

## Chapter 4

# Feedback Shift Register Sequences

Feedback shift register sequences have been widely used as synchronization codes, masking or scrambling codes, and for white noise signals in communication systems, signal sets in CDMA (code division multiple access) communications, key stream generators in stream cipher cryptosystems, random number generators in many cryptographic primitive algorithms, and for testing vectors in hardware design. S. Golomb's popular book "Shift Register Sequences", first published in 1967 and revised in 1982 is a pioneering book which discusses this type of sequences. In this chapter, we introduce this topic and discuss the synthesis and the analysis of periodicity of linear feedback shift register (LFSR) sequences. We give different (though equivalent) definitions and representations for LFSR sequences and point out which are most suitable for either implementation or analysis. This chapter contains seven sections, which are organized as follows. In Section 1, we give a general description for feedback shift registers at the gate level for the binary case and as a finite field configuration for  $q$ -ary case. In Sections 2-4, we introduce the definition of LFSR sequences from the point of view of polynomial rings and discuss their characteristic polynomials, minimal polynomials and periods. Then, we show the decomposition of LFSR sequences. We provide the matrix representation of LFSR sequences in Section 5 as another historic approach and discuss their trace representation for the irreducible case in detail in Section 6, which is a more modern approach. (The general case will be treated in Chapter 6.) LFSRs with primitive minimal polynomials are basic building blocks for nonlinear generators. The trace representation of LFSR sequences is a powerful tool for the analysis of unpredictability or randomness of pseudo-random sequences and for the design of pseudo-random sequences with desired properties. In Section 7, we present the generating function method for studying LFSR sequences.

## 4.1 Feedback Shift Registers

In this section, we give a definition and some of the basic terms for feedback shift register sequences. We denote  $F = GF(2) = \{0, 1\}$  and

$$F^n = \{(a_0, a_1, \dots, a_{n-1}) \mid a_i \in F\},$$

a vector space over  $F$  of dimension  $n$ . A function with  $n$  binary inputs and one binary output is called a *boolean function of  $n$  variables*, i.e.,  $f : F^n \rightarrow F$ , which can be represented as follows:

$$f(x_0, x_1, \dots, x_{n-1}) = \sum c_{i_1 i_2 \dots i_t} x_{i_1} x_{i_2} \dots x_{i_t}, c_{i_1 i_2 \dots i_t} \in F \quad (4.1)$$

where the sum runs through all subsets  $\{i_1, \dots, i_t\}$  of  $\{0, 1, \dots, n-1\}$ . This shows that there are  $2^{2^n}$  different boolean functions of  $n$  variables.

### A. Basic Concepts and Examples

An  $n$ -stage shift register is a circuit consisting of  $n$  consecutive 2-state storage units (flip-flops) regulated by a single clock. At each clock pulse, the state (1 or 0) of each memory stage is shifted to the next stage in line. A shift register is converted into a code generator by including a feedback loop, which computes a new term for the left-most stage, based on the  $n$  previous terms. In Figure 4.1, we see a diagram of a feedback shift register (FSR).

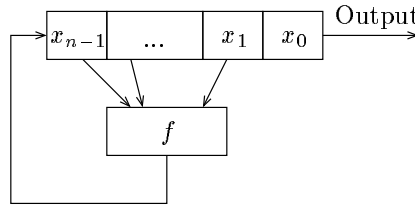
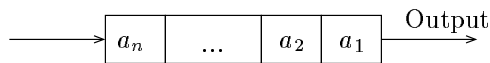


Figure 4.1: A Block Diagram for an FSR

Each of the squares is a 2-state storage unit. The  $n$  binary storage elements are called the *stages* of the shift register, and their contents (regarded as either a binary number or a binary vector,  $n$  bits in length) is called a *state* of the shift register.  $(a_0, a_1, \dots, a_{n-1}) \in F^n$  is called an *initial state of the shift register*. The feedback function  $f(x_0, x_1, \dots, x_{n-1})$  is a boolean function of  $n$  variables, defined in (4.1). At every clock pulse, there is a transition from one state to the next. To obtain a new value for stage  $n$ , we compute  $f(x_0, x_1, \dots, x_{n-1})$  of all the present terms in the shift register and use this in stage  $n$ . For example, the next state of the shift register in Figure 4.1 becomes where

$$a_n = f(a_0, a_1, \dots, a_{n-1}).$$



After the consecutive clock pulses, a feedback shift register outputs a sequence:

$$a_0, a_1, \dots, a_n, \dots \quad (4.2)$$

The sequence satisfies the following recursive relation

$$a_{k+n} = f(a_k, a_{k+1}, \dots, a_{k+n-1}), k = 0, 1, \dots \quad (4.3)$$

Any  $n$  consecutive terms of the sequence in (4.2)

$$a_k, a_{k+1}, \dots, a_{k+n-1}$$

represents a state of the shift register in Figure 4.1. A *state (or vector) diagram* is a diagram that is drawn based on the successors of each of the states. The output sequence is called a *feedback shift register sequence*. If the feedback function  $f(x_0, x_1, \dots, x_{n-1})$  is a linear function, then the output sequence is called a *linear feedback shift register (LFSR) sequence*. Otherwise, it is called a *nonlinear feedback shift register (NLFSR) sequence*. Sometimes, we also say that  $\mathbf{a}$  is *generated* by an LFSR (or NLFSR). Here linear means that the feedback function computes the modulo 2 sum of a subset of the stages of the shift registers.

**Example 4.1** In Figure 4.2, we see a 3-stage shift register with a (nonlinear) feedback function  $f(x_0, x_1, x_2) = x_0x_1$ .

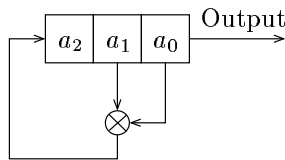


Figure 4.2: A 3-stage nonlinear feedback shift register

From this, we can compute the next-state function, as shown in the following table, a “successor table”, for each of the eight states of the shift register.

Succession of states

$x_0x_1x_2$	$x_0x_1x_2$
current state	next state
000	000
001	010
010	100
011	110
100	000
101	010
110	101
111	111

The state diagram of the eight states is shown in Figure 4.3. From the state diagram, we can directly observe the autonomous behavior of the device. For example, the initial state 100 leads to the output sequence 1000..., and this state diagram is shown in Figure 4.4.

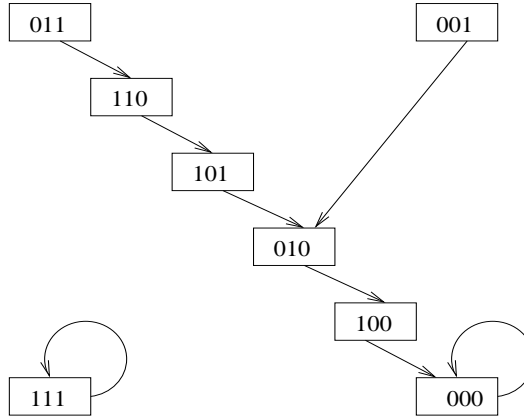


Figure 4.3: The state diagram of the FSR in Fig.4.2

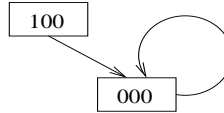


Figure 4.4: The state diagram with the initial state  $x_0x_1x_2 = 100$

**Example 4.2** A 3-stage LFSR is shown in Figure 4.5 with the linear feedback function  $f(x_0, x_1, x_2) = x_0 + x_1$ . The truth table of this feedback function is given in Table 4.1, and the state diagram of the corresponding LFSR is shown in Figure 4.6. The output sequence with the initial state 100 is

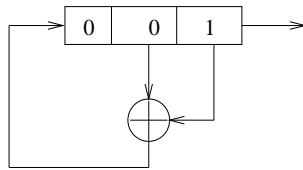


Figure 4.5: A 3-stage LFSR for example 4.2

Table 4.1: Truth table of  $f = x_0 + x_1$

Truth table of $f = x_0 + x_1$	
$x_0x_1x_2$	$f = x_0 + x_1$
000	0
001	0
010	1
011	1
100	1
101	1
110	0
111	0

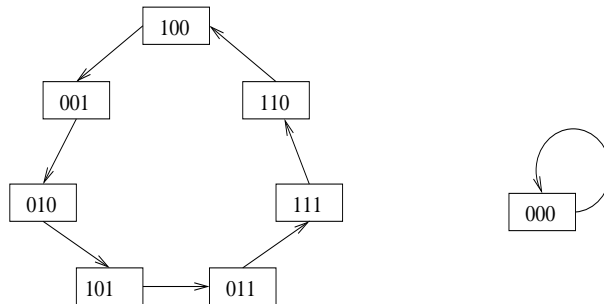


Figure 4.6: State diagram of the LFSR in Figure 4.5

10010111001011...

which is seen to repeat periodically with a period of 7.

**Example 4.3** A 3-stage LFSR with a nonlinear feedback function  $f(x_0, x_1, x_2) = x_0 + x_1x_2 + x_2 + 1$ , as shown in Figure 4.7. The truth table of this boolean function is given in Table 4.2. From the truth table, the state diagram of the NLFSR is easily obtained, as shown in Figure 4.8. Therefore, the output sequence with the initial state 100 is

1000101110001011....

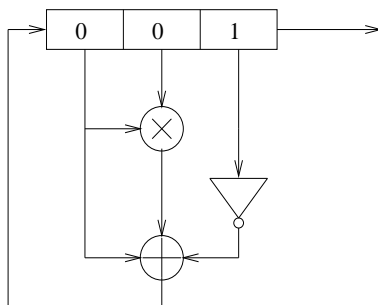


Figure 4.7: The 3-stage NLFSR for example 4.3

Table 4.2: Truth table of  $f = x_0 + x_1x_2 + x_2 + 1$ 

$x_0x_1x_2$	$x_0 + x_1x_2 + x_2 + 1$
000	1
001	0
010	1
011	1
100	0
101	1
110	0
111	0

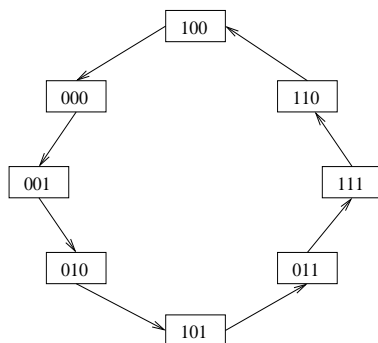


Figure 4.8: State diagram of Figure 4.7

**Example 4.4** A 4-stage NLFSR with a nonlinear feedback function  $f(x_0, x_1, x_2, x_3) = x_0 + x_1x_2x_3 + x_1 + 1$ . The output sequence with the initial state 1111 is given by

$$1111011001010000 \dots$$

which has a period of 8.

A *de Bruijn sequence* is an output sequence of an  $n$ -stage NLFSR having period  $2^n$  and satisfying that each  $n$ -tuple occurs exactly once in each period. The sequences given by Examples 4.3 and 4.4 are de Bruijn sequences with periods 8 and 16 respectively. For a general discussion of nonlinear feedback shift register sequences including the number of de Bruijn sequences and constructions for several subsets of de Bruijn sequences, see Chapter 6 in *Shift Register Sequences* [63].

We will see that the periods of LFSRs are completely determined by their feedback functions in a mathematically predictable way. However, for NLFSRs, there are only a few results on the period problem in the literature.

