Florian Greinecker, MSc.

Combinatorial and Number Theoretic Properties of Certain Automatic Sequences

DOCTORAL THESIS

to achieve the university degree of

Doktor der Naturwissenschaften

submitted to

Graz University of Technology

Supervisor

Univ.-Prof. Dipl.-Ing. Dr.techn. Peter Grabner

Department of Analysis and Computational Number Theory (Math A)

Graz, November 2015

AFFIDAVIT

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline is identical to the present doctoral thesis.

Date

Signature

Contents

Chapter 1

Introduction

In this work I study several combinatorial sequences and their connections to automata. It consists of two papers and supplementary material to each of them. The two papers are similar to their original form, hence there are minor overlaps with the remarks.

The first chapter gives an introduction to automatic and regular sequences to lay a basis for the second chapter. I elaborate on the connection of automatic sequences and deterministic finite automata with output. There are three common, equivalent definitions of "automatic sequence": via the kernel, via automata, and via uniform morphisms.

Proofs of the equivalence of these definitions and some other well known facts about automatic sequences will be provided. In order to clarify the definitions, there will be examples for each type of definition. Finally k-regular sequences are defined and three examples are given.

The second chapter follows the paper "On the 2-Abelian Complexity of the Thue–Morse Word"[35], about the integer sequence $(\mathcal{P}_n)_{n\geq 0}$. This sequence is the 2-abelian complexity sequence of the well known infinite Thue–Morse word $\mathbf{t} := 011010011001\ldots$ The 2-abelian complexity is a complexity measure for infinite words with a resolution between the abelian complexity and the factor complexity.

As the main result Theorem 3.1.1, I show that the Thue–Morse sequence is 2-regular. This solves an open conjecture from Elise Vandomme, Aline Parreau and Michel Rigo [47].

Section 3.3 examines subwords of the infinite Thue–Morse word $\mathbf{t}$. Some of these subwords have unique extensions. I give bounds for the lengths of unique extensions in Theorem 3.3.5 and provide an algorithm to calculate the unique extension of a given word w.

In Section 3.7 I show that $(\mathcal{P}_n)_{n\geq 0}$ is a concatenation of palindromes of increasing size and provide an alternative proof that the sequence is unbounded.

The fourth chapter is about the submitted paper "Spatial equidistribution of

combinatorial number schemes" [34]. Here I investigate some multidimensional combinatorial sequences.

These sequences can be used to define number schemes. A number scheme is a generalized matrix digital system together with coloring functions that assigns residue classes modulo a prime p to numbers and digits. A number scheme satisfies the generalized Lucas' congruence if the residue class of an entry at position n modulo p is the product of the residue classes of its digits.

The number schemes allow us to define an iterated function system and a sequence U_i of sets which converge to a limit fractal. The colorings are defined on the sequence U_i. It is not possible to define a coloring for the limit fractal.

The two main results of Chapter 4 are the equidistribution results of Theorem 4.2.4 and Theorem 4.2.5. For number schemes that satisfy the generalized Lucas' congruence I give sufficient conditions that for $\lim_{i\to\infty}$

- allmost all digits in U_i have color zero, and
- the nonzero colors are equidistributed in U_i modulo a prime p.

The Stirling numbers of the first and second kind (the original impetus of the paper "Spatial equidistribution of combinatorial number schemes") as well as the binomial coefficients, the Apéry numbers, the Gaussian q-nomial coefficients and the multinomial coefficients are number schemes that satisfy the generalized Lucas' congruence. The proof of Theorem 4.2.5 combines ideas from fractal measure theory with analytic number theory.

The last chapter is an addendum to chapter 4. I prove two additional results which would distract from the main line of chapter 4.

Theorem 5.1.2 estimates the speed of convergence towards equidistribution for binomial coefficients. It also applies to Stirling numbers of the first and second kind and the Gaussian q-nomial coefficients since they are affine transformations of the binomial coefficients. I use the Weil bound to prove this nontrivial estimate.

It would be possible to formulate the proofs in chapter 4 in the language of automata. As an example I show in Theorem 5.2.2 an equidistribution result modulo p for the p-free binomial coefficients. With the use of an automaton I can show an equidistribution result for this sequence which does not satisfy the conditions in chapter 4.

In the rest of this thesis I will use "we" to refer to the reader and the author.

Chapter 2

On automatic and regular sequences

In this chapter we will define k-automatic and k-regular sequences needed in the next chapter about the paper "On the 2-Abelian Complexity of the Thue-Morse Word"[35].

There will be three equivalent definitions of an "k-automatic sequence." The first definition uses deterministic finite automata with output, the second definition uses k-uniform morphisms, and the third definition uses the k-kernel. We will demonstrate the equivalence of the definitions and give some examples. Then we will define k-regular sequences and also provide some examples. For this chapter we mostly follow the seminal book "Automatic sequences" [2].

Let us recall that an *alphabet* Σ is a finite set of symbols. We write Σ_k for the alphabet $\{0, 1, \ldots, k-1\}$. We use Σ^* to denote the set of finite words over Σ. While 1-automatic sequences can be defined, we will always assume that k is an integer $k \geq 2$.

2.1 Automatic sequences

If we speak of an automaton in the context of automatic sequences we mean a *deterministic finite automaton with output*, short DFAO. This automaton maps stings over an input alphabet to letters of an output alphabet.

To be more precise it reads an input word w from left to right. Starting from an initial state, it moves between a finite number of states according to the transition function δ while reading the letters of w. After reading all symbols in w, the automaton is in state q. It then returns the letter $\tau(q)$ of the output alphabet which is assigned to the state q.

Formally a DFAO is a 6-tuple

$$M = (Q, \Sigma, \delta, q_0, \Delta, \tau).$$

Here

- Q is a finite set of states,
- Σ is the finite input alphabet,
- $\delta : Q \times \Sigma \to Q$ is the transition function,
- $q_0 \in Q$ is the initial state,
- Δ is the output alphabet
- and $\tau : Q \to \Delta$ is the output function.

For the empty word ϵ and the state $q \in Q$ we define $\delta(q, \epsilon) = q$. For a word $u \in \Sigma^*$ and $a \in \Sigma$ we define $\delta(q, ua) = \delta(\delta(q, u), a)$. This extends the domain of δ to $Q \times \Sigma^*$. So the automaton M defines a function $f_M(w) := \tau(\delta(q_0, w))$ from Σ^* to Δ. A DFAO with the input alphabet $\Sigma_k = \{0, \ldots, k-1\}$ is called a k-DFAO.

A convenient way to represent a DFAO is the *transition diagram*, which is a directed, edge-labeled multigraph. The vertex set of the graph represents the states $q \in Q$. Each state is labeled with the state name $q \in Q$ and the output symbol $\tau(q) \in \Delta$. There is a labeled arrow $p \xrightarrow{a} q$ if and only if $\delta(p, a) = q$ for $a \in \Sigma$ and $p, q \in Q$. An unlabeled arrow denotes q_0. To compute $\delta(q_0, a_0 a_1 \ldots a_i)$ you start at q_0 and follow the edge $q_0 \xrightarrow{a_0} q_1$, then $q_1 \xrightarrow{a_1} q_2$ until you reach the final state q_{i+1} with the output $\tau(q_{i+1})$.

Now we can define k-automatic sequences. A sequence $\mathbf{u}$ is k-automatic if there is a k-DFAO that reads the base-k representation of i (starting with the most significant digit) and has u_i as output.

Definition 2.1.1. *Let $k \geq 2$ be an integer. We can write any nonnegative integer n in the form*

$$n = \sum_{0 \leq i \leq s} c_i k^i,$$

with $c_i \in \Sigma_k$. We define the word $[n]_k := c_s \cdots c_0$ and call it a base-k representation *of n. If $c_s \neq 0$, we write $(n)_k$ instead of $[n]_k$ and call it the* canonical base-k representation.

With the use of an greedy algorithm it is easy to prove that the canonical base-k representation is unique. The representation $[n]_k$ is not unique since there may be some leading zeros.

Definition 2.1.2. *A sequence* $\mathbf{u} = (u_i)_{i\geq 0}$ *is* k-automatic *if there is a* k*-DFAO* M *that outputs* u_n *given a base* k*-representation* $[n]_k$ *of* n *as input. In this case we say* M generates $\mathbf{u}$.

Let us look at an example. One way to define the Thue–Morse sequence $\mathbf{t}$ is the following:

$$t_n = \begin{cases} 0 & \text{if the number of 1's in the base 2 expansion on } n \text{ is even,} \\ 1 & \text{if the number of 1's in the base 2 expansion on } n \text{ is odd.} \end{cases}$$

The infinite Thue–Morse word is $\mathbf{t} := t_0 t_1 t_2 \cdots$ This can be implemented as 2-automaton like in Figure 2.1.

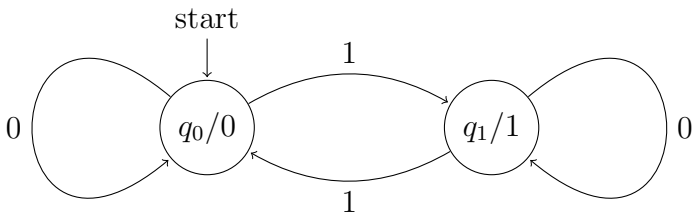

Figure 2.1: The 2-DFAO which generates the Thue–Morse sequence

The characterization of k-automatic sequences is quite robust. There are several variations to the definition of k-automatic sequences which describe the same class of sequences. We will prove two of them which we will need later. The first theorem (cited after [2, pp. 157–158]) tells us that instead of a base k-representation $[n]_k$ of n we can use the canonical base k-representation $(n)_k$.

Lemma 2.1.3. *A sequence* $\mathbf{u} = (u_i)_{i\geq 0}$ *is* k*-automatic if and only if there is a DFAO* M *with* $\delta(q_0, 0) = q_0$ *and* $u_n = \tau(\delta(q_0, (n)_k))$ *for all* $n \geq 0$.

Proof. One direction is trivial. For the other direction we modify a DFAO

$$M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$$

which generates $\mathbf{u}$ to a DFAO

$$M' = (Q \cup \{q_0'\}, \Sigma_k, \delta', q_0', \Delta, \tau')$$

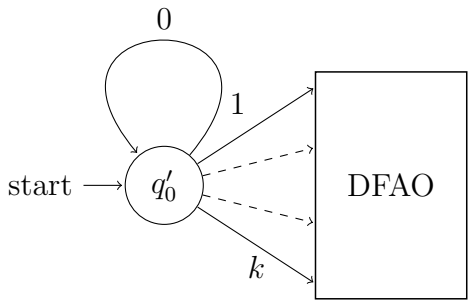

Figure 2.2: Illustration to the proof of Lemma 2.1.3

by adding a state q_0' with $\tau'(q_0') = \tau(q_0)$, $\delta'(q_0', 0) = q_0'$ and $\delta'(q_0', a) = \delta(q_0, a)$ for all $a \neq 0$. Otherwise M and M' are identical. Then

$$\tau'(\delta'(q_0', 0^i(n)_k)) = \tau(\delta(q_0, (n)_k)).$$

□

The next theorem (theorem and proof cited after [2, pp.139,159]). shows that instead of the most significant digit we can also start with the least significant digit. Given a word $w = w_0 w_1 \cdots w_t$ we define $w^R = w_t w_{t-1} \cdots w_0$.

Theorem 2.1.4. *A sequence* $\mathbf{u} = (u_i)_{i \geq 0}$ *is* k-automatic *if there is a DFAO* M *that outputs* u_n *given* $[n]_k^R$ *as input.*

Proof. We show if there is a DFAO M which computes the function $f_M(w)$ there is another DFAO M' which computes the function $f_{M'}(w) = f_M(w^R)$.

Let $M = (Q, \Sigma_k, \delta, q_0, \Delta, \tau)$ be a DFAO which computes $f(w)$. We define a new DFAO $M' = (S, \Sigma_k, \delta', q_0', \Delta, \tau')$ where $S = \Delta^Q$. So S contains all functions which map Q to Δ. We define q_0' as the function $q \mapsto \tau(q)$, $\tau'(g) = g(q_0)$ (for $g \in S$) and $\delta'(g, a) = h$ where $h(q) = g(\delta(q, a))$ for $a \in \Sigma$.

Now we use induction on $|w|$ to prove that $\delta'(q_0', w) = h$ where h is the function $h : \ q \mapsto \tau(\delta(q, w^R))$. The base case with $|w| = 0$ is trivial. Now we make the inductive step from a word $w = x$ with $|w| = n$ to a word $w = xa$ with $|w| = n+1$.

Then

$$\delta'(q_0', xa) = \delta'(\delta'(q_0', x), a) = \delta'(g, a) = h.$$

By induction we have $g : \ q \mapsto \tau(\delta(q, x^R))$. Then

$$h(q) = g(\delta(q, a)) = \tau(\delta(\delta(q, a), x^R)) = \tau(\delta(q, ax^R)) =$$

$$= \tau(\delta(q, (xa)^R)) = \tau(\delta(q, w^R)).$$

Since $\tau'(h) = h(q_0)$, we know that M' computes $f(w^R)$. □

2.2 The paperfolding sequence

Let us introduce the paperfolding sequence to have a second sequence as an example for the definitions. We will give three definitions for the paperfolding sequence and will hint at their equivalence. A rigorous mathematical proof can be found in [11], which provides a nice survey over the paperfolding sequence.

One way to generate the *paperfolding sequence* is to take a rectangular sheet of paper and fold the left edge onto the right edge. Continue the process ad infinitum and then unfold the paper. The sequence $(p_i)_{i\geq 1}$ of *down* $(+1)$ and *up* (-1) bends is the paperfolding sequence $\mathbf{p}$. It begins with

$$(p_i)_{i\geq 1} = +1\ +1\ -1\ +1\ +1\ -1\ -1\ +1 + 1\ +1\ -1 \cdots .$$

If we unfold the bends to 90°, we get the shapes in figure 2.3. As limit object we get the "dragon curve", a fractal. A better aproximation of the dragon curve can be seen in Figure 2.4. The dragon curve has the property that it never crosses itself [22].

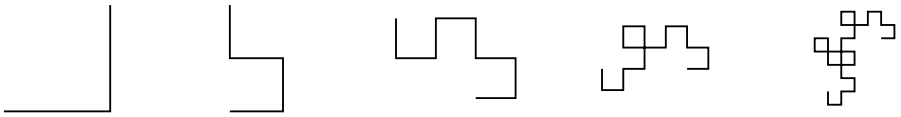

Figure 2.3: The first five steps of the paperfolding sequence

After n folds we have 2^n-1 bends. Let $D_n := p_1p_2\cdots p_{2^n-1}$ denote the sequence after n folds. Now we look at the sheet of paper after $n+1$ folds in two ways.

On the one hand we get

$$D_{n+1} = p_1\ p_2 \cdots p_{2^n-1}\ +1\ -p_{2^n-1} \cdots -p_2\ -p_1 = D_n\ +1\ -D_n^R. \qquad (2.1)$$

Here fold the sheet once (the "1" in the middle) and then n more times (the bends are symmetric to the middle where one is the inverse of the other, i.e. $p_i = -p_{2^n-i}$ for $i \leq 2^{n-1}$). The recursion in equation (2.1) describes the physical paperfolding procedure.

On the other hand we have

$$D_{n+1} = 1\ p_1\ -1\ p_2\ +1 \cdots +1\ p_{2^n-1}\ -1.$$

Here we fold the sheet n times (which constructs the bends in the even positions) and the once more (which gives an alternating sequence of 1 and -1 at the odd positions).

This is the idea behind the *Toeplitz construction*. We start with the infinite word w of period length 4 given by

$$w = +1, 0, -1, 0, +1, 0, -1, 0, \cdots$$

Then we replace all 0's in the sequence consecutively by the terms of w and get

$$w = +1, +1, -1, 0, +1, -1, -1, 0, \cdots$$

We iterate the process ad infinitum and get the paperfolding sequence as the limit object.

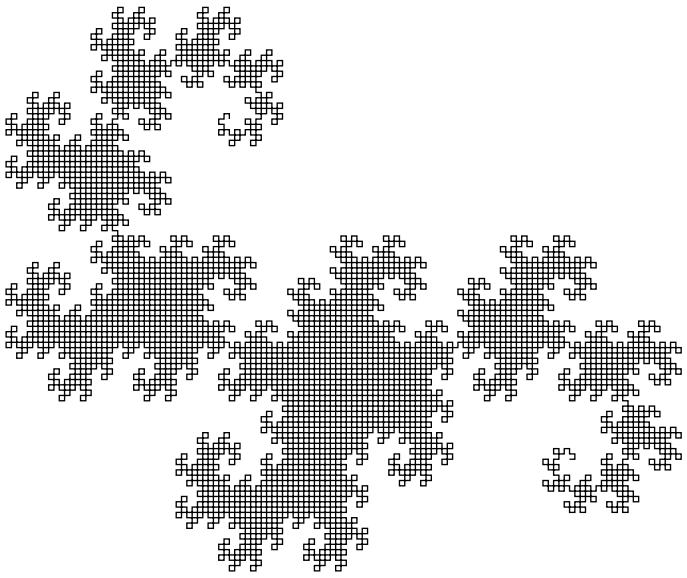

Figure 2.4: Thirteen iterations of the paperfolding sequence

The Toeplitz construction leads to another way to describe the paperfolding sequence $(p_i)_{i\geq 1}$. In the k-th step of the Toeplitz construction we write an alternating sequence of 1 and -1 on the positions of the form $p_{2^k\times u}$ with u odd. So let us write i in the form $i = 2^k \times u$ where u is an odd number. Then

$$p_i := \begin{cases} +1 & \text{if } u \equiv 1 \pmod 4, \\ -1 & \text{if } u \equiv 3 \pmod 4. \end{cases}$$

With this description it is straightforward to construct the 2-DFAO in Figure 2.5. We have to track of the last 1 and the digit before. If the word ends in 010^s, we are in state A, with 01 in B, with 11^s in C and with 110^s in D. Since we have binary input words, the paperfolding sequence is 2-automatic.

2.3 Uniform morphisms

A *morphism* is a map $\varphi : \Sigma^* \to \Delta^*$ such that $\varphi(xy) = \varphi(x)\varphi(y)$ for all $x, y \in \Sigma^*$. We will always let $\Sigma = \Delta$ and therefore have a semigroup. A morphism φ with $|\varphi(a)| = k$ for all $a \in \Sigma$ is called *k-uniform*. A 1-uniform morphism is a *coding*.

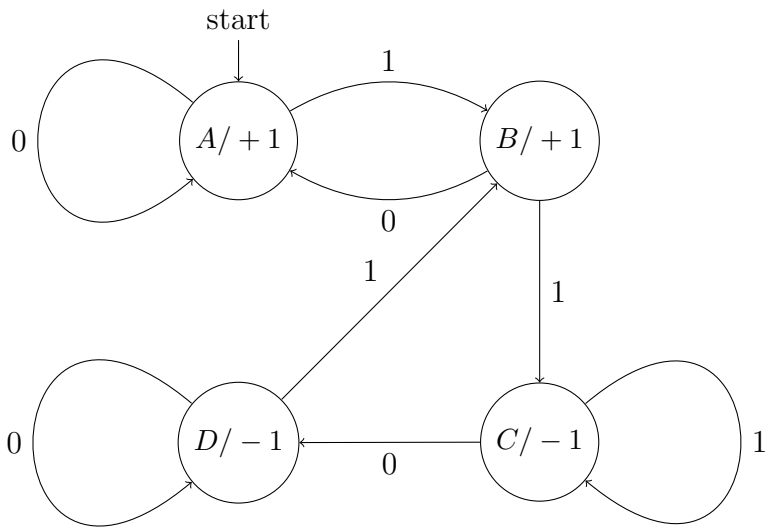

Figure 2.5: The 2-DFAO which generates the paperfolding sequence

The k-uniform morphism φ is *prolongable* on a letter a if $\varphi(a) = ax$ for a word $x \in \Sigma^*$ with $|x| = k - 1$. Then the word

$$\mathbf{v} = \varepsilon^\omega(a) := ax\varphi(x)\varphi^2(x)\varphi^3(x)\cdots$$

is the unique fixed point of the morphism φ starting with a.

There is an intimate connection between k-uniform morphisms and k-automatic sequences which is stated in Cobham's theorem [20] (theorem and proof cited after [2, pp. 175–176]).

Theorem 2.3.1 (Cobham)**.** *A sequence* $\mathbf{u} = (u_n)_{n\geq 0}$ *is* k*-automatic if and only if it is the image, under a coding, of a fixed point of a* k*-uniform morphism.*

Proof. $\Rightarrow$: If $(u_n)_{n\geq 0}$ is a k-automatic sequence, it can be generated by a k-DFAO $(Q, \Sigma, \delta, q_0, \Delta, \tau)$. According to Theorem 2.1.3 we can assume that $\delta(q_0, 0) = q_0$. We define the following morphism φ for each $q \in Q$ by

$$\varphi(q) = \delta(q, 0)\delta(q, 1)\ldots\delta(q, k-1).$$

Since $\delta(q_0, 0) = q_0$, the morphism has a fixed point, say $\mathbf{w} = w_0w_1w_2\ldots$ Now we use induction on the length of y to show that $\delta(q_0, y) = w_{[y]_k}$. In the base case

$|y| = 0$, we have $\delta(q_0, \epsilon) = q_0 = w_0$. Now we write $y = xa$ with $a \in \Sigma_k$. We get

$$\delta(q_0, y) = \delta(q_0, xa) = \delta(\delta(q_0, x), a) = \delta(w_{[x]_k}, a)$$

by the induction hypothesis. Then

$$\delta(w_{[x]_k}, a) = \varphi(w_{[x]_k})_a = w_{k[x]_k+a}$$

since $\varphi(w_0, \ldots, w_t) = \varphi(w_0, \ldots, w_{t-1})\varphi(w_t) = (w_0 \ldots w_{kt-1})(w_{kt} \ldots w_{kt+k-1})$ and so $\varphi(w_t) = (w_{kt} \ldots w_{kt+k-1})$. Finally we get

$$w_{k[x]_k+a} = w_{[xa]_k} = w_{[y]_k}.$$

Now we apply the coding, which is just $u_n = \tau(\delta(q_0, (n)_k)) = \tau(w_n)$.

$\Leftarrow$: For the other direction we reverse the construction and define a k-DFAO with $\delta(q, b) = \varphi(q)_b$. We use induction on n to prove $w_n = \delta(q_0, (n)_k)$ where $\mathbf{w} = w_0 w_1 w_2 \ldots$ is the fixed point of the morphism and $q_0 = w_0$. The base case with $n = 0$ is trivial. Write $(n)_k = n_1 n_2 \ldots n_t$ as $n = kn' + n_t$ with $0 \leq n_t < k$. We get

$$\delta(q_0, (n)_k) = \delta(q_0, kn' + n_t) = \delta(\delta(q_0, (n')_k, n_t) =$$

$$= \delta(w_{n'}, n_t) = \varphi(w_{n'})_{n_t} = w_{kn'+n_t} = w_n.$$

□

For the Thue–Morse sequence the 2-uniform morphism is just $\varphi(0) = 01$ and $\varphi(1) = 10$. Thus, we have

$$\mathbf{t} = \varphi^\omega(0) = 01101001 \cdots$$

With the paperfolding sequence we have to separate the coding and the morphism.

Example 2.3.2. *Using the 2-DFAO from Figure 2.5 we get* $\mathbf{p} = \tau(\varphi^\omega(A))$ *with*

$$\phi(A) = AB, \quad \tau(A) = +1,$$

$$\phi(B) = AC, \quad \tau(B) = +1,$$

$$\phi(C) = DC, \quad \tau(C) = -1,$$

$$\phi(D) = DB, \quad \tau(D) = -1$$

as morphism and coding for the paperfolding word.

2.4 The k-kernel

For an infinite sequence $\mathbf{u} = (u_n)_{n\geq 0}$ we define the *k-kernel* as the set of subsequences

$$K_k(\mathbf{u}) := \{(u_{k^i\times n+j})_{n\geq 0} \mid i \geq 0 \text{ and } 0 \leq j < k^i\}.$$

The kernel gives us another way to characterize k-automatic sequences. The relation between k-automatic sequences and the k-kernel is a famous theorem by Eilenberg [27] (theorem and proof cited after [2, pp. 185–186]).

Theorem 2.4.1 (Eilenberg)**.** *A sequence $(u_n)_{n\geq 0}$ is k-automatic sequence if and only if the k-kernel is finite.*

Proof. $\Rightarrow$: The base k-representation of $k^i \times n + j$ always ends with the same suffix of length i. Now we take a reverse DFAO $M = (Q, \Sigma, \delta, q_0, \Delta, \tau)$ which reads $[n]_k^R$ and outputs $\mathbf{u}$. Such a DFAO exists according to Theorem 2.1.4. The reversed base k-representation of $k^i \times n + j$ always starts with the same prefix of length i. For fixed i and j we will always reach the same state q after reading i letters of the reversed base k-representation of $k^i \times n + j$. The sequence $u_{k^i\times n+j}$ is generated by the DFAO $(Q, \Sigma, \delta, q_0, \Delta, \tau)$. Since we have a finite number of choices for q, there is a finite number of sequences in the kernel.

$\Leftarrow$: Let us define a equivalence relation on Σ_k^*. For two words $x, w \in \Sigma_k^*$ we have

$$x \equiv w \iff u_{k^{|w|}\times n+[w]_k} = u_{k^{|x|}\times n+[x]_k}.$$

We write $[x]$ to denote the equivalence class of x.

The following k-DFAO generates $(u_n)_{n\geq 0}$:

$$\begin{aligned} Q &= \{[x] : x \in \Sigma_k^*\}, \\ \delta([x], a) &= [ax], \\ \tau([x]) &= u_{[x]_k}, \\ q_0 &= [\varepsilon]. \end{aligned}$$

Now we have to show that this DFAO is independent of the class representative, i.e. if $[x] = [w]$ then $\delta([x], a) = \delta([w], a)$ and $\tau([x]) = \tau([w])$.

If we set $n = km + a$ in

$$u_{k^{|w|}\times n+[w]_k} = u_{k^{|x|}\times n+[x]_k}, \tag{2.2}$$

we get

$$u_{k^{|aw|}\times m+[aw]_k} = u_{k^{|ax|}\times m+[ax]_k}$$

and therefore

$$[ax] = [aw]\,.$$

To show that $[x] = [w] \Rightarrow u_{[x]_k} = u_{[w]_k}$ we just set $n = 0$ in equation (2.2). Since $\delta([x]\,, a) = [ax]$, an induction shows that $\delta(q_0, w^R) = [w]$. Hence

$$\tau(\delta(q_0, w^R)) = \tau([w]) = u_{[w]_k}.$$

□

An easy way to determine the k-kernel of a sequence $\mathbf{u}$ is to start a list with the sequence $\mathbf{u}$ and then split it into subsequences of the form $(u_{k\times n+j})_{n\geq 0}$ with $0 \leq j < k$. If this new subsequences are not in the list, we split them again until every sequence we could get by splitting is already in the list.

