Polynomial pseudo-noise sequences based on algebraic feedback shift registers

Mark Goresky[*] Andrew Klapper[†]

Abstract

We apply the framework of algebraic feedback shift registers to polynomial rings over finite fields. This gives a construction of new pseudorandom sequences (over non-prime finite fields), which satisfy Golomb's three randomness criteria.

1 Introduction

The purpose of this paper is twofold: to study properties of a class of sequence generators based on polynomial algebra that generalizes linear feedback shift registers, and to study statistical properties of sequences over non-prime fields. The two purposes intersect when we determine conditions under which the new generators produce sequences for which certain statistics are ideal.

Sequences with various statistical randomness properties are important for many areas including radar, error correction, CDMA, cryptography, and Monte Carlo simulation. Golomb [4] (Chapt. III Sect. 4), for example, identified several such properties for binary sequences: the uniform distribution of subsequences, a distribution of lengths of runs of ones and zeros that fits the expectation for random sequences, ideal autocorrelations, and the shift and add property. In modern systems there is often an advantage to using sequences over a nonbinary alphabet, typically of size 2^8 or 2^w where w is the word size of the architecture in use. It is thus natural to consider generalizations of these properties to sequences over more general alphabets. Such generalizations are discussed in Section 2 and in Section 3 we see how sequences that are ideal with respoect to these properties

[*] School of Mathematics, Inst. for Adv. Study, Princeton N.J. www.math.ias.edu/∼goresky. Research partially supported by NSF grant #0002693.

[†] Dept. of Computer Science, 779A Anderson Hall, University of Kentucky, Lexington, KY, 40506-0046. www.cs.uky.edu/∼klapper. Research partially supported by NSF grant #9980429.

can be generated by LFSRs. In Section 4 we see that Blackburn has completely characterized the sequences with the shift and add property [1], extending work by Gong, Di Porto, and Wolfowicz [5]. We show that this characterization can be extended to a characterization of the sequences with both the shift and add property and uniform distribution of subsequences.

In many cases it is important to have fast methods of generating sequences with the desirable statistical properties. Sometimes it suffices to use maximal period linearly recurrent sequences, or *m-sequences*. These sequences can be efficiently generated by linear feedback shift registers (LFSRs). Sometimes, however, it is desirable to have other efficiently generated pseudorandom sequences. Recently the authors have studied a class of efficient sequence generators called *feedback with carry shift registers* or FCSRs and more generally *algebraic feedback shift registers* or AFSRs. These generators have algebraic mechanisms for analysis that parallel those of LFSRs — where LFSR sequences are analyzed in terms of the sequence of coefficients in the power series expansion of certain rational functions, FCSR sequences are analyzed in terms of the coefficient sequence of certain p-adic numbers. We have also identified the maximal period FCSR sequences, called *ℓ-sequences*, and shown that they share many of m-sequences' desirable properties [15, 16, 7, 17]. These ℓ-sequences have excellent randomness properties that are nearly as good as those of m-sequences. In Section 5 we review the basic properties of FCSRs, AFSRs, and ℓ-sequences.

AFSRs are a class of sequence generators that subsumes both LFSRs and FCSRs. Each class of AFSRs depends on a choice of an algebraic ring R and an element $r \in R$. For LFSRs, $R = F[x]$, the polynomial ring in one variable, and $r = x$. For FCSRs, $R = \mathbf{Z}$, the integers. In previous work we have studied properties of AFSRs based on more general rings R with characteristic 0. The central purpose of this paper is to study AFSRs based on $R = F[x]$ again, but with $r \neq x$. The new sequences are generally distinct from m-sequences (see Section 9), but have many of the same statistical properties as m-sequences: the distribution of fixed size subsequences is as uniform as possible, the distribution of lengths of runs matches the expectation and, with an appropriate definition, they have ideal autocorrelations. The basic properties of AFSR sequences over polynomial rings are described in Section 6, and their randomness properties are described in Section 7. The relationships between these sequences and Blackburn's construction and m-sequences are described in Sections 8 and 9. The question as to whether such sequences exist in abundance is considered in Section 10, and an example is given in Section 11.

2 Pseudorandomness of Sequences

In this section we describe Golomb's randomness postulates for sequences of elements in a vector space over a prime field $\mathbf{F}_p$.

2.1 Distribution of subsequences

For applications to cryptography and pseudo-Monte Carlo simulation, it is important that there be no statistical bias in the occurrence of individual symbols or small blocks of symbols in a sequence. Let F be a vector space of dimension e over the prime field $\mathbf{F}_p$. Throughout this section we assume that A is a periodic sequence of elements of F with period N. Golomb's first randomness postulate is the balance property.

Definition 2.1 *If $N = |F|^k - 1$, then the sequence A is* balanced *if each element $a \in F$ occurs $|F|^{k-1}$ times except for a single element, which occurs $|F|^{k-1} - 1$ times.*

More generally Golomb's fourth randomness property says that all blocks of fixed length occur as equally often as possible. A *string* $b = (b_0, b_1, \cdots, b_{t-1})$ of length t is an ordered sequence of t elements $b_i \in F$. An *occurrence* of the string b in (a single period of) the sequence A is an index $i \leq N - 2$ such that $(a_i, a_{i+1}, \cdots, a_{i+t-1}) = b$.

Definition 2.2 *If $N = |F|^k$, then the sequence A is a* de Bruijn sequence of span k *if every string of length k occurs exactly once in (each period of) A.*

In particular, if A is a de Bruijn sequence of span k, then the string $b = (0, 0, \cdots, 0)$ of length k occurs exactly once in each period of A.

Definition 2.3 *If $N = |F|^k - 1$, then the sequence A is a* punctured de Bruijn sequence of span k *if A is obtained from a de Bruijn sequence by deleting a single 0 from the single occurrence of the string $(0, 0, \cdots, 0)$ of length k in each period of A.*

A *run of length t* in A is a string of t consecutive identical symbols that is not contained in a longer string of consecutive symbols. For random sequences, if a run of as, $a \in F$, begins at position i, then with probability $(|F| - 1)/|F|$ the next symbol differs from a, so the run has length 1. If to the contrary the next symbol is a, then the symbol at position $i+1$ differs from a with probability $(|F|-1)/|F|$, so the run starting at position i has length two with probability $(|F| - 1)/|F|^2$. Continuing in this way, we see that the probability that a run in a random sequence has length t is $(|F| - 1)/|F|^t$. Golomb's second randomness postulate says that the distribution of runs is as close as possible to the expected distribution for random sequences.

Definition 2.4 *If $N = |F|^k - 1$, then the sequence A satisfies* the run property *if the number of runs of length $m \le k-1$ is $|F|^{k-m-1}$, the number of runs of length k is $|F|-1$, and there are no runs of length greater than k.*

The following statistical properties of punctured de Bruijn sequences are well known.

Lemma 2.5 *Let A be a punctured de Bruijn sequence of span k over F. Then*

1. *For any $t \le k$ and for any string b of length t, the number of occurrences of b in (a single period of) the sequence A is $|F|^{k-t}$ except for the single string $(0, 0, \cdots, 0)$ of length t, and this string occurs $|F|^{k-t} - 1$ times. In particular, A is balanced.*

2. *A has the run property.*

Proof: Part (1) just counts the number of ways of completing b to a string of length k. For part (2), first consider the case $1 \le m \le k-2$. A run of length t is a string of the form $xyy\cdots yz$ where x and z are distinct from y. By part (1), for each such choice of x, y, z there are $|F|^{k-t-2}$ occurrences of this string, and there are $|F|(|F|-1)^2$ such strings. If $t = k-1$ then each string $xyy\cdots y$ with $x \neq y$ occurs once. If $y \neq 0$ then for exactly one of these values of x will there be another y immediately following this string. So two values of x are forbidden and this gives $|F|-2$ possible values for x for each nonzero value of y, accounting for $(|F|-1)(|F|-2)$ runs. If $y = 0$ then for no choice of x will this string $xyy\cdots y$ be immediately followed by another y, so there are $|F|-1$ allowable values for x. Therefore the total number of runs of length $k-1$ is $(|F|-1)^2$. Next, suppose $t = k$. For each $y \neq 0$ there is a single string of k consecutive $y's$ and it occurs once in (each period of) A. This gives $|F|-1$ such runs; there is no run consisting of k zeroes. Finally, if there were a run of length greater than k consisting of a single symbol $c \in F$ then there would be two (or more) occurrences of the string $(c, c, \cdots, c)$ of length k, which is a contradiction. □

Remark. A choice of basis for F over $\mathbf{F}_p$ gives a way of translating each $a \in F$ into a string $\psi(a)$ over $\mathbf{F}_p$ of length e. Applying ψ to each symbol of A gives a sequence $\psi(A)$ over $\mathbf{F}_p$ whose period is e times the period of A. If A has one of the randomness properties described in this section, then $\psi(A)$ will not, in general, have the same property (both because the period is wrong and we must be concerned with substrings that don't align with the ends of the $\mathbf{F}_p$-ary representations of elements.

2.2 Autocorrelations and shift and add

Autocorrelations are usually defined only for sequences over prime fields. To generalize this notion, we briefly recall some standard facts about characters of finite Abelian groups.

