and

$$G_q(\boldsymbol{\gamma}, 2k+2, L) = \sum_{i=1}^{L} G_q(\boldsymbol{\gamma} \oplus \boldsymbol{\epsilon}_i, 2k+1, L) \tag{A8}$$

where $\boldsymbol{\epsilon}_i \in Z_q^L$ has all components equal to zero except the ith component, which equals 1.

For the purposes of computation, it will be found to be more convenient to rewrite this in terms of weight distribution vectors. For example, in the binary case, the weight distribution of a binary vector $\boldsymbol{\alpha}$ is completely determined by its Hamming weight w, and we will therefore simply write

$$F_2(w, l, L) = F_2(\boldsymbol{\alpha}, l, L) \tag{A9}$$

for all $l = 1, 2, 3, \cdots$. In terms of this notation, the recursion in the binary case simplifies to

$$F_2(w, l, L) = (L - w) F_2(w+1, l-1, L) + w F_2(w-1, l-1, L) \tag{A10}$$

with $F_2(0,0,L) = 1$ serving to initialize the recursion. It is useful as a check on computation to note that

$$F_2(w, l, L) = 0 \tag{A11}$$

whenever any of the following conditions is not satisfied

$$w \geq 0, \qquad w \leq l, \qquad \text{and} \quad w = l(2).$$

This recursion relation in (10) can be used to show that the coefficients $F_2(0, 2k, L)$, $k = 1, 2, \cdots$, are polynomials in L of degree k. The first six polynomials so obtained are listed in Table I. The binary coefficients $F_2(w, k, L)$ have been studied extensively in connection with error-correction codes and other recursion relations may be found in the literature (see for example, Ch. 5 of [12] and Ch. 16 of [16]).

Fig. 1 provides a graphic description of the recursion in (10). From this graph, using induction, one can prove the following.

Lemma 1:

$$\frac{F_2(w, l, L)}{F_2(w, l-2, L)} > \frac{F_2(w', l, L)}{F_2(w', l-2, L)} \tag{A12}$$

whenever all four coefficients are nonzero, and $w > w'$.

As a result, one obtains the following.

Corollary 1:

$$\frac{F_2(w, l, L)}{F_2(w, l-2, L)} > \frac{F_2(0, l, L)}{F_2(0, l-2, L)} \tag{A13}$$

whenever $w > 0$ and all four coefficients are nonzero.

Using Corollary 1 and Fig. 1, one can also prove the following.

Lemma 2:

$$\frac{F_2(0, 2k+2, L)}{F_2(0, 2k, L)} > \frac{F_2(0, 2k, L)}{F_2(0, 2k-2, L)} \tag{A14}$$

whenever all four coefficients are nonzero.

Acknowledgment

Thanks are due to L. Welch for providing the closed-form expression for the coefficients $F_2(\cdot, \cdot, L)$ and to H. Chung for helping with the translation [2].

References

[1] L. R. Welch, "Lower bounds on the maximum cross correlation of signals," *IEEE Trans. Inform. Theory*, vol. IT-20, pp. 397–399, May 1974.

[2] V. M. Sidelnikov, "Cross correlation of sequences," *Probl. Kybern.*, vol. 24, pp. 15–42, 1971 (in Russian).

[3] V. M. Sidelnikov, "On mutual correlation of sequences," *Soviet Math. Dokl.*, vol. 12, pp. 197–201, 1971.

[4] R. Gold, "Optimal binary sequences for spread-spectrum multiplexing," *IEEE Trans. Inform. Theory*, vol. IT-13, pp. 619–621, Oct. 1967.

[5] R. Gold, "Maximal recursive sequences with 3-valued recursive cross-correlation functions," *IEEE Trans. Inform. Theory*, vol. IT-14, pp. 154–156, Jan. 1968.

[6] T. Kasami, "Weight distribution formula for some class of cyclic codes," Coordinated Science Laboratory, Univ. Illinois, Urbana, Tech. Rep. R-285 (AD632574), 1966.

[7] T. Kasami, "Weight distribution of Bose–Choudhary–Hochquengham codes," in *Combinatorial Mathematics and Its Applications*. Chapel Hill, NC: Univ. North Carolina Press, 1969.

[8] J. D. Olsen, R. A. Scholtz, and L. R. Welch, "Bent-function sequences," *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 858–864, Nov. 1982.

[9] D. V. Sarwate and M. B. Pursley, "Crosscorrelation properties of pseudorandom and related sequences," *Proc. IEEE*, vol. 68, May 1980, pp. 593–619.

[10] R. A. Scholtz and L. R. Welch, "Group characters: Sequences with good correlation properties," *IEEE Trans. Inform. Theory*, vol. IT-24, pp. 537–545, 1978.

[11] J. S. No and P. V. Kumar, "A new family of pseudorandom sequences having optimal correlation properties and large linear span," *IEEE Trans. Inform. Theory*, vol. 35, pp. 371–379, Mar. 1989.

