Lower Bounds for Modular Counting by Circuits with Modular Gates

David Mix Barrington[1] and Howard Straubing[2]

[1] COINS Department, University of Massachusetts
Amherst, Massachusetts
USA 01003

[2] Computer Science Department, Boston College
Chestnut Hill, Massachusetts
02167 USA

Abstract. We prove that constant depth circuits, with one layer of MOD_m gates at the inputs, followed by a fixed number of layers of MOD_p gates, where p is prime, require exponential size to compute the MOD_q function, if q is a prime that divides neither p nor q.

1 Introduction

This paper is a contribution to the complexity theory of constant-depth circuits with unbounded fan-in gates. The original work in this area was that of Furst, Saxe and Sipser ([4]) and Ajtai ([1]) establishing superpolynomial lower bounds on the size of constant-depth unbounded fan-in boolean circuit families that compute parity.

Much recent investigation has been devoted to the question of how far the power of such circuit families can be extended by permitting modular counting gates and threshold gates, as well as the usual AND, OR, and NOT gates. We are particularly interested here in the power of constant-depth circuit families built exclusively of gates that determine whether the sum of the input bits is divisible by q, where q is a fixed positive integer. Our knowledge here is still quite sketchy. We can prove fairly easily that if q is a prime power then such circuit families, regardless of size, cannot compute the AND function, but we know very little about the power of polynomial-size families of such circuits when q has at least two distinct prime divisors.

We believe that the class of languages recognized by such circuit families is strictly smaller than NC^1. We have conjectured ([2], [8]) that such circuit families can neither compute the AND function, nor count modulo p, where p is a prime that does not divide q. This conjecture has a number of equivalent natural-looking formulations in both semigroup theory ([6]) and logic ([8], [9]).

In earlier work done in collaboration with D. Thérien ([2]), we proved the first part of the conjecture for a subclass of such circuits. Our results apply to circuits in which there is a layer of MOD_m gates at the inputs, for some $m > 1$, followed by a fixed but arbitrary number of layers of MOD_p gates, where p is prime. We established an exponential lower bound on the size of such circuit families that compute the AND function. In the present paper, we shall prove the second part of the conjecture for the same subclass—we establish an exponential lower bound on the size of such circuits that compute the MOD_q function, where q is a prime that divides neither p nor m.

A similar result was proved by Krause and Pudlák ([5]), who also obtain lower bounds for thresholds of MOD_m gates. They use a probabilistic argument to reduce a circuit computing the MOD_q function to one in which each MOD_m gate at the input level has a relatively small fan-in, and thence derive a contradiction by algebraic means. Our method is entirely different, and, we think, of considerable independent interest. We explicitly compute the Fourier coefficients of the MOD_q function over a finite field and derive our circuit lower bounds by counting the number of nonzero coefficients.

In the next three sections we give precise definitions of the circuit model we work with and of the Fourier transform, and we show the connection between the two. Section 5 is devoted to the computation of the Fourier transform of the MOD_q function. In Section 6 we apply the results of this computation to obtain our circuit lower bounds.

Throughout the paper, we use some basic notions concerning finite fields. These are presented in most introductory textbooks on abstract algebra.

2 The Circuit Model

Let $k > 1$ be a positive integer. For each $n > 0$ we define a function

$$MOD_k^n : \{0,1\}^n \to \{0,1\}$$

by

$$MOD_k^n(u_1, \ldots, u_n) = \begin{cases} 1 \text{ if } k|(u_1 + \cdots + u_n) \\ 0 \text{ otherwise} \end{cases}$$

A MOD_k gate in a circuit is one that computes the MOD_k^r function of its inputs, where r is the fan-in of the gate. The language MOD_k is the set $\{w \in \{0,1\}^* : MOD_k^{|w|}(w) = 1\}$.

As is customary, we define a circuit with n inputs to be a directed acyclic graph with $2n$ source nodes (labelled $x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n$), gates at the non-source nodes, and a single sink node. The depth of the circuit is the length of the longest path from a source to the sink, and the size of the circuit is the number of nodes. Let $m > 1$ and let p be an *odd* prime. We define a *special* (m, p) *circuit* to be one in which on every path from a source node to a sink node, the first gate is either a MOD_m gate or a MOD_p gate, and every subsequent gate is a MOD_p gate. In other words, such circuits consist of a layer of MOD_m gates at the inputs, followed by arbitrarily many layers of MOD_p gates.

