On the Linear Complexity of Feedback Registers

(extended abstract)

A. H. Chan M. Goresky
A. Klapper

Northeastern University
College of Computer Science
360 Huntington Ave.
Boston, MA, 02115

ABSTRACT

In this paper, we study sequences generated by arbitrary feedback registers (not necessarily feedback shift registers) with arbitrary feedforward functions. 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. We prove that a feedback shift register whose feedback function is of the form $x_1 + h(x_2, \ldots, x_n)$ can generate long periodic sequences with high linear complexites only if its linear and quadratic terms have certain forms.

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 [1,2,4,6,7]. The *linear complexity* of a sequence is defined as the length of the shortest linear feedback shift register that

J.J. Quisquater and J. Vandewalle (Eds.): Advances in Cryptology - EUROCRYPT '89, LNCS 434, pp. 563-570, 1990.

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..., 0000..., 1000..., and 01000... These sequences have linear complexities 1, 0, 2, and 2, respectively, so 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

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 look for a supporting subspace of minimal dimension. The dimension of this subspace is an upper bound on the strong linear complexity of the original register.

Definition 1 *Let $\mathcal{F} = (F, g)$ be a linear register of length n, and let W be a subspace of $GF(2)^n$. W supports F or is $\mathcal{F}$-supporting if there is a subspace U in $GF(2)^n$, complimentary to W, such that*

1. $GF(2)^n = W + U$,
2. $F(U) \subseteq U$, *and*
3. *If $w \in W$ and $u \in U$, then $g(w + u) = g(w)$.*

Let w be in W and u be in U. For every i, $F^{(i)}$ is linear. By iterating condition 2, $F^{(i)}(u)$ is in U. It follows that $g{\circ}F^{(i)}(w + u) = g(F^{(i)}(w) + F^{(i)}(u)) = g{\circ}F^{(i)}(w)$. Thus the output from $\mathcal{F}$ can be completely determined from its action on W.

Lemma 1 *Suppose* $GF(2)^n$ *contains a* $\mathcal{F}$*-supporting subspace* W*. Then the strong linear complexity of* $\mathcal{F}$ *is less than or equal to the dimension of* W*.*

The strong linear complexity of a register is bounded from above by the length of any linear feedback register which can produce all the output sequences of the original register. For an arbitrary feedback register $\mathcal{F} = (F', g,)$ of length n, such a linear register $\mathcal{F}' = (F', g')$ of length $2^n - 1$ can be constructed as follows.

The Construction Let S be the set of nonempty subsets of $\{1, \ldots, n\}$. 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 $a^2 = a$, so 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 $\mathcal{F}'$. For each I in S, let $F_I(x_1, \ldots, x_n) = \prod_{i \in I} F_i(x_1, \ldots, x_n)$, and let $F'_I(x_1, \ldots, x_{\{1,\ldots,n\}})$ be the linear 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 $F' = (F'_{\{1\}}, \ldots, F'_{\{1,\ldots,n\}})$ defines a linear function from $V = GF(2)^{2^n - 1}$ to V. The feedforward function g' can be defined similarly as a linear combination of the monomials x_I, giving a linear function from V to $GF(2)$. $\mathcal{F}' = (F', g')$ defines a linear feedback register of length $2^n - 1$ with linear feedforward function.

To show that $\mathcal{F}'$ generates all the output sequences of $\mathcal{F}$, we consider the embedding $\theta : GF(2)^n \rightarrow V$ where the I-th coordinate of $\theta(x_1, \ldots, x_n)$ is $\prod_{i \in I} x_i$. We claim that $\theta \circ F = F' \circ \theta$ and $g = g' \circ \theta$. In other words, the diagram in figure 1 commutes. To see this, note first that $(\theta \circ F)_I(x_1, \ldots, x_n) = \prod_{i \in I} F_i(x_1, \ldots, x_n) = F_I(x_1, \ldots, x_n)$. On the other hand, $(F' \circ \theta)_I(x_1, \ldots, x_n) = F'_I(\ldots, \prod_{j \in J} x_j, \ldots)$, i.e., is derived from F'_I by replacing x_J by $\prod_{j \in J}$. But F'_I was derived from F_I by doing the opposite, so $(F' \circ \theta)_I = F_I = (\theta \circ F)_I$, so $F' \circ \theta = \theta \circ F$. The second claim is proved similarly.