Definition 2.6 *A* character *of a finite abelian group G is a group homomorphism from G to the multiplicative group $\mathbf{C}^* = \mathbf{C} - \{0\}$ of the complex numbers $\mathbf{C}$. That is, it is a function $\chi : G \to \mathbf{C}^*$ such that $\chi(a+b) = \chi(a)\chi(b)$ for all $a, b \in G$. Such a function is* nontrivial *if $\chi(a) \neq 1$ for some a.*

Lemma 2.7 *Let $\chi : G \to \mathbf{C}^*$ be a nontrivial character. Then $\sum_{g \in G} \chi(g) = 0$.*

Proof: Since χ is nontrivial, there exists $h \in G$ with $\chi(h) \neq 1$. Then

$$\chi(h) \sum_{g \in G} \chi(g) = \sum_{g \in G} \chi(gh) = \sum_{g' \in G} \chi(g')$$

so $(1 - \chi(h)) \sum_{g \in G} \chi(g) = 0$. □

Definition 2.8 *Let G be a finite abelian group.*

1. *Let A be a sequence of elements of G, with period N and let χ be a character of G. The* autocorrelation function *of A with respect to χ is the function*

$$\mathcal{A}_{A,\chi}(m) = \sum_{i=0}^{N-1} \chi(a_i)\overline{\chi}(a_{i+m}) = \sum_{i=0}^{N-1} \chi(a_i - a_{i+m}).$$

2. *A sequence A of elements of G, with period N has* ideal autocorrelations *if for every nontrivial character χ of G and every $m \not\equiv 0 \bmod N$ we have $|\mathcal{A}_{A,\chi}(m)| \leq 1$.*

Part 2 of Definition 2.8 generalizes Golomb's third randomness postulate.

Let $A = (a_0, a_1, \cdots)$ be a periodic sequence of elements from the vector space F and let $A_\tau = (a_\tau, a_{\tau+1}, \cdots)$ be its shift by τ steps. Let $A + A_\tau = (a_0 + a_\tau, a_1 + a_{\tau+1}, \cdots)$ be the sequence obtained from termwise addition of A and A_τ.

Definition 2.9 *The sequence A has the* shift-and-add property *if, for any shift τ, either (1) $A + A_\tau = 0$ (the all-zeroes sequence) or (2) there exists τ' such that $A + A_\tau = A_{\tau'}$.*

This generalizes Golomb's fifth randomness postulate. Similarly we can define the shift and subtract property. More generally, if F is a vector space over a field E containing $\mathbf{F}_p$, then A satisfies the shift and add property with coefficients in E if, for any $c, d \in E$ and for any shift τ, either $cA + dA_\tau = 0$ or else there exists a shift τ' such that $cA + dA_\tau = A_{\tau'}$.

Lemma 2.10 *The following are equivalent for a sequence A with characteristic p.*

1. *A has the shift and add property.*
2. *A has the shift and subtract property.*
3. *A has the shift and add property with coefficients in $\mathbf{F}_p$.*

Theorem 2.11 *If A is a balanced sequence over the $\mathbf{F}_p$-vector space F with period $N = |F|^k - 1$ and A has the shift and add property, then A has ideal autocorrelations.*

Proof: Let $\chi : F \to \mathbf{C}^*$ be a nontrivial (additive) character. By Lemma 2.7,

$$\sum_{i=0}^{N-1} \chi(a_i) = |F|^{k-1} \sum_{b \in F} \chi(b) - \chi(0) = -1$$

Since by Lemma 2.10 A satisfies the shift-with-subtract property, for any shift $\tau \not\equiv 0 \bmod T$, there is another shift τ' such that $a_i - a_{i+\tau} = a_{i+\tau'}$ for all i. Hence

$$\sum_{i=0}^{N-1} \chi(a_i)\overline{\chi}(a_{i+\tau}) = \sum_{i=0}^{N-1} \chi(a_i - a_{i+\tau}) = \sum_{i=0}^{N-1} \chi(a_{i+\tau'}) = \sum_{i=0}^{N-1} \chi(a_i) = -1. \qquad \square$$

3 LFSRs and M-sequences

Let F be a finite (Galois) field and let $q_1, q_2, \cdots, q_k \in F$. The *linearly recurrent sequence* of order k with multipliers $q_1, q_2, \cdots, q_k \in F$ and initial state $(a_0, a_1, \cdots, a_{k-1})$ is the unique solution to the equations

$$a_j = q_1 a_{j-1} + q_2 a_{j-2} + \cdots + q_k a_{j-k}$$

for $j \geq k$. Such a sequence may be described in three different ways. First, it is the output from a *linear feedback shift register* (LFSR) of length k with multipliers $q_i \in F$ and initial entries $a_0, a_1, \cdots, a_{k-1} \in F$, as illustrated in Figure 1. The $\oplus$ box denotes addition in F.

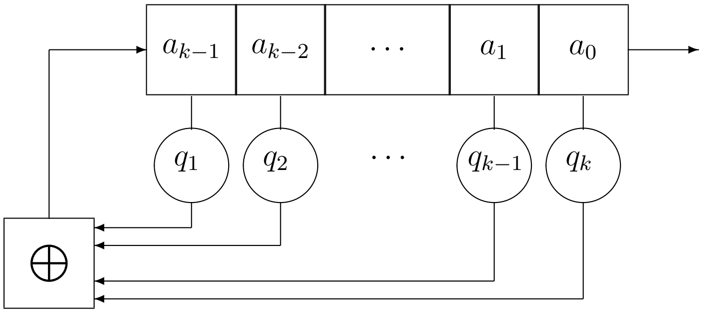

Figure 1: A Linear Feedback Shift Register

The *connection polynomial* $q \in F[x]$ associated with this recurrence or LFSR is the polynomial

$$q(x) = q_0 + \sum_{i=1}^{k} q_i x^i$$

where $q_0 = -1$. The second description is the well known fact [4] that the sequence $a_0, a_1, \cdots$ is also the coefficient sequence of the power series expansion

$$p(x)/q(x) = a_0 + a_1 x + a_2 x^2 + \cdots \tag{1}$$

of the rational function $p(x)/q(x)$ with denominator $q(x)$ and numerator

$$p(x) = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} x^j. \tag{2}$$

Finally, there are "exponential" representations for the sequence. In general we can express

$$a_j = T(ax^j) \tag{3}$$

for some $a \in L = F[x]/(q(x))$, where $T : L \to F$ is an F-linear function, the easiest of which is given by

$$T(\sum_{i=0}^{d-1} c_i x^i) = c_0$$

where $d = \deg(q)$. This function may also be expressed as $T(p) = p \bmod x$, for any polynomial $p(x) \in L$, so equation(3) becomes $a_j = ax^j \bmod x$. If $q(x)$ is irreducible then

L may be identified with the unique field extension of F having degree d. In this case equation (3) becomes the more familiar

$$a_i = Tr_{L/F}(a\alpha^i)$$

where $\alpha \in L$ is a choice of root of $q(x)$.

A linearly recurrent sequence of order k is eventually periodic and its period is at most $|F|^k - 1$. A linearly recurrent sequence of order k whose period is $|F|^k - 1$ is called a *maximal length sequence* or *m-sequence*. It is well known that this maximal period is achieved precisely when the connection polynomial $q(x)$ is a primitive polynomial (that is, any root of $q(x)$ is a generator for the multiplicative group of the Galois field with $|F|^k$ elements). These sequences are of great interest in part because they can be generated very efficiently, and in part because they have excellent randomness properties (cf. Chapter 8 of Lidl and Niederreiter's book [20]).

Theorem 3.1 *Let A be an m-sequence over the finite field F. Then A is a punctured de Bruijn sequence and has the shift and add property. Hence A is balanced, has the run property, and has ideal autocorrelations.*

4 Characterization of Shift and Add Sequences

Zierler [24] stated that the sequences over a finite field with the shift and add property are exactly the m-sequences. His proof is correct for sequences over a prime field $\mathbf{F}_p$, but is incorrect for sequences over non-prime fields. Gong, Di Porto, and Wolfowicz gave the first counterexamples [5]. Subsequently, Blackburn gave a correct characterization, which we now describe.

Let $A = (a_0, a_1, \cdots)$ be an m-sequence of span k with entries in a finite field $F = \mathbf{F}_{p^e}$ where p is a prime number. Then, as we have seen, there is a primitive element $\alpha \in \mathbf{F}_{p^{ek}}$ and a nonzero element $a \in \mathbf{F}_{p^{ek}}$ such that for every i we have

$$a_i = Tr_{p^e}^{p^{ek}}(a\alpha^i).$$

Conversely, every such sequence is an m-sequence. Note that the function $T(x) = Tr_{p^e}^{p^{ek}}(ax)$ is $\mathbf{F}_{p^e}$-linear. Every $\mathbf{F}_{p^e}$-linear function from $\mathbf{F}_{p^{ek}}$ to $\mathbf{F}_{p^e}$ is of the form $T(x) = Tr_{p^e}^{p^{ek}}(ax)$ for some constant $a \in \mathbf{F}_{p^{ek}}$, hence the m-sequences are exactly the sequences A of the form $a_i = T(\alpha^i)$ for some $\mathbf{F}_{p^e}$-linear function from $\mathbf{F}_{p^{ek}}$ to $\mathbf{F}_{p^e}$ and some primitive element $\alpha \in \mathbf{F}_{p^{ek}}$.

The fact that such a sequence A has the shift and add property can be easily proved from this representation. A shift of A by τ positions is the sequence whose ith element is $T(\alpha^{i+\tau})$. The term-by-term sum of A and this shift is the sequence whose ith term is

$$T(\alpha^i) + T(\alpha^{i+\tau}) = T((1+\alpha^\tau)\alpha^i).$$

If $1+\alpha^\tau \neq 0$, then $1+\alpha^\tau = \alpha^{\tau'}$ for some τ', so the ith term of the sum is $T(\alpha^{i+\tau'})$. That is, it is another shift of A. However, we have only used the $\mathbf{F}_p$-linearity of T, not the full $\mathbf{F}_{p^e}$-linearity. Thus we have the following theorem.