**Remark 4.1** The de Bruijn sequence in Example 4.3 can be obtained by inserting an extra 0 into the run of 2 consecutive zeros of the LFSR sequence in Example 4.2. But the feedback functions of these two sequences are completely different. A sequence generated by an  $n$ -stage LFSR with period  $2^n - 1$  is called a *maximal length sequence*, or  $m$ -sequence for short (we will formally define it later). From any  $m$ -sequence of period  $2^n - 1$ , we can get a de Bruijn sequence by inserting an extra 0 into the run of  $n - 1$  consecutive zeros of the  $m$ -sequence. However, this type of de Bruijn sequence is not secure for use in stream cipher cryptosystems.

### B. Finite Field Configuration for $q$ -ary Feedback Shift Register Sequences

Let  $F = GF(q)$  where  $q$  is a prime or a power of a prime (so  $|F| = q$ ). Considering Figure 4.1, if each stage is replaced by a  $q$ -state storage unit and the feedback function is replaced by a function from  $F^n$  to  $F$ , i.e.,

$$f(x_0, x_1, \dots, x_{n-1}) = \sum c_{i_0, i_1, \dots, i_{n-1}} x_{i_0}^{i_0} x_{i_1}^{i_1} \cdots x_{i_{n-1}}^{i_{n-1}}, \quad c_{i_0, i_1, \dots, i_{n-1}} \in F, \quad (4.4)$$

which is a polynomial function in  $n$  interdeterminates. Here the exponents  $i_0, i_1, \dots, i_{n-1} \in \{0, 1, \dots, q-1\}$ , since  $x^q = x$  for any  $x \in F$ . The output sequences of the shift register is called a  *$q$ -ary feedback shift register (FSR) sequence*. In other words, we define an abstract model for  $q$ -ary FSR sequences as follows. Let  $\mathbf{a} = \{a_i\}_{i \geq 0}$ ,  $a_i \in F$  be a  $q$ -ary sequence whose elements are given by

$$a_{n+k} = f(a_k, a_{k+1}, \dots, a_{k+n-1}), \quad k = 0, 1, \dots, \quad (4.5)$$

where  $f(x_0, x_1, \dots, x_{n-1})$  is a function from  $F^n$  to  $F$  defined by (4.4). Then  $\mathbf{a}$  is called a  *$q$ -ary feedback shift register (FSR) sequence*, and  $(a_0, a_1, \dots, a_{n-1})$  is still called an *initial state* of  $\mathbf{a}$ . The significant difference between  $q = 2$  and  $q > 2$  is that Figure 4.1 for binary FSR sequences is a device at the gate level, but Figure 4.1 for a  $q$ -ary FSR ( $q > 2$ ) is only a finite field configuration. The latter needs more circuits to implement arithmetics of finite fields for feedback function computation.

For whichever case  $q = 2$  or  $q > 2$ , the design of LFSR sequences with desired properties requires us to understand the functionality of three components of

an LFSR: the initial state, the feedback function and the output sequences. In other words, we want to understand how the behavior of the output sequences is completely determined by initial states and feedback functions. From now on, we treat the cases  $q = 2$  and  $q > 2$  together when we discuss their algebraic properties. From (4.4), the following result is immediate.

**Property 4.1** *There are  $q^{q^n}$  different functions from  $F^n$  to  $F$ .*

### C. Periodic Property

**Definition 4.1** *The sequence  $a_0, a_1, \dots$  is denoted as  $\mathbf{a}$  or  $\{a_i\}$ . If  $a_i \in F$ , then we say that  $\mathbf{a}$  is a  $q$ -ary sequence or a sequence over  $F$ . If there exist integers  $r > 0$  and  $u \geq 0$  such that*

$$a_{i+r} = a_i \quad \text{for all } i \geq u, \quad (4.6)$$

*then the sequence is said to be ultimately periodic with parameters  $(r, u)$ , and  $r$  is called a period of the sequence. The smallest number  $r$  satisfying (4.6) is called a (least) period of the sequence. If  $u = 0$ , then the sequence is said to be periodic. When the context is clear, we simply say a period of  $\mathbf{a}$  instead of the least period of  $\mathbf{a}$ .*

For example, the output sequence  $00011011011\dots$  of a 4-stage LFSR with the feedback function  $f(x_0, x_1, x_2, x_3) = x_2 + x_3$  and the initial state  $a_0a_1a_2a_3 = 0001$  is an ultimately periodic sequence, where  $u = 2$  and the period  $r$  is 3. The output sequence of the feedback shift register sequence in Figure 4.5 is a periodic sequence with period 7.

**Theorem 4.1** *Any  $q$ -ary feedback shift register sequence is ultimately periodic with period  $r \leq q^n$  where  $n$  is the number of the stages. In particular, if  $q = 2$ , then  $r \leq 2^n$ .*

*Proof.* In a  $q$ -ary feedback shift register with  $n$  stages, there are  $q^n$  possible states. Each state uniquely determines its successor. Hence, the first time a previous state is repeated, a period for the sequence is established. (If state  $S$  at time  $t_1$  is the same as state  $S'$  at time  $t_2$ , then the states at times  $t_1 + 1$  and  $t_2 + 1$  are the same, as are the states at times  $t_1 + 2$  and  $t_2 + 2$ , etc.) Thus the maximum possible period is  $q^n$ , the number of the different states. □

### D. Linear Feedback Shift Register Sequences

If the feedback function  $f(x_0, x_1, \dots, x_{n-1})$  is a linear function, i.e., if it can be expressed as

$$f(x_0, x_1, \dots, x_{n-1}) = c_0x_0 + c_1x_1 + \dots + c_{n-1}x_{n-1}, c_i \in F,$$

## 4.2. DEFINITION OF LFSR SEQUENCES IN TERMS OF POLYNOMIAL RINGS 89

then the recursive relation shown in (4.3) becomes the following linear recursive relation

$$a_{k+n} = \sum_{i=0}^{n-1} c_i a_{k+i}, k = 0, 1, \dots. \quad (4.7)$$

Thus an LFSR sequence is also called a *linear recursive sequence* (or *linear recurring sequence*) over  $F$  in the literature, where  $F$  could be  $GF(2)$  or any finite field  $GF(q)$ . Note that there are only  $q^n$  different  $n$ -stage LFSRs. In particular, for  $q = 2$ , we only have  $2^n$  different  $n$ -stage binary LFSRs.

**Theorem 4.2** *Let  $\mathbf{a}$  be a sequence generated by an  $n$ -stage LFSR over  $F$ . Then the period of  $\mathbf{a}$  is  $\leq q^n - 1$ . In particular, if  $q = 2$ , the period of any binary  $n$ -stage LFSR sequence is  $\leq 2^n - 1$ .*

*Proof.* Note that the successor state of  $00 \cdots 0$  ( $n$  times 0) of an  $n$ -stage LFSR is again  $00 \cdots 0$ . Using the same argument as in the proof of Theorem 1, we see that the period of  $\mathbf{a}$  is  $\leq q^n - 1$ , since the state  $00 \cdots 0$  cannot be part of any other period. □

In the rest of this chapter, we will restrict ourselves to LFSR sequences. It is worth pointing out that it is difficult to generate nonlinear sequences with the desired properties using an NLFSR directly. Most of the methods for generating nonlinear sequences make use of one or several LFSRs together with some control operations.

## 4.2 Definition of LFSR Sequences in Terms of Polynomial Rings

In order to characterize the periodicity of LFSR sequences, we introduce another equivalent definition of LFSR sequences over  $F$  in terms of the polynomial ring  $F[x]$ . We use the notation  $F = GF(q)$ . The basic mathematical tools of this section are from linear algebra.

### A. The Left Shift Operator

Let  $V(F)$  be a set consisting of all infinite sequences whose elements are taken from  $F$ , i.e.,

$$V(F) = \{\mathbf{a} = (a_0, a_1, \dots) \mid a_i \in F\}. \quad (4.8)$$

Let

$$\begin{aligned} \mathbf{a} &= (a_0, a_1, a_2, \dots), \\ \mathbf{b} &= (b_0, b_1, b_2, \dots) \end{aligned}$$

be two sequences in  $V(F)$  and let  $c \in F$ . We define addition and scalar multiplication on  $V(F)$  as follows:

$$\begin{aligned}\mathbf{a} + \mathbf{b} &= (a_0 + b_0, a_1 + b_1, a_2 + b_2, \dots), \\ c\mathbf{a} &= (ca_0, ca_1, ca_2, \dots).\end{aligned}$$

It is easy to verify that  $V(F)$  is a linear space (i.e., a vector space) over  $F$  under these two operations. We also denote the zero sequence  $\mathbf{0} = (0, 0, 0, \dots)$ . ( $\mathbf{0}$  represents an element in  $V(F)$  or in  $F$  depending on the context. Sometimes, we also use  $\mathbf{0}$  for the zero sequence.) Thus, an LFSR sequence is a sequence

$$\mathbf{a} = (a_0, a_1, a_2, \dots)$$

in  $V(F)$  whose elements satisfy the linear recursive relation

$$a_{n+k} = \sum_{i=0}^{n-1} c_i a_{k+i}, \quad k = 0, 1, \dots \quad (4.9)$$

For  $\mathbf{a} = (a_0, a_1, a_2, \dots) \in V(F)$ , we define a (left) shift operator  $L$  as follows:

$$L\mathbf{a} = (a_1, a_2, a_3, \dots).$$

Note that  $L$  is a linear transformation of  $V(F)$ . Generally, for any positive integer  $i$ , we have

$$L^i \mathbf{a} = (a_i, a_{i+1}, a_{i+2}, \dots).$$

By convention, we write  $L^0 \mathbf{a} = I\mathbf{a} = \mathbf{a}$ , where  $I$  is the identity transformation on  $V(F)$ . By using the left shift operation  $L$ , the formula (4.9) can be written as

$$L^n \mathbf{a} = \sum_{i=0}^{n-1} c_i L^i \mathbf{a},$$

or equivalently,

$$\left( L^n - \sum_{i=0}^{n-1} c_i L^i \right) \mathbf{a} = \mathbf{0}. \quad (4.10)$$

We write

$$\begin{aligned}f(x) &= x^n - (c_{n-1}x^{n-1} + \dots + c_0), \\ f(L) &= L^n - (c_{n-1}L^{n-1} + \dots + c_0I), \text{ and } f(L)\mathbf{a} = \mathbf{0}.\end{aligned}$$

From (4.10), the definition of LFSR sequences (or linear recursive sequences) is equivalent to the following definition.

**Definition 4.2** For any infinite sequence  $\mathbf{a}$  in  $V(F)$ , if there exists a non-zero monic polynomial  $f(x) \in F[x]$  such that

$$f(L)\mathbf{a} = \mathbf{0},$$

then  $\mathbf{a}$  is called a linear recursive sequence, or equivalently, an LFSR sequence. The polynomial  $f(x)$  is called the characteristic polynomial of  $\mathbf{a}$  over  $F$ . The reciprocal polynomial of  $f(x)$  is called the feedback polynomial of  $\mathbf{a}$ .

## 4.2. DEFINITION OF LFSR SEQUENCES IN TERMS OF POLYNOMIAL RINGS 91

Note that the definition of reciprocal polynomials has been introduced in Section 3 of Chapter 3. For any non-zero polynomial  $f(x) \in F[x]$ , we use  $G(f)$  to represent the set consisting of all sequences in  $V(F)$  with

$$f(L)\mathbf{a} = \mathbf{0}.$$

Since  $f(L)$  is also a linear transformation,  $G(f)$  is a subspace of  $V(F)$ .

