rijksuniversiteit groningen
faculteit Wiskunde en Natuurwetenschappen

The overlap distribution of paperfolding sequences

Bacheloronderzoek Wiskunde

Augustus 2010

Student: E. de Groote

Begeleider: Prof. dr. A.C.D. van Enter

Abstract

In this report we will discuss non-periodic bi-infinite sequences. The elements of these sequences have value -1 or $+1$. We will consider sequences of four different types. Sequences of these types can be constructed in a certain manner. Based on this manner of constructing, the overlap distribution of sequences of each of these four types can be determined. We will demonstrate known results about three types of sequences. These are about Thue-Morse sequences, Toeplitz sequences and Fibonacci sequences. For these three types of sequences a manner of constructing the sequences of these types and the overlap distribution is known. With the aid of these results the overlap distribution of paperfolding sequences will be determined.

Contents

Chapter 1

Introduction

In this report, we will discuss the overlap distribution of four different types of non-periodic bi-infinite sequences. These types are Thue-Morse, Toeplitz, Fibonacci and paperfolding. The overlap distributions of the first three types of sequences are known. These types will be studied and are used to determine the overlap distribution of paperfolding sequences, which is the aim of this report.

In Chapter 2 definitions and theorems relevant to the subject of this report will be given. Amongst other things we will describe what kind of sequences we will discuss later on in this report and introduce the terms overlap and overlap distribution.

In Chapter 3 three known results will be discussed. These results are about Thue-Morse sequences, Toeplitz sequences and Fibonacci sequences. This are three different types of sequences. We will describe how the sequences of each type can be constructed and what the overlap distribution of each of these types of sequences is.

In Chapter 4 we will give the definition of a paperfolding sequence. Also a manner of constructing paperfolding sequences will be introduced in this chapter.

In Chapter 5 the overlap distribution of paperfolding sequences will be determined. This will be done by using the manner of constructing paperfolding sequences given in the third chapter.

In Chapter 6 we will give a conclusion about the overlap distribution of paperfolding sequences.

Chapter 2

Model and definitions

2.1 Probability space

Later on in this report we will discuss sequences in a probability space, which is a particular instance of a measure space. Therefore we first define what a measure space is.

Definition 2.1. *A measure space is a triple* $(\Omega, \mathcal{F}, \mu)$*.* Ω *is a non-empty set, also referred to as sample space, which is the set of all possible outcomes.* $\mathcal{F}$ *is a* σ*-algebra and* μ *is the measure.*

A collection $\mathcal{F}$ *of subsets of* Ω *is a* σ*-algebra of subsets of* Ω *is* $\mathcal{F}$ *satisfies the following three conditions.*

1. $\Omega \in \mathcal{F}$,
2. $\forall\ F \in \mathcal{F} : F^c = \Omega \setminus F \in \mathcal{F}$,
3. *the union of any countable collection of sets in* $\mathcal{F}$ *is again in* $\mathcal{F}$*:*

$$\textit{if } F_n \in \mathcal{F} \textit{ with } n \geq 1 \textit{ then } \bigcup_{n=1}^{\infty} F_n \in \mathcal{F}.$$

The elements of $\mathcal{F}$ *are called measurable subsets of* Ω*. Thus* $F \in \mathcal{F}$ *and "*F *is a measurable subset of* Ω*" have the same meaning.*

The measure μ *is a function* $\mu : \mathcal{F} \to [0, \infty)$*, defined on a* σ*-algebra* $\mathcal{F}$ *of* Ω*, that satisfies the following three properties:*

1. $\mu(\varnothing) = 0$,
2. $\mu(F) \geq 0 \quad \forall F \in \mathcal{F}$,

3. *for all countable collections* $\{F_k\}_1^\infty$ *of elements* $F_k \in \mathcal{F}$ *which are pairwise disjoint subsets, so* $F_i \bigcap F_j = \varnothing$ *if* $i \neq j$ *is:*

$$\mu\left(\bigcup_{k=1}^{\infty} F_k\right) = \sum_{k=1}^{\infty} \mu(F_k).$$

From now on we will discuss measure spaces that are probability spaces. These are spaces with total measure $\mu(\Omega) = 1$, thus $\mu : \mathcal{F} \to [0,1]$. The measure in these spaces is called probability measure. In a probability space the σ-algebra $\mathcal{F}$ is called a set of events.

2.2 Model

Definition 2.2. *A sequence is a function which has domain* $\mathbb{N}$ *or* $\mathbb{Z}$*.*

We will discuss infinite sequences which have $\{-1, +1\}$ as range of the function. In later chapters we will sometimes write $-$ and $+$ instead of -1 and $+1$.

Definition 2.3. *An infinite sequence with domain* $\mathbb{N}$ *(or an bi-infinite sequence with domain* $\mathbb{Z}$*) and range* $\{-1, +1\}$ *is a map:*

$$\sigma : \mathbb{N}(or\, \mathbb{Z}) \to \{-1, +1\}.$$

Given such a map we denote $\sigma(n)$ *by* σ_n*.*

The sequences we will go through are bi-infinite. To construct a suitable model we use the one-dimensional lattice of integers $\mathbb{Z}$. In this model, each element of a sequence is labeled by an integer. The sample space Ω is the space of all bi-infinite sequences which have range $\{-1, +1\}$ and are defined on $\mathbb{Z}$: $\Omega = \{-1, +1\}^{\mathbb{Z}}$. σ_i is defined as the *ith* element of a sequence, so $\sigma_i \in \{-1, +1\}$. When we denote the sequences by X, they are given by:

$$\begin{aligned}\Omega &= \{-1, +1\}^{\mathbb{Z}} \\ &= \{X = (\ldots, \sigma_{-1}(X), \sigma_0(X), \sigma_1(X), \ldots) : \sigma_i(X) \in \{-1, +1\}\, \forall i \in \mathbb{Z}\}.\end{aligned}$$

By elements $F \in \mathcal{F}$ we can think for example the set of sequences which have $\sigma_0 = +1$ or $\sigma_5\sigma_6\sigma_7 = -1 + 1 - 1$.

In this report we will study the overlap distribution of non-periodic sequences. Therefore we first state the definition of a periodic sequence.

Definition 2.4. *A periodic sequence is a sequence* σ_n *satisfying the following property:*

$$\sigma_{n+p} = \sigma_n \qquad \forall n \in \mathbb{N}.$$

where p *is the period of* σ*.*

This means that the sequences we are discussing don't have a number p for which definition 2.4 holds.

2.3 System

We will consider maps between a sample space Ω and itself. First we give the general definitions of measurability and measure-preservingness of a map between sample spaces.

Definition 2.5. *Suppose* $(\Omega_1, \mathcal{F}_1, \mu_1)$ *and* $(\Omega_2, \mathcal{F}_2, \mu_2)$ *are probability spaces. Let a transformation* V *be a mapping*

$$V : \Omega_1 \to \Omega_2$$

from the sample space Ω_1 *to the sample space* Ω_2*.*

- V *is measurable if* $\forall\ F_2 \in \mathcal{F}_2\ :\ V^{-1}F_2 \in \mathcal{F}_1$*, with*

  $$V^{-1}F_2 := \{X \in \Omega_1 : VX \in F_2\}$$

  and hence $V^{-1}\mathcal{F}_2 \subset \mathcal{F}_1$*.*

- V *is measure-preserving if* V *is measurable and* V *preserves the measure:*

  $$\mu_1(V^{-1}F_2) = \mu_2(F_2) \qquad \forall F_2 \in \mathcal{F}_2.$$

When we have a sequence of a certain type (for example Thue-Morse, Toeplitz, Fibonacci, paperfolding; see chapter 3 and 4) all sequences of that type can be found by shifting the given sequence and including its limit-points. Shifting a sequence is done by using the shift S. The shift is defined in definition 2.6.

