Pseudo-noise Sequences based on Algebraic Feedback Shift Registers

Mark Goresky *Member* and Andrew Klapper *Senior Member*

***Abstract*— Over the past half century various statistical properties of pseudorandom sequences have played important roles in a variety of applications. Among these properties are Golomb's randomness conditions: (R1) balance, (R2) run property, and (R3) ideal autocorrelations, as well as the closely related properties (R4) shift and add, and (R5) de Bruin (uniform distribution of subblocks). The purpose of this paper is to describe the relations between these conditions, and to introduce a new method for generating sequences with all these properties, using algebraic feedback shift registers.**

***Index Terms*—Pseudo-random sequences, feedback shift registers, ideal autocorrelation, de Bruijn sequences, function fields.**

I. Introduction

The purpose of this paper is twofold: to review the basic properties of a class of pseudo-random sequences (punctured de Bruijn sequences with the shift and add property) over non-prime fields, and to describe their generation by algebraic feedback shift registers.

A. Generalities on sequences

Rapidly generated pseudo-random sequences with "good" statistical (randomness) properties are essential components in a wide variety of modern applications including radar, CDMA, error correction, cryptographic systems, and Monte Carlo simulations. Acceptable sequences should exhibit no statistical bias in the occurrence of individual symbols or small blocks of symbols. With these goals in mind, in his classic book S. Golomb [5] defined a *pseudonoise* sequence to be a periodic binary sequence that passes three statistical tests for randomness:

(R1) Balance,
(R2) Run property,
(R3) Ideal autocorrelation,

each of which is described below, in Section II. Golomb showed that linearly recurrent sequences of maximal period $2^r - 1$, or *binary m-sequences*, satisfy all three of these properties and moreover, such sequences can be rapidly generated using linear feedback shift registers (LFSRs).

It is still unknown whether there are any binary sequences, other than m-sequences, with all three of these properties. However there exist non-binary sequences, besides m-sequences, which pass these tests. Since 1982 a great many additional statistical tests for randomness have been studied (see, for example, [20] Chapter 3 or [22] Chapter 2). Although Golomb's list looks rather minimal by today's standards, it is an amazing fact that there are still only a handful of known techniques for constructing (nonbinary) sequences with all three of these properties. In Section VI of this paper we develop a new technique for the construction of sequences satisfying (R1)-R(3).

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, as we do in this paper, periodic sequences whose elements are taken from some finite dimensional vector space V over a (finite) Galois field F. The precise definitions of (R1)-(R3) in this setting are provided in Section II.

Following Golomb, let us consider the additional properties

(R4) shift and add (cf. Section II-B)
(R5) punctured de Bruijn (cf. Section II-A)

which are also enjoyed by m-sequences. Golomb showed, for binary periodic sequences, that condition (R1) (balance) together with condition (R4) (shift-and-add) is sufficient to guarantee condition (R3) (ideal correlation). The same holds, and the same proof works, in the non-binary setting (cf. Theorem 2).

Condition (R5) means: the period of A is $|V|^k - 1$ (for some k), and every block of length k occurs exactly once in each period of A except for the single block consisting of k zeroes; see Section II-A. The precise relation between conditions (R4) and (R5) is very interesting and it is still not completely understood. Even in the binary case, there exist punctured de Bruijn sequences that do not satisfy the shift-and-add condition. In [1], S. Blackburn (extending work of Gong, Di Porto and Wolfowicz [4]) gave a characterization of all sequences satisfying the shift-and-add property (R4). Using his result, in Theorem 3 we count the number of (cyclically) distinct shift-and-add sequences. Moreover we characterize (a) those shift-and-add sequences with ideal autocorrelations and (b) those shift-and-add sequences which also have the de Bruijn property (R5). (One such characterization, described in part (4) of Theorem 3 was suggested to us by an anonymous referee.) Despite these advances, we have not succeeded in counting the number of cyclically distinct shift-and-add sequences with the de Bruijn property.

B. The new sequences

Let F be a (finite) Galois field and let $q(x) \in F[x]$ be an irreducible polynomial of degree g. Then the quotient ring

M. Goresky is with the School of Mathematics, Inst. for Adv. Study, Princeton N.J. www.math.ias.edu/∼goresky. Research partially supported by DARPA grant HR0011-04-1-0031.

A. Klapper is with the Dept. of Computer Science, University of Kentucky, www.cs.uky.edu/∼klapper. Research partially supported by NSF grant #9980429.

$F[x]/(q)$ is a field, isomorphic to the Galois field with $|F|^g$ elements. Each element of $F[x]/(q)$ may be thought of as a polynomial of degree less than g. Let $r \in F[x]/(q)$ be a primitive element. Thus, $r = r(x) \in F[x]$ is a polynomial of degree less than g, and the various powers $r(x)^i \bmod q \in F[x]/(q)$ (for $0 \le i \le |F|^g - 1$) exactly account for the nonzero elements of $F[x]/(q)$. So we may also consider $r(x)^i \bmod q$ to be a polynomial of degree less than g. There are (at least) two things one might do with such a polynomial.

(i) Retain the constant term of the polynomial to obtain an element

$$a_i = (r^i \bmod q)(\bmod x) \in F \tag{1}$$

(ii) Reduce the polynomial modulo r to obtain an element

$$b_i = (r^i \bmod q)(\bmod r) \in F[x]/(r). \tag{2}$$

Let $e = \deg(r)$. If the polynomial $r(x)$ is irreducible then $F[x]/(r)$ is a field with $|F|^e$ elements. In general, all we can say is that $V = F[x]/(r)$ is a vector space over F of dimension e.

Using procedure (i), the resulting sequence $a_i \in F$ is an m-sequence, with period $|F|^e - 1$ and hence it satisfies conditions (R1)-(R5). The new sequences are those obtained from method (ii) when $g = \deg(q)$ is a multiple of $e = \deg(r)$. In Theorem 7 we prove, under these conditions, that the sequence (2) also satisfies conditions (R1)-(R5): it is a punctured de Bruijn sequence with the shift-and-add property and ideal autocorrelation. (For technical reasons, in the body of the paper, we consider the reverse sequence $(r^{-i} \bmod q)(\bmod r)$.)

One might ask whether the sequence (2) is perhaps just an m-sequence "in disguise". If $\hat{F}$ denotes the Galois field with the same number, $|F|^e$, of elements as $V = F[x]/(r)$, might there exist a vector space isomorphism between $\hat{F}$ and V which converts the sequence $b_i \in V$ into an m-sequence in $\hat{F}$? In Theorem 8 (in Section IX) we prove, in fact, that there exists *no set theoretic mapping* $\psi : V \to \hat{F}$ such that the sequence $\psi(b_i) \in \hat{F}$ is an m-sequence, provided $\deg(r) > 1$.

C. Algebraic feedback shift registers

Although the formula (2) gives a (relatively) explicit way to generate the sequence b_i, it is also possible to generate this sequence using an *algebraic feedback shift register* or AFSR. An AFSR is an LFSR that has been modified in two ways (cf. Figure 2):

1) The cell contents are allowed to be elements of a (fixed, finite) commutative ring S.
2) An additional "memory" or "carry" cell is incorporated in the feedback architecture.

The operation of the general AFSR is explained in detail in Section V. In Section VI we repeat this explanation for the special case in which the cell contents are elements in the ring $S = F[x]/(r)$ that contains the symbols b_i of equation (2). So Section VI describes the AFSR generation of the new sequences.

The AFSR construction is very general and it includes the case of LFSRs (where $S = \mathbf{F}_2$ and the memory cell is always 0) as well as the case of *feedback with carry shift registers*, or FCSRs, where $S = \mathbf{F}_2$ and the memory is an integer. The FCSR architecture, reviewed in Section V, and being a special case of the AFSR, is somewhat easier to understand than the general AFSR architecture, and its discovery predates the general AFSR. In a sequence of articles [7], [15]–[17] the authors have described the generation and analysis of maximal period FCSR sequences (which we refer to as ℓ-sequences): they exhibit many of the desirable randomness properties of m-sequences. The new AFSR sequences are, in some sense, a natural outgrowth of this line of investigation.

In Section VIII we estimate the cost of a software implementation of the AFSR architecture for the generation of the new sequences. But the real merit in having an AFSR description of the sequence, rather than an "exponential" implementation as in equation (2), lies in the possibility of implementing the generator in high speed hardware.

In Section X we show that parameters $r(x), q(x)$ giving rise to the new de Bruijn sequences are plentiful, although we have not succeeded in counting the number of distinct such sequences (for given period and symbol alphabet). Finally in Section XI we work out an example, which may help to clarify the discussions in Sections V and VII.

II. Pseudorandomness Properties of Sequences

In this section we describe, for nonbinary sequences, the randomness conditions (R1)-(R5) and the relations between them.