We also wish to define such special circuits in the case $p = 2$, but we will need a slightly different definition. If $p > 2$, we can simulate both a NOT gate and a fan-in two AND gate using MOD_p gates and constants. (To obtain $AND(x, y)$, we feed $p - 2$ copies of the constant 1 to the gate, as well as x

and y.) Furthermore, we can obtain the constant 1 by feeding p copies of any input bit to a MOD_p gate. Thus although our circuits are built exclusively with modular gates, we may assume that they contain bounded fan-in boolean gates as well. However, we cannot simulate a fan-in two AND gate using only MOD_2 gates. We therefore define a special $(m,2)$-circuit as above, except that we use MOD_4 gates in place of the MOD_2 gates. Alternatively, one can define special (m,p) circuits for any prime p as containing a layer of MOD_m gates, followed by arbitrarily many layers of of MOD_p gates and bounded fan-in boolean gates.

It is easy to show (see, for example, [8]) that for any $t > 0$ and $k > 2$, a MOD_{k^t} gate with n inputs can be simulated by a circuit of depth $O(t)$ and size $O(n^t)$ built exclusively of MOD_k gates. (If $k = 2$ we can still do the simulation, but we must build the circuit of MOD_4 gates.) In particular, special (m,p) circuits of bounded depth can count modulo any power of p.

3 The Fourier Transform

Let $r > 1$, and let F be a field whose characteristic does not divide r. Suppose F contains an element ω that is a primitive r^{th} root of unity. That is, $\omega^r = 1$, and for any $0 < s < r$, $\omega^s \neq 1$. Thus $\Omega = \{\omega^k : 0 \leq k < r\}$ contains exactly r elements.

The set V of functions from Ω^n to F forms a vector space of dimension r^n over F. We will consider a special basis for this vector space. Let

$$\mathbf{x} = (x_1, \ldots, x_n), \mathbf{w} = (w_1, \ldots, w_n) \in \Omega^n.$$

For $v \in \Omega$, let $\log v$ be the unique element of $\{0, 1, \ldots, r-1\}$ such that $\omega^{\log v} = v$. We define $P_{\mathbf{w}} \in V$ by

$$P_{\mathbf{w}}(\mathbf{x}) = w_1^{\log x_1} \cdots w_n^{\log x_n}.$$

Let us stress that these functions are defined relative to a particular choice of the field F and primitive r^{th} root ω.

We wish to show that the functions $P_{\mathbf{w}}$ form a basis for V. To do this, we first define a bilinear map

$$<,>: V \times V \to F$$

by

$$< f_1, f_2 >= \sum_{\mathbf{x} \in \Omega^n} f_1(\mathbf{x}) f_2(\mathbf{x}^{-1}),$$

where $\mathbf{x}^{-1}$ denotes $(x_1^{-1}, \ldots, x_n^{-1})$ when $\mathbf{x} = (x_1, \ldots, x_n)$.

Now observe that if $\mathbf{v} \neq \mathbf{w}$, then for some i, $u = v_i w_i \neq 1$. Thus there exists $z \in F$ such that

$$< P_{\mathbf{v}}, P_{\mathbf{w}} >= z \sum_{x \in \Omega} u^{\log x} = z(u^r - 1)(u-1)^{-1} = 0,$$

since $u^r = 1$. On the other hand, if $\mathbf{v} = \mathbf{w}$, then

$$< P_{\mathbf{v}}, P_{\mathbf{w}} >= \sum_{\mathbf{x} \in \Omega^n} 1 = r^n \cdot 1 \neq 0,$$

since the the characteristic of F does not divide r. These orthogonality relations imply immediately that the $P_{\mathbf{v}}$ are linearly independent. Since there are $r^n = dimV$ such functions, they form a basis for V. The orthogonality relations further imply that when $f \in V$ is expressed in terms of this basis, that is, when

$$f = \sum_{\mathbf{w} \in \Omega^n} c_{\mathbf{w}} P_{\mathbf{w}},$$

then

$$c_{\mathbf{w}} = \frac{1}{r^n} \cdot < f, P_{\mathbf{w}} > .$$

If $f \in V$, then the function $Tf \in V$ defined by

$$(Tf)(\mathbf{w}) =< f, P_{\mathbf{w}} >= \sum_{\mathbf{x} \in \Omega^n} f(x) P_{\mathbf{w}}(\mathbf{x}^{-1})$$

is called the *Fourier transform* of f. The value of the Fourier transform of f evaluated at $\mathbf{w}$ is thus, up to a constant non-zero multiple, the coefficient of $P_{\mathbf{w}}$ when f is written in terms of this basis.

