Large Period Nearly deBruijn FCSR Sequences (Extended Abstract)

Andrew Klapper

Dept. of Computer Science
University of Kentucky
Lexington, KY 40506-0046 USA
klapper@cs.engr.uky.edu

Mark Goresky

Dept. of Mathematics
Northeastern University
Boston, MA 03115 USA
goresky@nuhub.neu.edu

Abstract. Recently, a new class of feedback shift registers (FCSRs) was introduced, based on algebra over the 2-adic numbers. The sequences generated by these registers have many algebraic properties similar to those generated by linear feedback shift registers. However, it appears to be significantly more difficult to find maximal period FCSR sequences. In this paper we exhibit a technique for easily finding FCSRs that generate nearly maximal period sequences. We further show that these sequence have excellent distributional properties. They are balanced, and nearly have the deBruijn property for distributions of subsequences.

Index Terms – **Binary sequences, feedback with carry shift registers, deBruijn property, 2-adic numbers.**

1 Introduction

Pseudorandom sequences with a variety of statistical properties, such as large period, high linear span, and good statistical distributions, are important in many areas of communications and computing, such as cryptography, spread spectrum communications, error correcting codes, and Monte Carlo integration. Thus devices for generating sequences with such good properties are basic tools for the design of stream ciphers (as well as for other applications). While such properties alone are insufficient to make sequences useful for encryption, they are initial minimal requirements. Once we can generate sequences with these properties, various techniques can be used to further scramble sequences making them suitable for encryption, while perhaps retaining good statistical properties.

One class of sequences that has many nice properties is the class of linear feedback shift register (LFSR) sequences. Maximal period LFSR sequences (or m-sequences) are known to have large period and a balance of zeros and ones, and to become deBruijn sequences when a single zero is inserted [3]. These properties, as well as the availability of algebraic tools for their analysis, have led to their use in a number of constructions of key stream generators. Examples

L.C. Guillou and J.-J. Quisquater (Eds.): Advances in Cryptology - EUROCRYPT '95, LNCS 921, pp. 263-273, 1995.

include nonlinear feedforward functions [6], nonlinear combining functions [14], and clock controlled shift registers [2].

Recently a new class of binary sequence generators, *feedback with carry shift registers* (or FCSRs) has been described by Klapper and Goresky [8, 9]. They have many of the nice algebraic properties of LFSR sequences and, it is hoped, will serve as building blocks for stream ciphers in much the same way that LFSR sequences have in the past. In this paper we study some of the basic statistical properties of FCSR sequences. We show how to construct, in an effective manner, FCSRs with very large period. Previously described methods for doing this required the choice of a prime number q for which 2 is a primitive root. Unfortunately, there is no known effective way of testing this condition, nor is it even known whether there are infinitely many such primes. Here we show that if p is a prime number such that 2 is a primitive root modulo p and modulo p^2, then for any positive integer e, using $q = p^e$ as the connection integer of a FCSR results in an output sequence with period $\phi(q) = p^e - p^{e-1}$. The condition that 2 be primitive modulo p^2 is known to hold whenever 2 is primitive modulo p for $p < 2 \cdot 10^{10}$. We also give an explicit procedure for finding the initial settings of FCSRs with this period.

We further show that the sequences so constructed have excellent statistical properties in the sense that they are nearly deBruijn sequences. Recall that a deBruijn sequence is a sequence bfa of period N such that every sequence of length $\log(N)$ (N must be a power of 2) occurs precisely once in each period of bfa. In other words, the numbers of occurrences of any two subsequences of length $\log(N)$ are equal. For sequences whose period is not a power of 2, the best we can hope for in this regard is that the numbers of occurrences of any two sequences differ by at most one. This, in fact, is the case when q is prime and 2 is a primitive root modulo q [8]. When q is a power of a prime and 2 is a primitive root modulo q, we show that the numbers of occurrences of any two subsequences differ by at most two. This holds for subsequences of *any* length.

Finally, we consider an arithmetic, or "with carry," analog of the cross-correlation of two sequences. We show that for any two decimations of a FCSR sequence of the type described above, the arithmetic correlations are identically zero, except when the two sequences coincide.