*Note:* By convention, the constant polynomial 1 is the characteristic polynomial of the zero sequence  $00\dots$ .

**Theorem 4.3** *Let  $f(x) \in F[x]$  be a monic polynomial of degree  $n$ . Then  $G(f)$  is a linear space of dimension  $n$ . Hence it contains  $q^n$  different sequences. In particular, if  $q = 2$ ,  $G(f)$  contains  $2^n$  different binary sequences.*

*Proof.* For a sequence

$$\mathbf{a} = (a_0, a_1, \dots, a_{n-1}, a_n, \dots) \in G(f),$$

since  $\deg(f) = n$ , once the first  $n$  terms  $(a_0, a_1, \dots, a_{n-1})$  (or equivalently, an initial state) are given, the other terms of  $\mathbf{a}$  can be determined by the formula (4.9) starting from  $a_n$ . There are  $q^n$  different ways to choose an  $n$ -tuple  $(a_0, a_1, \dots, a_{n-1})$  in  $F^n$ . Therefore  $|G(f)| = q^n$ . □

Note that the sequences in  $V(F)$  may or may not be periodic. Definition 4.2 of LFSR sequences makes it easier to determine periodicity of the sequences. We will discuss this in the next section. We conclude this section with a summary of three definitions that we already encountered plus some examples. Note that these three definitions are equivalent. Thus when we say that a sequence  $\mathbf{a} = (a_0, a_1, a_2, \dots)$  is *generated by an  $n$ -stage LFSR*, we mean any one of the following three equivalent definitions.

1.  $\mathbf{a} = (a_0, a_1, a_2, \dots)$  is an output sequence of an LFSR with the linear feedback function

$$f(x_0, x_1, \dots, x_{n-1}) = \sum_{i=0}^{n-1} c_i x_i, c_i \in F$$

and an initial state

$$(a_0, a_1, \dots, a_{n-1}),$$

so that the elements of  $\mathbf{a}$  satisfy the following recursive relation:

$$a_{k+n} = \sum_{i=0}^{n-1} c_i a_{k+i}, k = 0, 1, \dots \quad (4.11)$$

2.  $\mathbf{a}$  is a linear recursive sequence which satisfies the above recursive relation (4.11).

3. There exists a monic polynomial  $f(x) = x^n - \sum_{i=0}^{n-1} c_i x^i \in F[x]$  of degree  $n$  such that

$$f(L)\mathbf{a} = 0,$$

or equivalently,  $\mathbf{a} \in G(f)$ .

The polynomial  $f(x)$  is called the *characteristic polynomial* of  $\mathbf{a}$  for each of these definitions; and the reciprocal polynomial of  $f(x)$  is called the *feedback polynomial* of  $\mathbf{a}$ .

**Example 4.5** A sequence  $\mathbf{a} = (00010011010111)$  of period 15 is generated by an LFSR with the feedback function  $f(x_0, x_1, x_2, x_3) = x_0 + x_1$ , and the LFSR implementation is shown in Figure 4.9. Thus,  $\mathbf{a}$  satisfies the linear recursive

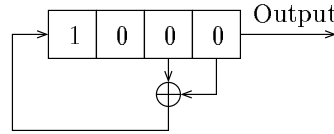


Figure 4.9: A 4-Stage LFSR

relation:

$$a_{4+k} = a_k + a_{1+k}, k = 0, 1, \dots$$

and its characteristic polynomial is given by

$$f(x) = x^4 + x + 1.$$

Equivalently,  $\mathbf{a}$  is a linear recursive sequence that satisfies the following linear recursive relation:

$$a_{4+k} = a_k + a_{1+k}, k = 0, 1, \dots,$$

or equivalently,  $\mathbf{a} \in G(f)$ , i.e.,

$$f(L)\mathbf{a} = (L^4 + L + I)\mathbf{a} = 0.$$

In particular, there are  $2^4 = 16$  different sequences in  $G(f)$ , since there are 16 ways to choose a 4-tuple  $(a_0, a_1, a_2, a_3)$ . We list all of these sequences in Table 4.3.

### 4.3 Minimal Polynomials and Periods

In the previous section, we associated LFSR sequences over  $F$  with polynomials over  $F$ . This enables us to use the extensive theory of periods of polynomials to investigate the periods of LFSR sequences. In this section, we discuss the minimal polynomials and the periods of LFSR sequences in terms of the polynomial ring  $F[x]$ .

Table 4.3: All Sequences in  $G(x^4 + x + 1)$ 

Initial State	Sequence
0000	0000...
0001	0001001101011111
0010	001001101011110
0011	001101011110001
0100	0100110201111100
0101	010111100010011
0110	011010111100010
0111	011110001001101
1000	100010011010111
1001	100110101111000
1010	101011110001001
1011	101111000100110
1100	110001001101011
1101	110101111000100
1110	111000100110101
1111	111100010011010

### A. Minimal Polynomials of LFSR Sequences

Let  $\mathbf{a}$  be an LFSR sequence. According to the definition, there is a nonzero monic polynomial  $f(x)$  such that

$$f(L)\mathbf{a} = 0. \quad (4.12)$$

In fact, for the fixed sequence  $\mathbf{a}$ , there are many polynomials for which (4.12) is satisfied. For example, given the LFSR sequence:

$$\mathbf{a} = 011011\dots,$$

then the polynomial  $f(x) = x^2 + x + 1$  satisfies the property  $f(L)\mathbf{a} = 0$ . But the polynomial  $x^3 + 1$  also has this property. In order to find relations among these polynomials, for the LFSR sequence  $\mathbf{a}$ , we define

$$A(\mathbf{a}) = \{f(x) \in F[x] \mid f(L)\mathbf{a} = 0\}.$$

In other words,  $A(\mathbf{a})$  is the set consisting of all polynomials satisfy the condition

$$f(L)\mathbf{a} = 0.$$

According to the definition of characteristic polynomials of a sequence,  $A(\mathbf{a})$  consists of all characteristic polynomials of  $\mathbf{a}$ .

**Theorem 4.4** *Let  $\mathbf{a}$  be an LFSR sequence and  $A(\mathbf{a})$  be defined as above. Then  $A(\mathbf{a})$  satisfies the following properties:*

- (a) *The zero polynomial belongs to  $A(\mathbf{a})$ .*
- (b) *If  $f(x), g(x) \in A(\mathbf{a})$ , then  $f(x) \pm g(x) \in A(\mathbf{a})$ .*
- (c) *If  $f(x) \in A(\mathbf{a})$  and  $h(x) \in F[x]$ , then  $h(x)f(x) \in A(\mathbf{a})$ .*

*Proof.* (a)  $0\mathbf{a} = 0 \implies 0 \in A(\mathbf{a})$ .

(b)

$$\begin{aligned} f(x), g(x) \in A(\mathbf{a}) &\implies f(L)\mathbf{a} = 0 \text{ and } g(L)\mathbf{a} = 0 \\ &\implies (f(L) \pm g(L))\mathbf{a} = f(L)\mathbf{a} \pm g(L)\mathbf{a} = 0 \\ &\implies f(x) \pm g(x) \in A(\mathbf{a}). \end{aligned}$$

(c)

$$\begin{aligned} f(x) \in A(\mathbf{a}) &\implies f(L)\mathbf{a} = 0 \\ &\implies (h(L)f(L))\mathbf{a} = h(L)(f(L)\mathbf{a}) = h(L)0 = 0. \end{aligned}$$

□

*Note.* Since  $A(\mathbf{a})$  is closed with respect to all of these operations,  $A(\mathbf{a})$  is not merely a linear space, but also an algebra.

**Definition 4.3** *A monic polynomial of the lowest degree in  $A(\mathbf{a})$  is called a minimal polynomial of  $\mathbf{a}$  over  $F$ .*

In other words, the minimal polynomial of a sequence represents the LFSR of shortest length which can generate the sequence.

**Remark 4.2** According to the definition of LFSR sequences, the constant polynomial 1 is the minimal polynomial of the zero sequence  $00\dots$ , and the polynomial  $x - 1$  is the minimal polynomial of any constant sequence  $(c, c, \dots)$ ,  $0 \neq c \in F$ .

**Theorem 4.5** *Let  $\mathbf{a} \in V(F)$  and  $m(x)$  be the minimal polynomial of  $\mathbf{a}$ . Then the minimal polynomial of  $\mathbf{a}$  is unique and satisfies the following two properties:*

- (a)  $m(L)\mathbf{a} = 0$ .
- (b) *For  $f(x) \in F[x]$ ,  $f(L)\mathbf{a} = 0$  if and only if  $m(x) \mid f(x)$ , i.e.,  $m(x)$  divides  $f(x)$ .*

*Proof.* We first establish the validity of these two assertions on minimal polynomials of sequences, and then we show their uniqueness. If  $\mathbf{a} = 00\dots$ , then it is clear that the results are true. Now we suppose that  $\mathbf{a}$  is a nonzero sequence. Since  $m(x) \in A(\mathbf{a})$ , then  $m(L)\mathbf{a} = 0$  is satisfied automatically, which gives (a). For the assertion (b), if  $m(x) \mid f(x)$ , we can write  $f(x) = m(x)g(x)$ . Since  $m(x)$

is the minimal polynomial of  $\mathbf{a}$ , then  $m(x) \in A(\mathbf{a})$ . By Proposition 4.4 - (c),  $f(x) \in A(\mathbf{a})$ . Next, we show that if  $f(x) \in A(\mathbf{a})$  then  $m(x)|f(x)$ . Applying the division algorithm to  $f(x)$  and  $m(x)$ , there exist  $q(x), r(x) \in F[x]$  such that

$$f(x) = q(x)m(x) + r(x), 0 \leq \deg(r(x)) < \deg(m(x)).$$

Again using Proposition 4.4 - (c), it follows that  $q(L)m(L)\mathbf{a} = 0$ . Hence

$$0 = f(L)\mathbf{a} = (q(L)m(L) + r(L))\mathbf{a} = q(L)m(L)\mathbf{a} + r(L)\mathbf{a} = r(L)\mathbf{a}.$$

So if  $r(x) \neq 0$ , then  $r(L)\mathbf{a} = 0 \implies r(x) \in A(\mathbf{a})$ . But  $\deg(r(x)) < \deg(m(x))$  which contradicts  $m(x)$  is a polynomial in  $A(\mathbf{a})$  with lowest degree. Therefore  $r(x) = 0 \implies f(x) = q(x)m(x) \implies m(x)|f(x)$ . If there is another polynomial  $m_1(x)$  in  $A(\mathbf{a})$  which is also a minimal polynomial of  $\mathbf{a}$ , then according to (b), we have both  $m(x)|m_1(x)$  and  $m_1(x)|m(x)$ . Since the polynomials in  $A(\mathbf{a})$  are monic,  $m(x) = m_1(x)$ . □

Note that for any  $\mathbf{a} \in G(f)$ ,  $f(x)$  need not be the minimal polynomial of  $\mathbf{a}$ . However, we have the following result.