If $\mathbf{w} \in \Omega^n$, we define a function

$$Q_{\mathbf{w}} : \{0,1\}^n \to F$$

by

$$Q_{\mathbf{w}}(u_1, \ldots, u_n) = w_1^{u_1} \cdots w_n^{u_n}.$$

In essence, we are identifying $\{0,1\}$ with $\{1, \omega\}$ and considering the restriction of $P_{\mathbf{w}}$ to $\{0,1\}^n$. The functions $Q_{\mathbf{w}}$ span the vector space of functions from $\{0,1\}^n$ into F, but do not form a basis for it. Thus a function in this space will have many different representations as an F-linear combination of the $Q_{\mathbf{w}}$.

For $\mathbf{v}, \mathbf{w} \in \Omega^n$ as above, we denote by $\mathbf{v} \cdot \mathbf{w}$ the componentwise product $(v_1 w_1, \ldots, v_n w_n)$. Obviously $P_{\mathbf{v} \cdot \mathbf{w}}$ is equal to the pointwise product $P_{\mathbf{v}} P_{\mathbf{w}}$. Similarly, $Q_{\mathbf{v} \cdot \mathbf{w}} = Q_{\mathbf{v}} Q_{\mathbf{w}}$.

4 Algebraic Representation of Circuit Behavior

Let us suppose that $\theta : \{0,1\}^n \to \{0,1\}$ is computed by a special (m,p) circuit of size s and depth d. We will suppose in this section that p does not divide m. Let r be a multiple of m such that p does not divide r. (In the applications we will choose $r = m$ or $r = 2m$.) For every $k > 0$, there is a finite field of characteristic p with p^k elements. Since p is a unit in the ring of integers mod r, we may choose k so that $r | p^k - 1$. $F^* = F \backslash \{0\}$ is a cyclic group of cardinality $p^k - 1$. Thus if we let g be a generator of F^*, and set $\omega = g^t$, where $t = \frac{|F-1|}{r}$, then ω is a primitive r^{th} root of unity.

We can thus define the functions $Q_{\mathbf{w}}$ and $P_{\mathbf{w}}$ with respect to F and ω. Since $\{0,1\} \subseteq F$, we can view θ as a function from $\{0,1\}^n$ into F. We will show how to use the circuit to express θ as a linear combination of the $Q_{\mathbf{w}}$. As we construct the expression, we will be concerned with bounding the number of distinct $Q_{\mathbf{w}}$ that occur.

We construct the expression by induction on the depth of the circuit. If $d = 0$, then $\theta(u_1, \ldots, u_n) = u_i$ or $\theta(u_1, \ldots, u_n) = 1 - u_i$. Observe that

$$u_i = (\omega^{u_i} - 1)/(\omega - 1).$$

Since both ω^{u_i} and the constant function 1 are among the $Q_{\mathbf{w}}$, we obtain expressions for θ with no more than two terms. Suppose now $d = 1$ and the output gate is a MOD_m gate. Then

$$\theta(u_1, \ldots, u_n) = MOD_m(v_1, \ldots, v_n),$$

where for each i, either $v_i = u_i$ or $v_i = 1 - u_i$. Since $z^{|F|-1} = 1$ for all $z \in F^*$,

$$\theta(u_1, \ldots, u_n) = 1 - (\omega^{\frac{r}{m}(v_1 + \cdots + v_n)} - 1)^{|F|-1}.$$

Now $\omega^{\frac{r}{m}(v_1 + \cdots + v_n)}$ is one of the functions $Q_{\mathbf{w}}$, and, as we noted in the last section, the product of two of the $Q_{\mathbf{w}}$ is another one of these functions. Thus we can expand the right-hand side of the above equation and express it as a linear combination of $|F|$ of the $Q_{\mathbf{w}}$. Finally, suppose $d > 0$ and the output gate is a MOD_p gate. In the case where p is odd, we have

$$\theta(u_1, \ldots, u_n) = 1 - \left(\sum_{i=1}^{R} g_i(u_1, \ldots, u_n) \right)^{|F|-1},$$

