcal{C}^n$ be a linear subspace of dimension D such that $f|_B = 0$. Notice that since $f|_B = 0$, we have

$$0 = S^d_m(L_1, ..., L_m)|_B = S^d_m(L_1|_B, ..., L_m|_B).$$

Since $dim(B) = D$, the non-triviality of f implies that there are D independent linear functions among $\{L_i|_B\}$. Hence there is a subspace of dimension D such that S^d_m vanishes on it (this subspace is the image of B under the linear transformation $(x_1, ..., x_n) \to (L_1, ..., L_m)$). According to our assumption we get

$$g(m,d) \geq D \ .$$

As a result we have

$$g(s_{\text{sym}}(f), \deg(f)) \geq D \ .$$

♣

Thus an upper bound on the dimension of a linear space A, such that $S^d_m|_A = 0$, will give lower bounds for symmetric circuits. The following theorems give such upper bounds.

Theorem 4.2 *For every affine subspace, A, such that $\deg(S^d_m|_A) < d$ we have that*

$$dim(A) \leq \max(m-d, d-1) \ .$$

This theorem is tight for $d > \frac{m}{2}$, for smaller values of d we have the following theorem.

Theorem 4.3 *For every $d \geq 2$ and every affine subspace, A, such that $\deg(S^d_m|_A) < d$ we have that*

$$dim(A) < \frac{m+d}{2} \ .$$

Combining lemma 4.1 with theorems 4.2, 4.3 yields the following corollary.

Corollary 4.1 *For every non-trivial polynomial f in n variables, that has a subspace $B \subset \mathcal{C}^n$ of dimension D satisfying $f|_B = 0$, we have:*

$$\max(S_{\text{sym}}(f) - \deg(f), \deg(f) - 1) \geq D \ ,$$

and

$$\frac{S_{\text{sym}}(f) + \deg(f)}{2} > D \ .$$

Therefore

$$S_{\text{sym}}(f) > 2D - \deg(f) \ ,$$

and if $\deg(f) - 1 < S_{\text{sym}}(f) - \deg(f)$ then we also have

$$S_{\text{sym}}(f) \geq D + \deg(f) \ .$$

Before proving the theorems we show how to use them to prove lower bounds.

4.1. Lower bounds

In this section we use lemma 4.1 to prove lower bounds for explicit functions such as the determinant, and the sum of two monomials. We also use theorem 4.3 to strengthen some results of [19].

4.1.1 Lower Bounds for the Symmetric Model

Let

$$DET_n(x_{1,1},...,x_{n,n}) = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^{n} x_{i,\sigma(i)} \ .$$

Theorem 4.4

$$s_{\text{sym}}(DET_{\sqrt{n}}) \quad \geq \quad 2n - 3\sqrt{n} \ . \tag{1}$$

$$s_{\text{sym}}(\prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i) \quad \geq \quad \frac{3}{2}n - 2 \ . \tag{2}$$

Proof: We first prove the theorem for $DET_{\sqrt{n}}$. Let B be the subspace of all $\sqrt{n} \times \sqrt{n}$ matrices whose first row is all zeros. Clearly $DET_{\sqrt{n}}|_B = 0$, $\mathrm{d}im(B) = n - \sqrt{n}$. According to corollary 4.1 we get that

$$S_{\text{sym}}(DET_{\sqrt{n}}) \geq 2(n - \sqrt{n}) - \sqrt{n} = 2n - 3\sqrt{n} \ .$$

We now prove the lower bound for $\prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i$. Let B be the subspace defined by $x_1 = x_n = 0$. Clearly $f|_B = 0$, and $\mathrm{d}im(B) = n - 2$. Plugging it into corollary 4.1 we get

$$\max(S_{\text{sym}}(\prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i) - \frac{n}{2}, \frac{n}{2} - 1) \geq n - 2 \ ,$$

hence $S_{\text{sym}}(\prod_{i=1}^{\frac{n}{2}} x_i + \prod_{i=\frac{n}{2}+1}^{n} x_i) \geq \frac{3}{2}n - 2$. ♣

Comment 2 *In order to apply lemma 4.1 to our functions, we have to show that they are non-trivial, since it is quite obvious we omit the proof.*

Although we can only prove linear lower bounds, the following argument shows, as usual, that almost all polynomials are hard for the symmetric model.

