JOHANNES KEPLER
UNIVERSITÄT LINZ

Eingereicht von
Richard Hofer

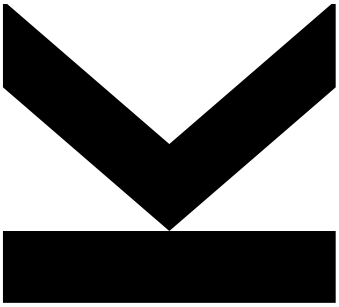
Angefertigt am
Johann Radon Institute for Computational and Applied Mathematics

Betreuer und Erstbeurteiler
Univ.-Doz. Dr. Arne Winterhof

Zweitbeurteiler
Prof. Dr. Gohar Kyureghyan

März 2017

New Bounds on Some Measures of Pseudorandomness



Dissertation

zur Erlangung des akademischen Grades

Doktor der technischen Wissenschaften

im Doktoratsstudium

Technische Wissenschaften

JOHANNES KEPLER
UNIVERSITÄT LINZ
Altenbergerstraße 69
4040 Linz, Österreich
www.jku.at
DVR 0093696

Abstract

Pseudorandom numbers are generated by deterministic algorithms and are not random at all. However, in contrast to truly random numbers they guarantee certain randomness properties. Their desirable features depend on the application area. For example, unpredictable sequences are needed for cryptography and uncorrelated sequences for wireless communication or radar. Some corresponding quality measures are linear complexity and expansion complexity for unpredictability and autocorrelation or more general correlation measure of order k.

The Legendre sequence possesses several desirable features of pseudorandomness in view of different applications such as a high linear complexity for cryptography and a small (aperiodic) autocorrelation for radar, GPS, or sonar. In this thesis we prove the first nontrivial bound on its arithmetic autocorrelation, another figure of merit coming from coding theory and introduced by Mandelbaum. Sequences with small arithmetic autocorrelation can be used to define good error-correcting codes over the integers.

Furthermore, we analyze the relation between arithmetic autocorrelation and correlation measures of higher orders. Roughly speaking, we show that any binary sequence with small correlation measure of order k up to a sufficiently large k cannot have a large arithmetic autocorrelation.

In 2012, Diem introduced a new figure of merit for cryptographic sequences called expansion complexity. Expansion complexity is essentially the same as linear complexity in the periodic case but finer in the aperiodic case. Sequences with small expansion complexity are predictable and thus not suitable in cryptography.

In this thesis we study the predictability of some number theoretic sequences over finite fields by analyzing their expansion complexity. Additionally, we consider the expansion complexity of some linear combinations of these sequences.

Zusammenfassung

Pseudozufallszahlen, welche von deterministischen Algorithmen erzeugt werden, sind nicht zufällig, garantieren aber im Gegensatz zu echten Zufallszahlen bestimmte Zufallsmerkmale. Diese gewünschten Zufallseigenschaften variieren mit der Art der Anwendung, zum Beispiel benötigt man unvorhersagbare Folgen in der Kryptographie, unkorrelierte Folgen in der drahtlosen Kommunikation oder bei Radar. Einige dazugehörige Qualitätsmaße sind die lineare Komplexität und die Expansionskomplexität für die Unvorhersagbarkeit und die Autokorrelation oder allgemeiner das Korrelationsmaß der Ordnung k.

Die Legendre Folge besitzt einige wünschenswerte Eigenschaften der Pseudozufälligkeit im Hinblick auf verschiedene Anwendungen, wie eine große lineare Komplexität für die Kryptographie und eine kleine (aperiodische) Autokorrelation für Radar, GPS oder Sonar. In dieser Arbeit beweisen wir die erste nichttriviale Schranke für ihre arithmetische Autokorrelation, eine weitere aus der Kodierungstheorie stammende Gütezahl, welche von Mandelbaum eingeführt wurde. Folgen mit kleiner arithmetischer Autokorrelation können zur Definition guter fehlerkorrigierender Codes über den ganzen Zahlen herangezogen werden.

Weiters analysieren wir den Zusammenhang zwischen der arithmetischen Autokorrelation und dem Korrelationsmaß der Ordnung k. Grob gesagt zeigen wir, dass binäre Folgen mit kleinem Korrelationsmaß der Ordnung k bis zu einem hinreichend großen k keine große arithmetische Autokorrelation haben können.

Im Jahr 2012 führte Diem die Expansionskomplexität als neue Gütezahl für kryptographische Folgen ein. Die Expansionskomplexität entspricht im periodischen Fall im Wesentlichen der linearen Komplexität, im aperiodischen Fall ist sie jedoch feiner als die lineare Komplexität. Folgen mit kleiner Expansionskomplexität sind vorhersagbar und somit nicht geeignet in der Kryptographie.

In dieser Arbeit studieren wir die Vorhersagbarkeit einiger zahlentheoretischer Folgen, indem wir ihre Expansionskomplexität analysieren. Zusätzlich betrachten wir die Expansionskomplexität von einigen Linearkombinationen dieser Folgen.

Acknowledgments

First of all, I want to express my thanks to Univ.-Doz. Arne Winterhof for supervising my thesis and for all his support throughout the last years.

I am also very thankful to Prof. Gohar Kyureghyan for her agreeing to review this thesis.

Furthermore, I want to thank all my colleagues and my family.

I have been partially supported by the Austrian Science Fund FWF Project 5511-N26 which is part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications."

Richard Hofer
Linz, March 2017

Contents

1 Introduction **7**

2 Correlation measures **10**
2.1 Preliminaries 10
2.1.1 Quadratic residues and Legendre symbol 10
2.1.2 Cyclotomic numbers of order 2 15
2.1.3 Pattern distribution of the Legendre sequence 19
2.2 Arithmetic autocorrelation of the Legendre sequence 22
2.2.1 Periodic autocorrelation 23
2.2.2 Arithmetic autocorrelation 24
2.2.3 A bound on the arithmetic autocorrelation of the Legendre sequence 26
2.3 Arithmetic autocorrelation and correlation measure 31
2.3.1 Correlation measure of order k 31
2.3.2 A bound on the arithmetic autocorrelation 32
2.3.3 Applications 35

3 Complexity measures **37**
3.1 Preliminaries 37
3.1.1 Linear complexity and expansion complexity 37
3.1.2 Growth of $L_N(\mathcal{S})$ and $E_N(\mathcal{S})$ 39
3.1.3 Purely periodic sequences 42
3.2 Linear complexity and expansion complexity of some number theoretic sequences 46
3.2.1 The characteristic sequence of the set of sums of three squares 46
3.2.2 Expansion complexity of p-periodic sequences over $\mathbb{F}_p$. 48
3.2.3 Expansion complexity of t-periodic sequences over $\mathbb{F}_q$ with $t \mid q-1$ 54

4 Outlook **58**
4.1 Arithmetic correlation measure 58
4.1.1 Arithmetic correlation measure of order k 58
4.1.2 A bound on the arithmetic correlation measure of order k of the Legendre sequence 60
4.2 Expansion complexity of tp^r-periodic sequences over $\mathbb{F}_q$ 64

Bibliography **70**

Chapter 1

Introduction

The Legendre sequence (of period p) satisfies several desirable features of pseudorandomness. For example, Turyn [42] proved that it has a high linear complexity (see also [12] and [9, Chapter 9.3]), which is necessary but not sufficient for cryptographic applications. It also provides a high linear complexity profile, see [39, Theorem 9.2].

Autocorrelation measures the similarity of a sequence and its shifts. A small (aperiodic) autocorrelation is important for radar and sonar. It is well known (see [37], [38]) that the (periodic) autocorrelation of the Legendre sequence is one-valued or two-valued depending on whether $p \equiv 3 \pmod 4$ or $p \equiv 1 \pmod 4$, and that the (absolute value of the) aperiodic autocorrelation is of order of magnitude at most $p^{1/2} \ln p$.

In this thesis we prove the first nontrivial bound on its arithmetic autocorrelation, another figure of merit introduced in [26] for error-correcting codes. More precisely, we show that the (absolute value of the) arithmetic autocorrelation of the Legendre sequence is of order of magnitude at most $p^{3/4}(\log_2 p)^{1/2}$.

Finding relations between different measures of pseudorandomness is an important goal. For example, the linear complexity provides essentially the same quality measure as certain lattice tests coming from the area of Monte Carlo methods, see [13, 33]. The correlation measure of order k is a rather general measure of pseudorandomness introduced by Mauduit and Sárközy [28]. A relation between linear complexity and the correlation measure of order k is given in [5]. Hence, we may roughly say that correlation measure is a stronger measure than linear complexity. A relation between the arithmetic autocorrelation and the correlation measure of order k is provided in this thesis.

Expansion complexity introduced in [10] is another measure which is essentially the same as linear complexity in the periodic case but finer in the

aperiodic case [31] (see also [32]).

In this thesis we analyze the linear complexity and expansion complexity of some number theoretic sequences over finite fields including the characteristic sequence of the set of sums of three squares and (linear combinations of) sequences of binomial coefficients.

Organization of the thesis

This thesis is organized as follows:

Chapter 2

- In Section 2.1 we start with some well-known preliminary results of elementary number theory. We define the Legendre sequence and show that it has the best possible distribution of patterns of length 2, see [11, Proposition 1].

- In Section 2.2 we prove that the (absolute value of the) arithmetic autocorrelation of the Legendre sequence is of order of magnitude at most $p^{3/4}(\log_2 p)^{1/2}$.

- In Section 2.3 we give a relation between arithmetic autocorrelation and the correlation measure of order k. Roughly speaking, we show that any binary sequence with small correlation measure of order k up to a sufficiently large k cannot have a large arithmetic autocorrelation. We apply our result to several classes of sequences including Legendre sequences defined with polynomials.

Chapter 3

- In Section 3.1 we provide some basic properties of the linear complexity and expansion complexity. In particular, we present the proof of [31, Theorem 1] for (purely) periodic sequences.

- In Section 3.2 we show that the characteristic sequence of the set of sums of three squares has a very small expansion complexity and thus is rather predictable. Moreover, we prove that some linear combinations of p-periodic sequences of binomial coefficients modulo p have a very small expansion complexity and are predictable despite of a high linear complexity. As an example, we consider the Legendre sequence and verify that it does not belong to this class of predictable sequences.

Parts of this thesis have already been published:

- R. Hofer, A. Winterhof, On the arithmetic autocorrelation of the Legendre sequence, Advances in Mathematics of Communications 11 (2017), no. 1, 237–244, see [20].
- R. Hofer, A. Winterhof, Linear complexity and expansion complexity of some number theoretic sequences. In: S. Duquesne, S. Petkova-Nikova (eds.), Arithmetic of Finite Fields (WAIFI 2016), 67–74, Lecture Notes in Computer Science 10064, Springer, Cham, 2016, see [21].
- R. Hofer, L. Mérai, A. Winterhof, Measures of pseudorandomness: Arithmetic autocorrelation and correlation measure. In: C. Elsholtz, P. Grabner (eds.), Number Theory – Diophantine Problems, Uniform Distribution and Applications, Festschrift in honour of Robert F. Tichy's 60th birthday, Springer, to appear, see [22].

Chapter 2

Correlation measures

In this chapter we study a different notion of autocorrelation, the arithmetic autocorrelation introduced by Mandelbaum [26]. Sequences with small arithmetic autocorrelation can be used to define good error-correcting codes over the integers (instead of finite fields). Also, see the recent monograph by Goresky and Klapper [17] for more background and results on arithmetic correlations.

Throughout this chapter let p be an odd prime number.

2.1 Preliminaries

In this section we provide some basic notions and results of elementary number theory including the Legendre symbol, 2-adic integers and cyclotomic numbers of order 2. For more details, see [9], [17], [24], [35], [36] and [40].

2.1.1 Quadratic residues and Legendre symbol

We start with a brief review of some well-known terminologies and results of elementary number theory.

Definition 2.1. Let $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$. Then a is called a *quadratic residue* modulo p if there exists $b \in \mathbb{Z}$ such that $a \equiv b^2 \pmod{p}$. Otherwise a is called a *quadratic nonresidue* modulo p.

Since $a + p$ is a quadratic residue or nonresidue modulo p, according as a is or is not, we consider as distinct residues or nonresidues only those that are distinct modulo p. Hence, we find all incongruent quadratic residues and nonresidues modulo p in the set $\{1, 2, \ldots, p-1\}$.

Remark 2.2. If a is a quadratic residue modulo p, then $a \equiv b^2 \pmod{p}$ for some $b \in \{1, 2, \ldots, p-1\}$. Then $r^2 \equiv a \equiv b^2 \pmod{p}$ implies that p divides $(r-b)(r+b)$ and so $r = b$ or $r = p-b$. Thus, the quadratic residues modulo p in $\{1, 2, \ldots, p-1\}$ can all be found by computing $b^2 \pmod{p}$ for $b = 1, 2, \ldots, (p-1)/2$ (since the remaining integers up to $p-1$ are all equal to $p-b$ for one of these b). Therefore, there are exactly $(p-1)/2$ quadratic residues modulo p in $\{1, 2, \ldots, p-1\}$, and consequently there are also exactly $(p-1)/2$ quadratic nonresidues modulo p in $\{1, 2, \ldots, p-1\}$.

Statements about quadratic residues can be formulated in an elegant manner by using the following notation.

Definition 2.3. Let $a \in \mathbb{Z}$. The *Legendre symbol* $\left(\frac{a}{p}\right)$ is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 0 & \text{if } p \mid a, \\ 1 & \text{if } a \text{ is a quadratic residue modulo } p, \\ -1 & \text{if } a \text{ is a quadratic nonresidue modulo } p. \end{cases}$$

The Legendre symbol is simply a way of identifying whether or not an integer is a quadratic residue modulo p.

Lemma 2.4 (Fermat's Little Theorem)**.** *Let $a \in \mathbb{Z}$ with $\gcd(a, p) = 1$. Then*

$$a^{p-1} \equiv 1 \pmod{p}.$$

Proof. (see [24]). We first claim that the integers $a, 2a, 3a, \ldots, (p-1)a$ are pairwise distinct modulo p. Otherwise, we would have $ia \equiv ja \pmod{p}$ for some $i, j \in \{1, 2, \ldots, p-1\}$. But this would mean that p divides $(i-j)a$, and since $\gcd(a, p) = 1$, we would have $p \mid (i-j)$. Since $1 \leq i, j < p$, the only way this can happen is if $i = j$.

We conclude that the integers $a, 2a, \ldots, (p-1)a$ are simply a rearrangement of $1, 2, \ldots, (p-1)$ when considered modulo p. Hence

$$a \cdot 2a \cdot \ldots \cdot (p-1)a \equiv 1 \cdot 2 \cdot \ldots \cdot (p-1) \pmod{p},$$

that is $a^{p-1}(p-1)! \equiv (p-1)! \pmod{p}$. Thus, p divides $(p-1)!(a^{p-1}-1)$. Since $(p-1)!$ is not divisible by p, we have $p \mid (a^{p-1}-1)$ and the result follows. □

Lemma 2.5 (Wilson's Theorem)**.** *We have*

$$(p-1)! \equiv -1 \pmod{p}.$$

Proof. (see [36]). If $p = 3$, the result is trivial. Thus we may assume that $p \geq 5$. Suppose that $i \in \{1, 2, \ldots, p-1\}$. Then there exists a unique $j \in \{1, 2, \ldots, p-1\}$ such that

$$ij \equiv ji \equiv 1 \pmod{p}.$$

We say i and j form a pair. If $i = j$, then $i^2 \equiv 1 \pmod{p}$ and so $i = 1$ or $i = p-1$. Hence, if $i \in \{2, 3, \ldots, p-2\}$ we have $i \neq j$ and it follows $2 \cdot 3 \cdot \ldots \cdot (p-2) \equiv 1 \pmod{p}$. Thus

$$(p-1)! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (p-2) \cdot (p-1) \equiv p-1 \equiv -1 \pmod{p}. \qquad \square$$

The Legendre symbol satisfies the following properties.

Proposition 2.6. *Let $a \in \mathbb{Z}$. Then*

$$\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}.$$

Proof. (see [36]). If p divides a, the congruence is easily verified. Thus we can assume that $\gcd(a, p) = 1$. If $\left(\frac{a}{p}\right) = 1$, then there exists a $b \in \mathbb{Z}$ such that $a \equiv b^2 \pmod{p}$ and by Fermat's Little Theorem we get

$$a^{(p-1)/2} \equiv b^{p-1} \equiv 1 \pmod{p}$$

since $\gcd(b, p) = \gcd(a, p) = 1$. To each $i \in \{1, 2, \ldots, p-1\}$ there exists some unique $j \in \{1, 2, \ldots, p-1\}$, so that $a \equiv ij \pmod{p}$. We say i and j form a pair. If $\left(\frac{a}{p}\right) = -1$, then there exists no $b \in \mathbb{Z}$ such that $a \equiv b^2 \pmod{p}$ and hence $i \neq j$. By Wilson's Theorem and the fact that there are exactly $(p-1)/2$ distinct ordered pairs i, j with $i \neq j$ and $a \equiv ij \pmod{p}$ we get

$$a^{(p-1)/2} \equiv (p-1)! \equiv -1 \pmod{p}. \qquad \square$$

Proposition 2.7. *Let $a, b \in \mathbb{Z}$. Then*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Proof. (see [35]). From Proposition 2.6 it follows that

$$\left(\frac{ab}{p}\right) \equiv (ab)^{(p-1)/2} \equiv a^{(p-1)/2} b^{(p-1)/2} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod{p}.$$

Now both extreme sides of this congruence have the value 0, 1 or -1 and so the congruence holds if and only if equality holds. $\square$

Example 2.8 (see [35])**.** Let $a = -1$. By Proposition 2.6 we obtain

$$\left(\frac{-1}{p}\right) \equiv (-1)^{(p-1)/2} \pmod{p}.$$

Both sides of this congruence have the value 1 or -1, and so we get the equality

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}.$$

Thus, -1 is a quadratic residue modulo p if and only if $p \equiv 1 \pmod 4$.

For sums of Legendre symbols we can state the following lemmas.

Lemma 2.9. *We have*

$$\sum_{a=0}^{p-1} \left(\frac{a}{p}\right) = 0.$$

Proof. (see [35]). Let $b \in \mathbb{Z}$ be a quadratic nonresidue modulo p. If a runs through $\{1, 2, \ldots, p-1\}$ in some order, then also $ab \pmod p$ does so. Hence

$$\sum_{a=0}^{p-1} \left(\frac{a}{p}\right) = \sum_{a=0}^{p-1} \left(\frac{ab}{p}\right) = \left(\frac{b}{p}\right) \sum_{a=0}^{p-1} \left(\frac{a}{p}\right) = -\sum_{a=0}^{p-1} \left(\frac{a}{p}\right)$$

and the result follows. □

Lemma 2.10. *Let $b \in \mathbb{Z}$ with $\gcd(b, p) = 1$. Then*

$$\sum_{a=0}^{p-1} \left(\frac{a}{p}\right) \left(\frac{a+b}{p}\right) = -1.$$

Proof. (see [35]). To each $a \in \{1, 2, \ldots, p-1\}$ there exists some unique $a^{-1} \in \{1, 2, \ldots, p-1\}$, such that $a^{-1}a \equiv 1 \pmod p$. If $\gcd(a, p) = 1$, then by Fermat's Little Theorem and Proposition 2.6 we get

$$\left(\frac{a}{p}\right)^2 \equiv \left(\frac{a^2}{p}\right) \equiv a^{p-1} \equiv 1 \equiv \left(\frac{1}{p}\right) \equiv \left(\frac{a^{-1}a}{p}\right) \equiv \left(\frac{a^{-1}}{p}\right)\left(\frac{a}{p}\right) \pmod{p}$$

and therefore $\left(\frac{a}{p}\right) = \left(\frac{a^{-1}}{p}\right)$. Thus

$$\sum_{a=0}^{p-1} \left(\frac{a}{p}\right) \left(\frac{a+b}{p}\right) = \sum_{a=1}^{p-1} \left(\frac{a^{-1}(a+b)}{p}\right) = \sum_{i=2}^{p-1} \left(\frac{i}{p}\right) = -\left(\frac{1}{p}\right) = -1.$$

by Lemma 2.9 and the observation that if a runs through $\{1, 2, \ldots, p-1\}$ in some order, then $a^{-1}(a+b) \pmod p$ runs through $\{0\} \cup \{2, 3, \ldots, p-1\}$ in some order. □

For $a \in \mathbb{Z}$ with $\gcd(a,p) = 1$, we see from Fermat's Little Theorem that there is some a^h, where $1 \leq h \leq p-1$, that is congruent to 1 modulo p.

Definition 2.11. For $a \in \mathbb{Z}$ with $\gcd(a,p) = 1$, the least positive integer h such that $a^h \equiv 1 \pmod{p}$ is called the *multiplicative order* of a modulo p. If the multiplicative order of a modulo p is equal to $p-1$, then a is called a *primitive root* modulo p.

Using the Legendre symbol we can define the following well-known pseudorandom sequence.

Definition 2.12. The *Legendre sequence* $(\ell_n)_{n\geq 0}$ is defined by

$$\ell_n = \begin{cases} 1 & \text{if } \left(\frac{n}{p}\right) = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.1}$$

Obviously, the Legendre sequence $(\ell_n)_{n\geq 0}$ is a p-periodic binary sequence. It has an optimal balance between zeros and ones. A period of $(\ell_n)_{n\geq 0}$ consists of exactly $(p+1)/2$ zeros and $(p-1)/2$ ones.

Definition 2.13. A 2-*adic integer* is a formal expression

$$\sum_{n=0}^{\infty} f_n 2^n,$$

where $f_n \in \{0,1\}$.