Definition 2.6. *The map* $S : \Omega \to \Omega$ *will be called the shift if for all* $X \in \Omega$ *a shifted sequence* $S(X)$ *is given by*

$$S(\sigma_i(X)) = \sigma_{i+1}(X) \qquad \forall i \in \mathbb{Z}$$

It is clear that the shift is a map from a set Ω to itself. For a map from a set to itself, measurability and measure-preservingness of a transformation are defined as follows:

Definition 2.7. *Suppose* $(\Omega, \mathcal{F}, \mu)$ *is a probability space. Let a transformation* W *be a mapping*

$$W : \Omega \to \Omega$$

from the sample space to itself.

- W *is measurable if* $\forall\ F \in \mathcal{F}\ :\ W^{-1}F \in \mathcal{F}$*, with*

  $$W^{-1}F := \{X \in \Omega : TX \in F\}$$

  and hence $W^{-1}\mathcal{F} \subset \mathcal{F}$*.*

- *W is measure-preserving if W is measurable and W preserves the measure:*

$$\mu(W^{-1}F) = \mu(F) \qquad \forall F \in \mathcal{F}.$$

Later on we will discuss the union of all sequences of the same type. Therefore we need definitions 2.8 and 2.9

Definition 2.8. *Let S be the shift. For a sequence $X \in \Omega$ the orbit of X, $Orb(X)$, is given by*

$$Orb(X) = \{S^k X : k \in \mathbb{Z}\}$$

The closure of the orbit of X, $\overline{Orb(X)}$, is the union of the orbit of X and its limitpoints.

Definition 2.9. *The 4-tuple $(\Omega, \mathcal{F}, \mu, T)$ with $(\Omega, \mathcal{F}, \mu)$ a measure space and T a measure-preserving transformation of $(\Omega, \mathcal{F}, \mu)$ is called a system.*

2.4 Ergodicity

Later on in this report we will introduce the term frequency. For this it will turn out to be useful to know something about ergodicity. This is addressed in theorem 2.10 and definition 2.11.

Theorem 2.10. *[2]. Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and suppose T is a measure-preserving transformation on $(\Omega, \mathcal{F}, \mu)$. Let $F \in \mathcal{F}$ and $X \in \Omega$. Define*

$$K_n(F, X) := number\{k : 0 \leq k < n, T^k(X) \in F\}$$

and

$$L_n(F, X) := \frac{1}{n} K_n(F, X)$$

Then for almost every $X \in \Omega$ (except possibly a set of X for which $\mu(X) = 0$) the limit

$$L(F, X) := \lim_{n \to \infty} L_n(F, X)$$

exists.

Definition 2.11. *[8]. The system $(\Omega, \mathcal{F}, \mu, T)$ is ergodic if the following holds:*

$$\forall F \in \mathcal{F} \text{ with } F = T^{-1}F \ : \ \mu(F) = 0 \text{ or } \mu(F) = 1$$

We will also use the terms "T is (μ-)ergodic" and "μ is (T-)ergodic".

If the measure-preserving transformation T is ergodic it then follows, see e.g. [2], that in theorem 2.10 holds that

$$L(F, X) = \mu(F).$$

Next, some general definitions will be given. These definitions will be applied later in this chapter to explain what kind of systems we will discuss in the remainder of this report.

Definition 2.12. *A subset $\Omega_S \subset \Omega$ is an T-invariant set if $\Omega_S = T\Omega_S$ and Ω_S is a closed set if $\Omega_S = \overline{\Omega_S}$.*

Definition 2.13. *A closed T-invariant subset $\Omega_S \subset \Omega$, $\Omega_S \neq \emptyset$, is minimal if for each closed T-invariant subset $\omega_S \subset \Omega_S$, $\omega_S \neq \emptyset$, holds that $\omega_S = \Omega_S$.*

Theorem 2.14. *[5]. A subset $\Omega_S \subset \Omega$, $\Omega_S \neq \emptyset$, is minimal $\Leftrightarrow \Omega_S = \overline{Orb(X)}$ $\forall\, X \in \Omega_S$.*

Definition 2.15. *[5]. Let μ be a probability measure on $\Omega_S \subset \Omega$, $\Omega_S \neq \emptyset$ and $\Omega_S = T\Omega_S = \overline{\Omega_S}$. μ is T-invariant if $\mu(F) = \mu(T^{-1}F)$ $\forall\, F \in \mathcal{F}$.*

From [5] we know that Ω_S allows one or more T-invariant probability measures.

Definition 2.16. *A subset $\Omega_S \subset \Omega$, $\Omega_S \neq \emptyset$ and $\Omega_S = T\Omega_S = \overline{\Omega_S}$, is uniquely ergodic if Ω_S admits a unique T-invariant probability measure μ.*

In the following definition $C(\Omega)$ and a function f will be mentioned. $C(\Omega)$ is the space of continuous functions on Ω and an example of such a function f can be given by $f = \sigma_0$ or $f = \sigma_0\sigma_5\sigma_{15}$.

Definition 2.17. *[5][6]. A sequence $X \in \Omega$ is strictly transitive if for each function $f \in C(\Omega)$ the limit*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} f(T^{i+k}X)$$

exists and is uniform in $k \in \mathbb{Z}$.

This means that a sequence $X \in \Omega$ is strictly transitive if given $\varepsilon > 0$ there exists an $N(\varepsilon)$ such that $\forall\, k \in \mathbb{Z}$ and $N > N(\varepsilon)$

$$\left| \frac{1}{N} \sum_{i=1}^{N} f(T^{i+k}X) - \mu(f) \right| < \varepsilon.$$

Theorem 2.18. *[5][6]. For $X \in \Omega$ the following holds:*

$$\overline{Orb(X)} \text{ is uniquely ergodic} \Leftrightarrow X \text{ is strictly transitive.}$$

Definition 2.19. *A subset $\Omega_S \in \Omega$, $\Omega_S \neq \emptyset$ and $\Omega_S = T\Omega_S = \overline{\Omega_S}$, is strictly ergodic if Ω_S is uniquely ergodic and minimal.*

2.5 Strictly ergodic systems

In the following chapters we discuss systems of sequences of a certain type. So if X is a sequence of the type we are interested in, then we consider the system $(\overline{Orb(X)}, \mathcal{F}_{\overline{Orb(X)}}, \mu_{\overline{Orb(X)}}, S)$. It is clear that $\overline{Orb(X)} \subset \Omega$.

To better clarify what kind of systems will be discussed, we introduce the term density of a finite pattern. A finite pattern is a finite sequence which has range $\{-1, +1\}$.

Definition 2.20. *We denote a finite pattern of a sequence by P and a segment $\{-l, +l\}$ of $\mathbb{Z}$ by l with $l \neq 0$. The number of times P occurs in l is denoted by $N_l(P)$. The size of l is $N(l) = 2l + 1$. The finite-interval density of P in l, $finite - intervaldensity(P, l)$, is given by*

$$finite - intervaldensity(P, l) = \frac{N_l(P)}{N(l)}.$$

Definition 2.21. *The density of a finite pattern P, density(P), is obtained by taking the limit of the finite-interval density.*

$$density(P) = \lim_{l \to \infty} \frac{N_l(P)}{N(l)}$$

$L_n(F, X)$ (theorem 2.10) and the finite-interval density are in essence the same. This is because the shift S is a measure-preserving transformation and we kan take F such that it contains the pattern. From theorem 2.10 follows that the density of P, $density(P)$, for almost every sequence exists (except possibly for a set of sequences with measure zero). Because S is ergodic $density(P) = \mu(P)$.

By definition 2.12 it is clear that $\overline{Orb(X)}$ is a closed S-invariant set. Futhermore $\overline{Orb(X)} \neq \emptyset$.

The systems we will discuss concern non-periodic sequences which are strictly transitive. From theorem 2.18 follows that $\overline{Orb(X)}$ is uniquely ergodic. By definition 2.16 we know that there exists a unique S-invariant probability measure $\mu_{\overline{Orb(X)}}$ on $\overline{Orb(X)}$. $\mu_{\overline{Orb(X)}}$ consist of the densities of all finite patterns because we know this probability measure exists by theorem 2.10 and definition 2.21. $\overline{Orb(X)}$ is also minimal (theorem 2.14) so by definition 2.19 we know that $\overline{Orb(X)}$ is strictly ergodic.