A. Distribution of blocks

Throughout this paper we fix a prime number p and let $\mathbf{F}_p$ denote the field with p elements. Let V be a vector space of dimension e over $\mathbf{F}_p$. Throughout this section we assume that A is a periodic sequence of elements from V, with period N. (There is no advantage in considering V to be a finite field of characteristic p, nor is any generality added by considering vector spaces over non-prime fields.)

Recall that a *block* $b = (b_0, b_1, \cdots, b_{k-1})$ of length k is an ordered sequence of k elements, $b_i \in V$. An *occurrence* of the block b in (a single period of) the sequence A is an index $i \le N - 1$ such that $(a_i, a_{i+1}, \cdots, a_{i+k-1}) = b$. A *run* of length k is a block of k consecutive identical symbols that is not contained in a longer block of consecutive symbols. That is, it is a block $(a_{i-1}, a, \cdots, a_{i+k})$ in A such that $a_{i-1} \ne a_i = a_{i+1} = \cdots = a_{i+k-1} \ne a_{i+k}$. The sequence A is a *de Bruijn sequence of span* k if every block of length k occurs exactly once in (each period of) A. The sequence A is a *punctured* de Bruijn sequence (R5) of span k if it is obtained from a de Bruijn sequence by deleting a single 0 (the zero vector of the vector space V) from the single occurrence of the block $(0, 0, \cdots, 0)$ of length k in each period of A. The period of a punctured de Bruijn sequence of span k is $N = |V|^k - 1$.

Suppose a sequence A of elements in V has period N. The sequence is said to satisfy the *balance* property (R1) if, for some integer t, within a single period every element $a \in V$ occurs t times or $t-1$ times. Thus we may take $t = \lceil N/|V| \rceil$. In particular, if $N = |V|^k$, then $t = |V|^{k-1}$ and every element $a \in |V|$ occurs t times. Similarly, if $N = |V|^k - 1$, then

$t = |V|^{k-1}$ and every element $a \in |V|$ except a single element occurs t times. The remaining element occurs $t-1$ times.

The sequence A satisfies the *run* property (R2) if it has period $N = |V|^k$ or $N = |V|^k - 1$ and if, for each $m \leq k-1$ the number of runs of length m is $|V|^{k-m-1}(|V|-1)^2$, the number of runs of length k is $|V|$ or $|V|-1$, and there are no runs of length greater than k. It is well known, and is easy to see that these are the closest integer approximations to the expected number of runs, averaged among all periodic sequences of period N. The original argument of [5] §4.2 shows:

Lemma 1: Every de Bruijn sequence A of span k, and every punctured de Bruijn sequence of span k, is balanced and has the run property.

In fact, such a sequence satisfies the following strong form of the balance condition:

(R1′) for any $t \leq k$ and for any block b of length t, the number of occurrences of b in (a single period of) a de Bruijn sequence A of span k is $|V|^{k-t}$. The same holds for a punctured de Bruijn sequence A except for the single block $(0, 0, \cdots, 0)$ of length t, which occurs $|V|^{k-t} - 1$ times.

a) Remark.: A choice of basis for V over $\mathbf{F}_p$ gives a way of translating each $a \in V$ into a block $\psi(a)$ 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 $\psi(A)$ does not, in general, have the same property (both because the period is wrong and because the relevant subblocks do not necessarily align with the ends of the $\mathbf{F}_p$-ary representations of elements.

B. Shift and add

Let V be a vector space over $\mathbf{F}_p$. Let $A = (a_0, a_1, \cdots)$ be a periodic sequence of elements from V 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 1: The sequence A has the *shift-and-add property* (R4) if, for any shift τ, either (1) $A + A_\tau = 0$ (the all-zeroes sequence) or (2) there exists a shift θ such that $A + A_\tau = A_\theta$.

Similarly we can define the shift and subtract property. More generally, we say that A satisfies the shift and add property with coefficients in the field $\mathbf{F}_p$ if, for any $c, d \in \mathbf{F}_p$ 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: The following statements are equivalent.

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

Proof: For any $v \in V$ the equation $pv = 0$ says that $-v = v + v + \ldots + v$ ($p-1$ times) so the shift and add property implies the shift and subtract property and vice versa. Similarly the shift and add property with coefficients in $\mathbf{F}_p$ follows from repeated application of the shift and add property. ■

It follows from part (3) of Lemma 2 that the set of shifts of a sequence A with the shift and add property, together with the all zero sequence, forms a vector space over $\mathbf{F}_p$. Such a vector space has cardinality p^n for some n. So if A is nonzero, the number of distinct cyclic shifts, and hence the period, equals $p^n - 1$. In particular a sequence with the shift and add property cannot be a de Bruijn sequence, but it might be a punctured de Bruijn sequence.

Suppose A is a punctured de Bruijn sequence over V (meaning that the symbols in the sequence are elements of V). Suppose $\hat{V}$ is another vector space and that $\phi : V \to \hat{V}$ is a (not necessarily linear) set theoretic mapping. Applying ϕ to each element of A gives a sequence $\phi(A)$ over $\hat{V}$. The sequence $\phi(A)$ will again be a punctured de Bruijn sequence if and only if (a) the vector space $\hat{V}$ has the same dimension as V, (b) the mapping ϕ is a one-to-one correspondence and (c) it satisfies $\phi(0) = 0$. The next theorem similarly characterizes those mappings $\phi : V \to \hat{V}$ which preserve both the punctured de Bruijn property and the shift-and-add property.

Theorem 1: Let V be a vector space over $\mathbf{F}_p$ and let $A = (a_0, a_1, \cdots)$ be a periodic sequence of elements in V. Suppose A is a punctured de Bruijn sequence with the shift-and-add property. Let $\hat{V}$ be another vector space over $\mathbf{F}_p$ and let $\phi : V \to \hat{V}$ be a set-theoretic mapping. Then the following conditions are equivalent:

1) The sequence $\phi(A)$ is a punctured de Bruijn sequence with the shift and add property.
2) The mapping $\phi : V \to \hat{V}$ is a (linear) isomorphism of vector spaces.

Proof: One implication is trivial: if ϕ is a linear isomorphism of vector spaces then $\phi(A)$ is a punctured de Bruijn sequence satisfying the shift and add property. Conversely, suppose $\phi : V \to \hat{V}$ is an arbitrary mapping such that $\phi(A)$ is a punctured de Bruijn sequence with the shift and add property. Since A and $\phi(A)$ have the same period, the mapping ϕ must be a one-to-one correspondence and $\dim(V) = \dim(\hat{V})$. In particular $\phi(0) = 0$ because 0 occurs in A (and in $\phi(A)$) fewer times than the other symbols. For every shift τ $(0 \leq \tau \leq N-1)$ there exists a unique shift $k = k(\tau)$ such that

$$\phi(A + A_\tau) - \phi(A) = \phi(A_{k(\tau)}). \tag{3}$$

This follows from the facts that A is a shift and add sequence and that $\phi(A)$ is a shift and subtract sequence. So for each i,

$$\phi(a_i + a_{i+\tau}) = \phi(a_i) + \phi(a_{i+k(\tau)}).$$

Suppose there exists an index τ such that $k(\tau) \neq \tau$. Then whenever i satisfies: $a_{i+\tau} = 0$ we obtain $\phi(a_{i+k(\tau)}) = 0$. In other words, if $a_\ell = 0$ then $a_{\ell+k(\tau)-\tau} = 0$.

The sequence A contains a unique largest block of zeroes (with $k-1$ zeroes, where k is the span of the de Bruijn sequence). Applying the above implication to each of these zeroes gives another (possibly overlapping) block of $k-1$ zeroes. This is a contradiction unless these two blocks coincide, meaning that $k(\tau) = \tau$. This combined with equation (3) proves that ϕ is linear, and so it is a linear isomorphism of vector spaces. ■

C. Autocorrelations

Golomb's third postulate (R3) is that a sequence should have an ideal autocorrelation function. The autocorrelation function of a sequence is usually defined for sequences whose symbols are taken from a cyclic group, whereas the sequences considered in this paper consist of symbols in some vector space V over the field $\mathbf{F}_p$. Before describing the appropriate notion of ideal autocorrelation in this context, we briefly review some standard facts about finite Abelian groups.

Definition 2: 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 χ necessarily takes values in the unit circle. It is said to be *nontrivial* if $\chi(a) \neq 1$ for some $a \in G$.

Lemma 3: 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 $a \in G$ with $\chi(a) \neq 1$. Then

$$\chi(a) \sum_{g\in G} \chi(g) = \sum_{g\in G} \chi(ag) = \sum_{g'\in G} \chi(g')$$

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

Definition 3: Let G be a finite Abelian group and let A be a periodic sequence of elements of G, with period N. Let χ be a character of G. The *autocorrelation* of A with shift τ, with respect to χ is the complex number

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

The sequence A has *ideal autocorrelations* if

(R3′) $|\mathcal{A}_{A,\chi}(\tau)| \leq 1$ for every nontrivial character χ of G and every $\tau \not\equiv 0 \bmod N$.

It is customary to consider the autocorrelation $\mathcal{A}_{A,\chi}$ to be a (periodic) function of τ.

