An Analysis of the Structure and Complexity of Nonlinear Binary Sequence Generators

EDWIN L. KEY, MEMBER, IEEE

Abstract—A method of analysis is presented for the class of binary sequence generators employing linear feedback shift registers with nonlinear feed-forward operations. This class is of special interest because the generators are capable of producing very long "unpredictable" sequences. The period of the sequence is determined by the linear feedback connections, and the portion of the total period needed to predict the remainder is determined by the nonlinear feed-forward operations. The linear feedback shift registers are represented in terms of the roots of their characteristic equations in a finite field, and it is shown that nonlinear operations inject additional roots into the representation. The number of roots required to represent a generator is a measure of its complexity, and is equal to the length (number of stages) of the shortest linear feedback shift register that produces the same sequence. The analysis procedure can be applied to any arbitrary combination of binary shift register generators, and is also applicable to the synthesis of complex generators having desired properties. Although the discussion in this paper is limited to binary sequences, the analysis is easily extended to similar devices that generate sequences with members in any finite field.

I. INTRODUCTION

SHIFT REGISTER generators are commonly used to produce binary sequences for various purposes. These devices are small, light, inexpensive, and offer a rich variety of sequences. For example, an r-stage register has \(2^r\) possible combinations of feedback/feed-forward connections. Most of these generators have nonlinear feedback, which is difficult to analyze, and they frequently exhibit undesirable properties, such as output sequences with very short periods. Linear feedback shift registers (LFSR's) are more readily analyzed and more commonly used. However there are only \(2^r\) linear feedback connections possible with an r-stage register, and only \(\phi(2^r - 1)/r\) of which have the maximum period \(2^r - 1\). The most serious shortcoming of linear feedback is that no more than \(2r\) successive outputs are needed to determine the feedback connections and initial state of an r-stage register, leaving no further doubt as to the entire sequence. For detailed background on shift register generators, [1] is an excellent source.

It is easy to see that any given periodic binary sequence can be generated by a family of LFSR's. The member of this family with the least number of stages is called the linear equivalent of whatever generator was actually used to generate the given periodic sequence. Although there are several ways to determine the linear equivalent, the algorithm described by Massey [2] seems almost ideal. For this paper we define the complexity of a sequence (or generator) as the length of its linear equivalent.

The inherent predictability of linear feedback generators has motivated designers to use nonlinear operations on generator outputs to increase greatly the equivalent length while maintaining a relatively short actual length. The remaining problem becomes one of predicting the resulting complexity.

This paper presents a method of analysis based on Galois field theory that enables one to predict the generator complexity resulting from nonlinear operations. Moreover the theory provides a conceptually simple basis for synthesizing devices with desirable characteristics, and is readily extendable to a larger class of more complex generators.

II. GALOIS FIELD REPRESENTATION

Consider a binary sequence \(\{a_n\}\) where \(a_n\) is the nth member of the sequence, \(n = 0,1,2, \ldots\). If the sequence is generated by an r-stage LFSR, it is completely specified by the initial loading \(a_0,a_1,a_2, \ldots, a_{r-1}\) and by the linear recursion that specifies the feedback,

\[
a_n + \sum_{i=1}^{r} c_i a_{n-i} = 0, \quad n \geq r
\]

where the sequence members \(a_n\) and the feedback constants \(c_i\) are members of \(GF(2)\). Also the operations of (1) are the defined operations of \(GF(2)\); namely, addition and multiplication modulo 2. In all that follows, the constant \(c_r\) is one, otherwise the register would have only \(r-1\) effective stages.

The linear recursion can be expressed as a linear difference equation

\[
(E^r + \sum_{i=1}^{r} c_i E^{r-i}) a_n = 0, \quad n \geq 0
\]

where \(E\) is the shifting operator which operates on \(a_n\) to give \(a_{n+1}\); i.e., \(Ea_n = a_{n+1}\). Associated with (2) is the characteristic equation

\[
x'^r + \sum_{i=1}^{r} c_i x'^{r-i} = 0.
\]

An equation such as (3) with coefficients \(c_i\) in \(GF(2)\) is said to be over \(GF(2)\), and is thoroughly treated in the literature. (For instance, see Albert [3] or Dickson [4].) Equation (3) is known to have roots in \(GF(2^m)\), where \(m\) is the least...
common multiple of the degrees of the irreducible factors\(^2\) of (3). Let \(\alpha\) be such a root. Then \(A\alpha^n\) is a solution of (2), where \(A\) is an arbitrary constant. Likewise, each distinct root of (3) will provide a linearly independent solution. If \(\alpha\) is a root of multiplicity \(q\), then \((\gamma)\alpha^n\) with \(i = 0,1,2, \cdots, q - 1\) provides \(q\) linearly independent solutions, where \((\gamma)\) is the binomial coefficient reduced modulo 2. Since (3) has \(r\) roots, there are \(r\) linearly independent solutions to (2), and the general solution is a linear combination of these solutions with \(r\) arbitrary constants determined by the initial values \(a_0, a_1, a_2, \cdots, a_{r - 1}\). A simple example of this representation will serve to clarify the concept.