where $R \leq s$, and each g_i is computed by a special (m, p) circuit of depth $d - 1$. If we express each g_i as a linear combination of no more than M of the $Q_{\mathbf{w}}$, we obtain an expression for θ as a linear combination of no more than $s^{|F|-1} M^{|F|-1}$ of the $Q_{\mathbf{w}}$. If $p = 2$, the output gate in the inductive step is a MOD_4 gate. We can replace the MOD_4 gate by a circuit with a layer of MOD_2 gates at the input, a layer of NOT gates, a layer of fan-in two AND gates, and a layer of MOD_2 gates. Observe that

$$NOT(g) = 1 - g,$$

and

$$AND(g_1, g_2) = g_1 \cdot g_2.$$

In particular, if g_1 and g_2 are both expressed as linear combinations of no more than K of the $Q_{\mathbf{w}}$, $NOT(g_1)$ can be expressed as a linear combination of not more than $K + 1$ of the $Q_{\mathbf{w}}$, and $AND(g_1, g_2)$ as a linear combination of not more than K^2 of the $Q_{\mathbf{w}}$.

It follows that in either case, θ can be written as a linear combination of the $Q_{\mathbf{w}}$ in which the number of terms is bounded by a polynomial in s. The polynomial depends only on d and $|F|$, and $|F|$ in turn depends only on p and r.

The remarks in this section establish the following result on the algebraic representation of special (m, p) circuits.

Lemma 1. *Let $d > 0$, $m > 1$, r a multiple of m, and p a prime that does not divide r. There exist a polynomial H, which depends only on d, p and r, and a finite field F of characteristic p with a primitive r^{th} root of unity ω, such that the following property holds: If $\theta : \{0,1\}^n \to \{0,1\}$ is computed by a special (m,p) circuit of size s and depth d, then*

$$\theta = \sum_{\mathbf{w} \in D} c_{\mathbf{w}} Q_{\mathbf{w}},$$

where each $c_{\mathbf{w}}$ belongs to F, and where $D \subseteq \Omega^n$ has no more than $H(s)$ elements.

A different version of this lemma is proved in [2]. In that paper we do not deal with circuits directly, and instead consider programs over finite solvable groups. We showed that if the solvable group has a special form—namely, if it is an extension of a p-group by an abelian group—then the boolean function computed by the program can be expressed as a linear combination of the $Q_{\mathbf{w}}$, where the size of the expression depends polynomially on the size of the program. In [8] it is proved that programs over solvable groups capture precisely the power of bounded-depth circuits with modular counting gates of a fixed modulus. In the construction of [8], special (m,p) circuits correspond to the solvable groups of the form described above. In the present paper we do not really need the program model, so we have chosen to do everything in terms of circuits.

5 The Fourier Transform of the MOD_q Function

Now let us suppose that p and q are distinct primes, $r \geq 2$, and that neither p nor q divides r. Let F be a field of characteristic p with a primitive r^{th} root of unity ω. We define

$$\theta_{q,n} : \{1, \omega, \ldots, \omega^{r-1}\}^n \to F$$

by

$$\theta_{q,n}(\mathbf{x}) = \begin{cases} MOD_q^n(\log x_1, \ldots, \log x_n) & \text{if } \mathbf{x} \in \{1,\omega\}^n \\ 0 & \text{otherwise} \end{cases},$$

where $\mathbf{x} = (x_1, \ldots, x_n)$.

In this section we will compute the Fourier coefficients of $\theta_{q,n}$. Our goal is to show that exponentially many of the coefficients are nonzero. This will be used in the next section to show that MOD_q^n cannot be written as a subexponential-size linear combination of the $Q_{\mathbf{w}}$.

Let $\mathbf{w} = (\omega^{c_1}, \ldots, \omega^{c_n})$.

$$\begin{aligned} (T\theta_{q,n})(\mathbf{w}) &= \sum_{\mathbf{x} \in \Omega^n} \theta_{q,n}(\mathbf{x}) \omega^{-(c_1 \log x_1 + \cdots + c_n \log x_n)} \\ &= \sum_{\substack{A \subseteq \{0,1\}^n \\ q \mid |A|}} \omega^{-\sum_{i \in A} c_i} \end{aligned}$$

$$= \sum_{j\geq 0} \sum_{\substack{A\subseteq\{0,1\}^n \\ |A|=jq}} \omega^{-\sum_{i\in A} c_i}.$$