Suppose that V is a vector space of dimension e over the field $\mathbf{F}_p$, and that $\chi : V \to \mathbf{C}^*$ is a character. Then χ satisfies $\chi(ax+by) = \chi(x)^a \chi(y)^b$ for any $x, y \in V$ and any $a, b \in \{0, 1, \ldots, p-1\}$. Moreover $\chi(px) = \chi(x)^p = 1$ for any $x \in V$, so χ takes values in the set

$$\mu_p = \{e^{2\pi i m/p} : 0 \leq m \leq p-1\}$$

of p-th roots of unity.

Theorem 2: If A is a periodic balanced sequence of elements taken from the vector space V (over the field $\mathbf{F}_p$) and if A has the shift and add property, then A has ideal autocorrelations in the sense of Definition 3.

Proof: Let N be the period of A. Let $\chi : V \to \mathbf{C}^*$ be a nontrivial character and let $\tau \in \mathbf{Z}$ be a shift. To compute the autocorrelation $\mathcal{A}_{A,\chi}(\tau)$, use Lemma 2 which says that A satisfies the shift-with-subtract property. So there exists a shift τ' with

$$\begin{aligned}\mathcal{A}_{A,\chi}(\tau) &= \sum_{i=0}^{N-1} \chi(a_i)\overline{\chi}(a_{i+\tau}) \\ &= \sum_{i=0}^{N-1} \chi(a_i - a_{i+\tau}) \\ &= \sum_{i=0}^{N-1} \chi(a_{i+\tau'}) \\ &= \sum_{i=0}^{N-1} \chi(a_i).\end{aligned}$$

Since A is balanced (§II), its period is either $|V|^k$ or $|V|^k \pm 1$ and each element $a \in V$ occurrs the same number, $|V|^{k-1}$ of times, except possibly for one single element. So the autocorrelation is

$$|V|^{k-1} \sum_{a\in V} \chi(a) + \varepsilon$$

where $\varepsilon = 0$ if the period $N = |V|^k$; otherwise $\varepsilon = \pm\chi(b)$ for a single element $b \in V$. By Lemma 3 the first term vanishes, leaving $|\mathcal{A}_{A,\chi}(\tau)| = |\varepsilon| \leq 1$ as claimed. ■

III. 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 (see Section IV for a review of the definition of m-sequences and related concepts). His proof is valid for sequences over a prime field $\mathbf{F}_p$, but it is incorrect for sequences over non-prime fields. Gong, Di Porto, and Wolfowicz gave the first counterexamples [4]. Subsequently, Blackburn gave a complete characterization of shift and add sequences [1]. In this section we describe and extend Blackburn's results.

Let V be a vector space of dimension e over $\mathbf{F}_p$. We consider periodic sequences of period $p^n - 1$ with entries in V. Let L be the Galois field with p^n elements. Let $\alpha \in L$ be a primitive element and let $T : L \to V$ be a set-theoretic mapping that is not identically 0. Let $A = (a_0, a_1, \ldots)$ be the sequence given by $a_i = T(\alpha^i)$.

We say that T is *balanced* if $n \geq e$ and if the set $T^{-1}(a)$ contains the same number, p^{n-e}, of elements, for every $a \in V$. If T is balanced, then it is surjective. If T is linear over $\mathbf{F}_p$, denote by $K = \ker(T)$ the kernel of T. If $u \in L$ then denote by

$$uK = \{ux \in L : x \in K\} = \{ux \in L : T(x) = 0\}$$

the translate of this subspace by the action of multiplication by u. We say that T has the *kernel property* if T is F_p-linear, if $n = ek$ for some k, and if

$$\bigcap_{i=0}^{k-1} \alpha^{-i} K = \{0\}. \tag{4}$$

If T has the kernel property then T is surjective (see the proof of part (3) of Theorem 3 in Section XII). In Theorem 3, we show that properties of the mapping $T : L \to V$ give rise to properties of the resulting sequence A according to the following table.

Properties of T	Properties of A
$\mathbf{F}_p$-linear	shift and add
balanced	ideal autocorrelations
kernel property	de Bruijn

In part (4) of Theorem 3 below, the kernel property is expressed in terms of a basis for V. We are grateful to an anonymous referee for suggesting part (4) of this theorem. In what follows, φ denotes Euler's function. We also recall that two periodic sequences of the same period are said to be *cyclically distinct* if the second sequence cannot be realized as a shift of the first sequence.

Theorem 3: Let $T : L \to V$ be a set-theoretic mapping that is not identically 0, where V is a vector space of dimension e over $\mathbf{F}_p$ and L is the field with p^n elements, as above. Fix a primitive element $\alpha \in L$. Let $A = (a_0, a_1, \cdots)$ denote the sequence $a_i = T(\alpha^i)$. Then the following statements hold.

(1) The mapping T is $\mathbf{F}_p$-linear if and only if the sequence A is a shift-and-add sequence, and in this case its (minimum) period is $p^n - 1$. There are

$$\frac{(p^{ne}-1)}{p^n-1}\frac{\varphi(p^n-1)}{n}$$

cyclically distinct non-zero sequences (of elements in V) with (minimum) period $p^n - 1$ which satisfy the shift and add property. Each of these arises from such a pair (T, α) (where $\alpha \in L$ is primitive and $T : L \to V$ is $\mathbf{F}_p$-linear).

(2) Suppose the mapping T is $\mathbf{F}_p$-linear. Then $T : L \to V$ is surjective if and only if it is balanced, which holds if and only if sequence A has ideal autocorrelations. There are

$$\frac{(p^n-p)(p^n-p^2)\cdots(p^n-p^{e-1})\varphi(p^n-1)}{n}$$

cyclically distinct shift-and-add sequences (of elements in V) with ideal autocorrelations and minimal period p^n-1. Each of these arises from such a pair (T, α) (where $\alpha \in L$ is primitive and $T : L \to V$ is $\mathbf{F}_p$- linear and balanced).

(3) Suppose $T : L \to V$ is $\mathbf{F}_p$-linear and surjective and let $n = ek$. Then T has the kernel property if and only if the sequence A is a punctured de Bruijn sequence (of elements in V, with minimal period $|V|^k - 1$, with ideal autocorrelations, which satisfies the shift-and-add property).

(4) Suppose the mapping T is $\mathbf{F}_p$-linear and suppose that $n = ek$ for some k. Choose a basis for V and write

$$T(x) = (T_1(x), T_2(x), \cdots, T_e(x)) \tag{5}$$

for the resulting coordinates of $T(x)$. Each $T_j : L \to \mathbf{F}_p$ is $\mathbf{F}_p$- linear so there exist (cf. Fact **(a)** in Section XII) unique elements $u_j \in L$ such that

$$T_j(x) = Tr^L_{\mathbf{F}_p}(u_j x). \tag{6}$$

Then the mapping T has the kernel property if and only if the following collection of ek elements

$$\{u_j \alpha^i : 1 \le j \le e,\ 0 \le i \le k-1\}$$

forms a basis for L over $\mathbf{F}_p$.

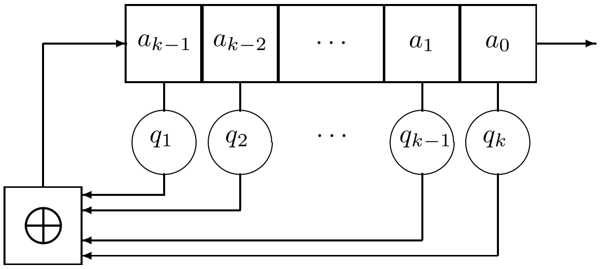

Fig. 1. A Linear Feedback Shift Register

The last statement of part (1) is Blackburn's characterization of sequences with the shift and add property [1]. The proof of Theorem 3 appears in Section XII.

Part (4) can be interpreted as saying that every sequence over a vector space V satisfying the shift-and-add and punctured de Bruijn properties can be obtained by selecting an m-sequence and a set of shifts that satisfy a certain linear-independence criterion and interlacing the shifted sequences.

b) Remark.: If $K = \ker(T)$ is preserved under multiplication by elements from the sub-field $F = \mathbf{F}_{p^e} \subset L$ then the kernel condition holds automatically. This occurs, for example, if $V = F = \mathbf{F}_{p^e}$ and if T is F-linear, in which case the resulting sequence is an m-sequence over F. More generally if $g \in Gal(L/\mathbf{F}_p)$ is an element of the Galois group and if $g(K)$ is preserved by multiplication by elements of F then T has the kernel property. The kernel property is somewhat mysterious and we do not know of a simple method for counting the number of linear mappings T with this property.

IV. Review of LFSRs and M-Sequences

In this section we review some basic properties of LFSRs to motivate the ensuing discussion of AFSRs.

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 $j \ge k$. 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.