Consider the three-stage LFSR illustrated in Fig. 1. The generator is specified by the linear recursion

\[
a_n + a_{n-2} + a_{n-3} = 0, \quad n \geq 3
\]

and by the initial state \(a_0 = 1, a_1 = 0, a_2 = 0\). The sequence satisfies the difference equation

\[
(E^3 + E + 1)a_n = 0
\]

with the characteristic equation

\[
x^3 + x + 1 = 0.
\]

The polynomial on the left in (6) is irreducible and its roots are elements of \(GF(2^3)\). Let \(\alpha\) be such a root. Then

\[
\alpha^3 + \alpha + 1 = 0.
\]

From (7) we can generate the powers of \(\alpha\): \(\alpha = \alpha, \alpha^2 = \alpha^2, \alpha^3 = \alpha + 1, \alpha^4 = \alpha^2 + \alpha, \alpha^5 = \alpha^2 + \alpha + 1, \alpha^6 = \alpha^2 + 1, \alpha^7 = 1\), which together with 0 constitute the eight elements of \(GF(2^3)\).

Now (7) has two other conjugate roots in addition to \(0, \alpha, \alpha^2\), and it can be shown that these are \(\alpha^2\) and \(\alpha^4 = \alpha + 1\). The existence of these conjugate roots follows from the fact that, for any polynomial \(f(x)\) over \(GF(2)\), \(f(x)^2 = f(x^2)\). Therefore

\[
a_n = A_1\alpha^n + A_2\alpha^{2n} + A_3(\alpha^2 + \alpha)^n.
\]

From the values of \(a_0, a_1, a_2\), we can determine that \(A_1 = A_2 = A_3 = 1\), giving

\[
a_n = \alpha^n + \alpha^{2n} + (\alpha^2 + \alpha)^n.
\]

Since \(\alpha^7 = 1\), we see that the sequence has a period of 7. If (9) is evaluated using the powers of \(\alpha\), the resulting sequence is \([1, 0, 0, 1, 0, 1, 1]\).

The method of solution in this example was to create \(GF(2^3)\) by adjoining \(\alpha\), a root of (6), to \(GF(2)\), and then identifying two additional conjugate roots of the form \(\alpha^k\) with \(k = 1, 2\). This gives a total of three distinct roots, which is sufficient for a general solution to (5). The method can be extended to any irreducible characteristic equation by applying the following theorem.

**Theorem:** If \(p(x)\) is an irreducible polynomial of degree \(d\) over \(GF(2)\), then the extension field \(GF(2^d)\) generated by adjoining a root \(\alpha\) of \(p(x)\) to \(GF(2)\) contains \(d\) distinct roots of \(p(x)\), and these roots are given by \(\alpha^k\) with \(k = 0, 1, 2, \cdots, d - 1\).

A proof of this theorem is contained in theorems 6.21 and 6.26 of [5].

We can apply the results of the theorem to each irreducible factor of (3). If the \(j\)th factor has degree \(d_j\), then it has \(d_j\) conjugate roots in \(GF(2^{d_j})\), the extension field \(GF(2^{d_j})\) being produced by the adjunction of a root of the factor. Each of the extension fields so produced from the factors of (3) is a subfield of \(GF(2^m)\) where \(m\) is the least common multiple of the \(d_j\). Since we are able to produce \(r\) roots of (3), all of which are elements of \(GF(2^m)\), and since each root provides a linearly independent solution to (3), we have the general solution for any sequence that satisfies a linear recursive relationship such as (1). Thus, any LFSR can be represented by a set of elements of a Galois field, the set being the roots of its characteristic equation.