Comment 3 *The set* $\{ \ f \ | \ \deg(f) = d \ , \ s_{\text{sym}}(f) \leq m \ \}$ *is a variety of dimension* $\leq mn$ *in the space of all degree* d *polynomials in n variables. Since this space is of dimension* $\binom{n+d-1}{d}$ *we have a lower bound of*

$$s_{\text{sym}}(f) \geq \frac{\binom{n+d-1}{d}}{n}$$

for a generic polynomial f *in* n *variables of degree* d*.*

4.1.2 Lower Bounds for $\Sigma\Pi\Sigma$ circuits

In [19] the following lower bound for S_n^d was proved (Theorem 4.1 there):

Theorem: For every $\log n \leq d \leq 2n/3$ we have $s_3(S_n^d) \geq max(\Omega(\frac{n^2}{d}) \ , \ \Omega(nd))$.

The next lower bound extends this result to a wider range of degrees.

Theorem 4.5 $s_3(S_n^d) = \Omega(d(n-d))$ *for every* $d \geq 2$*.*

Proof: Assume that C is a depth 3 circuit that computes S_n^d. We want to show that there are at least $\frac{n-d}{2}$ multiplication gates of degree $\geq d$ in C. This will certainly imply the theorem. So, suppose that there are less then $\frac{n-d}{2}$ such gates. We denote these gates with $M_1, ..., M_r$ for some $r < \frac{n-d}{2}$. The following claim, which is similar to lemma 3.3 of [19], enables us to use theorem 4.3 in the scenery of depth 3 circuits.

Claim 4.6 *There is an affine subspace $A \subset \mathcal{C}^n$, such that $\forall i \leq r \;\; M_i|_A = constant$, and $dim(A) > \frac{n+d}{2}$.*

Proof: We will find linear functions $\ell_1, ..., \ell_k$ (for some $k \leq r$) such that each ℓ_i divides a different gate M_i, (i.e. there is a gate in the first level that computes ℓ_i that fans into M_i), and such that all the gates $M_1, ..., M_r$ vanish or set to constants when restricted to the set of common zeros of all the ℓ_i's.

So let $\ell_1 = \ell_{i_1}$ be some linear function that divides M_1. Let ℓ_2 be a linear function that divides M_2 such that ℓ_{i_1} and ℓ_2 have a common zero (i.e. there is $x \in \mathcal{C}^n$ such that $\ell_{i_1}(x) = \ell_2(x) = 0$). If no such ℓ_2 exists then we move to M_3. Suppose we reached M_k and we already picked $\ell_{i_1}, \ell_{i_2}, ..., \ell_{i_t}$ ($t < k$ of course). We pick ℓ_k to be a linear function that divides M_k such that $\ell_{i_1}, \ell_{i_2}, ..., \ell_{i_t}, \ell_k$ have a common zero. If there is no such ℓ_k then we proceed to M_{k+1}. After r steps we stop. Thus we have a set $\ell_{i_1}, \ell_{i_2}, ..., \ell_{i_t}$ for some $t \leq r$ such that all the ℓ's have a common zero. Let A be their set of common zeros:

$$A = \{\ x \in \mathcal{C}^n \mid \ell_{i_1}(x) = \ell_{i_2}(x) = \ldots = \ell_{i_t}(x) = 0\ \}\ .$$

Since A is defined by at most $r < \frac{n-d}{2}$ linear equations we have that $dim(A) \geq n - r > \frac{n+d}{2}$. Clearly $M_{i_j}|_A = 0$ for every $1 \leq j \leq t$. In addition we have that any other multiplication gate M_i for $1 \leq i \leq r$ is set to constant when restricted to A. This is because either one of the ℓ_{i_j} divides M_i, or all the linear functions that divide M_i are set to constants when restricted to A (this follows immediately from the fact that we couldn't find a linear function that divides M_i that is linear independent from the linear functions we already chose). ♣

Let A be the subspace guaranteed by the claim, we consider $S_n^d|_A$. Notice that $C|_A$ computes $S_n^d|_A$. In $C|_A$ all the gates are either constants or of degree less than d. Therefore $\deg(S_n^d|_A) < d$. So according to theorem 4.3 we must have $dim(A) < \frac{n+d}{2}$, in contradiction. ♣

Corollary 4.2 *For every constant $0 < \alpha < 1$, $s_3(S_n^{\alpha n}) = \theta(n^2)$.*

Proof: Immediate from theorems 3.1, 4.5. ♣

5 Proof of theorems

We first introduce the algebraic tools that we will need for the proofs.