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

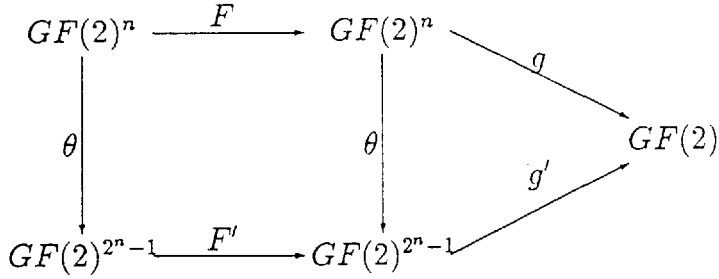

Figure 1: Linearizing a feedback register

Example Let $\mathcal{F} = (F, g)$ 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

$$F'(x_1, x_2, x_3, x_4, x_{1,2}, x_{1,3}, x_{1,4}, x_{2,3}, x_{2,4}, x_{3,4}, x_{1,2,3}, x_{1,2,4}, x_{1,3,4}, x_{2,3,4}, x_{1,2,3,4})$$

$$= (x_2, x_3, x_4, x_1 + x_{2,4} + x_{2,3,4}, x_{2,3}, x_{2,4}, x_{1,2} + x_{2,4} + x_{2,3,4}, x_{3,4}, x_{1,3},$$

$$x_{1,4} + x_{2,4} + x_{2,3,4}, x_{2,3,4}, x_{1,2,3}, x_{2,4} + x_{1,2,4} + x_{2,3,4}, x_{1,3,4}, x_{1,2,3,4}).$$

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

From the construction above we observe that, if the sequence of polynomials $g'(\bar{x})$, $g' \circ F'(\bar{x})$, $g' \circ F' \circ F'(\bar{x}), \ldots$ contains only terms in $\{x_I | I \in Q\}$ for some $Q \subseteq S$, then we need only those monomials in $\mathcal{F}'$ indexed by elements of Q. Hence a linear feedback register of length $|Q|$ can be constructed that generates the same sequences as $\mathcal{F}$. This shows that the strong linear complexity of $\mathcal{F}$ can be bounded above by $|Q|$. The determination of Q is given by the following theorem.

Theorem **1** *Let $F(x_1, \ldots, x_n)$ be the state change function of a register of length n with feedforward function $g(x_1, \ldots, x_n)$. Let $T = \{I \in S : \prod_{i \in I} x_i$ has a non-zero 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 $J \in Q$. Then the strong linear complexity of (F, g) is bounded above by the cardinality of Q.*

Corollary 1 *Let (F, g) be a feedback shift register with feedback function f. Let $T = \{I \in S : \prod_{i \in I} x_i$ has a non-zero coefficient in $g\}$, $R = \{I \in S : \prod_{i \in I} x_i$ has a non-zero 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 $J \in R$, $J \cup \{i+1 \leq n : i \in I\} \in Q$.*
2. *If $I \in Q$ and $n \notin I$, then $\{i+1 : i \in I\} \in Q$.*

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

We now treat the special case of a feedback shift register $\mathcal{F} = (F, g)$ of length n with feedback function $f(x_1, \ldots, x_n) = x_1 + h(x_2, \ldots, x_n)$ and standard feedforward function. Let T, R, and Q be as in corollary 1. Then $\{1\} \in T \subset Q$, so, by applying condition 2 repeatedly, $\{i\} \in Q$ for all i. In particular $\{n\} \in Q$. If J is the index set of a monomial that has a non-zero coefficient in $h(x_2, \ldots, x_n)$, then we can apply condition 1 with $I = \{n\}$, so $J \in Q$. Let I 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, $I = I_1, \ldots, I_n$. One more such application would give us I back again. Actually, we may return to I after a smaller number of applications of the conditions, but this number must divide n. If r is the cardinality of I, then r is the cardinality of each I_i and we call the set $\{I_1, \ldots, I_n\}$ a r-cycle, or simply a cycle if the cardinality is clear. For example, with $n = 4$, starting with $I = \{2, 3\}$ we get the 2-cycle $\{2, 3\}, \{3, 4\}, \{1, 4\}, \{1, 2\}$, whereas starting with $I = \{2, 4\}$, we get the 2-cycle $\{2, 4\}, \{1, 3\}$. These cycles are independent of $h(x_2, \ldots, x_n)$. The set S of all index sets decomposes into a disjoint union of such cycles, each cycle having cardinality dividing n (in fact, there is a relationship between this cycle decomposition and the decomposition of a finite field into cyclotomic cosets). If any one element of a cycle is in Q, then every element of that cycle must be in Q.

