A Note on MOD p - MOD m Circuits

Howard Straubing[1] and Denis Thérien[2]

[1] Computer Science Department, Boston College
Chestnut Hill, Massachusetts, USA 02467
[2] School of Computer Science, McGill University
Montréal, Québec
Canada

1 Introduction

An outstanding problem in circuit complexity concerns the computing power of constant-depth circuit families in which the output of each gate depends on the sum, modulo m, of its input bits. It is conjectured that such circuits require exponential size to compute the AND function of the inputs and to compute the sum, modulo q, of the inputs, where q is a prime that does not divide m.

Several papers have concentrated on a special subclass of these circuits—those in which there is a single layer of MOD_m gates connected to the inputs, followed by a fixed number of layers of MOD_{p^k}-gates, where p is prime. (We may always assume that p does not divide m, for if $p|m$, then we can construct an equivalent circuit with $MOD_{\frac{m}{p}}$-gates at the inputs.) Barrington, Straubing and Thérien [2] showed that such circuits require exponential size to compute AND; Krause and Pudlák [7], and Barrington and Straubing [1], showed that such circuits require exponential size to compute MOD_q, when q is a prime different from p that does not divide m. The definitive result in this direction was found by Grolmusz and Tardös [6], who showed that the only symmetric boolean functions computed by such circuits in subexponential size have a periodic spectrum with period mp^t, where $mp^t \leq n$. (By the *spectrum* of a symmetric function $f : \{0,1\}^n \rightarrow \{0,1\}$, we mean the map $\bar{f} : \{0,1,\ldots,n\} \rightarrow \{0,1\}$ such that $\bar{f}(k) = f(1^k 0^{n-k})$.) They further showed that any such periodic function with $t = O(\log\log n)$ can be computed by quasipolynomial-size circuits, thus completely characterizing the symmetric functions computable by quasipolynomial-size circuits of this special form.

The proofs in [2] and [1] use Fourier expansions over finite fields, while those in [7] and [6] are more combinatorial and rely on probabilistic arguments. In particular, [6] employs a new method, which is a kind of modular analogue of the random restriction techniques of Furst, Saxe and Sipser [5]. In the present note we show how to use the Fourier techniques to obtain different proofs of the new results of [6], both the lower and upper bounds. In fact, our lower bounds argument is only a slight modification of the proof given in [1], and actually simplifies the argument while strengthening the result. Along the way we obtain a surprising normal form result for circuits of this kind.

2 Background

2.1 Modular circuits

We adopt here the definitions and notations of Grolmusz and Tardös [6]: If $m > 0$ and $A \subseteq \{0, 1, \ldots, m-1\}$, then a MOD_m^A gate outputs 1 if and only if the sum, modulo m, of its inputs is in A, and outputs 0 otherwise. (This is more general than the traditional definition of such a gate in circuit complexity, which usually takes $A = \{0\}$.) Let p be prime. A $((MOD_{p^k})^d, MOD_m)$-circuit consists of a single layer of MOD_m^A-gates, for various A, connected to the input bits, followed by d layers of $MOD_{p^k}^B$-gates, for various B. We allow the same input bit to appear several times as an input to a single MOD_m^A-gate; of course, the number of times need never exceed $m-1$. If the number of inputs to the circuit is n, then, as usual, the circuit computes a function from $\{0,1\}^n$ into $\{0,1\}$. We define the *size* of the circuit to be the number of gates.

2.2 Discrete Fourier Transform

For a full account of the ideas in this subsection, see [2] or [1]; here we just cite the facts that we will need in the sequel.

