Fourier Transforms and the 2-adic Span of Periodic Binary Sequences

Mark Goresky* Andrew Klapper† Lawrence Washington‡

January 15, 2001

Abstract

An arithmetic or with-carry analog of Blahut's theorem is presented. This relates the length of the smallest feedback with carry shift register to the number of nonzero classical Fourier coefficients of a periodic binary sequence.

Keywords: Periodic binary sequence; feedback register; Fourier coefficients; Blahut's theorem; 2-adic numbers.

1 Introduction

The purpose of this paper is to develop an arithmetic analog of Blahut's theorem [1, 3], which relates the linear span of a sequence to its discrete Fourier transform. For comparison, let us recall this theorem. Let $S = a_0, a_1, \cdots$ be a periodic binary sequence with period L. The *linear span* of S, denoted $\lambda(S)$, is the length of the shortest linear recurrence satisfied by S or, equivalently, the size of the smallest linear feedback shift register that generates S. It is an important measure of the complexity of a sequence, and it is used in a number of engineering applications. For example, suppose that S is to be used as the key in a stream cipher. The Berlekamp-Massey algorithm can be used by a cryptanalyst to recover the sequence once $2\lambda(S)$ bits of S are known. Thus S is secure only if $\lambda(S)$ is large.

*Institute for Advanced Study, Princeton N.J., 08540

†763H Anderson Hall, Department of Computer Science, University of Kentucky, Lexington KY 40506-0046, klapper@cs.engr.uky.edu. Research supported in part by NSF Grant NCR-9400762.

‡Dept. of Mathematics, University of Maryland, College Park MD 20742, lcw@math.umd.edu. Research supported in part by NSA Grant MDA904-96-10036.

Let τ be a primitive L-th root of unity in some field extension F of $GF(2)$. (Such a τ exists if and only if L is odd. Various work has been done to extend Blahut's theorem to the case when L is even, [4].) The k-th discrete Fourier coefficient of S is

$$\bar{a}_k = \sum_{i=0}^{L-1} a_i \tau^{ki} \in F.$$

Blahut's theorem remarkably says that the linear span of S is equal to the number of nonzero discrete Fourier coefficients of S. It makes precise the common observation that a "complex" signal is one with many nonzero Fourier components.

Recently two of the authors have developed an approach to pseudorandom sequences which is based on *feedback with carry shift registers* (or *FCSRs*) [2]. It is in many ways analogous to the usual theory which is based on linear feedback shift registers. We include here just enough detail of the theory of FCSR's for our current purposes. An FCSR consists of a shift register and a small amount of auxiliary memory. For any periodic sequence S, the size of the smallest FCSR that generates S is called the *2-adic span* and is denoted by $\lambda_2(S)$. There is an algorithm which is an analog of the Berlekamp-Massey algorithm. Given approximately $2\lambda_2(S)$ bits of S, it outputs the smallest FCSR that generates S. FCSRs can be thought of as arithmetic or with-carry analogs of linear feedback shift registers, and are related to arithmetic codes in much the same way that linear feedback shift registers are related to error correcting codes.

The discrete Fourier transform (described above) does not appear to be the "correct" transform for the study of sequences generated by FCSR's, and it is not immediately obvious what the best replacement should be. In this paper, we show that the usual (complex) Fourier transform is an appropriate vehicle for the study of FCSR sequences, and we use it to describe an analog of Blahut's theorem for the 2-adic span of a sequence. Not only does this give a "shift register" interpretation for the usual (complex) Fourier transform of a periodic binary sequence, but it also adds another entry to the long and growing list of parallels between the theory of LFSR sequences and the theory of FCSR sequences.

2 Statement of the Result

In what follows we let $S = a_0, a_1, \cdots$ be a periodic binary sequence with period L and let $\zeta \in \mathbf{C}$ be a complex primitive L-th root of unity.

Definition 2.1 *For $k = 0, 1, \cdots, L-1$, the k-th Fourier coefficient is*

$$\hat{a}_k = \sum_{i=0}^{L-1} a_i \zeta^{ki} \in \mathbf{C}.$$

The set of Fourier coefficients is the Fourier transform of S. *We denote by* $\sigma(S)$ *the number of nonzero Fourier coefficients of* S.