On the set of 2-adic integers we can perform addition and multiplication with carry which make the set of 2-adic integers a ring. More precisely, the statement

$$\sum_{n=0}^{\infty} f_n 2^n + \sum_{n=0}^{\infty} g_n 2^n = \sum_{n=0}^{\infty} h_n 2^n$$

with $f_n, g_n, h_n \in \{0,1\}$ means that there exist integers $c_0, c_1, c_2, \ldots$ so that $c_0 = 0$ and for all $n \geq 0$

$$f_n + g_n + c_n = h_n + 2c_{n+1}.$$

The quantity c_n is called the *carry* and obviously $c_n \in \{0,1\}$. It is easy to see, by induction, that c_n, h_n are uniquely determined by f_n, g_n. In fact

$$h_n = (f_n + g_n + c_n) \pmod{2}$$

and

$$c_{n+1} = \lfloor (f_n + g_n + c_n)/2 \rfloor.$$

Similarly, the statement

$$\sum_{n=0}^{\infty} f_n 2^n \cdot \sum_{n=0}^{\infty} g_n 2^n = \sum_{n=0}^{\infty} h_n 2^n$$

means that there exist integers $c_0, c_1, c_2, \ldots$ so that $c_0 = 0$ and for all $n \geq 0$

$$\sum_{i=0}^{n} f_i g_{n-i} + c_n = h_n + 2c_{n+1},$$

although in this case the carry c_n may be greater than 1.

Remark 2.14. Addition and subtraction of 2-adic integers are not the same operation. We say that

$$\sum_{n=0}^{\infty} f_n 2^n - \sum_{n=0}^{\infty} g_n 2^n = \sum_{n=0}^{\infty} h_n 2^n,$$

if there are integers $c_0, c_1, c_2, \ldots$ so that $c_0 = 0$ and for all $n \geq 0$

$$f_n - g_n - c_n = h_n - 2c_{n+1}.$$

This implies $h_n = (f_n - g_n - c_n) \pmod 2$ and $c_{n+1} = -\lfloor (f_n - g_n - c_n)/2 \rfloor$.

2.1.2 Cyclotomic numbers of order 2

In this paragraph let g be a primitive root modulo p and let $f = (p-1)/2$.

Definition 2.15. For $a \in \mathbb{Z}$, the *cyclotomic classes D_a of order* 2 are defined by

$$D_a = \{g^{2u+a} \pmod p : u = 0, 1, \ldots, f-1\}.$$

The cyclotomic classes D_0 and D_1 of order 2 are exactly the sets of incongruent quadratic residues and nonresidues modulo p, respectively.

Lemma 2.16. *If $g^{2u+a} \equiv g^{2v+b} \pmod p$ with $u, v \in \{0, 1, \ldots, f-1\}$ and $a, b \in \{0, 1\}$, then $a = b$ and $u = v$.*

Proof. (see [40]). We have $g^{2u+a} \equiv g^{2v+b} \pmod p$ if and only if

$$2u + a \equiv 2v + b \pmod{p-1}, \tag{2.2}$$

that is $p-1$ divides $2(u-v) + (a-b)$. Since $2 \mid (p-1)$ it follows that $2 \mid (a-b)$ and thus $a = b$. Then (2.2) simplifies to $2u \equiv 2v \pmod{p-1}$, or equivalently $u \equiv v \pmod f$, that is $f \mid (u-v)$. But $0 \leq u, v < f$ and hence $u = v$. □

From Lemma 2.16 it follows that the cyclotomic classes D_a of order 2 are pairwise disjoint and $D_{a+2j} = D_a$ for $j \in \mathbb{Z}$. Hence the cyclotomic classes D_{2j} and D_{1+2j} of order 2 coincide with the cyclotomic classes D_0 and D_1 of order 2, respectively. Thus we have exactly two different cyclotomic classes of order 2 and by Remark 2.2 we obtain

$$|D_a| = |D_0| = |D_1| = (p-1)/2 = f.$$

For $j \in \mathbb{Z}$ put

$$D_a + j = \{a + j \pmod{p} : a \in D_a\}$$

and

$$jD_a = \{ja \pmod{p} : a \in D_a\}.$$

Lemma 2.17. *If $r \in D_0$, then*

$$rD_0 = D_0 \quad \textit{and} \quad rD_1 = D_1.$$

If $r \in D_1$, then

$$rD_0 = D_1 \quad \textit{and} \quad rD_1 = D_0.$$

Proof. (see [9]). If $r \in D_0$, then $r \equiv g^{2v} \pmod{p}$ for some $v = 0, 1, \ldots, f-1$. Thus we get

$$rD_0 = \{g^{2(u+v)} \pmod{p} : u = 0, 1, \ldots, f-1\} = D_0$$

and

$$rD_1 = \{g^{2(u+v)+1} \pmod{p} : u = 0, 1, \ldots, f-1\} = D_1.$$

If $r \in D_1$, then $r \equiv g^{2v+1} \pmod{p}$ for some $v = 0, 1, \ldots, f-1$. Hence we obtain

$$rD_0 = \{g^{2(u+v)+1} \pmod{p} : u = 0, 1, \ldots, f-1\} = D_1$$

and

$$rD_1 = \{g^{2(u+v+1)} \pmod{p} : u = 0, 1, \ldots, f-1\} = D_0. \qquad \square$$

We now define the cyclotomic numbers of order 2.

Definition 2.18. For $a, b \in \mathbb{Z}$, the *cyclotomic number* $(a, b)_2$ *of order* 2 is defined to be the number of solutions of the equation

$$u + 1 = v, \quad u \in D_a, v \in D_b.$$

That is, $(a,b)_2$ is the number of ordered pairs u, v such that

$$g^{2u+a}+1 \equiv g^{2v+b} \pmod{p}, \quad u,v \in \{0,1,\ldots,f-1\}, \tag{2.3}$$

or equivalently

$$(a,b)_2 = |(D_a+1) \cap D_b|.$$

By the definition of D_a and $(a,b)_2$ it follows immediately that

$$(a+2i, b+2j)_2 = (a,b)_2, \quad i,j \in \mathbb{Z}.$$

Thus there are at most four distinct cyclotomic numbers of order 2.

Lemma 2.19. *There exists a $k \in \{0,1,\ldots,f-1\}$ such that*

$$g^{2k+j} \equiv -1 \pmod{p},$$

where $j=0$ if f is even, and $j=1$ if f is odd.

Proof. (see [40]). By Fermat's Little Theorem we have $g^{p-1} \equiv 1 \pmod{p}$ and thus $g^{(p-1)/2} \equiv -1 \pmod{p}$ since g is a primitive root modulo p.

If f is even, then

$$-1 \equiv g^{(p-1)/2} \equiv g^{2(f/2)} \pmod{p},$$

that is $k = f/2$. If f is odd, then

$$-1 \equiv g^{(p-1)/2} \equiv g^{\frac{p-3}{2}+1} \equiv g^{2\left(\frac{f-1}{2}\right)+1} \pmod{p},$$

that is $k = (f-1)/2$. □

The cyclotomic numbers of order 2 satisfy the following properties.

Lemma 2.20. *We have*

$$(a,b)_2 = (2-a, b-a)_2 = \begin{cases} (b,a)_2 & \text{if } f \text{ is even}, \\ (b+1, a+1)_2 & \text{if } f \text{ is odd}, \end{cases}$$

Proof. (see [40]). By Fermat's Little Theorem we have $g^{2f} \equiv 1 \pmod{p}$. If we multiply both sides of (2.3) by $g^{2(f-u-1)+(2-a)}$, that is the inverse of g^{2u+a}, we get

$$g^{2(f-u-1)+(2-a)} + 1 \equiv g^{2(v-u)+(b-a)} \pmod{p},$$

whose number of solutions is $(2-a, b-a)_2$ by definition.

From Lemma 2.19 it follows that there exists a $k \in \{0,1,\ldots,f-1\}$ such that $g^{2k+j} \equiv -1 \pmod{p}$, where $j=0$ if f is even, and $j=1$ if f is odd. If we multiply both sides of (2.3) by $g^{2k+j} \equiv -1 \pmod{p}$ we obtain

$$g^{2(v+k)+(b+j)} + 1 \equiv g^{2(u+k)+(a+j)} \pmod{p},$$

whose number of solutions is $(b+j, a+j)_2$ by definition. □

Lemma 2.21. *We have*

$$(a,0)_2 + (a,1)_2 = f - \beta_a,$$

where

$$\beta_a = \begin{cases} 1 & \text{if } f \text{ is even and } a \equiv 0 \pmod 2, \\ 1 & \text{if } f \text{ is odd and } a \equiv 1 \pmod 2, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (see [40]). We have

$$(a,0)_2 + (a,1)_2 = |(D_a + 1) \cap D_0| + |(D_a + 1) \cap D_1|.$$

From Example 2.8 it follows that $p-1 \in D_0$ if and only if $p \equiv 1 \pmod 4$, or equivalently $0 \in D_0 + 1$ if and only if $p \equiv 1 \pmod 4$. Furthemore, we have $|D_a + 1| = |D_a| = f$, and the only element of $\{0, 1, \ldots, p-1\}$ not in some D_a is 0 which is neither a quadratic residue nor a quadratic nonresidue modulo p. Hence if f is even, that is $p \equiv 1 \pmod 4$, then

$$|(D_a + 1) \cap D_0| + |(D_a + 1) \cap D_1| = \begin{cases} f - 1 & \text{if } a \equiv 0 \pmod 2, \\ f & \text{if } a \equiv 1 \pmod 2. \end{cases}$$

If f is odd, that is $p \equiv 3 \pmod 4$, then

$$|(D_a + 1) \cap D_0| + |(D_a + 1) \cap D_1| = \begin{cases} f & \text{if } a \equiv 0 \pmod 2, \\ f - 1 & \text{if } a \equiv 1 \pmod 2. \end{cases}$$

Thus the result follows. □

The properties given in Lemma 2.20 and Lemma 2.21 are sufficient to determine the cyclotomic numbers of order 2.

Proposition 2.22. *If $p \equiv 1 \pmod 4$, the cyclotomic numbers of order 2 are given by*

$$(0,0)_2 = (p-5)/4, \quad (0,1)_2 = (1,0)_2 = (1,1)_2 = (p-1)/4.$$

If $p \equiv 3 \pmod 4$, they are given by

$$(0,0)_2 = (1,0)_2 = (1,1)_2 = (p-3)/4, \quad (0,1)_2 = (p+1)/4.$$

Proof. (see [9]). If f is even, that is $p \equiv 1 \pmod 4$, it follows from Lemma 2.20 and Lemma 2.21 that the four distinct cyclotomic numbers are related by

$$(0,0)_2 =: A, \quad (0,1)_2 = (1,0)_2 = (1,1)_2 =: B,$$

where A and B satisfy the equations

$$A + B = f - 1, \quad 2B = f.$$

Solving these equations we get

$$(0,0)_2 = A = (p-5)/4, \quad (0,1)_2 = (1,0)_2 = (1,1)_2 = B = (p-1)/4.$$

If f is odd, that is $p \equiv 3 \pmod 4$, it follows that

$$(0,0)_2 = (1,0)_2 = (1,1)_2 =: C, \quad (0,1)_2 =: D,$$

where

$$C + D = f, \quad 2C = f - 1.$$

Solving these equations we obtain

$$(0,0)_2 = (1,0)_2 = (1,1)_2 = C = (p-3)/4, \quad (0,1)_2 = D = (p+1)/4. \quad \square$$

2.1.3 Pattern distribution of the Legendre sequence

Employing the elementary facts about cyclotomic numbers of order 2 we can present the proof of [11, Proposition 1], which shows that Legendre sequences have an ideal distribution of patterns of length 2. But first we introduce the notion of a pattern.

Definition 2.23. Let s be a positive integer. A *pattern of length* s is a string

$$i_0 * \cdots * i_1 * \cdots * \cdots * i_{s-1},$$

where $i_0, i_1, \ldots, i_{s-1} \in \{0,1\}$ are fixed bits, the $*$'s indicate arbitrary bits that could be either 0 and 1, and the distances among $i_0, i_1, \ldots, i_{s-1}$ are fixed.

Ding [11] studied the pattern distribution of the Legendre sequence. More precisely, for $i_0, i_1, \ldots, i_{s-1} \in \{0,1\}$ and positive integers $d_1, d_2, \ldots, d_{s-1}$ with $0 < d_1 < d_2 < \ldots < d_{s-1} < p$, put

$$\mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(\ell_n) = |\{0 \le n \le p-1 : \ell_n = i_0, \ell_{n+d_1} = i_1, \ldots, \ell_{n+d_{s-1}} = i_{s-1}\}|.$$

The parameters $\mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(\ell_n)$ count the number of patterns distributed in a cycle of the Legendre sequence.

For the distribution of patterns of length 2 in Legendre sequences we have the following exact result, which means that Legendre sequences have the best possible distribution of patterns of length 2.

Proposition 2.24 (Ding, [11]). *If* $p \equiv 3 \pmod 4$, *then*

$$\mathcal{P}_{i_0,i_1}(\ell_n) = \begin{cases} (p-3)/4 & \text{for } i_0 = i_1 = 1, \\ (p+1)/4 & \text{otherwise.} \end{cases}$$

If $p \equiv 1 \pmod 4$, *then*

$$\mathcal{P}_{1,1}(\ell_n) = \begin{cases} (p-5)/4 & \text{for } (p-1)/2 \text{ elements } d_1 \text{ of } \{1,2,\ldots,p-1\}, \\ (p-1)/4 & \text{for the remaining elements,} \end{cases}$$

$$\mathcal{P}_{1,0}(\ell_n) = \begin{cases} (p+3)/4 & \text{for } (p-1)/2 \text{ elements } d_1 \text{ of } \{1,2,\ldots,p-1\}, \\ (p-1)/4 & \text{for the remaining elements,} \end{cases}$$

$$\mathcal{P}_{0,1}(\ell_n) = \begin{cases} (p+3)/4 & \text{for } (p-1)/2 \text{ elements } d_1 \text{ of } \{1,2,\ldots,p-1\}, \\ (p-1)/4 & \text{for the remaining elements,} \end{cases}$$

$$\mathcal{P}_{0,0}(\ell_n) = \begin{cases} (p+3)/4 & \text{for } (p-1)/2 \text{ elements } d_1 \text{ of } \{1,2,\ldots,p-1\}, \\ (p-1)/4 & \text{for the remaining elements.} \end{cases}$$

Proof. (see [11]). Recall that

$$(a,b)_2 = |(D_a + 1) \cap D_b| = |D_b \cap (D_a + 1)|.$$

Note that 0 is neither a quadratic residue nor a quadratic nonresidue modulo p and the cyclotomic classes are pairwise disjoint.

We have

$$\mathcal{P}_{1,1}(\ell_n) = |\{0 \le n \le p-1 : \ell_n = 1, \ell_{n+d_1} = 1\}| = |D_0 \cap (D_0 + d_1)|.$$

To each $d_1 \in \{1,2,\ldots,p-1\}$ there exists some unique $d_1^{-1} \in \{1,2,\ldots,p-1\}$, such that $d_1 d_1^{-1} \equiv 1 \pmod p$. Hence

$$|D_0 \cap (D_0 + d_1)| = |d_1^{-1} D_0 \cap (d_1^{-1} D_0 + 1)|$$

and by Lemma 2.17 we get

$$\mathcal{P}_{1,1}(\ell_n) = |d_1^{-1} D_0 \cap (d_1^{-1} D_0 + 1)| = |D_j \cap (D_j + 1)| = (j,j)_2,$$

where $d_1^{-1} \in D_j$ for some $j \in \{0,1\}$.

Similarly,

$$\begin{aligned}\mathcal{P}_{0,1}(\ell_n) &= |\{0 \le n \le p-1 : \ell_n = 0, \ell_{n+d_1} = 1\}| = |(D_1 \cup \{0\}) \cap (D_0 + d_1)| \\ &= |(d_1^{-1}D_1 \cup \{0\}) \cap (d_1^{-1}D_0 + 1)| = |(D_{j+1} \cup \{0\}) \cap (D_j + 1)| \\ &= |D_{j+1} \cap (D_j + 1)| + |\{0\} \cap (D_j + 1)| = (j, j+1)_2 + |\{0\} \cap (D_j + 1)|,\end{aligned}$$

where $d_1^{-1} \in D_j$ for some $j \in \{0, 1\}$.

With similar argument, we have

$$\begin{aligned}\mathcal{P}_{1,0}(\ell_n) &= |\{0 \le n \le p-1 : \ell_n = 1, \ell_{n+d_1} = 0\}| = |D_0 \cap ((D_1 \cup \{0\}) + d_1)| \\ &= |d_1^{-1}D_0 \cap ((d_1^{-1}D_1 \cup \{0\}) + 1)| = |d_1^{-1}D_0 \cap ((d_1^{-1}D_1 + 1) \cup \{1\})| \\ &= |d_1^{-1}D_0 \cap (d_1^{-1}D_1 + 1)| + |d_1^{-1}D_0 \cap \{1\}| \\ &= |D_j \cap (D_{j+1} + 1)| + |D_j \cap \{1\}| = (j+1, j)_2 + |D_j \cap \{1\}|\end{aligned}$$

and

$$\begin{aligned}\mathcal{P}_{0,0}(\ell_n) &= |\{0 \le n \le p-1 : \ell_n = 0, \ell_{n+d_1} = 0\}| \\ &= |(D_1 \cup \{0\}) \cap ((D_1 \cup \{0\}) + d_1)| \\ &= |(d_1^{-1}D_1 \cup \{0\}) \cap ((d_1^{-1}D_1 \cup \{0\}) + 1)| \\ &= |(d_1^{-1}D_1 \cup \{0\}) \cap ((d_1^{-1}D_1 + 1) \cup \{1\})| \\ &= |d_1^{-1}D_1 \cap ((d_1^{-1}D_1 + 1) \cup \{1\})| + |\{0\} \cap ((d_1^{-1}D_1 + 1) \cup \{1\})| \\ &= |d_1^{-1}D_1 \cap (d_1^{-1}D_1 + 1)| + |d_1^{-1}D_1 \cap \{1\}| + |\{0\} \cap (d_1^{-1}D_1 + 1)| \\ &= |D_{j+1} \cap (D_{j+1} + 1)| + |D_{j+1} \cap \{1\}| + |\{0\} \cap (D_{j+1} + 1)| \\ &= (j+1, j+1)_2 + |D_{j+1} \cap \{1\}| + |\{0\} \cap (D_{j+1} + 1)|,\end{aligned}$$

where $d_1^{-1} \in D_j$ for some $j \in \{0, 1\}$.

We have $1 \in D_a$ if $a \equiv 0 \pmod 2$ since 1 is always a quadratic residue modulo p and from Example 2.8 it follows that $0 \in D_0 + 1$ if and only if $p \equiv 1 \pmod 4$. Hence

$$|\{0\} \cap (D_a + 1)| = \begin{cases} a \pmod 2 & \text{if } p \equiv 3 \pmod 4, \\ a + 1 \pmod 2 & \text{if } p \equiv 1 \pmod 4, \end{cases}$$

and

$$|D_a \cap \{1\}| = \begin{cases} 1 & \text{if } a \equiv 0 \pmod 2, \\ 0 & \text{if } a \equiv 1 \pmod 2. \end{cases}$$

The result follows then from Proposition 2.22 and the above four formulae for $\mathcal{P}_{1,1}(\ell_n)$, $\mathcal{P}_{0,1}(\ell_n)$, $\mathcal{P}_{1,0}(\ell_n)$ and $\mathcal{P}_{0,0}(\ell_n)$. □

Example 2.25. For $p = 7$ the (incongruent) quadratic residues modulo 7 are 1, 2 and 4, the (incongruent) quadratic nonresidues modulo 7 are 3, 5 and 6. Hence, the 7-periodic Legendre sequence is given by

$$(\ell_n)_{n\geq 0} = (0\ 1\ 1\ 0\ 1\ 0\ 0\ 0\ 1\ 1\ 0\ 1\ 0\ 0\ \ldots)$$

with $\ell_{n+7} = \ell_n$ for all $n \geq 0$. If $i_0 = i_1 = 1$, the patterns

$$\begin{aligned}
1\ 1 &\text{ appear once as } \ell_1\ \ell_2 \text{ in a cycle of } (\ell_n)_{n\geq 0},\\
1\ *\ 1 &\text{ appear once as } \ell_2\ *\ \ell_4 \text{ in a cycle of } (\ell_n)_{n\geq 0},\\
1\ *\ *\ 1 &\text{ appear once as } \ell_1\ *\ *\ \ell_4 \text{ in a cycle of } (\ell_n)_{n\geq 0},\\
1\ *\ *\ *\ 1 &\text{ appear once as } \ell_4\ *\ *\ *\ \ell_1 \text{ in a cycle of } (\ell_n)_{n\geq 0},\\
1\ *\ *\ *\ *\ 1 &\text{ appear once as } \ell_4\ *\ *\ *\ *\ \ell_2 \text{ in a cycle of } (\ell_n)_{n\geq 0},\\
1\ *\ *\ *\ *\ *\ 1 &\text{ appear once as } \ell_2\ *\ *\ *\ *\ *\ \ell_1 \text{ in a cycle of } (\ell_n)_{n\geq 0}.
\end{aligned}$$

Thus $\mathcal{P}_{1,1}(\ell_n) = 1$ for all $d_1 \in \{1, 2, \ldots, 6\}$. By Proposition 2.24 we have $\mathcal{P}_{1,0}(\ell_n) = \mathcal{P}_{0,1}(\ell_n) = \mathcal{P}_{0,0}(\ell_n) = 2$, which means that all remaining patterns appear twice in a cycle of $(\ell_n)_{n\geq 0}$.

For the distribution of patterns of length $s \geq 3$ in Legendre sequences Ding proved the following proposition.

Proposition 2.26 (Ding, [11])**.** *For $s \geq 3$ we have*

$$\left|\mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(\ell_n) - \frac{p}{2^s}\right| \leq \frac{p^{1/2}(2^{s-1}(s-3)+2) + 2^{s-1}(s+1) - 1}{2^s}. \tag{2.4}$$

Proof. See [9] or [11, Proposition 2]. □

If $s = 3$, the bound (2.4) simplifies to

$$\left|\mathcal{P}_{i_0,i_1,i_2}(\ell_n) - \frac{p}{2^3}\right| \leq \frac{2p^{1/2} + 15}{2^3}.$$

Numerical computation shows that these lower and upper bounds for $s = 3$ are quite tight.