2 Feedback with Carry Shift Registers

In this section we review the operation of FCSRs and recall their basic algebraic properties. See [8, 9] for details. Let q be an odd positive integer, and let $q+1$ have the binary expansion $q+1 = \sum_{i=1}^{r} q_i 2^i$ with $q_i \in \{0, 1\}$. For convenience we also let $q_0 = -1$, so $q = \sum_{i=0}^{r} q_i 2^i$. The coefficients $q_1, \cdots, q_r$ are to be thought of as the taps on a feedback register. We can think of q as giving a recurrence with carry on the output sequence of this register.

Definition 1. The FCSR with connection integer q is a feedback register with r bits of storage plus additional memory for carry. If the contents of the register

at any given time are $(a_{r-1}, a_{r-2}, \ldots, a_1, a_0)$ and the memory is m, then the operation of the shift register is defined as follows:

A1. Form the integer sum $\sigma = \sum_{k=1}^{r} q_k a_{r-k} + m$.
A2. Shift the contents one step to the right, outputting the rightmost bit a_0.
A3. Place $a_r = \sigma \pmod 2$ into the leftmost cell of the shift register
A4. Replace the memory m with $(\sigma - a_r)/2$.

Such a register outputs an infinite binary sequence $\mathbf{a} = (a_0, a_1, a_2, \cdots)$. The analysis of FCSR sequences employs the 2-adic number associated with $\mathbf{a}$, i.e., the power series with indeterminate replaced by 2, $\alpha = \sum_{i=0}^{\infty} a_i 2^i$. This 2-adic number plays a role similar to that of the generating function in linear feedback shift register theory. See [11] for background on 2-adic numbers. For convenience from time to time we speak of α as being the output of a FCSR, or as being periodic, or as having any other property that should more properly be attributed to the sequence $\mathbf{a}$.

The following facts are known about FCSRs and their output sequences [8, 9].

1. A binary sequence $\mathbf{a}$ is eventually periodic if and only if its associated 2-adic number α is a rational number c/q. It is strictly periodic if and only if moreover $-q < c \leq 0$.
2. If $\mathbf{a}$ is the output sequence of a FCSR, and α is the associated 2-adic number, then $\mathbf{a}$ is eventually periodic and $\alpha = c/q$ where q is the connection number of a FCSR that outputs $\mathbf{a}$. In this case we can write

$$c = \sum_{i=0}^{r-1} \sum_{j=0}^{r-i-1} q_i a_j 2^{i+j} - m2^r, \tag{1}$$

   where m is the initial state of the extra memory.
3. Conversely, every eventually periodic binary sequence $\mathbf{a}$ whose associated 2-adic number can be written $\alpha = c/q$ for integers c, q, with q odd, is the output of a FCSR with connection number q.
4. Suppose $\mathbf{a}$ is the output sequence of a FCSR with connection integer q. Let γ be the inverse of 2 modulo q. Then there exists $A \in \mathbf{Z}/(q)$ such that for every i, $a_i = (A\gamma^i \bmod q) \bmod 2$. This composition of mod operations means first reduce modulo q to a number between 0 and $q-1$, then reduce the result modulo 2.
5. Adding b to the initial memory changes α by $-b2^r/q$.

As a consequence of the exponential representation of FCSR sequences, it is apparent that the period of the output of a FCSR with connection number q can be no more than the cardinality of the multiplicative group of the integers modulo q, $(\mathbf{Z}/(q))^*$.

Definition 2. An ℓ-sequence is a FCSR sequence with maximum possible period $T = |(\mathbf{Z}/(q))^*|$.

An ℓ-sequence is analogous to an m-sequence in LFSR theory. Such a sequence is generated by connection numbers q for which 2 is a *primitive root*. The best we can hope is that the period is $q-1$. This occurs when q is prime and 2 is a primitive root modulo q. The search for primes q such that 2 is a primitive root is related to a large body of contemporary number theory. It is believed that there are infinitely many primes q with this property [5]. However, finding such primes (and even finding large primes at all) is problematic.

In this paper we consider two fundamental questions about FCSR sequences:

1. How can we guarantee the output sequence has large period?
2. What are the statistical properties of large period FCSR sequences?

The first question can be divided into two parts:

1. How can we guarantee that the 2-adic number c/q has large period? Equivalently, how can we guarantee that 2 is a primitive root modulo q and c is relatively prime to q?
2. Given a rational number c/q, how can we efficiently construct the initial loading of a FCSR that outputs c/q?