In the case of the Thue–Morse sequence $\mathbf{t} = 01101001\cdots$ the 2-kernel consists of the original sequence $\mathbf{t}$ and its complement sequence $\overline{\mathbf{t}} = 10010110\cdots$

Example 2.4.2. *Take the 2-kernel of the paperfolding sequence*

$$\mathbf{p} = +1 + 1 - 1 + 1 + 1 - 1 - 1 + 1 + 1 + 1 - 1 - 1 \ldots$$

If we extract the subsequences in the even and odd positions, we get

$$\mathbf{p}^1 = +1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 + 1 - 1 \ldots$$

$$\mathbf{p}^2 = +1 + 1 - 1 + 1 + 1 - 1 - 1 + 1 + 1 + 1 - 1 - 1 \ldots$$

Since $\mathbf{p}^2 = \mathbf{p}$, we will get no new sequences by splitting $\mathbf{p}^2$. After splitting the alternating sequence $\mathbf{p}^1$ we have

$$\mathbf{p}^3 = +1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 + 1 \ldots$$

$$\mathbf{p}^4 = -1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 \ldots$$

Splitting these sequences will give us the same sequence again. Therefore, we know that the 2-kernel of the paperfolding sequence is

$$\{\mathbf{p}, \mathbf{p}^1, \mathbf{p}^3, \mathbf{p}^4\}.$$

We can also state this as recursions: $p_{2k} = p_k$, $p_{4k+1} = +1$ and $p_{4k+3} = -1$. Remember that we defined the paperfolding sequence as starting with $n = +1$.

2.5 Regular sequences

Eilenberg's theorem tells us that a sequence is k-automatic if and only if its k-kernel is finite. To generalize the concept of k-automatic sequences, Allouche and Shallit introduced k-regular sequences. A k-regular sequence may be unbounded while a k-automatic sequence can only assume a finite number of values.

Definition 2.5.1. *We say a sequence is k-regular if the $\mathbb{Z}$-module generated by its k-kernel is finitely generated.*

To get a better feeling for this definition, we will provide three examples from [1] and [3].

Example 2.5.2. *A simple example of a k-regular sequence is $s_2(n)$ which counts the 1's in the binary expansion of n. If $0 \leq b < 2^i$, then*

$$s_2(2^i n + b) = s_2(n) + s_2(b).$$

Therefore, every element of the k-kernel can be written as a $\mathbb{Z}$-linear combination of the sequence $(s_2(n))_{n\geq 0}$ and the constant sequence 1.

Example 2.5.3. *The famous Stern–Brocot tree (cf. [33]) is an arrangement of all positive rational numbers in an infinite binary tree. To calculate the Stern–Brocot tree start with the fractions $\frac{0}{1}$ and $\frac{1}{0}$ and insert the mediant $\frac{m+m'}{n+n'}$ between the two parent nodes $\frac{m}{n}$ and $\frac{m'}{n'}$.*

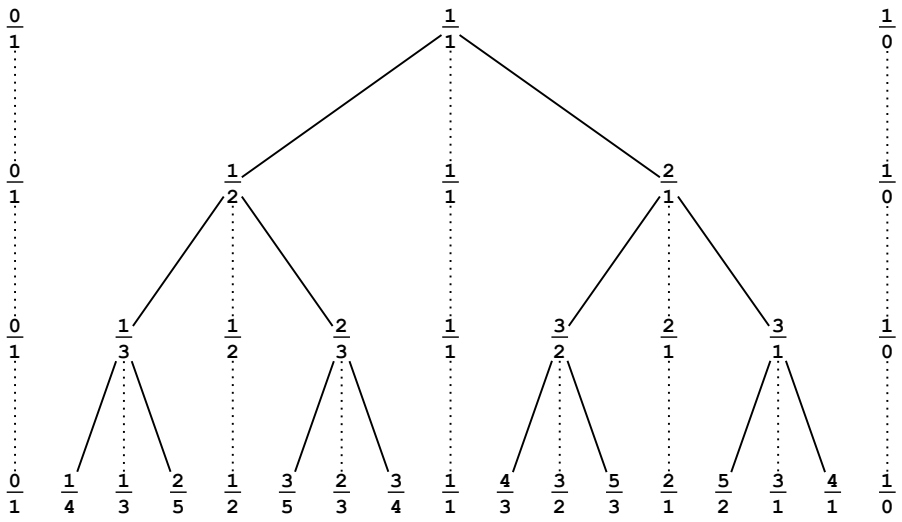

Figure 2.6: Four levels of the Stern–Brocot tree. Picture by [49].

The Stern–Brocot tree is closely related to continued fractions and can be used to compute the best rational approximation a real number x. A rational number

$\frac{m}{n}, n > 0$ is best rational approximation of x if it is closer to x than any rational number with a smaller or equal denominator. Brocot was a French clockmaker who developed the Stern–Brocot tree to find optimal gear ratios for his clocks.

The Christoffel tree is isomorphic to the Stern–Brocot tree (cf. [12]) and the Farey series is a subtree of the Stern–Brocot tree. Furthermore the Stern–Brocot tree is a number system for the rational numbers (cf. [33]).

The numerators of the Stern–Brocot tree are given by the recursion

$$d(0) = 0,\ d(1) = 1,\ d(2n) = d(n),\ d(2n+1) = d(n) + d(n+1).$$

The sequence $d(n)/d(n+1)$ runs through all reduced nonnegative rationals exactly once (cf. [15]). The recursion implies that

$$d(4n+1) = 2d(n) + d(n+1) \text{ and } d(4n+3) = 2d(n+1) + d(n).$$

Therefore, the sequence d is 2-regular.

Example 2.5.4. *A nice example is the Josephus problem (cf. [33]): There are n numbers arranged in a circle. Starting the count with 1 every second number is erased until only one number remains. This number is $J(n)$.*

For $n = 9$ we erase $2, 4, 6, 8, 1, 5, 9, 7$ (in this order) and hence $J(9) = 3$. The sequence starts with $(J(n))_{n\geq 1} = 1, 1, 2, 1, 3, 2, 4, 1, 5, 3, 6, \ldots$

Lemma 2.5.5. *The sequence $(J(n))$ is 2-regular with $J(2n) = 2J(n) - 1$ and $J(2n+1) = 2J(n) + 1$.*

Proof. In the beginning we will always erase all even numbers. If we start with $2n$ numbers after the first revolution in the circle, we are left with the odd numbers $1, 3, \ldots, 2n-1$. This is the same situation as a circle with n numbers after mapping $n \mapsto 2n - 1$, so $J(2n) = 2J(n) - 1$. In the odd case we start with $2n+1$ numbers. The number 1 is deleted after the number $2n$ and we are left with the numbers $3, 5, \ldots, 2n+1$ which is the same situation as a circle with n numbers after mapping $n \mapsto 2n + 1$, so $J(2n) = 2J(n) + 1$. □

Chapter 3

On the 2-Abelian Complexity of the Thue–Morse Word

3.1 Introduction

In this chapter we study the infinite Thue–Morse word $\mathbf{t} = 0110100110010110\cdots$ It is defined as

$$\mathbf{t} := \lim_{n\to\infty} \varrho^n(0)$$

where ϱ is the morphism

$$\varrho\colon 0 \mapsto 01, 1 \mapsto 10.$$

The set of all finite factors of the Thue–Morse word will be denoted by $\mathrm{Fac}(\mathbf{t})$, while $\mathrm{Fac}_n(\mathbf{t})$ stands for the set of Thue–Morse factors of length n.

In this chapter we will prove three theorems about $\mathbf{t}$. The first one is about extensions of factors of $\mathbf{t}$. If we want to prolongate a factor of $\mathbf{t}$ to a longer factor of $\mathbf{t}$ there is sometimes only one possible letter. For example after 00 the next letter has to be a 1. We will give upper and lower bounds for the length of such extensions in Theorem 3.3.5.

Then we take a look at the 2-abelian complexity sequence of the Thue–Morse word. We will prove that it is a concatenation of palindromes of increasing length (Theorem 3.7.3) and secondly, the following theorem which is the main theorem of this chapter.

Theorem 3.1.1. *The 2-abelian complexity of the Thue–Morse word* $\mathbf{t}$ *is 2-regular.*

To understand this theorem, let us take a look at the first concept in Theorem 3.1.1: the ℓ-abelian complexity.

The ℓ-abelian complexity is a complexity measure, which was first introduced in 1981 by Karhumäki [39]. The ℓ-abelian complexity $\mathcal{P}_w^{(\ell)}(n)$ of an infinite word w

builds a bridge between the abelian complexity which corresponds to $\ell = 1$ and the factor complexity which corresponds to $\ell = +\infty$ and allows for a finer resolution. The abelian complexity $\mathcal{P}_w^{(1)}(n)$ of an infinite word counts the anagrams of length n, while the factor complexity $\mathcal{P}_w^{(\infty)}(n)$ counts the factors in $\mathrm{Fac}_n(\mathbf{t})$. The abelian complexity of the infinite Thue–Morse word is $(\mathcal{P}_t^{(1)}(n))_{n\geq 0} = 1, (2,3)^\omega$. Here $(2,3)^\omega$ denotes the right-infinite concatenation $(2,3)^\omega = 232323\cdots$ So $1,(2,3)^\omega$ is the ultimately periodic word $1,(2,3)^\omega = 1232323\cdots$ The factor complexity [14][23] of the Thue–Morse word $\mathbf{t}$ is well known to be

$$\mathcal{P}_t^{(\infty)}(0) = 1, \quad \mathcal{P}_t^{(\infty)}(1) = 2, \quad \mathcal{P}_t^{(\infty)}(2) = 4,$$

$$\mathcal{P}_t^{(\infty)}(n) = \begin{cases} 4n - 2\cdot 2^m - 4, & \text{if } 2\cdot 2^m < n \leq 3\cdot 2^m; \\ 2n + 4\cdot 2^m - 2, & \text{if } 3\cdot 2^m < n \leq 4\cdot 2^m. \end{cases}$$

Before we define ℓ-abelian complexity we need some vocabulary. For a word $w = w_0 w_1 \cdots w_n$ the prefix of length ℓ is defined as $\mathrm{pref}_\ell(w) := w_0 \cdots w_{\ell-1}$ while the suffix of length ℓ is $\mathrm{suff}_\ell(w) := w_{n-\ell+1} \cdots w_n$.

We write $|w|$ to denote the *length* of a word w. If v is a factor of w, the number of occurrences of v in w is denoted by $|w|_v$. We write $\mathbb{N}_0$ for the natural numbers, including 0.

Definition 3.1.2. *For an integer $\ell \geq 1$, two words $u, v \in A^*$ are* ℓ-abelian equivalent*, for some alphabet A, if*

1. $\mathrm{pref}_{\ell-1}(u) = \mathrm{pref}_{\ell-1}(v)$ *and* $\mathrm{suff}_{\ell-1}(u) = \mathrm{suff}_{\ell-1}(v)$*, and*
2. *for all $w \in A^*$ with $|w| = \ell$ the number of occurrences of w in u and v is equal, i.e. $|u|_w = |v|_w$.*

We then write $u \equiv_\ell v$.

There are several equivalent definitions of ℓ-abelian equivalence (cf. [40]), we use the one from [47]. Note that the given definition is slightly redundant, it would suffice to use either $\mathrm{pref}_{\ell-1}(v)$ or $\mathrm{suff}_{\ell-1}(v)$.

It is easy to check that ℓ-abelian equivalence is indeed an equivalence relation. The first part of the definition, where we fix the prefix and suffix, guarantees that two ℓ-abelian equivalent words are also $(\ell-1)$-abelian equivalent.

Example 3.1.3. *Let us take two words $w = 001011$ and $v = 001101$. We see that $w \equiv_2 v$ since $|w|_{00} = 1$, $|w|_{01} = 2$, $|w|_{10} = 1$, $|w|_{11} = 1$ and we get the same values for v. Furthermore, both words have the same prefix and suffix. On the other hand $w \not\equiv_3 v$ since $|001011|_{010} = 1$ and $|001101|_{010} = 0$. Also the suffixes differ, $11 \neq 01$.*

Since $\equiv_\ell$ is an equivalence relation, it is natural to count equivalence classes.

We are interested in the number of 2-abelian equivalence classes for words of a given length:

$$\mathcal{P}_t^{(2)}(n) := \#(\mathrm{Fac}_n(\mathbf{t})/_{\equiv_2}).$$

Usually we write $\mathcal{P}_w^{(\ell)}(n)$ to denote the number of ℓ-abelian equivalence classes of factors of w of length n, where w is an infinite word. In the rest of this chapter, we will only consider the 2-abelian complexity of the Thue–Morse word $\mathbf{t}$. Therefore, we will use the simpler notation $\mathcal{P}_n := \mathcal{P}_t^{(2)}(n)$.

The sequence starts with

$$(\mathcal{P}_n)_{n\geq 0} = 1, 2, 4, 6, 8, 6, 8, 10, 8, 6, 8, 8, 10, 10, 10, 8, 8, 6, 8, 10, 10, 8, 10, 12, 12,$$
$$10, 12, 12, 10, 8, 10, 10, 8, 6, 8, 8, 10, 10, 12, 12, 10, 8, 10, 12, 14, 12, 12, 12, 12, \ldots.$$

Definition 3.1.4. *We assign to every word w its equivalence class. To denote the 2-abelian equivalence class of a word $w = w_0 \cdots w_n$, we use a 6-tuple.*

$$\mathrm{class}\colon \mathrm{Fac}(\mathbf{t}) \to \mathbb{N}_0^4 \times \{0,1\}^2,\ \ w \mapsto (|w|_{00}, |w|_{01}, |w|_{10}, |w|_{11}, w_0, w_n).$$

Example 3.1.5. *We have* $\mathrm{class}(w) = (1, 2, 3, 1, 1, 0)$ *for* $w = 10011010$.

Karhumäki, Saarela and Zamboni showed in [41] that for $n \geq 1$, $m \geq 0$ we have

$$\mathcal{P}_n = O(\log n), \quad \mathcal{P}_n((2\cdot 4^m + 4)/3) = \Theta(m) \quad \text{ and } \mathcal{P}_n(2^m + 1) \leq 8.$$

Actually, $\mathcal{P}_n(2^m + 1) = 6$, for $m \geq 1$ since $\mathcal{P}_n(3) = 6$ and Relation 1. of Theorem 3.1.1.* states $\mathcal{P}_n(4m+1) = \mathcal{P}_n(2m+1)$. The order of growth of the 2-abelian complexity follows also from a recent result in [19].

Theorem 3.1.1 combines two different concepts: the ℓ-abelian complexity and k-regular sequences. We just treated ℓ-abelian complexity, let us now look at k-regular sequences.

Allouche and Shallit introduced k-regular sequences in 1990 [1]. It is a well-known theorem by Eilenberg [28] that a sequence is k-automatic if and only if its k-kernel is finite.

Definition 3.1.6. *Let $k \geq 2$ be an integer. The* k-kernel *of a sequence $(a(n))_{n\geq 0}$ is the set of subsequences*

$$\{(a(k^e n + c))_{n\geq 0} \mid e \geq 0,\ \ 0 \leq c < k^e\}.$$

For example, the Thue-Morse sequence is 2-automatic. Allouche and Shallit [1] took this characterization of k-automatic sequences via the kernel and extended it to k-regular sequences.

Definition 3.1.7 (Allouche and Shallit)**.** *Let $k \geq 2$ be an integer. An integer sequence $(a(n))_{n\geq 0}$ is* k-regular *if the $\mathbb{Z}$-module generated by its k-kernel is finitely generated.*

Just recently research begun to investigate the regularity of the abelian complexity. Madill and Rampersad showed that the abelian complexity of the paperfolding word is 2-regular [44].

This chapter solves an open conjecture from Elise Vandomme, Aline Parreau and Michel Rigo [47], who conjectured that the 2-abelian complexity of the infinite Thue–Morse word is 2-regular. This is a special case of a more general conjecture by Rigo.

Conjecture 3.1.8. *The ℓ-abelian complexity sequence of any k-automatic word is a k-regular sequence.*

Shortly after the discovery of the proof in this chapter, they found an independent proof [46] of their own, which uses the palindromic structure of the sequence.

This chapter is organized as follows:

After some definitions in the rest of this section we will introduce *reading frames* in Section 3.2. Reading frames are a factorization of words into factors $v_1, v_2, \ldots, v_m$ of the form $v_i = \varrho^q(0)$ or $v_i = \varrho^q(1)$ for some $q \in \mathbb{N}$, plus a prefix and suffix of shorter length. Reading frames are a natural way to think about the Thue–Morse word since they preserve the morphism structure.

We use these reading frames in Section 3.3 to prove Theorem 3.3.5 on unique extensions of Thue–Morse factors. For a factor w in $\mathbf{t}$ there is sometimes only one possibility for the next (or previous) letters $x_1 \cdots x_n$ so that $wx_1 \cdots x_n$ (or $x_1 \cdots x_n w$) is again a factor of $\mathbf{t}$. We give lower and upper bounds for the lengths of such unique extensions. Section 3.3 can be skipped if one is only interested in the proof of the 2-regularity.

In Definition 3.1.4 we needed 6 values to describe the 2-abelian equivalence class of a factor w, 4 integer values and 2 binary values. We introduce the *odd frame* in Section 3.4 in order to simplify this 6-tuple of values. The odd frame is a shifted reading frame, which does not preserve the morphism structure but allows us to use only 3 values to represent the 2-abelian equivalence class of a factor w, 2 binary values and 1 integer value. Only the possible values of the integer $\mathrm{count}(w)$, the possible numbers of pairs in a factor w, are nontrivial to determine.

Besides the odd frame we also introduce a *short coding*. The short coding is a way to encode words in the odd frame so that the numbers of pairs in a factor w can be seen immediately.

In Section 3.5 we use the properties of pure odd words to prove a recursion (Theorem 3.5.2) on two types of sets, where $\mathbf{pairs}(n)$ is the set of all possible values of $\mathrm{count}(w)$ for factors w of length n. Once we have the recursion we can

use it to determine $\mathcal{P}_n(n)$ for all n (Theorem 3.5.5). These two theorems are used in all further proofs.

Equipped with Theorem 3.5.2 and Lemma 3.5.5, we prove the main Theorem 3.1.1 in Section 3.6 by verifying 13 linear relations given in Theorem 3.1.1.* which generate all sequences for the $\mathbb{Z}$-kernel. For each of the 13 relations the calculations are similar, but since we have to look at three cases for each of them, a bit lengthy.

Finally, we show some additional properties of $\mathcal{P}_n(n)$ in Section 3.7, most notably that $\mathcal{P}_n(n)$ is a concatenation of longer and longer palindromes. Again, we use Theorem 3.5.2 and Theorem 3.5.5 to do this. We also show that $\mathcal{P}_n(n)$ is unbounded.

Before we continue with the proof, we will need some definitions. We will use the fact that $\mathbf{t}$ is overlap-free (cf. [2], Theorem 1.6.1). An *overlap* is a word of the form $xwxwx$, where w is a word, possibly empty, and x is a single letter. The word $\mathbf{t}$ is also cube-free, i.e. it contains no word of the form www where w is a nonempty factor.

The Thue–Morse word is defined over the binary alphabet $A = \{0, 1\}$. A (literal) *pattern* is a word over the alphabet $E = \{\alpha, \beta\}$. Furthermore we define the involutive letter-to-letter morphisms $\overline{0} := 1$ and $\overline{1} := 0$ and similarly $\overline{\alpha} := \beta$ and $\overline{\beta} := \alpha$. If w is the word $w = x_1x_2 \dots x_n$, we call the word $\overline{w} := \overline{x}_1\overline{x}_2 \dots \overline{x}_n$ the *complement* of w. An *assignment* of a pattern is the image of a pattern under a bijective function from E to A. A pattern p and a word w are *equal* ($p = w$) if $p = w$ for one assignment of p. We introduce patterns to avoid case distinctions for complementary words.

Example 3.1.9. *The pattern* $\alpha\overline{\alpha}\alpha\overline{\alpha}\overline{\alpha}$ *can stand for* 01011 *or* 10100 *depending on the assignment of* α. *So does* $\overline{\alpha}\alpha\overline{\alpha}\alpha\alpha$.

If we want to create a pattern out of a word, we do this via the morphism

$$\mathrm{pat}\colon 0 \mapsto \alpha,\ 1 \mapsto \overline{\alpha}.$$

From now on, we will just write *word* if we mean a finite factor of $\mathbf{t}$.

3.2 Reading frames

From its definition via the morphism ϱ it is clear that the Thue–Morse word is composed of copies of its first 2^q letters, $q \in \mathbb{N}$, and their complements. To denote the special role of these words, we define $f_{2^q} := t_0t_1 \cdots t_{2^q-1}$.

Example 3.2.1. *We take* $q = 2$ *and get* $\mathrm{pat}(f_{2^2}) = \alpha\overline{\alpha\alpha}\alpha$ *which gives us*

$$\mathbf{t} = 01101001100101101001011 \cdots = \underline{0110}\underline{1001}\underline{1001}\underline{0110}\underline{1001}011 \cdots$$

A property that will turn out to be useful later is that the word f_{2^q} has the image $\varrho(f_{2^q}) = f_{2^{q+1}}$ and the preimage $\varrho^{-1}(f_{2^q}) = f_{2^{q-1}}$.

Definition 3.2.2. *A* 2^q-reading frame *of a word* $w \in \mathrm{Fac}(\mathbf{t})$ *is a factorization of* w *into words* $w = pv_1 \cdots v_m s$*, where* $v_1, \cdots, v_m$ *are words of length* 2^q *plus a prefix* p *and a suffix* s *with* $|p|, |s| < 2^q$*, so that* v_i *is a word with* $v_i = \mathrm{pat}(f_{2^q})$*, for the prefix* p *we have* $p = \mathrm{pat}(\mathrm{suff}_{|p|}(f_{2^q}))$ *and for the suffix* s *we have accordingly* $s = \mathrm{pat}(\mathrm{pref}_{|s|}(f_{2^q}))$.

We call p, s and the v_i *frame words*, especially the v_i are called *complete frame words*. The 1-reading frame is called the *trivial frame*. A word $w \in \mathrm{Fac}(\mathbf{t})$ may have several different 2^q-reading frames but at most 2^q. If we shift a 2^q-reading frame one letter to the right, we are in a new reading frame and after 2^q shifts we are in the original reading frame again. If there is only one 2^q-reading frame, it is called the *extensible* reading frame.

We said in the beginning of this section that $\mathbf{t}$ is composed of copies of its first 2^q letters and their complements. So the infinite Thue–Morse word $\mathbf{t}$ has a 2^q-reading frame for every q. Since any word w in $\mathrm{Fac}(t)$ is a factor of $\mathbf{t}$, it can be read in the same reading frame as $\mathbf{t}$. And if w has only one 2^q-reading frame, it has to be the 2^q-reading frame of $\mathbf{t}$.

We can get the previous and next letters of $\mathbf{t}$ if we fill up prefix and suffix to complete frame words in the 2^q-reading frame. We use a *gray* font for filled up letters. Since $\mathbf{t}$ is infinite, the extensible 2^q-reading frame of w can be extended to arbitrary length (possibly in different ways) but the filled up letters are unambiguous. So if the filled up letters in the prefix and suffix give a word which is not in $\mathrm{Fac}(\mathbf{t})$, we can not be in the extensible reading frame.

If there is an extensible 2^q-reading frame, we call the 2^{q-1}-reading frame that we get by splitting every complete 2^q-reading frame word into two complete 2^{q-1} reading frame words, extensible too.

Example 3.2.3. *The word* 0101 *has two* 2*-reading frames:* $\underline{01}|\underline{01}|$ *and* $\underline{0}|\underline{10}|\underline{1}$*, but the extensible* 2*-reading frame is* $\underline{01}|\underline{01}|$*. We can not extend the* $\underline{0}|\underline{10}|\underline{1}$ *reading frame since if we fill up the letters we get* $\underline{1}0|\underline{10}|\underline{10}|$*, but* $\mathbf{t}$ *is cube-free.*

And 0101 *has a unique* 4*-reading frame:* $\underline{01}|\underline{01}$*. We can get the* 4*-reading frame by merging two* 2*-reading frame words. Since every complete* 4*-reading frame word has the pattern* $\alpha\overline{\alpha\alpha}\alpha$*, it can not be* $|\underline{0101}|$ *so the extensible* 4*-reading frame is* $|\underline{10}01|\underline{01}10|$*. Thus,* 0101 *is extensible.*

Let us state some of the previous comments explicitly as lemma.

Lemma 3.2.4. *A word* $w \in \mathrm{Fac}(\mathbf{t})$ *has a* 2^q*-reading frame for every* $q \in \mathbb{N}_0$.

Proof. By definition w occurs somewhere in $\mathbf{t}$. But $\mathbf{t}$ can be read in a 2^q-reading frame for every q therefore w can be read in a 2^q-reading frame too. □

Corollary 3.2.5. *A word w is not in* Fac($\mathbf{t}$) *if there exists an integer q, so that w has no 2^q-reading frame.*

Example 3.2.6. *The words assigned to the patterns $\alpha\alpha\alpha$, $\alpha\alpha\overline{\alpha}\alpha\alpha$ and $\alpha\overline{\alpha}\alpha\overline{\alpha}\alpha$ are not in* Fac($\mathbf{t}$) *since* $\mathbf{t}$ *is overlap-free. We can also prove this with reading frames. The words assigned to the patterns $\alpha\alpha\alpha$ and $\alpha\alpha\overline{\alpha}\alpha\alpha$ are not in* Fac($\mathbf{t}$) *since they have no* 2*-reading frame. The words assigned to the pattern $\alpha\overline{\alpha}\alpha\overline{\alpha}\alpha$ are not in* Fac($\mathbf{t}$) *since they have no* 4*-reading frame.*

3.3 Maximal extensible reading frames

An *extension* of a word $w \in \mathrm{Fac}(\mathbf{t})$ is a pair of words (v_1, v_2) with $v_1, v_2 \in \mathrm{Fac}(\mathbf{t})$, $|v_1| + |v_2| > 0$ so that $v_1 w v_2 \in \mathrm{Fac}(\mathbf{t})$. An extension (v_1, v_2) is *unique* if for all pairs $(y_1, y_2) \in \{0,1\}^* \times \{0,1\}^*$ with $|v_1| = |y_1|, |v_2| = |y_2|$ and $(v_1, v_2) \neq (y_1, y_2)$ it follows $y_1 w y_2 \notin \mathrm{Fac}(\mathbf{t})$.