[12] F. J. MacWilliams and N. J. A. Sloane, *The Theory of Error-Correcting Codes*. Amsterdam: North Holland, 1977.

[13] R. A. Scholtz and S. W. Golomb, Unpublished lecture notes on information theory.

[14] L. R. Welch, private communication.

[15] S. W. Golomb, *Shift Register Sequences*. San Francisco, CA: Holden-Day, 1967; (revised edition, Laguna Hills, CA: Aegean Park, 1982).

[16] E. R. Berlekamp, *Algebraic Coding Theory*. New York: McGraw-Hill, 1968.

[17] F. S. Roberts, *Applied Combinatorics*. Englewood Cliffs, NJ: Prentice Hall, 1984, ch. 4.

[18] P. V. Kumar and C. M. Liu, "On lower bounds to the maximum correlation of complex, roots-of-unity sequences," CSI Tech. Rep. No. 87-02-03, Univ. Southern California, 1987; revised Sept. 1989.

On the Linear Complexity of Feedback Registers

AGNES HUI CHAN, MARK GORESKY,
AND ANDREW KLAPPER

Abstract —**Sequences generated by arbitrary feedback registers (not necessarily feedback shift registers) with arbitrary feedforward functions are studied. We generalize the definition of linear complexity of a sequence to the notions of strong and weak linear complexity of feedback registers. A technique for finding upper bounds for the strong linear complexities of such registers is developed. This technique is applied to several classes of registers. A feedback shift register in which the feedback function is of the form $x_1 + h(x_2, \cdots, x_n)$ can generate long periodic sequences with high linear complexities only if its linear and quadratic terms have certain forms is proven.**

I. Introduction

Periodic sequences generated by feedback shift registers have many applications in modern communications systems because of their desirable properties such as long period and balanced statistics. One measure of the strength (usefulness) of such a sequence is its linear complexity, as studied by various authors [1], [2], [4], [7], [8]. The linear complexity of a sequence is defined as the length of the shortest linear feedback shift register that generates it. If a sequence has small linear complexity, then the synthesis of a linear equivalent of the sequence generator (such as by the Berlekemp–Massey algorithm [6]) becomes computationally feasible. In this correspondence, we

Manuscript received December 12, 1988; revised September 11, 1989.

The authors are with the College of Computer Science, Northeastern University, Boston, MA 02115.

IEEE Log Number 8932931.

consider pseudorandom sequences generated by general feedback registers (not necessarily shift registers) with arbitrary feedforward functions and develop a new technique for finding upper bounds for the linear complexity of these sequences. We apply this technique to several classes of feedback registers. We prove that if the feedback function of a feedback shift register of length n and maximal linear complexity has the form $x_1 + h(x_2,\cdots,x_n)$ and its feedforward function is x_1 (recall [3] that binary feedback functions that are not of this form cannot generate maximal period sequences), then $h(x_2,\cdots,x_n)$ must either have linear terms or at least $(n-1)/2$ quadratic terms. A more general result is stated in Theorem 2. We also generalize a well-known result of Key [4] bounding the linear complexity of linear feedback shift registers with nonlinear feedforward functions.

In this section, we extend the definition of linear complexity of a sequence to the notion of linear complexity of a feedback register. The technique of establishing upper bounds is developed in Section II, and Section III generalizes the results to an arbitrary finite field GF(q), where q is a power of a prime.

Let GF(2) denote the finite field with two elements. A feedback register (or simply register) of length n is a pair (F,g), where $F=(F_1,\cdots,F_n)$ is a function from GF(2)^n to GF(2)^n (the state transition function), and g is a function from GF(2)^n to GF(2) (the output or feedforward function; see Fig. 1).

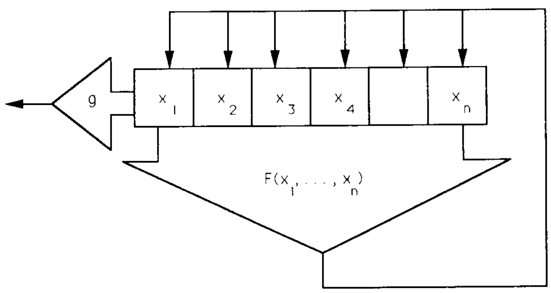

Fig. 1. Feedback register with state transition function F and feedforward function g.