It can be seen from the development of the bounds that they are usually tight for small s. Hence, Proposition 2.26 shows that Legendre sequences have a rather ideal distribution of patterns of length s when s is small.

2.2 Arithmetic autocorrelation of the Legendre sequence

In this section we show that the Legendre sequence of period p has a maximal (absolute value of the) arithmetic autocorrelation of order of magnitude at most $p^{3/4}(\log_2 p)^{1/2}$.

2.2.1 Periodic autocorrelation

We start with the calculation of the (periodic) autocorrelation of the Legendre sequence.

Definition 2.27. The *(periodic) autocorrelation function* $C(t)$ of a (purely) T-periodic binary sequence $(a_n)_{n\geq 0}$ is defined as

$$C(t) = \sum_{n=0}^{T-1} (-1)^{a_n - a_{n+t}}, \quad 1 \leq t \leq T-1.$$

The (periodic) autocorrelation $C(t)$ is a measure for the similarity of a sequence and its shifts by t positions. Note that $(-1)^{a_n - a_{n+t}} = (-1)^{a_n + a_{n+t}}$ since $(a_n)_{n\geq 0}$ is a binary sequence.

Proposition 2.28. *The (periodic) autocorrelation function of the p-periodic binary sequence $(\ell_n)_{n\geq 0}$ defined by* (2.1) *satisfies*

$$C(t) = \begin{cases} -1 & \text{if } p \equiv 3 \pmod 4, \\ -2\left(\frac{t}{p}\right) - 1 & \text{if } p \equiv 1 \pmod 4, \end{cases} \quad 1 \leq t \leq p-1.$$

Proof. (see [35]). If $n \not\equiv 0 \pmod p$, then

$$(-1)^{\ell_n} = -\left(\frac{n}{p}\right).$$

Hence for $1 \leq t \leq p-1$ we get

$$\begin{aligned} C(t) = \sum_{n=0}^{p-1} (-1)^{\ell_n + \ell_{n+t}} &= (-1)^{\ell_t} + (-1)^{\ell_{p-t}} + \sum_{\substack{n=1 \\ n\neq p-t}}^{p-1} (-1)^{\ell_n + \ell_{n+t}} \\ &= -\left(\frac{t}{p}\right) - \left(\frac{p-t}{p}\right) + \sum_{\substack{n=1 \\ n\neq p-t}}^{p-1} \left(\frac{n}{p}\right)\left(\frac{n+t}{p}\right) \\ &= -\left(\frac{t}{p}\right)\left(1 + \left(\frac{-1}{p}\right)\right) + \sum_{\substack{n=1 \\ n\neq p-t}}^{p-1} \left(\frac{n}{p}\right)\left(\frac{n+t}{p}\right) \\ &= -\left(\frac{t}{p}\right)\left(1 + (-1)^{(p-1)/2}\right) + \sum_{n=1}^{p-1} \left(\frac{n}{p}\right)\left(\frac{n+t}{p}\right) \end{aligned}$$

by Proposition 2.6 and the result follows from Lemma 2.10. □

Remark 2.29. The (absolute value of the) *aperiodic autocorrelation* of the Legendre sequence

$$\sum_{n=0}^{M-1} (-1)^{\ell_n - \ell_{n+t}}, \quad 1 \le t \le p-1,$$

is of order of magnitude at most $p^{1/2} \ln p$ for $1 \le M \le p-1$ (see for example [41, Theorem 3.1]).

2.2.2 Arithmetic autocorrelation

For an ultimately T-periodic binary sequence $(s_n)_{n\ge 0}$ with preperiod T_0, that is $s_{n+T} = s_n$ for all $n \ge T_0$, the *imbalance* $I(s_n)$ is defined by

$$I(s_n) = N_0 - N_1, \tag{2.5}$$

where

$$N_i = |\{T_0 \le n \le T_0 + T - 1 : s_n = i\}|, \quad i = 0, 1.$$

The *arithmetic autocorrelation function* $A(t)$ of a (purely) T-periodic binary sequence $(a_n)_{n\ge 0}$ is defined as follows. For $t \in \{1, 2, \ldots, T-1\}$ let $(a_{n+t})_{n\ge 0}$ be the shift of $(a_n)_{n\ge 0}$ by lag t. Put

$$x_t = \sum_{n=0}^{T-1} a_{n+t} 2^n \quad \text{and} \quad \alpha_t = \sum_{n=0}^{\infty} a_{n+t} 2^n, \quad 0 \le t < T. \tag{2.6}$$

Note that with respect to the 2-adic norm of $\mathbb{Q}$, that is

$$|x|_2 = 2^{-k} \quad \text{if } x = 2^k \frac{u}{v} \in \mathbb{Q} \setminus \{0\} \text{ with odd } u \text{ and } v,$$

the geometric series $\sum_{n=0}^{\infty} x^n$ converges for any even integer x to

$$\sum_{n=0}^{\infty} x^n = -\frac{1}{x-1}, \quad |x|_2 < 1.$$

In particular we have $\sum_{n=0}^{\infty} 2^n = -1$, or more general

$$\sum_{n=0}^{\infty} 2^{nk} = -\frac{1}{2^k - 1}, \quad k = 1, 2, \ldots$$

and therefore we get

$$\alpha_t = \sum_{n=0}^{T-1} a_{n+t} 2^n \sum_{m=0}^{\infty} 2^{mT} = -\frac{x_t}{2^T - 1}, \quad 0 \le t < T.$$

We write

$$\alpha_0 - \alpha_t = \sum_{n=0}^{\infty} s_{n,t} 2^n \tag{2.7}$$

with unique $s_{n,t} \in \{0, 1\}$.

If $x_0 \geq x_t$, note that $(s_{n,t})_{n\geq 0}$ is (purely) periodic with period T since

$$\sum_{n=0}^{\infty} s_{n,t} 2^n = (x_0 - x_t) \sum_{n=0}^{\infty} 2^{nT}.$$

If $x_0 < x_t$, note that

$$0 < \sum_{n=0}^{T-1} s_{n,t} 2^n = 2^T + \sum_{n=0}^{T-1} (a_n - a_{n+t}) 2^n = 2^T + x_0 - x_t < 2^T$$

and thus $(s_{n,t})_{n\geq 0}$ is ultimately periodic with period T from T on (see also Goresky and Klapper [14, Proposition 2]) since

$$\sum_{n=T}^{\infty} s_{n,t} 2^{n-T} = -1 + \sum_{n=0}^{\infty} (a_n - a_{n+t}) 2^n = (2^T - 1 + x_0 - x_t) \sum_{n=0}^{\infty} 2^{nT}. \tag{2.8}$$

In both cases we define

$$A(t) = I(s_{n,t}), \quad 1 \leq t \leq T - 1.$$

Remark 2.30. The arithmetic autocorrelation $A(t)$ is a with-carry analogue to the (periodic) autocorrelation $C(t)$. In Proposition 2.28 we showed that the (periodic) autocorrelation of the Legendre sequence $(\ell_n)_{n\geq 0}$ is one-valued if $p \equiv 3 \pmod 4$ and two-valued if $p \equiv 1 \pmod 4$. For the arithmetic autocorrelation of the Legendre sequence we prove that

$$|A(t)| \leq 4p^{3/4} (\log_2 p)^{1/2}, \quad 1 \leq t \leq p - 1.$$

For very small $\min\{t, p - t\}$ we improve this bound.

The arithmetic autocorrelation satisfies the following symmetry property.

Proposition 2.31. *The arithmetic autocorrelation function of a periodic binary sequence $(a_n)_{n\geq 0}$ of least period T satisfies*

$$A(t) = -A(T - t), \quad 1 \leq t \leq T - 1.$$

Proof. For $0 \le t < T$, let x_t and α_t be defined by (2.6). If $x_0 > x_t$, then we have

$$-2^{T+t} < \sum_{n=0}^{T+t-1} (a_n - a_{n+T-t})2^n = \sum_{n=0}^{t-1} (a_n - a_{n+T-t})2^n - 2^t(x_0 - x_t) < 0.$$

Hence

$$\alpha_0 - \alpha_{T-t} = \underbrace{2^{T+t} + \sum_{n=0}^{T+t-1} (a_n - a_{n+T-t})2^n}_{<\,2^{T+t}} + 2^t \sum_{n=T}^{\infty} (1 - s_{n,t})2^n = \sum_{n=0}^{\infty} s_{n,T-t}2^n$$

with $(s_{n,t})_{n\ge 0}$ and $(s_{n,T-t})_{n\ge 0}$ defined by (2.7). Both $(s_{n,t})_{n\ge 0}$ and $(s_{n,T-t})_{n\ge 0}$ are (ultimately) periodic with period T from T on and the number of ones in a period of $(s_{n,t})_{n\ge 0}$ equals the number of zeros in a period of $(s_{n,T-t})_{n\ge 0}$. Hence

$$A(T-t) = I(s_{n,T-t}) = -I(s_{n,t}) = -A(t), \quad t = 1, \ldots, T-1.$$

If $x_0 < x_t$, then we have

$$2^{T+t} > \sum_{n=0}^{T+t-1} (a_n - a_{n+T-t})2^n = \sum_{n=0}^{t-1} (a_n - a_{n+T-t})2^n - 2^t(x_0 - x_t) > 0$$

and thus

$$\alpha_0 - \alpha_{T-t} = \sum_{n=0}^{T+t-1} (a_n - a_{n+T-t})2^n + 2^t \sum_{n=T}^{\infty} (1 - s_{n,t})2^n = \sum_{n=0}^{\infty} s_{n,T-t}2^n$$

by (2.8) and the result follows as in the first case. □

2.2.3 A bound on the arithmetic autocorrelation of the Legendre sequence

For $t = 1$ the arithmetic autocorrelation of the Legendre sequence $(\ell_n)_{n\ge 0}$ is easy to determine. Then

$$x_0 - x_1 = x_0/2 = x_1 = \sum_{n=0}^{p-1} \ell_{n+1}2^n,$$

$N_0 = N_1 + 1 = (p+1)/2$ and thus

$$A(1) = 1 = -A(p-1). \tag{2.9}$$

Now we deal with any $1 \le t \le p-1$.

Theorem 2.32. *The arithmetic autocorrelation function of the p-periodic binary sequence $(\ell_n)_{n\geq 0}$ defined by (2.1) satisfies*

$$|A(t)| \leq \begin{cases} 4p^{3/4}(\log_2 p)^{1/2} & \text{if } r > m, \\ 2^r(4\log_2 p + 2(m^2 - r^2))p^{1/2} & \text{if } r \leq m, \end{cases}$$

where $m = \lfloor 1/4 \log_2 p - 1/2 \log_2 \log_2 p \rfloor$ and $r = \min\{t, p-t\}$ for $1 \leq t \leq p-1$.

Proof. By (2.9) and Proposition 2.31 we may assume $2 \leq t \leq (p-1)/2$. In the following we derive a lower bound on the number N_1 of ones in a period of the p-periodic sequence $(s_{n,t})_{n\geq 0}$ defined by (2.7).

If $p \leq 4p^{3/4}(\log_2 p)^{1/2}$ or $p \leq 2^t(4\log_2 p + 2(m^2 - t^2))p^{1/2}$, respectively, then the result follows immediately since the trivial bound $|A(t)| \leq p$ always holds. Thus it is enough to prove the inequality for $p^{1/4} > 4(\log_2 p)^{1/2}$ or $p^{1/2} > 2^t(4\log_2 p + 2(m^2 - t^2))$, respectively.

Note that $1 \leq m \leq 1/4 \log_2 p$. Take $a \in \{0, 1\}$. For some k and n with $0 \leq k < m$ and $p \leq n < 2p$ assume

$$\begin{aligned} &(\ell_{n-k-1}, \ell_{n-k-1+t}) = (a, 1-a), \\ &\ell_{n-k+j} = \ell_{n-k+j+t}, \quad j = 0, \ldots, k-1, \\ &(\ell_n, \ell_{n+t}) \in \{0, 1\}^2. \end{aligned} \tag{2.10}$$

We consider only patterns of length $4 \leq s = 2k + 4 \leq 1/2 \log_2 p + 2$ and therefore we can further estimate (2.4) by $sp^{1/2}/2$, that is

$$\begin{aligned} \left| \mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(\ell_n) - \frac{p}{2^s} \right| &\leq \frac{p^{1/2}(2^{s-1}(s-3)+2) + 2^{s-1}(s+1) - 1}{2^s} \\ &\leq \left(\frac{s-3}{2} + 2^{1-s} \right) p^{1/2} + \frac{s+1}{2} \\ &\leq \left(\frac{s}{2} - \frac{11}{8} \right) p^{1/2} + \frac{7}{4} \log_2 p \leq \frac{s}{2} p^{1/2} \end{aligned} \tag{2.11}$$

since $p^{1/4} > 4(\log_2 p)^{1/2}$ or $p^{1/2} > 2^t(4\log_2 p + 2(m^2 - t^2)) \geq 16 \log_2 p$, respectively.

First we assume $m + 1 \leq t \leq (p-1)/2$. From (2.11) we know that (for fixed a) the number of patterns

$$\begin{pmatrix} \ell_{n-k-1} & \ell_{n-k} & \ldots & \ell_{n-1} & \ell_n \\ \ell_{n-k-1+t} & \ell_{n-k+t} & \ldots & \ell_{n-1+t} & \ell_{n+t} \end{pmatrix} \tag{2.12}$$

satisfying the assumptions (2.10) in

$$\begin{array}{cccccccc} \ell_{p-k-1} & \ell_{p-k} & \ldots & \ell_{p-1} & \ell_p & \ldots & \ell_{2p-2} & \ell_{2p-1} \\ \ell_{t+p-k-1} & \ell_{t+p-k} & \ldots & \ell_{t+p-1} & \ell_{t+p} & \ldots & \ell_{t+2p-2} & \ell_{t+2p-1} \end{array} \tag{2.13}$$

is at least $p/2^{2k+4} - (k+2)p^{1/2}$. We have to distinguish between two cases.

If $a = 1$, then $(\ell_{n-k-1}, \ell_{n-k-1+t}) = (1, 0)$. The subtraction of 0 from 1 gives no carry, no matter if there was a carry in the previous step. Hence

$$s_{n,t} = \begin{cases} 1 & \text{if } \ell_n \neq \ell_{n+t}, \\ 0 & \text{if } \ell_n = \ell_{n+t}. \end{cases}$$

Since there are 2^{k+1} possible choices for the pattern (2.12) we count at least $p/2^{k+3} - (k+2)2^{k+1}p^{1/2}$ different $p \leq n < 2p$ with $s_{n,t} = 1$.

If $a = 0$, then $(\ell_{n-k-1}, \ell_{n-k-1+t}) = (0, 1)$. The subtraction of 1 from 0 gives a carry, no matter if there was a carry in the previous step. Hence

$$s_{n,t} = \begin{cases} 1 & \text{if } \ell_n = \ell_{n+t}, \\ 0 & \text{if } \ell_n \neq \ell_{n+t}. \end{cases}$$

Just as before there are 2^{k+1} possible choices for the pattern (2.12) and so we get at least $p/2^{k+3} - (k+2)2^{k+1}p^{1/2}$ additional n with $s_{n,t} = 1$.

Thus in total we have at least

$$p/2^{k+2} - (k+2)2^{k+2}p^{1/2}$$

different $p \leq n < 2p$ with $\ell_{n-k-1} \neq \ell_{n-k-1+t}$, $(\ell_{n-k+j}, \ell_{n-k+j+t}) \in \{(0,0), (1,1)\}$ for $j = 0, \ldots, k-1$ and $s_{n,t} = 1$.

Summing up all the contributions we get the formula

$$N_1 \geq \frac{1}{4}\left(\sum_{k=0}^{m-1} 2^{-k}\right) p - 2\left(\sum_{k=0}^{m-1} 2^{k+1}(k+2)\right) p^{1/2}.$$

The first sum on the right hand side of the inequality is a geometric series, hence we have

$$\frac{1}{4}\sum_{k=0}^{m-1} \frac{1}{2^k} = \frac{1}{2} - 2^{-m-1}.$$

The second sum can be estimated by

$$\sum_{k=0}^{m-1} 2^{k+1}(k+2) = m2^{m+1} \leq 2^{m-1}\log_2 p,$$

where we used $m \leq 1/4 \log_2 p$. Thus by the definition of m we get

$$\begin{aligned} N_1 &\geq \frac{1}{2}p - 2^{-m-1}p - 2^m p^{1/2}\log_2 p \\ &\geq \frac{p}{2} - p^{3/4}(\log_2 p)^{1/2} - p^{3/4}(\log_2 p)^{1/2} = \frac{p}{2} - 2p^{3/4}(\log_2 p)^{1/2}. \end{aligned}$$

Analogously N_0 can be bounded below by

$$N_0 \geq \frac{p}{2} - 2p^{3/4}(\log_2 p)^{1/2}$$

and therefore since $N_0 + N_1 = p$

$$|A(t)| = |N_0 - N_1| = |p - 2N_1| = |p - 2N_0| \leq 4p^{3/4}(\log_2 p)^{1/2}.$$

Now we assume $2 \leq t \leq m$, that means some indices in (2.12) coincide and so we have to deal with shorter patterns. From (2.11) we know that (for fixed a) the number of patterns (2.12) satisfying the assumptions (2.10) in (2.13) is at least

$$\begin{aligned} &p/2^{2k+4} - (k+2)p^{1/2}, \quad k \leq t-2, \\ &p/2^{k+t+2} - \frac{k+t+2}{2}p^{1/2}, \quad k \geq t-1. \end{aligned}$$

Similarly as before if $a = 1$, then

$$s_{n,t} = \begin{cases} 1 & \text{if } \ell_n \neq \ell_{n+t}, \\ 0 & \text{if } \ell_n = \ell_{n+t}, \end{cases}$$

and if $a = 0$, then

$$s_{n,t} = \begin{cases} 1 & \text{if } \ell_n = \ell_{n+t}, \\ 0 & \text{if } \ell_n \neq \ell_{n+t}. \end{cases}$$

For each case we have 2^{k+1} possible choices for the pattern (2.12) if $k \leq t-2$ and 2^{t-1} possible choices if $k \geq t-1$ and thus in total we count at least

$$\begin{aligned} &p/2^{k+2} - (k+2)2^{k+2}p^{1/2}, \quad k \leq t-2, \\ &p/2^{k+2} - (k+t+2)2^{t-1}p^{1/2}, \quad k \geq t-1, \end{aligned}$$

different $p \leq n < 2p$ with $\ell_{n-k-1} \neq \ell_{n-k-1+t}$, $(\ell_{n-k+j}, \ell_{n-k+j+t}) \in \{(0,0),(1,1)\}$ for $j = 0, \ldots, k-1$ and $s_{n,t} = 1$.

Put $m' = 2m - t + 1$. Summing up all the contributions we get

$$\begin{aligned} N_1 &\geq \frac{1}{4}\left(\sum_{k=0}^{m'-1} 2^{-k}\right)p - 2\left(\sum_{k=0}^{t-2} 2^{k+1}(k+2) + 2^{t-2}\sum_{k=t-1}^{m'-1}(k+t+2)\right)p^{1/2} \\ &= \frac{p}{2} - 2^{-m'-1}p - 2(2^t(t-1) + 2^{t-3}((m')^2 + (2t+3)m' + 2 + t - 3t^2))p^{1/2} \\ &= \frac{p}{2} - 2^{-m'-1}p - 2^{t-2}((m'+t-1)^2 + 5m' - 4t^2 + 11t - 7)p^{1/2} \\ &\geq \frac{p}{2} - 2^{-2m+t-2}p - 2^{t-2}(4m^2 - 4t^2 + 16m)p^{1/2} \\ &\geq \frac{p}{2} - 2^{-2m+t-2}p - 2^{t-1}(2(m^2 - t^2) + 2\log_2 p)p^{1/2}, \end{aligned}$$

where we used $m \leq 1/4 \log_2 p$. Thus by the definition of m

$$N_1 \geq \frac{p}{2} - 2^{t-1}(4 \log_2 p + 2(m^2 - t^2))p^{1/2}.$$

Analogously N_0 can be bounded below by

$$N_0 \geq \frac{p}{2} - 2^{t-1}(4 \log_2 p + 2(m^2 - t^2))p^{1/2}$$

and therefore

$$|A(t)| = |N_0 - N_1| \leq 2^t(4 \log_2 p + 2(m^2 - t^2))p^{1/2}.$$

Thus the result follows. □

Remark 2.33. Sequences with *ideal arithmetic autocorrelation* equal to zero for all nontrivial shifts t are known, see [14]. However, the maximum absolute value of the (periodic) autocorrelation of these so-called *ℓ-sequences* equals the period since the second half of a period is the bit-wise complement of the first half [14, Proposition 1]. Hence, these sequences are far away from looking random. In contrast to these sequences, the Legendre sequence of (almost) perfect (periodic) autocorrelation still guarantees a rather small arithmetic autocorrelation with respect to its period p if p is sufficiently large.

The following table of maximum absolute values of the arithmetic autocorrelation of the Legendre sequence of period p for all primes $p < 150$ may lead to the conjecture that it is bounded by $p^{1/2} \ln p$ which we actually checked for all primes $p < 1000$:

p	3	5	7	11	13	17	19	23	29	31	37	41
$\max_{1 \leq t < p} \lvert A(t) \rvert$	1	3	3	5	7	7	9	9	7	13	15	15
$\lfloor p^{1/2} \ln p \rfloor$	1	3	5	7	9	11	12	15	18	19	21	23

p	43	47	53	59	61	67	71	73	79	83	89	97
$\max_{1 \leq t < p} \lvert A(t) \rvert$	17	15	13	17	15	17	17	13	23	21	21	27
$\lfloor p^{1/2} \ln p \rfloor$	24	26	28	31	32	34	35	36	38	40	42	45

p	101	103	107	109	113	127	131	137	139	149
$\max_{1 \leq t < p} \lvert A(t) \rvert$	21	23	23	21	25	35	29	27	27	27
$\lfloor p^{1/2} \ln p \rfloor$	46	47	48	48	50	54	55	57	58	61

2.3 Arithmetic autocorrelation and correlation measure

In this section we prove a relation between the arithmetic autocorrelation and the correlation measure of order k. Roughly speaking, we show that any binary sequence with small correlation measure of order k up to a sufficiently large k cannot have a large arithmetic autocorrelation. We apply our result to several classes of sequences including Legendre sequences defined with polynomials.