Theorem 4.1 *If $\alpha \in \mathbf{F}_{p^{ek}}$ is primitive and T is an $\mathbf{F}_p$-linear function from $\mathbf{F}_{p^{ek}}$ to $\mathbf{F}_p^e$, then the sequence whose ith element is $T(\alpha^i)$ has the shift and add property.*

Blackburn then showed that in fact every sequence with period $p^{ek}-1$ and entries in $\mathbf{F}_p^e$ that has the shift and add property can be written this way.

Theorem 4.2 *Let A be a sequence with period $p^{ek}-1$ and entries in $\mathbf{F}_p^e$. Suppose that A has the shift and add property. Then there exists a primitive element $\alpha \in \mathbf{F}_{p^{ek}}$ and an $\mathbf{F}_p$-linear function T from $\mathbf{F}_{p^{ek}}$ to $\mathbf{F}_p^e$ so that the ith element of A is $T(\alpha^i)$.*

Thus each pair (T,α) as in Theorem 4.2 gives rise to a sequence over $\mathbf{F}_{p^e}$ with the shift and add property. We next ask when a second such pair (S,β) gives rise to the same sequence. That is, when

$$T(\alpha^i) = S(\beta^i) \tag{4}$$

The function $Tr_p^{p^e} \circ T$ is $\mathbf{F}_p$-linear, so there exists an element $u \in \mathbf{F}_{p^{ek}}$ so that $Tr_p^{p^e} \circ T(a) = Tr_p^{p^{ek}}(ua)$. Similarly, there is a $v \in \mathbf{F}_{p^{ek}}$ so that $Tr_p^{p^e} \circ T(a) = Tr_p^{p^{ek}}(va)$. If equation (4) holds, then also

$$Tr_p^{p^{ek}}(u\alpha^i) = Tr_p^{p^{ek}}(v\beta^i)$$

for all i. As i varies from 0 to ∞, the sequence of values on the left hand side forms an m-sequence whose minimal polynomial is the minimal polynomial of α over $\mathbf{F}_p$. Similarly, the values on the right hand side give an m-sequence whose minimal polynomial is the minimal polynomial of β over $\mathbf{F}_p$. Thus α and β are Galois conjugates, $\beta = \alpha^{p^j}$ for some j. Since α is primitive, it follows that $T(x) = S(x^{p^j})$ for all x. Equivalently, $S(x) = T(x^{p^{ek-j}})$ for all x. Conversely, if T is any $\mathbf{F}_p$-linear function from $\mathbf{F}_{p^{ek}}$ to $\mathbf{F}_{p^e}$, then $S(x) = T(x^{p^{ek-j}})$ is also $\mathbf{F}_p$-linear, and the pair (S, α^{p^j}) gives rise to the same sequence as the pair (T,α). Since the various powers α^{p^j} are all distinct for $j = 0, 1, \cdots, ek-1$, each pair (T,α) is one of ek pairs that give rise to the same sequence. This allows us to count the sequences with the shift and add property.

Theorem 4.3 *There are*

$$\frac{(p^{ek}-1)(p^{ek}-p)\cdots(p^{ek}-p^{e-1})\varphi(p^{ek}-1)}{ek}$$

nonzero sequences over $\mathbf{F}_{p^e}$ *with period* $p^{ek}-1$ *and the shift and add property, where* φ *is Euler's function.*

Proof: There are $\varphi(p^{ek}-1)$ primitive elements α. Thus there are $\varphi(p^{ek}-1)/(ek)$ Galois conjugacy classes of these elements. We can choose T by first fixing a bases for $\mathbf{F}_{p^{ek}}$ and $\mathbf{F}_{p^e}$ over $\mathbf{F}_p$, then picking an e by ek matrix with entries in $\mathbf{F}_p$ and rank e. Such a matrix is uniquely determined by a choice of a nonzero first row, for which there are $p^{ek}-1$ choices, then a choice of a second row that is not in the $\mathbf{F}_p$-span of the first row, for wich there are $p^{ek}-p$ choices, and so on. This gives the desired count. □

We can further identify the sequences with the shift and add property that are punctured de Bruijn sequences by writing elements in terms of a basis $\beta_1,\cdots,\beta_e$ for $\mathbf{F}_p^e$ over $\mathbf{F}_p$. If $A=(a_0,a_1,\cdots)$ is any sequence with the shift and add property, then $a_i=T(\alpha^i)$ for some primitive element $\alpha\in\mathbf{F}_{p^{ek}}$ and $\mathbf{F}_p$-linear function T. We have

$$a_i=\sum_{j=1}^{e}T_j(\alpha^i)\beta_j,$$

where $T_j:\mathbf{F}_{p^{ek}}\rightarrow\mathbf{F}_p$ is $\mathbf{F}_p$-linear. Thus there exist elements $u_j\in\mathbf{F}_{p^{ek}}$ so that $T_j(x)=Tr_p^{p^{ek}}(u_jx)$. As was suggested to us by an anonymous referee, we can characterize the punctured de Bruijn sequences with the shift and add property in terms of the u_j.

Theorem 4.4 *Let A have the shift and add property. Then A is a punctured de Bruijn sequence if and only if*

$$V=\{u_j\alpha^i:0\le j<k,0\le i<e\}$$

is a basis for $\mathbf{F}_{p^{ek}}$ *over* $\mathbf{F}_p$.

Proof: The sequence A is a punctured de Bruijn sequence if and only if each nonzero k-tuple of elements of $\mathbf{F}_p^e$ occurs exactly once in each period of A, and the zero k-tuple does not occur. Since the period of A is $p^{ek}-1$, this is equivalent to each such k-tuple occurring at most once in A, and the zero k-tuple not occurring. Since A has the shift and add property, each k-tuple occurs at most once if the zero k-tuple does not occur. Indeed, if the shift and add property holds, then so does the shift and subtract property. If any k-tuple occurs twice, then we can shift A by the distance between the two occurrences, then subtract A from this shift to obtain an occurrence of the all zero k-tuple.

The all-zero k-tuple occurs if and only if for some n we have $a_n = a_{n+1} = \cdots = a_{n+k-1} = 0$. That is,

$$Tr_p^{p^{ek}}(u_j \alpha^{i+n}) = 0$$

for $0 \le j < k$ and $0 \le i < e$. The set

$$\alpha^n V = \{u_j \alpha^{i+n} : 0 \le j < k, 0 \le i < e\}$$

is a basis for $\mathbf{F}_{p^{ek}}$ over $\mathbf{F}_p$ if and only if V is. A linear function is zero on a basis if and only if it is identically zero. But the trace function is not identically zero. Thus, if V is a basis, then A is a punctured de Bruijn sequence.

Conversely, if

$$\sum_{i=0}^{e-1} \sum_{j=0}^{k-1} c_{ij} u_j \alpha^i = 0$$

with each c_{ij} in $\mathbf{F}_p$ and not all zero, then for any n,

$$\sum_{i=0}^{e-1} \sum_{j=0}^{k-1} c_{ij} Tr_p^{p^{ek}}(u_j \alpha^{i+n}) = 0.$$

That is, the $\mathbf{F}_p$-coordinates of all k-tuples satisfy a common linear relation. Hence not all nonzero values of k-tuples can occur and A is not a punctured de Bruijn sequence. □

5 FCSRs and AFSRs

A class of pseudo-random sequences that is analogous to LFSR sequences but is based on addition with carry was developed by the authors of this paper and independently by Couture and L'Ecuyer [2, 3, 13, 14, 15]. Let M be a positive integer, and identify the ring $\mathbf{Z}/(M)$ with the integers $\{0, 1, 2, \cdots, M-1\}$. Fix multipliers $q_1, q_2, \cdots, q_k \in \mathbf{Z}/(M)$, an initial state $a_0, a_1, \cdots, a_{k-1} \in \mathbf{Z}/(M)$ and an initial memory (or "carry") $t_{k-1} \in \mathbf{Z}$. The *multiply with carry sequence* or *feedback with carry shift register (FCSR) sequence* is the unique solution to the *with-carry linear recurrence*

$$a_j + Mt_j = q_1 a_{j-1} + q_2 a_{j-2} + \cdots + q_k a_{j-k} + t_{j-1}$$

for $j \ge k$. This means that the right side of the equation is to be computed as an integer $\sigma \in \mathbf{Z}$. Then a_j is the remainder after dividing σ by M, and t_j is the whole number quotient $\lfloor \sigma/M \rfloor = (\sigma - a_j)/M$. We write $a_j = \sigma \bmod M$ and $t_j = \sigma \operatorname{div} M$. This psuedo-random sequence has three descriptions which are parallel to those of the LFSR sequence. First, it is the output of a *feedback with carry shift register* or FCSR (see [15]). The *connection integer* associated with this FCSR is the number

$$q = q_0 + \sum_{i=1}^{k} q_i M^i \in \mathbf{Z},$$

where $q_0 = -1$. Second, it is the coefficient sequence in the M-adic expansion (cf. [9, 15]) of the rational number

$$u/q = a_0 + a_1 M + a_2 M^2 + \cdots \tag{5}$$

with denominator q and with numerator

$$u = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} M^j - tM^k. \tag{6}$$

The sequence is strictly periodic if and only if $-q \le u \le 0$. Third, in analogy with equation (3) the sequence may be expressed as

$$a_j = (a\delta^j \bmod q) \bmod M \tag{7}$$

where $\delta = M^{-1}$ is the inverse of M in $\mathbf{Z}/(q)$ and $a \in \mathbf{Z}/(q)$ is an element which depends on the initial state [9, 15]. This notation means that the quantity $a\delta^j \bmod q$ is represented as an integer in the range $\{0, 1, \cdots, q-1\}$ and then this integer is reduced modulo M.