The functions F_i and g can always be written as polynomials in n variables $x_1,\cdots,x_n$ over GF(2), such that each variable has degree of, at most, one. We will write $F^{(i)}$ for the composition of F with itself i times. An initial loading of a register $\mathcal{F}=(F,g)$ is an element $\alpha \in \mathrm{GF}(2)^n$. $\mathcal{F}$, with initial loading α, generates the sequence $\mathcal{F}(\alpha)=(g(\alpha), g\cdot F(\alpha), g\cdot F^{(2)}(\alpha),\cdots)$. Several special cases are of interest. The standard feedforward function is $g(x_1,\cdots,x_n)=x_1$. A register (F,g) is a feedback shift register with feedforward function g if

$$F(x_1,\cdots,x_n)=(x_2,x_3,\cdots,x_n,f(x_1,\cdots,x_n))$$

for some function f from GF(2)^n to GF(2), which is called the feedback function. Such a register is simply called a feedback shift register if it has the standard feedforward function. This case is specified by F (or even f). A register is linear (resp., affine) if g and each F_i are linear polynomials (resp., affine polynomials, i.e., polynomials that have a degree of, at most, one). In case $\mathcal{F}$ is linear, it may be more convenient to think of F as a matrix and of g as a vector, acting by matrix multiplication and dot product, respectively. In this case, $F^{(i)}$ corresponds to the ith power of the matrix F.

We need to distinguish two notions of linear complexity. One, the traditional notion of linear complexity, concerns bit sequences and, by extension, feedback registers with fixed initial loadings. The other, introduced here, concerns feedback registers with no specific initial loadings. The latter notion thus bounds the linear complexities of all sequences generated by a register.

Definition 1: The linear complexity of an ultimately periodic sequence β of elements of GF(2) is the length of the shortest linear feedback shift register $\mathcal{F}$, which has an initial loading α with $\mathcal{F}(\alpha)=\beta$. The weak linear complexity of a register $\mathcal{F}$ is the maximum over-all initial loadings α of the linear complexities of the sequences $\mathcal{F}(\alpha)$.

Definition 2: The strong linear complexity of a register $\mathcal{F}=(F,g)$ is the length of the smallest linear feedback shift register $\mathcal{F}'$ such that for every initial loading α of $\mathcal{F}$, there is an initial loading α' of $\mathcal{F}'$ with $\mathcal{F}(\alpha)=\mathcal{F}'(\alpha')$.

In order to study the strong linear complexity of a register $\mathcal{F}$, we will consider the sequence of polynomials $g, g\cdot F, g\cdot F\cdot F,\cdots$. The output sequence generated by $\mathcal{F}$ with an initial loading α is found by evaluating this sequence of polynomials at α.

The strong linear complexity of a register is greater than or equal to its weak linear complexity, and equality holds for:

a) registers of length n whose output sequences are of maximal period 2^n (i.e., de Bruijn sequences [2]);
b) registers of length n whose state change and feedforward functions do not contain constant terms and whose output sequences are of period 2^n-1 (i.e., modified de Bruijn sequences);
c) linear feedback shift registers; and
d) linear feedback registers with linear feedforward functions (as will be seen by the remarks following Theorem 1 of Section II).

In general, however, these notions do not coincide. For example, the nonlinear feedback shift register $\mathcal{F}$ of length two with feedback function $f(x_1,x_2)=x_1x_2$ generates the sequences $1111\cdots$, $0000\cdots$, $1000\cdots$, and $01000\cdots$. These sequences have linear complexities 1, 0, 1, and 2, respectively; therefore, the weak linear complexity of $\mathcal{F}$ is two. The strong linear complexity of $\mathcal{F}$, however, is three since each of these sequences is generated by the linear feedback shift register of length three with feedback function x_3 and not by any shorter linear feedback shift register.

We also note that the strong linear complexity of a register $\mathcal{F}$ is equal to the degree of the least common multiple of the connection polynomials of the sequences generated by $\mathcal{F}$.

II. Upper Bounds

In this section, we derive a technique for computing bounds on the strong linear complexity of (linear and nonlinear) registers with arbitrary feedforward functions. The idea is to embed the given register into a linear register (of exponentially greater length N). For such a register, the state transition function is considered to be a linear transformation on a vector space of dimension N. We then bound the strong linear complexity of this large linear register. Our first theorem gives a characterization of the strong linear complexity of a register.

Theorem 1: Let $\mathcal{F}=(F,g)$ be a feedback register of length n. The strong linear complexity of $\mathcal{F}$ is the dimension of the span of $\{g\cdot F^{(i)}: i\geq 0\}$, that is, the largest k such that $\{g\cdot F^{(i)}: i=0,\cdots,k-1\}$ are linearly independent.

Proof: If k is as in the statement of the theorem, then $g \cdot F^{(k)}$ can be written as a linear combination of $\{g \cdot F^{(i)}\colon i = 0,\cdots,k-1\}$. Thus, there are elements $\{a_i\colon i = 0,\cdots,k-1\}$ of GF(2) such that

$$g \cdot F^{(k)} = \sum_{i=0}^{k-1} a_i g \cdot F^{(i)}.$$

It follows that for any $j \geq 0$

$$g \cdot F^{(k+j)} = \sum_{i=0}^{k-1} a_i g \cdot F^{(i+j)}.$$

Let $f'\colon \mathrm{GF}(2)^k \to \mathrm{GF}(2)$, and $\theta\colon \mathrm{GF}(2)^n \to \mathrm{GF}(2)^k$ be defined as

$$f'(y_0,\cdots,y_{k-1}) = \sum_{i=0}^{k-1} a_i y_i$$

and

$$\theta(x_1,\cdots,x_n) = \big(g(x_1,\cdots,x_n),\cdots,g \cdot F^{(k-1)}(x_1,\cdots,x_n)\big).$$