5.1. Algebraic tools

The fact that we will use in the proofs of both theorem 4.2 and theorem 4.3 is the basic theorem of the symmetric functions. This theorem says that every symmetric polynomial can be represented as a polynomial in some elementary symmetric polynomials (see [17]).

Theorem 5.1 (basic theorem of symmetric functions) *Every symmetric polynomial*

$$f \in \mathcal{C}[x_1, ..., x_m]$$

can be written in a unique way as a polynomial in $T_m^1, ..., T_m^m$, or as a polynomial in $S_m^1, ..., S_m^m$.

This uniqueness of representation shows that $T_m^1, ..., T_m^m$ are algebraically independent, and so are $S_m^1, ..., S_m^m$.

It is interesting to note that although our theorems speak about one polynomial we need to use the algebraic independence of all the elementary symmetric polynomials.

The main technical tool in our proof is the symmetrizing of polynomials. We now define a linear operator that symmetrize polynomials.

Definition 5.1 *For a polynomial $f \in \mathcal{C}[X_1, ..., X_k]$ define*

$$\Psi(f) = \frac{1}{k!} \sum_{\sigma \in S_k} f(x_{\sigma(1)}, ..., x_{\sigma(k)}).$$

Proposition 5.2 *We have some easy observations about Ψ:*

- $\Psi(f)$ *is a symmetric polynomial of degree* $\leq \deg(f)$.
- *If g is a symmetric polynomial then* $\Psi(gf) = g\Psi(f)$.
- Ψ *is a linear operator, i.e.* $\Psi(\alpha f + \beta g) = \alpha\Psi(f) + \beta\Psi(g)$ *(α, β are constants from the field).*

Now that we have the basic algebraic tools we can proof our theorems.

5.2. Zeros of symmetric polynomials

In order to prove theorem 4.2 we need the following lemma that reduces the case of an affine subspace to the case of a linear subspace.

Lemma 5.3 *Let $A = V + u$ be an affine subspace of $\mathcal{C}^m$, where $V \subset \mathcal{C}^m$ is a linear space and $u \in \mathcal{C}^m$, such that $\deg(S^d_m|_A) < d$. Then $S^d_m|_V = 0$.*

Proof: Let A be defined by

$$\begin{gathered} x_{k+1} = \ell_1(x_1, ..., x_k) \ , \\ x_{k+2} = \ell_2(x_1, ..., x_k) \ , \\ \ldots \\ x_m = \ell_{m-k}(x_1, ..., x_k) \ . \end{gathered}$$

Write $\ell_i = L_i + \alpha_i$ where L_i is a linear form, and $\alpha_i \in \mathcal{C}$. Since $\deg(S^d_m|_A) < d$ we get that the homogeneous part of degree d of $S^d_m|_A$ is zero. Since

$$S^d_m|_A = S^d_m(x_1, ..., x_k, \ell_1, ..., \ell_{m-k}) \ ,$$

the homogeneous part of degree d is

$$S^d_m(x_1, ..., x_k, L_1, ..., L_{m-k}) \ ,$$

which equals to zero according to our assumption. Since V is the linear space defined by the equations

$$\begin{gathered} x_{k+1} = L_1(x_1, ..., x_k) \ , \\ x_{k+2} = L_2(x_1, ..., x_k) \ , \\ \ldots \\ x_m = L_{m-k}(x_1, ..., x_k) \ , \end{gathered}$$

we have that $S^d_m|_V = 0$. Hence, the result follows. ♣

We can now turn to the proof of the Theorems.

5.3. Proof of theorem 4.2

We are going to assume by contradiction that there exist a subspace A of dimension k such that $k > \max(m - d, d - 1)$ and such that $\deg(S^d_m|_A) < d$. We will show that as a consequence of our assumption we can write S^d_k as a polynomial in $S^1_k, S^2_k, ..., S^{d-1}_k$. This is a contradiction because of the basic theorem of the symmetric functions.