The sequence S can be interpreted as the coefficients of a formal power series,

$$\alpha = \sum_{i=0}^{\infty} a_i 2^i.$$

Such a series does not converge in the usual sense, but it nevertheless defines a *2-adic number*, $\alpha \in \mathbf{Z_2}$. (An elementary review of the 2-adic numbers is presented in [2].) The ring $\mathbf{Z_2}$ of 2-adic numbers contains all the (usual) rational numbers with *odd denominator*. It turns out that, since S is periodic, the 2-adic number α is in fact a rational number, which we may write in lowest terms as

$$\alpha = \frac{-p}{q},$$

with p and q relatively prime, and $0 \le p < q$. The *2-adic complexity* is $\Phi_2(S) = \log_2(q)$. It differs from the 2-adic span by the number of memory bits in the FCSR which is, in any case, no more than $\log_2(\Phi_2(S))$ and is completely analyzed in [2]. Our main result is:

Theorem 2.2 *Let* S *be a periodic binary sequence of period* L. *Then the 2-adic complexity* $\Phi_2(S)$ *is bounded as follows,*

$$\Phi_2(S) < \sigma(S) + 2^{\pi(L)-1}$$

where $\pi(L)$ *denotes the number of distinct prime divisors of* L.

The "error" term $2^{\pi(L)-1}$ is sharp, as the following result shows. The obstacle to obtaining a corresponding lower bound is that it is very difficult to estimate the cancelation between the numerator and denominator in the expression $\alpha = -h(2)/u(2)$ in Section 3 below. However, it seems reasonable to expect a lower bound for $\Phi_2(S)$ that is something like $\sigma(S) - 2^{\pi(L)-1}$.

Proposition 2.3 *Let* $\epsilon > 0$. *Then there exist periodic binary sequences* S *with* $\pi(L)$ *arbitrarily large such that*

$$\Phi_2(S) > \sigma(S) + (1-\epsilon)2^{\pi(L)-1}.$$

Also, there exist periodic binary sequences S *with* $\pi(L)$ *arbitrarily large such that*

$$\Phi_2(S) < \sigma(S) - (1-\epsilon)2^{\pi(L)-1}.$$

The proof is given in Section 4.

3 Preliminaries to the Proof

There is a useful polynomial that connects the 2-adic interpretation of a sequence with the definition of the Fourier coefficients:

$$f(x) = \sum_{i=0}^{L-1} a_i x^i.$$

Here we think of $f(x)$ as a polynomial with complex coefficients. Thus the nth Fourier coefficient is $f(\zeta^n)$ and the associated 2-adic number is

$$\alpha = f(2)2^0 + f(2)2^L + f(2)2^{2L} + \ldots = \frac{-f(2)}{2^L - 1}. \tag{1}$$

Now reduce α to lowest terms, say $\alpha = -p/q$ so that $\Phi_2(S) = \log_2(q)$.

We need to find a relation between $\gcd(f(2), 2^L - 1)$ (the greatest common divisor) and the number of vanishing Fourier coefficients $f(\zeta^k)$. In fact, we will relate both of these to the polynomial $\gcd(f(x), x^L - 1)$ (which by Galois theory has integer coefficients).

Let $g(x) = \gcd(f(x), x^L - 1)$, $f(x) = g(x)h(x)$, and $x^L - 1 = u(x)g(x)$. Let us factor $g(2)$ out of the numerator and denominator in equation (1). Then

$$\alpha = \frac{-h(2)}{u(2)}.$$

If $h(2)$ and $u(2)$ were relatively prime then we would have $\Phi_2(S) = \log_2(u(2))$, however this is not always the case (even though the polynomials $h(x)$ and $u(x)$ are relatively prime, as polynomials with integer coefficients.) The best we can say is that

$$\Phi_2(S) \leq \log_2(u(2)).$$

Recall that the *n-th cyclotomic polynomial* $c_n(x)$ is the minimal polynomial (with integer coefficients) of the primitive n-th roots of unity (that is, the monic polynomial of smallest degree that vanishes on the complex number $e^{2\pi i/n}$). Its degree is Euler's phi function, $\phi(n)$, the number of integers $j \leq n$ that are relatively prime to n. Recall also ([5] prop. 13.2.2) that

$$x^L - 1 = \prod_{n|L} c_n(x)$$

(where $n|L$ means that n is a divisor of L).

