Substitutive structures in combinatorics, number theory, and discrete geometry

Jana Lepšová

THÈSE EN COTUTELLE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEURE DE L'UNIVERSITÉ DE BORDEAUX ET DE CZECH TECHNICAL UNIVERSITY

ECOLE DOCTORALE MATHÉMATIQUES et INFORMATIQUE

FACULTY of NUCLEAR SCIENCES and PHYSICAL ENGINEERING

Par **Jana LEPŠOVÁ**

Structures substitutives en combinatoire, théorie des nombres et géométrie discrète

Substitutive structures in combinatorics, number theory and discrete geometry

Sous la direction de : Sébastien LABBÉ
Co-directrice : L'ubomíra DVOŘÁKOVÁ

Soutenue le 9 mai 2024

Membres du jury :

L'ubomíra DVOŘÁKOVÁ	Associate professor	Czech Technical University	Directrice
Sébastien LABBÉ	Chargé de Recherche	Université de Bordeaux, CNRS	Directeur
Émilie CHARLIER	Chargée de Cours	Université de Liège	Rapportrice
Dirk FRETTLÖH	Doctor	Universität Bielefeld	Rapporteur
Shigeki AKIYAMA	Professor	University of Tsukuba	Examinateur
Nicolas BÉDARIDE	Maître de Conférences	Université d'Aix Marseille	Examinateur
Zuzana MASÁKOVÁ	Professor	Czech Technical University	Présidente
Pascal WEIL	Directeur de Recherche	Université de Bordeaux, CNRS	Examinateur

Structures substitutives en combinatoire, théorie des nombres, et géométrie discrète

Résumé : Ce travail vise à découvrir et à développer des liens entre trois domaines mathématiques liés mais distincts : la combinatoire des mots, la théorie des nombres et la géométrie discrète. Du point de vue de la combinatoire des mots, nous étudions les mots finis et infinis et les morphismes, qui agissent comme des fonctions sur les mots. Les substitutions sont des morphismes qui satisfaisant certaines propriétés supplémentaires. Nous nous concentrons sur les mots sturmiens, les mots d'Arnoux-Rauzy et sur les morphismes sturmiens. Nous fournissons une formule pour déterminer l'exposant critique et l'exposant critique asymptotique de mots d'Arnoux–Rauzy réguliers. À l'aide de cette formule, il est possible de prouver que l'exposant critique minimal et l'exposant critique asymptotique minimal parmi les mots d'Arnoux–Rauzy réguliers d-aires sont atteints par le mot de d-bonacci. Nous introduisons une représentation fidèle du monoïde spécial de Sturm par des matrices 3×3 avec des entrées entières positives (y compris 0) qui ont une matrice d'incidence correspondante dans le coin supérieur gauche. À l'aide de cette représentation, nous abordons la question des racines carrées des points fixes des morphismes dans le monoïde spécial de Sturm. En ce qui concerne la théorie des nombres, nous étudions les numérations de position pour les nombres entiers positifs et négatifs: nous définissons un analogue de la notation du complément à deux pour $\mathbb{Z}$ en nous basant sur l'algorithme du complément à deux basé sur la séquence des nombres de Fibonacci. Nous l'appelons la numération du complément de Fibonacci pour $\mathbb{Z}$ et nous démontrons ses propriétés en ce qui concerne l'addition. Nous retrouvons la numération du complément de Fibonacci dans un autre contexte de systèmes de numération qui décrivent des points fixes et périodiques de substitutions. Nous appelons ces systèmes de numération systèmes de numération de Dumont-Thomas pour $\mathbb{Z}$, nous montrons qu'ils sont caractérisés par un ordre total particulier et qu'ils peuvent être naturellement étendus à $\mathbb{Z}^d$, $d \geq 1$. La géométrie discrète est présente sous la forme de pavages de Wang. En utilisant le système de numération du complément de Fibonacci étendu à $\mathbb{Z}^2$, nous caractérisons un pavage particulier du plan comme une séquence automatique.
Mots-clés : langage, pavage, dynamique symbolique, combinatoire

---

Substitutive structures in combinatorics, number theory, and discrete geometry

Abstract: This work aims to discover and develop links between three related but distinct mathematical domains: combinatorics on words, number theory and discrete geometry. From the point of view of combinatorics on words, we study finite and infinite words and morphisms, which act as maps on words. Substitutions are morphisms satisfying some additional properties. Namely, we focus on Sturmian and Arnoux–Rauzy words and on Sturmian morphisms. We provide a formula to determine both critical and asymptotic critical exponent of regular Arnoux–Rauzy words. With the help of this formula, we prove that the minimal critical and minimal asymptotic critical exponent among regular d-ary Arnoux–Rauzy words is attained by the d-bonacci word. We introduce a faithful representation of the special Sturmian monoid by 3×3 matrices with nonnegative integer entries, which enables us to tackle the question of the square roots of fixed points of morphisms in the special Sturmian monoid. As for the number theory, we study positional numeration systems for both nonnegative and negative integers: we define an analogue of the two's complement notation for $\mathbb{Z}$ based on the sequence of Fibonacci numbers. We call it the Fibonacci complement numeration system and we study its properties with respect to addition. We recover this positional numeration system in another context of numeration systems which describe fixed and periodic points of substitutions as automatic sequences. We call these numeration systems Dumont–Thomas numeration systems for $\mathbb{Z}$, we show that they are characterized by a particular total order and they extend naturally to $\mathbb{Z}^d$, $d \geq 1$. The discrete geometry is present in the form of Wang tilings. Using the Fibonacci complement numeration system extended to $\mathbb{Z}^2$, we characterize a particular tiling of the plane as an automatic sequence.
Keywords: language, tilings, symbolic dynamics, combinatorics

---

Substitutivní struktury v kombinatorice, teorii čísel a diskrétní geometrii

Abstrakt: Cílem této práce je zkoumat a rozvíjet souvislosti mezi třemi příbuznými matematickými obory: kombinatorikou na slovech, teorií čísel a diskrétní geometrií. Z hlediska kombinatoriky na slovech zkoumáme konečná a nekonečná slova a morfismy, které zobrazují slova na slova. Speciálním případem morfismů jsou substituce, které splňují určité další vlastnosti. Konkrétně se zabýváme sturmovskými a Arnouxovými–Rauzyovými slovy a sturmovskými morfismy. Prezentujeme vzorec pro výpočet kritického a asymptotického kritického exponentu regulárních Arnouxových–Rauzyových slov. Pomocí tohoto vzorce lze ukázat, že minimální kritický i minimální asymptotický kritický exponent mezi regulárními Arnouxovými–Rauzyovými slovy nad abecedou kardinality d je nabýván pro d-bonacciho slovo. Představíme věrnou reprezentaci speciálního sturmovského monoidu pomocí matic rozměru 3×3 s nezápornými celočíselnými prvky, která umožňuje určit morfismus, jehož pevným bodem je tzv. odmocnina z pevného bodu morfismu. Navíc popíšeme algoritmus určení věrné reprezentace, díky čemuž objasníme vztah mezi věrnými reprezentacemi vzájemně konjugovaných morfismů. Co se týče teorie čísel, studujeme poziční číselné soustavy pro celá čísla: definujeme poziční číselnou soustavu, která je analogická ke dvojkovému doplňku, ale namísto mocnin čísla 2 užívá Fibonacciho čísla. Tuto číselnou soustavu nazýváme Fibonacciho doplněk a popisujeme její vlastnosti vzhledem ke sčítání. V jiném kontextu ukážeme, že Fibonacciho doplněk patří mezi číselné soustavy, které popisují pevné a periodické body substitucí jako automatické posloupnosti. Tyto soustavy nazýváme Dumontovy–Thomasovy numerační systémy pro $\mathbb{Z}$, ukážeme, že jsou charakterizovány určitým úplným uspořádáním a lze je přirozeně rozšířit na $\mathbb{Z}^d$, $d \geq 1$. Diskrétní geometrie je zastoupena především ve formě Wangových dláždění. S pomocí Fibonacciho doplňku rozšířeného na $\mathbb{Z}^2$ charakterizujeme určité Wangovo dláždění roviny jako automatickou posloupnost.
Klíčová slova: jazyk, dláždění, symbolické dynamické systémy, kombinatorika

This thesis was written within the scope of a "Cotutelle contract" between the Czech Technical University in Prague and the University of Bordeaux. According to this agreement, the long abstract is written in English, French and Czech.

Long abstract

This work aims to discover and develop links between three related but distinct mathematical domains: combinatorics on words, number theory and discrete geometry. The representation of these domains in this work is the following. From the point of view of combinatorics on words, we study finite and infinite words and morphisms, which act as maps on words. Substitutions are morphisms satisfying some additional properties. Both morphisms and substitutions provide what we call a substitutive structure. Namely, we focus on Sturmian and Arnoux–Rauzy words and on Sturmian morphisms. As for the number theory, we study positional numeration systems for both nonnegative and negative integers. The discrete geometry is present in the form of Wang tilings – coverings of the plane by squares with colors on the edges, called Wang tiles, so that colors on the adjacent edges match. As the highlight of this text, we study a particular set of Wang tiles $\mathcal{Z}$ and we show its close relation to two-dimensional morphisms and numeration systems. Characterizing a particular tiling of the plane by the tiles $\mathcal{Z}$ as an automatic sequence, we make an unconventional link between the three domains. This text is structured into five chapters containing results of five scientific papers (including 2 conference papers), which have either been published or are under review in international scientific journals. We present the list of the papers in Section 1.1. Moreover, each chapter contains additional results where the author develops or generalizes the published or submitted papers.

Positional numeration systems: Numeration systems enable us to represent numbers by finite words over a suitable alphabet. In this text, we only consider the numeration systems which represent integers by finite words over an alphabet consisting of nonnegative integers. For instance, in the classical binary numeration system, a nonnegative integer $n \in \mathbb{N}$ is expressed as a sum of powers of 2, which gives rise to its representation over the binary alphabet $\Sigma = \{\mathtt{0}, \mathtt{1}\}$. The value map $\mathrm{val}_2 : \Sigma^* \to \mathbb{N}$ of the binary numeration system evaluates a word $w = w_{k-1}w_{k-2}\cdots w_1 w_0$ of length k over the binary alphabet as the sum

$$\mathrm{val}_2(w) = \sum_{i=0}^{k-1} w_i 2^i.$$

For every nonnegative integer $n \in \mathbb{N}$, there exists a unique binary word w which does not start with leading zeroes such that $\mathrm{val}_2(w) = n$. This word w is called the binary representation of n and denoted $\mathrm{rep}_2(n)$. One of the ways to generalize the binary numeration system to all integers is the two's complement notation [Knu98, §4.1]. Its value map $\mathrm{val}_{2c} : \Sigma^* \to \mathbb{Z}$ assigns to a binary word $w = w_{k-1}w_{k-2}\cdots w_1 w_0 \in \Sigma^k$ the difference between $\mathrm{val}_2(w)$ and the most significant digit w_{k-1} multiplied by the k-th power of 2

$$\mathrm{val}_{2c}(w) = -w_{k-1}2^k + \sum_{i=0}^{k-1} w_i 2^i.$$

The numeration systems which are based on an increasing sequence of nonnegative integers and which evaluate representations by multiplying letters in w with the elements in the sequence in a corresponding position are called positional. The so far presented positional numeration systems use the sequence of powers of 2, however, we can use other strictly increasing sequences such as the sequence of Fibonacci numbers $(F_i)_{i=0}^{+\infty}$. This sequence is defined by a recurrence

relation such that the next number is obtained as the sum of the previous two numbers, starting from $F_0 = 1$ and $F_1 = 2$. The value map $\mathrm{val}_{\mathcal{F}}$ of the Fibonacci numeration system for $\mathbb{N}$ evaluates a binary word w as the sum

$$\mathrm{val}_{\mathcal{F}}(w) = \sum_{i=0}^{k-1} w_i F_i.$$

In Chapter 3, we generalize the Fibonacci numeration system to $\mathbb{Z}$ in an analogous way to the two's complement notation and we call it the Fibonacci complement numeration system for $\mathbb{Z}$. Its value map $\mathrm{val}_{\mathcal{F}c} : \Sigma^* \to \mathbb{Z}$ assigns to a binary word w the sum

$$\mathrm{val}_{\mathcal{F}c}(w) = -w_{k-1}F_k + \sum_{i=0}^{k-1} w_i F_i.$$

We show that for every $n \in \mathbb{Z}$ there exists a unique odd-length word such that $n = \mathrm{val}_{\mathcal{F}c}(w)$, w does not contain consecutive ones and $w \notin \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*$.

Proposition 3.2.1 *The map* $\mathrm{val}_{\mathcal{F}c} \colon \Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*) \to \mathbb{Z}$ *is a bijection.*

The representation map $\mathrm{rep}_{\mathcal{F}c}$ of the Fibonacci complement numeration system is thus defined as the inverse map of $\mathrm{val}_{\mathcal{F}c}$. We show the properties of the Fibonacci complement numeration system with respect to addition. Moreover, we describe a new class of numeration systems for $\mathbb{Z}$ which contains both the two's complement numeration system and the Fibonacci complement numeration system. This is why we call the class the complement numeration systems. We characterize the representation map of every complement numeration system as a bijection increasing with respect to a particular total order $\prec$.

n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$
-10	1000100	0	0	10	0010010
-9	1000101	1	001	11	0010100
-8	1001000	2	010	12	0010101
-7	1001001	3	00100	13	0100000
-6	1001010	4	00101	14	0100001
-5	10000	5	01000	15	0100010
-4	10001	6	01001	16	0100100
-3	10010	7	01010	17	0100101
-2	100	8	0010000	18	0101000
-1	1	9	0010001	19	0101001

Sturmian and Arnoux–Rauzy words and morphisms: Sturmian words are a class of infinite aperiodic words over the binary alphabet, which belong to the most explored objects in combinatorics on words. They may be defined as the infinite words which contain $n+1$ distinct finite words (factors) of length n, for every $n \in \mathbb{N}$. An example of Sturmian words is the Fibonacci word which starts with the following prefix

$$\mathbf{f} = abaababaabaababaabab \cdots .$$

Words that generalize Sturmian words to d-ary alphabets for any integer $d \geq 2$ are called episturmian [DJP01]. Episturmian words are such that their language is closed under reversal and they have at most one right special factor of every length. The d-ary Arnoux–Rauzy words form a subclass of the episturmian words, which are called strict episturmian. An example of Arnoux–Rauzy words is the Tribonacci word which starts with the following prefix

$$\mathbf{t} = abacabaabacababacabaabacab \cdots .$$

A particular subclass of the Arnoux–Rauzy words are the d-bonacci words. The Fibonacci word is a d-bonacci word for $d = 2$ and the Tribonacci word is a d-bonacci word for $d = 3$.

The critical exponent of a right-infinite word expresses the maximal repetition rate of factors in the word. Similarly, the asymptotic critical exponent of a right-infinite word expresses the maximal repetition rate of factors in the word when their length grows to infinity. In Chapter 4, we present our results concerning the critical exponents of d-ary Arnoux–Rauzy words. We provide a formula to determine both critical and asymptotic critical exponent of regular Arnoux–Rauzy words. With the help of this formula, it is possible to prove that the minimal critical exponent and the minimal asymptotic critical exponent among regular d-ary Arnoux–Rauzy words is attained by the d-bonacci word. Moreover, in the special case of d-bonacci words for $4 \leq d \leq 15$, we show that their critical exponent coincides with their asymptotic critical exponent.

A morphism μ is a map which assigns to every letter a in an alphabet A a finite word over A and which fulfills the condition that $\mu(uv) = \mu(u)\mu(v)$, for every pair of finite words u, v over A. It can be naturally extended to right-infinite words $\mathbf{u} \in A^{\mathbb{N}}$ by the rule $\mu(u_0u_1u_2\cdots) = \mu(u_0)\mu(u_1)\mu(u_2)\cdots$. A right-infinite word $\mathbf{u} \in A^{\mathbb{N}}$ is called a fixed point of μ if $\mathbf{u} = \mu(\mathbf{u})$. Similarly, a right-infinite word $\mathbf{u} \in A^{\mathbb{N}}$ is called a periodic point of μ if there exists an integer $p \geq 1$ (called a period) so that $\mathbf{u} = \mu^p(\mathbf{u})$. Left-infinite and two-sided fixed and periodic points are defined in an analogous way. A letter $a \in A$ is called growing with respect to μ if the length of $\mu^n(a)$ grows to infinity when n goes to infinity. If $\mathbf{u} \in A^{\mathbb{Z}}$ is a two-sided word, we separate by a vertical bar its elements u_{-1} and u_0 to indicate the origin. If the letters u_{-1} and u_0 are growing, then the periodic point $\mathbf{u} = \mu^p(\mathbf{u})$ is defined entirely by its seed $u_{-1}|u_0$. More precisely, $\mathbf{u} = \lim_{k\to+\infty} \mu^{pk}(u_{-1})|\mu^{pk}(u_0)$.

Morphisms which map Sturmian words to Sturmian words are called Sturmian morphisms and they form the so-called monoid of Sturm. A particular submonoid of the monoid of Sturm is called the special monoid of Sturm. Morphisms over the binary alphabet can be represented by 2×2 matrices with nonnegative integer entries which are called incidence matrices. Two distinct morphisms might be represented with the same incidence matrix, meaning that this representation is not faithful. In Chapter 5, we introduce a faithful representation of the special Sturmian monoid by 3×3 matrices with nonnegative integer entries which have a corresponding incidence matrix in the top-left corner. With the help of this representation, we tackle the question of the so-called square roots of fixed points of morphisms in the special Sturmian monoid. Moreover, we describe an algorithm to determine the faithful representation of morphisms in the special Sturmian monoid, which clarifies the relationship between the faithful representations of mutually conjugate morphisms.

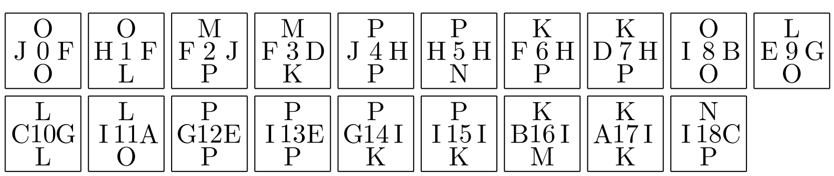
Automatic sequences: Cobham proved in 1972 that there is an equivalence between the right-infinite fixed points of k-uniform morphisms and the sequences obtained by feeding a deterministic finite automaton with output with the base-k representations of $\mathbb{N}$ [Cob72], [AS03, §6], where $k \geq 2$. A k-uniform morphism is such that $\mu(a)$ has length k, for every $a \in A$. More precisely, it follows from the Cobham's results that if $\mu : A^* \to A^*$ is a k-uniform morphism and $\mathbf{u}$ is the fixed point $\mathbf{u} = \mu(\mathbf{u})$ with growing letter $u_0 = a$, the letter at every position $n \in \mathbb{N}$ is obtained as $\mathcal{A}_{\mu,a}(\mathrm{rep}_k(n))$ where $\mathcal{A}_{\mu,a}$ is a deterministic finite automaton with output canonically associated with μ and a. The k-uniform morphisms fulfill naturally that they are not erasing and they have a growing letter. Morphisms having these two properties are called substitutions [Fog02]. The idea of Cobham was extended to the fixed points of all substitutions [RM02], using the deterministic finite automata with output canonically associated with substitutions and a much broader definition of numeration systems as regular languages. These numeration systems which might not be positional are called abstract numeration systems. Later on, a generalization for the fixed points of multidimensional morphisms was proposed [CKR10]. Another approach to define numeration systems for representing nonnegative integers, as well as real numbers in a certain interval, were proposed in [DT89] by Dumont and Thomas. Every Dumont–Thomas numeration system for $\mathbb{N}$ is based on a fixed point of a substitution.

In Chapter 6, we extend the Dumont–Thomas numeration systems for $\mathbb{N}$ to $\mathbb{Z}$. We show that every two-sided periodic point of a substitution with a growing seed is an automatic sequence.

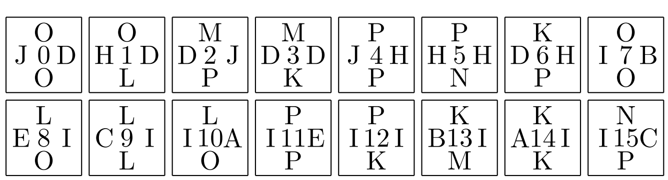
Theorem 6.3.1 *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u}$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Then, for every $n \in \mathbb{Z}$, we have $u_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_{\mathbf{u}}(n))$.*

The deterministic finite automaton emerges from the canonical automaton $\mathcal{A}_{\eta,a}$ by adding a new initial state `start` and two extra edges. Also, we prove that the Dumont–Thomas numeration systems for $\mathbb{Z}$ are bijections increasing with the total order $\prec$ defined in Chapter 3. Thanks to this property, we recover the two's complement numeration system and the Fibonacci complement numeration system as Dumont–Thomas numeration systems for $\mathbb{Z}$. As a new unpublished result, we show a sufficient condition for the Dumont–Thomas numeration systems to be positional, linking them to the positional numeration systems described in Chapter 3. Finally, we show that these numeration systems can be naturally extended to $\mathbb{Z}^d$ for every $d \geq 1$. The Fibonacci complement numeration system extended to $\mathbb{Z}^2$ is of particular interest as it may be used to describe a Wang tiling as an automatic sequence which we show in Chapter 7.

Wang tilings: A Wang tile is a unit square with a color on each edge. Given a finite set of Wang tiles, we assume that we have infinitely many copies of them and we are allowed to arrange two Wang tiles side by side (without rotating them) provided that the colors on their common edge match. In general, we are interested in such sets of Wang tiles which tile the plane, but which do not do so in a periodic way. Given a set of Wang tiles, the set of all tilings of the plane is called a Wang shift and a Wang shift is called aperiodic if it does not contain a periodic tiling. Wang shifts have been studied since 1961 and since then aperiodic Wang shifts with gradually smaller tile sets were discovered. Finally in 2021, Jeandel and Rao described an aperiodic Wang shift based on 11 tiles and they proved that every set with less than 11 Wang tiles admits a periodic tiling [JR21]. Shortly after, a set of 19 Wang tiles $\mathcal{U}$ which is aperiodic, minimal and self-similar was constructed based on the Jeandel–Rao Wang set [Lab21, Lab19].



In Chapter 7, we study a particular set $\mathcal{Z}$ of 16 Wang tiles which emerged from $\mathcal{U}$ by identifying some colors. As a new unpublished result, we show that the Wang shift $\Omega_{\mathcal{Z}}$ is topologically conjugate to the Wang shift $\Omega_{\mathcal{U}}$. As a corollary, we have that $\Omega_{\mathcal{Z}}$ is aperiodic, minimal and self-similar.



We show that the set $\mathcal{Z}$ admits a Wang tiling of the plane described by a deterministic finite automaton with output, which takes as input the representation of a position $\boldsymbol{n} \in \mathbb{Z}^2$ in the Fibonacci complement numeration system extended to $\mathbb{Z}^2$ and outputs a Wang tile.

Theorem 7.4.1 *There exists a deterministic finite automaton with output $\mathcal{A}$ such that the configuration*

$$\begin{array}{rccl} x: & \mathbb{Z}^2 & \to & \{0,1,\dots,15\} \\ & \boldsymbol{n} & \mapsto & \mathcal{A}(\mathrm{rep}_{\mathcal{F}c}(\boldsymbol{n})) \end{array}$$

satisfies the condition that $x \in \Omega_{\mathcal{Z}}$.

Résumé long

Ce travail vise à découvrir et développer des liens entre trois domaines mathématiques liés mais distincts: la combinatoire des mots, la théorie des nombres et la géométrie discrète. La représentation de ces domaines dans ce travail est la suivante. Du point de vue de la combinatoire des mots, nous étudions les mots finis et infinis et les morphismes, qui agissent comme des fonctions sur les mots. Les substitutions sont des morphismes satisfaisant certaines propriétés supplémentaires. Les morphismes et les substitutions fournissent ce qu'on appelle une structure substitutive. Nous nous concentrons sur les mots sturmiens, les mots d'Arnoux–Rauzy et sur les morphismes sturmiens. En ce qui concerne la théorie des nombres, nous étudions les numérations de position pour les nombres entiers. La géométrie discrète est présente sous la forme de pavages de Wang. Une tuile de Wang est un carré avec des couleurs sur les bords. Un pavage de Wang est un recouvrement du plan par des tuiles de Wang, de manière à ce que les couleurs des bords adjacents correspondent. Nous étudions un ensemble particulier de tuiles de Wang $\mathcal{Z}$ et nous démontrons son rapport avec les morphismes multidimensionnels et les systèmes de numération. Caractérisant un pavage particulier du plan par les tuiles $\mathcal{Z}$ comme une séquence automatique, nous établissons un lien non-conventionnel entre les trois domaines. Cette thèse est structurée en cinq chapitres contenant les résultats de cinq articles (dont 2 articles de conférence), qui ont été publiés ou sont en cours de révision. Nous présentons la liste des articles dans la Section 1.1. Chaque chapitre contient des résultats supplémentaires où l'auteur développe ou généralise les articles publiés ou soumis.

Numération de position: Les systèmes de numération nous permettent de représenter les nombres par des mots finis sur un alphabet approprié. Dans ce texte, nous ne considérons que les systèmes de numération qui représentent les entiers par des mots finis sur un alphabet composé d'entiers positifs (y compris 0). Par exemple, dans le système de numération binaire classique, un entier positif $n \in \mathbb{N}$ est exprimé comme une somme de puissances de 2, ce qui donne lieu à sa représentation sur l'alphabet binaire $\Sigma = \{\mathtt{0}, \mathtt{1}\}$. La fonction $\mathrm{val}_2 : \Sigma^* \to \mathbb{N}$ du système de numération binaire évalue un mot $w = w_{k-1}w_{k-2}\cdots w_1w_0$ de longueur k sur l'alphabet binaire comme la somme

$$\mathrm{val}_2(w) = \sum_{i=0}^{k-1} w_i 2^i.$$

Pour tout entier positif $n \in \mathbb{N}$, il existe un unique mot binaire w qui ne commence pas par des zéros initiaux, tel que $\mathrm{val}_2(w) = n$. Ce mot w est appelé représentation binaire de n et noté $\mathrm{rep}_2(n)$. Une des façons de généraliser le système de numération binaire à tous les entiers, y compris les négatifs, est la notation du complément à deux [Knu98, §4.1]. La fonction val_{2c} : $\Sigma^* \to \mathbb{Z}$ attribue à un mot binaire $w = w_{k-1}w_{k-2}\cdots w_1w_0 \in \Sigma^k$ la différence entre $\mathrm{val}_2(w)$ et le

chiffre le plus significatif w_{k-1} multiplié par la k-ième puissance de 2

$$\mathrm{val}_{2c}(w) = -w_{k-1}2^k + \sum_{i=0}^{k-1} w_i 2^i.$$

Les systèmes de numération qui sont basés sur une séquence croissante d'entiers positifs et qui évaluent les représentations en multipliant les lettres de w avec les éléments de la séquence à une position correspondante, sont appelés numérations de position. Les numérations de position présentées jusqu'à présent utilisent la séquence des puissances de 2. Cependant, on peut utiliser d'autres séquences strictement croissantes comme la suite des nombres de Fibonacci $(F_i)_{i=0}^{+\infty}$. Cette suite est définie par une relation de récurrence telle que chaque terme est obtenu comme la somme des deux termes précédents, en commençant par $F_0 = 1$ et $F_1 = 2$. La fonction $\mathrm{val}_{\mathcal{F}}$ de la numération de Fibonacci pour $\mathbb{N}$ évalue un mot binaire w comme la somme

$$\mathrm{val}_{\mathcal{F}}(w) = \sum_{i=0}^{k-1} w_i F_i.$$

Dans le chapitre 3, nous généralisons la numération de Fibonacci à $\mathbb{Z}$ de manière analogue à la notation du complément à deux et nous l'appelons la numération du complément de Fibonacci pour $\mathbb{Z}$. Notons la $\mathcal{F}c$. Sa fonction $\mathrm{val}_{\mathcal{F}c} : \Sigma^* \to \mathbb{Z}$ assigne à un mot binaire w la somme

$$\mathrm{val}_{\mathcal{F}c}(w) = -w_{k-1}F_k + \sum_{i=0}^{k-1} w_i F_i.$$

Nous démontrons que pour chaque $n \in \mathbb{Z}$ il existe un mot unique de longueur impaire tel que $n = \mathrm{val}_{\mathcal{F}c}(w)$, w ne contient pas de 1 consécutifs et $w \notin \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*$.

Proposition 3.2.1 *La fonction* $\mathrm{val}_{\mathcal{F}c} \colon \Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*) \to \mathbb{Z}$ *est une bijection.*

La fonction de représentation $\mathrm{rep}_{\mathcal{F}c}$ de la numération $\mathcal{F}c$ est donc définie comme la fonction inverse de $\mathrm{val}_{\mathcal{F}c}$. Nous démontrons les propriétés de la numération $\mathcal{F}c$ en ce qui concerne l'addition. De plus, nous décrivons une nouvelle classe de numération pour $\mathbb{Z}$ qui contient à la fois le complément à deux et la numération du complément de Fibonacci. C'est pourquoi nous appelons cette classe systèmes de numération complémentaire. Nous caractérisons la fonction de représentation de chaque système de numération complémentaire par une bijection croissante relativement à un ordre total particulier $\prec$.

n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$
-10	1000100	0	0	10	0010010
-9	1000101	1	001	11	0010100
-8	1001000	2	010	12	0010101
-7	1001001	3	00100	13	0100000
-6	1001010	4	00101	14	0100001
-5	10000	5	01000	15	0100010
-4	10001	6	01001	16	0100100
-3	10010	7	01010	17	0100101
-2	100	8	0010000	18	0101000
-1	1	9	0010001	19	0101001

Mots sturmiens, mots d'Arnoux–Rauzy, morphismes sturmiens Let mots sturmiens sont une classe de mots apériodiques infinis sur l'alphabet binaire, qui font partie des objets les plus explorés en combinatoire des mots. Ils peuvent être définis comme les mots infinis qui contiennent $n+1$ mots finis distincts (facteurs) de longueur $n+1$, pour chaque $n \in \mathbb{N}$. Un exemple de mots sturmiens est le mot de Fibonacci qui commence par le préfixe suivant

$$\mathbf{f} = abaababaabaababaabab \cdots .$$

Les mots généralisant les mots sturmiens aux alphabets d-aires pour tout entier $d \geq 2$ sont appelés épisturmiens [DJP01]. Les mots d'Arnoux–Rauzy d-aires forment une sous-classe des mots épisturmiens, qui sont appelés mots épisturmiens stricts. Un exemple de mots d'Arnoux–Rauzy est le mot de Tribonacci qui commence par le préfixe suivant

$$\mathbf{t} = abacabaabacababacabaabacab \cdots .$$

Une sous-classe particulière de mots d'Arnoux–Rauzy sont les mots de d-bonacci. Le mot de Fibonacci est un mot de d-bonacci pour $d = 2$ et le mot de Tribonacci est un mot de d-bonacci pour $d = 3$.

L'exposant critique d'un mot infini exprime le taux de répétition maximal des facteurs dans le mot. De même, l'exposant critique asymptotique d'un mot infini exprime le taux de répétition maximal des facteurs dans le mot lorsque leur longueur croît à l'infini. Dans le chapitre 4, nous présentons nos résultats concernant les exposants critiques de mots d'Arnoux–Rauzy d-aires. Nous fournissons une formule pour déterminer l'exposant critique et l'exposant critique asymptotique de mots d'Arnoux–Rauzy réguliers. À l'aide de cette formule, il est possible de prouver que l'exposant critique minimal et l'exposant critique asymptotique minimal parmi les mots d'Arnoux–Rauzy réguliers d-aires sont atteints par le mot de d-bonacci. De plus, dans le cas de mots de d-bonacci pour $4 \leq d \leq 15$, nous démontrons que leur exposant critique coïncide avec leur exposant critique asymptotique.

Un morphisme μ est une fonction qui assigne à chaque lettre a dans un alphabet A un mot fini sur A et qui remplit la condition que $\mu(uv) = \mu(u)\mu(v)$, pour toute paire de mots finis u, v sur A. Elle peut être naturellement étendue aux mots infinis vers la droite $\mathbf{u} \in A^{\mathbb{N}}$ par la règle $\mu(u_0u_1u_2\cdots) = \mu(u_0)\mu(u_1)\mu(u_2)\cdots$. Un mot infini $\mathbf{u} \in A^{\mathbb{N}}$ est appelé un point fixe de μ si $\mathbf{u} = \mu(\mathbf{u})$. De même, un mot infini $\mathbf{u} \in A^{\mathbb{N}}$ est appelé un point périodique de μ s'il existe un entier $p \geq 1$ (appelé période) de sorte que $\mathbf{u} = \mu^p(\mathbf{u})$. Les points fixes et les points périodiques infinis vers la gauche et bi-infinis sont définis d'une manière analogue. Une lettre $a \in A$ est appelé croissante par rapport à μ si la longueur de $\mu^n(a)$ croît à l'infini quand n tend vers l'infini. Si $\mathbf{u} \in A^{\mathbb{Z}}$ est un mot bi-infini, nous séparons par une barre verticale ses éléments u_{-1} et u_0 pour indiquer l'origine. Si les lettres u_{-1} et u_0 sont croissantes, alors le point périodique $\mathbf{u} = \mu^p(\mathbf{u})$ est défini entièrement par son germe $u_{-1}|u_0$. Plus précisément, $\mathbf{u} = \lim_{k\to+\infty} \mu^{pk}(u_{-1})|\mu^{pk}(u_0)$.

Les morphismes qui font correspondre des mots sturmiens à des mots sturmiens sont appelés morphismes sturmiens et forment ce que l'on appelle le monoïde de Sturm. Un sous-monoïde particulier du monoïde de Sturm est appelé le monoïde spécial de Sturm. Les morphismes sur l'alphabet binaire peuvent être représentés par des matrices 2×2 avec des entrées entières positives (y compris 0), appelées matrices d'incidence. Deux morphismes distincts peuvent être représentés par la même matrice d'incidence, ce qui signifie que cette représentation n'est pas fidèle. Dans le chapitre 5, nous introduisons une représentation fidèle du monoïde spécial de Sturm par des matrices 3×3 avec des entrées entières positives (y compris 0) qui ont une matrice d'incidence correspondante dans le coin supérieur gauche. À l'aide de cette représentation, nous abordons la question des racines carrées des points fixes des morphismes dans le monoïde spécial de Sturm. De plus, nous décrivons un algorithme pour déterminer la représentation fidèle, ce qui clarifie la relation entre les représentations fidèles de morphismes mutuellement conjugués.

Suites automatiques: Cobham a prouvé en 1972 qu'il existe une équivalence entre les points fixes infinis vers la droite de morphismes k-uniformes et les suites k-automatiques, où $k \geq 2$ [Cob72], [AS03, §6]. Un morphisme k-uniforme est tel que $\mu(a)$ a une longueur k, pour tout $a \in A$. Il est possible d'associer canoniquement à μ et a un automate fini déterministe avec une sortie noté $\mathcal{A}_{\mu,a}$. Il découle des résultats de Cobham que si $\mu : A^* \to A^*$ est un morphisme k-uniforme et que $\mathbf{u} \in A^{\mathbb{N}}$ est le point fixe $\mathbf{u} = \mu(\mathbf{u})$ avec la lettre croissante $u_0 = a$, la lettre à chaque position $n \in \mathbb{N}$ dans $\mathbf{u}$ est obtenue comme $\mathcal{A}_{\mu,a}(\mathrm{rep}_k(n))$. Les morphismes k-uniformes ont naturellement deux propriétés: ils ne sont pas effaçants et ils ont une lettre croissante. Les morphismes ayant ces deux propriétés sont appelés substitutions. L'idée de Cobham a été étendue aux points fixes de toutes les substitutions [RM02] avec une définition des systèmes de numération plus abstraite. Ces systèmes de numération, qui peuvent ne pas être des numérations de position, sont appelés systèmes de numération abstraits. Plus tard, une généralisation pour les points fixes des morphismes multidimensionnels a été proposée [CKR10]. Une autre approche pour définir des systèmes de numération pour représenter les entiers positifs, ainsi que les réels dans un certain intervalle, a été proposée dans [DT89] par Dumont et Thomas. Chaque système de numération Dumont–Thomas pour $\mathbb{N}$ est basé sur un point fixe d'une substitution.

Dans le chapitre 6, nous étendons les systèmes de numération de Dumont–Thomas pour $\mathbb{N}$ à $\mathbb{Z}$. Nous démontrons que chaque point périodique bi-infini d'une substitution avec un germe croissant est une suite automatique.

Theorem 6.3.1 *Soit $\eta : A^* \to A^*$ une substitution et $\mathbf{u}$ son point périodique bi-infini avec un germe croissant $s = u_{-1}|u_0$. Alors, pour tout $n \in \mathbb{Z}$, $u_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_{\mathbf{u}}(n))$.*

L'automate fini déterministe $\mathcal{A}_{\eta,s}$ émerge de l'automate canonique $\mathcal{A}_{\eta,a}$ en ajoutant un nouvel état initial `start` et deux arêtes supplémentaires. Nous prouvons également que les systèmes de numération de Dumont–Thomas pour $\mathbb{Z}$ sont des bijections croissantes avec l'ordre total $\prec$ défini au chapitre 3. Grâce à cette propriété, nous retrouvons le système de numération du complément à deux et le système de numération du complément de Fibonacci comme systèmes de numération de Dumont–Thomas pour $\mathbb{Z}$. Dans un nouveau résultat non publié, nous démontrons une condition suffisante pour que les systèmes de numération de Dumont–Thomas pour $\mathbb{Z}$ soient des numérations de position, en les reliant aux systèmes de numération de position décrits dans le chapitre 3. Enfin, nous démontrons que ces systèmes de numération peuvent être naturellement étendus à $\mathbb{Z}^d$ pour tout $d \geq 1$. Le système de numération du complément de Fibonacci étendu à $\mathbb{Z}^2$ est particulièrement intéressant, car il peut être utilisé pour décrire un pavage de Wang comme une suite automatique, ce que nous démontrons dans le chapitre 7.