For any initial value, the memory t will quickly enter a certain range $w^- \le t \le w^+$ (cf. [9, 15]) where it will remain thereafter. So an FCSR is a finite state machine and in particular, every FCSR sequence is eventually periodic. Its period is a divisor of the order of M modulo q and hence a divisor of $\varphi(q)$. (Here, φ denotes Euler's function; in particular, if p is prime then $\varphi(p) = p - 1$.) An FCSR sequence with maximal period $\varphi(q)$ is called an *ℓ-sequence*. A necessary and sufficient condition for the existence of an ℓ-sequence based on a given connection integer q is that q is a power of a prime, and M is a primitive root modulo q.

LFSR sequences and FCSR sequences admit a common generalization, the *algebraic feedback shift register (AFSR) sequences* [17]. A class of AFSR sequences is based on a triple (R, r, S), where R is an integral domain (that is, a ring with no zero divisors), $r \in R$, and S is a subset of R which is a complete set of representatives for $R/(r)$. An AFSR in this class is then determined by a choice of multipliers $q_0, q_1, \cdots, q_k \in R$ such that q_0 is invertible modulo r. The AFSR is a (not necessarily finite) state device whose

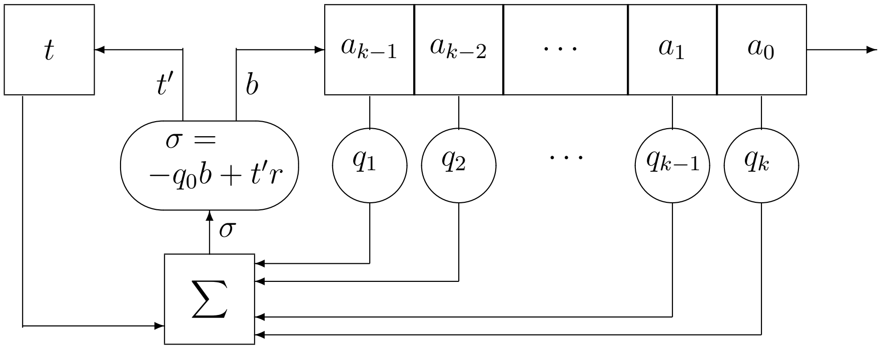

Figure 2: Algebraic Feedback Shift Register.

states are tuples $(a_0, a_1, \cdots, a_{k-1}; t)$ with each $a_i \in S$ (the "cell entries") and $t \in R$ (the "memory"). It changes states as follows. There are unique elements $a_k \in S$ and $t' \in R$ such that

$$-q_0 a_k + rt' = t + \sum_{i=1}^{k} q_i a_{k-i}. \tag{8}$$

(This fact is reproven below when R is a Euclidean domain.) Then the new state is $(a_1, a_2, \cdots, a_k; t')$. The resulting sequence $a_0, a_1, a_2, \cdots$ of elements in $R/(r)$ is called an AFSR sequence. We refer to equation (8) as a *linear recurrence with carry* over $R/(r)$. The element $q = \sum_{i=0}^{k} q_i r^i \in R$ is called the *connection element.* These ingredients may be expressed in terms of a (possibly infinite) state machine (see Figure 1) which is analogous to the LFSR and FCSR.

Even at this level of generality there is an analog to the power series representations (1) and (5). Let

$$R_r = \{\sum_{i=0}^{\infty} a_i r^i : a_i \in S, i = 0, 1, \cdots\}$$

be the *r-adic ring* of formal power series. There is a natural ring homomorphism from R to R_r which is one-to-one if

$$\bigcap_{i=1}^{\infty} (r^i) = (0) \tag{9}$$

that is, if no nonzero element of R is divisible by every power of r. This homomorphism extends to the set of fractions u/q with $u, q \in R$ and q invertible modulo r. We refer to the representation of an element u/q in R_r as its *r-adic expansion.* If equation (9) is satisfied, then this representation is unique. An AFSR sequence $a_0, a_1, \cdots$ with connection element

q is the sequence of coefficients in the r-adic expansion

$$u/q = \sum_{i=0}^{\infty} a_i r^i \tag{10}$$

of the fraction u/q where

$$u = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} r^j - t_{k-1} r^k. \tag{11}$$

The proof [17] of this fact is a calculation which goes back, originally, to the proof [4] of equation (2) in the case of LFSRs, to the proof [13, 15] of equation (6) in the case of FCSRs, and to [13] in the case of d-FCSRs.

The third expression for the AFSR sequence is a direct generalization of equations (3) and (7); see Theorem 3.1 and Theorem 10 of [17].

Theorem 5.1 *([17]) Let A be periodic. Let U denote the set of elements $v \in R$ such that v/q corresponds to a shift of A. Suppose no two elements of U are congruent modulo q and let V be a complete set of representatives modulo q containing U. Then*

$$a_i = v(sr^{-i} \bmod q) \bmod r, \tag{12}$$

for some $v \in R/(r)$ and $s \in R/(q)$. As in equation (7) this means that the element $sr^{-i} \in R/(q)$ is first lifted to the set V, then reduced modulo r, then multiplied by $v \in R/(r)$.

An LFSR over a field F is an AFSR with $R = F[x]$, $r = x$, $S = F$, $q_0 = 1$, each $q_i \in F$, and with initial memory $t = 0$. An FCSR is an AFSR with $R = \mathbf{Z}$, $r = M \in \mathbf{Z}$, $S = \{0, 1, \cdots, M-1\}$, $q_0 = 1$, and each $q_i \in S$. In both these cases the ring R is a Euclidean domain, so any element $\sigma \in R$ has a unique expression

$$\sigma = Ar + B \tag{13}$$

where $B \in S$, in which case we write $B = \sigma(\bmod r)$ and $A = \sigma(\operatorname{div} r)$. Therefore equation (8) may be rewritten

$$a_k = -\sigma \ (\bmod r) \text{ and } t' = \sigma \ (\operatorname{div} r) \tag{14}$$

where $\sigma = \sum_{i=1}^{k} q_i a_{k-i} + t \in R$. (This determines a_k since q_0 is invertible in $R/(r)$.) In these cases the memory remains within a certain finite set, so the AFSR in Figure 1 may be considered a finite state machine. With each clock cycle the entries in the cells shift one step to the right. The cell contents a_i may be thought of as elements of the ring $R/(r)$, but when computing the contents σ of the box Σ, (with each clock cycle) they should be thought of as elements of $S \subset R$. Then $a_k = -\sigma(\bmod r)$ is fed into the leftmost cell while $t' = (\sigma + a_k)/r$ is fed back into the memory.

There exist AFSRs (R, r, S) for which the output sequence $a_0, a_1, \cdots$ is aperiodic and for which the memory t does not remain bounded. The authors have studied several generalizations of the FCSR architecture, each of which may be described as an AFSR sequence for appropriate R, r, and S [6, 9, 10, 11, 12, 18, 19]. In many cases it is known that the resulting maximal length sequences have good correlation and distribution properties.

6 AFSRs Based on Polynomial Rings

Let F be a finite (Galois) field. Then there is a prime number p and an integer d such that $F \cong \mathbf{F}_{p^d}$. Let $R = F[x]$ be the polynomial ring in one variable, and let $r \in F[x]$ be a polynomial of some degree e. The division theorem for polynomial says that $F[x]$ is a Euclidean domain: for any polynomial $\sigma(x) \in F[x]$ there are unique polynomials $A(x), B(x)$ such that $\deg(B) < e$ and $\sigma(x) = A(x)r(x) + B(x)$. Let $S \subset F[x]$ be the collection of all polynomials of degree less than e, so $B(x) \in S$. The statement $B(x) = \sigma(x) \bmod r(x)$ (or simply, $B = \sigma \bmod r$) reflects the fact that the set S is a complete set of representatives for the quotient ring $F[x]/(r)$. The set S is closed under addition, but not under multiplication. For the remainder of this paper we study AFSR sequences based on (R, r, S).

Let $q(x) \in R = F[x]$ be a polynomial that is relatively prime to $r(x)$. Then the image of q in $R/(r)$ is invertible, and the image of r is invertible in $R/(q)$. Since R is a Euclidean domain we may write

$$q(x) = \sum_{i=0}^{k} q_i r^i \tag{15}$$

for some k, where $q_i \in S$ (for $i = 0, 1, \cdots, k$), where $q_k \neq 0$, and where q_0 is invertible modulo r. We consider the possible output sequences from an AFSR based on (R, r, S) with multipliers $q_0, q_1, \cdots, q_k \in S$. Thus, such a sequence is generated by the finite state machine illustrated in Figure 1.