**Corollary 4.1** *If  $f(x) \neq 0$ ,  $\mathbf{a} \in G(f)$ , then the minimal polynomial of  $\mathbf{a}$ , say  $m(x)$ , divides  $f(x)$ .*

*Proof.* According to the definition of  $G(f)$ ,  $\mathbf{a} \in G(f) \implies f(L)\mathbf{a} = 0$ . Applying Theorem 4.5-(b),  $m(x)|f(x)$ . □

**Example 4.6** Let  $F = GF(2)$  and  $f(x) = x^5 + x^4 + 1$ . Then the following sequences belong to  $G(f)$ :

$$\begin{array}{ll} \mathbf{a} & = 100101110010\dots & \text{of period 7} \\ \mathbf{b} & = 01101101\dots & \text{of period 3} \\ \mathbf{c} & = 10000111110101001100010000\dots & \text{of period 21} \end{array}$$

which have the minimal polynomials  $f_1(x) = x^3 + x + 1$ ,  $f_2(x) = x^2 + x + 1$ , and  $f(x)$ , respectively. Note that  $f(x) = f_1(x)f_2(x)$ . Thus the minimal polynomials of  $\mathbf{a}$  and  $\mathbf{b}$  are divisors of  $f(x)$ .

**Corollary 4.2** *With the same notation as in Corollary 4.1, if  $f(x)$  is irreducible, then  $f(x)$  is the minimal polynomial of any nonzero sequence in  $G(f)$ .*

*Proof.* Note that  $f(x)$  has only 1 and itself as its factors, and the sequence having 1 as its minimal polynomial is the zero sequence. The result follows immediately. □

**Example 4.7** Let  $f(x)$  be as given in Example 4.5. Since the characteristic polynomial  $f(x)$  is irreducible, then all the 15 nonzero sequences have  $f(x)$  as their minimal polynomial. (In fact, they are simply cyclic shifts of each other.)

According to the definition of minimal polynomials, the degree of the minimal polynomial of  $\mathbf{a}$  is equal to the length of the shortest LFSR which can generate  $\mathbf{a}$ . This is a very important security parameter for measuring unpredictability of pseudorandom sequences used as key stream sequences in stream cipher cryptosystems. The degree of the minimal polynomial of a sequence is called the *linear span* (or *linear complexity*) of the sequence. We will give a formal definition below.

### B. Periodicity

For any periodic sequence, we have the following result.

**Theorem 4.6** *If  $\mathbf{a}$  is an ultimately periodic sequence with parameters  $(u, r)$ , then the minimal polynomial of  $\mathbf{a}$  is  $m(x) = x^u m_1(x)$  with  $m_1(0) \neq 0$  and  $m_1(x) | (x^r - 1)$ . Hence, it can be generated by an LFSR.*

*Proof.* Note that

$$a_{k+r} = a_k, k = u, u+1, \dots$$

$\implies (L^r - 1)L^u(\mathbf{a}) = \mathbf{0} \implies m(x) | x^u(x^r - 1)$ . We write  $m(x) = x^u m_1(x)$ . Then  $m_1(0) \neq 0$  and  $m_1(x)$  divides  $x^r - 1$ . Therefore,  $\mathbf{a}$  can be generated by an LFSR with the characteristic polynomial  $m(x)$ . □

The following corollary follows immediately from Theorem 4.6.

**Corollary 4.3** *If  $\mathbf{a}$  is periodic with period  $r$ , then its minimal polynomial  $m(x)$  divides  $x^r - 1$ .*

From the proof of Theorem 4.6, if  $\mathbf{a}$  is ultimately periodic, then the minimal polynomial of  $\mathbf{a}$  can be written as  $m(x) = x^u m_1(x)$  with  $m_1(0) \neq 0$ .

**Definition 4.4** *Let  $\mathbf{a}$  be an ultimately periodic sequence over  $F$ . Then the degree of the minimal polynomial of  $\mathbf{a}$  is called the linear span or linear complexity of  $\mathbf{a}$ .*

In other words, the linear span of a periodic sequence is the length of the shortest LFSR which can generate the sequence.

Note. This definition can be extended to any finite segment of a sequence.

**Lemma 4.1** *Let  $r$  be the least period of  $\mathbf{a}$ . If  $l$  is a period of  $\mathbf{a}$ , then  $r | l$ .*

*Proof.*  $l$  is a period of  $\mathbf{a} \implies L^l \mathbf{a} = \mathbf{a}$ , i.e.,  $a_i = a_{i+l}, \forall i \geq 0$ . On the other hand, since  $r$  is the least period of  $\mathbf{a}$ , we also have  $L^r \mathbf{a} = \mathbf{a}$ . Applying the division algorithm for integers to  $l$  and  $r$ , there exist two integers  $q$  and  $t$  such that

$$l = qr + t, 0 \leq t < r.$$

Note that  $L^{qr} \mathbf{a} = \mathbf{a}$ . Thus  $L^l \mathbf{a} = L^{qr+t} \mathbf{a} = L^t(L^{qr} \mathbf{a}) = L^t \mathbf{a} = \mathbf{a}$ . Since  $r$  is the smallest number having this property,  $t = 0$ . Hence,  $r | l$ . □

Let us denote by  $per(s)$  a period of a sequence  $s$  or of a polynomial  $s$ .

**Theorem 4.7** Let  $\mathbf{a}$  be an LFSR sequence with minimal polynomial  $m(x)$ .

(a) If  $m(0) \neq 0$ , then  $\mathbf{a}$  is periodic. In this case,

$$\text{per}(\mathbf{a}) = \text{per}(m(x)).$$

In other words,

*the period of an LFSR sequence is equal to  
the period of the minimal polynomial of the sequence*

(b) The inverse of the first assertion is also true, i.e., if  $\mathbf{a}$  is periodic, then  $m(0) \neq 0$ .

*Proof.* Let  $\mathbf{a}$  be an ultimately periodic sequence over  $F$  with parameters  $(u, r)$  and  $m(x)$  be its minimal polynomial over  $F$ . According to Definition 4.1,  $\mathbf{a}$  is periodic if and only if  $u = 0$ . From Theorem 4.6,  $m(x) = x^u m_1(x)$  with  $m_1(0) \neq 0$  and  $m_1(x) | (x^r - 1)$ . Thus  $\mathbf{a}$  is periodic if and only if  $m(x) = m_1(x)$ . According to the definitions of periods of sequences and polynomials, it is immediate that  $\text{per}(\mathbf{a}) = \text{per}(m(x))$ . The assertion (a) is now established. Conversely, if  $\mathbf{a}$  is periodic, then  $u = 0$  and  $x^r - 1 \in A(\mathbf{a})$ . According to Theorem 4.5, we have  $m(x) | (x^r - 1) \implies m(0) \neq 0$ , which gives the assertion (b). □

Thus far, we have obtained a criterion for determining whether an ultimately periodic sequence is periodic in terms of the evaluation of its minimal polynomial (or any characteristic polynomial of the sequence) at 0. Next, we will show the relationships among the period of the sequence, the period of its minimal polynomial and the order of a root of the minimal polynomial when the minimal polynomial is irreducible.

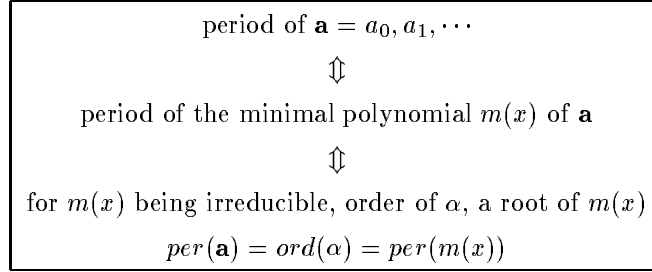
**Theorem 4.8** Let  $\mathbf{a}$  be an LFSR sequence with minimal polynomial  $m(x)$ . Assume that  $m(x)$  is an irreducible polynomial over  $F = GF(q)$  of degree  $n$ . Let  $\alpha$  be a root of  $m(x)$  in the extension field  $GF(q^n)$ . Then

$$\text{per}(\mathbf{a}) = \text{per}(m(x)) = \text{ord}(\alpha)$$

In other words, the period of the sequence  $\mathbf{a}$ , the period of the minimal polynomial of  $\mathbf{a}$ , and the order of a root of the minimal polynomial of  $\mathbf{a}$  are equal.

*Proof.* Note that  $m(x)$  is the minimal polynomial of  $\alpha$ . From Theorem 3.9 in Section 3.5 of Chapter 3, we have  $\text{per}(m(x)) = \text{ord}(\alpha)$ . According to Theorem 4.7, the assertion is established. □

This is an important discovery in the history of sequence design and analysis [67]. The period of an LFSR sequence is equal to the period of its minimal polynomial. If the minimal polynomial is irreducible, then the period of the sequence is equal to the order of a root (all roots have the same order) of the minimal polynomial in the extension field. The following diagram illustrates this relationship.



**Example 4.8** Let  $\mathbf{a} = (1000010101110110001111100110100)$ , generated by  $f(x) = x^5 + x^3 + 1$ . The period of  $f(x)$  is 31. So  $per(\mathbf{a}) = 31$ . Furthermore,  $f(x)$  is primitive over  $GF(2)$ . Let  $\alpha$  be a root of  $f(x)$ . Then  $\alpha$  is a primitive element in  $GF(2^5)$ . Thus, the order of  $\alpha$  is 31. Therefore, we have

$$per(\mathbf{a}) = per(f(x)) = ord(\alpha) = 31$$

We define the following sets:

- $S$ , the set consisting of all periodic LFSR sequences over  $GF(q)$ , and  $S_0$ , the subset of  $S$  in which the minimal polynomials of the sequences are irreducible over  $GF(q)$ ;
- $P$ , the set of all polynomials over  $GF(q)$  with nonzero constant terms, and  $P_0$ , the subset of  $P$  which are the irreducible polynomials over  $GF(q)$ ; and
- $F$ , the set consisting of all finite fields.

By combining the results of Theorems 4.7 and 4.8, we have the following one-to-one correspondences among these sets:

$$\begin{aligned} S &\leftrightarrow P \\ S_0 &\leftrightarrow P_0 \leftrightarrow F \end{aligned}$$

These relations are also illustrated in Figure 4.10.

*Note.* It is worth pointing out that all irreducible polynomials of degree  $n$  over  $GF(q)$  generate finite fields with order  $q^n$ , and all these fields are isomorphic, namely  $GF(q^n)$ . The above one-to-one correspondence  $P_0 \leftrightarrow F$  represents the different computations in  $GF(q^n)$  invoked by different irreducible polynomials.

For the rest of this chapter, we will restrict ourselves to periodic sequences.

### C. Structure of $G(f)$ for $f$ Irreducible

**Definition 4.5** Two periodic sequences  $\mathbf{a} = \{a_i\}$  and  $\mathbf{b} = \{b_i\}$  are called (cyclically) shift equivalent if there exists an integer  $k$  such that

$$a_i = b_{i+k}, \quad \forall i \geq 0. \quad (4.13)$$

In this case, we write  $\mathbf{a} = L^k(\mathbf{b})$ , or simply  $\mathbf{a} \sim \mathbf{b}$ . Otherwise, they are called (cyclically) shift-distinct.

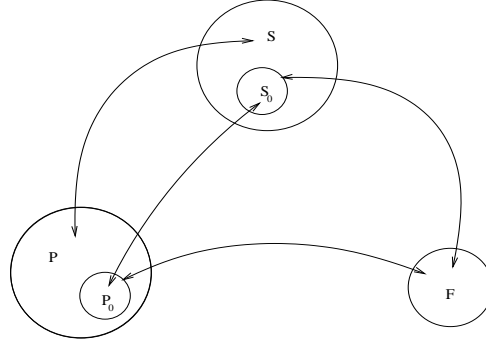


Figure 4.10: One-to-one correspondences among LFSR sequences, polynomials and finite fields

Note that  $\sim$  is an equivalent relation on  $V(F)$ . A set in which all sequences are shift-equivalent is called a *shift-equivalent class*. One shift-equivalent class of  $G(f)$  corresponds to one cycle of states in the state diagram of the LFSR with  $f(x)$ .