Pavages de Wang: Une tuile de Wang est un carré unitaire dont chaque bord est coloré. Étant donné un ensemble fini de tuiles de Wang, nous supposons que nous en avons une infinité de copies et nous sommes autorisés à disposer deux tuiles de Wang côte à côte (sans les faire pivoter) à condition que les couleurs de leur bord commun correspondent. En général, nous nous intéressons à de tels ensembles de tuiles de Wang qui tapissent le plan, mais qui ne le font pas de manière périodique. Étant donné un ensemble de tuiles de Wang, l'ensemble de touts les pavages du plan est appelé un sous-shift de Wang et celui-ci est dit apériodique s'il ne contient pas de pavage périodique. Les sous-shifts de Wang sont étudiés depuis 1961 et depuis lors, des sous-shifts de Wang apériodiques avec des ensembles de tuiles de plus en plus petits ont été découverts. Enfin, en 2021, Jeandel et Rao ont décrit un sous-shift de Wang apériodique basé sur un ensemble de 11 tuiles et ils ont prouvé que tout ensemble de moins de 11 tuiles de Wang admet un pavage périodique [JR21]. Peu de temps après, un ensemble de 19 tuiles de Wang $\mathcal{U}$ qui est apériodique, minimal et autosimilaire a été construit. Il est basé sur l'ensemble de Wang de Jeandel et Rao [Lab21, Lab19].

O J 0 F O	O H 1 F L	M F 2 J P	M F 3 D K	P J 4 H P	P H 5 H N	K F 6 H P	K D 7 H P	O I 8 B O	L E 9 G O
L C 10 G L	L I 11 A O	P G 12 E P	P I 13 E P	P G 14 I K	P I 15 I K	K B 16 I M	K A 17 I K	N I 18 C P	

Dans le chapitre 7, nous étudions un ensemble particulier $\mathcal{Z}$ de 16 tuiles de Wang qui a émergé de $\mathcal{U}$ en identifiant certaines couleurs. Dans un résultat non publié, nous démontrons que le sous-shift de Wang $\Omega_{\mathcal{Z}}$ est topologiquement conjugué au sous-shift de Wang $\Omega_{\mathcal{U}}$. En corollaire, le sous-shift $\Omega_{\mathcal{Z}}$ est apériodique, minimal et autosimilaire.

O J 0 D O	O H 1 D L	M D 2 J P	M D 3 D K	P J 4 H P	P H 5 H N	K D 6 H P	O I 7 B O
L E 8 I O	L C 9 I L	L I 10 A O	P I 11 E P	P I 12 I K	K B 13 I M	K A 14 I K	N I 15 C P

Nous démontrons que l'ensemble $\mathcal{Z}$ admet un pavage de Wang du plan décrit par un automate fini déterministe avec une sortie, qui prend en entrée la représentation d'une position $\boldsymbol{n} \in \mathbb{Z}^2$ dans le système de numération du complément de Fibonacci étendu à $\mathbb{Z}^2$ et produit ainsi une tuile de Wang.

Theorem 7.4.1 *Il existe un automate fini déterministe avec une sortie $\mathcal{A}$ tel que*

$$\begin{array}{rccl} x: & \mathbb{Z}^2 & \to & \{0,1,\ldots,15\} \\ & \boldsymbol{n} & \mapsto & \mathcal{A}(\mathrm{rep}_{\mathcal{F}c}(\boldsymbol{n})) \end{array}$$

satisfait à la condition $x \in \Omega_{\mathcal{Z}}$.

Rozšířený abstrakt

Cílem této práce je zkoumat a rozvíjet souvislosti mezi třemi příbuznými matematickými obory: kombinatorikou na slovech, teorií čísel a diskrétní geometrií. Zastoupení těchto oborů v této práci je následující. Z hlediska kombinatoriky na slovech zkoumáme konečná a nekonečná slova a morfismy, které zobrazují slova na slova. Speciálním případem morfismů jsou substituce, které splňují určité další vlastnosti. Jak morfismy, tak substituce poskytují takzvanou substitutivní strukturu. Konkrétně se zabýváme sturmovskými a Arnouxovými–Rauzyovými slovy a sturmovskými morfismy. Co se týče teorie čísel, studujeme poziční číselné soustavy pro celá čísla. Diskrétní geometrie je zastoupena především ve formě Wangových dláždění. Wangova dlaždice je čtverec o jednotkové ploše, který má na každé hraně barvu. Wangovo dláždění je pokrytí roviny Wangovými dlaždicemi tak, aby barvy na dotýkajících se hranách sousedních dlaždic byly stejné. Jeden z výrazných výsledků tohoto textu se týká konkrétní množiny Wangových dlaždic $\mathcal{Z}$, o níž ukážeme, že úzce souvisí s 2-dimenzionálními morfismy a číselnými soustavami. Popíšeme konkrétní dláždění roviny Wangovými dlaždicemi $\mathcal{Z}$ jako automatickou posloupnost, čímž získáme nekonvenční souvislost mezi kombinatorikou na slovech, teorií čísel a diskrétní geometrií. Tento text je členěn do 5 kapitol, z nichž každá obsahuje výsledky některého z 5 odborných článků (včetně dvou příspěvků v konferenčním sborníku), které byly publikovány v mezinárodních impaktovaných časopisech nebo jsou v recenzním řízení. Seznam těchto

článků uvádíme v Sekci 1.1. Každá kapitola obsahuje navíc další výsledky, které zobecňují nebo jinak rozvíjejí výsledky dosažené v článcích.

Poziční číselné soustavy: Číselné soustavy umožňují reprezentaci čísel konečnými slovy nad vhodnou abecedou. V tomto textu se zabýváme pouze číselnými soustavami, které reprezentují celá čísla konečnými slovy nad abecedami, které sestávají z nezáporných celých čísel. Například pro reprezentaci nezáporného celého čísla $n \in \mathbb{N}$ v binární (dvojkové) číselné soustavě je n vyjádřeno jako součet mocnin čísla 2. Binární reprezentace n je konečné slovo nad binární abecedou $\Sigma = \{\mathtt{0}, \mathtt{1}\}$. Zobrazení $\mathrm{val}_2 : \Sigma^* \to \mathbb{N}$ binární číselné soustavy vyčísluje konečné slovo $w = w_{k-1}w_{k-2}\cdots w_1w_0$ délky k nad binární abecedou jako součet

$$\mathrm{val}_2(w) = \sum_{i=0}^{k-1} w_i 2^i.$$

Pro každé nezáporné celé číslo $n \in \mathbb{N}$ existuje právě jedno slovo w nad binární abecedou, které nezačíná nulami a platí, že $\mathrm{val}_2(w) = n$. Toto slovo w nazýváme binární reprezentací čísla n a značíme $w = \mathrm{rep}_2(n)$. Jedním ze způsobů rozšíření binární číselné soustavy pro všechna celá čísla (včetně záporných celých čísel) je dvojkový doplněk či také doplňkový kód [Knu98, §4.1]. Zobrazení $\mathrm{val}_{2c} : \Sigma^* \to \mathbb{Z}$ přiřazuje slovu nad binární abecedou $w = w_{k-1}w_{k-2}\cdots w_1w_0 \in \Sigma^k$ rozdíl mezi $\mathrm{val}_2(w)$ a nejvýznamnější číslicí w_{k-1} vynásobenou k-tou mocninou čísla 2

$$\mathrm{val}_{2c}(w) = -w_{k-1}2^k + \sum_{i=0}^{k-1} w_i 2^i.$$

Číselné soustavy, které používají jako bázi ostře rostoucí posloupnost nezáporných celých čísel a které vyčíslují reprezentace jako součet násobků číslic ve w se členy posloupnosti na odpovídající pozici, jsou nazývány poziční číselné soustavy. Poziční číselné soustavy, které byly dosud představeny, používají jako bázi posloupnost mocnin čísla 2, nicméně lze použít i jiné ostře rostoucí posloupnosti, jako například posloupnost Fibonacciho čísel $(F_i)_{i=0}^{+\infty}$. Tato posloupnost je definována následujícím rekurentním vzorcem - další člen posloupnosti se spočte jako součet předchozích dvou členů. Počáteční podmínky jsou $F_0 = 1$ a $F_1 = 2$. Zobrazení $\mathrm{val}_{\mathcal{F}}$ Fibonacciho číselné soustavy pro $\mathbb{N}$ vyčísluje slovo w nad binární abecedou jako součet

$$\mathrm{val}_{\mathcal{F}}(w) = \sum_{i=0}^{k-1} w_i F_i.$$

V kapitole 3 rozšíříme Fibonacciho číselnou soustavu na $\mathbb{Z}$ analogickým způsobem k dvojkovému doplňku. Proto tuto číselnou soustavu nazýváme Fibonacciho doplňkem (Fibonacci complement numeration system). Zobrazení $\mathrm{val}_{\mathcal{F}c} : \Sigma^* \to \mathbb{Z}$ přiřazuje slovu w nad binární abecedou součet

$$\mathrm{val}_{\mathcal{F}c}(w) = -w_{k-1}F_k + \sum_{i=0}^{k-1} w_i F_i.$$

Ukážeme, že pro každé celé číslo $n \in \mathbb{Z}$ existuje právě jedno slovo nad binární abecedou liché délky takové, že $n = \mathrm{val}_{\mathcal{F}c}(w)$, w neobsahuje po sobě jdoucí jedničky a $w \notin \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*$.

Proposition 3.2.1 *Zobrazení* $\mathrm{val}_{\mathcal{F}c}\colon \Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*) \to \mathbb{Z}$ *je bijekce.*

Zobrazení $\mathrm{rep}_{\mathcal{F}c}$ Fibonacciho doplňku je pak definováno jako inverzní zobrazení k zobrazení $\mathrm{val}_{\mathcal{F}c}$. Ukážeme vlastnosti Fibonacciho doplňku vzhledem ke sčítání. Navíc popíšeme novou třídu pozičních číselných soustav pro $\mathbb{Z}$, která obsahuje jak dvojkový doplněk, tak Fibonacciho doplněk. Proto tyto číselné soustavy nazýváme doplňkové (complement numeration systems). Každou doplňkovou číselnou soustavu popíšeme jako jazyk uspořádaný vzhledem ke konkrétnímu úplnému uspořádání $\prec$.

n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$
-10	1000100	0	0	10	0010010
-9	1000101	1	001	11	0010100
-8	1001000	2	010	12	0010101
-7	1001001	3	00100	13	0100000
-6	1001010	4	00101	14	0100001
-5	10000	5	01000	15	0100010
-4	10001	6	01001	16	0100100
-3	10010	7	01010	17	0100101
-2	100	8	0010000	18	0101000
-1	1	9	0010001	19	0101001

Sturmovská a Arnouxova–Rauzyova slova a morfismy: Sturmovská slova jsou třídou nekonečných aperiodických slov nad binární abecedou, které patří k nejlépe prozkoumaným objektům v kombinatorice na slovech. Jedna z možných ekvivalentních definic je definuje jako nekonečná slova, která obsahují $n+1$ různých konečných slov (faktorů) délky n pro každé $n \in \mathbb{N}$. Příkladem sturmovských slov je Fibonacciho slovo, které začíná následujícím prefixem

$$\mathbf{f} = abaababaabaababaabab \cdots .$$

Slova, která zobecňují sturmovská slova pro abecedy kardinality $d \geq 2$, se nazývají episturmovská [DJP01]. Arnouxova–Rauzyova slova tvoří podtřídu episturmovských slov, která se nazývají také striktně episturmovská. Příkladem Arnouxových–Rauzyových slov je Tribonacciho slovo, které začíná následujícím prefixem

$$\mathbf{t} = abacabaabacababacabaabacab \cdots .$$

Speciální podtřídou Arnouxových–Rauzyových slov jsou d-bonacciho slova. Fibonacciho slovo je d-bonacciho slovo pro $d = 2$ a Tribonacciho slovo je d-bonacciho slovo for $d = 3$.

Kritický exponent nekonečného slova vyjadřuje maximální možnou míru opakování faktorů v daném nekonečném slově. Podobně asymptotický kritický exponent nekonečného slova vyjadřuje maximální možnou míru opakování faktorů v daném nekonečném slově tak, že délka faktorů roste limitně do nekonečna. V kapitole 4 shrnujeme naše výsledky týkající se kritických exponentů Arnouxových–Rauzyových slov. Prezentujeme vzorec pro výpočet kritického a asymptotického kritického exponentu regulárních Arnouxových–Rauzyových slov. Pomocí tohoto vzorce lze ukázat, že minimální kritický i minimální asymptotický kritický exponent mezi regulárními Arnouxovými–Rauzyovými slovy nad abecedou kardinality d je nabýván pro d-bonacciho slovo. Navíc pro speciální případ d-bonacciho slov pro $4 \leq d \leq 15$ ukážeme, že jejich kritický exponent splývá s jejich asymptotickým kritickým exponentem.

Morfismus μ je zobrazení, které přiřazuje písmenu a v abecedě A konečné slovo nad abecedou A a které splňuje, že $\mu(uv) = \mu(u)\mu(v)$ pro každou dvojici konečných slov u, v nad A. Morfismus můžeme přirozeně rozšířit na nekonečná slova $\mathbf{u} \in A^{\mathbb{N}}$ pomocí pravidla $\mu(u_0u_1u_2\cdots) = \mu(u_0)\mu(u_1)\mu(u_2)\cdots$. Nekonečné slovo $\mathbf{u} \in A^{\mathbb{N}}$ se nazývá pevný bod morfismu μ, pokud platí, že $\mathbf{u} = \mu(\mathbf{u})$. Podobně se nekonečné slovo $\mathbf{u} \in A^{\mathbb{N}}$ nazývá periodický bod morfismu μ, pokud platí, že existuje celé číslo $p \geq 1$ takové, že $\mathbf{u} = \mu^p(\mathbf{u})$. Dosud jsme uvažovali jednostranně nekonečná slova $\mathbf{u} \in A^{\mathbb{N}}$. Podobně lze uvažovat oboustranně nekonečná slova $\mathbf{u} \in A^{\mathbb{Z}}$ a definovat oboustranně nekonečné pevné a periodické body morfismů analogickým způsobem. Písmeno

$a \in A$ nazývá rostoucí vzhledem k μ jestliže délka $\mu^n(a)$ roste do nekonečna, pokud n roste do nekonečna. V případě oboustranně nekonečných slov $\mathbf{u} \in A^{\mathbb{Z}}$ oddělujeme vertikální čarou prvky u_{-1} a u_0 k označení počátku na ose $\mathbb{Z}$. Pokud jsou prvky u_{-1} a u_0 rostoucí vzhledem k μ, pak je periodický bod $\mathbf{u} = \mu^p(\mathbf{u})$ úplně definován pomocí "semínka" (seed) $u_{-1}|u_0$, nebo-li $\mathbf{u} = \lim_{k\to+\infty} \mu^{pk}(u_{-1})|\mu^{pk}(u_0)$.

Morfismy, které zobrazují sturmovská slova na sturmovská slova, se nazývají sturmovské morfismy a tvoří tzv. sturmovský monoid. Určitý submonoid sturmovského monoidu se nazývá speciální sturmovský monoid. Morfismy nad binární abecedou lze klasicky reprezentovat maticemi rozměru 2×2 s celočíselnými nezápornými prvky. Tyto matice se nazývají incidenční. Incidenční matice některých morfismů splývají, a tedy tato reprezentace není "věrná" (faithful). V kapitole 5 představíme věrnou reprezentaci speciálního sturmovského monoidu pomocí matic rozměru 3×3 s nezápornými celočíselnými prvky, které mají odpovídající incidenční matici v levém horním rohu. S pomocí této reprezentace dokážeme, jakým postupem určit morfismus, jehož pevným bodem je tzv. odmocnina z pevného bodu morfismu ze speciálního sturmovského monoidu. Navíc popíšeme algoritmus určení věrné reprezentace morfismů ve speciálním sturmovském monoidu, díky čemuž objasníme vztah mezi věrnými reprezentacemi vzájemně konjugovaných morfismů.

Automatické posloupnosti: Cobham dokázal v roce 1972 vztah mezi jednostranně nekonečnými pevnými body k-uniformních morfismů a posloupnostmi generovanými pomocí deterministického konečného automatu a reprezentací nezáporných celých čísel v číselných soustavách o základu k [Cob72], [AS03, §6], $k \geq 2$. Morfismus μ se nazývá k-uniformní, pokud $\mu(a)$ je délky k pro každé $a \in A$. Přesněji tedy plyne z Cobhamových výsledků, že pokud je $\mu : A^* \to A^*$ k-uniformní morfismus a $\mathbf{u}$ je jeho pevný bod $\mathbf{u} = \mu(\mathbf{u})$ s rostoucím $u_0 = a$, potom písmeno na každé pozici $n \in \mathbb{N}$ v $\mathbf{u}$ lze získat jako $\mathcal{A}_{\mu,a}(\mathrm{rep}_k(n))$, kde $\mathcal{A}_{\mu,a}$ je deterministický konečný automat kanonicky přiřazený morfismu μ a písmenu a. Každý k-uniformní morfismus splňuje, že zobrazuje písmena na neprázdná slova a má rostoucí písmeno. Morfismy splňující tyto dvě vlastnosti se nazývají substituce. Cobhamův přístup byl zobecněn pro pevné body všech substitucí [RM02], s pomocí deterministických konečných automatů kanonicky přiřazených substitucím a číselných systémů v mnohem obecnějším smyslu, než jsou poziční číselné soustavy. Tyto systémy se nazývají abstraktní numerační systémy (abstract numeration systems). Později byl tento přístup zobecněn pro pevné body morfismů ve vyšších dimenzích [CKR10]. Další způsob, jak reprezentovat nezáporná celá čísla na základě pevných bodů substitucí, byl představen autory Dumontem a Thomasem [DT89].

V kapitole 6 rozšíříme přístup autorů Dumonta a Thomase pro reprezentaci všech celých čísel, čímž definujeme Dumontovy–Thomasovy numerační systémy pro $\mathbb{Z}$. Ukážeme, že každý oboustranně nekonečný periodický bod určité substituce s rostoucím semínkem je automatická posloupnost.

Theorem 6.3.1 *Nechť $\eta : A^* \to A^*$ je substituce a $\mathbf{u}$ je oboustranně nekonečný periodický bod s rostoucím semínkem $s = u_{-1}|u_0$. Pak $u_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_{\mathbf{u}}(n))$ pro každé $n \in \mathbb{Z}$.*

Deterministický konečný automat vytvoříme z kanonického automatu $\mathcal{A}_{\eta,a}$ přidáním nového počátečního stavu `start` a dvou nových hran. Dále dokážeme, že Dumontovy–Thomasovy numerační systémy pro $\mathbb{Z}$ jsou bijekce rostoucí vzhledem k úplnému uspořádání $\prec$, které bylo definováno v kapitole 3. Díky této vlastnosti dokážeme, že dvojkový doplněk a Fibonacciho doplněk jsou Dumontovy–Thomasovy numerační systémy pro $\mathbb{Z}$. Novým dosud nepublikovaným výsledkem je postačující podmínka, aby Dumontovy–Thomasovy numerační systémy byly poziční, čímž vytváříme souvislost s třídou pozičních číselných soustav definovanou v kapitole 3.

Nakonec ukážeme, že Dumontovy–Thomasovy numerační systémy mohou být rozšířeny pro vyšší dimenze $\mathbb{Z}^d$ pro každé $d \geq 1$. Fibonacciho doplněk pro $\mathbb{Z}^2$ má zvláštní důležitost pro tento text, protože ho použijeme v kapitole 7 k popisu Wangova dláždění jako automatické posloupnosti.

Wangova dláždění: Wangova dlaždice je čtverec jednotkového obsahu, který má na každé hraně barvu. Pro danou konečnou množinu Wangových dlaždic předpokládáme, že máme nekonečně mnoho kopií těchto dlaždic. Dvě Wangovy dlaždice mohou sousedit, pokud sdílí stejnou barvu na společné hraně. Obecně nás zajímají množiny Wangových dlaždic, kterými lze vydláždit rovinu, ale není to možné periodickým způsobem. Množina všech dláždění roviny pomocí dané množiny Wangových dlaždic se nazývá Wangův posun (Wang shift) a Wangův posun se nazývá aperiodický, pokud neobsahuje periodické dláždění. Wangovy posuny jsou předmětem zájmu od roku 1961. Od té doby byly zkonstruovány různé aperiodické Wangovy posuny s čím dál menším počtem generujících Wangových dlaždic. Konečně v roce 2021 byl popsán aperiodický Wangův posun s generující množinou 11 Wangových dlaždic [JR21]. Autoři Jeandel a Rao navíc dokázali, že každá menší množina Wangových dlaždic umožňuje periodické dláždění roviny. Na základě těchto 11 Wangových dlaždic byla popsána množina 19 Wangových dlaždic $\mathcal{U}$, která generuje aperiodický, minimální a soběpodobný Wangův posun [Lab21, Lab19].

O J 0 F O	O H 1 F L	M F 2 J P	M F 3 D K	P J 4 H P	P H 5 H N	K F 6 H P	K D 7 H P	O I 8 B O	L E 9 G O
L C10G L	L I 11A O	P G12E P	P I 13E P	P G14 I K	P I 15 I K	K B16 I M	K A17 I K	N I 18C P	

V kapitole 7 studujeme konkrétní množinu 16 Wangových dlaždic $\mathcal{Z}$, která byla vytvořena z množiny $\mathcal{U}$ sloučením některých barev. Nově ukážeme, že Wangův posun $\Omega_\mathcal{Z}$ a Wangův posun $\Omega_\mathcal{U}$ jsou topologicky konjugované. V důsledku toho ukážeme, že je Wangův posun $\Omega_\mathcal{Z}$ aperiodický, minimální a soběpodobný.

O J 0 D O	O H 1 D L	M D 2 J P	M D 3 D K	P J 4 H P	P H 5 H N	K D 6 H P	O I 7 B O
L E 8 I O	L C 9 I L	L I 10A O	P I 11E P	P I 12 I K	K B13 I M	K A14 I K	N I 15C P

Ukážeme, že existuje Wangovo dláždění pomocí dlaždic $\mathcal{Z}$, které lze vygenerovat deterministickým konečným automatem, který akceptuje jazyk reprezentací Fibonacciho doplňku pro $\mathbb{Z}^2$.

Theorem 7.4.1 *Existuje deterministický konečný automat $\mathcal{A}$ takový, že*

$$\begin{array}{rccl} x: & \mathbb{Z}^2 & \to & \{0,1,\ldots,15\} \\ & \boldsymbol{n} & \mapsto & \mathcal{A}(\mathrm{rep}_{\mathcal{F}c}(\boldsymbol{n})) \end{array}$$

splňuje $x \in \Omega_\mathcal{Z}$.

Acknowledgements

I would like to thank everybody who supported me during my PhD studies. In the first place, I express my gratitude to both my advisors, Ľubomíra Dvořáková and Sébastien Labbé, who invested an enormous amount of time and patience to introduce me to their fields and to help me grow in independence. It is certain that I would have never gone so far without their support and I can never thank them enough. I am also grateful to Edita Pelantová who was the person linking me to Sébastien and Ľubomíra and who made a great effort to get me involved in the combinatorics on words community. In particular, she let me work on an article with her and Štěpán Starosta, which was a very valuable experience for a first year PhD student. Other thanks go to Zuzana Masáková for her high quality teaching. On the French side, it was thanks to my friends Quentin Manière, Jonathan Narboni and Govind from the office 123 that Bordeaux became my second home. I also want to thank Marthe Bonamy, Vincent Delecroix and Pascal Weil who were always there for me to offer advice, academic or other. I realize that this list is not exhaustive and many others contributed to the creation of this document in various ways.

I further acknowledge the financial support of the following projects

- The French Institute in Prague and the Czech Ministry of Education, Youth and Sports through the Barrande fellowship programme,
- Agence Nationale de la Recherche through the project Codys (ANR-18-CE40-0007),
- Grant Agency of Czech technical university in Prague, through the project SGS20/183/OHK4/3T/14.

This thesis would never come into existence without the infinite support of my parents Václav and Zdeňka, my sister Barunka, my best friend Helenka and my fiancé Élam.

Contents

Chapter 1

Introduction

This work aims to discover and develop links between three related but distinct mathematical domains: combinatorics on words, number theory and discrete geometry. The presence of these domains in this work is the following. From the point of view of combinatorics on words, we study finite and infinite strings of symbols (called words) and morphisms, which act as maps on words. Substitutions are morphisms satisfying some additional properties, however, both substitutions and morphisms provide what we call a substitutive structure. As for the number theory, we are interested in numeration systems and, in particular, in those numeration systems which one may use to represent both nonnegative and negative integers. The discrete geometry is present in the form of Wang tilings – coverings of the plane by squares with colors on the edges so that colors on the adjacent edges match. We study a particular set of such tiles and its close relation to two-dimensional morphisms and numeration systems.

The common thread of this work are objects related to the name Fibonacci. This Italian mathematician, also known as Leonardo of Pisa (1170 – 1250), studied a problem involving the growth of a population of rabbits, which led to the definition of Fibonacci numbers. These are defined by a recurrence relation such that the next number is obtained as the sum of the previous two numbers, starting from 1 and 2. This gives rise to the sequence of Fibonacci numbers:

$$(F_i)_{i=0}^{+\infty} = 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, \ldots$$

It belongs to the common knowledge that we can uniquely represent nonnegative integers $\mathbb{N} = \{0, 1, 2 \ldots\}$ using a sequence $(N^i)_{i=0}^{+\infty}$ of powers of an integer $N \geq 2$, which gives rise to the integer-base numeration systems. The most well-known examples include the decimal numeration system where $N = 10$ and the binary numeration system where $N = 2$. However, other sequences may be used to define more general numeration systems, such as the sequence of Fibonacci numbers $(F_i)_{i=0}^{+\infty}$. Every integer $n \in \mathbb{N}$ can be decomposed into a sum $\sum_{i=0}^{k-1} w_i F_i$ for some $k \in \mathbb{N}$ so that w_i are digits in the set $\{\mathtt{0}, \mathtt{1}\}$, for every $0 \leq i \leq k-1$. The concatenation of symbols $w_{k-1}w_{k-2}\cdots w_0$ is called a Fibonacci representation of n. There exists a unique Fibonacci representation of every integer n such that $w_{k-1}\cdots w_0$ does not contain consecutive ones and $w_{k-1} \neq \mathtt{0}$ [Zec72]. Such representations are called greedy and denoted $\mathrm{rep}_{\mathcal{F}}(n) = w_{k-1}\cdots w_0$. For instance, the greedy Fibonacci representations of integers $n \leq 6$ are:

$$\mathrm{rep}_{\mathcal{F}}(1) = \mathtt{1},\ \mathrm{rep}_{\mathcal{F}}(2) = \mathtt{10},\ \mathrm{rep}_{\mathcal{F}}(3) = \mathtt{100},\ \mathrm{rep}_{\mathcal{F}}(4) = \mathtt{101},\ \mathrm{rep}_{\mathcal{F}}(5) = \mathtt{1000},\ \mathrm{rep}_{\mathcal{F}}(6) = \mathtt{1001}, \ldots$$

Over the last 40 years, various generalizations of integer-base numeration systems such as the Pisot, Parry and Bertrand numeration systems [BM89, BHMV94, Hol98, MPR19, CCS22] were

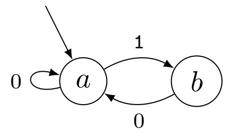


Figure 1.1: The deterministic finite automaton with output $\mathcal{A}_{\mathcal{F}}$.

extensively studied. The Fibonacci numeration system $\mathcal{F}$ is an example of all of these numeration systems, which have the following hierarchy

$$\text{Pisot} \subsetneq \text{Parry} \subsetneq \text{Bertrand}.$$

Another famous object carrying the Fibonacci name is the Fibonacci word. This right-infinite string of symbols 0 and 1 is created recursively so that its next prefix f_{n+2} is a concatenation of the previous two prefixes $f_{n+1} \cdot f_n$, starting from the prefixes $f_0 = a$ and $f_1 = ab$:

$$\begin{aligned} f_0 &= a, \\ f_1 &= ab, \\ f_2 = ab \cdot a &= aba, \\ f_3 = aba \cdot ab &= abaab, \\ f_4 = abaab \cdot aba &= abaababa. \end{aligned}$$

Thus, the Fibonacci word $\mathbf{f}$ is an infinite string over the binary alphabet beginning with the prefix

$$\mathbf{f} = abaababaabaababaabab \cdots .$$

The Fibonacci word $\mathbf{f}$ is an example of Sturmian words. Sturmian words are a class of infinite aperiodic words over the binary alphabet, which belong to the most explored objects in combinatorics on words. They may be defined as the infinite words which contain $n+1$ distinct finite words of length n, for every $n \in \mathbb{N}$. First, they appeared in the work of Morse and Hedlund [MH40] in 1940 and, since then, various equivalent definitions were proposed. For instance, Sturmian words are exactly the balanced words which are aperiodic. Balanced words have been studied recently with respect to their repetition rate of factors [DDP23, DPOS22]. The words that generalize Sturmian words to larger alphabets are called Arnoux–Rauzy words [AR91] or episturmian words [DJP01]. In particular, the words that generalize the Fibonacci word $\mathbf{f}$ to an alphabet of size $d \geq 2$ are called the d-bonacci words.

Among other interesting properties, the Fibonacci word $\mathbf{f}$ is fixed by the morphism

$$\varphi : a \mapsto ab, \ b \mapsto a,$$

which means that $\varphi(\mathbf{f}) = \mathbf{f}$. We refer to the morphism φ as to the Fibonacci morphism. The morphism φ belongs to the so-called monoid of Sturm, meaning that it maps every Sturmian word to a Sturmian word. Its second power $\varphi^2 : a \mapsto aba$, $b \mapsto ab$ belongs to a certain submonoid of the monoid of Sturm, which is called the special monoid of Sturm. Recent results on the Sturmian morphisms concern the properties of the derived words of their fixed points [KMPS18], [PS21].

The Fibonacci word $\mathbf{f}$ is also an automatic sequence – it can be generated by feeding the deterministic finite automaton with output in Figure 1.1 with the Fibonacci representations

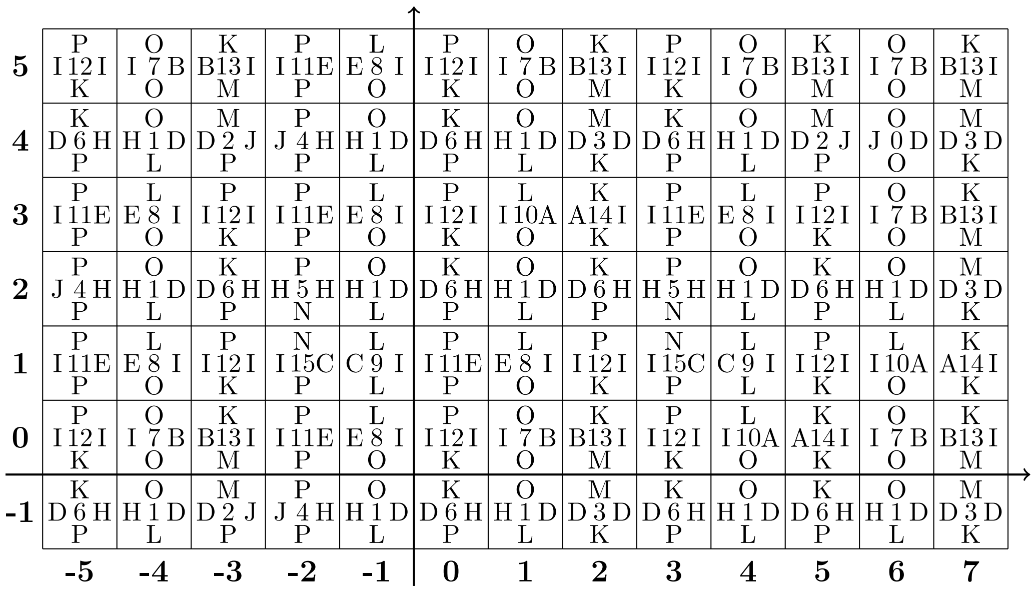


Figure 1.2: A partial valid tiling with a certain set of Wang tiles.

$\text{rep}_{\mathcal{F}}(n)$, for every $n \in \mathbb{N}$. For instance, the automaton $\mathcal{A}_{\mathcal{F}}$ reading $\text{rep}_{\mathcal{F}}(5) = \texttt{1000}$ starts at the initial state and it moves between states $a \xrightarrow{\texttt{1}} b \xrightarrow{\texttt{0}} a \xrightarrow{\texttt{0}} a \xrightarrow{\texttt{0}} a$, finishing at the state a and thus outputting the letter a. This connection between the fixed points of morphisms and automatic sequences was first observed by Cobham in 1972 [Cob72], [AS03, §6]. He proved that the automatic sequences with a k-state automaton are exactly the fixed points of morphisms having the images of all letters of length k, where $k \geq 2$. In this case, the classical base-k representations are inputted into the automaton. The idea was extended to the fixed points of all morphisms [RM02], using a much broader definition of the abstract numeration systems. This direction is broadly studied in the following book [Rig14]. Abstract numeration systems were generalized to describe the fixed points of multidimensional morphisms [CKR10]. Recently, the automatic sequences based on Parry and Bertrand numeration systems were studied [MPR19].

Last but not least, the Fibonacci morphism φ is also involved in a particular Wang tiling, which can be described as a fixed point of a 2-dimensional substitution closely related to φ. A Wang tile is a unit square with a color on each edge. Given a finite set of Wang tiles, we assume that we have infinitely many copies of them and we are allowed to arrange two Wang tiles side by side (without rotating them) provided that the colors on their common edge match; see a partial valid tiling in Figure 1.2. In general, we are interested in such sets of Wang tiles which tile the plane, but which do not do so in a periodic way. Given a set of Wang tiles, the set of all tilings of the plane is called a Wang shift and the Wang shift is called aperiodic if it does not contain a periodic tiling. Wang shifts have been studied since 1961 [Wan61], when Wang proposed the fundamental conjecture stating: if a Wang shift is not empty, then it contains a periodic Wang tiling. This conjecture was refuted by Berger, who constructed the first aperiodic Wang shift [Ber66] based on more than 20 000 distinct Wang tiles. During the next decades, the number of generating Wang tiles gradually decreased to 13 [Kar96, Cul96]. In 2021, Jeandel and Rao constructed a set of 11 Wang tiles, which generates an aperiodic Wang shift and they proved that every set of less than 11 Wang tiles admits a periodic tiling [JR21]. It is worth noting that the domain of tilings is much broader as the plane can be tiled by other shapes than unit squares; see [GS87] for more details. Until recently, the smallest known sets of tiles which admit only non-periodic tilings of the plane contained two tiles, the Penrose tiling being the most famous example [Pen79]. Furthermore, the Penrose tiling is a model of quasicrystals, linking tilings to another mathematical domain [Sen96]. During the work on this thesis, a set containing a single tile was described [SMKGS23], much to the delight of the tiling community.

1.1 Listing of the author's contributions

The five chapters of this text contain results of five papers (including 2 conference papers) that have either been published or are under review. Among the five papers, four are a common work of the author and her advisors. We present the list of the papers in the same order as the chapters in this text.

(I) Sébastien Labbé and J. L., *A Fibonacci analogue of the two's complement numeration system*, RAIRO - Theoretical Informatics and Applications, 57:12, 2023.

(II) Ľubomíra Dvořáková and J.L., *Critical exponents of regular Arnoux-Rauzy sequences.* In Combinatorics on words, volume 13899 of Lecture Notes in Comput. Sci., pages 130–142. Springer, 2023.

(III) J. L., Edita Pelantová and Štěpán Starosta, *On a faithful representation of Sturmian morphisms.* In European Journal of Combinatorics, 110:103707, 2023.

- Jana Lepšová is a co-author, worked mainly on the part concerning the square roots of fixed points of Sturmian morphisms, found the hypothesis of the main theorem concerning square roots of fixed points based on computer experiments, assisted with editing the paper.

(IV) Sébastien Labbé and J. L., *Dumont-Thomas numeration systems for* $\mathbb{Z}$. Under review in Integers – Electronic Journal of Combinatorial Number Theory. Preprint accessible at `https://arxiv.org/abs/2302.14481`.

(V) Sébastien Labbé and J. L., *A numeration system for Fibonacci-like Wang shifts.* In Combinatorics on words, volume 12847 of Lecture Notes in Comput. Sci., pages 104–116. Springer, Cham, 2021.

1.2 Structure of the thesis

In Chapter 3, we extend the Fibonacci numeration system to all integers including the negative ones, defining the Fibonacci complement numeration system $\mathcal{F}c$. We show its properties with respect to addition. As a new result not presented in (I), we describe a class of positional numeration systems for all integers, which contains the numeration system $\mathcal{F}c$, and we study their properties with respect to a certain order.

In Chapter 4, we study the repetitions of finite words appearing in Arnoux–Rauzy words. We provide a formula for the critical exponent and the asymptotic critical exponent of regular Arnoux-Rauzy words. Also, we show that, among regular Arnoux-Rauzy words over a d-ary alphabet, the minimal (asymptotic) critical exponent is reached by the d-bonacci word. As a new result not presented in (II), we show that the critical exponent coincides with the asymptotic critical exponent of d-bonacci words for every $4 \leq d \leq 15$.

In Chapter 5, we faithfully represent the elements of the special Sturmian monoid by 3×3 matrices with nonnegative entries. Using the faithful representation, we study the so-called square roots of fixed points of Sturmian morphisms. As a new result not presented in (III), we provide an algorithm to determine the faithful representation of a given Sturmian morphism, which enables us to answer an open question from (III) concerning the intercepts of the fixed points of mutually conjugate Sturmian morphisms.

In Chapter 6, we define numeration systems for $\mathbb{Z}$ based on substitutions and we show their properties with respect to automata and order. We call these numeration systems Dumont–Thomas numeration systems for $\mathbb{Z}$ as they extend the numeration systems for nonnegative integers described by Dumont and Thomas. The Fibonacci complement numeration system $\mathcal{F}c$ is recovered here as a numeration system related to the Fibonacci substitution φ. Unlike the numeration systems studied in Chapter 3, Dumont–Thomas numeration systems might not be positional. However, they are positional assuming a certain sufficient condition, which we provide as a new result not presented in (IV).