We already saw that we can get unique extensions of a word $w \in \mathrm{Fac}(\mathbf{t})$ if we fill up prefix and suffix to complete frame words in the extensible reading frame. In this section, we will prove that we get all unique extensions of a word $w \in \mathrm{Fac}(\mathbf{t})$ by filling up prefix and suffix in a certain 2^q-reading frame.

There is a *maximal extensible reading frame* (abbreviated as MERF), since if w is a factor of f_{2^q}, it can not uniquely determine a 2^{q+1}-reading frame.

Example 3.3.1. *The word $w = 011$ has the extensible* 2*-reading frame $\lfloor 01 \rfloor \lfloor 10 \rfloor$. The* 2*-reading frame is the MERF since there are two possible* 4*-reading frames; w is a factor of $\lfloor 0110 \rfloor$ and w is also a factor of $\lfloor 1001 \rfloor \lfloor 1001 \rfloor$.*

For short words it is easy to determine the maximal extensible reading frame by hand, while longer words can be reduced to short words as preimages under the morphism ϱ.

Pattern	MERF
$\lfloor \alpha \rfloor$	1-reading frame
$\lfloor \alpha \rfloor \lfloor \alpha \rfloor$	2-reading frame
$\lfloor \alpha \rfloor \lfloor \overline{\alpha} \rfloor$	1-reading frame
$\lfloor \alpha \rfloor \lfloor \alpha\overline{\alpha} \rfloor$	2-reading frame
$\lfloor \alpha \rfloor \lfloor \overline{\alpha} \rfloor \lfloor \alpha \rfloor$	1-reading frame
$\lfloor \alpha\overline{\alpha} \rfloor \lfloor \overline{\alpha} \rfloor$	2-reading frame

Pattern	MERF
$\lfloor \alpha\alpha\overline{\alpha} \rfloor \lfloor \alpha \rfloor$	4-reading frame
$\lfloor \alpha \rfloor \lfloor \alpha\overline{\alpha} \rfloor \lfloor \overline{\alpha} \rfloor$	2-reading frame
$\lfloor \alpha \rfloor \lfloor \overline{\alpha}\alpha\alpha \rfloor$	4-reading frame
$\lfloor \alpha\overline{\alpha} \rfloor \lfloor \alpha\overline{\alpha} \rfloor$	4-reading frame
$\lfloor \alpha\overline{\alpha} \rfloor \lfloor \overline{\alpha}\alpha \rfloor$	2-reading frame

Table 3.1: MERFs for all nonempty words in Fac($\mathbf{t}$) up to length 4.

The extensible reading frame of a factor v of w also determines the extensible reading frame of w. Therefore, every word in $\mathrm{Fac}(\mathbf{t})$ of length at least 4 has an extensible 2-reading frame. We can now formulate an algorithm to determine the MERF of a word w.

The algorithm determines the extensible 2-reading frame of the word and fills up the prefix and the suffix of w to complete frame words, then takes the preimage of the new word and repeats those steps until it reaches a word with no extensible 2-reading frame. In every step the reading frame size doubles and the algorithm will need q steps if the MERF has size 2^q.

In every step there will be at most two new letters before the word length is halved. So the words will get shorter in every step until they have a length of 4 or shorter. Since the algorithm terminates for all words in Table 3.1, it will terminate in general.

Algorithm 1 Determines the MERF of a word $w \in \mathrm{Fac}(\mathbf{t})$ and fills the MERF

procedure: $MERF(w)$
$q \leftarrow 0$
$w' \leftarrow w$
while w' has a nontrivial reading frame **do**
  $q \leftarrow q + 1$
  $w' \leftarrow$ **FillFrame**(w') {Determines and fills the extensible 2-reading frame}
  $w' \leftarrow \varrho^{-1}(w')$
end while
return $\varrho^q(w')$, 2^q "-reading frame."

To decide whether w has a nontrivial reading frame we can use Table 3.1 as lookup table, since there are only 6 words (3 patterns) with a trivial reading frame. For **FillFrame**(w') we use the same lookup table at the first 4 letters of w' to determine the 2-reading frame, then we find the frame prefix and suffix and fill them up. The original word w will occur only once as factor in $\varrho^q(w')$. Let us look at two examples.

Example 3.3.2. *Let us determine* MERF(0100) *"by hand". The only possibility for a* 2*-reading frame (so that each frame word consists of two different letters) is* $w = \underline{0␣10␣0}$. *We fill the frame to get* $\underline{␣10␣10␣01␣}$. *The only way to join different* 2*-reading frame words to a* 4*-reading frame word is* $\underline{10␣1001␣}$ *and after filling it up we get* $\underline{␣0110␣1001␣}$. *Now there are two different possibilities for an* 8*-reading frame, so we have to stop and conclude that* $\mathrm{MERF}(0100) = \underline{␣0110␣1001␣}$.

How does the algorithm solve the same task? The algorithm starts with the initial values $q = 0$ *and* $w = \underline{0␣10␣0}$. *Then it enters the while loop and calculates*

sequentially the following values

$$q = 1,\ \ FillFrame(w') = ⌊\underline{1}0⌋⌊\underline{1}0⌋⌊\underline{0}1⌋,\ \ \varrho^{-1}(w') = ⌊\underline{1}⌋⌊\underline{1}0⌋,$$

$$q = 2,\ \ FillFrame(w') = ⌊\underline{0}1⌋⌊\underline{1}0⌋,\qquad \varrho^{-1}(w') = ⌊\underline{0}1⌋.$$

The word $⌊\underline{0}⌋⌊\underline{1}⌋$ *has a trivial reading frame so the algorithm leaves the while loop and returns* $\mathrm{MERF}(0100)\colon\ ⌊\underline{0}110⌋⌊\underline{1}001⌋$ *4-reading frame.*

Example 3.3.3. *What is* $\mathrm{MERF}(0110010)$*?*
We start with $w = ⌊\underline{0}1⌋⌊\underline{1}0⌋⌊\underline{0}1⌋⌊\underline{1}0⌋$ *and* $q = 0$*. In the loop we get*

$$q = 1,\ \ FillFrame(w') = ⌊\underline{0}1⌋⌊\underline{1}0⌋⌊\underline{0}1⌋⌊\underline{1}0⌋,\ \ \varrho^{-1}(w') = ⌊\underline{0}⌋⌊\underline{1}0⌋⌊\underline{1}0⌋,$$

$$q = 2,\ \ FillFrame(w') = ⌊\underline{1}0⌋⌊\underline{1}0⌋⌊\underline{0}1⌋,\qquad \varrho^{-1}(w') = ⌊\underline{1}⌋⌊\underline{1}0⌋,$$

$$q = 3,\ \ FillFrame(w') = ⌊\underline{0}1⌋⌊\underline{1}0⌋,\qquad \varrho^{-1}(w') = ⌊\underline{0}⌋⌊\underline{1}⌋.$$

Since $⌊\underline{0}⌋⌊\underline{1}⌋$ *has a trivial reading frame, the algorithm leaves the loop and returns* $\mathrm{MERF}(0110010)\colon\ ⌊\underline{0}1101001⌋⌊\underline{1}0010110⌋$ *8-reading frame.*

The algorithm terminates at an extensible reading frame, so we have the following lemma:

Lemma 3.3.4. *A factor* w *of the Thue–Morse word of length* $2^q \le |w| < 2^{q+1}$ *has an extensible* 2^{q-1}*-reading frame.*

Proof. The output word of the algorithm is at least as long as the input word. The while loop of the algorithm will only end if it reaches a word with pattern $\mathrm{pat}(w) = \alpha\overline{\alpha}$ or $\mathrm{pat}(w) = \alpha\overline{\alpha}\alpha$. Hence the output word will have length $2 \cdot 2^i$ or $3 \cdot 2^i$ for $i \in \mathbb{N}_0$. But $3 \cdot 2^i \ge 2 \cdot 2^i \ge 2^q$ implies $i \ge q - 1$. □

Equipped with the algorithm, we are ready to prove the main theorem of this section. As usual, we define $a \bmod b := a - \lfloor \frac{a}{b} \rfloor b$.

Theorem 3.3.5. *Any factor* w *of the infinite Thue–Morse word with a given length* $n := |w| = 2^q + r$*, where* $r < 2^q$*, uniquely determines at least* $u_{\min}(n)$ *and at most* $u_{\max}(n)$ *letters where*

$$u_{\min}(n) := \begin{cases} 0 \text{ for } n = 1 \\ 0 \text{ for } n = 2 \\ 0 \text{ for } n = 3 \\ -n \bmod 2^{q-1} \text{ for } n > 3 \end{cases} \qquad \text{and } u_{\max}(n) := \begin{cases} 0 \text{ for } n = 1 \\ 2 \text{ for } n = 2 \\ 1 \text{ for } n = 3 \\ 2^{\lfloor \log_2(n-2) \rfloor + 2} - n \text{ for } n > 3. \end{cases}$$

These bounds are sharp.

Proof. Table 3.1 allows us to check the cases with $n \leq 3$. Then we take a look at the function $u_{\min}(n)$. According to Lemma 3.3.4 the word w has an extensible 2^{q-1}-reading frame and therefore determines at least $2^{q-1} \cdot i - n$ letters for some $i \in \mathbb{N}_0$. But the smallest positive value of $2^{q-1} \cdot i - n$ is exactly $-n \operatorname{Mod} 2^{q-1}$. To show that it is actually possible to obtain this value for $r \leq 2^{q-1}$ and $r > 2^{q-1}$, take the first n letters of $f_{2^{q-1}}\overline{f_{2^{q-1}}}f_{2^{q-1}}$ and $f_{2^q}\overline{f_{2^q}}$, respectively.

To analyze the function $u_{\max}(n)$, we insert a word $w \in \mathrm{Fac}_n(\mathbf{t})$ in a 2^{q-1}-reading frame, which exists according to Lemma 3.3.4. A word of length 2^q or $2^q + 1$ can determine up to 3 frame words in the extensible 2^{q-1}-reading frame. So after $q-1$ iterations of the while loop we have a word of length 3. We enter the while loop again, extend the word (in the best case) to length 4 and then map it via ϱ^{-1} to a word with pattern $\alpha\overline{\alpha}$. So after q iterations we determined $2 \cdot 2^q$ letters.

If the word w has length $2^q + 2 \leq |w| \leq 2^q + 2^{q-1} + 1$, it can determine 4 frame words. So after $q-1$ iterations we have a word of length 4 which (in the best case) has a 4-reading frame and gives therefore 2 further iterations before we end up in a word with pattern $\alpha\overline{\alpha}$. Here we determined $2 \cdot 2^{q+1}$ letters.

If $2^q + 2^{q-1} + 2 \leq |w| \leq 2^{q+1} - 1$, the word w can determine 5 frame words, so we have a word of length 5 after $q-1$ iterations, extend it to length 6, map it to length 3 and (in the best case) extend it to length 4, before it is mapped to a word with pattern $\alpha\overline{\alpha}$. Again we determined $2 \cdot 2^{q+1}$ letters.

In each of these cases we determined $2^{\lfloor \log_2(n-2) \rfloor + 2} - n$ new letters, but we always assumed a best case. What is left is to show that there is always a word w, with $|w| = 2^q + r$, $r < 2^q$, so that the best case occurs. The first n letters of $\mathrm{suff}_1(f_{2^{q-1}})f_{2^{q-1}}\overline{f_{2^{q-1}}f_{2^{q-1}}}f_{2^{q-1}}$ form such a word. □

Let us look at the relative length $\frac{|w'|}{|w|}$ of an extension, where w is the input and w' is the output of the algorithm.

Lemma 3.3.6. *Let w be a factor of the Thue–Morse word and let w' be its unique extension. The relative length taken over all $w \in \mathrm{Fac}(\mathbf{t})$ satisfies* $\inf \frac{|w'|}{|w|} = 1$ *and* $\sup \frac{|w'|}{|w|} = 4$.

Proof. We have $\frac{|w'|}{|w|} = 1$ for the words $v = f_{2^q}\overline{f_{2^q}}$ of length $2 \cdot 2^q$ and $w = f_{2^q}\overline{f_{2^q}}f_{2^q}$ of length $3 \cdot 2^q$.

For the upper limit we look at the words $w_q = \mathrm{suff}_1(f_{2^{q-1}})f_{2^{q-1}}\overline{f_{2^{q-1}}}1$. These words of length $2^q + 2$ have an unique extension w'_q of length $4 \cdot 2^q$. Therefore, we have $\lim_{q\to\infty} \frac{|w'_q|}{|w_q|} = \lim_{q\to\infty} \frac{4 \cdot 2^q}{2^q+2} = 4$. □

Example 3.3.7. *The word $w = \underline{0110}\underline{1001}$ is a word of length 8 without an extension ($\frac{|w'|}{|w|} = 1$), while $v = \underline{0}\underline{01101}$ is a word of length 6 which can be extended to the word $v' = \underline{10010110}\underline{01101001}$ of length 16 ($\frac{|v'|}{|v|} = 8/3$).*

3.4 The odd frame and the short coding

In this section we compress the information in class(w) from Definition 3.1.4. From now on we call the extensible 2-reading frame also *even frame*. We can get another reading frame if we shift the even frame one letter. This new reading fame is called *odd frame*.

We use these names since a given factor occurs an infinite number of times in the Thue–Morse word but always with the same parity (for $n > 3$) of the first letter. So a word is in the *even frame* if its first letter in the Thue–Morse word has even parity and it is in the *odd frame* otherwise.

Example 3.4.1. *A word* 01011 *can be read in the even frame* $\underline{01}|\underline{01}|\underline{1}$ *or in the odd frame* $\underline{0}|\underline{10}|\underline{11}$.

While the only two complete frame words in the even frame are 01 and 10, we have the four complete frame words $00, 01, 10$ and 11 in the odd frame. We call the odd frame words 00 and 11 *pairs*. The easiest way to find the odd frame of a word is to look for pairs, since pairs can only occur in the odd frame.

We define a *short coding* for odd frame words as: $\underline{01}|, |\underline{10}| \mapsto \mathbf{D}$(ifferent), $\underline{00}|, |\underline{11}| \mapsto \mathbf{E}$(qual) and finally $\underline{0}|, \underline{1}|, |\underline{0}, |\underline{1} \mapsto \mathbf{S}$(hort).

An odd word without a prefix and suffix in the odd frame is called *pure odd word*. The study of pure odd words will turn out to be crucial for the rest of the chapter. There is no **S** in the short coding of a pure odd word and all pure odd words have even length.

Example 3.4.2. *The word* $v = 1100$ *is a pure odd word since it has the odd frame* $v = |\underline{11}|\underline{00}|$ *and the short coding* **EE**. *On the other hand, the word* $w = 01001$ *is not a pure odd wort since the odd frame* $w = \underline{01}|\underline{00}|\underline{1}$ *has a single letter suffix and therefore the short coding* **DES**.

If an odd frame word ends with one letter, the next one starts with another letter since $\underline{\alpha}|\underline{\alpha}$ in the odd frame would be $|\alpha\alpha|$ in the even frame which can not occur. This fact allows us to recover a word w from its short coding if we know a single letter of w, and to recover pat(w) from the short coding too.

Example 3.4.3. *Take the word* $w = 1001011$. *It contains two pairs* $1|\underline{00}|10|\underline{11}|$ *and has therefore the odd frame* $w = \underline{1}|\underline{00}|\underline{10}|\underline{11}|$. *The short coding* **SEDE** *gives the pattern* $\underline{\alpha}|\overline{\alpha\alpha}|\alpha\overline{\alpha}|\alpha\alpha|$ *since* $\mathbf{t}$ *is cube-free. If we know* w_0, *we can recover* w *from its short coding.*

Thus, we can switch between patterns in the odd frame and the short coding. We will use this in the following proofs.

Lemma 3.4.4. *The odd frame of the infinite Thue–Morse word has following properties:*

- *The sequence* **DD** *can not occur.*
- *The sequence* **DEED** *can not occur.*
- *The sequence* **EEEE** *can not occur.*

Proof. We use Corollary 3.2.5 to prove the lemma.

- We showed in Example 3.2.3 that the word 0101 and therefore the pattern $\alpha\overline{\alpha}\alpha\overline{\alpha}$ is in the even frame.
- A word with pattern $\alpha\overline{\alpha}\alpha\alpha\overline{\alpha}\overline{\alpha}\alpha\overline{\alpha}$ has no 4-reading frame and is therefore not in $\mathrm{Fac}(\mathbf{t})$.
- A word with pattern $\alpha\alpha\overline{\alpha}\overline{\alpha}\alpha\alpha\overline{\alpha}\overline{\alpha}$ has no 8-reading frame and is therefore not in $\mathrm{Fac}(\mathbf{t})$.

□

So, at least every second letter in a short coding is an **E** and at most 3/4 of the letters are **E**. This gives an upper bound for the growth of the 2-abelian complexity $\mathcal{P}_n$ of $\mathbf{t}$. As consequence of Lemma 3.4.4 two consecutive **E** have either the form **EE** or **EDE** which corresponds to the patterns $\alpha\alpha\overline{\alpha}\overline{\alpha}$ and $\alpha\alpha\overline{\alpha}\alpha\overline{\alpha}\overline{\alpha}$. So, we just proved the next Lemma.

Lemma 3.4.5. *The pairs* 00 *and* 11 *alternate in the odd frame.*

Now, we will use the odd frame to compress the information from Definition 3.1.4. To achieve this, we define a function

$$\mathrm{count}\colon \mathrm{Fac}(\mathbf{t}) \to \mathbb{N}_0,\ \ w \mapsto |w|_{00} + |w|_{11},$$

which counts the pairs in a word w, and a function

$$\mathrm{frame}(w) := \begin{cases} 0, & \text{if } w_0w_1 \text{ is in the even frame;} \\ 1, & \text{if } w_0w_1 \text{ is in in the odd frame;} \end{cases}$$

which determines the reading frame of w. This means a word w is a pure odd word if $\mathrm{frame}(w) = 1$ and $|w|$ is even. It would be possible to use the short coding to define $\mathrm{frame}(w)$: $\mathrm{frame}(w) = 0$ if the short coding of w starts with **S** and $\mathrm{frame}(w) = 1$ otherwise.

With the two functions $\mathrm{count}(w)$ and $\mathrm{frame}(w)$ we can collect all information necessary to determine the 2-abelian equivalence class of a word w in a 3-tuple:

$$\mathrm{tup}\colon \mathrm{Fac}(\mathbf{t}) \to \mathbb{N}_0 \times \{0,1\}^2,\ \ w \mapsto (w_0, \mathrm{count}(w), \mathrm{frame}(w)).$$

Example 3.4.6. *Let us look at Example 3.1.5 again. For the word* $w = 10011010$ *we have now* $\mathrm{tup}(w) = (1, 2, 0)$.

In the next theorem, we show that we have all information of $\mathrm{class}(w)$ in $\mathrm{tup}(w)$, we can recover $\mathrm{class}(w)$ from $\mathrm{tup}(w)$. We will use the *XOR* operator $\oplus$ and the Iverson bracket $[\![S]\!]$ which is 1 if the statement S is true and 0 otherwise.

Lemma 3.4.7. *Let* w *be a word of length* n*. There is a function* h *so that*

$$h(\mathrm{tup}(w)) = \mathrm{class}(w).$$

For two words $v, w \in$ *with* $v \neq w$*, we have* $h(\mathrm{tup}(v)) = h(\mathrm{tup}(w))$ *if and only if* $v_0 = w_0$*,* $\mathrm{count}(v) = \mathrm{count}(w)$ *and* $\mathrm{count}(v)$ *is even.*

Proof. The basic idea is to use parity arguments. If we take w and erase one letter from every pair, we get a sequence of length $|w| - \mathrm{count}(w)$ which alternates between 0 and 1 and starts with w_0. Since the sequence has the same number of 01 and 10 as w, we can use it to determine $|w|_{01}$, $|w|_{10}$ and the last letter w_n.

The pairs already form an alternating sequence (Lemma 3.4.5), so we only need to identify the first pair. In the odd frame a word can start either with **E** which corresponds to the pattern $\alpha\alpha$ or with **DE** which corresponds to the pattern $\alpha\overline{\alpha}\alpha\alpha$. In both cases the first pair is w_0w_0.

In the even frame a word starts with **SE** which corresponds to the pattern $\alpha\overline{\alpha\alpha}$ or with **SDE** which corresponds to the pattern $\alpha\overline{\alpha}\alpha\overline{\alpha\alpha}$. In both cases the first pair is $\overline{w_0w_0}$. This allows us to determine $|w|_{00}$, $|w|_{11}$. We can also give these values in an explicit form as

$$h\colon (w_0, \mathrm{count}(w), \mathrm{frame}(w)) \mapsto (a, b, c, d, w_0, e)$$

with

$$\begin{aligned}
a &= \lfloor \tfrac{\mathrm{count}(w)}{2} \rfloor + [\![\mathrm{count}(w) \text{ odd}]\!][\![w_0 \neq \mathrm{frame}(w)]\!] \\
b &= \lfloor \tfrac{|w| - \mathrm{count}(w)}{2} \rfloor + [\![|w| - \mathrm{count}(w) \text{ even}]\!][\![w_0 = 0]\!] \\
c &= \lfloor \tfrac{|w| - \mathrm{count}(w)}{2} \rfloor + [\![|w| - \mathrm{count}(w) \text{ even}]\!][\![w_0 = 1]\!] \\
d &= \lfloor \tfrac{\mathrm{count}(w)}{2} \rfloor + [\![\mathrm{count}(w) \text{ odd}]\!][\![w_0 = \mathrm{frame}(w)]\!] \\
e &= [\![|w| - \mathrm{count}(w) \text{ even}]\!] \oplus w_0.
\end{aligned}$$

Let $v, w \in \mathrm{Fac}(\mathbf{t})$ with $v \neq w$ now be two words that belong to the same equivalence class. Then $v_0 = w_0$ and $\mathrm{count}(v) = \mathrm{count}(w)$, so they can only differ in the reading frame with $\mathrm{frame}(v) \neq \mathrm{frame}(w)$. The reading frame determines the first pair in the alternating pair sequence. If $\mathrm{count}(w)$ is even, the numbers $|w|_{00}$ and $|w|_{11}$ do not depend on $\mathrm{frame}(w)$. So $h(\mathrm{tup}(v)) = h(\mathrm{tup}(w))$ for $v \neq w$ if and only if $v_0 = w_0$, $\mathrm{count}(v) = \mathrm{count}(w)$ and $\mathrm{count}(w)$ is even. □

The idea of Lemma 3.4.7 is to gather more information in less memory. We need two boolean and four integer variables for class(w) while tup(w) uses only one integer and two boolean variables.

Example 3.4.8. *We have* class(w) = class(v) *for the two words* $w = 011001$ *and* $v = 001011$ *but* tup$(w) \neq$ tup(v) *since* frame$(w) = 0$ *and* frame$(v) = 1$. *So* tup *can distinguish more words than* class.

With Lemma 3.4.7 we can determine the 2-abelian complexity $\mathcal{P}_n$ (cf. page 21) of $\mathbf{t}$ if we know the possible values of tup(w), for $w \in \mathrm{Fac}_n(\mathbf{t})$. Which are the possible values of tup(w)?

The boolean variables can be 0 or 1 since a word $w \in \mathrm{Fac}_n(\mathbf{t})$ can start either with 0 or 1 and can be in the even frame or in the odd frame. The difficult part is to find the possible values of count(w). In the next section we will find a method to obtain them.

3.5 On pairs

We are interested in how many pairs can occur in a word in $\mathrm{Fac}_n(\mathbf{t})$. So we define

$$\mathbf{pairs}(n) := \{\mathrm{count}(w) \mid w \in \mathrm{Fac}_n(\mathbf{t})\}$$

where count(w) counts the number of pairs in the word w. It will emerge that we will need a second set

$$\mathbf{PAIRS}(2n) := \{\mathrm{count}(w) \mid \mathrm{frame}(w) = 1,\ w \in \mathrm{Fac}_{2n}(\mathbf{t})\}.$$

The value $\mathbf{PAIRS}(n)$ is undefined for odd n. The elements of $\mathbf{PAIRS}(2n)$ are the possible numbers of $\mathbf{E}$ in pure odd words of length $2n$.

Example 3.5.1. *Let us determine* $\mathbf{pairs}(6)$ *and* $\mathbf{PAIRS}(6)$ *with Lemma 3.4.4.*

Pattern	*Coding*	count(w)
$\alpha\alpha\overline{\alpha}\alpha\overline{\alpha}\overline{\alpha}$	**EDE**	2
$\alpha\alpha\overline{\alpha}\overline{\alpha}\alpha\alpha$	**EEE**	3
$\alpha\alpha\overline{\alpha}\overline{\alpha}\alpha\overline{\alpha}$	**EED**	2
$\alpha\overline{\alpha}\alpha\alpha\overline{\alpha}\alpha$	**DED**	1

Pattern	*Coding*	count(w)
$\alpha\overline{\alpha}\alpha\alpha\overline{\alpha}\overline{\alpha}$	**DEE**	2
$\alpha\overline{\alpha}\alpha\overline{\alpha}\overline{\alpha}\alpha$	**SDES**	1
$\alpha\overline{\alpha}\overline{\alpha}\alpha\alpha\overline{\alpha}$	**SEES**	2
$\alpha\overline{\alpha}\overline{\alpha}\alpha\overline{\alpha}\alpha$	**SEDS**	1

So $\mathbf{pairs}(6) = \mathbf{PAIRS}(6) = \{1, 2, 3\}$.

It is not always the case that $\mathbf{pairs}(2n)$ is equal to $\mathbf{PAIRS}(2n)$. For example $\mathbf{pairs}(8) = \{1, 2, 3\}$ while $\mathbf{PAIRS}(8) = \{2, 3\}$.

The following theorem is the main tool to prove results about the 2-abelian complexity $\mathcal{P}_n$ (cf. page 21), since all properties of $\mathcal{P}_n$ can be obtained from the properties of $\mathbf{pairs}(n)$ and $\mathbf{PAIRS}(n)$.

Theorem 3.5.2. *For $n \geq 4$ the sets $\mathbf{pairs}(n)$ and $\mathbf{PAIRS}(n)$ fulfill the recursions*

$$\mathbf{PAIRS}(2n) = n - \mathbf{pairs}(n+1) \tag{3.1}$$

$$\mathbf{pairs}(2n+1) = \mathbf{PAIRS}(2n) \tag{3.2}$$

$$\mathbf{pairs}(2n) = \mathbf{PAIRS}(2n) \cup \mathbf{PAIRS}(2n-2) \tag{3.3}$$

with $n - \mathbf{pairs}(n+1) := \{n - y \mid y \in \mathbf{pairs}(n+1)\}$.

Proof. The proof works only for $n \geq 4$ since we distinguish between even and odd frame and words of length $n < 4$ may not have a defined 2-reading frame.