Since $u(x)$ divides $x^L - 1$, it must be a product of distinct cyclotomic polynomials,

$$u(x) = \prod_{j=1}^{t} c_{n_j}(x), \tag{2}$$

with n_j dividing L. Thus a bound for $c_{n_j}(2)$ leads to a bound for $\Phi_2(S)$.

Proposition 3.1 *Let n be a positive integer and let*

$$m = \prod_{p|n} p = s(n)$$

(the product taken over all prime divisors p of n) denote the largest squarefree integer dividing n. Let $q = n/m$. Then

$$1 - 2^{-q} < \frac{c_n(2)}{2^{\phi(n)}} < 1$$

if m has an even number of distinct prime divisors, and

$$1 < \frac{c_n(2)}{2^{\phi(n)}} < (1 - 2^{-q})^{-1}$$

if m has an odd number of distinct prime divisors.

Proof: Recall the Möbius function μ, which is defined for positive integers n by

$$\begin{aligned} \mu(n) &= (-1)^k \text{ if } n \text{ is squarefree and has } k \text{ distinct prime factors, and} \\ &= 0 \text{ otherwise.} \end{aligned}$$

Following are four basic properties of μ (see [5]§2.2).

1. $$\sum_{j|n} \mu(j) = \begin{cases} 0 & \text{if } n > 1, \\ 1 & \text{if } n = 1. \end{cases}$$

2. $\mu(ab) = \mu(a)\mu(b)$ if a and b are relatively prime.

3. $\sum_{j|n} j\mu(n/j) = \phi(n)$.

4. $c_n(x) = \prod_{j|n}(x^j - 1)^{\mu(n/j)}$.

The fourth property is obtained by applying the Möbius inversion formula ([5], §2.2) to the relation $x^n - 1 = \prod_{d|n} c_d(x)$.

The fourth and third properties imply that

$$c_n(2) = \prod_{j|n}(2^j - 1)^{\mu(n/j)} = 2^{\phi(n)} \prod_{j|n}(1 - 2^{-j})^{\mu(n/j)},$$

so it remains to estimate the latter product. We first need the following.

Lemma 3.2 *Let $d \geq 1$ and let $0 < x \leq \frac{1}{2}$. Then*

$$\prod_{j=d+1}^{\infty} (1 - x^j) > 1 - x^d.$$

Proof: The logarithm of the left-hand side can be expanded as

$$\sum_{j=d+1}^{\infty} \ln(1 - x^j) = -\sum_{i=1}^{\infty} \sum_{j=d+1}^{\infty} \frac{x^{ij}}{i}.$$

Since

$$\sum_{j=d+1}^{\infty} x^{ij} \leq \sum_{k=id+1}^{\infty} x^k \leq x^{id},$$

with both inequalities being equalities only when $i = 1$ and $x = \frac{1}{2}$, we obtain

$$\sum_{j=d+1}^{\infty} \ln(1 - x^{j+1}) > -\sum_{i=1}^{\infty} \frac{x^{id}}{i} = \ln(1 - x^d).$$

This yields the lemma. $\square$

Returning to the proof of Proposition 3.1, assume m has an even number of distinct prime factors, so $\mu(m) = 1$. Let $q = n/m$. Recall that $\mu(n/j) \neq 0$ only when n/j is squarefree, which is equivalent to q dividing j. Write $j = qk$. The condition $j|n$ becomes $k|m$, and we have

$$\prod_{j|n} (1 - 2^{-j})^{\mu(n/j)} = \prod_{k|m} (1 - 2^{-qk})^{\mu(m/k)}.$$

The first factor in this product is $(1 - 2^{-q})^{\mu(m)} = 1 - 2^{-q}$. The remaining product is bounded above by

$$\prod_{k=2}^{\infty} (1 - 2^{-qk})^{-1} < (1 - 2^{-q})^{-1},$$

by Lemma 3.2. This yields

$$\frac{c_n(2)}{2^{\phi(n)}} = \prod_{j|n} (1 - 2^{-j})^{\mu(n/j)} < 1.$$

To get a lower bound, note that the first nontrivial factor in

$$\prod_{k|m} (1 - 2^{-qk})^{\mu(m/k)}$$

is when $k = 1$ and the second is when k is the smallest prime factor p of m. Since m has an even number of prime factors, m/p has an odd number of prime factors, so $\mu(m/p) = -1$. The product for the remaining values of k is bounded below by

$$\prod_{k>p}(1-2^{-qk}) > (1-2^{-qp}).$$

This yields

$$\frac{c_n(2)}{2^{\phi(n)}} > (1-2^{-q})(1-2^{-qp})^{-1}(1-2^{-qp}),$$

which yields the result.