Proof: Lemma 5.3 allows us to assume w.l.o.g. that A is a linear subspace and that $S^d_m|_A = 0$. So assume that $dim(A) = k$, and let A be defined by

$$\begin{aligned} x_{k+1} &= L_1(x_1, ..., x_k), \\ x_{k+2} &= L_2(x_1, ..., x_k), \\ \ldots &= \ldots \\ x_m &= L_{m-k}(x_1, ..., x_k). \end{aligned}$$

In other words,

$$A = \{\ (x_1, ..., x_m) \mid x_{k+i} = L_i,\ \ \forall 1 \le i \le m-k\ \}.$$

We have

$$0 = S_m^d|_A = S_m^d(x_1, ..., x_k, L_1, ..., L_{m-k}) =$$
$$S_k^d(x_1, ..., x_k) + S_{m-k}^d(L_1, ..., L_{m-k}) +$$
$$+ \sum_{i=1}^{d-1} S_k^{d-i}(x_1, ..., x_k) \cdot S_{m-k}^i(L_1, ..., L_{m-k})\ .$$

Assume by contradiction that $k > \max(m-d, d-1)$. Since $k > m-d$, we have that $d > m-k$ and therefore

$$S_{m-k}^d(L_1, ..., L_{m-k}) = 0.$$

So we have

$$0 = S_k^d(x_1, ..., x_k) +$$

$$+ \sum_{i=1}^{d-1} S_k^{d-i}(x_1, ..., x_k) \cdot S_{m-k}^i(L_1, ..., L_{m-k})\ . \tag{3}$$

We now apply Ψ (as defined in definition 5.1) to both sides of the equation and get:

$$0 = \Psi(S_k^d(x_1, ..., x_k) +$$

$$+ \sum_{i=1}^{d-1} S_k^{d-i}(x_1, ..., x_k) \cdot S_{m-k}^i(L_1, ..., L_{m-k}))\ .$$

From the linearity of Ψ and proposition 5.2 we get that

$$0 = \Psi\left(S_k^d(x_1, ..., x_k)\right) +$$

$$+ \sum_{i=1}^{d-1} \Psi\left(S_k^{d-i}(x_1, ..., x_k) \cdot S_{m-k}^i(L_1, ..., L_{m-k})\right) =$$

$$= S_k^d(x_1, ..., x_k) +$$

$$+ \sum_{i=1}^{d-1} S_k^{d-i}(x_1, ..., x_k) \cdot \Psi\left(S_{m-k}^i(L_1, ..., L_{m-k})\right)\ .$$

Each $\Psi\left(S_{m-k}^l(L_1, ..., L_{m-k})\right)$ (where $l < d$) is a symmetric polynomial of degree at most $d-1$ in the k variables $x_1, ..., x_k$. Theorem 5.1 shows that it is a polynomial in

$$S_k^1(x_1, ..., x_k), S_k^2(x_1, ..., x_k), ..., S_k^{d-1}(x_1, ..., x_k).$$

So after applying Ψ to both sides of equation 3 we can write the result as

$$0 = S_k^d(x_1, ..., x_k) + Q(S_k^1, S_k^2, ..., S_k^{d-1}),$$

(for some polynomial Q in $d-1$ variables) i.e.

$$S_k^1(x_1, ..., x_k), S_k^2(x_1, ..., x_k), ..., S_k^d(x_1, ..., x_k)$$

are algebraic dependent. But we assumed that $k > d-1$, so by theorem 5.1 we have a contradiction. ♣

5.4. Proof of theorem 4.3

As in the proof of theorem 4.2 we will use the linear operator Ψ that symmetrizes polynomials. In addition we will also have to consider partial derivatives of polynomials.

5.4.1 Partial derivatives with respect to vectors

Definition 5.2 *For $v \in \mathcal{C}^m$, and a polynomial $f \in \mathcal{C}[X_1, ..., X_m]$ define $\partial_v(f) = \sum_{i=1}^m v_i \frac{\partial f}{\partial x_i}$ (where $v = (v_1, ..., v_m)$).*

The following proposition shows some easy properties of ∂_v. It is similar in spirit to some propositions from [4, 12].