This means that we consider systems that are strictly ergodic. From [1] follows that in such systems for every finite pattern that occurs in a sequence belonging to the system holds that the uniformly defined density is strictly larger than zero.

2.6 Overlap distribution

We want to determine the overlap distribution of non-periodic bi-infinite sequences of a certain type. It will be helpful to first state the definitions of overlap and overlap distribution.

Definition 2.22. *For two sequences X and $Y \in \Omega$ the overlap, q_{XY}, is given by*

$$q_{XY} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(X)\sigma_i(Y) \quad \textit{with } N \in \mathbb{N}$$

if it exists.

The overlap distribution of q_{XY}, $p(q)$, is given by

$$p(q) = \int_{X,Y\in\Omega} p_{XY}(q)\delta(q - q_{XY})\mu \otimes \mu(dXdY)$$

where $p_{XY}(q)$ is the probability density of a certain overlap q_{XY}.

Parisi introduced these quantities in the theory of spin glasses. However we will not further discuss this background here.

Chapter 3

Known results

3.1 Thue-Morse sequences

Thue-Morse sequences are sequences which can be obtained by the following substitution:

$$\begin{cases} +1 & \rightarrow & +1 & -1 \\ -1 & \rightarrow & -1 & +1 \end{cases}$$

A possible beginning of a Thue-Morse sequence is shown in figure 3.1.

$$- + + - + - - + + - - + - + + - + - - + - + + - - + + - + - - + \ldots$$

Figure 3.1: Beginning of a Thue-Morse sequence.

The Thue-Morse system is the orbit closure of this Thue-Morse sequence.

It is also possible to construct a Thue-Morse sequence as described below. This idea is described e.g. in [1].

Take the sequence where all elements of $\mathbb{Z}$ have value $+1$.

$$+ + + + + + + + + + + + + + + +$$

Step 1. Multiply the value of every second element by -1.

$$- + - + - + - + - + - + - + - +$$

Step 2. Divide the sequence in blocks of two elements. Multiply the value of the elements in every second block by -1.

$$+ - - + + - - + + - - + + - - +$$

Step 3. Divide the sequence in blocks of $2^{3-1} = 4$ elements. Multiply the value of the elements in every second block by -1.

$$- + + - + - - + - + + - + - - +$$

Step n. Divide the sequence in blocks of 2^{n-1} elements. Multiply the value of the elements in every second block bij -1.

When we think of these steps as elements of a sequence, a limitpoint of this sequence of sequences is a non-periodic Thue-Morse sequence. So this Thue-Morse sequence is a non-periodic sequence which is a limitpoint of periodic sequences.

Theorem 3.1. *The overlap distribution $p(q)$ for μ_{TM} is a point measure concentrated on $q = 0$.[1]*

The proof uses the fact that the diffraction spectrum is continuous, see [1].

3.2 Toeplitz sequences

Toeplitz (or period-doubling) sequences are sequences which can be obtained by the following substitution:

$$\begin{cases} +1 & \rightarrow & -1 \quad -1 \\ -1 & \rightarrow & -1 \quad +1 \end{cases}$$

A possible beginning of a Toeplitz sequence is shown in figure 3.2.

$- + - - - + - + - + - - - + - - - + - - - + - + - + - - - + - + \ldots$

Figure 3.2: Beginning of a Toeplitz sequence.

The Toeplitz system is the orbit closure of this Toeplitz sequence.

It is also possible to construct a Toeplitz sequence as described below. This idea is described e.g. in [1].

Step 1. Occupy one of the two sublattices of $\mathbb{Z}$ of period 2, so the even or odd integers. Place a -1 on every site of the sublattice.

$$\begin{array}{rcccccccccccc} \text{even integers :} & - & \cdot & - & \cdot & - & \cdot & - & \cdot & - & \cdot & - & \cdot \\ \text{odd integers :} & \cdot & - & \cdot & - & \cdot & - & \cdot & - & \cdot & - & \cdot & - \end{array}$$

Step 2. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period $4\,(= 2^2)$. We can also speak of these two sublattices as the even and odd holes. Place a $(-1)^2 = +1$ on every site of the sublattice, so a $+1$ in the even or odd holes.

$$\begin{array}{ccccccccccccccccc} - & \cdot & - & \cdot & - & \cdot & - & \cdot & \Rightarrow & - & + & - & \cdot & - & + & - & \cdot \\ & & & & & & & & \Rightarrow & - & \cdot & - & + & - & \cdot & - & + \end{array}$$

$$\begin{array}{ccccccccccccccccc} \cdot & - & \cdot & - & \cdot & - & \cdot & - & \Rightarrow & + & - & \cdot & - & + & - & \cdot & - \\ & & & & & & & & \Rightarrow & \cdot & - & + & - & \cdot & - & + & - \end{array}$$

Step 3. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period $8\,(=2^3)$ and place a $-1=(-1)^3$ on every site of the sublattice. There are now 8 possible sequences.

Step n. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period 2^n and place a $(-1)^n$ on every site of the sublattice. This gives 2^n possible sequences.

The sequence that arises when $n \to \infty$ is a Toeplitz sequence.

The Toeplitz system consists of all possible sequences constructed as described above. So, all possible choices of sublattices of period 2^i, with $i \in \mathbb{N}$, give rise to different sequences. A Toeplitz sequence is a non-periodic sequence which is a limitpoint of periodic sequences.

Theorem 3.2. *The overlap distribution $p(q)$ for μ_T contains countably many points.[1]*

Proof. The idea of the proof comes from [1].
For determining the overlap distribution we can choose an arbitrary sequence and determine the overlap between the chosen sequence and each of the other sequences. Moreover we have to determine the probability of this overlap. This is an appropriate way to find the overlap distribution because the probability of a certain overlap q_{XY} between two arbitrary sequences X and Y is equal to the probability of the same overlap between the chosen sequence and an arbitrary other one. This holds because the probability measure is translation-invariant.

The sequence we choose to use to determine the overlap distribution is given in figure 3.3.

$$-+---+-+-+---+---+---+-+-+---+-+-+--$$

Figure 3.3: The sequence used to determine the overlap distribution, $(*)$

This sequence arises from choosing in each step the sublattice which represents the even holes. We will refer to this sequence by $(*)$.

Next, we will determine the overlap distribution. This will be done by using the steps of constructing a Toeplitz sequence.

Step 1. When constructing a Toeplitz sequence in the first step there are two possible options for sublattices of $\mathbb{Z}$ of period 2. These two possibilities both happen with probability $\frac{1}{2}$. In figure 3.4 these two possibilities are shown together with the first step of $(*)$.
We denote the overlap between $(*)$ and sequences for which the first step in the construction of these sequences is equal to (i) by $q_{(*)(i)}$.

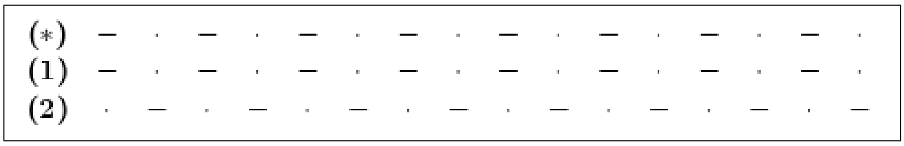

Figure 3.4: Toeplitz step 1.