3 Finding ℓ-Sequences for Prime Powers

In this section we give a method for generating ℓ-sequences based on FCSRs whose connection numbers are prime powers. Note that if q is not a prime power, then $\mathbf{Z}/(q)^*$ is not a cyclic group, so 2 cannot be primitive and we can have no ℓ-sequences. The following fact is well known, but we include a brief proof for completeness.

Theorem 3. *Let q be a power of a prime p, say $q = p^e$. If 2 is a primitive root modulo p^2, then 2 is a primitive root modulo q as well.*

Proof: The order of the multiplicative group of integers modulo q is $\phi(q) = p^{e-1}(p-1)$. Thus $2^{p^{e-1}(p-1)} \equiv 1 \bmod q$, and we must show that there is no prime number $t > 1$ dividing $p^{e-1}(p-1)$ such that $2^{p^{e-1}(p-1)/t} \equiv 1 \bmod q$. We do so by induction on e. That is, we may assume that the order of 2 modulo p^{e-1} is $p^{e-2}(p-1)$. Also note that if 2 is a primitive root modulo p^2, it must also be a primitive root modulo p.

Suppose t divides $p-1$. From the fact that $2^p \equiv 2 \bmod p$, it follows that $2^{p-1}/t \equiv 1 \bmod p$, contradicting the primitivity of 2 modulo p.

Thus we may assume $t = p$. Thus p^e divides $2^{p^{e-2}(p-1)} - 1$, but by induction p^{e-1} does not divide $2^{p^{e-3}(p-1)} - 1$. Also, p^{e-2} divides $2^{p^{e-3}(p-1)} - 1$. Thus $2^{p^{e-3}(p-1)} = 1 + p^{e-2}y$ for some y relatively prime to p. But then

$$2^{p^{e-2}(p-1)} - 1 = (1 + p^{e-2}y)^p - 1 \equiv p^{e-1}y \bmod p^e$$

when $e \geq 3$. Thus p^e does not divide $2^{p^{e-2}(p-1)} - 1$, a contradiction. □

Suppose we have a prime p for which 2 is primitive. Checking whether 2 is primitive modulo p^2 is quite easy. Suppose this is not the case. The order of the multiplicative group $\mathbf{Z}/(p^2)^*$ is $p(p-1)$, so for some divisor $t \neq 1$ of $p(p-1)$, the order of 2 is $p(p-1)/t$. That is, p^2 divides $2^{p(p-1)/t} - 1$. We may assume t is prime. Thus either $t = p$ or t is a divisor of $p-1$.

If t is a divisor of $p-1$, then p is a divisor of $2^{p(p-1)/t} - 1$. But $2^p \equiv 2 \bmod p$, so p divides $2^{(p-1)/t} - 1$. This contradicts the assumption that 2 is primitive modulo p. Thus $t = p$, so p^2 divides $2^{p-1} - 1$. It follows that to check whether 2 is a primitive root modulo p^2, it suffices to check whether p^2 divides $2^{p-1} - 1$.

Thus we can find ℓ-sequences as follows: choose a small prime p for which 2 is primitive (for small p this can be checked easily); check that p^2 does not divide $2^{p-1} - 1$; choose $e > 0$ and let $q = p^e$; choose an integer c relatively prime to p, with $0 < c < q$; and construct the FCSR with connection integer q and output $-c/q$. This gives us a FCSR whose output is strictly periodic with period $|\mathbf{Z}/(q)^*| = p^e - p^{e-1}$. In the next section we discuss the construction of the initial loading of the FCSR for a given $-c/q$.

One can ask about the abundance of primes p for which 2 is a primitive root modulo p^2. Hardy and Wright point out that the condition that p^2 divides $2^{p-1} - 1$ holds for only two primes p less than $3 \cdot 10^7$ [4, p. 73], and by computer search Bombieri has extended this limit to $2 \cdot 10^{10}$ [1]. (The two primes are 1093 and 3511.) In both cases 2 is not primitive modulo p. Thus for a large number of primes, we need only check the primitivity of 2 modulo p. In fact, it is not known whether there are *any* primes p such that 2 is primitive modulo p but not modulo p^2, though there is no compelling reason to believe there are no such primes.

4 Initial Loading of a FCSR

In this section we describe how an initial loading can be chosen for a FCSR that guarantees the output will be purely periodic and will have the maximum period for the given connection number.