To be pedantic, we use a paragraph to repeat the salient properties of the AFSR, in this special case. The machine has multipliers $(q_0, q_1, \cdots, q_k)$ and "state vector" $(a_0, a_1, \cdots, a_{k-1})$, where each $q_i, a_j \in R/(r) = F[x]/(r(x))$ may be identified with a polynomial in S of degree less than e. Given the initial state $(a_0, a_1, \cdots, a_{k-1})$ with initial memory $t_{k-1} \in F[x]$, the next state is computed from the linear recurrence with carry (8). That is, set $\sigma(x) = \sum_{i=1}^{k} q_i(x)a_{k-i}(x) + t_{k-1}(x) \in F[x]$. Then equation (8) can be rewritten

$$a_k = \gamma\sigma(\operatorname{mod} r) \;\text{ and }\; t_k = \frac{\sigma + q_0 a_k}{r} \tag{16}$$

where $\gamma = -q_0^{-1}(\operatorname{mod} r) \in F[x]/(r)$. By equations (10) and (11), the output sequence $a_0, a_1, \cdots$ is precisely the coefficient sequence of the r-adic expansion of the rational function

$$u(x)/q(x) = \sum_{i=0}^{\infty} a_i r^i \tag{17}$$

whose denominator $q(x)$ is determined as in equation (15) by the multipliers $q_0, q_1, \cdots, q_k$ and whose numerator

$$u = \sum_{j=0}^{k-1}\sum_{i=0}^{j} q_i a_{j-i} r^j - t_{k-1} r^k \tag{18}$$

is determined by the initial state vector $(a_0, a_1, \cdots, a_{k-1})$. (Conversely, every rational function $u(x)/q(x)$ with denominator q can be expressed uniquely as an r-adic number as in equation (17).)

Proposition 6.1 *Let $u(x) \in F[x]$. The coefficient sequence $A = (a_0, a_1, \cdots)$ of the r-adic expansion of $u(x)/q(x)$ in equation (17) is eventually periodic. It is strictly periodic if and only if the degree of u is less than the degree of q. In this case the (minimal) period of A is the multiplicative order of r modulo q, that is, the smallest positive integer N such that $r^N = 1$ in the finite (multiplicative) group $(R/(q))^*$ of invertible elements in $R/(q)$.*

Proof: If the state is $(a_{j-k}, \cdots, a_{j-1}; t_{j-1})$, then by equation (16), the degree of σ is at most $\max(2(e-1), \deg(t_{j-1}))$. The degree of $q_0 a_j$ is at most $2(e-1)$. Thus the quotient $t_j = (\sigma + q_0 a_j)/r$ has degree at most $\max(e-2, \deg(t_{j-1}) - e)$. Thus from any initial state with memory t_{k-1}, the degree of the memory decreases monotonically in at most $(\deg(t_{j-1}) - e + 2)/e$ steps until the degree of the memory is at most $e-2$, and this bound persists from then on. Thus A is eventually periodic.

Suppose that A is strictly periodic, say with period M. Then

$$\frac{u}{q} = (\sum_{i=0}^{M-1} a_i r^i)\sum_{i=1}^{\infty} r^{Mi} = \frac{\sum_{i=0}^{M-1} a_i r^i}{1 - r^M}.$$

The degree of the numerator in this last expression is strictly less than Me, the degree of the denominator. Thus the degree of u is less than the degree of q, which proves the first half of the first statement. Moreover, the equation

$$u(1 - r^M) = q \sum_{i=0}^{M-1} a_i r^i$$

implies that $r^M \equiv 1 \bmod q$, so the multiplicative order N of r divides the period M of A.

Conversely, suppose that $\deg(u) < \deg(q)$. Let N denote the multiplicative order of r modulo q, so $1 - r^N = sq$ for some polynomial s. It follows that $u/q = (su)/(1 - r^N)$, and $\deg(su) < Ne$. Thus we can write $su = \sum_{i=0}^{N-1} b_i r^i$ with $b_i \in S$. It follows that $a_j = b_{j \bmod N}$ for all j, so A is strictly periodic, of period N. In particular, the minimal period of A divides N. □

Corollary 6.2 *Given an AFSR with multipliers $q_0, q_1, \cdots, q_k$ and initial state vector $(a_0, a_1, \cdots, a_{k-1})$, there exists a value t of the memory such that the output sequence is strictly periodic. If $q_k \in F$ (that is, if $\deg(q_k) = 0$) then this value of t is unique. In this case, $t = 0$ if and only if*

$$\deg \sum_{i=0}^{k-1} a_i q_{k-i-1} \leq e - 1. \tag{19}$$

Proof: Given the initial state vector $(a_0, a_1, \cdots, a_{k-1})$ let us consider the effects of different values t of the memory on the degree of the polynomial $u(x)$ in equation (18). Let $H(x)$ denote the double sum in equation (18). By the division theorem for polynomials, there exists a unique polynomial $t \in F[x]$ such that

$$H(x) = t(x)r^k + J(x)$$

with $\deg(J) < \deg(r^k) = ke \leq \deg(q)$ since $q_k \neq 0$. Taking this $t = t_{k-1}$ for the memory gives a state of the AFSR whose output sequence is the r-adic expansion of u/q, where $u = H - tr^k = J$ has degree $< \deg(q)$. So by Proposition 6.1 the output sequence is strictly periodic. This proves that such a t always exists.

Now suppose q_k has degree 0. Then $\deg(q) = ek$. We wish to prove that t is unique. Given the initial state vector $(a_0, a_1, \cdots, a_{k-1})$ suppose there are two values, $t \neq t'$ for the memory such that the output sequence is strictly periodic. Let u, u' be the corresponding polynomials from equation (18). Then $\deg(u), \deg(u') < ek$ by Proposition 6.1. However $u - u' = (t' - t)r^k$ which has degree $\geq ek$ and this is a contradiction.

Now suppose equation (19) holds. The terms of highest degree in the double sum of equation (18) are

$$r^{k-1} \sum_{i=0}^{k-1} a_i q_{k-i-1}$$

which has degree

$$ek - e + \deg \sum_{i=0}^{k-1} a_i q_{k-i-1} < ek = \deg(q)$$

by assumption. However the term tr^k has degree ek. So any nonzero value for t will result in $\deg(u) \geq \deg(g)$ and the output sequence will fail to be strictly periodic, by Proposition 6.1. The converse is similar. □

It also follows from Proposition 6.1 that if u/q and v/q have periodic r-adic expansions, then u and v are not congruent modulo q. Hence by Theorem 5.1 there is an exponential representation for A.

Corollary 6.3 *If A is periodic, then*

$$a_i = v(sr^{-i} \bmod q) \bmod r, \tag{20}$$

for some $v \in R/(r)$ and $s \in R/(q)$. This means that the element $sr^{-i} \in R/(q)$ is first represented by an element of R with degree less than the degree of q, then reduced modulo r, then multiplied by $v \in R/(r)$.

Definition 6.4 *Let F be a finite field, and let $r, q \in F[x]$ be relatively prime. Denote by $\mathcal{S}_{r,q}$ the collection of sequences $A = (a_0, a_1, \cdots)$ of elements in $F[x]/(r)$ that are the coefficient sequences of the r-adic expansions of those rational functions $u(x)/q(x)$ such that* $\deg(u) < \deg(q)$.

7 Randomness of Polynomial Based AFSRs

Throughout this section $A = (a_0, a_1, \cdots)$ is an (r, q)-adic ℓ-sequence of the sort considered in Section 6. Thus $F = \mathbf{F}_{p^d}$ is a finite Galois field, $r(x) \in F[x]$ is a polynomial of degree e, and as in (15),

$$q(x) = \sum_{i=0}^{k} q_i(x) r(x)^i \tag{21}$$

is a polynomial of degree g which is relatively prime to $r(x)$. We may consider A to be a sequence of elements $a_i \in K = F[x]/(r)$.

According to Proposition 6.1 the period of the sequence A is the multiplicative order of r modulo q. This is maximal if $F[x]/(q)$ is a field (i.e., q is irreducible) and if r is a primitive element in this field (which is not the same as being a primitive polynomial in $F[x]$). To obtain a punctured de Bruijn sequence we also need to have a sequence of period $|K|^k - 1$ for some k, which implies that $|F|^{ek} = |F|^g$, or $k\deg(r) = \deg(q)$. By (21) we see that then $\deg(q_k) = 0$.

Definition 7.1 *Let $r, q \in F[x]$ be relatively prime, with degrees e and g respecitvely. A sequence $A \in \mathcal{S}_{r,q}$ is an (r, q)-adic ℓ-sequence if $g = ek$ for some integer k and if A has period $|F|^g - 1$, or equivalently, if q is irreducible and r is primitive modulo q.*

Such a sequence A is necessarily the r-adic expansion of a rational function $u(x)/q(x)$ with $\deg(u) < \deg(q) = ek$, and there are $|F|^{ek} - 1$ possible nonzero choices for u. Therefore, for any nonzero $u(x) \in F[x]$ with $\deg u < \deg q$ the coefficient sequence of the r-adic expansion of $u(x)/q(x)$ is a shift of the sequence A. Conversely, any shift of the sequence A is the coefficient sequence of the r-adic expansion of $u(x)/q(x)$ for some polynomial u with $\deg(u) < \deg(g)$.