Result step 1:

$$q_{(*)(2)} = -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots = -\frac{1}{3} \tag{3.1}$$

This is true because:

- Concerning $(*)$, in the first step is chosen for the sublattice consisting of the even integers. The elements of this sublattice have value -1. Concerning (2) in the first step is chosen for the sublattice of the odd integers. So each of the sublattices in the following steps will be in the even integers. This means that the even integers in (2) will consist of $+1$ for $\frac{1}{2}$ part, of -1 for $\frac{1}{4}$ part, of $+1$ for $\frac{1}{8}$ part etc. This means that the overlap in the even integers of $(*)$ and (2) is equal to:

$$\begin{aligned} q_{(*)(2)_{evenintegers}} &= \lim_{N\to\infty} \frac{2}{N} \sum_{i=1}^{N/2} \sigma_{2i}(*)\sigma_{2i}(2) \\ &= \lim_{N\to\infty} \frac{2}{N}(\frac{N}{4} \times (-1) + \frac{N}{8} \times (+1) + \frac{N}{16} \times (-1) + \ldots) \\ &= -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots \end{aligned}$$

  With this, $N \in \mathbb{N}$ should be even. Otherwise $\frac{N}{2} \notin \mathbb{N}$.

- This also holds for the odd integers, because then in (2) the odd integers have value -1 and the sublattices of $(*)$ from the second step are in the odd integers.

$$q_{(*)(2)_{oddintegers}} = \lim_{N\to\infty} \frac{2}{N} \sum_{i=1}^{N/2} \sigma_{2i-1}(*)\sigma_{2i-1}(2) = -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} \times (-1) + \ldots$$

  Again, N should be even.

- From these two findings follows that:

$$\begin{aligned} q_{(*)(2)} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(2) \\ &= \lim_{N\to\infty} \frac{1}{N}(\frac{N}{2} \times (-1) + \frac{N}{4} \times (+1) + \frac{N}{8} \times (-1) + \ldots) \\ &= -\frac{1}{2} + \frac{1}{4} + -\frac{1}{8} + \ldots \end{aligned}$$

For determining the overlap between $(*)$ and (1) it is necessary to know which sublattices are chosen in the following steps. However, (1) will not give a contribution to the probability of an overlap of $-\frac{1}{3}$. This holds because from the even integers $q_{(*)(1)}$ is already $\frac{1}{2}$, which can never become $-\frac{1}{3}$ because there are not enough elements left to realise that.

Thus with probability $\frac{1}{2}$ we find overlap $-\frac{1}{3}$.

Step 2. The overlap of $(*)$ and (2) is known, we are now concerned with (1). For these sequences in the second step there are two possible options for sublattices of $\mathbb{Z}$ of period 4. In this step both options occur with probability $\frac{1}{2}$ so the total probability of each one of these options is $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$. In figure 3.5 these two options are shown together with the first two steps of $(*)$.

$(*)$	−	+	−	·	−	+	−	·	−	+	−	·	−	+	−	·
(1.1)	−	+	−	·	−	+	−	·	−	+	−	·	−	+	−	·
(1.2)	−	·	−	+	−	·	−	+	−	·	−	+	−	·	−	+

Figure 3.5: Toeplitz step 2.

Result step 2:

$$q_{(*)(1.2)} = \frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + \ldots = \frac{1}{2} - \frac{1}{3} \times \frac{1}{2} = \frac{1}{3} \tag{3.2}$$

This is true because:
From the second step this case is similar to showing that $q_{(*)(2)} = -\frac{1}{3}$. However in this case we start with the choice between the sublattices of period 4.

- In $(*)$ the elements of the sublattice which consists of the even holes have value +1. These holes in (1.2) will consist of −1 for $\frac{1}{2}$ part, of +1 for $\frac{1}{4}$ part, of −1 for $\frac{1}{8}$ part etc. So the overlap in the even holes is equal to:

$$\begin{aligned} q_{(*)(1.2)_{evenholes}} &= \lim_{N\to\infty} \frac{4}{N} \sum_{i=1}^{N/4} \sigma_{1+4(i-1)}(*)\sigma_{1+4(i-1)}(1.2) \\ &= \lim_{N\to\infty} \frac{4}{N}\left(\frac{N}{8} \times (-1) + \frac{N}{16} \times (+1) + \frac{N}{32} \times (-1) + \ldots\right) \\ &= -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \ldots \end{aligned}$$

  With this, $N \in \mathbb{N}$ should be divisible by 4. Otherwise $\frac{N}{4} \notin \mathbb{N}$.

- This also holds for the odd holes, because then the odd holes have value +1 in (1.2). And the sublattices of $(*)$ from the second step are in the odd holes.

$$q_{(*)(1.2)_{oddholes}} = \lim_{N\to\infty} \frac{4}{N} \sum_{i=1}^{N/4} \sigma_{3+4(i-1)}(*)\sigma_{3+4(i-1)}(1.2)$$
$$= -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \dots$$

Again, N shoud be divisible by 4.

- Because $q_{(*)(1.2)_{evenholes}} = q_{(*)(1.2)_{oddholes}}$ it follows that

$$q_{(*)(1.2)_{holes}} = -\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \dots$$

- $q_{(*)(1.2)_{holes}}$ gives the contribution of the odd integers to $q_{(*)(1.2)}$. Together with $q_{(*)(1.2)_{evenintegers}}$ we get $q_{(*)(1.2)}$. With this we use that $\sigma_i(1.2) = \sigma_i(*) = -1$ als i even.

$$\begin{aligned} q_{(*)(1.2)} &= \frac{1}{2} \times q_{(*)(1.2)_{evenintegers}} + \frac{1}{2} \times q_{(*)(1.2)_{holes}} \\ &= \frac{1}{2} \times 1 + \frac{1}{2} \times \left(-\frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \dots\right) \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{8} - \frac{1}{16} + \dots = \frac{1}{2} + \left(\frac{1}{2} \times \left(-\frac{1}{3}\right)\right) = \frac{1}{3} \end{aligned}$$

For determining the overlap between (∗ and (1.1) it is required to know more about the sublattices that are chosen in the following steps.

Thus with probability $\frac{1}{4}$ we find overlap $\frac{1}{3}$.

Step 3. For (1.1) in the third step there are two possible options for sublattices of $\mathbb{Z}$ of period 8. In this step both options occur with probability $\frac{1}{2}$ so the total probability of each one of these options is $\frac{1}{2} \times \frac{1}{4} = \frac{1}{8}$. In figure 3.6 these two options are shown together with the first three steps of (∗).

(∗)	−	+	−	−	−	+	−	·	−	+	−	−	−	+	−	·
(1.1.1)	−	+	−	−	−	+	−	·	−	+	−	−	−	+	−	·
(1.1.2)	−	+	−	·	−	+	−	−	−	+	−	·	−	+	−	−

Figure 3.6: Toeplitz step 3.

Result step 3:

$$q_{(*)(1.1.2)} = \frac{1}{2} + \frac{1}{4} - \frac{1}{8} + \frac{1}{16} - \frac{1}{32} + \dots = \frac{1}{2} + \frac{1}{4} + \left(\frac{1}{4} \times \left(-\frac{1}{3}\right)\right) = \frac{2}{3} \quad (3.3)$$

This holds because of the same arguments we used above.

Thus with probability $\frac{1}{8}$ we find overlap $\frac{2}{3}$.

When all Toeplitz sequences are arranged like this the following overlap distribution arises.[1]

$$p(q) = \sum_{n=0}^{\infty} \left(\frac{1}{2^{n+1}} \right) \delta \left(q - \frac{3 \cdot 2^{n-2} - 1}{3 \cdot 2^{n-2}} \right) \tag{3.4}$$

It is clear that this overlap distribution contains countably many points. □

3.3 Fibonacci sequences

Fibonacci sequences are sequences that can be obtained by the following substitution:

$$\begin{cases} +1 & \rightarrow & +1 \quad -1 \\ -1 & \rightarrow & +1 \end{cases}$$

A possible beginning of a Fibonacci sequence is shown in figure 3.7.

+ − + + − + − + + − + + − + − + + − + − + ...

Figure 3.7: Beginning of a Fibonacci sequence.

The Fibonacci system is the orbit closure of this Fibonacci sequence.

It is also possible to construct a Fibonacci sequence as described below. This idea is described e.g. in [1].