In Chapter 7, we study a particular set of 16 Wang tiles $\mathcal{Z}$. We find an automatic characterization of a particular valid Wang configuration over $\mathcal{Z}$, using the Fibonacci complement numeration system $\mathcal{F}c$ extended to $\mathbb{Z}^2$. As a new result not presented in (V), we prove that the Wang shift $\Omega_{\mathcal{Z}}$ is topologically conjugate to another Wang shift $\Omega_{\mathcal{U}}$, which was derived from the Jeandel–Rao Wang shift [JR21].

Chapter 2

Preliminaries

In this part, we present the essential background and notation in combinatorics on words, automata theory, number theory and Wang tilings. We clarify that the notation $\mathbb{N}$ refers to the set of nonnegative integers $\{0, 1, 2, \ldots\}$, $\mathbb{Z}_{<0}$ refers to the set of negative integers $\{\ldots, -3, -2, -1\}$ and $\mathbb{Z}$ refers to the set of all integers $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$. For every pair of integers $i, j \in \mathbb{Z}$ such that $i \leq j$, we denote the interval of integers between i and j by $[\![i, j]\!] = \{i, i+1, \ldots, j-1, j\}$.

2.1 Words and languages

We introduce terminology concerning finite and infinite words from [Lot02]. An *alphabet* A is a finite set and the elements of an alphabet are called *letters.* A *finite word* u over A of *length* n is a finite string $u = u_0u_1 \cdots u_{n-1}$ of letters $u_i \in A$, $i \in [\![0, n-1]\!]$. The length of u is denoted by $|u|$, whereas $|u|_a$ denotes the number of occurrences of a letter $a \in A$ in u. The set of all finite words over A is denoted A^*. If $u = xyz$ is a concatenation of finite words $x, y, z \in A^*$, then x is a *prefix* of u, z is a *suffix* of u and y is a *factor* of u. A finite word $x \in A^*$ is a *proper* prefix of u if it is a prefix of u and $x \neq u$. A *reversed factor* $\tilde{u}$ of a factor $u = u_0u_1 \cdots u_{n-1}$ is the factor $\tilde{u} = u_{n-1} \cdots u_1u_0$. A finite word $u = u_0u_1 \cdots u_{n-1}$ is a *palindrome* if $u = \tilde{u}$. The set A^* with the operation of concatenation of finite words forms a monoid whose neutral element is the *empty word* ε.

A *right-infinite word* (or simply a *word*) $\mathbf{u}$ over A is a right-infinite string $\mathbf{u} = u_0u_1u_2 \cdots$ of letters $u_i \in A$, for all $i \in \mathbb{N}$. A *left-infinite word* $\mathbf{u}$ over A is a left-infinite string $\mathbf{u} = \cdots u_{-3}u_{-2}u_{-1}$ of letters $u_i \in A$, for all $i \in \mathbb{Z}_{<0}$. We call a biinfinite string $\mathbf{u} : \mathbb{Z} \to A$ a *two-sided word* over A and we separate its elements u_{-1} and u_0 by a vertical bar to indicate the origin, i.e.,

$$\mathbf{u} = \cdots u_{-3}u_{-2}u_{-1}|u_0u_1u_2 \cdots .$$

The set of all right-infinite (resp., left-infinite, two-sided) words over A is denoted by $A^{\mathbb{N}}$ (resp., $A^{\mathbb{Z}_{<0}}$, $A^{\mathbb{Z}}$).

Let $\mathbf{u} \in A^{\mathbb{N}}$. A finite word $v \in A^*$ such that $v = u_iu_{i+1}u_{i+2} \cdots u_{j-1}$ for some $i, j \in \mathbb{N}$, $i \leq j$, is called a *factor* of $\mathbf{u}$. The number i is called an *occurrence* of the factor v in the word $\mathbf{u}$. A factor $v \in A^*$ is a *prefix* of $\mathbf{u}$ if $i = 0$ is an occurrence of v in the word $\mathbf{u}$. We call $\mathbf{u}$

- *recurrent* if every factor of $\mathbf{u}$ has infinitely many occurrences in $\mathbf{u}$;
- *uniformly recurrent* if $\mathbf{u}$ is recurrent and the distances between the consecutive occurrences of each factor in $\mathbf{u}$ are bounded;

- *eventually periodic* if there exist $v, w \in A^*$ such that $\mathbf{u} = vwww\cdots = vw^\omega$;
- *periodic* if $\mathbf{u}$ is eventually periodic and $v = \varepsilon$;
- *aperiodic* if it is not eventually periodic.

Now, assume that $\mathbf{u}$ is recurrent and consider a factor v of $\mathbf{u}$. Let $i < j$ be two consecutive occurrences of v in $\mathbf{u}$. The word $u_i u_{i+1} \cdots u_{j-1}$ is called a *return word* to v in $\mathbf{u}$. The set of all return words to v in $\mathbf{u}$ is denoted by $\mathcal{R}_{\mathbf{u}}(v)$. If v is a prefix of $\mathbf{u}$, then $\mathbf{u}$ can be written as a concatenation $\mathbf{u} = r_{d_0} r_{d_1} r_{d_2} \cdots$ of the return words to v. The *derived word* of a word $\mathbf{u}$ to a factor v is the word $\mathbf{d_u}(v) = \mathtt{r}_{d_0}\mathtt{r}_{d_1}\mathtt{r}_{d_2}\cdots$ over the alphabet $\{\mathtt{r}_1, \mathtt{r}_2, \dots\}$ of cardinality $\#\mathcal{R}_{\mathbf{u}}(v)$.

Let $\mathbf{u} \in A^{\mathbb{N}}$. The *language* $\mathcal{L}(\mathbf{u})$ of is the set of factors occurring in $\mathbf{u}$. The *factor complexity* of $\mathbf{u}$ is the map $\mathcal{C}_{\mathbf{u}} : \mathbb{N} \to \mathbb{N}$ defined for every $n \in \mathbb{N}$ as $\mathcal{C}_{\mathbf{u}}(n) = \#\{w \in \mathcal{L}(\mathbf{u}) : |w| = n\}$. The language $\mathcal{L}(\mathbf{u})$ is called *closed under reversal* if, for every factor $w \in \mathcal{L}(\mathbf{u})$, it contains its reversed factor $\tilde{w} \in \mathcal{L}(\mathbf{u})$. A factor w of $\mathbf{u}$ is called *right special* if there exist two distinct letters $a, b \in A$ such that $wa, wb \in \mathcal{L}(\mathbf{u})$. Similarly, a factor w of $\mathbf{u}$ is called *left special* if there exist two distinct letters $a, b \in A$ such that $aw, bw \in \mathcal{L}(\mathbf{u})$. A factor is called *bispecial* if it is both left and right special.

2.2 Automata

We introduce terminology concerning automata theory from [BR10]. A *finite automaton* is a labeled graph given by a 5-tuple $\mathcal{A} = (Q, A, E, I, T\}$ where

- Q is a finite set of *states*,
- $E \subset Q \times A^* \times Q$ is a finite set of *edges* defining the *transition relation*,
- $I \subset Q$ is a set of *initial states*,
- and T is a set of *final states*.

A *path* in the automaton $\mathcal{A}$ is a sequence $(q_0, u_0, q_1, u_1, \dots, q_{k-1}, u_{k-1}, q_k)$ such that, for all $i \in [\![0, k-1]\!]$, $(q_i, u_i, q_{i+1}) \in E$, and the finite word $u_0 \cdots u_{k-1} \in A^*$ is called the *label* of the path. We visualize such a path as

$$q_0 \xrightarrow{u_0} q_1 \xrightarrow{u_1} \dots \xrightarrow{u_{k-1}} q_k.$$

A path is *successful* if $q_0 \in I$ and $q_k \in T$. The *language* $\mathcal{L}(\mathcal{A})$ *accepted* or *recognized* by $\mathcal{A}$ is the set of labels of all successful paths in $\mathcal{A}$. Moreover, we denote $\mathcal{L}_m(\mathcal{A}) = \mathcal{L}(\mathcal{A}) \cap A^m$, for every $m \in \mathbb{N}$. A state $q \in Q$ is *accessible* if there exists a path from an initial state to q and a state $q \in Q$ is *co-accessible* if there exists a path from q to a final state. If all states of $\mathcal{A}$ are both accessible and co-accessible, then $\mathcal{A}$ is said to be *trim*.

A finite automaton is said *deterministic* if it has only one initial state, edges are labeled by letters and for every pair $(q, a) \in Q \times A$, there is at most one state $r \in Q$ such that $(q, a, r) \in E$. Then E defines a transition function, which might be partial. We summarize that a *deterministic finite automaton* is a labeled graph given by a 5-tuple $\mathcal{A} = (Q, A, \delta, q_{\text{ini}}, T)$ where

- Q is a set of *states*,

- A is an alphabet,
- $\delta : Q \times A \to Q$ is called the *transition function*,
- $q_{\text{ini}} \in Q$ is the *initial state*,
- and $T \subset Q$ is a set of *final states*.

The transition function $\delta = \delta_{\mathcal{A}}$ is naturally extended to $Q \times A^*$ by $\delta(q, \varepsilon) = q$ for every $q \in Q$, and, for every $q \in Q$, $a \in A$ and $w \in A^*$, $\delta(q, aw) = \delta(\delta(q, a), w)$. For every $w \in A^*$, we denote $\mathcal{A}(w) = \delta(q_{\text{ini}}, w)$. Consequently, the language $\mathcal{L}(\mathcal{A})$ is the set $\{w \in A^* : \delta(q_{\text{ini}}, w) \in T\}$. We refer to a deterministic finite automaton as to the 5-tuple $(Q, A, \delta_{\mathcal{A}}, q_{\text{ini}}, T)$.

A set of finite words $X \subset A^*$ is *recognizable* or *regular* if there exists a finite automaton $\mathcal{A}$ so that $X = \mathcal{L}(\mathcal{A})$. It can be shown that finite automata and deterministic finite automata recognize exactly the same languages of finite words. Therefore, equivalently, $X \subset A^*$ is regular if there exists a deterministic finite automaton $\mathcal{A}$ so that $X = \mathcal{L}(\mathcal{A})$.

Let $X \subset A^*$ and $w \in A^*$. The *left quotient* [Lot02] is defined as the set of all finite words $u \in A^*$ which can be concatenated after $w \in A^*$ so that $wu \in X$:

$$w^{-1}X = \{u \in A^* : wu \in X\}.$$

We denote by $\mathcal{A}(X)$ the deterministic automaton $(Q(X), A, \delta_{\mathcal{A}(X)}, I(X), T(X))$ such that

- $Q(X) = \{w^{-1}X : w \in A^*\}$, i.e., every state is a set,
- $\delta_{\mathcal{A}(X)}(S, a) = a^{-1}S$, for every $S \in Q(X)$ and $a \in A$,
- $I(X) = X$,
- and $T(X) = \{S \in Q(X) : \varepsilon \in S\}$.

The automaton $\mathcal{A}(X)$ is called the *minimal automaton* of X and it recognizes X. Note that $\mathcal{A}(X)$ is not necessarily finite. However, a set $X \subset A^*$ is regular if and only if its minimal automaton $\mathcal{A}(X)$ is finite.

A *deterministic finite automaton with output* is a 6-tuple $\mathcal{A} = (Q, A, \delta, \{q_{\text{ini}}\}, B, \xi)$ where A and B are alphabets, Q, δ and q_{ini} are defined as in the case of a deterministic finite automaton and $\xi : Q \to B$ is the *output function*. In this text, we have always $B = Q$ and ξ is the identity map.

2.3 Morphisms

We introduce terminology concerning morphisms and substitutions from [BR10] and [Fog02]. Let A, B be alphabets. A *morphism* is a map $\psi : A^* \to B^*$ such that $\psi(uv) = \psi(u)\psi(v)$ for all words $u, v \in A^*$. In particular, if $A = B$, then we call ψ a morphism over the alphabet A. A morphism ψ is called

- *non-erasing* if $\psi(a) \in A^+$ is nonempty, for every $a \in A$;
- *growing* or *with a growing letter* if there exists $a \in A$ such that $\lim_{k \to +\infty} |\psi^k(a)| = +\infty$;
- k*-uniform* for some integer $k \geq 2$ if $|\psi(a)| = k$, for every $a \in A$;

- *primitive* if there exists $k \in \mathbb{N}$ such that, for every pair $a, b \in A$, a occurs in $\psi^k(b)$.

A morphism ψ that is non-erasing and growing is called a *substitution*. The *incidence matrix* $M_\psi \in \mathbb{N}^{|A| \times |A|}$ of a morphism ψ is defined as $(M_\psi)_{a,b} = |\psi(b)|_a$, for every $a, b \in A$, where $|A|$ denotes the cardinality of A. In the matrix theory, a matrix is called *primitive* if there exists $k \in \mathbb{N}$ such that all entries of M^k are positive. An important theorem treating primitive matrices is called the Perron–Frobenius theorem; see for instance [Fie08]. The eigenvalue from Theorem 2.3.1 is called the *Perron–Frobenius eigenvalue* of M. Clearly, a morphism is primitive if and only if its incidence matrix is primitive and, consequently, the Perron–Frobenius theorem is relevant for primitive morphisms.

Theorem 2.3.1 (Perron–Frobenius)**.** *Let $d \in \mathbb{N}$ and let $M \in \mathbb{N}^{d \times d}$ be primitive. Then*

- *M has a positive eigenvalue λ which fulfills the condition that $|\lambda'| < \lambda$, for every eigenvalue λ' of M such that $\lambda' \neq \lambda$;*

- *λ is algebraically simple, i.e., the eigenspace associated with λ is one-dimensional;*

- *the eigenvector corresponding to λ may be chosen as a vector with positive entries and no other eigenvalue has an eigenvector having this property.*

Morphisms can be naturally extended to $A^{\mathbb{N}}$, $A^{\mathbb{Z}_{<0}}$ and $A^{\mathbb{Z}}$ by setting

$$\begin{aligned}
\psi(u_0 u_1 u_2 \cdots) &= \psi(u_0)\psi(u_1)\psi(u_2)\cdots, \\
\psi(\cdots u_{-3} u_{-2} u_{-1}) &= \cdots \psi(u_{-3})\psi(u_{-2})\psi(u_{-1}), \\
\psi(\cdots u_{-3} u_{-2} u_{-1} | u_0 u_1 u_2 \cdots) &= \cdots \psi(u_{-3})\psi(u_{-2})\psi(u_{-1}) | \psi(u_0)\psi(u_1)\psi(u_2)\cdots.
\end{aligned}$$

Let $\mathbb{D} \in \{\mathbb{Z}, \mathbb{N}, \mathbb{Z}_{<0}\}$. A word $\mathbf{u} \in A^{\mathbb{D}}$ is called a *periodic point* of a morphism ψ if there exists an integer $p \geq 1$ such that $\psi^p(\mathbf{u}) = \mathbf{u}$. In this case, p is called *a period* of $\mathbf{u}$. The minimal integer $p \geq 1$ such that $\psi^p(\mathbf{u}) = \mathbf{u}$ is called *the* period of $\mathbf{u}$. A periodic point with the period $p = 1$ is called a *fixed point* of ψ. The set of periodic points of ψ is denoted $\mathrm{Per}_{\mathbb{D}}(\psi) = \{\mathbf{u} \in A^{\mathbb{D}} \colon \psi^p(\mathbf{u}) = \mathbf{u}$ for some $p \geq 1\}$. We omit the domain when $\mathbb{D} = \mathbb{Z}$ and we write $\mathrm{Per}(\psi) = \mathrm{Per}_{\mathbb{Z}}(\psi)$. If $\mathbf{u} \in \mathrm{Per}(\psi)$ is a two-sided periodic point of a substitution ψ, we say that the pair of letters $u_{-1}|u_0$ is the *seed* of $\mathbf{u}$ [BG13, §4.1]. If the seed-letters of a two-sided periodic point are growing, then the periodic point is defined entirely by its seed, which is called *growing*. More precisely, $\mathbf{u} = \lim_{j \to +\infty} \psi^{jp}(u_{-1})|\psi^{jp}(u_0)$ where p is a period of $\mathbf{u}$.

2.4 Beta-expansions

We use the terminology from [BR10, §2]. Let $\beta > 1$ be a real number. The *β-transformation* is a map $\tau_\beta : [0,1) \to [0,1)$ defined as $\tau_\beta(x) = \beta x - \lfloor \beta x \rfloor$. For a real number $x \in [0,1)$, let $(x_i)_{i=1}^{+\infty}$ be the sequence of integers such that $x_i = \lfloor \beta \tau_\beta^{i-1}(x) \rfloor$, for every $i \geq 1$. Then $x = \sum_{j=1}^{+\infty} x_j \beta^{-j}$.

Interpreting x_i as letters from the alphabet $\{0, 1, \ldots, \lceil \beta \rceil - 1\}$, the right-infinite word $(x_i)_{i=1}^{+\infty}$ is called the *β-expansion* of x and denoted $d_\beta(x)$. The set of β-expansions of all real numbers in $[0,1)$ is denoted D_β. The set of factors occurring in the β-expansions of the real numbers in $[0,1)$ is denoted $\mathrm{Fact}(D_\beta)$. The *β-expansion of unity* $d_\beta(1)$, sometimes called the Rényi expansion of unity, is defined as

$$d_\beta(1) = t_1 t_2 \cdots, \qquad \text{where } t_1 = \lfloor \beta \rfloor \text{ and } t_2 t_3 \cdots = d_\beta(\beta - \lfloor \beta \rfloor).$$

The *quasi-greedy β-expansion of unity* $d^*_\beta(1)$ is defined as

$$d^*_\beta(1) = \begin{cases} (t_1 \cdots t_{m-1}(t_m - 1))^\omega, & \text{if } d_\beta(1) = t_1 \cdots t_m \text{ is finite;} \\ d_\beta(1), & \text{otherwise.} \end{cases}$$

A *Parry number* is such a real $\beta > 1$ that $d_\beta(1)$ is eventually periodic. In particular, a Parry number β is *simple* if $d_\beta(1)$ is finite.

Chapter 3

Complement numeration systems

In this chapter, we present our work concerning the Fibonacci complement numeration system, which generalizes the classical Fibonacci numeration system for $\mathbb{N}$ to $\mathbb{Z}$. In particular, we show its properties with respect to addition. The above mentioned results were published in RAIRO – Theoretical Informatics and Applications [LL23a]. Also, we present our results, which do not make part of the publication. We extend numeration systems associated with simple Parry numbers to $\mathbb{Z}$, creating a class of numeration systems which contains both the two's complement notation and the Fibonacci complement numeration system. We show the properties of complement numeration systems with respect to a particular total order.

3.1 Introduction to positional numeration systems

First, we summarize properties of positional numeration systems [Lot02, §7]. A *positional numeration system* U is a strictly increasing sequence of integers $U = (U_n)_{n=0}^{+\infty}$ such that $U_0 = 1$ and $C_U := \sup_{n\geq 0}\lceil \frac{U_{n+1}}{U_n} \rceil$ is finite. The *value map* $\mathrm{val}_U : A^* \to \mathbb{N}$ of the numeration system U maps a word $w = w_{k-1}\cdots w_0 \in A^*$ over an integer alphabet A to the integer $\sum_{i=0}^{k-1} w_i U_i$

$$\mathrm{val}_U \colon w_{k-1}\cdots w_0 \mapsto \sum_{i=0}^{k-1} w_i U_i. \tag{3.1}$$

We denote $\Sigma_U = [\![0, C_U - 1]\!]$ the alphabet of U. A *U-representation* of an integer $n \in \mathbb{N}$ is a word $w \in \Sigma_U^*$ such that $\mathrm{val}_U(w) = n$. The *normal* or *greedy U-representation* of an integer $n \in \mathbb{N}$ is a U-representation $w = w_{k-1}\cdots w_0 \in \Sigma_U^k$ such that $w_{k-1} \neq 0$ and $\mathrm{val}_U(w_j \cdots w_0) < U_{j+1}$, for every $j \in [\![0, k-1]\!]$. The *normalization* on A^* is the partial function $A^* \to \Sigma_U^*$ that maps every word $w \in A^*$ such that $\mathrm{val}_U(w) \geq 0$ onto the normal U-representation of $\mathrm{val}_U(w)$. We omit the letter U in a U-representation if the numeration system is clear from context.

The *representation map* $\mathrm{rep}_U : \mathbb{N} \to \Sigma_U^*$ maps an integer $n \in \mathbb{N}$ to its greedy representation $\mathrm{rep}_U(n)$. In particular, $\mathrm{rep}_U(0) = \varepsilon$ is the empty word. The set $\mathtt{0}^* \mathrm{rep}_U(\mathbb{N})$ of all greedy representations preceded by leading zeroes is called the *numeration language*. For every $n \in \mathbb{N}$, it holds that

$$\mathrm{val}_U(\mathrm{rep}_U(n)) = n \tag{3.2}$$

and thus $\mathrm{rep}_U : \mathbb{N} \to \mathrm{rep}_U(\mathbb{N})$ is a bijection. Moreover, the set $\mathrm{rep}_U(\mathbb{N})$ of greedy representations is totally ordered with respect to the radix order. The radix order is a total order on an alphabet A such that, for every $u, v \in A^*$, $u <_{\mathrm{rad}} v$ if and only if $|u| < |v|$ or $|u| = |v|$ and $u <_{lex} v$. This order is sometimes called genealogical. The radix order on $\mathrm{rep}_U(\mathbb{N})$ is induced by the natural order of the integer alphabet Σ_U. We reformulate [Lot02, Proposition 7.3.5].

Proposition 3.1.1. [Lot02] *Let U be a positional numeration system. For all m, $n \in \mathbb{N}$,*

$$m < n \text{ if and only if } \mathrm{rep}_U(m) <_{\mathrm{rad}} \mathrm{rep}_U(n).$$

We illustrate the introduced notions on two important examples for our study.

Example 3.1.2 (Binary numeration system)**.** The sequence of integers $U = (2^n)_{n=0}^{+\infty}$ is strictly increasing, $U_0 = 1$ and $C_U = \sup_{n\geq 0}\lceil \frac{2^{n+1}}{2^n} \rceil = 2$ is finite. The positional numeration system U is called the *binary numeration system*, it has the binary alphabet $\Sigma_U = \Sigma = \{\mathtt{0}, \mathtt{1}\}$ and the value map denoted val_2 acting on a word $w = w_{k-1} \cdots w_1 w_0 \in \Sigma^*$ as

$$\mathrm{val}_2 : w_{k-1} \cdots w_1 w_0 \mapsto \sum_{i=0}^{k-1} w_i 2^i. \tag{3.3}$$

The numeration language is Σ^*. The representations $\mathrm{rep}_2(n)$ in the binary numeration system of the integers $n \in [\![0, 7]\!]$ are shown in Table 3.1.

Example 3.1.3 (Fibonacci numeration system)**.** The sequence of the Fibonacci numbers $(F_n)_{n=0}^{+\infty}$ is defined by the recurrence relation $F_n = F_{n-1} + F_{n-2}$, for all $n \geq 2$, with the initial conditions $F_0 = 1$, $F_1 = 2$. It holds that $\lim_{n\to+\infty} \frac{F_{n+1}}{F_n} = \tau$, where τ denotes the golden mean. Therefore the sequence $U = (F_n)_{n=0}^{+\infty}$ is strictly increasing, $U_0 = 1$ and $C_U = \sup_{n\geq 0}\lceil \frac{F_{n+1}}{F_n} \rceil = 2$ is finite. The positional numeration system U is called the *Fibonacci numeration system*, it has the binary alphabet $\Sigma_U = \Sigma$ and the value map acting on a word $w = w_{k-1} \cdots w_1 w_0 \in \Sigma^*$ as

$$\mathrm{val}_{\mathcal{F}} : w_{k-1} \cdots w_0 \mapsto \sum_{i=0}^{k-1} w_i F_i. \tag{3.4}$$

The numeration language is $\Sigma^* \setminus (\Sigma^* \mathtt{11} \Sigma^*)$ as proved in [Lek52,Day60,Car68,Zec72]. We reformulate the corresponding theorem, which is usually called Zeckendorf's theorem, in Theorem 3.1.4. The representations $\mathrm{rep}_{\mathcal{F}}(n)$ of the integers $n \in [\![0, 7]\!]$ are shown in Table 3.1.

Theorem 3.1.4. [Zec72] *The map* $\mathrm{val}_{\mathcal{F}} \colon \Sigma^* \setminus (\mathtt{0}\Sigma^* \cup \Sigma^*\mathtt{11}\Sigma^*) \to \mathbb{N}$ *is a bijection.*

Positional numeration systems are classically defined for nonnegative integers $\mathbb{N}$ and this theoretical framework is widely explored. To be able to perform arithmetical operations in a computer, various ways were described to extend the numeration systems to $\mathbb{Z}$. Let Σ denote an integer alphabet. A positional numeration system U for $\mathbb{Z}$ has a value map $\mathrm{val}_U : \Sigma^* \to \mathbb{Z}$. We aim to find a subset $\mathcal{L} \subset \Sigma^*$ which fulfills that $\mathrm{val}_U : \mathcal{L} \to \mathbb{Z}$ is a bijection. Then we call the inverse map val_U^{-1} the representation map and we denote it $\mathrm{rep}_U : \mathbb{Z} \to \mathcal{L}$. Clearly, then $\mathrm{val}_U(\mathrm{rep}_U(n)) = n$.

An intuitive approach to extend a numeration system for $\mathbb{N}$ to $\mathbb{Z}$ is interpreting the most significant digit as the sign. However, a disadvantage of this approach is that minus zero and plus zero have distinct representations. Examples that avoid this disadvantage include the two's complement notation, which generalizes the binary numeration system [Knu69, §4.1], and the negaFibonacci numeration system, which generalizes the Fibonacci numeration system [Knu11, §7.1.3].

Example 3.1.5 (Two's complement notation)**.** The sequence of integers $U = (2^n)_{n=0}^{+\infty}$ and the alphabet Σ_U is the same as for the binary numeration system presented in Example 3.1.2. However, its value map val_{2c} opposed to the binary value map val_2 has the term, which contains the most significant digit, multiplied by (-1):

$$\mathrm{val}_{2c} : w_{k-1} \cdots w_1 w_0 \mapsto \sum_{i=0}^{k-2} w_i 2^i - w_{k-1} 2^{k-1}. \tag{3.5}$$

Note that (3.6) can also be expressed in another form, which we use in the next parts

$$\mathrm{val}_{2c} : w_{k-1} \cdots w_1 w_0 \mapsto \sum_{i=0}^{k-1} w_i 2^i - w_{k-1} 2^{k}. \tag{3.6}$$

It holds that $\mathrm{val}_{2c} : \Sigma^+ \setminus (\mathtt{00}\Sigma^* \cup \mathtt{11}\Sigma^*) \to \mathbb{Z}$ is a bijection. We prove this in a broader framework in Section 3.3. We show the representations $\mathrm{rep}_{2c}(n)$ of the integers $n \in [\![-5, 7]\!]$ in Table 3.1. We observe that $\mathrm{rep}_{2c}(n) \in \mathtt{0}\Sigma^*$ if and only if $n \in \mathbb{N}$. We refer to the two's complement notation also as to the two's complement numeration system.

Example 3.1.6 (NegaFibonacci numeration system)**.** The sequence of negaFibonacci numbers $(F_n)_{-\infty}^{+\infty}$ is given by the same recurrence relation as the Fibonacci numbers $F_n = F_{n-1} + F_{n-2}$, but it is defined for all $n \in \mathbb{Z}$. Therefore we have

$$\ldots,\ F_{-6} = -3,\ F_{-5} = 2,\ F_{-4} = -1,\ F_{-3} = 1,\ F_{-2} = 0,\ F_{-1} = 1,\ F_0 = 1,\ F_1 = 2,\ \ldots$$

The terms $(F_n)_{n=-3}^{-\infty}$ can be used to define the negaFibonacci numeration system for $\mathbb{Z}$. It has the binary alphabet $\Sigma_U = \Sigma$ and the value map denoted $\mathrm{val}_{\mathrm{neg}\mathcal{F}}$ acting on a binary word $w = w_{k-1} \cdots w_1 w_0 \in \Sigma^*$ as $\mathrm{val}_{\mathrm{neg}\mathcal{F}} : w_{k-1} \cdots w_1 w_0 \mapsto \sum_{i=0}^{k-1} w_i F_{-i-3}$. The map $\mathrm{val}_{\mathrm{neg}\mathcal{F}}$ is a bijection $\Sigma^* \setminus (\mathtt{0}\Sigma^* \cup \Sigma^*\mathtt{11}\Sigma^*) \to \mathbb{Z}$. The representations $\mathrm{rep}_{\mathrm{neg}\mathcal{F}}(n)$ of the integers $n \in [\![-5, 7]\!]$ are shown in Table 3.1. We observe that $\mathrm{rep}_{\mathrm{neg}\mathcal{F}}(n)$ has odd length if and only if $n \in \mathbb{N}$.

n	$\mathrm{rep}_2(n)$	$\mathrm{rep}_{\mathcal{F}}(n)$	$\mathrm{rep}_{2c}(n)$	$\mathrm{rep}_{\mathrm{neg}\mathcal{F}}(n)$	$\mathrm{rep}_{\mathcal{G}}(n)$
7	111	1010	0111	10100	01010
6	110	1001	0110	10001	01001
5	101	1000	0101	10000	01000
4	100	101	0100	10010	0101
3	11	100	011	101	0100
2	10	10	010	100	010
1	1	1	01	1	01
0	ε	ε	0	ε	0
-1	-	-	1	10	1
-2	-	-	10	1001	10
-3	-	-	101	1000	100
-4	-	-	100	1010	1001
-5	-	-	1011	100101	1000

Table 3.1: The binary, the Fibonacci, the two's complement, the negaFibonacci numeration system and the numeration system $\mathcal{G}$.

3.2 Fibonacci complement numeration system

The two's complement notation presented in Example 3.1.5 extends the binary numeration system for $\mathbb{N}$ to all integers $\mathbb{Z}$. Similarly, the negaFibonacci numeration system presented in Example 3.1.6 extends the Fibonacci numeration system for $\mathbb{N}$ to all integers $\mathbb{Z}$. However, these two approaches differ significantly. In this section, we extend the Fibonacci numeration system to all integers using the approach of the two's complement notation.

An analogue of the two's complement notation emerges if we choose the value map to have the same structure as (3.6), using the Fibonacci numbers instead of powers of 2:

$$\mathrm{val}_{\mathcal{F}c} : w_{k-1} \cdots w_0 \mapsto \sum_{i=0}^{k-1} w_i F_i - w_{k-1} F_k. \tag{3.7}$$

The representations in the numeration system derived from (3.7) cannot include $\mathtt{0}$, $\mathtt{1}$ and representations of all lengths, because we have that $\mathrm{val}_{\mathcal{F}c}(\mathtt{0}) = 0$, $\mathrm{val}_{\mathcal{F}c}(\mathtt{1}) = -1$, $\mathrm{val}_{\mathcal{F}c}(\mathtt{10}) = -1$, $\mathrm{val}_{\mathcal{F}c}(\mathtt{11}) = 0$. Even though this might seem as a disadvantage at first glance, we will see later that the contrary is true. A choice was made in [LL23a] to only consider odd-length representations not containing consecutive $\mathtt{1}$s, which resulted in the following proposition. In this text, we carry out its proof in a broader framework in Section 3.3; see Proposition 3.3.3 and Example 3.3.7.

Proposition 3.2.1. *The map* $\mathrm{val}_{\mathcal{F}c} \colon \Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*) \to \mathbb{Z}$ *is a bijection.*

It is thus possible to define the representation map

$$\mathrm{rep}_{\mathcal{F}c} : \mathbb{Z} \to \Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*)$$

as the inverse of the map $\mathrm{val}_{\mathcal{F}c}$. We show the representations of integers $n \in [\![-10, 19]\!]$ in the numeration system $\mathcal{F}c$ in Table 3.2.

n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$	n	$\mathrm{rep}_{\mathcal{F}c}(n)$
−10	1000100	0	0	10	0010010
−9	1000101	1	001	11	0010100
−8	1001000	2	010	12	0010101
−7	1001001	3	00100	13	0100000
−6	1001010	4	00101	14	0100001
−5	10000	5	01000	15	0100010
−4	10001	6	01001	16	0100100
−3	10010	7	01010	17	0100101
−2	100	8	0010000	18	0101000
−1	1	9	0010001	19	0101001

Table 3.2: The Fibonacci complement numeration system $\mathcal{F}c$.

Remark 3.2.2. *It might seem more natural to construct an analogue of the two's complement notation with the value map having the same structure as* (3.5)*:*

$$\mathrm{val}_{\mathcal{G}} : w_{k-1} \cdots w_1 w_0 \mapsto \sum_{i=0}^{k-2} w_i F_i - w_{k-1} F_{k-1}. \tag{3.8}$$

Observe that (3.7) *and* (3.8) *do not coincide. It is indeed possible to construct a numeration system for* $\mathbb{Z}$ *with the value map* $\mathrm{val}_{\mathcal{G}}$*, which is a bijection* $\Sigma^+ \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{101}\Sigma^* \cup \mathtt{00}\Sigma^*) \to \mathbb{Z}$. *We omit the proof of this fact as this numeration system is not of interest for our study, which we explain further in Section 3.3, Chapter 6 and Chapter 7. However, we still show in Table 3.1 the representations, which might indeed look more natural and closer to the original two's complement numeration system than those in the numeration system* $\mathcal{F}c$.

We observe that the Fibonacci complement numeration system has the following interesting property. In Proposition 3.2.1, we excluded words starting with the prefixes $\texttt{000}$ and $\texttt{101}$ so that the value map $\mathrm{val}_{\mathcal{F}}$ is a bijection. Moreover, these prefixes can be used to pad words without changing their value, as stated in the following lemma. Its proof follows easily from our work in a broader framework in Section 3.3; see Example 3.3.7 and Equation (3.16).

Lemma 3.2.3. *For every word $w \in \Sigma^*$, we have*

$$\mathrm{val}_{\mathcal{F}c}(\texttt{000}w) = \mathrm{val}_{\mathcal{F}c}(\texttt{0}w) \quad \textit{and} \quad \mathrm{val}_{\mathcal{F}c}(\texttt{101}w) = \mathrm{val}_{\mathcal{F}c}(\texttt{1}w).$$

Hence, Lemma 3.2.3 leads to the definition of the *neutral prefix*. In Section 3.3, the concept of a neutral prefix is slightly modified to give rise to *neutral words*; see Equation (3.16).

Definition 3.2.4 (Neutral prefix of $\mathcal{F}c$)**.** *Let $w \in \Sigma^*$. We say that* $\texttt{00}$ *(resp.,* $\texttt{10}$*) is the* neutral prefix *of w if $w \in \texttt{0}\Sigma^*$ (resp., if $w \in \texttt{1}\Sigma^*$).*

We show in Section 3.5 that the numeration system $\mathcal{F}c$ has interesting properties with respect to addition. In Chapter 6, we recover the numeration system $\mathcal{F}c$ as a numeration system based on the Fibonacci substitution, and in Chapter 7, we use the numeration system $\mathcal{F}c$ extended to $\mathbb{Z}^2$ to describe a particular Wang subshift.

3.3 Complement numeration systems associated with simple Parry numbers

The classical binary and Fibonacci numeration system for $\mathbb{N}$ fall into a broader framework of numeration systems associated with simple Parry numbers. The two's complement numeration system extends the binary numeration system to $\mathbb{Z}$ and the Fibonacci complement numeration system extends the Fibonacci numeration system to $\mathbb{Z}$. We extend all numeration systems associated with simple Parry numbers to $\mathbb{Z}$ in an analogous way.

First, we summarize the background of numeration systems canonically associated with real numbers β. Let $\beta > 1$ be a real number and denote $d^*_\beta(1) = (x_i)_{i\geq 1}$ its quasi-greedy expansion of unity. The numeration system $U_\beta = (U_n)_{n=0}^{+\infty}$ canonically associated with β is defined by the relation

$$U_n = \sum_{i=1}^{n} x_i U_{n-i} + 1, \text{ for every } n \geq 0. \tag{3.9}$$

The canonical alphabet Σ_β and the representation map rep_β of the numeration system U_β are defined as in Section 3.1, i.e., $\Sigma_\beta = [\![0, C_U - 1]\!]$, where $C_U := \sup_{n\geq 0}\lceil \frac{U_{n+1}}{U_n}\rceil$ is finite, and the representation map $\mathrm{rep}_\beta : \mathbb{N} \to \Sigma^*_\beta$ maps every integer $n \in \mathbb{N}$ to its greedy representation. A greedy representation of an integer $n \in \mathbb{N}$ is a word $w = w_{k-1}\cdots w_0$ over the alphabet Σ_β, which satisfies the condition that

$$\sum_{i=0}^{k-1} w_i U_i = n, \quad w_{k-1} \neq 0, \quad \text{and} \quad \sum_{i=0}^{j-1} w_i U_i < U_j, \text{ for every } j \in [\![1, k]\!].$$

The value map val_β maps every word $w = w_{k-1}\cdots w_0 \in A^*$ over an integer alphabet A to the sum $\mathrm{val}_\beta(w) = \sum_{i=0}^{k-1} w_i U_i$. Consequently, $\mathrm{val}_\beta(\mathrm{rep}_\beta(n)) = n$ for every $n \in \mathbb{N}$ and U_β is a positional numeration system for $\mathbb{N}$. We recall that a Parry number is such a real $\beta > 1$ that its greedy expansion of unity $d_\beta(1)$ is eventually periodic and, in particular, a Parry number

β is called simple if $d_\beta(1)$ is finite. A numeration system canonically associated with a Parry number β is called a Parry numeration system. Parry numeration systems are a strict subset of Bertrand numeration systems, which are positional numeration systems called after Bertrand-Mathis [BM89, BH97]. They were studied recently in [CCS22].

