Lower Bounds for Polynomial Evaluation and Interpolation Problems*

Victor Shoup
Roman Smolensky

Abstract

We show that there is a set of points $p_1, p_2, \ldots, p_n$ such that any arithmetic circuit of depth d for polynomial evaluation (or interpolation) at these points has size

$$\Omega\left(\frac{n \log n}{\log(2 + d/\log n)}\right).$$

Moreover, for circuits of sub-logarithmic depth, we obtain a lower bound of $\Omega(dn^{1+1/d})$ on its size.

1 Introduction

To prove a superlinear lower bound for a natural problem is one of the greatest challenges of theoretical computer science. Algebraic complexity theory is the study of a restricted class of algorithms that can perform arithmetic operations on data (i.e., add, subtract, multiply and divide), but that do not care how the data is represented. This is a reasonable class of algorithms to consider when solving algebraic problems. We shall use *arithmetic circuits* as our model of computation. Two very natural measures of the complexity of such a circuit are its *size* (number of gates and wires) and *depth* (length of longest path from input to output). We formally define these notions in the next section.

*Appeared in *Proc. 31st Annual Symposium on Foundations of Computer Science* (**FOCS**), pp. 378–383, 1991; this work was done while both authors were postdoctoral fellows in the Dept. of Computer Science at the University of Toronto; first author's current address: IBM Zürich Research Laboratory, Säumerstrasse 4, 8803 Rüschlikon, Switzerland, sho@zurich.ibm.com.

Even though restricting the model of computation to arithmetic circuits allows one to obtain lower bounds that are not currently obtainable in a more general model, proving lower bounds on the algebraic complexity of natural problems is still extremely difficult.

In this paper, we consider the problem of evaluating a polynomial at a fixed set of points. This problem has a very simple definition: the input is the list of coefficients of a polynomial of degree $n-1$ over the complex numbers $\mathbf{C}$; the output is the list of values of this polynomial at n points $p_1, p_2, \ldots, p_n \in \mathbf{C}$. It is important to note that we *do not* view these points as part of the input. This problem is the same as computing the linear transformation on $\mathbf{C}^n$ defined by the Vandermonde matrix $V = (p_i^{j-1})$.

Multi-point evaluation is of central importance in algebraic computation, and has been well-studied (see, e.g., [AHU74] or [Bmu75]). The best arithmetic circuits for this problem are of size $O(n(\log n)^2)$ and depth $O((\log n)^2)$. When the points are the n-th roots of unity, then this problem is called the Discrete Fourier Transform (DFT). When n is a power of 2, the DFT can be computed by an arithmetic circuit of size $O(n(\log n))$ and depth $O(\log n)$.

Suppose we are given an $n \times n$ matrix A with entries in $\mathbf{C}$ and ask: what is the size of the smallest arithmetic circuit that computes the linear transformation on $\mathbf{C}^n$ defined by A? It is easy to use a transcendence degree argument to show that if A contains a superlinear number of entries that are algebraically independent over $\mathbf{Q}$, the rational numbers, then any circuit that computes the linear transformation defined by A must have superlinear size.

An $n \times n$ Vandermonde matrix V contains at most n algebraically independent points, and so the above transcendence degree argument does not apply. In fact, it is not known if there exists a set of points for which polynomial evaluation at these points requires circuits of superlinear size.

Since we are computing a set of linear forms, for the purpose of proving lower bounds, we can restrict ourselves to *linear* circuits. A linear circuit is an arithmetic circuit that that uses only multiplication by constants and addition. At the expense of increasing the size and depth by constant factors, we can replace an arbitrary arithmetic circuit for computing a set of linear forms by a linear circuit [St73b]. We allow unbounded fan-in of addition gates, so that linear circuits of sub-logarithmic depth are possible.

In this paper we consider circuits of small depth. We prove that there exists a set of points $p_1, \ldots, p_n$ such that for any circuit of size ℓ and depth

d for polynomial evaluation at these points, we have

$$\ell = \Omega\left(\frac{n \log n}{\log(2 + d/\log n)}\right).$$

Thus, for circuits of depth $(\log n)^{O(1)}$, we have $\ell = \Omega(n \log n / \log\log n)$, and for circuits of depth $O(\log n)$, we have $\ell = \Omega(n \log n)$. For circuits of sub-logarithmic depth, we can obtain an even better lower bound: $\ell = \Omega(dn^{1+1/d})$.