The case when m has an odd number of prime factors is similar. □

Proposition 3.3 *Fix a squarefree integer $m \geq 1$. Then*

$$\sum_{s(n)=m} \log_2\left(\frac{c_n(2)}{2^{\phi(n)}}\right) = -\mu(m),$$

and all terms in the sum have the same sign.

Proof: The fact that all terms have the same sign follows immediately from Proposition 3.1.

Observe that two integers m and n have exactly the same prime factors if and only if each divides a power of the other. For simplicity, we write $m|n^\infty$ to mean that m divides some power of n. For a squarefree integer m, this is equivalent to $m|n$.

Now fix a squarefree integer m. We have $s(n) = m$ if and only if $m|n$ and $n|m^\infty$. Moreover, an integer q occurs as n/m for some n with $s(n) = m$ if and only if $q|m^\infty$.

As in the proof of Proposition 3.1,

$$\frac{c_n(2)}{2^{\phi(n)}} = \prod_{j|m}(1-2^{-jq})^{\mu(m/j)},$$

where $q = n/m$. Therefore

$$\prod_{s(n)=m} \frac{c_n(2)}{2^{\phi(n)}} = \prod_{q|m^\infty}\prod_{j|m}(1-2^{-jq})^{\mu(m/j)}.$$

Note that an integer k occurs as a product jq, with $q|m^\infty$ and $j|m$, exactly when $k|m^\infty$. Fix such a k. The condition $k = jq$ automatically implies $q|m^\infty$. The values of j that yield k (i.e., $k = jq$ for some q) are therefore those that satisfy both $j|m$ and $j|k$. This is equivalent to $j|\gcd(m,k)$. Therefore the factor $(1-2^{-k})$ occurs in the last product with exponent

$$\sum_{j|\gcd(m,k)} \mu(m/j).$$

Write $m = \gcd(m,k)m'$. Since m is squarefree, $\gcd(m,k)$ and m' are relatively prime, so $\mu(m/j) = \mu(\gcd(m,k)/j)\mu(m')$, by Property 2 of μ. Therefore we obtain

$$\mu(m') \sum_{j|\gcd(m,k)} \mu(\gcd(m,k)/j) = 0 \text{ when } \gcd(m,k) \neq 1,$$

by Property 3 of μ, since $\gcd(m,k)/j$ runs though all the divisors of $\gcd(m,k)$. Since $k|m^\infty$, we have $\gcd(m,k) = 1$ if and only if $k = 1$, which corresponds to $q = j = 1$. Therefore the product reduces to

$$(1-2^{-1})^{\mu(m)} = 2^{-\mu(m)}.$$

This yields the result.

Finally, note that no factor $(1-2^{-k})$ occurs more than m times in the above calculations (actually, no more than the number of divisors of m). Moreover, $\prod_{k=1}^{\infty}(1-2^{-k})^m$ converges to a nonzero number. Therefore the above products converge "absolutely" (i.e., their logarithms are absolutely convergent series), so the rearrangement of the factors in the above infinite products is justified. This completes the proof of Proposition 3.3. □

Lemma 3.4 *Let $L > 1$ be squarefree. Then L has $2^{\pi(L)-1}$ squarefree divisors with an odd number of prime factors.*

Proof: Every squarefree divisor of L divides $s(L)$, so we may work with $s(L)$. Let E be the set of j dividing $s(L)$ with an even number of prime factors, and let O be the set of j dividing $s(L)$ with an odd number of prime factors. By Property 1 of μ,

$$0 = \sum_{j|s(L)} \mu(j) = \#(E) - \#(O).$$

Therefore half the divisors of $s(L)$ are in E and half are in O. Since the number of divisors of $s(L)$ is $2^{\pi(s(L))}$ (see [5], §2.2.2), the lemma follows. □

4 Proof of Theorem 2.2

It follows from equation (2) that

$$\begin{aligned}
\log_2(u(2)) &= \sum_{j=1}^{t} \log_2(c_{n_j}(2)) \\
&= \sum_{j=1}^{t} \left(\phi(n_j) + \log_2 \left(\frac{c_{n_j}(2)}{2^{\phi(n_j)}} \right) \right) \\
&= \deg(u) + \sum_{j=1}^{t} \log_2 \left(\frac{c_{n_j}(2)}{2^{\phi(n_j)}} \right).
\end{aligned}$$