**Theorem 4.9** *Let  $f(x)$  be an irreducible polynomial over  $GF(q)$  of degree  $n$ . Then the number of shift-equivalent classes of non-zero LFSR sequences in  $G(f)$  is given by*

$$(q^n - 1)/\text{per}(f).$$

*Proof.* Let  $0 \neq \mathbf{a} \in G(f)$ . According to Proposition 1 - (b),  $f(L)L^k \mathbf{a} = 0 \implies L^k \mathbf{a} \in G(f)$ . Since  $f$  is irreducible,  $f$  is the minimal polynomial of  $\mathbf{a}$ . Using Theorem 4.8,  $\text{per}(\mathbf{a}) = \text{per}(f(x)) = r$ . So,  $L^r \mathbf{a} = \mathbf{a}$ . Let  $G_1$  denote the set consisting of  $\mathbf{a}$  and all its shifts, i.e.,  $G_1 = \{\mathbf{a}, L\mathbf{a}, \dots, L^{r-1}\mathbf{a}\}$ . Then every sequence in  $G_1$  has period  $r$ , and all the sequences in  $G_1$  are shift-equivalent. Next we take  $\mathbf{b} \in G(f)$  which does not belong to  $G_1$ . By performing the shift operator, we obtain  $G_2 = \{L^i \mathbf{b} \mid 0 \leq i \leq r-1\}$ , which is a shift-equivalent class in  $G(f)$  with the same cardinal number  $r$  as that of  $G_1$ . Continuing the process in this manner, we get  $(q^n - 1)/r$  shift-equivalent classes.  $\square$

In the language of the state diagram, this theorem shows that for an LFSR with an irreducible polynomial, in its state diagram there are  $(q^n - 1)/\text{per}(f)$  cycles with length  $\text{per}(f)$  and one cycle of length 1, i.e., the zero sequence. For example, let  $q = 2$ . Then  $f(x) = x^4 + x^3 + x^2 + x + 1 \in GF(2)[x]$  is irreducible over  $GF(2)$ . In  $G(f)$ , there are three shift-equivalent classes,  $G_i$ ,  $i = 1, 2$ , and 3, in which each sequence has period 5, as shown in Table 4.4. Thus, we have

$$G(f) = \{0\} \cup G_1 \cup G_2 \cup G_3.$$

The state diagram of this LFSR is shown in Figure 4.11.

As a consequence of Theorem 4.9, we have the following assertion.

Table 4.4: Shift-equivalent classes of  $G(f)$ 

$G_1$	$G_2$	$G_3$
00011	01010	11110
00110	10100	11101
01100	01001	11011
11000	10010	10111
10001	00101	01111

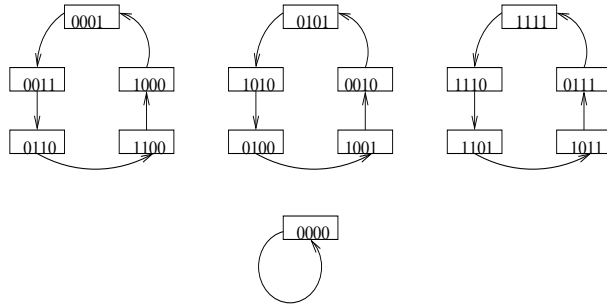


Figure 4.11: State diagram of the 4-stage LFSR

**Corollary 4.4** *With the notation of Theorem 4.9, if  $f(x)$  is primitive, then any nonzero sequence  $\mathbf{a}$  in  $G(f)$  has period  $q^n - 1$  and*

$$G(f) = \{L^i \mathbf{a} \mid 0 \leq i \leq q^n - 2\} \cup \{0\}.$$

*In particular, if  $q = 2$ , then any nonzero sequence  $\mathbf{a}$  in  $G(f)$  has period  $2^n - 1$  and*

$$G(f) = \{L^i \mathbf{a} \mid 0 \leq i \leq 2^n - 2\} \cup \{0\}.$$

**Definition 4.6** *A  $q$ -ary sequence generated by an  $n$ -stage LFSR is called a maximal length sequence if it has period  $q^n - 1$  (an  $m$ -sequence for short). In particular, if  $q = 2$ , a binary  $m$ -sequence is a sequence generated by an  $n$ -stage LFSR with period  $2^n - 1$ .*

**Example 4.9** For  $q = 2$ , (a)  $1001011 \dots$  is an  $m$ -sequence of period 7 with the minimal polynomial  $f(x) = x^3 + x + 1$ . (b) The sequences in Example 4.5 of Section 2 are  $m$ -sequences of period 15 with the minimal polynomial  $f(x) = x^4 + x + 1$ . (c) The sequence in Example 4.8 is an  $m$ -sequence of period 31 with the minimal polynomial  $f(x) = x^5 + x^3 + 1$ .

According to Corollary 4.4, in order to generate an  $m$ -sequence of period  $q^n - 1$  over  $F$  by an LFSR, we only need to select a primitive polynomial over  $F$  of degree  $n$  as the characteristic polynomial of this LFSR. In particular, for

$q = 2$ , we may sometimes choose primitive polynomials having the following form:

$$f(x) = x^n + x^k + 1.$$

This is called a *trinomial*. In this case, at the implementation level, we only need to use one  $n$ -stage shift register and one exclusive-or gate to generate an  $m$ -sequence with period  $2^n - 1$ . Thus, it is more efficient to use a primitive trinomial for generation of  $m$ -sequences than to use polynomials which have more non-zero coefficients. For primitive trinomials over  $GF(2)$ , researchers have computed all such polynomials of degree  $\leq 2000$ . However, the conjecture that infinitely many primitive trinomials over  $GF(2)$  exist is still open. (See the assignments at the end of this chapter.)

**Remark 4.3** For cryptographic applications of  $m$ -sequences, using trinomials (more general, by using primitive polynomials with low weights) to generate  $m$ -sequences are vulnerable to some types of correlation attacks. *So, trinomials are not recommended for this type of applications.*

## 4.4 Decomposition of LFSR Sequences

In this section, we will analyze the structures of  $G(f)$  when  $f(x)$  is a product of distinct irreducible polynomials over  $\mathbb{F}_q$ . The method used here can be easily generalized to the case for which  $f(x)$  satisfies  $f(0) \neq 0$ . Before we give the main result of this section, we establish the following lemma.

**Lemma 4.2** For any non-zero monic polynomials  $f(x), g(x) \in F[x]$ ,

- (a)  $G(f) \subset G(g)$  if and only if  $f(x) | g(x)$ .
- (b)  $G(f) \cap G(g) = G(d)$  where  $d = \gcd(f, g)$ .
- (c)  $G(f) + G(g) = G(h)$  where  $h = \text{lcm}[f, g]$ , the least common multiple of  $f$  and  $g$ .

*Proof.* (a) Assume that  $f(x) | g(x)$ . For any  $\mathbf{a} \in G(f)$ ,  $f(L)\mathbf{a} = 0$ .  $f(x) | g(x) \implies g(x) = t(x)f(x) \implies g(L)\mathbf{a} = t(L)f(L)\mathbf{a} = 0 \implies \mathbf{a} \in G(g) \implies G(f) \subset G(g)$ . Conversely, we choose  $\mathbf{a} \in G(f)$  such that the minimal polynomial of  $\mathbf{a}$  is  $f(x)$ . According to Theorem 4.5,  $g(L)\mathbf{a} = 0 \implies f(x) | g(x)$ .

(b) From (a),  $G(d) \subset G(f), G(d) \subset G(g) \implies G(d) \subset G(f) \cap G(g)$ . Assume that  $\mathbf{a} \in G(f) \cap G(g)$ , i.e.,

$$f(L)\mathbf{a} = 0 \text{ and } g(L)\mathbf{a} = 0.$$

Since  $d(x) = \gcd(f(x), g(x))$ , there exist two polynomials  $v(x)$  and  $u(x)$  such that

$$d(x) = u(x)f(x) + v(x)g(x).$$

Therefore  $d(L)\mathbf{a} = u(L)f(L)\mathbf{a} + v(L)g(L)\mathbf{a} = \mathbf{0} \implies G(f) \cap G(g) \subset G(d)$ . Together with  $G(d) \subset G(f) \cap G(g)$ , we get  $G(f) \cap G(g) = G(d)$ .

(c) From (a),  $G(f) \subset G(h)$ ,  $G(g) \subset G(h)$ . Thus

$$G(f) + G(g) \subset G(h).$$

Since the sets on both sides of the above inclusion are linear spaces, we only need to prove their dimensions are equal. Let  $\dim(V)$  denote the dimension of the linear space  $V$ . Notice that  $\text{lcm}[f, g]\text{gcd}(f, g) = fg$ . According to the dimension formula, we have

$$\begin{aligned} \dim(G(f) + G(g)) &= \dim(G(f)) + \dim(G(g)) - \dim(G(f) \cap G(g)) \\ &= \dim(G(f)) + \dim(G(g)) - \dim(G(d)) \\ &= \deg(f) + \deg(g) - \deg(d) \\ &= \deg(h) \\ &= \dim(G(h)). \end{aligned}$$

□

**Theorem 4.10** *Let  $f(x) = f_1(x) \cdots f_s(x)$  where the  $f_i$  are distinct irreducible polynomials over  $F$ ,  $s > 0$ . Then  $G(f)$  can be decomposed as a direct sum of subspaces  $G(f_i)$ ,  $1 \leq i \leq s$ , i.e.,*

$$G(f) = G(f_1) \oplus G(f_2) \oplus \cdots \oplus G(f_s).$$

*Proof.* We will use induction to prove this result. If  $s = 1$ , the result is true. Assume that the result is true for  $s = k - 1$ . For  $s = k$ , let  $f(x) = f_1(x) \cdots f_k(x)$  and  $g(x) = f_2(x) \cdots f_k(x)$ . Then  $f(x) = f_1(x)g(x)$  and  $\text{gcd}(f_1(x), g(x)) = 1$ . According to Lemma 4.2 (b) and (c),

$$\begin{aligned} G(f) &= G(f_1) + G(g), \\ G(f_1) \cap G(g) &= \{\mathbf{0}\}. \end{aligned}$$

Thus we have the decomposition of a direct sum:

$$G(f) = G(f_1) \oplus G(g). \tag{4.14}$$

Applying the induction hypothesis to  $G(g)$ , we get

$$G(g) = G(f_2) \oplus \cdots \oplus G(f_s).$$

Substituting this into (4.14), the result follows for  $s = k$ . Therefore the result is true for every  $s > 0$ .

□

We state the following result without proof.

**Fact 4.1** *Let  $f(x)$  be the same as in Theorem 4.10. Then the period of  $f(x)$  is equal to  $\text{lcm}[\text{per}(f_1), \dots, \text{per}(f_s)]$ , the least common multiple of the periods of the  $f_i(x)$ 's.*

From Fact 4.1 and Theorem 4.10, the following corollary is immediate.