From now on, we assume that $\beta > 1$ is a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 \cdots t_m$. We can derive from Equation (3.9) the following linear recurrence relation for the numeration system $U_\beta = (U_n)_{n=0}^{+\infty}$:

$$U_n = \sum_{i=1}^{m} t_i U_{n-i}, \text{ for every } n \geq m. \tag{3.10}$$

The numeration language preceded by leading zeroes $\mathtt{0}^* \operatorname{rep}_\beta(\mathbb{N})$ is a regular language [Hol98]. We denote $\mathcal{H}_{\beta,q_1}$ the trim minimal automaton which accepts the numeration language:

$$\mathcal{L}(\mathcal{H}_{\beta,q_1}) = \mathtt{0}^* \operatorname{rep}_\beta(\mathbb{N}). \tag{3.11}$$

It is shown in [Lot02, Theorem 7.2.13] that $\mathcal{H}_{\beta,q_1}$ has a particular form.

Remark 3.3.1. *Denote $d^*_\beta(1) = (x_1 x_2 \cdots x_m)^\omega$. The automaton $\mathcal{H}_{\beta,q_1}$ has the set of states $Q_\beta = \{q_1, \ldots, q_m\}$, all of which are final. The initial state is q_1. For every $j \in [\![1, m]\!]$, there are x_j edges $q_j \to q_1$ labeled $\mathtt{0}, \ldots, (x_j - 1)$, and for every $j \in [\![1, m-1]\!]$ there is one edge $q_j \to q_{j+1}$ labeled x_j. There is also one edge $q_m \to q_1$ labeled x_m.*

We show the automata associated with $\beta = 2$, the golden mean $\beta = \tau$ and β the real root of the irreducible polynomial $x^3 - 3x^2 + 2x - 2$ in Figure 3.1. We denote $\mathcal{L}_n(\mathcal{H}_{\beta,q_1}) = \{w \in \mathcal{L}(\mathcal{H}_{\beta,q_1}) \colon |w| = n\}$ and we denote $\mathcal{H}_{\beta,q_\chi}$ the automaton which arises from $\mathcal{H}_{\beta,q_1}$ by changing its initial state to the state q_χ, for a fixed state $\chi \in [\![1, m]\!]$.

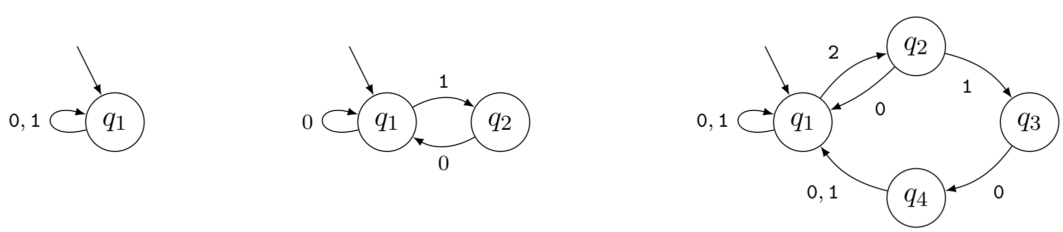


Figure 3.1: The automata $\mathcal{H}_{\beta,q_1}$ associated with the simple Parry numbers $\beta \in \{2, \tau, \beta'\}$, where β' is the real root of the irreducible polynomial $x^3 - 3x^2 + 2x - 2$. If $\beta = 2$, then $d_\beta(1) = 2$ and $d^*_\beta(1) = 1^\omega$. If $\beta = \tau$, then $d_\beta(1) = 11$ and $d^*_\beta(1) = (10)^\omega$. If $\beta = \beta'$, then $d_\beta(1) = 2102$ and $d^*_\beta(1) = (2101)^\omega$ [Bas02].

The following lemma can be derived from Equation (3.11).

Lemma 3.3.2. *The automaton $\mathcal{H}_{\beta,q_1}$ has the property that $U_n = \#\mathcal{L}_n(\mathcal{H}_{\beta,q_1})$, for every $n \in \mathbb{N}$.*

We aim to define a new value map that will enable us to evaluate words over an integer alphabet as both nonnegative and negative integers. Let $\chi \in [\![1, m]\!]$. We base the value map on the automaton $\mathcal{H}_{\beta,\chi}$, which arises from the automaton $\mathcal{H}_{\beta,q_1}$ by creating a new initial state $\mathtt{start}$ and adding two additional edges $\mathtt{start} \xrightarrow{\mathtt{0}} q_1$ and $\mathtt{start} \xrightarrow{\mathtt{1}} q_\chi$; compare Figure 3.1 and Figure 3.2. Clearly, then the language $\mathcal{L}(\mathcal{H}_{\beta,\chi}) = \mathtt{0}\mathcal{L}(\mathcal{H}_{\beta,q_1}) \cup \mathtt{1}\mathcal{L}(\mathcal{H}_{\beta,q_\chi})$. We denote $\mathtt{W}^\chi_{\min}$ the label of the path of length m in the automaton $\mathcal{H}_{\beta,\chi}$ starting at the state q_1 and following the

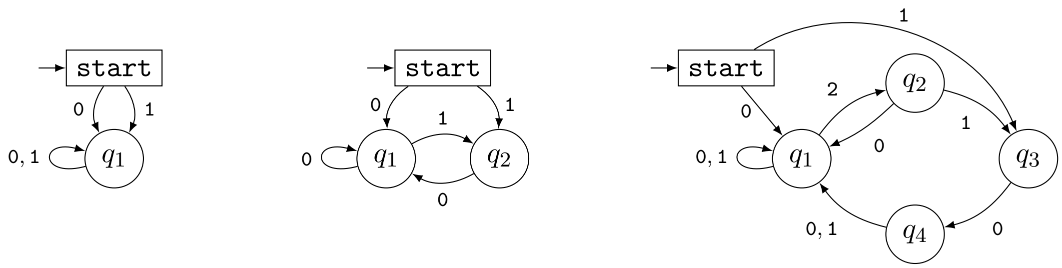


Figure 3.2: The automata $\mathcal{H}_{2,1}$, $\mathcal{H}_{\tau,2}$ and $\mathcal{H}_{\beta',3}$ which were created by modifying the automata in Figure 3.1. We recall that β' denotes the real root of the irreducible polynomial $x^3 - 3x^2 + 2x - 2$.

edges with minimal lexicographical value and we denote $\mathtt{W}_{\max}^{\chi}$ the label of the path of length m in the automaton $\mathcal{H}_{\beta,\chi}$ starting at the state q_χ and following the edges with maximal lexicographical value. Therefore we have

$$\mathtt{W}_{\min}^{\chi} = \mathtt{0}^m \quad \text{and} \quad \mathtt{W}_{\max}^{\chi} = t_\chi \cdots t_{m-1}(t_m - 1)t_1 \cdots t_{\chi-1}. \tag{3.12}$$

We observe that the words $\mathtt{W}_{\min}^{\chi}$ and $\mathtt{W}_{\max}^{\chi}$ fulfill the condition that, for every $w \in \mathcal{L}(\mathcal{H}_{\beta,\chi})$,

$$\mathcal{H}_{\beta,\chi}(w) = \begin{cases} \mathcal{H}_{\beta,\chi}(\mathtt{0}(\mathtt{W}_{\min}^{\chi})^* v), & \text{if } w = \mathtt{0}v; \\ \mathcal{H}_{\beta,\chi}(\mathtt{1}(\mathtt{W}_{\max}^{\chi})^* v), & \text{if } w = \mathtt{1}v. \end{cases} \tag{3.13}$$

We define, for every $n \in \mathbb{Z}$,

$$V_{n,\chi} = \#\mathcal{L}_n(\mathcal{H}_{\beta,q_\chi}). \tag{3.14}$$

As a consequence of Lemma 3.3.2, we have $(V_{n,1})_{n=0}^{+\infty} = (U_n)_{n=0}^{+\infty}$. Also, by definition, we have $V_{n,\chi} = 0$, for every $n < 0$. We define the value map $\mathrm{val}_{\beta,\chi}$ so that, for every word $w = w_{k-1} \cdots w_0 \in A^k$ over an integer alphabet, we have

$$\mathrm{val}_{\beta,\chi}(w) = \textstyle\sum_{i=0}^{k-2} w_i U_i - w_{k-1} V_{k-1,\chi}. \tag{3.15}$$

We prove the following proposition in Section 3.3.1.

Proposition 3.3.3. *Let β be a simple Parry number with the quasi-greedy expansion of unity $d_\beta^*(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$ and let $\Sigma = \Sigma_\beta$ be the canonical alphabet of U_β. Let $\chi \in [\![1, m]\!]$. Denote $\mathtt{W}_{\min}^{\chi} = \mathtt{0}^m$ and $\mathtt{W}_{\max}^{\chi} = t_\chi \cdots t_{m-1}(t_m - 1)t_1 \cdots t_{\chi-1}$. Then the map $\mathrm{val}_{\beta,\chi}$ is a bijection:*

$$\mathrm{val}_{\beta,\chi} \colon \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}_{\min}^{\chi}\Sigma^* \cup \mathtt{1}\mathtt{W}_{\max}^{\chi}\Sigma^*) \to \mathbb{Z}.$$

With the help of Proposition 3.3.3, we extend the numeration systems for $\mathbb{N}$ associated with simple Parry numbers to $\mathbb{Z}$. More precisely, a numeration system U_β associated with a simple Parry number β such that $d_\beta(1) = t_1 \cdots t_m$ can be extended to $\mathbb{Z}$ in m different ways. We call these numeration systems *complement numeration systems* as their value map has the same structure as the two's complement numeration system; compare Equations (3.6) and (3.15). Their representation map is defined as the inverse of the map $\mathrm{val}_{\beta,\chi}$.

Definition 3.3.4 (Complement numeration system)**.** *Let $\beta > 1$ be a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 t_2 \cdots t_m$. Let $\chi \in [\![1, m]\!]$. We define $\mathrm{rep}_{\beta,\chi} = \mathrm{val}_{\beta,\chi}^{-1}$.*

Remark 3.3.5. *The language* $\operatorname{rep}_{\beta,\chi}(\mathbb{Z})$ *contains words of length* $m\ell+1$*, for all* $\ell \in \mathbb{N}$*. In particular,* $\mathtt{0}$ *and* $\mathtt{1}$ *belong to the language* $\operatorname{rep}_{\beta,\chi}(\mathbb{Z})$ *and we have that*

$$\operatorname{rep}_{\beta,\chi}(0) = \mathtt{0} \quad \textit{and} \quad \operatorname{rep}_{\beta,\chi}(-1) = \mathtt{1}.$$

Indeed, from Equation (3.15) *we observe that* $\operatorname{val}_{\beta,\chi}(\mathtt{0}) = 0$ *and* $\operatorname{val}_{\beta,\chi}(\mathtt{1}) = -1$*.*

The two's complement numeration system and the Fibonacci complement numeration system $\mathcal{F}c$ belong to the class of complement numeration systems; see Example 3.3.6 and Example 3.3.7. The numeration system $\mathcal{G}$ from Remark 3.2.2 is not a complement numeration system; see Example 3.3.8. To illustrate a complement numeration system on a larger alphabet than the binary one, we show the complement numeration system associated with β the real root of the irreducible polynomial $x^3 - 3x^2 + 2x - 2$ in Example 3.3.9.

Example 3.3.6. Let $\beta = 2$. Then $d_\beta(1) = 2$ and $d^*_\beta(1) = 1^\omega$. From Equation (3.9), we have $U_0 = 1$, and from Equation (3.10), $U_n = 2U_{n-1}$ for every $n \geq 1$. Thus $U_\beta = (2^n)_{n=0}^{+\infty}$. As the greedy expansion of unity $d_\beta(1)$ has length 1, we have only one choice for $\chi = 1$. In Figure 3.2, we observe that $V_{n,\chi} = \#\mathcal{L}_n(\mathcal{H}_{\beta,q_1}) = U_n$, for every $n \in \mathbb{N}$. Then $\operatorname{val}_{\beta,\chi}(w) = \sum_{i=0}^{k-2} w_i 2^i - w_{k-1}2^{k-1}$, which is exactly the value map of the two's complement notation from Equation (3.5).

Example 3.3.7. Let $\beta = \tau$. Then $d_\beta(1) = 11$ and $d^*_\beta(1) = (10)^\omega$. From Equation (3.9), we have $U_0 = 1$ and $U_1 = 2$, and from Equation (3.10), $U_n = U_{n-1} + U_{n-2}$ for every $n \geq 2$. Thus $U_\beta = (F_n)_{n=0}^{+\infty}$, the sequence of Fibonacci numbers. As the greedy expansion of unity $d_\beta(1)$ has length 2, we have two possible choices of $\chi \in \{1, 2\}$. We choose $\chi = 2$. In Figure 3.2, we observe that $V_{n,\chi} = \#\mathcal{L}_n(\mathcal{H}_{\beta,q_2}) = \#\mathtt{0}\mathcal{L}_{n-1}(\mathcal{H}_{\beta,q_1}) = U_{n-1}$, for every $n \geq 1$. Therefore

$$\operatorname{val}_{\beta,\chi}(w) = \sum_{i=0}^{k-2} w_i F_i - w_{k-1}F_{k-2} = \sum_{i=0}^{k-1} w_i F_i - w_{k-1}(F_{k-2} + F_{k-1}) = \sum_{i=0}^{k-1} w_i F_i - w_{k-1}F_k,$$

which is the value map of the Fibonacci complement numeration system $\mathcal{F}c$ from Equation (3.7).

Example 3.3.8. Assume by contradiction that there exists a simple Parry number β such that $d_\beta(1) = t_1 \cdots t_m$ and there exists a state $\chi \in [\![1, m]\!]$ so that, for every $n \in [\![-5, 7]\!]$, we have $\operatorname{rep}_{\beta,\chi}(n) = \operatorname{rep}_{\mathcal{G}}(n)$ as shown in Table 3.1. As the $\mathcal{G}$-representations do not exclude words of any length, it follows from Remark 3.3.5 that $m = 1$. The simple Parry numbers with the greedy expansion of unity of length 1 are exactly the integers $\beta \geq 2$. As the $\mathcal{G}$-representations are words over the binary alphabet, the only option left is $\beta = 2$. But the only complement numeration system associated with $\beta = 2$ is the two's complement system, which does not coincide with $\mathcal{G}$.

Example 3.3.9. . Let β' be the real root of the irreducible polynomial $x^3 - 3x^2 + 2x - 2$. Then $d_{\beta'}(1) = 2102$ and $d^*_{\beta'}(1) = (2101)^\omega$ [Bas02]. From Equation (3.9), we have $U_0 = 1$, $U_1 = 3$, $U_2 = 8$, $U_3 = 20$, and from Equation (3.10), we have $U_n = 2U_{n-1} + U_{n-2} + 2U_{n-4}$, for every $n \geq 4$. As the greedy expansion of unity $d_{\beta'}(1)$ has length 4, we have 4 possible choices of $\chi \in [\![1, 4]\!]$. We choose $\chi = 3$. From Figure 3.2, we observe that $V_{0,\chi} = 1$, $V_{1,\chi} = 1$, and, for every $n \geq 2$,

$$V_{n,\chi} = \#\mathcal{L}_n(\mathcal{H}_{\beta,q_3}) = \#\mathtt{0}\{\mathtt{0},\mathtt{1}\}\mathcal{L}_{n-2}(\mathcal{H}_{\beta,q_1}) = 2U_{n-2}.$$

Then, for every word $w = w_{k-1} \cdots w_0$ of length $k \geq 2$ over an integer alphabet, we have

$$\operatorname{val}_{\beta',\chi}(w) = \sum_{i=0}^{k-2} w_i U_i - w_{k-1}V_{k-1,\chi} = \sum_{i=0}^{k-2} w_i U_i - w_{k-1}U_{k-3}.$$

We show in Table 3.3 all words of length $4\ell+1$, $\ell \in \{0,1\}$, accepted by the automaton $\mathcal{H}_{\beta',\chi}$ evaluated by the value map $\mathrm{val}_{\beta',\chi}$. We observe that all words have distinct values, expect for two cases. The words $\mathtt{0}$ and $\mathtt{0W}^{\chi}_{\mathrm{min}} = \mathtt{00000}$ have the same value and the words $\mathtt{1}$ and $\mathtt{1W}^{\chi}_{\mathrm{max}} = \mathtt{10121}$ have the same value. Indeed, $\mathtt{0W}^{\chi}_{\mathrm{min}}$ and $\mathtt{1W}^{\chi}_{\mathrm{max}}$ are excluded in Proposition 3.3.3.

w	$\mathrm{val}_{\beta',\chi}(w)$	w	$\mathrm{val}_{\beta',\chi}(w)$	w	$\mathrm{val}_{\beta',\chi}(w)$	w	$\mathrm{val}_{\beta',\chi}(w)$
10000	−16	0	0	00200	16	01112	33
10001	−15	00000	0	00201	17	01120	34
10002	−14	00001	1	00202	18	01121	35
10010	−13	00002	2	00210	19	01200	36
10011	−12	00010	3	01000	20	01201	37
10012	−11	00011	4	01001	21	01202	38
10020	−10	00012	5	01002	22	01210	39
10021	−9	00020	6	01010	23	02000	40
10100	−8	00021	7	01011	24	02001	41
10101	−7	00100	8	01012	25	02002	42
10102	−6	00101	9	01020	26	02010	43
10110	−5	00102	10	01021	27	02011	44
10111	−4	00110	11	01100	28	02012	45
10112	−3	00111	12	01101	29	02020	46
10120	−2	00112	13	01102	30	02021	47
10121	−1	00120	14	01110	31	02100	48
1	−1	00121	15	01111	32	02101	49

Table 3.3: A complement numeration system associated with β', the real root of $x^3 - 3x^2 + 2x - 2$, and $\chi = 3$.

We believe that the construction of complement numeration systems may be carried out with some modifications for the numeration systems associated with non-simple Parry numbers as well. For now, we leave this as an open question.

Question 3.3.10. *Can we construct complement numeration systems for $\mathbb{Z}$ associated with non-simple Parry numbers in analogy with Definition 3.3.4?*

3.3.1 Proof of Proposition 3.3.3

We summarize some existing results and prove several lemmas in order to prove Proposition 3.3.3. As U_β is a positional numeration system for $\mathbb{N}$, the language $\mathrm{rep}_\beta(\mathbb{N})$ is ordered with respect to the radix order; see Proposition 3.1.1. Combining Proposition 3.1.1 with Equation (3.11), we reformulate this property in the following lemma.

Lemma 3.3.11. *Let $\beta > 1$ be a simple Parry number and $U = U_\beta$ be the numeration system associated with β. Then the map $\mathrm{val}_\beta \colon (\mathcal{L}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0}\Sigma_\beta^*, <_{\mathrm{rad}}) \to (\mathbb{N}, <)$ is an increasing bijection.*

We recall that the automaton $\mathcal{H}_{\beta,q_1}$ plays an essential role for a numeration system U_β and the automaton $\mathcal{H}_{\beta,q_\chi}$ is created by changing the initial state of $\mathcal{H}_{\beta,q_1}$ to q_χ. We show that the words accepted by the automaton $\mathcal{H}_{\beta,q_\chi}$ are also accepted by the automaton $\mathcal{H}_{\beta,q_1}$.

Lemma 3.3.12. *Let β be a simple Parry number with the quasi-greedy expansion of unity $d^*_\beta(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$ and let $\Sigma = \Sigma_\beta$ be the canonical alphabet of U_β. Let $\chi \in [\![1, m]\!]$. For every word $w \in \Sigma^*$, we have that if $w \in \mathcal{L}(\mathcal{H}_{\beta,q_\chi})$, then $w \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$.*

Proof. If $w = \varepsilon$ is the empty word, the statement holds. Let $w = w_{k-1} \cdots w_0 \in \mathcal{L}(\mathcal{H}_{\beta,q_\chi})$ such that $|w| > 0$. Then $t_1 \cdots t_{\chi-1} w \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$. From Equation (3.11) and the fact that $t_1 \neq 0$, there exists a nonnegative integer $n \geq 0$ such that $\mathrm{rep}_\beta(n) = t_1 \cdots t_{\chi-1} w$. As this representation is greedy, we have that $\mathrm{val}_\beta(w_j \cdots w_0) < U_{j+1}$, for every $j \in [\![0, k-1]\!]$. We denote $u \in \Sigma^*$ the word of maximal length so that $w \in 0^* u$. Then $u = u_{\ell-1} \cdots u_0$ fulfills $u_{\ell-1} \neq 0$ and it is thus a greedy representation. Therefore $w \in 0^* \mathrm{rep}_\beta(\mathbb{N})$ and the statement holds from Equation (3.11). □

The following observation holds from the definition of the map $\mathrm{val}_{\beta,\chi}$ in Equation (3.15).

Lemma 3.3.13. *Let β be a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 t_2 \cdots t_m$. Let $\chi \in [\![1, m]\!]$ and let $w \in \Sigma_\beta^*$. Then $\mathrm{val}_{\beta,\chi}(\mathtt{0}^\ell w) = \mathrm{val}_\beta(w) \geq 0$, for every $\ell \geq 1$.*

The following lemma helps us to express the values of $V_{n,\chi}$ as a function of (U_n) and the greedy expansion of unity. Its proof is based on the definition of the automata $\mathcal{H}_{\beta,q_1}$ and $\mathcal{H}_{\beta,q_\chi}$.

Lemma 3.3.14. *Let β be a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 t_2 \cdots t_m$. Let $\chi \in [\![1, m]\!]$. Let $U = U_\beta$ denote the numeration system canonically associated with β and $V = (V_{n,\chi})_{n=0}^{+\infty}$ be the sequence defined in Equation (3.14). Then*

i) $V_{n,\chi} = \sum_{i=\chi}^m t_i U_{n+\chi-1-i}$, for every $n \geq m - \chi + 1$;

ii) $V_{m\ell,\chi} = U_{m\ell+\chi-1} - \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i}$, for every $\ell \in \mathbb{N}$.

Proof. i) Paths starting at q_χ in the automaton $\mathcal{H}_{\beta,q_\chi}$ pass through the state q_1 after at most $m - \chi + 1$ edges. From the form of the automaton $\mathcal{H}_{\beta,q_1}$, we have, for every $n \geq m - \chi + 1$,

$$\begin{aligned}
\mathcal{L}_n(\mathcal{H}_{\beta,q_\chi}) &= \bigsqcup_{i=\chi}^{m} t_\chi t_{\chi+1} \cdots t_{i-1} \{0, \ldots, t_i - 1\} \mathcal{L}_{n-(i-\chi+1)}(\mathcal{H}_{\beta,q_1}) \\
&= \bigsqcup_{i=\chi}^{m} t_\chi t_{\chi+1} \cdots t_{i-1} \{0, \ldots, t_i - 1\} \mathcal{L}_{n+\chi-1-i}(\mathcal{H}_{\beta,q_1}),
\end{aligned}$$

where the symbol $\bigsqcup$ denotes a disjoint union. As a consequence of Equation (3.14) and Lemma 3.3.2, we have

$$\begin{aligned}
V_{n,\chi} &= \#\mathcal{L}_n(\mathcal{H}_{\beta,q_\chi}) \\
&= \#\left(\bigsqcup_{i=\chi}^{m} t_\chi t_{\chi+1} \cdots t_{i-1} \{0, \ldots, t_i - 1\} \mathcal{L}_{n+\chi-1-i}(\mathcal{H}_{\beta,q_1}) \right) \\
&= \sum_{i=\chi}^{m} \#\Big(t_\chi t_{\chi+1} \cdots t_{i-1} \{0, \ldots, t_i - 1\} \mathcal{L}_{n+\chi-1-i}(\mathcal{H}_{\beta,q_1}) \Big) \\
&= \sum_{i=\chi}^{m} \#\Big(\{0, \ldots, t_i - 1\} \mathcal{L}_{n+\chi-1-i}(\mathcal{H}_{\beta,q_1}) \Big) \\
&= \sum_{i=\chi}^{m} t_i \#\mathcal{L}_{n+\chi-1-i}(\mathcal{H}_{\beta,q_1}) \Big) \\
&= \sum_{i=\chi}^{m} t_i U_{n+\chi-1-i}.
\end{aligned}$$

ii) Assume $\ell = 0$. Then we use Equation (3.9) for $n = \chi - 1$ to obtain

$$V_{m\ell,\chi} = V_{0,\chi} = 1 = U_{\chi-1} - \sum_{i=1}^{\chi-1} t_i U_{\chi-1-i} = U_{m\ell+\chi-1} - \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i}.$$

Assume $\ell > 0$. Then $m\ell \geq m - \chi + 1$ and we use part i) for $n = m\ell$ and Equation (3.10) for $n = m\ell + \chi - 1$ to obtain

$$\begin{aligned}
V_{m\ell,\chi} &= \sum_{i=\chi}^{m} t_i U_{m\ell+\chi-1-i} \\
&= U_{m\ell+\chi-1} - U_{m\ell+\chi-1} + \sum_{i=\chi}^{m} t_i U_{m\ell+\chi-1-i} \\
&= U_{m\ell+\chi-1} - \sum_{i=1}^{m} t_i U_{m\ell+\chi-1-i} + \sum_{i=\chi}^{m} t_i U_{m\ell+\chi-1-i} \\
&= U_{m\ell+\chi-1} - \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i}. \quad \square
\end{aligned}$$

In the next lemma, we show the neutral role of the word $\mathtt{W}^{\chi}_{\max}$ for the value map $\mathrm{val}_{\beta,\chi}$.

Lemma 3.3.15. *Let β be a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 t_2 \cdots t_m$. Let $\chi \in [\![0, m]\!]$ and $\ell \in \mathbb{N}$. Then for every $w \in \mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi})$, we have*

$$\mathrm{val}_{\beta,\chi}(\mathtt{1} t_\chi t_{\chi+1} \cdots t_{m-1}(t_m - 1) t_1 \cdots t_{\chi-1} w) = \mathrm{val}_{\beta,\chi}(\mathtt{1} w).$$

Proof. Let $\chi \in [\![0, m]\!]$, $\ell \in \mathbb{N}$ and let $w \in \mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi})$. Then we have

$$\begin{aligned}
&\mathrm{val}_{\beta,\chi}(\mathtt{1}w) - \mathrm{val}_\beta(w) = -V_{m\ell,\chi} \\
&= -U_{m\ell+\chi-1} + \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i} \\
&= -U_{m\ell+\chi-1} + \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i} - V_{m(\ell+1),\chi} + V_{m(\ell+1),\chi} \\
&= -U_{m\ell+\chi-1} + \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i} - V_{m(\ell+1),\chi} + \sum_{i=\chi}^{m} t_i U_{m(\ell+1)+\chi-1-i} \\
&= -V_{m(\ell+1),\chi} + \sum_{i=\chi}^{m} t_i U_{m(\ell+1)+\chi-1-i} - U_{m\ell+\chi-1} + \sum_{i=1}^{\chi-1} t_i U_{m\ell+\chi-1-i} \\
&= -V_{m(\ell+1),\chi} + t_\chi U_{m(\ell+1)-1} + \ldots + t_{m-1} U_{m\ell+\chi} + (t_m - 1) U_{m\ell+\chi-1} + t_1 U_{m\ell+\chi-2} + \ldots + t_{\chi-1} U_{m\ell} \\
&= \mathrm{val}_{\beta,\chi}(\mathtt{1} t_\chi t_{\chi+1} \cdots t_{m-1}(t_m - 1) t_1 \cdots t_{\chi-1} w) - \mathrm{val}_\beta(w),
\end{aligned}$$

where we used Lemma 3.3.14 ii) and Lemma 3.3.14 i) for $n = m(\ell + 1)$. $\square$

As a consequence of Lemma 3.3.13 and Lemma 3.3.15, we see that the words $\mathtt{W}^{\chi}_{\min}$ and $\mathtt{W}^{\chi}_{\max}$ fulfill the condition that

$$\mathrm{val}_{\beta,\chi}(w) = \begin{cases} \mathrm{val}_{\beta,\chi}(\mathtt{0}(\mathtt{W}^{\chi}_{\min})^* v), & \text{if } w = \mathtt{0}v; \\ \mathrm{val}_{\beta,\chi}(\mathtt{1}(\mathtt{W}^{\chi}_{\max})^* v), & \text{if } w = \mathtt{1}v. \end{cases} \tag{3.16}$$

That is why we call the words $\mathtt{W}^{\chi}_{\min}$ and $\mathtt{W}^{\chi}_{\max}$ *neutral*. We can insert a corresponding neutral word right after the first letter $\mathtt{0}$ (or $\mathtt{1}$) without changing the value of the original word. In the following lemma, we show that excluding words with a corresponding neutral word inserted enables us to determine, in which interval a word is evaluated.

Lemma 3.3.16. *Let β be a simple Parry number with the quasi-greedy expansion of unity $d^*_\beta(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$ and let $\Sigma = \Sigma_\beta$ be the canonical alphabet of U_β. Let $\chi \in [\![1, m]\!]$. Denote $\mathtt{W}^\chi_{\min} = \mathtt{0}^m$ and $\mathtt{W}^\chi_{\max} = t_\chi \cdots t_{m-1}(t_m - 1)t_1 \cdots t_{\chi-1}$. Let $u \in \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}^\chi_{\min}\Sigma^* \cup \mathtt{1}\mathtt{W}^\chi_{\max}\Sigma^*)$ and $\ell \in \mathbb{N}$. Then we have*

(1) $u \in \mathtt{0}\Sigma^{m\ell}$ if and only if $U_{m(\ell-1)} \le \mathrm{val}_{\beta,\chi}(u) < U_{m\ell}$;

(2) $u \in \mathtt{1}\Sigma^{m\ell}$ if and only if $-V_{m\ell,\chi} \le \mathrm{val}_{\beta,\chi}(u) < -V_{m(\ell-1),\chi}$.

Proof. First we prove the implication from left to right for both cases.

(1) We denote $u = \mathtt{0}w$. Assume $|w| = 0$. Then $U_{-1} = \mathrm{val}_{\beta,\chi}(\mathtt{0}) = 0 < 1 = U_0$. Now, assume that $|w| = m\ell$ for $\ell \ge 1$. By Lemma 3.3.12, the word w fulfills the condition that $w \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$. From Equation (3.11), we have that w is a greedy representation possibly preceded by leading zeroes. Thus

$$\mathrm{val}_{\beta,\chi}(\mathtt{0}w) = \mathrm{val}_\beta(w) < U_{m\ell}.$$

The lexicographically smallest word in the language $\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0}^m\Sigma^*$ is $w_{\min} = \mathtt{0}^{m-1}\mathtt{1}\mathtt{0}^{m(\ell-1)}$. Then using Lemma 3.3.13 and Lemma 3.3.11, we have

$$\mathrm{val}_{\beta,\chi}(\mathtt{0}w) = \mathrm{val}_\beta(w) \ge \mathrm{val}_\beta(w_{\min}) = \mathrm{val}_\beta(\mathtt{0}^{m-1}\mathtt{1}\mathtt{0}^{m(\ell-1)}) = U_{m(\ell-1)}.$$

(2) We denote $u = \mathtt{1}w$. Assume $|w| = 0$. Then $-V_{0,\chi} = \mathrm{val}_{\beta,\chi}(\mathtt{1}) = -1 < 0 = -V_{-m,\chi}$. Now, assume that $|w| = m\ell$ for $\ell \ge 1$. We observe that $\mathtt{0}^{m\ell} <_{lex} w$ and therefore $\mathtt{0}^{m\ell} <_{rad} w$. Thus using Lemma 3.3.11, we have

$$\mathrm{val}_{\beta,\chi}(\mathtt{1}w) = -V_{m\ell,\chi} + \mathrm{val}_\beta(w) \ge -V_{m\ell,\chi} + \mathrm{val}_\beta(\mathtt{0}^{m\ell}) = -V_{m\ell,\chi}.$$

Let $w_{\max}$ denote the lexicographically largest word in the language $\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{W}^\chi_{\max}\Sigma^*$. By Lemma 3.3.12, the word $w_{\max}$ fulfills the condition that $w_{\max} \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$. From Equation (3.11), we have that $w_{\max}$ is a greedy representation possibly preceded by leading zeroes and there exists a word $w' \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$ such that $\mathrm{val}_\beta(w') = \mathrm{val}_\beta(w_{\max}) + 1$. More precisely, $w' = \mathtt{W}^\chi_{\max}\mathtt{0}^{m(\ell-1)}$. We observe that $w' \in \mathcal{L}(\mathcal{H}_{\beta,q_\chi})$. Using Lemma 3.3.11 and Lemma 3.3.15, we have

$$\begin{aligned}
\mathrm{val}_{\beta,\chi}(\mathtt{1}w) &= -V_{m\ell,\chi} + \mathrm{val}_\beta(w) \\
&\le -V_{m\ell,\chi} + \mathrm{val}_\beta(w_{\max}) \\
&= -V_{m\ell,\chi} + \mathrm{val}_\beta(\mathtt{W}^\chi_{\max}\mathtt{0}^{m(\ell-1)}) - 1 \\
&= \mathrm{val}_{\beta,\chi}(\mathtt{1}\mathtt{W}^\chi_{\max}\mathtt{0}^{m(\ell-1)}) - 1 \\
&= \mathrm{val}_{\beta,\chi}(\mathtt{1}\mathtt{0}^{m(\ell-1)}) - 1 \\
&= -V_{m(\ell-1),\chi} - 1 \\
&< -V_{m(\ell-1),\chi}.
\end{aligned}$$

The converse follows from the following observation. Let $X = \bigcup_{\ell=0}^{+\infty} X^{(1)}_\ell \cup \bigcup_{\ell=0}^{+\infty} X^{(2)}_\ell$ be a disjoint union. Let $F\colon X \to Y$ be a map, such that $F(X^{(i_1)}_{i_2}) \cap F(X^{(j_1)}_{j_2}) = \emptyset$ for every $i_1, j_1 \in \{1, 2\}$ and $i_2, j_2 \in \mathbb{N}$ such that $X^{(i_1)}_{i_2} \ne X^{(j_1)}_{j_2}$. Then it holds that, for every $x \in X$, $F(x) \in F(X^{(i_1)}_{i_2})$ implies $x \in X^{(i_1)}_{i_2}$. Indeed, if $x \in X^{(j_1)}_{j_2}$ such that $X^{(i_1)}_{i_2} \ne X^{(j_1)}_{j_2}$, then $F(x) \in F(X^{(i_1)}_{i_2}) \cap F(X^{(j_1)}_{j_2})$, a contradiction. It suffices to use this observation with $F = \mathrm{val}_{\beta,\chi}$ and $X = \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}^\chi_{\min}\Sigma^* \cup \mathtt{1}\mathtt{W}^\chi_{\max}\Sigma^*)$. ∎

Now, we turn to the proof of Proposition 3.3.3. We will need a simple observation following from Equation (3.15). For every word $w = w_{k-1} \cdots w_0 \in A^k$ over an integer alphabet, we have

$$\mathrm{val}_{\beta,\chi}(w) = \mathrm{val}_\beta(w) - w_{k-1}(V_{k-1,\chi} + U_{k-1}). \tag{3.17}$$

Proof of Proposition 3.3.3. Denote $\Sigma = \Sigma_\beta$ and $D = \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0W}^\chi_{\min}\Sigma^* \cup \mathtt{1W}^\chi_{\max}\Sigma^*)$.

(Injectivity): Let $u, v \in D$ be such that $\mathrm{val}_{\beta,\chi}(u) = \mathrm{val}_{\beta,\chi}(v)$. By Lemma 3.3.16, u and v have the same length $\ell \in \mathbb{N}$ and the same first digit $\mathtt{d} \in \{\mathtt{0}, \mathtt{1}\}$. Using Equation (3.17), we have

$$\mathrm{val}_\beta(u) = \mathrm{val}_{\beta,\chi}(u) + \mathtt{d}(V_{\ell-1,\chi} + U_{\ell-1}) = \mathrm{val}_{\beta,\chi}(v) + \mathtt{d}(V_{\ell-1,\chi} + U_{\ell-1}) = \mathrm{val}_\beta(v).$$

Denote $u = \mathtt{d}u'$ and $v = \mathtt{d}v'$. Denote u'', v'' the words of minimal length so that $u' \in \mathtt{0}^*u''$ and $v' \in \mathtt{0}^*v''$. Thus $u'', v'' \notin \mathtt{0}\Sigma^*$ and $\mathrm{val}_\beta(u'') = \mathrm{val}_\beta(v'')$. If $\mathtt{d} = \mathtt{0}$, then $u', v' \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$. If $\mathtt{d} = \mathtt{1}$, then $u', v' \in \mathcal{L}(\mathcal{H}_{\beta,q_\chi})$ and from Lemma 3.3.12, $u', v' \in \mathcal{L}(\mathcal{H}_{\beta,q_1})$. Consequently, in both cases, $u'', v'' \in \mathcal{L}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0}\Sigma^*$ and $\mathrm{val}_\beta(u'') = \mathrm{val}_\beta(v'')$. By Lemma 3.3.11, $u = v$.

(Surjectivity): It holds that $\mathbb{Z} = \bigcup_{\ell=0}^{+\infty} [U_{m(\ell-1)}, U_{m\ell}) \cup [-V_{m\ell,\chi}, -V_{m(\ell-1),\chi})$, a disjoint union, and

$$D = \bigcup_{\ell=0}^{+\infty} \mathtt{0}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0W}^\chi_{\min}\Sigma^* \cup \bigcup_{\ell=0}^{+\infty} \mathtt{1}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{1W}^\chi_{\max}\Sigma^*,$$

a disjoint union. As $\mathrm{val}_{\beta,\chi}$ is an injective map $D \to \mathbb{Z}$, it suffices to show for every $\ell \in \mathbb{N}$, that

- $\mathtt{0}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0W}^\chi_{\min}\Sigma^*$ has the same cardinality as $[U_{m(\ell-1)}, U_{m\ell})$, and
- $\mathtt{1}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{1W}^\chi_{\max}\Sigma^*$ has the same cardinality as $[-V_{m\ell,\chi}, -V_{m(\ell-1),\chi})$.