Consider the linear feedback shift register $\mathscr{F}'$ of length k with feedback function f' and standard feedforward function. For any initial loading $\alpha \in \mathrm{GF}(2)^n$ of $\mathscr{F}$ and any $i \geq 0$, $g \cdot F^{(i)}(\alpha) = g' \cdot F'^{(i)}(\theta(\alpha))$, that is, $\mathscr{F}(\alpha) = \mathscr{F}'(\theta(\alpha))$. Thus, $\theta(\alpha)$ is an initial loading of $\mathscr{F}'$ giving the same output sequence as $\mathscr{F}$ with initial loading α. It follows that the strong linear complexity of $\mathscr{F}$ is, at most, k.

To show equality, let $\mathscr{F}' = (F', g')$ be any linear feedback shift register of length r (g' is therefore the standard feedforward function) that produces all output sequences that $\mathscr{F}$ produces, and suppose r is the strong linear complexity of $\mathscr{F}$. Then, there is a function $\theta\colon \mathrm{GF}(2)^n \to \mathrm{GF}(2)^r$ such that for every $\alpha \in \mathrm{GF}(2)^n$, $\mathscr{F}(\alpha) = \mathscr{F}'(\theta(\alpha))$. $\mathscr{F}'$ is a linear feedback shift register; therefore, there exist elements $\{a_i\colon i = 0,\cdots,r-1\}$ of GF(2) such that

$$g' \cdot F'^{(r)} = \sum_{i=0}^{r-1} a_i g' \cdot F'^{(i)} \tag{1}$$

(the coefficients of the feedback function define a linear recurrence for the output sequence). For any $\alpha \in \mathrm{GF}(2)^n$, $\mathscr{F}(\alpha) = \mathscr{F}'(\theta(\alpha))$; hence, for every i, $g \cdot F^{(i)}(\alpha) = g' \cdot F'^{(i)}(\theta(\alpha))$. Composing (1) with θ, we see that

$$g \cdot F^{(r)} = \sum_{i=0}^{r-1} a_i g \cdot F^{(i)}.$$

By hypothesis, $\{g \cdot F^{(i)}\colon i = 0,\cdots,k-1\}$ are linearly independent; therefore, k is, at most, r. It follows that k equals the strong linear complexity of $\mathscr{F}$. □

It is a direct consequence of Theorem 1 that the strong linear complexity of a linear register is, at most, its length (the dimension of the space of linear functions on n variables is n), whereas the strong linear complexity of an affine register is at most one greater than its length (the dimension of the space of affine functions on n variables is $n+1$). Next, we show that for an arbitrary feedback register $\mathscr{F} = (F, g)$ of length n, an affine register $\mathscr{F}' = (F', g')$ of length $2^n - 1$ can be constructed such that $\mathscr{F}'$ generates every output sequence generated by $\mathscr{F}$. The register $\mathscr{F}'$ will be linear if both F and g have no constant terms. We will then be able to use Theorem 1 to bound the linear complexity of $\mathscr{F}'$, and hence of $\mathscr{F}$.

A. The Construction

Let S be the set of nonempty subsets of $\{1,\cdots,n\}$. For every I in S, we construct a new variable x_I and identify it with the monomial $\prod_{i \in I} x_i$. Recall that every element a in GF(2) satisfies $a^2 = a$; therefore, all high-degree terms such as x_i^k, $k \geq 1$ appear as x_i. S has cardinality $2^n - 1$ and is used as the index set for the $2^n - 1$ variables in $\mathscr{F}'$. For each I in S, let $F_I(x_1,\cdots,x_n) = \prod_{i \in I} F_i(x_1,\cdots,x_n)$, and let $F_I'(x_{\{1\}},\cdots,x_{\{1,\cdots,n\}})$ be the affine function derived from F_I by replacing each monomial $\prod_{j \in J} x_j$ by the variable x_J, where J is in S. Then, $F' = (F_{\{1\}}',\cdots,F_{\{1,\cdots,n\}}')$ defines an affine function from $\mathrm{GF}(2)^{2^n-1}$ to $\mathrm{GF}(2)^{2^n-1}$. The feedforward function g' can be defined similarly as a linear combination of the monomials x_I and the constant function 1, giving an affine function from $\mathrm{GF}(2)^{2^n-1}$ to GF(2). $\mathscr{F}' = (F', g')$ defines an affine feedback register of length $2^n - 1$. $\mathscr{F}'$ is linear if neither F nor g has constant terms.