Observe that the inner summation in the last line is the coefficient of y^{jq} in the polynomial

$$(1+\omega^{-c_1}y)\cdots(1+\omega^{-c_n}y),$$

so that $(T\theta_{q,n})(\mathbf{w})$ is the sum of the coefficients of degree divisible by q in this polynomial. Now suppose that n is a multiple of r, so that $n = rs$, and choose $\mathbf{w}$ so that for each $k \in \{0,\ldots,r-1\}$, exactly s of the c_i are equal to k. The number of such $\mathbf{w}$ is the multinomial coefficient

$$\frac{n!}{(s!)^r}.$$

By Stirling's formula for factorials, this is at least $\frac{r^n}{Q(n)}$, where Q is a polynomial. In this case $(T\theta_{q,n})(\mathbf{w})$ is the sum of the coefficients of degree divisible by q in the polynomial

$$\left[(1+y)(1+\omega y)\cdots(1+\omega^{r-1}y)\right]^s.$$

The coefficient of y^k in $\prod_{i=0}^{r-1}(1+\omega^i y)$ is the k^{th} elementary symmetric polynomial in $1,\omega,\ldots,\omega^{r-1}$, which is $(-1)^k$ times the coefficient of y^{n-k} in $\prod_{i=0}^{r-1}(y-\omega^i) = y^r - 1$. Thus,

$$\left[(1+y)(1+\omega y)\cdots(1+\omega^{r-1}y)\right]^s = (1+(-1)^{r-1}y^r)^s.$$

Since r is relatively prime to q, rt is divisible by q if and only if t is divisible by q. This gives

$$(T\theta_{q,n})(\mathbf{w}) = \begin{cases} \binom{s}{0} - \binom{s}{q} + \binom{s}{2q} - \cdots & \text{if } r \text{ is even and } q \text{ is odd} \\ \binom{s}{0} + \binom{s}{q} + \binom{s}{2q} + \cdots & \text{otherwise} \end{cases}$$

Let us denote the alternating sum on the right-hand side of the above equation by $K_{s,q}$, and the positive sum by $L_{s,q}$. We stress that the two sides of these equations are elements of F. Since F has characteristic p, the binomial coefficients on the right-hand side are to be taken modulo p. Thus $K_{(s,q)}$ and $L_{(s,q)}$ should be thought of as elements of the ring $\mathbf{Z}_p$ of integers modulo p.

Lemma 2. *$K_{s,q}$ and $L_{s,q}$ are eventually periodic in s. That is, there exist $s_0, s_1, s_2 >$ 0 such that for all $s \geq s_0$, $K_{s,q} = K_{s+s_1,q}$, and $L_{s,q} = L_{s+s_1,q}$.*

Proof. For $0 \leq j < q$, let $K_{s,q}^{(j)}$ denote the sum of the coefficients of degree congruent to j modulo q in the polynomial $(1-y)^s$. Thus $K_{s,q} = K_{s,q}^{(0)}$. Since $(1-y)^{s+1} = (1-y)^s(1-y)$, we have

$$K_{s+1,q}^{(j)} = K_{s,q}^{(j)} - K_{s,q}^{(j-1)}.$$

(We compute the difference $j-1$ modulo q, so that $0-1=q-1$.) Thus the vectors

$$(K_{s,q}^{(0)},\ldots,K_{s,q}^{(q-1)})$$

are obtained by repeatedly applying a q-dimensional linear transformation $\mathcal{T}$ over the integers modulo p to the vector

$$(K_{0,q}^{(0)},\ldots,K_{0,q}^{(q-1)})=(1,0,\ldots,0).$$

Since the space $\mathbf{Z}_p^q$ is finite, we will have $\mathcal{T}^{s_0+s_1}=\mathcal{T}^{s_0}$ for some $s_0,s_1>0$, which gives the desired result in the first case. The argument for $L_{s,q}$ is essentially identical; we simply consider the polynomial $(1+y)^s$ in place of $(1-y)^s$.