2.3.1 Correlation measure of order k

We start with the definition of a more general notion of the (periodic) autocorrelation.

Definition 2.34. The *(periodic) correlation measure of order* $k \geq 1$ of a (purely) T-periodic binary sequence $(a_n)_{n\geq 0}$ is defined as

$$C_k(a_n) = \max_{0<d_1<\ldots<d_{k-1}<T} \left| \sum_{n=0}^{T-1} (-1)^{a_n+a_{n+d_1}+\cdots+a_{n+d_{k-1}}} \right|.$$

Note that for any (purely) T-periodic binary sequence $(a_n)_{n\geq 0}$ we have $C_1(a_n) = |I(a_n)|$ and $C_2(a_n)$ is simply the maximum over all $0 < d_1 < T$ of the absolute value of the (periodic) autocorrelation of $(a_n)_{n\geq 0}$.

For the distribution of patterns of length k

$$\mathcal{P}_{i_0,i_1,\ldots,i_{k-1}}(a_n) = |\{0 \leq n \leq T-1 : a_n = i_0, a_{n+d_1} = i_1, \ldots, a_{n+d_{k-1}} = i_{k-1}\}|$$

in (purely) T-periodic binary sequences $(a_n)_{n\geq 0}$ Mauduit and Sárközy [29] proved the aperiodic analogue of the following proposition.

Proposition 2.35. *For $k \geq 1$ we have*

$$\left| \mathcal{P}_{i_0,i_1,\ldots,i_{k-1}}(a_n) - \frac{T}{2^k} \right| \leq \frac{1}{2^k} \sum_{l=1}^{k} \binom{k}{l} C_l(a_n).$$

Proof. (see [29]). Put $d_0 = 0$. Then

$$\begin{aligned}
\mathcal{P}_{i_0,i_1,\ldots,i_{k-1}}(a_n) &= \sum_{n=0}^{T-1} \prod_{l=1}^{k} \frac{(-1)^{a_{n+d_{l-1}}+i_{l-1}}+1}{2} \\
&= \sum_{n=0}^{T-1} \frac{1}{2^k} \left(1 + \sum_{l=1}^{k} \sum_{1\leq j_1<\ldots<j_l\leq k} \prod_{r=j_1}^{j_l} (-1)^{a_{n+d_{r-1}}+i_{r-1}} \right) \\
&= \frac{1}{2^k} \left(T + \sum_{l=1}^{k} \sum_{1\leq j_1<\ldots<j_l\leq k} \prod_{r=j_1}^{j_l} (-1)^{i_{r-1}} \sum_{n=0}^{T-1} (-1)^{a_{n+d_{j_1-1}}+\cdots+a_{n+d_{j_l-1}}} \right).
\end{aligned}$$

Since $(a_n)_{n\geq 0}$ is a (purely) T-periodic binary sequence it follows

$$\left|\mathcal{P}_{i_0,i_1,\ldots,i_{k-1}}(a_n) - \frac{T}{2^k}\right| \leq \frac{1}{2^k}\left(\sum_{l=1}^{k}\sum_{1\leq j_1<\ldots<j_l\leq k}\left|\sum_{n=0}^{T-1}(-1)^{a_{n+d_{j_1-1}}+\cdots+a_{n+d_{j_l-1}}}\right|\right)$$
$$\leq \frac{1}{2^k}\sum_{l=1}^{k}\binom{k}{l}C_l(a_n). \qquad \square$$

2.3.2 A bound on the arithmetic autocorrelation

Now we estimate the arithmetic autocorrelation of a binary sequence of period T in terms of correlation measures.

Theorem 2.36. *Put*

$$\Gamma_s = \max_{1\leq l\leq s} C_l(a_n).$$

Then the arithmetic autocorrelation function of a T-periodic binary sequence $(a_n)_{n\geq 0}$ satisfies

$$A(t) \ll \min\left\{T^{1/2}\Gamma_{\lfloor \log T\rfloor}^{1/2}, 2^r\Gamma_{\lfloor \log T\rfloor}\log T\right\},$$

where $r = \min\{t, T-t\}$ for $1 \leq t \leq T-1$.

Proof. The proof is very similar to the proof of Theorem 2.32. By the symmetry $A(t) = -A(T-t)$ of the arithmetic autocorrelation (see Proposition 2.31) we may assume $1 \leq t \leq \lfloor T/2\rfloor$. In the following we derive a lower bound on the number N_1 of ones in a period of the T-periodic sequence $(s_{n,t})_{n\geq 0}$ defined by (2.7).

Take $c \in \{0,1\}$. For some k and n with $1 \leq k < m$ and $T \leq n < 2T$ assume

$$\begin{aligned} &(a_{n-k}, a_{n-k+t}) = (c, 1-c),\\ &a_{n-k+j} = a_{n-k+j+t}, \quad j = 1,\ldots,k-1,\\ &(a_n, a_{n+t}) \in \{0,1\}^2. \end{aligned} \tag{2.14}$$

We consider only patterns of length $4 \leq s = 2k+2$ and therefore it follows from Proposition 2.35 that

$$\left|\mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(a_n) - \frac{T}{2^s}\right| \leq \frac{1}{2^s}\sum_{l=1}^{s}\binom{s}{l}C_l(a_n) \leq \max_{1\leq l\leq s} C_l(a_n) = \Gamma_s. \tag{2.15}$$

First we assume $m+1 \leq t \leq \lfloor T/2\rfloor$. From (2.15) we know that (for fixed c) the number of patterns

$$\begin{pmatrix} a_{n-k} & a_{n-k+1} & \ldots & a_{n-1} & a_n\\ a_{n-k+t} & a_{n-k+t+1} & \ldots & a_{n-1+t} & a_{n+t}\end{pmatrix} \tag{2.16}$$

satisfying the assumptions (2.14) in

$$\begin{array}{cccccccc} a_{T-k} & a_{T-k+1} & \dots & a_{T-1} & a_T & \dots & a_{2T-2} & a_{2T-1} \\ a_{t+T-k} & a_{t+T-k+1} & \dots & a_{t+T-1} & a_{t+T} & \dots & a_{t+2T-2} & a_{t+2T-1} \end{array} \tag{2.17}$$

is at least $T/2^{2k+2} - \Gamma_{2k+2}$. We have to distinguish between two cases.

If $c = 1$, then $(a_{n-k}, a_{n-k+t}) = (1, 0)$. The subtraction of 0 from 1 gives no carry, no matter if there was a carry in the previous step. Hence

$$s_{n,t} = \begin{cases} 1 & \text{if } a_n \neq a_{n+t}, \\ 0 & \text{if } a_n = a_{n+t}. \end{cases}$$

Since there are 2^k possible choices for the pattern (2.16) we count at least $T/2^{k+2} - 2^k\Gamma_{2k+2}$ different $T \leq n < 2T$ with $s_{n,t} = 1$.

If $c = 0$, then $(a_{n-k}, a_{n-k+t}) = (0, 1)$. The subtraction of 1 from 0 gives a carry, no matter if there was a carry in the previous step. Hence

$$s_{n,t} = \begin{cases} 1 & \text{if } a_n = a_{n+t}, \\ 0 & \text{if } a_n \neq a_{n+t}. \end{cases}$$

Just as before there are 2^k possible choices for the pattern (2.16) and so we get at least $T/2^{k+2} - 2^k\Gamma_{2k+2}$ additional n with $s_{n,t} = 1$.

Thus in total we have at least $T/2^{k+1} - 2^{k+1}\Gamma_{2k+2}$ different $T \leq n < 2T$ with $a_{n-k} \neq a_{n-k+t}$, $(a_{n-k+j}, a_{n-k+j+t}) \in \{(0,0), (1,1)\}$ for $j = 1, \dots, k-1$ and $s_{n,t} = 1$.

Summing up all the contributions we get

$$\begin{aligned} N_1 &\geq \frac{1}{2}\left(\sum_{k=1}^{m-1} 2^{-k}\right) T - 2\left(\sum_{k=1}^{m-1} 2^k\Gamma_{2k+2}\right) \\ &\geq \frac{1}{2}\left(\sum_{k=1}^{m-1} 2^{-k}\right) T - 2\Gamma_{2m}\left(\sum_{k=1}^{m-1} 2^k\right) \geq \frac{T}{2} - 2^{-m}T - 2^{m+1}\Gamma_{2m}, \end{aligned}$$

where we used $\Gamma_s = \Gamma_{2k+2} \leq \Gamma_{2m}$ since $k \leq m-1$. Analogously N_0 can be bounded below by

$$N_0 \geq \frac{T}{2} - 2^{-m}T - 2^{m+1}\Gamma_{2m}$$

and therefore

$$|A(t)| = |N_0 - N_1| \leq 2^{-m+1}T + 2^{m+2}\Gamma_{2m}$$

since

$$\begin{aligned} N_0 - N_1 &= T - 2N_1 \leq 2^{-m+1}T + 2^{m+2}\Gamma_{2m}, \\ N_0 - N_1 &= 2N_0 - T \geq -(2^{-m+1}T + 2^{m+2}\Gamma_{2m}). \end{aligned}$$

Now we assume $1 \le t \le m$, that means some indices in (2.16) coincide and so we have to deal with shorter patterns. From (2.15) we know that (for fixed c) the number of patterns (2.16) satisfying the assumptions (2.14) in (2.17) is at least

$$\begin{aligned} &T/2^{2k+2} - \Gamma_{2k+2}, \quad k \le t-1, \\ &T/2^{k+t+1} - \Gamma_{k+t+1}, \quad k \ge t. \end{aligned}$$

Similarly as before if $c = 1$, then

$$s_{n,t} = \begin{cases} 1 & \text{if } a_n \ne a_{n+t}, \\ 0 & \text{if } a_n = a_{n+t}, \end{cases}$$

and if $c = 0$, then

$$s_{n,t} = \begin{cases} 1 & \text{if } a_n = a_{n+t}, \\ 0 & \text{if } a_n \ne a_{n+t}. \end{cases}$$

For each case we have 2^k possible choices for the pattern (2.16) if $k \le t-1$ and 2^{t-1} possible choices if $k \ge t$ and thus in total we count at least

$$\begin{aligned} &T/2^{k+1} - 2^{k+1}\Gamma_{2k+2}, \quad k \le t-1, \\ &T/2^{k+1} - 2^{t}\Gamma_{k+t+1}, \quad k \ge t, \end{aligned}$$

different $T \le n < 2T$ with $a_{n-k} \ne a_{n-k+t}$, $(a_{n-k+j}, a_{n-k+j+t}) \in \{(0,0),(1,1)\}$ for $j = 1, \dots, k-1$ and $s_{n,t} = 1$.

Put $m' = 2m - t$. Summing up all the contributions we get

$$\begin{aligned} N_1 &\ge \frac{1}{2}\left(\sum_{k=1}^{m'-1} 2^{-k}\right) T - 2\left(\sum_{k=1}^{t-1} 2^k \Gamma_{2k+2} + 2^{t-1}\sum_{k=t}^{m'-1} \Gamma_{k+t+1}\right) \\ &\ge \frac{1}{2}\left(\sum_{k=1}^{m'-1} 2^{-k}\right) T - 2\Gamma_{2m}(2^t - 2) - 2^t\Gamma_{m+t}(m' - t) \\ &\ge \frac{T}{2} - 2^{-m'}T - 2^{t+1}\Gamma_{2m} - 2^t\Gamma_{2m}(m' - t) \\ &\ge \frac{T}{2} - 2^{-2m+t}T - 2^{t+1}(m - t + 1)\Gamma_{2m}, \end{aligned}$$

where we used $\Gamma_{2k+2} \le \Gamma_{2m}$ or $\Gamma_{k+t+1} \le \Gamma_{m+t} \le \Gamma_{2m}$, respectively. Analogously N_0 can be bounded below by

$$N_0 \ge \frac{T}{2} - 2^{-2m+t}T - 2^{t+1}(m - t + 1)\Gamma_{2m}$$

and therefore

$$|A(t)| = |N_0 - N_1| \leq 2^{-2m+t+1}T + 2^{t+2}(m-t+1)\Gamma_{2m}.$$

Choosing

$$m = \left\lfloor \frac{1}{2} \log \frac{T}{\Gamma_{\lfloor \log T \rfloor}} \right\rfloor$$

we obtain the result. (Note that we may assume $\Gamma_{\lfloor \log T \rfloor} = o(T)$ and thus $m \geq 2$ since otherwise the result is trivial.) □

2.3.3 Applications

For a squarefree polynomial $f(x) \in \mathbb{F}_p[x]$ with positive degree d let the p-periodic sequences $(l_n)_{n\geq 0}$ be defined by

$$l_n = \begin{cases} 1 & \text{if } \left(\frac{f(n)}{p}\right) = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.18}$$

For $f(n) = n$ these sequences are the Legendre sequences.

Corollary 2.37. *If $d < 0.5 \log p / \log\log p$ or 2 is a primitive root modulo p, then the arithmetic autocorrelation function of the p-periodic sequences $(l_n)_{n\geq 0}$ defined by* (2.18) *satisfies*

$$A(t) \ll \min\left\{ d^{1/2} p^{3/4} (\log p)^{1/2}, 2^r d p^{1/2} (\log p)^2 \right\},$$

where $r = \min\{t, p-t\}$ for $1 \leq t \leq p-1$.

This result immediately follows from Theorem 2.36 and from the following proposition.

Proposition 2.38. *If $f(x)$ has no multiple zeros in the algebraic closure of $\mathbb{F}_p$ and*

(i) $k < p$ and 2 is a primitive root modulo p, or

(ii) $(4k)^d < p$,

then the (periodic) correlation measure of order k satisfies $C_k(l_n) \ll k d p^{1/2}$.

Proof. Similar to the proof of Theorem 1 in [18] (since we consider the periodic correlation measure of order k instead of the aperiodic one we lose the $\log p$ term). □

Let q be the power of an odd prime, let g be a primitive element of $\mathbb{F}_q$, and let η denote the quadratic character of $\mathbb{F}_q$, that is

$$\eta(a) = \begin{cases} 0 & \text{if } a = 0, \\ 1 & \text{if } a \text{ is a square of an element of } \mathbb{F}_q^*, \\ -1 & \text{if } a \text{ is not a square of an element of } \mathbb{F}_q^*, \end{cases}$$

where $a \in \mathbb{F}_q$. We denote by $(u_n)_{n\geq 0}$ the $(q-1)$-periodic *Sidelnikov-Lempel-Cohn-Eastman sequence* defined by

$$u_n = \begin{cases} 1 & \text{if } \eta(g^n + 1) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Similar to the proof in [5, Lemma 1] it follows that the (periodic) correlation measure of order k satisfies $C_k(u_n) \ll kq^{1/2}$. From Theorem 2.36 we get

$$A(t) \ll \min\left\{q^{3/4}(\log q)^{1/2}, 2^r q^{1/2}(\log q)^2\right\},$$

where $r = \min\{t, q-1-t\}$ for $1 \leq t \leq q-2$.

Let $\lambda \in \mathbb{F}_p^*$ be of multiplicative order T and let $f(x) \in \mathbb{F}_p[x]$ be a polynomial of positive degree d not of the form $bx^\beta(g(x))^2$ with $b \in \mathbb{F}_p$, $g(x) \in \mathbb{F}_p[x]$ and β a positive integer. Define the T-periodic sequence $(v_n)_{n\geq 0}$ by

$$v_n = \begin{cases} 1 & \text{if } \left(\frac{f(\lambda^n)}{p}\right) = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Similar to the proof in [19, Theorem 2] it follows that if T is a prime and either $(4k)^d \leq T$ or 2 is a primitive root modulo T, then the (periodic) correlation measure of order k satisfies $C_k(v_n) \ll kdp^{1/2}$. From Theorem 2.36 we get that if $d \leq \log p/\log\log p$ or 2 is a primitive root modulo T, then

$$A(t) \ll \min\left\{d^{1/2}p^{1/4}T^{1/2}(\log T)^{1/2}, 2^r dp^{1/2}(\log T)^2\right\},$$

where $r = \min\{t, T-t\}$ for $1 \leq t \leq T-1$.

Remark 2.39. For fixed $1 \leq t < T$, Goresky and Klapper [15, 16] proved that the expected arithmetic autocorrelation, averaged over all binary sequences of period T, is

$$\frac{T}{2^{T-\gcd(t,T)}}.$$

Actually, the correlation measure of order k was defined for finite sequences. Analogs of our results for finite sequences can be easily obtained with the obvious definition of arithmetic autocorrelation. Moreover, for a truly random sequence of length T, Alon et al. [2] showed that the correlation measure of order k is of order of magnitude $k^{1/2}T^{1/2}(\log T)^{1/2}$ and thus its arithmetic autocorrelation is at most of order of magnitude $T^{3/4}(\log T)^{1/2}$.

Chapter 3

Complexity measures

Expansion complexity introduced by Diem [10] and linear complexity are both measures for the unpredictability of a sequence. Sequences with small linear complexity or expansion complexity are predictable and thus not suitable in cryptography. Expansion complexity is essentially the same as linear complexity in the periodic case but finer in the aperiodic case [31]. For more background and results on linear complexity and related measures see [30, 34, 41, 43].

Throughout this chapter let q be a power of the prime number p.

3.1 Preliminaries

In this section we provide some basic properties of the linear complexity and expansion complexity. A large linear complexity or expansion complexity is necessary but not sufficient for cryptographic applications. For more details, see [9], [23], [31] and [43].

3.1.1 Linear complexity and expansion complexity

First of all we have to define the linear complexity and expansion complexity.

Definition 3.1. Let k be a positive integer. A sequence $\mathcal{S} = (s_n)_{n\geq 0}$ over $\mathbb{F}_q$ is called a *linear recurring sequence* if it satisfies a linear recurrence relation

$$s_{n+k} + \alpha_{k-1}s_{n+k-1} + \alpha_{k-2}s_{n+k-2} + \cdots + \alpha_0 s_n = 0, \quad n \geq 0,$$

for some $\alpha_0, \alpha_1, \ldots, \alpha_{k-1} \in \mathbb{F}_q$.

For more information on linear recurring sequences see [25, Chapter 8].

Definition 3.2. For a positive integer N, the *Nth linear complexity* $L_N(\mathcal{S})$ of a sequence $\mathcal{S} = (s_n)_{n\geq 0}$ over $\mathbb{F}_q$ is defined as the length of a shortest linear recurrence relation

$$s_{n+L_N(\mathcal{S})} + \sum_{\ell=0}^{L_N(\mathcal{S})-1} \alpha_\ell s_{n+\ell} = 0, \quad 0 \leq n \leq N - L_N(\mathcal{S}) - 1,$$

for some $\alpha_\ell \in \mathbb{F}_q$, which is satisfied by the first N elements of the sequence. We use the convention

$$\begin{aligned} &L_N(\mathcal{S}) = 0 \quad \text{if } s_n = 0 \text{ for } 0 \leq n \leq N-1, \\ &L_N(\mathcal{S}) = N \quad \text{if } s_n = 0 \text{ for } 0 \leq n \leq N-2 \text{ and } s_{N-1} \neq 0. \end{aligned} \tag{3.1}$$

The *linear complexity* $L(\mathcal{S})$ is

$$L(\mathcal{S}) = \sup_{N\geq 1} L_N(\mathcal{S}).$$

It is well-known, see for example [25], that the linear complexity $L(\mathcal{S})$ is finite if and only if $\mathcal{S}$ is ultimately periodic, that is, $\mathcal{S}$ is a linear recurring sequence.

Definition 3.3. The *generating function* $G_{\mathcal{S}}(x)$ of a sequence $\mathcal{S} = (s_n)_{n\geq 0}$ over $\mathbb{F}_q$ is

$$G_{\mathcal{S}}(x) = \sum_{n=0}^{\infty} s_n x^n.$$

We call a generating function $G_{\mathcal{S}}(x)$ *algebraic* over $\mathbb{F}_q[x]$ if there exists a nonzero polynomial $h(x, y) \in \mathbb{F}_q[x, y]$ such that $h(x, G_{\mathcal{S}}(x)) = 0$.

Definition 3.4 (Diem, [10])**.** For a positive integer N, the *Nth expansion complexity* $E_N(\mathcal{S})$ of a sequence $\mathcal{S} = (s_n)_{n\geq 0}$ over $\mathbb{F}_q$ is defined as the least total degree of a nonzero polynomial $h(x, y) \in \mathbb{F}_q[x, y]$ with

$$h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^N}.$$

We use the convention

$$E_N(\mathcal{S}) = 0 \quad \text{if } s_n = 0 \text{ for } 0 \leq n \leq N-1.$$

The *expansion complexity* $E(\mathcal{S})$ is

$$E(\mathcal{S}) = \sup_{N\geq 1} E_N(\mathcal{S}).$$

Note that $E_N(\mathcal{S})$ depends only on the first N elements of $\mathcal{S}$. By a famous result of Christol [7, 8] the expansion complexity $E(\mathcal{S})$ is finite, that is $G_{\mathcal{S}}(x)$ is algebraic, if and only if $\mathcal{S}$ is an *automatic sequence*. For more details on automatic sequences we refer to the monograph of Allouche and Shallit [1].

3.1.2 Growth of $L_N(\mathcal{S})$ and $E_N(\mathcal{S})$

We start with the description of the possible growth of the nondecreasing function $N \mapsto E_N(\mathcal{S})$.