Let $w \in \mathrm{Fac}_{n+1}(\mathbf{t})$. We have $|w|_{01} + |w|_{10} = n - \mathrm{count}(w)$ since w has n factors of length 2. The image $\varrho(w)$ has length $2n+2$ and is in the even frame. We remove the first and last letter of $\varrho(w)$ to get a pure odd word w' of length $2n$. We have a pair 00 in w' if and only if there is a 10 in w and a pair 11 in w' if and only if there is a 01 in w. Thus, $\mathrm{count}(w') = n - \mathrm{count}(w)$. Since the steps to get from w to w' are bijective, the two sets $\mathbf{PAIRS}(2n)$ and $n - \mathbf{pairs}(n+1)$ are equal.

All words $w \in \mathrm{Fac}_{2n+1}(\mathbf{t})$ are of the form $w'\mathbf{S}$ or $\mathbf{S}w'$ where w' is a pure odd word of length $2n$. Since $\mathrm{count}(w') = \mathrm{count}(w'\mathbf{S}) = \mathrm{count}(\mathbf{S}w')$, the bijection $w' \mapsto w'\mathbf{S}$ proves $\mathbf{pairs}(2n+1) = \mathbf{PAIRS}(2n)$.

Every word in $w \in \mathrm{Fac}_{2n}(\mathbf{t})$ is either of the form w' or $\mathbf{S}w''\mathbf{S}$, where w' is a pure odd word of length $2n$ and w'' is a pure odd word of length $2n-2$. Again, adding and removing $\mathbf{S}$ is a bijection which does not change the number of pairs. Therefore, $\mathbf{pairs}(2n) = \mathbf{PAIRS}(2n) \cup \mathbf{PAIRS}(2n-2)$. □

Example 3.5.3. *In Example 3.5.1 we determined the values of $\mathbf{pairs}(6)$ and $\mathbf{PAIRS}(6)$ directly. Now we can use Theorem 3.5.2 and Table 3.2 to determine these values.*

$$\mathbf{PAIRS}(6) = 3 - \mathbf{pairs}(4) = 3 - \{0, 1, 2\} = \{1, 2, 3\},$$

$$\mathbf{pairs}(6) = \mathbf{PAIRS}(6) \cup \mathbf{PAIRS}(4) = \{1, 2, 3\} \cup \{1, 2\} = \{1, 2, 3\}.$$

We write $[a, b]$ to denote the interval of all integers between a and b, including both. The integer interval $[a, b]$ has cardinality $\#[a, b] = b - a + 1$. For a set S we define $\#_2 S$, the number of even elements in S, as

$$\#_2 S := \#\{s \in S \mid s \equiv 0 \pmod 2\}.$$

Lemma 3.5.4. *The sets* $\mathbf{pairs}(n)$ *and* $\mathbf{PAIRS}(2n)$ *are integer intervals.*

Proof. This is true for $n < 10$ (cf. Table 3.2). All other values can be calculated using Theorem 3.5.2. In cases (1) and (2) of Theorem 3.5.2 it is obvious that integer intervals are mapped to integer intervals, we just have to show that the set $\mathbf{PAIRS}(2n) \cup \mathbf{PAIRS}(2n-2)$ is an integer interval. This is true, since we know from the definition of $\mathbf{PAIRS}(2n)$ that the upper and lower limit of two consecutive sets can differ only by 1. □

n	$\mathbf{PAIRS}(n)$	$\mathbf{pairs}(n)$	$\mathcal{P}_n$
0	{0}	{0}	1
1		{0}	2
2	{0,1}	{0,1}	4
3		{0,1}	6
4	{1,2}	{0,1,2}	8
5		{1,2}	6
6	{1,2,3}	{1,2,3}	8
7		{1,2,3}	10
8	{2,3}	{1,2,3}	8
9		{2,3}	6

Table 3.2: Small values of the functions

In the next theorem we determine the number of 2-abelian equivalence classes $\mathcal{P}_n$ (cf. page 21).

Lemma 3.5.5. *For* $n \geq 4$ *the number of* 2*-abelian equivalence classes* $\mathcal{P}_n$ *is given by the two formulas*

$$\begin{aligned}\mathcal{P}_{2n} &= 2(\#\mathbf{PAIRS}(2n) + \#\mathbf{PAIRS}(2n-2) \\ &\qquad - \#_2(\mathbf{PAIRS}(2n) \cap \mathbf{PAIRS}(2n-2))) \\ \mathcal{P}_{2n+1} &= 2(2\#\mathbf{PAIRS}(2n) - \#_2\mathbf{PAIRS}(2n)).\end{aligned}$$

Proof. This is an immediate consequence of Lemma 3.4.7 and Theorem 3.5.2. First we look at $\mathcal{P}_{2n}$. We want to find all possible values of $\mathrm{tup}(w)$ for $w \in \mathrm{Fac}_{2n}(\mathbf{t})$. We have $\#\mathbf{PAIRS}(2n)$ possibilities to choose $\mathrm{count}(w)$ in the odd frame and we have $\#\mathbf{PAIRS}(2n-2)$ possibilities to choose $\mathrm{count}(w)$ in the even frame. If there is an even value $\mathrm{count}(w)$ in both frames, the odd frame and the even frame give the same equivalence class so we subtract $\#_2(\mathbf{PAIRS}(2n) \cap \mathbf{PAIRS}(2n-2))$. Finally we multiply by 2 since w_0 can be 0 or 1.

We use the same argument for $\mathcal{P}_{2n+1}$, but we have $\mathbf{PAIRS}(2n)$ possible pairs in both frames and therefore also in the intersection. □

Now we look at several sets, because for even numbers the recursion needs two sets. If we know $\mathbf{pairs}(n)$ and $\mathbf{pairs}(n+1)$, we can use Theorem 3.5.2 to

determine $\mathbf{PAIRS}(2n-2)$ and $\mathbf{PAIRS}(2n)$ and thus $\mathbf{pairs}(2n-1)$, $\mathbf{pairs}(2n)$ and $\mathbf{pairs}(2n+1)$. With Lemma 3.5.5 we can then determine $\mathcal{P}_{2n-1}$, $\mathcal{P}_{2n}$ and $\mathcal{P}_{2n+1}$.

It is clear from the definition of $\mathbf{pairs}(n)$ that the sequence $\min \mathbf{pairs}(n)$ is monotonically increasing in steps of 0 or 1:

$$\min \mathbf{pairs}(n+1) - \min \mathbf{pairs}(n) \in \{0, 1\}.$$

This also holds for $\max \mathbf{pairs}(n)$, $\min \mathbf{PAIRS}(n)$ and $\max \mathbf{PAIRS}(n)$. So if $\mathbf{pairs}(n) = [a, b]$, we have four possibilities for $\mathbf{pairs}(n+1)$:

$$[a, b], \ \ [a+1, b], \ \ [a, b+1] \text{ and } [a+1, b+1].$$

Now we use Theorem 3.5.2 to make Table 3.3.

Example 3.5.6. *Let us take a look at the first case in Table 3.3. We start with the two sets* $\mathbf{pairs}(n) = [a, b]$ *and* $\mathbf{pairs}(n+1) = [a, b]$. *Now we use Equation (3.1) of Theorem 3.5.2 and get*

$$\mathbf{PAIRS}(2n) = n - \mathbf{pairs}(n+1) = n - [a, b] = [n-b, n-a]$$

and

$$\mathbf{PAIRS}(2n-2) = n - 1 - \mathbf{pairs}(n) = n - 1 - [a, b] = [n-b-1, n-a-1].$$

Then we use Theorem 3.5.2 Equation (3.2) to get

$$\mathbf{pairs}(2n-1) = \mathbf{PAIRS}(2n-2) = [n-b-1, n-a-1]$$

and

$$\mathbf{pairs}(2n+1) = \mathbf{PAIRS}(2n) = [n-b, n-a].$$

Finally we use use Theorem 3.5.2 Equation (3.3) and get

$$\mathbf{pairs}(2n) = \mathbf{PAIRS}(2n-2) \cup \mathbf{PAIRS}(2n) = [n-b-1, n-a-1] \cup [n-b, n-a] =$$
$$= [n-b-1, n-a].$$

We continue in the same manner to complete Table 3.3.

In Case *IV.* in Table 3.3 we have $\mathbf{PAIRS}(2n-2) = \mathbf{PAIRS}(2n)$ (and also $\mathbf{pairs}(n+1) = 1 + \mathbf{pairs}(n)$). If we observe small values of $\mathbf{PAIRS}(2n)$ in Table 3.2, we see that Case *IV.* does not occur. The column $\mathbf{PAIRS}(2n+i)$ of Table 3.3 shows that Case *IV.* can also not occur as image of other values. Therefore, the Case *IV.* can not occur anywhere and we have just proved:

$$\mathbf{PAIRS}(2n-2) \neq \mathbf{PAIRS}(2n) \text{ and } \mathbf{pairs}(n+1) \neq 1 + \mathbf{pairs}(n). \tag{3.4}$$

From the way we made Table 3.3 we can also deduce the next lemma about consecutive values of $\mathbf{pairs}(n)$, $\mathbf{PAIRS}(n)$ and $\mathcal{P}_n$ since we can calculate $\mathcal{P}_n$ with Lemma 3.5.5.

Case	i	**pairs**$(n+i)$	**PAIRS**$(2n+i)$	**pairs**$(2n+i)$	**PAIRS**$(4n+i)$	**pairs**$(4n+i)$
I.	-2		$[n-b-1, n-a-1]$		$[n+a-1, n+b]$	
	-1			$[n-b-1, n-a-1]$		$[n+a-1, n+b]$
	0	$[a, b]$	$[n-b, n-a]$	$[n-b-1, n-a]$	$[n+a, n+b]$	$[n+a-1, n+b]$
	$+1$	$[a, b]$		$[n-b, n-a]$		$[n+a, n+b]$
II.	-2		$[n-b-1, n-a-1]$		$[n+a, n+b]$	
	-1			$[n-b-1, n-a-1]$		$[n+a, n+b]$
	0	$[a, b]$	$[n-b, n-a-1]$	$[n-b-1, n-a-1]$	$[n+a+1, n+b]$	$[n+a, n+b]$
	$+1$	$[a+1, b]$		$[n-b, n-a-1]$		$[n+a+1, n+b]$
III.	-2		$[n-b-1, n-a-1]$		$[n+a-1, n+b]$	
	-1			$[n-b-1, n-a-1]$		$[n+a-1, n+b]$
	0	$[a, b]$	$[n-b-1, n-a]$	$[n-b-1, n-a]$	$[n+a, n+b+1]$	$[n+a-1, n+b+1]$
	$+1$	$[a, b+1]$		$[n-b-1, n-a]$		$[n+a, n+b+1]$
IV.	-2		$[n-b-1, n-a-1]$		$[n+a, n+b]$	
	-1			$[n-b-1, n-a-1]$		$[n+a, n+b]$
	0	$[a, b]$	$[n-b-1, n-a-1]$	$[n-b-1, n-a-1]$	$[n+a+1, n+b+1]$	$[n+a, n+b+1]$
	$+1$	$[a+1, b+1]$		$[n-b-1, n-a-1]$		$[n+a+1, n+b+1]$

Table 3.3: Behavior of integer intervals under the maps **pairs**(n) and **PAIRS**(n).

Lemma 3.5.7. *If we know* **pairs**(n) *for all* $n \in [a, b]$*, we can determine* **pairs**(n') *and* $\mathcal{P}_{n'}$ *for* $n' \in [2a-1, 2b-1]$ *and* **PAIRS**(n'') *for* $n'' \in [2a-2, 2b-2]$*.*

Remark: *An especially interesting case of Lemma 3.5.7 is to go from* **pairs**(n)*,* $n \in [2^{q-1}+1, 2^q+1]$ *to* **pairs**(n')*,* $n' \in [2^q+1; 2^{q+1}+1]$*. This is the idea behind the proof of Theorem 3.7.3.*

3.6 The sequence $\mathcal{P}_n$ is 2-regular

We will prove Theorem 3.1.1 by proving the 13 relations of Theorem 3.1.1.* which generate all sequences for the $\mathbb{Z}$-kernel. If Theorem 3.1.1.* is true, the $\mathbb{Z}$-kernel is finitely generated and Theorem 3.1.1 follows. Please remember that $\mathcal{P}_n$ stands for the 2-abelian complexity sequence (cf. page 21).

Theorem 3.1.1.* *The $\mathbb{Z}$-kernel of the Thue-Morse word is generated by the* 13 *relations:*

1. $\mathcal{P}_{4n+1} = \mathcal{P}_{2n+1}$
2. $\mathcal{P}_{8n+4} = \mathcal{P}_{8n+3} + \mathcal{P}_{4n+3} - \mathcal{P}_{4n+2}$
3. $\mathcal{P}_{16n} = \mathcal{P}_{8n}$

4. $\mathcal{P}_{16n+2} = \mathcal{P}_{8n+2}$
5. $\mathcal{P}_{16n+6} = -\mathcal{P}_{16n+3} + \mathcal{P}_{8n+3} + 3\mathcal{P}_{8n+2} + \mathcal{P}_{4n+3} - 2\mathcal{P}_{4n+2} - \mathcal{P}_{2n+1}$
6. $\mathcal{P}_{16n+7} = -\mathcal{P}_{16n+3} + \mathcal{P}_{8n+3} + 3\mathcal{P}_{8n+2} + 2\mathcal{P}_{4n+3} - 3\mathcal{P}_{4n+2} - \mathcal{P}_{2n+1}$
7. $\mathcal{P}_{16n+8} = \mathcal{P}_{8n+2} + \mathcal{P}_{4n+3} - \mathcal{P}_{2n+1}$
8. $\mathcal{P}_{16n+10} = \mathcal{P}_{8n+2} + \mathcal{P}_{4n+3} - \mathcal{P}_{2n+1}$
9. $\mathcal{P}_{16n+11} = -\mathcal{P}_{16n+3} + 3\mathcal{P}_{8n+2} + \mathcal{P}_{4n+3} - 2\mathcal{P}_{2n+1}$
10. $\mathcal{P}_{16n+14} = \mathcal{P}_{16n+3} + \mathcal{P}_{8n+7} - \mathcal{P}_{8n+3} - \mathcal{P}_{8n+2} - \mathcal{P}_{4n+3} + 3\mathcal{P}_{4n+2} - \mathcal{P}_{2n+1}$
11. $\mathcal{P}_{16n+15} = \mathcal{P}_{16n+3} + 2\mathcal{P}_{8n+7} - 3\mathcal{P}_{8n+6} - 2\mathcal{P}_{8n+3} + 6\mathcal{P}_{4n+2} - 3\mathcal{P}_{2n+1}$
12. $\mathcal{P}_{32n+3} = \mathcal{P}_{8n+3}$
13. $\mathcal{P}_{32n+19} = -\mathcal{P}_{16n+3} + \mathcal{P}_{8n+3} + 3\mathcal{P}_{8n+2} + 2\mathcal{P}_{4n+3} - 3\mathcal{P}_{4n+2} - \mathcal{P}_{2n+1}$

Proof. If we observe the right hand side of these 13 relations, we see that every sequence $\mathcal{P}_{2^q n+r}$ is a linear combination of $\mathcal{P}_{2n+1}$, $\mathcal{P}_{4n+2}$, $\mathcal{P}_{4n+3}$, $\mathcal{P}_{8n}$, $\mathcal{P}_{8n+2}$, $\mathcal{P}_{8n+3}$, $\mathcal{P}_{8n+6}$, $\mathcal{P}_{8n+7}$ and $\mathcal{P}_{16n+3}$.

To see that these 13 relations generate all sequences for the $\mathbb{Z}$-kernel, we just have to check that the left hand side of this 13 relations cover all residue classes modulo 32. Please note that we have to write all relations modulo 32, so e.g. instead of $\mathcal{P}_{16n+7}$ we write $\mathcal{P}_{16(2n)+7} = \mathcal{P}_{32n+7}$ and $\mathcal{P}_{16(2n+1)+7} = \mathcal{P}_{32n+23}$. □

These relations were found by computer experiments by Parreau, Rigo and Vandomme. An algorithm for finding such relations for an ℓ-abelian sequence is presented in Chapter 6 of [3]. We will prove the relations one by one by a four step approach.

In this four step approach Table 3.4 will be used in the second and third step[1]. Table 3.4 was generated in the same manner as Table 3.3 starting with all three possibilities (because of equation (3.4)) for the relative sizes of **pairs**$(n+1)$ and **pairs**$(n+2)$.

So let us now describe the four steps in detail, we will need them over and over again for all thirteen relations.

First step: We move all terms of a relation to the right hand side and replace them there using Lemma 3.5.5. We then divide the whole equation by 2 (as consequence of Lemma 3.5.5 all $\mathcal{P}_n$ are even for $n > 0$).

[1]For Relation 3 we need a different table.

Example 3.6.1 (Relation 1)**.** *To make the steps clearer, we will prove Relation 1 as example. In order to prove $\mathcal{P}_{4n+1} = \mathcal{P}_{2n+1}$, we move all terms to the right side.*

$$0 = -\mathcal{P}_{4n+1} + \mathcal{P}_{2n+1}$$

Now we use Lemma 3.5.5 to replace them and get

$$0 = -2(2\#\mathbf{PAIRS}(4n) - \#_2\mathbf{PAIRS}(4n)) + 2(2\#\mathbf{PAIRS}(2n) - \#_2\mathbf{PAIRS}(2n)).$$

Finally divide this equation by 2 to get

$$0 = -(2\#\mathbf{PAIRS}(4n) - \#_2\mathbf{PAIRS}(4n)) + (2\#\mathbf{PAIRS}(2n) - \#_2\mathbf{PAIRS}(2n)).$$

Second step: We use Table 3.4 to substitute the term $\#\mathbf{PAIRS}(.)$ by cardinalities of suitably chosen integer intervals (we keep $\#_2\mathbf{PAIRS}(.)$).

Since there are three cases, we will now have three equations. We calculate the cardinalities on the right hand side in all three cases, simplify until we have a single integer and put the result into a triple $(\mathrm{rhs}_1, \mathrm{rhs}_2, \mathrm{rhs}_3)$, where rhs_i is the integer on the right hand side in the i-th case.

Example 3.6.2 (Relation 1, continued)**.** *We start with*

$$0 = -(2\#\mathbf{PAIRS}(4n) - \#_2\mathbf{PAIRS}(4n)) + (2\#\mathbf{PAIRS}(2n) - \#_2\mathbf{PAIRS}(2n)).$$

We look at Case I of Table 3.4 and substitute $\#\mathbf{PAIRS}(2n) = \#[n-b, n-a]$ and $\#\mathbf{PAIRS}(4n) = \#[n+a, n+b]$ to get

$$0 = -2\#[n+a, n+b] + \#_2\mathbf{PAIRS}(4n) + 2\#[n-b, n-a] - \#_2\mathbf{PAIRS}(2n).$$

Now we calculate the cardinalities. Since $\#[a, b] = b - a + 1$, we get

$$0 = -2(b - a + 1) + \#_2\mathbf{PAIRS}(4n) + 2(b - a + 1) - \#_2\mathbf{PAIRS}(2n).$$

This can be simplified to

$$0 = \#_2\mathbf{PAIRS}(4n) - \#_2\mathbf{PAIRS}(2n)$$

so the integer in Case I. is 0. The other two cases are identical and we get

$$0 = \#_2\mathbf{PAIRS}(4n) - \#_2\mathbf{PAIRS}(2n) + (0, 0, 0).$$

This is a shorthand for three (identical) equations. Most times two or even all three cases will be identical.

Third step: We will use the Table 3.4 again to substitute the cardinalities of the sets $\#_2\mathbf{PAIRS}(.)$ with integer intervals in all three cases and intersect them if necessary.

Example 3.6.3 (Relation 1, continued)**.** *In all three cases we get*

$$0 = \#_2[n{+}a, n{+}b]{-}\#_2[n{-}b, n{-}a].$$

Fourth step: Now we deal with the number of even elements in integer intervals. If two intervals of the same size with different signs contain the same number of even elements, they cancel. If they do not cancel (because their borders have the wrong parity) or have different sizes, we split off Iverson brackets from the beginning or end of the larger set until both sets have the same parity of borders and the same cardinality

$$\#_2[a, b] = [\![a \text{ even}]\!] + \#_2[a{+}1, b]; \quad \#_2[a, b] = \#_2[a, b{-}1] + [\![b \text{ even}]\!].$$

In order to check that two intervals of the same size with different signs contain the same number of even elements, we use the following procedure, which we call *normalization*:

- Replace all even numbers by 0 and all odd numbers by 1.
- Change all "−" signs to "+" signs.
- If an a occurs in the upper border of a set, we change the upper and lower border.

It is easy to see that parity of the interval borders does not chance during the procedure, basically we calculate modulo 2. It does, however, change the size of the intervals so it is important to use normalization only on intervals of the same size.

Example 3.6.4. *We want to show that*

$$0 = -\#_2[5n + a + 1, 5n + b + 2] + \#_2[3n - b, 3n - a + 1].$$

Both sides have the same cardinality $b - a + 2$, therefore we can normalize them:

$$\begin{aligned} -\#_2[5n{+}a{+}1, 5n{+}b{+}2] + \#_2[3n{-}b, 3n{-}a{+}1] &= \\ = -\#_2[n{+}a{+}1, n{+}b] + \#_2[n{-}b, n{-}a{+}1] &= \\ = -\#_2[n{+}a{+}1, n{+}b] + \#_2[n{+}b, n{+}a{+}1] &= \\ = -\#_2[n{+}a{+}1, n{+}b] + \#_2[n{+}a{+}1, n{+}b] &= 0. \end{aligned}$$

Example 3.6.5 (Relation 1, continued)**.** *In all three cases we have*

$$0 = \#_2[n{+}a, n{+}b]{-}\#_2[n{-}b, n{-}a].$$

Since the sets have the same cardinality $b-a+1$, we normalize and get

$$0 = \#_2[n+a, n+b] - \#_2[n-b, n-a] =$$
$$= \#_2[n+a, n+b] - \#_2[n+b, n+a] =$$
$$= \#_2[n+a, n+b] - \#_2[n+a, n+b]$$

which is true.

The only non-mechanical step in the whole procedure is step four. If we have to split intervals before we can normalize, we demonstrate the whole calulation of step four. After these four steps we will have an equation which is trivially true. Since even the fourth step is straightforward, it should be possible to let a computer verify linear equations in $\mathcal{P}_n$ like the relations in Theorem 3.1.1.*

Proof of Theorem 3.1.1.

1. This is shown in the example.

2. We will show this relation in more detail. We want to prove

$$0 = -\mathcal{P}_{8n+4} + \mathcal{P}_{8n+3} + \mathcal{P}_{4n+3} - \mathcal{P}_{4n+2}.$$

In the first step we use Lemma 3.5.5 and divide by 2 to get

$$\begin{aligned} 0 =& -\#\mathbf{PAIRS}(8n+4) - \#\mathbf{PAIRS}(8n+2) + \#_2(\mathbf{PAIRS}(8n+4) \cap \mathbf{PAIRS}(8n+2)) \\ &+ 2\#\mathbf{PAIRS}(8n+2) - \#_2\mathbf{PAIRS}(8n+2) + 2\#\mathbf{PAIRS}(4n+2) - \#_2\mathbf{PAIRS}(4n+2) \\ &- \#\mathbf{PAIRS}(4n+2) - \#\mathbf{PAIRS}(4n) + \#_2(\mathbf{PAIRS}(4n+2) \cap \mathbf{PAIRS}(4n)). \end{aligned}$$

For the second step we substitute the intervals from Table 3.4. With the shorthand $\mathcal{C} := b-a+1$ we get in the first case (where $\mathbf{pairs}(n+1) = [a, b]$ and $\mathbf{pairs}(n+2) = [a, b]$)

$$\begin{aligned} 0 =& -2\mathcal{C} - 2 + \#_2(\mathbf{PAIRS}(8n+4) \cap \mathbf{PAIRS}(8n+2)) + 2\mathcal{C} + 2 - \#_2\mathbf{PAIRS}(8n+2) \\ &+ 2\mathcal{C} + 2 - \#_2\mathbf{PAIRS}(4n+2) - 2\mathcal{C} - 1 + \#_2(\mathbf{PAIRS}(4n+2) \cap \mathbf{PAIRS}(4n)). \end{aligned}$$

Most terms cancel and we have 1 remaining on the right hand side. It turns out this is also true in the two other cases, so our triple is $(1, 1, 1)$ and we get

$$\begin{aligned} 0 =& + \#_2(\mathbf{PAIRS}(8n+4) \cap \mathbf{PAIRS}(8n+2)) - \#_2\mathbf{PAIRS}(8n+2) \\ &- \#_2\mathbf{PAIRS}(4n+2) + \#_2(\mathbf{PAIRS}(4n+2) \cap \mathbf{PAIRS}(4n)) + (1, 1, 1). \end{aligned}$$

This is an abbreviation for three equations. Now we apply the third step, intersect the sets and get in all three cases

$$\begin{aligned}0 = &+ \#_2[3n{-}b{+}1, 3n{-}a{+}1] - \#_2[3n{-}b, 3n{-}a{+}1]\\ &- \#_2[n{+}a, n{+}b{+}1] + \#_2[n{+}a, n{+}b] + 1.\end{aligned}$$

In the fourth step we can not normalize, so we split off Iverson brackets and get

$$\begin{aligned}0 = &+ \#_2[3n{-}b{+}1, 3n{-}a{+}1] - [\![3n - b \text{ even}]\!] - \#_2[3n{-}b{+}1, 3n{-}a{+}1]\\ &- \#_2[n{+}a, n{+}b] - [\![n + b + 1 \text{ even}]\!] + \#_2[n{+}a, n{+}b] + 1.\end{aligned}$$

The intervals cancel and we are left with the true equation

$$0 = -[\![3n - b \text{ even}]\!] - [\![n + b + 1 \text{ even}]\!] + 1,$$

in all three cases.

3. We can get the values needed if we extend Table 3.3.

i	**PAIRS**$(8n{+}i)$	**PAIRS**$(16n{+}i)$
-2	$[3n{-}b{-}1, 3n{-}a]$	$[5n{+}a{-}1, 5n{+}b]$
0	$[3n{-}b, 3n{-}a]$	$[5n{+}a, 5n{+}b]$
-2	$[3n{-}b{-}1, 3n{-}a{-}1]$	$[5n{+}a, 5n{+}b]$
0	$[3n{-}b, 3n{-}a{-}1]$	$[5n{+}a{+}1, 5n{+}b]$
-2	$[3n{-}b{-}2, 3n{-}a]$	$[5n{+}a{-}1, 5n{+}b{+}1]$
0	$[3n{-}b{-}1, 3n{-}a]$	$[5n{+}a, 5n{+}b{+}1]$

The triple is $(0, 0, 0)$. The rest is a straightforward. In the fourth step we just have to normalize.

