Cryptographic Counter Generators

Cunsheng Ding

Turku Centre for Computer Science
TUCS Dissertations, No 4

Cryptographic Counter Generators

Cunsheng Ding

To be presented, with the permision of the Faculty of Mathematics and Natural Sciences of the University of Turku, for public criticism in the Auditorium of the Computer Science Department on April 26th, 1997, at 12 noon.

University of Turku
Department of Mathematics
FIN-20014 Turku, Finland

January 1997

ISBN 951-650-929-0
ISSN 1239-1883
Painosalama Oy

Acknowledgements

First of all, I would like to thank Professor Arto Salomaa for giving me the opportunity to do my Ph.D under his supervision. I have benefited much from reading some of his papers and books, and have learned personally many things from him. Arto has already had an important influence on my life, and more of his influence is bound to come.

It is my great pleasure to thank the Turku Centre for Computer Science, officially listed as one of the top twelve research centers in Finland, for providing me with excellent working conditions and financial support during my study. I thank all staff members in TUCS who have helped me.

Special thanks are due to Professors Helmut Jürgensen and Valtteri Niemi for reviewing my thesis and giving detailed comments and suggestions that have much improved it.

I would like to thank Professor Aimo Tietäväinen for discussions on sequences. I have also benefited from seminars on sequences, coding theory and function fields organized by his group.

I gratefully acknowledge discussions with Ross Anderson, Tom Cusick, Andrew Klapper, Tor Helleseth, Juhani Karhumäki, Torleiv Kløve, Tero Laihonen, Jyrki Lahtonen, James Massey, Alexandru Mateescu, Kaisa Nyberg, Gheorghe Păun, Ari Renvall, and others I have forgotten to mention here. Thanks are also due to Elisa Mikkola for her help during my study.

Again, I thank Weijuan and Xiang for their support and for sharing both the good and bad times in my life.

January, 1997

Cunsheng Ding

Contents

Chapter 1

Introduction

Ciphers, which are usually divided into block and stream ciphers, have been used for millennia to safeguard military and diplomatic communications. In our information-sensitive society, the widespread recognition of an urgent need for means to provide for the integrity of information has led to much of the public activity of research into cryptology. The seminal idea of public-key cryptography has put new blood into cryptography[16, 59, 60].

In spite of the active research and progress in cryptology, it may be said that the theory of cryptography is still not mature. The main problem is the measure of security [40, 41]. As far as security is concerned, the best one can do is to analyze the security of a cipher system with respect to known specific attacks.

Stream ciphers belong to the classical system, they still remain to be the most employed ones in practice. The main reason may be that the theory of stream ciphers is much more analytical, while only a few aspects about block ciphers can be mathematically measured and analyzed. Another important reason may be that stream ciphers usually destroy statistical properties of the plaintext, while block ciphers do not. The most popular stream ciphers are the additive ones, since they are much more analytical. The key problem for additive stream ciphers is the design and analysis of the keystream generators.

The main topic of this thesis is the counter generators, which are conceptually simple. Some counter generators are cryptographically attractive since many of their security aspects are consistent by which we mean that there is no trade-off within those aspects. This can be illustrated by the natural sequence generator analyzed in this thesis.

The aim of this thesis is to give a detailed study of some counter generators, and to point out some connections between number theory and stream ciphers. The main topic is the natural sequence generator. The thesis is organized as follows.

Chapter 2 presents some basic notions and notations about stream ciphers and sequences. A brief classification of attacks on stream ciphers is given. Some general security problems of stream ciphers are discussed.

Chapter 3 introduces some counter generators and presents the differential cryptanalysis on the natural sequence generator. One of the main tasks of this chapter is to prove the equivalence among seven cryptographic aspects of the natural sequence generator. Another main task is to present a basic theorem, which serves as a bridge between some number-theoretic problems and some problems of stream ciphers.

Chapter 4 is mainly concerned with the relations between difference and almost difference sets and a number of cryptographic aspects of cryptographic functions such as the difference property, nonlinearity, information stability, and linear approximation. Those relations are essential not only to counter generators, but also to other generators.

Chapter 5 is about the cyclotomic generators of order $2k$. The analysis of those generators serves also as an example of how to prove some cryptographic attributes of other cyclotomic generators described in later chapters.

Chapter 6 presents two-prime generators and relates some cryptographic aspects of those generators to the generalized cyclotomic numbers. The most promising two-prime generator is the one based on twin primes. Two-prime generators of order 2 and 4 are the main topic of this chapter.

Chapter 7 first introduces a kind of new cyclotomic numbers and then presents the prime-square generator. Finally, the linear and sphere complexity of the generator are discussed.

Chapter 8 first points out a number of cryptographic ideas behind the cyclotomic generators, then mentions some of the relations between some number-theoretic problems and cryptographic problems. The presentation here is to provide information, not to formulate the relation in detail, since describing the relations in detail must involve quite many number-theoretic notions and theorems. As a summary of the cyclotomic generators presented in the foregoing chapters, a number of cryptographic attributes of those generators are collected.

This thesis is also planned to bridge the gaps between some cryptographic problems and some number-theoretic problems. After the set-up of some bridges, the application of those number-theoretic problems in cryptographic generators is investigated. On the other hand, new number-theoretic problems are proposed from the cryptographic point of view. It is seen that many number-theoretic problems are related to cryptographic sequences.

A large part of the material of the thesis comes from my papers [23, 24, 21, 29, 30], and is related with my paper [19] and my monograph [22] jointly with W. Shan and G. Xiao as well as my book [25] jointly with G. Xiao. Other materials are from my recent work.

Chapter 2

Stream Ciphers

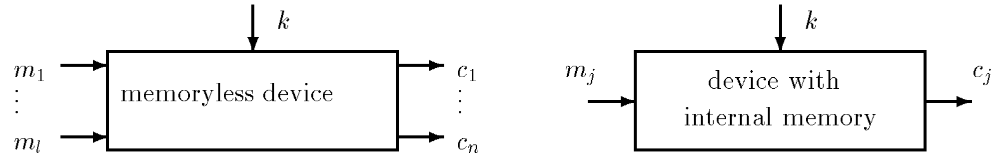
In this chapter an introduction to stream ciphering is first given. Then some cryptographic factors of sequences, such as linear complexity, weight complexity, sphere complexity, autocorrelation and crosscorrelation functions, pattern distribution, quadratic span and maximum order complexity, are introduced. The relation between some key stream generators and finite state automata is also formulated. Finally, a general discussion on the security of and attacks on stream ciphers is made. The aim of this chapter is to introduce necessary notions and notations which will be needed in later chapters.

Contributions of this chapter include a nonadditive synchronous stream ciphering scheme described in Section 2.1.3, and the introduction of complexities for sequences, i.e., weight and sphere complexity, outlined in Section 2.3.4 which is based on the papers [19, 21] and the monograph [22].

2.1 Stream Ciphering

Cryptographic systems are generally classified into block and stream ciphers, in analogy to error-correcting codes which are subdivided into block and convolutional codes. The clear distinction between block and stream ciphers is the memory, which can be depicted by Figures 2.1(a) and 2.1(b).

A block cipher breaks each plaintext message into successive blocks and encrypts each block M under the control of a key k into a ciphertext block $C = (c_1, \cdots, c_n)$, where the message text and ciphertext alphabet are usually identical. Each block is typically several characters long. Simple substitution and homophonic substitution ciphers are examples of block ciphers, even though the unit of encipherment is a single character. This is because the same key is used for each character. A stream cipher specifies a device with internal memory that encrypts the jth digit m_j of the message stream into the jth digit c_j of the ciphertext stream by means of a function which depends on both the secret key k and the internal state of the stream cipher



$C = E_k(M)$

(a) Block ciphers

$c_j = E_{z_j}(m_j)$; $z_j = f(k, c_{j-1})$

(b) Stream ciphers

Figure 2.1: The difference between block and stream ciphers

at time j. The sequence $z^\infty = z_0 z_1 \cdots$ which controls the enciphering is called the *key stream* or *running key.* The deterministic automaton which produces the key stream from the actual key k and the internal state is called the *running-key generator* or *keystream generator.* A special case is that the key consists only of the initial state. A stream cipher is *periodic* if the keystream repeats after d characters for some fixed d; otherwise it is nonperiodic.

There are two different approaches to stream encryption: synchronous methods and self-synchronous methods. In a *synchronous stream cipher*, the next state depends only on the previous state and not on the input so that the succession of states is independent of the message stream, the key stream is therefore generated independently of the message stream. Consequently, the enciphering transformation is memoryless, but time-varying. Thus, if a ciphertext character is lost during transmission, the sender and receiver must resynchronize their generators before they proceed further. Furthermore, this must be done in a way which ensures that no part of the key stream is repeated (thus the keystream generator should not be reset to an earlier state). It is therefore natural, in a synchronous stream cipher, to separate the enciphering transformation from the generation process of time-varying parameters which control the deciphering transformation.

In a *self-synchronous stream cipher*, each key character is derived from a fixed number n of preceding cipher characters. Thus, if a ciphertext character is lost or altered during transmission, the error propagates forward for n characters, but the cipher resynchronizes itself after n correct ciphertext characters have been received. Self-synchronous stream ciphers are nonperiodic because each key character is functionally dependent on the entire preceding message stream.

2.1.1 Additively Synchronous Stream Ciphering

As mentioned above, in a synchronous stream cipher, the key stream $z^\infty = z_0 z_1 \cdots$ is independent of the message stream. The algorithm that generates the stream

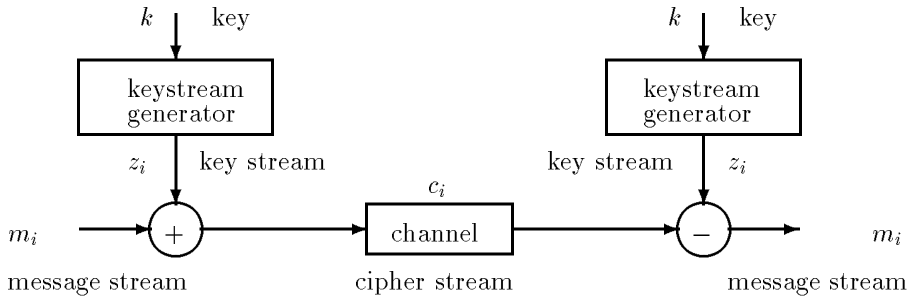


Figure 2.2: Additively synchronous stream ciphers

must be deterministic so that the stream can be reproduced for decipherment. One important kind of synchronous stream ciphers is the *additively synchronous stream ciphers* depicted by Figure 2.2, where the characters of the key stream are from an abelian group $(G, +)$ and the ciphertext character c_i is the addition of the key stream character z_i and message stream character m_i, and "$-$" denotes the inverse operation of "$+$".

2.1.2 Additively Self-Synchronous Stream Ciphering

In a self-synchronous stream cipher each keystream character is derived from a fixed number n of preceding ciphertext characters. The idea of this kind of ciphers traces back to the time of Vigenère in the 16th Century. Autokey ciphers and cipher feedback systems are examples of the additively self-synchronous stream ciphers.

An *autokey cipher* is one in which the key is derived from the message it encrypts. Another important class of self-synchronous stream ciphers consists of those where the cipher is fed back to the keystream generator as depicted in Figure 2.3. The main problems concerning this kind of stream ciphers are the design of the keystream generator and the way with which the feedback ciphertext character is used in the keystream generator. This kind of stream ciphers is fairly difficult to design and analyze because of the feedback approach.

2.1.3 Nonadditive Synchronous Stream Ciphering

There are advantages and disadvantages in both block ciphering and additive stream ciphering. Additive synchronous stream ciphers have the disadvantage that a ciphertext-plaintext character pair gives out immediately the corresponding keystream character under which the plaintext character is encrypted. This makes possible various kinds of key-recovering attacks such as correlation attacks,

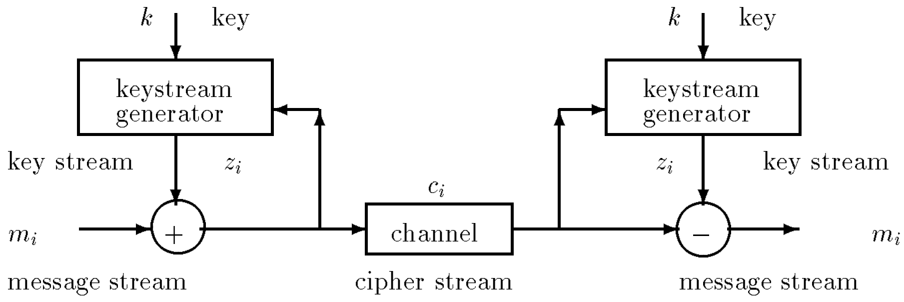


Figure 2.3: Additively self-synchronous stream ciphers

equivalent-machine attacks such as the one based on the Berlekamp-Massey algorithm, approximate-machine attacks such as those based on linear approximations. One of their advantages is that the keystream is time-varying, which ensures that the same plaintext character usually corresponds to different ciphertext characters at different times. This usually conceals some statistical properties of the plaintext. Block ciphers have the disadvantage that their keys cannot be changed very frequently due to the problem of key management. In addition, the same block of message corresponds always to the same ciphertext block if one key is selected and fixed. This may make many attacks such as differential attacks on some block ciphers applicable. One of their advantages is that the detection of the modification of messages may be possible owing to the fact that messages are encrypted block by block.

To keep the merits of both additive stream ciphering and block ciphering, but to get rid of the fault of both approaches, a dynamic block ciphering approach is described as follows, With which a keystream generator and a conventional (one-key) block cipher are combined in the way that some output characters of the keystream generator are employed to serve as the dynamic key of the block cipher for each message block.

For a block cipher of plaintext block length n, let $E_k(.)$ and $D_k(.)$ denote respectively the encryption and decryption transformation specified by a key k. To use the block cipher to encipher and decipher dynamically, a dynamic key k_i for the block cipher is produced by a sequence generator as $(z_{ti}, z_{ti+1}, \cdots, z_{ti+t-1})$, where t is a positive integer, and z^∞ denotes the sequence produced by the sequence generator. The parameter t could be 1 or another designed constant. Thus, the encryption and decryption are done respectively by

$$c_i = E_{k_i}(m_i),$$
$$m_i = D_{k_i}(c_i),$$

where m_i is the plaintext block, c_i the ciphertext block at time i. Since the key k_i is time-varying, this is a dynamic block ciphering, and therefore a nonadditive synchronous stream ciphering approach. The key of the system consists of that of the keystream generator.

In this ciphering system it is not necessary to require a large linear complexity for the output sequence of the generator if the underlying block cipher is properly designed. The sequence generator and the block cipher should be designed so that they can protect each other. This kind of ciphering approach is intended to thwart as many attacks on block and/or additive synchronous stream ciphers as possible. Indeed, if the system is well designed, it seems that the known attack approaches to additive stream ciphers and block ciphers do not apply to the system. To attack the system, one needs to develop new approaches.

Another aim of this system is to get fast ciphering algorithms. It is possible to use fast sequence generators and fast block ciphers in this system to get fast and secure ciphering algorithms.

Additive synchronous stream ciphers and all block ciphers can be regarded as special cases of the system. If the block ciphering algorithm of the above dynamic ciphering system is chosen as the termwise addition of the key and the plaintext block, then the system is a usual additive synchronous stream cipher. In this case, we use one of the worst block ciphering algorithms in the system. If the sequence generator is chosen such that the keystream is a constant sequence, then it is the usual block ciphering approach. In this case, we employ the worst keystream generator. Thus, the usual block ciphers and additive stream ciphers are special cases of the approach, and in detail, are the two extreme cases of the system.

Another promising nonadditive synchronous stream ciphering system is the cooperatively distributed ciphering introduced in [27], in which several block ciphers work in a cooperative and distributive way.

2.2 Keystream Generators and Finite Automata

Finite state machines constitute important mathematical objects for modeling electronic hardware specified above the register transfer level. Furthermore, due to their recursiveness finite state machines are convenient means for realizing infinite wordfunctions built over finite alphabets. Many keystreams can be modeled by finite state machines [17]. In a synchronous stream cipher, the running-key generator may be generally viewed as an autonomous finite state machine as depicted in Figure 2.4.

The keystream generator as a finite state machine consists of an output alphabet and a state set, together with two functions and an initial state. The *next state function* f_s maps the current state S_j into a new state S_{j+1}, and the output function

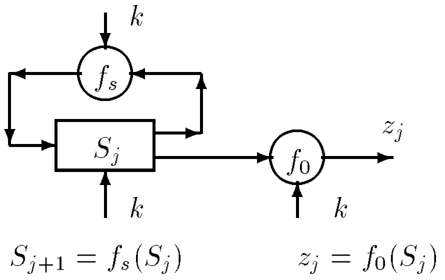


Figure 2.4: Keystream generators as autonomous finite state machines

f_0 maps the current state S_j into an output symbol z_j from the output alphabet. The key may determine the next state function and the output function as well as the initial state. It is noted that for the autonomous finite state machines the next state function is only a function of the previous state.

The fundamental problem of designing a keystream generator is to find the next state function f_s and the output functions f_0 which are guaranteed to produce a running key z^∞ that satisfies certain cryptographic requirements such as large linear complexity, good linear complexity stability, large period, good autocorrelation property, uniform pattern distribution property, etc. In some cases, the output function f_0 should possess good difference property with respect to some binary operation of the state vector space and good nonlinearity with respect to the binary operation of the state vector space and that of the output alphabet space. These binary operations depend on the realization of the next state function f_s. The actual specific requirements for the next state and output function depend on the system in which the generator is used.

In order to meet certain requirements, special classes of finite state machines have been employed as running-key generators. Unfortunately, the theory of autonomous automata whose change of state function is nonlinear has not been well developed. There are many kinds of proposed keystream generators. Some are easy to implement, but their security may be difficult to control; some are secure against certain kinds of attacks, but may have a relatively slow implementation.

2.3 Cryptographic Parameters of Sequences

Sequences for different stream ciphering purposes can be very different. It is often the case that sequences used in one stream cipher are required to have some properties which are different from those required for some other sequences employed in another stream cipher. Thus, cryptographic sequences may be different in some aspects. However, for keystream sequences for additive synchronous stream ciphers there are some common cryptographic measures on their strength such as the linear complexity (linear span or linear equivalence), weight complexity, sphere complexity, pattern distribution, autocorrelation function. This section introduces some of these measures and illustrates their cryptographic importance. Here only sequences over finite fields are discussed.

2.3.1 Linear Complexity

To introduce the linear complexity, we need the shift operator on sequences. A left shift operator E is defined by $Es_i = s_{i-1}$ for all possible i. In this way we can define recursively the operators E^l for $l > 1$. Thus, for a polynomial $f(x)$ of $GF(q)[x]$ the polynomial operator $f(E)$ is well defined, if we write $E^0 = 1$, the identity operator. If a sequence is over a finite field $GF(q)$ and $f(x)$ is a polynomial with

$$f(x) = c_0 + c_1 x + \cdots + c_{L-1} x^{L-1},$$

then we define

$$f(E)s_j = c_0 s_j + c_1 s_{j-1} + \cdots + c_{L-1} s_{j-L+1}.$$

Let s^n denote a sequence $s_0 s_1 \cdots s_{n-1}$ of length n over a finite field $GF(q)$. For a finite sequence, the n is finite; for a semi-infinite sequence the n is ∞. A polynomial $f(x) \in GF(q)[x]$ of degree $\leq l$ with $c_0 \neq 0$ is called a *zero polynomial* of the sequence s^n if

$$f(E)s_j = 0, \text{ for all } j \text{ with } j \geq l. \tag{2.1}$$

If the above equations hold for l, then they hold also for $l+1$. Thus, for every zero polynomial there is a least $l \geq \deg(f)$ such that the above equations hold. We call the smallest l the *associated recurrence length* of $f(x)$ with respect to the sequence. It is easy to see that there are zero polynomials of a sequence such that their associated recurrence length is minimal. Such a zero polynomial is called a *minimal polynomial* of the sequence, and the associated recurrence length is called the *linear span* or *linear complexity* of the sequence, which is denoted as $\mathrm{L}(s^n)$ here and hereafter. It follows immediately from this definition that $\mathrm{L}(s^n) = 0$ iff $s^n = 0^n$, where 0^n denotes the all-zero sequence of length n. Another immediate consequence is that, for $0 < n < \infty$, $\mathrm{L}(s^n) = n$ iff $s^n = 0^{n-1}$ and $s_{n-1} \neq 0$. If a semi-infinite s^∞ is periodic, then its minimal polynomial is unique if we require that $c_0 = 1$. The linear complexity of periodic sequences is equal to the degree of their minimal polynomials.

The engineering interpretation of linear complexity is as the length of the shortest linear feedback shift register that generates the sequence, see Figure 2.5. Such

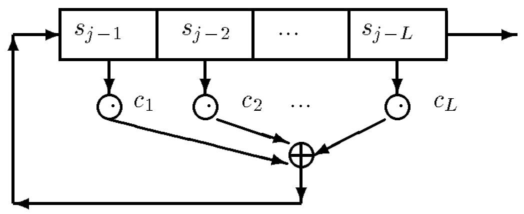


Figure 2.5: The linear feedback shift register interpretation of recursion (2.1)

a linear feedback shift register is said to be non-singular when $c_1 \neq 0$, i.e., when it corresponds to a linear recursion of order $\mathrm{L}(s^n)$. The minimal polynomial is actually a feedback polynomial of the linear feedback shift register.

The cryptographic significance of the linear complexity of keystream sequences is well known due to the Berlekamp-Massey algorithm [48]. If the linear complexity of a key stream is L, then $2L$ consecutive characters of the sequence could be used to construct the whole key stream with the Berlekamp-Massey algorithm. Thus, it is cryptographically necessary but not sufficient to require keystream sequences for additive stream ciphers to have large linear complexity. This will be illustrated when we introduce the sphere complexity. It should be mentioned, however, that for some nonadditive stream ciphers a large linear complexity of the keystream is not a necessary cryptographic requirement (see Section 2.1.3).

Recall that in the definition of linear complexity above, the coefficients of the feedback polynomial and the entries of the sequence are required to be in the same field. Let $s^n = s_0 s_1 \cdots s_{n-1}$ be a sequence over the field $GF(q^m)$. The linear complexity of s^n with respect to the subfield $GF(q)$, here and hereafter denoted as $\mathrm{L}_{GF(q)}(s^n)$, is defined as the smallest nonnegative integer L such that there exist $c_1, c_2, \cdots, c_L \in GF(q)$ for which

$$s_j + c_1 s_{j-1} + \cdots + c_L s_{j-L} = 0, \text{ for all } L \leq j < n. \tag{2.2}$$

In this definition [22], the coefficients $c_1, c_2, \cdots, c_L$ are required to be in $GF(q)$. If $m = 1$, then the two complexities are identical. Thus, the linear complexity $\mathrm{L}_{GF(q)}$ is a generalization of the usual linear complexity. The following inequality apparently holds:

$$\mathrm{L}(s^n) \leq \mathrm{L}_{GF(q)}(s^n). \tag{2.3}$$

The cryptographic importance of this generalized linear complexity may be seen as follows. It is well known that $GF(q^m)$ can be regarded as a linear space of dimension m over $GF(q)$. Let $u_1, u_2, \cdots, u_m$ be a basis of $GF(q^m)$ over $GF(q)$, then each u of $GF(q^m)$ can be expressed as

$$u = \sum_{i=1}^{m} a_i u_i, \ a_1, \cdots, a_m \in GF(q).$$

Assume that for every j we have

$$s_j = \sum_{i=1}^{m} s_{i,j} u_i, \quad s_{i,j} \in GF(q),$$

then recursion (2.2) is equivalent to

$$\begin{array}{l} s_{i,j} + c_1 s_{i,j-1} + \cdots + c_L s_{i,j-L} = 0, \\ \text{for all } L \le j \le n, \text{ and for all } i = 1, \cdots, m. \end{array} \tag{2.4}$$

This means that the linear complexity $\mathrm{L}_{GF(q)}(s^n)$ is the shortest length of the linear feedback shift registers which can generate the lower field sequences $s_{i,0}s_{i,1}\cdots s_{i,n-1}$, $i = 1, 2, \cdots, m$, at the same time only with different initial states. The determination of the shortest linear feedback shift register which generates the m sequences is called the linear feedback shift register synthesis of multisequences, which is important in decoding cyclic codes. Several algorithms for the linear feedback shift register synthesis of multisequences have been developed [8, 33, 20, 22]. If we use multisequences to encipher a message stream in parallel or to use a matrix sequence to encipher, then this kind of generalized linear complexity is cryptographically important.

Now we turn to the usual linear complexity of periodic sequences. As already mentioned, the linear complexity of periodic sequences over finite fields is precisely the degree of their minimal polynomials. Cryptographically, we need to know not only the linear complexity of a sequence, but the minimal polynomial also. To introduce some results about the minimal polynomials of periodic sequences, we need the concept of generating functions.

The *formal power series* or *generating function* of a semi-infinite sequence s^∞ over $GF(q)$ is defined by

$$s(x) = \sum_{i=0}^{\infty} s_i x^i.$$

If s^∞ is periodic with period N, then we have

$$(1 - x^N)s(x) = s^N(x) = \sum_{i=0}^{N-1} s_i x^i.$$

It follows that the following proposition holds.

Proposition 2.1 *The generating function of each periodic sequence s^∞ can be expressed as*

$$s(x) = g(x)/f(x) \tag{2.5}$$

where $g(x)$ and $f(x)$ are polynomials with $f(0) \neq 0$ and $deg(g) < deg(f)$.

There "ultimately periodic" means that for a sequence $\{s_i\}$ there are integers i_0 and N such that $s_{i+N} = s_i$ for all $i \geq i_0$. The expression in (2.5) is called a *rational form* of the generating function $s(x)$ and of the sequence s^∞. If $\gcd(g(x), f(x)) = 1$, then it is called a *reduced rational form*. Let us stipulate that f_s will denote the minimal polynomial of a sequence s^∞. The following two classic propositions are very useful, whose proofs can be found, for example, in [44, 64, 22].

Proposition 2.2 *Let s^∞ be a periodic sequence over $GF(q)$ and*

$$s(x) = r(x)/f(x), \; f(0) = 1$$

a rational form of the generating function of s^∞. Then $f(x)$ is the minimal polynomial of the sequence iff $\gcd(r(x), f(x)) = 1$.

Concerning the minimal polynomial of the sum of two periodic sequences we have the following conclusion, which follows easily from Proposition 2.2.

Proposition 2.3 *Let the reduced rational forms of two periodic sequences s^∞ and t^∞ be respectively*

$$s(x) = r_s(x)/f_s(x), \; t(x) = r_t(x)/f_t(x).$$

Then the minimal polynomial of the sum sequence of the two sequences is given by

$$f_{s+t} = f_s f_t / \gcd(f_s f_t, r_s f_t + r_t f_s).$$

2.3.2 Pattern Distribution

Let s^∞ be a sequence of period N over $GF(q)$, where N is not necessary the least period. The vector $(s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}})$ is called a *pattern of length k* with distances $(\tau_1, \tau_2 - \tau_1, \cdots, \tau_{k-1} - \tau_{k-2})$. A pattern $(s_t, s_{t+\tau})$ was also called a τ-*bigram* by Selmer [64]. τ-bigrams were first introduced by Zierler under another name for the purpose of studying the autocorrelation of maximum-length sequences [71].

Consider the following sequence s^∞ of period 7:

$$s^\infty = \underbrace{0111001}\underbrace{0111001}\ldots$$

and the pattern $0 * 1 * *0$, where $*$ denotes an arbitrary bit, possibly different ones. It is easily seen that this pattern appears only once the periodic circle of the sequence.

The notion of a multiplier was first introduced by Carmichael [7]. If s^∞ is a sequence of least period N, an element $M \neq 1$ in $GF(q), q > 2$, such that

$$M \cdot s^\infty = s_\tau^\infty, \ 0 < \tau < N, \tag{2.6}$$

is called a *multiplier* of the sequence, where s_τ^∞ is the τ-shift version of s^∞. The multiplier is related to the multiplier of residue difference sets [3]. The τ here was called the *span* of M by Ward [68]. For maximum-length sequences over $GF(q)$, Zierler showed that the τ-bigrams are evenly distributed if τ is not the span of some multiplier $M \neq 1$. It is important to observe that no such multiplier exist for binary sequences.

Bigram, trigram and τ-gram are terms used in linguistic studies. We will use the term pattern instead. The distribution of some special patterns was investigated by Golomb [35], i.e., the runs of 0's and 1's, which were called *gap* and *block* respectively. He proved that in a binary maximum-length sequence of period $2^n - 1$, there are 2^n runs. Half the runs has length 1, one fourth has length 2, one-eighth has length 3, etc., until two runs of length $n-2$; for each of these lengths, there are equally many gaps and blocks. Finally, there is one gap of length $n-1$ and one block of length n. These are Golomb's three randomness postulates.

Why is the pattern distribution property of a key stream cryptographically important? Some intuitive facts can only give us some explanations superficially. For example, the pattern distribution of length 1 is in fact the distribution of the elements of the field over the sequence. Thus, if there are much more 1's than 0's in a binary sequence, then the sequence is not cryptographically good. To see the cryptographic importance of a roughly equally likely distribution for certain patterns, we first prove a conservation law of patterns.

Consider now the patterns of length k with distances $(\tau_1, \tau_2 - \tau_1, \cdots, \tau_{k-1} - \tau_{k-2})$. For a sequence with least period N over $GF(q)$, the vector variable $(s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}})$ takes on vectors of $GF(q)^k$ when t runs from 0 to $N-1$. Let $n((s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}}) = a)$ denote the number of times with which the vector variable $(s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}})$ takes on $a \in GF(q)^k$ when t runs from 0 to $N-1$. It is straightforward to see that the following theorem holds.

Theorem 2.4 (The conservation law of patterns) *Let the symbols be the same as before, then*

$$\sum_{a \in GF(q)^k} n((s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}}) = a) = N.$$

Clearly, this theorem means that these $n((s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}}) = a)$ are conservative. The constant $n((s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}}) = a)/N$, denoted as $p(a)$, is the probability that $(s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}})$ takes on a. It follows that

$$\sum_{a \in GF(q)^k} p(a) = 1. \tag{2.7}$$

In general, *bad patterns* refer to those which appear with small probability in the key streams. If in a sequence of period N over $GF(q)$ the almost equally likely distribution of a pattern of length k with distances $(\tau_1, \tau_2 - \tau_1, \cdots, \tau_{k-1} - \tau_{k-2})$ is required, this means that $n((s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}}) = a) = N/q^k$ or it is almost a constant. Now the problem is why such a uniform pattern distribution should be cryptographically required. This may be shown by the differential attack on the natural sequence generator of Figure 3.1(b) outlined in Section 3.2. The idea behind the attack is that bad patterns give much more information about the key than other patterns.

2.3.3 Correlation Functions

Many problems in ranging systems, radar systems, spread-spectrum communication systems, multiple-terminal system identification, and code-division multiple-access communications systems require sets of signals which have one or both of the following properties:

- Each signal in the set is easy to distinguish from a time-shifted version of itself;
- each signal in the set is easy to distinguish from (a possibly time-shift version of) every other signal in the set.

This leads to an intensive study of the periodic autocorrelation and crosscorrelation functions of sequences.

To introduce correlation functions, we need group characters. Let G be a finite abelian group (written multiplicatively) of order $|G|$ with identity element 1_G. A *character* χ of G is a homomorphism from G into the multiplicative group U of complex numbers of absolute value 1.

Proposition 2.5 *Let χ be a character of a finite abelian group G as above, and $g \in G$. Then*
(a) $\chi(1_G) = 1$.
(b) $\chi(g)$ is a $|G|$th root of unity.
(c) $\chi(g^{-1}) = \chi(g)^{-1} = \overline{\chi(g)}$,
where the bar denotes the conjugate.

For every G we have the *trivial* character χ_0 defined by $\chi_0 = 1$ for all $g \in G$. For each character χ the character $\overline{\chi}$ defined by $\overline{\chi} = \overline{\chi(g)}$, is called the *conjugate character* of χ. Given finitely many characters $\chi_1, \cdots, \chi_n$ of G, the product character $\chi_1 \cdots \chi_n$ is defined by

$$(\chi_1 \cdots \chi_n)(g) = \chi_1(g) \cdots \chi_n(g), \text{ for all } g \in G.$$

Thus, the meaning of χ^n is clear. Under this multiplication of characters, the set G^C of characters of G forms an abelian group. It follows immediately from part (b) of the above proposition that G^C is abelian.

Let G be a finite cyclic group of order n, and let g be a generator of G. For a fixed integer j, $0 \le j \le n-1$, the function

$$\chi_j(g^k) = e^{2\pi ijk/n}, \quad k = 0, 1, \cdots, n-1,$$

defines a character of G, and it is easy to see that G^C consists exactly of the characters $\chi_0, \chi_1, \cdots, \chi_{n-1}$.

Let p be the characteristic of $GF(q)$, and $q = p^m$. We denote the prime field of $GF(q)$ with Z_p. The *absolute trace function* $\mathrm{Tr}(a)$ from $GF(q)$ to $GF(p)$ is defined by

$$\mathrm{Tr}(a) = a + a^p + a^{p^2} + \cdots + a^{p^{m-1}}, \ a \in GF(q). \tag{2.8}$$

With the absolute trace function we can now define the function χ_1 by

$$\chi_1(c) = e^{2\pi i \ \mathrm{Tr}(c)/p} \ \text{ for all } c \in GF(q), \tag{2.9}$$

which is a character of the additive group $(GF(q), +)$. Following [44] and [61], we call the characters of the group $(GF(q), +)$ *additive characters*, the above character χ_1 the *canonical additive character* of $GF(q)$. This character is important due to the following proposition:

Proposition 2.6 *For $b \in GF(q)$, the function χ_b with $\chi_b = \chi_1(bc)$ for all $c \in GF(q)$ is an additive character of $GF(q)$, and every additive character of $GF(q)$ is obtained in this way.*

In particular, for the finite field $GF(p)$, where p is prime, we know that

$$\chi_j(a) = e^{2\pi ija/p}, \ j = 0, 1, \cdots, p-1 \tag{2.10}$$

are all the additive characters of $GF(p)$, since $(GF(p), +)$ is cyclic.

Characters of the multiplicative group $GF(q)^*$ are called *multiplicative characters* of $GF(q)$. Since $GF(q)^*$ is a cyclic group of order $q-1$, its characters can be easily determined.

Proposition 2.7 *Let g be a fixed primitive element of $GF(q)$. For each $j = 0, 1, \cdots, q-2$, the function ψ_j with*

$$\psi_j(g^k) = e^{2\pi ijk/(q-1)} \quad k = 0, 1, \cdots, q-2$$

defines a multiplicative character of $GF(q)$, and every multiplicative character of $GF(q)$ is obtained in this way.

Let $GF(q)$ be a finite field, χ be an additive character of $GF(q)$, and s^∞ and t^∞ be two sequences of period respectively N and M and $P = \mathrm{lcm}\{M, N\}$. Then the *periodic crosscorrelation function* of the two sequences is defined by

$$\mathrm{CC}_{s,t}(l) = \sum_{i=0}^{P-1} \chi(s_i - t_{i+l}) = \sum_{i=0}^{P-1} \chi(s_i)\overline{\chi(t_{i+l})}. \tag{2.11}$$

If the two sequences are identical, then $P = M = N$ and the crosscorrelation function is the so-called *periodic autocorrelation function* of s^∞ described by

$$\mathrm{AC}_s(l) = \sum_{i=0}^{N-1} \chi(s_i - s_{i+l}) = \sum_{i=0}^{N-1} \chi(s_i)\overline{\chi(s_{i+l})}. \tag{2.12}$$

If $q = 2$, then $\chi(a) = (-1)^a$ is an additive character of $GF(2)$, here we identify $GF(2)$ with Z_2. Then (2.11) and (2.12) are the usual crosscorrelation and autocorrelation functions of binary sequences.

For two sequences s^∞ and t^∞ of period respectively N and M and $P = \mathrm{lcm}\{M, N\}$, the *aperiodic crosscorrelation function* of the two sequences is defined by

$$\mathrm{ACC}_{s,t}(l, u, v) = \sum_{i=u}^{v} \chi(s_i - t_{i+l}) = \sum_{i=u}^{v} \chi(s_i)\overline{\chi(t_{i+l})}, \tag{2.13}$$

where $\overline{x}$ denotes the complex conjugate of x. If the two sequences are identical, then $P = M = N$ and the crosscorrelation function is the so-called *aperiodic autocorrelation function* of s^∞ described by

$$\mathrm{AAC}_s(l, u, v) = \sum_{i=u}^{v} \chi(s_i - s_{i+l}) = \sum_{i=u}^{v} \chi(s_i)\overline{\chi(s_{i+l})}. \tag{2.14}$$

Here our definitions of aperiodic functions are slightly different from those used for communication purposes, which are defined to be $\mathrm{ACC}_{s,t}(l, 0, v)$ and $\mathrm{AAC}_s(l, 0, v)$. Aperiodic autocorrelation property may be much more important cryptographically than the periodic autocorrelation property, since the former reflects the local randomness property and the latter the global randomness property and what we actually use for stream ciphering is only a small part of periodic sequences. Generally speaking, the periodic autocorrelation property is relatively much easier to control than the aperiodic autocorrelation property.

The cryptographic importance of pattern distribution has already been shown in the above section. We now see the connection between pattern distribution and autocorrelation property. For the sake of simplicity, we take binary sequences as an example. Recall that for binary sequence of period N the periodic autocorrelation function is defined by

$$\begin{aligned}\mathrm{AC}_s(l) &= \sum_{i=0}^{N-1}(-1)^{s_i+s_{i+l}} \\ &= [n((s_i, s_{i+l}) = (0,0)) + n((s_i, s_{i+l}) = (1,1)) \\ &\quad -n((s_i, s_{i+l}) = (0,1)) - n((s_i, s_{i+l}) = (1,0))].\end{aligned}$$

This shows clearly the connection between the distribution property of patterns of length 2 and the autocorrelation property. Thus, if a binary sequence has a uniform distribution of all patterns of length 2, then all of its out-of-phase autocorrelation values are approximately 0.

2.3.4 Weight and Sphere Complexity

Let x be a finite sequence of length n over $GF(q)$, the *weight complexity* [19, 22] is defined by

$$\mathrm{WC}_u(x) = \min_{\mathrm{WH}(y)=u} \mathrm{L}(x+y), \tag{2.15}$$

where $\mathrm{WH}(y)$ denotes the Hamming weight of y, i.e., the number of components of y that are different from zero.

Consider now the space $GF(q)^n$ with Hamming distance d_H. Denoting $S(x,u) = \{y : \ d_H(x,y) = u\}$, by definition we have

$$\mathrm{WC}_u(x) = \min_{y \in S(x,u)} \mathrm{L}(y).$$

This means that the weight complexity is the maximum lower bound of linear complexities of all the sequences of length n on the sphere surface $S(x,u)$.

Let $O(x,y) = \{y : \ 0 < d_H(x,y) \leq u\}$ be the sphere without center x. The *sphere complexity* is defined in [19, 22] by

$$\mathrm{SC}_u(x) = \min_{y \in O(x,u)} \mathrm{L}(y) = \min_{0<v\leq u} \mathrm{WC}_v(x). \tag{2.16}$$

Similarly, let s^∞ be a sequence of period N (not necessarily the least period) over $GF(q)$. The weight and sphere complexity of periodic sequences with respect to N are defined respectively by

$$\mathrm{WC}_u(s^\infty) = \min_{\mathrm{WH}(t^N)=u, per(t^\infty)=N} \mathrm{L}(s^\infty + t^\infty) \tag{2.17}$$

$$\mathrm{SC}_u(s^\infty) = \min_{0<v\leq u} \mathrm{WC}_v(s^\infty), \tag{2.18}$$

where $per(t^\infty) = N$ denotes that t^∞ has a period N.

These two complexities were introduced to measure the stability of the linear complexity function. The cryptographic background of these complexities is that some key streams with large linear complexity can be approximated with some sequences with much lower linear complexity [19, 22]. The sphere and weight complexities are based on the linear-feedback shift-register approximation model. In contrast to the linear complexity which is based on the shortest linear feedback shift register that produces a sequence, the sphere complexity $\mathrm{SC}_k(s^\infty)$ is based on the shortest linear feedback shift register that produces another sequence with a probability of agreement no less than $(1 - k/N)$, where N is a period of the sequence s^∞ with which the sphere complexity is concerned. The weight complexity $\mathrm{WC}_k(s^\infty)$ is based on the shortest linear feedback shift register that produces another sequence with a probability of agreement being equal to $(1 - k/N)$.