**III. NONLINEAR GENERATORS**

Groth [6] has described a class of nonlinear generators with linear feedback determined by a primitive\(^3\) characteristic polynomial and with nonlinear logic on only two stages to provide the output. This class serves well to introduce the analysis of nonlinear devices based on the Galois field representation. From our prior discussion, we have seen that the output sequence of an LFSR with irreducible polynomial is given by

\[
a_n = \sum_{i=0}^{r-1} A_i (\alpha^{2i})^n
\]

where \(\alpha\) is a root of the characteristic polynomial. Let \(a^*\) be the sequence from a different stage. Then

\[
a^*_n = \sum_{i=0}^{r-1} A^*_i (\alpha^{2i})^n.
\]

Since any nonlinear operation on two binary digits can be reduced to one product modulo 2 with a possible sum modulo 2 and a possible addition of 1 (complementation), it is sufficient to consider the product sequence \(a_n a^*_n\), which is given by

\[
a_n a^*_n = \sum_{i=0}^{r-1} \sum_{j=0}^{r-1} A_i A^*_j (\alpha^{2(i+j)2})^n.
\]

The exponent of \(\alpha^{2i+2j}\) can be thought of as the binary number representation of an integer \(\mu\). Since \(\alpha^{2^\mu-1} = 1\), any

\[^2\]Irreducible polynomials over a field are polynomials that cannot be separated into factors of positive degree over the field.

\[^3\]A primitive polynomial is a polynomial with roots of maximum order for its degree. If the polynomial is of degree \(r\), the roots have order \(2^r - 1\). Every primitive polynomial is irreducible, but not conversely.
μ with the same residue modulo $2^{r-1}$ gives the same element. If $i = j$, then μ has only one nonzero bit, and there are $r$ such integers, all different mod $2^{r-1}$. Conversely, if $i \neq j$, the integer μ has two nonzero bits, and there are $r(r - 1)/2$ such integers, all different modulo $2^{r-1}$. Thus there are $r + \lfloor (r - 1)/2 \rfloor = r(r + 1)/2$ distinct powers of $\alpha$ present in (12). Provided the coefficients of the terms do not vanish, the generator would have complexity $r(r + 1)/2$.

For the case under consideration these coefficients do not vanish as can be seen from the following argument. The sequence $[a^{*n}]$ is simply a different phase (a delayed or advanced version) of the sequence $[a_n]$, i.e.,

$$a^{*n} - a_{n+6}, \quad -r \leq \delta \leq r.$$  

(13)

By substituting $n + 6$ for $n$ in (10) and then comparing (10) with (11), we see that

$$A_i(\alpha^2)^i = A_i^*.$$  

(14)

If $i \neq j$, then the coefficient $A_iA_j^* = A_iA_j(\alpha^2)^i$, and the coefficient $A_jA_i^* = A_jA_i(\alpha^2)^j$. Because of the symmetry of $2^i + 2^j$ in $i$ and $j$, both coefficients are applied to the same power of $\alpha$, and they must be equal if that power is to vanish. However, since neither $A_iA_j$ nor any power of $\alpha$ is zero, these terms all survive. If $i \neq j$, then the coefficient $A_iA_j^* = A_iA_j(\alpha^2)^i$, and the coefficient $A_jA_i^* = A_jA_i(\alpha^2)^j$. Therefore, the total number of distinct roots $\lambda_{N_m}$ that can be present in $r$-stage LFSR's (having primitive characteristic polynomials) with $m$ stages combined in multiplication is given by

$$rN_m = \sum_{i=1}^{m} \binom{m}{i}.$$  

(18)

Provided none of the terms has zero coefficients, the length of the linear equivalent will be $r N_m$. Unlike the case for $m = 2$, these coefficients can vanish. Fig. 3 presents an example of such a degeneracy.

The four-stage LFSR has a characteristic polynomial that is primitive. Three stages are combined in multiplication, but the linear equivalent is only ten stages long rather than the expected fourteen. However, exceptions of this sort are few and present neither a serious theoretical difficulty nor a practical limitation.

The result of (18) is stated without proof in a paper by Ristenbatt et al. [7] and has been confirmed (with certain exceptions noted) by computer simulation using the Massey algorithm.

We see that the three additional roots $\alpha^3$, $\alpha^6$, and $\alpha^5$ have been introduced, thereby increasing the complexity to six.

The three additional roots are the conjugate roots of the irreducible polynomial $x^3 + x^2 + 1 = 0$. The linear equivalent of this generator is also shown in Fig. 2. The resulting linear sequence is [0,0,0,0,1,1]. We can extend Groth's approach to generators having more than two stages combined in a nonlinear fashion.