4. Similar to relation 3. but with values from Table 3.4. The triple is $(0, 0, 0)$ and we just have to normalize.

5. After the first two steps we get the triple $(0, 1, 0)$. Please note that the sets in Cases I. and III. are identical, so we can ignore Case III. In the first case we have

$$\begin{aligned}0 = &- 3\overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2} + 2\overbrace{\#_2[n{+}a, n{+}b]}^{4n+2} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1}\\ &- \overbrace{\#_2[n{+}a, n{+}b{+}1]}^{4n+3} + \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3}\\ &+ \overbrace{\#_2[5n{+}a{+}1, 5n{+}b{+}2]}^{16n+6} - \overbrace{\#_2[3n{-}b, 3n{-}a{+}1]}^{8n+3},\end{aligned}$$

Case	i	**pairs**$(n+i)$	**PAIRS**$(2n+i)$	**PAIRS**$(4n+i)$	**PAIRS**$(8n+i)$	**PAIRS**$(16n+i)$	**PAIRS**$(32n+i)$
	0		$[n-b, n-a]$	$[n+a, n+b]$	$[3n-b, 3n-a]$	$[5n+a, 5n+b]$	$[11n-b, 11n-a]$
	1	$[a, b]$					
	2	$[a, b]$	$[n-b+1, n-a+1]$	$[n+a, n+b+1]$	$[3n-b, 3n-a+1]$	$[5n+a, 5n+b+1]$	$[11n-b, 11n-a+1]$
	4			$[n+a+1, n+b+1]$	$[3n-b+1, 3n-a+2]$	$[5n+a+1, 5n+b+2]$	$[11n-b+1, 11n-a+2]$
	6				$[3n-b+2, 3n-a+3]$	$[5n+a+1, 5n+b+3]$	$[11n-b+1, 11n-a+3]$
I.	8				$[3n-b+3, 3n-a+3]$	$[5n+a+2, 5n+b+3]$	$[11n-b+2, 11n-a+3]$
	10					$[5n+a+2, 5n+b+4]$	$[11n-b+2, 11n-a+4]$
	12					$[5n+a+3, 5n+b+4]$	$[11n-b+3, 11n-a+5]$
	14					$[5n+a+4, 5n+b+5]$	$[11n-b+4, 11n-a+6]$
	16					$[5n+a+5, 5n+b+5]$	$[11n-b+5, 11n-a+6]$
	18						$[11n-b+5, 11n-a+7]$
	0		$[n-b, n-a]$	$[n+a, n+b]$	$[3n-b, 3n-a]$	$[5n+a, 5n+b]$	$[11n-b, 11n-a]$
	1	$[a, b]$					
	2	$[a+1, b]$	$[n-b+1, n-a]$	$[n+a+1, n+b+1]$	$[3n-b, 3n-a+1]$	$[5n+a, 5n+b+1]$	$[11n-b, 11n-a+1]$
	4			$[n+a+2, n+b+1]$	$[3n-b+1, 3n-a+1]$	$[5n+a+1, 5n+b+2]$	$[11n-b+1, 11n-a+2]$
	6				$[3n-b+2, 3n-a+2]$	$[5n+a+2, 5n+b+3]$	$[11n-b+1, 11n-a+3]$
II.	8				$[3n-b+3, 3n-a+2]$	$[5n+a+3, 5n+b+3]$	$[11n-b+2, 11n-a+3]$
	10					$[5n+a+3, 5n+b+4]$	$[11n-b+2, 11n-a+4]$
	12					$[5n+a+4, 5n+b+4]$	$[11n-b+3, 11n-a+4]$
	14					$[5n+a+5, 5n+b+5]$	$[11n-b+4, 11n-a+5]$
	16					$[5n+a+6, 5n+b+5]$	$[11n-b+5, 11n-a+5]$
	18						$[11n-b+5, 11n-a+6]$
	0		$[n-b, n-a]$	$[n+a, n+b]$	$[3n-b, 3n-a]$	$[5n+a, 5n+b]$	$[11n-b, 11n-a]$
	1	$[a, b]$					
	2	$[a, b+1]$	$[n-b, n-a+1]$	$[n+a, n+b+1]$	$[3n-b, 3n-a+1]$	$[5n+a, 5n+b+1]$	$[11n-b, 11n-a+1]$
	4			$[n+a+1, n+b+2]$	$[3n-b+1, 3n-a+2]$	$[5n+a+1, 5n+b+2]$	$[11n-b+1, 11n-a+2]$
	6				$[3n-b+1, 3n-a+3]$	$[5n+a+1, 5n+b+3]$	$[11n-b+1, 11n-a+3]$
III.	8				$[3n-b+2, 3n-a+3]$	$[5n+a+2, 5n+b+3]$	$[11n-b+2, 11n-a+3]$
	10					$[5n+a+2, 5n+b+4]$	$[11n-b+2, 11n-a+4]$
	12					$[5n+a+3, 5n+b+5]$	$[11n-b+3, 11n-a+5]$
	14					$[5n+a+4, 5n+b+6]$	$[11n-b+4, 11n-a+6]$
	16					$[5n+a+5, 5n+b+6]$	$[11n-b+5, 11n-a+6]$
	18						$[11n-b+5, 11n-a+7]$

Table 3.4: Values for the proof of Theorem 3.1.1.

where all terms in a single line normalize to zero.

In the second case we have

$$\begin{aligned}0 = &+\overbrace{\#_2[5n{+}a{+}2, 5n{+}b{+}2]}^{16n+6} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1} - 2\overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2}\\ &- \overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2} + 2\overbrace{\#_2[n{+}a{+}1, n{+}b]}^{4n+2} - \overbrace{\#_2[3n{-}b, 3n{-}a{+}1]}^{8n+3}\\ &- \overbrace{\#_2[n{+}a{+}1, n{+}b{+}1]}^{4n+3} + \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3} + 1,\end{aligned}$$

where all terms in the first line normalize to zero. We then split off Iverson brackets to get

$$\begin{aligned}0 = &- \#_2[3n{-}b, 3n{-}a] + \#_2[5n{+}a, 5n{+}b] + [\![5n{+}b{+}1 \text{ even}]\!]\\ &- \#_2[n{+}a{+}1, n{+}b] - [\![n{+}b{+}1 \text{ even}]\!] + \#_2[n{+}a{+}1, n{+}b]\\ &- \#_2[3n{-}b, 3n{-}a{+}1] + \#_2[n{+}a{+}1, n{+}b] + 1.\end{aligned}$$

The first two lines cancel. We split the first interval two times and get

$$-\#_2[3n{-}b, 3n{-}a{-}1] - [\![3n{-}a \text{ even}]\!] - [\![3n{-}a{+}1 \text{ even}]\!] + \#_2[n{+}a{+}1, n{+}b] + 1,$$

and finally

$$0 = -[\![3n{-}a \text{ even}]\!] - [\![3n{-}a{+}1 \text{ even}]\!] + 1.$$

6. We already know that Relation 5. is true so we subtract it from Relation 6. to get:

$$\mathcal{P}_{16n+7} - \mathcal{P}_{16n+6} = \mathcal{P}_{4n+3} - \mathcal{P}_{4n+2}.$$

Now we follow the usual procedure and get the triple $(0, 0, 0,)$. Cases I. and III. are identical and we get

$$\begin{aligned}0 = &+\overbrace{\#_2[5n{+}a{+}1, 5n{+}b{+}3]}^{16n+7} - \overbrace{\#_2[5n{+}a{+}1, 5n{+}b{+}2]}^{16n+6}\\ &- \overbrace{\#_2[n{+}a, n{+}b{+}1]}^{4n+3} + \overbrace{\#_2[n{+}a, n{+}b]}^{4n+2}\end{aligned}$$

which we split to

$$\begin{aligned}0 = &+ \#_2[5n{+}a{+}1, 5n{+}b{+}2] + [\![5n{+}b{+}3 \text{ even}]\!] - \#_2[5n{+}a{+}1, 5n{+}b{+}2]\\ &- \#_2[n{+}a, n{+}b] - [\![n{+}b{+}1 \text{ even}]\!] + \#_2[n{+}a, n{+}b].\end{aligned}$$

So we have the true equation

$$0 = [\![5n{+}b{+}3 \text{ even}]\!] - [\![n{+}b{+}1 \text{ even}]\!].$$

In Case II. we have

$$0 = + \overbrace{\#_2[5n{+}a{+}2, 5n{+}b{+}3]}^{16n+7} - \overbrace{\#_2[5n{+}a{+}2, 5n{+}b{+}2]}^{16n+6} - \overbrace{\#_2[n{+}a{+}1, n{+}b{+}1]}^{4n+3} + \overbrace{\#_2[n{+}a{+}1, n{+}b]}^{4n+2}.$$

We split to the true equation

$$0 = + \#_2[5n{+}a{+}2, 5n{+}b{+}2] + [\![5n{+}b{+}3 \text{ even}]\!] - \#_2[5n{+}a{+}2, 5n{+}b{+}2] - \#_2[n{+}a{+}1, n{+}b] - [\![n{+}b{+}1 \text{ even}]\!] + \#_2[n{+}a{+}1, n{+}b].$$

7. Straightforward. The triple is $(0, 0, 0,)$ and we can normalize to zero.

8. We subtract Relation 7. from Relation 8. we get $\mathcal{P}_{16n+10} - \mathcal{P}_{16n+8} = 0$. Again the triple is $(0, 0, 0,)$ and we can normalize to zero.

9. If we subtract Relation 8. from Relation 9. we get:

$$\mathcal{P}_{16n+11} - \mathcal{P}_{16n+10} = -\mathcal{P}_{16n+3} + 2\mathcal{P}_{8n+2} - \mathcal{P}_{2n+1}.$$

Our triple is $(-1, -1, -1)$. In the identical Cases I. and III. we get

$$0 = + \overbrace{\#_2[5n{+}a{+}2, 5n{+}b{+}4]}^{16n+11} + \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3} - \overbrace{\#_2[5n{+}a{+}2, 5n{+}b{+}3]}^{16n+10} - 2\overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1} - 1.$$

After splitting off the upper borders in the first line we have

$$0 = [\![5n{+}b{+}4 \text{ even}]\!] + [\![5n{+}b{+}1 \text{ even}]\!] - 1$$

as final result.

In the second case we start with

$$0 = + \overbrace{\#_2[5n{+}a{+}3, 5n{+}b{+}4]}^{16n+11} + \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3} - \overbrace{\#_2[5n{+}a{+}3, 5n{+}b{+}3]}^{16n+10} - 2\overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1} - 1.$$

and split off the upper borders in the first line. We get the same final result

$$0 = [\![5n{+}b{+}4 \text{ even}]\!] + [\![5n{+}b{+}1 \text{ even}]\!] - 1.$$

10. The triple is $(0,-1,0)$. The rest of the calculation is lengthy, since for the first time all three cases are different, but not too hard. We have to split in the second and third case. In the second case we have

$$\begin{aligned}0 =& \overbrace{\#_2[5n{+}a{+}5, 5n{+}b{+}4]}^{16n+14} - \overbrace{\#_2[n{+}a{+}1, n{+}b]}^{4n+2}\\ &- \overbrace{\#_2[3n{-}b{+}2, 3n{-}a{+}2]}^{8n+7} + \overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2}\\ &+ \overbrace{\#_2[3n{-}b, 3n{-}a{+}1]}^{8n+3} - \overbrace{\#_2[n{+}a{+}1, n{+}b]}^{4n+2}\\ &- \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1}\\ &+ \overbrace{\#_2[n{+}a{+}1, n{+}b{+}1]}^{4n+3} - \overbrace{\#_2[n{+}a{+}1, n{+}b]}^{4n+2} - 1.\end{aligned}$$

The first two lines normalize to zero and we split to get

$$\begin{aligned}0 =& + \#_2[3n{-}b, 3n{-}a{-}1] + [\![3n{-}a \text{ even}]\!] + [\![3n{-}a{+}1 \text{ even}]\!] - \#_2[n{+}a{+}1, n{+}b]\\ &- \#_2[5n{+}a, 5n{+}b] - [\![5n{+}b{+}1 \text{ even}]\!] + \#_2[n{-}b, n{-}a]\\ &+ \#_2[n{+}a{+}1, n{+}b] + [\![n{+}b{+}1 \text{ even}]\!] - \#_2[n{+}a{+}1, n{+}b] - 1.\end{aligned}$$

Finally we get

$$0 = [\![3n{-}a \text{ even}]\!] + [\![3n{-}a{+}1 \text{ even}]\!] - 1.$$

In the third case we have

$$\begin{aligned}0 =& \overbrace{\#_2[5n{+}a{+}4, 5n{+}b{+}5]}^{16n+14} - \overbrace{\#_2[5n{+}a, 5n{+}b{+}1]}^{16n+3}\\ &+ \overbrace{\#_2[3n{-}b, 3n{-}a]}^{8n+2} - 2\overbrace{\#_2[n{+}a, n{+}b]}^{4n+2} + \overbrace{\#_2[n{-}b, n{-}a]}^{2n+1}\\ &- \overbrace{\#_2[3n{-}b{+}1, 3n{-}a{+}3]}^{8n+7} - \overbrace{\#_2[n{+}a, n{+}b]}^{4n+2}\\ &+ \overbrace{\#_2[n{+}a, n{+}b{+}1]}^{4n+3} + \overbrace{\#_2[3n{-}b, 3n{-}a{+}1]}^{8n+3}.\end{aligned}$$

The first two lines normalize to zero and we split to get

$$\begin{aligned}0 =& - \#_2[3n{-}b{+}1, 3n{-}a{+}1] - [\![3n{-}a{+}2 \text{ even}]\!] - [\![3n{-}a{+}3 \text{ even}]\!] - \#_2[n{+}a, n{+}b]\\ &+ \#_2[n{+}a, n{+}b] + [\![n{+}b{+}1 \text{ even}]\!] + [\![3n{-}b \text{ even}]\!] + \#_2[3n{-}b{+}1, 3n{-}a{+}1].\end{aligned}$$

This normalizes to the true equation

$$0 = -[\![3n{-}a{+}2 \text{ even}]\!] - [\![3n{-}a{+}3 \text{ even}]\!] + [\![n{+}b{+}1 \text{ even}]\!] + [\![3n{-}b \text{ even}]\!].$$

11. Another long calculation. Our triple is $(0,-2,-1)$. We have to split in the second and third case. For the second case we get

$$\begin{aligned}0 =& -2\overbrace{\#_2[3n{-}b{+}2,3n{-}a{+}2]}^{8n+7}+2\overbrace{\#_2[n{-}b,n{-}a]}^{2n+1}\\ &+3\overbrace{\#_2[3n{-}b{+}2,3n{-}a{+}1]}^{8n+6}-3\overbrace{\#_2[n{+}a{+}1,n{+}b]}^{4n+2}\\ &+2\overbrace{\#_2[3n{-}b,3n{-}a{+}1]}^{8n+3}-3\overbrace{\#_2[n{+}a{+}1,n{+}b]}^{4n+2}+\overbrace{\#_2[5n{+}a{+}5,5n{+}b{+}5]}^{16n+15}\\ &+\overbrace{\#_2[n{-}b,n{-}a]}^{2n+1}-\overbrace{\#_2[5n{+}a,5n{+}b{+}1]}^{16n+3}-2.\end{aligned}$$

The first two lines cancel and we can start to split intervals.

$$\begin{aligned}0 =& \#_2[5n{+}a{+}5,5n{+}b{+}5]-[\![5n{+}a \text{ even}]\!]-\#_2[5n{+}a{+}1,5n{+}b{+}1]\\ &+\#_2[n{-}b,n{-}a{-}1]+[\![n{-}a \text{ even}]\!]-\#_2[n{+}a{+}1,n{+}b]\\ &+2\#_2[3n{-}b,3n{-}a{+}1]-2\#_2[n{+}a{+}1;n{+}b]-2.\end{aligned}$$

The first two lines cancel and after splitting the first interval in the last line twice we get the true equation

$$0 = +2[\![3n{-}a{+}1 \text{ even}]\!]+2[\![3n{-}a \text{ even}]\!]-2.$$

In the third case we have

$$\begin{aligned}0 =& +3\overbrace{\#_2[n{-}b,n{-}a]}^{2n+1}-3\overbrace{\#_2[n{+}a,n{+}b]}^{4n+2}\\ &+\overbrace{\#_2[3n{-}b{+}1,3n{-}a{+}2]}^{8n+6}-\overbrace{\#_2[5n{+}a,5n{+}b{+}1]}^{16n+3}\\ &+2\overbrace{\#_2[3n{-}b{+}1,3n{-}a{+}2]}^{8n+6}-2\overbrace{\#_2[3n{-}b{+}1,3n{-}a{+}3]}^{8n+7}\\ &+2\overbrace{\#_2[3n{-}b,3n{-}a{+}1]}^{8n+3}-2\overbrace{\#_2[n{+}a,n{+}b]}^{4n+2}\\ &+\overbrace{\#_2[n{+}a,n{+}b]}^{4n+2}+\overbrace{\#_2[5n{+}a{+}4,5n{+}b{+}6]}^{16n+15}-1.\end{aligned}$$

The first two lines cancel and after splitting we get

$$\begin{aligned}0 =& +2\#_2[3n{-}b{+}1,3n{-}a{+}2]-2\#_2[3n{-}b{+}1,3n{-}a{+}2]-2[\![3n{-}a{+}3 \text{ even}]\!]\\ &+2\#_2[3n{-}b,3n{-}a]+2[\![3n{-}a{+}1 \text{ even}]\!]-2\#_2[n{+}a,n{+}b]\\ &+\#_2[n{+}a,n{+}b]+\#_2[5n{+}a{+}4,5n{+}b{+}4]+[\![5n{+}b{+}5 \text{ even}]\!]+[\![5n{+}b{+}6 \text{ even}]\!]-1\end{aligned}$$

which reduces to

$$0 = [\![5n{+}b{+}5 \text{ even}]\!]+[\![5n{+}b{+}6 \text{ even}]\!]-1.$$

12. Straightforward. The triple is $(0,0,0)$ and we can normalize to zero.

13. We subtract Relation 6. from Relation 13. to get

$$\mathcal{P}_{32n+19} - \mathcal{P}_{16n+7} = 0.$$

The triple is $(0,0,0)$. We normalize to zero.

□

3.7 Properties of $\mathcal{P}_n$

In this section we show three additional properties of the 2-abelian complexity $\mathcal{P}_n$ (cf. page 21). In the first lemma we show that $\mathcal{P}_n$ changes in steps of 2 (for $n > 0$).

Lemma 3.7.1. *For $n \geq 4$ we have*

$$\mathcal{P}_{n+1} - \mathcal{P}_n \in \{-2, 0, 2\}.$$

Proof. Due to equation (3.4) we know that there are three possibilities for the relative sizes of $\mathbf{PAIRS}(2n-2)$ and $\mathbf{PAIRS}(2n)$:

1. $\begin{matrix}[a,b]\\[a+1,b]\end{matrix}$ 2. $\begin{matrix}[a,b]\\[a,b+1]\end{matrix}$ 3. $\begin{matrix}[a,b]\\[a+1,b+1]\end{matrix}$

We use Lemma 3.5.5 and with $\mathcal{C} := \#(\mathbf{PAIRS}(2n) \cap \mathbf{PAIRS}(2n-2))$ and $\mathcal{E} := \#_2(\mathbf{PAIRS}(2n) \cap \mathbf{PAIRS}(2n-2))$ we get

Case	$\mathcal{P}_{2n-1}$	$\mathcal{P}_{2n}$	$\mathcal{P}_{2n+1}$
1.	$2(2(\mathcal{C}{+}1){-}\mathcal{E}{-}[\![a \text{ even}]\!])$	$2((\mathcal{C}{+}1){+}\mathcal{C}{-}\mathcal{E})$	$2(2\mathcal{C}{-}\mathcal{E})$
2.	$2(2\mathcal{C}{-}\mathcal{E})$	$2((\mathcal{C}{+}1){+}\mathcal{C}{-}\mathcal{E})$	$2(2(\mathcal{C}{+}1){-}\mathcal{E}{-}[\![b{+}1 \text{ even}]\!])$
3.	$2(2(\mathcal{C}{+}1){-}\mathcal{E}{-}[\![a \text{ even}]\!])$	$2((\mathcal{C}{+}1){+}(\mathcal{C}{+}1){-}\mathcal{E})$	$2(2(\mathcal{C}{+}1){-}\mathcal{E}{-}[\![b{+}1 \text{ even}]\!])$

In all three cases, regardless if a and $b+1$ are even or odd, Lemma 3.7.1 is true. □

Now we want to prove that the sequence $\mathcal{P}_n$ is unbounded.

Lemma 3.7.2. *If* $\mathbf{pairs}(n) = [a,b]$, $\mathbf{pairs}(n+1) = [a,b+1]$ *with n and b odd and a even and then the sequence $a_0 = n$, $a_{i+1} = 16a_i - 5$ satisfies*

$$\mathcal{P}_{a_n} = \mathcal{P}_{a_0} + 6n.$$

Proof. If we start with two sets $\mathbf{pairs}(n) = [a, b]$ and $\mathbf{pairs}(n+1) = [a, b+1]$, we can use Theorem 3.5.2 to get $\mathbf{pairs}(16n-5) = [5n+a-3, 5n+b-1]$ and $\mathbf{pairs}(16n-4) = [5n+a-3, 5n+b]$.

So if we start in Case III. with $\mathbf{pairs}(n)$ and $\mathbf{pairs}(n+1)$, we will be in Case III. with $\mathbf{pairs}(16n-5)$ and $\mathbf{pairs}(16n-4)$ again. Therefore, we can repeat the whole process.

Furthermore, if n and b are odd and a is even, then $16n-5$ and $5n+b-1$ are odd and $5n+a-3$ is even. Hence the sets

$$\#\mathbf{pairs}(n) = \#[a, b] \text{ and } \#\mathbf{pairs}(16n-5) = \#[a', b']$$

contain an even number of elements and we have:

$$\#[a, b] = 2s, \;\; \#[a', b'] = 2s+2, \;\; \#_2[a, b] = s, \;\; \#_2[a', b'] = s+1.$$

Since the values n and $16n-5$ are both odd, we know from Theorem 3.5.2 that $\mathbf{PAIRS}(n-1) = \mathbf{pairs}(n)$ and $\mathbf{PAIRS}(16n-6) = \mathbf{pairs}(16n-5)$. Now we apply Lemma 3.5.5 and get $\mathcal{P}_n = 6s$ and $\mathcal{P}_{16n-5} = 6s+6$. □

The sequence $\mathcal{P}_3 = 6, \mathcal{P}_{43} = 12, \mathcal{P}_{683} = 18, \ldots$ is one example of such an unbounded sequence. In [41] they show $\mathcal{P}_t^{(2)}((2 \cdot 4^m + 4)/3) = \Theta(m)$ which also proves that the sequence $\mathcal{P}_n$ is unbounded.

The last theorem reveals some symmetries of $\mathcal{P}_n$.

Case	i	$\mathbf{pairs}(n+i)$		$\mathbf{pairs}(2n+i)$	$\#\mathbf{pairs}(n+i)$		$\#\mathbf{pairs}(2n+i)$
	-1			$[n-b-1, n-a-1]$			$\mathcal{C}$
I.	0	$[a, b]$	$\longmapsto$	$[n-b-1, n-a]$	$\mathcal{C}$	$\longmapsto$	$\mathcal{C}+1$
	$+1$	$[a, b]$		$[n-b, n-a]$	$\mathcal{C}$		$\mathcal{C}$
	-1			$[n-b-1, n-a-1]$			$\mathcal{C}+1$
II.	0	$[a, b]$	$\longmapsto$	$[n-b-1, n-a-1]$	$\mathcal{C}+1$	$\longmapsto$	$\mathcal{C}+1$
	$+1$	$[a+1, b]$		$[n-b, n-a-1]$	$\mathcal{C}$		$\mathcal{C}$
	-1			$[n-b-1, n-a-1]$			$\mathcal{C}$
III.	0	$[a, b]$	$\longmapsto$	$[n-b-1, n-a]$	$\mathcal{C}$	$\longmapsto$	$\mathcal{C}+1$
	$+1$	$[a, b+1]$		$[n-b-1, n-a]$	$\mathcal{C}+1$		$\mathcal{C}+1$

Table 3.5: The action of Theorem 3.5.2 on $\mathbf{pairs}(n)$ and $\mathbf{pairs}(n+1)$

Theorem 3.7.3. *The 2-abelian complexity of the Thue–Morse word is a concatenation of palindromes of increasing size since the sequence*

$$\mathcal{P}_{2^q+1}\mathcal{P}_{2^q+2}\mathcal{P}_{2^q+3}\cdots\mathcal{P}_{2^{q+1}+1}$$

is a palindrome, or equivalently

$$\mathcal{P}_{2^q+1+i} = \mathcal{P}_{2^{q+1}+1-i}$$

for $0 \leq i \leq 2^{q-1}$.

Proof. We will prove the theorem with three inductions. With the first induction, we show that the sequence

$$\#\mathbf{pairs}(2^q+1), \#\mathbf{pairs}(2^q+2), \ldots, \#\mathbf{pairs}(2^{q+1}+1)$$

is a palindrome. The base case is

$$\#\mathbf{pairs}(3) = 2, \#\mathbf{pairs}(4) = 3, \#\mathbf{pairs}(5) = 2.$$

Two sets $\mathbf{pairs}(2^q+1+i)$ and $\mathbf{pairs}(2^{q+1}+1-i)$ with $0 \leq i \leq 2^{q-1}$ are called *corresponding sets*. If a consecutive pair of sets is mapped to a consecutive triple of sets

$$\begin{array}{lll} & & \mathbf{pairs}(2^{q+1}+1+2i) \\ \mathbf{pairs}(2^q+1+i) & \mapsto & \mathbf{pairs}(2^{q+1}+2+2i) \\ \mathbf{pairs}(2^q+1+(i+1)) & & \mathbf{pairs}(2^{q+1}+3+2i) \end{array}$$

with $1 \leq i \leq 2^{q-1}$, then the corresponding pair of consecutive sets is mapped to a consecutive triple of sets

$$\begin{array}{lll} & & \mathbf{pairs}(2^{q+2}+1-(2i+2)) \\ \mathbf{pairs}(2^{q+1}+1-(i+1)) & \mapsto & \mathbf{pairs}(2^{q+2}+2-(2i+2)) \\ \mathbf{pairs}(2^{q+1}+1-i) & & \mathbf{pairs}(2^{q+2}+3-(2i+2)). \end{array}$$

As next step we calculate Table 3.5. This is a simplified extract of Table 3.3 where we write $\mathcal{C}$ for $\#[a,b] = b-a+1$.

Now we look at Table 3.5 and see that it suffices to know the relative sizes of consecutive pairs of sets to determine in which case we are. So if a consecutive pair of sets is mapped to a consecutive triple of sets via Case II., the corresponding consecutive pair of sets is mapped to a corresponding consecutive triple of sets via Case III. and vise versa. If we have a Case I. map for the consecutive pair of sets, we also have a Case I. map for the corresponding consecutive pair of sets. In all three cases the palindromic structure of the set cardinality is preserved.