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 ([5] Section 2.5) that the sequence $a_0, a_1, \cdots$ is also the coefficient sequence of the power series expansion

$$\frac{p(x)}{q(x)} = a_0 + a_1 x + a_2 x^2 + \cdots \tag{7}$$

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

$$p(x) = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} x^j. \tag{8}$$

Finally, there is an "exponential" representation for the sequence. Let $L = F[x]/(q(x))$. Elements of L can be represented as polynomials of degree less than $k = \deg(q)$. The polynomial $x \in L$ is invertible with

$$\alpha = x^{-1} = q_1 + q_2 x + \ldots + q_k x^{k-1}$$

since $x\alpha = 1 + q$. Let $\phi : L \to F$ be the F-linear mapping $\phi(h) = h \bmod x$, that is,

$$\phi(\sum_{i=0}^{k-1} h_i x^i) = h_0. \tag{9}$$

Then the sequence $a_0, a_1, \ldots$ is given by

$$a_j = \phi(p\alpha^j) = (p(x)x^{-j} (\bmod q))(\bmod x) \tag{10}$$

where $p(x)$ is given by equation (8). If $q(x)$ is irreducible then L may be identified with the unique field extension of F having degree d in such a way that α becomes identified with the inverse of a root of $q(x)$ or equivalently, α becomes identified with a root of the reciprocal polynomial $x^k q(1/x)$. In this case equation (10) becomes the more familiar

$$a_i = Tr_F^L (b\alpha^i)$$

for an appropriate choice of $b \in L$.

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 interest in part because they can be generated efficiently, and in part because they have the following randomness properties (cf. [21] Chapter 8):

Theorem 4: Let A be an m-sequence over the finite field F. Then A is a punctured de Bruijn sequence and it has the shift and add property. Hence A is balanced, has the run property, and has ideal autocorrelations.

V. FCSRs and AFSRs

A class of pseudo-random sequences that is analogous to LFSR sequences but is based on addition with carry was developed [13], [14], [16] by the authors of this paper and independently by Couture and L'Ecuyer [2], [3]. 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* $A = (a_0, a_1, \cdots)$ is the unique solution to the *with-carry linear recurrence*

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

for $j \geq k$. 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 pseudo-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 [16]). 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. [8], [16]) of the rational number

$$\frac{u}{q} = a_0 + a_1 M + a_2 M^2 + \cdots \tag{11}$$

with denominator q and with numerator

$$u = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} M^j - t_{k-1} M^k. \tag{12}$$

The sequence is strictly periodic if and only if $-q \leq u \leq 0$. Third, in analogy with equation (10) the sequence may be expressed as

$$a_j = (u\delta^j \bmod q) \bmod M \tag{13}$$

where $\delta = M^{-1}$ is the inverse of M in $\mathbf{Z}/(q)$ [8], [16]. This notation means that the quantity $u\delta^j \bmod q$ is represented as an integer in the range $\{0, -1, -2, \cdots, -q+1\}$ and then this integer is reduced modulo M.

For any initial value, the memory t will quickly enter a certain range $w^- \leq t \leq w^+$ (cf. [8], [16]) 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. 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]. Let R be an integral domain (that is, a ring with no zero divisors). Recall that two elements $v, w \in R$ are *relatively prime* if there exist elements $a, b \in R$ so that $av + bw = 1$, or equivalently, if v is invertible modulo w (or vice versa). Fix an element $r \in R$ and let $S \subset R$ be a complete set of representatives for the elements of $R/(r)$. A class of AFSRs is based on the triple (R, r, S). An AFSR in this class is 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 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

$$-q_0 a_k + rt' = t + q_1 a_{k-1} + q_2 a_{k-2} + \cdots + a_k a_0. \tag{14}$$

(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 (14) as a *linear recurrence with carry* over $R/(r)$. The element

$$q = \sum_{i=0}^{k} q_i r^i \in R \tag{15}$$

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 (7) and (11). 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

$$\bigcap_{i=1}^{\infty} (r^i) = (0) \tag{16}$$

that is, if no non-zero element of R is divisible by every power of r. This homomorphism extends to the set of fractions u/q (with $u, q \in R$) such that q is relatively prime to r. (By 15, this holds if and only if q_0 is relatively prime to r.) We refer to the representation of an element u/q in R_r as its *r-adic expansion*. If equation (16) is satisfied, then this representation is unique and we may unambiguously write

$$\mathbf{seq}_r(u/q) = (a_0, a_1, \cdots)$$

meaning that

$$\frac{u}{q} = \sum_{i=0}^{\infty} a_i r^i \in R_r. \tag{17}$$

The following theorem states that $\mathbf{seq}_r(u/q)$ is the output sequence of an AFSR with connection element q.

Theorem 5: Given (R, r, S) as above, with R an integral domain, $r \in R$, $S \subset R$ a complete set of representatives for $R/(r)$, such that (16) holds. Choose $q_0, q_2, \cdots, q_k \in S$ and set

$$q = \sum_{i=0}^{k} q_i r^i \in R.$$

Assume that the image of q_0 is invertible in $R/(r)$. For any $u \in R$ there exists unique elements $a_i \in S$ $(0 \le i \le k-1)$ and $t_{k-1} \in R$ such that

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

Then the output sequence of the AFSR with multipliers q_i $(1 \le i \le k)$ and initial state $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$ is the sequence $\mathbf{seq}_r(u/q) = a_0, a_1, \cdots$ of coefficients in the r-adic expansion (17) for the fraction u/q.

The proof [17] of this fact is a calculation which goes back, originally, to the proof [5] Section 2.5 of equation (8) in the case of LFSRs, to the proof [14], [16] of equation (12) in the case of FCSRs, and to the proof [14] in the case of d-FCSRs.

The third expression for the AFSR sequence is a direct generalization of equations (10) and (13); see Theorem 3.1 and Theorem 10 of [17].

Theorem 6: ([17]) Given (R, r, S) as in Theorem 5, choose "multipliers" q_0, q_1, $\cdots$, $q_k \in S$ so that q_0 is invertible in $R/(r)$. Let $q = \sum_{i=0}^{k} q_i r^i \in R$ be the resulting connection element and set $w = q_0^{-1} \in R/(r)$. Let $u \in R$ and suppose the sequence $\mathbf{seq}(u/q) = (a_0, a_1, \cdots)$ is strictly periodic. Let $V \subset R$ be a complete set of representatives for the elements of $R/(q)$ and assume that V contains the set

$$\{v \in R : \mathbf{seq}_r(v/q) \text{ is a shift of } \mathbf{seq}_r(u/q)\}.$$

Then

$$a_i = w(ur^{-i} \bmod q) \bmod r. \tag{19}$$

As in equation (13) this equation means that the element $ur^{-i} \in R/(q)$ is first lifted to the set V, then reduced modulo r, then multiplied by $w \in R/(r)$.

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

$$\sigma = Ar + B \tag{20}$$

where $B \in S$, in which case we write $B = \sigma(\bmod r)$ and $A = \sigma(\operatorname{div} r)$. Therefore equation (14) may be rewritten

$$a_k = \sigma \ (\bmod r) \text{ and } t' = \sigma \ (\operatorname{div} r) \tag{21}$$

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], [8]–[11], [18], [19]. In many cases it is known that the resulting maximal length sequences have good correlation and distribution properties.

VI. 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 polynomials 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

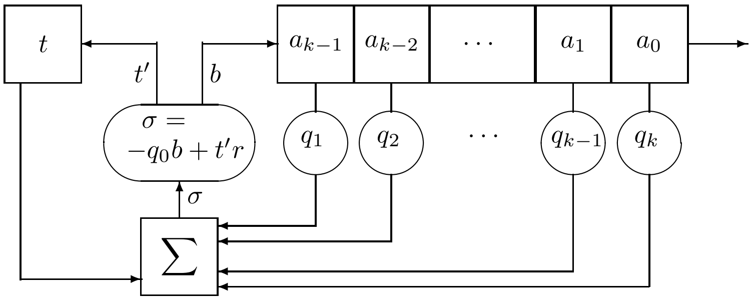

Fig. 2. Algebraic Feedback Shift Register.

$B(x) \in S$. 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 $(F[x], r, S)$.

Fix elements $q_0, q_1, \cdots, q_k \in S$ such that q_0 is invertible modulo r. Such a choice of multipliers gives rise to an AFSR with connection element

$$q(x) = \sum_{i=0}^{k} q_i r^i. \tag{22}$$

Then q is relatively prime to r. Conversely, if $q(x) \in R = F[x]$ is any polynomial that is relatively prime to $r(x)$, then, since $F[x]$ is a Euclidean domain, we may write in a unique way as $q(x) = \sum_{i=0}^{k} q_i r^i$ 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. Throughout this section we consider the possible output sequences from an AFSR based on $(F[x], r, S)$ with multipliers $q_0, q_1, \cdots, q_k \in S$.