Proposition 3.5 (Mérai, Niederreiter, Winterhof, [31]). *Let $\mathcal{S} = (s_n)_{n\geq 0}$ be a sequence over $\mathbb{F}_q$. Then*

$$E_N(\mathcal{S}) \leq E_{N+1}(\mathcal{S}) \leq E_N(\mathcal{S}) + 1.$$

Proof. (see [31]). Evidently the function $N \mapsto E_N(\mathcal{S})$ is nondecreasing since $h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^{N+1}}$ implies $h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^N}$.

If $h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^N}$, then

$$xh(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^{N+1}},$$

from which the second inequality follows. □

For comparison, we state the corresponding result on the possible growth of the nondecreasing function $N \mapsto L_N(\mathcal{S})$ called the *linear complexity profile* of $\mathcal{S}$. But first we prove the following lemma.

Lemma 3.6. *Let $\mathcal{S} = (s_n)_{n\geq 0}$ and $\mathcal{T} = (t_n)_{n\geq 0}$ be sequences over $\mathbb{F}_q$. Then*

$$L_N(\mathcal{S} + \mathcal{T}) \leq L_N(\mathcal{S}) + L_N(\mathcal{T}).$$

Proof. (see [23]). Put $U = L_N(\mathcal{S})$ and $V = L_N(\mathcal{T})$. Let

$$s_{n+U} + \sum_{i=0}^{U-1} \alpha_i s_{n+i} = 0, \quad 0 \leq n \leq N - U - 1,$$

be a shortest linear recurrence relation satisfied by the first N elements of $\mathcal{S}$ and let

$$t_{n+V} + \sum_{j=0}^{V-1} \beta_j t_{n+j} = 0, \quad 0 \leq n \leq N - V - 1,$$

be a shortest linear recurrence relation satisfied by the first N elements of $\mathcal{T}$. Then $\mathcal{S}$ and $\mathcal{T}$ satisfy the linear recurrence relations

$$s_{n+U+V} + \sum_{i=0}^{U-1} \alpha_i s_{n+i+V} = 0, \quad 0 \leq n \leq N - U - V - 1,$$

and

$$t_{n+U+V} + \sum_{j=0}^{V-1} \beta_j t_{n+j+U} = 0, \quad 0 \leq n \leq N - U - V - 1,$$

as well, respectively. Adding the last two equations we get

$$\begin{aligned}
&s_{n+U+V} + t_{n+U+V} + \sum_{i=0}^{U-1} \alpha_i (s_{n+i+V} + t_{n+i+V}) + \sum_{j=0}^{V-1} \beta_j (s_{n+j+U} + t_{n+j+U}) \\
&= \sum_{i=0}^{U-1} \alpha_i t_{n+i+V} + \sum_{j=0}^{V-1} \beta_j s_{n+j+U} = -\sum_{i=0}^{U-1} \alpha_i \sum_{j=0}^{V-1} \beta_j t_{n+i+j} - \sum_{j=0}^{V-1} \beta_j \sum_{i=0}^{U-1} \alpha_i s_{n+i+j} \\
&= -\sum_{i=0}^{U-1} \sum_{j=0}^{V-1} \alpha_i \beta_j (s_{n+i+j} + t_{n+i+j}), \quad 0 \le n \le N-U-V-1,
\end{aligned}$$

that is a linear recurrence relation (but not necessarily the shortest one) of length $U+V$ for the first N elements of $\mathcal{S}+\mathcal{T}$. Thus the result follows. □

The proof of the following proposition is constructive and provides the well-known Berlekamp-Massey algorithm [4, 27] for the calculation of the linear complexity profile of $\mathcal{S}$.

Proposition 3.7. *Let $\mathcal{S} = (s_n)_{n\ge 0}$ be a sequence over $\mathbb{F}_q$. If $L_N(\mathcal{S}) > N/2$, then*

$$L_{N+1}(\mathcal{S}) = L_N(\mathcal{S}).$$

If $L_N(\mathcal{S}) \le N/2$, then

$$L_{N+1}(\mathcal{S}) \in \{L_N(\mathcal{S}), N+1-L_N(\mathcal{S})\}.$$

Proof. (see [43]). Let

$$s_{n+L_N(\mathcal{S})} + \sum_{\ell=0}^{L_N(\mathcal{S})-1} \alpha_\ell s_{n+\ell} = 0, \quad 0 \le n \le N - L_N(\mathcal{S}) - 1, \tag{3.2}$$

be a shortest linear recurrence relation satisfied by the first N elements of $\mathcal{S}$. Put

$$\mu = s_N + \alpha_{L_N(\mathcal{S})-1} s_{N-1} + \cdots + \alpha_0 s_{N-L_N(\mathcal{S})}.$$

If $\mu = 0$, then the linear recurrence relation (3.2) holds for $n = N - L_N(\mathcal{S})$ as well and we have $L_{N+1}(\mathcal{S}) = L_N(\mathcal{S})$.

Otherwise if $\mu \ne 0$ we define the first $N+1$ elements of a sequence $\mathcal{T} = (t_n)_{n\ge 0}$ over $\mathbb{F}_q$ by

$$t_n = s_n, \quad n = 0, \ldots, N-1 \quad \text{and} \quad t_N = s_N - \mu.$$

Then $\mathcal{T} = (t_n)_{n\ge 0}$ satisfies the linear recurrence relation

$$t_{n+L_N(\mathcal{T})} + \sum_{\ell=0}^{L_N(\mathcal{T})-1} \alpha_\ell t_{n+\ell} = 0, \quad 0 \le n \le N - L_N(\mathcal{T}),$$

and we have $L_{N+1}(\mathcal{T}) = L_N(\mathcal{T})$. Hence we get

$$\begin{aligned} N+1 &= L_{N+1}(\mathcal{S}-\mathcal{T}) \le L_{N+1}(\mathcal{S}) + L_{N+1}(-\mathcal{T}) \\ &= L_{N+1}(\mathcal{S}) + L_{N+1}(\mathcal{T}) = L_{N+1}(\mathcal{S}) + L_N(\mathcal{T}) = L_{N+1}(\mathcal{S}) + L_N(\mathcal{S}), \end{aligned}$$

where we used (3.1), Lemma 3.6 and the fact that the sequences $\mathcal{T}$ and $-\mathcal{T}$ trivially have the same linear complexity profile. Since the linear complexity profile of $\mathcal{S}$ is nondecreasing, that is $L_{N+1}(\mathcal{S}) \ge L_N(\mathcal{S})$, we obtain

$$L_{N+1}(\mathcal{S}) \ge \max\{L_N(\mathcal{S}), N+1-L_N(\mathcal{S})\}.$$

We show equality by induction on N. We may assume that there is a positive integer $M \le N-1$ such that

$$L_N(\mathcal{S}) = L_{N-1}(\mathcal{S}) = \ldots = L_{M+1}(\mathcal{S}) > L_M(\mathcal{S}).$$

Then by induction hypothesis we have

$$L_{M+1}(\mathcal{S}) = \max\{L_M(\mathcal{S}), M+1-L_M(\mathcal{S})\} = M+1-L_M(\mathcal{S}),$$

that is $L_M(\mathcal{S}) = M+1-L_{M+1}(\mathcal{S})$. Let

$$s_{n+M+1-L_{M+1}(\mathcal{S})} + \sum_{\ell=0}^{M-L_{M+1}(\mathcal{S})} \beta_\ell s_{n+\ell} = 0, \quad 0 \le n \le L_{M+1}(\mathcal{S}) - 2,$$

be a linear recurrence relation satisfied by the first M elements of $\mathcal{S}$ and put

$$\nu = s_M + \beta_{M-L_{M+1}(\mathcal{S})} s_{M-1} + \cdots + \beta_0 s_{L_{M+1}(\mathcal{S})-1} \ne 0.$$

If $L_N(\mathcal{S}) > N/2$, then

$$s_{n+L_N(\mathcal{S})} + \sum_{\ell=0}^{L_N(\mathcal{S})-1} \alpha_\ell s_{n+\ell} - R_1 = 0, \quad 0 \le n \le N - L_N(\mathcal{S}),$$

with

$$R_1 := \mu\nu^{-1} \left(s_{n+M-N+L_{M+1}(\mathcal{S})} + \sum_{\ell=0}^{M-L_{M+1}(\mathcal{S})} \beta_\ell s_{n+\ell-N+2L_{M+1}(\mathcal{S})-1} \right),$$

is a linear recurrence relation of length $L_N(\mathcal{S})$, and if $L_N(\mathcal{S}) \le N/2$, then

$$s_{n+N+1-L_N(\mathcal{S})} + \sum_{\ell=0}^{L_N(\mathcal{S})-1} \alpha_\ell s_{n+\ell+N-2L_N(\mathcal{S})+1} - R_2 = 0, \quad 0 \le n \le L_N(\mathcal{S}) - 1,$$

with

$$R_2 := \mu\nu^{-1} \left(s_{n+M+1-L_{M+1}(\mathcal{S})} + \sum_{\ell=0}^{M-L_{M+1}(\mathcal{S})} \beta_\ell s_{n+\ell} \right),$$

is a linear recurrence relation of length $N+1-L_N(\mathcal{S})$ for the first $N+1$ elements of $\mathcal{S}$. This completes the proof. □

3.1.3 Purely periodic sequences

In this paragraph let $\mathcal{S} = (s_n)_{n\geq 0}$ be a (purely) T-periodic sequence over $\mathbb{F}_q$. Then $L(\mathcal{S})$ is finite and obviously $L(\mathcal{S}) \leq T$. In fact, the linear complexity can be computed explicitly as the following proposition shows.

Proposition 3.8. *Put*

$$S^T(x) = \sum_{n=0}^{T-1} s_n x^n.$$

Then

$$L(\mathcal{S}) = T - \deg(\gcd(S^T(x), 1 - x^T)).$$

Proof. (see [9]). We have

$$(1 - x^T) \sum_{n=0}^{\infty} s_n x^n = \sum_{n=0}^{\infty} s_n x^n - \sum_{n=T}^{\infty} s_{n-T} x^n = S^T(x)$$

since $\mathcal{S}$ is (purely) T-periodic and hence

$$\frac{(1 - x^T)}{\gcd(S^T(x), 1 - x^T)} \sum_{n=0}^{\infty} s_n x^n = \frac{S^T(x)}{\gcd(S^T(x), 1 - x^T)}. \tag{3.3}$$

Comparing the coefficients of both sides of this equation we get a linear recurrence relation of length $T - \deg(\gcd(S^T(x), 1 - x^T))$, which is satisfied by $\mathcal{S}$.

It remains to show that this is the shortest linear recurrence relation satisfied by $\mathcal{S}$. Therefore let

$$s_{n+L(\mathcal{S})} + \sum_{\ell=0}^{L(\mathcal{S})-1} \alpha_\ell s_{n+\ell} = 0, \quad n \geq 0,$$

and put

$$f(x) = \left(1 - \sum_{\ell=0}^{L(\mathcal{S})-1} \alpha_\ell x^{L(\mathcal{S})-\ell}\right) \sum_{n=0}^{\infty} s_n x^n \in \mathbb{F}_q[x]. \tag{3.4}$$

Note that $f(x)$ is a polynomial of degree at most $L(\mathcal{S}) - 1$. Using (3.3) and (3.4) we obtain

$$f(x) \frac{(1 - x^T)}{\gcd(S^T(x), 1 - x^T)} = \left(1 - \sum_{\ell=0}^{L(\mathcal{S})-1} \alpha_\ell x^{L(\mathcal{S})-\ell}\right) \frac{S^T(x)}{\gcd(S^T(x), 1 - x^T)}.$$

Since

$$\frac{(1-x^T)}{\gcd(S^T(x),1-x^T)} \quad \text{and} \quad \frac{S^T(x)}{\gcd(S^T(x),1-x^T)}$$

are relatively prime, $S^T(x)/\gcd(S^T(x),1-x^T)$ divides $f(x)$. Thus

$$L(\mathcal{S})-1 \geq \deg(f) \geq T-1-\deg(\gcd(S^T(x),1-x^T))$$

and the result follows. □

Next we want to prove a relation between linear complexity and expansion complexity for (purely) periodic sequences.

Lemma 3.9. *The generating function $G_{\mathcal{S}}(x)$ of $\mathcal{S}$ is a rational function*

$$G_{\mathcal{S}}(x) = \frac{f(x)}{g(x)}, \quad f,g \in \mathbb{F}_q[x],\ g \neq 0, \tag{3.5}$$

where $\deg(f) < L(\mathcal{S})$ *and* $\deg(g) = L(\mathcal{S})$.

Proof. (see [25]). Let

$$\sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell s_{n+\ell} = 0, \quad n \geq 0, \quad \alpha_{L(\mathcal{S})} = 1,$$

be a shortest linear recurrence relation satisfied by $\mathcal{S}$. Choose

$$g(x) = \sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell x^{L(\mathcal{S})-\ell} \in \mathbb{F}_q[x]$$

and

$$\begin{aligned} f(x) = g(x)G_{\mathcal{S}}(x) &= \sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell x^{L(\mathcal{S})-\ell} \sum_{n=0}^{\infty} s_n x^n = \sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell \sum_{n=0}^{\infty} s_n x^{n+L(\mathcal{S})-\ell} \\ &= \sum_{j=0}^{L(\mathcal{S})-1} \left(\sum_{\ell=L(\mathcal{S})-j}^{L(\mathcal{S})} \alpha_\ell s_{j-L(\mathcal{S})+\ell} \right) x^j + \sum_{j=L(\mathcal{S})}^{\infty} \left(\sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell s_{j-L(\mathcal{S})+\ell} \right) x^j \\ &= \sum_{j=0}^{L(\mathcal{S})-1} \left(\sum_{\ell=L(\mathcal{S})-j}^{L(\mathcal{S})} \alpha_\ell s_{j-L(\mathcal{S})+\ell} \right) x^j \in \mathbb{F}_q[x] \end{aligned}$$

since

$$\sum_{\ell=0}^{L(\mathcal{S})} \alpha_\ell s_{j-L(\mathcal{S})+\ell} = 0, \quad j \geq L(\mathcal{S}).$$

Obviously $\deg(f) < L(\mathcal{S})$ and $g(x) \neq 0$ with $\deg(g) \leq L(\mathcal{S})$. But $\mathcal{S}$ is (purely) T-periodic and therefore $\alpha_0 \neq 0$ since otherwise

$$\sum_{\ell=0}^{L(\mathcal{S})-1} \alpha_{\ell+1} s_{n+\ell} = 0, \quad n \geq 0,$$

would be a shorter linear recurrence relation satisfied by $\mathcal{S}$. Hence we have $\deg(g) = L(\mathcal{S})$. □

Lemma 3.10 (Mérai, Niederreiter, Winterhof, [31])**.** *Let $G_{\mathcal{S}}(x)$ in (3.5) be not identically zero and let $h(x, y) \in \mathbb{F}_q[x, y]$ be a nonzero polynomial of local degree d in y. Put*

$$H(x) = g(x)^d h(x, G_{\mathcal{S}}(x)).$$

If $H(x)$ is the zero polynomial, then the total degree of $h(x, y)$ satisfies

$$\deg(h) \geq L(\mathcal{S}) + 1.$$

Proof. (see [31]). We write

$$h(x, y) = \sum_{i=0}^{d} h_i(x) y^i \in \mathbb{F}_q[x, y]$$

with $h_d(x) \neq 0$. Then $H(x) = 0$ implies

$$g(x)^d \sum_{i=0}^{d} h_i(x) G_{\mathcal{S}}(x)^i = \sum_{i=0}^{d} h_i(x) f(x)^i g(x)^{d-i} = 0,$$

that is

$$\sum_{i=0}^{d-1} h_i(x) f(x)^i g(x)^{d-i} = -h_d(x) f(x)^d,$$

and $d \geq 1$ since otherwise if $d = 0$ we would have $h_d(x) = 0$. Hence, $h_d(x)$ is divisible by $g(x)$ and thus of degree at least $\deg(g) = L(\mathcal{S})$. Finally, we obtain

$$\deg(h) = \deg\left(\sum_{i=0}^{d} h_i y^i\right) \geq \deg(h_d y^d) = \deg(h_d) + d \geq L(\mathcal{S}) + 1. \qquad \square$$

The following proposition shows that expansion complexity and linear complexity are essentially the same for (purely) periodic sequences.

Proposition 3.11 (Mérai, Niederreiter, Winterhof, [31]). *If $G_{\mathcal{S}}(x)$ in (3.5) is not identically zero, then*

$$E_N(\mathcal{S}) \geq \begin{cases} L(\mathcal{S})+1 & \text{for } N > L(\mathcal{S})(L(\mathcal{S})+1), \\ \lceil N/(L(\mathcal{S})+1) \rceil & \text{otherwise,} \end{cases}$$

and

$$E_N(\mathcal{S}) \leq L(\mathcal{S})+1.$$

Proof. (see [31]). Take

$$h(x,y) = \sum_{i=0}^{d} h_i(x) y^i \in \mathbb{F}_q[x,y], \quad h_d(x) \neq 0,$$

and put $H(x) = g(x)^d h(x, G_{\mathcal{S}}(x))$. We may assume $\deg(h) < N/(L(\mathcal{S})+1)$ since otherwise $\deg(h) \geq N/(L(\mathcal{S})+1)$ implies

$$E_N(\mathcal{S}) \geq \lceil N/(L(\mathcal{S})+1) \rceil.$$

Then we have

$$\begin{aligned} \deg(H) &= \deg(g^d) + \deg(h) = L(\mathcal{S})d + \deg(h) \\ &\leq L(\mathcal{S})(\deg(h) - \deg(h_d)) + \deg(h) \leq \deg(h)(L(\mathcal{S})+1) < N \end{aligned}$$

since $\deg(h) \geq \deg(h_d) + d$ and therefore

$$h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^N}$$

is equivalent to $H(x) = 0$. Now the lower bound follows by Lemma 3.10.

Choosing the polynomial

$$h(x,y) = g(x)y - f(x) \in \mathbb{F}_q[x,y]$$

of degree

$$\deg(h) = \max\{\deg(f), \deg(g)+1\} = L(\mathcal{S})+1,$$

which satisfies

$$h(x, G_{\mathcal{S}}(x)) \equiv 0 \pmod{x^N},$$

we get the upper bound. □

Remark 3.12. If $N > L(\mathcal{S})(L(\mathcal{S})+1)$ the lower and upper bound in Proposition 3.11 coincide and we obtain $E_N(\mathcal{S}) = L(\mathcal{S})+1$. However, Proposition 3.11 shows that

$$E_N(\mathcal{S}) \leq E(\mathcal{S}) = L(\mathcal{S})+1$$

for any (purely) periodic sequence $\mathcal{S}$ over $\mathbb{F}_q$.

Proposition 3.11 can be easily extended to ultimately periodic sequences, see [31, Theorem 1], but then the expansion complexity becomes finer by the preperiod than the linear complexity.

3.2 Linear complexity and expansion complexity of some number theoretic sequences

In this section we study the predictability of some number theoretic sequences over finite fields and thus their suitability in cryptography. First we analyze the (not ultimately periodic) binary automatic sequence $\mathcal{T} = (t_n)_{n\geq 0}$ with $t_n = 1$ whenever n is the sum of three integer squares. We show that it has a large Nth linear complexity, which is necessary but not sufficient for unpredictability. However, it also has a very small expansion complexity and thus is rather predictable.

Next we study p-periodic sequences of binomial coefficients over $\mathbb{F}_p$. In particular, we prove that some linear combinations of p-periodic sequences of binomial coefficients modulo p have a very small expansion complexity and are predictable despite of a high linear complexity. As an application we consider the Legendre sequence and verify that it does not belong to this class of predictable sequences.

Finally, we analyze the expansion complexity of t-periodic sequences over $\mathbb{F}_q$ where $t \mid q-1$.

3.2.1 The characteristic sequence of the set of sums of three squares

We define the (not ultimately periodic) automatic sequence $\mathcal{T} = (t_n)_{n\geq 0}$ over $\mathbb{F}_2$ by

$$t_n = \begin{cases} 1 & \text{if } n = u^2+v^2+w^2 \text{ for some integers } u,v,w, \\ 0 & \text{otherwise.} \end{cases}$$

By the Three-Square Theorem, see for example [3], this is equivalent to

$$t_n = \begin{cases} 0 & \text{if there exist nonnegative integers } a,k \\ & \text{such that } n = 4^a(8k+7), \\ 1 & \text{otherwise.} \end{cases}$$

Theorem 3.13. *We have*

$$E(\mathcal{T}) \leq 12,$$

and

$$L_N(\mathcal{T}) \geq (N-7)/4.$$

Proof. The generating function $G_{\mathcal{T}}(x)$ of $\mathcal{T}$ is

$$G_{\mathcal{T}}(x) = \sum_{n=0}^{\infty} t_n x^n = \sum_{n=0}^{\infty} x^n + \sum_{a=0}^{\infty} \sum_{k=0}^{\infty} x^{4^a(8k+7)}$$
$$= \frac{1}{x+1} + \sum_{a=0}^{\infty} \left(x^7 \sum_{k=0}^{\infty} x^{8k} \right)^{4^a} = \frac{1}{x+1} + \sum_{a=0}^{\infty} \left(\frac{x^7}{(x+1)^8} \right)^{4^a}.$$

We have

$$G_{\mathcal{T}}(x) + G_{\mathcal{T}}(x)^4 = \frac{1}{x+1} + \frac{1}{(x+1)^4} + \frac{x^7}{(x+1)^8},$$

which is equivalent to

$$(x+1)^8 (G_{\mathcal{T}}(x) + G_{\mathcal{T}}(x)^4) + x^6 + x^5 + x^3 + x^2 + x = 0. \tag{3.6}$$

Hence the nonzero polynomial

$$h(x,y) = (x+1)^8 (y + y^4) + x^6 + x^5 + x^3 + x^2 + x \in \mathbb{F}_2[x,y]$$

satisfies $h(x, G_{\mathcal{T}}(x)) = 0$ and we obtain $E(\mathcal{T}) \leq \deg(h) = 12$.