Let us fix a prime p and $m > 0$ such that p does not divide m. There is a finite field F of characteristic p that contains a primitive m^{th} root of unity ω. We set

$$\Omega = \{1, \omega, \omega^2, \ldots, \omega^{m-1}\},$$

and consider the F-vector space V of maps from Ω^n into F. For $v \in \Omega$, we denote by $\log v$ the unique $c \in \{0, 1, \ldots, m-1\}$ such that $v = \omega^c$. If $\mathbf{x} = (x_1, \ldots, x_n) \in \Omega^n$, then we set $\mathbf{x}^{-1} = (x_1^{-1}, \ldots, x_n^{-1})$. If $\mathbf{v} = (v_1, \ldots, v_n), \mathbf{w} = (w_1, \ldots, w_n) \in \Omega^n$ then we define

$$P_{\mathbf{v}}(\mathbf{w}) = \prod_{i=1}^{n} w_i^{\log v_i}.$$

The set $\{P_{\mathbf{v}} : v \in \Omega^n\}$ forms a basis for V. If $f \in V$, then the coefficient of $P_{\mathbf{v}}$ in the expansion of f in terms of this basis is $m^{-n}(Tf)(\mathbf{v})$, where

$$(Tf)(\mathbf{v}) = \sum_{\mathbf{x} \in \Omega^n} f(\mathbf{x}) P_{\mathbf{v}}(\mathbf{x}^{-1}).$$

Tf is called the *Fourier transform* of f, and its values (scaled by m^{-n}) are called the *Fourier coefficients* of f.

Note that $P_{\mathbf{v}} \cdot P_{\mathbf{w}} = P_{\mathbf{v} \cdot \mathbf{w}}$, where on the left-hand side of the equation we have the pointwise product in F of the two functions, and on the right-hand side the componentwise product of the two vectors.

3 Representation of Circuit Behavior

We define, for $\mathbf{v} = (v_1, \ldots, v_n) \in \Omega^n$, a function $Q_\mathbf{v} : \{0,1\}^n \to F$ by

$$Q_\mathbf{v}(x_1, \ldots, x_n) = \omega^{x_1 \log v_1 + \cdots + x_n \log v_n}.$$

The maps $Q_\mathbf{v}$ span the vector space of functions from $\{0,1\}^n$ into F, but unless $|F| = 2$, they do not form a basis for this space. We define the *weight* of an element f of this space to the smallest r such that f is a linear combination of r of the $Q_\mathbf{v}$. As with the Fourier basis, we have $Q_\mathbf{v} \cdot Q_\mathbf{w} = Q_{\mathbf{v} \cdot \mathbf{w}}$. This implies that the weight is submultiplicative; that is, the weight of the product of two functions is no more than the product of their weights.

Lemma 1. *Let $f : \{0,1\}^n \to \{0,1\}$ be computed by a $((MOD_{p^k})^d, MOD_m)$-circuit of size s, where p is prime and p does not divide m. Then there is a polynomial h depending only on m, p, k and r, such that f has weight no more than $h(s)$.*

Proof. If $\{u_1, \ldots, u_r\}$ is a subset of the n input variables (with, possibly, some repeated values among the u_i), then

$$\begin{aligned} MOD_m^B(u_1, \ldots, u_r) &= 1 - \prod_{b \in B} (1 - MOD_m^{\{b\}}(u_1, \ldots, u_r)) \\ &= 1 - \prod_{b \in B} (\omega^{u_1 + \cdots + u_r} - \omega^b)^{|F|-1} \\ &= \{Q_{(1,\ldots,1)} - \prod_{b \in B} (Q_\mathbf{v} - \omega^b Q_{(1,\ldots,1)})^{|F|-1}\}(x_1, \ldots, x_n), \end{aligned}$$

where the i^{th} component of $\mathbf{v}$ is the number of times x_i appears in $(u_1, \ldots, u_r)$. Thus each MOD_m^B-gate computes a function of the inputs of weight no more than $2^{m|F|}$. We now apply the fact that a $(MOD_{p^k})^d$-circuit of s inputs can be represented as a polynomial over $\mathbf{Z}_p$ of degree D in s variables, where D only depends on k and d. (See, for example, Beigel and Tarui [3].) As $\mathbf{Z}_p$ is a subfield of F, we can compose this polynomial with the representations of the MOD_m^B-gates to obtain a representation for the circuit. When we compose a monomial of degree D with functions of weight $2^{m|F|}$, we obtain a function of weigh no more than $2^{m|F|D}$. As there no more than s^D monomials of degree D or less, the resulting representation of the circuit has weight no more than $s^D 2^{m|F|D}$, which we take as our polynomial $h(s)$.