To show that $\mathscr{F}'$ generates all the output sequences of $\mathscr{F}$, we consider the embedding $\theta\colon \mathrm{GF}(2)^n \to \mathrm{GF}(2)^{2^n-1}$, where the Ith coordinate of $\theta(x_1,\cdots,x_n)$ is $\prod_{i \in I} x_i$. We claim that $\theta \cdot F = F' \cdot \theta$ and $g = g' \cdot \theta$. In other words, the diagram in Fig. 2 commutes.

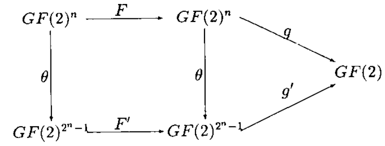

Fig. 2. Linearizing feedback register.

To see this, note first that

$$(\theta \cdot F)_I(x_1,\cdots,x_n) = \prod_{i \in I} F_i(x_1,\cdots,x_n) = F_I(x_1,\cdots,x_n).$$

On the other hand, $(F' \cdot \theta)_I(x_1,\cdots,x_n) = F_I'(\cdots,\prod_{j \in J} x_j,\cdots)$, i.e., $(F' \cdot \theta)_I$ is derived from F_I' by replacing x_J by $\prod_{j \in J} x_j$. However, F_I' was derived from F_I by doing the opposite; therefore, $(F' \cdot \theta)_I = F_I = (\theta \cdot F)_I$, and $F' \cdot \theta = \theta \cdot F$. The second claim is proved similarly.

It follows that for any $\alpha \in \mathrm{GF}(2)^n$ and any k, $g \cdot F^{(k)}(\alpha) = g' \cdot F'^{(k)}(\theta(\alpha))$. Thus, the initial loading $\theta(\alpha)$ of $\mathscr{F}'$ gives the same output sequence as the initial loading α of $\mathscr{F}$.

Example: Let $\mathscr{F} = (F, g)$ be a feedback shift register of length 4 with $g(x_1,x_2,x_3,x_4) = x_1$ and feedback function

$$f(x_1,x_2,x_3,x_4) = x_1 + x_2x_4 + x_2x_3x_4.$$

Then

$$\begin{aligned} F'(&x_1,x_2,x_3,x_4,x_{1,2},x_{1,3},x_{1,4},x_{2,3},x_{2,4},x_{3,4},\\ &x_{1,2,3},x_{1,2,4},x_{1,3,4},x_{2,3,4},x_{1,2,3,4})\\ = (&x_2,x_3,x_4,x_1+x_{2,4}+x_{2,3,4},x_{2,3},x_{2,4},\\ &x_{1,2}+x_{2,4}+x_{2,3,4},x_{3,4},x_{1,3},x_{1,4}+x_{2,4}+x_{2,3,4},x_{2,3,4},\\ &x_{1,2,3},x_{2,4}+x_{1,2,4}+x_{2,3,4},x_{1,3,4},x_{1,2,3,4}). \end{aligned}$$

The output sequence obtained from $\mathscr{F}$ with the initial loading $(1,1,0,1)$ is obtained from $\mathscr{F}'$ with the initial loading $(1,1,0,1,1,0,1,0,1,0,0,1,0,0,0)$.

From this construction, we observe that if the set of polynomials $\{g' \cdot F'^{(i)}\colon i \geq 0\}$ contains only terms in $\{x_I | I \in Q\}$ for some $Q \subseteq S$, then we need only those monomials in $\mathscr{F}'$ indexed by elements of Q. Hence, an affine feedback register of length $|Q|$ (linear if neither F nor g has constant terms) can be constructed that generates the same sequences as $\mathscr{F}$. This shows that the strong linear complexity of $\mathscr{F}$ is bounded above by $|Q|+1$ (by

$|Q|$ if neither F nor g has constant terms). The determination of such a Q is given by the following corollary.

Corollary 1: Let $F(x_1,\cdots,x_n)$ be the state change function of a register of length n with feedforward function $g(x_1,\cdots,x_n)$. Let $T=\{I\in S\colon \prod_{i\in I}x_i$ has a nonzero coefficient in $g\}$, and let Q be the smallest subset of S containing T such that if $I\in Q$ and the coefficient of x_J in F_I' is nonzero, then $J\in Q$.

1) If F or g has constant terms, then the strong linear complexity of (F,g) is bounded above by $|Q|+1$.

2) If neither F nor g has constant terms, then the strong linear complexity of (F,g) is bounded above by $|Q|$.

Proof: In the first case, the space spanned by $\{x_I\colon I\in Q\}\cup\{1\}$ contains the space W spanned by $\{g\cdot F^{(i)}\}$. In the second case, W is spanned by $\{x_I\colon I\in Q\}$. The corollary follows from Theorem 1. □

In the case where $\mathscr{F}$ is a shift register, the determination of Q is given by shifting the corresponding indices, as is given by the next corollary.