Recall again that each monomial in $x_1, \ldots, x_n$ corresponds to an index set, so $\mathcal{F}$ 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 non-zero coefficients.

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

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

This theorem makes precise the folklore belief that shift registers with only high degree terms are not good.

If the output sequence $(z_0, z_1, \ldots)$ from a register $\mathcal{F}$ of length n has maximal period $2^n - 1$, then any set of $2^n - 1$ consecutive bits contains 2^{n-1} ones and $2^{n-1} - 1$ zeros. Therefore the sequence satisfies the relation $z_i + z_{i+1} + \cdots + z_{i+2^n-2} = 0$ for every i. The linear complexity is thus at most $2^n - 2$, and there are registers of length n with linear complexity $2^n - 2$ (for example, the sequence consisting of $2^{n-1} - 1$ zeros followed by 2^{n-1} ones can be generated by such a register.) Note that in the case of a register that outputs a maximal period sequence, the strong and weak linear complexities of the register and the linear complexity of the output sequence all coincide.

In particular, if $\mathcal{F}$ and r are as in the previous theorem, then $\mathcal{F}$ cannot generate a maximal period, maximal linear complexity sequence unless h has quadratic terms and for every 2-cycle C there is an I in C whose corresponding monomial in $h(x_1, \ldots, x_n)$ has non-zero coefficient, or $h(x_1, \ldots, x_n)$ has linear terms.

Corollary **2** *Let $\mathcal{F} = (F, g)$ be a feedback shift register of length n, with feedback function $x_1 + h(x_2, \ldots, x_n)$, and standard feedforward function. If $\mathcal{F}$ 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 1, we generalize a theorem due to Key.

Proposition **1** *(Key [4]) If every term of the feedback function of a feedback shift register with feedforward function has degree* 1 *(resp.* ≤ 1*), and every term of the feedforward function has degree* $\leq k$*, then the strong linear complexity of the register is bounded by* $\sum_{i=1}^{k} \binom{n}{i}$ *(resp.* $\sum_{i=0}^{k} \binom{n}{i}$*).*

We also prove several similar results.

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

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 is of the form* $b_{m+1}x_{m+1} + \cdots + b_n x_n$ *(resp.* $a + b_{m+1}x_{m+1} + \cdots + b_n x_n$*), then the strong linear complexity of the register is bounded above by* $n - m + \sum_{i=k}^{n} \binom{n}{i}$ *(resp.* $1 + n - m + \sum_{i=k}^{n} \binom{n}{i}$*).*

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 $GF(q)$, 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 degree up to $q-1$. The remaining definitions (output sequence, weak and strong linear complexity, etc.) carry over verbatim. The counting techniques can then be generalized using multi-sets, and the main results are modified as follows: 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.

REFERENCES

[1] A.H. Chan, R.A.Games and E.L. Key. *On the complexity of deBruijn sequences.* Journal of Combinatorial Theory, Series **A 33-3**, pp. 233-246, 1982.

[2] H. Fredricksen. *A Survey of Full Length Nonlinear Shift Register Cycle Algorithms.* SIAM Review **24**, pp. 195-221, 1982.

[3] S. Golomb, "Shift Register Sequences", Aegean Park Press, Laguna Hills, CA, 1982.

[4] E.L. Key. *An Analysis of the structure and complexity of nonlinear binary sequence generators.* IEEE Trans. Inform. Theory **IT-22** no. 6, pp. 732-736, Nov. 1976.

[5] J.L. Massey. *Shift Register Synthesis and BCH Decoding.* IEEE Trans. Inform. Theory **IT-15**, page 122-127, 1969.

[6] R.A. Rueppel. *New approaches to stream ciphers.* Ph.D. Thesis, Swiss Federal Instituite of Technology, Zurich, Switzerland. 1984.

[7] R.A. Rueppel and O.J. Staffelbach. *Products of Linear Recurring Sequences with Maximum Complexity* IEEE Trans. Inform. Theory **IT-33** no. 1, pp.124-131, 1987.