We shall also need a converse to Lemma 1.

Lemma 2. *Suppose a function $f : \{0,1\}^n \to \{0,1\}$ has weight $K > 0$. Then f is realized by a $(MOD_p^{\{1\}}, MOD_m^A)$-circuit of size no more than pK.*

Proof. Since F is a vector space over the subfield $\mathbf{Z}_p$, we can choose a basis for F that includes the field identity 1 as one of the basis elements. If $a \in F$, then we define the *1-component* of a with respect to this basis as the coefficient of 1 in the expansion of a as a linear combination of the basis elements. Consider now a term $c \cdot Q_{\mathbf{v}}$ in the representation of f. We have

$$Q_{\mathbf{v}}(x_1, \ldots, x_n) = \omega^{k_1 x_1 + \cdots k_n x_n}$$

for some $k_1, \ldots, k_n \in \{0, 1, \ldots, m-1\}$. We realize this term with gates $MOD_m^{B_i}$, $i = 1, \ldots, p-1$, each of them connected to k_1 copies of x_1, k_2 copies of x_2, etc., where $q \in B_i$ if and only if the 1-component of $c \cdot \omega^q$ is greater than or equal to i. Thus for any given input sequence $(x_1, \ldots, x_n)$, the number of these gates that output 1 is exactly the 1-component of $c \cdot Q_{\mathbf{v}}(x_1, \ldots, x_n)$. We now take the sum, modulo p of these $(p-1)K$ gates, which gives the 1-component of $f(x_1, \ldots, x_n)$. Since f takes values in $\{0, 1\}$, its value is completely determined by this 1-component. Observe that we used only a single $MOD_p^{\{1\}}$-gate, and $(p-1) \cdot K$ MOD_m-gates. Thus the total number of gates is $(p-1) \cdot K + 1 < pK$.

The two lemmas above have the following curious consequence: Every $((MOD_{p^k}^A)^d, MOD_m^B)$-circuit is equivalent to a $(MOD_p^{\{1\}}, MOD_m)$-circuit, with a polynomial blowup in size. (A simpler approach, using the polynomial representation of the MOD_p portion of the circuit, gives a layer of constant fan-in AND-gates between the MOD_m-layer and the MOD_p-gate; the surprising fact is that these AND-gates are unnecessary.)

4 The Fourier Coefficients of a Symmetric Function

Let $f : \{0,1\}^n \to F$ be a symmetric function. We associate to f its *spectrum* $\bar{f} : \{0, \ldots, n\} \to F$ defined by $\bar{f}(k) = f(1^k 0^{n-k})$. Observe that, since f is symmetric, $\bar{f}$ completely determines f. We will be concerned with the case where $\bar{f}$ is periodic.

We also define $\phi_f : \Omega^n \to F$ by

$$\phi_f(\omega^{c_1}, \ldots, \omega^{c_n}) = \begin{cases} f(c_1, \ldots, c_n) & \text{if } (c_1, \ldots, c_n) \in \{0,1\}^n \\ 0 & \text{otherwise} \end{cases}$$

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

$$\begin{aligned} (T\phi_f)(\mathbf{w}) &= \sum_{\mathbf{x} \in \Omega^n} \phi_f(\mathbf{x}) \omega^{-c_1 \log x_1 - \cdots - c_n \log x_n} \\ &= \sum_{j=0}^{n} \bar{f}(j) \sum_{\substack{A \subseteq \{1, \ldots, n\} \\ |A| = j}} \omega^{-\sum_{i \in A} c_i}. \end{aligned}$$