It has been shown that, for a given rational number c/q, the initial loading for an FCSR that gives output c/q can be found by the following procedure [8, 9].

B1. Set $m_{-1} = c$.

B2. For each $i = 0, 1, \ldots, r-1$ compute the following numbers:

$$\sigma_i = \sum_{k=0}^{i-1} q_{i-k} a_k + m_{i-1} \in \mathbf{Z} \tag{2}$$

$$a_i = \sigma_i \pmod 2 \in \mathbf{Z}/(2) \tag{3}$$

$$m_i = \frac{\sigma_i - a_i}{2}. \tag{4}$$

If we use the initial loading $(a_{r-1}, a_{r-2}, \ldots, a_1, a_0)$ and initial memory $m_{r-1} \in R$, then the resulting FCSR outputs the 2-adic expansion of c/q. If c is relatively prime to q, then the period of the sequence is $T = \mathrm{ord}_q(2)$. However if c and q have a common factor then the period may be smaller but at least it will divide $\mathrm{ord}_q(2)$. Thus for $q = p^e$ with 2 primitive modulo p^2 (and hence also modulo q), if we randomly choose c, check that $\gcd(c, p) = 1$, and then find an initial loading using the above procedure, we will find an initial loading that gives a period $p^e - p^{e-1}$ strictly period output sequence. The expected number of random choices of c needed to achieve this is $p/(p-1)$, since the probability that c is relatively prime to p is $(p-1)/p$.

Alternatively, we may want more control over the initial setting of the register. We can choose the initial contents of the register, then attempt to find an initial value of the memory that gives the desired output sequence. We proceed as follows.

1. Randomly choose bits $a_0, \cdots, a_{r-1} \in \{0, 1\}$.
2. Compute

$$z = \sum_{i=0}^{r-1} \sum_{j=0}^{r-i-1} q_i a_j 2^{i+j} \tag{5}$$

$$= \sum_{k=0}^{r-1} \sum_{i=0}^{k} q_i a_{k-i} 2^k. \tag{6}$$

3. Let

$$m = \left\lceil \frac{z}{2^r} \right\rceil.$$

4. Check $\gcd(2^r m - z, q) = 1$. If so, use $a_0, \cdots, a_{r-1}$ as the initial loading, and m as the initial memory. If not, repeat (1) – (4). In some cases $\lfloor (z+q)/2^r \rfloor = \lceil z/2^r \rceil + 1$ and can also be tried as the the initial memory.

To see that this gives a maximal period purely periodic sequence for q, it suffices to check that $0 \leq 2^r m - z < q$, since the output from the q-FCSR with these initial values is $z - 2^r m/q$. But this follows immediately from the choice of m.

The big question is the time complexity. First observe that in any given repetition of (1) – (4), the probability of success is at least $(p^e - p^{e-1})/p^e = (p-1)/p$. Thus the expected number of trials is only $p/(p-1)$.

The most costly part of this algorithm is step (2). This can be done quickly using a divide and conquer algorithm similar to divide and conquer multiplication. For $r-1$ bit integers q and a, we define the operation

$$\mathrm{semimult}(q, a, r) = \sum_{k=0}^{r-1} \sum_{i=0}^{k} q_i a_{k-i} 2^k.$$

If we write $q = q' + 2^{[r/2]}q''$ and $a = a' + 2^{[r/2]}a''$, then we have

$$\text{semimult}(q, a, r) =$$

$$q'a' + 2^{[r/2]}(\text{scmimult}(q', a'', r - [r/2]) + \text{scmimult}(q'', a', r - [r/2])).$$

Thus the time it takes to compute semimult(q, a, r) satisfies a recurrence

$$T(r) = 2T([r/2]) + S([r/2]) + \mathcal{O}(r),$$

where $S(r)$ is the time it takes to multiply two r bit numbers. Thus $T(r) \leq S(r) + cr$ for some constant c. Furthermore, if we use the Schönhage-Strassen algorithm [15], then $S(r) = \mathcal{O}(r \log r \log\log r)$. This can be improved to $S(r) \sim r \log r$ using Pollard's nonasymptotic algorithm for $r < 2^{37}$ on a 32 bit machine or $r < 2^{70}$ on a 64 bit machine [13].

Finally, observe that $\gcd(2^r m - z, q) = 1$ if and only if $\gcd(2^r m - z, p) = 1$. This can be checked using the Euclidean algorithm in $\mathcal{O}(r \log(p)^2)$ bit operations.