Proposition 5.4
- *$\partial_v(L) = L(v) = L(v_1, ..., v_m)$ for every linear function L.*
- *$\partial_v(fg) = g\partial_v(f) + f\partial_v(g)$, for every pair of functions f, g.*
- *$\partial_v(\prod_{i=1}^t L_i) = \sum_{i=1}^t (L_i(v) \prod_{j \neq i} L_j)$, for every set of linear functions $L_1, ..., L_t \in \mathcal{C}[X_1, ..., X_m]$ (follows immediately from the previous properties).*

The following lemma shows that for a vector v of a special structure, $\partial_v(S_k^r(x_1, ..., x_k))$ is equal to $S_{d-1}^{r-1}(x_1, ..., x_{d-1})$, when restricted to a certain subspace of the inputs.

Lemma 5.5 *Let U be the following subspace*

$$U = \{\ (x_1, ..., x_k) \in \mathcal{C}^k \mid x_d = x_{d+1} = ... = x_k = 0\}.$$

Let $v \in \mathcal{C}^k$ be a vector satisfying:

1. *$v_1 = v_2 = ... = v_{d-1} = 1$*
2. *$\sum_{i=1}^k v_i = 0$*

Then for every $r \leq d$ we have

$$\partial_v\ (S_k^r(x_1, ..., x_k))\ |_U = (1-r)S_{d-1}^{r-1}(x_1, ..., x_{d-1})\ .$$

Proof: From the linearity of the derivative we have

$$\partial_v(S_k^r(x_1, ..., x_k))|_U = \sum_{\substack{T \subset [k] \\ |T| = r}} \partial_v(\prod_{i \in T} x_i)|_U\ .$$

According to proposition 5.4 the right hand side is equal to

$$\sum_{\substack{T \subset [k] \\ |T| = r}} (\sum_{i \in T} v_i \prod_{\substack{j \in T \\ j \neq i}} x_j)|_U =$$

$$= \sum_{\substack{T \subset [k] \\ |T| = r-1}} (\sum_{i \notin T} v_i)(\prod_{i \in T} x_i)|_U\ .$$

From the definition of U, we have that $x_d, ..., x_k = 0$. Therefore the right hand side is equal to

$$\sum_{\substack{T \subset [d-1] \\ |T| = r-1}} (\sum_{i \notin T} v_i)(\prod_{i \in T} x_i)\ .$$

Since v satisfies that

$$v_1 = v_2 = ... = v_{d-1} = 1 \ , \ and$$
$$\sum_{i=1}^{k} v_i = 0 \ ,$$

we have that for $r \leq d-1$,

$$\sum_{i=r}^{k} v_i = -(r-1) = 1-r \ .$$

So we get

$$\partial_v (S_k^r(x_1, ..., x_k))|_U = \sum_{\substack{T \subset [d-1] \\ |T| = r-1}} (1-r)(\prod_{i \in T} x_i) = \tag{4}$$
$$(1-r) S_{d-1}^{r-1}(x_1, ..., x_{d-1}) \ .$$

♣

The following basic lemma shows the conditions under which we can find a vector satisfying the properties of lemma 5.5.

Lemma 5.6 *In every vector space of dimension r there is a vector with at least r coordinates that are all 1.*

5.4.2 The proof

The proof of Theorem 4.3 has the same spirit as the proof of theorem 4.2, but in order to write S_k^d as a polynomial in $S_k^1, S_k^2, ..., S_k^{d-1}$ we need to use partial derivatives as well.

Proof of Theorem 4.3 As before, lemma 5.3 allows us to assume w.l.o.g. that A is a linear subspace and that $S_m^d|_A = 0$. So assume that $dim(A) = k$, and let A be defined by

$$\begin{aligned} x_{k+1} &= L_1(x_1, ..., x_k), \\ x_{k+2} &= L_2(x_1, ..., x_k), \\ ... &= \ ... \\ x_m &= L_{m-k}(x_1, ..., x_k). \end{aligned}$$

We have:

$$0 = S_m^d|_A = S_m^d(x_1, ..., x_k, L_1, ..., L_{m-k}) =$$
$$S_k^d(x_1, ..., x_k) + S_{m-k}^d(L_1, ..., L_{m-k}) +$$
$$\sum_{i=1}^{d-1} S_k^i(x_1, ..., x_k) \cdot S_{m-k}^{d-i}(L_1, ..., L_{m-k}) \ . \tag{5}$$