The inner summation is the coefficient of y^j in the polynomial $\prod_{i=1}^{n}(1+\omega^{-c_i}y)$. Let us restrict attention to the case where $n = mp^k$ for some $k > 0$, and where $\mathbf{w} = (\omega^{c_1}, \ldots, \omega^{c_n})$ is *balanced*, that is, each $j \in \{0, 1, \ldots, m-1\}$ appears exactly p^k times among the c_i. The number of balanced vectors is given by the multinomial coefficient

$$\frac{n!}{(p^k!)^r},$$

which, by Stirling's formula, is bounded below by $m^n/u(n)$, where u is a polynomial that depends on m. With this restriction we have

$$\begin{aligned}\prod_{i=1}^{n}(1+\omega^{-c_i y}) &= (\prod_{j=1}^{m}(1+\omega^{-j}y))^{p^k} \\ &= ((-1)^{m-1}y^m + 1)^{p^k}.\end{aligned}$$

The coefficient of y^j in this polynomial is thus 0 unless j is a multiple of m. We conclude that

$$(T\phi_f)(\mathbf{w}) = \begin{cases} \sum_{i=0}^{p^k} \bar{f}(mi)(-1)^i\binom{p^k}{i} & \text{if } m \text{ is even} \\ \sum_{i=0}^{p^k} \bar{f}(mi)\binom{p^k}{i} & \text{if } m \text{ is odd} \end{cases}$$

Observe, however, that this sum is in a field of characteristic p. Since $p|\binom{p^k}{i}$ for $1 < i < p^k$, we obtain:

Lemma 3. *Let $m, p > 0$, where p is a prime that does not divide m. If $n = mp^k$ for some $k > 0$ and $f : \{0,1\}^n \to \{0,1\}$ is symmetric, then for balanced $\mathbf{w} \in \Omega^n$,*

$$(T\phi_f)(\mathbf{w}) = \begin{cases} \bar{f}(0) - \bar{f}(mp^k) & \textit{if } m \textit{ is even} \\ \bar{f}(0) + \bar{f}(mp^k) & \textit{if } m \textit{ is odd} \end{cases}$$

5 The Circuit Lower Bounds

In this section we prove the following theorem, which first appears in [6]:

Theorem 4. *Let $m > 0$, $p > 0$, where p is a prime that does not divide m. Let $f : \{0,1\}^n \to \{0,1\}$ be a function computed by a $((MOD^A_{p^k})^d, MOD^B_m)$-circuit. Then either $\bar{f}$ is periodic of period mp^t for some t such that $mp^t \leq n$, or the size of the circuit is at least c^n, where c is a constant depending on m, p, k, and d.*

The theorem has as immediate corollaries the earlier results of Barrington, Straubing, Thérien, Krause and Pudlák cited in the Introduction.

Let t be the largest power of p such that $mp^t \leq n$. If $\bar{f}$ is not periodic of period mp^k for any $k \leq t$, then there exist $0 \leq i < j = i + mp^t \leq n$ such that $\bar{f}(i) \neq \bar{f}(j)$. Let us define a symmetric function $g : \{0,1\}^{mp^t} \to \{0,1\}$

by setting $\bar{g}(r) = \bar{f}(i+r)$. Then $\bar{g}(0) \neq \bar{g}(mp^t)$. Suppose f is computed by a $((MOD_{p^k}^A)^d, MOD_m^B)$-circuit of size s. Then g is computed by a circuit whose size is no larger than s, and thus by Lemma 1, $g = \sum_{\mathbf{v}\in D} c_{\mathbf{v}} Q_{\mathbf{v}}$, where $D \subseteq \Omega^{mp^t}$ has cardinality less than $h(s)$, where h is a polynomial depending on m, p, k and d. We define

$$\gamma = \sum_{\mathbf{v}\in D} c_{\mathbf{v}} P_{\mathbf{v}}.$$

Suppose first that m is even. Let $\alpha : \Omega^{mp^t} \rightarrow \{0,1\}$ be the characteristic function of $\{1,\omega\}^n$. Then, as in Section 4,

$$\begin{aligned}(T\alpha)(\omega^{c_1},\ldots,\omega^{c_{mp^t}}) &= \sum_{A\subseteq\{1,\ldots,mp^t\}} \omega^{-\sum_{i\in A} c_i} \\ &= \prod_{i=1}^{mp^t} (1+\omega^{-c_i}).\end{aligned}$$