In proving this lower bound, we show that we can take any set of points $\{p_1, \ldots, p_n\}$ that is algebraically independent over $\mathbf{Q}$. In fact, this condition is stronger than necessary; it suffices that the set $\{p_i^{e_i} : 0 \leq e_i \leq n-1\}$ is *linearly* independent over $\mathbf{Q}$. So for example, if $q_1, \ldots, q_n$ are distinct primes larger than n, and if we let p_i be a primitive q_i-th root of unity, then our lower bound holds.

Moreover, we show that we can take a any set of integer points that grow fast enough so as to "look" algebraically independent. For example

$$p_1 = 2,\ p_2 = 2^n,\ \ldots,\ p_n = 2^{n^{n-1}}.$$

These lower bounds are not entirely satisfying. A much more challenging problem is to prove a lower bound, assuming that the points are small integers, bounded, say, by $n^{O(1)}$, or even $2^{n^{O(1)}}$.

One should understand our result as follows. Any arithmetic circuit for multi-point polynomial evaluation that does not use some specific algebraic relations among the points is subject to our lower bound.

We also prove analogous results for the inverse problem of polynomial interpolation, i.e., computing the linear transformation on $\mathbf{C}^n$ defined by the matrix V^{-1}.

Related work

Valiant [Val77] also considered circuits of small depth for computing linear transformations. Valiant proved that if the matrix of a linear transformation is "rigid" then a linear circuit of small depth requires superlinear size in order to compute the transformation.

A matrix has *high rigidity* if its rank is high and remains high if the values of a small number of entries are changed. It was conjectured by Valiant that all the Vandermonde matrices have high rigidity, but this conjecture remains unproven. Moreover, it is not known if *any* Vandermonde matrix has high rigidity. Some estimations of rigidity for specific matrices were made by Razborov [Raz] and Friedman [Fri93]. However all the known bounds on the rigidity of specific matrices are still not high enough to imply complexity bounds.

We also point out that even though we do not prove any rigidity conditions, the lower bounds we obtain are similar to what one would obtain if these matrices could be shown to be highly rigid.

Other restrictions on circuits besides depth are possible. For example, Morgenstern [Mor73] showed that if the constants used in the linear circuit have absolute value at most one, then a circuit that computes the DFT has size $\Omega(n \log n)$. In fact, the best known circuits for this problem satisfy this property.

In this paper, we view the points $p_1, \ldots, p_n$ as fixed, but one can also consider the problem of multi-point polynomial evaluation in which the points are viewed as part of the input. In this situation, the functions being computed are no longer linear forms. Using a degree argument involving tools from algebraic geometry, Strassen [St73a] obtained an $\Omega(n \log n)$ lower bound for this problem on circuit size—with no depth restriction.

Our result complements Strassen's, in that it says that no amount of precomputation based on the points will yield a linear size, small depth circuit.

2 Basic Notation and Definitions

An *arithmetic circuit* P over a field K is a directed acyclic graph. Each node is labeled as one of five types: *input*, *constant*, *addition*, *multiplication*, and *division*. The edges are also labeled with constants from K called *edge weights*. The input nodes and constant nodes have in-degree 0, and the constant nodes are labeled with constants from K. The multiplication and division nodes all have in-degree 2. We allow addition nodes to have arbitrary in-degree. Some nodes are also distinguished as *output* nodes.

P computes a function in the usual way; we only remark that the interpretation given to the edge weights is that a value is multiplied by the weight on that edge before being fed to the target node of that edge.

The *size* of P is defined as the number of edges in the graph. The *depth* of P is the length of the longest path from an input to an output. It will also be useful to define the *level* of an edge as the length of the longest path from an input to the target node of that edge.

A *linear circuit* over K is a special type of arithmetic circuit in which the only allowed node types are input and addition. A linear circuit P with n input nodes and m output nodes computes the linear transformation from K^n to K^m defined by the matrix $A = (a_{ij})$, where a_{ij} is the sum over all paths from input node j to output node i of the product of the edge weights on that path. Conversely, it is known that for infinite K, if an arithmetic circuit P computes a linear transformation, then there is a linear circuit P' that computes the same function as P, and the size and depth of P' are within a constant factor of the size and depth of P [St73b].