We use a few paragraphs to repeat the salient properties of the AFSR in this special case. The elements of $F[x]/(r)$ are represented by polynomials of degree less than $e = \deg(r)$, the collection of which is denoted S. The AFSR sequence $A = (a_0, a_1, \cdots)$ is generated by the finite state machine as illustrated in Figure 1. The machine has fixed "multipliers" $(q_0, q_1, \cdots, q_k)$ where $q_i \in S$. The state $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$ consists of a "state vector" $(a_0, a_1, \cdots, a_{k-1})$ (with $a_i \in S$) and a "memory" $t_{k-1} \in F[x]$, that is, a polynomial of any degree. Given this initial state, the next state is computed from the linear recurrence with carry (14). 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 (14) can be rewritten

$$a_k = \gamma\sigma(\bmod r) \quad \text{and} \quad t_k = \frac{\sigma + q_0 a_k}{r} \tag{23}$$

where $\gamma = q_0^{-1}(\bmod r) \in F[x]/(r)$. By equations (17) and (18), the output sequence $A = \mathbf{seq}_r(u/q) = (a_0, a_1, \cdots)$ is precisely the coefficient sequence of the r-adic expansion of the rational function

$$\frac{u(x)}{q(x)} = \sum_{i=0}^{\infty} a_i r^i \tag{24}$$

whose denominator $q(x)$ is determined as in equation (22) by the multipliers $q_0, q_1, \cdots, q_k$ and whose numerator

$$u = \sum_{j=0}^{k-1} \sum_{i=0}^{j} q_i a_{j-i} r^j - t_{k-1} r^k \tag{25}$$

is determined by the initial state $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$. Conversely, every polynomial $u(x) \in F[x]$ corresponds to a unique state $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$ under equation (25), and the sequence $\mathbf{seq}_r(u/q)$ is precisely the output sequence of the AFSR.

Proposition 1: Let $u(x) \in F[x]$. The sequence $A = \mathbf{seq}_r(u/q)$ 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 $(F[x]/(q))^*$ of invertible elements in $F[x]/(q)$.

Proof: If the state is $(a_{j-k}, \cdots, a_{j-1}; t_{j-1})$, then by equation (23), 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 1: 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.

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 (25). Let $H(x)$ denote the double sum in equation (25). 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 $\deg(J) < \deg(r^k) = ke \le \deg(q)$ since $q_k \ne 0$. Taking this $t = t_{k-1}$ for the memory gives a state of the AFSR whose output sequence is $\mathbf{seq}_r(u/q)$, where $u = H - tr^k = J$ has degree $< \deg(q)$. So by Proposition 1 the output sequence is strictly periodic. This proves that such a t always exists.

Now suppose q_k has degree 0. Then $\deg(q) = ek$ since $q = \sum_{i=0}^{k} q_i r^i$. We wish to prove that the memory value t is unique. Given the initial state vector $(a_0, a_1, \cdots, a_{k-1})$ suppose there are two values, $t \ne t'$ for the memory such that the output sequence is strictly periodic. Let u, u' be the corresponding polynomials from equation (25). Then $\deg(u), \deg(u') < ek$ by Proposition 1. However $u - u' = (t' - t)r^k$ which has degree $\ge ek$ and this is a contradiction. ■

Corollary 2: Consider an AFSR with multipliers $q_0, q_1, \cdots, q_k$ and $\deg(q_k) = 0$. Suppose $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$ is a (strictly) periodic state of the AFSR, and let $u \in F[x]$ be the corresponding element defined by equation (25). Then $\deg(u) < ek$ and we may write in a unique way, $u = \sum_{i=0}^{k-1} u_i r^i$ with $\deg(u_i) < e$. Then

1) $a_i = 0$ for $0 \le i \le k-2$ if and only if $u_i = 0$ for $0 \le i \le k-2$.
2) The memory vanishes, $t_{k-1} = 0$, if and only if

$$\deg(\sum_{i=0}^{k-1} a_i q_{k-i-1}) \le e - 1. \qquad (26)$$

Proof: First suppose $a_i = 0$ for $0 \le i \le k-2$. By equation (25),

$$u = q_0 a_{k-1} r^{k-1} - t_{k-1} r^k = (q_0 a_{k-1} - t_{k-1} r) r^{k-1}.$$

If v denotes the polynomial within the parentheses, then $\deg(v) + e(k-1) = \deg(u) < ek$ which gives $\deg(v) < e$. In other words, $v = u_{k-1}$. The converse is a bit harder. Suppose $u = vr^{k-1}$ and $\deg(v) < e$. Then $u \bmod r^m = 0$ for $1 \le m \le k-1$. By equation (25), $u \bmod r = q_0 a_0 = 0$ which implies that $a_0 = 0$. Then by equation (25) again, $u \bmod r^2 = q_0 a_1 r^1 = 0$ which implies $a_1 = 0$. Continuing in this way we obtain $a_i = 0$ for $0 \le i \le k-2$.

Now suppose equation (26) holds. The terms of highest degree in the double sum of equation (25) 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 non-zero value for t_{k-1} will result in $\deg(u) \ge \deg(g)$ and, by Proposition 1 the output sequence will fail to be strictly periodic. The converse is similar. ■

$$\begin{tikzcd}
F[x]/(q) \arrow[loop above, "S"] \arrow[rr, "i"] \arrow[dr, "\phi"'] & & \Sigma \arrow[loop above, "T"] \arrow[dl, "\mathtt{OUT}"] \\
& F[x]/(r) &
\end{tikzcd}$$

Fig. 3. Algebraic model for AFSR

c) Notation: Let $\mathcal{S}_{r,q}$ denote the collection of all strictly periodic AFSR sequences based on $(F[x], r, S)$ with connection element q. By Proposition 1, $\mathcal{S}_{r,q}$ is the collection of all coefficient sequences $\mathbf{seq}_r(u/q)$ of the r-adic expansions of fractions $u(x)/q(x)$ such that $\deg(u) < \deg(q)$.

In Theorem 6 we may take V to be the set of all polynomials $u \in F[x]$ such that $\deg(u) < \deg(q)$. Then (cf. Theorem 6) no two elements of V are congruent modulo q so there is an exponential representation for every sequence in $\mathcal{S}_{r,q}$.

Corollary 3: Let $q = \sum_{i=0}^{k} q_i r^i \in F[x]$ be the connection element of an AFSR, where $q_i \in S$ and where q_0 is invertible in $F[x]/(r)$. Let $w = q_0^{-1} \in F[x]/(r)$. Fix $u \in F[x]$ and let $A = \mathbf{seq}_r(u/q) = (a_0, a_1, \cdots)$ be the resulting sequence. Assume that $A \in \mathcal{S}_{r,q}$ is strictly periodic. Then

$$a_i = w(ur^{-i} \bmod q) \bmod r, \qquad (27)$$

for all i.
Equation (27) means that the element $ur^{-i} \in F[x]/(q)$ is first represented by an element of $F[x]$ with degree less than $\deg(q)$, then reduced modulo r, then multiplied by $w \in F[x]/(r)$.

Corollary 3 can be expressed by saying that the diagram in Figure 3 commutes. Here, Σ denotes the set of all periodic states of the AFSR. The function $\mathtt{OUT} : \Sigma \to F[x]/(r)$ is the output function which assigns to a state $(a_0, a_1, \cdots, a_{k-1}; t_{k-1})$ the contents a_0 of the rightmost cell. The mapping $T : \Sigma \to \Sigma$ is the state change mapping. On the left side of the diagram the mapping $\phi : F[x]/(q) \to F[x]/(r)$ is given by $\phi(h) = wh(\bmod r)$, where $w = q_0^{-1} \in F[x]/(r)$. That is, $\phi(q_0 \sum_{i=0}^{k-1} z_i \pi^i) = z_0$ if each z_i has degree less than e. The mapping $S : F[x]/(q) \to F[x]/(q)$ is multiplication by r^{-1}. Finally the mapping $i : F[x]/(q) \to \Sigma$ assigns to any $u \in F[x]/(q)$ the state given by equation (25).

VII. Maximal Length AFSR Sequences

Throughout this section $A = (a_0, a_1, \cdots) \in \mathcal{S}_{r,q}$ is a strictly periodic AFSR sequence of the sort considered in Section VI.

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 equation (22),

$$q(x) = \sum_{i=0}^{k} q_i(x) r(x)^i \tag{28}$$

is a polynomial of degree g which is relatively prime to $r(x)$, and $A = \mathbf{seq}_r(v/q)$ is the AFSR sequence corresponding to some $v \in F[x]$ with $\deg(v) < \deg(q)$. We may consider A to be a sequence of elements $a_i \in K = F[x]/(r)$.

According to Proposition 1 the period of the sequence A is the multiplicative order of r modulo q. This is greatest 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 the sequence to have period $|K|^k - 1$ for some k, which implies that $|F|^{ek} = |F|^g$, or $k\deg(r) = \deg(q)$. By equation (28) we see that $\deg(q_k) = 0$.

Definition 4: The 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.