Assume by contradiction that

$$k \geq \frac{m+d}{2}$$

Define

$$L_0(x_1, ..., x_k) = x_1 + ... + x_k \ (= S_k^1(x_1, ..., x_k)) \ .$$

Let V be the vector space:

$$V = \{ \ v \in \mathcal{C}^k \mid L_i(v) = 0 \ \ \forall \, 0 \leq i \leq m-k \ \}.$$

According to our assumption, the dimension of V is at least

$$dim(V) \geq k - (m-k+1) \geq d-1.$$

According to Lemma 5.6 there is a vector $v \in V$ s.t, w.l.o.g. , its first $d-1$ coordinates are 1. Since $L_i(v) = 0 \;\; \forall 0 \leq i \leq m-k$, proposition 5.4 gives

$$\partial_v \left(S_k^1(x_1,...,x_k) \right) = \partial_v (L_0(x_1,...,x_k)) = L_0(v) = 0 \ ,$$

$$and \;\; \partial_v \left(S_{m-k}^l (L_1,...,L_{m-k}) \right) = 0 \;\; \forall 1 \leq l \leq d \ .$$

Therefore by applying ∂_v to both sides of Equation 5 we get by proposition 5.4 that

$$0 = \partial_v \left(S_k^d(x_1,...,x_k) \right) +$$

$$+ \sum_{i=2}^{d-1} \partial_v \left(S_k^i(x_1,...,x_k) \right) \cdot S_{m-k}^{d-i}(L_1,...,L_{m-k}) \ . \tag{6}$$

Let U be the subspace

$$U = \{ \; (x_1,...,x_k) \in V \mid x_d = x_{d+1} = ... = x_k = 0 \}.$$

According to Lemma 5.5 and proposition 5.2 we have that for every $r \leq d$

$$\partial_v \left(S_k^r(x_1,...,x_k) \right) |_U = (1-r) S_{d-1}^{r-1}(x_1,...,x_{d-1}) \ . \tag{7}$$

Therefore by restricting both sides of Equation 6 to U we get:

$$0 = \partial_v \left(S_k^d(x_1,...,x_k) \right) |_U +$$

$$+ \sum_{i=2}^{d-1} \left(\partial_v \left(S_k^i(x_1,...,x_k) \right) \cdot S_{m-k}^{d-i}(L_1,...,L_{m-k}) \right) |_U \ ,$$

and according to equation 7 this is equal to

$$0 = (1-d) S_{d-1}^{d-1}(x_1,...,x_{d-1}) +$$

$$+ \sum_{i=2}^{d-1} (1-i) S_{d-1}^{i-1}(x_1,...,x_{d-1}) \cdot (S_{m-k}^{d-i}(L_1,...,L_{m-k})|_U) \ .$$

We now apply Ψ to both sides of the equation

$$0 = \Psi((1-d) S_{d-1}^{d-1}(x_1,...,x_{d-1}) +$$

$$+ \sum_{i=2}^{d-1} (1-i) S_{d-1}^{i-1}(x_1,...,x_{d-1}) \cdot (S_{m-k}^{d-i}(L_1,...,L_{m-k})|_U))$$

$$= (1-d) S_{d-1}^{d-1}(x_1,...,x_{d-1}) +$$

$$\sum_{i=2}^{d-1} (1-i) S_{d-1}^{i-1}(x_1,...,x_{d-1}) \cdot \Psi \left(S_{m-k}^{d-i}(L_1,...,L_{m-k})|_U \right) .$$

For $1 \leq i \leq d-2$, $\Psi \left(S_{m-k}^i(L_1,...,L_{m-k})|_U \right)$ is a symmetric polynomial of degree at most $d-2$ in $x_1,...,x_{d-1}$ and therefore can be written as a polynomial in $S_{d-1}^1(x_1,...,x_{d-1}),...,S_{d-1}^{d-2}(x_1,...,x_{d-1})$. So we have

$$0 = (1-d) S_{d-1}^{d-1} + Q(S_{d-1}^1,...,S_{d-1}^{d-2}),$$

for some polynomial Q. But, according to theorem 5.1, $S_{d-1}^1,...,S_{d-1}^{d-2},S_{d-1}^{d-1}$ are algebraically independent, and $(1-d)$, which is the coefficient of S_{d-1}^{d-1}, is not 0, so we have a contradiction. Thus $k < \frac{m+d}{2}$. ♣

5.5. Tightness of results

In this subsection we show that theorems 4.2, 4.3 are tight in some cases.