In summary, a desired initial loading can be found in less than expected $2r \log r + \mathcal{O}(r \log(p)^2)$ time for $r < 2^{37}$ on a 32 bit machine, or $r < 2^{70}$ on a 64 bit machine.

5 Distributional Properties

In this section we show that the sequences constructed above have excellent distributional properties. First we note that they are balanced.

Proposition 4. *Let q be a power of a prime p, say $q = p^e$, and suppose that 2 is primitive modulo q. Let* **a** *be any maximal period FCSR sequence, generated by a FCSR with connection integer q. The number of zeros and the number of ones in one period of* **a** *are equal.*

Furthermore we can consider higher order distributions. We show next that these sequences are close to having the deBruijn property that each subsequence of length log of the period occurs exactly once in each period. We show that for any two such subsequences, their numbers of occurrences can differ by at most two.

Theorem 5. *Let q be a power of a prime p, say $q = p^e$, and suppose that 2 is primitive modulo q. Let s be any nonnegative integer, and let A and B be s bit subsequences. Let* **a** *be any maximal period, purely periodic FCSR sequence, generated by a FCSR with connection integer q. Then the numbers of occurrences of A and B in* **a** *with their starting positions in a fixed period of* **a** *differ by at most 2.*

Proof: The purely periodic FCSR sequences with connection integer q are precisely the 2-adic expansions of rational numbers $-x/q$, with $0 \leq x < q$ [8, 9]. Such a sequence has maximum period if and only if p does not divide x. Since 2

is primitive modulo q, the cyclic shifts of bfa correspond to the set of all rational numbers $-x/q$, with $0 \leq x < q$. Thus an s bit subsequence A occurs in bfa if and only if it occurs as the first s bits in the 2-adic expansion of some rational numbers $-x/q$ with $0 \leq x < q$ and p not dividing x. Two rational number $-x_1/q$ and $-x_2/q$ have the same first s bits if and only if $-x_1/q \equiv -x_2/q \bmod 2^s$, if and only if $x_1 \equiv x_2 \bmod 2^s$. Thus we want to count the number of x with a given first s bits, $0 \leq x < q$, and x not divisible by p.

Let $2^r < q < 2^{r+1}$. If $s > r$, there are either zero or one such x, so the result follows. Thus we may assume $s \leq r$.

We first count the number of x with the first s bits fixed and $0 \leq x < q$, ignoring the divisibility condition. If $A = a_0, \cdots, a_{s-1}$, we let $\alpha = \sum_{i=0}^{s-1} a_i 2^i$. Let $q = \sum_{i=0}^{r} q_i 2^i$, and $q' = \sum_{i=0}^{s-1} q_i 2^i$. If $\alpha < q'$, then every choice of $a_s, \cdots, a_r$ with $\sum_{i=s}^{r} a_i 2^i \leq \sum_{i=s}^{r} q_i 2^i$ gives a unique x in the right range. If $\alpha \geq q'$, then every choice of $a_s, \cdots, a_r$ with $\sum_{i=s}^{r} a_i 2^i < \sum_{i=s}^{r} q_i 2^i$ gives a unique x in the right range. Thus for different choices of A, the numbers of such x differ by at most one.

Next we consider those x for which $0 \leq x < q$ and p divides x. That is, $x = py$ for some y, and $0 \leq y < q/p = p^{e-1}$. As above, $x_1 = py_1$ and $x_2 = py_2$ have the same first s bits if and only if the same is true of y_1 and y_2. The preceding paragraph shows that the numbers of such y for different choices of the first s bits differ by at most one. But if $x = py$, then $y \equiv A \bmod 2^s$ if and only if $x \equiv pA \bmod 2^s$, so for any B and C, the number of xs divisible by p with first s bits equal to B differs from the number of xs divisible by p with first s bits equal to C by at most 1. We have

$$|\{x : 0 \leq x < q,\ p \not| x, \wedge x \equiv \alpha \bmod 2^s\}|$$
$$= |\{x : 0 \leq x < q \wedge x \equiv \alpha \bmod 2^s\}| - |\{x : 0 \leq x < q,\ p|x, \wedge x \equiv \alpha \bmod 2^s\}|.$$

As α varies the two terms on the right hand side vary by at most one from their values for any fixed choice of α. Thus the difference varies by at most 2. □

It is easy to check that the difference can be as large as 2.