To illustrate the difference between the linear complexity and sphere complexity, we consider the binary sequence of period N:

$$s^\infty = \underbrace{0....01}_{N}{}^\infty.$$

The linear complexity of the sequence is N by definition since no linear feedback shift register of length less than N can produce it. However its sphere complexity $SC_1(s^\infty) = 1$ by definition since there is a linear feedback shift register of length one that produces the all-zero sequence having the probability $(1 - 1/N)$ of agreement with the sequence s^∞.

These linear-feedback shift-register approximation-model complexities are cryptographically important only if there is an efficient algorithm to find the linear feedback shift register for approximating the original generator. To see the cryptographic importance of these complexities, we develop an attack on all synchronous additive stream ciphers.

Suppose that a cryptanalyst has got a number of consecutive ciphertext-plaintext pairs of a synchronous additive stream cipher which enable him/her to derive a piece of key stream, say $z_0 z_1 \cdots z_n$. Suppose also that he/she knows nothing else but the source codes of enemy's messages. What can he/she do under these assumptions if he/she wants to decipher the enemy's ciphertext? The best the cryptanalyst can do may be the construction of a new generator which produces a sequence with a large probability of agreement with the original key stream.

To make things simple, we assume that the key stream is binary. Under the assumption that the linear complexity of the enemy's key stream is very unstable, the cryptanalyst can try to construct a linear feedback shift register to approximate the original keystream generator according to the following procedure:

Step 1 Use the Berlekamp-Massey algorithm to construct a linear feedback shift register which produces the sequence $z^n = z_0 z_1 \cdots z_{n-1}$.

Step 2 Use the linear feedback shift register to decipher the known ciphertext. If only ϵ% (this constant can be flexible, say less than 15) of the deciphered ciphertext makes no sense, then accept the linear feedback shift register and stop; otherwise, go to Step 3.

Step 3 For $i = 0$ to $n-1$, change z_i into $z_i \oplus 1$, then go to Step 1. If an acceptable linear feedback shift register is found, then stop; otherwise, go to Step 4.

Step 4 For all $i = 0$ to $n-1$ and $j = 0$ to $n-1$ with $i < j$, change z_i into $z_i \oplus 1$ and z_j into $z_j \oplus 1$. Then repeat Step 1. If an acceptable linear feedback shift register is found, then write the linear feedback shift register and stop; otherwise, print "fail" and stop.

Since the complexity of the Berlekamp-Massey algorithm for sequences of length n is $O(n^2)$, the complexity of this attack is $O(n^4)$. Thus, if s^∞ is a key stream such that its linear complexity is very large (say, for example, 2^{40}) and $\mathrm{SC}_k(s^\infty)$ is small enough (say less than 1000 for example) for some k with k/N being very small, then this attack must succeed. The basic idea of this attack is that we expect that the keystream sequence can be expressed as

$$z^\infty = u^\infty + v^\infty$$

such that u^∞ and v^∞ are of period N, the $\mathrm{WH}(v^N)/N$ is very small, and the sequence u^∞ has small linear complexity. This can be done when the linear complexity of the key stream is very unstable. In this case, we expect that the known key stream z^n can be expressed as

$$z^n = u^n + v^n$$

with $\mathrm{WH}(v^n) \leq 2$ if $n < 2N/k$. Furthermore, we may also use the regular decimation sequences of z^n to replace z^n, then derive the minimal polynomial of u^n from the decimated sequences. Thus, it follows that the designer of a synchronous additive stream cipher must ensure that for very small k's, the sphere complexity $\mathrm{SC}_k(s^\infty)$ is large enough. In other words, the designer of an additive synchronous stream cipher should make sure that his/her key stream cannot be well approximated by a sequence with small linear complexity, since the above polynomial-time algorithm can be used to find a linear feedback shift register to approximate the original keystream sequence if its linear complexity is very unstable. This is why sphere complexity is cryptographically important. For the purpose of measuring the linear complexity stability of sequences, fixed-complexity distance and variable-complexity distance were introduced. Connection between these measures and some lower bounds on these measures for some sequences can be found in [22].

2.3.5 Higher Order Complexities

Linear complexity (also called *linear span*) of a sequence is defined to be the length of the shortest linear feedback shift register that generates the sequence. A sequence with very large linear span may be generated by a much shorter feedback shift register if nonlinear terms are allowed in the feedback function. If quadratic terms and linear terms are only allowed, then we have the *quadratic span*. If general terms are allowed, then we have the *maximum order complexity* or generally *span*.

The quadratic span and other nonlinear spans are cryptographically important only when there are efficient algorithms for finding the shortest nonlinear feedback shift registers or a feedback shift register that is short enough. Suppose there is an efficient algorithm for finding the shortest quadratic feedback shift register that generates any given sequence, then we need further to investigate the relations between the linear span and quadratic span. It is obvious that the linear span of any sequence is greater than or equal to the quadratic span of the sequence. However, if the relation

$$\mathrm{L}(s^\infty) \geq \mathrm{Q}(s^\infty) \geq \sqrt{\mathrm{L}(s^\infty) + a},$$

holds for every periodic binary sequence, where $\mathrm{Q}(s^\infty)$ denotes the quadratic span of the sequence s, and a is a constant, then the control of the linear span results also in that of the quadratic span.

2.4 Security and Attacks

Without attacks to cipher systems there would be no problem of security. Security is associated with attacks and is usually relative to attacks. Attacks are also relative to the assumptions about a cryptanalyst's knowledge of a cipher system.

Attacks can be classified according to the assumed available information about a cipher that the cryptanalyst has. This kind of classification results in three types of attacks: (A) *ciphertext-only attacks* under the assumption that only pieces of ciphertext are known to a cryptanalyst; (B) *known-plaintext attacks* under the assumption that a piece of ciphertext with corresponding plaintext is known; (C) *chosen-plaintext attacks* under the assumption that a cryptanalyst has chosen a piece of plaintext and obtained the corresponding ciphertext.

Suppose that a cryptanalyst has got a piece of keystream sequence, there are two further assumptions concerning attacks on stream ciphers:

B1: It is assumed that the cryptanalyst only knows a piece of keystream. In this case there are the following possible attacks: (1) *equivalent-machine attacks*, which make use of the piece of key stream to construct a new generator which produces the same key stream. For example, if the linear complexity of the key stream is not very large, the Berlekamp-Massey algorithm can be used to construct a linear feedback shift register which produces the same key stream. (2) *Approximate-machine attacks*, which make use of the known piece of key stream to construct another generator to approximate the original generator. One example of these attacks is the attack based on the linear complexity stability of the key stream described in Section 2.3.4.

B2: Apart from a piece of keystream it is assumed that the cryptanalyst also knows the type of generator and the cryptographic algorithm. Under these assumptions attacks are flexible in forms and in techniques, for example, those in [22, 23]. Attacks under these assumptions depend on specific systems and on the technique a cryptanalyst uses. In this case, there is one more type of attacks than in the case (B1), i.e., *key recovering attacks*, which aim to recover the original key or equivalent keys.

For ciphertext-only attacks, some other information about a cipher is usually assumed to be known to a cryptanalyst. For example, under the assumption that the structure of the keystream generator is known there are several kinds of correlation attacks [65]. These are also key-recovering attacks.

The general idea of the key-recovering attacks is to make use of known materials about a cipher to get information about the key. But techniques of attacks vary with the structure of ciphers and with the known materials as well as with cryptanalysts. The following "mother problem" may illustrate the flexibility of attacks.

Suppose that there are ten children in such an order $z_1 z_2 \cdots z_{10}$ that z_i is older than or has the same birthday as z_{i+1}. We are told that these children have the same father F who is known and also the same mother who could be any one of the 100 mothers in the set $\{M_1, \cdots, M_{100}\}$. If we have further information about the pairs (F, M_i), then we can make use of it to get the set of mothers smaller or to determine the mother. For instance, by studying the known information about all (F, M_i), suppose that we know that having twins is only possible for (F, M_1) to have twins. Then if twins are found in the 10 children, the mother must be M_1. Of course, it may be technically difficult to find this character of (F, M_1). In fact we can visualize the keys as mothers, the algorithm as the father, the ciphertext as the children.

To determine the mother, one cryptanalyst may try to study the distribution of sexes among the children because he/she has some technique to study the sex properties of each (F, M_i). Another one may try to study the color of the hair of the mothers, father and children. Others may study the blood types of father, mothers and children and age distribution of the mothers and children. Every method must use the idea of "consistency" or "correlation" either in a direct way or in a hidden way , i.e., whether the analyzed property p of the set $\{(F, M_i) : 1 \leq i \leq 100\}$ is consistent with the analyzed property p of the children or what is the extent of correlation between the properties of the two sets. Thus, perhaps each cryptanalyst has his/her own technique for getting information about the mother[1].

Keystream generators are flexible and diverse. Each generator may have its own special security aspects, although some common requirements exist (e.g., linear complexity and linear complexity stability). Though it may be generally said that cryptographic gains and losses usually go together, there are differences between cipher systems. Some are easy to implement, but may have tradeoffs between known security parameters; some are relatively difficult to implement, but their security may be easy to control; others may have both an easy implementation and ideal security, but be slow. Of course, fewer tradeoffs make the design easier. In designing secure cipher systems the most important problems are: 1.) How can we build systems which have as few security tradeoffs as possible? 2.) What are the tradeoffs or conflicts in a given system? 3.) How do we manage tradeoffs and conflicts? 4.) How do we coordinate security and performance? It is possible to build some stream ciphers such that many of their security aspects are in harmony. The binary additive synchronous stream cipher based on some natural sequence generators may show the possibility.

[1] It could be good if we exchange the roles of mother and father in this example, as suggested by Valtteri Niemi.

Chapter 3

Counter Generators and Their Analysis

This chapter introduces some counter generators, and is mainly concerned with a special counter generator—the *natural sequence generator*. The differential cryptanalysis of the binary natural sequence generator is presented in detail. The equivalence between seven analyses for the binary natural sequence generator is shown. Another topic of this chapter is the natural sequence generator realization of periodic sequences. Finally, a basic theorem about the linear and sphere complexities is proved. It serves as one of the theoretical bases of the keystream generators concerned in this thesis.

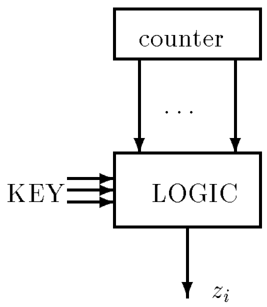
The main contributions of this chapter are the differential cryptanalysis of the natural sequence generator described in Section 3.2 including several conservation laws outlined by Theorems 3.1 and 3.2, the proof of the equivalence between the seven analyses of the natural sequence generator described in Section 3.3, and a basic theorem and bridge between number theory and stream ciphers as described in Theorem 3.5. It is worth noting that the natural sequence generator seems to be the only kind of generators known whose seven analyses are proven to be equivalent. This chapter is mainly based on the papers [23, 24].

3.1 Generators Based on Counters

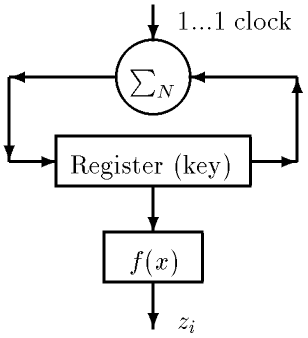
A counter, one of the simplest automata, has maximal period, which is usually q^n, where q is a positive integer. A counter generator of period N counts the integers from 0 to $N-1$ cyclically, and has a register whose content $c(t)$ is updated with $c(t)+1 \bmod N$, i.e.,

$$c(t+1) = c(t) + 1 \bmod N,$$

where $c(i)$ is the content of the register of the counter. Diffie and Hellman suggested to apply a nonlinear logic to a counter to construct keystream generators [17], which is depicted in Figure 3.1(a). In this kind of generators, the key is used to control the



(a) The Diffie-Hellman generator



(b) The natural sequence generator

Figure 3.1: Some counter generators

logic. The initial values of the counter may be taken as part of the key or a random value sent as an indicator. A specific proposal given by Diffie and Hellman is to use a fixed component of a block cipher algorithm as the logic for the generators of Figure 3.1(a) [17].

If we consider counters of arbitrary period and apply a fixed logic to a counter, we have a generator of Figure 3.1(b), where $\sum_N$ denotes the integer modulo N addition, $f(x)$ is a mapping from the residue ring Z_N to an abelian group G, and the output sequence can be expressed as

$$z_i = f((i + k) \bmod N),$$

There are slight differences between the two generators. In the generator of Figure 3.1(a), the key or part of the key is used to control the logic, while in the generator of Figure 3.1(b) the logic is fixed, the key consists only of the initial values of the register. The generator of Figure 3.1(b) was called the *natural sequence generator* because every periodic sequence can be realized by this generator in a natural way and many security aspects of the generator are consistent. Synchronous additive stream ciphers based on this kind of generators are called *additively natural stream ciphers* [23].

In [23] the differential cryptanalysis and design of the additively natural stream ciphers were developed and a theoretical basis for the analysis is established. It was shown that an improperly designed natural keystream generator could be broken by a differential attack. Other possible attacks, such as key determining attacks based on decision trees, partial-key attacks, linear approximation attacks with respect to the additions of Z_N and the abelian group over which the key stream is built and key (key stream) correlation attacks, were also studied in [23]. This thesis is mainly concerned with the design and analysis of the natural sequence generators.

3.2 Differential Cryptanalysis of Binary Natural Generators

The differential cryptanalysis for block ciphers was introduced by Biham and Shamir [4]. It is possible to develop similar attacks on some stream ciphers. This can be illustrated by the differential cryptanalysis of the additive synchronous stream ciphers based on a natural sequence generator. Because of the additive structure its differential cryptanalysis becomes that of its keystream generator [23].

For the binary natural sequence generator of Figure 3.1.b, assume that $f(x)$, N and a piece of key stream $z^t = z_0 z_1 \cdots z_{t-1}$ are known to a cryptanalyst, who wishes to recover the time-varying key k at the time when z_0 was produced.

3.2.1 Differential Cryptanalysis

Let $C_i = \{x : x \in Z_N,\ f(x) = i\}$ for $i = 0, 1$. Then the procedure of this attack can be described as follows:

Step 1: Find parameters $(i, j; w)$'s with $(i, j) \in G \times G$, $w \in Z_N$ such that

$$d_f(i, j; w) = |C_i \cap (C_j - w)| \tag{3.1}$$

is as small as possible, and determine the corresponding sets

$$D_f(i, j; w) = C_i \cap (C_j - w). \tag{3.2}$$

Step 2: Look at the known sequence z^t and find patterns $i * * \cdots * j$ of length $w + 1$ in the known keystream sequence. If such a pattern is found and i is in the k'th position of the known piece of sequence, then

$$k \in D_f(i, j; w) - k'.$$

If $d_f(i, j; w)$ is small, then search the set for k; otherwise, choose another $(i', j'; w')$ and find the corresponding $D_f(i', j'; w')$ $-k''$. Then

$$k \in (D_f(i, j; w) - k') \cap (D_f(i', j'; w') - k'').$$

Continue in this way until the cardinality of the set which contains k is small enough.

There are some relations among the above parameters $d_f(i, j; w)$ which can be used to simplify the differential cryptanalysis. Let $|C_1| = t$. In the binary case the following relations hold

$$
\begin{aligned}
&d_f(0,0;w) + d_f(0,1;w) = N - t;\\
&d_f(0,0;w) + d_f(1,0;w) = N - t;\\
&d_f(1,1;w) + d_f(0,1;w) = t;\\
&d_f(1,1;w) + d_f(1,0;w) = t,
\end{aligned}
$$

where w is any nonzero element of Z_N. These conservation relations tell us that if $d_f(i, j; w)$ is large (resp. small), then $d_f(i + 1, j; w)$ and $d_f(i, j + 1; w)$ must be small (resp. large). Thus, finding triples $(i, j; w)$ such that $d_f(i, j; w)$ is small, is equivalent to finding those such that $d_f(i, j; w)$ is large. Furthermore, by definition and the above conservation relations we have the following relations

$$
\begin{aligned}
d_f(i,i;w) &= d_f(i,i;-w), \quad w \in Z_N;\\
d_f(1,0;w) &= d_f(0,1;w), \quad w \in Z_N.
\end{aligned}
$$

The first of these two relations shows that to find triples $(i, j; w)$ with small w is equivalent to finding those with large w. This is of importance to our cryptanalysis. Thus, the analysis of $d_f(0, 0; w)$ or $d_f(1, 1; w)$ suffices. This reduces the complexity of the differential cryptanalysis.

3.2.2 Form of the Attack

In Step 1 above, the parameter $d_f(i, j; w)$ is in fact the number of ways in which the element w of Z_N can be written as the difference of the elements of C_j and C_i, i.e., it is the number of solutions of the difference equation

$$w = x_j - x_i, \quad x_j \in C_j, \; x_i \in C_i.$$

That is why (3.1) and (3.2) are called difference parameters. The main part of the attack is based on analyzing the difference parameters: the idea is to find small w's or large w's which can be written as differences of some C_i and C_j, as few as possible.

To see why the analysis of the difference parameters can be regarded as a differential analysis, we take $(G, +) = (Z_2, +)$. Consider the input pairs (x, y) such that $x - y = a$, and consider the difference of the corresponding output pairs. Then we have the following expressions

$$
\begin{aligned}
\frac{|\{(x,y): f(x)-f(y)=1,\ x-y=a\}|}{|\{(x,y): x-y=a\}|} &= \frac{d_f(0,1;a)}{N} + \frac{d_f(1,0;a)}{N}\\
\frac{|\{(x,y): f(x)-f(y)=0,\ x-y=a\}|}{|\{(x,y): x-y=a\}|} &= \frac{d_f(0,0;a)}{N} + \frac{d_f(1,1;a)}{N}.
\end{aligned}
$$

3.2.3 Theoretical Basis of the Attack

Theoretically every bit of a keystream can give information about the initial state and key of a generator. Thus a basic requirement for stream ciphers is that every bit of a keystream gives approximately the same amount of information. In our case, this yields balance requirements for the filter function $f(x)$. This *single bit analysis* is apparently suitable to all synchronous stream ciphers. If $n = \log_2 N$, we can write

$$I(k; z_0 = 0) = n - \log_2 |C_0| \quad \text{bits},$$
$$I(k; z_0 = 1) = n - \log_2 |C_1| \quad \text{bits}.$$

Noticing that $|C_0| + |C_1| = N$, we get

$$2^{n-I(k;z_0=0)} + 2^{n-I(k;z_0=1)} = N.$$

This is the theoretical basis for keeping the mutual information stability of a keystream as flat as possible.

If we now consider two bits z_i and z_j separately or arbitrarily, we may not obtain $I(k; z_i) + I(k; z_j)$ bits of information about the key. If the cipher is not properly designed, some combinations of bits may give much more information about the key than others. We call such combinations with their length $(z_i, z_j, |i-j|)$ *bad patterns*. The idea behind the differential attack is to look for bad patterns, and in particular for triples $(i, j; w)$ which gives as much information about the key as possible.

One may argue that we should design our cipher so that the mutual information $I(k; (i, j; w))$ is as small as possible for all $(i, j; w) \in Z_2 \times Z_2 \times Z_N$, but in fact we cannot achieve this: one pattern $(i, j; w) \in Z_2 \times Z_2 \times Z_N$ gives us

$$I = n - \log_2 d_f(i, j; w) = n - \log_2 |C_i \cap (C_j - w)| \text{ bits}$$

of information about the key. Now consider the following theorems:

Theorem 3.1 (Conservation Law for Difference Parameters) *With the symbols as above, we have*

$$\sum_j d_f(i, j; w) = |C_i|, \; i \in Z_2, \; w \in Z_N;$$
$$\sum_i d_f(i, j; w) = |C_j|; \; j \in Z_2, \; w \in Z_N;$$
$$\sum_{(i,j) \in Z_2 \times Z_2} d_f(i, j; w) = N, \; w \in Z_N.$$

These are the laws of conservation between difference parameters which appear in three forms. It follows that the following theorem holds.

Theorem 3.2 (Conservation Law of Mutual Information) *With the symbols as above, we have*

$$\sum_j 2^{n-I(k;(i,j;w))} = |C_i|, \ i \in Z_2, \ w \in Z_N;$$

$$\sum_i 2^{n-I(k;(i,j;w))} = |C_j|; \ j \in Z_2, \ w \in Z_N;$$

$$\sum_{(i,j)\in Z_2\times Z_2} 2^{n-I(k;(i,j;w))} = N, \ w \in Z_N;$$

$$\sum_{(i,j;w)} 2^{n-I(k;(i,j;w))} = N^2.$$

It is not difficult to prove the above theorems, which provide the theoretical basis for the mutual information stability between two-bit patterns and the key. Generalization of the above theorems to the case of an arbitrary finite G is also ready.

In particular, similar results hold for three-bit patterns, and attacks based on three-bit pattern analysis can also be developed. From the above two theorems we see that the differential cryptanalysis outlined in Section 3.2.1 is in fact a kind of mutual information stability analysis. So it may be better to call the differential cryptanalysis the stability cryptanalysis of the natural additive stream ciphers.

3.2.4 The Attack and Pattern Distribution

Many tests and measures of sequence randomness have been proposed in the literature, including Golomb's three postulates. One may ask why these postulates are important; this is shown clearly by differential attacks. Indeed, a stronger randomness postulate is indicated:

Randomness requirements: In a sequence of least period N over $GF(q)$ for each k with $1 \le k \le \lfloor \log_q N \rfloor$, the pattern $(s_t, s_{t+\tau_1}, \cdots, s_{t+\tau_{k-1}} = a)$ of length k appears approximately $\lfloor N/q^k \rfloor$ times when t runs from 0 to $N-1$.

If some keystream sequence has bad patterns with significantly *lower* than random probability, then searching for these patterns may enable us to break it more quickly than by exhaustive search.

3.2.5 Practical Attacks

As shown, the key part of this attack is the first step, i.e., finding some of the parameters of (3.1) and (3.2). In practice we do not need to determine all of these, as a selection of them may be enough to determine the key. To this end, we may first try to determine the C_i for some i, and then use (3.1) and (3.2) to determine the $d_f(i,j;w)$ and $D_f(i,j;w)$. Another method is to introduce the function

$$g_{i,j;w}(x) = (f(x)+1+i)(f(x+w)+1+j),$$

where $f(x)$ is a binary function. It is easy to see that

$$\{x \in Z_N : g_{i,j;w}(x) = 1\} = D_f(i,j;w).$$

This means that the determination of the differential parameters of (3.1) and (3.2) is equivalent to that of the characteristic set of $g_{i,j;w}(x)$. Since the algorithm is known, there are two ways to do this: computation and theoretical analysis.

The complexity of the computational method depends partially on N and $f(x)$. If N is not large enough, it is possible to break the system. For example, if $N = 19$, $f(x) = (x^2 \bmod 19) \bmod 2$, we have

$$\begin{aligned}
&C_0 = \{0,2,17,4,15,5,14\},\\
&D_f(0,0;1) = \{4,14\},\ D_f(0,0;2) = \{17,0,15,2\},\\
&D_f(0,0;3) = \{14,2\},\ D_f(0,0;4) = \{0,17,15\},\\
&D_f(0,0;5) = \{0,14\},\ D_f(0,0;7) = \{14,17\}.
\end{aligned}$$

Thus, one pattern 00 in the output sequences makes the set containing the key to be $\{4,14\}$. Similarly, $0**0$ gives $\{14, 2\}$, $0****0$ gives $\{0,14\}$, etc. Furthermore, two patterns 00 and $0**0$ determine the key 14. Thus, if we can compute difference parameters exhaustively, then we can break such a cipher unless $f(x)$ is extremely well chosen. However, the computational complexity of this task needs to be further investigated.

The other approach, theoretical analysis, means determining the differential parameters algebraically. A poor choice of $f(x)$ may make it easy to find C_i's and thus the differential parameters. Let us illustrate this by the following example.

We choose for the additive natural sequence generator of Figure 3.1.b $N = 4t+1$ a prime with t also prime or $N = 4t-1$ with $2t-1$ also prime, and filter function $f(x) = x \bmod 2$. It is not difficult to show that there is no nontrivial affine function from $(Z_N, +)$ to $(Z_2, +)$ when N is odd. It follows that $f(x)$ is nonlinear with respect to the additions of Z_N and Z_2. It can be proven that $\mathrm{L}(s^\infty) = N-1$ and $\mathrm{SC}_k(s^\infty) \geq N-1$ if $0 \leq k \leq (N-3)/2$, where $\mathrm{L}(s)$ denotes the linear complexity of s, and $\mathrm{SC}_k(s)$ is the sphere complexity of s. By analysis we have

$C_0 = \{0, 2, 4, \cdots, N-3, N-1\}$, $C_1 = \{1, 3, 5, \cdots, N-4, N-2\}$ and the differential results are as follows:

$$\begin{aligned}
&C_0 \cap (C_0 - 1) = \{N-1\} \\
&C_0 \cap (C_0 - 3) = \{N-1, N-3\} \\
&C_0 \cap (C_0 - 5) = \{N-1, N-3, N-5\} \\
&\vdots \\
&C_0 \cap (C_0 - (N-2)) = \{N-1, N-3, N-5, \cdots, 2\}.
\end{aligned}$$

This means that if we have two bits 00 in the known keystream, then the key that generated the first 0 was $k = N-1$. We have

$$\begin{aligned}
&00 \to k = N-1 \\
&0**0 \to k = N-1, \text{ or } N-3 \\
&0****0 \to k = N-1, N-3, \text{or } N-5.
\end{aligned}$$

Similar arguments show that

$$\begin{aligned}
&C_0 \cap (C_1 - 2) = \{N-1\} \\
&C_0 \cap (C_1 - 4) = \{N-1, N-3\} \\
&C_0 \cap (C_1 - 6) = \{N-1, N-3, N-5\} \\
&\vdots \\
&C_0 \cap (C_1 - (N-1)) = \{N-1, N-3, N-5, \cdots, 2\}.
\end{aligned}$$

$$\begin{aligned}
&0*1 \to k = N-1 \\
&0***1 \to k = N-1, \text{ or } N-3 \\
&0*****1 \to k = N-1, N-3, \text{or } N-5.
\end{aligned}$$

$$\begin{aligned}
&C_1 \cap (C_0 - 2) = \{N-2\} \\
&C_1 \cap (C_0 - 4) = \{N-2, N-4\} \\
&C_1 \cap (C_0 - 6) = \{N-2, N-4, N-6\} \\
&\vdots \\
&C_1 \cap (C_0 - (N-1)) = \{N-2, N-4, N-6, \cdots, 1\}.
\end{aligned}$$

$$\begin{aligned}
&1*0 \to k = N-2 \\
&1***0 \to k = N-2, \text{ or } N-4 \\
&1*****0 \to k = N-2, N-4, \text{or } N-6.
\end{aligned}$$

$$
\begin{aligned}
&C_1 \cap (C_1 - 3) = \{N-2\} \\
&C_1 \cap (C_1 - 5) = \{N-2, N-4\} \\
&C_1 \cap (C_1 - 7) = \{N-2, N-4, N-6\} \\
&\vdots \\
&C_1 \cap (C_1 - (N-2)) = \{N-2, N-4, N-6, \cdots, 1\}.
\end{aligned}
$$

$$
\begin{aligned}
&1**1 \to k = N-2 \\
&1****1 \to k = N-2, \text{ or } N-4 \\
&1******1 \to k = N-2, N-4, \text{or } N-6.
\end{aligned}
$$

The above analysis shows that the bad patterns in the keystream are:

- $0*\cdots*0$ and $1*\cdots*1$ of length $2l+1$ with l small enough
- $0*\cdots*1$ and $1*\cdots*0$ of length $2l$ with l small enough.

Thus keys which are very near to $N-1$ (k or $N-k$ for small k) are very weak under this analysis. Those keys which are far away from 0 in both directions are of course strong keys.

Of course, the output sequence of the natural sequence generator of Figure 3.1.b is in this case cryptographically very bad. This is because the nonlinearity (and the difference property) of $f(x)$ is poor. The purpose of the example is to show that if the characteristic sets C_i can be determined algebraically, then the cipher can often be broken, provided that the $f(x)$ has bad difference property. Thus there are two rules a designer of the natural sequence generator must follow:

1. make sure that the characteristic sets of your $f(x)$ cannot be determined (neither computationally nor algebraically); or
2. choose f so that its characteristic sets have good difference property.

If a designer chooses the second option, then the cipher may be secure against this attack even if the characteristic sets are given to an attacker. It is worthy to mention that the characteristic sets of some cryptographic functions are easy to know. One example is the strict majority logic functions used for the nonlinear combiner [6].

3.3 Consistency of the Binary Natural Sequence Generators

The natural sequence generator of Figure 3.1.b is cryptographically attractive because not only every periodic sequence can be produced with a proper choice of the parameters of the generator, but also many of its security aspects are consistent. For the binary natural sequence generator the following cryptographic analyses are equivalent:

1. differential analysis of the cryptographic function $f(x)$;
2. nonlinearity analysis of the cryptographic function $f(x)$;
3. autocorrelation analysis of the cryptographic function $f(x)$;
4. autocorrelation analysis of the output sequence;
5. two-bit pattern distribution analysis of the output sequence;
6. stability analysis of the mutual information $I(i;\ z_i z_{i+t-1})$, here and hereafter z^∞ denotes the output sequence of the natural sequence generator.
7. transdensity analysis of the additive stream cipher system with this natural sequence generator as the keystream generator by which we mean the analysis of the probability of agreement between two encryption resp. decryption transformations specified by two encryption resp. decryption keys [23].

Equivalence is understood in the sense that one analysis results in another analysis.

Proof of the equivalence:

We now prove the equivalence between the above seven analyses and show that an "ideal difference property" of the cryptographic function $f(x)$, ensures automatically

- an ideal nonlinearity of the cryptographic function $f(x)$,
- an ideal autocorrelation property of $f(x)$,
- an ideal autocorrelation property of the output sequence z^∞,
- an ideal two-bit pattern distribution property of the output sequence z^∞, and
- an ideal balance between the mutual information $I(i;\ z_i z_{i+t-1})$ for all possible $(z_i, z_{i+t-1}) \in Z_2 \times Z_2$, where t is arbitrary.

In what follows Z_N denotes the residue class ring modulo an integer N.

Consider now the natural sequence generator of Figure 3.1. Assume that $(G, +)$ is the abelian group over which the keystream sequence is constructed, and $|G| = n$. For each $i \in G$ let $C_i = \{x \in Z_N : f(x) = i\}$. The ordered set $\{C_0, C_1, \cdots, C_{n-1}\}$ is called the *characteristic class*. For any ordered partition $\{C_0, C_1, \cdots, C_{n-1}\}$ of Z_N, there exists a function $f(x)$ with this partition as its characteristic class. Recall that the differential analysis of the system is the analysis of the following *difference parameters*:

$$d_f(i,j;w) = |C_i \cap (C_j + w)|, (i,j) \in G \times G, \ w \in Z_N.$$

We say that f has an *ideal difference property* if the values $d_f(i,j;w)$ are approximately the same for all possible $(i,j;w)$.

In what follows we prove the equivalence between the above seven analyses for the binary natural sequence generator.

Between differential and nonlinearity analysis:

The nonlinearity analysis of the cryptographic function $f(x)$ refers to the analysis of the probability $p(f(x+a) - f(x) = b)$. It can be easily seen that

$$\begin{aligned} Np(f(x) - f(x-a) = 1) &= d_f(0,1;a) + d_f(1,0;a), \\ Np(f(x) - f(x-a) = 0) &= d_f(0,0;a) + d_f(1,1;a) \end{aligned} \tag{3.3}$$

and

$$\begin{aligned} 2d_f(0,0;-a) &= |C_0| - |C_1| + Np(f(x+a) - f(x) = 0), \\ 2d_f(1,1;-a) &= |C_1| - |C_0| + Np(f(x+a) - f(x) = 0), \\ 2d_f(1,0;-a) &= 2d_f(0,1;-a) = N - Np(f(x+a) - f(x) = 0). \end{aligned} \tag{3.4}$$

Then formulae (3.3) and (3.4) show the equivalence.

Between differential and autocorrelation analysis:

The autocorrelation analysis of $f(x)$ refers to the analysis of the autocorrelation function

$$\mathrm{AC}_f(a) = \frac{1}{N} \sum_{x \in Z_N} (-1)^{f(x+a)-f(x)},$$

here we use normalized autocorrelation function. It is easily verified

$$N\mathrm{AC}_f(a) = N - 4d_f(1,0;a) \tag{3.5}$$

and

$$\begin{aligned} 4d_f(0,0;a) &= 4|C_0| - N + N\mathrm{AC}_f(a), \\ 4d_f(1,1;a) &= 4|C_1| - N + N\mathrm{AC}_f(a), \\ 4d_f(1,0;a) &= 4d_f(0,1;a) = N - N\mathrm{AC}_f(a). \end{aligned} \tag{3.6}$$

Combining formulae (3.5) and (3.6) proves the equivalence between the differential and autocorrelation analysis of $f(x)$.

The autocorrelation analysis of the output binary sequence z^∞ refers to the analysis of the autocorrelation function

$$\mathrm{AC}_z(a) = \frac{1}{N} \sum_{i \in Z_N} (-1)^{z_{i+a} - z_i}.$$

Apparently by the definition of the natural sequence generator we have

$$\mathrm{AC}_z(a) = \mathrm{AC}_f(a), \quad \text{for each } a.$$

Thus, the above formulae (3.5) and (3.6) are also true if we replace $C_f(a)$ with $C_z(a)$. This fact shows the equivalence between the differential analysis and the autocorrelation analysis of the output sequence z^∞.

Between differential and two-bit pattern distribution analysis:

The two-bit pattern distribution analysis of z^∞ is concerned with how the two-bit patterns are distributed in a circle of length N in the sequence. For each fixed t with $0 < t \le N-1$ the vector (z_i, z_{i+t}) takes on elements of $Z_2 \times Z_2$ when i runs from 0 to $N-1$. Let $n[(z_i, z_{i+t}) = (a, b)]$ denote the number of times with which the vector (z_i, z_{i+t}) takes on $(a, b) \in Z_2 \times Z_2$ when i runs from 0 to $N-1$. Then we have obviously

$$n[(z_i, z_{i+t}) = (a, b)] = d_f(a, b; -t). \tag{3.7}$$

Thus, for the binary natural sequence generator each difference parameter represents in fact the number of times with which a two-bit pattern appears in a circle of length N of the binary output sequence z^∞.

Between differential and mutual information analysis:

We are given two bits z_i and z_{i+t} of the output sequence of the binary natural sequence generator. It is cryptographically interesting to know how much information these two bits give to the value of the register of the counter in the binary natural sequence generator at the time the output bit z_i was produced. It is easy to verify that

$$I(i; z_i z_{i+t}) = \log_2 N - \log_2 d_f(z_i, z_{i+t}; -t) \quad \text{bits} \tag{3.8}$$

and

$$d_f(z_i, z_{i+t}; -t) = N 2^{-I(i; z_i z_{i+t})}, \tag{3.9}$$

where the mutual information $I(i; z_i z_{i+t})$ is measured in bits. Formulae (3.8) and (3.9) show clearly the equivalence. In addition they show that the difference parameters are in fact a measure of uncertainty.

Between differential and transdensity analysis:

In a cipher system it is possible for two keys to determine the same encryption (resp. decryption) transformation. Even if the two transformations are distinct, it is cryptographically interesting to know the probability of agreement between the two transformations. The control of this probability of agreement may prevent a cipher from a key approximation attack, that is, to use one key to decrypt the message encrypted by another key. Let E_k (resp. D_k) denote the encryption (resp. decryption) transformation specified by the key k. The analysis of the density (briefly, transdensity analysis) of a cipher system refers to the analysis of the probability of agreement $p(E_k(m) = E_{k'}(m))$, where m can be restricted to plaintext blocks or without restriction [23].

For the additive binary stream cipher with the binary natural sequence generator as its keystream generator this probability can be expressed easily as

$$p(E_k = E_{k'}) = \mathrm{AC}_z(k - k' \bmod N) = \mathrm{AC}_f(k - k' \bmod N), \tag{3.10}$$

because of the additive structure of the additive stream cipher and the fact that the keystream sequences specified by all keys are shift versions of each other. Thus, the equivalence follows easily from formula (3.10).

So far we have proved the equivalence between differential analysis and the other six analyses. Thus, the equivalence among the seven analyses follow.

In addition, there is no trade-off between all the above seven aspects and the linear complexity and its stability aspects for this generator (we will see this fact in later sections). This means that it is possible to design the natural sequence generator so that it has not only an ideal property for all the seven aspects in the usual sense, but also a large linear complexity and ideal linear complexity stability for the output sequence. It is because of these facts and because every periodic sequence can be produced by the natural sequence generator that the generator was called a natural one [23].

Formulae (3.3)–(3.9) clearly show that to ensure an ideal property for all the seven aspects, it suffices to control the difference property of the cryptographic function $f(x)$. Thus, in what follows we will concentrate on the control of the difference property of $f(x)$, of the linear complexity, and of the sphere complexity of the output sequence of each specific natural sequence generator.

3.4 Natural Sequence Generator Realizations

There are many ways to generate sequences, for instance, the nonlinear combining of some linear feedback shift registers, the nonlinear filtering of linear feedback shift registers, the clock controlling of linear feedback shift register sequences, to mention only a few. In spite of the flexibility of the ways to generate sequences, it is easily seen that every sequence generator is equivalent to the kind of natural sequence generators depicted in Figure 3.1.b. By equivalence of two generators we mean that they have the same output sequence. In this section we search for the natural sequence generators which can realize some sequences and for the equivalent natural sequence generators of some known generators. To this end, we need theorems about the trace representation of sequences.

It is well known that every periodic sequence in $K = GF(q)$ has a trace representation shown by the following two propositions [44].

Proposition 3.3 *Let s^∞ be a periodic sequence in $K = GF(q)$ whose characteristic polynomial $f(x)$ of degree k is irreducible over K. Let α be a root of $f(x)$ in the extension field $F = GF(q^k)$. Then there exists a uniquely determined $\theta \in F$ such that*

$$s_n = \mathrm{Tr}_{F/K}(\theta\alpha^n), \ \ n = 0, 1, ...,$$

This theorem gives a trace representation only for periodic sequences whose characteristic polynomial is irreducible over K. Generally we have the following conclusion [44].

Proposition 3.4 *Let s^∞ be a periodic sequence in $K = GF(q)$ with characteristic polynomial $f(x) = f_1(x) \cdots f_r(x)$, where the $f_i(x)$ are distinct irreducible polynomials over K. For $i = 1, \cdots, r$, let α_i be a root of $f_i(x)$ in its splitting field F_i over K. Then there exist uniquely determined elements $\theta_1 \in F_1, \cdots, \theta_r \in F_r$ such that*

$$s_n = \mathrm{Tr}_{F_1/K}(\theta_1\alpha_1^n) + \cdots + \mathrm{Tr}_{F_r/K}(\theta_r\alpha_r^n), \ \ n = 0, 1, ...,$$

Now we describe the natural sequence generator realization of periodic sequences in the finite field $K = GF(q)$. Let the sequence s^∞ be one depicted in Proposition 3.3, then one of its natural sequence generator realizations is depicted by Figure 3.2. For the sequence s^∞ of Proposition 3.4 we have a natural sequence generator realization of Figure 3.3. The natural sequence generator realization of the maximum-length sequences is easy with the above two theorems. If one has a characteristic polynomial or minimal polynomial of a sequence, then it is possible to give the natural sequence generator realization. Thus, the natural sequence generator realization of many of the known sequence generators may be done.

3.5 A Basic Theorem and Main Bridge

The cryptographic importance of linear and sphere complexity was explained in Sections 2.3.1 and 2.3.4 respectively. Now we prove a theorem which can be used to set up bounds for the linear and sphere complexity of any sequences.