Corollary 2: Let (F,g) be a feedback shift register with feedback function f. Let $T=\{I\in S\colon \prod_{i\in I}x_i$ has a nonzero coefficient in $g\}$, $R=\{I\in S\colon \prod_{i\in I}x_i$ has a nonzero coefficient in $f\}$. Let Q be the smallest subset of S containing T such that

1) If $I\in Q$ and $n\in I$, then for each $J\in R$, $J\cup\{i+1\le n\colon i\in I\}\in Q$.

2) If $I\in Q$ and $n\notin I$, then $\{i+1\colon i\in I\}\in Q$.

Then, the strong linear complexity of (F,g) is bounded by

1) $|Q|+1$ if f or g has constant terms;

2) $|Q|$ if neither f nor g has constant terms.

We now treat the special case of a feedback shift register $\mathscr{F}=(F,g)$ of length n with feedback function $f(x_1,\cdots,x_n)=x_1+h(x_2,\cdots,x_n)$ and standard feedforward function. Let T, R, and Q be as in Corollary 2; therefore, $T=\{\{1\}\}$, $\{1\}\in R$, and no other element of R contains 1. Since $\{1\}\in T\subset Q$, we may apply condition 2 repeatedly to obtain $\{i\}\in Q$ for all i. In particular, $\{n\}\in Q$. If J is the index set of a monomial that has a nonzero coefficient in $h(x_2,\cdots,x_n)$, then we can apply condition 1 with $I=\{n\}$; therefore, $J\in Q$. Let I_1 be any element of Q. Then, applying either condition 1 with $J=\{1\}$ or condition 2 (only one condition is applicable to a given index set) $n-1$ times, we get a sequence of elements of Q: $I_1,\cdots,I_n$. One more such application would give us I_1 back again. Actually, we may return to I_1 after a smaller number of applications of the conditions, but this number must divide n. If r is the cardinality of I_1, then r is the cardinality of each I_i, and we call the set $\{I_1,\cdots,I_n\}$ an *r cycle*, or simply a cycle, if the cardinality is clear. Thus, an r cycle is a set $I_1\subseteq\{1,\cdots,n\}$ together with those sets obtained from I_1 by cyclic permutation of the indices $(1,\cdots,n)$. For example, with $n=4$, starting with $I_1=\{2,3\}$, we get the two-cycle $\{2,3\}$, $\{3,4\},\{1,4\},\{1,2\}$, whereas starting with $I_1=\{2,4\}$, we get the two-cycle $\{2,4\},\{1,3\}$. These cycles are independent of $h(x_2,\cdots,x_n)$. The set S of all index sets decomposes into a disjoint union of such cycles with each cycle having cardinality dividing n. If any one element of a cycle is in Q, then every element of that cycle must be in Q.

Remark: There is an interesting relationship between this cycle decomposition and the decomposition of the finite field GF(2^n) into cyclotomic cosets (the orbits under the action of the Galois group over GF(2) [5]). Let α be a primitive element of GF(2^n), $I=\{i_1,\cdots,i_k\}$ be an index set, and $r=\sum_{j=1}^{k}2^{i_j}$. Then, we can identify I with the element α^r of GF(2^n). Under this identification, the cycle containing I corresponds to the cyclotomic coset containing α^r.

Recall again that each monomial in $x_1,\cdots,x_n$ corresponds to an index set; therefore, $\mathscr{F}$ can have high linear complexity only if Q contains many index sets. As seen by the following theorem, this means that the feedback function must have many nonzero coefficients.

Theorem 2: Let $\mathscr{F}=(F,g)$ be a feedback shift register of length n with feedback function $f(x_1,\cdots,x_n)=x_1+h(x_2,\cdots,x_n)$ and standard feedforward function. Let r be the smallest integer such that $h(x_2,\cdots,x_n)$ has a term of degree r with a nonzero coefficient. For any collection of r cycles $C_1,\cdots,C_k$, each of whose corresponding monomials has a zero coefficient in $h(x_1,\cdots,x_n)$, the strong linear complexity of $\mathscr{F}$ is, at most

$$2^n-1-\sum_{i=2}^{r-1}\binom{n}{i}-\sum_{i=1}^{k}|C_i|$$

if h has a constant term and, at most

$$2^n-2-\sum_{i=2}^{r-1}\binom{n}{i}-\sum_{i=1}^{k}|C_i|$$

if h has no constant term.

Proof: Let $P=\{I\colon |I|=1\}\cup\{I\colon \forall i\colon I\notin C_i, |I|=r\}\cup\{I\colon r+1\le|I|\le n-1\}$. We will show that P satisfies the conditions of Corollary 2 and thus contains the set Q of that corollary. P contains the set T and satisfies condition 2 by the observations preceding this theorem. We claim that P satisfies condition 1 as well. Let R be as in Corollary 2. Then $R\subset\{\{1\}\}\cup\{I\colon \forall i\colon I\notin C_i, |I|=r\}\cup\{I\colon r+1\le|I|\le n-1\}\subseteq P$. We have two types of elements of P to which condition 1 applies.

1) $\{n\}\in P$. Condition 1 is satisfied because $R\subseteq P$.