3 Lower Bounds for Polynomial Evaluation

The motivation for our proof comes from the theory of algebraic dimension, which measures the amount of algebraic independence of a set $x_1, \ldots, x_r$ of elements in some algebra over a field in terms of the function $D(m)$ defined as the dimension of the vector space spanned by homogeneous polynomials of degree m in $x_1, \ldots, x_r$. Informally, we show that from the point of view of a small-depth circuit, the entries of a Vandermonde matrix "look" more algebraically independent than they "really" are.

Let $A = (a_{ij})$ be an $n \times n$ matrix with entries in $\mathbf{C}$. Denote by $\Gamma_A(m)$ the set of all monomials of degree m in $\{a_{ij}\}$. Let $D_A(m)$ be the dimension of the vector space over $\mathbf{Q}$ spanned by $\Gamma_A(m)$.

Lemma 1. *Suppose we have a linear circuit P over $\mathbf{C}$ of size ℓ and depth d that computes a linear transformation on $\mathbf{C}^n$, and let A be the associated matrix. Then*

$$D_A(n) \leq \binom{n+r}{n}^d,$$

where $r = \lceil \ell/d \rceil$.

Proof. Let $\Gamma = \Gamma_A(n)$ and $D = D_A(n)$.

Consider the graph defining the circuit P. For $1 \leq i \leq d$, let ℓ_i be the number of edges at level i, and let T_i be the set of edge weights that appear at level i.

Each matrix entry can be expressed as a sum of products of the form $\alpha_1 \cdots \alpha_d$, where $\alpha_i \in T_i \cup \{1\}$; we include 1 here because we might skip some levels in the graph. Each element in Γ is obtained by multiplying together

n matrix entries, and therefore can be expressed as a sum of products of the form

$$(\alpha_1^{(1)} \cdots \alpha_1^{(n)}) \cdots (\alpha_d^{(1)} \cdots \alpha_d^{(n)}),$$

where $\alpha_i^{(j)} \in T_i \cup \{1\}$. We let S be the collection of all such products.

It is clear that each element in Γ is in the $\mathbf{Z}$-span of S. Since $D \leq \mathrm{Card}(S)$, we now want to bound $\mathrm{Card}(S)$. The number of products of the form

$$\alpha_i^{(1)} \cdots \alpha_i^{(n)}$$

is no more than the number of monomials in ℓ_i variables of degree at most n, which is $\binom{n+\ell_i}{\ell_i}$. Therefore,

$$\begin{aligned} D &\leq \prod_{i=1}^{d} \binom{n+\ell_i}{n} \\ &= \prod_{i=1}^{d} \prod_{j=1}^{n} \frac{n+\ell_i-j+1}{j} \\ &= \prod_{j=1}^{n} j^{-d} \prod_{i=1}^{d} (n+\ell_i-j+1). \end{aligned}$$

By the arithmetic-geometric mean inequality,

$$\prod_{i=1}^{d} (n+\ell_i-j+1) \leq (n+\ell/d-j+1)^d.$$

Therefore,

$$D \leq \prod_{j=1}^{n} \frac{(n+r-j+1)^d}{j^d} = \binom{n+r}{n}^d.$$

□

Theorem 1. *Let $p_1, \ldots, p_n$ be complex numbers, algebraically independent over $\mathbf{Q}$, with $n > 1$. Consider a linear circuit of size ℓ and depth d for polynomial evaluation at these points. There exists an absolute constant C such that*

$$\ell > C \frac{n \log n}{\log(2 + d/\log n)}.$$

Moreover, if $d \leq \log n / \log 3$, then

$$\ell > C d n^{1+1/d}.$$

Proof. Consider just those products of elements in the corresponding matrix V taking one element from each row. This is the set of products of the form

$$p_1^{e_1} p_2^{e_2} \cdots p_n^{e_n},$$

where the e_i's range over all integers between 0 and $n-1$. These products are linearly independent over $\mathbf{Q}$ since the p_i's are algebraically independent, and there are n^n of them. Combining this with Lemma 1, we have

$$n^n \leq D_V(n) \leq \binom{n+r}{n}^d,$$

where $r = \lceil \ell/d \rceil$. From Stirling's approximation, we have

$$\binom{n+r}{n} = O((1+r/n)^n (1+n/r)^r).$$

Taking logarithms, we obtain

$$n \log(1+r/n) + r \log(1+n/r) \geq \frac{n \log n}{d} + O(1).$$

We prove the second assertion of the theorem first. Assume that $d \leq \log n / \log 3$. Since $\log(1+x) < x$ for all $x > 0$, $r \log(1+n/r) < n$. Therefore,

$$\begin{aligned} n \log(1+r/n) + n &\geq \frac{n \log n}{d} + O(1), \\ \log(1+r/n) &\geq \frac{\log n}{d} - 1 + O(1/n), \\ \frac{r}{n} &\geq n^{1/d} e^{-1+O(1/n)} - 1. \end{aligned}$$