Indeed,

$$\begin{aligned}
\#\mathtt{0}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{0W}^\chi_{\min}\Sigma^* &= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{W}^\chi_{\min}\Sigma^* \\
&= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) \setminus \mathtt{W}^\chi_{\min}\mathcal{L}_{m(\ell-1)}(\mathcal{H}_{\beta,q_1}) \\
&= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_1}) - \#\mathcal{L}_{m(\ell-1)}(\mathcal{H}_{\beta,q_1}) \\
&= U_{m\ell} - U_{m(\ell-1)}
\end{aligned}$$

and

$$\begin{aligned}
\#\mathtt{1}\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{1W}^\chi_{\max}\Sigma^* &= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{W}^\chi_{\max}\Sigma^* \\
&= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) \setminus \mathtt{W}^\chi_{\max}\mathcal{L}_{m(\ell-1)}(\mathcal{H}_{\beta,q_\chi}) \\
&= \#\mathcal{L}_{m\ell}(\mathcal{H}_{\beta,q_\chi}) - \#\mathcal{L}_{m(\ell-1)}(\mathcal{H}_{\beta,q_\chi}) \\
&= V_{m\ell} - V_{m(\ell-1)}. \quad \square
\end{aligned}$$

3.4 Total order $\prec$

A positional numeration system U for $\mathbb{N}$ has the language $\mathrm{rep}_U(\mathbb{N})$ ordered with respect to the radix order; see Proposition 3.1.1. Equivalently, this fact may be reformulated in the sense that the value map of a positional numeration system is an increasing bijection with respect to the radix order between the language $\mathrm{rep}_U(\mathbb{N})$ and $\mathbb{N}$. In this part, we introduce a total order suitable to characterize positional numeration systems for $\mathbb{Z}$ such as the two's complement notation and its Fibonacci analogue.

First, we define the reversed-radix order as a total order such that $u <_{\mathrm{rev}} v$ if and only if $|u| > |v|$ or $|u| = |v|$ and $u <_{lex} v$. Next, we define a total order on $\{\mathtt{0}, \mathtt{1}\}A^*$ as follows.

Definition 3.4.1 (total order $\prec$). *For every $u, v \in \{\mathtt{0}, \mathtt{1}\}A^*$, we define $u \prec v$ if and only if*

- *$u \in \mathtt{1}A^*$ and $v \in \mathtt{0}A^*$, or*
- *$u, v \in \mathtt{0}A^*$ and $u <_{\mathrm{rad}} v$, or*
- *$u, v \in \mathtt{1}A^*$ and $u <_{\mathrm{rev}} v$.*

For instance, ordering the language $\{\mathtt{0}, \mathtt{1}\}^*$, we get

$$\cdots \prec \mathtt{100} \prec \mathtt{101} \prec \mathtt{110} \prec \mathtt{111} \prec \mathtt{10} \prec \mathtt{11} \prec \mathtt{1} \prec \mathtt{0} \prec \mathtt{00} \prec \mathtt{01} \prec \mathtt{000} \prec \mathtt{001} \prec \mathtt{010} \prec \mathtt{011} \prec \cdots.$$

We prove that the value map of every complement numeration system associated with a simple Parry number is increasing with respect to the order $\prec$.

Proposition 3.4.2. *Let β be a simple Parry number with the quasi-greedy expansion of unity $d_\beta^*(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$. Let $\chi \in [\![1, m]\!]$. Then the map $\mathrm{val}_{\beta,\chi}$ is an increasing bijection:*

$$\mathrm{val}_{\beta,\chi} \colon (\mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}^\chi_{\min}\Sigma^* \cup \mathtt{1}\mathtt{W}^\chi_{\max}\Sigma^*), \prec) \to (\mathbb{Z}, <),$$

where $\Sigma = \Sigma_\beta$, $\mathtt{W}^\chi_{\min} = \mathtt{0}^m$ and $\mathtt{W}^\chi_{\max} = t_\chi \cdots t_{m-1}(t_m - 1)t_1 \cdots t_{\chi-1}$.

Proof. Denote $D = \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}^\chi_{\min}\Sigma^* \cup \mathtt{1}\mathtt{W}^\chi_{\max}\Sigma^*)$. It follows from Proposition 3.3.3 that $\mathrm{val}_{\beta,\chi} \colon D \to \mathbb{Z}$ is a bijection. We prove that it is increasing with respect to the order $\prec$.

Let $u, v \in \mathcal{L}$ be such that $u \prec v$. Let $k, \ell \in \mathbb{N}$ be such that $|u| = mk + 1$ and $|v| = m\ell + 1$.

- If $u \in \mathtt{1}\Sigma^*$ and $v \in \mathtt{0}\Sigma^*$, then by Lemma 3.3.16,
$$\mathrm{val}_{\beta,\chi}(u) < -V_{m(k-1),\chi} < 0 \le U_{m(\ell-1)} \le \mathrm{val}_{\beta,\chi}(v).$$
- Assume that $u, v \in \mathtt{0}\Sigma^*$ and $|u| < |v|$. Then $k \le \ell - 1$ and, by Lemma 3.3.16,
$$\mathrm{val}_{\beta,\chi}(u) < U_{mk} \le U_{m(\ell-1)} \le \mathrm{val}_{\beta,\chi}(v).$$
- Assume that $u, v \in \mathtt{1}\Sigma^*$ and $|u| > |v|$. Then $k - 1 \ge \ell$ and, by Lemma 3.3.16,
$$\mathrm{val}_{\beta,\chi}(u) < -V_{m(k-1),\chi} \le -V_{m\ell,\chi} \le \mathrm{val}_{\beta,\chi}(v).$$
- Assume that $u, v \in \mathtt{d}\Sigma^*$ for some $\mathtt{d} \in \{\mathtt{0}, \mathtt{1}\}$ and $|u| = |v|$. In this case, we have $k = \ell$ and $u <_{lex} v$. Thus $u <_{\mathrm{rad}} v$ and from Lemma 3.3.11, $\mathrm{val}_\beta(u) < \mathrm{val}_\beta(v)$. From Equation (3.17),
$$\mathrm{val}_{\beta,\chi}(u) = \mathrm{val}_\beta(u) - \mathtt{d}(V_{k-1,\chi} + U_{k-1}) < \mathrm{val}_\beta(v) - \mathtt{d}(V_{k-1,\chi} + U_{k-1}) = \mathrm{val}_{\beta,\chi}(u). \quad \square$$

We prove that the map $\mathrm{rep}_{\beta,\chi}$ is characterized by the fact of being an increasing bijection.

Proposition 3.4.3. *Let β be a simple Parry number with $d_\beta^*(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$. Let $\chi \in [\![1, m]\!]$. Let $f : \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}\Sigma^*$ be a map. The following statements are equivalent:*

- *$f = \mathrm{rep}_{\beta,\chi}$,*
- *f is increasing with respect to the order $\prec$, its image is $f(\mathbb{Z}) = \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0}\mathtt{W}^\chi_{\min}\Sigma^* \cup \mathtt{1}\mathtt{W}^\chi_{\max}\Sigma^*)$ and $f(0) = \mathtt{0}$.*

Proof. Suppose that $f = \mathrm{rep}_{\beta,\chi}$. As f is the inverse map $\mathrm{val}_{\beta,\chi}^{-1}$, from Proposition 3.4.2 it is increasing and its image is $\mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0W}_{\min}^{\chi}\Sigma^* \cup \mathtt{1W}_{\max}^{\chi}\Sigma^*)$. Also, $\mathrm{rep}_{\beta,\chi}(0) = \mathtt{0}$ from Remark 3.3.5.

Let $f : \mathbb{Z} \to \{\mathtt{0}, \mathtt{1}\}\Sigma^*$. Suppose f is increasing, its image is $\mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0W}_{\min}^{\chi}\Sigma^* \cup \mathtt{1W}_{\max}^{\chi}\Sigma^*)$ and $f(0) = \mathtt{0}$. The map $\mathrm{rep}_{\beta,\chi}$ satisfies the same properties. Since there is a unique increasing bijection $\mathbb{Z} \to f(\mathbb{Z})$ such that $f(0) = \mathtt{0}$, we conclude that $f = \mathrm{rep}_{\beta,\chi}$. □

As a corollary, the value maps val_{2c} and $\mathrm{val}_{\mathcal{F}c}$ are increasing bijective functions with respect to the total order $\prec$ on the corresponding languages of representations; see Example 3.3.6 and Example 3.3.7. Hence, Proposition 3.4.3 provides a unified proof of results proved in [LL23a].

Corollary 3.4.4. [LL23a] *The map* rep_{2c} *is the unique increasing bijection*

$$\mathrm{rep}_{2c} \colon (\mathbb{Z}, <) \to \left(\Sigma^+ \setminus (\mathtt{11}\Sigma^* \cup \mathtt{00}\Sigma^*), \prec\right)$$

such that $0 \mapsto \mathtt{0}$.

Corollary 3.4.5. [LL23a] *The map* $\mathrm{rep}_{\mathcal{F}c}$ *is the unique increasing bijection*

$$\mathrm{rep}_{\mathcal{F}c} \colon (\mathbb{Z}, <) \to \left(\Sigma(\Sigma\Sigma)^* \setminus (\Sigma^*\mathtt{11}\Sigma^* \cup \mathtt{000}\Sigma^* \cup \mathtt{101}\Sigma^*), \prec\right)$$

such that $0 \mapsto \mathtt{0}$.

The numeration system $\mathcal{G}$ seems to be increasing with respect to the order $\prec$ as well, even though it is not a complement numeration system associated with a simple Parry number; see Table 3.1 and Example 3.3.8. On the other hand, the value map of the negaFibonacci numeration system is not increasing with respect to the order $\prec$. In fact, it does not even preserve the radix order on $\mathbb{N}$; see Table 3.1. For instance, we have

$$\mathrm{rep}_{\mathrm{neg}\mathcal{F}}(5) = \mathtt{10000} <_{\mathrm{rad}} = \mathtt{10010} = \mathrm{rep}_{\mathrm{neg}\mathcal{F}}(4).$$

It seems convenient that the order $\prec$ extends the radix order, which characterizes the positional numeration systems for $\mathbb{N}$. In this light, the numeration systems such as the two's complement notation and the Fibonacci complement numeration system seem to be natural extensions of the classical numeration systems to $\mathbb{Z}$.

3.5 Addition in the Fibonacci complement numeration system

Integers represented in the two's complement notation [Knu69, §4.1] are added using the same algorithm, which applies for the addition of nonnegative integers represented in the binary numeration system; see Example 3.5.1. Due to this property, the two's complement notation is nowadays used in computers to represent signed (meaning both nonnegative and negative) integers[1] The Fibonacci complement numeration system is an analogue of the two's complement notation, using the Fibonacci numbers instead of powers of 2. In this part, we show that these two numeration systems share the same properties with respect to addition. More precisely, the addition of integers represented in the Fibonacci complement numeration system is performed with (almost) the same algorithm as the addition of nonnegative integers represented in the Fibonacci numeration system.

[1] "All operations on signed integers assume a two's complement representation." Section 4.2.1 in Intel 64 and IA-32 Architectures Software Developer's Manual, Volume 1: Basic Architecture, retrieved September 2023

Example 3.5.1. We illustrate the addition in the binary and in the two's complement numeration system. We add two representations `01011` and `10001` with the classical carry method, starting from the right and transferring a carry at each step. The resulting representation is `11100`, which is not greedy.

$$\begin{array}{lcl} 11 & \texttt{01011} & +11 \\ +17 & \texttt{10001} & -15 \\ \hline +28 & \texttt{11100} & -4 \end{array}$$

If we interpret all three binary words as the binary representations, we obtain

$$\mathrm{val}_2(\texttt{01011})+\mathrm{val}_2(\texttt{10001}) = (2^3+2^1+2^0)+(2^4+2^0) = 11+17 = 28 = 2^4+2^3+2^2 = \mathrm{val}_2(\texttt{11100}).$$

But if we interpret all three binary words as the two's complement representations, we obtain

$$\mathrm{val}_{2c}(\texttt{01011})+\mathrm{val}_{2c}(\texttt{10001}) = (2^3+2^1+2^0)+(-2^4+2^0) = 11+(-15) = -4 = -2^4+2^3+2^2 = \mathrm{val}_2(\texttt{11100}).$$

3.5.1 Addition in the Fibonacci numeration system

Addition of integers represented in the binary numeration system can be performed using a pencil-and-paper algorithm shown in Example 3.5.1. Addition of integers represented in the Fibonacci numeration system is less elementary as the traditional carry method clearly does not work. In [Ber86, p. 22], Berstel introduced a Mealy machine $\mathcal{B}$ called *adder*, which can perform addition of the Fibonacci representations.

A *Mealy machine* M is a labeled directed graph. Its vertices are called *states* and its edges are called *transitions* [Lin12, Appendix A]. Let A and B be finite alphabets. The transitions are labeled by pairs a/b of letters $a \in A$, $b \in B$. The first letter $a \in A$ is called the *input symbol*, the second letter $b \in B$ is called the *output symbol* and A (resp. B) is called the *input* (resp. *output*) *alphabet*. The empty word ε is sometimes included in the output alphabet B. For every state s and every letter a, there is at most one transition starting from the state s with the input symbol a. One distinguished state is called the *initial state*.

A machine M computes a function $M : A^* \to B^*$. Let $x = x_0x_1 \cdots x_{k-1} \in A^*$ and $y = y_0y_1 \cdots y_{k-1} \in B^*$ be two words of length $k \in \mathbb{N}$ over the input and output alphabet. The word y is the output of x under the machine M if and only if there is a sequence $\{s_i\}_{0\leq i\leq k}$ of states of M such that s_0 is the initial state and for every i with $0 \leq i < k$, there is a transition from s_i to s_{i+1} labeled by x_i/y_i. The output word y is denoted by $M(x)$. The last state s_k is denoted by $M_{\mathrm{last}}(x)$ and an extra output word depending on the last state is denoted by $M_{\downarrow}(x)$.

The Berstel adder is the Mealy machine $\mathcal{B} = (Q, \texttt{000.0}, \{\texttt{0},\texttt{1},\texttt{2}\}, \{\texttt{0},\texttt{1}\}, \delta_{\mathcal{B}}, \eta_{\mathcal{B}}, \phi_{\mathcal{B}})$ with the set of 10 states

$$Q = \left\{ \begin{array}{ccccc} \texttt{000.0}, & \texttt{001.1}, & \texttt{010.3}, & \texttt{100.5}, & \texttt{101.6}, \\ \texttt{000.1}, & \texttt{001.2}, & \texttt{010.4}, & \texttt{100.6}, & \texttt{101.7} \end{array} \right\}$$

with the initial state `000.0`, the input alphabet $\{\texttt{0},\texttt{1},\texttt{2}\}$, the output alphabet $\{\texttt{0},\texttt{1}\}$, the transition function $\delta_{\mathcal{B}}$ and the output function $\eta_{\mathcal{B}}$ as shown in Figure 3.3. The states of the Berstel adder is a subset of $S \times \{0, 1, \ldots, 7\}$, where $S = \{\texttt{000}, \texttt{001}, \texttt{010}, \texttt{100}, \texttt{101}\}$. The function $\phi_{\mathcal{B}}$ is the canonical projection $S \times \{0, 1, \ldots, 7\} \to S$. Note that our notation of the 10 states is not the same as the 10 states provided by Berstel, however, it is equivalent.

Reading an input word $u \in \{\texttt{0},\texttt{1},\texttt{2}\}^*$ from left to right, the Berstel adder $\mathcal{B}$ produces an output word $\mathcal{B}(u) \in \{\texttt{0},\texttt{1}\}^*$ concatenated with a three-letter word $\mathcal{B}_{\downarrow}(u) \in \{\texttt{000}, \texttt{001}, \texttt{010}, \texttt{100}, \texttt{101}\}$,

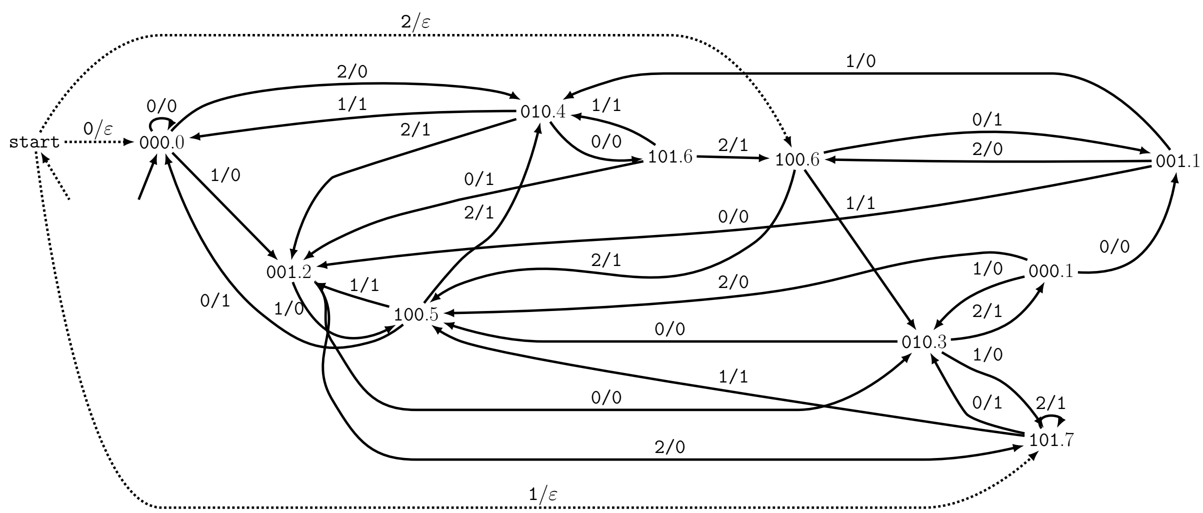


Figure 3.3: The Berstel adder $\mathcal{B}$ is a 10-state Mealy machine with 30 transitions illustrated as solid edges with the initial state 000.0. The Mealy machine $\mathcal{T}$ is obtained by adding a new state `start` that replaces 000.0 as initial state and adding three additional transitions shown with dashed edges.

which depends on the final state reached after reading u. We illustrate this in Example 3.5.3. The resulting concatenation $\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u)$ is a binary word which has the same Fibonacci value as the input word u:

$$\mathrm{val}_{\mathcal{F}}(u) = \mathrm{val}_{\mathcal{F}}(\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u)).$$

We reformulate this result in Theorem 3.5.2.

Theorem 3.5.2. [Ber86] *The Berstel adder $\mathcal{B}$ has the property that for every input $u \in \{0,1,2\}^*$, it outputs a word $\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u) \in \{0,1\}^*$ with same value in the Fibonacci numeration system:*

$$\mathrm{val}_{\mathcal{F}}(u) = \mathrm{val}_{\mathcal{F}}(\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u)).$$

Example 3.5.3. Feeding the Berstel adder $\mathcal{B}$ in Figure 3.3 with the word $u = 2220121$ gives

$$000.0 \xrightarrow{2/0} 010.4 \xrightarrow{2/1} 001.2 \xrightarrow{2/0} 101.7 \xrightarrow{0/1} 010.3 \xrightarrow{1/0} 101.7 \xrightarrow{2/1} 101.7 \xrightarrow{1/1} 100.5;$$

therefore the last state is $\mathcal{B}_{\text{last}}(u) = 100.5$ and we obtain $\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u) = 0101011\ \cdot 100$. Thus

$$\mathrm{val}_{\mathcal{F}}(2220121) = 42 + 26 + 16 + 3 + 4 + 1 = 92 = 55 + 21 + 8 + 5 + 3 = \mathrm{val}_{\mathcal{F}}(0101011100).$$

Consequently, we can perform addition of integers represented in the Fibonacci numeration system by summing the Fibonacci representations digit by digit and feeding the resulting word over the alphabet $\{0,1,2\}$ to the Berstel adder $\mathcal{B}$. The word produced by the Berstel adder $\mathcal{B}$ is a binary word with the correct Fibonacci value. However, it might not be its greedy representation. Indeed, it is known [Ber86, Sak87, Fro91] that no single right-to-left and no single left-to-right transducer can normalize all words $u \in \{0,1,2\}^*$. We illustrate this in Example 3.5.4.

Example 3.5.4. We use the Berstel adder to compute the sum $33 + 25$. First, we express 33 and 25 by their Fibonacci representation (in general, if representations do not have the same length, the shorter one is padded with leading $\mathtt{0}$'s). Then, we add them digit by digit to obtain a ternary word in $\{\mathtt{0},\mathtt{1},\mathtt{2}\}^*$:

$$\begin{array}{rr} 33 & \mathtt{1010101} \\ +25 & \mathtt{1000101} \\ \hline 58 & \mathtt{2010202} \end{array}$$

Reading from left to right and giving the word $u = \mathtt{2010202}$ as input to the Berstel adder (see Figure 3.3), we obtain the following path from the initial state $\mathtt{000.0}$:

$$\mathtt{000.0} \xrightarrow{\mathtt{2/0}} \mathtt{010.4} \xrightarrow{\mathtt{0/0}} \mathtt{101.6} \xrightarrow{\mathtt{1/1}} \mathtt{010.4} \xrightarrow{\mathtt{0/0}} \mathtt{101.6} \xrightarrow{\mathtt{2/1}} \mathtt{100.6} \xrightarrow{\mathtt{0/1}} \mathtt{001.1} \xrightarrow{\mathtt{2/0}} \mathtt{100.6}.$$

Therefore, the output word is $\mathcal{B}(u) = \mathtt{0010110}$ and the path ends in state $\mathcal{B}_{\text{last}}(u) = \mathtt{100.6}$. Removing the last digit (6) of the state $\mathcal{B}_{\text{last}}(u)$, which we can ignore for now, we obtain the three-letter extra output word $\mathcal{B}_{\downarrow}(u) = \mathtt{100}$. Concatenating the two words gives $\mathtt{0010110} \cdot \mathtt{100}$, which has the correct Fibonacci value 33+25=58

$$\text{val}_{\mathcal{F}}\left(\mathcal{B}(u) \cdot \mathcal{B}_{\downarrow}(u)\right) = \text{val}_{\mathcal{F}}\left(\mathtt{0010110} \cdot \mathtt{100}\right) = 3 + 8 + 13 + 34 = 58.$$

However, it is not the greedy representation as it contains leading $\mathtt{0}$s and consecutive $\mathtt{1}$s.

3.5.2 Extension to $\mathbb{Z}$

In Section 3.5.1, we recalled an algorithm for addition of nonnegative integers represented in the Fibonacci numeration system with the use of Berstel adder $\mathcal{B}$. A simple counterexample suffices to see that it is not possible to use the Berstel adder in its original form to add integers represented in the Fibonacci complement numeration system; see Example 3.5.5.

Example 3.5.5. We compute the sum $-1+(-1)$ using the Berstel adder $\mathcal{B}$. The representation of -1 in the Fibonacci complement numeration system is $\text{rep}_{\mathcal{F}c}(-1) = \mathtt{1}$ and the digit-by-digit addition gives $\text{rep}_{\mathcal{F}c}(-1) + \text{rep}_{\mathcal{F}c}(-1) = \mathtt{2}$. The Berstel adder follows the path

$$\mathtt{000.0} \xrightarrow{\mathtt{2/0}} \mathtt{010.4}$$

and therefore the output word is $\mathcal{B}(\mathtt{2}) = \mathtt{0}$, the path ends in state $\mathcal{B}_{\text{last}}(\mathtt{2}) = \mathtt{010.4}$ and the extra output word is $\mathcal{B}_{\downarrow}(\mathtt{2}) = \mathtt{010}$. Concatenating the two words gives $\mathtt{0} \cdot \mathtt{010}$, which *does not* have the correct Fibonacci complement value

$$\text{val}_{\mathcal{F}c}\left(\mathtt{0} \cdot \mathtt{010}\right) = 2 \neq -2 = -1 + (-1).$$

Nevertheless, it is possible to adapt the Berstel adder $\mathcal{B}$ to perform addition of integers represented in the Fibonacci complement numeration system. We do so by adding a new initial state $\mathtt{start}$ and three additional transitions

$$\mathtt{start} \xrightarrow{\mathtt{0}/\varepsilon} \mathtt{000.0}, \quad \mathtt{start} \xrightarrow{\mathtt{1}/\varepsilon} \mathtt{101.7}, \quad \mathtt{start} \xrightarrow{\mathtt{2}/\varepsilon} \mathtt{100.6},$$

creating a new Mealy machine which we denote $\mathcal{T}$; see Figure 3.3. The following theorem is an analogue of Theorem 3.5.2 for the Fibonacci complement numeration system.

Theorem 3.5.6. [LL23a] *The Mealy machine $\mathcal{T}$ has the property that for every nonempty input $u \in \{\mathtt{0},\mathtt{1},\mathtt{2}\}^+$, it outputs a word $\mathcal{T}(u) \cdot \mathcal{T}_\downarrow(u) \in \{\mathtt{0},\mathtt{1}\}^+$ with the same value in the Fibonacci complement numeration system:*

$$\mathrm{val}_{\mathcal{F}c}(u) = \mathrm{val}_{\mathcal{F}c}(\mathcal{T}(u) \cdot \mathcal{T}_\downarrow(u)).$$

In analogy with Example 3.5.3, we illustrate Theorem 3.5.6 in the following two examples.

Example 3.5.7. We feed the modified Berstel adder $\mathcal{T}$ in Figure 3.3 with the same word $u = \mathtt{2220121}$ as in Example 3.5.3, this time obtaining

$$\mathtt{start} \xrightarrow{2/\varepsilon} 100.6 \xrightarrow{2/1} 100.5 \xrightarrow{2/1} 010.4 \xrightarrow{0/0} 101.6 \xrightarrow{1/1} 010.4 \xrightarrow{2/1} 001.2 \xrightarrow{1/0} 100.5;$$

therefore the last state is $\mathcal{T}_{\mathrm{last}}(u) = 100.5$ and we obtain $\mathcal{T}(u) \cdot \mathcal{T}_\downarrow(u) = \mathtt{110110} \cdot \mathtt{100}$. Interpreting the results in the Fibonacci complement numeration system, we observe

$$\mathrm{val}_{\mathcal{F}c}(\mathtt{2220121}) = (-42)+42+16+3+4+1 = 24 = (-34)+34+13+8+3 = \mathrm{val}_{\mathcal{F}c}(\mathtt{110110100}).$$

Example 3.5.8. We use the Mealy machine $\mathcal{T}$ to compute the sum $-1+(-9)$. First, we express (-1) and (-9) by their Fibonacci complement representation and we pad the shorter word with an appropriate neutral prefix ($\mathtt{00}$'s if it starts with $\mathtt{0}$ or $\mathtt{10}$'s if it starts with $\mathtt{1}$, see Definition 3.2.4) so that they have the same length. Then, we add them digit by digit to obtain a ternary word in $\{\mathtt{0},\mathtt{1},\mathtt{2}\}^*$:

$$\begin{array}{rr} -1 & \mathtt{1010101} \\ +(-9) & \mathtt{1000101} \\ \hline -10 & \mathtt{2010202} \end{array}$$

Note that the resulting word $u = \mathtt{2010202}$ coincides with the one in Example 3.5.4, where we show properties of addition in Fibonacci numeration system. Reading from left to right and giving the word $u = \mathtt{2010202}$ as input to the Mealy machine $\mathcal{T}$ (see Figure 3.3), we obtain the following path from the initial state $\mathtt{start}$:

$$\mathtt{start} \xrightarrow{2/\varepsilon} 100.6 \xrightarrow{0/1} 001.1 \xrightarrow{1/0} 010.4 \xrightarrow{0/0} 101.6 \xrightarrow{2/1} 100.6 \xrightarrow{0/1} 001.1 \xrightarrow{2/0} 100.6.$$

Therefore, the output word is $\mathcal{T}(u) = \mathtt{100110}$ and the path ends in state $\mathcal{T}_{\mathrm{last}}(u) = 100.6$. Removing the last digit (6) of the last state, we obtain the three-letter extra output word $\mathcal{T}_\downarrow(u) = \mathtt{100}$. Concatenating the two words gives $\mathtt{100110} \cdot \mathtt{100}$, which is a Fibonacci complement representation of the sum $-1+(-9)$. We confirm that its Fibonacci complement value is correct:

$$\mathrm{val}_{\mathcal{F}c}(\mathcal{T}(u) \cdot \mathcal{T}_\downarrow(u)) = \mathrm{val}_{\mathcal{F}c}(\mathtt{100110} \cdot \mathtt{100}) = 3+8+13-34 = -10.$$

It would be interesting to construct transducers to perform addition in a similar manner for all complement numeration systems associated with simple Parry numbers, which we defined in Section 3.3. For now, we leave this as an open question.

Question 3.5.9. *Can we construct a Mealy machine, which performs addition in a given complement numeration system associated with a simple Parry number?*

Chapter 4

Critical exponents of Arnoux–Rauzy words

In this chapter, we summarize properties of Sturmian words [Lot02, §2] and Arnoux–Rauzy words, which are a generalization of Sturmian words to larger alphabets [AR91]. We present our work concerning the repetitions of factors in Arnoux–Rauzy words [DL23], which includes a formula for the critical exponents of regular Arnoux–Rauzy words and a proof that the minimal critical exponents among regular Arnoux–Rauzy words are attained by the d-bonacci words. It is known that the critical and asymptotic critical exponent of a d-bonacci word coincide for $d = 2$ and $d = 3$. We present a new result stating that, for $d \in [\![4, 15]\!]$, the critical and asymptotic critical exponent of a d-bonacci word coincide as well. During the course of finalizing this text, a significantly stronger result was proved – for every $d \geq 2$, the critical and asymptotic critical exponent of a d-bonacci word coincide and, moreover, the minimal critical exponent among all d-ary episturmian words is attained by the d-bonacci word [DP23].

4.1 Introduction to Sturmian and Arnoux–Rauzy words

4.1.1 Sturmian words

Sturmian words belong to the core subjects of combinatorics on words. Named after the French mathematician Charles François Sturm (1803-1855), the term Sturmian words (trajectories) first appeared in the work of Morse and Hedlund [MH40] in 1940. Sturmian words can be defined in multiple ways. Morse and Hedlund proposed a definition connected to the solutions of second order differential equations. However, for our purposes, a definition using the factor complexity function is more appropriate. A Sturmian word is an infinite word $\mathbf{u}$ over the binary alphabet with the factor complexity function fulfilling the condition that $\mathcal{C}_{\mathbf{u}}(n) = n + 1$, for every $n \in \mathbb{N}$. In other words, a Sturmian word contains exactly $n + 1$ distinct factors of every length n. The Fibonacci word belongs to the well-known examples of Sturmian words.

Example 4.1.1. The Fibonacci word may be defined in various ways, but for now we choose the one closest to the notion of Fibonacci numbers. Let $(f_n)_{n=0}^{+\infty}$ be the sequence of finite words defined recurrently by $f_{n+2} = f_{n+1}f_n$, for every $n \in \mathbb{N}$, with the initial conditions $f_0 = 0$ and $f_1 = 01$. Then

$$\mathbf{f} = \lim_{n \to +\infty} f_n = 0100101001001010010100100101001001\ldots$$

defines the Fibonacci word as the infinite word, which has f_n as a prefix, for every $n \in \mathbb{N}$.

Recall that a binary word $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ is balanced if for every pair of words $x, y \in \mathcal{L}(\mathbf{u})$ such that $|x| = |y|$ it holds that $||x|_1 - |y|_1| \leq 1$, i.e., every pair of factors x, y occurring in $\mathbf{u}$ of the same length fulfills that x and y have the number of occurrences of the symbol 1 differing by at most one. Balanced words have well-defined slopes – the slope of a balanced word is the limit $\lim_{n\to+\infty} \frac{|u^{(n)}|_1}{n}$, where $u^{(n)}$ denotes the prefix of $\mathbf{u}$ of length n, for every $n \in \mathbb{N}$. The slope of a balanced infinite word $\mathbf{u}$ is denoted $\pi(\mathbf{u})$. The Fibonacci word is a balanced aperiodic word and we compute its slope in Example 4.1.2; see also [Lot02, Example 2.1.9].

Example 4.1.2. We compute the slope of the Fibonacci word $\mathbf{f}$. It suffices to compute the limit $\lim_{n\to+\infty} \frac{|f_n|_1}{|f_n|}$. For every $n \in \mathbb{N}$, we have $|f_n| = F_n$ and $|f_n|_1 = F_{n-2}$. Therefore $\pi(\mathbf{f}) = \lim_{n\to+\infty} \frac{|f_n|_1}{|f_n|} = \lim_{n\to+\infty} \frac{F_{n-2}}{F_n} = \frac{1}{\tau^2}$, where τ denotes the golden mean.

Sturmian words coincide with the set of binary balanced aperiodic words. It follows from [Lot02, Theorem 1.3.13] that an infinite word $\mathbf{u}$ over an alphabet A is aperiodic if and only if its factor complexity function fulfills the condition that $\mathcal{C}_{\mathbf{u}}(n) \geq n + \#A - 1$. In this sense, Sturmian words are aperiodic words with the minimal factor complexity. As, according to [Lot02, Proposition 2.1.2], the factor complexity function of a balanced word $\mathbf{v}$ satisfies the condition that $\mathcal{C}_{\mathbf{v}}(n) \leq n + 1$, Sturmian words are balanced words with the maximal factor complexity. Also, the language of a Sturmian word is closed under reversal [Lot02, Proposition 2.1.19].

Sturmian words coincide with mechanical words with irrational slopes. We recall the definition of a lower and upper mechanical word. Let $\alpha \in [0,1]$ and $\delta \in \mathbb{R}$ be real numbers. The word $\mathbf{s}_{\alpha,\delta} : \mathbb{N} \to \{0,1\}$ and the word $\mathbf{s}'_{\alpha,\delta} : \mathbb{N} \to \{0,1\}$, defined for every $n \in \mathbb{N}$ by the relation

$$\mathbf{s}_{\alpha,\delta}(n) := \lfloor \alpha(n+1) + \delta \rfloor - \lfloor \alpha n + \delta \rfloor \quad \text{and} \quad \mathbf{s}'_{\alpha,\delta}(n) := \lceil \alpha(n+1) + \delta \rceil - \lceil \alpha n + \delta \rceil,$$

is called the *lower* and the *upper*, respectively, *mechanical word* with the *slope* α and the *intercept* δ. It is easy to see that the words $\mathbf{s}_{\alpha,\delta}$ and $\mathbf{s}'_{\alpha,\delta}$ are indeed binary. Also, the slope α of an upper (lower) mechanical word $\mathbf{u}$ coincides with the slope $\pi(\mathbf{u})$ [Lot02], and thus

$$\alpha = \lim_{n\to+\infty} \frac{|u^{(n)}|_1}{n}, \tag{4.1}$$

where $u^{(n)}$ denotes the prefix of $\mathbf{u}$ of length n, for every $n \in \mathbb{N}$.

Mechanical words with the irrational slope α and the intercept $\delta = 0$ fulfill the condition that $\mathbf{s}_{\alpha,0}(0) = 0$, $\mathbf{s}'_{\alpha,0}(0) = 1$, and $\mathbf{s}_{\alpha,0}(n) = \mathbf{s}'_{\alpha,0}(n)$, for every $n \geq 1$. The word $\mathbf{c}_\alpha$ such that $\mathbf{s}_{\alpha,0} = 0\mathbf{c}_\alpha$ and $\mathbf{s}'_{\alpha,0} = 1\mathbf{c}_\alpha$ is called the *characteristic word* of the slope α. Equivalently, $\mathbf{c}_\alpha = \mathbf{s}_{\alpha,\alpha} = \mathbf{s}'_{\alpha,\alpha}$. We interpret the Fibonacci word $\mathbf{f}$ as a mechanical word in Example 4.1.3. More precisely, the Fibonacci word $\mathbf{f}$ is a characteristic Sturmian word.

Example 4.1.3. The Fibonacci word $\mathbf{f}$ interpreted as a mechanical word is shown in Figure 4.1. As this word is the characteristic word of slope $\frac{1}{\tau^2}$, it is equal to both the lower and the upper mechanical word with the slope $\alpha = \frac{1}{\tau^2}$ and the intercept $\delta = \frac{1}{\tau^2}$:

$$\mathbf{f} = \mathbf{s}_{\frac{1}{\tau^2},\frac{1}{\tau^2}} = \mathbf{s}'_{\frac{1}{\tau^2},\frac{1}{\tau^2}}, \tag{4.2}$$

where $\tau = \frac{1+\sqrt{5}}{2}$ denotes the golden mean.

Mechanical words can be interpreted also as the so-called cutting sequences. Again, we consider a straight line $y = \beta x + \rho$ for some $\beta > 0$, which is not restricted from above this time,

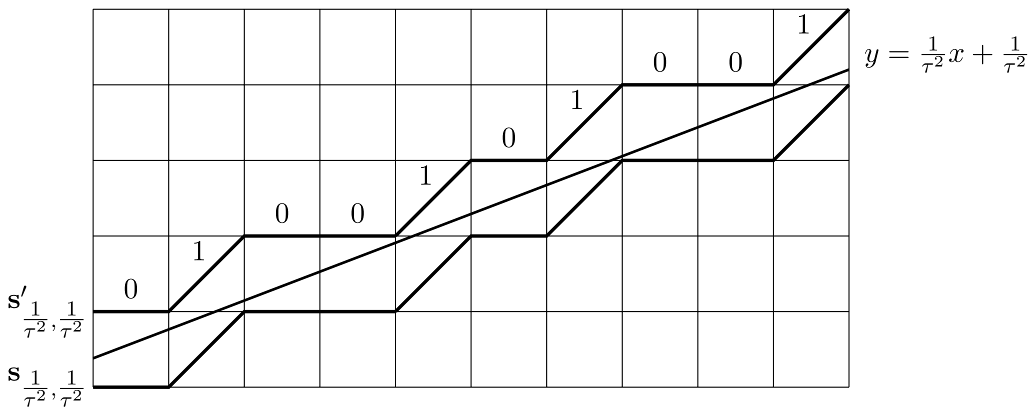


Figure 4.1: Lower and upper mechanical words associated with the line $y = \frac{1}{\tau^2}x + \frac{1}{\tau^2}$, where $\tau = \frac{1+\sqrt{5}}{2}$ denotes the golden mean.

and some $\rho \in \mathbb{R}$ not restricted to be positive. We consider the intersections of this line with a grid with integer coordinates. If the line $y = \beta x + \rho$ intersects a vertical line of the grid, we call such an intersection *vertical*. Otherwise the intersection is called *horizontal*. If an intersection is vertical and horizontal simultaneously, we consider this intersection as two intersections, horizontal and vertical, in this particular order. Writing a letter 0 for every vertical intersection and a letter 1 for every horizontal intersection, we obtain an infinite binary word $K_{\beta,\rho}$, which is called the lower cutting sequence. We obtain the upper cutting sequence $K'_{\beta,\rho}$ if we choose the opposite order in treating the intersections that are both horizontal and vertical. It can be shown that a lower cutting sequence coincides with a lower mechanical sequence with the slope and intercept transformed in the following way

$$K_{\beta,\rho} = \mathbf{s}_{\beta/(1+\beta),\rho/(1+\beta)}.$$

Thus, there are two different notions of slope. The slope of a mechanical word expresses the frequency of the letter 1 in the word, see Equation (4.1), whereas the slope of a cutting sequence $\mathbf{u}$ is the ratio of the frequencies of letters 1 and 0:

$$\beta = \lim_{n\to+\infty} \tfrac{|u^{(n)}|_1}{|u^{(n)}|_0}, \tag{4.3}$$

where $u^{(n)}$ denotes the prefix of $\mathbf{u}$ of length n, for every $n \in \mathbb{N}$. Further in the text, we always specify whether the slope is in the sense of a mechanical word or a cutting sequence. We interpret the Fibonacci word as a cutting sequence in Example 4.1.4. In the graphical illustration, we mark vertical intersections with white points and horizontal intersections with black points; see Figure 4.2.