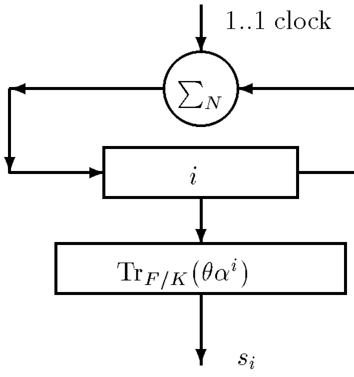


Figure 3.2: The natural sequence generator realization of some sequences

Basic Theorem 3.5 *Let $N = p_1^{e_1} \cdots p_t^{e_t}$, where $p_1, \cdots, p_t$ are t pairwise distinct primes, q a positive integer such that* $\gcd(q, N) = 1$. *Then for each nonconstant sequence s^∞ of period N over $GF(q)$, we have*

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{p_1}(q), \cdots, \mathrm{ord}_{p_t}(q)\}$$

and

$$\mathrm{SC}_k(s^\infty) = \begin{cases} \geq & \min\{\mathrm{ord}_{p_1}(q), \cdots, \mathrm{ord}_{p_t}(q)\}, \\ & \textit{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}, \\ = 0, & \textit{otherwise.} \end{cases}$$

where $\mathrm{L}(s^\infty)$ *and* $\mathrm{SC}_k(s^\infty)$ *denote the linear and sphere complexity of the sequence respectively,* $\mathrm{WH}(x)$ *the Hamming weight of x,* $\mathrm{ord}_p(q)$ *denotes the least positive integer m such that $p^m \equiv 1 \pmod{q}$, and $s^N = s_0 s_1 \cdots s_{N-1}$.*

Proof: Let K be a field of characteristic p, n a positive integer not divisible by p, and ξ a primitive nth root of unity over K. The nth cyclotomic polynomial is defined by

$$Q_n(x) = \prod_{s=1, \gcd(s,n)=1}^{n} (x - \xi^s).$$

To prove the theorem, we need the following properties of the cyclotomic polynomial (see Lidl and Niederreiter [44] for proof).

1. $Q_n(x)$ is independent of the choice of ξ.
2. $\deg(Q_n(x)) = \phi(n)$.

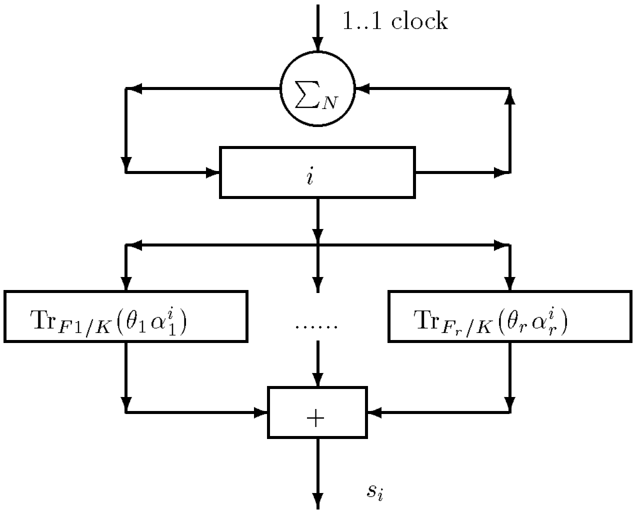


Figure 3.3: The natural sequence generator realization of some general sequences

3. The coefficients of $Q_n(x)$ belong to the prime subfield of K.
4. $x^n - 1 = \prod_{d|n} Q_d(x)$.
5. If $K = \mathrm{GF}(q)$ with $\gcd(q, n) = 1$, then Q_n factors into $\phi(n)/d$ distinct monic irreducible polynomials in $K[x]$ of the same degree d, where d is the least positive integer such that $q^d = 1 \bmod n$, i.e., d is the order (or exponent) of q modulo n, denoted as $\mathrm{ord}(q)$ modulo n or $\mathrm{ord}_n(q)$.

It is easily seen that $\mathrm{ord}_{p^k}(a) \geq \mathrm{ord}_p(a)$ for any prime p and any positive integer a with $\gcd(a, p) = 1$. By assumptions and the above basic property 5 the polynomial $x^n - 1$ is equal to the product of $\phi(n)/d$ distinct monic irreducible polynomials over $GF(q)[x]$ of the same degree d, where d is the least positive integer such that $q^d = 1 \bmod n$. Since the minimum polynomial of each sequence of period N over $GF(q)$ divides $x^N - 1$ and $x^N - 1 = \prod_{n|N} Q_n(x)$, we consider the orders $\mathrm{ord}_n(q)$ for each possible divisor n of N.

If n divides N, there are integers $h_{i_1}, ..., h_{i_s}$, where $1 \leq h_{i_j} \leq e_{i_j}$ for $1 \leq j \leq s$ and $1 \leq s \leq t$, such that $n = p_{i_1}^{e_{i_1}} \cdots p_{i_s}^{e_{i_s}}$. By the Chinese Remainder Theorem and the above conclusions

$$\begin{aligned}
\mathrm{ord}_n(q) &= \mathrm{lcm}\{\mathrm{ord}_{p_1^{e_1}}(q), \cdots, \mathrm{ord}_{p_t^{e_t}}(q)\} \\
&\geq \max\{\mathrm{ord}_{p_1^{e_1}}(q), \cdots, \mathrm{ord}_{p_t^{e_t}}(q)\} \\
&\geq \min\{\mathrm{ord}_{p_1}(q), \cdots, \mathrm{ord}_{p_t}(q)\}.
\end{aligned}$$

Thus, the conclusion of the first part of this theorem follows.

When $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$, changing any k positions in corresponding positions of each period does not give a constant sequence. The conclusion of the second part then follows. □

One can see that the above lower bound is optimal. If $t = 1$ and $e_1 = 1$, then we have the general lower bound for sequences of a prime period. If $t = 1$, then it gives a lower bound for the linear complexity and sphere complexity of sequences with period being a prime power. Many theorems and corollaries in this thesis are special cases of the above basic theorem, that is why we call it a basic theorem. We say that it is a bridge between number theory and stream ciphers because it makes a clear connection between the linear and sphere complexity of sequences and quite many number-theoretic problems such as primes of special forms (e.g., Sophie German primes, i.e., primes p with $(p-1)/2$ being also prime, and twin primes) and their distributions, primality testing, primitive roots and their distributions, and primitivity testing. Some of these connections will be made clear in later chapters.

This basic theorem shows that it is usually quite easy to control the global linear and sphere complexity. However, it seems fairly difficult to control the local linear and sphere complexity. It is worthy to note that here we use specialities of the period to control the linear and sphere complexity, while some cryptographic functions are traditionally used to control the global linear complexity. Thus, we stress the importance of the period. Basic Theorem 3.5 will be used many times in the sequel.

Chapter 4

Difference and Cryptographic Functions

Every cryptographic function is directly or indirectly responsible for combining plaintext digits and keystream digits. It is the cryptographic functions that constitute the encryption and decryption algorithms. So the security of every cryptosystem should depend *mainly* on the design of the cryptographic functions.

Cryptographic requirements for those functions vary from system to system. They depend on the known attacks on a system. So it may be very difficult to say which functions are cryptographically good ones without specific cryptographic contexts. It may be concluded that the cryptographic goodness and badness of functions depend not only on the structure of the function, but also on the specific system in which they are used and how they are used. This is apparent from the role of linear functions in cryptology. This chapter is mainly concerned with the relation between the difference and other properties of cryptographic functions.

The main contributions of this chapter are the following:

1. The connections among the difference property, correlation functions, nonlinearity, information stability, and linear approximation described by Theorems 4.2, 4.4, 4.5, 4.6, 4.7, 4.8, 4.9, 4.10, and Corollary 4.2;

2. the introducing of almost difference sets and the searching for them, including three classes of almost difference sets described in Theorems 4.15, 4.16, and 4.17;

3. the connections among the almost difference set property, autocorrelation property, nonlinearity, and linear approximation, as described by Theorems 4.18 and 4.19.

This chapter is based partly on the papers [23, 24], and partly on my recent results.

4.1 Rudiments of Difference Sets

Since the autocorrelation function of the difference set characterized sequences is two-level, difference sets have been extensively studied from the engineering point of view. The application of difference sets in securing communications has also been investigated.

Let us recall that an (N, k, λ) difference set $D = \{d_1, \cdots, d_k\}$ is a subset of an abelian group $(G, +)$ such that, for each $a \neq 0$, the equation

$$d_i - d_j = a$$

has exactly λ solutions (d_i, d_j) with d_i and d_j in D, where $|G| = N$. They can also be defined in nonabelian groups [3]. Difference sets are called *cyclic* or *noncyclic* according to whether the group G is cyclic.

So far two types of difference sets have been intensively studied: the residue difference sets, i.e., difference sets of $(Z_N, +)$'s and difference sets of $(GF(q)^n, +)$, which are cyclic and noncyclic respectively. Difference sets in $(GF(2)^{2n}, +)$ correspond to the so-called bent functions. A survey of bent Boolean functions, and thus of difference sets of $(GF(2)^{2n}, +)$, was given by Dillon [18]. Difference sets of both types are very useful in cryptography. However, for our applications, we are mainly interested here in cyclic difference sets. Therefore in most of the sections in this chapter residue difference sets are mainly discussed.

Difference sets are not easy to construct. On the other hand, for $N = 4t+1$ there is no $(N, (N-1)/2, \lambda)$ difference set, because 4 cannot divide $(N-3)$. However, there are some sets which have almost the same properties as difference sets. Let N be an odd number, and $D = \{d_1, \cdots, d_k\}$ a subset of an abelian group $(G, +)$. If for each of half of the nonzero elements a of Z_N, the equation

$$d_i - d_j = a$$

has exactly λ solutions (d_i, d_j) with d_i and d_j in D, and for each of the other half exactly $\lambda + 1$ solutions, we call D an (N, k, λ) *almost difference set.*

Given the difference set $D = \{d_1, \cdots, d_k\}$ of Z_N, then for any s in Z_N the set $s + D = \{d_1 + s, \cdots, d_k + s\}$ is also a difference set of Z_N and is called a *shift* (or *cyclic shift*) of the original set D. Again, for any t in Z_N with $\gcd(t, N) = 1$, the set $tD = \{td_1, \cdots, td_k\}$ is also a difference set with the same parameters (N, k, λ), where the product td_i takes the multiplication of the ring $(Z_N, +, \cdot)$. If D_i and D_j are two difference sets having the same parameters and if $D_i = tD_j + s$ for any t and s in Z_N with $\gcd(t, N) = 1$, then D_i and D_j are said to be *equivalent difference sets.*

If D_1 is a difference set of Z_N, then $D_2 = Z_N \setminus D_1$ is also a difference set of Z_N, and is called the *complementary difference set* of Z_N. If D has parameters (N, k, λ), then its complement D^* has parameters $(N^*, k^*, \lambda^*) = (N, N-k, N-2k+\lambda)$. Set $n = k-\lambda$, then both N and n are invariant under complement of the difference sets. For most purposes it is sufficient to consider only one of a pair of complementary difference sets. This is frequently done by insisting that k be less than $N/2$.

For applications the most important problems concerning difference sets are the existence and the construction problem. As an immediate consequence of the definition the relation

$$k(k-1) = (N-1)\lambda$$

necessarily holds among the parameters N, k, λ. Given any positive integer N there are certain obvious difference sets of Z_N. These are the following:

1. the null set $D = \phi$;
2. all singletons $D = \{i\}$, $0 \le i \le N-1$;
3. $D = Z_N$;
4. $D = Z_N \setminus \{i\}$, $0 \le i \le N-1$.

These difference sets are called *trivial* and are often ignored.

There are some general results about the parameters of difference sets which can be summarized in the following proposition.

Proposition 4.1 *Let D be an (N, k, λ) difference set in a group G. Then*

1. *If N is even, then n is a square.*
2. *If N is odd, then the equation*

$$x^2 = ny^2 + (-1)^{\frac{N-1}{2}} \lambda z^2$$

has a solution in integers x, y, z, not all zero.

Generally, for two sets D_1 and D_2 of an abelian group $(G, +)$, if there exists an automorphism α of G such that

$$D_1^\alpha = D_2 + g$$

for some $g \in G$, where D_1^α is the image of D_1 under the automorphism, then D_1 and D_2 are said to be equivalent. Particularly, α is said to be a *multiplier* of D if $D^\alpha = D + g$ for some $g \in G$. Multipliers are useful in constructing difference sets.

For any binary sequence $A^N = a_0 a_1 \cdots a_{N-1}$ of length N, we call the set $C_A = \{i : \ a_i = 1, \ 0 \le i \le N-1\}$ the *characteristic set* of the finite sequence A^N. On the other hand, every subset of Z_N defines a binary sequence of length N, i.e., for each i, $a_i = 1$ if i belongs to the set; otherwise $a_i = 0$. We call the sequence the *characteristic sequence* of the subset. Thus, there is a one-to-one correspondence between the set of all binary sequences with length N and the set of all subsets of Z_N. For any subset S of Z_N, denoting s^N as its characteristic sequence, we call the infinite sequence $s^\infty = s^N s^N \cdots$, the *periodic characteristic sequence* of the set S.

From the cryptographic point of view, we are mostly concerned with the cryptographic attributes of the periodic characteristic sequence of difference sets. In an (N, k, λ) difference set if $N/k \approx 1/2$, then its periodic characteristic sequence has good balance between 0's and 1's, and furthermore, the out-of-phase value of the autocorrelation function is approximately $\pm 1/N$. However, to employ these sequences as keystream sequences for some binary additive stream ciphers, we must be able to control the linear complexity and other statistical properties of the sequences.

In cryptography we are interested not only in the characteristic sequences, but also in the characteristic functions of difference and almost difference sets. The nonlinearity and difference property of those characteristic functions are essential for some cryptographic applications.

4.2 Difference Sets and Correlation Functions

Let D be a subset of Z_N, the *characteristic function* $f_D(x) : Z_N \to \{0,\ 1\}$ of D is defined by

$$f_D(x) = \begin{cases} 1, & \text{if } x \in D, \\ 0, & \text{otherwise.} \end{cases}$$

On the other hand, for every function $g(x) : Z_N \to \{0,\ 1\}$, let $S = \{x : g(x) = 1, x \in Z_N\}$, then $g(x) = f_S(x)$. The set S is called the *characteristic set* of the function $g(x)$. Consequently, the difference set property must be characterized by their characteristic functions. To this end, we need the autocorrelation function of functions from Z_N to $\{0,\ 1\}$ which is given by

$$\mathrm{AC}_f(w) = \sum_{x \in Z_N} (-1)^{f(x)+f(x+w)}/N.$$

Let D be a subset of Z_N. For any $w \in Z_N$, let

$$d_D(w) = |(w + D) \cap D|,$$

then we have $d_D(w) = d_D(-w)$ and

$$\sum_{0 \neq w \in Z_N} d_D(w) = |D|(|D| - 1).$$

It is also easy to know that $d_D(w)$ is the number of times with which w appears in the set $D^- = \{d_1 - d_2 : d_1, d_2 \in D, d_1 \neq d_2\}$. So we call the function $d_D(w)$ the *difference function* of the subset D. This difference function is closely related to the autocorrelation function of the characteristic function of the subset D.

Theorem 4.2 *Let D be a subset of Z_N with $|D| = k$, and $f_D(x)$ be the characteristic function of D and s_D^∞ be the periodic characteristic sequence of D. Then we have*

$$\mathrm{AC}_{f_D}(w) = \mathrm{AC}_{s_D}(w) = [N - 4(k - d_D(w))]/N.$$

Proof: By definition we have

$$\begin{aligned}\mathrm{AC}_{f_D}(w) = \mathrm{AC}_{s_D}(w) &= [(\sum_{x\in D} + \sum_{x\in D^*})(-1)^{f(x)+f(x+w)}]/N \\ &= [|(-w+D)\cap D| - |(-w+D)\cap D^*| \\ &\quad +|(-w+D^*)\cap D^*| - |(-w+D^*)\cap D|]/N,\end{aligned}$$

where $D^* = Z_N \setminus D$.

It is not difficult to see that

$$\begin{aligned}&|(-w+D)\cap D^*| = k - d_D(w),\\ &|(-w+D^*)\cap D| = k - d_D(w),\\ &|(-w+D^*)\cap D^*| = N - 2k + d_D(w).\end{aligned}$$

Combining the above results, we get the conclusion of the theorem. □

The following autocorrelation function characterization of difference sets follows immediately from Theorem 4.2.

Corollary 4.3 *Let D be a subset of Z_N with $|D| = k$, and $f_D(x)$ be the characteristic function of D. Then D is an $(N,\ k,\ \lambda)$ difference set of Z_N if and only if for each $w \neq 0$,*

$$\mathrm{AC}_{f_D}(w) = \mathrm{AC}_{s_D}(w) = [N - 4(k - \lambda)]/N.$$

This theorem shows that D is a difference set if and only if the autocorrelation function of its characteristic function and that of its characteristic sequence are two valued. If $f(x)$ is a cryptographic function, its autocorrelation function means the extent to which $f(x)$ can be approximated by its delayed phase $f(x + a)$.

4.3 Difference Sets and Nonlinearity

Before investigating the application of the characteristic functions of difference sets to stream ciphers, we need to analyze their nonlinearity. For cryptographic purposes the nonlinearity of functions has been extensively investigated. There are two related concepts for nonlinearity of functions from a finite abelian group $(G, +)$ to another finite abelian group $(H, +)$: *local nonlinearity* and *global nonlinearity*. Let $f(x): (G, +) \to (H, +)$ be a function, then

$$p_f(\alpha) = \max_{r \in H} p(f(x+\alpha) - f(x) = r), \quad 0 \neq \alpha \in G$$

is the measure of the local nonlinearity at α and

$$p_f = \max_{\alpha \neq 0} p_f(\alpha)$$

is that of the global nonlinearity *with respect to* the pair of operations of the two groups, here and hereafter $p(\cdot)$ denotes the probability. Obviously, the function $p_f(\alpha)$ can be viewed as a kind of differential. It should be noted that in many cases it makes no sense to use the minimum distance between a mapping and all linear mappings as a kind of nonlinearity measure, since it often happens that there are only trivial linear mappings from an abelian group to another one.

For the nonlinearity of the characteristic functions of subsets of Z_N, we have the following conclusion.

Theorem 4.4 *Let D be a subset of Z_N with $|D| = k$, and $f(x)$ the characteristic function of D. Then for any $\alpha \neq 0$,*

$$p(f(x) - f(y) = \beta \mid x - y = \alpha) = \begin{cases} [N - 2(k - d_D(\alpha))]/N, & \beta = 0, \\ 2(k - d_D(\alpha))/N, & \beta = 1. \end{cases}$$

Proof: Let D^* be the complement of D with respect to Z_N. By assumptions, for each $\alpha \neq 0$ we have

$$|\{(x, y) : \ x, \ y \in Z_N, \ x - y = \alpha\}| = N$$

and

$$\begin{aligned} &|\{(x, y) : \ x, \ y \in D^*, \ x - y = \alpha\}| = N - 2k + d_D(\alpha), \\ &|\{(x, y) : \ (x, y) \in D \times D^* \cup D^* \times D, x - y = \alpha\} \\ &= |(D^* + \alpha) \cap D| + |(D + \alpha) \cap D^*| = 2(k - d_D(\alpha)). \end{aligned}$$

It follows that

$$p(f(x) - f(y) = 0 | x - y = \alpha) = \frac{N - 2(k - d_D(\alpha))}{N}$$

and

$$p(f(x) - f(y) = 1 | x - y = \alpha) = \frac{2(k - d_D(\alpha))}{N}.$$

This proves the theorem. □

Theorem 4.5 *Let D be an $(N,\ k,\ \lambda)$ difference set of Z_N, and $f(x)$ be the characteristic function of D. Then for any $\alpha \neq 0$,*

$$p(f(x) - f(y) = \beta \mid x - y = \alpha) = \begin{cases} [N - 2(k - \lambda)]/N, & \beta = 0, \\ 2(k - \lambda)/N, & \beta = 1. \end{cases}$$

If f is the characteristic function of an (N, k, λ) difference set, this theorem shows that for each $\alpha \neq 0$,

$$p_f(\alpha) = p_f = \max\{[N - 2(k - \lambda)]/N,\ 2(k - \lambda)/N\}.$$

Noticing that $[N - 2(k - \lambda)]/N + 2(k - \lambda)/N = 1$, we see that, to ensure good nonlinearity of the characteristic function $f(x)$, it suffices to make $k - \lambda \approx N/4$. This is actually a way to keep the *stability of local nonlinearities*. Conversely, the theorem indicates that if a function from Z_N to Z_2 has good nonlinearity, then its characteristic set must have good difference property.

4.4 Difference Set and Information Stability

Let $z = f(x)$ be a function from Z_N to $GF(2)$ and C_f be its characteristic set. Assume that x is a random variable which takes on each element of Z_N equally likely, then z is also a random variable. For the binary NSG it is important to keep a balance between the mutual information $I(x; z = 1)$ and $I(x; z = 0)$ if $f(x)$ is used as the cryptographic function of the generator [23], where $z = f(x)$. For the sake of simplicity, let $k = |C_f|$, i.e., the number of elements of C_f.

Since x takes on each element of Z_N equally likely, the self information of x is

$$I(x) = \log_2 N \quad \text{bits}.$$

On the other hand, it can be easily seen that

$$I(x;\ z = 1) = -\log_2 k/N;$$
$$I(x;\ z = 0) = -\log_2 (N - k)/N.$$

Naturally, we want to have both $I(x;\ z = 1)$ and $I(x;\ z = 0)$ as small as possible. However, note that $k + (N - k) = N$. Thus, if one of $I(x;\ z = 1)$ and $I(x;\ z = 0)$ is very small, the other must be very large. Hence the best we can do is to balance them. Thus, if N is odd, the best k is $(N-1)/2$ or $(N+1)/2$; if N is even, the best k is $N/2$. This means that k should be as close to $N/2$ as possible.

Note that

$$\frac{1}{2^{I(x;\ z=1)}} + \frac{1}{2^{I(x;\ z=0)}} = 1,$$

we say that the mutual information $I(x;\ f(x))$ is *optimum*, if $|I(x;\ f(x)=0) - I(x;\ f(x)=1)|$ is minimal. Thus, $I(x;\ f(x))$ is optimum if and only if $k = (N-1)/2$ or $(N+1)/2$ when N is odd and $k = N/2$ when N is even.

In some stream cipher systems it is important to ensure that two bits of outputs $[f(x), f(x+a)]$ give the minimal information about the input x, where a is any fixed element of Z_N. To simplify notations, let us use C in place of C_f. Since

$$|\{x \in Z_N : (f(x), f(x+a)) = (1,1)\} = |C \cap (C-a)| = d_C(-a),$$

we have

$$I(x; (f(x), f(x+a)) = (1,1)) = \log_2 N - \log_2 d_C(-a).$$

Similarly, we have

$$\begin{aligned} I(x; (f(x), f(x+a)) = (1,0)) &= \log_2 N - \log_2(k - d_C(-a)), \\ I(x; (f(x), f(x+a)) = (0,1)) &= \log_2 N - \log_2(k - d_C(-a)), \\ I(x; (f(x), f(x+a)) = (0,0)) &= \log_2 N - \log_2(N - 2k + d_C(-a)). \end{aligned}$$

Assume that $k = (N-1)/2$. To balance the above four information quantities, $d_C(-a)$, $k - d_C(-a)$, and $N - 2k + d_C(-a)$ should be about the same for all $a \neq 0$. When C is an $(N, (N-1)/2, (N-3)/4)$ difference set, the balance among the four information quantities is optimum.

Let $C_0 = C_f$, $C_1 = Z_N \setminus C_0$. To analyze the mutual information

$$I(x; (f(x+a_0), \cdots, f(x+a_{t-1})) = (l_0, \cdots, l_{t-1}))$$

is equivalent to analyzing the parameter

$$d_{l_0 \cdots l_{t-1}}(a_0, \cdots, a_{t-1}) = \left| \cap_{i=0}^{t-1} (C_{l_i} - a_i) \right|.$$

To stabilize the above mutual information, it is necessary and sufficient to require that the parameter be as flat (stable) as possible, i.e., it is approximately $N/2^t$ for all $l_0, \cdots, l_{t-1} \in Z_2$, and pairwise distinct elements $a_0, \cdots, a_{t-1}$ of Z_N. This requirement is much stronger than that for difference sets when $t \geq 3$.

4.5 Difference Set and Linear Approximation

In some cases it is possible to construct a deciphering machine to approximate the original one [22]. To this end, linear approximation approaches are often employed. In many cipher systems the linearity with respect to some operations is fatal. However, employing linear functions may be desirable (for example, to achieve diffusion [49], and to control the transdensity [23]). The requirement depends on the system.

Let D be an (N, k, λ) difference set of Z_N. We now investigate the best affine approximations of the characteristic function $f_D(x)$. Here the best affine approximation of $f_D(x)$ is with respect to the additions of the rings Z_N and Z_2. Let $L_{Z_N:Z_2}$ denote the set of all affine functions from Z_N to Z_2, here the affine functions are defined to be the $g(x)$'s such that $g(x) - l$ is linear for some $l \in Z_2$. If $l(x) \in L_{Z_N:Z_2}$ and

$$p(f_D(x) = l(x)) = \max_{g(x) \in L_{Z_N:Z_2}} p(f_D(x) = g(x)),$$

then $l(x)$ is said to be a *best affine approximation* of $f_D(x)$.

Before studying the problem, let us have a look at the affine functions from $(Z_N, +)$ to $(Z_2, +)$. Apparently, $l(x) = 0$ and $l(x) = 1$ are two trivial affine functions from Z_N to Z_2. Assume that $l(x)$ is a linear function. Let $V_0 = \{x : f(x) = 0, x \in Z_N\}$ and $V_1 = Z_N \setminus V_0$. Then it is easy to see that $(V_0, +)$ is a subgroup of $(Z_N, +)$. If $V_1 \neq \Phi$ and $V_0 \neq \Phi$, let a be one element of V_1. Then it follows easily that $V_0 + a \subseteq V_1$, so $|V_1| \geq |V_0|$. Similarly, we can prove that $|V_0| \geq |V_1|$. It follows therefore $|V_0| = |V_1|$. On the other hand, if N is even, then $l(x) = x \bmod 2$, is a nontrivial linear function. Hence, we have proved the following theorem.

Theorem 4.6 *There is a nontrivial affine function from $(Z_N, +)$ to $(Z_2, +)$ if and only if N is even. And if $l(x)$ is a nontrivial affine function, then $|f^{-1}(0)| = |f^{-1}(1)| = N/2$, where $f^{-1}(i) = \{x : f(x) = i, x \in Z_N\}$ for $i = 0, 1$.*

This theorem means there is no sense to study the affine approximation problem for the functions from Z_N to Z_2 with respect to the additions of the two rings when N is odd. So we discuss only the affine approximation of the characteristic functions of difference sets in Z_N with even N.

Let $f(x)$ be a function from Z_N to Z_2, $a(x) = l(x) + u$ be an affine function, where $l(x)$ is linear. From the definition of autocorrelation function it follows immediately that the following theorem is true.

Theorem 4.7 *Let $f(x)$ be a function from Z_N to Z_2, $h(x) = f(x)+l(x)+u$, where $l(x)$ is linear and u is a constant. Then for each $w \neq 0$*

$$\mathrm{AC}_h(w) = (-1)^{l(w)}\mathrm{AC}_f(w).$$

Theorem 4.8 *Let D be an (N, k, λ) difference set of Z_N, $f_D(x)$ the characteristic function of D, $l(x)$ any nontrivial affine function from $(Z_N, +)$ to $(Z_2, +)$ and $h(x) = f_D(x) + l(x)$. Let H be the characteristic set of $h(x)$, $s = |H|$ and $c = [N - 4(k - \lambda)]/N$. Then*

1. $s = |H| = \left[N + \sqrt{N - (-1)^{l(0)}cN}\right]/2$ *or* $\left[N - \sqrt{N - (-1)^{l(0)}cN}\right]/2$;
2. $d_H(w) = s - N[1 - (-1)^{l(w)}c)]/4$ *for each* $w \neq 0$;
3. $p(f_D(x) = l(x)) = \frac{1}{2} \pm \frac{\sqrt{N-(-1)^{l(0)}cN}}{2N} = \frac{1}{2} \pm \frac{\sqrt{1-(-1)^{l(0)}c}}{2\sqrt{N}}$,

where $p(f_D(x) = l(x))$ denotes the probability of agreement between $f_D(x)$ and $l(x)$.

Proof: It follows from Theorems 4.2 and 4.6 as well as Corollary 4.3 that

$$[N - 4(s - d_H(w))]/N = \mathrm{AC}_h(w) = (-1)^{l(w)}\mathrm{AC}_{f_D}(w) = (-1)^{l(w)}c$$

for each $w \neq 0$. This gives

$$d_H(w) = s - [N(1 - (-1)^{l(w)}c)]/4$$

for each $w \neq 0$. On the other hand, by employing the conservation law of difference

$$\sum_{0 \neq w \in Z_N} d_H(w) = s(s - 1)$$

we get further

$$s(s - 1) = \sum_{w \neq 0} d_H(w) = s(N - 1) - N[N - 1 - \sum_{w \neq 0}(-1)^{l(w)}c]/4.$$

This fact together with Theorem 4.6 gives the following equation

$$s^2 - Ns + N[N - 1 + (-1)^{l(0)}c]/4 = 0.$$

The solutions of the equation prove the remaining parts of the theorem. □

Note that for every difference set of Z_N, the constant c satisfies $-1 < c < 1$. It follows that for any nontrivial affine function $h(x)$ and the characteristic function $f(x)$ of any difference set of Z_N , we have

$$\frac{1}{2} - \frac{1}{\sqrt{2N}} < p(f(x) = h(x)) < \frac{1}{2} + \frac{1}{\sqrt{2N}}.$$

This rather surprising result means that nontrivial affine approximation of the characteristic function of any difference set of Z_N is relatively bad. The above lower bound for c can be improved by making use of $k(k-1) = (N-1)\lambda$.

Another important fact, which follows from the above theorem, is that if the constant c is approximately zero (i.e., $k - \lambda \approx N/4$), then the set H has also good difference property.

Difference sets with parameters of the form $(4m, k, \lambda)$ in a group were also called *Menon difference sets*, *Hadamard difference sets* or H-sets. A rather remarkable result about the parameters of Menon difference sets proved by Menon is that a Menon difference set has parameters of one of the forms

$$(N, k, \lambda, n) = (4m^2, 2m^2 \mp m, m^2 \mp m, m^2),$$

where n is defined to be $k - \lambda$.

The characteristic functions of the Menon difference sets have the worst probability of agreement with all affine functions. Equivalently, they have the best nonlinearity with respect to the additions of Z_{4m^2} and Z_2.

4.6 Almost Difference Sets

The definition of almost difference sets was given in Section 4.1. To search for cryptographic functions, we shall investigate almost difference sets and their importance in cryptography further in the following sections. Let us stipulate that the integer N in this section and Sections 4.7, and 4.8 is odd.

From the definition of (N, k, λ) almost difference sets of Z_N, it follows immediately that the following necessary condition

$$k(k-1) = (2\lambda + 1)(N-1)/2$$

holds for all (N, k, λ) almost difference sets of Z_N. It is obvious that every odd integer (≥ 3) must be of one of the two forms $4t+1$ and $4t-1$. If $N = 4t-1$, then $(N-1)/2 = 2t-1$ is odd, it follows that $(2\lambda+1)(N-1)/2$ must be odd. This proves the following result.

Theorem 4.9 *Let N be an odd integer. If Z_N has an (N, k, λ) almost difference set, then N must be of the form $4t+1$.*

Now we search for cryptographic almost difference sets of Z_N. It is known that if $N = 4t+1$, is a prime, then the quadratic residues modulo N form an $(N, (N-1)/2, (N-5)/4)$ almost difference set, which can be seen from the following proposition, which is easy to prove.

Proposition 4.10 *Let $N = 2f + 1$ be prime. Then the cyclotomic numbers of order 2 are given by*

$$(0,0) = (f-2)/2; \ (0,1) = (1,0) = (1,1) = f/2$$

if f is even, and by

$$(0,0) = (1,0) = (1,1) = (f-1)/2; \ (0,1) = (f+1)/2$$

otherwise.

The reader is referred to Section 5.1 for the definition of cyclotomic numbers. Now a natural question is whether the biquadratic residues form an $(N, t, (t-3)/4)$ almost difference set, if $N = 4t + 1$ is a prime. The following necessary condition is easy to see.

Theorem 4.11 *Let $N = 4t + 1$. If the biquadratic residues modulo N form an $(N, t, (t-3)/4)$ almost difference set, then t must be of the form $4u + 3$.*

Taking $N = 13$ as an example, we have the set of biquadratic residues $D = \{1, 3, 9\}$, $\lambda = 0$. Another positive example is the case $N = 29$. A negative example is the case $N = 61$. For this example we have $t = 15$ and

$$D = (2) = \{1, 16, 12, 9, 22, 47, 20, 15, -4, -3, -48, -36, -27, -5, -19\}.$$

This example shows that the biquadratic residues modulo N may not form an almost difference set, if $N = 4t + 1$ with t being a prime of the form $4u + 3$. However, we have the following general conclusion:

Theorem 4.12 *Let $N = 4t + 1$, be a prime with $t \equiv 3 \pmod 4$, and D be the biquadratic residues modulo N. Then the following equation*

$$a = x - y, \quad (x, y) \in D \times D, \ x \neq y$$

has u solutions for half of the nonzero elements a of Z_N, and v solutions for the other half, where $u \geq 0$, $v \geq 0$ and $u + v = t(t-1)/2$.

Proof: Let α be a primitive root modulo N and $D = (\alpha^4)$ the multiplicative subgroup of Z_N^*. Then D is the set of biquadratic residues modulo N. Let

$$D^- = \{d_i - d_j : \ (d_i, d_j) \in D \times D, i \neq j\}.$$

By assumptions we have $-1 = \alpha^{2t} = \alpha^{8u-2}$. This gives $\alpha^2 D = -D$ and $\alpha^3 D = -\alpha D$. Hence, we get

$$|D^- \cap \alpha^2 D| = |D^- \cap D| = u,$$
$$|D^- \cap \alpha^3 D| = |D^- \cap \alpha D| = v.$$

Furthermore, we have

$$\begin{aligned} t(t-1) &= |D^-| \\ &= |D^- \cap D| + |D^- \cap \alpha D| + |D^- \cap \alpha^2 D| + |D^- \cap \alpha^3 D| \\ &= 2u + 2v. \end{aligned}$$

Thus, the conclusion of the theorem follows. □

As already mentioned above, the set of biquadratic residues D is a multiplicative subgroup of the group $(Z_N^*, \cdot)$, where $Z_N^* = Z_N \setminus \{0\}$. This property makes the set of biquadratic residues cryptographically attractive if they have good difference property, because the multiplicative group structure can lead to an easy realization of their characteristic functions. The following theorem gives necessary and sufficient conditions which ensure the almost difference set property of the biquadratic residues.

Theorem 4.13 *Let a prime* $N = 4t + 1 = x^2 + 4y^2$ *with* $x = 1 \bmod 4$ *and* t *being odd. Then the biquadratic residues modulo* N *form an* $(N, t, (t-3)/4)$ *almost difference set if and only if* $x = 5$ *or* -3.

Proof: We consider the binary cyclotomic numbers of order 4. Let $N = 4t + 1$ be a chosen prime. Then N can be expressed as $N = x^2 + 4y^2$, $x = 1 \bmod 4$, here y is two-valued, depending on the choice of the primitive root [14]. There are five possible different cyclotomic numbers in the case f is even; i.e., (0,0), (1,3)=(2,3)=(1,2), (1,1)=(0,3), (2,2)=(0,2), (3,3)=(0,1) and

$$\begin{aligned} &(0,0) = (N - 11 - 6x)/16, \\ &(0,1) = (N - 3 + 2x + 8y)/16, \\ &(0,2) = (N - 3 + 2x)/16, \\ &(0,3) = (N - 3 + 2x - 8y)/16, \\ &(1,2) = (N + 1 - 2x)/16. \end{aligned}$$

For the case f being odd, there are at most five distinct cyclotomic numbers, which are

$$\begin{aligned} &(0,0) = (2,2) = (2,0) = (N - 7 + 2x)/16, \\ &(0,1) = (1,3) = (3,2) = (N + 1 + 2x - 8y)/16, \\ &(1,2) = (0,3) = (3,1) = (N + 1 + 2x + 8y)/16, \\ &(0,2) = (N + 1 - 6x)/16, \\ &\text{the rest} = (N - 3 - 2x)/16. \end{aligned}$$

It follows from these cyclotomic formulae that

$$\frac{2x-7}{16}+\frac{3+2x}{16}=\frac{4x-4}{16}=\frac{x-1}{4}=\pm 1$$

if and only if $x = 5$ or -3. This completes the proof. □

Let $N = 8t + 1$. It is possible for the set of octic residues $D_0 = (\alpha^8)$ to form an almost difference set of Z_N, where α is a primitive root of N. Since $|D_0| = t$, a necessary condition for D_0 to be an almost difference set is $t(t-1) = (2\lambda+1)(N-1)/2$. It follows that $t = 8\lambda + 5$ and therefore

$$N = 8t + 1 = 64\lambda + 41 = 16(4\lambda + 2) + 9.$$

Under these necessary conditions the cyclotomic numbers of order 8 are given in two sets of formulae according to whether 2 is a quartic residue or not, in terms of N, x, y, a and b which are determined by

$$N = x^2 + 4y^2 = a^2 + 2b^2, \quad (x = a = 1 \bmod 4). \tag{4.1}$$

The cyclotomic constants of order 8 were calculated by Lehmer [42]. The eight cyclotomic numbers we need are the following:

$$\begin{aligned}
(1,1) = (3,3) = (5,5) = (7,7) &= (N-7+2x+4a)/64,\\
(2,2) = (6,6) &= (N-7-2x-8a)/64,\\
(0,0) = (4,4) &= (N-15-2x)/64.
\end{aligned}$$

By definitions of almost difference sets and cyclotomic numbers, D_0 is an almost difference set if and only if $(2,2) = (6,6) = (0,0) = (4,4)$ and $(1,1) = (0,0) \pm 1$. Obviously, $(0,0) = (2,2)$ if and only if $a = 1$. Under the condition $a = 1$, $(1,1) = (0,0) \pm 1$ if and only if $x = -19$ or 13. Since both of -19 and 13 are congruent to 1 modulo 4, they are solutions. Thus, we have proved the following conclusion.

Theorem 4.14 *Let $N = 8t + 1$ and $t = 8\lambda + 5$, where λ is a positive integer. Assume that 2 is a quartic residue modulo N. Then the set of octic residues D_0 forms an almost difference set if and only if N admits the simultaneous representations*

$$N = 19^2 + 4y^2 = 1 + 2b^2$$

or

$$N = 13^2 + 4y^2 = 1 + 2b^2.$$

For the case 2 is not a quartic residue we have the following result.

Theorem 4.15 *Let $N = 8t + 1$ and $t = 8\lambda + 5$, where λ is a positive integer such that 2 is not a quartic residue. Then the set of octic residues D_0 forms an almost difference set if and only if $N = 41$.*

Proof: The proof is similar to that of the foregoing theorem. We still assume the partition of (4.1). The eight cyclotomic numbers we need are the following [42]:

$$\begin{aligned}
(0,0) &= (4,4) = (N-15-10x-8a)/64, \\
(1,1) &= (5,5) = (N-7+2x+4a-16y)/64, \\
(2,2) &= (6,6) = (N-7+6x)/64 \\
(3,3) &= (7,7) = (N-7+2x+4a+16y)/64.
\end{aligned}$$

Since the equations $(1,1) = (3,3)$ and $(0,0) = (2,2)$ have no solution, there are only two sets of conditions in which the set D_0 can form an almost difference set. They are given by

$$(1,1) = (0,0), \ (2,2) = (3,3), \ (1,1) = (2,2) \pm 1$$

and

$$(1,1) = (2,2), \ (0,0) = (3,3), \ (1,1) = (0,0) \pm 1.$$

Solving the two sets of equations gives the only solution $(x, y, a) = (5, 2, -3)$ and $(x, y, a) = (5, -2, -3)$ respectively. Thus, we get the only prime 41. It is easily checked that 2 is indeed not a quartic residue modulo 41. This proves the conclusion. □

So far we have found all the primes of the form $8t + 1$ such that the set of octic residues forms an almost difference set.

4.7 Almost Difference Sets and Autocorrelation Functions

The characteristic function of a set of Z_N and the autocorrelation function of a function from Z_N to Z_2 were defined in Section 4.2. It was also proven in Section 4.2 that the difference sets of Z_N can be characterized by the autocorrelation function of their characteristic functions. Let D be an (N, k, λ) almost difference set of Z_N, f_D its characteristic function and s_D^∞ its periodic characteristic sequence. Similarly, we have the following result.