Since m is even, $\omega^{m/2} = -1$, and thus $T\alpha$ is nonzero at exactly $(m-1)^{mp^t}$ elements of Ω^{mp^t}. It follows (using the fact that $P_{\mathbf{v}} \cdot P_{\mathbf{w}} = P_{\mathbf{v}\cdot\mathbf{w}}$) that the Fourier expansion of $\phi_g = \alpha\gamma$ has at most $h(s)\cdot(m-1)^{mp^t}$ nonzero terms. But since $g(0) - g(mp^t) \neq 0$, Lemma 3 implies that the expansion has at least $m^{mp^t}/u(mp^t)$ nonzero terms, and thus

$$h(s) > \left(\frac{m}{m-1}\right)^{mp^t} / u(mp^t) > \left(\frac{m}{m-1}\right)^{mp^t/2},$$

provided n is sufficiently large. Since $h(s) < s^f$ for some positive integer f, we have

$$s > \left(\frac{m}{m-1}\right)^{mp^t/2} > \left(\frac{m}{m-1}\right)^{bn}$$

for some $b > 0$, since $p^t > n/(p\log_p m)$.

Now suppose m is odd. Then Ω does not contain -1. Let $h : \{0,1\}^{mp^t} \rightarrow F$ be the symmetric function such that $\bar{h}(j) = (-1)^j$. Then

$$\begin{aligned}(T\phi_h)(\omega^{c_1},\ldots,\omega^{c_{mp^t}}) &= \sum_{A\subseteq\{1,\ldots,mp^t\}} (-1)^{|A|}\omega^{-\sum_{i\in A} c_i} \\ &= \prod_{i=1}^{mp^t} (1-\omega^{-c_i}).\end{aligned}$$

Thus $T\phi_h$ is nonzero at exactly $(m-1)^{mp^t}$ elements of Ω^{mp^t}. It follows that the Fourier expansion of $\gamma\phi_h$ has at most $(m-1)^{mp^t}$ nonzero terms. Observe, however, that $\gamma\phi_h = \phi_v$, where $v : \{0,1\}^{mp^t} \rightarrow F$ is the symmetric function defined by $\bar{v}(j) = (-1)^j\bar{g}(j)$. Thus $\bar{v}(0)+\bar{v}(mp^t) = \bar{g}(0)+(-1)^{mp^t}\bar{g}(mp^t)$, which is nonzero, since $\bar{g}(0) \neq \bar{g}(mp^t)$, and $\bar{g}$ takes values in $\{0,1\}$. So we conclude as in the even case that s is exponential in n. This completes the proof.

6 Upper Bounds

Let p be prime. As is well known, every function $f:\{0,1\}^n \to \{0,1\}$ can be represented by a polynomial over $\mathbf{Z}_p$ in n variables, where no variable appears in any monomial to a power higher than 1. (See, for example, Smolensky [8].) The variable x_i itself represents a function from $\{0,1\}^n$ into $\{0,1\}$ that is also given by the expression

$$\frac{\omega^{x_i}-1}{\omega-1},$$

which has weight 2. It follows that a monomial over $\mathbf{Z}_p$ of degree d has weight no more than 2^d, and since there are fewer than n^d monomials of degree no more than d, every boolean function represented by a polynomial over $\mathbf{Z}_p$ of degree d has weight no more than $(2n)^d$.

Now consider a boolean-valued symmetric function $f(x_1,\ldots,x_n)$ whose spectrum is periodic of period mp^t for some $t>0$, where $m>0$ is not divisible by p. Thus f is identical to $MOD^A_{mp^t}$ for some $A \subseteq \{0,1,\ldots,mp^t-1\}$. Considering these as functions with values in F, we have

$$\begin{aligned} MOD^A_{mp^t} &= \sum_{a\in A} MOD^{\{a\}}_{mp^t} \\ &= \sum_{a\in A} MOD^{\{a \bmod m\}}_{m} \cdot MOD^{\{a \bmod p^t\}}_{p^t} \end{aligned}$$