Assume $G_{\mathcal{T}}(x)$ is a rational function, that is

$$G_{\mathcal{T}}(x) = \frac{f(x)}{g(x)}, \quad f, g \in \mathbb{F}_2[x],\ g \neq 0,$$

with $\gcd(f,g) = 1$. Then from (3.6) we get

$$(x+1)^8 (f g^3 + f^4) + (x^6 + x^5 + x^3 + x^2 + x) g^4 = 0.$$

Hence $(x+1)^8 \mid g^4$, that is $(x+1)^2 \mid g$. Also $g^3 \mid (x+1)^8$ since $\gcd(f,g) = 1$. This is only possible if $g(x) = x^2 + 1$. Now $(x^2+1) G_{\mathcal{T}}(x) = f(x)$ implies $t_{n+2} = t_n$ for $n \geq \deg(f)$. However, if $n \equiv 7 \pmod 8$ and thus $n + 2 \equiv 1 \pmod 8$, we have $1 = t_{n+2} \neq t_n = 0$. Consequently, $G_{\mathcal{T}}(x)$ is not rational. Moreover, the four zeros of $h(x,y)$ are obviously $y = G_{\mathcal{T}}(x) + \gamma$ with $\gamma \in \mathbb{F}_4$ and none of them is rational.

Let

$$\sum_{\ell=0}^{L_N(\mathcal{T})} \alpha_\ell t_{n+\ell} = 0, \quad 0 \leq n \leq N - L_N(\mathcal{T}) - 1, \quad \alpha_{L_N(\mathcal{T})} = 1,$$

be a shortest linear recurrence relation satisfied by the first N elements of $\mathcal{T}$. Choosing

$$g(x) = \sum_{\ell=0}^{L_N(\mathcal{T})} \alpha_\ell x^{L_N(\mathcal{T}) - \ell} \in \mathbb{F}_2[x]$$

we get

$$g(x)G_{\mathcal{T}}(x) \equiv f(x) \pmod{x^N}$$

for some polynomial $f(x) \in \mathbb{F}_2[x]$ of degree at most $L_N(\mathcal{T}) - 1$. Then

$$(x+1)^8(fg^3 + f^4) + (x^6 + x^5 + x^3 + x^2 + x)g^4 = K(x)x^N$$

with $K(x) \neq 0$ since $h(x, y)$ has no rational zero. Comparing the degrees of both sides we get

$$4L_N(\mathcal{T}) + 7 \geq N.$$

Thus $L_N(\mathcal{T}) \geq (N-7)/4$. □

Note that lower bounds on the Nth linear complexity of many other automatic sequences including the *Thue-Morse sequence*, the *Rudin-Shapiro sequence*, and the *regular paper-folding sequence* were obtained in [32]. Roughly speaking, for the class of (not ultimately periodic) automatic sequences the linear complexity is a much weaker measure for the unpredictability of a sequence than the expansion complexity.

3.2.2 Expansion complexity of p-periodic sequences over $\mathbb{F}_p$

For $0 \leq k \leq p-1$ we study the p-periodic sequence $\mathcal{A}_k = (a_{m,k})_{m \geq 0}$ of binomial coefficients over $\mathbb{F}_p$ defined by

$$a_{m,k} = \binom{m+k}{k} \pmod{p}.$$

Proposition 3.14 (Mérai, Niederreiter, Winterhof, [31])**.** *We have*

$$G_{\mathcal{A}_k}(x) = \frac{1}{(1-x)^{k+1}},$$

and

$$L(\mathcal{A}_k) = k + 1.$$

Proof. (see [31]). First verify that

$$\binom{p-1-k}{m}(-1)^m \equiv \prod_{j=1}^{m} \frac{k+j}{j} \equiv \binom{m+k}{m} \equiv \binom{m+k}{k} \pmod{p}.$$

Then we get

$$(1-x)^p G_{\mathcal{A}_k}(x) = (1-x^p) G_{\mathcal{A}_k}(x) = \sum_{m=0}^{p-1} a_{m,k} x^m = \sum_{m=0}^{p-1-k} \binom{m+k}{k} x^m$$
$$= \sum_{m=0}^{p-1-k} \binom{p-1-k}{m} (-x)^m = (1-x)^{p-1-k},$$

where we used that

$$\binom{m+k}{k} \equiv 0 \pmod{p}, \quad m = p-k, \ldots, p-1.$$

Thus

$$G_{\mathcal{A}_k}(x) = \frac{1}{(1-x)^{k+1}}.$$

By Proposition 3.8 we obtain

$$L(\mathcal{A}_k) = p - \deg\left(\gcd\left(\sum_{m=0}^{p-1} a_{m,k} x^m, 1-x^p\right)\right)$$
$$= p - \deg(\gcd((1-x)^{p-1-k}, (1-x)^p)) = k+1. \qquad \square$$

The following theorem shows that the p-th expansion complexity of sequences of binomial coefficients can be very small if the linear complexity is large with respect to p.

Theorem 3.15 (Mérai, Niederreiter, Winterhof, [31])**.** *For $(k+1)(k+2) < p$ we have*

$$E_p(\mathcal{A}_k) = k+2,$$

and for $(k+1)(k+2) \geq p$ we have

$$\left\lceil \frac{p}{k+2} \right\rceil \leq E_p(\mathcal{A}_k) \leq \max\left\{ \left\lceil \frac{p}{k+2} \right\rceil, (k+1)\left\{ \frac{p}{k+1} \right\} \right\},$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x.

Proof. (see [31]). By Proposition 3.11 and Proposition 3.14 we get

$$E_p(\mathcal{A}_k) = k+2$$

if $(k+1)(k+2) < p$ and

$$\left\lceil \frac{p}{k+2} \right\rceil \leq E_p(\mathcal{A}_k)$$

if $(k+1)(k+2) \geq p$.

Put

$$d = \min\left\{\left\lfloor \frac{p}{k+1} \right\rfloor, \left\lceil \frac{p}{k+2} \right\rceil\right\}$$

and take

$$h(x,y) = y^d - (1-x)^{p-d(k+1)} \in \mathbb{F}_p[x,y].$$

By Proposition 3.14 we have

$$G_{\mathcal{A}_k}(x) = \frac{1}{(1-x)^{k+1}}$$

and thus

$$\begin{aligned} h(x, G_{\mathcal{A}_k}(x)) &= \frac{1}{(1-x)^{d(k+1)}} - (1-x)^{p-d(k+1)} = \frac{1-(1-x)^p}{(1-x)^{d(k+1)}} \\ &= \frac{x^p}{(1-x)^{d(k+1)}} \equiv 0 \pmod{x^p} \end{aligned}$$

since $\gcd(x,(1-x)) = 1$. Hence

$$E_p(\mathcal{A}_k) \leq \deg(h) = \max\{d, p-d(k+1)\} = \begin{cases} d & \text{if } d = \left\lceil \frac{p}{k+2} \right\rceil, \\ p-d(k+1) & \text{otherwise,} \end{cases}$$

and the result follows. □

Next we study p-periodic sequences $\mathcal{A}_{u,v} = (A_m)_{m\geq 0}$ over $\mathbb{F}_p$ of the form

$$A_m = \sum_{k=u}^{v} \lambda_k a_{m,k} \pmod{p} \tag{3.7}$$

with $\lambda_u \lambda_v \neq 0$, $\lambda_k \in \mathbb{F}_p$ and $0 \leq u < v \leq p-1$.

Lemma 3.16. *We have*

$$G_{\mathcal{A}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{(1-x)^{k+1}},$$

and

$$L(\mathcal{A}_{u,v}) = v+1.$$

Proof. By Proposition 3.14 we get

$$G_{\mathcal{A}_{u,v}}(x) = \sum_{m=0}^{\infty} A_m x^m = \sum_{k=u}^{v} \lambda_k \sum_{m=0}^{\infty} a_{m,k} x^m = \sum_{k=u}^{v} \frac{\lambda_k}{(1-x)^{k+1}}.$$

Since

$$\gcd\left(1-x^p, \sum_{m=0}^{p-1} A_m x^m\right) = \gcd\left((1-x)^p, (1-x)^p \sum_{k=u}^{v} \frac{\lambda_k}{(1-x)^{k+1}}\right)$$
$$= \gcd\left((1-x)^p, (1-x)^p \frac{\sum_{k=u}^{v} \lambda_k (1-x)^{v-k}}{(1-x)^{v+1}}\right) = (1-x)^{p-1-v}$$

it follows from Proposition 3.8 that

$$L(\mathcal{A}_{u,v}) = p - \deg((1-x)^{p-1-v}) = v+1. \qquad \square$$

Remark 3.17. Note that any p-periodic sequence can be written in the form (3.7). More precisely, any p-periodic sequence $\mathcal{S} = (s_m)_{m\ge 0}$ over $\mathbb{F}_p$ can be defined by

$$s_m = f(m), \quad m \ge 0,$$

with a unique polynomial $f(x)$ over $\mathbb{F}_p$ of degree at most $p-1$. Now the polynomials

$$f_k(x) = (k!)^{-1}(x+k)(x+k-1)\cdots(x+1) = \binom{x+k}{k}, \quad k = 0, \ldots, p-1,$$

of degree k are a basis of the linear space of polynomials over $\mathbb{F}_p$ of degree at most $p-1$. Hence, the sequences

$$\mathcal{A}_k = (a_{m,k})_{m\ge 0}, \quad k = 0, \ldots, p-1,$$

are a basis of the linear space of p-periodic sequences over $\mathbb{F}_p$ and any p-periodic sequence is a linear combination of these basis sequences.

The p-th expansion complexity of $\mathcal{A}_{u,v}$ has the following lower and upper bound.

Theorem 3.18. *The p-th expansion complexity of $\mathcal{A}_{u,v} = (A_m)_{m\ge 0}$, $u < v$, of the form* (3.7) *can be bounded by*

$$\min\left\{\left\lceil \frac{p}{v+2}\right\rceil, v+2\right\} \le E_p(\mathcal{A}_{u,v})$$
$$\le \min\left\{(u+1)\left\{\frac{p}{v+1}\right\} + (v-u)\frac{p}{v+1}, v+2\right\},$$

where $\{x\} = x - \lfloor x\rfloor$ is the fractional part of x.

Proof. The bound

$$\min\left\{\left\lceil\frac{p}{v+2}\right\rceil, v+2\right\} \le E_p(\mathcal{A}_{u,v}) \le v+2$$

follows from Proposition 3.11 and Lemma 3.16.

Recall that

$$G_{\mathcal{A}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{(1-x)^{k+1}} = \frac{1}{(1-x)^{v+1}} \sum_{k=u}^{v} \lambda_k (1-x)^{v-k}.$$

Put

$$d = \left\lfloor\frac{p}{v+1}\right\rfloor$$

and take

$$h(x,y) = y^d - \left(\sum_{k=u}^{v} \lambda_k (1-x)^{v-k}\right)^d (1-x)^{p-d(v+1)} \in \mathbb{F}_p[x,y].$$

Then

$$\begin{aligned} h(x, G_{\mathcal{A}_{u,v}}(x)) &= \frac{\left(\sum\limits_{k=u}^{v} \lambda_k (1-x)^{v-k}\right)^d - \left(\sum\limits_{k=u}^{v} \lambda_k (1-x)^{v-k}\right)^d (1-x)^p}{(1-x)^{d(v+1)}} \\ &= \frac{\left(\sum\limits_{k=u}^{v} \lambda_k (1-x)^{v-k}\right)^d x^p}{(1-x)^{d(v+1)}} \equiv 0 \pmod{x^p} \end{aligned}$$

since $\gcd(x,(1-x)) = 1$. Hence

$$\begin{aligned} E_p(\mathcal{A}_{u,v}) &\le \deg(h) = \max\{d, d(v-u) + p - d(v+1)\} \\ &= \max\{d, p - d(u+1)\} = p - d(u+1) \end{aligned}$$

and the result follows. □

Remark 3.19. In Theorem 3.18 we proved that

$$E_p(\mathcal{A}_{u,v}) \le \min\left\{(u+1)\left\{\frac{p}{v+1}\right\} + (v-u)\frac{p}{v+1}, v+2\right\} =: Z(\mathcal{A}_{u,v}). \quad (3.8)$$

On the one hand the bound can be very small if v is large with respect to p and $v-u$ is small. For the case $u = v$ see Theorem 3.15. On the other hand we have $L(\mathcal{A}_{u,v}) = v+1$ by Lemma 3.16. Hence, there are many p-periodic sequences over $\mathbb{F}_p$ of large linear complexity but small p-th expansion complexity and we have the following hierarchy of complexity measures for p-periodic sequences

$$E_p(\mathcal{A}_{u,v}) \le Z(\mathcal{A}_{u,v}) \le L(A_{u,v}) + 1 = v+2.$$

As an application we provide some examples of p-periodic sequences over $\mathbb{F}_p$ for which the bound (3.8) is not small.

Example 3.20. For a positive integer d consider the p-periodic sequence $\mathcal{S} = (s_m)_{m\geq 0}$ over $\mathbb{F}_p$ defined by

$$s_m \equiv m^d \pmod{p},$$

that is $s_m = 0$ for $m \equiv 0 \pmod{p}$ and $s_m \neq 0$ otherwise.

Assume $\mathcal{S} = \mathcal{A}_{u,v}$. Then we need $v = d$ and $u = 0$ since

$$A_m = \sum_{k=u}^{v} \lambda_k \binom{m+k}{k} \equiv 0 \pmod{p}, \quad m = p-u, \ldots, p-1.$$

Thus, v is large with respect to p if and only if $v - u$ is large and consequently the bound (3.8) cannot be small.

Example 3.21. Consider the Legendre sequence $\mathcal{L} = (\ell_m)_{m\geq 0}$ over $\mathbb{F}_p$ of period p defined by (see also Chapter 2)

$$\ell_m = \begin{cases} 1 & \text{if } m \text{ is a quadratic residue modulo } p, \\ 0 & \text{otherwise,} \end{cases}$$

or equivalently

$$\ell_m = \frac{m^{p-1} + m^{\frac{p-1}{2}}}{2} \pmod{p}.$$

We prove

$$Z(\mathcal{L}) = p + O\left(p^{\frac{1}{4\sqrt{e}}+\varepsilon}\right) \quad \text{for any } \varepsilon > 0.$$

Assume $\mathcal{L} = \mathcal{A}_{u,v}$. Then we need $v = p-1$. If $p \equiv 1 \pmod 4$, then $p-1$ is a quadratic residue modulo p, thus $\ell_{p-1} = 1$, and hence we must have $u = 0$ since otherwise $\ell_{p-1} = A_{p-1} = 0$. If $p \equiv 3 \pmod 4$, then $p-1$ is not a quadratic residue modulo p and also $p-u, \ldots, p-1$ must be quadratic nonresidues modulo p since $\ell_m = A_m = 0$ for $m = p-u, \ldots, p-1$. This simply means that $1, \ldots, u$ are quadratic residues modulo p since the product of two quadratic nonresidues is a quadratic residue. Thus

$$u = O\left(p^{\frac{1}{4\sqrt{e}}+\varepsilon}\right) \quad \text{for any } \varepsilon > 0$$

by the Burgess bound [6, Theorem 2].

Hence, the Legendre sequence does not belong to the class of sequences for which the bound (3.8) is small.

3.2.3 Expansion complexity of t-periodic sequences over $\mathbb{F}_q$ with $t \mid q-1$

Let $g \in \mathbb{F}_q^*$ be an element of order t, where $t \mid q-1$. For $0 \leq k \leq t-1$ we study the t-periodic sequence $\mathcal{B}_k = (b_{i,k})_{i \geq 0}$ over $\mathbb{F}_q$ defined by

$$b_{i,k} = (-1)^{t-1} g^{-(i+1)k}.$$

Proposition 3.22. *We have*

$$G_{\mathcal{B}_k}(x) = \frac{1}{g^k - x}.$$

Proof. First verify that

$$\prod_{\substack{j=0 \\ j \neq k}}^{t-1} g^j = g^{\frac{(t-1)t}{2} - k} = (-1)^{t-1} g^{-k} = b_{0,k}.$$

Put

$$f_k(x) = \sum_{i=0}^{t-1} b_{i,k} x^i \in \mathbb{F}_q[x]$$

and note that $b_{i,k} = b_{0,k} g^{-ik}$. Then

$$f_k(g^j) = b_{0,k} \sum_{i=0}^{t-1} g^{i(j-k)} = \begin{cases} 0 & \text{if } j \not\equiv k \pmod{t}, \\ b_{0,k} t & \text{if } j \equiv k \pmod{t}. \end{cases}$$

Hence $f_k(g^j) = 0$ for all $0 \leq j \leq t-1$ with $j \neq k$ and since $f_k(x)$ has degree $t-1$ and

$$f_k(0) = b_{0,k} = \prod_{\substack{j=0 \\ j \neq k}}^{t-1} g^j,$$

we can write $f_k(x)$ as

$$f_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{t-1} (g^j - x).$$

This implies

$$(1 - x^t) G_{\mathcal{B}_k}(x) = \sum_{i=0}^{t-1} b_{i,k} x^i = f_k(x) = \prod_{\substack{j=0 \\ j \neq k}}^{t-1} (g^j - x) = \frac{1 - x^t}{g^k - x}$$

and thus

$$G_{\mathcal{B}_k}(x) = \frac{1}{g^k - x}. \qquad \square$$

The (Nth) linear complexity and t-th expansion complexity of $\mathcal{B}_k$ can be computed exactly as the following theorem shows.

Theorem 3.23. *We have*

$$L(\mathcal{B}_k) = L_N(\mathcal{B}_k) = 1,$$

and

$$E_t(\mathcal{B}_k) = \begin{cases} 1 & \text{if } t \le 2, \\ 2 & \text{otherwise.} \end{cases}$$

Proof. Since $b_{i+1,k} = g^{-k} b_{i,k}$ the linear recurrence relation $b_{i+1,k} + \alpha_0 b_{i,k} = 0$ of length 1 is fulfilled with $\alpha_0 = -g^{-k} \in \mathbb{F}_q$. This implies $L(\mathcal{B}_k) = L_N(\mathcal{B}_k) = 1$.

By Proposition 3.11 we get

$$\min\left\{\left\lceil \frac{t}{2} \right\rceil, 2\right\} \le E_t(\mathcal{B}_k) \le 2$$

and the result follows for $t \ge 3$. For $t = 1, 2$ take

$$h(x, y) = y - \sum_{i=0}^{t-1} b_{i,k} x^i \in \mathbb{F}_q[x, y]$$

of degree 1 which satisfies

$$h(x, G_{\mathcal{B}_k}(x)) \equiv 0 \pmod{x^t}.$$

Thus $E_t(\mathcal{B}_k) = 1$ for $t = 1, 2$. □

Next we study t-periodic sequences $\mathcal{B}_{u,v} = (B_i)_{i \ge 0}$ over $\mathbb{F}_q$ of the form

$$B_i = \sum_{k=u}^{v} \lambda_k b_{i,k} \tag{3.9}$$

with $\lambda_u \lambda_v \ne 0$, $\lambda_k \in \mathbb{F}_q$ and $0 \le u < v \le t - 1$.

Lemma 3.24. *We have*

$$G_{\mathcal{B}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{g^k - x},$$

and

$$L(\mathcal{B}_{u,v}) = |\{u \le k \le v : \lambda_k \ne 0\}|.$$

Proof. By Proposition 3.22 we get

$$G_{\mathcal{B}_{u,v}}(x) = \sum_{i=0}^{\infty} B_i x^i = \sum_{k=u}^{v} \lambda_k \sum_{i=0}^{\infty} b_{i,k} x^i = \sum_{k=u}^{v} \frac{\lambda_k}{g^k - x},$$

that is

$$G_{\mathcal{B}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{g^k - x} = \frac{\sum\limits_{k=u}^{v} \lambda_k \prod\limits_{\substack{j=u \\ j\neq k, \lambda_j \neq 0}}^{v} (g^j - x)}{\prod\limits_{\substack{k=u \\ \lambda_k \neq 0}}^{v} (g^k - x)}.$$

Then

$$\gcd\left(1 - x^t, \sum_{i=0}^{t-1} B_i x^i\right) = \gcd\left(1 - x^t, (1 - x^t) \sum_{k=u}^{v} \frac{\lambda_k}{g^k - x}\right) = \frac{1 - x^t}{\prod\limits_{\substack{k=u \\ \lambda_k \neq 0}}^{v} (g^k - x)}$$

and by Proposition 3.8 we obtain

$$L(\mathcal{B}_{u,v}) = t - \deg\left(\frac{1 - x^t}{\prod\limits_{\substack{k=u \\ \lambda_k \neq 0}}^{v} (g^k - x)}\right) = |\{u \leq k \leq v : \lambda_k \neq 0\}|. \qquad \square$$

Remark 3.25. Note that any t-periodic sequence can be written in the form (3.9). Assume

$$\sum_{k=0}^{t-1} \lambda_k b_{i,k} = 0, \quad i = 0, \ldots, t-1,$$

which is equivalent to

$$\sum_{k=0}^{t-1} \lambda_k g^{-(i+1)k} = 0, \quad i = 0, \ldots, t-1.$$

This leads to the system of equations

$$M \begin{pmatrix} \lambda_0 \\ \lambda_1 \\ \vdots \\ \lambda_{t-1} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{3.10}$$

where

$$M = \begin{pmatrix} 1 & g^{-1} & (g^{-1})^2 & \dots & (g^{-1})^{t-1} \\ 1 & g^{-2} & (g^{-2})^2 & \dots & (g^{-2})^{t-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & g^{-t} & (g^{-t})^2 & \dots & (g^{-t})^{t-1} \end{pmatrix}$$

is a Vandermonde matrix. Since the determinant

$$\prod_{1 \le i < k \le t} (g^{-k} - g^{-i})$$

of the Vandermonde matrix M is nonzero if $g^{-i} \neq g^{-k}$ for $i \neq k$ it follows that M is regular. Thus the system of equations (3.10) is solvable and $\lambda_k = 0$ for $k = 0, \dots, t-1$. Hence, the sequences

$$\mathcal{B}_k = (b_{i,k})_{i \ge 0}, \quad k = 0, \dots, t-1,$$

are a basis of the linear space of t-periodic sequences over $\mathbb{F}_q$ with $t \mid q-1$ and any t-periodic sequence is a linear combination of these basis sequences.