Example 4.1.4. The Fibonacci word $\mathbf{f}$ interpreted as a cutting sequence is shown in Figure 4.2. The slope of the Fibonacci word as a cutting sequence is $\frac{1}{\tau}$.

We return to the concept of mechanical words. Denote $\{\alpha n+\delta\}$ the fractional part of $\alpha n+\delta$. We observe that $\mathbf{s}_{\alpha,\delta}(n) = 0$ if and only if $\{\alpha n + \delta\} \in [0, 1-\alpha)$ and $\mathbf{s}'_{\alpha,\delta}(n) = 0$ if and only if $\{\alpha n + \delta\} \in (0, 1-\alpha]$. This leads to an equivalent definition of mechanical words by rotations on a circle. The *rotation* of angle α is the map $R_\alpha : [0,1) \to [0,1)$ defined by $R_\alpha(x) = \{x + \alpha\}$.

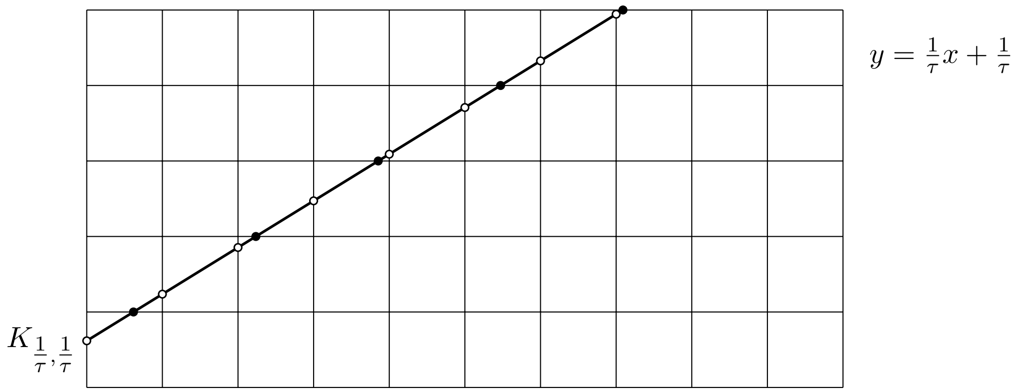


Figure 4.2: Lower cutting sequence $K_{\frac{1}{\tau},\frac{1}{\tau}} = 0100101001001\cdots$ associated with the line $y = \frac{1}{\tau}x + \frac{1}{\tau}$, where $\tau = \frac{1+\sqrt{5}}{2}$ denotes the golden mean.

Consequently, $R_\alpha^n(x) = \{n\alpha + x\}$ for every $n \in \mathbb{N}$. Note that, in this setup, the points 0 and 1 are identified. Defining a partition of the interval $[0, 1)$, we get an equivalent definition of mechanical words

$$\mathbf{s}_{\alpha,\delta}(n) = \begin{cases} 0, & \text{if } R_\alpha^n(\delta) \in [0, 1-\alpha); \\ 1, & \text{if } R_\alpha^n(\delta) \in [1-\alpha, 1); \end{cases} \quad \text{and} \quad \mathbf{s}'_{\alpha,\delta}(n) = \begin{cases} 0, & \text{if } R_\alpha^n(\delta) \in (0, 1-\alpha]; \\ 1, & \text{if } R_\alpha^n(\delta) \in (1-\alpha, 1]. \end{cases} \tag{4.4}$$

Note that the intervals are left-closed right-open for the lower and left-open right-closed for the upper mechanical word. We interpret the Fibonacci word as a coding of a rotation in Example 4.1.5.

Example 4.1.5. Interpreting the Fibonacci word as a coding of a rotation on a circle is shown in Figure 4.3. The partition of the circle (1-dimensional torus) is $I_0 = [0, 1-\alpha)$ and $I_1 = [1-\alpha, 1)$. Starting at the initial point $\delta = \frac{1}{\tau^2}$ and rotating by the angle $\alpha = \frac{1}{\tau^2}$ up to 7 times gives the iterations $R_\alpha^n(\delta)$ for $n \in [\![0, 7]\!]$. The iterations are coded as 01001010, which is the prefix of the Fibonacci word $\mathbf{f}$. In particular, it is easy to see that the factor 11 does not occur in $\mathbf{f}$.

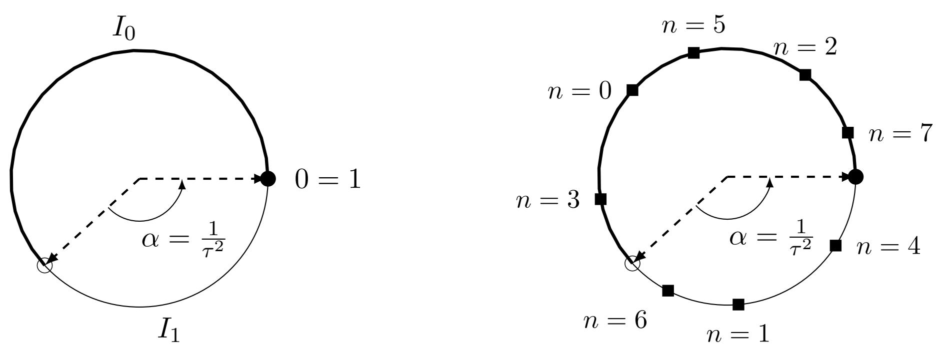


Figure 4.3: Rotation on a 1-dimensional torus with identified points 0 and 1 by the angle $\alpha = \frac{1}{\tau^2}$, where τ denotes the golden mean. The n-th iteration of the initial point δ is coded as 0 if and only if $R_\alpha^n(\delta) \in [0, 1-\alpha)$. On the right, observe the iterations $R_\alpha^n(\delta)$ of the initial point $\delta = \frac{1}{\tau^2}$ for $n \in [\![0, 7]\!]$.

Sturmian words can also be defined using the 2 interval exchange transformation, which generalizes the rotation on a circle. We discuss this approach in more detail in Section 5.4.

4.1.2 Arnoux–Rauzy words

As shown in the previous section, Sturmian words possess various interesting properties which inspired researchers to try and generalize them to larger alphabets. Rauzy [Rau82] introduced a word on the ternary alphabet, which can be interpreted as a coding of a rotation on a 2-dimensional torus. This word is called the Rauzy word, or, as it shares plenty of properties with the Fibonacci word and it is ternary, the Tribonacci word. Later on, Arnoux and Rauzy introduced certain words of the factor complexity $2n+1$ [AR91] as an attempt to generalize Sturmian words to the alphabet of size 3. Their definition led to a geometrical interpretation of these words as codings of 6 interval exchange transformation on a circle. The construction described in [AR91] may be extended even further. Every d-ary word with the factor complexity $(d-1)n+1$ having language closed under reversal and exactly one right special factor of every length prolongable by all letters $a \in A$, may be represented as a coding of $2d$ interval exchange transformation on a circle. Such words are classically called d-ary Arnoux–Rauzy words. Later in this text, we define the d-ary Arnoux–Rauzy words in a different manner, which is equivalent.

Another generalization of Sturmian words to d-ary alphabets for any integer $d \geq 2$ was done in [DJP01], where the words were called episturmian. Episturmian words are such that their language is closed under reversal and they have at most one right special factor of every length. Note that, in particular, episturmian words over an alphabet A include trivial examples such as a^ω for every letter $a \in A$. The d-ary Arnoux–Rauzy words form a subclass of the episturmian words, which are called strict episturmian.

A word $\mathbf{u}$ over the alphabet A of cardinality d is *Arnoux–Rauzy* if the language $\mathcal{L}(\mathbf{u})$ is closed under reversal and if there exists exactly one right special factor w of every length such that $wa \in \mathcal{L}(\mathbf{u})$ for every letter $a \in A$. An Arnoux–Rauzy word $\mathbf{u}$ is called *standard* if each of its prefixes is a left special factor of $\mathbf{u}$. For every Arnoux–Rauzy word, there exists a unique standard Arnoux–Rauzy word with the same language. As Sturmian words have languages closed under reversal and they have exactly one right special factor of every length, it is readily seen that Arnoux–Rauzy words over the binary alphabet coincide with the Sturmian words. Moreover, it follows from [Lot02, Proposition 2.1.23] that the set of left special factors of a Sturmian word is the set of prefixes of the characteristic word with the same slope. Consequently, standard Arnoux–Rauzy words over the binary alphabet coincide with the characteristic Sturmian words.

Example 4.1.6. In analogy with the Fibonacci word, we define the Tribonacci word which generalizes the Fibonacci word to the ternary alphabet. It was first introduced in [Rau82] as the fixed point of the morphism $0 \mapsto 01$, $1 \mapsto 02$, $2 \mapsto 0$, but we can define it in analogy to the definition of the Fibonacci word in Example 4.1.1. Let $(t_n)_{n=0}^{+\infty}$ be the sequence of finite words defined recurrently by the relation $t_{n+3} = t_{n+2}t_{n+1}t_n$, for every $n \in \mathbb{N}$, with the initial conditions $t_0 = 0$, $t_1 = 01$ and $t_2 = 0102$. The Tribonacci word $\mathbf{t}$ is the following limit

$$\mathbf{t} = \lim_{n\to+\infty} t_n = 010201001020101020100102001 \cdots .$$

We summarize some properties of the Arnoux–Rauzy words. A d-ary Arnoux–Rauzy word $\mathbf{u}$ has the factor complexity function $\mathcal{C}_\mathbf{u}(n) = (d-1)n+1$. Consequently, we observe that $\mathcal{C}_\mathbf{u}(n) \geq n+d-1$, for every $n \in \mathbb{N}$, which implies that $\mathbf{u}$ is aperiodic (see [Lot02, Theorem 1.3.13]). It can be derived easily that a word is uniformly recurrent if and only if its set of return words $\mathcal{R}_\mathbf{u}(y)$ to every factor $y \in \mathcal{L}(\mathbf{u})$ is finite. A d-ary Arnoux–Rauzy word $\mathbf{u}$ is uniformly recurrent and, moreover, its set of return words $\mathcal{R}_\mathbf{u}(y)$ to every factor $y \in \mathcal{L}(\mathbf{u})$ has exactly d elements [JV00]. Similarly, the derived word to a prefix of a d-ary Arnoux–Rauzy word is again a d-ary Arnoux–Rauzy word [Med19].

Standard Arnoux–Rauzy words can be described by their S-adic representation based on the morphisms $\varphi_a : A^* \to A^*$ defined for every $a \in A$ and every $x \in A$ so that

$$\varphi_a(x) = \begin{cases} a, & \text{if } x = a; \\ ax, & \text{otherwise.} \end{cases} \tag{4.5}$$

For every standard d-ary Arnoux–Rauzy word $\mathbf{u}$ there exists a unique sequence of morphisms $\Delta = (\psi_n)_{n=1}^{+\infty}$ such that $\psi_i \in \{\varphi_a : a \in A\}$, for every $i \in \mathbb{N}$, and a unique sequence of standard Arnoux–Rauzy words $(\mathbf{u}^{(n)})_{n=1}^{+\infty}$ such that $\mathbf{u} = \psi_1\psi_2\psi_3 \cdots \psi_n(\mathbf{u}^{(n)})$, for every $n \geq 1$ [JP02]. The sequence Δ is called the *directive sequence* of $\mathbf{u}$.

In Section 4.2, we study regular d-ary Arnoux–Rauzy words which were first studied by Glen [Gle07] and, more recently, by Peltomäki [Pel21], who was the first to call them by the name regular. A d-ary Arnoux–Rauzy word on the alphabet $\{1, 2, \ldots, d\}$ is called *regular* if its directive sequence Δ written in the form $\psi_1^{a_1}\psi_2^{a_2}\cdots$ with $\psi_N \neq \psi_{N+1}$ and $a_N > 0$ for all $N \geq 1$ fulfills that the sequence $(\psi_i)_{i=1}^{+\infty}$ is periodic with the period $\varphi_1\varphi_2\cdots\varphi_{\mathtt{d}}$; see Example 4.1.7 and Example 4.1.8 The *slope* of such a regular Arnoux–Rauzy word is defined as $\theta = (a_N)_{N\geq 0}$ with $a_0 = 0$. The name slope is chosen for the following reason. All Sturmian words are regular and the slope $\theta = (a_N)_{N\geq 0}$ of a Sturmian word $\mathbf{u}$ is closely related to the slope β of $\mathbf{u}$ interpreted as a cutting sequence – the continued fraction of β fulfills the condition that $\beta = [0; a_1, a_2, a_3, \ldots]$, which means that

$$\beta = 0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}$$

Observe that as the slope θ fulfills by definition the condition that $\theta < 1$, we describe only cutting sequences with the slope $\beta < 1$. Such cutting sequences have a frequency of the letter 1 smaller than the frequency of the letter 0; see Equation (4.3). We could recover the remaining Sturmian words by exchanging the letters 0 and 1.

Example 4.1.7. To clarify the definition of regular Arnoux–Rauzy words, we illustrate that the classes of regular Arnoux–Rauzy words and fixed points of morphisms are not a subset of one another. The fixed point $\mathbf{u}$ of the morphism $\varphi_1\varphi_2\varphi_1\varphi_3$ is not regular, even though its directive sequence $\Delta = (\varphi_1\varphi_2\varphi_1\varphi_3)^\omega$ is periodic. On the other hand, the word $\mathbf{u}$ with the directive sequence $\Delta = \varphi_1\varphi_2^2\varphi_3^3\varphi_1^4\varphi_2^5\varphi_3^6\cdots$ is regular, even though its directive sequence is not periodic.

Example 4.1.8. The Fibonacci word $\mathbf{f}$ has the directive sequence $\Delta = (\varphi_0\varphi_1)^\omega$, which results in the slope $\theta = (0, \overline{1})$. Simultaneously, $\mathbf{f}$ interpreted as a cutting sequence has the slope $\frac{1}{\tau}$, where τ denotes the golden mean and $\frac{1}{\tau} = [0; \overline{1}]$; see Example 4.1.4. The Tribonacci word $\mathbf{t}$ has the directive sequence $\Delta = (\varphi_0\varphi_1\varphi_2)^\omega$, which results in the slope $\theta = (0, \overline{1})$. Therefore the Fibonacci and Tribonacci word have the same slope, even though the cardinality of their alphabet differs.

We observed in Example 4.1.8 that the Fibonacci and Tribonacci word have the same slope $\theta = (0, \overline{1})$. This leads to a generalization called the d-bonacci word $\mathbf{u}_d$, for every integer $d \geq 2$.

Definition 4.1.9 (d-bonacci words)**.** *Let $d \geq 2$ be an integer. A regular d-ary Arnoux–Rauzy word $\mathbf{u}$ with the slope $\theta = (0, \overline{1})$ is called a d-bonacci word and denoted $\mathbf{u}_d$.*

4.2 Critical exponents of regular Arnoux–Rauzy words

In this part, we present results on the critical and asymptotic critical exponent of regular Arnoux–Rauzy words. The critical exponent of an infinite word expresses the maximal repetition rate of factors in the word. Similarly, the asymptotic critical exponent of an infinite word expresses the maximal repetition rate of factors in the word when their length grows to infinity.

If u is a non-empty word of length $\ell \in \mathbb{N}$ and $e \in \mathbb{Q}$ is a positive rational number of the form n/ℓ, then u^e denotes the prefix of length n of the infinite periodic word $uuu\cdots = u^\omega$. The rational exponent e describes the repetition rate of u in the string u^e. The *critical exponent* $\mathrm{E}(\mathbf{u})$ of an infinite word $\mathbf{u}$ is defined as

$$\mathrm{E}(\mathbf{u}) = \sup\{e \in \mathbb{Q} :\ u^e \in \mathcal{L}(\mathbf{u}) \ \text{ for some } \ u \neq \varepsilon\}\,.$$

The *asymptotic critical exponent* $\mathrm{E}^*(\mathbf{u})$ of $\mathbf{u}$ is defined as

$$\mathrm{E}^*(\mathbf{u}) = \begin{cases} +\infty, & \text{if } \mathrm{E}(\mathbf{u}) = +\infty,\\ \limsup\limits_{n\to\infty} \max\{e \in \mathbb{Q} : u^e \ \in \mathcal{L}(\mathbf{u}) \ \text{ and } \ |u| = n\}, & \text{otherwise.} \end{cases}$$

Clearly, $\mathrm{E}^*(\mathbf{u}) \leq \mathrm{E}(\mathbf{u})$ and the equality holds whenever $\mathrm{E}(\mathbf{u})$ is irrational. By the term *critical exponents*, we mean both the critical and the asymptotic critical exponent together. Let us remark that the terminology around critical exponents is not unified. The critical exponent is called the index in [Cas08] and the free index in [TW07]. The asymptotic critical exponent is called the asymptotic index in [Cas08] and it is called the critical exponent in [JP02, GJ09].

In [DDP23], Dolce, Dvořáková and Pelantová deduced a formula to compute the critical exponents of a uniformly recurrent aperiodic word based on its bispecial factors and return words. More precisely, if $\mathbf{u}$ is a uniformly recurrent aperiodic word, $(b_n)_{n\in\mathbb{N}}$ is the sequence of all bispecial factors in $\mathbf{u}$ ordered by length, [1] and, for every $n \in \mathbb{N}$, r_n is the shortest return word to the bispecial factor b_n, then the critical exponents of $\mathbf{u}$ satisfy the formula

$$\mathrm{E}(\mathbf{u}) = 1 + \sup_{n\in\mathbb{N}} \left\{\frac{|b_n|}{|r_n|}\right\} \qquad \text{and} \qquad \mathrm{E}^*(\mathbf{u}) = 1 + \limsup_{n\to\infty} \frac{|b_n|}{|r_n|}. \tag{4.6}$$

As d-ary Arnoux–Rauzy words are uniformly recurrent and aperiodic, Equation (4.6) served as a base for the formula derived in [DL23] for a subset of regular words.

Theorem 4.2.1. [DL23] *The critical exponent and the asymptotic critical exponent of a regular d-ary Arnoux–Rauzy word $\mathbf{u}$ with the slope $\theta = (a_N)_{N\geq 0}$ satisfy*

$$\mathrm{E}(\mathbf{u}) = \frac{d}{d-1} + \sup_{N\geq 1} \left\{ a_N + \frac{1}{d-1} \frac{\sum_{i=2}^{d} \left((d-i)a_{N-i+1} + 1\right) Q_{N-i} - d}{Q_{N-1}} \right\}; \tag{4.7}$$

$$\mathrm{E}^*(\mathbf{u}) = \frac{d}{d-1} + \limsup_{N\to+\infty} \left(a_N + \frac{1}{d-1} \frac{\sum_{i=2}^{d} \left((d-i)a_{N-i+1} + 1\right) Q_{N-i}}{Q_{N-1}} \right),$$

where Q_N fulfills $Q_N = \sum_{i=1}^{d-1} a_{N-i+1} Q_{N-i} + Q_{N-d}$, for every $N \geq 1$, with the initial conditions $a_i = 0$ and $Q_i = 1$ for every $i \in \{-d+1, \ldots, -1, 0\}$.

[1] An arbitrary order may be chosen for the bispecial factors having the same length.

Choosing $d = 2$ in Theorem 4.2.1 gives the formula for the critical exponent of Sturmian words derived independently by Damanik and Lenz [DL02] and Carpi and de Luca [CdL00].

We show in [DL23] that the minimal (asymptotic) critical exponent among regular Arnoux–Rauzy words is attained for the d-bonacci word; see Definition 4.1.9 and Theorem 4.2.2. A stronger result is known for the critical exponent in the binary case – the minimal value of the critical exponent among all Sturmian words is attained by the Fibonacci word [CdL00, DL02].

Theorem 4.2.2. [DL23] *Let $d \geq 2$. The minimal (asymptotic) critical exponent for regular d-ary Arnoux–Rauzy words is attained by the d-bonacci word $\mathbf{u}_d$ with the slope $\theta = (a_N) = (0, \overline{1})$. Moreover, denoting t the dominant real root of the polynomial $x^d - \sum_{i=1}^{d} x^{d-i}$, the asymptotic critical exponent $\mathrm{E}^*(\mathbf{u}_d)$ can be computed as*

$$\mathrm{E}^*(\mathbf{u}_d) = 2 + \frac{1}{t-1}. \tag{4.8}$$

The critical exponent of the Fibonacci and Tribonacci word is irrational [CdL00], [TW07]. Therefore the Fibonacci and Tribonacci word fulfill the condition that their critical exponent and asymptotic critical exponent coincide. In [DL23], we conjectured that every d-bonacci word has this property. In the next part, we provide a proof of this conjecture for $d \leq 15$.

Conjecture 4.2.3. [DL23] *Let $d \geq 4$. The d-bonacci word $\mathbf{u}_d$ satisfies $\mathrm{E}(\mathbf{u}_d) = \mathrm{E}^*(\mathbf{u}_d)$.*

During the finalization of this thesis, a significantly stronger result was proved. For every $d \geq 2$, the critical and asymptotic critical exponent of a d-bonacci word coincide and, moreover, the minimal critical exponent among all d-ary episturmian words is attained by the d-bonacci word [DP23].

4.3 Critical exponents of d-bonacci words

In this part, we prove Conjecture 4.2.3 for $d \in [\![4, 15]\!]$. First, we summarize important facts about the d-bonacci words, which have the slope $\theta = (0, \overline{1})$. Thus, the sequence (a_N) in Theorem 4.2.1 fulfills the condition that

$$a_N = \begin{cases} 0, & \text{for every } N \leq 0; \\ 1, & \text{for every } N \geq 1, \end{cases} \tag{4.9}$$

and the sequence (Q_N) satisfies the condition that

$$Q_N = \begin{cases} 1, & \text{for every } N \in [\![-d+1, -1]\!]; \\ 2^N, & \text{for every } N \in [\![0, d-1]\!]; \\ \sum_{i=1}^{d} Q_{N-i}, & \text{for every } N \geq d. \end{cases} \tag{4.10}$$

Also, we recall that the polynomial $x^d - \sum_{i=1}^{d} x^{d-i}$ has d roots, which we denote $t_1, \ldots, t_d \in \mathbb{C}$. One of the roots, say t_1, is called dominant as it fulfills the condition that $t_1 > |t_i|$, for every $i \geq 2$, and moreover $t_1 \in \mathbb{R}$.

The critical exponents $\mathrm{E}(\mathbf{u})$ and $\mathrm{E}^*(\mathbf{u})$ of an infinite word $\mathbf{u}$ fulfill the condition that $\mathrm{E}^*(\mathbf{u}) \leq \mathrm{E}(\mathbf{u})$. Therefore they coincide if and only if $\mathrm{E}(\mathbf{u}) \leq \mathrm{E}^*(\mathbf{u})$. Considering Equations (4.7) and (4.8), it suffices to prove the following proposition to prove Conjecture 4.2.3 for $d \leq 15$.

Proposition 4.3.1. *Let $d \in \mathbb{N}$ be such that $4 \leq d \leq 15$. We denote t the dominant real root of the polynomial $x^d - \sum_{i=1}^{d} x^{d-i}$. The d-bonacci word $\mathbf{u}_d$ fulfills that, for every integer $N \geq 1$,*

$$\tfrac{d}{d-1} + 1 + \tfrac{1}{d-1} \tfrac{\sum_{i=2}^{d}((d-i)a_{N-i+1}+1)Q_{N-i}-d}{Q_{N-1}} \leq 2 + \tfrac{1}{t-1}, \tag{4.11}$$

where the sequences (a_N) and (Q_N) are given in Equations (4.9) *and* (4.10)*.*

We summarize the results presented in [DD14] in the following lemma.

Lemma 4.3.2. [DD14] *Let $d \geq 4$ be an integer. Denote $t_1, \ldots, t_d$ the roots of the polynomial $x^d - \sum_{i=1}^{d} x^{d-i}$ and, in particular, denote $t = t_1 \in \mathbb{R}$ the dominant root. It holds that*

i) $2 - \frac{1}{d} < t < 2$;

ii) there exist $c_1 \in \mathbb{R}$ and $c_2, \ldots, c_d \in \mathbb{C}$ such that, for every $N \geq 0$, we have $Q_N = \sum_{i=1}^{d} c_i t_i^N$;

iii) we have $\lim_{N \to +\infty} e_N = 0$, where we denote

$$e_N = Q_N - c_1 t^N = \sum_{i=2}^{d} c_i t_i^N, \quad \text{for every } N \geq 0. \tag{4.12}$$

We prove a lemma inspired by [DD14]. It has an important corollary stated below.

Lemma 4.3.3. *Let $d \geq 4$ be an integer. Let $(e_N)_{N \geq 0}$ be the sequence defined by Equation* (4.12)*. Let $K \geq 0$ and let $M \geq d$ be an integer such that $|e_{M-i}| \leq K$, for every $i \in [\![1, d]\!]$. Then*

$$|e_M| \leq K.$$

Proof. Let $K \geq 0$ and let $M \geq d$ be an integer such that $|e_{M-i}| \leq K$, for every $i \in [\![1, d]\!]$. We assume by contradiction that $|e_M| > K$. Consequently, we obtain

$$|e_M| - |e_{M-i}| > 0, \text{ for every } i \in [\![1, d]\!]. \tag{4.13}$$

By definition, (e_N) fulfills the linear recurrence relation $e_N = \sum_{i=1}^{d} e_{N-i}$, for every $N \geq d$; see Equation (4.12). Hence, we have for every $N \geq d$ that

$$e_{N+1} - e_N = \sum_{i=1}^{d} e_{N+1-i} - \sum_{i=1}^{d} e_{N-i} = e_N - e_{N-d}. \tag{4.14}$$

We prove that $|e_{N+1}| > |e_N|$, for every $N \geq M$, in two steps 1) and 2). This will be a contradiction to Lemma 4.3.2 iii) and finish the proof.

1) We prove that $|e_{N+1}| > |e_N|$ for every integer N such that $M \leq N \leq M + d - 1$. If $N = M$, then using Equation (4.14) and Equation (4.13) with $i = d$, we have

$$|e_{M+1}| = |2e_M - e_{M-d}| \geq 2|e_M| - |e_{M-d}| > |e_M|.$$

Assume that for an integer N such that $M + 1 \leq N \leq M + d - 1$, it holds that $|e_N| > |e_M|$. Note that we have $M - N + d \in [\![1, d-1]\!]$. Then using Equation (4.14) and Equation (4.13) with $i = M - N + d$, we have

$$|e_{N+1}| = |2e_N - e_{N-d}| \geq 2|e_N| - |e_{N-d}| > |e_N| + |e_M| - |e_{M-(M-N+d)}| > |e_N|.$$

So far, we proved that for $N = M + d$, we have

$$|e_N| > |e_{N-1}| > \ldots > |e_{N-d+1}| > |e_{N-d}| = |e_M|. \tag{4.15}$$

2) We prove by induction that $|e_{N+1}| > |e_N|$, for every $N \geq M + d$. Induction hypothesis: Equation (4.15) holds for an integer $N \geq M + d$. Using Equation (4.14) and Equation (4.15), we have

$$|e_{N+1}| = |2e_N - e_{N-d}| \geq 2|e_N| - |e_{N-d}| > |e_N|,$$

and thus Equation (4.15) holds for $N + 1$. □

Corollary 4.3.4. *Let $K \geq 0$ and let $M \geq d$ be an integer such that $|e_{M-i}| \leq K$, for every $i \in [\![1, d]\!]$. Then $|e_N| \leq K$, for every $N \geq M$.*

We need another lemma to prove Proposition 4.3.1.

Lemma 4.3.5. *Let $d \geq 4$ be an integer. Let $(Q_N)_{N\geq 0}$ and $(e_N)_{N\geq 0}$ be the sequences defined by Equation (4.10) and Equation (4.12), respectively. Let $N \geq d$. The following inequalities are equivalent*

$$(t-1)\left(Q_N + \sum_{i=2}^{d-1}(d-i)Q_{N-i} - d\right) \leq (d-1)Q_{N-1}, \tag{4.16}$$

$$(t-1)\left(e_N + \sum_{i=2}^{d-1}(d-i)e_{N-i} - d\right) \leq (d-1)e_{N-1}. \tag{4.17}$$

Proof. We start by deriving a relation, which will be useful in the proof. We denote t the dominant real root of the polynomial $x^d - \sum_{i=1}^{d} x^{d-i}$. For every $m \in \mathbb{N}$, summing up m powers of t gives $\sum_{j=0}^{m} t^j = \frac{t^{m+1}-1}{t-1}$, which we can differentiate with respect to t to obtain

$$\sum_{j=1}^{m} j t^{j-1} = \frac{(m+1)t^m(t-1) - (t^{m+1}-1)}{(t-1)^2}. \tag{4.18}$$

Substituting Equation (4.12) into Inequality (4.16), we have

$$(t-1)\left(c_1 t^N + e_N + \sum_{i=2}^{d-1}(d-i)c_1 t^{N-i} + \sum_{i=2}^{d-1}(d-i)e_{N-i} - d\right) = (d-1)\left(c_1 t^{N-1} + e_{N-1}\right).$$

Thus, to prove the equivalence between Inequalities (4.16) and (4.17), it suffices to show that

$$(t-1)c_1 t^N + (t-1)\sum_{j=1}^{d-2} j c_1 t^{N+j-d} = (d-1)c_1 t^{N-1}. \tag{4.19}$$

We modify Equation (4.19) in the following steps, using twice Equation (4.18)

$$\begin{aligned}
(t-1)t^N + (t-1)\sum_{j=1}^{d-2} jt^{N+j-d} &= (d-1)t^{N-1},\\
(t-1)t^N + (t-1)t^{N-d+1}\sum_{j=1}^{d-2} jt^{j-1} &= (d-1)t^{N-1},\\
(t-1)t^{d-1} + (t-1)\sum_{j=1}^{d-2} jt^{j-1} &= (d-1)t^{d-2},\\
(t-1)t^{d-1} + (t-1)\frac{(d-1)t^{d-2}(t-1)-(t^{d-1}-1)}{(t-1)^2} &= (d-1)t^{d-2},\\
(t-1)t^{d-1} + (d-1)t^{d-2} - \frac{(t^{d-1}-1)}{t-1} &= (d-1)t^{d-2},\\
(t-1)t^{d-1} - \frac{(t^{d-1}-1)}{t-1} &= 0,\\
(t-1)t^{d-1} - \sum_{j=0}^{d-2} t^j &= 0,\\
t^d - \sum_{j=0}^{d-1} t^j &= 0,
\end{aligned}$$

which holds as t is a root of the polynomial $x^d - \sum_{j=0}^{d-1} x^j$. □

We may now prove Proposition 4.3.1.

Proof of Proposition 4.3.1. Let $d \in [\![4, 15]\!]$. We observe that Equation (4.11) can be modified into the following equivalent forms

$$\frac{1}{d-1}\frac{Q_{N-1} + \sum_{i=2}^{d}\left((d-i)a_{N-i+1}+1\right)Q_{N-i} - d}{Q_{N-1}} \leq \frac{1}{t-1},$$

$$(t-1)\left(Q_{N-1} + \sum_{i=2}^{d}\left((d-i)a_{N-i+1}+1\right)Q_{N-i} - d\right) \leq (d-1)Q_{N-1}. \tag{4.20}$$

Also, from Equation (4.10) and Lemma 4.3.2 i), we have that, for every $i \in [\![0, d-2]\!]$,

$$tQ_i = t2^i < 2 \cdot 2^i = 2^{i+1} = Q_{i+1}. \tag{4.21}$$

1) Assume $N \in [\![1, d-1]\!]$. Then we have that

$$2 - (d-N+1) - N(t-1) \leq 2 - 2 - N = -N < 0. \tag{4.22}$$

The left-hand side of Inequality (4.20) can be gradually rewritten using Equation (4.10), Inequality (4.21) and Inequality (4.22), which finishes the proof of part 1):

$$\begin{aligned}
&(t-1)\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}+\sum_{i=N+1}^{d}Q_{N-i}-d\right)\\
&=(t-1)\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}+(d-N)-d\right)\\
&=t\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}\right)-\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}\right)-N(t-1)\\
&\le\left(Q_{N}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i+1}\right)-\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}\right)-N(t-1)\\
&=\left(Q_{N}+\sum_{i=1}^{N-1}((d-i-1)+1)Q_{N-i}\right)-\left(Q_{N-1}+\sum_{i=2}^{N}((d-i)+1)Q_{N-i}\right)-N(t-1)\\
&=Q_N+(d-2)Q_{N-1}-\sum_{i=2}^{N-1}Q_{N-i}-(d-N+1)Q_0-N(t-1)\\
&=2^N+(d-2)2^{N-1}-\sum_{i=1}^{N-2}2^i-(d-N+1)-N(t-1)\\
&=2^N+(d-2)2^{N-1}-(2^{N-1}-1)+1-(d-N+1)-N(t-1)\\
&=(d-1)2^{N-1}+2-(d-N+1)-N(t-1)\\
&\le(d-1)2^{N-1}\\
&=(d-1)Q_{N-1}.
\end{aligned}$$

2) Assume $N \ge d$. We modify the left-hand side of Inequality (4.20), obtaining

$$\begin{aligned}
&(t-1)\left(Q_{N-1}+\sum_{i=2}^{d}\left((d-i)+1\right)Q_{N-i}-d\right)\\
=&(t-1)\left(Q_{N-1}+\sum_{i=2}^{d}(d-i)Q_{N-i}+\sum_{i=2}^{d}Q_{N-i}-d\right)\\
=&(t-1)\left(Q_{N-1}+\sum_{i=2}^{d}(d-i)Q_{N-i}+Q_N-Q_{N-1}-d\right)\\
=&(t-1)\left(Q_{N}+\sum_{i=2}^{d-1}(d-i)Q_{N-i}-d\right).
\end{aligned}$$

Then, Inequality (4.20) becomes Inequality (4.16), which holds if and only if Inequality (4.17) holds; see Lemma 4.3.5. We prove Inequality (4.17) to finish the proof.