Now we show that

$$\#_2\mathbf{pairs}(2^q+1), \#_2\mathbf{pairs}(2^q+2), \ldots, \#_2\mathbf{pairs}(2^{q+1}+1)$$

is a palindrome too. We do this by showing that for two corresponding sets $\mathbf{pairs}(2^q+1+i) = [a,b]$ and $\mathbf{pairs}(2^{q+1}+1-i) = [a',b']$ we have

$$a \equiv b' \pmod 2 \text{ and } a' \equiv b \pmod 2. \tag{3.5}$$

Since two corresponding sets have the same cardinality, we can conclude that $\#_2[a,b] = \#_2[a',b']$.

With the second induction we show that equation (3.5) is true for

$$\mathbf{pairs}(2^q+1) = [a,b] \text{ and } \mathbf{pairs}(2^{q+1}+1) = [a',b'].$$

We use Theorem 3.5.2 to get

$$[a',b'] = \mathbf{PAIRS}(2^{q+1}+1) = 2^q - \mathbf{pairs}(2^q+1) = [2^q-b, 2^q-a]$$

and obviously

$$a \equiv 2^q - a \pmod 2 \text{ and } 2^q - b \equiv b \pmod 2.$$

We use a third induction to go from $\mathbf{pairs}(2^q+1+i)$ and $\mathbf{pairs}(2^{q+1}+1-i)$ to $\mathbf{pairs}(2^q+1+(i+1))$ and $\mathbf{pairs}(2^{q+1}+1-(i+1))$. We have to check the three cases from Table 3.5 again:

In the first case nothing changes, we go from

$$\mathbf{pairs}(2^q+1+i) = [a,b] \text{ and } \mathbf{pairs}(2^{q+1}+1-i) = [a',b']$$

to

$$\mathbf{pairs}(2^q+1+(i+1)) = [a,b] \text{ and } \mathbf{pairs}(2^{q+1}+1-(i+1)) = [a',b']$$

and Equation 3.5 is trivially fulfilled.

In the second case we go from

$$\mathbf{pairs}(2^q+1+i) = [a,b] \text{ and } \mathbf{pairs}(2^{q+1}+1-i) = [a',b']$$

to

$$\mathbf{pairs}(2^q+1+(i+1)) = [a+1,b] \text{ and } \mathbf{pairs}(2^{q+1}+1-(i+1)) = [a',b'-1]$$

and Equation 3.5 is fulfilled again.

In the third case the step is from

$$\mathbf{pairs}(2^q+1+i) = [a,b] \text{ and } \mathbf{pairs}(2^{q+1}+1-i) = [a',b']$$

to

$$\mathbf{pairs}(2^q+1+i) = [a,b+1] \text{ and } \mathbf{pairs}(2^{q+1}+1-i) = [a'-1,b']$$

and Equation 3.5 holds.

Since

$$\#\mathbf{pairs}(2^q+1+i) = \#\mathbf{pairs}(2^{q+1}+1-i)$$

and also

$$\#_2\mathbf{pairs}(2^q + 1 + i) = \#_2\mathbf{pairs}(2^{q+1} + 1 - i)$$

with $0 \leq i \leq 2^{q-1}$, we can use Theorem 3.5.2 to get

$$\#\mathbf{PAIRS}(2^q + 2i) = \#\mathbf{PAIRS}(2^{q+1} - 2i)$$

and also

$$\#_2\mathbf{PAIRS}(2^q + 2i) = \#_2\mathbf{PAIRS}(2^{q+1} - 2i)$$

with $0 \leq i \leq 2^{q-2}$. Now we use Lemma 3.5.5 and get $\mathcal{P}_{2^q+1+i} = \mathcal{P}_{2^{q+1}+1-i}$ with $0 \leq i \leq 2^{q-1}$. □

3.8 Outlook

In this chapter we developed a series of tools to capture the structure of the 2-abelian complexity sequence $\mathcal{P}_n$. These tools are custom made for the sequence $\mathcal{P}_n$ and rely heavily on the structure of Thue–Morse sequence $\mathbf{t}$.

On one hand, this means that the methods of this chapter cannot be adapted that easily for other sequences. The most obvious generalization is to the 2^q-abelian complexity sequences $\mathcal{P}_t^{(2^q)}(n)$ of $\mathbf{t}$ since we can still use the same alternation properties as in the short coding.

On the other hand, I believe that the essence of the 2-abelian complexity sequence $\mathcal{P}_n$ is described by Theorem 3.5.2 and Lemma 3.5.5. They should suffice to prove most statements about $\mathcal{P}_n$, of which we can see a glimpse in Section 3.7.

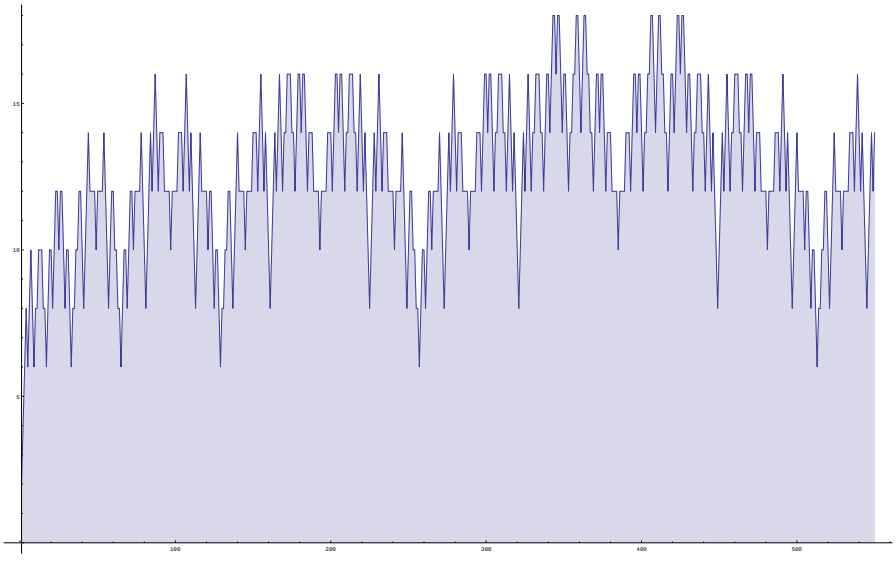

Figure 3.1: The sequence $\mathcal{P}_n$

Chapter 4

Spatial equidistribution of combinatorial number schemes

4.1 Introduction

Let us start with an investigation of the binomial coefficients, as well as the Stirling numbers of the first and second kind. If we color them according to their residue classes modulo a prime p, they have a self-similar, fractal structure, which we can observe in Figure 4.1. The most prominent member of this class of fractals is the Sierpinski triangle, which we obtain if we color the binomial coefficients according to their residue class modulo 2.

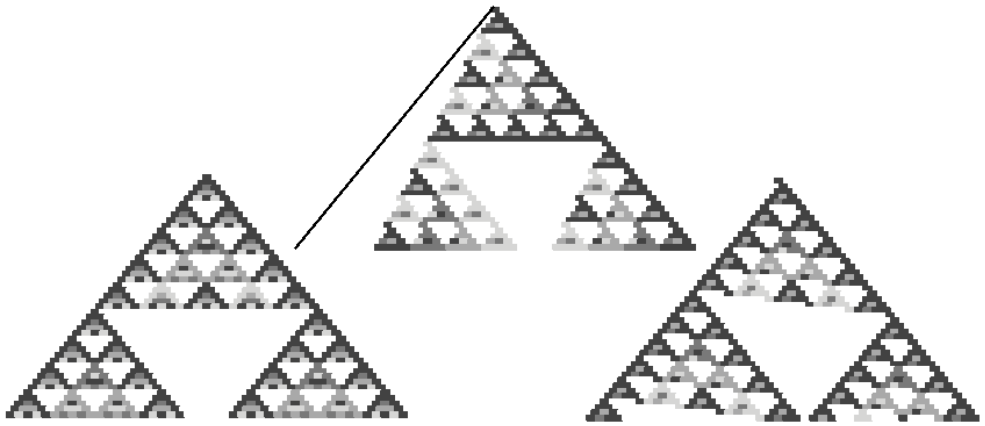

Figure 4.1: Binomial coefficients, Stirling numbers of the first and second kind modulo 5

The binomial coefficients and Stirling numbers of the first and second kind share some properties; they have a combinatorial interpretation, they have a recursive

structure and they can be easily expressed by generating functions.

We will focus on two further common properties:

The p-divisibility property For a fixed prime p, almost all binomial coefficients, Stirling numbers of the first and the second kind are divisible by p.

The equidistribution property The binomial coefficients, Stirling numbers of the first and second kind in the nonzero residue classes modulo a prime p approach an equidistribution in the first n lines, if n goes to infinity.

As main result of this chapter, we will prove Theorem 4.2.4 and Theorem 4.2.5 which are generalizations of these two properties.

The first property is very well studied. Regarding the binomial coefficients, Fine [30] gave a formula for the number of zeros in a row modulo p. Carlitz proved the corresponding results for columns [18]. Singmaster showed in [51] that "any integer divides almost all binomial coefficients" with respect to four different definitions of "almost all". He also gives a nice survey of known results in [52]. An approach via cellular automata is described in [54]. This approach may be extended since in [4] and [5] it is shown that the binomial coefficients, unsigned Stirling numbers of the first kind modulo a prime p and the Gaussian q-binomials modulo m (if $\gcd(m, q) = 1$) are automatic sequences. The ideas of this chapter can be expressed in terms of automatic sequences (cf. [2]).

The p-divisibility of the Stirling numbers of the first kind has been studied by Carlitz [16] and by Peele, Radcliffe and Wilf [48]. For the Stirling numbers of the second kind there are results by Carlitz [17] and Lundell [43]. A result for the q-binomial coefficients is given in Howard [36].

In this chapter we will use a unified approach and show that a prime p divides almost all numbers in any number scheme which satisfies the generalized Lucas' congruence. All previous examples turn out to share this property given below.

The equidistribution property is not so well studied. There are results on the equidistribution of the binomial coefficients from Garfield and Wilf [31] and by Barbolosi and Grabner [9] (with a generalization in [8]).

The original motivation of this chapter was to extend this results to Stirling numbers of the first and second kind. We show that in all number schemes which satisfy the generalized Lucas' congruence the nonzero residue classes modulo p are spatially equidistributed. This proves the equidistribution property for Stirling numbers and also generalizes the result for binomial coefficients.

The classic Lucas' congruence for binomial coefficients is stated in the following Theorem by Lucas' (cf. [25], p.271).

Theorem 4.1.1 (Èdouard Lucas 1878)**.** *The binomial coefficients satisfy*

$$\binom{n}{k} \equiv \prod_{i=0}^{\ell} \binom{n_i}{k_i} \pmod{p}$$

with $n = \sum_{i=0}^{\ell} n_i p^i$ *and* $k = \sum_{i=0}^{\ell} k_i p^i$ *where* $0 \leq n_i, k_i \leq p-1$.

Lucas' theorem allows us to calculate the binomial coefficients modulo p digit wise if we write n and k in base p. We generalize this idea in the following way.

Instead of p-adic representatives, we use matrix digital systems. They allow us to define an iterated function system and a sequence U_N of sets which converge to a limit fractal. Then we assign "colors", their residue classes modulo p, to the points in U_N.

The primary idea of this chapter is to define a *generalized Lucas' congruence* (cf. Definition 4.2.2) for a number scheme. A number scheme is a matrix digital system together with coloring functions that assign residue classes to numbers and digits. A number scheme satisfies the generalized Lucas' congruence if the residue class of a number $\vec{n}$ modulo p is the product of the residue classes of the digits of $\vec{n}$.

The binomial coefficients and Stirling numbers of the first and second kind all have number schemes which satisfy the generalized Lucas' congruence. We will show that for all number schemes which satisfy the generalized Lucas' congruence the sequence U_N is p-divisible and has a generalized equidistribution property. We can even replace the equidistribution property with a stronger property and show that we have an equidistribution modulo p for all μ-continuity sets of a normalized Hausdorff measure μ. The chapter is organized in the following way:

In Section 4.2 we define *matrix digital systems* and the *generalized Lucas' congruence*. Then we state our two main results, which are stronger, formalized versions of the two properties stated in this section.

We use Section 4.3 to point out the fractal structure of number schemes which satisfy the generalized Lucas' congruence.

Then we recall properties of Dirichlet characters in Section 4.4 and use them to prove the two main results.

Finally we show in Section 4.5 that our results are applicable to a range of well known number schemes before we finish with some concluding remarks in Section 4.6.

4.2 Results

First, we define *matrix digital systems* (based on *matrix number systems* in [42]). To avoid confusion between 2-dimensional vectors and binomial coefficients we

write vectors with an arrow above them. So $\overrightarrow{\binom{5}{3}}$ denotes a vector while $\binom{5}{3}$ denotes a binomial coefficient.

For every dimension $d \geq 1$ we have the d-dimensional Euclidean space $\mathbb{R}^d$ and the embedded ring of integer vectors $\mathbb{Z}^d$. Let $A \in \mathbb{Z}^{d\times d}$ be a $d \times d$ matrix whose eigenvalues all have modulus greater then 1. Then $\mathcal{L} = A\mathbb{Z}^d$ is a subgroup of $\mathbb{Z}^d$ and the factor group $\mathbb{Z}^d/A\mathbb{Z}^d$ has order $\mathrm{ord}(\mathbb{Z}^d/A\mathbb{Z}^d) = |\det A| > 1$.

Definition 4.2.1. *Let the* digit set $\mathcal{D} \subseteq \mathbb{Z}^d$ *be a complete residue system mod* $\mathcal{L}$ *with* $\overrightarrow{0} \in \mathcal{D}$. *We call the pair* $(A, \mathcal{D})$ *a* matrix digital system.

We define the set T *of all vectors* $\vec{n} \in \mathbb{Z}^d$ *with a representation of the form*

$$\vec{n} = \vec{\varepsilon_0} + A\vec{\varepsilon_1} + A^2\vec{\varepsilon_2} + \cdots + A^\ell\vec{\varepsilon_\ell} \tag{4.1}$$

as

$$T := \left\{ \sum_{k=0}^{\ell} A^k\vec{\varepsilon_k} \;\middle|\; \ell \in \mathbb{N},\ \vec{\varepsilon_k} \in \mathcal{D} \right\} \subseteq \mathbb{Z}^d.$$

Since $\mathcal{D}$ *is a complete residue system, the representation of the form* (4.1) *is unique. The set of vectors having a representation of the form* (4.1) *with* $\ell + 1$ *digits is denoted by* T_ℓ. *We write* $\vec{n}$ *also as* $\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\ldots\vec{\varepsilon_\ell})$.

If $T = \mathbb{Z}^d$, we have a *matrix number system* in the sense of [42]. We are not interested in whether all $\vec{n} \in \mathbb{Z}^d$ have such a representation.

Definition 4.2.2. *A function* $f : T \to \mathbb{Z}/p\mathbb{Z}$ *satisfies the* generalized Lucas' congruence *modulo a prime* p *if there is a function* f *with* $f : \mathcal{D} \to \mathbb{Z}/p\mathbb{Z}$ *and a matrix digital system (with* $\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\ldots\vec{\varepsilon_\ell})$*) so that*

$$f(\vec{n}) \equiv \prod_{i=0}^{\ell} f(\vec{\varepsilon_i}) \pmod{p}. \tag{4.2}$$

This implies that $f(\overrightarrow{0}) = 1$. *We say that* f *is a* coloring *and the elements of* $\mathbb{Z}/p\mathbb{Z}$ *are* colors. *The generalized Lucas' congruence extends the domain of* f *from digits to numbers.*

Definition 4.2.3. *Let* f *be a function that satisfies the generalized Lucas' congruence. We say that* f uniformly generates $(\mathbb{Z}/p\mathbb{Z})^*$ *if there is a* $k \in \mathbb{N}$ *with*

$$(\mathbb{Z}/p\mathbb{Z})^* = \{f(\vec{n}) \mid \vec{n} \in T_k\}.$$

If f uniformly generates $(\mathbb{Z}/p\mathbb{Z})^*$ for a $k_0 \in \mathbb{N}$ then f also uniformly generates $(\mathbb{Z}/p\mathbb{Z})^*$ for all $k > k_0$.

The following two theorems apply to all functions which satisfy the generalized Lucas' congruence. Binomial coefficients and Stirling numbers of the first and

second kind are the best known examples of functions which satisfy the generalized Lucas' congruence. Numerous other examples can be found in Section 4.5 and in the comprehensive paper [45].

Theorem 4.2.4. *Let f be a function which satisfies the generalized Lucas' congruence. If there is an $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod{p}$, then*

$$\lim_{\ell \to \infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod{p}\}}{\#\{\vec{n} \in T_\ell\}} \to 0.$$

If there is a matrix digital system with a coloring function which satisfies the generalized Lucas' congruence, we speak of a *number scheme*. Let us recall that a *continuity set* of a measure μ is any Borel set B with a boundary set ∂B of measure zero $\mu(\partial B) = 0$.

Theorem 4.2.5. *Let $(A, \mathcal{D})$ be a matrix digital system in $\mathbb{Z}^d$ with a matrix A of the form $A = gQ$ where Q is an orthonormal matrix and $g \in \mathbb{N}$. Let f be a function which uniformly generates $(\mathbb{Z}/p\mathbb{Z})^*$ and let $\mu := \mathcal{H}^s|\mathcal{F}$ be the normalized Hausdorff measure of dimension*

$$s = \frac{\log \#\{\vec{\varepsilon} \in \mathcal{D} \mid f(\vec{\varepsilon}) \neq 0\}}{\log g}$$

restricted to the fundamental domain $\mathcal{F}$. If $\mathcal{B}$ is a μ-continuity set and $a \not\equiv 0 \pmod{p}$, then

$$\lim_{\ell \to \infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \equiv a \pmod{p},\ \vec{n} \in A^{\ell+1}\mathcal{B}\}}{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod{p}\}} \to \frac{\mu(\mathcal{B})}{p-1}.$$

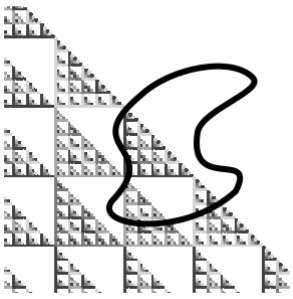

Figure 4.2: A μ-continuity set on the binomial coefficients.

Theorem 4.2.5 is a spatial distribution result. One way to visualize theorem 4.2.5 is to take a μ-continuity set $\mathcal{B}$ in $\mathbb{R}^d$, that is $\mu(\partial\mathcal{B}) = 0$. We assign

colors to numbers in T_ℓ and represent each number by a colored dot. If the assumptions of Theorem 4.2.5 are satisfied, the nonzero colors inside $\mathcal{B}$ converge to an equidistribution. The set $A^{-\ell-1}T_\ell \cap \mathcal{B}$ is either empty or the colors occur with the same asymptotic frequency if $\ell \to \infty$. Figure 4.2 is a graphic representation of this idea.

4.3 Fractals

Matrix number systems are connected to fractals. The set $\mathcal{F}$ of all numbers $\vec{x} \in \mathbb{R}^d$ which can be written as

$$\vec{x} = \sum_{i=0}^{\infty} A^{-i}\vec{\varepsilon_i}$$

with $\vec{\varepsilon_i} \in \mathcal{D}$ is called the *fundamental region* with respect to $(A, \mathcal{D})$. Since

$$\mathcal{F} = \bigcup_{\vec{\varepsilon} \in \mathcal{D}} (A^{-1}\varepsilon + A^{-1}\mathcal{F}),$$

the set is self-similar and we can write it as a fractal in the sense of Barnsley [10].

We will use the term *fractal* for sets generated by an iterated function system (short IFS). An *iterated function system* is a finite collection $(w_1, w_2, \ldots, w_t)$ of contractions on a compact metric space $(\mathbf{X}, d)$. We will denote the IFS by

$$\mathcal{W} = \{\mathbf{X}; w_1, w_2, \ldots, w_t\}.$$

The main theorem about IFS is from Hutchinson [38].

Theorem 4.3.1 (Hutchinson 1981)**.** *Let $\mathcal{W} = \{\mathbf{X}; w_1, w_2, \ldots, w_t\}$ be an IFS on the compact metric space $(\mathbf{X}, d)$. The transformation $W : \mathcal{H}(\mathbf{X}) \to \mathcal{H}(\mathbf{X})$ defined by*

$$W(U) := \cup_{i=1}^{t} w_i(U) \quad \textit{(Hutchinson operator)}$$

has a unique fixed point $\mathcal{F} \in \mathcal{H}(\mathbf{X})$ for all $U_0 \in \mathcal{H}(\mathbf{X})$, which obeys $\mathcal{F} = W(\mathcal{F})$. It can be constructed as $\mathcal{F} = \lim_{n\to\infty} U_N$ where $U_N = W(U_{N-1})$. This fixed point is called the attractor of the IFS.

We will follow the books form Egdar [26] and Falconer [29] to define measures on the IFS. If $(\mathbf{X}, d)$ is a compact metric space, let $\mathcal{P}(\mathbf{X})$ denote the space of normalized Borel measures on $\mathbf{X}$. To every contraction w_i in $\{\mathbf{X}; w_1, w_2, \ldots, w_t\}$ we assign a weight $p_i > 0$ so that $\sum_{i=1}^{t} p_i = 1$.

In our case, for a self-similar IFS which satisfies the open set condition there is a privileged choice of weights.

Definition 4.3.2. *Be $\mathcal{F}$ the attractor of the IFS $\{\mathbf{X}; w_1, w_2, \ldots, w_t\}$ with similarities $w_1, w_2, \ldots, w_t$. We say $\mathcal{F}$ fulfills the* open set condition *if there is a nonempty open set $\mathcal{O}$ satisfying*

- $w_i(\mathcal{O}) \cap w_j(\mathcal{O}) = \emptyset$ *for any* $i \neq j$;
- $w_i(\mathcal{O}) \subset \mathcal{O}$ *for all* $1 \leq i \leq t$.

The maps are similarities since the matrix A is orthonormal. The open set condition is fulfilled since the matrix A is an integer matrix and $\mathcal{D}$ is a complete residue system (cf. [13]).

Let the maps w_i be similarities with ratios r_i. The *similarity dimension* s is the solution of the equation

$$\sum_{i=1}^{t} r_i^s = 1.$$

The weights $p_i = r_i^s$ are called the *uniform weights*.

There is a unique probability measure $\mu \in \mathcal{P}(\mathbf{X})$, so that for all Borel sets $\mathcal{B}$ we have

$$\mu(\mathcal{B}) := \sum_{i=1}^{t} r_i^s \mu(w^{-1}(\mathcal{B})). \tag{4.3}$$

Since

$$\mu(\mathcal{F}) = \sum_{i=1}^{t} r_i^s \mu(w^{-1}(\mathcal{F})),$$

the measure μ is an invariant measure on the IFS with weights. It is known as the *uniform measure*. In our examples μ is always a Hausdorff measure. The support of μ is the fractal set $\mathcal{F}$.

Later we will use two different IFSs. In the proof of Theorem 4.2.4 we use an IFS $\mathcal{W} = \{X; w_{1,2}, \ldots, w_t\}$ where w_i is an affine map of the form

$$w_i : \vec{x} \mapsto A^{-1}(\vec{x} - \vec{\varepsilon_i})$$

with $\vec{\varepsilon_i} \in \mathcal{D}$. For this IFS we have $U_n = U_0$ if we take $U_0 = T_0$. The invariant measure is the usual Lebesgue measure λ. An example is visualized in Figure 4.3.

The other IFS we use for the proof of Theorem 4.2.5. Again we take an IFS $\mathcal{W} = \{X; w_{1,2}, \ldots, w_t\}$ where w_i is an affine map of the form

$$w_i : \vec{x} \mapsto A^{-1}(\vec{x} - \vec{\varepsilon_i}),$$

but this time with $\vec{\varepsilon_i} \in \mathcal{D} \setminus \{\vec{\varepsilon_i} \mid f(\vec{\varepsilon_i}) \equiv 0 \pmod{p}\}$. The invariant measure μ in this case is the normalized Hausdorff measure of dimension

$$s = \frac{\log \#\{\vec{\varepsilon} \in \mathcal{D} \mid f(\vec{\varepsilon}) \neq 0\}}{\log g}$$

restricted to $\mathcal{F}$.

Please note that the colorings are defined on the sequence U_n. It is not possible to define a coloring for the limit fractal.

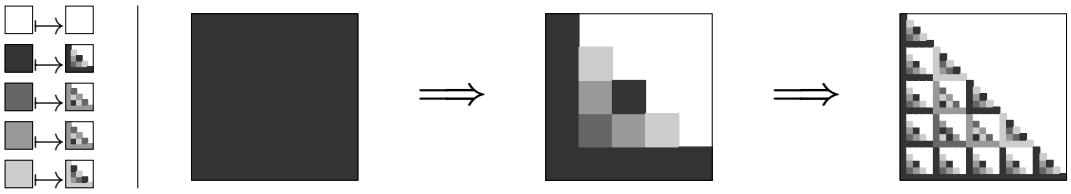

Figure 4.3: On the left we see the substitutions for $p = 5$, while on the right we see the first two steps of their action

4.4 Character sums and measures

The definition of p-colorings leads us to the question as to how many numbers have a certain color, i.e. belong to a certain residue class modulo p. To count the numbers in one residue class, we use *Dirichlet characters*.

Definition 4.4.1. *A Dirichlet character modulo k is a function $\chi : \mathbb{Z} \to \mathbb{C}$, so that*

(i) $\chi(n) = \chi(n + k)$,

(ii) $\chi(n) = 0$ *if and only if* $\gcd(n, k) > 1$ *and*

(iii) $\chi(nm) = \chi(n)\chi(m)$.

Characters take roots of unity as values. The *principal character* $\chi_0(n)$ is given by $\chi_0(n) = 1$ if $\gcd(n, k) = 1$.

Proof of Theorem 4.2.4. Each vector $\vec{n} \in T_\ell$ has a unique representation of the form $\vec{n} = (\vec{\varepsilon_0}\vec{\varepsilon_1}\dots\vec{\varepsilon_\ell})$. There are $\#\mathcal{D}$ ways to choose each digit. Therefore, $\#T_\ell = (\#\mathcal{D}_2)^{\ell+1}$.