Theorem 7.2 *Let $r, q \in F[x]$ be relatively prime with degrees e and $g = ek$ respectively, suppose q is irreducible and r is primitive modulo q. Then the resulting (r, q)-adic ℓ-sequence A is a punctured de Bruijn sequence. Consequently this sequence satisfies the first two of Golomb's randomness postulates.*

Proof: The sequence $A = (a_0, a_1, \cdots)$ is the output of an AFSR with multipliers $q_0, q_1, \cdots, q_k$. Suppose a string $b = (b_0, b_1, \cdots, b_{k-1})$ of length k occurs in A after some number of iterations. Consider the state of the AFSR at this point. The values $b_0, b_1, \cdots, b_{k-1}$ are the contents of the registers. By Corollary 6.2 there is a unique value t for the memory such that the output of the AFSR with this initial state vector b and initial memory t is a strictly periodic sequence. Since the sequence is, in fact, periodic from this point, the memory must have this value t. It follows that the string b can occur at most once in any period of A — otherwise the sequence would repeat upon the next occurrence of b, and its period would be less than $|F|^{ek} - 1$. However, there are $|F|^{ek}$ possible strings b, and the string $b = (0, 0, \cdots, 0)$ cannot occur in A (otherwise A would consist only of zeroes). Consequently every nonzero string b of length k occurs exactly once in a single period of A. □

In [4], Golomb used the shift-and-add property of binary m-sequences to show they have ideal autocorrelations, and thus satisfy his third randomness postulate. His methods have since been extended to many different sorts of pseudo-random sequences. Next we do the same for (r, q)-adic ℓ-sequences, verifying Golomb's third randomness postulate.

Theorem 7.3 *Let $r, q \in F[x]$ be relatively prime polynomials as in Section 6. Let A be an (r, q)-adic ℓ-sequence. Then A has the shift and add property with coefficients in F.*

Proof: The sequence A is the coefficient sequence in the r-adic expansion of some rational function $u(x)/q(x)$ where $\deg(u) < \deg(q)$. According to the comments preceding Theorem 7.2, for any integer τ, the τ-shift of A, A_τ, is the coefficient sequence for some rational function $u'(x)/q(x)$, where $\deg(u') < \deg(q)$. Let $v(x) = cu(x) + du'(x)$. Since $\deg(v) < \deg(q)$ the coefficient sequence for $v(x)/q(x)$, which is $cA + dA_\tau$, is a shift of the sequence A. □

Corollary 7.4 *The sequence A has ideal autocorrelations.*

Proof: By Theorem 7.2 and Lemma 2.5, A is balanced. By Theorem 7.3, A has the shift and add property. Thus by Theorem 2.11, A has ideal autocorrelatons.

□

Moreover, since every punctured de Bruijn sequence with the shift and add property arises from Blackburn's construction, we have the following corollary to Theorem 7.3.

Corollary 7.5 *There exists a primitive element $\alpha \in \mathbf{F}_{p^{dek}}$ and an $\mathbf{F}_p$-linear function $T : \mathbf{F}_{p^{dek}} \to \mathbf{F}_{p^d}[x]/(r)$ so that $a_i = T(\alpha^i)$.*

In fact this is also a corollary to Corollary 6.3. In this setting $\mathbf{F}_{p^d}[x]/(q)$ is isomorphic to $\mathbf{F}_{p^{dek}}$. In this field r is primitive so it plays the role of α. The function that maps a to $v(sa \bmod q) \bmod r$ is $\mathbf{F}_p$-linear, so this palys the role of T.

It is natural then to ask whether all punctured de Bruijn sequences with the shift and add property are (r, q)-adic ℓ-sequences. We believe that they are not. However, in another paper we consider AFSRs based on rings of the form $\mathbf{F}_{p^d}[x_1, \cdots, x_n]/I$ where I is an ideal. It is shown there that in fact all punctured de Bruijn sequences with the shift and add property are indeed ℓ-sequences in this setting.

8 Implementation Issues

For many applications it is essential that the pseudorandom sequences used be generated quickly. In this section we study the complexity of generating punctured de Bruijn sequences with the shift and add property.

Suppose we have such a sequence $A = (a_0, a_1, \cdots)$ over $\mathbf{F}_{p^e}$ with period $p^{ek} - 1$. We can realize A as $a_i = T(\alpha^i)$ where $T : \mathbf{F}_{p^{ek}} \to \mathbf{F}_{p^e}$ is $\mathbf{F}_p$-linear and α is a primitive element of $\mathbf{F}_{p^{ek}}$. Suppose that also A is an (r, q)-adic ℓ-sequence with $r, q \in \mathbf{F}_p[x]$, $\deg(r) = e$, $\deg(q) = ek$, and $q = \sum_{i=0}^{k} q_i r^i$ with $\deg(q_i) < e$, q_0 invertible modulo r, and $q_k = 1$.

We assume that r is a primitive element in $\mathbf{F}_p[x]/(q)$. Hence in particular $\mathbf{F}_p[x]/(q)$ is a field, so can be identified with $\mathbf{F}_{p^{ek}}$.

We think of addition and multiplication in $\mathbf{F}_p$ as atomic operations. For any n we let $M(n)$ denote the worst case time complexity of multiplication of polynomials over $\mathbf{F}_p$ of degree less than n. Then $M(n)$ is also the worst case time complexity of multiplication in $\mathbf{F}_{p^n}$. Using divide and conquer gives $M(n) \in O(n^{\log_2(3)})$. Using fast Fourier transforms gives $M(n) \in O(n \log(n))$. The worst case time complexity of addition in $\mathbf{F}_{p^n}$ is $O(n)$.

We compare three methods for generating punctured de Bruijn sequences over an $\mathbf{F}_p$-vector space F.

LFSR with Linear Output: We can use an LFSR with length ek and entries in $\mathbf{F}_p$, or an LFSR with length k and entries in $\mathbf{F}_{p^e}$ to generate powers of α and apply T to the successive states of the LFSR. In the first case the state change operation takes ek multiplications and $ek - 1$ additions in $\mathbf{F}_p$. The function T is realized by an ek by e matrix over $\mathbf{F}_p$, so takes e^2k multiplications and $e(e-1)k$ additions. Thus it takes a total of $2e^2k + O(ek)$ operations to generate one symbol of A.

In the second case, the state change takes k multiplications in $\mathbf{F}_{p^e}$. The cost of computing T is the same as in the previous paragraph since we have to interpret the state as a vector over $\mathbf{F}_p$ in general. Thus the cost of generating one symbol is $2e^2k + O(M(e)k)$, which is slightly worse.

Interleaving: By choosing a basis for $\mathbf{F}_{p^e}$ over $\mathbf{F}_p$, we can think of A as the interleaving of e m-sequences of span ek over $\mathbf{F}_p$. Each m-sequence can be generated by a LFSR of length ek with entries in $\mathbf{F}_p$. The state change for such an LFSR takes ek multiplications and $ek - 1$ additions in $\mathbf{F}_p$, and the output takes one operation (output the rightmost cell). Thus the total cost from all the LFSRs for generating one symbol of A is $2e^2k$. This is essentially the same complexity as in the previous case.

(r, q)**-Adic** ℓ**-Sequences:** We can generate A with an AFSR of length k based on $\mathbf{F}_p[x]$ and r with connection element q. The state change requires at most k multiplications of polynomials over F_p of degree less than e, plus $2k$ additions of polynomials over F_p of degree less than e. Then the total cost is $M(e)k + 2ek$.

The first and third methods can be sped up by precomputing tables for small chunks. E.g., in the first method think of a vector of length k as a vector of $k/8$ bytes of length 8 and precompute the inner products of all pairs ofl bytes. In the third method think of each polynomial of degree e as a sum of polynomials of degree less than 8 times appropriate powers of x^8 and precompute all products all pairs of polynomials of degree less than 8. This gives the same speedup for both methods.

It's possible that we can save some of the redundant work of the parallel LFSRs in the second method (all LFSRs are the same, they just have different start states). But this appears possible only if the phases of the LFSRs are close. Otherwise the storage costs become large.

In general all methods are faster in special cases. In the first method T may have many entries in $\mathbf{F}_p$ or even many zero entries. In the second method the LFSRs may have many zero coefficients or the phases may be close. In the third method the AFSR may have many zero coefficients or more generally the degrees of the coefficients $q_i(x)$ may be low. It is not clear to what extent we can force these things to happen.

If the sequence generation is to be implemented in software and $p = 2$, then we can speed up the second method as long as e is at most the word size (typically 32 bits or 64 bits). We use ek words and store the state of the first LFSR in the least significant bits of the words, the state of the second LFSR in the next least significant bits, and so on. Since the state change is the same for all LFSRs and the coefficients are zeros and ones, the new bit for each LFSR is computed as the exclusive or of some fixed set of state bits. Thus we can compute all the new bits simultaneously by taking the bitwise exclusive or of a fixed set of words. We then shift the words by one position. The total time required is apparently at most $2ek$ word operations. However, this analysis is not always correct. In some architectures the bitwise exclusive or of words is not actually implemented as an atomic operation in the hardware and its actual cost must be considered.

9 Relation with M-sequences

The (r, q)-adic ℓ-sequences shares many of the properties of m-sequences. In this section we show that, except in trivial cases, such a sequence A is never an m-sequence, and we give sufficient conditions to guarantee that A cannot be obtained from an m-sequence by a linear change of variable.

Let $F = \mathbf{F}_{p^d}$. Fix $r(x), q(x) \in F[x]$ relatively prime, of degrees e and $g = ek$ respectively, with q irreducible and with r primitive modulo q. The resulting (r, q)-adic ℓ-sequence has period $|F|^{ek} - 1$. Its entries are in the ring $K = F[x]/(r)$, which has p^{de} elements. By equation (12), there exists $s \in F[x]/(q)$ and $v \in F[x]/(r)$ so that the sequence is given by

$$a_i = v(sr^{-i} \bmod q) \bmod r, \tag{22}$$

appropriately interpreted, as in Theorem 5.1. (Different choices of s correspond to cyclic shifts of the sequence.) On the other hand, there exist m-sequences of the same period.