We show in Table 4.1 that there exists an integer $M \ge d$ such that $|e_{M-i}| \le \frac{1}{d-2}$, for every $i \in [\![1,d]\!]$. Then we have from Corollary 4.3.4 that $|e_N| \le \frac{1}{d-2}$, for every $N \ge M$.

a) In the case that $N \ge M+d-1$, we can modify the left-hand side of Inequality (4.17) in the following way

$$(t-1)\left(e_N + \sum_{i=2}^{d-1}(d-i)e_{N-i} - d\right) \leq (t-1)\left(\tfrac{1}{d-2} + \tfrac{1}{d-2}\sum_{i=1}^{d-2} j - d\right)$$
$$= (t-1)\left(\tfrac{1}{d-2} + \tfrac{(d-1)(d-2)}{2(d-2)} - d\right)$$
$$= (t-1)\left(\tfrac{1}{d-2} - \tfrac{d+1}{2}\right)$$

and the right-hand side of Inequality (4.17) in the following way

$$(d-1)e_{N-1} \geq (d-1)\left(-\tfrac{1}{d-2}\right) = -\tfrac{d-1}{d-2}.$$

To prove Inequality (4.17), it suffices to prove

$$(t-1)\left(\tfrac{1}{d-2} - \tfrac{d+1}{2}\right) \leq -\tfrac{d-1}{d-2},$$

which is equivalent to the following series of inequalities:

$$\tfrac{d-1}{d-2} \leq (t-1)\left(\tfrac{-2+(d+1)(d-2)}{2(d-2)}\right),$$
$$2(d-1) \leq (t-1)(d^2-d-4).$$

As by Lemma 4.3.2 i) we have $t > 2 - \frac{1}{d}$, we strengthen the condition to

$$2(d-1) \leq (1-\tfrac{1}{d})(d^2-d-4),$$

which holds for $d \geq 4$.

b) It remains to verify Equation (4.17) for $N \in [\![d, M+d-2]\!]$. This finite number of cases was treated by computer experiments. □

d	4	5	6	7	8	9	10	11	12	13	14	15
$\frac{1}{d-2}$	0.500	0.333	0.250	0.200	0.167	0.143	0.125	0.111	0.100	0.091	0.083	0.077
$\lvert e_0 \rvert$	0.092	0.058	0.035	0.021	0.012	0.007	0.004	0.002	0.001	0.001	0.000	0.000
$\lvert e_1 \rvert$	0.104	0.079	0.053	0.033	0.020	0.012	0.007	0.004	0.002	0.001	0.001	0.000
$\lvert e_2 \rvert$	0.056	0.087	0.072	0.050	0.032	0.020	0.012	0.007	0.004	0.002	0.001	0.001
$\lvert e_3 \rvert$	0.182	0.035	0.078	0.068	0.049	0.032	0.020	0.012	0.007	0.004	0.002	0.001
$\lvert e_4 \rvert$	-	0.203	0.023	0.072	0.066	0.048	0.032	0.020	0.012	0.007	0.004	0.002
$\lvert e_5 \rvert$	-	-	0.218	0.014	0.068	0.065	0.048	0.031	0.020	0.012	0.007	0.004
$\lvert e_6 \rvert$	-	-	-	**0.229**	0.009	0.066	0.064	0.047	0.031	0.020	0.012	0.007
$\lvert e_7 \rvert$	-	-	-	0.030	**0.236**	0.006	0.065	0.063	0.047	0.031	0.020	0.012
$\lvert e_8 \rvert$	-	-	-	0.040	0.021	**0.241**	0.003	0.064	0.063	0.047	0.031	0.020
$\lvert e_9 \rvert$	-	-	-	0.046	0.029	0.014	**0.245**	0.002	0.063	0.063	0.047	0.031
$\lvert e_{10} \rvert$	-	-	-	0.042	0.038	0.020	0.009	**0.247**	0.001	0.063	0.063	0.047
$\lvert e_{11} \rvert$	-	-	-	0.015	0.043	0.028	0.013	0.005	**0.248**	0.001	0.063	0.063
$\lvert e_{12} \rvert$	-	-	-	0.042	0.038	0.037	0.020	0.009	0.003	**0.249**	0.000	0.063
$\lvert e_{13} \rvert$	-	-	-	0.098	0.010	0.042	0.028	0.013	0.005	0.002	**0.249**	0.000
$\lvert e_{14} \rvert$	-	-	-	-	0.048	0.036	0.036	0.020	0.009	0.003	0.001	**0.250**
$\lvert e_{15} \rvert$	-	-	-	-	0.105	0.007	0.041	0.028	0.013	0.005	0.002	0.001
$\lvert e_{16} \rvert$	-	-	-	-	-	0.053	0.034	0.036	0.020	0.009	0.003	0.001
$\lvert e_{17} \rvert$	-	-	-	-	-	0.111	0.004	0.040	0.028	0.013	0.005	0.002
$\lvert e_{18} \rvert$	-	-	-	-	-	-	0.056	0.033	0.036	0.020	0.009	0.003
$\lvert e_{19} \rvert$	-	-	-	-	-	-	0.115	0.003	0.040	0.027	0.013	0.005
$\lvert e_{20} \rvert$	-	-	-	-	-	-	-	0.058	0.032	0.035	0.020	0.009
$\lvert e_{21} \rvert$	-	-	-	-	-	-	-	**0.119**	0.002	0.039	0.027	0.013
$\lvert e_{22} \rvert$	-	-	-	-	-	-	-	0.010	0.060	0.032	0.035	0.020
$\lvert e_{23} \rvert$	-	-	-	-	-	-	-	0.014	**0.121**	0.001	0.039	0.027
$\lvert e_{24} \rvert$	-	-	-	-	-	-	-	0.019	0.006	0.061	0.032	0.035
$\lvert e_{25} \rvert$	-	-	-	-	-	-	-	0.024	0.009	**0.122**	0.001	0.039
$\lvert e_{26} \rvert$	-	-	-	-	-	-	-	0.028	0.014	0.004	0.061	0.031
$\lvert e_{27} \rvert$	-	-	-	-	-	-	-	0.029	0.019	0.006	**0.123**	0.000
$\lvert e_{28} \rvert$	-	-	-	-	-	-	-	0.022	0.024	0.009	0.003	0.062
$\lvert e_{29} \rvert$	-	-	-	-	-	-	-	0.003	0.028	0.014	0.004	**0.124**
$\lvert e_{30} \rvert$	-	-	-	-	-	-	-	0.026	0.028	0.018	0.006	0.002
$\lvert e_{31} \rvert$	-	-	-	-	-	-	-	0.055	0.021	0.024	0.009	0.003
$\lvert e_{32} \rvert$	-	-	-	-	-	-	-	0.053	0.002	0.028	0.013	0.004
$\lvert e_{33} \rvert$	-	-	-	-	-	-	-	-	0.028	0.028	0.018	0.006
$\lvert e_{34} \rvert$	-	-	-	-	-	-	-	-	0.058	0.020	0.024	0.009
$\lvert e_{35} \rvert$	-	-	-	-	-	-	-	-	0.055	0.001	0.028	0.013
$\lvert e_{36} \rvert$	-	-	-	-	-	-	-	-	-	0.029	0.028	0.018
$\lvert e_{37} \rvert$	-	-	-	-	-	-	-	-	-	0.059	0.020	0.024
$\lvert e_{38} \rvert$	-	-	-	-	-	-	-	-	-	0.058	0.001	0.027
$\lvert e_{39} \rvert$	-	-	-	-	-	-	-	-	-	-	0.030	0.028
$\lvert e_{40} \rvert$	-	-	-	-	-	-	-	-	-	-	0.060	0.020
$\lvert e_{41} \rvert$	-	-	-	-	-	-	-	-	-	-	0.059	0.001
$\lvert e_{42} \rvert$	-	-	-	-	-	-	-	-	-	-	-	0.030
$\lvert e_{43} \rvert$	-	-	-	-	-	-	-	-	-	-	-	0.061
$\lvert e_{44} \rvert$	-	-	-	-	-	-	-	-	-	-	-	0.060

Table 4.1: The values $|e_i|$ for $d \in [\![4, 15]\!]$ and $i \in [\![0, 44]\!]$. If the values $|e_i|$ satisfy for some integer $M \geq d$ and for every $i \in [\![1, d]\!]$ the condition that $|e_{M-i}| \leq \frac{1}{d-2}$, then the values $|e_i|$ for $i \geq M$ are omitted. The values $|e_i|$ such that $|e_i| > \frac{1}{d-2}$ are shown in bold.

Chapter 5

Faithful representation of Sturmian morphisms

In this chapter, we present our results on a faithful representation of the special Sturmian monoid and on the so-called square roots of fixed points of Sturmian morphisms [LPS23]. Results which do not make part of the publication include an algorithm to determine the faithful representation of morphisms in the special Sturmian monoid, which helps to clarify the relationship between the representations of mutually conjugate morphisms. Based on this algorithm, we describe the relationship between intercepts of fixed points of mutually conjugate morphisms in the special Sturmian monoid, which provides an answer to an open question in [LPS23].

5.1 Introduction to Sturmian morphisms

A morphism is called *Sturmian* if it maps Sturmian words to Sturmian words. It is readily seen that a composition of Sturmian morphisms is a Sturmian morphism, and thus Sturmian morphisms form a submonoid of the monoid of morphisms on the binary alphabet. This submonoid is called the monoid of Sturm and it is denoted St. In particular, a Sturmian morphism is called *standard* if it maps every characteristic Sturmian word to a characteristic Sturmian word. Two trivial examples of Sturmian morphisms are the identity morphism $\mathrm{Id} : 0 \mapsto 0,\ 1 \mapsto 1$ and the morphism $E : 0 \mapsto 1,\ 1 \mapsto 0$. The following morphisms

$$\varphi : \begin{cases} 0 \mapsto 01, \\ 1 \mapsto 0, \end{cases} \quad \widetilde{\varphi} : \begin{cases} 0 \mapsto 10, \\ 1 \mapsto 0, \end{cases}$$

map mechanical words to mechanical words, see [Lot02, Lemma 2.2.18]. Consequently, φ and $\widetilde{\varphi}$ are Sturmian morphisms; see Section 4.1.1. Moreover, it can be shown that the morphisms E, φ and $\widetilde{\varphi}$ generate the monoid of Sturm, which is denoted as

$$\mathrm{St} = \langle E, \varphi, \widetilde{\varphi} \rangle.$$

In other words, every Sturmian morphism is a composition of the morphisms E, φ and $\widetilde{\varphi}$. The morphism φ is a substitution and we refer to it as to the Fibonacci substitution due to the following example.

Example 5.1.1. , The Fibonacci word $\mathbf{f}$ defined in Example 4.1.1 is the fixed point of the morphism φ. Indeed, it follows from [Lot02, Lemma 2.2.18] that the morphism φ maps every lower mechanical word $\mathbf{s}_{\alpha,\rho}$ to the upper mechanical word $\mathbf{s}'_{\frac{1-\alpha}{2-\alpha},\frac{1-\rho}{2-\alpha}}$. Together with Equation (4.2), we have that $\varphi(\mathbf{f}) = \mathbf{s}'_{\frac{1}{\tau^2},\frac{1}{\tau^2}} = \mathbf{f}$.

Moreover, it follows from [Lot02, Theorem 2.3.12] that a morphism is standard if and only if it maps a characteristic word to a characteristic word. In other words, it suffices that the morphism maps one characteristic word to a characteristic word. The Fibonacci word $\mathbf{f}$ is characteristic and $\varphi(\mathbf{f}) = \mathbf{f}$. Hence, the morphism φ is standard.

The generators of the monoid of Sturm enable defining other morphisms as follows

$$G = \varphi \circ E : \begin{cases} 0 \mapsto 0, \\ 1 \mapsto 01, \end{cases} \quad \widetilde{G} = \widetilde{\varphi} \circ E : \begin{cases} 0 \mapsto 0, \\ 1 \mapsto 10, \end{cases} \quad D = E \circ \varphi : \begin{cases} 0 \mapsto 10, \\ 1 \mapsto 1, \end{cases} \quad \widetilde{D} = E \circ \widetilde{\varphi} : \begin{cases} 0 \mapsto 01, \\ 1 \mapsto 1. \end{cases}$$

Remark 5.1.2. *Observe that choosing a binary alphabet* $\{0,1\}$ *in Equation* (4.5)*, we obtain that* $\varphi_0 = G$ *and* $\varphi_1 = D$*. However, the notation chosen in Chapter 4 is more suitable in the context of d-ary Arnoux–Rauzy words, whereas the notation* G *and* D *is typical in the context of Sturmian morphisms.*

The morphisms G, D, $\widetilde{G}$ and $\widetilde{D}$ generate the submonoid $\mathcal{M} = \langle G, D, \widetilde{G}, \widetilde{D} \rangle$ which is called the special Sturmian monoid. Every morphism $\psi \in \mathcal{M}$ can be written as a concatenation of the generators G, D, $\widetilde{G}$ and $\widetilde{D}$. However, this decomposition might not be unique as the special Sturmian monoid $\mathcal{M}$ is not free – for every $k \in \mathbb{N}$, it holds that

$$GD^k\widetilde{G} = \widetilde{G}\widetilde{D}^kG \quad \text{and} \quad DG^k\widetilde{D} = \widetilde{D}\widetilde{G}^kD. \tag{5.1}$$

Moreover, Equations (5.1) are a presentation of the monoid $\mathcal{M}$, which is to say that no other non-trivial independent relation can be found. Consequently, all decompositions of a Sturmian morphism have the same number of generators.

Example 5.1.3. The morphism $\eta : 0 \mapsto 010, 1 \mapsto 01010$ has two decompositions

$$\eta = GD\widetilde{G} \text{ and } \eta = \widetilde{G}\widetilde{D}G.$$

5.1.1 Mutually conjugate Sturmian morphisms

In this part, we recall the relation of conjugacy on Sturmian morphisms and its properties. A morphism η is a right conjugate of a morphism ψ, denoted by $\psi \triangleleft \eta$, if there exists a word $w \in A^*$ such that, for every letter $a \in A$,

$$w\psi(a) = \eta(a)w. \tag{5.2}$$

We say that a morphism ψ is a left conjugate of a morphism η if and only if $\psi \triangleleft \eta$.

Remark 5.1.4. *Note that there is some confusion around the notation of right conjugate morphisms. We choose notation close to the one used in [LPS23] to remain consistent while presenting our results. However, in [Lot02], the fact that* η *is the right conjugate of* ψ *is defined with Equation* (5.2) *symmetrically modified into* $w\eta(a) = \psi(a)w$*.*

Due to Remark 5.1.4, we reformulate [Lot02, Proposition 2.3.18] and [Lot02, Proposition 2.3.19] into our notation. They imply that the monoid of Sturm St is the closure under left conjugacy of the monoid of standard morphisms.

Proposition 5.1.5. [Lot02] *A morphism is Sturmian if and only if it is a left conjugate of a standard morphism.*

Proposition 5.1.6. [Lot02] *Let ψ and ψ' be standard morphisms such that $\psi \triangleleft \psi'$. Then $\psi = \psi'$.*

We call two morphisms η and ψ mutually conjugate (or conjugate) if $\eta \triangleleft \psi$ or $\psi \triangleleft \eta$. Clearly, this relation is an equivalence relation. It follows from Proposition 5.1.6 that the equivalence class of a Sturmian morphism has finitely many distinct elements and we can order the elements with respect to the relation $\triangleleft$ from the leftmost conjugate morphism to the rightmost conjugate morphism. The rightmost conjugate morphism is standard. Also, assuming that η and ψ are mutually conjugate and $\mathbf{u} \in \{0,1\}^{\mathbb{N}}$ is an infinite word such that 0 and 1 occur in $\mathbf{u}$, we have that if $\mathbf{u} = \psi(\mathbf{u}) = \eta(\mathbf{u})$, then $\psi = \eta$. Thus mutually conjugate morphisms which are distinct fix distinct words.

Clearly, mutually conjugate morphisms have the same incidence matrix. Moreover, a matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}(\mathbb{N}, 2)$ is the incidence matrix of $a+b+c+d-1$ mutually conjugate Sturmian morphisms [Lot02, Proposition 2.3.21], where we used the notation $\mathrm{Sl}(\mathbb{N}, n) = \{R \in \mathbb{N}^{n \times n} : \det R = 1\}$.

Example 5.1.7. The second power of the Fibonacci substitution $\varphi^2 : \mathtt{0} \mapsto 010$, $1 \mapsto 01$ has the incidence matrix $M = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$. All conjugate morphisms having the incidence matrix M are

$$\eta^{(0)} : \begin{cases} 0 \mapsto 010, \\ 1 \mapsto 10, \end{cases} \quad \eta^{(1)} : \begin{cases} 0 \mapsto 001, \\ 1 \mapsto 01, \end{cases} \quad \eta^{(2)} : \begin{cases} 0 \mapsto 100, \\ 1 \mapsto 10, \end{cases} \quad \eta^{(3)} : \begin{cases} 0 \mapsto 010, \\ 1 \mapsto 01, \end{cases}$$

and $\eta^{(0)} \triangleleft \eta^{(1)} \triangleleft \eta^{(2)} \triangleleft \eta^{(3)}$. The rightmost conjugate morphism $\eta^{(3)} = \varphi^2$ is standard.

Given a morphism $\eta^{(0)} \in \langle \widetilde{D}, \widetilde{G} \rangle$, the following algorithm produces all conjugate morphisms in the equivalence class of $\eta^{(0)}$ ordered from the leftmost morphism to the rightmost morphism. We call $||\eta^{(0)}|| = |\eta^{(0)}(0)| + |\eta^{(0)}(1)|$ the length of the morphism $\eta^{(0)}$.

Algorithm 5.1.8. *Put $\eta^{(0)} \in \langle \widetilde{D}, \widetilde{G} \rangle$ and denote $L = ||\eta^{(0)}||$. For $i \in [\![0, L-2]\!]$ do this:*

(1) Let $\psi_j \in \{G, D, \widetilde{G}, \widetilde{D}\}$, for every $j \in [\![0, k-1]\!]$, be such that $\eta^{(i)} = \psi_0 \psi_1 \cdots \psi_{k-1}$.

(2) Find minimal $j \in [\![0, k-1]\!]$ such that $\psi_j \in \{\widetilde{G}, \widetilde{D}\}$.

(3) If $\psi_j = \widetilde{D}$, then:

- *(a) Replace every morphism ψ_i such that $i \in [\![0, j-1]\!]$ and $\psi_i = G$ by the morphism $\widetilde{G}$.*
- *(b) Replace ψ_j by D.*

(4) If $\psi_j = \widetilde{G}$, then:

- *(a) Replace every morphism ψ_i such that $i \in [\![0, j-1]\!]$ and $\psi_i = D$ by the morphism $\widetilde{D}$.*
- *(b) Replace ψ_j by G.*

(5) Put $\eta^{(i+1)}$ equal to the resulting morphism.

We call the sequence of all Sturmian morphisms in the same equivalence class ordered with respect to the relation $\triangleleft$ a *chain of conjugate Sturmian morphisms*. Algorithm 5.1.8 produces a chain of conjugate Sturmian morphisms [Pel], which we illustrate in the following example.

i	$\eta^{(i)}$	$\eta^{(i)}(0)$	$\eta^{(i)}(1)$
5	DGG	10	10101
4	$\widetilde{D}GG$	01	01011
3	$DG\widetilde{G}$	10	10110
2	$\widetilde{D}G\widetilde{G}$	01	01101
1	$D\widetilde{G}\widetilde{G}$	10	11010
0	$\widetilde{D}\widetilde{G}\widetilde{G}$	01	10101

Table 5.1: A chain of Sturmian morphisms $(\eta^{(i)})_{i=0}^{5}$ such that $\eta^{(0)} = \widetilde{D}\widetilde{G}\widetilde{G}$.

Example 5.1.9. Let $\eta^{(0)} = \widetilde{D}\widetilde{G}\widetilde{G} : 0 \mapsto 01,\ 1 \mapsto 10101$. Then $||\eta^{(0)}|| = |01| + |10101| = 7$ and there are $||\eta^{(0)}|| - 1 = 6$ morphisms in the equivalence class of $\eta^{(0)}$, including itself. Observe in Table 5.1 that the sequence of morphisms $(\eta^{(i)})_{i=0}^{5}$ produced by Algorithm 5.1.8 fulfills that, for every $i \in [\![0, 4]\!]$, we have $\eta^{(i)} \triangleleft \eta^{(i+1)}$.

Given a particular morphism $\eta \in \mathcal{M}$, we can determine all morphisms in the equivalence class of η with the help of Algorithm 5.1.8. Denote $L = ||\eta||$ the length of η and consider the morphism $\eta^{(0)} \in \langle \widetilde{D}, \widetilde{G} \rangle$, which is created from the morphism η by replacing $G \mapsto \widetilde{G}$ and $D \mapsto \widetilde{D}$ in its decomposition. We denote this operation as

$$\eta^{(0)} = \mathrm{sub}_{G \mapsto \widetilde{G}}\, \mathrm{sub}_{D \mapsto \widetilde{D}}(\eta).$$

Then, Algorithm 5.1.8 produces a chain of conjugate Sturmian morphisms $(\eta^{(i)})_{i=0}^{L-1}$ such that $\eta^{(i)} = \eta$ for an integer $i \in [\![0, L-1]\!]$.

5.2 Faithful representation of the special Sturmian monoid

Elements of the special Sturmian monoid $\mathcal{M}$ can be represented by 3×3 matrices with non-negative entries. Although closely related to the classical representation by incidence matrices, this representation introduced in [LPS23] has an important property of faithfulness, i.e., distinct morphisms are represented by distinct matrices. We assign the following matrices to the morphisms $G, \widetilde{G}, D, \widetilde{D}$:

$$R_G = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},\ R_{\widetilde{G}} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},\ R_D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix},\ R_{\widetilde{D}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{5.3}$$

These matrices preserve the presentation of the special Sturmian monoid (5.1)

$$R_{\widetilde{G}} R_{\widetilde{D}}^k R_G = R_G R_D^k R_{\widetilde{G}} \quad \text{and} \quad R_{\widetilde{D}} R_{\widetilde{G}}^k R_D = R_D R_G^k R_{\widetilde{D}}. \tag{5.4}$$

Following the notation of [LPS23], we denote by $\mathcal{E}$ the monoid generated by the matrices (5.3)

$$\mathcal{E} = \langle R_G, R_{\widetilde{G}}, R_D, R_{\widetilde{D}} \rangle. \tag{5.5}$$

We can assign a matrix to every element in the special Sturmian monoid $\mathcal{M}$.

Definition 5.2.1. *Let $\mathcal{R} : \mathcal{M} \mapsto \mathbb{Z}^{3\times 3}$ be defined for $\psi \in \mathcal{M}$ by*

$$\mathcal{R}(\psi) = R_{\psi_0} R_{\psi_1} \cdots R_{\psi_n},$$

where $\psi = \psi_0 \circ \psi_1 \circ \cdots \circ \psi_n$ and $\psi_i \in \{G, \widetilde{G}, D, \widetilde{D}\}$, for every $i \in [\![0, n]\!]$.

Definition 5.2.1 is correct because, thanks to Equations (5.1) and (5.4), it does not depend on the decomposition of ψ into the elements $\{G, \widetilde{G}, D, \widetilde{D}\}$. We summarize the main results proved in [LPS23] in the following theorem. We recall the notation $\mathrm{Sl}(\mathbb{Z}, n) = \{R \in \mathbb{Z}^{n\times n} : \det R = 1\}$.

Theorem 5.2.2. [LPS23] *The monoid $\mathcal{M}$ and $\mathcal{E}$ are connected in the following way.*

i) Let $R \in \mathrm{Sl}(\mathbb{Z}, 3)$. Then $R \in \mathcal{E}$ if and only if there exist $a, b, c, d, e, f \in \mathbb{N}$ such that

$$R = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix}$$

and

$$ad - bc = 1, \tag{5.6}$$

$$e < a + c, \quad f < b + d \tag{5.7}$$

$$-c \leq cf - de < d\,. \tag{5.8}$$

ii) The monoid $\mathcal{E}$ coincides with $\mathcal{R}(\mathcal{M}) = \left\langle R_G, R_{\widetilde{G}}, R_D, R_{\widetilde{D}} \right\rangle$.

iii) The map $\mathcal{R}$ is a bijection $\mathcal{R} : \mathcal{M} \to \mathcal{E}$.

Remark 5.2.3. *Observe that the 2×2 matrix in the upper left corner of the matrix R_G (resp., $R_{\widetilde{G}}$, R_D, $R_{\widetilde{D}}$) is the incidence matrix of the morphism G (resp., $\widetilde{G}$, D, $\widetilde{D}$); see Equation* (5.3).

We can derive from Theorem 5.2.2 and Remark 5.2.3 that, for a given morphism $\psi \in \mathcal{M}$, the matrix $\mathcal{R}(\psi)$ is of the form

$$\mathcal{R}(\psi) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix},$$

where $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ is the incidence matrix of ψ and $a, b, c, d, e, f \in \mathbb{N}$ satisfy Equations (5.6), (5.7) and (5.8). Note that sometimes we refer to the parameters a, b, c, d, e, f of a morphism ψ as to a_ψ, b_ψ, c_ψ, d_ψ, e_ψ and f_ψ. Also, we have the following simple observation, which follows from the fact that the incidence matrices of the morphisms D and $\widetilde{D}$ coincide, and the incidence matrices of the morphisms G and $\widetilde{G}$ coincide.

Remark 5.2.4. *Let $\eta, \psi \in \mathcal{M}$ be such that $\eta = \eta_0\eta_1 \cdots \eta_k$ and $\psi = \psi_0\psi_1 \cdots \psi_k$ with $\eta_i, \psi_i \in \{G, \widetilde{G}, D, \widetilde{D}\}$, for every $i \in [\![0, k]\!]$. Moreover, assume that for every $i \in [\![0, k]\!]$ it holds that $\psi_i \in \{G, \widetilde{G}\}$ if and only if $\eta_i \in \{G, \widetilde{G}\}$. Then η and ψ have the same incidence matrix.*

We determine the faithful representation of a morphism fixing the Fibonacci word $\mathbf{f}$.

Example 5.2.5. The Fibonacci word $\mathbf{f}$ is the fixed point of the Fibonacci substitution φ; see Example 5.1.1. Clearly, for every $k \geq 1$, the morphism φ^k fixes $\mathbf{f}$. The Fibonacci substitution $\varphi \in \mathrm{St}$ is not an element of $\mathcal{M}$, however, $\varphi^2 = G \circ D \in \mathcal{M}$. Thus, we can determine

$$\mathcal{R}(\varphi^2) = \mathcal{R}(G)\mathcal{R}(D) = R_G R_D = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

In other words, $e_{\varphi^2} = 1$ and $f_{\varphi^2} = 0$. Indeed, the morphism $\varphi^2 : 0 \mapsto 010,\ 1 \mapsto 01$ has the incidence matrix $\left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ and, denoting $a = a_{\varphi^2}, b = b_{\varphi^2}, c = c_{\varphi^2}, d = d_{\varphi^2}, e = e_{\varphi^2}, f = f_{\varphi^2}$, we have

$$e = 1 < 2 + 1 = a + c, \qquad f = 0 < 1 + 1 = b + d, \qquad -c = -1 \leq cf - de < 1 = d.$$

Let us stress the advantage of the representation $\mathcal{R}$. We already mentioned that, for example, the generators G and $\widetilde{G}$ have the same incidence matrix. Similarly, any mutually conjugate morphisms share the same incidence matrix. The representation $\mathcal{R}$, on the other hand, assigns distinct matrices to distinct morphisms. However, there is also a disadvantage of the representation $\mathcal{R}$. Given a morphism ψ, it is not that easy to determine the parameters $e_\psi, f_\psi \in \mathbb{N}$. In the next part, a relatively simple algorithm to determine e_ψ, f_ψ is introduced.

5.3 Algorithm to determine the faithful representation

We start this part by illustrating the relations between the faithful representations of some mutually conjugate morphisms on an example.

Example 5.3.1 (Continuation of Example 5.1.7)**.** We use the notation from Example 5.1.7. We have that $\eta^{(0)} = \widetilde{G}\widetilde{D}$, $\eta^{(1)} = G\widetilde{D}$, $\eta^{(2)} = \widetilde{G}D$, $\eta^{(3)} = GD$ and therefore

$$\mathcal{R}(\eta^{(0)}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad \mathcal{R}(\eta^{(1)}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{R}(\eta^{(2)}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 2 & 1 & 1 \end{pmatrix}, \quad \mathcal{R}(\eta^{(3)}) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

Observe that denoting $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ the incidence matrix of the morphisms $(\eta^{(i)})_{i=0}^{3}$, we have

$$\begin{aligned}
e_{\eta^{(0)}} &= 1 = (a - 1 + 0a) \bmod (a + c), & f_{\eta^{(0)}} &= 1 = (b + 0b) \bmod (b + d), \\
e_{\eta^{(1)}} &= 0 = (a - 1 + 1a) \bmod (a + c), & f_{\eta^{(1)}} &= 0 = (b + 1b) \bmod (b + d), \\
e_{\eta^{(2)}} &= 2 = (a - 1 + 2a) \bmod (a + c), & f_{\eta^{(2)}} &= 1 = (b + 2b) \bmod (b + d), \\
e_{\eta^{(3)}} &= 1 = (a - 1 + 3a) \bmod (a + c), & f_{\eta^{(3)}} &= 0 = (b + 3b) \bmod (b + d).
\end{aligned}$$

In the following theorem, we extend the observation stated at the end of Example 5.3.1 to all elements in the special Sturmian monoid.

Theorem 5.3.2. *Let $(\eta^{(i)})_{i=0}^{L-1}$ be a chain of conjugate Sturmian morphisms. Denote $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}(\mathbb{N}, 2)$ their incidence matrix. Then, for every $i \in [\![0, L-1]\!]$,*

$$e_{\eta^{(i)}} = (a(i+1) - 1) \bmod (a + c) \quad \text{and} \quad f_{\eta^{(i)}} = (b(i+1)) \bmod (b + d).$$

To prove Theorem 5.3.2, we show some properties of the elements of the monoid $\mathcal{E}$. We note that, in particular, the identity matrix belongs to all four cases of Lemma 5.3.3.

Lemma 5.3.3. *Let* $R = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{smallmatrix}\right) \in \mathcal{E}$.

i) If $R \in \langle R_{\widetilde{D}}, R_{\widetilde{G}} \rangle$, *then* $e = a - 1$ *and* $f = b$;

ii) if $R \in \langle R_{\widetilde{D}}, R_G \rangle$, *then* $e = 0$ *and* $f = 0$;

iii) if $R \in \langle R_D, R_{\widetilde{G}} \rangle$, *then* $e = a + c - 1$ *and* $f = b + d - 1$;

iv) if $R \in \langle R_D, R_G \rangle$, *then* $e = c$ *and* $f = d - 1$.

Proof. We can check that all the cases i), ii), iii) and iv) are true for the identity matrix.

i) The statement holds for $R \in \{R_{\widetilde{D}}, R_{\widetilde{G}}\}$. We show that if $R \in \mathcal{E}$ fulfills the statement, then $RR_{\widetilde{D}}$ and $RR_{\widetilde{G}}$ fulfill the statement. Indeed,

$$RR_{\widetilde{D}} = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ a-1 & b & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a+b & b & 0 \\ c+d & d & 0 \\ a+b-1 & b & 1 \end{smallmatrix}\right) \quad \text{and} \quad RR_{\widetilde{G}} = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ a-1 & b & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & a+b & 0 \\ c & c+d & 0 \\ a-1 & a+b & 1 \end{smallmatrix}\right).$$

ii) The statement holds for $R \in \{R_{\widetilde{D}}, R_G\}$. We show that if $R \in \mathcal{E}$ fulfills the statement, then $RR_{\widetilde{D}}$ and RR_G fulfill the statement. Indeed,

$$RR_{\widetilde{D}} = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a+b & b & 0 \\ c+d & d & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) \quad \text{and} \quad RR_G = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & a+b & 0 \\ c & c+d & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right).$$

iii) The statement holds for $R \in \{R_D, R_{\widetilde{G}}\}$. We show that if $R \in \mathcal{E}$ fulfills the statement, then RR_D and $RR_{\widetilde{G}}$ fulfill the statement. Indeed,

$$RR_D = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ a+c-1 & b+d-1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a+b & b & 0 \\ c+d & d & 0 \\ a+b+c+d-1 & b+d-1 & 1 \end{smallmatrix}\right)$$

and

$$RR_{\widetilde{G}} = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ a+c-1 & b+d-1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & a+b & 0 \\ c & c+d & 0 \\ a+c-1 & a+b+c+d-1 & 1 \end{smallmatrix}\right).$$

iv) The statement holds for $R \in \{R_D, R_G\}$. We show that if $R \in \mathcal{E}$ fulfills the statement, then RR_D and RR_G fulfill the statement. Indeed,

$$RR_D = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ c & d-1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a+b & b & 0 \\ c+d & d & 0 \\ c+d & 0 & 1 \end{smallmatrix}\right) \quad \text{and} \quad RR_G = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ c & d-1 & 1 \end{smallmatrix}\right) \left(\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a & a+b & 0 \\ c & c+d & 0 \\ c & c+d-1 & 1 \end{smallmatrix}\right). \quad \square$$

We illustrate Lemma 5.3.3 on an example.

Example 5.3.4 (Continuation of Example 5.3.1)**.** We use the notation from Example 5.3.1. We check easily that as $\eta^{(0)} = \widetilde{G}\widetilde{D} \in \langle \widetilde{D}, \widetilde{G} \rangle$, we have that $\mathcal{R}(\eta^{(0)}) \in \langle R_{\widetilde{D}}, R_{\widetilde{G}} \rangle$ and $\mathcal{R}(\eta^{(0)})$ satisfies the condition that

$$e = a - 1 = 2 - 1 = 1 \quad \text{and} \quad f = b = 1,$$

which is in correspondence with Lemma 5.3.3 i). Similarly, it is possible to show that the parameters of the morphism $\eta^{(1)} = G\widetilde{D}$ are in correspondence with Lemma 5.3.3 ii), the parameters of the morphism $\eta^{(2)} = \widetilde{G}D$ are in correspondence with Lemma 5.3.3 iii), and the parameters of the morphism $\eta^{(3)} = GD$ are in correspondence with Lemma 5.3.3 iv).

We prove the following proposition.

Proposition 5.3.5. *Let $(\eta^{(i)})_{i=0}^{L-1}$ be a chain of conjugate Sturmian morphisms. Denote $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{Sl}(\mathbb{N}, 2)$ their incidence matrix. Then, for every $i \in [\![0, L-2]\!]$, we have*

$$e_{\eta^{(i+1)}} = (e_{\eta^{(i)}} + a) \bmod (a+c) \quad \textit{and} \quad f_{\eta^{(i+1)}} = (f_{\eta^{(i)}} + b) \bmod (b+d).$$

Proof. Let $i \in [\![0, L-2]\!]$. It follows from Algorithm 5.1.8 that there exist unique morphisms $\mu \in \langle D, G\rangle$, $\tilde{\xi} \in \{\widetilde{D}, \widetilde{G}\}$ and $\psi \in \langle D, G, \widetilde{D}, \widetilde{G}\rangle$ such that $\eta^{(i)} = \mu\tilde{\xi}\psi$. We denote

$$\mathcal{R}(\mu) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ c & d-1 & 1 \end{pmatrix} \text{ and } \mathcal{R}(\psi) = \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix},$$

where we used Lemma 5.3.3 iv) to obtain that $e_\mu = c$ and $f_\mu = d-1$.

(I) Assume $\tilde{\xi} = \widetilde{G}$. Then it follows from Algorithm 5.1.8 that, denoting $\tilde{\mu} = \mathrm{sub}_{D \mapsto \widetilde{D}}(\mu)$ and $\xi = G$, we have $\eta^{(i+1)} = \tilde{\mu}\xi\psi$. As $\tilde{\mu} \in \langle \widetilde{D}, G\rangle$, it holds by Lemma 5.3.3 ii) and Remark 5.2.4 that $\mathcal{R}(\tilde{\mu}) = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{smallmatrix}\right)$. Then we have

$$\mathcal{R}(\eta^{(i)}) = \mathcal{R}(\mu\tilde{\xi}\psi) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ c & d-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix} = \begin{pmatrix} aA+(a+b)C & aB+(a+b)D & 0 \\ cA+(c+d)C & cB+(c+d)D & 0 \\ cA+(c+d)C+E & cB+(c+d)D+F & 1 \end{pmatrix}$$

and

$$\mathcal{R}(\eta^{(i+1)}) = \mathcal{R}(\tilde{\mu}\xi\psi) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix} = \begin{pmatrix} aA+(a+b)C & aB+(a+b)D & 0 \\ cA+(c+d)C & cB+(c+d)D & 0 \\ E & F & 1 \end{pmatrix}.$$

Consequently, we obtain

$$\begin{aligned} e_{\eta^{(i+1)}} = E &= ((a+c)A + (a+b+c+d)C + E) \bmod ((a+c)A + (a+b+c+d)C) \\ &= (e_{\eta^{(i)}} + a_{\eta^{(i)}}) \bmod (a_{\eta^{(i)}} + c_{\eta^{(i)}}) \end{aligned}$$

and

$$\begin{aligned} f_{\eta^{(i+1)}} = F &= ((a+c)B + (a+b+c+d)D + F) \bmod ((a+c)B + (a+b+c+d)D) \\ &= (f_{\eta^{(i)}} + b_{\eta^{(i)}}) \bmod (b_{\eta^{(i)}} + d_{\eta^{(i)}}), \end{aligned}$$

where we used the pair of inequalities

$$\begin{aligned} E < A + C &\le (a+c)A + (a+b+c+d)C, \\ F < B + D &\le (a+c)B + (a+b+c+d)D. \end{aligned}$$

(II) Assume $\tilde{\xi} = \widetilde{D}$. Then it follows from Algorithm 5.1.8 that, denoting $\tilde{\mu} = \mathrm{sub}_{G \mapsto \widetilde{G}}(\mu)$ and $\xi = D$, we have $\eta^{(i+1)} = \tilde{\mu}\xi\psi$. As $\tilde{\mu} \in \langle D, \widetilde{G}\rangle$, it holds by Lemma 5.3.3 iii) and Remark 5.2.4 that $\mathcal{R}(\tilde{\mu}) = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ a+c-1 & b+d-1 & 1 \end{smallmatrix}\right)$. Then we have

$$\mathcal{R}(\eta^{(i)}) = \mathcal{R}(\mu\tilde{\xi}\psi) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ c & d-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix} = \begin{pmatrix} (a+b)A+bC & (a+b)B+bD & 0 \\ (c+d)A+dC & (c+d)B+dD & 0 \\ (c+d-1)A+(d-1)C+E & (c+d-1)B+(d-1)D+F & 1 \end{pmatrix}$$

and

$$\mathcal{R}(\eta^{(i+1)}) = \mathcal{R}(\tilde{\mu}\xi\psi) = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ a+c-1 & b+d-1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} A & B & 0 \\ C & D & 0 \\ E & F & 1 \end{pmatrix} = \begin{pmatrix} (a+b)A+bC & (a+b)B+bD & 0 \\ (c+d)A+dC & (c+d)B+dD & 0 \\ \begin{matrix}(a+b+c+d-1)A \\ +(b+d-1)C+E\end{matrix} & \begin{matrix}(a+b+c+d-1)B \\ +(b+d-1)D+F\end{matrix} & 1 \end{pmatrix}.$$

Consequently, we obtain

$$\begin{aligned}e_{\eta^{(i+1)}} &= (a+b+c+d-1)A + (b+d-1)C + E\\ &= ((a+b+c+d-1)A + (b+d-1)C + E) \bmod ((a+b+c+d)A + (b+d)C)\\ &= (e_{\eta^{(i)}} + a_{\eta^{(i)}}) \bmod (a_{\eta^{(i)}} + c_{\eta^{(i)}})\end{aligned}$$

and

$$\begin{aligned}f_{\eta^{(i+1)}} &= (a+b+c+d-1)B + (b+d-1)D + F\\ &= ((a+b+c+d-1)B + (b+d-1)D + F) \bmod ((a+b+c+d)B + (b+d)D)\\ &= (f_{\eta^{(i)}} + b_{\eta^{(i)}}) \bmod (b_{\eta^{(i)}} + d_{\eta^{(i)}}),\end{aligned}$$

where we used $E < A + C$ and $F < B + D$. $\square$

We turn to the proof of Theorem 5.3.2.

Proof of Theorem 5.3.2. Let $(\eta^{(i)})_{i=0}^{L-1}$ be a chain of conjugate Sturmian morphisms. Denote $\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) \in \mathrm{Sl}(\mathbb{N}, 2)$ their incidence matrix. We carry out the proof by finite induction.