Equation (4.2) allows us to write $f(\vec{n})$ as $f(\vec{n}) \equiv f(\vec{\varepsilon_0})f(\vec{\varepsilon_1})\dots f(\vec{\varepsilon_\ell}) \pmod p$. Since $f(\vec{\varepsilon})$ is in the finite field $\mathbb{Z}/p\mathbb{Z}$, we have $f(\vec{n}) \not\equiv 0 \pmod p$ if and only if $f(\vec{\varepsilon_i}) \not\equiv 0 \pmod p$ for $0 \le i \le \ell$. So we count $D := \sum_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon}))$. There is an $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod p$ and therefore $D < \#\mathcal{D}$ and we get

$$\lim_{\ell\to\infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod p\}}{\#\{\vec{n} \in T_\ell\}} = \lim_{\ell\to\infty} \frac{D^{\ell+1}}{(\#\mathcal{D})^{\ell+1}} = 0.$$

□

Remark: *Of course we could state Theorem 4.2.4 in the same form as Theorem 4.2.5.*

$$\lim_{\ell\to\infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \not\equiv 0 \pmod p,\ \vec{n} \in A^{\ell+1}\mathcal{B}\}}{\#\{\vec{n} \in T_\ell\}} \to 0.$$

Since the set of all colored points is a measure-zero set on the whole fractal (with respect to the Lebesgue measure), it is a measure-zero set for every μ-continuity set $\mathcal{B}$.

Lemma 4.4.2. *If and only if f uniformly generates $(\mathbb{Z}/p\mathbb{Z})^*$ then for all $\chi \neq \chi_0$ we have*

$$\left|\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi(f(\vec{\varepsilon}))\right| < \sum_{\substack{\vec{\varepsilon}\in\mathcal{D}\\ f(\vec{\varepsilon})\neq 0}} 1.$$

Proof. As first step we show that for every $\chi \neq \chi_0$ we have

$$\left|\sum_{\vec{n}\in T_k} \chi(f(\vec{n}))\right| < \left|\sum_{\vec{n}\in T_k} \chi_0(f(\vec{n}))\right| = \sum_{\substack{\vec{n}\in T_k\\ f(\vec{n})\neq 0}} 1. \tag{4.4}$$

There can be no equality in equation (4.4) since that would imply that all characters, which are roots of unity, have the same argument. On the one hand $\vec{0} \in T_k$ and $1 = f(\vec{0}) = \chi(f(\vec{0}))$. On the other hand, since the values of f generate $(\mathbb{Z}/p\mathbb{Z})^*$, there is a $\vec{n} \in T_k$ with $\chi(f(\vec{n})) \neq 1$.

Now we use the generalized Lucas' property and the multiplicativity of characters to write equation (4.4) as

$$\left|\prod_{i=0}^{k}\left(\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi(f(\vec{\varepsilon}))\right)\right| < \prod_{i=0}^{k}\left(\sum_{\substack{\vec{n}\in\mathcal{D}\\ f(\vec{\varepsilon})\not\equiv 0}} 1\right)$$

which proves Lemma 4.4.2. □

In the proof of Theorem 4.2.4 we just used the principal character as the indicator function for the non-zero residue classes. For the next proof we will need a well known identity for characters (cf. [7])

$$\frac{1}{p-1}\sum_{\chi} \overline{\chi(a)}\chi(b) = \begin{cases} 1 & \text{if } a \equiv b \\ 0 & \text{else} \end{cases} \tag{4.5}$$

where the sum runs over all characters.

Proof of Theorem 4.2.5. If we define a function

$$F_\chi(\vec{x}) := \frac{\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi(f(\vec{\varepsilon}))e(\langle\vec{\varepsilon},\vec{x}\rangle)}{\sum_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon}))},$$

Lemma 4.4.2 tells us that

$$\left|F_\chi(\vec{0})\right| < 1 \tag{4.6}$$

for $\chi \neq \chi_0$.

Now we define a sequence of measures on T_ℓ which, as we will show, converges weakly to a measure μ.

$$\frac{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\neq 0}} \delta_{A^{-\ell-1}\vec{n}}}{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\neq 0}} 1} \rightharpoonup \mu. \tag{4.7}$$

The measure μ satisfies equation (4.3). Then we define another sequence of measures on T_ℓ which converges weakly to a measure μ_a, as we will show later.

$$\frac{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\equiv a}} \delta_{A^{-\ell-1}\vec{n}}}{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\neq 0}} 1} \rightharpoonup \mu_a = \frac{1}{p-1}\mu \tag{4.8}$$

The measure μ_a also satisfies equation (4.3).

Then we look at the characteristic function $\hat{\mu}_a(\vec{x})$ of μ_a

$$\hat{\mu}_a(\vec{x}) = \lim_{\ell\to\infty} \frac{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\equiv a}} e(\langle A^{-\ell-1}\vec{n},\vec{x}\rangle)}{\sum\limits_{\substack{\vec{n}\in T_\ell \\ f(\vec{n})\neq 0}} 1}.$$

As usual, we have $e(x) := e^{2\pi i x}$.

We use Dirichlet characters to rewrite $\hat{\mu}_a(\vec{x})$. In the denominator we use the principal character as indicator function and get

$$\#\{\vec{n}\in T_\ell \mid f(\vec{\varepsilon}) \not\equiv 0 \pmod p\} = \sum_{\vec{n}\in T_\ell} \chi_0(f(\vec{n})).$$

In the numerator we can use equation (4.5) to get

$$\#\{\vec{n}\in T_\ell \mid f(\vec{\varepsilon}) \equiv a \pmod p\} = \frac{1}{p-1}\sum_\chi \overline{\chi(a)} \sum_{\vec{n}\in T_\ell} \chi(f(\vec{n})).$$

Hence, we have

$$\hat{\mu}_a(\vec{x}) = \lim_{\ell\to\infty} \frac{1}{p-1} \sum_{\chi} \overline{\chi(a)} \frac{\sum\limits_{\vec{n}\in T_\ell} \chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))}.$$

$$= \lim_{\ell\to\infty} \frac{1}{p-1} \frac{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} + \tag{4.9}$$

$$+ \lim_{\ell\to\infty} \frac{1}{p-1} \frac{\sum\limits_{\chi\neq\chi_0} \overline{\chi(a)} \sum\limits_{\vec{n}\in T_\ell} \chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))}. \tag{4.10}$$

The first part of the sum (4.9) comes from the principal character, while the second part (4.10) contains all other characters. The summand in equation (4.9) is just the characteristic function $\hat{\mu}(\vec{x})$ of μ

$$\lim_{\ell\to\infty} \frac{1}{p-1} \frac{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \frac{1}{p-1}\hat{\mu}(\vec{x}).$$

We use the conjugate transpose and write it as

$$\frac{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \prod_{k=0}^{\ell} \left(\frac{\sum\limits_{\vec{\varepsilon}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon_k}))e(\langle \vec{\varepsilon}, (A^T)^{-k}\vec{x}\rangle)}{\sum\limits_{\vec{n}\in\mathcal{D}} \chi_0(f(\vec{\varepsilon}))} \right) =$$

$$= \prod_{k=0}^{\ell} F_{\chi_0}(\vec{\varepsilon}, (A^T)^{-k}\vec{x}).$$

$F_{\chi_0}(\vec{0}) = 1$ and F_{χ_0} is continuously differentiable since it is a composition of continuously differentiable functions. Therefore, it is also Lipschitz continuous and since $A = gQ$ where Q is a orthonormal matrix, we get

$$\left|F_{\chi_0}((A^T)^{-k}\vec{x}) - 1\right| \leq L\|(A^T)^{-k}\vec{x}\| \leq Lg^{-k}\|\vec{x}\|.$$

Hence,

$$\prod_{k=0}^{\ell} F_{\chi_0}(\vec{\varepsilon}, (A^T)^{-k}\vec{x})$$

converges uniformly on every compact subset of $\mathbb{R}^d$ since $\|\vec{x}\|$ is bounded on a compact subset. Therefore, we can use the uniform convergence theorem and

know that $\hat{\mu}(\vec{x})$ is continuous on every compact subset and especially continuous in 0. We can use Lèvy's continuity theorem for characteristic functions and prove equation (4.7).

Lèvy's continuity theorem (cf. [21, Theorem 2.6.9]) states:

A sequence of random variables X_j in $\mathbb{R}^d$ convergences weakly to a random variable X if and only if the sequence of characteristic functions φ_{X_j} convergences pointwise to a function φ which is continuous in 0. Then φ is the characteristic function of X.

Now we focus on the nonprincipal characters in equation (4.10). We use the same idea as before and write

$$\frac{\sum\limits_{\vec{n}\in T_\ell} \chi(f(\vec{n}))e(\langle A^{-\ell-1}\vec{n}, \vec{x}\rangle)}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = \prod_{k=0}^{\ell} F_\chi(\vec{\varepsilon_k}, (A^T)^{-k}\vec{x}).$$

Because of equation (4.6) we know that there exist $\epsilon, \delta > 0$ so that

$$|F_\chi(\vec{x})| \leq 1 - \epsilon \text{ for } \|\vec{x}\| \leq \delta.$$

For $R > 0$ there is a $k_0 \in \mathbb{N}_0$, so that for all $k \geq k_0$

$$\|(A^T)^{-k}\vec{x}\| \leq \delta \text{ for } \|\vec{x}\| \leq R.$$

Thus,

$$\left|\prod_{k=0}^{\ell} F_\chi((A^T)^{-k}\vec{x})\right| \leq \prod_{k=k_0}^{\ell}(1-\epsilon) = (1-\epsilon)^{\ell-k_0+1}$$

and

$$\lim_{\ell\to\infty}(1-\epsilon)^{\ell-k_0+1} \to 0.$$

Therefore, the limit in equation (4.10) is zero and equation (4.8) is proved. □

4.5 Applications

In this section we give several examples of combinatorially defined number schemes which satisfy the generalized Lucas' congruence. Many more such examples (including some infinite classes) can be found in [45].

For each of these examples an analogon to Lucas' theorem can be found in the literature. Each time we will give an extended matrix digital system and a coloring function so that the generalized Lucas' congruence is satisfied. Therefore, Theorem 4.2.4 and Theorem 4.2.5 apply to the binomial coefficients, the Stirling

numbers of the first and second kind, the Gaussian q-nomial coefficients as well as to the multinomial coefficients.

In Figure 4.1 we see the binomial coefficients and Stirling numbers of the first and second kind modulo $p = 5$. All triangles have the same size but only the binomial coefficients have a bilateral symmetry. For the Stirling numbers of the first kind the leftmost numbers are all zero. The Stirling numbers of the second kind have a slanted structure.

With exception of the Apéry numbers it is very easy to check the conditions of Theorem 4.2.4 and Theorem 4.2.5 because all examples involve binomial coefficients. As usual, we define $a \bmod p := a - \lfloor \frac{a}{p} \rfloor p$.

Lemma 4.5.1. *The binomial coefficients satisfy the conditions of Theorem 4.2.4 and Theorem 4.2.5.*

Proof. The binomial coefficients fulfill Theorem 4.1.1. There is a corresponding matrix digital system $(A, \mathcal{D})$ with

$$A = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix}, \mathcal{D} = \left\{ \overrightarrow{\binom{v}{w}} \;\middle|\; 0 \le v, w \le p-1 \right\}$$

and coloring function $f(\overrightarrow{\binom{v}{w}}) := \binom{v}{w} \bmod p$ where p is a prime.

We have $f(\overrightarrow{\binom{v}{w}}) = 0$ for the $\frac{p(p-1)}{2}$ digits with $v < w$. Since there are digits $\vec{\varepsilon} \in \mathcal{D}$ so that $f(\vec{\varepsilon}) \equiv 0 \pmod p$, according to Theorem 4.2.4 almost all numbers are 0 modulo p.

For every $0 \le v \le p-1$ we have $f(\overrightarrow{\binom{v}{1}}) = f(\overrightarrow{\binom{v}{v-1}}) = v$. Thus, the binomial coefficients modulo p generate the whole group $(\mathbb{Z}/p\mathbb{Z})^*$ and we have equidistribution in the non-zero residue classes modulo p according to Theorem 4.2.5. □

4.5.1 Apéry numbers

The Apéry numbers $A_1(n)$ and $A_2(n)$ were introduced by Apéry in his 1979 proof that $\zeta(3)$ is irrational. They are defined as

$$A_1(n) := \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2, \quad A_2(n) := \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}.$$

Both of them satisfy the Lucas' congruence and for $j \in \{1, 2\}$ we have

$$A_j(n) \equiv \prod_{i=0}^{\ell} A_j(n_i) \pmod p.$$

The proof that the Apéry numbers $A_1(n)$ satisfy the Lucas' congruence has been published in [32], while the proof for the numbers $A_2(n)$ can be found in [24]. Here we have the usual p-adic representation which can be treated as a matrix digital system with the degenerate matrix $A = (p)$ and digit set $\mathcal{D} = \{v \mid 0 \le v \le p-1\}$.

The Apéry numbers differ from the other examples as there is no obvious way to check the conditions of Theorem 4.2.4 and Theorem 4.2.5.

In order to apply Theorem 4.2.4, there has to be an $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$. Often this condition is not satisfied. For instance, there is no $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$ for both Apéry numbers and $p = 13$. If there is no $\varepsilon \in \mathcal{D}$ with $f(\varepsilon) = 0$, the measure in Theorem 4.2.5 is the Lebesgue measure.

Based on calculations for the primes $p < 500$, I conjecture that the conditions of Theorem 4.2.5 are always satisfied for $p > 3$.

4.5.2 Binomial coefficients

As stated in the introduction, the binomial coefficients satisfy Lucas' original Theorem 4.1.1. We can treat the linear equation

$$\overrightarrow{\binom{np+r}{kp+s}} = \begin{pmatrix} p & 0 \\ 0 & p \end{pmatrix} \overrightarrow{\binom{n}{k}} + \overrightarrow{\binom{r}{s}} \tag{4.11}$$

as a matrix digital system $(A, \mathcal{D})$ with $A = \left(\begin{smallmatrix} p & 0 \\ 0 & p \end{smallmatrix}\right)$ and $\mathcal{D} = \{\overrightarrow{\binom{v}{w}} \mid 0 \le v, w \le p-1\}$.

If we interpret the vectors as binomial coefficients and use an induction, Lucas' theorem tells us that they satisfy the generalized Lucas' congruence with the coloring function

$$f\left(\overrightarrow{\binom{n}{k}}\right) = \binom{n}{k} \operatorname{Mod} p.$$

Lemma 4.5.1 shows that Theorems 4.2.4 and 4.2.5 are satisfied.

The binomial coefficients are an archetype for all other examples. We always look for a generalized Lucas' congruence, the corresponding matrix digital system and coloring functions that satisfy the congruence. Then we show that the digital function fulfills the hypothesis of Theorem 4.2.4 and Theorem 4.2.5.

4.5.3 Stirling numbers of the first kind

Theorem 4.5.2 is an analogon to Lucas' theorem for Stirling numbers of the first kind (cf. [48]). This is one of the more interesting examples and the first one with an extended matrix digital system, so we will look at it in more detail.

The Stirling numbers of the first kind can be defined by the recurrence relation

$$\begin{bmatrix} n+1 \\ k \end{bmatrix} = n \begin{bmatrix} n \\ k \end{bmatrix} + \begin{bmatrix} n \\ k-1 \end{bmatrix}$$

for $k > 0$ and with initial conditions

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = 1 \text{ and } \begin{bmatrix} n \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ n \end{bmatrix} = 0$$

for $n > 0$.

Theorem 4.5.2. *The Stirling numbers of the first kind satisfy the following congruence*

$$\begin{bmatrix} n \\ k \end{bmatrix} \equiv \begin{bmatrix} r \\ t \end{bmatrix} \binom{m}{s} (-1)^{m-s} \pmod{p} \tag{4.12}$$

with $n = mp + r$ *and* $0 \leq r \leq p-1$. *The numbers* s *and* t *are defined as* $k - m = s(p-1) + t$ *with* $0 \leq t < p-1$ *if* $r = 0$ *and* $0 < t \leq p-1$ *if* $r > 0$.

Again we use a linear equation

$$\overrightarrow{\binom{n}{k}} = \begin{pmatrix} p & 0 \\ 1 & p-1 \end{pmatrix} \overrightarrow{\binom{m}{s}} + \overrightarrow{\binom{r}{t}}$$

to define a matrix digital system $(B, \mathcal{D}_1)$ with the matrix $B = \left(\begin{smallmatrix} p & 0 \\ 1 & p-1 \end{smallmatrix}\right)$ and the digit set

$$\mathcal{D}_1 = \left\{ \overrightarrow{\binom{v}{w}} \;\middle|\; 1 \leq v, w \leq p-1 \right\} \cup \left\{ \overrightarrow{\binom{0}{w}} \;\middle|\; 0 \leq w \leq p-2 \right\}.$$

After we split off the least significant digit, let us call it $\overrightarrow{\varepsilon_{-1}}$, we are left with an expression of the form $\binom{m}{s}(-1)^{m-s}$. We already know that the binomial coefficients satisfy the generalized Lucas' congruence if we write m and s as p-adic numbers. Since $(-1)^p = -1$ for odd p and $-1 \equiv 1 \pmod 2$, we have

$$(-1)^{(m_\ell p^\ell + \cdots + m_1 p + m_0) - (s_\ell p^\ell + \cdots + s_1 p + s_0)} = \prod_{i=0}^{\ell} (-1)^{m_i - s_i}$$

so $(-1)^{m-s}$ also satisfies the generalized Lucas' congruence if we write m and s as p-adic numbers. Therefore, we use the matrix digital system $A = \left(\begin{smallmatrix} p & 0 \\ 0 & p \end{smallmatrix}\right)$ and $\mathcal{D} = \{\overrightarrow{\binom{v}{w}} \mid 0 \leq v, w \leq p-1\}$ from the binomial coefficients again to write $\overrightarrow{\binom{m}{s}}$ as

$$f\left(\overrightarrow{\binom{m}{s}}\right) \equiv \prod_{i=0}^{\ell} f(\vec{\varepsilon_i}) \pmod{p}$$

with coloring function

$$f\left(\overrightarrow{\binom{m}{s}}\right) = \binom{m}{s} (-1)^{m-s} \operatorname{Mod} p.$$

What we finally get is an expression of the form

$$\overrightarrow{\varepsilon_{-1}} + B(\vec{\varepsilon_0} + A\vec{\varepsilon_1} + A^2\vec{\varepsilon_2} + \cdots + A^\ell\vec{\varepsilon_\ell}). \tag{4.13}$$

Let us now look at the set V of all numbers of the form (4.13). We call this construction *nested matrix digital system*. Inside the brackets we have a matrix digital system with a similarity A. This matrix digital system satisfies a generalized Lucas' congruence and its fundamental domain $\mathcal{F}$ is the unit square. Outside the brackets we have $\#\mathcal{D}_1 = (p-1)p$ affine maps. It is easy to see that the interior of the unit square is mapped to nonoverlapping images.

If we define coloring functions

$$f_1\left(\overrightarrow{\binom{n}{k}}\right) = \begin{bmatrix} n \\ k \end{bmatrix} \operatorname{Mod} p,\ f\left(\overrightarrow{\binom{m}{s}}\right) = \binom{m}{s}(-1)^{m-s} \operatorname{Mod} p$$

equation (4.12) allows us to calculate $f_1\left(\overrightarrow{\binom{n}{k}}\right)$ as

$$f_1\left(\overrightarrow{\binom{n}{k}}\right) \equiv f_1(\overrightarrow{\varepsilon_{-1}}) \prod_{i=0}^{\ell} f(\vec{\varepsilon_i}) \pmod{p}. \tag{4.14}$$

Equation (4.14) and the fact that the images of the affine maps are nonoverlapping allows us to extend the generalized Lucas' congruence and the measures μ and μ_a to all numbers in V. The affine maps will deform the measures, but they do not change their equidistribution properties. Therefore, Theorem 4.2.4 and Theorem 4.2.5 apply to the Stirling numbers of the first kind.

4.5.4 Stirling numbers of the second kind

The Stirling numbers of the first kind can be defined by the recurrence relation

$$\left\{ {n+1 \atop k} \right\} = k \left\{ {n \atop k} \right\} + \left\{ {n \atop k-1} \right\}$$

for $k > 0$ and with initial conditions

$$\left\{ {0 \atop 0} \right\} = 1 \text{ and } \left\{ {n \atop 0} \right\} = \left\{ {0 \atop n} \right\} = 0$$

for $n > 0$.

The analogon to Lucas' theorem for Stirling numbers of the second kind can be found in Howard [37].

Theorem 4.5.3. *If p is prime and $n-(r+1)(p-1)=h$, then*

$$\left\{ n \atop hp \right\} \equiv \binom{r}{h-1} \bmod p. \tag{4.15}$$

If $n-(p-1)r-i=h$ and $1 \le m \le i \le p-1$, then

$$\left\{ n \atop hp+m \right\} \equiv \binom{r}{h} \left\{ i \atop m \right\} \bmod p. \tag{4.16}$$

If we take a look at the actual proof by Howard, we see that in all cases not covered by these equations we have $\left\{ n \atop k \right\} \equiv 0 \bmod p$.

To find a more convenient form for equation (4.15), we use the recursion formula for Stirling numbers of the second kind

$$\left\{ n \atop hp \right\} = \left\{ n-1 \atop hp-1 \right\} + \left\{ n-1 \atop hp \right\} \cdot hp \equiv \left\{ n-1 \atop (h-1)p+(p-1) \right\} \bmod p$$

and apply equation (4.16). We can use the linear equation

$$\overrightarrow{\binom{n}{k}} = \begin{pmatrix} p-1 & 1 \\ 0 & p \end{pmatrix} \overrightarrow{\binom{r}{h}} + \overrightarrow{\binom{i}{m}}$$

with matrix $B = \begin{pmatrix} p-1 & 1 \\ 0 & p \end{pmatrix}$ and digit set

$$\mathcal{D}_1 = \left\{ \overrightarrow{\binom{v}{w}} \;\middle|\; 1 \le v, w \le p-1 \right\} \cup \left\{ \overrightarrow{\binom{v}{p}} \;\middle|\; 2 \le w \le p \right\}$$

to define a nested matrix digital system. The matrix A and the digit set $\mathcal{D}$ are the same as in the previous example and we use the same argumentation to show that Theorem 4.2.4 and Theorem 4.2.5 apply to the Stirling numbers of the second kind.

4.5.5 Gaussian q-nomial coefficients

Here the analogon to Lucas‘ theorem is given by M. Sved in [53]. The Gaussian q-nomial coefficients are defined as

$$\begin{bmatrix} a \\ b \end{bmatrix}_q = \begin{cases} \frac{(1-q^a)(1-q^{a-1})\cdots(1-q^{a-b+1})}{(1-q)(1-q^2)\cdots(1-q^b)} & b \le a \\ 0 & b > a \end{cases}$$

for nonnegative integers a and b.

The Gaussian q-nomial coefficients have a regular, bilateral symmetric structure, which resembles that of the binomial coefficients. This is to be expected since the matrix digital systems are also very similar.

Theorem 4.5.4. *Let p be a prime, $q > 1$ a positive integer not divisible by p and let $a \neq 1$ be the minimal exponent for which $q^a \equiv 1 \pmod{p}$; then by Fermat's little theorem it follows that $a|(p-1)$. Furthermore, if $0 \leq r, s < a$, then*

$$\begin{bmatrix} na+r \\ ka+s \end{bmatrix}_q \equiv \binom{n}{k} \begin{bmatrix} r \\ s \end{bmatrix}_q \pmod{p}.$$

The linear equation

$$\overrightarrow{\binom{na+r}{ka+s}} = \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} \overrightarrow{\binom{n}{k}} + \overrightarrow{\binom{r}{s}} \tag{4.17}$$

gives us the nested matrix digital system with matrix $B = \left(\begin{smallmatrix} a & 0 \\ 0 & a \end{smallmatrix}\right)$ and digit set $\mathcal{D}_1 = \{\overrightarrow{\binom{v}{w}} \mid 0 \leq v, w \leq a-1\}$. Now we are left with the binomial coefficients for the infinite part and can use same arguments as for the Stirling numbers.

4.5.6 Multinomial coefficients

There is a d-dimensional analogon of Lucas' theorem for the multinomial coefficients (cf. [25], p.273):

Theorem 4.5.5. *The multinomial coefficients satisfy*

$$\binom{n}{k^{(1)}, \ldots, k^{(d)}} \equiv \prod_{d=0}^{L} \binom{n_\ell}{k_\ell^{(1)}, \ldots, k_\ell^{(d)}} \pmod{p}$$

with $n = \sum_{\ell=0}^{L} n_\ell p^\ell$ and $k^{(i)} = \sum_{\ell=0}^{L} k_\ell^{(i)} p^\ell$ where $0 \leq n_\ell, k_\ell^{(i)} \leq p-1$.

$$\overrightarrow{\begin{pmatrix} np+r \\ k^{(1)}p+s^{(1)} \\ \vdots \\ k^{(d)}p+s^{(d)} \end{pmatrix}} = \begin{pmatrix} p & 0 & \cdots & 0 \\ 0 & p & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & p \end{pmatrix} \overrightarrow{\begin{pmatrix} n \\ k^{(1)} \\ \vdots \\ k^{(d)} \end{pmatrix}} + \overrightarrow{\begin{pmatrix} r \\ s^{(1)} \\ \vdots \\ s^{(d)} \end{pmatrix}} \tag{4.18}$$

Again, we have a matrix digital system $(A, \mathcal{D})$ with $A = pI_{d+1}$ where I_d is the d-dimensional identity matrix and $\mathcal{D} = \{(v_1, \ldots, v_{d+1})^T \mid 0 \leq v_i \leq p-1\}$. We can use Lemma 4.5.1 since the multinomial coefficients contain the binomial coefficients.

4.6 Concluding Remarks

In this chapter we studied several examples of combinatorial defined number schemes. For these examples we showed that almost all entries in the number

scheme are divisible by a given prime p and the nonzero residue classes are equidistributed modulo p. It turned out that all multidimensional examples of combinatoric functions with the Lucas' property are based on the binomial coefficients, for which the matrix A is just the p-fold identity matrix.

Theorems 4.2.4 and 4.2.5 allow us to treat more complicated examples. For Theorem 4.2.4 we only need an extended matrix digital system with a coloring function which satisfies the generalized Lucas' congruence. The proof will work with any two integer matrices A and B which are affine expansions.

For the proof of Theorem 4.2.5 we required the matrix B to be a similarity. If B is a similarity, the measure μ is a uniform measure with identical weights $p_i = \frac{1}{t}$ and we can determine the dimension s and describe the measure μ, which is the Hausdorff measure.

If B is not a similarity map but any affine map given by an expanding integer matrix, the whole proof of Theorem 4.2.5 and the arguments with Dirichlet characters and Fourier transforms of measures work in exactly the same way. We still get an equidistribution result with respect to a probability measure μ, though μ will not be the Hausdorff measure but a less explicit measure.