The sequence $A = \mathbf{seq}_r(v/q)$ is the coefficient sequence of the r-adic expansion of a rational function $v(x)/q(x)$ with $\deg(v) < \deg(q) = ek$. The period of A is $|F|^{ek} - 1$ which coincides with the number of non-zero polynomials $u \in F[x]$ such that $\deg(u) < \deg(q)$. Therefore, for any such u, the sequence $\mathbf{seq}_r(u/q)$ 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: 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. Let $v \in F[x]$ with $\deg(v) < \deg(q)$. Then the resulting (r,q)-adic ℓ-sequence $A = \mathbf{seq}(v/q)$ is a punctured de Bruijn sequence and it satisfies the shift and add property with coefficients in F. Consequently this sequence satisfies all three 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 block $b = (b_0, b_1, \cdots, b_{k-1})$ 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 1 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 block 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 blocks b, and the block $b = (0, 0, \cdots, 0)$ cannot occur in A (otherwise A would consist only of zeroes). Consequently every non-zero block b of length k occurs exactly once in a single period of A. Hence A is a punctured de Bruijn sequence.

According to the comments preceding Theorem 7, for any shift τ, there exists $u \in F[x]$ whose degree is less than $\deg(q)$ such that $A_\tau = \mathbf{seq}_r(u/q)$. Let $c, d \in F$. Then the sequence $cA + dA_\tau = \mathbf{seq}_r((cv + du)/q)$. But $\deg(cv + du) < \deg(q)$ so this sequence is some other shift of A, or else it is zero. Hence A has the shift and add property. By Theorem 7 and Lemma 1, A is balanced. Thus by Theorem 2, A has ideal autocorrelatons. ■

Since by Theorem 3 every punctured de Bruijn sequence defined over a finite vector space and having the shift and add property arises from Blackburn's construction, we have the following corollary to Theorem 7.

Corollary 4: 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. Let $v \in F[x]$ with $\deg(v) < \deg(q)$. Let $A = \mathbf{seq}(v/q)$ be the resulting (r,q)-adic ℓ-sequence. Then 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 also follows from Corollary 3. In this setting $F[x]/(q)$ 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 $w(ua \bmod q) \bmod r$ is $\mathbf{F}_p$-linear, so this plays 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 a separate paper [12] the second author considered AFSRs based on rings of the form $\mathbf{F}_{p^d}[x_1, \cdots, x_n]/I$ where I is an ideal. It was shown there that in fact all punctured de Bruijn sequences with the shift and add property are indeed ℓ-sequences in this setting.

VIII. 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 of 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 products of 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.

IX. Relation with M-Sequences

The (r,q)-adic ℓ-sequences share 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 $q=\sum_{i=0}^{k} q_i r^i$ irreducible, with $\deg(q_i)<e=\deg(r)$, and with r primitive modulo q. In particular, as observed in Section VII, $\deg(q_k)=0$. A choice of $u\in F[x]/(q)$ corresponds to an initial state of the AFSR and the resulting output sequence $A=\mathbf{seq}_r(u/q)=(a_0,a_1,\cdots)$ is an (r,q)-adic ℓ-sequence with period $p^{dek}-1$. Now suppose we have an m-sequence $\hat{A}=(\hat{a}_0,\hat{a}_1,\cdots)$ of the same period with symbols drawn from an alphabet of the same size. It is most convenient to describe the sequence $\hat{A}$ as an AFSR sequence (with memory equal to zero) as in Section IV. Let $\hat{F}=\mathbf{F}_{p^{de}}$ be the field containing the symbols $\hat{a}_i$. The m-sequence $\hat{A}$ satisfies a linear recurrence of degree k over $\hat{F}$, corresponding to a primitive polynomial $\hat{q}\in\hat{F}[y]$. Let $\hat{\phi}:\hat{F}[y]/(q)\to\hat{F}$ be the mapping $\hat{\phi}(\hat{h})=\hat{h}\bmod y$. That is,

$$\hat{\phi}\left(\sum_{i=0}^{k-1}\hat{h}_i y^i\right)=\hat{h}_0$$

(where $\hat{h}_i\in\hat{F}$) as in equation (9). Then $y\in\hat{F}[y]/(\hat{q})$ is primitive and invertible, and up to a shift, the sequence $\hat{A}$ is given by $\hat{a}_j=\phi(y^{-j}\bmod\hat{q})$, as in equation (10).

Theorem 8: If $e=\deg(r)>1$ then there does not exist any set-theoretic mapping $\psi:\hat{F}\to F[x]/(r)$ such that $\psi(\hat{A})=A$.

Proof: Suppose such a mapping exists. By Lemma 1 the mapping ψ is $\mathbf{F}_p$-linear. As in Corollary 3 let $\phi:F[x]/(q)\to F[x]/(r)$ be the mapping $\phi(h)=wh(\bmod r)$ where $w=q_0^{-1}\in F[x]/(r)$. We claim there exists a unique mapping $\Psi:\hat{F}[y]/(\hat{q})\to F[x]/(q)$ so that Figure 4 "commutes".

In Figure 4, the triangles on the ends are just repeats of Figure 3, where Σ represents the set of strictly periodic states of the first AFSR (and similarly for $\hat{\Sigma}$). Each periodic state of $\hat{\Sigma}$ is uniquely determined by the contents $(\hat{a}_0,\hat{a}_1,\cdots,\hat{a}_{k-1})$ of the registers (and memory $\hat{t}_{k-1}=0$) of the right hand AFSR, which must therefore be mapped by ψ to $(a_0,a_1,\cdots,a_{k-1})$ of the left AFSR. By Corollary 1 this determines a unique value t_{k-1} for the memory of the left AFSR. So there is a uniquely determined bijection $\hat{\Sigma}\to\Sigma$ which commutes with the mappings $T,\hat{T},\phi,\hat{\phi}$, $\texttt{OUT}$, and ψ. Using i and $\hat{i}$ (each of which is a bijection) this mapping becomes a bijection $\Psi:\hat{F}[y]/(\hat{q})\to F[x]/(q)$.

$$\begin{tikzcd}
\Sigma \arrow[loop above, "T"] \arrow[dr, "\mathtt{OUT}"'] & F[x]/(q) \arrow[l, "i"'] \arrow[loop above, "S"] \arrow[d, "\phi"] & \hat{F}[y]/(\hat{q}) \arrow[l, "\Psi"'] \arrow[loop above, "\hat{S}"] \arrow[d, "\hat{\phi}"] \arrow[r, "\hat{\imath}"] & \hat{\Sigma} \arrow[loop above, "\hat{T}"] \arrow[dl, "\mathtt{OUT}"] \\
 & F[x]/(r) & \hat{F} \arrow[l, "\psi"'] &
\end{tikzcd}$$

Fig. 4. Comparing two AFSRs

Next we modify Ψ slightly so as to obtain an isomorphism of fields. First note that Ψ is $\mathbf{F}_p$ linear since ψ is $\mathbf{F}_p$-linear. Moreover, for any $\hat{v} \in \hat{F}[y]/(\hat{q})$, the following equation holds: $\Psi(\hat{S}(\hat{v})) = S(\Psi(\hat{v}))$ or

$$\Psi(y^{-1}\hat{v}) = r^{-1}\Psi(\hat{v}) \tag{29}$$

from which it follows that $\Psi(y\hat{v}) = r\Psi(\hat{v})$ for every $\hat{v}$ (because $(y^{-1})^N = y \in \hat{F}[y]/(\hat{q})$ and $(r^{-1})^N = r \in F[x]/(q)$ where $N = p^{ked} - 2$). The only problem is that $\Psi(1)$ may fail to equal 1.

Define $\Phi : \hat{F}[y]/(\hat{q}) \to F[x]/(q)$ by $\Phi(\hat{v}) = \Psi(1)^{-1}\Psi(\hat{v})$. Then $\Phi(y^i) = r^i$ so Φ is both multiplicative and additive. Hence Φ is an isomorphism of fields. To see that this leads to a contradiction we consider the subfield $\hat{F} \subset \hat{F}[y]/(\hat{q})$ which may be realized as the collection of all polynomials of degree 0 (and with coefficients in $\hat{F}$). For any $\hat{s} \in \hat{F}$ consider the element $\hat{s}y^{k-1} \in \hat{F}[y]/(\hat{q})$. The corresponding periodic state

$$(\hat{b}_0, \hat{b}_1, \cdots, \hat{b}_{k-1}) = \hat{\imath}(\hat{s}y^{k-1}) \in \hat{\Sigma}$$

has $\hat{b}_i = 0$ for $0 \le i \le k-2$ by Corollary 2. Therefore the corresponding state $i(\Psi(\hat{s}y^{k-1})) = (b_0, b_1, \cdots, b_{k-1}; t_{k-1}) \in \Sigma$ has $b_i = 0$ for $0 \le i \le k-1$ also. Again by Corollary 2 this implies that $\Psi(\hat{s}y^{k-1}) = b_{k-1}r^{k-1}$ for some $b_{k-1} \in F[x]$ of degree less than e. It follows from equation (29) that $\Psi(\hat{s}) = b_{k-1}$. That is, Ψ maps $\hat{F}$ bijectively to the collection of all polynomials of degree less than e.