**Corollary 4.5** *Let  $f(x)$  be the same as in Theorem 4.10, and let  $\mathbf{a} \in G(f)$  whose minimal polynomial is  $f(x)$ .*

(a)  $\mathbf{a}$  can be decomposed as

$$\mathbf{a} = \mathbf{a}_1 + \cdots + \mathbf{a}_s, 0 \neq \mathbf{a}_i \in G(f_i), i = 1, \dots, s.$$

(b) The period of  $\mathbf{a}$  is equal to  $\text{lcm}[\text{per}(\mathbf{a}_1), \dots, \text{per}(\mathbf{a}_s)]$  where  $\text{per}(\mathbf{a}_i) = \text{per}(f_i)$ .

(c) The linear span of  $\mathbf{a}$  is given by  $\sum_{i=1}^s \text{deg}(f_i)$ .

**Example 4.10** Let  $F = GF(2)$ , and let  $f_1(x) = x^4 + x + 1$ ,  $f_2(x) = x^2 + x + 1$ , and  $f(x) = f_1(x)f_2(x) = x^6 + x^5 + x^4 + x^3 + 1$ . Since  $f_1(x)$  and  $f_2(x)$  are coprime, according to Theorem 4.10 we have

$$G(f) = G(f_1) \oplus G(f_2) = G(x^4 + x + 1) \oplus G(x^2 + x + 1).$$

We take two nonzero sequences, say  $\mathbf{a}$  from  $G(x^4 + x + 1)$  and  $\mathbf{b}$  from  $G(x^2 + x + 1)$ . Notice that both  $f_1(x)$  and  $f_2(x)$  are primitive. Thus, any nonzero binary sequence in  $G(f)$  can be written as

$$cL^i\mathbf{a} + dL^j\mathbf{b}, 0 \leq i \leq 14, 0 \leq j \leq 2, c, d \in GF(2).$$

The number of all the nonzero sequences in  $G(f)$  is  $2^6 - 1 = 63$ , which can be classified according to their linear spans as follows:

- three shift-distinct classes of period 15 with linear span 6:  $\mathbf{a} + L^i\mathbf{b}$ ,  $0 \leq i \leq 2$ ;
- one shift-distinct class of period 15 with linear span 4:  $\mathbf{a}$ , an  $m$ -sequence of period 15; and
- one shift-distinct class of period 3 with linear span 2:  $\mathbf{b}$ , an  $m$ -sequence of period 3.

The state cycles for these sequences in  $G(f)$  are illustrated in Figure 4.12. For

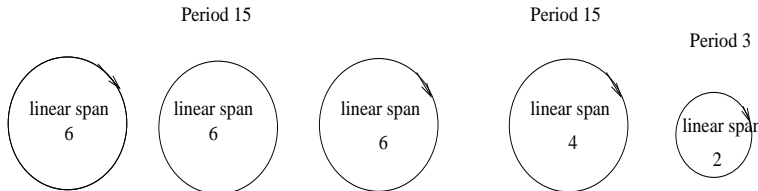


Figure 4.12: Cycle Structure of  $G(f)$

example, we have the following decomposition:

$$\begin{array}{r} \mathbf{a} = 100010011010111 \\ + \mathbf{b} = 011011011011011 \\ \hline \mathbf{c} = 111001000001100 \end{array}$$

where  $\mathbf{c} = \mathbf{a} + \mathbf{b}$  can be implemented by two LFSR implementations. (See Figures 4.13 and 4.14, which are equivalent.)

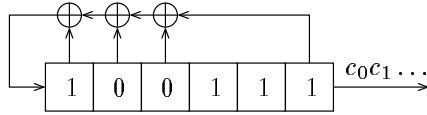


Figure 4.13: A 6-stage LFSR

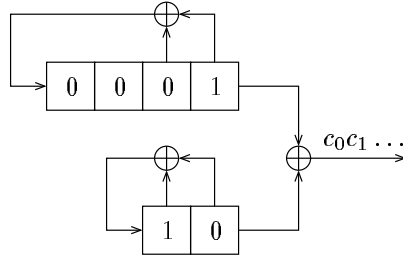


Figure 4.14: Decomposition of the LFSR in Figure 4.13

From Theorem 4.10, we understand that any LFSR sequence can be decomposed into a sum of LFSRs with irreducible characteristic polynomials (which may or may not be primitive polynomials). This implies that LFSRs with primitive or irreducible characteristic polynomials can be used as basic blocks for building more complicated *nonlinear* pseudorandom sequence generators.

## 4.5 The Matrix Representation

In Section 1, we saw that each state of an  $n$ -stage LFSR is a vector in the  $n$ -dimensional space  $F^n$ . The shift register is then a linear operator which changes the current state into its successor vector according to the feedback. In other words, the transformation of each non-zero sequence in  $G(f)$ , from state

$$(a_k, a_{k+1}, \dots, a_{k+n-1})$$

to its successor

$$(a_{k+1}, a_{k+2}, \dots, a_{k+n})$$

can be regarded as a linear operator on  $F^n$ . It is a familiar fact that a linear operator, operating on an  $n$ -dimensional vector space  $F^n$ , is conveniently studied when it is represented by an  $n \times n$  matrix. We know that

$$a_{n+k} = c_0 a_k + c_1 a_{k+1} + \cdots + c_{n-1} a_{k+n-1}, k \geq 0.$$

Hence, a shift register matrix takes the form

$$M = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & c_0 \\ 1 & 0 & 0 & \cdots & 0 & c_1 \\ 0 & 1 & 0 & \cdots & 0 & c_1 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & 1 & c_{n-1} \end{bmatrix}$$

and

$$\begin{aligned} (a_{k+1}, a_{k+2}, \dots, a_{k+n}) &= (a_k, a_{k+1}, \dots, a_{k+n-1})M \\ &= (a_{k-1}, a_k, \dots, a_{k-1+n-1})M^2 \\ &= \cdots \\ &= (a_0, a_1, \dots, a_{n-1})M^{k+1}. \end{aligned}$$

We called the matrix  $M$  a *state transform matrix* of the LFSR. Note that  $\det(M) = (-1)^n c_0$ . Thus  $M$  is invertible if and only if  $c_0 \neq 0$ .

As an example, the state transform matrix of the 3-stage LFSR in Example 4.2 in Section 4.1 is given by

$$M = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

In particular,

$$\begin{aligned} (001) &= (100) \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \\ (010) &= (001) \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}. \end{aligned}$$

The *characteristic equation* of the matrix  $M$  is defined by

$$\begin{aligned} f(x) &= \det |M - xI| = \begin{vmatrix} x & 0 & 0 & \cdots & 0 & -c_0 \\ -1 & x & 0 & \cdots & 0 & -c_1 \\ 0 & -1 & x & \cdots & 0 & -c_2 \\ & & \vdots & & & \\ 0 & 0 & 0 & \cdots & -1 & x - c_{n-1} \end{vmatrix} \\ &= x^n - c_{n-1}x^{n-1} - \cdots - c_1x - c_0 \end{aligned}$$

where  $I$  is the identity matrix. Notice that the characteristic polynomial  $f(x)$  is the same as the characteristic polynomial of the LFSR. Under this representation, determining the periodicity of the LFSR is equivalent, except for degenerate cases, to finding the smallest positive integer  $r$  such that  $M^r = I$ . (Note.  $r$  is also called the *order of the matrix*  $M$  in the group consisting of all invertible matrices whose entries are taken from  $F$ .)

A well-known theorem of matrix theory (the ‘‘Cayley-Hamilton Theory’’) asserts that every matrix formally satisfies its characteristic equation: thus  $f(M) = 0$ , where ‘‘0’’ here is the matrix of all zeros. If  $f(x)$  divides  $x^r - 1$ , then  $M$  is a root of  $x^r - I = 0$ . In other words, if  $f(x)$  divides  $x^r - 1$ , then  $M^r = I$ . Conversely, if  $f(x)$  is irreducible, it divides *every* polynomial which has the root  $M$  in common with it, and will divide  $x^r - 1$  if  $M^r = I$ . Roughly, this is a matrix theory proof of Theorem 4.7. The rest of the theory of LFSR sequences can be done entirely by matrix theory, and is frequently done so in the literature.

## 4.6 Trace Representation of LFSRs

In this section, we present trace representations for non-zero sequences in  $G(f)$  when  $f(x)$  is irreducible. We will discuss trace representations for NLFSR sequences with period  $N$  where  $N$  is a factor of  $q^n - 1$  in Chapter 6 in terms of the Fourier transforms of periodic sequences. Starting with this section, we will simply denote the finite field  $GF(Q) = \mathbb{F}_Q$  and  $Tr_{\mathbb{F}_{q^n}/\mathbb{F}_q}(x) = Tr(x) = x + x^q + \cdots + x^{q^{n-1}}$ , the trace function from  $\mathbb{F}_{q^n}$  to  $\mathbb{F}_q$ , if the context is clear. (For more properties of the trace functions, see Section 5 of Chapter 3. )

### A. Trace Representation

Let

$$f(x) = x^n + c_{n-1}x^{n-1} + \cdots + c_1x + c_0$$

be an irreducible polynomial over  $\mathbb{F}_{q^n}$  of degree  $n$ , and let  $\alpha$  be a root of  $f(x)$ , i.e.,

$$\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0 = 0. \quad (4.15)$$

Then we can construct a finite field  $\mathbb{F}_{q^n}$  with  $f(x)$  as a defining polynomial.

**Lemma 4.3** *Let  $\mathbf{a} = \{a_i\}$  whose elements are given by*

$$a_i = Tr(\beta\alpha^i), \quad i \geq 0, \beta \in \mathbb{F}_{q^n} \quad (4.16)$$

*Then  $\mathbf{a} \in G(f)$ .*

*Proof.* If  $\beta = 0$ , then  $\mathbf{a}$  is a zero sequence, so  $\mathbf{a} \in G(f)$ . If  $\beta \neq 0$ , then

$$\begin{aligned} & a_{k+n} + c_{n-1}a_{k+n-1} + \cdots + c_1a_{k+1} + c_0a_k \\ = & Tr(\beta\alpha^{k+n}) + c_{n-1}Tr(\beta\alpha^{k+n-1}) + \cdots + c_1Tr(\beta\alpha^{k+1}) + c_0Tr(\beta\alpha^k) \end{aligned}$$

$$\begin{aligned}
& \text{(by (4.16))} \\
& = \operatorname{Tr}(\beta\alpha^k(\alpha^n + c_{n-1}\alpha^{n-1} + \cdots + c_1\alpha + c_0)) \\
& = \operatorname{Tr}(\beta\alpha^k\mathbf{0}) = 0, \forall k \geq 0 \quad \text{(by (4.15))} \\
\Rightarrow & f(L)\mathbf{a} = \mathbf{0} \\
\Rightarrow & \mathbf{a} \in G(f).
\end{aligned}$$