The t-th expansion complexity of $\mathcal{B}_{u,v}$ has the following lower and upper bound.

Theorem 3.26. *The t-th expansion complexity of $\mathcal{B}_{u,v} = (B_i)_{i \ge 0}$, $u < v$, of the form* (3.9) *can be bounded by*

$$\min\left\{ \left\lceil \frac{t}{L(\mathcal{B}_{u,v})+1} \right\rceil, L(\mathcal{B}_{u,v})+1 \right\} \le E_t(\mathcal{B}_{u,v}) \le \min\{t-1, L(\mathcal{B}_{u,v})+1\}.$$

Proof. The bound

$$\min\left\{ \left\lceil \frac{t}{L(\mathcal{B}_{u,v})+1} \right\rceil, L(\mathcal{B}_{u,v})+1 \right\} \le E_t(\mathcal{B}_{u,v}) \le L(\mathcal{B}_{u,v})+1$$

follows from Proposition 3.11.

Recall that

$$G_{\mathcal{B}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{g^k - x}.$$

Take

$$h(x,y) = y - \sum_{k=u}^{v} \lambda_k \sum_{i=0}^{t-1} b_{i,k} x^i \in \mathbb{F}_q[x,y]$$

of degree $\max\{1, t-1\} = t-1$ which satisfies

$$h(x, G_{\mathcal{B}_{u,v}}(x)) \equiv 0 \pmod{x^t}.$$

Hence

$$E_t(\mathcal{B}_{u,v}) \le \deg(h) = t-1. \qquad \square$$

Chapter 4

Outlook

In this chapter we provide some additional results which may be of interest for further research. In particular, we study tp^r-periodic sequences over $\mathbb{F}_q$, where $p \nmid t$ and p is the characteristic of $\mathbb{F}_q$.

4.1 Arithmetic correlation measure

In this section we introduce a new arithmetic correlation measure which may be seen in some sense as a generalization of the arithmetic autocorrelation to higher orders. Proper estimates of this new measure may be used to estimate the 2-*adic span* (the arithmetic analogue of the linear complexity). Unfortunately, bounds on the Legendre sequence, though nontrivial, are too weak for this purpose. However, as first result we prove a nontrivial bound on this arithmetic correlation measure of order k of the Legendre sequence.

4.1.1 Arithmetic correlation measure of order k

The *arithmetic correlation measure of order* $k \geq 1$ of a (purely) T-periodic binary sequence $(a_n)_{n\geq 0}$ is defined as

$$A_k(a_n) = \max_{0<d_1<\ldots<d_{k-1}<T} |I(s_{n,d_1,\ldots,d_{k-1}})|,$$

where $(s_{n,d_1,\ldots,d_{k-1}})_{n\geq 0}$ is the ultimately p-periodic binary sequence defined by

$$s_{n,d_1,\ldots,d_{k-1}} = \begin{cases} 1 & \text{if } (-1)^{a_n+a_{n+d_1}+\cdots+a_{n+d_{k-1}}+z_n} = 1, \\ 0 & \text{otherwise,} \end{cases} \tag{4.1}$$

with $z_0 = 0$ and for all $n \geq 1$

$$z_n = \left\lfloor \frac{a_{n-1} + a_{n-1+d_1} + \cdots + a_{n-1+d_{k-1}} + z_{n-1}}{2} \right\rfloor.$$

We have $A_1(a_n) = |I(a_n)|$ and $z_n \in \{0, 1, \ldots, k-1\}$ since, by induction,

$$z_n \leq \lfloor (2k-1)/2 \rfloor = \lfloor k - 1/2 \rfloor = k - 1.$$

Remark 4.1. In contrast to the relation between the (periodic) autocorrelation and the correlation measure of order 2, the arithmetic correlation measure of order 2 is not the maximum over all $0 < d_1 < T$ of the absolute value of the arithmetic autocorrelation of $(a_n)_{n\geq 0}$ since addition and subtraction of 2-adic integers are not the same operation (see Remark 2.14). Thus we cannot say, that $A_k(a_n)$ is a direct generalization of the arithmetic autocorrelation to higher orders.

Let i be a positive integer. In the following put $k = 2^i + 1$ and choose

$$K = \lceil \log_2 k \rceil = i + 1.$$

Let $\mathcal{B}^{k\times K}$ denote the set of binary $k \times K$ matrices, that is

$$\mathcal{B}^{k\times K} = \left\{ \begin{pmatrix} b_{0,0} & b_{0,1} & \cdots & b_{0,K-1} \\ b_{1,0} & b_{1,1} & \cdots & b_{1,K-1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k-1,0} & b_{k-1,1} & \cdots & b_{k-1,K-1} \end{pmatrix} : b_{i,j} \in \{0,1\} \right\}.$$

Obviously, we have $|\mathcal{B}^{k\times K}| = 2^{kK}$.

Algorithm 4.2. *Let $B \in \mathcal{B}^{k\times K}$ and $c \in \{0, k-1\}$.*

1. *Input:* $z_{c,0} = c$
2. *For* $j = 1, 2, \ldots, K$ *do*

$$w_{j-1} = z_{k-1,j-1} - z_{0,j-1}$$
$$z_{c,j} = \left\lfloor \frac{b_{0,j-1} + b_{1,j-1} + \cdots + b_{k-1,j-1} + z_{c,j-1}}{2} \right\rfloor \in \{0, 1, \ldots, k-1\}$$

3. *Output:* $z_{c,K} \in \{0, 1, \ldots, k-1\}$, $w_{K-1} \in \{0, 1\}$

If we identify the i-th row of a matrix $B \in \mathcal{B}^{k\times K}$ with the 2-adic integer

$$\sum_{j=0}^{\infty} b_{i,j}2^j,$$

where $b_{i,j} = 0$ for $j \geq K$, then Algorithm 4.2 is simply addition of k 2-adic integers with input carry $z_{c,0}$ and output carry $z_{c,K}$, where $c \in \{0, k-1\}$. The value w_{j-1} computes the difference between the carries $z_{k-1,j-1}$ and $z_{0,j-1}$ in each step of the algorithm. If $z_{0,K} = z_{k-1,K}$ for some $B \in \mathcal{B}^{k\times K}$, then we know that B has the same output carry for all possible inputs $0, 1, \ldots, k-1$.

Put

$$\mathcal{Y}^{k\times K} = \{B \in \mathcal{B}^{k\times K} : z_{0,K} = z_{k-1,K}\} \tag{4.2}$$

and

$$\mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K} = \{B \in \mathcal{B}^{k\times K} : z_{0,K} \neq z_{k-1,K}\} \tag{4.3}$$

Proposition 4.3. *We have*

$$|\mathcal{Y}^{k\times K}| = |\mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}| = 2^{kK-1}.$$

Proof. For $k = 2^i + 1$ Algorithm 4.2 gives

$$w_{K-1} = z_{k-1,K-1} - z_{0,K-1} = 1$$

for all $B \in \mathcal{B}^{k\times K}$ since $w_{j-1} = w_{j-2}/2$ for $j = 2, 3, \ldots, K$. Computing $z_{c,K}$, where $c \in \{0, k-1\}$, the result follows. □

4.1.2 A bound on the arithmetic correlation measure of order k of the Legendre sequence

Exchanging $(1, 0)$ by $(0, 0)$ and $(0, 1)$ by $(1, 1)$ in the proof of Theorem 2.32 we can show for the Legendre sequence $(\ell_n)_{n\geq 0}$ that

$$A_2(\ell_n) \leq 4p^{3/4}(\log_2 p)^{1/2}.$$

Theorem 4.4. *Put $k = 2^i + 1$. Then the arithmetic correlation measure of order k of the p-periodic binary sequence $(\ell_n)_{n\geq 0}$ defined by (2.1) satisfies*

$$A_k(\ell_n) \leq 2^{\frac{k-1}{kK_1+1}+1} p^{1-\frac{1}{2(kK_1+1)}} (\log_2 p)^{1-\frac{kK_1}{kK_1+1}},$$

where $K_1 = \log_2 k + 1$.

Proof. Put

$$m = \left\lfloor \frac{1/2 \log_2 p - \log_2 \log_2 p - k + 1}{kK_1 + 1} \right\rfloor$$

and note that $K \le K_1$. In the following we derive a lower bound on the number N_1 of ones in a period of the p-periodic sequence $(s_{n,d_1,\ldots,d_{k-1}})_{n\ge 0}$ defined by (4.1).

If

$$p \le 2^{\frac{k-1}{kK_1+1}+1} p^{1-\frac{1}{2(kK_1+1)}} (\log_2 p)^{1-\frac{kK_1}{kK_1+1}},$$

then the result follows immediately since the trivial bound $A_k(\ell_n) \le p$ always holds. Thus it is enough to prove the inequality for

$$p^{\frac{1}{2(kK_1+1)}} > 2^{\frac{k-1}{kK_1+1}+1} (\log_2 p)^{\frac{1}{kK_1+1}},$$

that is $p^{1/2} > 2^{k(K_1+1)} \log_2 p$.

Note that $1 \le m \le (2kK)^{-1} \log_2 p$. For some r and n with $0 \le r < m$ and $p \le n < 2p$ assume

$$Y = \begin{pmatrix} \ell_{n-Kr-K} & \ell_{n-Kr-K+1} & \cdots & \ell_{n-Kr-1} \\ \ell_{n-Kr-K+d_1} & \ell_{n-Kr-K+1+d_1} & \cdots & \ell_{n-Kr-1+d_1} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n-Kr-K+d_{k-1}} & \ell_{n-Kr-K+1+d_{k-1}} & \cdots & \ell_{n-Kr-1+d_{k-1}} \end{pmatrix} \in \mathcal{Y}^{k\times K}, \quad (4.4)$$

$$X_j \in \mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}, \quad j = 0, \ldots, r-1, \quad (4.5)$$

where

$$X_j = \begin{pmatrix} \ell_{n-Kr+Kj} & \ell_{n-Kr+Kj+1} & \cdots & \ell_{n-Kr+Kj+K-1} \\ \ell_{n-Kr+Kj+d_1} & \ell_{n-Kr+Kj+1+d_1} & \cdots & \ell_{n-Kr+Kj+K-1+d_1} \\ \vdots & \vdots & \ddots & \vdots \\ \ell_{n-Kr+Kj+d_{k-1}} & \ell_{n-Kr+Kj+1+d_{k-1}} & \cdots & \ell_{n-Kr+Kj+K-1+d_{k-1}} \end{pmatrix}$$

and

$$(\ell_n, \ell_{n+d_1}, \ldots, \ell_{n+d_{k-1}}) \in \mathcal{B}^{k\times 1}. \quad (4.6)$$

We consider only patterns of length

$$k(K+1) \le s = k(K + Kr + 1) \le 1/2 \log_2 p + k$$

and therefore we can further estimate (2.4) by $sp^{1/2}/2$, that is

$$\begin{aligned}\left|\mathcal{P}_{i_0,i_1,\ldots,i_{s-1}}(\ell_n)-\frac{p}{2^s}\right| &\le \frac{p^{1/2}(2^{s-1}(s-3)+2)+2^{s-1}(s+1)-1}{2^s}\\ &\le \left(\frac{s-3}{2}+2^{1-k(K+1)}\right)p^{1/2}+\frac{2k+3}{4}\log_2 p \qquad (4.7)\\ &\le \left(\frac{s-3+(2k+11)2^{-(kK+k+1)}}{2}\right)p^{1/2}\le\frac{s}{2}p^{1/2}\end{aligned}$$

since $p^{1/2}>2^{k(K_1+1)}\log_2 p\ge 2^{k(K+1)}\log_2 p$.

From (4.7) we know that for fixed $Y\in\mathcal{Y}^{k\times K}$ the number of patterns

$$\begin{pmatrix} & & & & \ell_n\\ Y & X_0 & \ldots & X_{r-1} & \vdots\\ & & & & \ell_{n+d_{k-1}}\end{pmatrix} \qquad (4.8)$$

satisfying the assumptions (4.4)-(4.6) in

$$\begin{array}{llllll}\ell_{p-r-1} & \ldots & \ell_{p-1} & \ell_p & \ldots & \ell_{2p-1}\\ \ell_{d_1+p-r-1} & \ldots & \ell_{d_1+p-1} & \ell_{d_1+p} & \ldots & \ell_{d_1+2p-1}\\ \vdots & & \vdots & \vdots & & \vdots\\ \ell_{d_{k-1}+p-r-1} & \ldots & \ell_{d_{k-1}+p-1} & \ell_{d_{k-1}+p} & \ldots & \ell_{d_{k-1}+2p-1}\end{array} \qquad (4.9)$$

is at least $p/2^{k(K+Kr+1)}-k/2(K+Kr+1)p^{1/2}$. Note that if the integers $d_1<d_2<\ldots<d_{k-1}$ do not fulfill

$$\begin{aligned}&d_1\ge Km,\\ &d_{j+1}-d_j\ge Km,\quad j=1,2,\ldots,k-2,\\ &d_{k-1}\le p-Km,\end{aligned}$$

then some indices in (4.8) coincide and we would deal with shorter patterns which would lead to sharper bounds than $p/2^{k(K+Kr+1)}-d/2(K+Kr+1)p^{1/2}$ on the number of patterns (4.8) satisfying the assumptions (4.4)-(4.6) in (4.9). But all these bounds are at least $p/2^{k(K+Kr+1)}-k/2(K+Kr+1)p^{1/2}$ and therefore we can continue the proof without taking care of shorter patterns.

By (4.1) we have

$$s_{n,d_1,\ldots,d_{k-1}}=\begin{cases}1 & \text{if } \sum_{i=1}^{k-1}\ell_{n+d_i}+z_n \text{ is even},\\ 0 & \text{if } \sum_{i=1}^{k-1}\ell_{n+d_i}+z_n \text{ is odd}.\end{cases}$$

Since there are $2^{(kK-1)r+k-1}$ possible choices for the pattern (4.8) we count at least $p/2^{kK+r+1}-k(K+Kr+1)2^{(kK-1)r+k-2}p^{1/2}$ different $p\le n<2p$ with $s_{n,d_1,\ldots,d_{k-1}}=1$.

By Proposition 4.3 we have $|\mathcal{Y}^{k\times K}| = 2^{kK-1}$. Thus in total there are at least

$$p/2^{r+2} - k(K + Kr + 1)2^{(kK-1)r+kK+k-3}p^{1/2}$$

different $p \le n < 2p$ with $Y \in \mathcal{Y}^{k\times K}$, $X_j \in \mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}$ for $j = 0, \dots, r-1$ and $s_{n,d_1,\dots,d_{k-1}} = 1$.

Summing up all the contributions we get the formula

$$N_1 \ge \frac{1}{4}\left(\sum_{r=0}^{m-1} 2^{-r}\right)p - \frac{1}{4}\left(\sum_{r=0}^{m-1} 2^{r(kK-1)+kK+k-1}(rkK + kK + k)\right)p^{1/2}.$$

The second sum on the right hand side of the inequality can be estimated by

$$\begin{aligned}
&\sum_{r=0}^{m-1} 2^{r(kK-1)+kK+k-1}(rkK + kK + k) \le \sum_{r=0}^{m-1} 2^{rkK+kK+k-1}(rkK + kK + k)\\
&\le \sum_{r=0}^{m-1} 2^{rkK_1+kK_1+k-1}(rkK_1 + kK_1 + k) = \left(\sum_{r=0}^{m-1} 2^{rkK_1+kK_1+k}\right)'\\
&= \frac{((mdK_1 + k)(2^{kK_1} - 1) - kK_1)2^{mkK_1+kK_1+k-1} + k(K_1 + 1 - 2^{kK_1})2^{kK_1+k-1}}{(2^{kK_1} - 1)^2}\\
&\le \frac{((mkK_1 + k)(2^{kK_1} - 1) - kK_1)2^{mkK_1+k-1} + k(K_1 + 1 - 2^{kK_1})2^{k-1}}{2^{kK_1-1}}\\
&\le \frac{(mkK_1 + k)2^{mkK_1+kK_1+k-1}}{2^{kK_1-1}} = (mkK_1 + k)2^{mkK_1+k} \le 2^{mkK_1+k}\log_2 p,
\end{aligned}$$

where we used $(2^{kK_1} - 1)^2 \ge 2^{2kK_1-1}$ and

$$m \le \frac{1/2\log_2 p - \log_2\log_2 p - k + 1}{dK_1} \le \frac{\log_2 p}{kK_1} - \frac{k}{kK_1}.$$

Thus by the definition of m we get

$$\begin{aligned}
N_1 &\ge \frac{p}{2} - 2^{-m-1}p - 2^{mkK_1+k-2}p^{1/2}\log_2 p\\
&\ge \frac{p}{2} - 2^{\frac{k-1}{kK_1+1}}p^{1-\frac{1}{2(kK_1+1)}}(\log_2 p)^{1-\frac{kK_1}{kK_1+1}}.
\end{aligned}$$

Analogously N_0 can be bounded below by

$$N_0 \ge \frac{p}{2} - 2^{\frac{k-1}{kK_1+1}}p^{1-\frac{1}{2(kK_1+1)}}(\log_2 p)^{1-\frac{kK_1}{kK_1+1}}$$

and therefore since $N_0 + N_1 = p$

$$A_k(\ell_n) = |N_0 - N_1| \le 2^{\frac{k-1}{kK_1+1}+1}p^{1-\frac{1}{2(kK_1+1)}}(\log_2 p)^{1-\frac{kK_1}{kK_1+1}}. \qquad \square$$

Remark 4.5. In the proof of Theorem 4.4 we need that the sets $\mathcal{Y}^{k\times K}$ and $\mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}$ defined by (4.2) and (4.3), respectively, have the same cardinality. Unfortunately, this is only fulfilled if k is the next larger integer of a power of 2 (see Proposition 4.3). However, for $k \neq 2^i + 1$ we choose $K = \lceil \log_2 k \rceil + 1$. Then it can be proven that

$$|\mathcal{Y}^{k\times K}| = 2^{kK-1} + 2^{k-1}|\mathcal{Y}^{k\times K-1}|.$$

Thus, modifying the sets $\mathcal{Y}^{k\times K}$ and $\mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}$ by

$$\mathcal{Y}^{k\times K} \setminus S =: \mathcal{Y}_*^{k\times K}$$

and

$$(\mathcal{B}^{k\times K} \setminus \mathcal{Y}^{k\times K}) \cup S =: \mathcal{B}^{k\times K} \setminus \mathcal{Y}_*^{k\times K},$$

respectively, where

$$S = \{(Y, b) : Y \in \mathcal{Y}^{k\times K-1}, b \in \mathcal{B}^{k\times 1}\} \in \mathcal{B}^{k\times K},$$

and choosing

$$K_1 = \log_2 k + 2,$$

we can prove Theorem 4.4 for $k \neq 2^i + 1$ as well (since the new sets $\mathcal{Y}_*^{k\times K}$ and $\mathcal{B}^{k\times K} \setminus \mathcal{Y}_*^{k\times K}$ satisfy $|\mathcal{Y}_*^{k\times K}| = |\mathcal{B}^{k\times K} \setminus \mathcal{Y}_*^{k\times K}| = 2^{kK-1}$ for $k \neq 2^i + 1$).

4.2 Expansion complexity of tp^r-periodic sequences over $\mathbb{F}_q$

Let $n = n_0 + n_1 t$ and $k = k_0 + k_1 t$ with $0 \leq n_0 < t$, $0 \leq k_0 < t$ and $0 \leq n_1 < p^r$, $0 \leq k_1 < p^r$, respectively. For $0 \leq k \leq tp^r - 1$ we study the tp^r-periodic sequence $\mathcal{C}_k = (c_{n,k})_{n\geq 0}$ over $\mathbb{F}_q$ defined by

$$c_{n,k} = (-1)^{t-1} \binom{n_1 + k_1}{k_1} g^{-(n_0+1)k_0}.$$

Proposition 4.6. *We have*

$$G_{\mathcal{C}_k}(x) = \frac{1}{(1 - x^t)^{k_1}(g^{k_0} - x)},$$

and

$$L(\mathcal{C}_k) = tk_1 + 1.$$

Proof. It follows from Proposition 3.22 that

$$\sum_{n_0=0}^{t-1} b_{n_0,k_0} x^{n_0} = \sum_{n_0=0}^{t-1} (-1)^{t-1} g^{-(n_0+1)k_0} x^{n_0} = \frac{1-x^t}{g^{k_0}-x}.$$

Furthermore we have

$$\begin{aligned}\sum_{n_1=0}^{p^r-1} \binom{n_1+k_1}{k_1} x^{tn_1} &= \sum_{n_1=0}^{p^r-1-k_1} \binom{n_1+k_1}{k_1} x^{tn_1} \\ &= \sum_{n_1=0}^{p^r-1-k_1} \binom{p^r-1-k_1}{n_1} (-x^t)^{n_1} = (1-x^t)^{p^r-1-k_1},\end{aligned}$$

where we used that

$$\binom{n_1+k_1}{k_1} \equiv \binom{n_1+k_1-p^r}{k_1}\binom{1}{0} \equiv 0 \pmod{p}, \quad n_1 = p^r-k_1,\ldots,p^r-1,$$

by Lucas Congruence (see for example [35]) since $0 \le n_1+k_1-p^r < k_1 < p^r$. Then we get

$$\begin{aligned}(1-x^t)^{p^r} G_{\mathcal{C}_k}(x) &= (1-x^{tp^r}) G_{\mathcal{C}_k}(x) = \sum_{n=0}^{tp^r-1} c_{n,k} x^n \\ &= \sum_{n_0=0}^{t-1} \sum_{n_1=0}^{p^r-1} \left((-1)^{t-1} \binom{n_1+k_1}{k_1} g^{-(n_0+1)k_0} \right) x^{n_0+n_1 t} \\ &= \sum_{n_1=0}^{p^r-1} \binom{n_1+k_1}{k_1} x^{n_1 t} \sum_{n_0=0}^{t-1} (-1)^{t-1} g^{-(n_0+1)k_0} x^{n_0} = \frac{(1-x^t)^{p^r-k_1}}{g^{k_0}-x}.\end{aligned}$$

Thus

$$G_{\mathcal{C}_k}(x) = \frac{1}{(1-x^t)^{k_1}(g^{k_0}-x)}.$$

By Proposition 3.8 we obtain

$$\begin{aligned}L(\mathcal{C}_k) &= tp^r - \deg\left(\gcd\left(\sum_{n=0}^{tp^r-1} c_{n,k} x^n, 1-x^{tp^r}\right)\right) \\ &= tp^r - \deg\left(\gcd\left(\frac{(1-x^t)^{p^r-k_1}}{g^{k_0}-x}, (1-x^t)^{p^r}\right)\right) = tk_1+1. \qquad \square\end{aligned}$$

The following theorem shows that the tp^r-th expansion complexity of $\mathcal{C}_k$ can be very small if the linear complexity is large with respect to tp^r.