Now, since we are assuming that $n^{1/d} \geq 3$, for all sufficiently large n, we have

$$\begin{aligned} \frac{r}{n} &= \Omega(n^{1/d}), \\ \ell &= \Omega(d n^{1+1/d}). \end{aligned}$$

This proves the second assertion of the theorem.

To prove the first assertion, we can assume that $\log n / \log 3 < d < \sqrt{n}$, since, if $d \geq \sqrt{n}$, the lower bound $\Omega(n)$ is trivial, and if $d \leq \log n / \log 3$, the bound in the second assertion implies the bound in the first assertion.

Moreover, we can assume that $n/r > e - 1$, since otherwise $\ell = \Omega(n \log n)$. Then we have

$$\begin{aligned} n\log(1 + r/n) + r\log(1 + n/r) &\leq r + r\log(1 + n/r), \\ &\leq 2r\log(1 + n/r). \end{aligned}$$

Therefore,

$$2r\log(1 + n/r) \geq \frac{n \log n}{d} + O(1).$$

Put $t = n/r$. Then

$$\frac{2\log(1+t)}{t} \geq \frac{\log n}{d} + O(1/n),$$

and so

$$t/\log t = O(d/\log n).$$

This implies that

$$t = O((d/\log n)\log(2 + d/\log n)),$$

and so

$$\ell = \Omega\left(\frac{n \log n}{\log(2 + d/\log n)}\right).$$

This proves the first assertion of the theorem. □

Theorem 2. *Let* $p_1 = 2, p_2 = 2^n, \ldots, p_n = 2^{n^{n-1}}$, *with* $n > 1$. *Consider a linear circuit of size* ℓ *and depth* d *for polynomial evaluation at these points. There exists an absolute constant* C *such that*

$$\ell > C\frac{n \log n}{\log(2 + d/\log n)}.$$

Moreover, if $d \leq \log n/\log 3$, *then*

$$\ell > Cdn^{1+1/d}.$$

Proof (sketch). To prove this theorem, we replace the preceding dimension argument by a counting argument. Instead of the dimension $D_V(n)$, we consider the number $\hat{D}$ of distinct elements that can be expressed as sums of distinct elements in $\Gamma_V(n)$. Included among these are all integers between 0 and $2^{n^n} - 1$, so $\hat{D} \geq 2^{n^n}$. On the other hand, each element of $\Gamma_V(n)$ can be expressed as a sum $\sum_{\alpha \in S} c(\alpha)\alpha$, where S is as in the proof of Lemma 1, and each $c(\alpha)$ is a positive integer bounded by $2^{n^{O(1)}}$. Thus, one finds that

$$n^n \leq n^{O(1)} \binom{n+r}{n}^d,$$

where $r = \lceil \ell/d \rceil$.

The theorem is now proved using an argument similar to that used in the proof of Theorem 1. $\Box$

4 Lower Bounds for Polynomial Interpolation

To prove corresponding lower bounds for polynomial interpolation, we observe that for a Vandermonde matrix $V = (p_i^{j-1})$, $V^T V$ is the Hankel matrix

$$H = \begin{pmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2(n-1)} \end{pmatrix},$$

where s_j is the j-th power sum, $s_j = \sum_i p_i^j$. We also have

$$V^T = (V^T V)V^{-1} = HV^{-1}.$$

It then follows that

$$D_V(n) = D_{V^T}(n) \leq D_H(n) \cdot D_{V^{-1}}(n).$$

But, since H has only $2n-1$ distinct entries,

$$D_H(n) \leq \binom{3n-2}{n},$$

and so

$$D_V(n) \leq \binom{3n-2}{n} D_{V^{-1}}(n).$$

From this, one can easily prove the following analog of Theorem 1.

Theorem 3. *Let $p_1, \ldots, p_n$ be complex numbers, algebraically independent over $\mathbf{Q}$, with $n > 1$. Consider a linear circuit of size ℓ and depth d for polynomial interpolation at these points. There exists an absolute constant C such that*

$$\ell > C \frac{n \log n}{\log(2 + d/\log n)}.$$

Moreover, if $d \leq \log n/\log 3$, then

$$\ell > C d n^{1+1/d}.$$

By replacing the dimension argument with a counting argument, we obtain the following analog of Theorem 2.