Lemma 3. *$K_{s,q}$ and $L_{s,q}$ are nonzero for infinitely many values of s.*

Proof. Let us first recall a few facts about arithmetic in $\mathbf{Z}_p$: If $0<k<p$, the binomial coefficient $\binom{p}{k}$ is zero modulo p, and thus for any $a,b\in\mathbf{Z}_p$, $(a+b)^p=a^p+b^p$. The nonzero elements of $\mathbf{Z}_p$ form a cyclic group of order $p-1$ under multiplication, and thus for all elements a of $\mathbf{Z}_p$, $a^p=a$. We thus have

$$(1+y)^{ps}=((1+y)^s)^p=\left(\sum_{j=0}^{s}\binom{s}{j}y^j\right)^p=\sum_{j=0}^{s}\binom{s}{j}y^{jp}.$$

Equating the coefficients of two ends of this equation, we find $\binom{ps}{j}=0$, if p does not divide j, and $\binom{ps}{pt}=\binom{s}{t}^p=\binom{s}{t}$.

Thus,

$$\begin{aligned}
K_{ps,q}&=\binom{ps}{0}-\binom{ps}{q}+\binom{ps}{2q}-\cdots\\
&=\binom{ps}{0}+(-1)^p\binom{ps}{pq}+(-1)^{2p}\binom{ps}{2pq}+\cdots\\
&=\binom{s}{0}+(-1)^p\binom{s}{q}+(-1)^{2p}\binom{s}{2q}+\cdots\\
&=\binom{s}{0}-\binom{s}{q}+\binom{s}{2q}-\cdots\\
&=K_{s,q}.
\end{aligned}$$

The next-to-last equality follows because either p is odd, in which case $(-1)^p=-1$, or $p=2$, in which case $-1=1$. Thus

$$1=K_{1,q}=K_{p,q}=K_{p^2,q}=\cdots.$$

An analogous argument establishes the result for $L_{s,q}$.

Lemma 4. *There is a polynomial M such that*

$$|\{\mathbf{w}:(T\theta_{q,n})(\mathbf{w})\neq 0\}|\geq\frac{r^n}{M(n)},$$

for all sufficiently large n.

Proof. First suppose that r is even and q is odd. By Lemma 2 for sufficiently large values of s, $K_{s,q}$ is periodic with period s_1, and thus by Lemma 3, every interval of s_1 successive values of $K_{s,q}$ contains a nonzero element. Thus, there is a nonnegative integer $a < rs_1$ such that $n-a$ is divisible by r, and, if $t = \frac{n-a}{r}$, $K_{t,q}$ is nonzero.

It follows from the above remarks that there is a polynomial Q such that $T\theta_{q,n-a}$ has at least $\frac{r^{n-a}}{Q(n-a)}$ nonzero values. Let $\mathbf{w} = (w_1, \ldots, w_n) \in \Omega^{n-a}$, and let $\mathbf{w}' \in \Omega^n$ denote the vector obtained by padding $\mathbf{w}$ with a 1's. Then

$$\begin{aligned}(T\theta_{q,n})(\mathbf{w}') &= r^a \sum_{\mathbf{x} \in \Omega^{n-a}} \theta_{q,n-a}(\mathbf{x}) w_1^{-\log x_1} \cdots w_{n-a}^{-\log x_{n-a}} \\ &= r^a (T\theta_{q,n-a})(\mathbf{w}).\end{aligned}$$

Since $r^a \neq 0$, this shows that $T\theta_{q,n}$ has at least $\frac{r^{n-a}}{Q(n-a)}$ nonzero values. We can thus choose $M(n)$ to be a polynomial greater that r^{rs_1} times the maximum of $Q(n)$, $Q(n-1)$,...,$Q(n-rs_1)$ to obtain the result.

In the case where r is odd or $q = 2$, we argue identically, using the corresponding results for the sums $L_{s,q}$.

6 The Circuit Lower Bounds

Theorem 5. *Let $m \geq 2$, $d > 0$, and let p and q be distinct primes, where q does not divide m. There is a constant $c > 1$, depending on m, p, q, and d, such that any special (m,p) circuit computing the MOD_q^n function has size at least c^n.*

Proof. Of course, we need only show this for sufficiently large values of n, since we, can, if necessary, adjust the value of c to account for a finite number of exceptions. We can also assume that p does not divide m : If p divides m, then we can write $m = p^t m'$, where p does not divide m'. Since m' and p are relatively prime, we can simulate a MOD_m gate by the AND of a MOD_{p^t} gate and a $MOD_{m'}$ gate. As remarked earlier, we can simulate the MOD_{p^t} and the fan-in 2 AND gate by a fixed number of levels of MOD_p gates (or MOD_4 gates in case $p = 2$). All this entails at worst a polynomial blowup in the size of the circuit, so an exponential lower bound obtained for the size of the special (m', p) circuit obtained in this manner gives the desired exponential lower bound for the original special (m,p) circuit.