- Denote the rotations over the circle by an amount of $2\pi\gamma$ by T. With $\gamma = \frac{2}{1+\sqrt{5}} \in [0, 1)$.
- When $\alpha \in [0, 1)$ there corresponds a Fibonacci sequence F to every angle $2\pi\alpha \in [0, 2\pi)$. Thus to every angle in the circle.
- If $T^n\alpha \in [0, 2\pi\gamma)$ then $F(n) = 1$, otherwise $F(n) = -1$.
- $\gamma \approx 0.618034...$, so the possibility to get a $+1$ is larger than the possibility to get a -1. This agrees with the substitution.

The Fibonacci system consists of all possible constructions of Fibonacci sequences. Thus all possible angles $2\pi\alpha$.

Theorem 3.3. *The overlap distribution $p(q)$ for μ_F has a continuous part.[1]*

The proof uses the representation of Fibonacci sequences as orbits of an irrational circle rotation, see [1].

From [10] follows that every Sturmian sequence is a rotation sequence. Fibonacci sequences are examples of Sturmian sequences. This means that the overlap distributions of Sturmian sequences also have a continuous part.

Chapter 4

Paperfolding sequences

The paperfolding sequence is also a sequence $\in \Omega = \{-1,+1\}^{\mathbb{Z}}$. The name comes from (un)folding a sheet of paper. A sheet of paper can be folded into two directions, right and left. When a paper which has been folded infinitely many times is unfolded, the paperfolding sequence arises from unfolding this paper. The paper then shows left and right folds, which correspond to respectively +1 and −1.

Definition 4.1. *[3][4]. A paperfolding sequence $\sigma \in \Omega = \{-1,+1\}^{\mathbb{Z}}$ can be obtained by*

$$\begin{aligned} \sigma_{4in+j} &= +1 \\ \sigma_{4in+\frac{4i}{2}+j} &= -1 \end{aligned}$$

with $i = 2^k$, $j = 2^k - 1$ and $n \in \mathbb{Z}$, $k \in \mathbb{N}$.

Thus for $k = 0$ we have

$$\begin{aligned} \sigma_{4n} &= +1 \\ \sigma_{4n+2} &= -1 \end{aligned}$$

The paperfolding system consists of the orbitclosure of this paperfolding sequence. All sequences in the paperfolding system are paperfolding sequences.

With studying the overlap distribution of paperfolding sequences it will turn out to be helpful to notice that the paperfolding sequence can also be constructed in the following way. This idea is described in [3].

Take an arbitrary sequence $s_n \in \Omega = \{-1,+1\}^{\mathbb{Z}}$ with $n \in \mathbb{Z}$.

Step 1. Occupy one of the two sublattices of $\mathbb{Z}$ of period 2, so the even or odd integers. Place alternately s_1 and $\overline{s_1}$ on the sites of the sublattice. With this we use that $\overline{-1} = +1$ and $\overline{+1} = -1$.

$$\begin{array}{rccccccccccccc} \text{even integers :} & s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} & \cdot \\ \text{odd integers :} & \cdot & s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} \end{array}$$

Step 2. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period $4(=2^2)$. We can also speak of these two sublattices as the even and odd holes. Place alternately s_2 and $\overline{s_2}$ on the sites of the sublattice.

$$
\begin{array}{cccccccccccccccccc}
s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} & \cdot & \Rightarrow & s_1 & s_2 & \overline{s_1} & \cdot & s_1 & \overline{s_2} & \overline{s_1} & \cdot \\
 & & & & & & & & \Rightarrow & s_1 & \cdot & \overline{s_1} & s_2 & s_1 & \cdot & \overline{s_1} & \overline{s_2}
\end{array}
$$

$$
\begin{array}{cccccccccccccccccc}
\cdot & s_1 & \cdot & \overline{s_1} & \cdot & s_1 & \cdot & \overline{s_1} & \Rightarrow & s_2 & s_1 & \cdot & \overline{s_1} & \overline{s_2} & s_1 & \cdot & \overline{s_1} \\
 & & & & & & & & \Rightarrow & \cdot & s_1 & s_2 & \overline{s_1} & \cdot & s_1 & \overline{s_2} & \overline{s_1}
\end{array}
$$

Step 3. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period $8(=2^3)$. Place alternately s_3 and $\overline{s_3}$ on the sites of the sublattice. This gives rise to eight possible sequences.

Step n. Occupy one of the two empty sublattices of $\mathbb{Z}$ of period 2^n. Place alternately s_n and $\overline{s_n}$ on the sites of the sublattice. There are now 2^n possible sequences.

The sequence that arises when $n \to \infty$ is a paperfolding sequence.

This means that a paperfolding sequence is a non-periodic sequence which is a limitpoint of periodic sequences.

Notice that when constructing a sequence in step i, $i \in \mathbb{N}_{\geq 1}$, there are four options because there is a choice between two sublattices and s_i can be -1 and $+1$.

Chapter 5

The overlap distribution of paperfolding sequences

To determine the overlap distribution of paperfolding sequences, we can use a strategy similar to the one used for Toeplitz sequences. Similar to what is done in the proof of theorem 3.2 for determining the overlap distribution of Toeplitz sequences, we will choose an arbitrary sequence and then determine the overlap between this chosen sequence and each of the other sequences. The probability of a certain overlap can be found by determining the probability of sequences that give rise to that overlap. This is also an appropriate way to find the overlap distribution of paperfolding sequences because the probability measure is translation-invariant.

The sequence we choose to use to determine the overlap distribution is given in figure 5.1.

$$+ + - + + - - + + + - - + - - + + + - + + - - - + + - - + - - +$$

Figure 5.1: The sequence used to determine the overlap distribution, $(*)$.

This sequence arises from choosing the sublattice consisting of the even integers in the first step of constructing paperfolding sequences (see chapter 4) and in the following steps the sublattices consisting of the even holes. To fill in these sublattices we have chosen $s_n = +1 \ \forall n \in \mathbb{Z}$. This is possible because $(+1)^{\mathbb{Z}} \in \Omega$. We will refer to the sequence in figure 5.1 by $(*)$.

Subsequently, we will derive the overlap distribution. This will be done by using the steps of constructing a paperfolding sequence, because in this way all sequences are considered.

5.1 Step 1.

When constructing a paperfolding sequence, there are four possible options in the first step. In figure 5.2 these four options are shown together with the first step of $(*)$.

(∗)	+	·	−	·	+	·	−	·	+	·	−	·	+	·	−	·
(1)	+	·	−	·	+	·	−	·	+	·	−	·	+	·	−	·
(2)	−	·	+	·	−	·	+	·	−	·	+	·	−	·	+	·
(3)	·	+	·	−	·	+	·	−	·	+	·	−	·	+	·	−
(4)	·	−	·	+	·	−	·	+	·	−	·	+	·	−	·	+

Figure 5.2: Possible choices in the first step.

Each of these four options occur with probability $\frac{1}{4}$:

$$P(1) = P(2) = P(3) = P(4) = \frac{1}{4}. \tag{5.1}$$

We denote the overlap between $(*)$ and sequences for which the first step in the construction of these sequences is equal to (i) by $q_{(*)(i)}$.

Result step 1. There is one result following from the first step:

> **1R1.** The probability of an overlap of 0 is $\frac{1}{2}$: $p(0) = \frac{1}{2}$.

Proof. This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.2.

These three properties are:

> **1R1(1).** $q_{(*)(3)} = 0$
> **1R1(2).** $q_{(*)(4)} = 0$
> **1R1(3).** $p\left(q_{(*)(1)} = 0\right) = 0$ and $p\left(q_{(*)(2)} = 0\right) = 0$

First we will demonstrate that these three properties are true:

1R1(1). Concerning $(*)$, in the first step is chosen for the sublattice consisting of the even integers. The elements of this sublattice have alternately value $+1$ and -1, starting with $+1$ in $n = 0$. Because the sublattice is of period 2 both $+1$ and -1 are of period 4. Concerning (3), in the first step is chosen for the sublattice consisting of the odd integers. So each of the sublattices in the following steps will be in the even integers. These sublattices consist alternately of $+1$ and -1 and are of period 2^i $i \geq 2$. This means that all these sublattices have a period which is a multiple of 4. From this follows that each element of a certain sublattice in the overlap between $(*)$

and (3) will be multiplied by the same value, this can be $+1$ or -1. So the elements belonging to the same sublattice cancel each other in the overlap. For example this is shown in figure 5.3 for the four options for the sublattice of period 4 of (3).