Theorem 4.16 *Let D be an (N, k, λ) almost difference set of Z_N. Then*

$$\begin{aligned}
\mathrm{AC}_{f_D} &= \mathrm{AC}_{s_D} \\
&= \begin{cases} 1, \ w = 0; \\ 1 - 4(k-\lambda)/N; \ \textit{for half of the nonzero elements of } \ Z_N^*; \\ 1 - 4(k-\lambda-1)/N; \ \textit{for the other half.} \end{cases}
\end{aligned}$$

Proof: It follows from the definition of autocorrelation function that

$$\begin{aligned} N \cdot \mathrm{AC}_{f_D}(w) = N \cdot \mathrm{AC}_{s_D}(w) &= (\sum_{x \in D} + \sum_{x \in D^*})(-1)^{f(x)+f(x+w)} \\ &= |(-w+D) \cap D| - |(-w+D) \cap D^*| \\ &\quad +|(-w+D^*) \cap D^*| - |(-w+D^*) \cap D|, \end{aligned}$$

where $D^* = Z_N \setminus D$. Denoting $|(w+D) \cap D| = d(w)$ for $w \neq 0$, we obtain

$$\begin{aligned} &|(-w+D) \cap D^*| = k - d(w), \\ &|(-w+D^*) \cap D| = k - d(w), \\ &|(-w+D^*) \cap D^*| = N - 2k + d(w). \end{aligned}$$

On the other hand, we have $|(w+S) \cap S| = |(-w+S) \cap S|$ for any subset of Z_N. Combining the above results, we get

$$\mathrm{AC}_{f_D}(w) = \mathrm{AC}_{s_D}(w) = [N - 4(k - d(w))]/N.$$

Note $d(w) = \lambda$ for half of the elements of Z_N^* and $d(w) = \lambda + 1$ for the other half, we know that the conclusions of the theorem are true. □

This theorem shows that there is little difference between the autocorrelation property of the characteristic functions of difference sets and that of almost difference sets. It shows also that the difference property of the characteristic sequences of almost difference sets is almost the same as that of difference sets. These facts indicate to some extent the cryptographic significance of the almost difference sets.

4.8 Almost Difference Sets, Nonlinearity and Approximation

In Section 4.3 it has been proven that the characteristic functions of some difference sets are good cryptographic functions with best nonlinearity with respect to the additions of Z_N and Z_2. For the characteristic function of some almost difference sets, we have similar results.

Theorem 4.17 *Let D be an (N, k, λ) almost difference set of Z_N, and $f(x)$ be the characteristic function of D. Then*

$$p_f(\alpha, \beta) = \begin{cases} \left.\begin{cases} \frac{N-2(k-\lambda)}{N}, & \textit{for half of } 0 \neq \alpha\textit{'s}; \\ \frac{N-2(k-\lambda-1)}{N}, & \textit{for the other half.} \end{cases}\right\} \beta = 0; \\ \left.\begin{cases} \frac{2(k-\lambda)}{N}; & \textit{for half of } 0 \neq \alpha\textit{'s}; \\ \frac{2(k-\lambda-1)}{N}; & \textit{for the other half.} \end{cases}\right\} \beta = 1, \end{cases}$$

where $p_f(\alpha, \beta) = p(f(x) - f(y) = \beta | x - y = \alpha)$.

This theorem shows that the characteristic functions of the almost difference sets with parameters (N, k, λ) have good nonlinearity with respect to the additions of Z_N and Z_2, provided that $|N/4 - k + \lambda|$ is small enough.

Since there are only almost difference sets in Z_N with N being odd, and there are only trivial affine functions from Z_N to Z_2 for odd N, there are no affine approximation problems for the characteristic functions of almost difference sets with respect to the additions of the two rings Z_N and Z_2. However, the characteristic functions of almost difference sets can be regarded as functions from $GF(2)^m$ to $GF(2)$, where $m = \lceil \log_2 N \rceil$. The affine approximation problem of them remains therefore to be investigated.

Chapter 5

Cyclotomic Generator of Order $2k$

The relation between cyclotomic numbers and sequences is considered first in this chapter. Then the binary cyclotomic generator of order $2k$ is described, and the linear and sphere complexity of the output sequences of the generator are controlled. As an example, some security aspects of the difference set characterized generator are analyzed. Cryptographic properties of the sum of difference set characterized sequences are also investigated. It is shown that cyclotomic generators of order 2 are cryptographically attractive from a number of points of views.

The main contributions of this chapter are the following:

1. The discovery of connections between cyclotomy and cryptographic sequences, as shown by results of Sections 5.1 and 5.3;

2. the linear and sphere complexity analysis of the cyclotomic generators of order $2k$ described by Theorems 5.1, 5.2, 5.3, 5.4, 5.5, and 5.12;

3. the correlation analysis (Section 5.3.1), decimation analysis (Section 5.3.2), linear complexity analysis (Section 5.3.3) of cyclotomic sequences of order 2;

4. the security analysis of the cyclotomic generator of order 2 against a decision tree attack, including one of the main results Proposition 5.16 which is superior to known ones since we have generalized and made use of the Davenport reduction theorem (Section 5.3.4); and

5. the analysis of the sum of difference set characterized sequences including all results of this section.

The materials of Sections 5.1 and 5.2 are based on my papers [23, 24], and those of Sections 5.3 and 5.4 were obtained recently.

5.1 Cyclotomic Constants and Sequences

The motivation of the investigation of cyclotomic numbers is related to the outstanding Waring problem, difference sets, and the solution of equations over finite fields [13, 23, 56]. Cyclotomic numbers invented already by Gauss [34] turn out to be quite valuable in the design and analysis of some keystream generators.

Let $N = df+1$, be a odd prime and let θ be a fixed primitive element of Z_N. Denoting the multiplicative subgroup (θ^d) as D_0, we see that the coset decomposition of Z_N^* with respect to the subgroup D_0 is in this case

$$Z_N^* = \cup_{i=0}^{d-1} D_i,$$

where $D_i = \theta^i D_0$ for $0 \le i \le d-1$. The coset D_l is called the *index class* l [3] or *cyclotomic class* l [66]. Let $(l, m)_d$ denote the number of solutions (x, y) of the equation

$$1 = x - y, \quad (x, y) \in D_l \times D_m,$$

which were called *cyclotomic numbers* of order d [2, 3, 14, 42], or equivalently,

$$(l, m)_d = |(D_l + 1) \cap D_m|.$$

Apparently, there are at most d^2 distinct cyclotomic numbers of order d and these numbers depend not only on N, d, l, m, but also on which of the $\phi(N-1)$ primitive elements of Z_N is chosen.

The following are some elementary facts which are very important to our cryptographic applications, because they indicate several kinds of invariances between the cyclotomic numbers. They are the theoretical bases for the necessity of keeping the stability of local nonlinearities of some cryptographic functions. Proofs of these elementary facts can be found in [14, 3].

$$(A) \quad (l, m)_d = (d-l, m-l)_d = \begin{cases} (m, l)_d, & f \text{ even} \\ (m + d/2, l + d/2)_d, & f \text{ odd} \end{cases}$$

$$(B) \quad \sum_{m=0}^{d-1} (l, m)_d = f - n_l,$$

where

$$n_l = \begin{cases} 1, & l = 0 \bmod d, \ f \text{ even} \\ 1, & l = d/2 \bmod d, \ f \text{ odd} \\ 0, & \text{otherwise} \end{cases}$$

$$(C)\quad \sum_{l=0}^{d-1}(l,m)_d = f - k_m,$$

where

$$k_m = \begin{cases} 1, & \text{if } m = 0; \\ 0, & \text{otherwise} \end{cases}$$

$$(D)\quad \sum_{l=0}^{d-1}\sum_{m=0}^{d-1}(l,m)_d = df - 1 = N - 2.$$

$$(E)\quad (l,m)_{d'} = (sl, sm)_d,$$

where $(l,m)_{d'}$ is based on the primitive root $\theta' = \theta^s \bmod N$; necessarily then s is prime to $N-1$.

The meaning of the cyclotomic numbers can be considered from another viewpoint. From the definition we know the set $\{(l,m)_d : l = 0, 1, \cdots, d-1\}$ represents how the set $D_m + 1$ is distributed among the cyclotomic classes. Note that

$$|D_l \cap (D_m + \theta^k)| = |D_{l+N-1-k} \cap (D_{m+N-1-k} + 1)|$$

for each k, we see that the d sets of numbers $\{(l,m)_d : l = 0, 1, \cdots, d-1\}$ for $m = 0, 1, \cdots, d-1$, represent also the distribution of the elements of any set $D_m + w$ with $w \neq 0$.

As seen above, cyclotomic numbers represent in fact the difference property of the partition $\{D_0, D_1, \cdots, D_{d-1}\}$ of Z_N'. So they should have connections with difference sets. Actually, the investigation of residue difference sets is the main motivation of the calculation of cyclotomic numbers [2, 66].

Let the notation be the same as before. We now construct cryptographic functions from Z_N to an abelian group $(G, +)$ of d elements, where $G = \{g_0 = 0, g_1, \cdots, g_{d-1}\}$. Let

$$C_0 = D_0 \cup \{0\}, \quad C_i = D_i, \; i = 1, \cdots, d-1.$$

Without concerning the implementation problem, we define a function from Z_N to $(G, +)$ as: $f(x) = g_i$ iff $x \in C_i$.

If $i \cdot j \neq 0$, then we have

$$d_f(g_i, g_j; \theta^k) = (i + N - 1 - k, j + N - 1 - k)_d.$$

On the other hand, we have

$$d_f(g_0, g_0; \theta^k) = |(D_{N-1-k} \cup \{0\}) \cap (D_{N-1-k} \cup \{0\} + 1)|.$$

It follows that

$$0 \leq d_f(g_0, g_0; \theta^k) - (N-1-k, N-1-k)_d \leq 2.$$

Similarly, we have

$$0 \leq d_f(g_0, g_1; \theta^k) - (N-1-k, N-k)_d \leq 1.$$

and

$$0 \leq d_f(g_1, g_0; \theta^k) - (N-k, N-1-k)_d \leq 1.$$

Thus, we arrive at the conclusion that the difference parameters are almost the same as the cyclotomic numbers.

5.2 The Generator

Binary sequences with prime period are cryptographically attractive due to the following theorems about the linear and sphere complexity, which follow easily from Basic Theorem 3.5. They can be used to control the linear and sphere complexity of not only those cyclotomic sequences, but also many other sequences.

Theorem 5.1 *If N is prime, then for any nonconstant sequence s^∞ of period N over $GF(2)$ and over $GF(2^s \bmod N)$ with $\gcd(s, N-1) = 1$ and with $2^s \bmod N$ being a power of a prime,*

1. $\mathrm{L}(s^\infty) \geq \mathrm{ord}_N(2)$;
2. $\mathrm{SC}_k(s^\infty) = \begin{cases} \geq \mathrm{ord}_N(2), & \text{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}; \\ 0, & \text{otherwise.} \end{cases}$

Proof: Note that $\gcd(s, N-1) = 1$, we have $\mathrm{ord}_N(2^s) = \mathrm{ord}_N(2)$. Setting $t = 1$ and $e_1 = 1$ in Basic Theorem 3.5 proves the theorem. □

Theorem 5.2 *If $N = 4t + 1$ is prime and t is a odd prime, then for any nonconstant sequence s^∞ of period N over $GF(2)$ and over $GF(2^s \bmod N)$ with $\gcd(s, N-1) = 1$,*

1. $\mathrm{L}(s^\infty) = N$ *or* $N-1$;
2. $\mathrm{SC}_k(s^\infty) = \begin{cases} N \text{ or } N-1, & \text{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}; \\ 0, & \text{otherwise.} \end{cases}$

Proof: Recall that a is a primitive root modulo an integer N if and only if $\mathrm{ord}_N(a) = \phi(N)$, where $\phi(x)$ is the Euler function. Since both $N = 4t+1$ and t are primes, it is not hard to prove that 2 is a primitive root of N. Then the conclusion of this theorem follows from Theorem 5.1. □

Theorem 5.3 *Let $N = 4t - 1$ be a prime with t odd. If $(N-1)/2$ is prime (i.e., it is a Sophie Germain prime), then for any nonconstant sequence s^∞ of period N over $GF(2)$ and over $GF(2^s \bmod N)$ with $\gcd(s, N-1) = 1$,*

1. $\mathrm{L}(s^\infty) = N$ *or* $N - 1$;
2. $\mathrm{SC}_k(s^\infty) = \begin{cases} N \textit{ or } N-1, & \textit{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}; \\ 0, & \textit{otherwise.} \end{cases}$

Proof: By the speciality of the prime N it is easy to see that 2 is a primitive root modulo N. Then the conclusion of this theorem follows from Theorem 5.1. □

Theorem 5.4 *Let $N = 4t + 1$ be a prime with t odd and $t = t_1 t_2$, where t_1 and t_2 are primes. If*

$$2^{t_1 t_2} \neq \pm \bmod N,$$

then for any nonconstant sequence s^∞ of period N over $GF(2^s \bmod N)$ (especially over $GF(2)$) with $\gcd(s, N-1) = 1$,

1. $\mathrm{L}(s^\infty) = N$ *or* $N - 1$;
2. $\mathrm{SC}_k(s^\infty) = \begin{cases} N \textit{ or } N-1, & \textit{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}; \\ 0, & \textit{otherwise.} \end{cases}$

Proof: Note that $\mathrm{ord}_N(2)$ divides $\phi(N) = N - 1 = 4t_1t_2$ and that t_1 and t_2 are primes. It then follows by the two inequalities in the assumptions of the theorem that the order of 2 modulo N must be equal to $N - 1$. Combining this fact and Theorem 5.1 proves the theorem. □

Theorem 5.5 *Let $N = 4t - 1$ be a prime with t odd and $2t - 1 = t_1 t_2$, where t_1 and t_2 are primes. If*

$$2^{t_1} \neq -1 \bmod N, \quad 2^{t_2} \neq -1 \bmod N, \; 2^{t_1 t_2} \neq 1 \bmod N,$$

then for any nonzero sequence s^∞ of period N over $GF(2)$ and over $GF(2^s \bmod N)$ with $\gcd(s, N-1) = 1$,

1. $\mathrm{L}(s^\infty) = N$ *or* $N - 1$;

2. $\mathrm{SC}_k(s^\infty) = \begin{cases} N \textit{ or } N-1, & \textit{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}; \\ 0, & \textit{otherwise.} \end{cases}$

Proof: With the same arguments as in Theorem 5.4 we see that the order of 2 modulo N is $N-1$. Combining this fact and Theorem 5.1 yields the conclusion of this Theorem. □

After having proved the above theorems, which are needed to control the linear and sphere complexity of the cyclotomic sequences of order $2k$, we now describe the binary cyclotomic generator of order $2k$. Let prime $N = 2kf + 1$, and D_0, D_1, $\cdots$, D_{2k-1} be the cyclotomic classes of order $2k$ defined as before. If we choose the mapping

$$H(x) = (x^{(N-1)/2k} \bmod N) \bmod 2$$

as the cryptographic function for the natural sequence generator of Figure 3.1.b, then we have the *binary cyclotomic generator of order* $2k$.

It is not difficult to verify that each difference parameter of the above cryptographic function $H(x)$ for the binary cyclotomic generator of order $2k$ can be expressed as the sum of k^2 cyclotomic numbers of order $2k$. Thus, if the cyclotomic numbers of order $2k$ have an ideal stability, the formulae in Section 3.3 show that an ideal property for the seven aspects described in Section 3.3 is guaranteed if k is small enough. Even if the cyclotomic numbers of order $2k$ are stable to an ideal extent, the size of the k may lead to a relatively bad difference property of the cryptographic function. Thus, only those generators derived from small k are cryptographically attractive.

Theorems 5.1–5.5 clearly show how to control the linear complexity and its stability for the output sequence of the binary cyclotomic generator of order $2k$. By Theorem 5.1, to control the linear complexity and its stability of the output sequences of the binary cyclotomic generators, it suffices to choose the prime N such that $\mathrm{ord}_N(2)$ is large enough. The best choices for the primes are Sophie Germain primes, and Stern primes, i.e., primes $p = 4t + 1$ with t also prime.

If $k = 1$, then it is called the *binary cyclotomic generator* of order 2. The binary cyclotomic generators of order 2 can be further classified into difference set characterized and almost difference set characterized generators which corresponds to the cases $N = -1 \bmod 4$ and $N = 1 \bmod 4$, respectively [23]. In the case $k = 1$ the output sequence of the difference set characterized and almost difference set characterized generators are the 0-1 version of the Legendre sequences with a slight modification of the values for $\left(\frac{pi}{p}\right)$ for $i = 0, 1, \cdots$, and a proper choice of the prime p.

Hence, we can state the main results of this section. A difference set characterized generator with a Sophie Germain prime has the following cryptographic attributes: its output sequences have maximum linear complexity by Theorem 5.3; best autocorrelation property by the formulae for cyclotomic numbers of order 2 or the difference set property of the set $f^{-1}(1)$ and formula (3.5), where $f(x) = (x^{(N-1)/2} \bmod N) \bmod 2$; best linear complexity stability by Theorem 5.3; the cryptographic function $f(x)$ has best nonlinearity with respect to the additions of Z_N and Z_2 by the difference set property of $f^{-1}(1)$ and formula (3.3). The almost difference set characterized generator with a Stern prime has almost the same cryptographic attributes.

5.3 Analysis of Cyclotomic Generators of Order 2

Recall from last section that the cryptographic function $f(x)$ for both of the difference set characterized and almost difference set characterized generators is

$$f(x) = (x^{(N-1)/2} \bmod N) \bmod 2,$$

where N is a prime. If $N = 3 \bmod 4$, then the corresponding generator is the difference set characterized generator, otherwise it is the almost difference set characterized generator. We make such a distinction since the characteristic set of $f(x)$ is a difference set of Z_N if $N = 3 \bmod 4$, and an almost difference set otherwise. The nonlinearity of the above cryptographic function was analyzed in Sections 4.3 and 4.8. The autocorrelation analysis of the output sequences of difference set characterized and almost difference set characterized generators was done in Sections 4.2 and 4.7. We have also seen that the linear and affine approximations of the cryptographic function with respect to the addition of Z_N makes no sense. It is also clear that the differential attack does not work computationally for the cyclotomic generators of order 2, for the above cryptographic function $f(x)$ has optimal difference property with respect to the addition of Z_N. In this section we describe some other attributes precisely and do some security analysis for the cyclotomic generators of order 2. Our analysis for the cyclotomic generators of order 2 are roughly applicable to other cyclotomic generators.

5.3.1 Crosscorrelation Property

The output sequences of both difference set characterized and almost difference set characterized generators are the 0-1 version of modified Legendre sequences. Since Legendre sequences are based on either difference sets or almost difference sets, their autocorrelation property is optimal. This means that it is impossible to approximate a Legendre sequence with its delayed versions. Now a cryptographically interesting question is: What is the extent of correlation between two Legendre sequences of different prime periods?

Let s^∞ be a Legendre sequence[1] of prime period p defined by the modified Legendre symbol

$$s_i = \left(\frac{i}{p}\right)' = \begin{cases} 1, & \text{if } p|i, \\ \left(\frac{i}{p}\right), & \text{otherwise,} \end{cases}$$

where $\left(\frac{i}{p}\right)$ is the usual Legendre symbol that takes 1 if i is a quadratic residue and -1 if i is a quadratic nonresidue. Let t^∞ be another Legendre sequence of period $q \neq p$ defined by the same modified Legendre symbol. Then the correlation of the two sequences with respect to period pq is measured by

$$\begin{aligned} \mathrm{CC}_{s,t} &= \left[\sum_{i=0}^{pq-1} \left(\frac{i}{p}\right)' \left(\frac{i}{q}\right)'\right] / pq \\ &= \left[\sum_{i=0}^{pq-1} \left(\frac{i}{p}\right)\left(\frac{i}{q}\right) + \sum_{i=1}^{q-1} \left(\frac{p}{q}\right)\left(\frac{i}{q}\right) + \sum_{i=1}^{p-1} \left(\frac{q}{p}\right)\left(\frac{i}{p}\right) + 1\right] / pq. \end{aligned}$$

It is easy to prove that

$$\sum_{x=0}^{q-1} \left(\frac{ux+v}{q}\right) = q\left(\frac{v}{q}\right)\left[1 - \left(\frac{u}{q}\right)^2\right].$$

It follows that

$$\sum_{i=1}^{q-1} \left(\frac{i}{q}\right) = \sum_{i=1}^{p-1} \left(\frac{i}{p}\right) = 0,$$

and

$$\begin{aligned} \sum_{x=0}^{pq-1} \left(\frac{x}{p}\right)\left(\frac{x}{q}\right) &= \sum_{j=0}^{p-1}\sum_{i=0}^{q-1} \left(\frac{ip+j}{p}\right)\left(\frac{ip+j}{q}\right) \\ &= \sum_{j=0}^{p-1} \left(\frac{j}{p}\right) \sum_{i=0}^{q-1} \left(\frac{ip+j}{q}\right) \\ &= \sum_{j=0}^{p-1} \left(\frac{j}{p}\right) q \left(\frac{j}{q}\right)\left[1 - \left(\frac{p}{q}\right)^2\right] \\ &= 0. \end{aligned}$$

Hence, we have $C_{s,t} = 1/pq$. This result means that the crosscorrelation between two Legendre sequences defined by the modified Legendre symbol is almost minimal. Thus it is also impossible to approximate a Legendre sequence with another one.

[1] Here we use the $\{1, -1\}$-version of Legendre sequence. But we usually use the $\{0, 1\}$-version.

5.3.2 Decimation Property

The decimation of sequences is also cryptographically interesting. Let s^∞ be a Legendre sequence defined by the extended Legendre symbol. If a is a quadratic residue modulo the prime p then it is easy to see that $s_{ai}^\infty = s_i^\infty$; otherwise we have

$$s_{ai}^\infty + s_i^\infty = \underbrace{011\cdots1}\underbrace{011\cdots1}\cdots,$$

which is a sequence of period p whose first least period has the first bit being 0 and other bits being 1. Thus each decimation of a Legendre sequence is either the original one or almost the complement sequence of the original one. It follows that the multipliers of Legendre sequences are the quadratic residues. This fact also shows that most Legendre sequences based on Mersenne primes can not be maximum-length sequences.

5.3.3 Linear Complexity

The results about the linear and sphere complexity of sequences presented in Sections 3.5 and 5.2 are naturally true for Legendre sequences (we refer to the 0-1 version of Legendre sequences). But those results only depend on the special properties of the period. For Legendre sequences we have a much better result, as described by the following Theorem 5.6 [31].

Theorem 5.6 *Let s^∞ be the Legendre sequence of period p as before. Then*

1. *if $p = 8t - 1$ for some t, then $\mathrm{L}(s^\infty) = (p+1)/2$;*
2. *if $p = 8t + 1$ for some t, then $\mathrm{L}(s^\infty) = (p-1)/2$;*
3. *if $p = 8t + 3$ for some t, then $\mathrm{L}(s^\infty) = p$; and*
4. *if $p = 8t + 5$ for some t, then $\mathrm{L}(s^\infty) = p - 1$.*

Proof: For simplicity we use L_p to denote the linear complexity of the Legendre sequence with period p. Let Q denote the set of quadratic residues q with $0 < q \leq p-1$ and N the set of quadratic nonresidues modulo p. Define

$$S^p(x) = \sum_{i \in Q} x^i$$

and let β be a pth root of unity over the field $GF(2^m)$ that is the splitting field of $x^p - 1$. Then by Lemma 6.14 we have

$$\begin{aligned} L_p &= \deg[(x^p - 1)/\gcd(x^p - 1, S^p(x))] \\ &= p - |\{j : S^p(\beta^j) = 0,\ 0 \leq j \leq p-1\}|. \end{aligned} \tag{5.1}$$

Before calculating the actual linear complexity, we have to mention the following basic facts:

B1: $(Q, \cdot)$ is a group with $|Q| = (p-1)/2$ and $q \cdot N = N$ for any $q \in Q$, where "$\cdot$" denotes integer multiplication modulo p.

B2: $S^p(\beta^q) = S^p(\beta)$ for any $q \in Q$ and $S^p(\beta^n) = 1 + S^p(\beta)$ for any $n \in N$.

B3: Let the symbols be as before. Then

$$S^p(\beta) \in \{0, 1\} \text{ iff } (S^p(\beta))^2 = S^p(\beta) \text{ iff } 2 \in Q.$$

B4: $2 \in Q$ if and if $p = 8t \pm 1$ for some t.

Basic fact B1 is straightforward. Since $(Q, \cdot)$ is a group, we have $qQ = Q$ and $q^{-1} \in Q$ for any $q \in Q$. Hence,

$$S^p(\beta^q) = \sum_{x \in Q} \beta^{qx} = \sum_{y \in Q} \beta^y = S^p(\beta).$$

Because $n^{-1} \in N$ for any $n \in N$ and $nQ = N$, we have

$$S^p(\beta^n) = \sum_{x \in Q} \beta^{nx} = \sum_{y \in N} \beta^y = 1 + S^p(\beta).$$

This completes the proof of B2. Note that the characteristic of the field $GF(2^m)$ is 2, the first part of B3 is obvious. Note that $(S^p(\beta))^2 = S^p(\beta^2)$, it then follows from B2 that $(S^p(\beta))^2 = S^p(\beta)$ if and only if $2 \in Q$. B4 is proved by the Law of Quadratic Reciprocity.

Thus, we have completed the proofs of the above basic facts that will be used frequently in the sequel.

The proof of this theorem is then completed by considering two cases depending on whether 2 is a quadratic residue.

We first consider the case that $2 \in Q$, which happens if and only if $p = 1$ or $7 \bmod 8$ by basic fact B4. It follows from basic fact B3 that in this case we have $S^p(\beta) \in \{0, 1\}$ and either $S^p(\beta^q) = 0$ for all $q \in Q$ or $S^p(\beta^n) = 0$ for all $n \in N$. Since $S^p(1) = [(p-1)/2 \bmod 2]$, it follows that $S^p(1) = 0$ if $p = 1 \bmod 8$ and $S^p(1) = 1 \neq 0$ if $p = 7 \bmod 8$. Hence if $p = 1 \bmod 8$ then by (5.1)

$$\begin{aligned} L_p &= p - |\{j : S^p(\beta^j) = 0,\ 0 \le j \le p-1\}| \\ &= p - (p-1)/2 - 1 = (p-1)/2 \end{aligned}$$

and if $p = 7 \bmod 8$ then by (5.1)

$$L_p = p - (p-1)/2 = (p+1)/2.$$

Finally, we consider the case $2 \notin Q$, which happens if and only if $p = 3$ or $5 \bmod 8$ by basic fact B4. By basic fact B3 we have $S^p(\beta) \notin \{0, 1\}$. Since $-1 = 1$, it follows from basic fact B2 that $S^p(\beta^j) \neq 0$ for all j with $0 < j \leq p-1$. Since $S^p(1) = [(p-1)/2 \bmod 2]$, it follows that $S(1) = 1$ if $p = 3 \bmod 8$ and $S^p(1) = 0$ if $p = 5 \bmod 8$. Thus, if $p = 3 \bmod 8$ then by (5.1)

$$L_p = p - |\{j : S^p(\beta^j) = 0,\ 0 \leq j \leq p-1\}| = p$$

and if $p = 5 \bmod 8$ then by (5.1)

$$L_p = p - |\{j : S^p(\beta^j) = 0,\ 0 \leq j \leq p-1\}| = p - 1.$$

Hence, we have completed the proof of this theorem. □

Now we determine the feedback polynomials of Legendre sequences.

In the case that $2 \in Q$, let β be a primitive pth root over $GF(2^m)$ as before. Since $S^p(\beta)^2 = S^p(\beta)$, we have $S^p(\beta) = 0$ or 1. By basic fact B3 we can choose the primitive root β such that $S^p(\beta) = 1$. Because $2 \in Q$, the set $Q = 2Q$. Thus by basic fact B3

$$q(x) = \prod_{q \in Q} (x - \beta^q) \text{ and } n(x) = \prod_{n \in N} (x - \beta^n)$$

have coefficients from $GF(2)$. The polynomial $q(x)$ and $p(x)$ depends on the choice of β, but there are only two possibilities. In the sequel, we shall fix the β as above, i.e. $S^p(\beta) = 1$.

Theorem 5.7 *Let s^∞ be the Legendre sequence of period p as before and $m(x)$ its feedback polynomial. Then*

1. *if $p = 8t - 1$ for some t, then $m(x) = (x-1)q(x)$;*
2. *if $p = 8t + 1$ for some t, then $m(x) = q(x)$;*
3. *if $p = 8t + 3$ for some t, then $m(x) = x^p - 1$; and*
4. *if $p = 8t + 5$ for some t, then $m(x) = (x^p - 1)/(x - 1)$.*

Proof: Recall the proof of Theorem 5.6 in the previous section. We consider the first case that $2 \in Q$, which is equivalent to $p = 8t \pm 1$ for some t. As pointed out before, the polynomials $q(x)$ and $n(x)$ have coefficients from $GF(2)$. Obviously, we have

$$x^p - 1 = (x-1)q(x)n(x).$$

Recall also that $2 \in Q$ is equivalent to $p = 1$ or $7 \bmod 8$. In the case that $p = 7 \bmod 8$, the proof of Theorem 1 have shown that $S^p(1) = 1$. By our choice of

β and the proof of Theorem 1 $S^p(\beta^q) = S(\beta) = 1 \neq 0$ for all $q \in Q$ and therefore $S^p(\beta^n) = 0$ for all $n \in N$. We obtain that

$$\gcd(x^p - 1, S^p(x)) = n(x).$$

Hence, by (6.2) we have

$$m(x) = \frac{x^p - 1}{\gcd(x^p - 1, S^p(x)} = (x-1)q(x).$$

Similarly, if $p = 1 \bmod 8$, the proof of Theorem 1 have shown that $S^p(1) = 0$. And by our choice of β we get $S^p(\beta^q) = 1$ for all $q \in Q$ and $S^p(\beta^n) = 0$ for all $n \in N$. It follows that

$$\gcd(x^p - 1, S^p(x)) = (x-1)n(x).$$

Hence, by (6.2) we have

$$m(x) = \frac{x^p - 1}{\gcd(x^p - 1, S^p(x))} = q(x).$$

Now we consider the case $2 \notin Q$. The proof of Theorem 1 has shown that in this case we have

$$S^p(\beta^j) \neq 0 \text{ for all } 0 < j \leq p-1$$

and also $S^p(1) = 1$ if $p = 3 \bmod 8$, and $S^p(1) = 0$ if $p = 5 \bmod 8$. It follows that

$$\gcd(x^p - 1, S^p(x)) = \begin{cases} 1, & p = 3 \bmod 8 \\ x-1, & p = 5 \bmod 8. \end{cases}$$

Hence, by (6.2)

$$m(x) = \frac{x^p - 1}{\gcd(x^p - 1, S^p(x))} = \begin{cases} x^p - 1, & p = 3 \bmod 8 \\ \frac{x^p-1}{x-1}, & p = 5 \bmod 8. \end{cases}$$

□

Theorem 5.6 shows that the linear complexity of Legendre sequences is quite ideal. It is possible that the local linear complexity and its stability of many Legendre sequences are also ideal. But a rigorous proof seems to be open.

5.3.4 Security against A Decision Tree Attack

In this section we analyze the security of the difference set characterized generator with respect to a decision-tree based key-recovering attack. With a little modification the same analysis can be carried out for the almost difference set characterized generator.

Assume that a cryptanalyst knows the algorithm, and a number of successive plaintext-ciphertext bit pairs, then he/she has got a number of successive keystream bits $z_j z_{j+1} \cdots z_{j+t-1} = z^t$. The purpose of key-determining attacks is to recover the key or an equivalent key by making use of the information about the key, which is implied in the obtained ciphertext or known plaintext-ciphertext pairs. Under the assumption that the cryptographic algorithm is known to the enemy, it is cryptographically beneficial to require that the mutual information

$$I(k = a; c_1 c_2 \cdots c_n = b_1 b_2 \cdots b_n)$$

be approximately the same for as many a's in the key space K as possible, where $b_1 \cdots b_n$ is a given piece of ciphertext. At least the average mutual information

$$I(k = a; c_1 c_2 \cdots c_n)$$

for fixed n should be about the same for as many a's in K as possible. Similar requirements for the mutual information between key and known plaintext-ciphertext pairs are also imperative. We now introduce a decision-tree based attack which is based on the analysis of the mutual information between the key and the keystream sequence.

Decision-tree based attacks determine the present key i by making use of the information about i contained in the keystream sequence z^t. For the difference set characterized generator all the keys in Z_N are equally likely. We have already analyzed the relation between the mutual information stability and the difference property of the cryptographic function in Section 4.4, where it has been shown that the mutual information between the key and any two-bit pattern in the output sequences of the difference set characterized generator is stable. Hence the two pattern distributions of the difference set characterized sequences are optimal.

The stability analysis of the mutual information

$$I(i; z_i z_{i+1} \cdots z_{i+t-1} = i_0 i_1 \cdots i_{t-1})$$

is not easy when $t \geq 3$. To carry out the stability analysis, we need to develop some bounds on the number of occurrences of patterns of quadratic residues and nonresidues.

Let D_1 be the set of quadratic residues modulo N, $D_0 = Z_N \backslash D_1$, $r_0, r_1, \cdots, r_{s-1}$ be s pairwise distinct elements of Z_N, and

$$D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = \cap_{k=0}^{s-1} (D_{i_k} + r_k),$$

$$d_{i_0\cdots i_{s-1}}(r_0,\cdots,r_{s-1}) = |\cap_{k=0}^{s-1}(D_{i_k}+r_k)|,$$

where $i_0,\cdots,i_{s-1}\in Z_2$. Then $\{D_{i_0\cdots i_{s-1}}(r_0,\cdots,r_{s-1}) : i_0,\cdots,i_{s-1}\in Z_2\}$ forms a partition of Z_N.

Generally, the mutual information $I(i;z_iz_{i+1}\cdots z_{i+t-1}=i_0i_1\cdots i_{t-1})$ depends on $|\cap_{j=0}^{t-1}(D_{i_j}-j)|$. Thus, the analysis of the mutual information is equivalent to that of the parameters $d_{i_0\cdots i_{t-1}}(r_0,\cdots,r_{t-1})$. This is also equivalent to the distribution analysis of patterns of length t in the difference set characterized sequences.

To ensure ideal pattern distributions of patterns of length s in the difference set characterized sequences and to ensuring a stable mutual information $I(i;z_iz_{i+1}\cdots z_{i+t-1}=i_0i_1\cdots i_{t-1})$, it is necessary and sufficient to require

$$d_{i_0\cdots i_{t-1}}(r_0,\cdots,r_{t-1}) \approx \frac{N}{2^t}$$

for all pairwise distinct nonzero elements $r_0,r_1,\cdots,r_{t-1}$ of Z_N.

The analysis for the case $t=1$ and 2 shows that this condition is satisfied in these two cases. The calculation of $d_{i_0\cdots i_{t-1}}(r_0,\cdots,r_{t-1})$ becomes quite difficult for $t\geq 3$. We now derive some bounds on the parameters. Before doing so, let us observe some relations among the parameters.

There are many relations among these parameters which can be described by

$$\sum_{i_{j_1},\cdots,i_{j_u}} d_{i_0\cdots i_{s-1}}(r_0,\cdots,r_{s-1}) = d_{i_{j'_1}\cdots i_{j'_{s-u}}}(r_{i_{j'_1}},\cdots,r_{i_{j'_{s-u}}}),$$

where

$$\{j'_1\cdots j'_{s-u}\} = \{0,1,\cdots,s-1\}\setminus\{j_1,\cdots,j_u\}.$$

These equations describe the conservation between these parameters.

Basing on the Jacobsthal formulae [39, 62]

$$\sum_{x=0}^{N-1}\left(\frac{ax+b}{N}\right) = N\left(\frac{a}{N}\right)\left[1-\left(\frac{a}{N}\right)^2\right],$$
$$\sum_{x=0}^{N-1}\left(\frac{ax^2+bx+c}{N}\right) = \begin{cases} (N-1)\left(\frac{a}{N}\right), \text{ if } b^2-4ac=0 \bmod N \\ -\left(\frac{a}{N}\right), \text{ if } b^2-4ac\neq 0 \bmod N, \end{cases}$$

where $(\frac{x}{N})$ denotes the Legendre symbol, we can simply derive the following results:

Proposition 5.8 *If* $N=3 \bmod 4$, *then we have*

$$d_{ij}(r_0,r_1) = \begin{cases} (N-3)/4, & \textit{for } (i,j)=(1,1), r_0\neq r_1\in Z_N \\ (N+1)/4, & (i,j)\neq(1,1), r_0\neq r_1\in Z_N \end{cases}$$

If $N = 1 \bmod 4$, *then*

$$d_{11}(r_0, r_1) = \begin{cases} (N-5)/4, & \text{for } (N-1)/2 \text{ nonzero } r\text{'s of } Z_N \\ (N-1)/4, & \text{for the other half nonzero elements} \end{cases}$$

$$d_{10}(r_0, r_1) = \begin{cases} (N+3)/4, & \text{for } (N-1)/2 \text{ nonzero } r\text{'s of } Z_N \\ (N-1)/4, & \text{for the other half nonzero elements} \end{cases}$$

$$d_{01}(r_0, r_1) = \begin{cases} (N+3)/4, & \text{for } (N-1)/2 \text{ nonzero } r\text{'s of } Z_N \\ (N-1)/4, & \text{for the other half nonzero elements} \end{cases}$$

$$d_{00}(r_0, r_1) = \begin{cases} (N-5)/4, & \text{for } (N-1)/2 \text{ nonzero } r\text{'s of } Z_N \\ (N-1)/4, & \text{for the other half nonzero elements} \end{cases}$$

where $r = r_1 - r_0$*; and in both cases, we have*

$$d_1 = |D_1| = (N-1)/2, d_0 = |D_0| = (N+1)/2.$$

It is easy to see that the study of the parameters $d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1})$ of the quadratic residue partition is actually that of the number of occurrences of patterns of quadratic residues and nonresidues. In the literature much attention has been paid to the least quadratic nonresidue of a prime N, which is similar to the investigation of the least primitive root of a prime N. Both of these investigations have cryptographic values. However, cryptographically much more important is the analysis of the number of occurrences of arbitrary patterns of quadratic residues and nonresidues, i.e., the calculation of the parameters $d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1})$. The cases for $s = 1, 2$ have already been solved, as described above. However, they are very difficult to calculate for $s \geq 3$. According to Davenport [12], in the case $N = -1 \bmod 4$, Jacobsthal proved

$$d_{000}(r, r+1, r+2),\ d_{111}(r, r+1, r+2) = \frac{N}{2^3} + O(1),$$

as $N \to \infty$; and in the case $N = 1 \bmod 4$, it holds

$$d_{000}(r, r+1, r+2),\ d_{111}(r, r+1, r+2) = \frac{N}{2^3} + O(\sqrt{N}).$$

Hasse also gave tight bounds on $d_{i_0 i_1 i_2}(r_0, r_1, r_2)$ [36, pp. 165-167]. Some bounds about $d_{0000}(r, r+1, r+2, r+3)$ and $d_{1111}(r, r+1, r+2, r+3)$ were given by Dörg [32] and Hopf [37], which are of little value for our cryptographic purpose due to the fact that they are not tight at all. Davenport [12] developed upper bounds of the form $N/2^t + N^{c_t}$ on both $d_{00\cdots 0}(r, r+1, \cdots, r+t-1)$ and $d_{11\cdots 1}(r, r+1, \cdots, r+t-1)$ for t with $4 \leq t \leq 9$. For each of these bounds, $c_t \geq \frac{2}{3}$. These bounds are also too loose to have cryptographic values. The case $s = 4$ was also discussed by Hudson [38]. Certain patterns of values of totally multiplicative functions were investigated by

Walum [67]. An important development in the topic is a general lower and upper bound $\frac{N}{2^t} \pm t(3+\sqrt{N})$ on $d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})$ developed by Peralta [53], which will be referred to as Peralta bounds. Peralta treated zero as a quadratic residue, while we take it as a nonresidue. These two kinds of treatment result at most in a difference of t on $d_{i_0 \cdots i_t}(r_0, r_1, \cdots, r_{t-1})$. The Peralta bounds can be used to estimate the pattern distribution of the difference set characterized sequences and the above mutual information to some extent. However, they are still too loose. We now derive a bound on $d_{i_0 \cdots i_{t-1}}(r_0, r_1, \cdots, r_{t-1})$, which is superior to Peralta's.