We first consider the case where p and q are both odd. Let $r = 2m$. Thus p does not divide r, so we may choose a field F as in Lemma 1. Let $\mathcal{C}$ be a special (m,p) circuit of depth d that computes MOD_q^n. Let us denote by $|\mathcal{C}|$ the size of $\mathcal{C}$. By Lemma 1, there is a polynomial H, depending only on m, p, and d, such that MOD_q^n is an F-linear combination of $H(|\mathcal{C}|)$ of the functions $Q_{\mathbf{w}}$. That is,

$$MOD_q^n = \sum_{\mathbf{w} \in D} c_{\mathbf{w}} Q_{\mathbf{w}},$$

where $|D| \le H(|\mathcal{C}|)$. Now consider the function

$$\beta = \sum_{\mathbf{w}\in D} c_{\mathbf{w}} P_{\mathbf{w}}.$$

Observe that β and $\theta_{q,n}$ are equal on elements of $\{1,\omega\}^n$. Let $\gamma : \Omega^n \to F$ be defined by

$$\gamma(\mathbf{v}) = \begin{cases} 1 \text{ if } \mathbf{v} \in \{1,\omega\}^n \\ 0 \text{ otherwise} \end{cases}$$

We have

$$\begin{aligned}(T\gamma)(\mathbf{w}) &= \sum_{\mathbf{x}\in\{1,\omega\}^n} P_{\mathbf{w}}(\mathbf{x}^{-1}) \\ &= \sum_{A\subseteq\{0,1\}^n} \prod_{i\in A} w_i^{-1} \\ &= \prod_{i=1}^{n}(1+w_i^{-1}).\end{aligned}$$

Since $r = 2m$, $\omega^m = -1$. Thus $T\gamma$ is nonzero at exactly $(r-1)^n$ elements of Ω^n. We can consequently write $\gamma = \sum_{\mathbf{w}\in E} d_{\mathbf{w}} P_{\mathbf{w}}$, where $|E| = (r-1)^n$. We now have

$$\theta_{q,n} = \beta\gamma = \sum_{(\mathbf{v},\mathbf{w})\in D\times E} c_{\mathbf{v}} d_{\mathbf{w}} P_{\mathbf{v}\cdot\mathbf{w}}.$$

Thus $T\theta_{q,n}$ is nonzero at at most $H(|\mathcal{C}|)\cdot(r-1)^n$ elements of Ω^n. By Lemma 4, there is a polynomial M such that

$$H(|\mathcal{C}|) \ge \left(\frac{r}{r-1}\right)^n / M(n)$$

for sufficiently large n. Since $H(|\mathcal{C}|)) \le \mathcal{C}^f$for some positive integer f, we have

$$\mathcal{C} > \left\{ \left(\frac{r}{r-1}\right)^{\frac{1}{2f}} \right\}^n,$$

for sufficiently large n.

In the case where $p = 2$, we choose $r = m$. We still have $-1 \in \Omega$, because $1 = -1$, so we may argue precisely as above.

In the case where $q = 2$, we must again choose $r = m$ in order to insure that q does not divide r. In this case -1 is not in Ω, so we must change the argument slightly. Let $\delta : \Omega^n \to F$ be the function defined by

$$\delta(\mathbf{v}) = \begin{cases} 1 & \text{if } \mathbf{v} \in \{1,\omega\}^n \text{ contains an even number of 1's} \\ -1 & \text{if } \mathbf{v} \in \{1,\omega\}^n \text{ contains an odd number of 1's} \\ 0 & \text{if } \mathbf{v} \notin \{1,\omega\}^n \end{cases}$$

Now $\beta\delta = \theta_{2,n}$, and

$$(T\delta)(\mathbf{w}) = \sum_{A\subseteq\{0,1\}^n} (-1)^{|A|} \prod_{i\in A} w_i^{-1} = \prod_{i=1}^{n}(1-w_i^{-1}).$$

Thus $T\delta$ is nonzero at exactly $(r-1)^n$ elements of Ω^n, and we may argue as in the first case.