6 Arithmetic Correlations

Traditionally, the shifted cross-correlations of two sequences have been used as a measure of the extent to which the sequences are independent. These values are small if corresponding bits in one sequence are as likely to be equal as they are to be different. In the case of FCSR sequences, it appears quite difficult to compute cross-correlations in the usual sense. There is, however, an arithmetic (or "with carry") analog of the cross-correlation. This has been studied previously in the case of autocorrelation functions by Mandelbaum [12].

Definition 6. Let $\mathbf{a}$ and $\mathbf{b}$ be two eventually periodic sequences with period N, and let $0 \leq \tau < N$. Let $\mathbf{b}^\tau$ be the sequence formed by shifting $\mathbf{b}$ by τ positions, $\mathbf{b}_i^\tau = \mathbf{b}_{i+\tau}$. Then the *shifted arithmetic cross-correlation* $\Theta_{\mathbf{a},\mathbf{b}}(\tau)$ of $\mathbf{a}$ and $\mathbf{b}$ is the difference between the number of zeros and the number of ones in a complete

period of the periodic part of the sequence formed by subtracting $\mathbf{b}^\tau$ from $\mathbf{a}$ *with carry*. When $\mathbf{a} = \mathbf{b}$, the cross-correlation is called the *autocorrelation* of $\mathbf{a}$.

This corresponds to forming the 2-adic numbers α and β associated with $\mathbf{a}$ and $\mathbf{b}$, computing $\gamma = \alpha - 2^{-\tau}\beta$, and taking the difference between the number of coefficients that are zero and the number of coefficients that are one in a single period of γ. Even when $\mathbf{a}$ and $\mathbf{b}$ are purely periodic, there may be a transient prefix before γ becomes periodic. However, in this case the purely periodic part of γ is guaranteed to begin after at most N bits.

In the case of m-sequences and standard correlations the cross-correlations must be at least one in absolute value, simply because the periods of the sequences are odd. In general, the larger a family of sequences, the larger the maximum cross-correlations in the family. Remarkably, we can exhibit families of sequences in which all arithmetic correlations are identically zero. This generalizes a result of Mandelbaum showing that shifted autocorrelations of ℓ-sequences based on prime connection integers are identically zero [12]. Recall that a sequence $\mathbf{b}$ is a *k-fold decimation* of a sequence $\mathbf{a}$ if $\mathbf{b}$ is formed by taking every kth term of $\mathbf{a}$. That is, $b_i = a_{ki}$.

Theorem 7. *Let* $\mathbf{a}$ *be an* ℓ*-sequence based on connection integer* $q = p^e$, p *prime. Let* k *and* m *be integers that are relatively prime to the period* $p^e - p^{e-1}$ *of* $\mathbf{a}$. *Let* $\mathbf{b}$ *and* $\mathbf{c}$ *be* k*-fold and* m*-fold decimations of* $\mathbf{a}$, *respectively. Let* τ *be any shift. If* $\mathbf{c}$ *is a shift of* $\mathbf{b}$, *then there is one value of* τ *for which* $\Theta_{\mathbf{b},\mathbf{c}}(\tau) = p^e - p^{e-1}$. *In all other cases (whether or not* $\mathbf{c}$ *is a shift of* $\mathbf{b}$*),* $\Theta_{\mathbf{b},\mathbf{c}}(\tau) = 0$.

One can see, for example, that if we choose $q = 25$, then the decimations of the sequence of bits in the 2-adic expansion of $-1/q$ give eight cyclically distinct sequences of period 20 with ideal pairwise arithmetic correlations. In the classical theory of cross-correlations, any family of five or more sequence with this period must have maximum cross-correlation at least 5.

7 Conclusions

We have demonstrated that a large class of FCSR sequences are ℓ-sequences. This means that their periods are exponentially larger than the amount of initial information (taps on the register, initial register contents, and initial memory) required to generate the sequences. We have further shown that these sequences have excellent statistical properties, being nearly deBruijn sequences.