To this end, we need the calculation and estimation of the sum

$$\phi_r(a_1, \cdots, a_r) = \sum_{x=0}^{N-1} \left(\frac{(x+a_1)\cdots(x+a_r)}{N} \right),$$

where $a_1, \cdots, a_r$ are pairwise distinct elements of Z_N. It is already clear for the cases $r = 1$ and 2. For $r \geq 3$ it seems that there is no known exact formulas. According to Hopf [37], Jacobsthal deduced the bound

$$\left| \sum_x \left(\frac{x(x^2+a)}{N} \right) \right| < 2\sqrt{N}$$

in his dissertation of 1906. According to Hudson [38], Weil deduced the following general bound in 1945 [61]:

$$|\phi_3(a,b,c)| = \left| \sum_{x=0}^{N-1} \left(\frac{(x+a)(x+b)(x+c)}{N} \right) \right| < 2\sqrt{N},$$

where a, b, c are distinct elements of Z_N.

The Peralta bounds are based mainly on the following estimate:

Proposition 5.9 *Let the symbols be as before. Then*

$$|\phi_r(a_1, \cdots, a_r)| < (r-1)\sqrt{N},$$

where the a_i's are pairwise distinct elements of Z_N.

This result is a special case of a more general result about multiplicative characters due to Weil (see Schmidt [61, Theorem 2C, p. 43]). Following Peralta [53], we call this inequality *Weil bound.*

$\phi_2(a_1, a_2)$ can be evaluated exactly, since it can be reduced to the calculation of $\phi_1(a)$. This fact was generalized into that $\phi_r(a_1, \cdots, a_r)$ reduces to $\phi_{r-1}(0, 1, b_1, \cdots, b_{r-3})$ if r is even, to $\phi_r(0, 1, b_1, \cdots, b_{r-2})$ if r is odd. The proof of the case $r = 4$ was given by Davenport. We now show that Davenport's proof for the case $r = 4$ is in fact suitable to the general case of r being even.

Proposition 5.10 *Let r be even, and $a_1, \cdots, a_r$ be r pairwise distinct elements of Z_N. Then*

$$\phi_r(a_1, a_2, \cdots, a_r) = -1 + \left(\frac{c}{N}\right) \phi_{r-1}(d_2, \cdots, d_r),$$

where

$$\begin{aligned} d_i &= \frac{(a_i - a_2)(a_1 - a_3)}{(a_i - a_1)(a_2 - a_3)}, \ i \geq 2, \\ c &= (a_1 - a_2)(a_1 - a_3) \prod_{i=2}^{r} \frac{(a_i - a_2)(a_1 - a_3)}{(a_i - a_1)(a_2 - a_3)}. \end{aligned}$$

Proof: The trick for such a reduction is the use of the transformation

$$x = \frac{uy + v}{gy + h},$$

with $uh - gv \neq 0 \bmod N$, which gives

$$\sum_{x \neq u/g} F(x) = \sum_{y \neq -h/g} F\left(\frac{uy + v}{gy + h}\right),$$

where $F(x)$ is a polynomial of $Z_N[x]$. As Davenport did, we choose now

$$\begin{aligned} u &= -a_1(a_2 - a_3), \quad & v &= -a_2(a_1 - a_3), \\ g &= a_2 - a_1, & h &= a_1 - a_3. \end{aligned}$$

Then applying this transformation to $\phi_r(a_1, \cdots, a_r)$ completes the proof. It can be easily proven that $d_2, \cdots, d_r$ are pairwise distinct. □

We will refer to this result as the *Davenport reduction theorem*, which will play an important role in developing the new bounds. To derive the bounds, we make use of the function

$$G(x) = \frac{1}{2^s} \prod_{k=0}^{s-1} \left[1 + (-1)^{i_k} \left(\frac{x - r_k}{N}\right)\right].$$

It is easy to see

$$\begin{aligned} & 2^s \sum_x G(x) \\ = & \sum_{t=3}^{s-1} \sum_{0 \leq k_1, \cdots, k_t \leq s-1} (-1)^{\sum_{j=1}^{t} i_{k_j}} \sum_x \prod_{j=1}^{t} \left(\frac{x - r_{k_j}}{N}\right) + N - \binom{s}{2}. \end{aligned}$$

Applying the Davenport reduction theorem and the Weil bound to this equation gives, in the case s being odd,

$$\left|2^s \sum_x G(x) - N + \binom{s}{2}\right| < \{\textstyle\sum_{i=0}^{s} \binom{s}{i} i - \sum_{i=0}^{(s-1)/2} \binom{s}{2i} - \sum_{i=0}^{s} \binom{s}{i} + 2\}\sqrt{N} - \frac{s-3}{2}.$$

Similarly, for the case s being even, we have

$$\left|2^s \sum_x G(x) - N + \binom{s}{2}\right| < \{\textstyle\sum_{i=0}^{s} \binom{s}{i} i - \sum_{i=0}^{(s)/2} \binom{s}{2i} - \sum_{i=0}^{s} \binom{s}{i} + 2\}\sqrt{N} - \frac{s-2}{2}.$$

By making use of the combinatorial formulas

$$\sum_{i=0}^{s/2} \binom{s}{2i} = 2^{s-1}, s \text{ even},$$

$$\sum_{i=0}^{(s-1)/2} \binom{s}{2i} = 2^{s-1}, s \text{ odd},$$

$$\sum_{i=0}^{s} \binom{s}{i} i = s2^{s-1},$$

we get in both cases

$$\left|2^s \sum_x G(x) - N + \binom{s}{2}\right| < \{(s-3)2^{s-1} + 2\}\sqrt{N} - \left\lfloor \frac{s-2}{2} \right\rfloor.$$

Now let $\mathrm{WH}(i_0 \cdots i_{s-1})$ denote the Hamming weight of vector $(i_0 \cdots i_{s-1})$, i.e., the number of ones in the binary vector. Assume that $\mathrm{WH}(i_0 \cdots i_{s-1}) = t$. Then there are at most t elements in $\{r_0, r_1, \cdots, r_s\}$ which belong to the set $D_{i_0 \cdots i_{s-1}}$, where $s \le t$. We now assume that there are exactly u such elements, say $r_0, r_1, \cdots, r_u$, $u \le t \le s$. Let $B = \{r_0, r_1, \cdots, r_u\}$ and $A = D_{i_0 \cdots i_{s-1}} \setminus B$. Then we have

$$\sum_{x \in A} G(x) = d_{i_0 \cdots i_{s-1}} - u,$$

$$\sum_{x \in B} G(x) = u/2.$$

It follows that

$$d_{i_0 \cdots i_{s-1}} - \frac{t}{2} \le \sum_{x \in D_{i_0 \cdots i_{s-1}}} G(x) \le d_{i_0 \cdots i_{s-1}}.$$

Hence, for each $(i_0 \cdots i_{s-1})$ we have

$$d_{i_0 \cdots i_{s-1}} - \frac{s}{2} \leq \sum_{x \in D_{i_0 \cdots i_{s-1}}} G(x) \leq d_{i_0 \cdots i_{s-1}}.$$

If the Hamming distance between the vector $(j_0 \cdots j_{s-1})$ and $(i_0 \cdots i_{s-1})$ is greater than or equal to two, it is not difficult to see that

$$\sum_{x \in D_{j_0 \cdots j_{s-1}}} G(x) = 0.$$

Furthermore, if the Hamming distance between the vector $(j_0 \cdots j_{s-1})$ and $(i_0 \cdots i_{s-1})$ is one, it is apparent that

$$0 \leq \sum_{x \in D_{j_0 \cdots j_{s-1}}} G(x) \leq \frac{1}{2}.$$

It follows that

$$d_{i_0 \cdots i_{s-1}} - \frac{s}{2} \leq \sum_{x \in Z_N} G(x) \leq d_{i_0 \cdots i_{s-1}} + \frac{s}{2}.$$

Combining this and the estimation about $\sum_x G(x)$ proves the following proposition.

Proposition 5.11 *Let symbols and assumptions be as before. We have*

$$-\frac{s(s-1)}{2^{s+1}} - A < d_{i_0 \cdots i_{s-1}} - \frac{N}{2^s} < -\frac{s(s-1)}{2^{s+1}} + A,$$

where

$$A = \frac{\{(s-3)2^{s-1} + 2\}\sqrt{N} - \left\lfloor \frac{s-2}{2} \right\rfloor}{2^s}.$$

It is easy to see that the following holds generally:

$$A < \frac{s-2}{2}\sqrt{N}.$$

Thus, if N is very large, the bounds here are much tighter than the Peralta bounds. It is worthy to mention that the Davenport reduction theorem contributes much to the new bounds.

For the case $s = 3$, the above proof shows that

$$\frac{-3 - 2\sqrt{N}}{2^3} < d_{i_0 i_1 i_2} - \frac{N}{2^3} < \frac{-3 + 2\sqrt{N}}{2^3}.$$

Some numerical examinations show that this lower and upper bounds are quite tight. It can be seen from the development of the bounds that the new bounds are usually tight for small s. But they seem loose for large s. For cryptographic purposes, we need better bounds.

Now we are ready to discuss the stability of the mutual information

$$I(i; z_{r_0} z_{r_{i_1}} \cdots z_{r_{t-1}} = i_0 i_1 \cdots i_{t-1}),$$

where $r_0 = 0$ and $r_1, \cdots, r_{t-1}$ are arbitrary distinct elements of Z_N. Owing to the bounds of Proposition 5.11 on the parameters $d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})$ it could be concluded that the stability of this mutual information is ideal for small t. This means also that the pattern distribution of length t in the difference set characterized sequences is ideal for small t. Further conclusions about the stability of this mutual information and about the pattern distribution need tight bounds on the parameters $d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})$ when t is large.

In developing the bounds on the parameters $d_{i_0 \cdots i_{t-1}}(r_0, \cdots, r_{t-1})$, we treat arbitrary $r_0, r_1, \cdots, r_{t-1}$. For the set of special values

$$(r_0, r_1, \cdots, r_{t-1}) = (r, r+1, \cdots, r+t-1),$$

where r is arbitrary, it might be possible to develop better bounds for these special parameters.

We now introduce a decision-tree based attack on the difference set characterized generator. The decision is based on the mutual information analysis. To illustrate this, let us take the example of $N = 4t-1 = 19$ with $t = 5$. The difference set and its complement are respectively

$$D = \{1, 4, 5, 6, 7, 9, 11, 16, 17\},$$
$$D^* = \{0, 2, 3, 8, 10, 12, 13, 14, 15, 18\}.$$

We show now how to determine the key i by the following procedure. Let $s_j = z_{i+j}$ for each $j \geq 0$. The procedure can be described as:

Step 1:

$$s_0 = 1 \Rightarrow i \in D_1 = D \ (s_0 = 0 \Rightarrow i \in D_1^* = D^*)$$

Step 2:

$$s_0 s_1 = \left\{ \begin{array}{l} 10 \Rightarrow i \in \{1, 7, 9, 11, 17\} = D_{10} \\ 11 \Rightarrow i \in \{4, 5, 6, 16\} = D_{11} \end{array} \right\} D_1$$

Step 3:

$$s_0 s_1 s_2 = \left\{ \begin{array}{l} \left. \begin{array}{l} 100 \Rightarrow i \in \{1, 11\} = D_{100} \\ 101 \Rightarrow i \in \{7, 9, 17\} = D_{101} \end{array} \right\} D_{10} \\ \left. \begin{array}{l} 110 \Rightarrow i \in \{6, 16\} = D_{110} \\ 111 \Rightarrow i \in \{4, 5\} = D_{111} \end{array} \right\} D_{11} \end{array} \right\} D_1$$

Step 4:

$$s_0s_1s_2s_3 = \left\{ \left. \begin{array}{l} \left. \begin{array}{l} \left.\begin{array}{l} 1000 \Rightarrow i \in \{11\} = D_{1000} \\ 1001 \Rightarrow i \in \{1\} = D_{1001} \end{array}\right\} D_{100} \\ \left.\begin{array}{l} 1010 \Rightarrow i \in \{7, 9\} = D_{1010} \\ 1011 \Rightarrow i \in \{17\} = D_{1011} \end{array}\right\} D101 \end{array}\right\} D_{10} \\ \left. \begin{array}{l} \left.\begin{array}{l} 1100 \Rightarrow i \in \{16\} = D_{1100} \\ 1101 \Rightarrow i \in \{6\} = D_{1101} \end{array}\right\} D_{110} \\ \left.\begin{array}{l} 1110 \Rightarrow i \in \{5\} = D_{1110} \\ 1111 \Rightarrow i \in \{4\} = D_{1111} \end{array}\right\} D_{111} \end{array}\right\} D_{11} \end{array}\right\} D_1 \right.$$

Step 5:

$$s_0s_1s_2s_3s_4 = \left\{ \left. \begin{array}{l} 10100 \Rightarrow i \in \{9\} = D_{10100} \\ 10101 \Rightarrow i \in \{7\} = D_{10101} \end{array}\right\} D_{1010} \right.$$

Thus, we obtain the minimal number of consecutive bits of key stream

$$z^u = z_i z_{i+1} \cdots z_{i+u+1},$$

which specifically determine the key $i \in D$. For each key $i \in D^*$, we can similarly determine the sequence. The minimal sequences which determine the keys in D are

$$\begin{array}{lllllllll} 1 & = & 1001, & 4 & = & 1111, & 5 & = & 1110, \\ 6 & = & 1101, & 11 & = & 1000, & 9 & = & 10100, \\ 16 & = & 1100, & 17 & = & 1011, & 7 & = & 10101. \end{array}$$

The foregoing discussion shows the difference set characterized generator is theoretically breakable if a cryptanalyst knows enough consecutive bits of plaintext-ciphertext-bit pairs. We analyze now whether the generator is *computationally breakable*. Before doing so, we observe first the lower bound in Proposition 5.11. For stream ciphering purpose the modulus N is usually very large. So if s is also large, the lower bound in Proposition 5.11 is approximately

$$\frac{N}{2^s} - \frac{(s-1)\sqrt{N}}{2}.$$

Thus determining each key requires at least $M = \lceil A \rceil$ consecutive bits of key stream, where A is the solution of the following equation

$$\frac{N}{2^s} - \frac{(s-1)\sqrt{N}}{2} = 1.$$

Suppose that a cryptanalyst gets the keystream segment

$$z_i \cdots z_{i+u} = s_0 \cdots s_u,$$

where $u > \lfloor \log_2 N \rfloor$. To determine the present key i with the above method, he/she has to use an algorithm to compute $f(i)$ for arbitrary $i \in Z_N$. Assume that the cryptanalyst uses the fast exponentiation algorithm and the modulo N as well as the modulo 2 operations to determine each $D_{s_0 \cdots s_v}$ for each v with $0 \le v \le \lfloor \log_2 N \rfloor$. If we take each integer multiplication, each integer addition, each modulo 2 and each modulo N as one unit of computation, then the number of operations needed to determine the key i is at least

$$\begin{aligned} & 2M \sum_{i=0}^{M-1} N/2^i \quad \text{(from the part of exponentiation mod 2)} \\ + \; & 2 \sum_{i=0}^{M-1} N/2^i \quad \text{(from the part of } (\textstyle\sum)_N) \\ + \; & \sum_{i=0}^{M-1} N/2^i \quad \text{(from the part of } y \text{ mod 2)} \\ = \; & 2N(M+3)(1 - \frac{1}{2^M}) \\ = \; & \Theta(NM), \end{aligned}$$

here M is the one described above. If N is large enough, it is apparently impossible to determine the key computationally.

Generally, let $U_A(N)$ be the minimal number of operations to compute $f(i)$ for each $i \in Z_N$ by an algorithm A. Then we can similarly prove that the minimal number of operations of this kind of key-determining attack based on the algorithm A is at least

$$2U_A(N) \sum_{i=0}^{M-1} N/2^i = 4NU_A(N)(1 - \frac{1}{2^M}).$$

Note that for any algorithm A we have $U_A(N) \ge 1$, we know that for any such a key-determining attack, the minimum number of operations needed is at least

$$4N(1 - \frac{1}{2^M}).$$

If N is chosen to be large enough, for example, say about 2^{100}, any attack of this kind is computationally infeasible for the time being. On the other hand, it seems that the storage space needed is at least $O(N/2)$ with this procedure.

5.4 Sums of Difference Set Characterized Sequences

Since difference set characterized sequences are cryptographically attractive in many aspects, we analyze the bitwise-XOR of two difference set characterized sequences. For almost difference set characterized sequences the analysis is almost the same.

Let $N_1 = 4t_1 - 1$ and $N_2 = 4t_2 - 1$ be two distinct large primes with t_1 and t_2 being odd, z_1^∞ and z_2^∞ be the corresponding difference set characterized sequences.

5.4.1 Linear Complexity Analysis

The generating functions of the two sequences can be written as

$$\begin{aligned} z_1^\infty &= z_1^{N_1}(x)/(x^{N_1}+1), \\ z_2^\infty &= z_2^{N_2}(x)/(x^{N_2}+1). \end{aligned}$$

Hence

$$z^\infty(x) = (z_1+z_2)^\infty(x) = \frac{z_1^{N_1}(x)(x^{N_2}+1) + z_2^{N_2}(x)(x^{N_1}+1)}{(x^{N_1}+1)(x^{N_2}+1)}.$$

Since both N_1 and N_2 are primes and $N_1 \neq N_2$, we get that $\gcd(x^{N_1}+1,\ x^{N_2}+1) = (x+1)$. On the other hand, it follows from $\mathrm{L}(z_1^\infty) = N_1$ and $\mathrm{L}(z_2^\infty) = N_2$ that

$$\gcd(z_1^{N_1}(x),\ x^{N_1}+1) = \gcd(z_2^{N_2}(x),\ x^{N_2}+1) = 1.$$

This ensures that

$$\gcd((x^{N_1}+1)(x^{N_2}+1),\ z_1^{N_1}(x)(x^{N_2}+1) + z^{N_2}(x)(x^{N_1}+1)) = 1.$$

Thus, the minimal polynomial of the sequence z^∞ is $(x^{N_1}+1)(x^{N_2}+1)$, and therefore

$$\mathrm{L}(z_1^\infty + z_2^\infty) = N_1 + N_2.$$

5.4.2 Balance Analysis

It is easily seen that

$$\begin{aligned} p(z_{1i}=0) &= (N_1+1)/2N_1, \quad p(z_{1i}=1) = (N_1-1)/2N_1, \\ p(z_{2i}=0) &= (N_2+1)/2N_2, \quad p(z_{2i}=1) = (N_2-1)/2N_2. \end{aligned}$$

Note that z_{1i} and z_{2i} are statistically independent, we have

$$\begin{aligned} p(z_i=1) &= p(z_{1i}+z_{2i}=1) \\ &= p(z_{1i}=0,\ z_{2i}=1) + p(z_{1i}=1,\ z_{2i}=0) \\ &= \frac{N_1+1}{2N_1}\frac{N_2-1}{2N_2} + \frac{N_1-1}{2N_2}\frac{N_2+1}{2N_2} \\ &= \frac{N_1N_2-1}{2N_1N_2}. \end{aligned}$$

This means that the sequence z^∞ is almost balanced, with almost the same balance property as the sequences z_1^∞ and z_2^∞.

5.4.3 Correlation Analysis

It is easily verified that

$$p(z_i = z_{1i}) = p(z_{2i} = 0) = (N_2 + 1)/2N_2,$$
$$p(z_i = z_{2i}) = p(z_{1i} = 0) = (N_1 + 1)/2N_1.$$

This means that any correlation attack by making use of the correlation between z_i and z_{1i}, z_i and z_{2i} is impossible.

5.4.4 Differential Analysis

It is easy to see the equivalence between the autocorrelation analysis and the difference analysis of binary sequences. For the correlation property of the sequence z^∞, we have the following conclusion.

Theorem 5.12 *The autocorrelation function of the sequence z^∞ is 4-valued, i.e.,*

$$\mathrm{AC}_z(j) = \begin{cases} 1, \quad j = 0; \\ -\frac{1}{N_2}, \quad j = kN_1,\ k \neq 0; \\ -\frac{1}{N_1}, \quad j = kN_2,\ k \neq 0; \\ \frac{1}{N_1N_2}, \quad \gcd(j,\ N_1N_2) = 1. \end{cases}$$

Proof: If $j = kN_1,\ k \neq 0$, then

$$\begin{aligned} \mathrm{AC}_z(j) &= \sum_{i=0}^{N_1N_2-1} (-1)^{z_{2,i}+z_{2,i+j}}/N_1N_2 \\ &= C_{z_2}(j \bmod N_2) = -1/N_2. \end{aligned}$$

Similarly, we can prove that

$$\mathrm{AC}_z(j) = -1/N_1, \quad \text{for } j = kN_2,\ k \neq 0.$$

If $\gcd(j,\ N_1N_2) = 1$, then by definition we have

$$\begin{aligned} \mathrm{AC}_z(j) &= [2|\{z_{1,i} + z_{1,i+j} = 0\} \cap \{z_{2,i} + z_{2,i+j} = 0\}| \\ &\quad +2|\{z_{1,i} + z_{1,i+j} = 1\} \cap \{z_{2,i} + z_{2,i+j} = 1\}| - N_1N_2]/N_1N_2 \\ &= 2p(z_{1,i} + z_{1,i+j} = 0)p(z_{2,i} + z_{2,i+j} = 0) \\ &\quad +2p(z_{1,i} + z_{1,i+j} = 1)p(z_{2,i} + z_{2,i+j} = 1) - 1 \\ &= (C_{z_1}(j \bmod N_1) + 1)(C_{z_2}(j \bmod N_2) + 1)/2 \\ &\quad +(1 - C_{z_1}(j \bmod N_1))(1 - C_{z_2}(j \bmod N_2))/2 - 1 \\ &= 1/N_1N_2. \end{aligned}$$

This completes the proof. □

Now we calculate the difference parameters $d_s(g, g'; j)$'s. Let D be the characteristic set of z^∞, then $|D| = (N_1N_2 - 1)/2$. On the other hand, we have

$$\begin{aligned}
&|(D+j)\cap D| + |(D+j)\cap D^*| = (N_1N_2-1)/2,\\
&|(D^*+j)\cap D^*| + |(D^*+j)\cap D| = (N_1N_2+1)/2,\\
&|(D+j)\cap D| + |(D^*+j)\cap D| = (N_1N_2-1)/2,
\end{aligned}$$

where $D^* = Z_N \setminus D$. Consequently, we obtain

$$\begin{aligned}
\mathrm{AC}_z(j) &= [|(D+j)\cap D| + |(D^*+j)\cap D^*| \\
&\quad -|(D+j)\cap D^*| - |(D^*+j)\cap D|]/N_1N_2 \\
&= [4|(D+j)\cap D| + 2 - N_1N_2]/N_1N_2,
\end{aligned}$$

whence for $j \neq 0$,

$$|(D+j)\cap D| = \begin{cases} (N_1N_2 - N_1 - 2)/4, & j = kN_1,\ k \neq 0; \\ (N_1N_2 - N_2 - 2)/4, & j = kN_2,\ k \neq 0; \\ (N_1N_2 - 1)/4, & \gcd(j;\ N_1N_2) = 1 \end{cases}$$

and

$$|(D+j)\cap D^*| = \begin{cases} (N_1N_2 + N_1)/4, & j = kN_1,\ k \neq 0; \\ (N_1N_2 + N_2)/4, & j = kN_2,\ k \neq 0; \\ (N_1N_2 - 1)/4, & \gcd(j;\ N_1N_2) = 1. \end{cases}$$

These results show that D is not a difference set of $Z_{N_1N_2}$, but has a relatively good difference property. Speaking specifically, for all $j \in Z_N$ with $\gcd(j,\ N_1N_2) = 1$, the equations

$$j = d_1 - d_2, \quad d_1,\ d_2 \in D$$

have the same number of solutions. Only for those j's with $j \bmod N_1 = 0$ and $j \bmod N_2 = 0$, the above equation has a different number of solutions. It is necessary to choose two primes N_1 and N_2 such that $|N_1 - N_2|$ is small, in order to get a sum sequence with better autocorrelation and difference property.

Chapter 6

Two-Prime Generators

Based on the generalized cyclotomic numbers, two kinds of generators are described in this chapter. Especially interesting are the twin-prime generators which are special two-prime generators. It is shown that the generalized cyclotomic numbers are of significance to the two-prime generators in a number of cryptographic aspects. All the seven cryptographic aspects of the generators described in Section 3.3 are controllable.

Main contributions of this chapter are the following:

1. The discovery of the connections between the generalized cyclotomy and the two-prime generator;

2. a detailed analysis of the two-prime and twin-prime generators based on generalized cyclotomic numbers, including one of the main results— Theorem 6.9;

3. control of the linear and sphere complexity of the output sequences of the generators, as described by Theorem 6.10 and Corollaries 6.11, 6.12, 6.13, 6.14, and 6.15.

This chapter is mainly based on the paper [24, 29].

6.1 Generalized Cyclotomic Numbers

Let p and q be two distinct odd primes with $\gcd(p-1, q-1) = 2$, and

$$R = \{0\}, \quad P = \{p, 2p, \cdots, (q-1)p\}, \quad Q = \{q, 2q, \cdots, (p-1)q\}.$$

Whiteman's approach is first to get a proper representation of the elements of the multiplicative group Z_{pq}^* as described in the following proposition [70].

Proposition 6.1 *Let g be a fixed common primitive root of both primes p and q; let $d = \gcd(p-1, q-1)$ and $de = (p-1)(q-1)$. Then there exists an integer x such that*

$$Z_{pq}^* = \{g^s x^i : s = 0, 1, \cdots, e-1;\ i = 0, 1, \cdots, d-1\}.$$

The set Z_{pq}^* is also called the reduced residue system modulo $N = pq$. In Whiteman's generalized cyclotomy the *index class* or *cyclotomic class* D_i consists of e numbers and is defined by

$$D_i = \{g^s x^i : s = 0, 1, \cdots, e-1\}$$

and the generalized cyclotomic number $(i, j)_d$ by

$$(i, j)_d = |(D_i + 1) \cap D_j|.$$

There are d cyclotomic classes $D_0, \cdots, D_{d-1}$, which form a partition of Z_{pq}^*.

The proof of the proposition, which is a constructive one, can be found in [3, p. 131] or [70]). Here it is necessary to mention how the x is constructed. Whiteman's construction of the x is based on the solution of the following simultaneous congruences

$$\begin{aligned} x &= g \bmod p, \quad y = 1 \bmod p \\ x &= 1 \bmod q, \quad y = g \bmod q. \end{aligned}$$

The existence and uniqueness of such x, y are guaranteed by the Chinese Remainder Theorem. One important relation between x and g for our applications is

$$x^d = g^u \bmod N$$

for some fixed u such that $0 \le u \le e-1$.

For cryptographic purposes another fact about the foregoing g has to be mentioned, that is, the order of g modulo $N = pq$ is the e=lcm$(p-1, q-1)$. We now analyze the relation between the difference property of the partition of Z_{pq}^* and the generalized cyclotomic numbers. It is obvious that $x \in Z_{pq}^*$. It can be proven that the order of x modulo N is $p-1 \ge d$. Let $w \in Z_{pq}^*$. Then there must exist two integers s and t with $0 \le s \le e-1$, $0 \le t \le d-1$ such that $w = g^s x^t$. Because $x^d = g^u$ for some fixed u such that $0 \le u \le e-1$, the difference parameter can be expressed as

$$\begin{aligned} d(i, j; w) &= |(D_i + g^s x^t) \cap D_j| \\ &= |(D_{(p-1-t+i) \bmod d} + 1) \cap D_{(p-1-t+j) \bmod d}| \\ &= ((p-1-t+i) \bmod d, (p-1-t+j) \bmod d)_d, \end{aligned}$$

where $0 \leq i,\ j \leq d-1;\ w \in Z_N^*$. This means that for each $(i,j;w)$ with $0 \leq i,j \leq d-1$, $w \in Z_{pq}^*$, the difference parameter $d(i,j;w)$ is in fact one cyclotomic number. We will discuss the case for $w \notin Z_{pq}^*$ later.

Similar to the cyclotomy in prime fields Z_p's, there are some elementary properties of this generalized cyclotomy. Apparently, d is even. Let $p-1 = df$, $q-1 = df'$, $e = dff'$ for some relatively prime integers f, f' (in particular f, f' are not both even). Then it is easy to prove

$$-1 = \begin{cases} g^{e/2} \bmod N & \text{when } ff' \text{ is odd;} \\ g^v x^{d/2} \bmod N & \text{when } ff' \text{ is even,} \end{cases}$$

where v is some fixed integer, $0 \leq v \leq e-1$. The following properties about generalized cyclotomic numbers are fundamental:

$$(A) \quad (l,m)_d = (l',m')_d \text{ when } l = l' \bmod d \text{ and } m = m' \bmod d;$$

$$(B) \quad (l,m)_d = (d-l, m-l)_d = \begin{cases} (m,l)_d, & ff' \text{ odd} \\ (m+d/2, l+d/2)_d, & ff' \text{ even} \end{cases}$$

$$(C) \quad \sum_{m=0}^{d-1} (l,m)_d = \frac{(p-2)(q-2)-1}{d} + n_l,$$

where

$$n_l = \begin{cases} 1, & l = 0 \bmod d, \quad ff' \text{ odd,} \\ 1, & l = d/2 \bmod d, \quad ff' \text{ even,} \\ 0, & \text{otherwise.} \end{cases}$$

$$(D) \quad \sum_{l=0}^{d-1} (l,m)_d = \frac{(p-2)(q-2)-1}{d} + k_m,$$

where

$$k_m = \begin{cases} 1, & \text{if } m = 0 \bmod d; \\ 0, & \text{otherwise.} \end{cases}$$

$$(E) \quad \sum_{l=0}^{d-1} \sum_{m=0}^{d-1} (l,m)_d = (p-2)(q-2).$$

The above elementary properties can be easily proved by definition [70, 3]. Properties (C-E) show the conservation property of the generalized cyclotomic numbers.

As seen above, the index classes $D_0, \cdots, D_{d-1}$ constitute a partition of Z_N^*. Since

$$R = \{0\}, \quad P = \{p, 2p, \cdots, (q-1)p\}, \quad Q = \{q, 2q, \cdots, (p-1)q\},$$

the sets $D_0, \cdots, D_{d-1}$; R; P; Q form a partition of Z_N. To extend the partition of Z_N^* into one of Z_N having ideal difference property, we have to study the difference property among the above sets. The following conclusion [70, p. 112] is useful for our application.

Proposition 6.2 *For any $r \in P \cup Q$,*

$$d(0, 1; r) = |(D_0 + r) \cap D_1| = (p-1)(q-1)/d^2.$$

For our cryptographic purposes, we need to know the value of $d(i, j; r)$. Actually, the following general conclusion is true.

Theorem 6.3 *For any $r \in P \cup Q$ and any $1 \leq k \leq d-1$,*

$$d(0, k; r) = |(D_0 + r) \cap D_k| = (p-1)(q-1)/d^2.$$

The proof for Proposition 6.2 given by Whiteman can be easily generalized to prove Theorem 6.3. Here it should be mentioned that the condition that $1 \leq k \leq d-1$ is necessary, as shown by the following conclusion [70].

Proposition 6.4 *Let the notations be the same as before, then*

$$\begin{aligned} d(0, 0; r) &= |(D_0 + r) \cap D_0| \\ &= \begin{cases} (p-1)(q-1-d)/d^2, & r \in P, r \notin Q; \\ (q-1)(p-1-d)/d^2, & r \in Q, r \notin P. \end{cases} \end{aligned}$$

Since $x \in Z_N^*$ and $x = g^u$ for some u with $0 \leq u \leq e-1$, for each r we have

$$\begin{aligned} d(i, j; r) &= |(D_i + r) \cap D_j| \\ &= |x^{d-i}(D_i + r) \cap x^{d-i} D_j| \\ &= d(0, (j + d - i) \bmod d; x^{d-i} r \bmod N). \end{aligned}$$

If $r \in P$ (resp. Q), then $x^{d-i} r \in P$ (resp. Q). Combining all the foregoing results in this section, we have completed the proof of the following theorem.

Theorem 6.5 *Let the notations be the same as before. Then for the partition* $D_0, \cdots, D_{d-1}$ *of* Z_N^* *and* $w \neq 0$*, we have*

$$d(i,j;w) = \begin{cases} (p-1)(q-1)/d^2, & i \neq j,\ w \in P \cup Q; \\ (p-1)(q-1-d)/d^2, & i = j,\ w \in P,\ w \notin Q; \\ (q-1)(p-1-d)/d^2, & i = j,\ w \in Q,\ w \notin P; \\ (i',j')_d \text{ for some } (i',j'), & \text{otherwise.} \end{cases}$$

In order to put the elements of R, P, Q into some of the D_i's to get a partition of Z_N with good difference property, we need the following result [70].

Proposition 6.6 *Let the notations be the same as before. If* $r \notin Q \cup R$*, then*

$$|D_0 \cap (Q \cup \{0\} + r)| = (p-1)/d.$$

With the help of Propositions 6.2, 6.4, and 6.6, we are ready to prove the following result [70].

Proposition 6.7 *Let the notations be the same as before. Then the set* $D_0 \cup R \cup Q$*, which is*

$$\{1, g, g^2, \cdots, g^{e-1}; 0, q, 2q, \cdots, (p-1)q\},$$

forms a difference set with parameters $N = pq$, $k = (N-1)/d$, $\lambda = (N-1-d)/d^2$ *if and only if* ff' *is even and the following conditions are satisfied:*

$$q = (d-1)p + 2$$
$$(i,0)_d = (d-1)[(p-1)/d]^2 \quad (i = 0, 1, \cdots, d-1).$$

It is important to note that $n = k - \lambda = (p-f)^2$ by the first condition above. It follows that $\gcd(n, N) = 1$. By a theorem in [23] the linear complexity of the characteristic sequences of these difference sets is N or $N-1$.

6.2 Two-Prime Generator of Order 2

In what follows in this section we assume that $d = \gcd(p-1, q-1) = 2$. To describe and analyze the cryptographic function for the generator, we now consider the characteristic function of the partition $\{R \cup Q \cup D_0,\ P \cup D_1\} = \{C_0, C_1\}$ of Z_{pq}.

The generalized cyclotomic numbers of order 2 obtained by Whiteman [70] are described by the following proposition.

Proposition 6.8 *Let the notations be as before. If ff' is even, we have $(0,0) = (1,0) = (1,1)$ and two different cyclotomic numbers*

$$(0,0) = \frac{(p-2)(q-2)+1}{4}, \quad (0,1) = \frac{(p-2)(q-2)-3}{4}.$$

If ff' is odd, we have $(0,1) = (1,0) = (1,1)$ and

$$(0,0) = \frac{(p-2)(q-2)+3}{4}, \quad (0,1) = \frac{(p-2)(q-2)-1}{4}.$$

We now analyze the difference property of the partition $\{C_0, C_1\}$ of Z_{pq}. Note that

$$d_C(0,0;r) = |[(R+r) \cup (Q+r) \cup (D_0+r)] \cap [R \cup Q \cup D_0]|.$$

Setting

$$\begin{aligned} a(0,0;r) &= |(Q+r) \cap Q| + |(Q+r) \cap D_0| \\ &+ |(D_0+r) \cap Q| + |(D_0+r) \cap D_0|, \end{aligned}$$

we can prove

$$0 \le d_C(0,0;r) - a(0,0,r) \le 2.$$

So our task now is to estimate the $a(0,0;r)$ with $r \ne 0$. One simple fact is

$$|(Q+r) \cap Q| = \begin{cases} p-2, & r \in Q; \\ 0, & r \in P \cup Z_{pq}^*; \end{cases}$$

Note that if $r \in P$, then it is possible to have $Q + r \subset D_0$. Thus, for each r we have the following two apparent facts:

$$\begin{aligned} &0 \le |(Q+r) \cap D_0| \le p-1; \\ &0 \le |Q \cap (D_0+r)| \le p-1. \end{aligned}$$

It follows that

$$|(D_0+r) \cap D_0| \le a(0,0;r) \le 3p - 4 + |(D_0+r) \cup D_0|.$$

Setting

$$B = \max\{\frac{(p-2)(q-2)+3}{4}, \frac{(p-1)(q-3)}{4}, \frac{(p-3)(q-1)}{4}\}$$

and

$$C = \min\{\frac{(p-2)(q-2)-3}{4}, \frac{(p-1)(q-3)}{4}, \frac{(p-3)(q-1)}{4}\},$$

we get

$$C \leq a(0,0;r) \leq 3p - 4 + B,$$

and therefore

$$C \leq d_C(0,0;r) \leq 3p - 2 + B.$$

We can similarly prove that for each $r \neq 0$, it holds

$$C \leq d_C(1,1;r) \leq 3q - 4 + B.$$

In what follows we analyze $d_C(1,0;r)$ and $d_C(0,1;r)$. By definition we have

$$\begin{aligned}
& d_C(1,0;r) \\
= \; & |(C_1 + r) \cap C_0| = |[(P+r) \cup (D_1 + r)] \cap (R \cup Q \cup D_0)| \\
= \; & |(P+r) \cap R| + |(P+r) \cap Q| + |(P+r) \cap D_0| \\
& + |(D_1 + r) \cap R| + |(D_1 + r) \cap Q| + |(D_1 + r) \cap D_0|.
\end{aligned}$$

If $r \in P$, then by Proposition 6.2 we have

$$|(D_1 + r) \cap D_0| = (p-1)(q-1)/4.$$

In addition we have apparently

$$\begin{aligned}
& |(P+r) \cap Q| = |(P+r) \cap D_0| = |(D_1 + r) \cap R| = 0 \\
& |(P+r) \cap R| = 1, \;\; 0 \leq |(D_1 + r) \cap Q| \leq p - 1.
\end{aligned}$$

Hence, we obtain in the case $r \in P$

$$1 + \frac{(p-1)(q-1)}{4} \leq d_C(1,0;r) \leq \frac{(p-1)(q-1)}{4} + p.$$

If $r \in Q$, we can similarly prove

$$\frac{(p-1)(q-1)}{4} \leq d_C(1,0;r) \leq \frac{(p-1)(q-1)}{4} + q - 1.$$

If $r \in Z_{pq}^* \setminus P \cup Q \cup R$, then by Proposition 6.8 we get

$$\frac{(p-2)(q-2) - 3}{4} \leq |(D_1 + r) \cap D_0| \leq \frac{(p-2)(q-2) + 3}{4}.$$

In addition we have apparently

$$\begin{aligned}
& |(P+r) \cap R| = 0 \\
& 0 \leq |(P+r) \cap Q| \leq \min\{p-1, q-1\} \\
& 0 \leq |(P+r) \cap D_0| \leq q - 1 \\
& 0 \leq |(D_1 + r) \cap Q| \leq p - 1 \\
& 0 \leq |(D_1 + r) \cap R| \leq 1.
\end{aligned}$$

It follows in this case that

$$\begin{aligned}\tfrac{(p-2)(q-2)-3}{4} &\leq d_C(1,0;r)\\ &\leq \tfrac{(p-1)(q-1)}{4} + \min\{p-1, q-1\} + \tfrac{3}{4}(p+q) + \tfrac{1}{2}.\end{aligned}$$

Combining the results for the three cases, we obtain

$$\frac{(p-2)(q-2)-3}{4} \leq d_C(1,0;r) \leq \frac{(p-1)(q-1)}{4} + E,$$

where

$$E = \max\{p, q-1, \min\{p-1, q-1\} + \frac{3}{4}(p+q) + \frac{1}{2}\}.$$

Similarly, one can prove

$$\frac{(p-2)(q-2)-3}{4} \leq d_C(0,1;r) \leq \frac{(p-1)(q-1)}{4} + E,$$

Thus, we have proved the following theorem.