Claim 5.7 *For $d \geq \frac{m+1}{2}$ the result of Theorem 4.2 is tight.*

Proof: According to theorem 4.2 the smallest dimension of an affine subspace A such that $S_m^d|_A = 0$ is at most $\max(m - d, d-1) = d-1$ (since $d \geq \frac{m+1}{2}$). Let A be the subspace of $\mathcal{C}^m$ defined by

$$A = \{ \ (x_1, \ldots, x_m) \ \mid \ x_d = x_{d+1} = \ldots = x_m \ \} \ .$$

Since every monomial of S_m^d contains at least one variable of $x_d, ..., x_m$ we clearly have $S_m^d|_A = 0$. Since $dim(A) = d-1$ we are done. ♣

Claim 5.8 *For $d = 2$ the result of theorem 4.3 is tight (in [14] the exact behavior of S_m^2 is determined).*

Proof: Since $d = 2$, Theorem 4.3 implies that the co-dimension of every subspace B such that $S_m^2|_B = 0$, is at least $\frac{m}{2}$. To prove a lower bound on the co-dimension write S_m^2 as

$$S_m^2 = x_m(\sum_{i=1}^{m-1} x_i) + \frac{1}{2}((\sum_{i=1}^{m-1} x_i)^2 - \sum_{i=1}^{m-1} {x_i}^2).$$

When we restrict our polynomial to

$$A = \{ \ (x_1, ..., x_m) \mid \sum_{i=1}^{m-1} x_i = 0 \ \}$$

we get:

$$S_m^2|_A = \frac{1}{2}(-\sum_{i=1}^{m-2} {x_i}^2 - (\sum_{i=1}^{m-2} x_i)^2)$$

We now have two cases, m odd and m even.

- m even. We restrict our polynomial to:

  $$\tilde{A} = \{ \ (x_1, ..., x_m) \in A \mid \sum_{l=1}^{m-2} x_l = 0, \ \text{and}$$

  $$x_{2j-1} = i \cdot x_{2j}, \ \ j = 1, ..., \frac{m-2}{2} \ \}$$

  ($i^2 = -1$), and we get

  $$S_m^2|_{\tilde{A}} = 0.$$

  Since we had $2 + \frac{m-2}{2}$ restrictions, we see that the dimension of $\tilde{A}$ is $\frac{m}{2} - 1$. Hence theorem 4.3 is tight for $d = 2$ and m even.

- m odd. In a similar way we get that the dimension of A is at least $\lceil \frac{m}{2} \rceil + 1$.

♣

As one can see, combining theorems 4.2, 4.3 with lemma 4.1 we can only get lower bounds of the form $s_{\text{sym}}(f) \geq 2n$ for polynomials in n variables. Proving a better lower bounds using lemma 4.1 seems a difficult task because of the next theorem.

Theorem 5.9 *[18] There exists a linear subspace A of dimension $\lfloor m/l \rfloor$ s.t $S_m^d|_A = 0$, where l is the smallest integer such that $l \not| d$.*

Therefore, for odd values of d we have a space A of dimension $\lfloor m/2 \rfloor$ s.t $S_m^d|_A = 0$. So lemma 4.1 cannot give a lower bound better than $2n$ for polynomials of odd degree in n variables.

The subspace A, of theorem 5.9, is constructed in such a way that in the representation of S_m^d as a polynomial in $\{T_m^i\}$, every monomial of this representation, i.e. every product of the form

$$\prod_i T_m^{d_i} \ \ where \ \ \sum d_i = d \ ,$$

is equal to zero when restricted to A.