Since Φ is a field isomorphism, the set

$$\Phi(\hat{F}) = \left\{\Psi(1)^{-1}b(x) : \ \deg(b) < e\right\} \subset F[x]/(q)$$

is a subfield of $F[x]/(q)$. Taking $b(x) = 1$ shows that $\Psi(1)^{-1} \in \Phi(\hat{F})$ so we even conclude that

$$\Phi(\hat{F}) = \{b(x) : \ \deg(b) < e\}$$

is a subfield of $F[x]/(q)$. But this set is not closed under multiplication unless $e = 1$, which is a contradiction. ■

In [4], Gong, Di Porto, and Wolfowicz constructed pseudonoise sequences by applying an invertible $\mathbf{F}_p$-linear map to each element in an m-sequence over $\mathbf{F}_{p^f}$. Theorem 8 gives sufficient conditions that an (r, q)-adic ℓ-sequence cannot be so obtained.

X. 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 punctured de Bruijn sequence we 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

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

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 $e \le 42$ there exist r with $\deg(r) = e$ and r primitive mod q. If $g \ge 75$ 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 $g \le 22$. Suppose $r \in F[x]$ is a polynomial of degree $e < g$ and suppose r is not a power of a polynomial $r \ne h^n$ 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 $|N(r, F, g) - 6| \le 64$ 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.

XI. 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

$$1, 1, x, x, x+1, x, 0, x, 1, x+1, x+1, 1, 0, x+1, 0. \quad (30)$$

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 (30) 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. By Theorem 8 there is no set theoretic isomorphism $\phi : \mathbf{F}_4 \to \mathbf{F}_4$ so that $\phi(A)$ is an m-sequence.

XII. Proof of Theorem 3

We need to recall several standard facts before giving the proof of Theorem 3. Suppose $L = \mathbf{F}_{p^d}$ is a finite field of characteristic p and degree d.

(a.) If $A, B : L \to \mathbf{F}_p$ are non-zero $\mathbf{F}_p$ linear mappings then there exits a unique non-zero element $u \in L$ such that $B(x) = A(ux)$ for all $x \in L$.
(b.) If $p(x)$ is an irreducible polynomial with coefficients in $\mathbf{F}_p$, whose degree equals the degree of L (over $\mathbf{F}_p$), and if $\alpha, \beta \in L$ are roots of this polynomial, then they are Galois conjugate and there exists an integer m such that $\beta = \alpha^{p^m}$.

Lemma 4: Suppose $\alpha, \beta \in L$ are primitive elements. Suppose $A : L \to \mathbf{F}_p$ is a non-zero $\mathbf{F}_p$-linear mapping. Define the mapping $B : L \to \mathbf{F}_p$ by $B(0) = 0$ and

$$B(\beta^i) = A(\alpha^i)$$

for $0 \le i \le |L| - 2$. Then B is $\mathbf{F}_p$-linear if and only if α and β are Galois conjugates.

Proof: There exists t such that $\alpha = \beta^t$. Therefore $B(\beta^i) = A(\alpha^i) = A(\beta^{ti})$ so $B(x) = A(x^t)$ for all $x \in L$. If α and β are Galois conjugates then t is a power of p by **(b)** above, so the mapping $x \mapsto x^t$ is $\mathbf{F}_p$-linear. Therefore B is $\mathbf{F}_p$-linear. Conversely, suppose B is $\mathbf{F}_p$-linear. Let $p(x) = \sum_{i=0}^{d-1} a_i x^i$ be an irreducible polynomial with coefficients $a_i \in \mathbf{F}_p$ such that $p(\alpha) = 0$. We need to show that $p(\beta) = 0$. It suffices to show that $B(\beta^t p(\beta)) = 0$ for all $t \ge 0$. But

$$\begin{aligned} B(\beta^t p(\beta)) &= \sum_{i=0}^{d-1} a_j B(\beta^{t+j}) \\ &= \sum_{i=0}^{d-1} a_j A(\alpha^{i+j}) \\ &= A(\alpha^t p(\alpha)) \\ &= 0. \end{aligned}$$

So by Fact **(b)** above, α and β are Galois conjugate. ■

d) Proof of part (1).: Suppose T is $\mathbf{F}_p$ linear. If τ is a shift with $0 \le \tau < p^n - 1 = |L| - 1$, then

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

Since α is primitive, there exists θ with $1+\alpha^\tau = \alpha^\theta$. Therefore $a_i + a_{i+\tau} = a_{i+\theta}$, so A is a shift and add sequence. Let $R : V \to \mathbf{F}_p$ be a non-zero $\mathbf{F}_p$-linear mapping. Then the composition $RT : L \to \mathbf{F}_p$ is $\mathbf{F}_p$-linear so the sequence $RT(\alpha^i) = R(a_i) \in \mathbf{F}_p$ is an m-sequence and has minimum period $p^n - 1$. Hence the sequence A has minimum period $p^n - 1$ also. The converse is due to Blackburn who proved [1] the remarkable fact that for any shift and add sequence $A = (a_0, a_1, \ldots)$ with entries in V and with period $p^n - 1$, there exists a pair (T, α) (with $\alpha \in L$ primitive and $T : L \to V$ a non-zero $\mathbf{F}_p$-linear mapping), such that $a_i = T(\alpha^i)$.

To count the number of shift and add sequences we first count the number of pairs (T, α) where $\alpha \in L$ is a primitive element and $T : L \to V$ is $\mathbf{F}_p$-linear. Then we determine when two such pairs define the same sequence.

The number of primitive elements $\alpha \in L$ is $\varphi(p^n - 1)$. To count the number of $\mathbf{F}_p$-linear mappings $T : L \to V$ choose bases for both, as vector spaces of dimension n and e respectively, over $\mathbf{F}_p$. Each linear mapping T then corresponds to a unique $n \times e$ matrix with entries in $\mathbf{F}_p$, and there are p^{ne} such matrices. So there are $p^{ne} - 1$ non-zero linear mappings T.

Now consider decomposing the collection of pairs (T, α) into equivalence classes, with two pairs belonging to the same class if the resulting sequences are the same. We show that each class contains exactly n pairs by showing that (T, α) and (S, β) belong to the same class if and only if α and β are Galois conjugates. Hence $\beta = \alpha^{p^t}$ for some t, and $S(x^{p^t}) = T(x)$. In other words, the mapping S is uniquely determined by T, α, and β.

Suppose two pairs (T, α) and (S, β) give rise to the same sequence. That is,

$$S(\beta^i) = T(\alpha^i) \quad (31)$$

for all i. Then the images of S and T coincide. Let $v \ne 0 \in V$ be in the image of S and T. Choose any linear mapping R from V to F_p such that $R(v) \ne 0$. Then R is surjective and both compositions RT and RS are non-zero.

Now we have the equation $RT(\alpha^i) = RS(\beta^i)$, for all i. Since both RT and RS are non-zero $\mathbf{F}_p$-linear mappings, Lemma 4 implies that α and β are Galois conjugates with $\beta = \alpha^{p^t}$ for some t. Therefore $S(x^{p^t}) = T(x)$ by equation (31).

Conversely, given (T, α) let $\beta \in L$ be a Galois conjugate to α with $\beta = \alpha^{p^t}$. Let $S(x^{p^t}) = T(x)$. Then $S : L \to V$ is $\mathbf{F}_p$-linear and $S(\beta^i) = T(\alpha^i)$ for all i. Consequently (T, α) and (S, β) give rise to the same sequence.

To summarize, there are $p^{ne}-1$ choices for T and $\varphi(p^n-1)$ choices for α. Such a pair determines a class consisting of the n Galois conjugates β of α, and uniquely determined $\mathbf{F}_p$-linear mappings S to go with them. This counts the total number of sequences; the cyclically distinct sequences are counted by dividing by the period, $p^n - 1$. This completes the proof of part (1) of Theorem 3.

e) Proof of part (2).: If T is balanced then it is surjective. If T is surjective then it is balanced because $T^{-1}(a)$ is the set of solutions to a system of inhomogeneous linear equations, which is therefore a translate of the set $T^{-1}(0)$. Theorem 2 implies that A has ideal autocorrelations.

On the other hand, suppose that A has ideal autocorrelations, but T is not surjective. (Since $T \neq 0$ this implies that $k > 1$.) Let $I \subset V$ denote the image of the mapping $T : L \to V$. Choose a complementary subspace $J \subset V$ so that $V \cong I \oplus J$. Let $\chi_1 : J \to \mathbf{C}^*$ be any nontrivial character, and define $\chi : V \to \mathbf{C}^*$ to by $\chi(a, b) = \chi_1(b)$. Then χ is a character of V, and $\chi(T(x)) = 1$ for all $x \in L$. Let $\tau \neq 0$ be a non-zero shift. The autocorrelation of shift τ with respect to the character χ is

$$\sum_{i=0}^{|L|-1} \chi(T(\alpha^i - \alpha^{i+\tau})) = \sum_{i=0}^{|L|-1} 1 = p^n - 1$$

which is greater than 1. This is a contradiction, hence T is surjective.