With an $m$th order nonlinearity, the expression analogous to (12) is

$$a_n \cdot a_{n} \cdot a_{n} \cdots m a_n = \sum_{i_1=0}^{r-1} \sum_{i_2=0}^{r-1} \cdots \sum_{i_m=0}^{r-1} A_{i_1} \cdots A_{i_m}$$

$$\cdots A_{i_m} (\alpha^{2i_1 + 2i_2 + \cdots + 2i_m} \cdots)^n$$  

(17)

where $a_{k,n}$ is the sequence of the $k$th output and is associated with the summation involving $i_k$. We can, as before, identify a binary number $\mu = 2^1 + 2^2 + \cdots + 2^m$ and advance similar arguments. In general there are $2^m$ distinct modulo $2^{r-1}$ with only $i$ nonzero bits. Therefore, the total number of distinct roots $\lambda_{N_m}$ that can be present in $r$-stage LFSR's (having primitive characteristic polynomials) with $m$ stages combined in multiplication is given by

$$rN_m = \sum_{i=1}^{m} \binom{m}{i}.$$  

(18)

Provided none of the terms has zero coefficients, the length of the linear equivalent will be $r N_m$. Unlike the case for $m = 2$, these coefficients can vanish. Fig. 3 presents an example of such a degeneracy.

The four-stage LFSR has a characteristic polynomial that is primitive. Three stages are combined in multiplication, but the linear equivalent is only ten stages long rather than the expected fourteen. However, exceptions of this sort are few and present neither a serious theoretical difficulty nor a practical limitation.

The result of (18) is stated without proof in a paper by Ristenbatt et al. [7] and has been confirmed (with certain exceptions noted) by computer simulation using the Massey algorithm.

We see that, as $m \to r$, $rN_m \to 2^r - 1$; i.e., all of the nonzero elements of $GF(2^r)$ are included as roots. Thus the complexity can be made to approach the period of the se-
sequence, and since the period can be extremely long, large complexity can be achieved.

The nonlinear generators studied thus far utilize only the products of different phases of the same sequence. Consequently, the roots multiplied are in the same field, as are their products. We have shown that a sufficient number of such multiplications will produce and include in the generator every element (except zero) of the field. To create more complexity, we must introduce elements of another field. Let us then consider the complexity that results from the product of two distinctly different LFSR's.

Consider two LFSR's, both with irreducible characteristic polynomials, having degrees \( r \) and \( s \), respectively. Let \( r \) and \( s \) be relatively prime. The roots of the characteristic polynomials are in \( GF(2^r) \) and \( GF(2^s) \), respectively. Since \( r \) and \( s \) are relatively prime, the intersection of \( GF(2^r) \) and \( GF(2^s) \) contains only \( GF(2) \). This follows from the fact that every subfield of \( GF(2^m) \) is a field \( GF(2^m) \) where \( m \) is a divisor of \( n \).

Suppose that \( \alpha \) and \( \beta \) are elements of \( GF(2^r) \) and \( GF(2^s) \), respectively, and that neither is in \( GF(2) \). Then their product \( \alpha \beta \) is in neither \( GF(2^r) \) nor \( GF(2^s) \) as we can easily demonstrate. Since \( GF(2^r) \) contains \( \alpha \), it also contains \( \alpha^{-1} \), and if it contains \( \alpha \beta \), it also contains \( \alpha^{-1} \alpha \beta = \beta \). However, since \( \beta \) is not in \( GF(2^r) \), neither is \( \alpha \beta \). Likewise, \( \alpha \beta \) is not in \( GF(2^s) \). The product \( \alpha \beta \) is contained in the superfield \( GF(2^r) \). The situation is illustrated in Fig. 4.

One of our LFSR's has \( r \) conjugate roots in \( GF(2^r) \), and the other has \( s \) conjugate roots in \( GF(2^s) \). If their outputs are multiplied, the product sequence from the above argument has \( rs \) distinct roots in \( GF(2^r) \). It can be shown that these \( rs \) roots are the conjugate roots of an irreducible polynomial of degree \( rs \).

Since there is no danger of any of these roots vanishing (as happened in the previous example), the complexity of the resulting sequence is exactly \( rs \). The period of the product sequence is the least common multiple of the periods of the two generators. An example of this process is shown in Fig. 5. The two-stage generator of Fig. 5 has a primitive root \( \beta \) for which \( \beta^2 + \beta + 1 = 0 \). Then \( \beta = \beta, \beta^2 = \beta + 1, \) and \( \beta^3 = 1 \). The generated sequence \( b_n \) is given by

\[
b_n = (\beta + 1)\beta^n + \beta\beta^{2n}.
\]  

(19)

It has a period of 3 and is the sequence \( [1,0,1] \). The sequence generated by the three-stage generator is given by (9). The product sequence is given by

\[
a_n b_n = (\beta + 1)(\alpha\beta)^n + \beta(\alpha\beta^2)^n + (\beta + 1)(\alpha^2\beta)^n + \beta(\alpha^2\beta^2)^n + (\beta + 1)(\alpha^4\beta) + \beta(\alpha^4\beta^2)^n.
\]  

(20)

The product sequence has the roots \( \alpha\beta, \alpha\beta^2, \alpha^3\beta, \alpha^4\beta \), and \( \alpha^4\beta^2 \), which are the six conjugate roots of the irreducible polynomial

\[
x^6 + x^4 + x^2 + x + 1 = 0.
\]  

(21)

The sequence generated is

\[
[1,0,0,1,0,1,1,0,0,0,0,1,0,0,1,0,0,1,0,0,1,0,0,1]
\]

which has period 21, the least common multiple of the periods of the two LFSR's.