Proof of theorem 5.9 Let w be a l-primitive root of unity. Take

$$A = \{(x_1, wx_1, w^2x_1, ..., w^{l-1}x_1, x_2, wx_2, \ldots$$
$$\ldots, w^{l-1}x_{\lfloor m/l \rfloor}, 0, ..., 0) \mid x_i \in \mathcal{C}\}.$$

Notice that if $l \not| \, k$ then $T_m^k|_A = 0$. Writing S_m^d using the Newton identities, we see that in every product $\prod_{k=1}^d (T_m^k)^{m_k}$ (s.t $\sum_{k=1}^d km_k = d$) there is some k with $l \not| \, k$ and $m_k \neq 0$. Therefore every such product is zero when restricted to A. Hence $S_m^d|_A = 0$. ♣

Thus for all odd d the function $g(m, d)$ that appears in lemma 4.1 is at most $\frac{m}{2}$. Therefore one must consider the arithmetic properties of the degree, d, when trying to answer problem 1.1.

6. Acknowledgement

I would like to thank Avi Wigderson for suggesting this line of research, and for his unstoppable encouragement in the last couple of years. I would also like to thank Michael Ben-Or, Dima Grigoriev, Jaikumar Radhakrishnan and Mike Saks for many helpful discussions during various stages of this work. I thank Jaikumar Radhakrishnan, Pranab Sen and Sundar Vishwanathan for sending me a preliminary version of their paper [14].

References

[1] M. Ajtai. Σ_1^1-formulae on finite structures. *Annals of Pure and Applied Logic*, 1983.
[2] W. Baur and V. Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22:317–330, 1983.
[3] M. Ben-Or. Private communication.
[4] R. Ehrenborg and G. C. Rota. Apolarity and canonical forms for homogeneous polynomials. *European J. Combinatorics*, 14:157–181, 1993.
[5] M. Furst, J. Saxe, and M. Sipser. Parity, circuits, and the polynomial-time hierarchy. In *FOCS*, pages 260–270, 1981.
[6] D. Grigoriev and M. Karpinski. An exponential lower bound for depth 3 arithmetic circuits. In *STOC*, pages 577–582, 1998.
[7] D. Grigoriev and A. A. Razborov. Exponential complexity lower bounds for depth 3 arithmetic circuits in algebras of functions over finite fields. In *FOCS*, pages 269–278, 1998.
[8] A. Hajnal, W. Maass, P. Pudlak, M. Szegedy, and G. Turan. Threshold circuits of bounded depth. In *FOCS*, pages 99–110, 1987.
[9] J. Håstad. Almost optimal lower bounds for small depth circuits. In *STOC*, pages 6–20, 1986.
[10] R. Impagliazzo, R. Paturi, and M. Saks. Size-depth trade-offs for threshold circuits. *SIAM Journal on Computing*, 26:693–707, 1997.
[11] N. Nisan. Lower bounds for non-commutative computation. In *STOC*, pages 410–418, 1991.
[12] N. Nisan and A. Wigderson. Lower bound on arithmetic circuits via partial derivatives. *Computational Complexity*, 6:217–234, 1996.
[13] P. Pudlak. Communication in bounded depth circuits. *Combinatorica*, 14(2):203–216, 1994.
[14] J. Radhakrishnan, P. Sen, and S. Vishwanathan. Depth-3 arithmetic circuits for $s_n^2(x)$ and extensions of the graham-pollack theorem. To appear, 2000.
[15] R. Raz and A. Shpilka. Lower bounds for matrix product, in bounded depth circuits with arbitrary gates. In *ECCC*, 2000. TR029.
[16] A. A. Razborov. Lower bounds for the size of circuits with bounded depth with basis $\{\wedge, \oplus\}$. *Mat. Zametki*, 1987.
[17] B. E. Sagan. *Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Chapman & Hall, 1991.
[18] M. Saks. Private communication.
[19] A. Shpilka and A. Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. In *CCC*, volume 14, pages 87–96, 1999.

[20] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *STOC*, pages 77–82, 1987.
[21] V. Strassen. Die berechnungskomplexiät von elementarsymmetrischen funktionen und von interpolationskoeffizienten. *Numer. Math*, 20:238–251, 1973.
[22] L. G. Valiant. Completeness classes in algebra. In *STOC*, pages 249–261, 1979.
[23] L. G. Valiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8:189–201, 1979.
[24] L. G. Valiant. Reducibility by algebraic projections. *Logic and Algorithms. L'ensignment Mathmatique*, 30:365–380, 1982.
[25] A. C. Yao. Separating the polynomial hierarchy by oracles. In *FOCS*, pages 1–10, 1985.
[26] A. C. Yao. On acc and threshold circuits. In *FOCS*, pages 619–627, 1990.