Theorem 6.9 *Let*

$$a = \frac{(p-2)(q-2)-3}{4}, \qquad d = \frac{(p-1)(q-1)}{4},$$

$$b = \max\{\frac{(p-2)(q-2)+3}{4}, \frac{(p-1)(q-3)}{4}, \frac{(p-3)(q-1)}{4}\}$$

and

$$c = \min\{\frac{(p-2)(q-2)-3}{4}, \frac{(p-1)(q-3)}{4}, \frac{(p-3)(q-1)}{4}\}.$$

then we have

$$\begin{aligned}&c \leq d_C(0,0;r) \leq 3p-2+b,\\ &c \leq d_C(1,1;r) \leq 3q-4+b,\\ &a \leq d_C(1,0;r) \leq d+E,\\ &a \leq d_C(0,1;r) \leq d+E,\end{aligned}$$

for each $r \neq 0 \bmod N$*, where*

$$E = max\{p, q-1, \min\{p-1, q-1\} + \frac{3}{4}(p+q) + \frac{1}{2}\}.$$

This theorem tells us that if $|p-q|$ is small enough, the cryptographic function $F_C(x)$ has an ideal difference property. Thus, the other six aspects are automatically ensured due to our formulae in Section 3.3. In this case, the facts that $|C_0| = (p-1)(q-1)/2 + q$ and $|C_1| = (p-1)(q-1)/2 + p - 1$, show that the function has also good balance. It is called a cyclotomic generator because the difference property and the nonlinearity of the above cryptographic function depend on the generalized cyclotomy developed by Whiteman.

It is not difficult to see that the characteristic function of the partition $\{C_0, C_1\}$ can be expressed by

$$F_C(j) = \begin{cases} 0, & j \in R \cup Q; \\ 1, & j \in P; \\ (1 - (\frac{j}{p})(\frac{j}{q}))/2, & \text{otherwise.} \end{cases}$$

With this cryptographic function the binary two-prime cyclotomic generator of order 2 is defined by

$$s(k)_i = F_C(k+i) \quad \text{for all } i \geq 0, \tag{6.1}$$

where $0 \leq k \leq pq - 1$ is the key of this generator.

On the other hand, the two primes should be chosen properly for the purpose of controlling the linear complexity and the linear complexity stability of the output sequence. In fact we have generally the following result, which is a special case of Basic Theorem 3.5.

Theorem 6.10 *Let $N = rs$ be a product of two distinct primes, u an integer with* $\gcd(u, N) = 1$. *Then for any nonconstant sequence s^∞ of period N over $GF(u)$, it holds*

$$\begin{array}{lll} \mathrm{L}(s^\infty) & \geq & \min\{\mathrm{ord}_r(u), \mathrm{ord}_s(u)\}; \\ \mathrm{SC}_k(s^\infty) & \geq & \min\{\mathrm{ord}_r(u), \mathrm{ord}_s(u)\}, \\ & & \textit{if } k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}, \end{array}$$

where $\mathrm{SC}_k(s^\infty)$ denotes the sphere complexity of the sequence, $\mathrm{WH}(s^N)$ the Hamming weight of the finite sequence, $ord_r(u)$ the order of u modulo r.

Proof: Setting $t = 2$ and $e_1 = e_2 = 1$ in Basic Theorem 3.5 proves this theorem. □

We have the following corollaries:

Corollary 6.11 *Let $r = 4t_1 + 1$, $s = 4t_2 + 1$, $r \neq s$. If r, s, t_1 and t_2 are odd primes, then for any nonconstant binary sequence of period $N = rs$,*

1. $L(s^\infty) \geq \min\{r-1, s-1\}$

2. $\mathrm{SC}_k(s^\infty) \geq \min\{r-1, s-1\}$, *if* $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$.

Proof: Note that the proof of Theorem 5.2 shows that $\mathrm{ord}_r(2) = r - 1$ and $\mathrm{ord}_s(2) = s-1$, that is, 2 is a primitive root of both r and s. Combining these two facts and Theorem 6.10 proves this corollary. □

Corollary 6.12 *Let* $r = 4r_1 - 1$, $s = 4s_1 - 1$, $(r-1)/2$ *and* $(s-1)/2$ *are all odd primes. If* $r > 5$ *and* $s > 5$, *then for each nonconstant binary sequence* s^∞ *of period* $N = rs$, *we have*

1. $\mathrm{L}(s^\infty) \geq \min\{r-1, s-1\}$;
2. $\mathrm{SC}_k(s^\infty) \geq \min\{r-1, s-1\}$; *if* $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$.

Proof: The Proof of Theorem 5.3 shows that 2 is a common primitive root of r and s. Combining this and Theorem 6.10 yields the conclusion of this corollary.
□

Corollary 6.13 *Let* $r = 4r_1 + 1$, $s = 4s_1 - 1$. *If* $r, r_1, s, (s-1)/2$ *are odd primes, then for each nonconstant binary sequence* s^∞ *of period* $N = rs$, *we have*

1. $\mathrm{L}(s^\infty) \geq \min\{r-1, s-1\}$;
2. $\mathrm{SC}_k(s^\infty) \geq \min\{r-1, s-1\}$; *if* $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$.

Proof: The proofs of Theorems 5.2 and 5.3 show that 2 is common primitive root of both r and s. Combining this and Theorem 6.10 proves this corollary. □

The above theorem and its three corollaries clearly show how to control the linear complexity and its stability for the output sequence of the binary two-prime generator of order 2.

Summarizing the above results, we see that the parameters should be chosen such that

1. p and q both are large enough with $\gcd(p-1, q-1) = 2$;
2. $|p - q|$ is small enough, compared with pq;
3. both of $\mathrm{ord}_p(2)$ and $\mathrm{ord}_q(2)$ are large enough.

It should be made clear that the speciality of the primes p and q determines partly the quadratic partition of the primes, and thus contributes to the stability of the cyclotomic numbers, and consequently to the difference property and other six aspects described in Section 3.3.

Generally speaking, the two-prime generator is more flexible, due to the fact that we have much freedom to select the primes.

In the remaining part of this section, under the condition that $\gcd(p-1, q-1) = 2$ we compute the linear complexity of the output sequences of the two-prime generator. Our results are based on [29] and show that their linear complexity is quite good. Therefore there is no need to use the speciality of primes to control the linear complexity.

Under the condition that $\gcd(p-1, q-1) = 2$, recall the definition of Whiteman's generalized cyclotomy of order 2 described in the previous section. We now compute the linear complexity of the sequences $s(k)^\infty$. They are shift versions of each other, so we need only to compute $s(0)^\infty$. For simplicity, we use s^∞ to denote $s(0)^\infty$ in the sequel.

The following lemma is straightforward by Propositions 2.1 and 2.2.

Lemma 6.14 *Let s^∞ be a sequence of period n over a field F, and*

$$S^n(x) = s_0 + s_1 x + \cdots + s_{n-1} x^{n-1}.$$

Then

1. *the minimal polynomial of s^∞ is given by*

$$(x^n - 1)/\gcd(x^n - 1, S^n(x)); \tag{6.2}$$

2. *the linear complexity of s^∞ is given by*

$$n - \deg(\gcd(x^n - 1, S^n(x))). \tag{6.3}$$

To compute the linear complexity of the sequence s^∞, we need a number of lemmas.

Lemma 6.15 *Let g, e, d, D_0 and D_1 be the same as before. Then*

1. $\mathrm{ord}_N(g) = e$, *where* $\mathrm{ord}_N(g)$ *denotes the order of g modulo N.*
2. *D_0 is a group with respect to the integer multiplication modulo pq.*
3. *If $a \in D_0$ then $aD_1 = D_1$ and $aD_0 = D_0$; if $a \in D_1$ then $aD_1 = D_0$ and $aD_0 = D_1$.*

Proof: Since g is a common primitive root of both p and q, by the Chinese Remainder Theorem

$$\begin{aligned} \mathrm{ord}_N(g) &= \mathrm{lcm}\{\mathrm{ord}_p(g), \mathrm{ord}_q(g)\} \\ &= \mathrm{lcm}\{p-1, q-1\} \\ &= (p-1)(q-1)/d = e. \end{aligned}$$

This proves part one.

The second part follows easily from part one and the definition of D_0.

Since $x \in Z_N^*$, there must exist an integer u with $0 \leq u \leq e-1$ such that $x^2 = g^u$. If $a \in D_1$, there must exist a v such that $a = g^v x$. It follows that

$$\begin{aligned} aD_1 &= \{g^{s+v}x^2 : s = 0, 1, \cdots, e-1\} \\ &= \{g^{s+v+2} : s = 0, 1, \cdots, e-1\} \\ &= D_0. \end{aligned}$$

The remaining parts can be similarly proved. □

Let m be the order of 2 modulo N. Then the field $GF(2^m)$ has an Nth primitive root of unity. Let θ denote such a primitive root and define

$$\begin{aligned} S(x) &= \sum_{i \in C_1} x^i \\ &= \sum_{i \in P} x^i + \sum_{i \in D_1} x^i \in GF(2)[x]. \end{aligned}$$

By (6.2) we now compute $\gcd(x^N - 1, S(x))$. To this end, we need some auxiliary results.

Note that θ is an Nth primitive root of unity, we have

$$0 = \theta^N - 1 = (\theta^p)^q - 1 = (\theta^p - 1)(1 + \theta^p + \theta^{2p} + \cdots + \theta^{(q-1)p}).$$

It follows that

$$\theta^p + \theta^{2p} + \cdots + \theta^{(q-1)p} = 1. \tag{6.4}$$

By symmetry we get

$$\theta^q + \theta^{2q} + \cdots + \theta^{(p-1)q} = 1. \tag{6.5}$$

Lemma 6.16 *Let the symbols be the same as before. Then*

$$\sum_{i \in D_1} \theta^{ai} = \begin{cases} (\frac{p-1}{2} \bmod 2), & \textit{if } a \in P; \\ (\frac{q-1}{2} \bmod 2), & \textit{if } a \in Q; \end{cases}$$

Proof: Suppose that $a \in Q$. Since g is a common primitive root of both p and q and the order of g modulo N is e, by the definition of x we have

$$\begin{aligned} D_i \bmod p &= \{g^s x^i \bmod p : s = 0, 1, \cdots, e-1\} \\ &= \{g^{s+i} \bmod p : s = 0, 1, \cdots, e-1\} \\ &= \{1, 2, \cdots, p-1\}. \end{aligned}$$

When s ranges over $\{0, 1, \cdots, e-1\}$, $g^s x^i \bmod p$ takes on each element of $\{1, 2, \cdots, p-1\}$ $(q-1)/2$ times. It follows from (6.4) that

$$\begin{aligned}\sum_{i \in D_1} \theta^{ai} &= \left(\frac{q-1}{2} \bmod 2\right) \sum_{i \in Q} \theta^i \\ &= \left(\frac{q-1}{2} \bmod 2\right).\end{aligned}$$

Similarly, we have

$$\begin{aligned}D_i \bmod q &= \{g^s x^i \bmod q : s = 0, 1, \cdots, e-1\} \\ &= \{g^{s+i} \bmod p : s = 0, 1, \cdots, e-1\} \\ &= \{1, 2, \cdots, q-1\}.\end{aligned}$$

The rest of the conclusion of this lemma can be similarly proved. □

Lemma 6.17 *Let the symbols be the same as before.*

$$S(\theta^a) = \begin{cases} S(\theta), & a \in D_0; \\ S(\theta) + 1, & a \in D_1; \\ 1 + (\frac{p-1}{2} \bmod 2), & a \in P; \\ (\frac{q-1}{2} \bmod 2), & a \in Q. \end{cases}$$

Proof: By Lemma 6.15, $aD_0 = D_0$ if $a \in D_0$. If $a \in D_0$, $aP = P$ since $\gcd(a, q) = 1$. Hence

$$\begin{aligned}S(\theta^a) &= \sum_{i \in P} \theta^{ai} + \sum_{i \in D_1} \theta^{ai} \\ &= \sum_{i \in P} \theta^i + \sum_{j \in aD_1} \theta^j \\ &= \sum_{i \in P} \theta^i + \sum_{j \in D_1} \theta^j \\ &= S(\theta).\end{aligned}$$

If $a \in D_1$, by Lemma 6.15 $aD_1 = D_0$. Note that

$$\left(\sum_{i \in D_0} + \sum_{i \in D_1} + \sum_{i \in P} + \sum_{i \in Q}\right) \theta^i + 1 = \sum_{i=0}^{N-1} \theta^i = 0.$$

By (6.4) and (6.5) together with Lemma 6.16 we obtain

$$S(\theta^a) = \sum_{i \in P} \theta^{ai} + \sum_{i \in D_1} \theta^{ai}$$

$$
\begin{aligned}
&= \sum_{i\in P}\theta^i + \sum_{j\in aD_1}\theta^j \\
&= \sum_{i\in P}\theta^i + \sum_{j\in D_0}\theta^j \\
&= S(\theta) + 1 + \sum_{i\in Q}\theta^i + \sum_{i\in P}\theta^i \\
&= S(\theta) + 1.
\end{aligned}
$$

If $a \in P$, then $aP = P$ since $\gcd(p, q) = 1$. Then by Lemma 6.16

$$
\begin{aligned}
S(\theta^a) &= \sum_{i\in P}\theta^{ai} + \sum_{i\in D_1}\theta^{ai} \\
&= \sum_{i\in P}\theta^i + \sum_{i\in D_1}\theta^{ai} \\
&= 1 + \sum_{i\in D_1}\theta^{ai} \\
&= 1 + \left(\frac{p-1}{2} \bmod 2\right).
\end{aligned}
$$

If $a \in Q$, then $aP = \{0\}$. Then by Lemma 6.16

$$
\begin{aligned}
S(\theta^a) &= \sum_{i\in P}\theta^{ai} + \sum_{i\in D_1}\theta^{ai} \\
&= [(q-1) \bmod 2] + \sum_{i\in D_1}\theta^{ai} \\
&= \sum_{i\in D_1}\theta^{ai} \\
&= \left(\frac{q-1}{2} \bmod 2\right).
\end{aligned}
$$

This completes the proof of this lemma. □

Note also that

$$S(1) = [p - 1 + (p-1)(q-1)/2] \bmod 2 = 0. \tag{6.6}$$

Lemma 6.18 $2 \in D_0$ *if and only if* $S(\theta) \in \{0, 1\}$.

Proof: By (6.4) and definition

$$
\begin{aligned}
S(\theta) &= \sum_{i\in P}\theta^i + \sum_{i\in D_1}\theta^i \\
&= 1 + \sum_{i\in D_1}\theta^i.
\end{aligned}
$$

Since the field $GF(2^m)$ has characteristic 2, if $2 \in D_0$ then we have

$$\begin{aligned} S(\theta)^2 &= S(\theta^2) = 1 + \sum_{i \in D_1} \theta^{2i} \\ &= 1 + \sum_{j \in 2D_1} \theta^j \\ &= 1 + \sum_{j \in D_1} \theta^j \\ &= S(\theta). \end{aligned}$$

Hence, $S(\theta) = 0$ or 1.

If $2 \in D_1$, then we have similarly

$$\begin{aligned} S(\theta)^2 &= S(\theta^2) = 1 + \sum_{i \in D_1} \theta^{2i} \\ &= 1 + \sum_{j \in 2D_1} \theta^j \\ &= 1 + \sum_{j \in D_0} \theta^j \\ &= 1 + \left(\sum_{j \in D_0} + \sum_{j \in D_1} + \sum_{j \in P} + \sum_{j \in Q} \right) \theta^j + \left(S(\theta) + \sum_{j \in Q} \theta^j \right) \\ &= 1 + S(\theta). \end{aligned}$$

Hence, $S(\theta) \notin \{0, 1\}$.

Since $2 \in D_0 \cup D_1$, we have completed the proof. □

In the sequel, we need the following Generalized Chinese Remainder Theorem.

Lemma 6.19 *Let m be the least common multiple of two positive integers m_1 and m_2. The system of congruences*

$$x \equiv a_1 \pmod{m_1}, \quad x \equiv a_2 \pmod{m_2} \tag{6.7}$$

has solutions if and only if

$$\gcd(m_1, m_2) | a_1 - a_2, \tag{6.8}$$

where $a|b$ means that a divides b. When the condition (6.8) holds, the system of congruences of (6.7) has only one solution modulo m.

For a proof of this theorem, we refer to [26].

Lemma 6.20 $2 \in D_0$ *if and only if* $p \equiv \pm 1 \pmod 8$ *and* $q \equiv \pm 1 \pmod 8$ *or* $p \equiv \pm 3 \pmod 8$ *and* $q \equiv \pm 3 \pmod 8$.

Proof: Assume that $2 \in D_0$. By definition there is an integer s with $0 \le s \le e-1$ such that $2 = (g^s \bmod pq)$. It follows that

$$g^s \equiv 2 \pmod p, \quad g^s \equiv 2 \pmod q.$$

If s is even, then 2 is a quadratic residue modulo both p and q. Hence, by the Law of Quadratic Reciprocity $p \equiv \pm 1 \pmod 8$ and $q \equiv \pm 1 \pmod 8$.

If s is odd, then 2 is a quadratic nonresidue modulo both p and q. Hence, by the Law of Quadratic Reciprocity $p \equiv \pm 3 \pmod 8$ and $q \equiv \pm 3 \pmod 8$. This proves the necessity.

If $p \equiv \pm 1 \pmod 8$ and $q \equiv \pm 1 \pmod 8$. Then 2 is a quadratic residue modulo both p and q. Thus, there are even s_1 and even s_2 with $0 \le s_1 \le p-1$ and $0 \le s_2 \le q-1$ such that

$$g^{s_1} \equiv 2 \pmod p, \quad g^{s_2} \equiv 2 \pmod q. \tag{6.9}$$

Note that $\gcd(p-1, q-1) = 2$ and s_1 and s_2 both are even, by the Generalized Chinese Remainder Theorem described in Lemma 6.19, there is an integer s with $0 \le s \le e-1$ such that

$$s \equiv s_1 \pmod{p-1}, \quad s \equiv s_2 \pmod{q-1}.$$

This s is unique. Hence, $g^s \equiv 2 \pmod{pq}$, and $2 \in D_0$.

If $p \equiv \pm 3 \pmod 8$ and $q \equiv \pm 3 \pmod 8$, we can similarly prove that $2 \in D_0$. □

Let θ be the same as before. Among the pq pqth roots of unity θ^i, where $0 \le i \le pq-1$, the q elements θ^i, $i \in P \cup R$, are qth roots of unity, the p elements θ^i, $i \in Q \cup R$, are pth roots of unity. Hence,

$$x^p - 1 = \prod_{i \in Q \cup R} (x - \theta^i), \quad x^q - 1 = \prod_{i \in P \cup R} (x - \theta^i).$$

Let

$$d(x) = \prod_{i \in D_0 \cup D_1} (x - \theta^i).$$

It follows that

$$x^{pq} - 1 = \prod_{i=0}^{pq-1} (x - \theta^i) = \frac{(x^p - 1)(x^q - 1)}{x - 1} d(x),$$

where $d(x) \in GF(2)[x]$.

In the sequel let L and $m(x)$ denote the linear complexity and minimal polynomial of our generalized cyclotomic sequence of order 2 with respect to the two primes p and q.

Theorem 6.21 **(I)** *If* $p \equiv 1 \pmod 8$ *and* $q \equiv 3 \pmod 8$ *or* $p \equiv -3 \pmod 8$ *and* $q \equiv -1 \pmod 8$, *then*

$$L = pq - 1, \quad m(x) = \frac{x^{pq} - 1}{x - 1}.$$

(II) *If* $p \equiv -1 \pmod 8$ *and* $q \equiv 3 \pmod 8$ *or* $p \equiv 3 \pmod 8$ *and* $q \equiv -1 \pmod 8$, *then*

$$L = (p-1)q, \quad m(x) = \frac{x^{pq} - 1}{x^q - 1}.$$

(III) *If* $p \equiv -1 \pmod 8$ *and* $q \equiv -3 \pmod 8$ *or* $p \equiv 3 \pmod 8$ *and* $q \equiv 1 \pmod 8$, *then*

$$L = pq - p - q + 1, \quad m(x) = \frac{(x^{pq} - 1)(x - 1)}{(x^p - 1)(x^q - 1)}.$$

Proof: Note that $\gcd(p-1, q-1) = 2$, By Lemma 6.20 these six cases described in this theorem are the only ones such that $\gcd(p-1, q-1) = 2$ and $2 \notin D_0$.

In the two cases of (I), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6))}; \\ \neq 0, & a \in D_0 \cup D_1 \quad \text{(By Lemma 6.20)}; \\ 1, & a \in P \cup Q. \end{cases}$$

Hence, $\gcd(x^{pq} - 1, S(x)) = x - 1$. It follows that

$$\begin{aligned} m(x) &= \frac{x^{pq} - 1}{\gcd(x^{pq} - 1, S(x))} = \frac{x^{pq} - 1}{x - 1}, \\ L &= \deg(m(x)) = pq - 1. \end{aligned}$$

In the two cases of (II), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6))}; \\ \neq 0, & a \in D_0 \cup D_1 \quad \text{(By Lemma 6.20)}; \\ 0, & a \in P; \\ 1, & a \in Q. \end{cases}$$

Hence, $\gcd(x^{pq} - 1, S(x)) = x^q - 1$. It follows that

$$\begin{aligned} m(x) &= \frac{x^{pq} - 1}{\gcd(x^{pq} - 1, S(x))} = \frac{x^{pq} - 1}{x^q - 1}, \\ L &= \deg(m(x)) = pq - q = (p-1)q. \end{aligned}$$

In the two cases of (III), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6));} \\ \neq 0, & a \in D_0 \cup D_1 \quad \text{(By Lemma 6.20);} \\ 0, & a \in P \cup Q; \end{cases}$$

Hence, $\gcd(x^{pq}-1, S(x)) = (x^p-1)(x^q-1)/(x-1)$. It follows that

$$\begin{aligned} m(x) &= \frac{x^{pq}-1}{\gcd(x^{pq}-1, S(x))} = \frac{(x^{pq}-1)(x-1)}{(x^p-1)(x^q-1)}, \\ L &= \deg(m(x)) = pq - p - q + 1. \end{aligned}$$

□

Define

$$d_a(x) = \prod_{i \in D_a} (x - \theta^i), \quad a = 0, 1.$$

In the case $2 \in D_0$, by Lemma 6.15 we have $2D_0 = D_0$ and

$$\begin{aligned} d_0(x)^2 &= \prod_{i \in D_0} (x^2 - \theta^{2i}) \\ &= \prod_{j \in 2D_0} (x^2 - \theta^j) \\ &= \prod_{j \in D_0} (x^2 - \theta^j) \\ &= d_0(x^2). \end{aligned}$$

Hence, $d_0(x) \in GF(2)[x]$. Similarly, we can prove that $d_1(x) \in GF(2)[x]$.

By definition

$$d(x) = d_0(x) d_1(x).$$

Thus, in the case $2 \in D_0$ we get

$$x^{pq} - 1 = \frac{(x^p-1)(x^q-1)d_0(x)d_1(x)}{x-1}. \tag{6.10}$$

Note that $d_0(x)$ and $d_1(x)$ depend on the choice of θ. However, by Lemmas 6.17 and 6.18, exactly one of $S(\theta)$ and $S(\theta^a)$ is zero, where a is an element of D_1. Thus, we can choose our θ such that $S(\theta) = 0$. With this choice we can fix the polynomials $d_0(x)$ and $d_1(x)$.

Theorem 6.22 (IV) *If* $p \equiv 1 \pmod 8$ *and* $q \equiv -1 \pmod 8$ *or* $p \equiv -3 \pmod 8$ *and* $q \equiv 3 \pmod 8$, *then*

$$L = \frac{pq + p + q - 3}{2}, \quad m(x) = \frac{x^{pq}-1}{(x-1)d_0(x)}.$$

(V) *If* $p \equiv -1 \pmod 8$ *and* $q \equiv 1 \pmod 8$ *or* $p \equiv 3 \pmod 8$ *and* $q \equiv -3 \pmod 8$*, then*

$$L = \frac{(p-1)(q-1)}{2}, \quad m(x) = d_1(x).$$

(VI) *If* $p \equiv -1 \pmod 8$ *and* $q \equiv -1 \pmod 8$ *or* $p \equiv 3 \pmod 8$ *and* $q \equiv 3 \pmod 8$*, then*

$$L = \frac{(p-1)(q+1)}{2}, \quad m(x) = \frac{(x^p-1)d_1(x)}{x-1}.$$

Proof: By Lemma 6.20 there are eight cases such that $2 \in D_0$, but two of them do not satisfy $\gcd(p-1, q-1) = 2$. It is easy to check that the six cases described in this theorem are the only ones such that $2 \in D_0$ and $\gcd(p-1, q-1) = 2$.

In the two cases of (IV), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6));} \\ 0, & a \in D_0 \quad \text{(By the choice of } \theta\text{);} \\ 1, & a \in D_1 \quad \text{(By the choice of } \theta\text{);} \\ 1, & a \in P \cup Q \quad \text{(By Lemma 6.17).} \end{cases}$$

Hence,

$$\begin{aligned} &\gcd(x^{pq}-1, S(x)) = (x-1)d_0(x), \\ &m(x) = \frac{x^{pq}-1}{\gcd(x^{pq}-1, S(x))} = \frac{x^{pq}-1}{(x-1)d_0(x)}, \\ &L = \deg(m(x)) = pq - 1 - (p-1)(q-1)/2 = (pq+p+q-3)/2. \end{aligned}$$

In the two cases of (V), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6));} \\ 0, & a \in D_0 \quad \text{(By the choice of } \theta\text{);} \\ 1, & a \in D_1 \quad \text{(By the choice of } \theta\text{);} \\ 0, & a \in P \cup Q \quad \text{(By Lemma 6.17).} \end{cases}$$

Hence,

$$\begin{aligned} &\gcd(x^{pq}-1, S(x)) = \frac{(x^p-1)(x^q-1)d_0(x)}{x-1}, \\ &m(x) = \frac{x^{pq}-1}{\gcd(x^{pq}-1, S(x))} = d_1(x), \\ &L = \deg(m(x)) = \frac{(p-1)(q-1)}{2}. \end{aligned}$$

In the two cases of (VI), by Lemma 6.17

$$S(\theta^a) = \begin{cases} 0, & a = 0 \quad \text{(By (6.6));} \\ 0, & a \in D_0 \quad \text{(By the choice of } \theta\text{);} \\ 1, & a \in D_1 \quad \text{(By the choice of } \theta\text{);} \\ 0, & a \in P \quad \text{(By Lemma 6.17);} \\ 1, & a \in Q \quad \text{(By Lemma 6.17);} \end{cases}$$

Hence,

$$\begin{aligned} &\gcd(x^{pq} - 1, S(x)) = (x^q - 1)d_0(x), \\ &m(x) = \frac{x^{pq} - 1}{\gcd(x^{pq} - 1, S(x))} = \frac{(x^p - 1)d_1(x)}{x - 1}, \\ &L = \deg(m(x)) = pq - \frac{(p-1)(q-1)}{2} - p + 1 = \frac{(p-1)(q+1)}{2}. \end{aligned}$$

□

By the Chinese Remainder Theorem the two-prime generator of order 2 can be implemented in Hardware as in Figure 6.1, where CC1 and CC2 denote two cyclic counters that count the numbers $\{0, 1, 2, \cdots, p-1\}$ and $\{0, 1, 2, \cdots, q-1\}$ cyclically, respectively, and within CC1 and CC2 there are registers R1 and R2 that store the current counted number. The initial contents k_1 and k_2 of the two registers form the key of this generator, i.e., $k = (k_1, k_2)$, where $0 \le k_1 \le p-1$ and $0 \le k_2 \le q-1$. Cyclic counters are very efficient and frequently seen in modern electronic devices. In Figure 6.1 MEC1 and MEC2 are two special chips for modular exponentiation with respect to p and q respectively. They are similar to RSA chips, and can also be made relatively efficient as the two primes here are much smaller than those for RSA public-key cryptosystem. Here we use primes having about 46 bits, while in RSA at least 512-bit primes are needed. MEC1 and MEC2 compute $x^{(p-1)/2} \bmod p$ and $y^{(q-1)/2} \bmod q$ respectively. The u_0 and v_0 denote the least significant bits of the output numbers of MEC1 and MEC2 respectively, and u_1 and v_1 the next bits of u_0 and v_0 respectively. Finally, $\otimes$ and $\oplus$ denotes the binary multiplier and adder that realize the multiplication and addition of $GF(2) = \{0, 1\}$.

The correctness of this implementation is proved as follows. For each $0 \le j \le pq - 1$ let

$$\begin{aligned} u &= j^{(p-1)/2} \bmod p \in \{0, 1, p-1\}, \\ v &= j^{(q-1)/2} \bmod q \in \{0, 1, q-1\}. \end{aligned}$$

By the definitions of u_0, u_1, v_0, v_1, and s_j, we have the following correspondence depicted by Table 6.1. It is easily seen that

$$s_j = (u_1 \oplus 1) \otimes v_1 \oplus (u_0 \oplus 1) \otimes v_0.$$

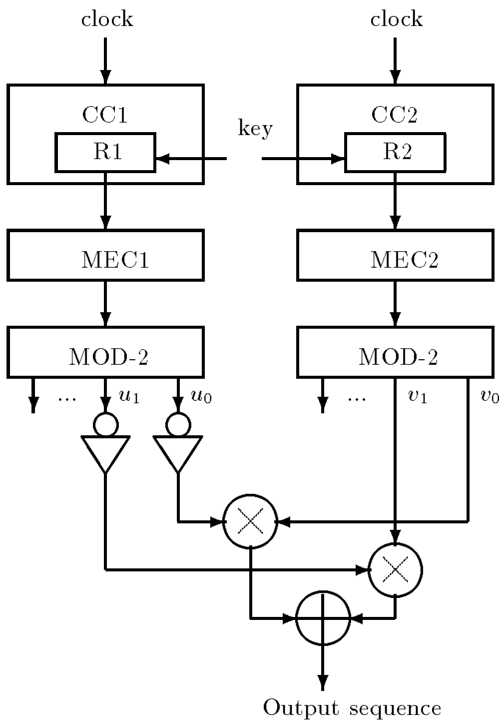


Figure 6.1: Hardware implementation of the two-prime generator of order 2.

(u, v)	(u_1, u_0, v_1, v_0)	s_j
$(1, 1)$	$(0, 1, 0, 1)$	0
$(1, q-1)$	$(0, 1, 1, 0)$	1
$(p-1, 1)$	$(1, 0, 0, 1)$	1
$(p-1, q-1)$	$(1, 0, 1, 0)$	0
$(0, 1)$	$(0, 0, 0, 1)$	1
$(0, q-1)$	$(0, 0, 1, 0)$	1
$(1, 0)$	$(0, 1, 0, 0)$	0
$(p-1, 0)$	$(1, 0, 0, 0)$	0
$(0, 0)$	$(0, 0, 0, 0)$	0

Table 6.1: The relations.

Then the correctness of this implementation follows from the Chinese Remainder Theorem.

The additive synchronous stream cipher based on this generator is as usual. For this purpose we suggest to use two 48-bit primes, then the keysize is 96 bits that should be large enough as far as brute-force attack is concerned. Twin primes might be better than others. With current chips for modular exponentiation with respect to such primes, this specific cipher should be able to encrypt and decrypt at least at 30 KBytes per second. Note that one page of English text (A4 size, ASCII) is about 3 Kbytes. Thus, such a cipher could encrypt 10 pages of English text per second. This performance may be slow for multimedia applications, but certainly reasonable in military and diplomatic communications, where the data in communication is usually not very large, say less than 600 pages of English text in each communication. Note that ciphering a 600-page document takes only one minute. It can also be used to encrypt and decrypt classified large data base where performance is not so important, but secrecy is the primary concern.

6.3 Two-Prime Generator of Order 4

To design cryptographic functions from Z_{pq} to Z_4, we follow the same approach as in the foregoing sections. Let $p = 4f + 1$, $q = 4f' + 1$ with $\gcd(f, f') = 1$, and $N = pq$. Then $d = \gcd(p-1, q-1) = 4$ and $e = 4ff'$. By a well-known theorem [43, p. 128] there are exactly two representations of N in the form $N = a^2 + b^2$ with $a = 1 \bmod 4$ and the sign of b indeterminate. Let

$$N = a^2 + 4b^2, \quad N = a'^2 + 4b'^2 \quad (a = a' = 1 \bmod 4)$$

denote these two representations. Let g be a common primitive root of p and q, and x be selected by the Chinese Remainder Theorem as in the foregoing section, (i, j) be the cyclotomic number defined as before. Whiteman proved that the sixteen cyclotomic constants depend solely upon one of the two representations. He also showed, for ff' even,

$$\begin{array}{ll} 8(0,0)_4 = -a + 2M + 3, & 8(0,1)_4 = -a - 4b + 2M - 1 \\ 8(0,2)_4 = 3a + 2M - 1, & 8(0,3)_4 = -a + 4b + 2M - 1 \\ 8(0,0)_4 = -a + 2M + 3, & \end{array}$$

the remaining (i, j)'s being equal to one of these cyclotomic numbers above. Here $M = [(p-2)(q-2) - 1]/4$. For the case ff' odd, he proved

$$\begin{array}{ll} 8(0,0)_4 = 3a + 2M + 5, & 8(0,1)_4 = -a + 4b + 2M + 1 \\ 8(0,2)_4 = -a + 2M + 1, & 8(0,3)_4 = -a - 4b + 2M + 1 \\ 8(1,2)_4 = a + 2M - 1, & \end{array}$$

the remaining (i,j)'s being equal to one of these cyclotomic numbers above. Note that

$$1 \le |a| \le \sqrt{pq-4}, \quad 1 \le |b| \le \sqrt{(pq-1)/4}$$

and that M is relatively much larger than a and b. We can conclude that the cyclotomic numbers have an ideal stability in both cases.

Define the function

$$F(j) = \begin{cases} 1, & \text{if } j \in \{0, q, 2q, \cdots, (p-1)q\}; \\ 0, & \text{if } j \in \{p, 2p, \cdots, (q-1)p\}; \\ ((j^{(q-1)/4} \bmod q) \bmod 2) \oplus ((j^{(p-1)/4} \bmod p) \bmod 2) \oplus 1, & \\ \quad \text{if } j \in Z_{pq}^*. & \end{cases}$$

With this $F(x)$ we describe a generator based on the generalized cyclotomy of order 4 as follows:

$$s(k)_i = F(k+i) \text{ for all } i \ge 0,$$

where $0 \le k \le pq-1$ is the key of this generator.

If we define the function $F^*(x)$ from Z_{pq}^* to Z_2 by

$$\begin{aligned} F^*(j) &= ((j^{(q-1)/4} \bmod q) \bmod 2) \\ &\oplus ((j^{(p-1)/4} \bmod p) \bmod 2) \oplus 1, \ j \in Z_{pq}^*. \end{aligned}$$

Then it is easy to show that $F^*(x)$ has the characteristic set $C_1 = D_i \cup D_j \cup Q$, where $Q = \{0, q, 2q, \cdots, (p-1)q\}$, D_i and D_j are two of the four generalized cyclotomic classes developed by Whiteman [70]. Since the above generalized cyclotomic numbers of order 4 are roughly stable, it follows from the formulae in Section 3.3 and the relation between cyclotomic numbers and the difference parameters described in Section 5.1 that we have ensured an ideal property for all the seven aspects described in Section 3.3.

The control of the linear complexity and its stability for the output sequence of the binary two-prime cyclotomic generator of order 4 is the same as that of order 2. By Theorems 6.9 and 6.10 the parameters should be chosen such that

1. p and q both are large enough with $\gcd(p-1, q-1) = 4$;
2. $|p-q|$ is small enough, compared with pq;
3. both of $\mathrm{ord}_p(2)$ and $\mathrm{ord}_q(2)$ are large enough.

Chapter 7

Prime-Square Generator

This chapter is concerned with a new kind of cyclotomy modulo p^2, and another kind of natural sequence generators. By controlling the difference property of the cryptographic function we shall control the seven aspects of the generator described in Section 3.3. Also the linear and sphere complexity of the output sequences of the generator are controlled.

The main contributions of this chapter are the introduction of a new kind of cyclotomy described in Section 7.1, and the analysis of the prime-square generator described by Theorems 7.2, 7.3, and 7.4. The materials of this chapter come mainly from [24, 30].

7.1 The Generator and Cyclotomy Modulo p^2

If we choose one of the following functions

$$F_C(x) = \begin{cases} 1, & x \in R; \\ (x^{p(p-1)/2} \bmod p^2) \bmod 2, & \text{otherwise} \end{cases}$$

and

$$F_C(x) = \begin{cases} 0, & x \in R; \\ (x^{p(p-1)/2} \bmod p^2) \bmod 2, & \text{otherwise} \end{cases}$$

for the cryptographic function of Figure 3.1.b, where $N = p^2$ and $R = \{0, p, 2p, \cdots, (p-1)p\}$, then we have the binary prime-square cyclotomic generator.

To describe the difference property of the above two cryptographic functions, we give a brief introduction to the cyclotomic numbers modulo p^2. Let p be a odd prime. There is a common primitive root α modulo both p and p^2. Setting $D_0 = (\alpha^2)$, a multiplicative subgroup of $Z_{p^2}^*$, and $D_1 = \alpha D_0$, then the cyclotomic numbers of order 2 modulo p^2 are defined by

$$(l, m)_2 = |D_l \cap (D_m + 1)|, \ 0 \le l, m \le 1.$$

We need the following theorem, which was conjectured by the author and proved by D. Pei [52], to ensure an ideal property for all the seven aspects described in Section 3.3.

Theorem 7.1 *Let the symbols be as before. If* $p = 3 \bmod 4$*, we have*

$$(0,1) = (0,0) = (1,1) = \frac{p(p-3)}{4}, \ (1,0) = \frac{p(p-3)}{4} + p.$$

If $p = 1 \bmod 4$*, we have*

$$(0,1) = (1,0) = (1,1) = \frac{p(p-1)}{4}, \ (0,0) = \frac{p(p-1)}{4} - p.$$

Let $\{C_0, C_1\}$ be the characteristic class of the above function $F_C(x)$, then similar to Section 5.1 it is easily verified that the difference parameters of the above function are approximately the same as the four cyclotomic numbers defined for the modulus $N = p^2$. Thus, an ideal difference property and therefore an ideal property for the other six aspects are ensured by the formulae described in Section 3.3.

7.2 Linear and Sphere Complexity

Sequences with period of the square of an odd prime are cryptographically attractive due to the following theorems about their linear and sphere complexity which follow easily from Basic Theorem 3.5.

Theorem 7.2 *Let r be an odd prime, $N = r^e$ and q an integer with* $\gcd(q, N) = 1$*. Then for any nonconstant sequence of period N over GF(q),*

1. $\mathrm{L}(s^\infty) \geq \ \mathrm{ord}_r(q)$ *;*
2. $\mathrm{SC}_k(s^\infty) \geq \ \mathrm{ord}_r(q)$*, if* $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$.

Proof: Setting $t = 1$ in Basic Theorem 3.5 proves this theorem. □

Theorem 7.3 *Let r be a odd prime, $N = r^2$ and q a primitive root modulo r and assume that r^2 does not divide $q^{r-1} - 1$, then for any nonzero sequence of period N over GF(q),*

1. $\mathrm{L}(s^\infty)$ *must be equal to one of* $\{\sqrt{N}, \sqrt{N}-1, N-\sqrt{N}, N-\sqrt{N}+1, N-1, N\}$*;*
2. $\mathrm{SC}_k(s^\infty) \geq \sqrt{N} - 1$*, if* $k < \min\{\mathrm{WH}(s^N), N - \mathrm{WH}(s^N)\}$.

Proof: Since q is a primitive root of r and r^2 does not divide $q^{r-1} - 1$ by assumptions, q must be a primitive root of r^2 (for proof, see [1]). Thus, by the basic properties of cyclotomic polynomials presented in the proof of Basic Theorem 3.5 we know that the cyclotomic polynomials $Q_r(x)$ and $Q_{r^2}(x)$ are irreducible over $GF(q)$. Again from the properties of cyclotomic polynomials it follows that

$$x^N - 1 = (x-1)Q_r(x)Q_{r^2}(x).$$

Note that $\deg(Q_r(x)) = r - 1$ and $\deg(Q_{r^2}(x)) = r(r-1)$ since q is a common primitive root of r and r^2. Combining these facts and the fact that the minimum polynomial of each sequence of period N over $GF(q)$ divides $x^N - 1$ proves this theorem. □

In general the linear and sphere complexity and the nonlinearity of the cryptographic function can be also controlled, provided that a large prime p is chosen such that $\mathrm{ord}_p(2)$ is large enough, as shown by Basic Theorem 3.5. In this case $\mathrm{ord}_p(2)$ is the lower bound for both the linear and sphere complexity.

In fact, we do not need to use the speciality of primes to control the linear complexity, as it can be computed exactly. In the remaining part of this section, we compute the linear complexity of the sequences. Our results are from [30].