To count the number of sequences with ideal autocorrelations, the same argument as in the proof of part (1) works, but we must count only those pairs (T, α) such that T has rank equal to k. Choosing bases for V and L over $\mathbf{F}_p$, the mapping T may be represented as an $n \times e$ matrix of elements of $\mathbf{F}_p$. The matrices of rank k are counted by choosing the first row to be any non-zero vector ($p^n - 1$ choices), the second row to be any vector that is not in the span of the first vector ($p^n - p$ choices), the third row to be any vector that is not in the span of the first two vectors ($p^n - p^2$ choices), and so on. As in the proof of part (1) above, this counts the total number of sequences; the cyclically distinct sequences are counted by dividing this number by the period, $p^n - 1$. This completes the proof of part (2).

f) Proof of part (3).: Consider the mapping $\Phi : L \to V^k$ given by $\Phi(x) = (T(x),\ T(\alpha x),\ \cdots,\ T(\alpha^{k-1}x))$. These k symbols form a block of the sequence A. Therefore A is a (punctured) de Bruijn sequence of rank k if and only if the mapping Φ is surjective, that is, if every non-zero k-tuple of vectors in V appears at some point in the sequence. Since $|L| = |V|^k$, the mapping Φ is surjective if and only if it is injective. But the kernel of Φ is exactly the intersection in equation (4). This completes the proof of part (3). This also shows that if T has the kernel property then T is surjective (because Φ is surjective).

g) Proof of part (4).: Having chosen a basis for V, the mapping $\Phi : L \to V^k$ of the preceding paragraph may be expressed as a $k \times e$ matrix $\Phi(x) = [T_j(\alpha^i x)]$ with $0 \leq i \leq k-1$ and $1 \leq j \leq e$. Using equation (6) this becomes the matrix $\Phi(x) = [Tr^L_{\mathbf{F}_p}(u_j \alpha^i x)]$. This mapping is an isomorphism if and only if the collection of linear functions

$$L_{ij}(x) = Tr^L_{\mathbf{F}_p}(u_j \alpha^i x),$$

with $0 \leq i \leq k-1$ and $1 \leq j \leq e$, forms a basis of the dual space $L^* = \mathrm{Hom}_{\mathbf{F}_p}(L, \mathbf{F}_p)$. However, the collection of vectors $\{u_j \alpha^i\}$ is linearly independent (and hence forms a basis of L) if and only if the collection of linear functions L_{ij} is linearly independent. This completes the proof of Theorem 3. ■

Acknowledgement

The authors would like to thank an anonymous referee for her or his many helpful comments on an earlier draft of this paper, and for suggesting part 4 of Theorem 3 to us. The first author is grateful to the Institute for Advanced Study for its hospitality and support.

References

[1] S. Blackburn, “A note on sequences with the shift and add property,” *Designs, Codes, and Crypt.*, vol. 9, pp. 251–256, 1996.

[2] R. Couture and P. L'Ecuyer, “‘on the lattice structure of certain linear congruential sequences related to awc/swb generators,” *Math. Comp.*, vol. 62, pp. 799–808, 1994.

[3] R. Couture and P. L'Écuyer, “Distribution properties of multiply-with-carry random number generators,” *Math. Comp.*, vol. 66, pp. 591–607, 1997.

[4] A. D. P. G. Gong and W. Wolfowicz, “Galois linear group sequences,” *La Comm., Note Rec. Not.*, vol. XLII, pp. 83–89, 1993.

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

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

[7] ——, “Arithmetic cross-correlations of fcsr sequences,” *IEEE Trans. Info. Theory*, vol. 43, pp. 1342–1346, 1997.

[8] ——, “Efficient multiply-with-carry random number generators with optimal distribution properties,” *ACM TOMACS*, vol. 13, pp. 1–12, 2003.

[9] ——, “Periodicity and correlations of d-fcsr sequences,” *Designs, Codes, and Cryptography*, vol. 33, pp. 123–148, 2004.

[10] A. Klapper, “Feedback with carry shift registers over finite fields,” in *Proceedings of Leuven Algorithms Workshop*, ser. Lecture Notes in Computer Science, vol. 1008. New York: Springer Verlag, 1994, pp. 170–178.

[11] ——, “Distributional properties of d-fcsr sequences,” *J. Complexity*, vol. 20, pp. 305–317, 2004.

[12] ——, “Randomness and register synthesis for afsrs based on function fields,” in *Sequences and Their Applications - SETA 2004*, ser. Lecture Notes in Computer Science, H.-Y. S. T. Helleseth, D. Sarwate and K. Yang, Eds., vol. 3486. New York: Springer Verlag, 2005, pp. 282–297.

[13] A. Klapper and M. Goresky, “2-adic shift registers,” in *Fast Software Encryption, Cambridge Security Workshop, Cambridge UK, December, 1993*, ser. Lecture Notes in Computer Science, R. Anderson, Ed., vol. 809. New York: Springer Verlag, 1993, pp. 174–178.

[14] ——, “Feedback shift registers, combiners with memory, and arithmetic codes,” Dept. of Computer Science, University of Kentucky, Tech. Rep., 1993.

[15] ——, “Large period nearly de bruijn fcsr sequences,” in *Advances in Cryptology - Eurocrypt 1995*, ser. Lecture Notes in Computer Science, vol. 921. New York: Springer Verlag, 1995, pp. 263–273.

[16] ——, “Feedback shift registers, combiners with memory, and 2-adic span,,” *J. Cryptology*, vol. 10, pp. 111–147, 1997.

[17] A. Klapper and J. Xu, “Algebraic feedback shift registers,” *Theoretical Comp. Sci.*, vol. 226, pp. 61–93, 1999.

[18] ——, “Feedback with carry shift registers over $\mathbf{z}/(n)$,” in *Proceedings of International Conference on Sequences and their Applications, Singapore, December 1998*. New York: Springer Verlag, 1999.

[19] ——, "Register synthesis for algebraic feedback shift registers based on non-primes," *Designs, Codes, and Crypt.*, vol. 31, pp. 227–25, 2004.

[20] D. Knuth, *The Art of Computer Programming, Volume 2: Seminumerical Algorithms.* Reading, MA: Addison-Wesley, 1997.

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

[22] H. Niederreiter, *Random Number Generation and Quasi-Monte Carlo Methods*. Philadelphia PA: SIAM, 1992.

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

[24] N. Zierler, "Linear recurring sequences," *J. Soc. Indust. Appl. Math.*, vol. 7, pp. 31–48, 1959.

PLACE
PHOTO
HERE

Mark Goresky received his Ph.D. in Mathematics from Brown University in 1976. He has held academic positions at the Massachusetts Institute of Technology, the University of British Columbia, and Northeastern University as well as visiting positions at various universities and research institutes in France, Germany, Italy, Canada and the USA. He is currently a "member" at the Institute for Advanced Study in Princeton NJ.

Dr. Goresky was awarded an Alfred P. Sloan Postdoctoral Fellowship in 1980. He is a Fellow of the Royal Society of Canada and he was awarded the Jeffrey-Williams prize of the Canadian Mathematical Society in 1986. Together with his colleague Robert MacPherson, he received the Leroy P. Steele Prize for a Seminal Contribution to Research, from the American Mathematical Society in 2002. Dr. Goresky is an associate editor for both the Journal and the Bulletin of the American Mathematical Society. His interest in pseudo-random sequences is supplemented by his mathematical interests in representation theory and automorphic forms.

PLACE
PHOTO
HERE

Andrew M. Klapper was born in White Plains, New York, in 1952. He received the A.B. degree in mathematics from New York University, New York, NY, in 1974, the M.S. degree in applied mathematics from SUNY at Binghamton, Binghamton, NY, in 1975, the M.S. degree in mathematics from Stanford University, Stanford, CA, in 1976, and the Ph.D. degree in mathematics from Brown University, Providence, RI, in 1982. His thesis, in the area of arithmetic geometry, concerned the existence of canonical subgroups in formal grouplaws.

From 1981 to 1984 he was a Postdoc in the Department of Mathematics and Computer Science at Clark University. From 1984 to 1991 he was an Assistant Professor in the College of Computer Science at Northeastern University. From 1991 to 1993 he was an Assistant Professor in the Computer Science Department at the University of Manitoba. Currently he is a Professor in the Department of Computer Science at the University of Kentucky. He was awarded a University Research Professorship for 2002-03. His past research has included work on algebraic geometry over p-adic integer rings, computational geometry, modeling distributed systems, structural complexity theory, and cryptography. His current interests include statistical properties of pseudo-random sequences with applications in cryptography and CDMA; covering properties of codes; and morris dancing.

Dr. Klapper is a Senior Member of the IEEE Information Theory Society. He was the general chair of the Crypto '98 conference and was the Associate for Sequences for the IEEE Transactions on Information Theory from 1999 to 2002.