Chapter 5

Remarks on "Spatial equidistribution of combinatorial number schemes"

5.1 Speed of convergence

The main theorem of the last chapter is Theorem 4.2.5. With a closer look at the proof of Theorem 4.2.5 we see that the speed of convergence depends entirely on the convergence of

$$\lim_{\ell\to\infty} \frac{\sum\limits_{\vec{n}\in T_\ell} \chi(f(\vec{n}))}{\sum\limits_{\vec{n}\in T_\ell} \chi_0(f(\vec{n}))} = 0. \tag{5.1}$$

If we specify a matrix digital system and a coloring function which satisfies the generalized Lucas' congruence, we can estimate the speed of convergence. In our examples (with the exception of the Apéry numbers) the speed of convergence depends on the binomial coefficients.

These are a number scheme with a corresponding matrix digital system $(A, \mathcal{D})$ with

$$A = \left(\begin{smallmatrix} p & 0 \\ 0 & p \end{smallmatrix}\right) \text{ and } \mathcal{D} = \{\overrightarrow{\binom{v}{w}} \mid 0 \le v, w \le p-1\}.$$

As coloring functions we take the functions

$$f\left(\overrightarrow{\binom{n}{k}}\right) := \binom{n}{k} \operatorname{Mod} p$$

with $a \operatorname{Mod} p := a - \lfloor \frac{a}{p} \rfloor p$ for $\overrightarrow{\binom{n}{k}} \in \mathbb{N}_0^2$ and the function

$$f_1\left(\overrightarrow{\binom{n}{k}}\right) := \binom{n}{k} \operatorname{Mod} p$$

for $\overrightarrow{\binom{n}{k}} \in \mathcal{D}$. As stated in the last chaper, the binomial coefficients satisfy Lucas' original theorem.

If we define

$$F_\chi(p) := \sum_{0 \le k < p} \chi\left(f_1\left(\binom{p-1}{k}\right)\right),$$

equation (5.1) becomes now

$$\lim_{\ell \to \infty} \left(\frac{F_{\dot\chi}(p)}{F_{\chi_0}(p)}\right)^\ell = 0$$

where $\dot\chi \neq \chi_0$ is the character for which $F_\chi(p)$ assumes the maximum.

Since the binomial coefficients $\binom{n}{k}$ are polynomials in n for a fixed k, we can use the Weil bound to estimate the speed of convergence. The Weil bound can be found as Theorem 2C' in [50] and states (paraphrased):

Theorem 5.1.1 (Weil bound)**.** *Let χ be of order $d > 1$. Suppose $f \in \mathbb{F}_p[X]$ has m distinct roots and f is not a d-th power. Then*

$$|\sum_{x \in \mathbb{F}_p} \chi(f(x))| \le (m-1)p^{1/2}.$$

In the paper we used the Weil bound for $k = 0$ which gives

$$\sum_{n=0}^{p-1} \chi\left(f_1\left(\binom{n}{1}\right)\right) = 0$$

and therefore

$$F_{\dot\chi}(p) \le F_{\chi_0}(p) - p.$$

But we can find a much better estimate.

Theorem 5.1.2. *The binomial coefficients satisfy*

$$|F_{\dot\chi}(p)| \le |h(p, \bar s)| \le \frac{p(p+1)}{2} - 2p + 4.$$

Here $h(p,s) := \frac{p(p+1)}{2} - s(2p - 3s - 1 - \sqrt{p}(s-1))$, $s := \frac{-1+\sqrt{p}+2p}{6+2\sqrt{p}}$ *and* $\bar s := \min(h(p, \lfloor s \rfloor), h(p, \lceil s \rceil))$.

p	$F_{\chi_0}(p) = \frac{p(p+1)}{2}$	$\max_{\chi \neq \chi_0} \lvert F_\chi(p)\rvert$	$h(p, \lfloor s \rfloor)$	$h(p, \lceil s \rceil)$	$h(p,1)$
2	3	–	–	–	–
3	6	4	6	4	4
5	15	9	9	13.472	9
7	28	14.731	18	19.291	18
11	66	28.872	48	42.633	48
13	91	29.411	60.211	64.633	69
17	153	52.393	107.246	105.739	123
19	190	68.113	136.718	132.153	156
23	276	91.028	196.775	201.550	234

Figure 5.1: The improved estimate

Proof. Since for a fixed k the binomial coefficients are polynomials in n we can use Theorem 5.1.1 to estimate the k-th column $\sum_{k \leq n \leq p-1} \dot{\chi}\left(\binom{n}{k}\right)$ of the triangle $F_{\dot{\chi}}(p) = \sum_{0 \leq k \leq n < p} \dot{\chi}\left(\binom{n}{k}\right)$. The k-th column has k distinct roots and we get

$$\left| \sum_{k \leq n \leq p-1} \dot{\chi}\left(\binom{n}{k}\right) \right| \leq (k-1)\sqrt{p}.$$

Since $\binom{n}{k} = \binom{n}{n-k}$, we can use this symmetry property to estimate the "inverse column" $\sum_{k \leq n \leq p-1} \dot{\chi}\left(\binom{n}{n-k}\right)$ and get

$$\left| \sum_{k \leq n \leq p-1} \dot{\chi}\left(\binom{n}{n-k}\right) \right| \leq (k-1)\sqrt{p}.$$

Now we want to use both of these estimates simultaneously for the s columns and "inverse columns" $k \in \{1, \ldots, s\}$. But since we count s^2 roots of unity twice, we have to add s^2 and get

$$\left| \sum_{\substack{1 \leq k \leq s \\ k \leq n \leq p-1}} \dot{\chi}\left(\binom{n}{k}\right) \right| + \left| \sum_{\substack{1 \leq k \leq s \\ k \leq n \leq p-1}} \dot{\chi}\left(\binom{n}{n-k}\right) \right| \leq 2\sqrt{p}\tfrac{(s-1)s}{2} + s^2.$$

For all other binomial coefficients in $F_{\dot{\chi}}(p)$ we use the trivial estimate. For $k = 0$ and $k = n$ there are altogether $2p - 2s - 1$ summands we have not dealt with. And there is a remaining triangle with $\frac{(p-2-2s)(p-1-2s)}{2}$ elements. Putting the estimates together, we define a function

$$h(p,s) := 2p - 2s - 1 + 2\sqrt{p}\tfrac{(s-1)s}{2} + s^2 + \tfrac{(p-2-2s)(p-1-2s)}{2}$$

and know that $|F_{\chi}(p)| \leq h(p,s)$ for $s \in \{1, 2, \ldots \frac{p-1}{2}\}$. To find the best s we differentiate and solve for s which gives us $s = \frac{-1+\sqrt{p}+2p}{6+2\sqrt{p}}$. Since s is an integer, we have to use either $\lfloor s \rfloor$ or $\lceil s \rceil$.

It is not obvious that this is indeed an improvement. But if we take $s = 1$ in theorem 5.1.1 we get

$$h(p,1) = \tfrac{p(p+1)}{2} - 2p + 4$$

which is better than the trivial estimate. □

Remark: *In order to simplify the calculation of $\bar{s}$, one can use $s = \lfloor \sqrt{p} \rfloor$ since $1 \leq s < \sqrt{p}$ and $s \sim \sqrt{p}$. The asymptotic analysis for s also yields that for the optimal s we will get*

$$h(p,s) \sim F_{\chi_0}(p) - p^{\frac{3}{2}} + 2p + p^{\frac{1}{2}}.$$

5.2 The p-free binomial coefficients

We mentioned in the introduction that we could prove the same results with the use of automata. We will give one example to show the general idea. Take the following theorem by Anton [6].

Theorem 5.2.1. *Suppose that p^{ℓ} is the highest power of p dividing $\binom{n}{k}$. Then*

$$\frac{1}{p^{\ell}}\binom{n}{k} \equiv (-1)^{\ell} \frac{n_0!}{k_0! r_0!} \frac{n_1!}{k_1! r_1!} \cdots \frac{n_d!}{k_d! r_d!} \pmod{p} \tag{5.2}$$

with $n = n_0 + n_1 p + \cdots + n_d p^d$ in base p, $k = k_0 + k_1 p + \cdots + k_d p^d$ and

$$r = n - k = r_0 + r_1 p + \cdots + r_d p^d$$

where $0 \leq n_i, k_i, r_i \leq p-1$ for $0 \leq i \leq d$. Note that ℓ is the number of "carries", when adding k and $r = n - k$ in base p.

Proof. We write n and k as $n = vp + a$ and $k = wp + b$ with $a, b < p$. If $b \leq a$, there is no carry and Lucas' theorem tells us that

$$\binom{n}{k} \equiv \binom{v}{w}\binom{a}{b} \pmod{p}.$$

For $b > a$ we will show that

$$\binom{n}{k} \equiv v\binom{v-1}{w}\binom{a+p}{b} \pmod{p}. \tag{5.3}$$

The main ingredient of the proof is the following congruence

$$\frac{n!}{p^v} \equiv (p-1)^v a!v! \pmod{p}. \tag{5.4}$$

This is because of

$$n! \equiv (vp+a)! \equiv (p-1)!\ p\ (p-1)!\ 2p \ldots (p-1)!\ vp\ a! \equiv p^v (p-1)^v a!v! \pmod{p}.$$

If we use equation (5.4) on $k = wp + b$ and $n - k = (v - w - 1)p + (p + a - b)$, we get $\frac{k!}{p^w} \equiv (p-1)^w b!w! \pmod{p}$ and

$$\frac{(n-k)!}{p^{v-w-1}} \equiv (p-1)^{v-w-1}(p+a-b)!(v-w-1)! \pmod{p}.$$

Since $\binom{n}{k} k!(n-k)! = n!$, we have

$$\frac{\binom{n}{k}}{p} \frac{k!}{p^w} \frac{(n-k)!}{p^{v-w-1}} \equiv \frac{n!}{p^v} \pmod{p}$$

and after inserting the expressions for $n!, k!$ and $(n-k)!$ we get

$$\frac{\binom{n}{k}}{p}(p-1)^w b!w!(p-1)^{v-w-1}(p+a-b)!(v-w-1)! \equiv (p-1)^v a!v! \pmod{p}.$$

We substitute

$$\frac{(p+a)!}{p} \equiv (p-1)!a! \equiv -a! \pmod{p} \tag{5.5}$$

(which is equation (5.4) with $v = 1$ the use of Wilson's theorem) on the right hand side and divide by

$$(p-1)^w (p-1)^{v-w-1} = (p-1)^{v-1}$$

to get

$$\frac{\binom{n}{k}}{p} b!w!(p+a-b)!(v-w-1)! \equiv \frac{(p+a)!}{p}(v-1)!v \pmod{p}.$$

Since $b < p+a$ and $w \leq v-1$, the numbers on the right hand side of equation (5.6) are integers and so

$$\frac{\binom{n}{k}}{p} \equiv \frac{1}{p} \frac{(p+a)!}{b!(p+a-b)!} \frac{(v-1)!}{w!(v-w-1)!} v \pmod{p} \tag{5.6}$$

is the same as equation (5.3).

If we substitute equation 5.5 in equation 5.6, we get

$$\frac{\binom{n}{k}}{p} \equiv \frac{-a!}{b!(p+a-b)!}\frac{v!}{w!(v-w-1)!} \pmod{p}$$

and Theorem 5.2.1 follows by induction. □

Theorem 5.2.1 looks very similar to a generalized Lucas' congruence but we cannot apply the theorems from the last chapter directly since we have carries in the calculation of $r = n - k$. We can circumvent this difficulty with the use of an automaton.

The input of this automaton are the digits $\overrightarrow{\binom{n_j}{m_j}}$, starting with the least significant digit $\overrightarrow{\binom{n_0}{m_0}}$. The automaton in Figure 5.2 handles the carries by switching between the states "non carry" (NC) and "carry" (C).

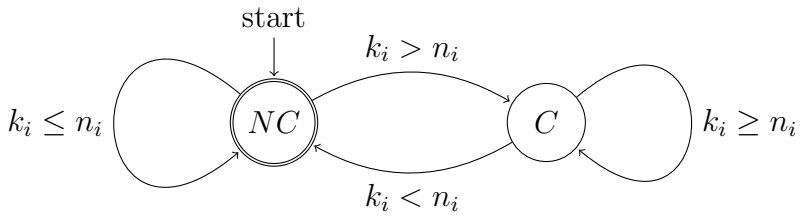

Figure 5.2: An automaton to handle the carries

Once we got the idea to let an automaton handle the carries we use 2×2 matrices to keep track of the actual values. The positions of the 2×2 matrices correspond to the transitions in the automaton.

$$\begin{pmatrix} NC\to NC & NC\to C \\ C\to NC & C\to C \end{pmatrix}.$$

We distinguish three cases and define

$$M(n_i,k_i) := \begin{cases} \begin{pmatrix} \binom{n_i}{k_i} & 0 \\ \frac{n_i!}{k_i!(n_i-k_i-1)!} & 0 \end{pmatrix} & \text{if } k_i < n_i, \\ \begin{pmatrix} \binom{n_i}{k_i} & 0 \\ 0 & \frac{-n_i!}{k_i!(p-1)!} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } k_i = n_i, \\ \begin{pmatrix} 0 & \frac{-n_i!}{k_i!(n_i+p-k_i)!} \\ 0 & \frac{-n_i!}{k_i!(n_i-k_i+p-1)!} \end{pmatrix} & \text{if } k_i > n_i. \end{cases}$$

The entries of the matrices are just $\frac{n_i!}{k_i!r_i!}$. In order to calculate r_i, we take $n_i - k_i$ and subtract one if we have a carry from the previous position. If we have

a carry to the next position, we add p. Instead of multiplying the final result by $(-1)^\ell$ there is a $-n!$ in the numerator for each carry.

Thus, we can write Theorem 5.2.1 as

$$\frac{1}{p^\ell}\binom{n}{k} \equiv (1\ 0)M(n_0,k_0)M(n_1,k_1)\dots M(n_d,k_d)\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right) \pmod p.$$

With this new formulation of Theorem 5.2.1, we can show an equidistribution result for the p-free binomial coefficients. As before, T_ℓ denotes the set of all vectors which have at most ℓ digits when written in base p. As coloring function we take the function

$$f\left(\overrightarrow{\binom{n}{k}}\right) := \frac{1}{p^\ell}\binom{n}{k} \operatorname{Mod} p$$

with $a \operatorname{Mod} p := a - \lfloor \frac{a}{p} \rfloor p$, where p^ℓ is the highest power of p dividing $\binom{n}{k}$.

Theorem 5.2.2. *The numbers $\frac{1}{p^\ell}\binom{n}{k}$, where p^ℓ is the highest power of p dividing $\binom{n}{k}$, are equidistributed in the residue classes modulo p. That is, if the set $\mathcal{B}$ is a λ-continuity set of the Lebesgue measure λ, then*

$$\lim_{\ell\to\infty} \frac{\#\{\vec{n} \in T_\ell \mid f(\vec{n}) \equiv a \pmod p,\ \vec{n} \in \mathcal{B}\}}{\#\{\vec{n} \in T_\ell\}} \to \frac{\mu(\mathcal{B})}{p-1}.$$

Proof. Like in the last chapter we will use Dirichlet characters again. We sum over the characters of all digits in the fundamental domain and define

$$M_\chi := \begin{pmatrix} \sum\limits_{0\le k\le n\le p-1} \chi\left(\binom{n}{k}\right) & \sum\limits_{0\le n<k\le p-1} \chi\left(\frac{-n!}{k!(n+p-k)!}\right) \\ \sum\limits_{0\le k<n\le p-1} \chi\left(\frac{-n!}{k!(n-k-1)!}\right) & \sum\limits_{0\le n\le k\le p-1} \chi\left(\frac{-n!}{k!(n-k+p-1)!}\right) \end{pmatrix}.$$

For the main character we get

$$M_{\chi_0} = \begin{pmatrix} \frac{p(p+1)}{2} & \frac{p(p-1)}{2} \\ \frac{p(p-1)}{2} & \frac{p(p+1)}{2} \end{pmatrix}$$

and thus

$$T_\ell = (1\ 0)M_{\chi_0}^\ell \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right).$$

Then we show that for all other characters $\chi \neq \chi_0$ the absolute value of every entry in the matrix is smaller than the corresponding entry in the matrix M_{χ_0}.

The following four equations show that the entries of the matrix assume every value in $(\mathbb{Z}/n\mathbb{Z})^*$. Note that if a assumes every value in $(\mathbb{Z}/n\mathbb{Z})^*$, so does $\frac{1}{a}$ and $\frac{1}{-a}$. We use again Wilson's theorem that

$$(p-1)! \equiv -1 \pmod p \text{ and } (p-2)! \equiv 1 \pmod p$$

respectively.

$$\binom{n}{1} = n \text{ for } 1 \le n \le p-1$$
$$\frac{n!}{k!(n+p-k)!} \equiv \frac{1}{-(n+1)} \pmod{p} \text{ for } k = n+1 \text{ and } 1 \le k \le p-1$$
$$\frac{n!}{k!(n-k-1)!} \equiv n \pmod{p} \text{ for } 1 \le n \le p-1 \text{ and } k = 0$$
$$\frac{n!}{k!(n+p-k-1)!} \equiv \frac{1}{(n+1)} \pmod{p} \text{ for } k = n+1 \text{ and } 1 \le k \le p-1$$

Since we sum over all nonzero residue classes modulo p and

$$\sum_{1 \le a \le p-1} \chi(a) = 0$$

for $\chi \neq \chi_0$, we have for every entry $m(i,j)$ of the matrices

$$|m_\chi(i,j)| \le m_{\chi_0}(i,j) - (p-1).$$

Now we use any matrix norm and get the equidistribution result. □

Example 5.2.3. *We calculate* $\binom{1356}{433}$ *modulo* 5. *Since we have* $1356 = [20411]_5$ *and* $433 = [3213]_5$, *it follows that* $r = [20411]_5 - [3213]_5 = [12143]_5$ *so* $\ell = 3$. *Hence Theorem 5.2.1 states*

$$\frac{1}{5^3}\binom{[20411]_5}{[03213]_5} \equiv (-1)^3 \frac{2!}{0!1!}\frac{0!}{3!2!}\frac{4!}{2!1!}\frac{1!}{1!4!}\frac{1!}{3!3!} \pmod{5},$$

which is the same as

$$2 \equiv (-1)^3 \times 2 \times 3 \times 2 \times 4 \times 1 \pmod{5}.$$

With the matrices we get

$$(1\ 0)\begin{pmatrix} 0 & \frac{-1!}{3!(1+5-3)!} \\ 0 & \frac{-1!}{3!(1-3+5-1)!} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \binom{4}{2} & 0 \\ \frac{4!}{2!(4-2-1)!} & 0 \end{pmatrix}\begin{pmatrix} 0 & \frac{-0!}{3!(0+5-3)!} \\ 0 & \frac{-0!}{3!(0-3+5-1)!} \end{pmatrix}\begin{pmatrix} \binom{2}{0} & 0 \\ \frac{2!}{0!(2-0-1)!} & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv$$

$$(1\ 0)\begin{pmatrix} 0 & 4 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv 2 \pmod{5}.$$

Bibliography

[1] J.-P. Allouche and J. Shallit. The ring of k-regular sequences. *Theoret. Comput. Sci.*, 98(2):163–197, 1992.

[2] J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, 2003.

[3] J.-P. Allouche and J. Shallit. The ring of k-regular sequences, II. *Theoret. Comput. Sci.*, 307(1):3–29, 2003.

[4] J.-P. Allouche, F. von Haeseler, H.-O. Peitgen, and G. Skordev. Linear cellular automata, finite automata and Pascal's triangle. *Discrete Appl. Math.*, 66:1–22, 1996.

[5] J.-P. Allouche, F. von Haeselerb, H.-O. Peitgen, A. Petersen, and G. Skordev. Automaticity of double sequences generated by one-dimensional linear cellular automata. *Theoret. Comput. Sci.*, 188:195–209, 1997.

[6] H. Anton. Die Elferprobe und die Proben für die Moduln 9, 13 und 101. *Archiv Math. Physik*, 49:241–308, 1869.

[7] T. M. Apostol. *Introduction to Analytic Number Theory*. Springer, 1976.

[8] G. Barat and P. Grabner. Digital functions and distribution of binomial coefficients. *J. London Math. Soc*, 64:523–547, 2001.

[9] D. Barbolosi and P. Grabner. Distribution des coefficients multinomiaux et q-binomiaux modulo p. *Indag. Math.*, 7:129–135, 1996.

[10] M. Barnsley. *Fractals Everywhere*. Academic Press Professional, Inc., San Diego, CA, USA, second edition, 1993.

[11] B. Bates, M. Bunder, and K. Tognetti. Mirroring and Interleaving in the Paperfolding Sequence. *Appl. Anal.Discrete Math.*, 4(1):96–118, 2010.

[12] J. Berstel, A. Lauve, C. Reutenauer, and F. V. Saliola. *Combinatorics on Words: Christoffel Words and Repetitions in Words*. American Mathematical Soc., 2008.

[13] C. Brandt. Self-Similar Sets 5. Integer Matrices and Fractal Tilings of $\mathbb{R}^n$. *Proc. Amer. Math. Soc.*, 112(2):549–562, 1991.

[14] S. Brlek. Enumeration of factors in the Thue–Morse word. *Discrete Appl. Math.*, 24(13):83–96, 1989.

[15] N. Calkin and H. Wilf. Recounting the Rationals. *Amer. Math. Monthly*, 107(4):360–363, 2000.

[16] L. Carlitz. A note on Stirling numbers of the first kind. *Math. Mag.*, 37(5):318–321, 1964.

[17] L. Carlitz. Some partition problems related to the Stirling numbers of the second kind. *Acta Arith.*, 10(4):409–422, 1965.

[18] L. Carlitz. The number of binomial coefficients divisible by a fixed power of a prime. *Rend. Circ. Mat. Palermo*, 16:299–320, 1967.

[19] J. Cassaigne, J. Karhumäki, and A. Saarela. On Growth and Fluctuation of k-Abelian Complexity. In *Computer Science – Theory and Applications*, volume 9139 of *Lecture Notes in Computer Science*, pages 109–122. Springer International Publishing, 2015.

[20] A. Cobham. Uniform tag sequences. *Math. Syst. Theory*, 6(1-2):164–192, 1972.

[21] R. Cuppens. *Decomposition of Multivariate Probabilities*. Elsevier, 1975.

[22] C. Davis and D. E. Knuth. Number representations and dragon curves. *J. Recreat. Math.*, 3(3):133–149, 1970.

[23] A. de Luca and S. Varricchio. Some combinatorial properties of the Thue-Morse sequence and a problem in semigroups. *Theoret. Comput. Sci.*, 63(3):333–348, 1989.

[24] E. Delaygue. Arithmetic properties of Apéry-like numbers. arXiv:1310.4131v1, 2013.

[25] L. E. Dickson. *History of the Theory of Numbers*, volume 1. Carnegie Institute of Washington, 1919.

[26] G. Edgar. *Integral, Probability, and Fractal Measures*. Springer, 1998.

[27] S. Eilenberg. *Automata, Languages, and Machines*. Academic Press, Inc., Orlando, FL, USA, 1974.

[28] S. Eilenberg. *Automata, languages, and machines.*, volume A. Academic Press, New York and London, 1974.

[29] K. Falconer. *The geometry of fractal sets*. Cambridge University Press, 1985.

[30] N. J. Fine. Binomial coefficients modulo a prime. *Amer. Math. Monthly*, 54:589–592, 1947.

[31] R. Garfield and H. S. Wilf. The distribution of the binomial coefficients modulo p. *J. Number Theory*, 41:1–5, 1992.

[32] I. Gessel. Some congruences for Apéry numbers. *J. Number Theory*, 14(3):362 – 368, 1982.

[33] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2nd edition, 1994.

[34] F. Greinecker. Spatial equidistribution of combinatorial number schemes. submitted.

[35] F. Greinecker. On the 2-abelian complexity of the Thue-Morse word. *Theoret. Comput. Sci.*, 593:88–105, 2015.

[36] F. T. Howard. Prime divisors of q-binomial coefficients. *Rend. Sem. Mat. Univ. Padova*, 48:181–188, 1972.

[37] F. T. Howard. Congruences for the Stirling numbers and associated Stirling numbers . *Acta Arith.*, 55(1):29–41, 1990.

[38] J. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30:713–747, 1981.

[39] J. Karhumäki. Generalized Parikh mappings and homomorphisms. In *Automata, languages and programming*, volume 115 of *Lect. Notes in Comput. Sci.*, pages 324–332. Springer, Berlin-New York, 1981.

[40] J. Karhumäki, A. Saarela, and L. Zamboni. On a generalization of abelian equivalence and complexity of infinite words. *J. Comb. Theory Ser. A*, 120(8):2189–2206, Nov. 2013.

[41] J. Karhumäki, A. Saarela, and L. Zamboni. Variations of the Morse–Hedlund theorem for k-abelian equivalence. In *Developments in Language Theory*, volume 8633 of *Lecture Notes in Computer Science*, pages 203–214. Springer International Publishing, 2014. http://dx.doi.org/10.1007/978-3-319-09698-8_18.

[42] I. Kátai. Generalized number systems in Euclidian spaces. *Math. Comput. Modelling*, 38:883–892, 2003.

[43] A. T. Lundell. A divisibility property for Stirling numbers. *J. Number Theory*, 10:35–54, 1978.

[44] B. Madill and N. Rampersad. The abelian complexity of the paperfolding word. *Discrete Math.*, 313(7):831– 838, 2013.

[45] R. Meštrović. Lucas' theorem: its generalizations, extensions and applications (1878–2014). arXiv:1409.3820v1, 2014.

[46] A. Parreau, M. Rigo, E. Rowland, and E. Vandomme. A new approach to the 2-regularity of the ℓ-abelian complexity of 2-automatic sequences. *Electr. J. Comb.*, 22(1):1–27, 2015.

[47] A. Parreau, M. Rigo, and E. Vandomme. A conjecture on the 2-abelian compelexity of the Thue–Morse word. Talk at the conference "Representing Streams II", 2014. http://orbi.ulg.ac.be/handle/2268/162740.

[48] R. Peele, A.J. Radcliffe, and H. Wilf. Congruence problems involving Stirling numbers of the first kind. *Fibonacci Quart.*, 31:27–34, 1993.

[49] A. Rotenberg. SternBrocotTree. Licensed under CC BY-SA 3.0 via Commons. https://commons.wikimedia.org/wiki/File:SternBrocotTree.svg.

[50] W. Schmidt. *Equations over finite fields: an elementary approach.* Kendrick Press, 2004.

[51] D. Singmaster. Notes on binomial coefficients III-any integer divides almost all binomial coefficients. *J. London Math. Soc.*, 8:555–560, 1974.

[52] D. Singmaster. Divisibility of binomial and multinomial coefficients by primes and prime powers. In *in A collection of manuscripts related to the Fibonacci sequence*, pages 98–113. The Fibonacci Association, 1980.

[53] M. Sved. Divisibility – with visibility. *Math. Intelligencer*, 10(2):56–64, 1988.

[54] S. Wolfram. Geometry of binomial coefficients. *Amer. Math. Monthly*, 91(9):566–571, 1984.