Theorem 4. *Let $p_1 = 2, p_2 = 2^n, \ldots, p_n = 2^{n^{n-1}}$, with $n > 1$. Consider a linear circuit of size ℓ and depth d for polynomial interpolation at these points. There exists an absolute constant C such that*

$$\ell > C \frac{n \log n}{\log(2 + d/\log n)}.$$

Moreover, if $d \leq \log n/\log 3$, then

$$\ell > C d n^{1+1/d}.$$

We remark that the idea of writing $V^T = HV^{-1}$ comes from the paper [CKY89], which uses this relation to prove an upper bound for computing the linear transformation defined by the matrix V^T.

5 Open Problems

Question 1. *Can we remove the depth restriction on our lower bounds?*

Without a depth restriction, we still know that the entries in the Vandermonde matrix V are expressible as multi-linear polynomials over $\mathbf{Q}$ in the edge weights. If the points $p_1, \ldots, p_n$ are algebraically independent, this could very well imply a superlinear lower bound on the number of edge weights that appear in a linear circuit for evaluating a polynomial at these points.

For example, the following conjecture would imply an affirmative answer to Question 1.

Conjecture. *Consider a transcendental number, say* π. *Suppose that there exist complex numbers* $c_1, \ldots, c_\ell$ *such that the powers* $\pi, \pi^2, \pi^4, \ldots, \pi^{2^{k-1}}$ *can be expressed as multi-linear polynomials over* $\mathbf{Q}$ *in the* c_i*'s. Then* $\ell = \Omega(k)$.

Our lower bounds do not exploit the structure of the matrix V, in the sense that they hold even if we rearrange the matrix entries in an arbitrary way. In particular, our techniques could never yield superlinear lower bounds for DFT, or for a matrix with 0-1 entries.

Question 2. *Can we exploit some constraints on the expressions for matrix entries in terms of the edge weights that arise from the structure of the matrix?*

An example of one such constraint is the following. We know that each matrix entry p_i^{j-1} of V can be written as

$$p_i^{j-1} = f_{ij}(c_1, \ldots, c_\ell)$$

where the c_i's are the edge weights in the circuit, and $f_{ij} \in \mathbf{Q}[X_1, \ldots, X_\ell]$. Each matrix V_k with entries in $\mathbf{Q}(X_1, \ldots, X_n)$ defined by

$$V_k = (\frac{\partial f_{ij}}{\partial X_k})$$

has rank 1.

Acknowledgments

Thanks to Russell Impagliazzo for his active participation in discussions that led to these results, and to Avi Wigderson, Yuri Rabinovich, and Igor Shparlinsky for their ideas which improved upon the results in an earlier draft of this paper. Finally, thanks to Allan Borodin and Erich Kaltofen for help and encouragement.

References

[AHU74] Aho, A. V, Hopcroft, J. E., Ullman, J. D., *The Design and Analysis of Computer Algorithms,* Addison-Wesley, 1974.

[BMu75] Borodin, A., Munro, I., *The Computational Complexity of Algebraic and Numeric problems,* American Elsevier publishing Co., 1975.

[**CKY89**] Canny, J. F., Kaltofen, E., Yagati, Y., Solving systems of non-linear polynomial equations faster, *Proc. 1989 International Symposium on Symbolic and Algebraic Computation,* pp. 121–128.

[**Fri93**] Friedman, J., Note on matrix rigidity, *Combinatorica*, 13:235–239, 1993.

[**Mor73**] Morgenstern, J., Note on a lower bound of the linear complexity of the Fast Fourier Transform, *JACM*, 20(2):305–306, 1973.

[**Raz**] Razborov, A. A., On rigid matrices, preprint, 1991.

[**St73a**] Strassen, V., Die Berechnungskomplexetat von Elementary Symmetrischen Funktionen und von Interpolationskoeffizienten, *Numerische Mathematik*, 20(3):238–251, 1973.

[**St73b**] Strassen, V., Vermeidung von Divisionen, *J. reine u. angew. Math.,* 264:182–202, 1973.

[**Val77**] Valiant, L.G., Graph-theoretic arguments in low-Level complexity, *MFCS 1977, Springer-Verlag LNCS*, pp. 162–176.