2) Let $n\in I\in P$ and $|I|\ge r$. Then, all other elements of the cycle containing I are in P. Let $J\in R$, and let $K=J\cup\{i+1\le n\colon i\in I\}$. We must show that $K\in P$. If $J=\{1\}$, then K is in the cycle determined by I; therefore, suppose $J\ne\{1\}$. If K has cardinality r, then $K=J\in P$, since J has cardinality at least r. If K has cardinality greater than r, and $K\ne\{1,\cdots,n\}$, then $K\in P$ by definition. Suppose $K=\{1,\cdots,n\}$. We cannot have $1\in\{i+1\le n\colon i\in I\}$; therefore, $1\in J$. It follows that $J=\{1\}$, and hence that $\{2,\cdots,n\}=\{i+1\le n\colon i\in I\}$. Therefore, $K=I$, but $\{1,\cdots,n\}\notin P$, so this is impossible.

P thus contains the set Q of Corollary 2 and has cardinality

$$2^n-2-\sum_{i=2}^{r-1}\binom{n}{i}-\sum_{i=1}^{k}|C_i|$$

proving the theorem. □

This theorem makes precise the belief that shift registers with only high-degree terms are not good. In the example following the construction, we have $r=2$; therefore, the corollary shows that the strong linear complexity of the given register can be, at most, 10.

If the output sequence $(z_0,z_1,\cdots)$ from a feedback shift register with standard feedforward function $\mathscr{F}$ of length n has maximal period 2^n, then any set of 2^n consecutive bits contains 2^{n-1} ones and 2^{n-1} zeros. Therefore, the sequence satisfies the relation $z_i+z_{i+1}+\cdots+z_{i+2^n-1}=0$ for every i. The linear complexity is thus, at most, 2^n-1, and there are registers of length n with linear complexity 2^n-1 [1]. For registers with no constant terms, the maximum possible linear complexity is $2^n - 2$. Note that in these cases, the strong and weak linear complexities of the register and the linear complexity of the output sequence all coincide.

In particular, if $\mathscr{F}$ and r are as in the previous theorem, then $\mathscr{F}$ cannot generate a maximal period, maximal linear complexity sequence unless at least one of the following conditions holds:

1) h has quadratic terms, and for every two-cycle C, there is an I in C whose corresponding monomial in $h(x_1,\cdots,x_n)$ has nonzero coefficient.

2) $h(x_1,\cdots,x_n)$ has linear terms.

Corollary 3: Let $\mathscr{F}=(F,g)$ be a feedback shift register of length n, with feedback function $x_1+h(x_2,\cdots,x_n)$ and standard feedforward function. If $\mathscr{F}$ generates a maximal period, maximal linear complexity sequence, then either h contains some linear terms, or it has at least $\lceil (n-1)/2 \rceil$ quadratic terms.

By a similar application of Corollary 2, we can prove a generalization of a theorem of Key [4].

Proposition 1 (Key [4]): If $\mathscr{F}$ is a feedback register with affine (resp. linear) state change function, every term of whose feedforward function has degree at most k (resp. at most k and at least 1), then its strong linear complexity is bounded above $\Sigma_{i=0}^{k}\binom{n}{i}$ (resp., $\Sigma_{i=1}^{k}\binom{n}{i}$).

Proof: Let $P=\{\{i_1,\cdots,i_l\}\colon\ 1\le l\le k \text{ and } i_1<\cdots<i_l\}$. Then, P satisfies conditions 1 and 2 of Corollary 2 and hence contains the set Q. The cardinality of P is $\Sigma_{i=1}^{k}\binom{n}{i}$. □

The remaining propositions are proved similarly.

Proposition 2: If every term of the feedback function and feedforward function of a feedback shift register with feedforward function has a degree greater than or equal to k, then the strong linear complexity of the register is bounded above by $\Sigma_{i=k}^{n}\binom{n}{i}$.

Proposition 3: If every term of the feedback function of a feedback shift register with feedforward function has degree $\geq k$ and the feedforward function has the form $b_{m+1}x_{m+1}+\cdots+b_nx_n$ (resp., $a+b_{m+1}x_{m+1}+\cdots+b_nx_n$), then the strong linear complexity of the register is bounded above by $n-m+\Sigma_{i=k}^{n}\binom{n}{i}$ (resp., $1+n-m+\Sigma_{i=k}^{n}\binom{n}{i}$).

Proposition 3 says that if the feedback function of a feedback register contains only high-degree terms, then the linear complexity is low.

III. Generalization to Arbitrary Finite Fields

The results of the previous section can be generalized to GF(q), which is the finite field of q elements, where q is a power of an arbitrary prime. The definitions of feedback registers and their various special cases are the same, with 2 replaced by q. The only change is that now every element a of GF(q) satisfies $a^q=a$ so that when we consider functions as polynomials, we must include monomials in which each variable has a degree up to $q-1$. The remaining definitions (output sequence, weak and strong linear complexity, etc.) carry over verbatim, and Theorem 1 still holds.