It follows from a result in Beigel and Tarui [3] that $MOD^{\{a \bmod p^t\}}_{p^t}$ is represented by a polynomial over $\mathbf{Z}_p$ of degree p^t-1. (Actually, for our purposes, any polynomial representation of degree $2^{O(d)}$ will suffice.) Thus $MOD^{\{a \bmod p^t\}}_{mp^t}$ has weight no more than $(2n)^{p^t}$. Since $MOD^{\{a \bmod m\}}_{m}$ has weight no more than $2^{|F|}$, we find that the weight of f is bounded by $2mp^t\cdot 2^{|F|}\cdot(2n)^{p^t}$. Thus by Lemma 2, f is realized by a $(MOD^{\{1\}}_p, MOD^A_m)$-circuit of size $2mp^{t+1}\cdot 2^{|F|}\cdot(2n)^{p^t}$. In particular, if $t=O(\log\log n)$, then f is realized by a circuit of size $n^{(\log n)^{O(1)}}$, that is, of *quasipolynomial* size.

Conversely, suppose we have a $((MOD_{p^k})^d, MOD_m)$ circuit of size $s = n^{(\log n)^{O(1)}}$ computing a symmetric function f. Let t be the smallest integer such that $\bar{f}$ is periodic of period mp^t. Then $\bar{f}$ is not periodic of period mp^{t-1}, so, as in the proof of Theorem 4, we can create a circuit with mp^{t-1} inputs, of size no more than s, computing $g:\{0,1\}^{mp^{t-1}} \to \{0,1\}$, where $\bar{g}(0)\neq\bar{g}(mp^{t-1})$. Thus by Theorem **??**, we have

$$n^{(\log n)^{O(1)}} > c^{mp^{t-1}}$$

for some constant c depending on m,p,r and d, which imples

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

This proves the following theorem of [6]:

Theorem 5. *A symmetric boolean function f is computed by a quasipolynomial-size $((MOD_{p^k})^d, MOD_m)$ circuit if and only if $\bar{f}$ is periodic of period mp^t, where $t = O(\log\log n)$.*

It is interesting to compare this fact to the following theorem of Fagin, Klawe, Pippenger and Stockmeyer [4]: It is possible, in AC^0, to count the number of 1's in an input string up to a threshold of t, as long as $t = (\log n)^{O(1)}$. The analogous statement for our modular circuits would be that *polynomial*-size $((MOD_{p^k})^d, MOD_m)$-circuits can count modulo p^t for $t = O(\log\log n)$. However the question of whether polynomial-size circuit families of this type can count modulo $p^{t(n)}$ for some $t(n) \to +\infty$ remains open.

7 Bibliography

References

1. 'Lower Bounds for Modular Counting by Circuits with Modular Gates', in Proceedings of the 2nd Latin American Symposium on Theoretical Computer Science, *Lecture Notes in Computer Science* **911** Springer, Berlin (1995), 60-71.
2. D. Mix Barrington, H. Straubing, and D. Thérien, "Nonuniform Automata over Groups", *Information and Computation* **89** (1990) 109-132.
3. R. Beigel and J. Tarui, "On ACC", ???
4. R. Fagin, M. Klawe, N. Pippenger, and L. Stockmeyer, "Bounded-depth, Polynomia-size Circuits for Symmetric Functions", *Theoret. Comput. Sci* **36** (1985) 239-250.
5. M. Furst, J. Saxe, and M. Sipser, "Parity, Circuits, and the Polynomial Time Hierarchy", *J. Math Systems Theory* **17** (1984) 13–27.
6. V. Grolmusz and G. Tardos, "Lower Bounds for (MOD p-MOD m) Circuits", *Proc. 39th IEEE FOCS,* 1998.
7. M. Krause and P. Pudlak, "On the Computational Power of Depth 2 Circuits with Threshold and Modular Gates", *Proc. 26th ACM STOC,* 1994.
8. R. Smolensky, "Algebraic Methods in the Theory of Lower Bounds for Boolean Circuit Complexity", *Proc. 19th ACM STOC* (1987) 77–82.