Let $s(n_j) = m_j$. Let E and O denote the sets of squarefree integers with $\mu(m_j) = 1$ and $\mu(m_j) = -1$, respectively. We obtain

$$\log_2(u(2)) \leq \deg(u) + \sum_{m_j \in O} \log_2 \left(\frac{c_{n_j}(2)}{2^{\phi(n_j)}} \right),$$

since the terms with $m_j \in E$ are negative by Proposition 3.1. If there are two or more n_j with the same m_j, then group them together. We obtain a subsum of the sum in Proposition 3.3. Since all the terms in the sum of Proposition 3.3 are positive, the proposition gives the upper bound of 1 for our subsum. By Lemma 3.4, there are at most $2^{s(L)-1}$ distinct values of $m_j \in O$. Therefore

$$\log_2(u(2)) < \deg(u) + 2^{\pi(L)-1}.$$

We claim that $\sigma(S) = \deg(u)$. Note that $\hat{a}_n \neq 0 \iff f(\zeta^n) \neq 0 \iff g(\zeta^n) \neq 0 \iff u(\zeta^n) = 0$. The last equivalence is because $u(x)g(x) = x^L - 1$, so each L-th root of unity ζ^n appears as a root of exactly one of $u(x)$ and $g(x)$. Note also that $u(x)$ has exactly $\deg(u)$ complex roots, and all are of the form ζ^n. Hence the number of nonzero Fourier coefficients satisfies $\sigma(S) = \deg(u)$. □

We finish by proving Proposition 2.3. Let N be a natural number and let $p_1 < p_2 < \ldots < p_N$ be primes such that $1 - 2^{1-p_1} > 2^{-\epsilon}$. Let $L = p_1 p_2 \cdots p_N$ and let $n_1, \ldots, n_t$ run through all divisors of L such that $\mu(n_i) = -1$. Then $t = 2^{N-1}$ by Lemma 3.4. Let

$$\alpha = \frac{-1}{c_{n_1}(2) \cdots c_{n_t}(2)} = \frac{-p}{q}$$

in the notation of Section 2, and let S be the corresponding periodic binary sequence. Then

$$\sigma(S) = \sum_i \deg c_{n_i}(X) = \sum_i \phi(n_i)$$

and

$$\Phi_2(S) = \sum_i \log_2(c_{n_i}(2)).$$

We have

$$\begin{aligned}
\frac{c_{n_i}(2)}{2^{\phi(n_i)}} &= \prod_{j|n_i}(1-2^{-j})^{\mu(n_i/j)} \\
&> (1-2^{-1})^{-1}\prod_{j\geq p_1}(1-2^{-j}) \\
&> 2(1-2^{1-p_1}) \text{ by Lemma 3.2} \\
&> 2^{1-\epsilon}.
\end{aligned}$$

Therefore

$$\Phi_2(S) > \sum_{i=1}^{t}(\phi(n_i)+(1-\epsilon)) = \sigma(S)+(1-\epsilon)2^{N-1}.$$

Since $N=\pi(L)$, we are done. The proof of the second half of Proposition 2.3 is obtained in a similar way by using n_i with $\mu(n_i)=1$. $\square$

References

[1] R. Blahut, Transform techniques for error-control codes, *IBM J. Res. Develop.* **23** (1979), 299-315.

[2] A. Klapper and M. Goresky, Feedback shift registers, 2-adic span, and combiners with memory, *J. Crypt.* **10** (1997), 111-147.

[3] J. Massey, Cryptography and system theory, *Proc. 24th Allerton Conf. Commun., Control, Comput.*, Oct 1-3, 1986.

[4] J. Massey and S. Serconek, Linear complexity of periodic sequences: a general theory, in N. Koblitz, ed., *Advances in Cryptology - CRYPTO 97.* (Springer Lecture Notes in Computer Science vol. **1109**, Springer Verlag, N. Y., 1997.) pp. 358-371.

[5] K. Ireland and M. Rosen, “A Classical Introduction to Modern Number Theory (second edition)”, Graduate Texts in Mathematics vol. 84, Springer Verlag, N.Y., 1990.