The picture for FCSRs whose connection number is not a prime power is more complicated. If $q = \prod p_i^{e_i}$, then the cardinality of $\mathbf{Z}/(q)$ is $\prod_{i=1}^{k} p_i^{e_i-1}(p_i - 1)$. However, this group is the product of cyclic groups of order $p_i^{e_i-1}(p_i - 1)$, so the order of its maximal order element is the least common multiple of the $p_i^{e_i-1}(p_i - 1)$. Since each of the $p_i - 1$ are even, the order of 2 cannot be the order of the full group. By the Chinese remainder theorem, the order of 2 modulo q is the least common multiple of the orders of 2 modulo $p_i^{e_i}$, $i = 1, \cdots, k$. Thus

if the p_i are chosen so that 2 is primitive modulo each p_i, then its order modulo q is

$$\mathrm{lcm}\{p_i^{e_i-1}(p_i-1)\} = \prod_{i=1}^{k} p_i^{e_i-1}\mathrm{lcm}\{p_i-1\}.$$

In fact, if we choose the p_i so that the greatest common divisor of any two of the $p_i - 1$ is 2, then the order of 2 modulo q is $\prod_{i=1}^{k} p_i^{e_i-1}\mathrm{lcm}\{p_i-1\}/2^{k-1}$. This is the largest we can hope for for the period of a FCSR whose connection number has k distinct prime factors.

The question of distributional properties of general FCSR sequences also remains. It would be nice to have bounds of the form in Section 5 in more general cases, or even just in the case described in the preceding paragraph.

Various extensions of the notion of FCSR have been suggested, based on complete valued fields other than the 2-adic numbers [10, 7]. The questions discussed in this paper can be asked in these settings as well.

Finally, now that we have established that certain FCSR sequences have good statistical properties, it remains to show that they can be modified (say with nonlinear feedforward functions, or by using nonlinear combiners) so they have large linear span and large 2-adic span. This would give us sequences with good statistics and resistance to the Berlekamp-Massey and 2-adic rational approximation algorithms, and thus good candidates for use in stream ciphers.

References

1. E. Bombieri, personal communication.
2. D. Gollman, Pseudo Random Properties of Cascade Connections of Clock Controlled Shift Registers, *Advances in Cryptology, Proceedings of Eurocrypt 84*, ed. T. Beth, N. Cot, and I. Ingemarsson, *Springer-Verlag LNCS* vol. 209, 1985, pp. 93-98.
3. S. Golomb, *Shift Register Sequences.* Aegean Park Press, Laguna Hills CA, 1982.
4. G. Hardy and E. Wright, *An Introduction to the Theory of Numbers.* Oxford University Press, Oxford UK, 1979.
5. C. Hooley, On Artin's conjecture. *J. Reine Angew. Math.* vol. 22, 1967 pp. 209-220.
6. E. L. Key, "An Analysis of the structure and complexity of nonlinear binary sequence generators," *IEEE Trans. Info. Theory,* vol. IT-22 no. 6, pp. 732-736, Nov. 1976.
7. A. Klapper, Feedback with Carry Shift Registers over Finite Fields, *Proceedings of Leuven Algorithms Workshop*, Leuven, Belgium, December, 1994.
8. A. Klapper and M. Goresky, Feedback Shift Registers, Combiners with Memory, and Arithmetic Codes, *Univ. of Kentucky, Dept. of Comp. Sci. Tech. Rep. No. 239-93.*
9. A. Klapper and M. Goresky, 2-Adic Shift Registers, *Fast Software Encryption: Proceedings of 1993 Cambridge Security Workshop,* ed. R. Anderson, *Springer-Verlag LNCS,* vol. 809, 1994, pp. 174-178.
10. A. Klapper, and M. Goresky, Feedback Registers Based on Ramified Extensions of the 2-Adic Numbers, to appear, *Proceedings, Eurocrypt 1994,* Perugia, Italy,

11. N. Koblitz, *p-Adic Numbers, p-Adic Analysis, and Zeta Functions.* Graduate Texts in Mathematics Vol. 58, Springer Verlag, N.Y. 1984.
12. D. Mandelbaum, Arithmetic codes with large distance. *IEEE Trans. Info. Theory,* vol. IT-13, 1967 pp. 237-242.
13. J. Pollard The Fast Fourier Transform in a Finite Field, *Math. Comp.,* vol. 25, 1971, pp. 365-374.
14. R. Rueppel, *Analysis and Design of Stream Ciphers.* Springer Verlag, New York, 1986.
15. A. Schönhage and V. Strassen, Schnelle Multiplikation Grosser Zahlen, *Computing,* vol. 7, 1971, pp. 281-292.