(∗)	+	·	−	·	+	·	−	·	+	·	−	·	+	·	−	·
(3.1)	+	+	·	−	−	+	·	−	+	+	·	−	−	+	·	−
(3.2)	−	+	·	−	+	+	·	−	−	+	·	−	+	+	·	−
(3.3)	·	+	+	−	·	+	−	−	·	+	+	−	·	+	−	−
(3.4)	·	+	−	−	·	+	+	−	·	+	−	−	·	+	+	−

Figure 5.3: Possible choices in the first step in (3).

We can see that the pluses and minuses (red color) in all four options cancel each other in the overlap between $(*)$ and (3).

Considering (3), in the first step is chosen for the sublattice consisting of the odd integers and in $(*)$ each of the sublattices from step two is in the odd integers. Because of the same reason as with the even integers the elements of a certain sublattice of $(*)$ with period ≥ 4 will cancel each other in the overlap.

From these two findings follows that: $q_{(*)(3)} = 0$.

1R1(2). This holds because of the same reasons why $q_{(*)(3)} = 0$.

1R1(3). For determining the overlap between $(*)$ and the sequences (1) and (2) it is necessary to know which sublattices are chosen in the following steps and which value s_i has in step i.

However, the sequences (1) and (2) will not give a contribution to the probability of an overlap of 0:

- The overlap between $(*)$ and (1) is 0 if and only if from the second step in (1) every time is chosen for the same sublattice as in $(*)$ and $s_i(1) = s_i(*) = -1 \quad \forall$ steps $i \in \mathbb{Z}_{\geq 2}$. The probability of this sequence is $\lim_{n\to\infty} \left(\frac{1}{4}\right)^n = 0$. So this sequence gives no contribution to the probability of an overlap of 0.

- The overlap between $(*)$ and (2) is 0 if and only if from the second step in (2) every time is chosen for the same sublattice as in $(*)$ and $s_i(2) = s_i(*) = +1 \quad \forall$ steps $i \in \mathbb{Z}_{\geq 2}$. The probability of this sequence is $\lim_{n\to\infty} \left(\frac{1}{4}\right)^n = 0$. So this sequence gives no contribution to the probability of an overlap of 0.

From $1R1(1)$, $1R1(2)$, $1R1(3)$ and equation 5.1 follows that $p(0) = \frac{1}{2}$. □

5.2 Step 2.

For (1) and (2) in the second step there are for four options each. In figure 5.4 these options are shown together with the first two steps of $(*)$.

(∗)	+	+	−	·	+	−	−	·	+	+	−	·	+	−	−	·
(1.1)	+	+	−	·	+	−	−	·	+	+	−	·	+	−	−	·
(1.2)	+	−	−	·	+	+	−	·	+	−	−	·	+	+	−	·
(1.3)	+	·	−	+	+	·	−	−	+	·	−	+	+	·	−	−
(1.4)	+	·	−	−	+	·	−	+	+	·	−	−	+	·	−	+

(a)

(∗)	+	+	−	·	+	−	−	·	+	+	−	·	+	−	−	·
(2.1)	−	+	+	·	−	−	+	·	−	+	+	·	−	−	+	·
(2.2)	−	−	+	·	−	+	+	·	−	−	+	·	−	+	+	·
(2.3)	−	·	+	+	−	·	+	−	−	·	+	+	−	·	+	−
(2.4)	−	·	+	−	−	·	+	+	−	·	+	−	−	·	+	+

(b)

Figure 5.4: Possible choices in the second step.

We will demonstrate that the probability of each of these options is equal to $\frac{1}{16}$:

In the second step there are both for (1) and (2) four options, so the following equation holds.

$$P(1.j|1) = P(2.j|2) = \frac{1}{4},\ j = 1, \dots, 4$$

It is obvious that

$$P(1.j \bigcap 1) = P(1.j) \text{ and } P(2.j \bigcap 2) = P(2.j).$$

Then follows using

$$P(1.j|1) = \frac{P(1.j \bigcap 1)}{P(1)} = \frac{P(1.j)}{P(1)} \text{ and } P(2.j|2) = \frac{P(2.j \bigcap 2)}{P(2)} = \frac{P(2.j)}{P(2)}$$

that

$$P(1.j) = P(2.j) = \frac{1}{4} \times \frac{1}{4} = \frac{1}{16} \tag{5.2}$$

Results step 2. There are two results following from the second step:

2R1. $p(\frac{1}{2}) = \frac{1}{8}$

2R2. $p(-\frac{1}{2}) = \frac{1}{8}$

Proof. (**2R1.**) This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.4a.

These three properties are:

2R1(1). $q_{(*)(1.3)} = \frac{1}{2}$

2R1(2). $q_{(*)(1.4)} = \frac{1}{2}$

2R1(3). $p\left(q_{(*)(1.1)} = \frac{1}{2}\right) = 0$ and $p\left(q_{(*)(1.2)} = \frac{1}{2}\right) = 0$

We will demonstrate that these three properties are true:

From the second step the proof of $2R1(1)$ is similar to the proof of $1R1(1)$. This implies that only the first sublattice of $(*)$ and (1.3) has a contribution to the overlap. This sublattice consists of the elements $\sigma_i(*)$ and $\sigma_i(1.3)$ for which i even. Moreover $\sigma_i(1.3) = \sigma_i(*)$ if i even.

$$\begin{aligned} q_{(*)(1.3)} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(1.3) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N/2} \sigma_{2i}(*)\sigma_{2i}(1.3) \\ &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N/2} 1 = \lim_{N\to\infty} \frac{1}{N}\frac{N}{2} = \frac{1}{2} \end{aligned}$$

$2R1(2)$ holds because of the same reasons why $q_{(*)(1.3)} = \frac{1}{2}$. For $2R1(3)$ see $1R1(3)$.

From $2R1(1)$, $2R1(2)$, $2R1(3)$ and equation 5.2 follows that $p(\frac{1}{2}) = \frac{1}{8}$. Thus we have proved the first result. □

Proof. (**2R2**) This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.4b.

These three properties are:

2R2(1). $q_{(*)(2.3)} = -\frac{1}{2}$

2R2(2). $q_{(*)(2.4)} = -\frac{1}{2}$

2R2(3). $p\left(q_{(*)(2.1)} = -\frac{1}{2}\right) = 0$ and $p\left(q_{(*)(2.2)} - \frac{1}{2}\right) = 0$

We will demonstrate that these three properties are true:

By showing that $2R2(1)$ holds we can use that in accordance with the overlap between $(*)$ and (1.3) only the first sublattice of $(*)$ and (2.3) has a contribution to the overlap. For the elements of this sublattice holds: $\sigma_i(2.3) = -\sigma_i(*)$ if i even.

$$q_{(*)(2.3)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(2.3) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N/2} \sigma_{2i}(*)\sigma_{2i}(2.3)$$
$$= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N/2}(-1) = \lim_{N\to\infty} \frac{1}{N}\frac{N}{2} \times (-1) = -\frac{1}{2}$$

$2R2(2)$ holds because of the same reasons why $q_{(*)(2.3)} = -\frac{1}{2}$. For $2R1(3)$ see $1R1(3)$.