Let g be a primitive root of p^2. Then g is also a primitive root of p. Recall that the generalized cyclotomic classes of order 2 with respect to p^2 are defined by

$$D_0 = (g^2), \quad D_1 = gD_0,$$

where (g^2) denotes the subgroup of Z_{p^2} generated by g^2. By definition the order of g modulo p^2 is $p(p-1)$. Hence,

$$D_0 \cap D_1 = \Phi, \quad D_0 \cup D_1 = Z^*_{p^2},$$

where Φ denotes the empty set.

The generalized cyclotomic binary sequence s^∞ of order 2 with respect to p^2 is defined by

$$s_i = \left\{ \begin{array}{ll} 0, & \text{if } (i \bmod p^2) \in D_0; \\ 1, & \text{if } (i \bmod p^2) \in D_1 \cup R; \end{array} \right\}, \quad i \geq 0,$$

where R is defined to be $R = \{0, p, 2p, \cdots, (p-1)p\}$. It is not hard to see that the output sequences of the prime-square generator are shift versions of the above sequence s^∞. So we need only to compute the linear complexity of this sequence.

To compute the linear complexity of the sequence, we need a number of lemmas.

Lemma 7.4 *1.* $\mathrm{ord}_{p^2}(g) = p(p-1)$.

2. D_0 *is a subgroup of* $Z^*_{p^2}$ *with* $|D_0| = p(p-1)/2$.

3. $aD_0 = \begin{cases} D_0, & a \in D_0; \\ D_1, & a \in D_1, \end{cases} \quad aD_1 = \begin{cases} D_1, & a \in D_0; \\ D_0, & a \in D_1. \end{cases}$

Proof: The first two parts follows from the definitions of g and D_0. We need only to prove the third conclusion.

If $a \in D_0$, by definition there is an integer s such that $a = g^{2s}$. It follows that

$$\begin{aligned} aD_0 &= \{g^{2s+2t} : t = 0, 1, \cdots, p(p-1)-1\} = D_0; \\ aD_1 &= \{g^{2s+2t+1} : t = 0, 1, \cdots, p(p-1)-1\} = D_1. \end{aligned}$$

The remaining part can be similarly proved. □

Let m be the order of 2 modulo p^2, where p is an odd prime, then there is a p^2th primitive root of unity over $GF(2^m)$, say θ.

Lemma 7.5 $\sum_{i \in R} \theta^i = 0$.

Proof: Note that

$$(1-\theta^p)\sum_{i \in R} \theta^i = 1 - \theta^{p^2} = 0.$$

Since θ is a p^2th primitive of unity, $\theta^p \neq 1$. The conclusion then follows. □

To compute the linear complexity and minimal polynomial for our generalized cyclotomic sequences, we define

$$S(x) = \sum_{i \in R} x^i + \sum_{i \in D_1} x^i \in GF(2)[x].$$

To finish our work, we need to calculate $\gcd(x^{p^2} - 1, S(x))$.

Since

$$\left(\sum_{i \in R} + \sum_{i \in D_1} + \sum_{i \in D_0}\right) \theta^i = \sum_{i=0}^{p^2-1} \theta^i = 0,$$

by Lemma 7.5 we obtain

$$S(\theta) = \sum_{i \in D_1} \theta^i = \sum_{i \in D_0} \theta^i. \tag{7.1}$$

Lemma 7.6 $S(\theta) = 0$ *or 1.*

Proof: By (7.1) and Lemma 7.4

$$\begin{aligned} S(\theta)^2 &= \sum_{i \in D_1} \theta^{2i} = \sum_{j \in 2D_1} \theta^j \\ &= \begin{cases} \sum_{j \in D_1} \theta^j = S(\theta), & \text{if } 2 \in D_0 \\ \sum_{j \in D_0} \theta^j = S(\theta), & \text{if } 2 \in D_1. \end{cases} \end{aligned}$$

Thus, $S(\theta)(S(\theta)-1)=0$. □

We now calculate $S(\theta)$. To this end, we use generalized cyclotomic numbers of order 2 with respect to p^2, which are defined to be

$$(i,j)=|D_i\cap(D_j+1)|,\quad i,j=0,1.$$

Thus, there are four cyclotomic numbers and some of them may be equal. The values for the four constants are given before, but we shall only use some facts about these cyclotomic numbers without using their exact values.

Lemma 7.7 *$-1\in D_0$ if and only if $p\equiv 1 \pmod 4$.*

Proof: If $p=4t+1$ for some integer t, then $g^{p-1}=g^{4t}=1$. It follows that $(g^{2t}-1)(g^{2t}+1)=g^{4t}-1=0$. Because g is a primitive root of p, $-1=g^{2t}\in D_0$.

If $p=4t+3$ for some integer t, then $g^{p-1}=g^{4t+2}=1$. It follows that $(g^{2t+1}-1)(g^{2t+1}+1)=g^{4t+2}-1=0$. Because g is a primitive root of p, $-1=g^{2t+1}\in D_1$. □

Lemma 7.8 *1. If $-1\in D_0$, then $|D_1\cap(D_0+bp)|$ is a constant when b ranges over $\{1,2,\cdots,p-1\}$.*

2. If $-1\in D_1$, then $|D_1\cap(D_1+bp)|$ is a constant when b ranges over $\{1,2,\cdots,p-1\}$.

Proof: We first consider the case $-1\in D_0$. Define $v=|D_1\cap(D_0+p)|$. For any $b\in\{1,2,\cdots,p-1\}$, there is an integer b' with $1\le b'\le p-1$ such that $bb'\equiv 1 \pmod p$.

For those $b\in\{1,2,\cdots,p-1\}$ with their $b'\in D_0$, by Lemma 7.4

$$\begin{aligned}|D_1\cap(D_0+bp)| &= |b'D_1\cap(b'D_0+b'bp)|\\ &= |D_1\cap(D_0+p)|=v.\end{aligned}$$

For those $b\in\{1,2,\cdots,p-1\}$ with their $b'\in D_1$, by Lemma 7.4

$$\begin{aligned}|D_1\cap(D_0+bp)| &= |-b'D_1\cap(-b'D_0-b'bp)|\\ &= |D_0\cap(D_1-p)|\\ &= |D_1\cap(D_0+p)|=v.\end{aligned}$$

Now we consider the case $-1\in D_1$. For those $b\in\{1,2,\cdots,p-1\}$ with their $b'\in D_0$, by Lemma 7.4

$$\begin{aligned}|D_1\cap(D_1+bp)| &= |b'D_1\cap[b'(D_1+bp)]|\\ &= |b'D_1\cap(b'D_1+bb'p)|\end{aligned}$$

$$
\begin{aligned}
&= |D_1 \cap (D_1 + p)| \\
&= |-D_1 \cap (-D_1 - p)| \\
&= |D_0 \cap (D_0 - p)| \\
&= |D_0 \cap (D_0 + p)|.
\end{aligned}
$$

For those $b \in \{1, 2, \cdots, p-1\}$ with their $b' \in D_1$, by Lemma 7.4

$$
\begin{aligned}
|D_1 \cap (D_1 + bp)| &= |b'D_1 \cap [b'(D_1 + bp)]| \\
&= |b'D_1 \cap (b'D_1 + b'bp)| \\
&= |D_0 \cap (D_0 + p)|.
\end{aligned}
$$

This completes the proof. □

Lemma 7.9 $S(\theta) = \begin{cases} 0, & \textit{if } p \equiv 1 \pmod 4; \\ 1, & \textit{if } p \equiv -1 \pmod 4. \end{cases}$

Proof: Assume that $p \equiv 1 \bmod 4$. By Lemma 7.7 we have $-1 \in D_0$. Then by Lemma 7.4

$$
\begin{aligned}
(i, j) &= |D_i \cap (D_j + 1)| \\
&= |(D_i - 1) \cap D_j)| \\
&= |D_j \cap (D_i - 1)| \\
&= |(-D_j) \cap (-D_i + 1)| \\
&= |D_j \cap (D_i + 1)| \\
&= (j, i).
\end{aligned}
$$

Since $-1 \in D_0$, by Lemma 7.4 we have $-D_1 = D_1$. Let $u = S(\theta)$. By Lemma 7.6, $u = S(\theta) \in \{0, 1\}$ and

$$
\begin{aligned}
u = u^2 &= \left(\sum_{i \in D_1} \theta^i\right)\left(\sum_{j \in D_0} \theta^j\right) \\
&= \left(\sum_{i \in D_1} \theta^{-i}\right)\left(\sum_{j \in D_0} \theta^j\right) \\
&= \sum_{i \in D_1} \sum_{j \in D_0} \theta^{j-i} \\
&= \sum_{a=1}^{p^2-1} |D_1 \cap (D_0 + a)| \theta^a
\end{aligned}
$$

$$\begin{aligned}
&= \sum_{b=1}^{p-1} |D_1 \cap (D_0 + bp)|\theta^{pb} + \\
&\quad \sum_{a \in D_0} |D_1 \cap (D_0 + a)|\theta^a + \\
&\quad \sum_{a \in D_1} |D_1 \cap (D_0 + a)|\theta^a \\
&= v \sum_{b=1}^{p-1} \theta^{pb} \quad \text{(By Lemma 7.8)} + \\
&\quad \sum_{a \in D_0} |a^{-1}D_1 \cap (a^{-1}D_0 + 1)|\theta^a + \\
&\quad \sum_{a \in D_1} |a^{-1}D_1 \cap (a^{-1}D_0 + 1)|\theta^a \\
&= -(p-1)v \quad \text{(By the fact that } \textstyle\sum_{b=1}^{p-1} \theta^{pb} = -1) + \\
&\quad \sum_{a \in D_0} |D_1 \cap (D_0 + 1)|\theta^a + \\
&\quad \sum_{a \in D_1} |D_0 \cap (D_1 + 1)|\theta^a \\
&= (1,0) \sum_{a \in D_0} \theta^a + (0,1) \sum_{a \in D_1} \theta^a \\
&= [(1,0) + (0,1)]u \qquad \text{(By (7.1))},
\end{aligned}$$

where v is defined to be $|D_1 \cap (D_0 + p)|$. Since we have proved that $(0,1) = (1,0)$, we have $u = [(1,0) + (0,1)]u = 0$. This completes the proof of the first part.

Assume now that $-1 \in D_1$, then $p \equiv 3 \pmod 4$. We have similarly

$$\begin{aligned}
u = u^2 &= \left(\sum_{i \in D_1} \theta^i\right)\left(\sum_{j \in D_0} \theta^j\right) \\
&= \left(\sum_{i \in D_1} \theta^i\right)\left(\sum_{j \in D_1} \theta^{-j}\right) \quad \text{(By the fact } -D_1 = D_0) \\
&= \sum_{i \in D_1} \sum_{j \in D_1} \theta^{i-j} \\
&= |D_1| + \sum_{a=1}^{p^2-1} |D_1 \cap (D_1 + a)|\theta^a \\
&= |D_1| + \sum_{b=1}^{p-1} |D_1 \cap (D_1 + bp)|\theta^{pb} +
\end{aligned}$$

$$
\begin{aligned}
& \sum_{a\in D_0} |D_1 \cap (D_1+a)|\theta^a + \\
& \sum_{a\in D_1} |D_1 \cap (D_1+a)|\theta^a \\
= \;& p(p-1)/2 + v\sum_{b=1}^{p-1}\theta^{pb} \quad \text{(By Lemma 7.8)} + \\
& (1,1)\sum_{a\in D_0}\theta^a + (0,0)\sum_{a\in D_1}\theta^a \\
= \;& 1 + [(1,1)+(0,0)]u,
\end{aligned}
$$

where v is defined to be $|D_1 \cap (D_1+p)|$. Hence, $u \neq 0$, and therefore $u = 1$. □

In what follows we need cyclotomic classes of order 2 with respect to p. Let $g_1 = (g \bmod p)$. Then g_1 is a primitive root of p. The cyclotomic classes D_0' and D_1' of order 2 with respect to p are defined by

$$D_0' = (g_1^2), \quad D_1' = g_1 D_0',$$

where (g_1^2) denotes the subgroup generated by g_1^2 of Z_p^*. It is easily seen that

$$D_0' \cup D_1' = Z_p^*, \quad D_0' \cap D_1' = \Phi,$$

where Φ denotes the empty set.

Let $\theta_1 = \theta^p$. Then θ_1 is a pth primitive root of unity. Define

$$S'(x) = \sum_{i\in D_1'} x^i.$$

Lemma 7.10 *$S'(\theta_1) \in \{0,1\}$ if and only if $2 \in D_0'$ if and only if $p \equiv \pm 1 \pmod 8$.*

Proof: It is well-known that 2 is a quadratic residue if and only if $p \equiv \pm 1 \pmod 8$. This can be proved by the Law of Quadratic Reciprocity.

Similar to Lemma 7.4, we can easily prove the following:

$$aD_0' = \begin{cases} D_0', & a \in D_0'; \\ D_1', & a \in D_1', \end{cases} \qquad aD_1' = \begin{cases} D_1', & a \in D_0'; \\ D_0', & a \in D_1'. \end{cases} \tag{7.2}$$

Note that

$$\sum_{j\in D_1'} \theta_1^j + \sum_{j\in D_0'} \theta_1^j + 1 = \sum_{i=0}^{p-1} \theta_1^i = 0. \tag{7.3}$$

If $2 \in D_0'$, then

$$\begin{aligned} S'(\theta_1)^2 &= \sum_{i \in D_1'} \theta_1^{2i} \\ &= \sum_{i \in 2D_1'} \theta_1^{i} \\ &= \sum_{i \in D_1'} \theta_1^{i} \\ &= S'(\theta_1). \end{aligned}$$

Hence, $S'(\theta_1) \in \{0, 1\}$.

If $2 \in D_1'$, then by (7.2) and (7.3)

$$\begin{aligned} S'(\theta_1)^2 &= \sum_{i \in D_1'} \theta_1^{2i} \\ &= \sum_{i \in 2D_1'} \theta_1^{i} \\ &= \sum_{i \in D_0'} \theta_1^{i} \\ &= S'(\theta_1) + 1. \end{aligned}$$

Hence, $S'(\theta_1) \notin \{0, 1\}$. □

Lemma 7.11

$$S(\theta) = \begin{cases} \frac{p+1}{2} \bmod 2, & \textit{if } a = 0; \\ S'(\theta_1) + 1, & \textit{if } a = a_1 p,\ a_1 \in D_0,\ 1 \le a_1 \le p-1; \\ S'(\theta_1), & \textit{if } a = a_1 p,\ a_1 \in D_1,\ 1 \le a_1 \le p-1; \\ S(\theta), & \textit{if } a \in Z_{p^2}^*. \end{cases}$$

Proof: By definition

$$S(1) = p + p(p-1)/2 \bmod 2 = \frac{p+1}{2} \bmod 2.$$

Since every $a \in R \setminus \{0\}$ is of the form $a_1 p$ with some a_1, where $1 \le a_1 \le p-1$, we have

$$\begin{aligned} S(\theta^a) &= \sum_{i \in R} \theta^{a_1 p i} + \sum_{i \in D_1} \theta^{a_1 p i} \\ &= p + \sum_{i \in a_1 D_1} \theta^{p i} \end{aligned}$$

$$
\begin{aligned}
&= \begin{cases} 1+\sum_{i\in D_1}\theta^{pi}, & a_1 \in D_0; \\ 1+\sum_{i\in D_0}\theta^{pi}, & a_1 \in D_1 \end{cases} \\
&= \begin{cases} 1+p\sum_{i\in D_1'}\theta_1^{i}, & a_1 \in D_0; \\ 1+p\sum_{i\in D_0'}\theta_1^{i}, & a_1 \in D_1 \end{cases} \\
&= \begin{cases} 1+S'(\theta_1), & a_1 \in D_0; \\ S'(\theta_1), & a_1 \in D_1, \end{cases}
\end{aligned}
$$

where we have made use of the following facts:

$$D_1 \bmod p = D_1', \quad D_0 \bmod p = D_0'$$

and each element of D_i' is obtained p times.

If $a \in Z_{p^2}^*$, we have $aR = R$ and

$$
\begin{aligned}
S(\theta^a) &= \sum_{i\in R}\theta^{ai} + \sum_{i\in D_1}\theta^{ai} \\
&= \sum_{i\in aR}\theta^{i} + \sum_{i\in aD_1}\theta^{i} \\
&= \sum_{i\in R}\theta^{i} + \sum_{i\in aD_1}\theta^{i} \\
&= \begin{cases} \sum_{j\in D_1}\theta^{j}, & a \in D_0; \\ \sum_{j\in D_0}\theta^{j}, & a \in D_1 \end{cases} \\
&= S(\theta),
\end{aligned}
$$

where we have made use of Lemma 7.5. □

Lemma 7.12 $2 \in D_i$ *if and only if* $2 \in D_i'$.

Proof: Obviously, $2 \in D_i$ implies that $2 \in D_i'$. Suppose that $2 \in D_0'$ and $2 \notin D_0$. Then $2 \in D_1$. It follows that $2 \in D_1'$. It then follows that

$$2 \in D_0' \cap D_1' = \Phi.$$

This is impossible. Thus, $2 \in D_0'$ implies that $2 \in D_0$. Similarly, $2 \in D_1'$ implies that $2 \in D_1$. □

For $i = 1$ and 2 define

$$
\begin{aligned}
d_i'(x) &= \prod_{a\in D_i'}(x-\theta_1^a), \\
d_i(x) &= \prod_{a\in D_i}(x-\theta^a).
\end{aligned}
$$

It is easily seen that

$$x^p - 1 = (x-1)d'_0(x)d'_1(x). \tag{7.4}$$

and

$$x^{p^2} - 1 = (x-1)d'_0(x)d'_1(x)d_0(x)d_1(x). \tag{7.5}$$

Lemma 7.13 *$d'_i(x) \in GF(2)[x]$ or $d_i(x) \in GF(2)[x]$ if and only if $p \equiv \pm 1 \pmod 8$.*

Proof: By Lemmas 7.10 and 7.12, $2 \in D'_0$ if and only if $2 \in D_0$ if and only if $p \equiv \pm 1 \pmod 8$. If $p \equiv \pm 1 \pmod 8$, then $2 \in D_0 \cap D'_0$. Hence, $2D_a = D_a$ and

$$\begin{aligned} d_a(x)^2 &= \prod_{j \in D_a} (x^2 - \theta^{2j}) \\ &= \prod_{j \in 2D_a} (x^2 - \theta^j) \\ &= \prod_{i \in D_a} (x^2 - \theta^j) \\ &= d_a(x^2). \end{aligned}$$

It follows that $d_a(x) \in GF(2)[x]$.

If $p \equiv \pm 3 \pmod 8$, then $2 \in D_1 \cap D'_1$. Hence, $2D_a = D_{(a+1) \bmod 2}$ and

$$\begin{aligned} d_a(x)^2 &= \prod_{j \in D_a} (x^2 - \theta^{2j}) \\ &= \prod_{j \in 2D_a} (x^2 - \theta^j) \\ &= \prod_{j \in D_{a+1 \bmod 2}} (x^2 - \theta^j) \\ &= d_{(a+1) \bmod 2}(x^2) \\ &\neq d_a(x^2). \end{aligned}$$

It follows that $d_a(x) \notin GF(2)[x]$. □

Note that $d'_i(x)$ and $d_i(x)$ depend on the choice of θ, but any new choice leads to at most an exchange between $d'_0(x)$ and $d'_1(x)$ and between $d_0(x)$ and $d_1(x)$.

Theorem 7.14 *Let L and $m(x)$ denote the linear complexity and the minimal polynomial of our generalized cyclotomic sequence.*

1. *If* $p \equiv 1 \pmod 8$, *then* $L = (p+1)/2$ *and*

$$m(x) = \begin{cases} (x-1)d_0'(x), & S'(\theta_1) = 0; \\ (x-1)d_1'(x), & S'(\theta_1) = 1. \end{cases}$$

2. *If* $p \equiv -1 \pmod 8$, *then* $L = p^2 - (p+1)/2$ *and*

$$m(x) = \begin{cases} d_1'(x)d_0(x)d_1(x), & S'(\theta_1) = 0; \\ d_0'(x)d_0(x)d_1(x), & S'(\theta_1) = 1. \end{cases}$$

3. *If* $p \equiv 3 \pmod 8$, *then* $L = p^2 - 1$ *and*

$$m(x) = \frac{x^{p^2} - 1}{x - 1}.$$

4. *If* $p \equiv -3 \pmod 8$, *then* $L = p$ *and*

$$m(x) = x^p - 1.$$

Proof: Similar to the proof of Lemma 7.12, we can prove that

$$a_1 \in D_i' \text{ if and only if } a_1 \in D_i, \quad i = 0, 1, \tag{7.6}$$

where $1 \leq a_1 \leq p-1$.

If $p \equiv 1 \pmod 8$, then $2 \in D_0 \cap D_0'$ by Lemmas 7.10 and 7.12. By Lemma 7.13, $d_0'(x) \in GF(2)[x]$ and $d_1'(x) \in GF(2)[x]$. By Lemma 7.10, $S'(\theta_1)$ or $S'(\theta_1)+1$ is zero. Then by Lemmas 7.9 and 7.11 as well as (7.6), we obtain

$$S(\theta^a) = \begin{cases} 1, & a = 0; \\ S'(\theta_1) + 1, & a = a_1 p, a_1 \in D_0'; \\ S'(\theta_1), & a = a_1 p, a_1 \in D_1'; \\ 0, & a \in Z_{p^2}^*. \end{cases}$$

It follows that

$$\gcd(x^{p^2} - 1, S(x)) = \begin{cases} d_1'(x)d_0(x)d_1(x), & S'(\theta_1) = 0; \\ d_0'(x)d_0(x)d_1(x), & S'(\theta_1) = 1. \end{cases}$$

Hence,

$$\begin{aligned} m(x) &= \frac{x^{p^2} - 1}{\gcd(x^{p^2} - 1, S(x))} \\ &= \begin{cases} (x-1)d_0'(x), & S'(\theta_1) = 0; \\ (x-1)d_1'(x), & S'(\theta_1) = 1 \end{cases} \end{aligned}$$

and

$$L = \deg(m(x)) = \frac{p+1}{2}.$$

If $p \equiv -1 \pmod 8$, then $2 \in D_0 \cap D_0'$ by Lemmas 7.10 and 7.12. By Lemma 7.13, $d_0'(x) \in GF(2)[x]$ and $d_1'(x) \in GF(2)[x]$. By Lemma 7.10, $S'(\theta_1)$ or $S'(\theta_1)+1$ is zero. Then by Lemmas 7.9 and 7.11 as well as (7.6), we obtain

$$S(\theta^a) = \begin{cases} 0, & a = 0; \\ S'(\theta_1)+1, & a = a_1 p, a_1 \in D_0'; \\ S'(\theta_1), & a = a_1 p, a_1 \in D_1'; \\ 1, & a \in Z_{p^2}^*. \end{cases}$$

It follows that

$$\gcd(x^{p^2}-1, S(x)) = \begin{cases} (x-1)d_1'(x), & S'(\theta_1) = 0; \\ (x-1)d_0'(x), & S'(\theta_1) = 1. \end{cases}$$

Hence,

$$\begin{aligned} m(x) &= \frac{x^{p^2}-1}{\gcd(x^{p^2}-1, S(x))} \\ &= \begin{cases} d_0'(x)d_0(x)d_1(x), & S'(\theta_1) = 0; \\ d_1'(x)d_0(x)d_1(x), & S'(\theta_1) = 1 \end{cases} \end{aligned}$$

and

$$L = \deg(m(x)) = p^2 - \frac{p+1}{2}.$$

If $p \equiv 3 \pmod 8$, then $2 \notin D_0' \cup D_0$ by Lemmas 7.10 and 7.12. Again by Lemma 7.10, both $S'(\theta_1)$ and $S'(\theta_1)+1$ are nonzero. Then by Lemmas 7.11 and 7.9, we obtain

$$S(\theta^a) = \begin{cases} 0, & a = 0; \\ \neq 0, & a = a_1 p, a_1 \in D_0' \cup D_1'; \\ 1, & a \in Z_{p^2}^*. \end{cases}$$

It follows that

$$\gcd(x^{p^2}-1, S(x)) = x-1.$$

Hence,

$$m(x) = \frac{x^{p^2}-1}{\gcd(x^{p^2}-1, S(x))} = \frac{x^{p^2}-1}{x-1}$$

and

$$L = \deg(m(x)) = p^2 - 1.$$

If $p \equiv -3 \pmod 8$, then $2 \notin D_0' \cup D_0$ by Lemmas 7.10 and 7.12. Again by Lemma 7.10, both $S'(\theta_1)$ and $S'(\theta_1) + 1$ are nonzero. Then by Lemmas 7.11 and 7.9, we obtain

$$S(\theta^a) = \begin{cases} 1, & a = 0; \\ \neq 0, & a = a_1 p, a_1 \in D_0' \cup D_1'; \\ 0, & a \in Z_{p^2}^*. \end{cases}$$

It follows that

$$\gcd(x^{p^2} - 1, S(x)) = d_0(x) d_1(x).$$

Hence,

$$m(x) = \frac{x^{p^2} - 1}{\gcd(x^{p^2} - 1, S(x))} = x^p - 1.$$

and

$$L = \deg(m(x)) = p.$$

$\square$

Chapter 8

Related Problems

As seen in some of the forgoing chapters, cyclotomic generators are cryptographically attractive from at least the seven aspects described in Section 3.3. Those generators are related not only to cyclotomic numbers, but also to a large number of other number-theoretic problems. There are also a number of cryptographic ideas behind those generators which seem interesting. This chapter points out some number-theoretic problems relating to and some cryptographic ideas behind those cyclotomic generators. The interaction between number theory and stream ciphers is also investigated.

The main contributions of this chapter are the following:

1. The analysis of the cryptographic ideas behind the cyclotomic generators which are also useful for the design of other generators, including Theorem 8.1;

2. the relations between stream-ciphering problems and some number-theoretic problems, especially the introduction of sexes to twin primes and the analysis of twin primes, including Theorems 8.2 and 8.3 (Section 8.2); and

3. a number of lower bounds on the linear and sphere complexity of sequences over the residue class rings Z_m, as described by the theorems and corollaries of Section 8.4.

This chapter is mainly based on the paper [24].

8.1 Behind the Cyclotomic Generators

There are several cryptographic ideas behind the construction of these cyclotomic generators. The first one is the order of choosing the design parameters for the generator. Contrary to traditional approaches, we first control the period of the output sequence. This will automatically ensure the linear complexity and its stability aspect only with the condition that the sequence is not a constant sequence. Then we choose the cryptographic function for other purposes. This approach is intended to avoid unnecessary trade-offs.

The second cryptographic idea behind the design and analysis of cyclotomic generators is the idea of introducing good "partners", in order to get a stable system. In particular, we search for pairs consisting of a period and finite field so that it is easy to control the linear complexity and its stability for those sequences over those fields with corresponding partner periods. This has been shown clearly by the theorems and corollaries concerning linear and sphere complexity. We say that such pairs work in harmony with respect to the aspects of linear complexity and its stability. For example, some Mersenne and Fermat primes are not good partners of the field $GF(2)$, since it is difficult to control the linear complexity and its stability for binary sequences with period of some Fermat and Mersenne primes. This is because the order of 2 modulo these primes is quite small. From Basic Theorem 3.5 and its corollaries it is rational to use $\mathrm{ord}_N(q)$ as a measure on the partnership between a positive integer N and an integer q with respect to the linear and sphere complexity aspect when designing sequences of period N over $GF(q)$, where $\gcd(N, q) = 1$. We say that q and N are the *best partners* when q is a primitive root modulo N.

Another kind of partnership is to find an integer r which is a power of a prime such that $\min\{\mathrm{ord}_{p_1}(r), \cdots, \mathrm{ord}_{p_h}(r)\}$ is large enough when designing sequences of period $N = p_1 \cdots p_h$ over $GF(r)$, where $p_1, \cdots, p_h$ are distinct primes (see Basic Theorem 3.5). We say that r is a *best common partner* of $p_1, \cdots, p_h$ if r is a common primitive root of these primes.

The third cryptographic idea is to use some techniques of ensuring "good + bad = good". With a simple argument each cryptographic function employed in the generators described in the foregoing chapters can be expressed as

$$F(x) = H(G(x)),$$

where $G(x)$ is a mapping from Z_N to U which is a subgroup of the group $(Z_N^*, \cdot)$ with order d, and $H(x)$ a mapping from U to Z_d. The nonlinearity of $G(x)$ with respect to $(Z_N, +)$ and $(U, \cdot)$ is determined mainly by the (generalized) cyclotomic numbers of order d, which usually have an ideal stability; while the function $H(x)$ is almost linear (or with good linearity) with respect to $(U, \cdot)$ and $(Z_d, +)$. Thus, it is clear that one cryptographic idea behind the cyclotomic generators is

"GOOD + BAD = GOOD".

The fourth cryptographic idea is to use cryptographic functions $f(x)$ from an abelian group $(G_1, +)$ to another abelian group $(G_2, +)$ such that $|G_2|$ does not divide $|G_1|$. We say that such a function is *linearly non-approximatable*, since there is no linear mapping other than the zero constant function from $(G_1, +)$ to $(G_2, +)$. This technique makes any linear approximation attack with respect to the two operations futile. It should be pointed out that the nonlinearity definition based on the minimum correlation (measured in probability of agreement, or distance [50] between a function and all affine functions) is not rational in many cases. This definition makes no sense for the above cryptographic functions.

The fifth cryptographic idea is to make use of the relativity about nonlinearity and linearity. It is a common fact that the nonlinearity and linearity are relative to the operations considered, and that both linear components and nonlinear components should be employed in many cipher systems. To find out some cryptographic functions with good nonlinearity with respect to some operations, one may try to find some linear cryptographic function with respect to some other operations and use them in the context of the former operations. This is to say that bad things in one sense may be good ones in another sense, and one way to get goodness is to use badness in a proper way and proper context. To illustrate this philosophy, we first take the corresponding function $G(x) = x^{(p-1)/d} \bmod p$ used to construct the cyclotomic generator of order $2k$. Then $G(x)$ is linear with respect to $(Z_p^*, \cdot)$ and $(U, \cdot)$, where U is the multiplicative subgroup of Z_p^* with order d. But $G(x)$ has ideal nonlinearity with respect to $(Z_p, +)$ and $(U, +)$ if we define $G(0)$ to be any fixed element of U. And we use $G(x)$ in the context of the later pair of operations exactly. The same idea has been used for other generators.

To illustrate the relativity of linearity and nonlinearity, we prove the following theorem.

Theorem 8.1 *For every nonzero linear function $L(x)$ from $F = GF(q^m)$ to $K = GF(q)$ with respect to the additions of the two fields, its nonlinearity with respect to $(F^*, \times)$ and $(K, +)$ is optimal, as described by*

$$p(L(x) - L(x/\alpha) = b) = p(L(x(1-\alpha^{-1})) = b) = \begin{cases} \frac{q^{m-1}-1}{q^m-1}, & \text{if } b = 0, \\ \frac{q^{m-1}}{q^m-1}, & \text{if } b \neq 0, \end{cases}$$

which holds for $\alpha \in F$ with $\alpha \neq 1$, and each $b \in K$.

Proof: For any nontrivial linear mapping $L(x)$ from F to K the kernel $L^{-1}(0)$ is an abelian subgroup of $(F, +)$ and thus $L(x)$ takes on each element of K equally likely, that is, q^{m-1} times. Confining the linear mapping on $F^* = F \setminus \{0\}$ and using the above fact proves the theorem. □

This theorem clearly shows the cryptographic importance of linear functions from $F = GF(q^m)$ to $K = GF(q)$ with respect to the additions of the two fields, especially the trace functions. Actually all the cryptographic functions for the cyclotomic generators are composition functions of linear functions with respect to different pairs of operations. It may be possible to employ this idea to design some block ciphers.

8.2 Some Related Number-Theoretic Problems

In this thesis at least two bridges between some number-theoretic problems and the design and analysis of the natural sequence generator have been established. These bridges may play an important role in the interactions between number theory and stream ciphers, and particularly in the design and analysis of cyclotomic generators.

The first one is the bridge supported by Basic Theorem 3.5 and the theorems and corollaries concerning linear and sphere complexity. The reason for calling it a bridge was already made clear in Section 3.5. Obviously, there are more results which can be derived from Basic Theorem 3.5. If we go across this bridge from the stream cipher side, we shall encounter at least the following two basic number-theoretic problems when we are designing sequences of period N over $GF(q)$ with ideal linear and sphere complexity (see Basic Theorem 3.5).

Basic Problem 1 *Find large positive integers N and positive integers q which are powers of primes such that*

1. $\gcd(N, q) = 1$;
2. *For any factor n of N, $ord_n(q) = \phi(n)$.*

Basic Problem 2 *Find large positive integers N and positive integers q which are powers of primes such that*

1. $\gcd(N, q) = 1$;
2. *N has few factors;*
3. *For any factor n of N, $ord_n(q)$, a factor of $\phi(n)$, is as large as possible.*

Attacking these two basic problems and many of their subproblems and variants will involve many number-theoretic problems. Among them are the searching for special primes, such as Fermat primes, Mersenne primes, Stern primes, prime repunits, and twin primes, the distribution of special primes, primitivity testing, the distribution of primitive roots. These facts have already been shown clearly by all the theorems concerning the linear and sphere complexity proved in this thesis. We need special primes for these generators, and have to know whether they exist. And if they exist, we have to know how to find large special primes.

We need not only twin primes for our twin-prime generator, but also special twin primes. The most interesting twin primes are those having a common primitive root 2 (the best common partner of the twins). However, some twins have, others don't. Problems as to which twins have the common primitive root 2 and how to find them are naturally important to the design of the twin-prime generators. By introducing sexes to twin we can solve one cryptographic problem for the twin-prime generator.

Let $(p, p+2)$ be a pair of twin primes and $p = \Xi(p) \bmod 4$, where $\Xi(p) = \pm 1$. Then we call $\Xi(p)$ the *sex characteristic* of the twins. If the twins $(p, p+2) = (4t-1, 4t+1)$ for some t, then we say that the twins have the same sex; otherwise, we say that they have different sexes.

In the above definitions, we say that twin primes $(p, p+2)$ have the same sex, because in the expression of the form $4u \pm 1$, the u's for both p and $p+2$ are the same, and have therefore the same parity, if $p = 4t-1$. If $p = 4t+1$, then $p+2 = 4(t+1)-1$ and t and $t+1$ have different parities. That is why we call them twins with different sexes. This discussion has also proved the following property of twins.

Theorem 8.2 (The Sex Principle of Twins) *If the smaller of the twins has sex characteristic* -1*, then the twins have the same sex; otherwise, they have different sexes.*

Theorem 8.3 *If* p *and* $p+2$ *have the same sex, then it is possible for them to have the common primitive root* 2 *(a common best partner); otherwise, they never have.*

Proof: With the help of the Law of Quadratic Reciprocity it is easy to see that a necessary condition for 2 to be a primitive root of a prime N is $N = 8k \pm 3$ for some k. Combining this fact and the definition of sexes proves the theorem. $\square$

Thus, the cryptographic importance of classifying twin primes into two classes according to their sexes for the design of the twin-prime generator clearly follows from the above theorem. Before searching for twins with the same sex, we have to know the distribution of the two classes. Solving this problem and searching for twin primes with the same sex are important design problems for the twin-prime generator.

The second bridge we set up when we design these cyclotomic generators is supported by the relations between the difference parameters of the cryptographic function $f(x)$, and the nonlinearity measure, the autocorrelation functions, mutual information, and two-bit pattern distributions, as described by the formulae in Section 3.3.

At one side of the second bridge are the autocorrelation property, the two-digit pattern distribution property of the output sequences, and the difference property and nonlinearity of the cryptographic functions of the natural sequence generator; while at the other side are cyclotomy-related problems and the Riemann Hypothesis for Curves over Finite Fields, which was proven to be true by Weil in 1948 [61]. We call it another bridge between number theory and stream ciphers because it makes a clear connection between the above set of cryptographic problems and the set of number-theoretic problems.

Related to cyclotomic numbers and their stability are the quadratic partition of primes and of some integers, the theory of quadratic forms, genus theory, class field theory, residue difference sets, group character, and character sums. In what follows we give a brief explanation concerning why these problems are related to the design of the cyclotomic generators. It is a pity that we cannot show here how some of these problems are related to the design and analysis of the cyclotomic generators since doing so must involve quite many number-theoretic concepts. However, pointing out some relations between design problems of cyclotomic generators and many number-theoretic problems might be helpful for those who are interested in these generators. We begin with the stability of cyclotomic numbers.

Although the formulae for the cyclotomic numbers of order 4 described in Section 4.6 show a roughly ideal stability of those cyclotomic constants, the actual stability depends on the quadratic partition $p = x^2 + 4y^2$ and $x = 1 \bmod 4$. If there is a big difference between $|x|$ and $|y|$, then there is a considerable difference between the cyclotomic numbers. And it is not difficult to give examples to show there do exist primes such that there is a considerable difference between x^2 and y^2 in their quadratic partitions. Thus, choosing large special primes to ensure better stability of cyclotomic numbers is cryptographically necessary.

There are two approaches to do this. The first approach is to use traditional methods to get first large primes. Then use some special algorithm to get the partition and see whether it results in an ideal stability of cyclotomic numbers. This special algorithm should be related to classical problems about quadratic partitions such as the number of solutions of such a quadratic partition. Another approach is to consider directly for given n the set

$$\{x^2 + ny^2 : x, y \in Z\}.$$

If there are infinitely many primes in the set, we can search for the primes which give an ideal stability of cyclotomic numbers. With this approach the first problem we have to solve is whether there are infinitely many primes in the above set. The Dirichlet theorem about primes in arithmetic progression does not help here. However, with some results from class field theory we can get a positive answer for each n [10].

If one has a quick look at each set of cyclotomic formulae known today, one will find that they have the same form. Behind this uniformity of all cyclotomic numbers is the Riemann Hypothesis for Curves over Finite Fields, which can be described as follows.

Proposition 8.4 (Weil) *Suppose that $F(x, y)$ is a polynomial of total degree d, with coefficients in $GF(q)$ and with N zeros $(x, y) \in GF(q) \times GF(q)$. Suppose that $F(x, y)$ is absolutely irreducible, i.e., irreducible not only over $GF(q)$, but also over every algebraic extension thereof. Then*

$$|N - q| \leq 2g\sqrt{q} + c_1(d),$$

where g is the genus of the curve $F(x, y) = 0$ and $c_1(d)$ is a constant depending on d.

This theorem, proven by Weil [61], not only indicates the uniformity of the form of cyclotomic formulae, but also can be used to set up bounds for pattern distributions in the output sequences of the cyclotomic generators. With the Weil Theorem one can see some cryptographic meanings of the genus of curves. For details about the Weil Theorem, the reader is referred to [61]. Since cyclotomic numbers are related to characters and character sums [14, 15, 42], the cyclotomic generators are naturally related to character and character sums. In fact many of the cryptographic functions $f(x)$ employed here are group characters.

One part of the design and analysis of these cyclotomic generators is to solve the following basic problem at the number-theory side of the second bridge.

Basic Problem 3 *Let N be a positive integer, and Z_N denote the residue ring modulo N. Find partitions $\{C_0, C_1\}$ of Z_N, i.e.,*

$$C_0 \cap C_1 = \Phi, \; C_0 \cup C_1 = Z_N,$$

such that

$$\begin{aligned} &|C_0| \approx N/2, \; |C_1| \approx N/2, \\ &|C_i \cap (C_j + r)| \approx N/4 \end{aligned} \tag{8.1}$$

for each nonzero r of Z_N, and each $i, j \in Z_2 = \{0, 1\}$. Here and hereafter $A \approx B$ means that $A = B \pm O(B^{\frac{1}{2}})$.

For our cryptographic purpose we hope that $A \approx B$ means that A is as near to B as possible. But for the flexibility of the solutions of the above basic problem, we give such a definition of $A \approx B$.

The characteristic function of a partition $\{C_0, C_1\}$ of Z_N satisfying (8.1) ensures that it has an ideal difference property, and therefore an ideal property for the other six aspects described in Section 3.3. But to ensure an ideal pattern distribution property for various pattern lengths means to solve the following more general basic problem.