The nonlinear process we have discussed can be extended in many directions and can serve as a basis for the synthesis of extremely complex generators. As a final example of the theory, we consider the nonlinear generator suggested by Geffe [8].

Geffe's generator consists of three LFSR's connected as shown in Fig. 6. The concept is to use LFSR \#2 as a control generator to connect either LFSR \#1 or LFSR \#3, but not both, to the output. If the control generator produces a 1, then LFSR \#1 is connected; if it produces a 0, then LFSR \#2 is connected.
In a sense there are really four LFSR's being employed. The operation of complementing the output of the control generator in its relationship to LFSR \# 3 is equivalent to modulo 2 addition of a single-stage LFSR as shown in Fig. 7. This device has the characteristic equation \( x + 1 = 0 \), which has the single root 1.

Suppose that the three LFSR's have distinct primitive characteristic polynomials of degree \( r \), \( s \), and \( t \), respectively. The resulting generator would then have complexity \( rs + (s + 1)t \), the 1 in \( s + 1 \) arising from the process of complementing. The period would be the least common multiple of \( 2^r - 1 \), \( 2^s - 1 \), and \( 2^t - 1 \).

Although the complexity of this device could be greater in a different configuration of the stages, the generator does have some desirable attributes. For instance, it has a balanced distribution of zeros and ones in its output. It also offers the advantage of being useful as a module of a superstructure of similar arrangements, i.e., the entire generator of Fig. 6 could play the role of LFSR \# 1 in the same arrangement with like generators. The complexity would escalate accordingly. A generator of the type described by Groth could also be used for the LFSR's of Fig. 6, and the resulting complexity could be very great. An example of a superstructure of these generators is shown in Fig. 8. The complexity of the first level of LFSR's is shown in circles, and the complexity at each subsequent point is also indicated.

The generator in Fig. 8 achieves a complexity of over one million with nine LFSR's having a total of 156 stages. If all the LFSR's have maximum period, the period of the output sequence is extremely long, and the balance between zero and one outputs should be very good. No attempt has been made to maximize the complexity, although there are ways to achieve far greater complexity with the same number of stages.

IV. SUMMARY AND CONCLUSIONS

LFSR's can be represented by the roots of their characteristic equation in a Galois field. This representation is well suited to the analysis of nonlinear, feed-forward operations on LFSR's. Such nonlinear operations increase the complexity of the resulting sequence by the process of introducing new roots. The representation is also a powerful aid to the synthesis of complex generators.

The product of two phases of the same sequence from an \( r \)-stage primitive LFSR gives a linear equivalent of \( (1) + (2) \) stages. If more stages are combined, additional binomial coefficients such as \( (3), (4), \) etc., can usually, but not always, be added. This process can be continued until all nonzero elements of \( GF(2^r) \) are present, at which point no further complexity is achievable without introducing a different sequence.

The product of two sequences whose complexities are \( r \) and \( s \), and whose determining roots are in different subfields of a finite field, has a complexity \( rs \).

Combining LFSR's that have a modest number of stages employing nonlinear operations can lead to extremely complex generators of very long periods. Although not discussed in this paper, there are methods of assuring that these sequences have certain desirable random properties such as a balance of zeros and ones. However, the question of how to control the random properties of these sequences needs further investigation. Other investigations might well include the properties of sequences whose members are elements of other finite fields or even rings.

ACKNOWLEDGMENT

The author gratefully acknowledges the contributions of R. W. Jacobus, Dr. J. D. R. Kramer, and Dr. H. Gelman who reviewed the paper and made valuable suggestions for its improvement.

REFERENCES

学霸图书馆
www.xuebalib.com

本文献由“学霸图书馆-文献云下载”收集自网络，仅供学习交流使用。

学霸图书馆（www.xuebalib.com）是一个“整合众多图书馆数据库资源，
提供一站式文献检索和下载服务”的24小时在线不限IP图书馆。

图书馆致力于便利、促进学习与科研，提供最强文献下载服务。

图书馆导航：
图书馆首页 文献云下载 图书馆入口 外文数据库大全 疑难文献辅助工具