Let $E \cong \mathbf{F}_{p^{de}}$ be the degree e extension of the field $F \cong \mathbf{F}_{p^d}$. Let $L \cong \mathbf{F}_{p^{dek}}$ be the degree k extension of E. Let

$$q'(y) = -1 + q'_1 y + \cdots + q'_k y^k \in E[y]$$

be an primitive polynomial of degree k with coefficients $q'_i \in E$. That is, any root r' of q' is a primitive element of L). Let $Tr_{L/E} : L \to E$ denote the trace mapping and let $s' \in L$. Then the sequence $B = (b_0, b_1, \cdots)$ given by

$$b_i = Tr_{L/E}(s'(r')^{-i}) \tag{23}$$

is a typical m-sequence of span k (and period $|E|^k - 1$) with entries in E.

There exists an isomorphism of fields, $L \cong E[y]/(q')$, although there is not a unique such choice of isomorphism, or even a best such. If such a choice of isomorphism is made then r' may be interpreted as a polynomial with $\deg(r') = 1 < \deg(q')$ and equation (23) becomes

$$b_i = Tr_{L/E}(s'(r')^{-i} \bmod q'). \tag{24}$$

There also exists an isomorphism of fields $L \cong F[x]/(q)$ where $q(x)$ is given by (21), but there is no best such choice of isomorphism. In any case, equations (22) and (23) or (24) look very similar. One important difference between them is that the symbols $b_i \in E$ are in the field E while the symbols $a_i \in K = F[x]/(r)$ are in a ring. Both E and K are vector spaces over F of the same dimension, e. If $r(x) \in F[x]$ is chosen to be irreducible then K is also a field, isomorphic to E, but again there is no best choice of isomorphism.

In any case, one is led to the following questions: Given the (r, q)-adic ℓ-sequence $A = (a_0, a_1, \cdots)$, does there exist an m-sequence $B = (b_0, b_1, \cdots)$ with entries in E and a (set theoretic) mapping $\phi : K \to E$, so that $B = \phi(A)$? (This means that $b_i = \phi(a_i)$ for all i.) Such a mapping ϕ, if it exists, is necessarily a one to one correspondence. One might ask the same question, but requiring $\phi : K \to E$ to be an F-linear vector space isomorphism. We are unable at his time to answer the first question, but we next answer the second question in the negative. In fact, an m-sequence is just a special of an ℓ-sequence in which the parameter r' has degree 1. Thus we prove a more general result without this assumption, then specialize to m-sequences.

Theorem 9.1 *Let $r, q \in R = \mathbf{F}_{p^d}[x]$ with $\deg(r) = e$, $\deg(q) = g = ek$, and with r primitive modulo q. Let A be a sequence with connection element q so that A is an (r, q)-adic ℓ-sequence. Let $r', q' \in R' = \mathbf{F}_{p^{d'}}[x]$ with $\deg(r') = e'$, $\deg(q') = g' = e'k'$, and with r' primitive modulo q'. Let A' be a sequence with connection element q' so that A' is an (r', q')-adic ℓ-sequence. If $e \neq e'$ and $R'/(r')$ is a field, then A' is not the image of A by an $\mathbf{F}_p$-linear isomorphism with $\phi(1) = 1$.*

Proof: We have $|R/(r)| = p^{de}$, $|R'/(r')| = p^{d'e'}$, $|R/(q)| = p^{dek}$, and $|R'/(q') = p^{d'e'k'}$. Suppose there is an $\mathbf{F}_p$-linear isomorphism ϕ from $R/(r)$ to $R'/(r')$ that maps A to A' and satisfies $\phi(1) = 1$. Then $d'e' = de$. Moreover, A and A' must have the same period, so $d'e'k' = dek$, and hence $k' = k$. We define the function $\rho : R/(q) \to R/(r)$ by

$$\rho(\sum_{i=0}^{k-1} h_i r^i) = h_0$$

if each h_i is in $S = \{a : \deg(a) < e\}$. This function is not a homomorphism, but is $\mathbf{F}_{p^d}$-linear.

Let E be the set of periodic sates of the AFSR with connection element q, and let Γ be the state change operation on this AFSR. We define the function $\sigma : R/(q) \to E$ as follows. For any $h \in R/(q)$, there is a unique $t \in R$ so that $(\rho(h), \rho(r^{-1}h), \cdots, \rho(r^{1-k}h); t)$ is a periodic state. Let $\sigma(h) = (\rho(h), \rho(r^{-1}h), \cdots, \rho(r^{1-k}h); t)$. Then for any h we have $\sigma(r^{-1}h) = \Gamma(\sigma(h))$. Also, if $\nu : E \to R/(r)$ is the output function for the AFSR, then $\nu(\sigma(h)) = \rho(h)$. This is an *injective model* in the language of our earlier paper [8].

We have another such setup for the AFSR with connection element q': a function $\rho' : R'/(q') \to R'/(r')$; E', the set of periodic states of the AFSR; Γ', the state change operation; a function $\sigma' : R'/(q') \to E'$ defined by

$$\sigma'(h) = (\rho'(h), \rho'((r')^{-1}h), \cdots, \rho((r')^{1-k'}h); t')$$

for some t' and satisfying $\sigma'((r')^{-1}h) = \Gamma'(\sigma'(h))$ for any $h \in R'/(q')$, and $\nu'(\sigma'(h)) = \rho'(h)$ where $\nu' : E' \to R'/(r')$ is the output function for the AFSR.

There is also a function $\pi : E \to E'$ such that the output starting at state $\pi(\alpha) \in E'$ is the image under ϕ of the output starting at state $\alpha \in E$. Therefore $\pi(a_0, \cdots, a_{k-1}; t) = (\phi(a_0), \cdots, \phi(a_{k-1}); t')$ for some t'.

Define $\mu : R/(q) \to R'/(q')$ by $\mu = (\sigma')^{-1} \circ \pi \circ \sigma$. Then for every h, $\mu(r^{-1}h) = (r')^{-1}\mu(h)$, from which it follows that $\mu(rh) = r'\mu(h)$. By induction we have $\mu(r^i) = (r')^i\mu(1)$ for every i. Also, μ is $\mathbf{F}_p$-linear since each of the component functions is.

Let $S = \{u \in R : \deg(u) < e\}$ and let $S' = \{v \in R' : \deg(v) < e'\}$. Let $u \in S$. Then we have $\sigma(ur^{k-1}) = (0, \cdots, 0, \rho(u); t)$ for some t. Thus

$$\begin{aligned} \pi(\sigma(ur^{k-1})) &= (0, \cdots, 0, \phi(\rho(u)); t') \text{ for some } t'. \\ &= \sigma'(v(r')^{k-1}), \end{aligned} \tag{25}$$

where $v \in S'$ and $\rho'(v) = \phi(\rho(u))$ (such a v must exist). Thus $\mu(ur^{k-1}) = v(r')^{k-1}$. Suppose $u = 1$. Then $\rho'(v) = \phi(\rho(u)) = 1$ so $v = 1$ and $\mu(r^{k-1}) = (r')^{k-1}$. But $\mu(r^{k-1}) = (r')^{k-1}\mu(1)$, so $\mu(1) = 1$ and $\mu(r^i) = (r')^i$ for all i. For any nonzero $u, v \in$

$R/(q)$, we have $u = r^i$ and $v = r^j$ for some i, j (r is primitive), so $\mu(uv) = \mu(r^i r^j) = (r')^{i+j} = \mu(r^i)\mu(r^j) = \mu(u)\mu(v)$. Therefore μ is a field isomorphism. Also, it follows from the above discussion that if $u \in S$ and v is as in equation (25), then $\mu(ur^{k-1}) = v(r')^{k-1}$, so $\mu(u) = v$. In particular, $\mu(S) = S'$.

Now suppose that $d' > d$. Since d' divides dek, the field $\mathbf{F}_{p^{d'}}$ is contained in $\mathbf{F}_{p^{dek}} = R/(q)$. The former field is also contained in S'. By uniqueness, we have $\mathbf{F}_{p^{d'}} = \mu^{-1}(\mathbf{F}_{p^{d'}}) \subseteq S$. Therefore there exists $u \in S - \mathbf{F}_{p^d}$ such that every power of u is also in S. But any such u is a polynomial of degree at least 1, and some power of it has degree greater than or equal to e and less than ek. Such a power is not in S, a contradiction. Similarly, $d > d'$ is impossible. $\square$

Corollary 9.2 *Let* $r, q \in R = \mathbf{F}_{p^d}[x]$ *with* $\deg(r) = e > 1$, $\deg(q) = g = ek$, *and with* r *primitive modulo* q. *Let* A *be a sequence with connection element* q *so that* A *is an* (r, q)*-adic* ℓ*-sequence. Then* A *is not the* $\mathbf{F}_p$*-linear image of an m-sequence.*

Proof: Suppose that $A = \phi(A')$ (term by term) where A' is an m-sequence. Suppose that $\phi(b) = 1$. define $\phi'(c) = \phi(bc)$. Then ϕ' is an $\mathbf{F}_p$-linear isomorphism since by definition, the elements of an m-sequence are in a field. We have $\phi'(1) = 1$. The sequence B obtained by multiplying each element of A' by b^{-1} is an m-sequence since it satisfies the same linear recurrences as A'. Thus A is the $\mathbf{F}_p$ linear image of an m-sequence by an isomorphism that maps 1 to 1, contradicting Theorem 9.1. $\square$

In [5], Gong, Di Porto, and Wolfowicz constructed pseudo-noise sequences by applying an invertible $\mathbf{F}_p$-linear map to each element in an m-sequence over $\mathbf{F}_{p^f}$. Corollary 9.2 gives sufficient conditions that an (r, q)-adic ℓ-sequence cannot be so obtained.