□

**Theorem 4.11** *With the above notation, for any sequence  $\mathbf{a} = \{a_i\} \in G(f)$ , there exists some  $\beta \in \mathbb{F}_{q^n}$  such that*

$$a_i = \operatorname{Tr}(\beta\alpha^i), \quad \forall i \geq 0. \quad (4.17)$$

*Proof.* Since there are  $q^n$  different sequences in  $G(f)$ , from Lemma 4.3, we have  $\mathbf{a} = \{a_i\}$ , with  $a_i = \operatorname{Tr}(\beta\alpha^i)$ , belonging to  $G(f)$  for any  $\beta \in \mathbb{F}_{q^n}$ . So, we only need to prove that two sequences related to two different elements in  $\mathbb{F}_{q^n}$  are different. Since the trace function is a linear function from  $\mathbb{F}_{q^n}$  to  $\mathbb{F}_q$ , this reduces to proving that if  $\beta \neq 0$ , then  $\mathbf{a} \neq \mathbf{0}$ . If not, we have  $\mathbf{a} = \mathbf{0}$ . We may write down the first  $n$  terms as follows:

$$\begin{aligned}
0 = \operatorname{Tr}(\beta) &= \beta + \beta^q + \cdots + \beta^{q^{n-1}} \\
0 = \operatorname{Tr}(\beta\alpha) &= \beta\alpha + \beta^q\alpha^q + \cdots + \beta^{q^{n-1}}\alpha^{q^{n-1}} \\
&\vdots \\
0 = \operatorname{Tr}(\beta\alpha^{n-1}) &= \beta\alpha^{n-1} + \beta^q\alpha^{(n-1)q} + \cdots + \beta^{q^{n-1}}\alpha^{(n-1)q^{n-1}}.
\end{aligned}$$

This is a system of  $n$  linear equations in  $n$  unknowns  $\beta, \beta^q, \dots, \beta^{q^{n-1}}$ . We rewrite the above  $n$  linear equations in the following matrix form

$$\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ \alpha & \alpha^q & \alpha^{q^2} & \cdots & \alpha^{q^{n-1}} \\ \alpha^2 & \alpha^{2q} & \alpha^{2q^2} & \cdots & \alpha^{2q^{n-1}} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \alpha^{n-1} & \alpha^{(n-1)q} & \alpha^{(n-1)q^2} & \cdots & \alpha^{(n-1)q^{n-1}} \end{bmatrix} \begin{bmatrix} \beta \\ \beta^q \\ \beta^{q^2} \\ \vdots \\ \beta^{q^{n-1}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Since  $f(x)$  is irreducible,

$$\alpha, \alpha^q, \alpha^{q^2}, \dots, \alpha^{q^{n-1}}$$

are all the  $n$  roots of  $f(x)$ , which are distinct. Therefore, the above matrix is a Vandemonde matrix. Hence, it is invertible. Thus  $\beta = 0$ , which is a contradiction to the assumption.

□

The formula (4.17) is called the *trace representation* for LFSR sequences with irreducible characteristic (or feedback) polynomials.

From Theorem 4.11, the following corollary is immediate.

**Corollary 4.6** *Let  $\mathbf{a}$  be a sequence over  $\mathbb{F}_q$ . Then  $\mathbf{a}$  is an  $m$ -sequence with period  $q^n - 1$  if and only if the elements of  $\mathbf{a}$  can be represented by,*

$$a_i = \text{Tr}(\beta\alpha^i), \forall i \geq 0, 0 \neq \beta \in \mathbb{F}_{q^n}$$

where  $\alpha$  is a primitive element in  $\mathbb{F}_{q^n}$ .

**Example 4.11** Let  $q = 2$ .

(a) Let  $f(x) = x^4 + x^3 + x^2 + x + 1$ , which defines  $\mathbb{F}_{2^4}$ , and let  $\alpha$  be a root of  $f(x)$  in  $\mathbb{F}_{2^4}$ , i.e.,  $\alpha^4 + \alpha^3 + \alpha^2 + \alpha + 1 = 0$ . Then for  $\mathbf{a} = (10001) \in G(f)$ , the trace representation of  $\mathbf{a}$  is given by

$$a_i = \text{Tr}(\beta\alpha^i), i = 0, 1, \dots$$

where  $\beta = 1 + \alpha$ .

(b) Let  $f(x) = x^3 + x + 1$ , and  $\mathbb{F}_{2^3}$  be defined by  $f(\alpha) = 0$ . Then

$$\mathbf{a} = (1001011) \in G(f) \leftrightarrow \text{the trace representation } a_i = \text{Tr}(\alpha^i), i = 0, 1, \dots$$

(c) Let  $f(x) = x^4 + x + 1$ , and  $\mathbb{F}_{2^4}$  be defined by  $f(\alpha) = 0$ . Then

$$\begin{aligned} \mathbf{a} &= (10011010111000) \in G(f) \leftrightarrow \\ &\text{the trace representation } a_i = \text{Tr}(\alpha^3\alpha^i), i = 0, 1, \dots \end{aligned}$$

(See Example 3.14 in Chapter 3 for the computation of  $\text{Tr}(x)$  in  $\mathbb{F}_{2^3}$  and  $\mathbb{F}_{2^4}$ .)

## B. $S$ -Decimation

**Definition 4.7** *Let  $f(x)$  be an irreducible polynomial over  $\mathbb{F}_q$  of degree  $n$  and  $s$  be a positive integer. For  $\mathbf{a} \in G(f)$ , if the elements of a sequence  $\mathbf{b} = \{b_i\}$  are defined by*

$$b_i = a_{si} \forall i \geq 0,$$

then  $\mathbf{b}$  is called an  $s$ -decimation sequence of  $\mathbf{a}$ , denoted by  $\mathbf{b} = \mathbf{a}^{(s)}$ .

For example, let  $\mathbf{a} = (1001011)$ . For  $\mathbf{b} = \mathbf{a}^{(3)}$ , i.e.,  $b_i = a_{3i}, i = 0, 1, \dots$ , we have  $\mathbf{b} = (1110100)$ .

**Theorem 4.12** *Let  $f(x)$  be an irreducible polynomial over  $\mathbb{F}_q$  of degree  $n$  and let  $s$  be a positive integer. Let  $\alpha$  be a root of  $f(x)$  in the extension  $\mathbb{F}_{q^n}$ . For  $0 \neq \mathbf{a} \in G(f)$ , if the  $s$ -decimation  $\mathbf{a}^{(s)} \neq \mathbf{0}$ , then the minimal polynomial of  $\mathbf{a}^{(s)}$  is equal to the minimal polynomial of  $\alpha^s$  over  $\mathbb{F}_q$ .*

*Proof.* Since  $\mathbf{a} \in G(f)$ , then there exists  $\beta \in \mathbb{F}_{q^n}$  such that  $a_i = \text{Tr}(\beta\alpha^i)$ ,  $\forall i \geq 0$ . Therefore

$$\mathbf{a}^{(s)} = (\text{Tr}(\beta), \text{Tr}(\beta\alpha^s), \text{Tr}(\beta\alpha^{2s}), \dots).$$

Let  $\gamma = \alpha^s$ . Then

$$\mathbf{a}^{(s)} = (Tr(\beta), Tr(\beta\gamma), Tr(\beta\gamma^2), \dots).$$

Let  $g(x)$  be the minimal polynomial of  $\gamma$ . From Lemma 4.3,  $\mathbf{a}^{(s)} \in G(g)$ . Since  $g(x)$  is irreducible, if  $\mathbf{a}^{(s)} \neq 0$ , then  $g(x)$  is the minimal polynomial of  $\mathbf{a}^{(s)}$ . □

According to Theorem 4.8,  $per(\mathbf{a}) = per(f) = ord(\alpha)$ . Then the period of  $s$ -decimation  $\mathbf{a}^{(s)}$  is equal to  $ord(\alpha^s) = ord(\alpha)/(s, ord(\alpha))$ . Hence

$$per(\mathbf{a}^{(s)}) = per(\mathbf{a})/(s, per(\mathbf{a})).$$

**Corollary 4.7** *With the above notation, if  $f(x)$  is primitive, then  $0 \neq \mathbf{a} \in G(f)$  is an  $m$ -sequence. In this case, an  $s$ -decimation of  $\mathbf{a}$  is also an  $m$ -sequence if and only if  $\gcd(s, q^n - 1) = 1$ . Moreover, the number of shift-distinct  $m$ -sequences over  $\mathbb{F}_q$  of period  $q^n - 1$  is given by*

$$\phi(q^n - 1)/n$$

*which is equal to the number of the primitive polynomials over  $\mathbb{F}_q$  of degree  $n$ . (Here  $\phi(\cdot)$  is Euler's phi-function.)*

**Example 4.12** Let  $\mathbb{F}_{2^4}$  be defined by  $f(x) = x^4 + x + 1$ , and let  $\alpha$  be a root of  $f(x)$  in  $\mathbb{F}_{2^4}$ . Let  $f_{\alpha^s}(x)$  be the minimal polynomial of  $\alpha^s$  over  $\mathbb{F}_2$ . We select  $\mathbf{a} \in G(f)$  as follows:

$$\mathbf{a} = (100010011010111) = (Tr(\alpha^{14}), Tr(\alpha^{14}\alpha), Tr(\alpha^{14}\alpha^2), \dots, Tr(\alpha^{14}\alpha^{14})).$$

All possible distinct decimated sequences of  $\mathbf{a}$  (up to shift-equivalence) and their corresponding minimal polynomials and periods are listed as follow.

All decimations of  $\mathbf{a}$  up to shift-equivalence

$s$	$\mathbf{a}^{(s)}$	Period $\mathbf{a}^{(s)}$	Minimal polynomial $f_{\alpha^s}(x)$	Period $f_{\alpha^s}(x)$	Order $\alpha^s$
1	100010011010111	15	$x^4 + x + 1$	15	15
3	10001	5	$x^4 + x^3 + x^2 + x + 1$	5	5
5	101	3	$x^2 + x + 1$	3	3
7	111010110010001	15	$x^4 + x^3 + 1$	15	15

## 4.7 Generating Functions of LFSRs

In this section, we will introduce another method to represent a periodic LFSR sequence, called a *generating function* of the LFSR sequence. This method is frequently used in studying the combinatorial structure of objects. Assume that  $\mathbf{a} = \{a_i\}$  is a periodic sequence over  $\mathbb{F}_q$ , generated by  $f(x) = x^n + \sum_i c_i x^i$  with  $c_0 \neq 0$ ,  $c_i \in \mathbb{F}_q$ . We associate the sequence  $\mathbf{a}$  with a polynomial  $a(x) = \sum_{i=0}^{\infty} a_i x^i$ . Then we have

$$a_{n+k} + \sum_i c_i a_{k+i} = 0, k = 0, 1, \dots. \tag{4.18}$$

Next we rewrite the linear recursive relation given by (4.18) into polynomial form. We denote the product of  $a(x)$  and  $f^{-1}(x)$ , the reciprocal of  $f(x)$ , by  $d(x) = \sum_{i=0}^{\infty} d_i x^i$ , i.e.,

$$a(x)f^{-1}(x) = \sum_{i=0}^{\infty} d_i x^i \quad (4.19)$$

From (4.18), we have

$$d_j = a_j + \sum_i c_i a_{j-n+i} = 0, j = n, n+1, \dots$$

Thus the right-hand side of (4.19) becomes

$$\sum_{i=0}^{\infty} d_i x^i = \sum_{i=0}^{n-1} d_i x^i \quad (4.20)$$

where

$$\begin{aligned} d_0 &= c_0 a_0 \\ d_1 &= c_0 a_1 + c_1 a_0 \\ &\vdots \\ d_i &= \sum_{k+j=i} a_k a_j \\ &\vdots \\ d_{n-1} &= \sum_{k+j=n-1} a_k a_j. \end{aligned}$$

In other words, the  $d_i$ 's can be determined by the initial state  $(a_0, a_1, \dots, a_{n-1})$  of the LFSR and the coefficients of the characteristic polynomial. We may write the above relation in the following matrix form.