From $2R2(1)$, $2R2(2)$, $2R2(3)$ and equation 5.2 follows that $p(-\frac{1}{2}) = \frac{1}{8}$. Thus we have proved the second result. □

5.3 Step 3.

For (1.1), (1.2), (2.1) and (2.2) in the third step there are four options each. In figure 5.5 these options are shown together with the first three steps of (∗).

$$P(1.1.j) = P(1.2.j) = P(2.1.j) = P(2.2.j) = \left(\frac{1}{4}\right)^3 = \frac{1}{64} \qquad j = 1, \ldots, 4 \tag{5.3}$$

Results step 3. There are four results following from the third step:

- **3R1.** $p(\frac{1}{2} + \frac{1}{4}) = p(\frac{3}{4}) = \frac{1}{32}$
- **3R2.** $p(\frac{1}{2} - \frac{1}{4}) = p(\frac{1}{4}) = \frac{1}{32}$
- **3R3.** $p(-\frac{1}{2} + \frac{1}{4}) = p(-\frac{1}{4}) = \frac{1}{32}$
- **3R4.** $p(-\frac{1}{2} - \frac{1}{4}) = p(-\frac{3}{4}) = \frac{1}{32}$

Proof. (**3R1.**) This will be proven by using three properties of the overlap between (∗) and the sequences shown in figure 5.5a.

These three properties are:

- **3R1(1).** $q_{(*)(1.1.3)} = \frac{1}{2} + \frac{1}{4}$
- **3R1(2).** $q_{(*)(1.1.4)} = \frac{1}{2} + \frac{1}{4}$
- **3R1(3).** $p\left(q_{(*)(1.1.1)} = \frac{1}{2} + \frac{1}{4}\right) = 0$ and $p\left(q_{(*)(1.1.2)} = \frac{1}{2} + \frac{1}{4}\right) = 0$

(∗)	+	+	−	+	+	−	−	·	+	+	−	−	+	−	−	·
(1.1.1)	+	+	−	+	+	−	−	·	+	+	−	−	+	−	−	·
(1.1.2)	+	+	−	−	+	−	−	·	+	+	−	+	+	−	−	·
(1.1.3)	+	+	−	·	+	−	−	+	+	+	−	·	+	−	−	−
(1.1.4)	+	+	−	·	+	−	−	−	+	+	−	·	+	−	−	+

(a)

(∗)	+	+	−	+	+	−	−	·	+	+	−	−	+	−	−	·
(1.2.1)	+	−	−	+	+	+	−	·	+	−	−	−	+	+	−	·
(1.2.1)	+	−	−	−	+	+	−	·	+	−	−	+	+	+	−	·
(1.2.3)	+	−	−	·	+	+	−	+	+	−	−	·	+	+	−	−
(1.2.4)	+	−	−	·	+	+	−	−	+	−	−	·	+	+	−	+

(b)

(∗)	+	+	−	+	+	−	−	·	+	+	−	−	+	−	−	·
(2.1.1)	−	+	+	+	−	−	+	·	−	+	+	−	−	−	+	·
(2.1.2)	−	+	+	−	−	−	+	·	−	+	+	+	−	−	+	·
(2.1.3)	−	+	+	·	−	−	+	+	−	+	+	·	−	−	+	−
(2.1.4)	−	+	+	·	−	−	+	−	−	+	+	·	−	−	+	+

(c)

(∗)	+	+	−	+	+	−	−	·	+	+	−	−	+	−	−	·
(2.2.1)	−	−	+	+	−	+	+	·	−	−	+	−	−	+	+	·
(2.2.2)	−	−	+	−	−	+	+	·	−	−	+	+	−	+	+	·
(2.2.3)	−	−	+	·	−	+	+	+	−	−	+	·	−	+	+	−
(2.2.4)	−	−	+	·	−	+	+	−	−	−	+	·	−	+	+	+

(d)

Figure 5.5: Possible choices in the third step.

We will demonstrate that these three properties are true:

From the third step the proof of $3R1(1)$ is similar to the proof of $1R1(1)$. This means that only the first and second sublattice, the even integers and even holes, have a contribution to the overlap. For the elements of these sublattices holds:

- $\sigma_j(1.1.3) = \sigma_j(*)$ if $j \in 2\mathbb{Z}$
- $\sigma_j(1.1.3) = \sigma_j(*)$ if $j \in 1 + 4\mathbb{Z}$

$$
\begin{aligned}
q_{(*)(1.1.3)} &= \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(1.1.3) \\
&= \lim_{N\to\infty} \frac{1}{N} \left(\sum_{i=1}^{N/2} \sigma_{2i}(*)\sigma_{2i}(1.1.3) + \sum_{i=1}^{N/4} \sigma_{1+4(i-1)}(*)\sigma_{1+4(i-1)}(1.1.3) \right) \\
&= \lim_{N\to\infty} \frac{1}{N} \left(\sum_{i=1}^{N/2} 1 + \sum_{i=1}^{N/4} 1 \right) = \lim_{N\to\infty} \frac{1}{N} \left(\frac{N}{2} + \frac{N}{4} \right) = \frac{1}{2} + \frac{1}{4}
\end{aligned}
$$

$3R1(2)$ holds because of the same reasons why $q_{(*)(1.1.3)} = \frac{1}{2} + \frac{1}{4}$. For $3R1(3)$ see $1R1(3)$.

From $3R1(1)$, $3R1(2)$ and $3R1(3)$ and equation 5.3 follows that $p(\frac{1}{2}+\frac{1}{4}) = \frac{1}{32}$. Thus we have proved the first result. □

Proof. (**3R2.**) This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.5b.

These three properties are:

3R2(1). $q_{(*)(1.2.3)} = \frac{1}{2} - \frac{1}{4}$
3R2(2). $q_{(*)(1.2.4)} = \frac{1}{2} - \frac{1}{4}$
3R2(3). $p\left(q_{(*)(1.2.1)} = \frac{1}{2} - \frac{1}{4}\right) = 0$ and $p\left(q_{(*)(1.2.2)} = \frac{1}{2} - \frac{1}{4}\right) = 0$

We will demonstrate that these three properties are true:

By showing that $3R2(1)$ holds we can use that in accordance with the overlap between $(*)$ and $(1.1.3)$ only the the first and second sublattice, the even integers and even holes, have a contribution to the overlap. For the elements of these sublattices holds:

- $\sigma_j(1.2.3) = \sigma_j(*)$ if $j \in 2\mathbb{Z}$
- $\sigma_j(1.2.3) = -\sigma_j(*)$ if $j \in 1 + 4\mathbb{Z}$

$$q_{(*)(1.2.3)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(1.2.3) = \lim_{N\to\infty} \frac{1}{N}\left(\frac{N}{2} - \frac{N}{4}\right) = \frac{1}{2} - \frac{1}{4}$$

$3R2(2)$ holds because of the same reasons why $q_{(*)(1.2.3)} = \frac{1}{2} - \frac{1}{4}$. For $3R2(3)$ see $1R1(3)$.

From $3R2(1)$, $3R2(2)$, $3R2(3)$ and equation 5.3 follows that $p(\frac{1}{2} - \frac{1}{4}) = \frac{1}{32}$. Thus we have proved the second result. □

Proof. (**3R3.**) This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.5c.

These three properties are:

3R3(1). $q_{(*)(2.1.3)} = -\frac{1}{2} + \frac{1}{4}$
3R3(2). $q_{(*)(2.1.4)} = -\frac{1}{2} + \frac{1}{4}$
3R3(3). $p\left(q_{(*)(2.1.1)} = -\frac{1}{2} + \frac{1}{4}\right) = 0$ and $p\left(q_{(*)(2.1.2)} = -\frac{1}{2} + \frac{1}{4}\right) = 0$

We will demonstrate that these three properties are true:

By showing that $3R3(1)$ holds we can use that in accordance with the overlap between $(*)$ and $(1.1.3)$ only the the first and second sublattice, the even integers and even holes, have a contribution to the overlap. For the elements of these sublattices holds:

- $\sigma_j(2.1.3) = -\sigma_j(*)$ if $j \in 2\mathbb{Z}$
- $\sigma_j(2.1.3) = \sigma_j(*)$ if $j \in 1 + 4\mathbb{Z}$

$$q_{(*)(2.1.3)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(2.1.3) = \lim_{N\to\infty} \frac{1}{N}\left(-\frac{N}{2} + \frac{N}{4}\right) = -\frac{1}{2} + \frac{1}{4}$$

$3R3(2)$ holds because of the same reasons why $q_{(*)(2.1.3)} = -\frac{1}{2} + \frac{1}{4}$. For $3R3(3)$ see $1R1(3)$.