Basic Problem 4 *Let* $C = \{C_0, C_1\}$ *be a partition of* Z_N, $r_0, r_1, \cdots, r_{s-1}$ *be* s *pairwisely distinct elements of* Z_N, *and*

$$D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = \cap_{k=0}^{s-1}(C_{i_k} + r_k),$$
$$d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) = |\cap_{k=0}^{s-1}(C_{i_k} + r_k)|,$$

where $i_0, \cdots, i_{s-1} \in Z_2$. *Then* $\{C_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) : i_0, \cdots, i_{s-1} \in Z_2\}$ *forms a partition of* Z_N.

Find partitions $\{C_0, C_1\}$ *of* Z_N *such that*

$$d_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1}) \approx \frac{N}{2^s} \tag{8.2}$$

holds for each s *with* $1 \leq s \leq \lfloor \log_2 N \rfloor$, *and for each set of pairwisely distinct elements* $r_0, \cdots, r_{s-1}$ *of* Z_N.

The conditions here include those of Basic Problem 3. They ensure that the output sequence of the natural sequence generator employing the characteristic function of the partition has an ideal pattern distribution for each pattern with length t satisfying $1 \leq t \leq \lfloor \log_2 N \rfloor$, and also an ideal mutual information stability $I(i; z_{i_1} z_{i_2} \cdots z_{i_t})$. It is noted that these *higher order difference parameters* $D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1})$ are exactly measures on patterns distributions of binary sequences.

One class of solutions to Basic Problem 3 is those partitions $C = \{C_0, C_1\}$ of Z_N such that C_0 is a residue difference set of Z_N with $|C_1| \approx |C_0|$. However, difference sets exist only for those $N = 3 \bmod 4$. For $N = 1 \bmod 4$ there may exist some partitions $\{C_0, C_1\}$ of Z_N with almost the same difference property. These are partitions $\{C_0, C_1\}$ of Z_N with C_1 being an almost difference set.

For other power residues it is possible to see whether they form almost difference sets by employing the cyclotomic numbers of various orders. In general, it is not easy to find almost difference residue sets. It is noted that difference sets and almost difference sets are employed in the cyclotomic generator of order 2 and twin-prime generator. So they are closely related to cyclotomic generators.

Another solution to Basic Problem 3 is to make use of power residues and cyclotomic numbers generally, as done in some of the foregoing chapters. One advantage of the sets of power residues is that they form multiplicative groups, and this makes the implementation problem of characteristic functions of the corresponding partitions simple. There may exist other solutions to Basic Problem 3, which remains to be investigated.

The condition (8.1) can only give a very rough guarantee for other conditions in (8.2). The cyclotomic numbers and the Weil Theorem seem to indicate that partitions based on power residues give ideal solutions to Basic problem 4. In fact we have already set up bounds of order $N \pm O(\sqrt{N})$ for these higher order difference parameters $D_{i_0 \cdots i_{s-1}}(r_0, \cdots, r_{s-1})$ for $s \geq 3$ based on the Weil Theorem when N is a prime in Section 5.3.4, but the bounds becomes more and more loose with the increase of s though the bounds remain in order $N \pm O(\sqrt{N})$.

We pointed out some number-theoretic problems here, since they are essential to the design and analysis of cyclotomic generators. For other types of generators there may be other related number-theoretic problems, such as the class numbers for imaginary quadratic fields which are related to properties of some number-theoretic generators [11].

8.3 A Summary of Cyclotomic Generators

Since we have controlled the difference property of the cryptographic functions and the linear and sphere complexities of the output sequences of the binary cyclotomic generators, the formulae in Section 3.3 and theorems and corollaries regarding the linear and sphere complexity show that these generators have the following properties:

1. the cryptographic function $f(x)$ has an ideal difference property;
2. the cryptographic function $f(x)$ has an ideal nonlinearity with respect to the additions of Z_N and Z_2;
3. the cryptographic function $f(x)$ has an ideal autocorrelation property;
4. the affine approximation of $f(x)$ with respect to $(Z_N, +)$ and $(Z_2, +)$ makes no sense, since there are only two trivial affine functions from Z_N to Z_2 for odd N;
5. the output sequence has an ideal autocorrelation property;
6. the output sequence has an ideal two-bit pattern distribution property;
7. the output sequence has an ideal linear complexity and its stability;
8. the mutual information $I(i;\ z_i z_{i+t-1})$ has an ideal stability, here z^∞ denotes the output sequence of the natural sequence generator; and
9. the additive stream cipher system with this natural sequence generator as the keystream generator has an ideal density of encryption (resp. decryption) transformations, i.e., the probability of agreement between two encryption (resp. decryption) transformations specified by two keys is approximately 1/2.

In fact we can calculate exact values of measures (such as autocorrelation values, the mutual information) for the above aspects based on the formulae in Section 3.3 if we have formulae for the difference parameters. For example measures for the above aspects for the cyclotomic generator of order 2 can be expressed exactly in terms of N, the modulus for the modulo N ring counter. If we have bounds for the difference parameters, then the use of the formulae in Section 3.3 gives bounds for measures on the above aspects.

In addition, the Weil Theorem and the formulae of cyclotomic numbers seem to indicate that the output sequences of these cyclotomic generators have a roughly ideal distribution property for any pattern with length $1 \leq l \leq \lfloor \log_2 N \rfloor$.

In this thesis we consider only some binary cyclotomic generators. It is possible to extend the results for binary cyclotomic generators to cyclotomic generators over other fields. For the linear and sphere complexity of periodic sequences over other fields we have similar results.

The performances of these cyclotomic generators are basically of the same order, since they are all based on the exponentiation-modulo-N operation. It seems that these generators are not fast, but with a fast exponentiation algorithm it is possible to get a reasonable performance if the moduli are not too large. For the time being moduli having about 64 bits seems enough for the generators. Though the generators are not fast, they might have an ideal security. Thus, trading performance for security might be possible.

8.4 Implementation and Application

All the cyclotomic generators described in this thesis can be implemented in hardware similar to the hardware implementation of the two-prime cyclotomic generator as depicted in Figure 6.1. The performance of such an implementation depends on the development of chip techniques for modular multiplication. It is important to note that the moduli used in these cyclotomic generators are in the range of 80–100 bits. Thus with the current chips for modular multiplication, the binary additive synchronous stream ciphers based on these cyclotomic generators should be able to encrypt and decrypt at 30Kbytes per second. Thus, they can encrypt 10 pages (A4, ASCII) of English text per second. As mentioned in Section 6.2, this performance is certainly reasonable in military and diplomatic communications and in ciphering large data base where the performance of ciphering is not much concern.

8.5 Linear Complexity of Sequences over Z_m

The definition of the linear complexity of sequences over fields was given before. The linear complexity for sequences over rings can be similarly defined as follows.

Let R be a ring with multiplicative identity 1, and $s^N = s_0 s_1 \cdots s_{N-1}$ be a sequence of length N over R, where $s_i \in R$. If s^N satisfies a linear recurrence relation

$$s_i = a_1 s_{i-1} + a_2 s_{i-2} + \cdots + a_l s_{i-l}, \quad i \geq l, \ a_i \in R,$$

then there exists such a shortest linear recurrence relation, and the shortest l is called the *linear complexity* or *linear span* of the sequence and is denoted by $\mathrm{L}(s^N)$. The linear complexity of a finite sequence s^N is defined to be N if s^N does not satisfy such a linear recurrence relation. For semi-infinite sequences the linear complexity is defined to be $+\infty$ if they satisfy no finite linear recurrence relation. For ultimately periodic sequences the linear complexity is finite. If the linear complexity of a sequence over a field is l, then $2l$ successive characters of the sequence can be used to determine a linear recurrence relation of length l satisfied by the sequence with the Berlekamp-Massey algorithm [48], which has complexity $O(l^2)$. Consequently, $2l$ successive characters of the sequence are sufficient to determine the whole sequence. Thus, sequences over fields for additive stream ciphering and for some code-division multiple-access systems should have a large linear complexity.

So far some results about the period of linear recurrence sequences over Z_m are known, but few results about the linear complexity of sequences over Z_m has been achieved. Sequences over fields are easy to construct, and their properties are easy to control. But it looks hard to do so for sequences over residue class rings. In this section we show how to construct sequences over Z_m from those over finite fields Z_p, where p are primes, and how to control their linear complexity with the help of the Chinese Remainder Theorem.

An important result we need is the following theorem whose second part was implied in the work of Reeds and Sloane [57] without giving proofs.

Lemma 8.5 *Let s be a sequence over Z_m, where $m = m_1 m_2 \cdots m_t$ and m_i are pairwise relatively prime, and let*

$$s(i) = s \bmod m_i, \quad i = 1, 2, ..., t,$$

i.e., $s(i)_j = s_j \bmod m_i$ for all possible j.

1. *If s is (ultimately) periodic, then each sequence $s(i)$ must be (ultimately) periodic, and $\mathrm{per}(s) = \mathrm{lcm}\{\mathrm{per}(s(1)), \cdots, \mathrm{per}(s(t))\}$, where $\mathrm{per}(s)$ denotes the least period.*

2. *$\mathrm{L}(s) = \max\{\mathrm{L}(s(1)), \cdots, \mathrm{L}(s(t))\}$, where $\mathrm{L}(s)$ denotes the linear complexity.*

Proof: Let φ be the mapping from Z_m to $Z_{m_1} \times \cdots \times Z_{m_t}$ given by

$$\varphi : x \bmod m \mapsto (x \bmod m_1, ..., x \bmod m_t).$$

By the Chinese remainder theorem φ is an isomorphism.

Assume that s is periodic and $N = \mathrm{per}(s)$. Then $\varphi(x_{N+j}) = \varphi(x_j)$ for all $j \geq 0$. It follows that $s(i)_{N+j} = s(i)_j$ for all $j \geq 0$ and all $i = 1, ..., t$. Thus, each $s(i)$ is periodic and N divides $N' = \mathrm{lcm}\{\mathrm{per}(s(1), \cdots, \mathrm{per}(s(t)\}$. On the other hand, since $s(i)_{N'+j} = s(i)_j$ for all j and i, we have $\varphi(s_{j+N'}) = \varphi(s_j)$ for all j. Since φ is one-to-one, $s_{j+N'} = s_j$ for all j. It follows that N divides N'. Combining the above results gives $N = N'$.

Now we prove part two. Let $l = \mathrm{L}(s)$ and

$$s_j = a_1 s_{j-1} + a_2 s_{j-2} + \cdots + a_l s_{j-l}, \quad j \geq l \tag{8.3}$$

be a shortest linear recurrence relation of s, where $a_i \in Z_m$. Let $\varphi(a_i) = (a(1)_i, \cdots, a(t)_i)$ for $i = 1, ..., t$. Applying the isomorphism φ to Equation (8.3) gives

$$\varphi(s_j) = \varphi(a_1)\varphi(s_{j-1}) + \cdots + \varphi(a_l)\varphi(s_{j-l}), \quad j \geq l$$

from which it follows that for each i with $1 \leq i \leq t$

$$s(i)_j = a(i)_1 s(i)_{j-1} + \cdots + a(i)_l s(i)_{j-l}, \quad j \geq l,$$

where $a(i)_j \in Z_{m_i}$. Thus, $l_i = \mathrm{L}(s(i)) \leq l$.

On the other hand, let $l' = \max\{l_1, ..., l_t\}$. Assume that

$$s(i)_j = a(i)_1 s(i)_{j-1} + \cdots + a(i)_{l_i} s(i)_{j-l_i}, \quad j \geq l_i$$

is a shortest linear recurrence relation the sequence $s(i)$ satisfies. Define $a(i)_j = 0$ for all j with $l_{i+1} \leq j \leq l'$, where $1 \leq i \leq t$. Then

$$\begin{array}{rcl} (s(1)_j, \cdots s(t)_j) & = & (a(1)_1, \cdots a(t)_1)(s(1)_{j-1}, \cdots s(t)_{j-1}) + \cdots + \\ & + & (a(1)_{l'}, \cdots a(t)_{l'})(s(1)_{j-l'}, \cdots s(t)_{j-l'}) \end{array} \tag{8.4}$$

holds for each $j \geq l'$. Let $a_j = \varphi^{-1}(a(1)_j, \cdots, a(t)_j) \in Z_m$. Applying the inverse isomorphism φ^{-1} to Equation (8.4) gives

$$s_j = a_1 s_{j-1} + a_2 s_{j-2} + \cdots + a_{l'} s_{j-l'}, \quad j \geq l'.$$

Thus, $l \leq l'$. Hence $l = l'$. This proves part two. $\square$

From now on in this section we stipulate that the period mentioned does not necessarily refer to the least one, i,e., the period is a multiple of the least one. To set up lower bounds on the linear complexity of sequences over Z_m, we need Basic Theorem 3.5.

A general lower bound for sequences over Z_m is the following.

Theorem 8.6 *Let $N = N_1 \cdots N_r$, $m = p_1 \cdots p_t$, where $N_1, ..., N_r, p_1, ..., p_t$ are pairwise distinct primes. If s^∞ is a nonconstant sequence of period N over Z_m, then*

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_1), ..., \mathrm{ord}_{N_r}(p_1), ..., \mathrm{ord}_{N_1}(p_t), ..., \mathrm{ord}_{N_r}(p_t)\}.$$

Proof: Let $s(i)^\infty = s^\infty \bmod p_i$ for $i = 1, 2, ..., t$. Then $s(i)^\infty$ is a sequence of period N over Z_{p_i}. Since s^∞ is nonconstant, there must exist an integer j such that $s(j)^\infty$ is nonconstant by the Chinese remainder theorem. Since $\gcd(p_j, N) = 1$, by Lemmas 3.5 and 8.5

$$\begin{aligned} \mathrm{L}(s^\infty) &\geq \min\{\mathrm{ord}_{N_1}(p_j), ..., \mathrm{ord}_{N_r}(p_j)\} \\ &\geq \min\{\mathrm{ord}_{N_1}(p_1), ..., \mathrm{ord}_{N_r}(p_1), ..., \mathrm{ord}_{N_1}(p_t), ..., \mathrm{ord}_{N_r}(p_t)\}. \end{aligned}$$

□

A better bound is described by the following theorem.

Theorem 8.7 *Let $N = N_1 N_2 \cdots N_r$, $m = p_1 \cdots p_t$, where $N_1, ..., N_r, p_1, ..., p_t$ are pairwise distinct primes. Assume that s^∞ is a nonconstant sequence of period N over Z_m. Let $s(i)^\infty = s^\infty \bmod p_i$ for $i = 1, ..., t$, and $i_1, ..., i_u$ be integers such that $s(i_1)^\infty, ..., s(i_u)^\infty$ are nonconstant, where $1 \leq i_1 < i_2 < \cdots < i_u \leq t$. Then L is greater than or equal to*

$$\begin{aligned} &\max\{\min\{\mathrm{ord}_{N_1}(p_{i_1}), ..., \mathrm{ord}_{N_r}(p_{i_1})\}, ..., \\ &\qquad \min\{\mathrm{ord}_{N_1}(p_{i_u}), ..., \mathrm{ord}_{N_r}(p_{i_u})\}\}, \end{aligned}$$

where L denotes $\mathrm{L}(s^\infty)$.

Proof: Since s^∞ is nonconstant, one of $s(i)^\infty$ must be nonconstant. By assumptions and Lemma 3.5

$$\mathrm{L}(s(i_k)^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_{i_k}), ..., \mathrm{ord}_{N_r}(p_{i_k})\}$$

for $k = 1, 2, ..., u$. The conclusion then follows from Lemma 8.5. □

Theorems 8.6 and 8.7 show that to control the linear complexity of sequences s^∞ over Z_m, we need only to ensure that $\min\{\mathrm{ord}_{N_1}(p_i), ..., \mathrm{ord}_{N_r}(p_i)\}$ is large enough for only one nonconstant sequence $s(i)^\infty$.

A number of practical tight bounds are described by the following five theorems.

Theorem 8.8 *Let $N, p_1, ..., p_t$ be pairwise distinct primes, and $m = p_1 \cdots p_t$. Assume $p_i^2 + 1 < N$ for $i = 1, 2, ..., t$, and $(N-1)/4$ is an odd prime. For any nonconstant sequence of period N over Z_m,*

1. *$\mathrm{L}(s^\infty) \geq (N-1)/4$;*

2. *if* $p_1, ..., p_t$ *are quadratic nonresidues modulo* N, *then* $\mathrm{L}(s^\infty) = N - 1$ *or* N.

Proof: By Theorem 8.6

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_N(p_1), ..., \mathrm{ord}_N(p_t)\}.$$

Since $\mathrm{ord}_N(p_i)$ divides $N-1$, $\mathrm{ord}_N(p_i)$ must be one of 4, k, $2k$, and $4k$, where $k = (N-1)/4$. Since $N > p_i^2+1 > p_i^2-1$, we have $p_i^4 - 1 = (p_i^2+1)(p_i^2-1) \neq 0 \bmod N$. Thus, $\mathrm{ord}_N(p_i) \geq k = (N-1)/4$. This proves part one.

If $p_1, ..., p_t$ are quadratic nonresidues modulo N, then $(\frac{p_i}{N}) = p_i^{(N-1)/2} \bmod N = -1$, where $(\frac{\cdot}{\cdot})$ is the Legendre symbol. It then follows that $p_i^{2k} = -1 \bmod N$, and that $p_i^k \neq 1 \bmod N$. Thus, $\mathrm{ord}_N(p_i) = 4k = N-1$, and $\mathrm{L}(s^\infty) \geq N-1$. This proves part two. □

Theorem 8.9 *Let* $N, p_1, ..., p_t$ *be pairwise distinct primes, and* $m = p_1 \cdots p_t$. *Assume* $p_i + 1 < N$ *for* $i = 1, 2, ..., t$, *and* $(N-1)/2$ *is an odd prime. For any nonconstant sequence of period* N *over* Z_m,

1. $\mathrm{L}(s^\infty) \geq (N-1)/2$;

2. *if* $p_1, ..., p_t$ *are quadratic nonresidues modulo* N, *then* $\mathrm{L}(s^\infty) = N - 1$ *or* N.

Proof: By Theorem 8.6

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_N(p_1), ..., \mathrm{ord}_N(p_t)\}.$$

Let $k = (N-1)/2$. Since k is prime and $\mathrm{ord}_N(p_i)$ divides $N-1 = 2k$, $\mathrm{ord}_N(p_i)$ must be one of 2, k and 2k. We first see that $\mathrm{ord}_N(p_i)$ does not equal 2, since $p_i^2 - 1 = (p_i+1)(p_i-1) \neq 0 \bmod N$ by the assumptions that $p_i + 1 < N$. Thus, $\mathrm{ord}_N(p_i) \geq k = (N-1)/2$. This proves part one. If $p_1, ..., p_t$ are quadratic nonresidues, then $(\frac{p_i}{N}) = p_i^{(N-1)/2} \bmod N = -1$, so $\mathrm{ord}_N(p_i)$ does not equal k. It follows that $\mathrm{ord}_N(p_i) = N-1$. Thus, $\mathrm{L}(s^\infty) = N-1$ or N. □

In practice the case $N = N_1N_2$ is also interesting. Sequences of period N_1N_2 over Z_m could also have large linear complexity if the primes $N_1, N_2, p_1, ..., p_t$ are properly designed, as shown by the following three theorems.

Theorem 8.10 *Let* $N_1, N_2, p_1, ..., p_t$ *be pairwise distinct primes,* $m = p_1 \cdots p_t$, *and* $N = N_1N_2$. *Assume that* $\max\{p_i^2 + 1 : i = 1, 2, ..., t\} < \min\{N_1, N_2\}$, *and* $(N_1 - 1)/4$ *and* $(N_2 - 1)/4$ *are odd primes. For any nonconstant sequence* s^∞ *of period* N *over* Z_m,

1. $\mathrm{L}(s^\infty) \geq [\min\{N_1, N_2\} - 1]/4$;

2. *if $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By Theorem 8.6

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_1), \mathrm{ord}_{N_2}(p_1), ..., \mathrm{ord}_{N_1}(p_t), \mathrm{ord}_{N_2}(p_t)\}.$$

By the proof of Theorem 8.8, $\mathrm{ord}_{N_i}(p_j) \geq [\min\{N_1, N_2\} - 1]/4$. The conclusion of part one then follows.

If $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, by the proof of Theorem 8.8 we obtain $\mathrm{ord}_{N_i}(p_j) = N_i - 1$. Thus, the conclusion of part two follows. □

Theorem 8.11 *Let $N_1, N_2, p_1, ..., p_t$ be pairwise distinct primes, $m = p_1 \cdots p_t$ and $N = N_1 N_2$. Assume that $\max\{p_i+1 : i = 1, 2, ..., t\} < \min\{N_1, N_2\}$, and $(N_1-1)/2$ and $(N_2-1)/2$ are odd primes. For any nonconstant sequence s^∞ of period N over Z_m,*

1. $\mathrm{L}(s^\infty) \geq [\min\{N_1, N_2\} - 1]/2$;

2. *if $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By Theorem 8.6

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_1), \mathrm{ord}_{N_2}(p_1), ..., \mathrm{ord}_{N_1}(p_t), \mathrm{ord}_{N_2}(p_t)\}.$$

By the proof of Theorem 8.9, $\mathrm{ord}_{N_i}(p_j) \geq [\min\{N_1, N_2\} - 1]/2$. The conclusion of part one then follows.

If $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, by the proof of Theorem 8.9 we obtain $\mathrm{ord}_{N_i}(p_j) = N_i - 1$. Thus, the conclusion of part two follows. □

Theorem 8.12 *Let $N_1, N_2, p_1, ..., p_t$ be pairwise distinct primes, $m = p_1 \cdots p_t$, and $N = N_1 N_2$. Assume that $\max\{p_i + 1 : i = 1, 2, ..., t\} < N_1$, $\max\{p_i^2 + 1 : i = 1, 2, ..., t\} < N_2$ and $(N_1-1)/2$ and $(N_2-1)/4$ are odd primes. For any nonconstant sequence s^∞ of period N over Z_m,*

1. $\mathrm{L}(s^\infty) \geq \min\{(N_1 - 1)/2, (N_2 - 1)/4\}$;

2. *if $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By Theorem 8.6

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_1), \mathrm{ord}_{N_2}(p_1), ..., \mathrm{ord}_{N_1}(p_t), \mathrm{ord}_{N_2}(p_t)\}.$$

By the proof of Theorem 8.8, $\mathrm{ord}_{N_2}(p_j) \geq [N_2 - 1]/4$. By the proof of Theorem 8.9, $\mathrm{ord}_{N_1}(p_j) \geq [N_1 - 1]/2$. The conclusion of part one then follows.

If $p_1, ..., p_t$ are quadratic nonresidues modulo both N_1 and N_2, by the proofs of Theorems 8.8 and 8.9 we obtain $\mathrm{ord}_{N_i}(p_j) = N_i - 1$. Thus, the conclusion of part two follows. □

Lower bounds on the linear complexity of sequences over Z_{p^k} can be developed as follows. Let s^∞ be a sequence of period N over Z_{p^k}, and $s(p)^\infty = s^\infty \bmod p$. Assume that $\mathrm{L}(s^\infty) = l$ and

$$s_i = a_1 s_{i-1} + a_2 s_{i-2} + \cdots + a_l s_{i-l}, \quad i \geq l$$

is a shortest linear recurrence relation of s^∞, then

$$s(p)_i = a(p)_1 s(p)_{i-1} + a(p)_2 s(p)_{i-2} + \cdots + a(p)_l s(p)_{i-l}, \quad i \geq l,$$

where $a(p)_i = a_i \bmod p$, and $s(p)_i = s_i \bmod p$. It follows that

$$\mathrm{L}(s^\infty) \geq \mathrm{L}(s(p)^\infty). \tag{8.5}$$

This inequality will provide a bridge for transferring bounds on the linear complexity of sequences over Z_p to those of sequences over Z_{p^k}.

Theorem 8.13 *Let $N = N_1^{n_1} N_2^{n_2} \cdots N_r^{n_r}$, and $m = p^k$, where $N_1, ..., N_r, p$ are pairwise distinct primes. For any sequence s^∞ of period N over Z_m, if $s(p)^\infty$ is not a constant sequence, then*

$$\mathrm{L}(s^\infty) \geq \{\mathrm{ord}_{N_1}(p), ..., \mathrm{ord}_{N_r}(p)\}.$$

Proof: By Equation (8.5) $\mathrm{L}(s^\infty) \geq \mathrm{L}(s(p)^\infty)$. The conclusion then follows from Theorem 3.5 and the assumption that $s(p)^\infty$ is nonconstant. □

Since there are many ways to control $\mathrm{ord}_{N_i}(p)$ [24], in many cases the linear complexity of sequences over Z_{p^k} is easy to control. As mentioned before, it is necessary to control the linear complexity of sequences over Z_{p^k} due to the Reeds-Sloane algorithm [57].

Since the arithmetic of Z_{2^e} can be efficiently implemented on standard processors, we describe some tight bounds on the linear complexity of sequences over Z_{2^e}. These bounds have already been set up for sequences over fields [24].

Theorem 8.14 *Let $N = 8k+3$ and $(N-1)/2$ both be odd primes, and e a positive integer. For any sequences s^∞ of period N over Z_{2^e}, if the binary sequence $s(2)^\infty$ is a nonconstant sequence, then*

$$\mathrm{L}(s^\infty) = N - 1 \text{ or } N.$$

Proof: Let $N = 2t + 1$. By assumption t is prime. Since $N = 8k + 3$, by the Law of Quadratic Reciprocity $2^{(N-1)/2} = -1 \bmod N$. The order of 2 modulo N must be one of 2, t and $2t$. Because $t \geq 3$, so $N \geq 7$. Thus, the order of 2 modulo N must be $2t = N - 1$. The conclusion then follows from Theorem 8.13. □

Theorem 8.15 *Let $N = 8k - 3$ and $(N - 1)/4$ both be primes, and e a positive integer. For any sequences s^∞ of period N over Z_{2^e}, if the binary sequence $s(2)^\infty$ is a nonconstant sequence, then*

$$\mathrm{L}(s^\infty) = N - 1 \text{ or } N.$$

Proof: Let $N = 4t + 1$. By assumption t is prime. Since $N = 8k - 3$, by the Law of Quadratic Reciprocity $2^{(N-1)/2} = -1 \bmod N$. The order of 2 modulo N must be one of 4, t, $2t$, and $4t$. Since both t and N are prime, $N \geq 13$. Thus, $\mathrm{ord}_N(2) \neq 4$, since $2^4 - 1 = 3 \times 5 \neq 0 \bmod N$. Since $2^{(N-1)/2} = 2^{2t} = -1$, the order of 2 modulo N is not equal to t or $2t$. Hence, the order of 2 modulo N must be $4t = N - 1$. The conclusion then follows from Theorem 8.13. □

Theorem 8.16 *Let $N_1 = 8k_1 + 3$ and $N_2 = 8k_2 + 3$, where $4k_1 + 1$ and $4k_2 + 1$ are odd primes. For any sequences s^∞ of period N_1N_2 over Z_{2^e}, if the binary sequence $s(2)^\infty$ is a nonconstant sequence, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By the proof of Theorem 8.14 we have

$$\mathrm{ord}_{N_1}(2) = N_1 - 1, \quad \mathrm{ord}_{N_2}(2) = N_2 - 1.$$

Then the conclusion follows from Theorem 8.13. □.

Theorem 8.17 *Let $N_1 = 8k_1 - 3$ and $N_2 = 8k_2 - 3$, where $2k_1 - 1$ and $2k_2 - 1$ are odd primes. For any sequences s^∞ of period N_1N_2 over Z_{2^e}, if the binary sequence $s(2)^\infty$ is a nonconstant sequence, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By the proof of Theorem 8.15 we have

$$\mathrm{ord}_{N_1}(2) = N_1 - 1, \quad \mathrm{ord}_{N_2}(2) = N_2 - 1.$$

Then the conclusion follows from Theorem 8.13. □.

Theorem 8.18 *Let $N_1 = 8k_1 + 3$ and $N_2 = 8k_2 - 3$, where $4k_1 + 1$ and $2k_2 - 1$ are odd primes. For any sequences s^∞ of period N_1N_2 over Z_{2^e}, if the binary sequence $s(2)^\infty$ is a nonconstant sequence, then*

$$\mathrm{L}(s^\infty) \geq \min\{N_1, N_2\} - 1.$$

Proof: By the proof of Theorems 8.14 and 8.15 we have

$$\mathrm{ord}_{N_1}(2) = N_1 - 1, \quad \mathrm{ord}_{N_2}(2) = N_2 - 1.$$

Then the conclusion follows from Theorem 8.13. □.

The bounds of Theorems 8.13, 8.14, 8.15, 8.16, 8.17, and 8.18 show how to control the linear complexity of sequences over Z_{2^e}. For sequences over Z_{p^k}, one can develop similar bounds.

A more general bound on the linear complexity of sequences over Z_m is described by the following theorem.

Theorem 8.19 *Let $N = N_1^{n_1} N_2^{n_2} \cdots N_r^{n_r}$, where N_i are distinct primes, and let $m = p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$, where p_i are pairwise distinct primes such that $\gcd(N, m) = 1$. For any sequence s^∞ of period N over Z_m, if one of the sequences $s(p_j)^\infty$ is nonconstant, then*

$$\mathrm{L}(s^\infty) \geq \min\{\mathrm{ord}_{N_1}(p_1), ..., \mathrm{ord}_{N_r}(p_1), ..., \mathrm{ord}_{N_1}(p_t), ..., \mathrm{ord}_{N_r}(p_t)\}.$$

Proof: Combining Lemma 8.5 and Theorem 8.13 proves this theorem.

All of the bounds before are special cases of this more general bound. Whether this bound is tight depends on the parameters N_i, n_i, p_i, e_i. By choosing proper values for these parameters one can easily control the linear complexity of sequences over Z_m. Here we use specialities of these parameters to control the linear complexity, instead of using some cryptographic functions to do so. These bounds could be quite useful in designing additive synchronous stream ciphers [28].

Bibliography

[1] T. M. Apostol, *Introduction to Analytic Number Theory*, Springer-Verlag, 1976.

[2] L. D. Baumert, H. Fredricksen, *The cyclotomic number of order eighteen with applications to difference sets*, Math. Comp. **21** (1967), 204-219.

[3] L. D. Baumert, *Cyclic Difference Sets*, Lecture Notes in Mathematics **182**, Springer-Verlag, 1971.

[4] E. Biham, A. Shamir, *Differential Cryptanalysis of the Data Encryption Standard,* Springer-Verlag, 1993.

[5] R. H. Bruck, *Computational aspects of certain combinatorial problems*, Proceedings of Symposia in Applied Mathematics **6**, McGraw-Hill, 1956, 31-43.

[6] J. O. Brüer, *On pseudorandom sequences as crypto generators*, Proceed. of Int. Zürich Sem. on Digital Communic., Zürich, Switzerland, 1984.

[7] R. D. Carmichael, *On sequences of integers defined by recurrence relations,* Quart. J. Pure Appl. Math. **48** (1920), 343-372.

[8] P. H. Chen, *Multisequence linear shift register synthesis and its application to BCH decoding,* IEEE Trans. Commun. **24** (1976), 438-440.

[9] S. Chowla, H. J. Ryser, *Combinatorial problems,* Canad. J. Math. **2** (1950), 93-99.

[10] D. A. Cox, *Primes of the Form $x^2 + ny^2$: Fermat, Class Field Theory, and Complex Multiplication,* John Wiley & Sons, 1989.

[11] T. W. Cusick, *Properties of the X^2 mod N generator,* to appear in IEEE Trans. Inform. Theory.

[12] H. Davenport, *On the distribution of quadratic residues (mod p),* J. London Math. Soc. **6** (1931), 49-54.

[13] J.-M. Deshouillers, *Waring's problem and the circle-method,* in R. A. Mollin (Ed.): Number Theory and Applications, Kluwer Academic Publishers, 1989, 37-44.

[14] L. E. Dickson, *Cyclotomy, higher congruences, and Waring's problem*, Amer. J. Math. **57** (1935), 391-424, 463-474.

[15] L. E. Dickson, *Solution of Waring's problem,* Amer. J. Math. **58** (1936), 530-535.

[16] W. Diffie, M. E. Hellman, *New directions in cryptography*, IEEE Trans. Inform. Theory **IT-22**(6) (1976), 644-654.

[17] W. Diffie, M. Hellman, *Privacy and authentication: An introduction to cryptography*, Proc. IEEE **67**(3) (1979), 397-427.

[18] J. F. Dillon, *A survey of bent functions*, NSA Technical Journal, Special Issue, 1972, 191-215.

[19] C. Ding, G. Xiao, W. Shan, *New measure indexes on the security of stream ciphers,* in Chinese, Proc. Third Chinese National Workshop on Cryptology, Xian, China, 1988, 5-15.

[20] C. Ding, *Proof of Massey's conjectured algorithm,* Proc. Eurocrypt'88, LNCS **330**, Springer-Verlag, 345-349.

[21] C. Ding, *Lower Bounds on the weight complexity of cascaded binary sequences*, Auscrypt'90, LNCS **453**, Springer-Verlag, 1991, 39-43.

[22] C. Ding, G. Xiao, W. Shan, *The Stability Theory of Stream Ciphers*, LNCS **561**, Springer-Verlag, 1991.

[23] C. Ding, *The differential cryptanalysis and design of the natural stream ciphers*, Fast Software Encryption, LNCS **809**, Springer-Verlag, 1994, 101-115.

[24] C. Ding, *Binary cyclotomic generators,* Fast Software Encryption, LNCS **1008**, Springer-Verlag, 1995, 29-60.

[25] C. Ding, G. Xiao, *Stream Ciphers and Their Application* (In Chinese), National Defense Industry Press, 1994.

[26] C. Ding, P. Pei, A. Salomaa, *Chinese Remainder Theorem: Applications in Computing, Coding, Cryptography,* World Scientific, 1996.

[27] C. Ding, A. Salomaa, *On cooperatively distributed ciphering and hashing,* Computers and Artificial Intelligence **15**(2) (1996), 233-245.

[28] C. Ding, V. Niemi, A. Renvall, A. Salomaa, *TWOPRIME: a fast stream ciphering algorithm,* Proceedings of FSE'97 (to appear).

[29] C. Ding, *Linear complexity of generalized cyclotomic binary sequences of order* 2, Finite Fields and Their Application (to appear).

[30] C. Ding, *Linear complexity of some generalized cyclotomic sequences,* submitted, July 1996.

[31] C. Ding, T. Helleseth, W. Shan, *On the linear Complexity of Legendre sequences,* Submitted.

[32] K. Dörge, *Zur Verteilung des Quadratischen Reste*, Jahresbericht d. Deutschen Math. Vereinigung **88** (1929), 41-49.

[33] G. L. Feng, K. K. Tzeng, *A generalization of the Berlekamp-Massey algorithm for multisequence shift-register synthesis with applications to cyclic codes,* IEEE Trans. Inform. Theory **IT-37**(5) (1991), 1274-1287.

[34] C. F. Gauss, *Disquisitiones Arith.,* Leipzig, 1801. English translation, Yale, New Haven, 1966. (Reprint by Springer-Verlag, Berlin, Heidelberg, and New York, 1986).

[35] S. W. Golomb, *Shift Register Sequences*, Aegean Park Press, Laguna Hills, California, 1982.

[36] H. Hasse, *Vorlesungen über Zahlentheorie,* Springer-Verlag, 1964.

[37] H. Hopf, *Über die Verteilung quadratischer Reste,* Math. Zeitschrift **32** (1930), 222-231.

[38] R. H. Hudson, *On the first occurrence of certain patterns of quadratic residues and non-residues,* Israel J. Math. **44** (1983), 23-32.

[39] E. Jacobsthal, *Über die Darstellung der Primzahlen der Form* $4n+1$ *als Summe zweier Quadrate,* J. für die reine und angewandte Mathematik **132** (1907), 238-245.

[40] H. Jürgensen, D. Matthews, *Some results on the information theoretic analysis of cryptosystems,* in D. Chaum (Ed.): Advances in Cryptology — Proc. Crypto'83, Plenum Press, 1984, 303-356.

[41] H. Jürgensen, L. Robbins, *Towards foundations of cryptography: investigation of perfect secrecy,* J. Universal Computer Science **2**(5) (1996), 347-379.

[42] E. Lehmer, *On the number of solutions of* $u^k + D = w^2 \bmod p$, Pacific J. Math. **5** (1955), 103-118.

[43] W. J. LeVeque, *Topics in Number Theory, vol. 1,* Addison-Wesley, 1956.

[44] R. Lidl, H. Niederreiter, *Finite Fields*, in Encyclopedia of Mathematics and Its Applications, vol. 20, Addison-Wesley, 1983.

[45] H. B. Mann, *Difference sets in elementary Abelian groups*, J. Math. **9** (1965), 212-219.

[46] H. B. Mann, *Recent advances in difference sets,* Amer. Math. Monthly **74** (1967), 229-235.

[47] Jr. H. Marshall, *Combinatorial Theory*, Blaisdell Publishing Company, 1967.

[48] J. L. Massey, *Shift-register synthesis and BCH decoding*, IEEE Trans. Inform. Theory **IT-15** (1969), 122-127.

[49] J. L. Massey, *SAFER K-64: a byte-oriented block-ciphering algorithm,* Fast Software Encryption, LNCS **809**, Springer-Verlag, 1994, 1-17.

[50] W. Meier, O. Staffelbach, *Nonlinearity criteria for cryptographic functions*, LNCS **434**, Advances in Cryptology, Springer-Verlag, 1990, 549-562.

[51] K. Nyberg, L. R. Knudsen, *Provable security against differential cryptanalysis*, LNCS, Advances in Cryptology: Eurocrypt'92, Springer-Verlag, 1993.

[52] D. Pei, *Personal communications,* Jan. 1994.

[53] R. Peralta, *On the distribution of quadratic residues and nonresidues modulo a prime number,* Math. Comp. **58** (1992), 433-440.

[54] F. R. Pichler, *Finite state machine modeling of cryptographic systems in Loops,* Proc. Eucrocrypt'87, D. Chaum and W. L. Price Eds., LNCS **304**, Springer-Verlag, 1988, 65-73.

[55] J. P. Pieprzyk, *Nonlinearity of exponent permutations*, Advances in Cryptology, Proc. Eurocrypt'89, J. J. Quisquater and J. Vandewalle Eds., Springer-Verlag, 1990, 80-92.

[56] S. Pillai. *On Waring's problem,* I. Ind. Math. Soc. **2** (1933), 16-44.

[57] J. A. Reeds, N. J. A. Sloane, *Shift-register synthesis (modulo m)*, SIAM J. Comput. **14**(3) (1985), 505-513.

[58] H. J. Ryser, *Combinatorial Mathematics*, Carus Mathematical Monograph, No. 14, 1963.

[59] A. Salomaa, *Computation and Automata,* the Cambridge University Press, 1985.

[60] A. Salomaa, *Public-key Cryptography*, EATCS Monographs on Theoretical Computer Science **23**, Springer-Verlag, 1990.

[61] W. M. Schmidt, *Equations over Finite Fields: An Elementary Approach,* Lecture Notes in Mathematics **536**, Springer-Verlag, 1976.

[62] L. von Schrutka, *Eine Beweis für die Zerlegbarkeit der Primzahlen von der Form* $6n+1$ *in ein einfaches und ein dreifaches Quadrat,* J. für die reine und angewandte Mathematik **140** (1911), 252-265.

[63] M. P. Schützenberger, *A nonexistence theorem for an infinite family of symmetrical block design,* Ann. Eugenics **14** (1949), 286-287.

[64] E. S. Selmer, *Linear Recurrence Relations over Finite Fields,* Department of Mathematics, University of Bergen, Norway, 1966.

[65] T. Siegenthaler, *Decrypting a class of stream ciphers using ciphertext only*, IEEE Trans. Comput. **C-34** (1985), 81-85.

[66] T. Storer, *Cyclotomy and Difference Sets*, Marham, Chicago, 1967.

[67] H. Walum, *A recurrent pattern in the list of quadratic residues mod a prime in the values of the Liouville lambda function,* J. Number Theory **12** (1980), 53-56.

[68] M. Ward, *The distribution of residues in a sequence satisfying a linear recursion relation,* Trans Amer. Math. Soc. **33** (1931), 166-190.

[69] M. Ward, *The arithmetical theory of linear recurring sequences*, Transactions of Americ. Math. Soc. **35** (1933), 600-628.

[70] A. L. Whiteman, *A family of difference sets,* Illinois J. Math. **6** (1962), 107-121.

[71] N. Zierler, *Linear recurring sequences,* J. Soc. Ind. Appli. Math. **7** (1959), 31-48.