$$\begin{bmatrix} d_0 \\ d_1 \\ d_2 \\ \vdots \\ d_{n-1} \end{bmatrix} = \begin{bmatrix} c_0 & 0 & 0 & \cdots & 0 \\ c_1 & c_0 & 0 & \cdots & 0 \\ c_2 & c_1 & c_0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{n-1} & c_{n-2} & c_{n-3} & \cdots & c_0 \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ \vdots \\ a_{n-1} \end{bmatrix} \quad (4.21)$$

Consequently, (4.19) and (4.20) yield

$$a(x) = \frac{d(x)}{f^{-1}(x)} \quad (4.22)$$

where  $\deg(d(x)) \leq n-1$ . Usually,  $\frac{d(x)}{f^{-1}(x)}$  is called a *generating function* of the sequence  $\mathbf{a}$ . If an initial state of the LFSR is given, then  $d(x)$  can be determined

by (4.21). Let the matrix in (4.21) be  $A$ . Then  $A$  is invertible. Thus, if  $d(x)$  is any polynomial over  $GF(q)$  of degree less than  $n$ , then the corresponding initial state of  $\mathbf{a}$  can be determined by

$$\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} = A^{-1} \begin{bmatrix} d_0 \\ d_1 \\ \vdots \\ d_{n-1} \end{bmatrix}.$$

Therefore, in terms of the generating function method, the set  $G(f)$  which consists of all sequences generated by  $f(x)$  is given by

$$G(f) = \left\{ \frac{d(x)}{f^{-1}(x)} \text{ such that } d(x) \in GF(q)[x] \text{ and } \deg(d) < n \right\}.$$

**Example 4.13** Let  $q = 2$ ,  $f(x) = x^4 + x + 1$  and  $d(x) = x^3 + 1$ . Then  $f^{-1}(x)$ , the reciprocal polynomial of  $f(x)$ , is given by  $f^{-1}(x) = x^4 + x^3 + 1$ . The fraction

$$\frac{d(x)}{f^{-1}(x)} = \frac{x^3 + 1}{x^4 + x^3 + 1} = 1 + x^4 + x^7 + x^8 + x^{10} + x^{12} + x^{13} + x^{14} + \dots$$

gives a sequence 100010011010111 with an initial state 1000 and period 15.

**Note.**

In the literature, the maximal length sequences or  $m$ -sequences are also called *pseudo-noise (PN) sequences* when  $q = 2$ . S. Golomb studied these sequences in his popular book [63], from which our discussion of binary LFSR sequences is taken. For general  $q$ , a power of a prime, most of the results on  $m$ -sequences appeared in Zierler's work [191]. Using the set  $G(f)$  to discuss the structure of LFSR sequences is also due to Zierler. For this approach, see also Shishung Ding's book which was published in 1982 in Chinese [37], McEliece's book [133], and Lidl and Niederreiter's book [123]. Another early reference for  $q$ -ary  $m$ -sequences was Selmer's book [167]. For correlation attacks on pseudo-random sequence generators for which  $m$ -sequences are generated by primitive polynomials with low weights, see [137].

## Exercises for Chapter 4

1. Given a 3-stage shift register with the boolean feedback function  $f(x_0, x_1, x_2) = x_0 + x_1x_2$ :
  - (a) draw the state diagram of the FSR .
  - (b) if the initial state is set as  $(a_0, a_1, a_2) = (011)$ , determine the output sequence and the period of the sequence.
2. Let  $f(x_0, x_1, \dots, x_{n-1})$  be a boolean function in  $n$  variables which is employed as the feedback function of a shift register. Prove that the cycles in the state diagram have no branch points if and only if the feedback function can be decomposed into

$$f(x_0, x_1, \dots, x_{n-1}) = x_0 + g(x_1, \dots, x_{n-1})$$

where  $g(x_1, \dots, x_{n-1})$  is a boolean function in  $n - 1$  variables. (Hint: The cycles in the state diagram have no branch points if and only if two distinct state vectors have distinct successors. If  $(a_0, a_1, \dots, a_{n-1})$  and  $(b_0, b_1, \dots, b_{n-1})$  differ in any component other than the first, then their successors  $(a_1, \dots, a_n)$  and  $(b_1, \dots, b_n)$  are still distinct. So, one only needs to consider whether  $(a_0, a_1, \dots, a_{n-1})$  and  $(a_0 + 1, b_1, \dots, b_{n-1})$  have distinct successors. )

3. Design an LFSR over  $GF(2)$  for implementation of the linear recurrence relation

$$a_{5+k} = a_{3+k} + a_k, k = 0, 1, \dots, .$$

Determine the characteristic polynomial  $f(x)$  of the sequence and the number of sequences in  $G(f)$ . Write the first 50 bits of the output sequence with a nonzero initial state.

4. Design an LFSR over  $GF(2)$  for implementation of the linear recurrence relation

$$a_{6+k} = a_{1+k} + a_k, k = 0, 1, \dots.$$

Determine the characteristic polynomial  $f(x)$  of the sequence and the number of sequences in  $G(f)$ .

5. Design an LFSR over  $GF(2)$  for implementation of the linear recurrence relation

$$a_{7+k} = a_{1+k} + a_k, k = 0, 1, \dots.$$

Determine the characteristic polynomial  $f(x)$  of the sequence and the number of sequences in  $G(f)$ .

6. Construct two different (shift-distinct) de Bruijn sequences with period 16.

7. Let  $f(x) = x^5 + x^4 + 1$  over  $\mathbb{F}_2$  be the characteristic polynomial of a 5-stage LFSR.
- Write the first 50 bits of the output sequence with the initial state 00101 and determine the period and the minimal polynomial of the sequence. (Hint:  $f(x) = (x^3 + x + 1)(x^2 + x + 1)$ .)
  - Write the first 50 bits of the output sequence with the initial state 01000 and determine the period and the minimal polynomial of the sequence.
  - Determine the number of sequences in  $G(f)$  and draw the state diagram.
8. Design an LFSR over  $GF(2)$  which generates a binary  $m$ -sequence with period 1023.
9. Determine the number of LFSRs over  $GF(2)$  which generate a binary  $m$ -sequence with period  $2^8 - 1 = 255$ .
10. Let  $\alpha$  be a primitive element of  $\mathbb{F}_{2^n}$ . Let  $\mathbf{a} = \{a_i\}$  be a binary  $m$ -sequence of degree  $n$  of period  $2^n - 1$  whose elements are given by  $a_i = Tr(\beta\alpha^i)$ ,  $\forall i \geq 0$  where  $Tr(x) = x + x^2 + \cdots + x^{2^{n-1}}$ , the trace function from  $\mathbb{F}_{2^n}$  to  $\mathbb{F}_2$ ,  $\beta \in F$ ,  $\forall i$ . Prove that  $\mathbf{a}$  has the following property:  $a_{2i} = a_i$ ,  $\forall i \geq 0$ , if and only if  $\beta = 1$ . (This property is also referred to as *constant-on-cosets*).
11. Find the initial state of an  $m$ -sequence which is generated by the primitive polynomial  $f(x) = x^7 + x + 1$  which satisfies the property of being constant-on-cosets.
12. An  $m$ -sequence of period 31 with the minimal polynomial  $f(x) = x^5 + x^3 + 1$  is given by:

$$\mathbf{a} = (1000010101110110001111100110100).$$

Determine its 7-decimation sequence  $\mathbf{a}^{(7)}$  and compute the minimal polynomial of this sequence.

13. Let  $\mathbf{a}$  be an  $m$ -sequence over  $GF(2)$  of period 511. Determine the periods and linear spans of the following decimation sequences.

$$\mathbf{a}^{(r)}, \text{ where } r = 2, 3, 5, 14, 146.$$

(It is not necessary to generate the  $m$ -sequences.)

14. An  $m$ -sequence  $\mathbf{a}$  of period 127 with the minimal polynomial  $f(x) = x^7 + x + 1$  is given by:

```

1 0 0 0 0 0 0 1 0 0 0 0 0 1 1 0 0 0 0 1 0 1 0 0 0 1 1 1 1 0
0 1 0 0 0 1 0 1 1 0 0 1 1 1 0 1 0 1 0 0 1 1 1 1 1 0 1 0 0 0

```

```

0 1 1 1 0 0 0 1 0 0 1 0 0 1 1 0 1 1 0 1 0 1 1 0 1 1 1 1 0 1
1 0 0 0 1 1 0 1 0 0 1 0 1 1 1 0 1 1 1 0 0 1 1 0 0 1 0 1 0 1
0 1 1 1 1 1 1

```

- (a) The set  $\Gamma$  consisting all the coset leaders modulo 127 is given by

$$\Gamma = \{1, 3, 5, 7, 9, 11, 13, 15, 19, 21, 23, 27, 29, 31, 43, 47, 55, 63\}.$$

Find the individual terms  $a_i, \forall i \in \Gamma$ .

- (b) Verify that the sequence  $\mathbf{a}$  is constant-on-cosets.

The following three unsolved problems and conjectures related to shift register sequences are proposed by S.W. Golomb.

15. (**Golomb's Conjecture**) There exist infinitely many  $n$  such that  $f(x) = x^n + x^k + 1$  is a primitive polynomial over  $GF(2)$ , where  $1 \leq k < n$  and  $k$  may differ for different  $n$ .

*Notes Regarding Golomb's Conjecture:*

- 1) It is easy to show that there are infinitely many irreducible trinomials over  $GF(2)$ . For example,  $x^{2 \cdot 3^k} + x^{3^k} + 1$  is irreducible for every  $k = 0, 1, 2, \dots$ , with period  $3^{k+1}$ , but it is primitive only for  $k = 0$ .
  - 2) Can you prove that there are infinitely many primitive polynomials over  $GF(2)$  which have no more than  $t$  terms, for any fixed integer  $t$ ? This would be a new result.
  - 3) It seems to be true for every degree  $n \geq 5$  that there are primitive 5-term polynomials (pentanomials) of degree  $n$  over  $GF(2)$ . This would be a far stranger result than 2) above.
16. It is known that an  $n$ -stage shift register with the boolean feedback function  $f(x_0, x_1, \dots, x_{n-1})$  produce "pure cycles" (without "branches") if and only if we can write  $f(x_0, x_1, \dots, x_{n-1}) = x_0 + g(x_1, \dots, x_{n-1})$  (see Problem 2). It is also known that for  $n > 2$ , the number of (pure) cycles, for this case, is *even* (*odd*) if and only if the number of *ones* in the truth table for  $g(x_1, \dots, x_{n-1})$  is *even* (*odd*). Are there other general, qualitative results about the cycles of a nonlinear shift register that can be similarly and simply stated in terms of the boolean functions  $f(x_0, x_1, \dots, x_{n-1})$  or  $g(x_1, \dots, x_{n-1})$ ?
17. If  $f(x_0, x_1, \dots, x_{n-1}) = x_0 + g(x_1, \dots, x_{n-1})$ , and  $g(0, \dots, 0) = 0$ , then the "all zero state" forms a pure cycle by itself. What further conditions on  $g$  will guarantee that the remaining  $2^n - 1$  states lie on a single cycle of the shift register? (It is not necessary to find conditions for all  $2^{2^n - 1} - n$  such sequences. It would be very interesting to find conditions for even a small family of *nonlinear* sequences of period  $2^n - 1$ .)