From $3R3(1)$, $3R3(2)$, $3R3(3)$ and equation 5.3 follows that $p(-\frac{1}{2} + \frac{1}{4}) = \frac{1}{32}$. Thus we have proved the third result. □

Proof. (**3R4.**) This will be proven by using three properties of the overlap between $(*)$ and the sequences shown in figure 5.5d.

These three properties are:

3R4(1). $q_{(*)(2.2.3)} = -\frac{1}{2} - \frac{1}{4}$
3R4(2). $q_{(*)(2.2.4)} = -\frac{1}{2} - \frac{1}{4}$
3R4(3). $p\left(q_{(*)(2.2.1)} = -\frac{1}{2} - \frac{1}{4}\right) = 0$ and $p\left(q_{(*)(2.2.2)} = -\frac{1}{2} - \frac{1}{4}\right) = 0$

We will demonstrate that these three properties are true:

By showing that $3R4(1)$ holds we can use that in accordance with the overlap between $(*)$ and (2.2.3) only the first and second sublattice, the even integers and even holes, have a contribution to the overlap. For the elements of these sublattices holds:

- $\sigma_j(2.2.3) = -\sigma_j(*)$ if $j \in 2\mathbb{Z}$
- $\sigma_j(2.2.3) = -\sigma_j(*)$ if $j \in 1 + 4\mathbb{Z}$

$$q_{(*)(2.2.3)} = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \sigma_i(*)\sigma_i(2.2.3) = \lim_{N\to\infty} \frac{1}{N} \left(-\frac{N}{2} - \frac{N}{4} \right) = -\frac{1}{2} - \frac{1}{4}$$

$3R4(2)$ holds because of the same reasons why $q_{(*)(2.2.3)} = -\frac{1}{2} - \frac{1}{4}$. For $3R4(3)$ see $1R1(3)$.

From $3R4(1)$, $3R4(2)$, $3R4(3)$ and equation 5.3 follows that $p(-\frac{1}{2} - \frac{1}{4}) = \frac{1}{32}$. Thus we have proved the fourth result. $\square$

5.4 Results

A summary of the results of the previous sections of chapter 4 and two added results is given in figure 5.6.

$$p(q) = \tfrac{1}{2} \quad \Leftrightarrow \quad q = 0$$

$$p(q) = \tfrac{1}{8} \quad \Leftrightarrow \quad q = \pm\tfrac{1}{2}$$

$$p(q) = \tfrac{1}{32} \quad \Leftrightarrow \quad q = \pm\tfrac{3}{4} \quad \vee \quad q = \pm\tfrac{1}{4}$$

$$p(q) = \tfrac{1}{128} \quad \Leftrightarrow \quad q = \pm\tfrac{7}{8} \quad \vee \quad q = \pm\tfrac{5}{8} \quad \vee \quad q = \pm\tfrac{3}{8} \quad \vee \quad q = \pm\tfrac{1}{8}$$

$$p(q) = \tfrac{1}{512} \quad \Leftrightarrow \quad q = \pm\tfrac{15}{16} \quad \vee \quad q = \pm\tfrac{13}{16} \quad \vee \quad q = \pm\tfrac{11}{16} \quad \vee \quad q = \pm\tfrac{9}{16} \quad \vee$$

$$q = \pm\tfrac{7}{16} \quad \vee \quad q = \pm\tfrac{5}{16} \quad \vee \quad q = \pm\tfrac{3}{16} \quad \vee \quad q = \pm\tfrac{1}{16}$$

Figure 5.6: Results step 1-5.

We can see that there is a pattern in which overlaps have the same probability. When all paperfolding sequences are arranged like this, the overlap distribution of paperfolding sequences arrises. This overlap distribution is given in theorem 5.1.

Theorem 5.1. *The overlap distribution of paperfolding sequences is given by:*

$$p(q) = \sum_{n=0}^{\infty} \frac{1}{2} \left(\frac{1}{4}\right)^n \delta\left(q \pm \frac{m}{2^n}\right) \tag{5.4}$$

with $n, m \in \mathbb{N}$ and m must satisfy the following two conditions:

- *if $n = 0$ then $m = 0$*
- *if $n > 0$ then m is odd and $m < 2^n$*

Notice that when $X = \{q : p(q) \neq 0\}$ then $\overline{X} = [-1, +1]$. This means that X is dense in $[-1, +1]$. The elements of X are called dyadic rationals.

Chapter 6

Conclusion

In this report we have discussed some non-periodic sequences in the sample space $\Omega = \{-1, +1\}^{\mathbb{Z}}$. The sequences we considered belong to four different types. We first discussed three known results. These results where about Thue-Morse sequences, Toeplitz sequences and Fibonacci sequences. For each of these three types there exists a manner of constructing all sequences of that type. Based on these manners the overlap distributions of these types of sequences could be determined.

We can define one particular paperfolding sequence. The paperfolding system consists of the orbit closure of this paperfolding sequence. We also gave a general construction of all paperfolding sequences. To determine the overlap distribution of paperfolding sequences it turned out that it was possible to use a strategy similar to the one used for Toeplitz sequences. This means that it is possible to determine the overlap distribution of paperfolding sequences by chosing an arbitrary sequence and then determine the overlap between this chosen sequence and each of the other sequences. To derive the overlap distribution, the steps to construct a paperfolding sequence were used.

The overlap distribution of paperfolding sequences is given by:

$$p(q) = \sum_{n=0}^{\infty} \frac{1}{2} \left(\frac{1}{4}\right)^n \delta\left(q \pm \frac{m}{2^n}\right)$$

with $n, m \in \mathbb{N}$ and m must satisfy the following two conditions:

- if $n = 0$ then $m = 0$
- if $n > 0$ then m is odd and $m < 2^n$

Bibliography

[1] A.C.D. van Enter, A. Hof and Jacek Miękisz, *Overlap distributions for deterministic systems with many pure states*, J. Phys. A: Math. Gen. **25**, 1992, L1133-L1137.

[2] T.J. Bedford, M.S. Keane and C. Series, *Ergodic theory, symbolic dynamics and hyperbolic spaces*, Oxford University Press, Oxford, 1991.

[3] J. Allouche, *The number of factors in a paperfolding sequence*, Bull. Austral. Math. Soc. **46**, 1992, 23-32.

[4] M. Mendès France and A.J. van der Poorten, *Arithmetic and analytic properties of paper folding sequences*, Bull. Austral. Math. Soc. **24**, 1981, 123-131.

[5] F.M. Dekking, *Substitutions*, Mathematisch Instituut, Katholieke Universiteit van Nijmegen, 1980.

[6] C. Grillenberger, *Constructions of Strictly Ergodic Systems*, Z. Wahrs. verw. Geb.**25**, 1973, 323-334.

[7] P. Walters, *An Introduction to Ergodic Theory*, Springer-Verlag Berlin Heidelberg, 1992

[8] M. Queffélec, *Substitution Dynamical Systems - Spectral Analysis (Lecture Notes in Mathematics 1294)*, Springer-Verlag Berlin Heidelberg, Germany, 1987.

[9] Marc Mezard, Giorgio Parisi, Miguel Angel Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore, 1987.

[10] K. Lü and J. Wang, *Construction of Sturmian sequences*, J. Phys. A: Math. Gen. **38**, 2005, 2891-2897.

[11] M. Baake, R.V. Moody, C. Richard and B. Sing, *Which distributions of matter diffract? - Some answers*, arXiv:math-ph/031019v1 15 Jan 2003.

[12] Manfred Schroeder, *Fractals, Chaos, Power Laws*, W.H. Freeman and Company, New York, 1991.