7 Related Problems

It is quite easy to show that constant-depth circuit families built entirely from MOD_p gates, where p is prime, cannot compute the AND function, regardless of size: We can represent the function $g : \{0,1\}^n \to \{0,1\}$ computed by such a circuit as a polynomial over $\mathbf{Z}_p$ in the variables $x_1, \ldots, x_n$. Observe that if $g_1, \ldots, g_r$ are such polynomials, then

$$MOD_p^r(g_1, \ldots, g_r) = 1 - (g_1 + \cdots + g_r)^{p-1},$$

which is a polynomial of degree no more than $p-1$ times the maximum of the degrees of the g_i. Thus a circuit of depth d is represented by a polynomial of degree no more than $(p-1)^d$, that is, of constant degree, independent of n. On the other hand, every function from $\{0,1\}^n$ into $\mathbf{Z}_p$ has a unique representation as a linear combination of the monomials $x_{i_1} \cdots x_{i_t}$, $1 \le i_1 < i_2 < \cdots i_t \le n$. Since the unique representation of the AND function in this form is the monomial $x_1 \cdots x_n$, which has degree n, we obtain the result. See [7] for a discussion of the representation of circuit behavior in this form. In the language of programs over groups, this result says that the AND function cannot be computed by any family of programs over a p-group. In [2] this is extended to nilpotent groups.

In [8] it is shown that if p and q are distinct primes, then we cannot recognize MOD_q by any constant-depth family of circuits built from MOD_p gates alone. The proof there uses a group-theoretic argument, and appeals to a result in [2]. Here we use our results on the Fourier representation of the MOD_q function to obtain a direct proof of this fact.

Theorem 6. *Let p and q be distinct primes, and let $d > 0$. If n is sufficiently large, then no circuit of depth d built entirely from MOD_p gates can compute the function MOD_q^n.*

Proof. If such a circuit exists, then MOD_q^n is represented as a polynomial of degree $t = (p-1)^d$ over $\mathbf{Z}^p$. Let us choose a finite field F of characteristic p such that q does not divide $r = |F^*|$. Then F contains a primitive r^{th} root of unity ω. As before, we observe that

$$x_i = (\omega^{x_i} - 1)/(\omega - 1),$$

and thus every monomial of degree t is a linear combination of no more that 2^t of the functions $Q_{\mathbf{w}}$. Since there are $O(n^t)$ monomials of degree no more than

t, we obtain a representation of MOD_q^n as a linear combination of $O(n^t)$ of the $Q_{\mathbf{w}}$. But in the proof of Theorem 5, we showed that any such representation requires exponentially many terms, a contradiction.

Krause and Pudlák also obtain lower bounds for the size of circuits consisting of a layer of MOD_m gates, followed by a single threshold gate: such circuits cannot compute MOD_q^n, where q is a prime that does not divide n, in subexponential size. We believe that the Fourier methods we use in the present paper can be adapted to obtain similar results; in this case we want to choose the primitive root ω to be an element of the field $\mathbf{C}$ of complex numbers and consider the Fourier transform over $\mathbf{C}$. Some results in this vein for the case $m = 2$ (where the primitive root is -1 and one can take the ground field to be the field of real numbers) were obtained by Bruck [3].

8 Bibliography

References

1. M. Ajtai, "Σ_1^1 formulae on finite structures", *Annals of Pure and Applied Logic* **24** (1983) 1–48.
2. D. Mix Barrington, H. Straubing, and D. Thérien, "Nonuniform Automata over Groups", *Information and Computation* **89** (1990) 109–132.
3. J. Bruck, "Harmonic Analysis of Polynomial Threshold Functions", *SIAM Journal of Discrete Mathematics* **3** (1990) 168-177.
4. M. Furst, J. Saxe, and M. Sipser, "Parity, Circuits, and the Polynomial Time Hierarchy", *J. Math Systems Theory* **17** (1984) 13–27.
5. M. Krause and P. Pudlák, "On the Computational Power of Depth 2 Circuits with Threshold and Modulo Gates", *Proc. 26th ACM STOC* (1994) 48-57.
6. P. McKenzie, P. Péladeau, and D. Thérien, NC^1 : The Automata-Theoretic Approach", to appear in *Theoretical Computer Science.*
7. R. Smolensky, "Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity", *Proc. 19th ACM STOC* (1987) 77–82.
8. H. Straubing, "Constant-depth Periodic Circuits", *International J. Algebra and Computation* **1** (1991) 49–88.
9. H. Straubing, *Finite Automata, Formal Logic and Circuit Complexity,* Birkhäuser, Boston, 1994.