10 Existence

It is not immediately apparent that (r, q)-adic ℓ-sequences that are not m-sequences are abundant. In order to find such sequences we fix the field $F = \mathbf{F}_{p^d}$ and search for a pair of polynomials $r, q \in F[x]$ such that q is irreducible and r is primitive modulo q. In order to get a de Bruijn sequence we will also require that $g = \deg(q)$ is a multiple of $e = \deg(r)$.

First recall the theorem of Pappalardi and Shparlinski [23]: Let $\overline{F}$ be an algebraic closure of F. Suppose r is not a k-th power of a function $h \in \overline{F}[x]$, for any k which divides $|F|^g - 1$. Then the number $N(r, F, g)$ of irreducible polynomials $q \in F[x]$ of degree g for which r is primitive satisfies

$$\left| N(r, F, g) - \frac{\varphi(M-1)}{g} \right| \leq 3eg^{-1} 2^{\nu(M-1)} \sqrt{M}$$

where $M = |F|^g$, where φ denotes Euler's φ function and where $\nu(k)$ denotes the number of distinct prime divisors of k. This implies the existence of many pairs (r, q) such that r is primitive mod q. For example, if $F = \mathbf{F}_2$ and $g = 13$ it says that for any $e \leq 42$ there exist r with $\deg(r) = e$ and r primitive mod q. If $g \geq 75$ then for every divisor e of g there exist polynomials r of degree e that are primitive mod g.

In fact, primitive polynomial pairs (r, q) are considerably more abundant than the above estimates predict. By computer search we have found the following for $F = \mathbf{F}_2$: Fix $g \leq 22$. Suppose $r \in F[x]$ is a polynomial of degree $e < g$ and suppose r is not a power of a polynomial $r \neq h^n$ where n divides g. Then there exists an irreducible polynomial q of degree g such that r is primitive mod q unless $r = x^4 + x$ and $g = 6$. In other words, there is a single unacceptable pair (r, g) in this range! (In this case, the above estimate says $|N(r, F, g) - 6| \leq 64$ so $N = 0$ is, indeed, a possibility.)

A class of examples which may be easily analyzed is the following. Let $q(x) \in F[x]$ be a primitive polynomial of degree $g = ke$. Let $r(x) = x^e$. Then r is primitive modulo q if and only if e is relatively prime to $|F[x]/(q)| - 1 = |F|^g - 1$. This is satisfied, for example, if g is relatively prime to $|F|^g - 1$. For example, if $F = \mathbf{F}_2$ and $r(x) = x^2$ we may take q to be any primitive polynomial of even degree. If such a q contains any terms of odd degree then some q_i has positive degree, so the resulting (r, q)-adic ℓ-sequence A is not an m-sequence. If $F = \mathbf{F}_2$ and $r(x) = x^3$ we may take q to be any primitive polynomial whose degree is an odd multiple of 3. If such a q contains any terms of degree not divisible by 3, then some q_i has positive degree, so the sequence A is not an m-sequence.

11 Example

In this section we let $p = 2$ and $d = 1$. If $\deg(r) = 1$, then we obtain m-sequences. The case $r(x) = x$ amounts to the standard analysis of m-sequences by power series. The case $r(x) = x + 1$ is equivalent by a change of basis.

Suppose that r has degree 2. Then for any choice of q we obtain sequences with elements in $K = \mathbf{F}_2[x]/(r) = \{0, 1, x, x+1\}$. If $r(x) = x^2 + x + 1$, which is irreducible over $\mathbf{F}_2$, we have $K = \mathbf{F}_4$, but for all other rs of degree two the ring K is not a field. If we let $r(x) = x^2 + x + 1$ and use the connection element $q(x) = x^4 + x^3 + 1 = r^2 + xr + x$, then it can be shown that r is primitive modulo q and one period of the (r, q)-adic ℓ-sequence A we obtain is a cyclic shift of

$$1, 1, x, x, x+1, x, 0, x, 1, x+1, x+1, 1, 0, x+1, 0. \tag{26}$$

All other (r, q)-adic ℓ-sequences obtained by different choices of r of degree 2 and q of degree 4 with r primitive modulo q are obtained from the sequence (26) by some combination of shifts, reversals, and permutations of the alphabet $\{0, 1, x, x+1\}$.

However, the sequence with one period equal to

$$1, 1, x, 1, 0, x+1, x+1, 1, x+1, 0, x, x, x+1, x, 0$$

is an m-sequence over $\mathbf{F}_4$, and all other m-sequences of span 2 over $\mathbf{F}_4$ are obtained from this sequence by some combination of shifts, reversals, and switching x and $x+1$. This illustrates the fact that the new set of sequences is disjoint from the set of m-sequences. In fact there is no set theoretic isomorphism $\phi : \mathbf{F}_4 \to \mathbf{F}_4$ so that $\phi(A)$ is an m-sequence, for the sequence A contains a string $x, 0, 1, 0, x$ and there is no analogous string in any of these m-sequences.

References

[1] S. Blackburn: A Note on Sequences with the Shift and Add Property. Designs, Codes, and Crypt. **9** (1996) pp. 251-256.

[2] R. Couture and P. L'Ecuyer, On the lattice structure of certain linear congruential sequences related to AWC/SWB generators, *Math. Comp.* **62** (1994), pp. 799–808.

[3] R. Couture and P. L'Écuyer, Distribution properties of multiply-with-carry random number generators. *Math. Comp.* **66**, pp. 591–607.

[4] S. Golomb, *Shift Register Sequences,* Aegean Park Press: Laguna Hills, CA, 1982.

[5] G. Gong, A. Di Porto, and W. Wolfowicz, Galois linear group sequences, *La Comm., Note Rec. Not.* **XLII** (1993) pp. 83-89.

[6] M. Goresky and A. Klapper, Feedback registers based on ramified extensions of the 2-adic numbers, *Advances in Cryptology - Eurocrypt 1994*. Lecture Notes in Computer Science **718**, Springer Verlag, New York, 1994, pp. 215-222.

[7] M. Goresky and A. Klapper, Arithmetic Cross-Correlations of FCSR Sequences, *IEEE Trans. Info. Theory.* **43** (1997) pp. 1342-1346.

[8] M. Goresky and A. Klapper: Fibonacci and Galois Mode Implementation of Feedback with Carry Shift Registers. IEEE Trans. Info. Thy. **48** (2002) 2826-2836.

[9] M. Goresky and A. Klapper, Efficient multiply-with-carry random number generators with optimal distribution properties, *ACM TOMACS* **13** (2003), pp. 1-12.

[10] M. Goresky and A. Klapper: Periodicity and Correlations of of d-FCSR Sequences. Designs, Codes, and Crypt. **33** (2004) 123-148.

[11] A. Klapper, Feedback with carry shift registers over finite fields, *Proceedings of Leuven Algorithms Workshop*. Lecture Notes in Computer Science **1008** Springer Verlag, New York, 1994, pp. 170-178.

[12] A. Klapper: Distributional properties of d-FCSR sequences. J. Complexity **20** (2004) 305-317.

[13] Klapper, A. and Goresky, M. 1993. Feedback shift registers, combiners with memory, and arithmetic codes. *Univ. of Kentucky Dept. of Comp. Sci. Tech. Rep. No. 239-93.*

[14] A. Klapper and M. Goresky, 2-adic shift registers. In *Fast Software Encryption, Cambridge Security Workshop, Cambridge UK, December, 1993*, R. Anderson, Ed. *Lecture Notes in Computer Science 809*, Springer Verlag, N.Y., pp. 174-178.

[15] A. Klapper and M. Goresky, Feedback Shift Registers, Combiners with Memory, and 2-Adic Span, *J. Cryptology* **10** (1997) pp. 111-147.

[16] A. Klapper and M. Goresky, Large period nearly de Bruijn FCSR sequences, *Advances in Cryptology - Eurocrypt 1995*. Lecture Notes in Computer Science **921**, Springer Verlag, New York, 1995, pp. 263-273.

[17] A. Klapper and J. Xu, Algebraic feedback shift registers, *Theoretical Comp. Sci.* **226** (1999) pp. 61-93.

[18] A. Klapper and J. Xu, Feedback with carry shift registers over $\mathbf{Z}/(N)$, *Proceedings of International Conference on Sequences and their Applications, Singapore, December 1998*. Springer Verlag, New York, to appear.

[19] A. Klapper and J. Xu: Register synthesis for algebraic feedback shift registers based on non-primes. Designs, Codes, and Crypt. **31** (2004) 227-25.

[20] R. Lidl and H. Niederreiter: Finite Fields, Encycl. Math. Appl. **20**. Addision Wesley, Reading, MA (1983).

[21] G. Marsaglia, The mathematics of random number generators, *The Unreasonable Effectiveness of Number Theory*, Amer. Math. Soc., Providence R. I., 1992. pp. 73-90.

[22] G. Marsaglia and A. Zaman, A new class of random number generators, *Annals of Appl. Prob.* **1** (1991) pp. 462-480.

[23] F. Pappalardi and I. Shparlinski, On Artin's conjecture over function fields, *Finite fields and their applications* **1** (1995), pp. 399–404.

[24] N. Zierler, Linear recurring sequences, *J. Soc. Indust. Appl. Math.* **7** (1959) 31-48.