Theorem 4.7. *For $(tk_1+1)(tk_1+2) < tp^r$ we have*

$$E_{tp^r}(\mathcal{C}_k) = tk_1 + 2,$$

and for $(tk_1+1)(tk_1+2) \geq tp^r$ we have

$$\left\lceil \frac{tp^r}{tk_1+2} \right\rceil \leq E_{tp^r}(\mathcal{C}_k) \leq \max\left\{ 2\left\lceil \frac{tp^r}{tk_1+2} \right\rceil, tk_1 \left\{ \frac{p^r}{k_1} \right\} \right\},$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x.

Proof. By Proposition 3.11 and Proposition 4.6 we get

$$E_{tp^r}(\mathcal{C}_k) = tk_1 + 2$$

if $(tk_1+1)(tk_1+2) < tp^r$ and

$$\left\lceil \frac{tp^r}{tk_1+2} \right\rceil \leq E_{tp^r}(\mathcal{C}_k)$$

if $(tk_1+1)(tk_1+2) \geq tp^r$.

Put

$$d = \min\left\{ \left\lfloor \frac{p^r}{k_1} \right\rfloor, \left\lceil \frac{tp^r}{tk_1+2} \right\rceil \right\}$$

and take

$$h(x,y) = y^d(g^{k_0} - x)^d - (1-x^t)^{p^r - dk_1} \in \mathbb{F}_q[x,y].$$

By Proposition 4.6 we have

$$G_{\mathcal{C}_k}(x) = \frac{1}{(1-x^t)^{k_1}(g^{k_0}-x)}$$

and thus

$$\begin{aligned} h(x, G_{\mathcal{C}_k}(x)) &= \frac{(g^{k_0}-x)^d}{(1-x^t)^{dk_1}(g^{k_0}-x)^d} - (1-x^t)^{p^r-dk_1} \\ &= \frac{1-(1-x^t)^{p^r}}{(1-x^t)^{dk_1}} = \frac{x^{tp^r}}{(1-x^t)^{dk_1}} \equiv 0 \pmod{x^{tp^r}} \end{aligned}$$

since $\gcd(x,(1-x)) = 1$. Hence

$$E_{tp^r}(\mathcal{C}_k) \leq \deg(h) = \max\{2d, tp^r - tdk_1\} = \begin{cases} 2d & \text{if } d = \left\lceil \frac{tp^r}{tk_1+2} \right\rceil, \\ t(p^r - dk_1) & \text{otherwise,} \end{cases}$$

and the result follows. □

Next we study tp^r-periodic sequences $\mathcal{C}_{u,v} = (C_n)_{n\geq 0}$ over $\mathbb{F}_q$ of the form

$$C_n = \sum_{k=u}^{v} \lambda_k c_{n,k} \tag{4.10}$$

with $\lambda_u \lambda_v \neq 0$, $\lambda_k \in \mathbb{F}_q$ and $0 \leq u < v \leq tp^r - 1$.

Lemma 4.8. *We have*

$$G_{\mathcal{C}_{u,v}}(x) = \sum_{k=u}^{v} \frac{\lambda_k}{(1-x^t)^{k_1}\,(g^{k_0}-x)},$$

and

$$L(\mathcal{C}_{u,v}) = \begin{cases} tv_1 + |\{u \leq k \leq v : \lambda_k \neq 0\}| & \text{if } u_1 = v_1, \\ tv_1 + |\{tv_1 \leq k \leq v : \lambda_k \neq 0\}| & \text{if } u_1 < v_1. \end{cases}$$

Proof. By Proposition 4.6 we get

$$G_{\mathcal{C}_{u,v}}(x) = \sum_{n=0}^{\infty} C_n x^n = \sum_{k=u}^{v} \lambda_k \sum_{n=0}^{\infty} c_{n,k} x^n = \sum_{k=u}^{v} \frac{\lambda_k}{(1-x^t)^{k_1}\,(g^{k_0}-x)}.$$

If $u_1 = v_1$, that is $u_0 \leq v_0$, we can write $G_{\mathcal{C}_{u,v}}(x)$ as

$$\frac{1}{(1-x^t)^{v_1}} \sum_{k=u}^{v} \frac{\lambda_k}{g^{k_0}-x} = \frac{\displaystyle\sum_{k=u}^{v} \lambda_k \prod_{\substack{j=u \\ j\neq k, \lambda_j \neq 0}}^{v} (g^{j_0}-x)}{(1-x^t)^{v_1} \displaystyle\prod_{\substack{k=u \\ \lambda_k \neq 0}}^{v} (g^{k_0}-x)} =: F_1$$

and otherwise if $u_1 < v_1$ as

$$\frac{\displaystyle\sum_{k=u}^{tv_1-1} \frac{\lambda_k (1-x^t)^{v_1}}{(1-x^t)^{k_1+1}} \prod_{\substack{j=0 \\ j\neq k_0}}^{t-1} (g^j - x) \prod_{\substack{i=tv_1 \\ \lambda_i \neq 0}}^{v} (g^{i_0}-x) + \sum_{k=tv_1}^{v} \lambda_k \prod_{\substack{j=tv_1 \\ j\neq k, \lambda_j \neq 0}}^{v} (g^{j_0}-x)}{(1-x^t)^{v_1} \displaystyle\prod_{\substack{k=tv_1 \\ \lambda_k \neq 0}}^{v} (g^{k_0}-x)} =: F_2.$$

Hence

$$\sum_{n=0}^{tp^r-1} C_n x^n = (1-x^t)^{p^r} \sum_{k=u}^{v} \frac{\lambda_k}{(1-x^t)^{k_1}(g^{k_0}-x)} = \begin{cases} (1-x^t)^{p^r} F_1 & \text{if } u_1 = v_1, \\ (1-x^t)^{p^r} F_2 & \text{if } u_1 < v_1. \end{cases}$$

Then

$$\gcd\left(1-x^{tp^r}, \sum_{n=0}^{tp^r-1} C_n x^n\right) = \begin{cases} (1-x^t)^{p^r-v_1} \left(\displaystyle\prod_{\substack{k=u \\ \lambda_k \neq 0}}^{v} (g^{k_0}-x)\right)^{-1} & \text{if } u_1 = v_1, \\ (1-x^t)^{p^r-v_1} \left(\displaystyle\prod_{\substack{k=tv_1 \\ \lambda_k \neq 0}}^{v} (g^{k_0}-x)\right)^{-1} & \text{if } u_1 < v_1, \end{cases}$$

and by Proposition 3.8 we obtain

$$L(\mathcal{C}_{u,v}) = tp^r - \deg\left(\gcd\left(1 - x^{tp^r}, \sum_{n=0}^{tp^r-1} C_n x^n\right)\right)$$
$$= \begin{cases} tv_1 + |\{u \le k \le v : \lambda_k \neq 0\}| & \text{if } u_1 = v_1, \\ tv_1 + |\{tv_1 \le k \le v : \lambda_k \neq 0\}| & \text{if } u_1 < v_1. \end{cases}$$

□

For the tp^r-th expansion complexity of $\mathcal{C}_{u,v}$ we can prove the following theorem.

Theorem 4.9. *Let $\Lambda = |\{u \le k \le v : \lambda_k \neq 0\}|$ be the number of nonzero linear coefficients λ_k. For $L(\mathcal{C}_{u,v}) + L(\mathcal{C}_{u,v})^2 < tp^r$ we have*

$$E_{tp^r}(\mathcal{C}_{u,v}) = L(\mathcal{C}_{u,v}) + 1,$$

and otherwise the tp^r-th expansion complexity of the sequence $\mathcal{C}_{u,v} = (C_n)_{n\ge 0}$, $u < v$, of the form (4.10) has the lower bound

$$\left\lceil \frac{tp^r}{L(\mathcal{C}_{u,v}) + 1} \right\rceil \le E_{tp^r}(\mathcal{C}_{u,v}),$$

and can be upper bounded by the maximum of either

$$(\Lambda + 1)\left\lceil \frac{tp^r\Lambda}{tu_1 + 2} \right\rceil,$$

or

$$tu_1\left\{\frac{p^r\Lambda}{v_1}\right\} + (v_1 - u_1)\frac{tp^r\Lambda}{v_1} + (\Lambda - 1)\left\lfloor \frac{p^r\Lambda}{v_1} \right\rfloor,$$

where $\{x\} = x - \lfloor x \rfloor$ is the fractional part of x.

Proof. By Proposition 3.11 and Lemma 4.8 we get

$$E_{tp^r}(\mathcal{C}_{u,v}) = L(\mathcal{C}_{u,v}) + 1$$

if $L(\mathcal{C}_{u,v}) + L(\mathcal{C}_{u,v})^2 < tp^r$ and

$$\left\lceil \frac{tp^r}{L(\mathcal{C}_{u,v}) + 1} \right\rceil \le E_{tp^r}(\mathcal{C}_{u,v})$$

if $L(\mathcal{C}_{u,v}) + L(\mathcal{C}_{u,v})^2 \ge tp^r$.

Put

$$d = \min\left\{\left\lfloor \frac{p^r\Lambda}{v_1}\right\rfloor, \left\lceil \frac{tp^r\Lambda}{tu_1+2}\right\rceil\right\}$$

and take

$$h(x,y) = \left(y\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}(g^{k_0}-x)\right)^d - \sum_{|D|=d}\binom{d}{D}\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}(g^{k_0}-x)^{\sum_{j=u,\, j\neq k,\lambda_j\neq 0}^{v} d_j}\frac{\lambda_k^{d_k}(1-x^t)^{p^r}}{(1-x^t)^{d_k k_1}}$$

with multiindex $D=(d_u,\dots,d_v)$. Recall that

$$G_{\mathcal{C}_{u,v}}(x) = \sum_{k=u}^{v}\frac{\lambda_k}{(1-x^t)^{k_1}(g^{k_0}-x)}.$$

Then we have

$$G_{\mathcal{C}_{u,v}}(x)\left(\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}(g^{k_0}-x)\right)^d = \sum_{|D|=d}\binom{d}{D}\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}\frac{\lambda_k^{d_k}(g^{k_0}-x)^{\sum_{j=u,\, j\neq k,\lambda_j\neq 0}^{v} d_j}}{(1-x^t)^{d_k k_1}}$$

and therefore

$$h(x, G_{\mathcal{C}_{u,v}}(x)) = (1-(1-x^{tp^r})^\Lambda)\sum_{|D|=d}\binom{d}{D}\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}\frac{\lambda_k^{d_k}(g^{k_0}-x)^{\sum_{j=u,\, j\neq k,\lambda_j\neq 0}^{v} d_j}}{(1-x^t)^{d_k k_1}}$$

$$= \left(\sum_{k=1}^{\Lambda}\binom{\Lambda}{k}(x^{tp^r})^k\right)\sum_{|D|=d}\binom{d}{D}\prod_{\substack{k=u\\ \lambda_k\neq 0}}^{v}\frac{\lambda_k^{d_k}(g^{k_0}-x)^{\sum_{j=u,\, j\neq k,\lambda_j\neq 0}^{v} d_j}}{(1-x^t)^{d_k k_1}} \equiv 0 \pmod{x^{tp^r}}$$

since $\gcd(x,(1-x))=1$. Hence

$$\begin{aligned} E_{tp^r}(\mathcal{C}_{u,v}) &\le \deg(h) = \max\left\{d+d\Lambda, \max_{|D|=d}\left(tp^r\Lambda + d(\Lambda-1) - t\sum_{\substack{k=u\\ \lambda_k\neq 0}}^{v} d_k k_1\right)\right\} \\ &\le \max\{d(\Lambda+1), tp^r\Lambda + d(\Lambda-1) - tdu_1\} \\ &= \begin{cases} d(\Lambda+1) & \text{if } d = \left\lceil \frac{tp^r\Lambda}{tu_1+2}\right\rceil, \\ tp^r\Lambda + d(\Lambda-1) - tdu_1 & \text{otherwise,}\end{cases} \end{aligned}$$

since $u_1 \le k_1 \le v_1$ for all $k = u,\dots,v$ with $\lambda_k \neq 0$. Thus the result follows. □

Bibliography

[1] J. P. Allouche, J. Shallit: *Automatic Sequences. Theory, Applications, Generalizations*. Cambridge University Press, Cambridge, 2003.

[2] N. Alon, Y. Kohayakawa, C. Mauduit, C. G. Moreira, V. Rödl, Measures of pseudorandomness for finite sequences: typical values, Proc. Lond. Math. Soc. 95 (2007), 778–812.

[3] N. C. Ankeny, Sums of three squares, Proc. Amer. Math. Soc. 8 (1957), 316–319.

[4] E. R. Berlekamp: *Algebraic Coding Theory (Revised Edition)*. World Sci. Publ., Singapore, 2015.

[5] N. Brandstätter, A. Winterhof, Linear complexity profile of binary sequences with small correlation measure, Period. Math. Hungar. 52 (2006), 1–8.

[6] D. A. Burgess, The distribution of quadratic residues and non-residues. Mathematika 4 (1957), 106–112.

[7] G. Christol, Ensembles presque periodiques k-reconnaissables, Theoret. Comput. Sci. 9 (1979), 141–145.

[8] G. Christol, T. Kamae, M. Mendés France, G. Rauzy, Suites algébriques, automates et substitutions, Bull. Soc. Math. France 108 (1980), 401–419.

[9] T. W. Cusick, C. Ding, A. Renvall: *Stream Ciphers and Number Theory (Revised Edition)*. North-Holland Mathematical Library, 66. Elsevier Science B.V., Amsterdam, 2004.

[10] C. Diem, On the use of expansion series for stream ciphers, LMS J. Comput. Math. 15 (2012), 326–340.

[11] C. Ding, Pattern distributions of Legendre sequences, IEEE Trans. Inform. Theory 44 (1998), 1693–1698.

[12] C. Ding, T. Helleseth, W. Shan, On the linear complexity of Legendre sequences, IEEE Trans. Inform. Theory 44 (1998), 1276–1278.

[13] G. Dorfer, A. Winterhof, Lattice structure and linear complexity profile of nonlinear pseudorandom number generators, Appl. Algebra Engrg. Comm. Comput. 13 (2003), 499–508.

[14] M. Goresky, A. Klapper, Arithmetic crosscorrelations of feedback with carry shift register sequences, IEEE Trans. Inform. Theory 43 (1997), 1342–1345.

[15] M. Goresky, A. Klapper, Some results on the arithmetic correlation of sequences (extended abstract). In: Sequences and their Applications—SETA 2008, 71–80, Lecture Notes in Comput. Sci. 5203, Springer, Berlin, 2008.

[16] M. Goresky, A. Klapper, Statistical properties of the arithmetic correlation of sequences, Internat. J. Found. Comput. Sci. 22 (2011), 1297–1315.

[17] M. Goresky, A. Klapper: *Algebraic Shift Register Sequences*. Cambridge University Press, Cambridge, 2012.

[18] L. Goubin, C. Mauduit, A. Sárközy, Construction of large families of pseudorandom binary sequences, J. Number Theory 106 (2004), 56–69.

[19] K. Gyarmati, A. Pethő, A. Sárközy, On linear recursion and pseudorandomness, Acta Arith. 118 (2005), 359–374.

[20] R. Hofer, A. Winterhof, On the arithmetic autocorrelation of the Legendre sequence, Adv. Math. Commun. 11 (2017), no. 1, 237–244.

[21] R. Hofer, A. Winterhof, Linear complexity and expansion complexity of some number theoretic sequences. In: S. Duquesne, S. Petkova-Nikova (eds.), Arithmetic of Finite Fields (WAIFI 2016), 67–74, Lecture Notes in Comput. Sci. 10064, Springer, Cham, 2016.

[22] R. Hofer, L. Mérai, A. Winterhof, Measures of pseudorandomness: Arithmetic autocorrelation and correlation measure. In: C. Elsholtz, P. Grabner (eds.), Number Theory – Diophantine Problems, Uniform Distribution and Applications, Festschrift in honour of Robert F. Tichy's 60th birthday, Springer, to appear.

[23] D. Jungnickel: *Finite Fields, Structure and Arithmetics*. Bibliographisches Institut, Mannheim, 1993.

[24] N. Koblitz: *A Course in Number Theory and Cryptography (Second Edition)*. Springer, New York, 1987.

[25] R. Lidl, H. Niederreiter: *Finite Fields*. Encyclopedia of Mathematics and its Applications, 20. Addison-Wesley Publishing Company, Advanced Book Program, Reading, MA, 1983.

[26] D. Mandelbaum, Arithmetic codes with large distance, IEEE Trans. Inform. Theory 13 (1967), 237–242.

[27] J. L. Massey, Shift-register synthesis and BCH decoding, IEEE Trans. Inform. Theory IT-15 (1969), 122–127.

[28] C. Mauduit, A. Sárközy, On finite pseudorandom binary sequences. I. Measure of pseudorandomness, the Legendre symbol, Acta Arith. 82 (1997), 365–377.

[29] C. Mauduit, A. Sárközy, On finite pseudorandom sequences of k symbols, Indag. Math. 13 (2002), 89–101.

[30] W. Meidl, A. Winterhof, Linear complexity of sequences and multisequences. In: Handbook of Finite Fields, 324–336, Discrete Mathematics and Its Applications, CRC Press, Boca Raton, FL, 2013.

[31] L. Mérai, H. Niederreiter, A. Winterhof, Expansion complexity and linear complexity of sequences over finite fields, Cryptogr. Commun. 9 (2017), 501–509.

[32] L. Mérai, A. Winterhof, On the Nth linear complexity of p-automatic sequences over $\mathbb{F}_p$. Preprint 2016.

[33] H. Niederreiter, A. Winterhof, Lattice structure and linear complexity of nonlinear pseudorandom numbers, Appl. Algebra Engrg. Comm. Comput. 13 (2002), 319–326.

[34] H. Niederreiter, Linear complexity and related complexity measures for sequences. In: Progress in Cryptology—INDOCRYPT 2003, 1–17, Lecture Notes in Comput. Sci. 2904, Springer, Berlin, 2003.

[35] H. Niederreiter, A. Winterhof: *Applied Number Theory*. Springer, Cham, 2015.

[36] I. Niven, H. S. Zuckerman, H. L. Montgomery: *An Introduction to the Theory of Numbers (Fifth Edition)*. John Wiley & Sons, New York, 1991.

[37] R. E. A. C. Paley, On orthogonal matrices, J. Math. Physics 12 (1933), 311–320.

[38] O. Perron, Bemerkungen über die Verteilung der quadratischen Reste, Math. Z. 56 (1952), 122–130.

[39] I. Shparlinski: *Cryptographic Applications of Analytic Number Theory. Complexity Lower Bounds and Pseudorandomness*. Progress in Computer Science and Applied Logic, 22. Birkhäuser Verlag, Basel, 2003.

[40] T. Storer: *Cyclotomy and Difference Sets*. Lectures in Advanced Mathematics, 2. Markham Publishing Co., Chicago, 1967.

[41] A. Topuzoğlu, A. Winterhof, Pseudorandom sequences. In: Topics in Geometry, Coding Theory and Cryptography, 135–166, Algebr. Appl. 6, Springer, Dordrecht, 2007.

[42] R. J. Turyn, The linear generation of Legendre sequence, J. Soc. Indust. Appl. Math. 12 (1964), 115–116.

[43] A. Winterhof, Linear complexity and related complexity measures. In: Selected Topics in Information and Coding Theory, 3–40, Ser. Coding Theory Cryptol. 7, World Sci. Publ., Singapore, 2010.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Dissertation selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe. Die vorliegende Dissertation ist mit dem elektronisch übermittelten Textdokument identisch.

Linz, März 2017

_____

Richard Hofer

Curriculum Vitae

Name: Richard Hofer

Nationality: Austria

Date of Birth: 30 July 1989

Place of Birth: Linz, Austria

Education:

1999–2007: Georg von Peuerbach Gymnasium (grammar school), Linz

2007–2012: Bachelor Programme in Technical Mathematics, JKU Linz

2012–2015: Masters Programme in Industrial Mathematics, JKU Linz

Since 2015: Doctoral Programme in Technical Sciences, JKU Linz

Publications:

R. Hofer, A. Winterhof, On the arithmetic autocorrelation of the Legendre sequence, AMC 11 (2017), no. 1, 237–244.

R. Hofer, A. Winterhof, Linear complexity and expansion complexity of some number theoretic sequences. In: S. Duquesne, S. Petkova-Nikova, Arithmetic of Finite Fields (WAIFI 2016), 67–74, LNCS 10064, Springer, Cham, 2016.

R. Hofer, L. Mérai, A. Winterhof, Measures of pseudorandomness: Arithmetic autocorrelation and correlation measure. In: C. Elsholtz, P. Grabner, Number Theory – Diophantine Problems, Uniform Distribution and Applications, Festschrift in honour of Robert F. Tichy's 60th birthday, Springer, to appear.