Recall that a multiset is a set I such that every member a has associated with it a nonnegative integer $\text{mult}_I(a)$, which is called the multiplicity of a in I. If I and J are multisets and k is a nonnegative integer, we then define the multisets I', I^k, $I\cup J$, and $red(I)$ by

1) $\text{mult}_{I'}(1)=0$ and $\text{mult}_{I'}(i)=\text{mult}_I(i-1)$ if $2\le i\le n$.
2) $\text{mult}_{I^k}(i)=k\cdot\text{mult}_I(i)$.
3) $\text{mult}_{I\cup J}(i)=\text{mult}_I(i)+\text{mult}_J(i)$.
4) $\text{mult}_{red(I)}(i)$
$$=\begin{cases}0, & \text{if } \text{mult}_I(i)=0\\ \text{mult}_I(i)-1\ (\text{mod}\ q-1)+1, & \text{otherwise.}\end{cases}$$

In other words, if $\text{mult}_I(i)$ is nonzero, then $\text{mult}_{red(I)}(i)$ is its residue mod $q-1$ in the set $\{1,\cdots,q-1\}$.

Let S be the set of multisets contained in $\{1,\cdots,n\}$ such that each element has multiplicity at most $q-1$, and some element has positive multiplicity. For $I\in S$, we construct a new variable x_I and identify it with the monomial $\prod_{i\in I}x_i^{\text{mult}_I(i)}$. S has cardinality q^n-1. Every function from $\text{GF}(q)^n$ to GF(q) can be written as a linear combination of the x_I and the constant function 1. For $I\in S$, we define

$$F_I(x_1,\cdots,x_n)=\prod_{i\in I}F_i(x)^{\text{mult}_I(i)},$$

reduced using the identities $x_j^q=x_j$, $j=1,\cdots,n$. Thus, each variable appears with a degree at most $q-1$. We then define the affine function F_I' by replacing each monomial $\prod_{i\in I}x_i^{\text{mult}_I(i)}$ in F_I by the corresponding variable x_I. We similarly define the affine function g' from g and combine these functions into an affine feedback register of length q^n-1 over GF(q) that generates all the output sequences of the original register as before.

With these definitions, Corollary 1 holds verbatim. Corollary 2 holds with conditions 1 and 2 replaced by the following: If $I\in Q$ and $J\in R$, then $red(J^{\text{mult}_I(n)}\cup I')\in Q$. Theorem 2 holds with the upper bound

$$q^n-\sum_{j=2}^{r-1}\binom{n}{j}(q-1)^j-\sum_{i=1}^{k}|C_i|(q-1)^r-(q-1)^n$$

in the first case, and

$$q^n-1-\sum_{j=2}^{r-1}\binom{n}{j}(q-1)^j-\sum_{i=1}^{k}|C_i|(q-1)^r-(q-1)^n$$

in the second.

Let $\#(n,i)$ be the number of monomials of degree i in n variables in which each variable has degree at most $q-1$. Proposition 1 then holds with $\binom{n}{i}$ replaced by $\#(n,i)$. In Proposition 2, we must require that each term of the feedback and feedforward functions contain at least k variables, and replace $\binom{n}{i}$ by $\#(n,i)$ in the conclusion. Similarly, in Proposition 3, we must require that each term of the feedback function contain at least k variables and replace $\binom{n}{i}$ by $\#(n,i)$ in the conclusion.

Acknowledgment

We would like to thank an anonymous referee for a number of helpful suggestions, which have resulted in substantial simplifications of this correspondence.

References

[1] A. H. Chan, R. A. Games, and E. L. Key, "On the complexity of deBruijn sequences," *J. Combinatorial Theory*, Series A 33-3, pp. 233–246, 1982.
[2] H. Fredricksen, "A Survey of full length nonlinear shift register cycle algorithms," *SIAM Rev.*, vol. 24, pp. 195–221, 1982.
[3] S. Golomb, *Shift Register Sequences*. Lagina Hills, CA: Aegean Park, 1982.
[4] E. L. Key, "An analysis of the structure and complexity of nonlinear binary sequence generators," *IEEE Trans. Inform. Theory*, vol. IT-22, no. 6, pp. 732–736, Nov. 1976.
[5] S. Lang, *Algebra*. Reading, MA: Addison Wesley, 1971.
[6] J. L. Massey, "Shift register synthesis and BCH decoding," *IEEE Trans. Inform. Theory*, vol. IT-15, pp. 122–127, 1969.
[7] R. A. Rueppel, "New approaches to stream ciphers," Ph.D. dissertation, Swiss Federal Institute of Technology, 1984.
[8] R. A. Rueppel and O. J. Staffelbach, "Products of linear recurring sequences with maximum complexity," *IEEE Trans. Inform. Theory*, vol. IT-33, no. 1, p. 124–131, 1987.