Assume $i = 0$. Then $\eta^{(0)} \in \langle \widetilde{D}, \widetilde{G}\rangle$ and we have from Lemma 5.3.3 i) that

$$e_{\eta^{(0)}} = a - 1 = (a-1) \bmod (a+c) \quad \text{and} \quad f_{\eta^{(0)}} = b = b \bmod (b+d).$$

Induction hypothesis: for an integer $i \in [\![1, L-1]\!]$, it holds that $e_{\eta^{(i-1)}} = (ai - 1) \bmod (a+c)$ and $f_{\eta^{(i-1)}} = bi \bmod (b+d)$. Then from Proposition 5.3.5, we have

$$\begin{aligned}e_{\eta^{(i)}} &= (e_{\eta^{(i-1)}} + a) \bmod (a+c)\\ &= (a(i+1) - 1) \bmod (a+c)\end{aligned}$$

and

$$\begin{aligned}f_{\eta^{(i)}} &= (f_{\eta^{(i-1)}} + b) \bmod (b+d)\\ &= (b(i+1)) \bmod (b+d). \quad \square\end{aligned}$$

We illustrate how Proposition 5.3.5 eases the determination of the faithful representation of a given morphism.

Example 5.3.6. Determining the faithful representation of the morphism $\eta = DGG\widetilde{D}G$ based solely on Definition 5.2.1 requires multiplying 5 matrices to obtain

$$\begin{aligned}\mathcal{R}(\eta) &= R_D R_G R_G R_{\widetilde{D}} R_G\\ &= \begin{pmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & 0 & 0\\ 1 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix} 1 & 1 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\\ &= \begin{pmatrix} 3 & 5 & 0\\ 4 & 7 & 0\\ 3 & 5 & 1\end{pmatrix}.\end{aligned}$$

Now, we determine the faithful representation with the knowledge of Algorithm 5.1.8, Lemma 5.3.3 and Proposition 5.3.5. The incidence matrix of the morphism is obtained by multiplying 2×2 matrices, giving $\left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 & 5\\ 4 & 7\end{smallmatrix}\right)$. Denote $i \in \mathbb{N}$ the integer such that $\eta = \eta^{(i)}$ in the chain of Sturmian morphisms. From Algorithm 5.1.8, we notice that $\eta^{(i-1)} = \widetilde{D}GG\widetilde{D}G$. From Lemma 5.3.3, we have $e_{\eta^{(i-1)}} = 0$ and $f_{\eta^{(i-1)}} = 0$. From Proposition 5.3.5, we have $e_{\eta^{(i)}} = 3$ and $f_{\eta^{(i)}} = 5$.

Let us remark that it is possible to assign a 3×3 matrix with integer entries to the morphism $E\colon 0 \mapsto 1,\, 1 \mapsto 0$. However, this matrix clearly cannot have nonnegative entries, because $E^2 = \mathrm{Id}$. As the proofs in [LPS23] exploit Perron–Frobenius theorem, which is applicable for matrices with nonnegative entries, we did not explore this direction. We claim, however, that the natural choice for the representation of the morphism E is

$$R_E = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & -1 \end{pmatrix}.$$

Then it is possible to extend the representation $\mathcal{R}$ of the special Sturmian monoid $\mathcal{M}$ to a representation of the Sturmian monoid St. For more details, see [LPS23, Remark 10].

We finish this part with an open question. As we observed in Chapter 4, Sturmian words present a binary case of words on d-ary alphabets which are called episturmian and there is a monoid of morphisms that map episturmian words to episturmian words [JP02]. It seems natural to ask whether it is possible to faithfully represent this monoid.

Question 5.3.7. *Can we faithfully represent the monoid of morphisms which map episturmian words to episturmian words?*

5.4 Intercepts of fixed points of conjugate Sturmian morphisms

In Section 4.1.1, we explained the description of Sturmian words as irrational mechanical words, irrational rotations on a circle and cutting sequences. Sturmian words can also be described with the help of two interval exchange transformation and it is this description, which was used to derive the faithful representation of special Sturmian monoid.

The *two interval exchange transformation* (2iet) generalizes the rotation on a circle. For given parameters $\ell_0, \ell_1 > 0$, consider an interval I_0 of length ℓ_0 and an interval I_1 of length ℓ_1. Sometimes, it is convenient that the intervals are left-closed right-open

$$I_0 = [0, \ell_0) \text{ and } I_1 = [\ell_0, \ell_0 + \ell_1),$$

whereas, sometimes, the opposite case is considered. In both cases, we define the map $T : I_0 \cup I_1 \to I_0 \cup I_1$ as

$$T(x) = \begin{cases} x + \ell_1, & \text{if } x \in I_0; \\ x - \ell_0, & \text{if } x \in I_1. \end{cases}$$

Setting an initial point $\rho \in I_0 \cup I_1$, the word $\mathbf{u} = u_0 u_1 u_2 \cdots \in \{0,1\}^{\mathbb{N}}$ defined by the relation

$$u_n = \begin{cases} 0, & \text{if } T^n(\rho) \in I_0; \\ 1, & \text{if } T^n(\rho) \in I_1, \end{cases}$$

is a *2iet word* with the *vector of parameters* $\vec{v}(\mathbf{u}) = (\ell_0, \ell_1, \rho)$.

We show a connection between 2iet words and mechanical words; see Section 4.1.1.

Lemma 5.4.1. *A 2iet word* $\mathbf{u}$ *with the parameters* (ℓ_0, ℓ_1, ρ) *coincides with the lower mechanical word* $\mathbf{s}_{\alpha,\delta}$ *with the slope* $\alpha = \frac{\ell_1}{\ell_0+\ell_1}$ *and the intercept* $\delta = \frac{\rho}{\ell_0+\ell_1}$.

Proof. Let $\mathbf{u}$ be a 2iet word with the vector of parameters $\vec{v} = (\ell_0, \ell_1, \rho)$ and let $I_0 = [0, \ell_0)$, $I_1 = [\ell_0, \ell_0 + \ell_1)$ be half-open intervals of the same type (the procedure would be analogous for left-open right-closed intervals). For every $n \in \mathbb{N}$, we have that

$$T^n(\rho) \in I_0 \quad \text{if and only if} \quad T^n(\tfrac{1}{\ell_0+\ell_1}\rho) \in \tfrac{1}{\ell_0+\ell_1} I_0.$$

Consequently, the 2iet word $\mathbf{v}$ with the rescaled parameters $\frac{1}{\ell_0+\ell_1}(\ell_0, \ell_1, \rho)$ satisfies the condition that $\mathbf{v} = \mathbf{u}$. Simultaneously, the word $\mathbf{v}$ fulfills the condition that the length of the interval $I_0 \cup I_1$ is equal to 1 and therefore the map T coincides with the rotation by the angle $\alpha = \frac{\ell_1}{\ell_0+\ell_1}$; compare with Equation (4.4). As a result, the 2iet word $\mathbf{u}$ is the lower mechanical word $\mathbf{s}_{\alpha,\delta}$ with the slope $\alpha = \frac{\ell_1}{\ell_0+\ell_1}$ and the intercept $\delta = \frac{\rho}{\ell_0+\ell_1}$. $\square$

We observe that Lemma 5.4.1 enables us to describe Sturmian words as 2iet words with a vector of parameters closely related to the slope and intercept of the word interpreted as a mechanical word. The following lemma, which is closely related to the Perron–Frobenius theorem, was proved in [LPS23, Corollary 9].

Lemma 5.4.2. [LPS23] *Let $\psi \in \mathcal{M}$ be a primitive morphism. The matrix $\mathcal{R}(\psi)$ has eigenvalues Λ, 1 and $\frac{1}{\Lambda}$, where $\Lambda > 1$ is a quadratic unit. An eigenvector corresponding to Λ can be found in the form $(x, y, z) \in \left(\mathbb{Q}(\Lambda)\right)^3$ with $x > 0$, $y > 0$ and $z \geq 0$. No other eigenvalue has an eigenvector with the first two components positive.*

We call Λ from Lemma 5.4.2 the dominant eigenvalue. The following proposition gives a simple criterion on whether a Sturmian word is fixed by a certain morphism from the special Sturmian monoid based on its vector of parameters.

Proposition 5.4.3. [LPS23] *Let $\psi \in \mathcal{M}$ be a primitive morphism and $\mathbf{u}$ be a Sturmian word with the vector of parameters $\vec{v}(\mathbf{u})$. The word $\mathbf{u}$ is fixed by ψ if and only if $\vec{v}(\mathbf{u})$ is an eigenvector to the dominant eigenvalue of $\mathcal{R}(\psi)$.*

We illustrate Proposition 5.4.3 on an example.

Example 5.4.4. The Fibonacci word $\mathbf{f}$ is a Sturmian word with the slope $\alpha = \frac{1}{\tau^2}$ and the intercept $\delta = \frac{1}{\tau^2}$, where the term slope refers to the slope of a mechanical word; see Example 4.1.3. We denote its vector of parameters $\vec{v}(\mathbf{f}) = (\ell_0, \ell_1, \rho)$. From Lemma 5.4.1, we have that $\frac{1}{\tau^2} = \frac{\ell_1}{\ell_0+\ell_1} = \frac{\rho}{\ell_0+\ell_1}$. Therefore, we have $\vec{v}(\mathbf{f}) = (\tau, 1, 1)$. In Example 5.2.5, we observed that the word $\mathbf{f}$ is fixed by the morphism $GD \in \mathcal{M}$ and

$$\mathcal{R}(GD) = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

We see that $\vec{v}(\mathbf{f}) = (\tau, 1, 1)$ is the eigenvector to the dominant eigenvalue $\Lambda = \tau^2$:

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} \tau \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2\tau + 1 \\ \tau + 1 \\ \tau + 1 \end{pmatrix} = \begin{pmatrix} \tau^2 + \tau \\ \tau^2 \\ \tau^2 \end{pmatrix} = \tau^2 \begin{pmatrix} \tau \\ 1 \\ 1 \end{pmatrix}.$$

The following lemma follows easily from Proposition 5.4.3.

Lemma 5.4.5. *Let $\psi \in \mathcal{M}$ be a primitive morphism and $\mathbf{u}$ be a Sturmian word with the vector of parameters $\vec{v}(\mathbf{u}) = (\ell_0, \ell_1, \rho)$, which is fixed by ψ. Let $\Lambda > 1$ denote the dominant eigenvalue of $\mathcal{R}(\psi) = \left(\begin{smallmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{smallmatrix}\right)$. Then we have*

$$\rho = \frac{e\ell_0 + f\ell_1}{\Lambda - 1}.$$

Proof. Using Proposition 5.4.3, we have that

$$\begin{pmatrix} a\ell_0 + b\ell_1 \\ c\ell_0 + d\ell_1 \\ e\ell_0 + f\ell_1 + \rho \end{pmatrix} = \begin{pmatrix} a & b & 0 \\ c & d & 0 \\ e & f & 1 \end{pmatrix} \begin{pmatrix} \ell_0 \\ \ell_1 \\ \rho \end{pmatrix} = \Lambda \begin{pmatrix} \ell_0 \\ \ell_1 \\ \rho \end{pmatrix},$$

which implies that $e\ell_0 + f\ell_1 + \rho = \Lambda\rho$, and therefore $\rho = \frac{e\ell_0 + f\ell_1}{\Lambda - 1}$. □

Combining Lemma 5.4.5 with Theorem 5.3.2, we can compute the intercepts of the Sturmian words fixed by conjugate Sturmian morphisms, which we illustrate on the following example.

Example 5.4.6. Let $\eta^{(0)} = \widetilde{D}\widetilde{G}\widetilde{G}\widetilde{D}\widetilde{G}$ and denote $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} 3 & 5 \\ 4 & 7 \end{smallmatrix}\right)$ its incidence matrix. Then $||\eta^{(0)}|| = 19$ and there are $||\eta^{(0)}|| - 1 = 18$ morphisms in the equivalence class of $\eta^{(0)}$, including itself. We denote $(\eta^{(i)})_{i=0}^{17}$ the chain of Sturmian morphisms produced by Algorithm 5.1.8 such that $\eta^{(0)} = \widetilde{D}\widetilde{G}\widetilde{G}\widetilde{D}\widetilde{G}$. For every $i \in [\![0, 16]\!]$, we determine the parameters $(e_{\eta^{(i)}}, f_{\eta^{(i)}})$ from Theorem 5.3.2; see Table 5.2. We denote $e_i = e_{\eta^{(i)}}$ and $f_i = f_{\eta^{(i)}}$ for simplification.

We have that, for every $i \in [\![0, 17]\!]$, the faithful representation

$$\mathcal{R}(\eta^{(i)}) = \begin{pmatrix} 3 & 5 & 0 \\ 4 & 7 & 0 \\ e_i & f_i & 1 \end{pmatrix}$$

has the same dominant eigenvalue $\Lambda = 5 + 2\sqrt{6}$. Denote $v_i = (\ell_0, \ell_1, \rho_i)$ the corresponding eigenvector. Choosing $\ell_0 = 1$, we have $\ell_1 = \frac{\Lambda - 3}{5}$ and we use Lemma 5.4.5 to obtain

$$\rho_i = \frac{e_i + f_i\ell_1}{\Lambda - 1}.$$

Consequently, we determine the value of the intercept δ_i from Lemma 5.4.1 as the rescaled value $\delta_i = \frac{\rho_i}{1+\ell_1}$. The approximate values are shown in Table 5.2.

5.5 Square roots of fixed points of Sturmian morphisms

We recall that a factor $v \in \{0,1\}^+$ of the form $v = ww = w^2$ for some $w \in \{0,1\}^+$ is called a *square*. A square is called *minimal* if none of its proper prefixes is a square. Saari [Saa10] showed that, for every Sturmian word $\mathbf{u}$, there exist 6 minimal squares $v_1, \ldots, v_6$ such that

$$\mathbf{u} = w_{i_1}^2 w_{i_2}^2 w_{i_3}^2 \cdots, \tag{5.9}$$

where $v_i = w_i^2$, for every $i \in [\![1, 6]\!]$, and $i_k \in [\![1, 6]\!]$, for every $k \in \mathbb{N}$. Inspired by this result, Peltomäki and Whiteland introduced the square root $\sqrt{\mathbf{u}}$ of the Sturmian word $\mathbf{u}$ as

$$\sqrt{\mathbf{u}} = w_{i_1} w_{i_2} w_{i_3} \cdots$$

and they proved that if $\mathbf{u}$ is a Sturmian word with the slope α and the intercept δ, then the word $\sqrt{\mathbf{u}}$ is a Sturmian word with the same slope α and the intercept $\frac{1-\alpha+\delta}{2}$ [PW17].

i	$\eta^{(i)}$	$(e_{\eta^{(i)}}, f_{\eta^{(i)}})$	ρ_i	δ_i
17	$DGGDG$	$(4,6)$	1.37980	0.57980
16	$\widetilde{D}GGDG$	$(1,1)$	0.26742	0.11237
15	$DG\widetilde{G}DG$	$(5,8)$	1.80227	0.75732
14	$\widetilde{D}G\widetilde{G}DG$	$(2,3)$	0.68990	0.28990
13	$D\widetilde{G}\widetilde{G}DG$	$(6,10)$	2.22475	0.93485
12	$DGG\widetilde{D}G$	$(3,5)$	1.11237	0.46742
11	$\widetilde{D}GG\widetilde{D}G$	$(0,0)$	0	0
10	$DGGD\widetilde{G}$	$(4,7)$	1.53485	0.64495
9	$\widetilde{D}GGD\widetilde{G}$	$(1,2)$	0.42247	0.17753
8	$DG\widetilde{G}D\widetilde{G}$	$(5,9)$	1.95732	0.82247
7	$\widetilde{D}G\widetilde{G}D\widetilde{G}$	$(2,4)$	0.84495	0.35505
6	$D\widetilde{G}\widetilde{G}D\widetilde{G}$	$(6,11)$	2.37980	1
5	$DGG\widetilde{D}\widetilde{G}$	$(3,6)$	1.26742	0.53258
4	$\widetilde{D}GG\widetilde{D}\widetilde{G}$	$(0,1)$	0.15505	0.06515
3	$DG\widetilde{G}\widetilde{D}\widetilde{G}$	$(4,8)$	1.68990	0.71010
2	$\widetilde{D}G\widetilde{G}\widetilde{D}\widetilde{G}$	$(1,3)$	0.57753	0.24268
1	$D\widetilde{G}\widetilde{G}\widetilde{D}\widetilde{G}$	$(5,10)$	2.11237	0.88763
0	$\widetilde{D}\widetilde{G}\widetilde{G}\widetilde{D}\widetilde{G}$	$(2,5)$	1	0.42020

Table 5.2: A chain of Sturmian morphisms $(\eta^{(i)})_{i=0}^{17}$ such that $\eta^{(0)} = \widetilde{D}\widetilde{G}\widetilde{G}\widetilde{D}\widetilde{G}$. The dominant eigenvalue of $\mathcal{R}(\eta^{(i)})$ for every $i \in [\![0, 17]\!]$ is $\Lambda = 5 + 2\sqrt{6}$ and the vector of parameters $\vec{v}(\mathbf{u}) = (\ell_0, \ell_1, \rho)$ has values $\ell_0 = 1$ and $\ell_1 = \frac{2(\sqrt{6}+1)}{5}$.

Example 5.5.1. We show the square root of the Fibonacci word $\mathbf{f}$. As we can write

$$\begin{array}{rcllllllllllll} \mathbf{f} & = & 010 & 010 & 100 & 100 & 10 & 10 & 01 & 01 & 0 & 0 & 10010 & 10010 & \cdots \\ & = & w_1 & w_1 & w_2 & w_2 & w_3 & w_3 & w_4 & w_4 & w_5 & w_5 & w_6 & w_6 & \cdots, \end{array}$$

the square root of the Fibonacci word is defined as the word

$$\begin{array}{rcllllllll} \sqrt{\mathbf{f}} & = & 010 & 100 & 10 & 01 & 0 & 10010 & \cdots \\ & = & w_1 & w_2 & w_3 & w_4 & w_5 & w_6 & \cdots. \end{array}$$

It follows from [PW17] that the word $\sqrt{\mathbf{f}}$ is Sturmian with the slope $\alpha = \frac{1}{\tau^2}$ and the intercept $\delta = \frac{1}{2}$. Note that all square roots of characteristic words have the intercept $\delta = \frac{1}{2}$.

Additional properties were proved in [LPS23] for those square roots of Sturmian words, which emerged from the fixed points of elements in the special Sturmian monoid.

Theorem 5.5.2. [LPS23] *Let $\mathbf{u} \in \{\mathtt{0}, \mathtt{1}\}^{\mathbb{N}}$ be a Sturmian word fixed by a primitive morphism $\psi \in \mathcal{M}$. The square root $\sqrt{\mathbf{u}}$ is fixed by a morphism $\widetilde{\psi} \in \mathcal{M}$, which is a conjugate of one of the morphisms ψ, ψ^2, ψ^3 or ψ^4.*

With the knowledge of some details from the proof of Theorem 5.5.2, it is possible to precisely determine the minimal power $k \in \{1, 2, 3, 4\}$ such that the square root is fixed by a conjugate of the k-th power of the original morphism. We summarize important steps from the proof of Theorem 5.5.2 in Remark 5.5.3.

Remark 5.5.3. *Let $\psi \in \mathcal{M}$ be a primitive morphism and denote M its incidence matrix. In the proof of Theorem 5.5.2, we search for an integer $k \in \mathbb{N}$, such that the term $\sum_{i=0}^{k-1} M^i$ has all elements even. As $M \in \mathrm{Sl}(\mathbb{N}, 2)$, it is equal* mod 2 *to one of the following 6 matrices*

$$\begin{aligned}
(a):& \quad \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \\
(b):& \quad \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \\
(c):& \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.
\end{aligned}$$

It follows that there exists $k \in \{2, 3, 4\}$ such that the term $\sum_{i=0}^{k-1} M^i$ has all elements even. We choose

$$k = \begin{cases} 2, & \text{if } M \text{ is of type (a);} \\ 3, & \text{if } M \text{ is of type (b);} \\ 4, & \text{if } M \text{ is of type (c).} \end{cases}$$

Example 5.5.4. The morphism $\psi = D\widetilde{G}$ has the incidence matrix $M = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 2 \end{smallmatrix}\right)$. As M is of type (b), the smallest possible integer $k \leq 4$ such that a conjugate of ψ^k fixes $\sqrt{\mathbf{u}}$ is $k = 3$.

The morphism $\psi = DG\widetilde{G}$ has the incidence matrix $M = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 3 \end{smallmatrix}\right)$. As M is of type (c), the smallest possible integer $k \leq 4$ such that a conjugate of ψ^k fixes $\sqrt{\mathbf{u}}$ is $k = 4$.

In the special case of characteristic Sturmian words, it suffices to search for $k \in \{1, 2, 3\}$ and the desired morphism fixing the square root has palindromic letter images of odd length.

Corollary 5.5.5. [LPS23] *Let $\mathbf{u} \in \{0, 1\}^{\mathbb{N}}$ be a characteristic Sturmian word fixed by a primitive morphism $\psi \in \mathcal{M}$ having the incidence matrix M. Let k be the minimal positive integer such that $(1, 1)M^k = (1, 1)$* mod 2*. Then $k \leq 3$ and the square root $\sqrt{\mathbf{u}}$ is fixed by a morphism $\widetilde{\psi} \in \mathcal{M}$, which is a conjugate of ψ^k. Moreover, $\widetilde{\psi}(0)$ and $\widetilde{\psi}(1)$ are palindromes of odd length.*

We can deduce even more from Corollary 5.5.5. As every pair of morphisms which are conjugate satisfies the condition that their fixed points are distinct, there is exactly one morphism $\widetilde{\psi} \in \mathcal{M}$ which is a conjugate of ψ^k and which fixes the square root $\sqrt{\mathbf{u}} = \widetilde{\psi}(\sqrt{\mathbf{u}})$. Thus, if we find a morphism η which is a conjugate of ψ^k and which has palindromic images of letters, then η fixes the square root $\sqrt{\mathbf{u}}$. We illustrate Corollary 5.5.5 on several examples.

Example 5.5.6. Let $\psi = DG^2 : 0 \mapsto 10,\ 1 \mapsto 10101$ be an element of the special Sturmian monoid $\mathcal{M}$. The fixed point $\mathbf{u} = \psi(\mathbf{u})$ can be expressed as a concatenation of the squares of the factors 10, 1, 0110101, 101, 01, 01101.

$$\mathbf{u} = \overline{1010}\,\underline{11}\,\overline{0101}\,\underline{01101010110101}\,\overline{1010}\,\underline{101101}\,\overline{0101}\,\underline{1010}\,\overline{101101}\cdots$$

The incidence matrix of ψ is $M = \left(\begin{smallmatrix} 1 & 2 \\ 1 & 3 \end{smallmatrix}\right)$. We have $(1, 1)M = (2, 5) \neq (1, 1)$ mod 2. Then we see that $(1, 1)M^2 = (7, 19) = (1, 1)$ mod 2. Hence, it suffices to find the palindromic conjugate of ψ^2. The square root

$$\sqrt{\mathbf{u}} = \overline{10}\,\underline{1}\,\overline{01}\,\underline{0110101}\,\overline{10}\,\underline{101}\,\overline{01}\,\underline{10}\,\overline{101}\,0110101011010110101 01\cdots$$

is fixed by the palindromic conjugate of ψ^2

$$0 \mapsto 1010101, 1 \mapsto 1010101101011010101.$$

To illustrate Corollary 5.5.5 more thoroughly, we show other examples, where a conjugate of ψ, ψ^2 or ψ^3 is used to fix the corresponding square root.

ψ	k	conjugate of ψ^k
D^2G^2	1	$0 \mapsto 101, 1 \mapsto 1011101$
GDG	1	$0 \mapsto 010, 1 \mapsto 01010$
D^2G	2	$0 \mapsto 10111011101, 1 \mapsto 101110111011101$
DG	3	$0 \mapsto 1010110110101, 1 \mapsto 1010110110101101101$
GD	3	$0 \mapsto 0101001001010010010, 1 \mapsto 0101001001010$

Chapter 6

Dumont–Thomas numeration systems for $\mathbb{Z}$

In this chapter, we mostly summarize our results [LL23b], which are currently under review in Integers – Electronic Journal of Combinatorial Number Theory. We generalize certain substitution-based numeration systems for $\mathbb{N}$ to $\mathbb{Z}$. Let us remark that substitutions form a subset of morphisms. We refer to the newly defined numeration systems as to the Dumont–Thomas numeration systems for $\mathbb{Z}$ and we describe their properties with respect to automata and a particular total order. As a matter of fact, this total order coincides with the one presented in Chapter 3. Also, we show how these numeration systems can be extended naturally to $\mathbb{Z}^d$, for $d \geq 1$. In particular, we recover the two's complement notation and the Fibonacci complement numeration system, which we studied in a different context in Chapter 3. As a new result not included in [LL23b], we provide a sufficient condition for the Dumont–Thomas numeration systems for $\mathbb{Z}$ to be positional.

6.1 Dumont–Thomas numeration systems for $\mathbb{N}$

Numeration systems for representing nonnegative integers, as well as real numbers in a certain interval, were introduced in [DT89] by Dumont and Thomas. They are based on right-infinite fixed points of substitutions and we refer to them as to the Dumont–Thomas numeration systems for $\mathbb{N}$. Note that, as opposed to the previous chapters, in this chapter we distinguish between right-infinite, left-infinite and two-sided fixed points. Also, we only treat substitutions, which are a special case of morphisms. We illustrate the main idea of admissible sequences, which form a base for the Dumont–Thomas numeration systems for $\mathbb{N}$, on the following example.

Definition 6.1.1 ([DT89], admissible sequence)**.** *Let $\eta : A^* \to A^*$ be a substitution. Let $a \in A$ be a letter, k an integer and, for each integer i, $0 \leq i \leq k$, (m_i, a_i) be an element of $A^* \times A$. We say that the finite sequence $(m_i, a_i)_{i=0,\ldots,k}$ is* admissible with respect to η *if and only if, for all i, $1 \leq i \leq k$, $m_{i-1}a_{i-1}$ is a prefix of $\eta(a_i)$. We say that this sequence is a-*admissible with respect to η *if it is admissible with respect to η and, moreover, $m_k a_k$ is a prefix of $\eta(a)$.*

Example 6.1.2. Consider the substitution $\eta : a \mapsto abc$, $b \mapsto baa$, $c \mapsto cbb$ and its right-infinite fixed point $\mathbf{u} = \eta(\mathbf{u}) = \lim_{k\to+\infty} \eta^k(a)$. The first 4 images of a under the substitution η are illustrated in a tree in Figure 6.1. The letters in the tree are connected by lines labeled by `0`, `1` or `2` according to the rule $d \xrightarrow{i} e$ if and only if $\eta(d)$ has letter e at position i, for every $d, e \in \{a, b, c\}$

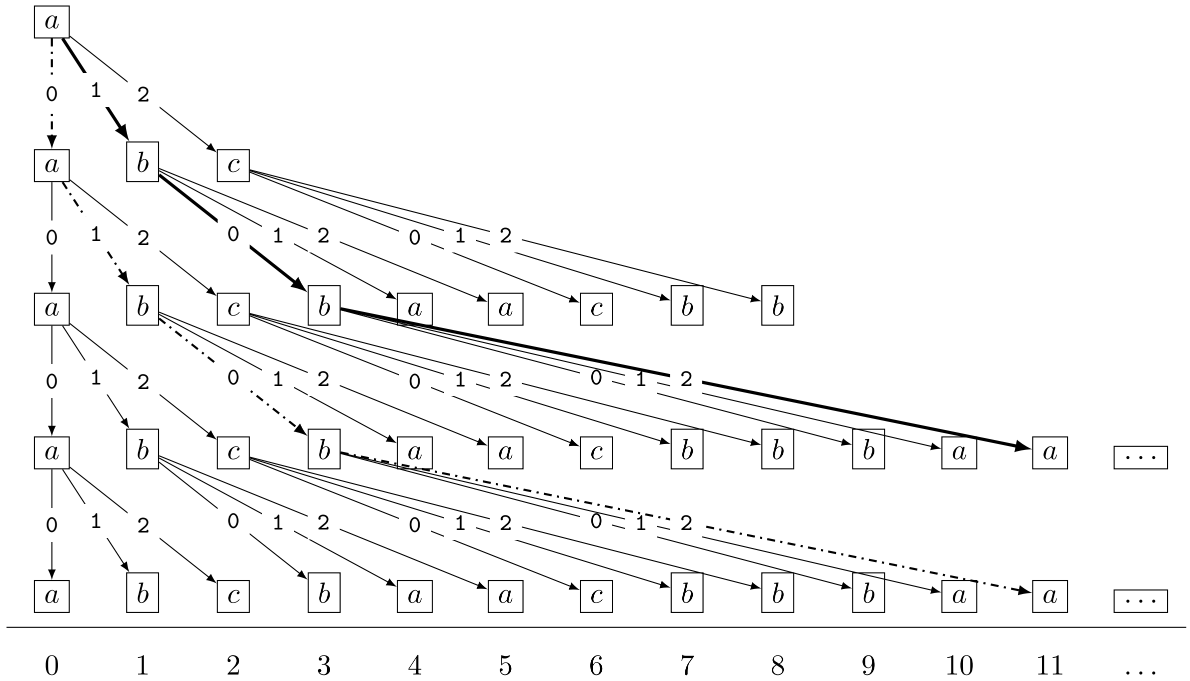


Figure 6.1: Illustration of a-admissible sequences of length up to 4 for the substitution $a \mapsto abc, b \mapsto baa, c \mapsto cbb$.

and every $i \in \{\mathtt{0}, \mathtt{1}, \mathtt{2}\}$. The lines connecting letters form paths, such as the path shown in bold

$$a \xrightarrow{\mathtt{1}} b \xrightarrow{\mathtt{0}} b \xrightarrow{\mathtt{2}} a,$$

which starts at the top of the tree and ends with a letter a at position 11 in the right-infinite fixed point $\mathbf{u}$. Using the notion of admissible sequences, we can interpret this path as the sequence $(m_i, a_i)_{i=0,\dots,2}$ with the terms

$$(m_2, a_2) = (a, b),$$
$$(m_1, a_1) = (\varepsilon, b),$$
$$(m_0, a_0) = (ba, a).$$

We note immediately that, starting again at the top of the tree, the path $a \xrightarrow{\mathtt{0}} a \xrightarrow{\mathtt{1}} b \xrightarrow{\mathtt{0}} b \xrightarrow{\mathtt{2}} a$ shown in dash-dotted style also brings us to the letter a at position 11, and so does the path $b \xrightarrow{\mathtt{2}} a$, which does not start at the top of the tree. Simply speaking, paths in the tree illustrate admissible sequences and those paths which start at the top of the tree illustrate a-admissible sequences.

We saw in Example 6.1.2 that every letter at position $n \in \mathbb{N}$ in the fixed point is reached by multiple paths. However, every position $n \geq 1$ is reached by exactly one path, which starts at the top of the tree and which does not start with $\mathtt{0}$. This is the idea of the following theorem proved by Dumont and Thomas, which enabled them to represent all nonnegative integers in a unique way.

Theorem 6.1.3. [DT89, Theorem 1.5] *Let $a \in A$ and let $\eta : A^* \to A^*$ be a substitution. Let $\mathbf{u} = \eta(\mathbf{u})$ be a right-infinite fixed point of η with a growing letter $u_0 = a$. For every integer $n \geq 1$, there exists a unique integer $k = k(n)$ and a unique sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that*

- *this sequence is a-admissible and* $m_{k-1} \neq \varepsilon$,
- $u_0 u_1 \cdots u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2}) \cdots \eta^0(m_0)$.

From Theorem 6.1.3, given a right-infinite fixed point of a substitution η with a growing letter $a \in A$, every integer $n \geq 1$ is represented with a unique a-admissible sequence. We refer to the map which assigns the unique a-admissible sequence to every position $n \geq 1$ and the empty word to the position $n = 0$ as to a Dumont–Thomas numeration system for $\mathbb{N}$.

Definition 6.1.4 (Numeration system for $\mathbb{N}$)**.** *Let $a \in A$ and let $\eta : A^* \to A^*$ be a substitution. Let $\mathbf{u} = \eta(\mathbf{u})$ be a right-infinite fixed point of η with a growing letter $u_0 = a$. Let $\mathcal{D} = \{0, ..., \max_{c \in A} |\eta(c)| - 1\}$. We define*

$$\begin{array}{rccl} \mathrm{rep}_{\eta,a} : & \mathbb{N} & \to & \mathcal{D}^* \\ & n & \mapsto & \begin{cases} |m_{k-1}| \cdot |m_{k-2}| \cdot \ldots \cdot |m_0|, & \text{if } n \geq 1; \\ \varepsilon, & \text{if } n = 0; \end{cases} \end{array}$$

where $k = k(n)$ is the unique integer and $(m_i, a_i)_{i=0,...,k-1}$ is the unique sequence obtained from Theorem 6.1.3.

Remark 6.1.5. *The notation $(\cdot)$ throughout this text stands for the concatenation of words within the monoid $\mathcal{D}^*$, not the multiplication of integers.*

Let $K \geq 2$ be an integer. Dumont–Thomas numeration systems for $\mathbb{N}$ based on K-uniform substitutions coincide with the classical K-ary numeration systems, which we illustrate in the following example.

Example 6.1.6 (Continuation of Example 6.1.2)**.** Using the greedy algorithm, the ternary representation of the number 11 is $\mathrm{rep}_3(11) = \mathtt{102}$ as

$$\begin{aligned} 11 &= 9 + 2 \\ &= \mathtt{1} \cdot 3^2 + \mathtt{0} \cdot 3^1 + \mathtt{2} \cdot 3^0. \end{aligned}$$

Now, we represent $n = 11$ in the Dumont–Thomas numeration system for $\mathbb{N}$ based on the 3-uniform substitution $\eta : a \mapsto abc$, $b \mapsto baa$, $c \mapsto cbb$. We apply Theorem 6.1.3 on the prefix of length 11 of the right-infinite fixed point $u = \eta(u)$ with a growing letter a. We obtain

$$\begin{aligned} abcbaacbbba &= abcbaacbb \cdot \varepsilon \cdot ba \\ &= \eta^2(a) \cdot \eta^1(\varepsilon) \cdot \eta^0(ba). \end{aligned}$$

Hence we have $\mathrm{rep}_{\eta,a}(11) = |a| \cdot |\varepsilon| \cdot |ba| = \mathtt{102}$.

Note that as $m_{k-1} \neq \varepsilon$ in Theorem 6.1.3, we have that $|m_{k-1}| \neq \mathtt{0}$ and thus the Dumont–Thomas representations do not start with leading zeroes. This is why, from now on, we draw the paths starting with leading zeroes in the figures illustrating a-admissible sequences in a dashed style; see Figure 6.2. We show a Dumont–Thomas numeration system for $\mathbb{N}$ based on the Tribonacci substitution, which is not K-uniform.

Example 6.1.7. Consider the Tribonacci substitution $\Theta : a \mapsto ab$, $b \mapsto ac$, $c \mapsto a$ and its right-infinite fixed point $\mathbf{t} = \Theta(\mathbf{t}) = \lim_{k \to +\infty} \Theta^k(a)$. We show the first 3 images of a under the substitution Θ in a tree in Figure 6.2 and we show the representations of small integers in the Dumont–Thomas numeration system associated with Θ and a in the following table.

n	0	1	2	3	4	5	6
$\mathrm{rep}_{\Theta,a}(n)$	ε	1	10	11	100	101	110

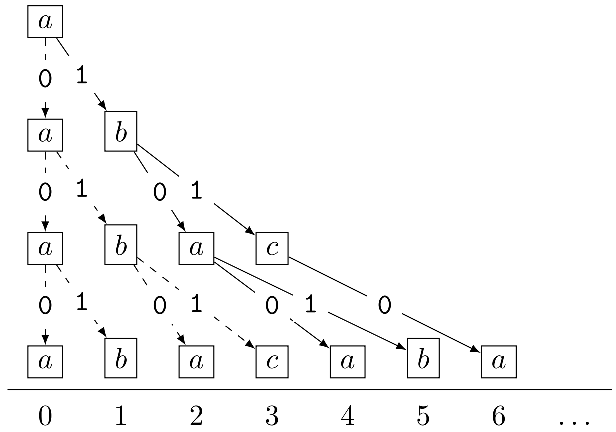


Figure 6.2: Illustration of all a-admissible sequences of length up to 3 for the Tribonacci substitution.

Let us stress the fact that, as opposed to the numeration systems in Chapter 3, the Dumont–Thomas numeration systems for $\mathbb{N}$ might not be positional, which we show in the following example.

Example 6.1.8. We slightly modify the Tribonacci substitution from Example 6.1.7. Let η be the substitution $\eta\colon a \mapsto aba$, $b \mapsto ac$, $c \mapsto a$. We observe that a is a growing letter as $\eta(a)$ starts with a and $|\eta(a)| \geq 2$. In analogy with Example 6.1.7, we show the first 2 images of the growing letter a in a tree; see Figure 6.3. By contradiction, we assume that there exists a positional numeration system $U = (U_n)_{n=0}^{+\infty}$ such that, for every $n \in \mathbb{N}$, the representation $\mathrm{rep}_{\eta,a}(n) = w_{k-1} \cdots w_0 \in \{\mathtt{0},\mathtt{1},\mathtt{2}\}^*$ is mapped to the correct position n by the value map $\mathrm{val}_U\colon w \mapsto \sum_{i=0}^{k-1} w_i U_i$; see Equation (3.1). The representations of the first five positions are:

$$\mathrm{rep}_{\eta,a}(1) = \mathtt{1}, \quad \mathrm{rep}_{\eta,a}(2) = \mathtt{2}, \quad \mathrm{rep}_{\eta,a}(3) = \mathtt{10}, \quad \mathrm{rep}_{\eta,a}(4) = \mathtt{11}, \quad \mathrm{rep}_{\eta,a}(5) = \mathtt{20}.$$

Clearly, $U_0 = 1$ and we have $\mathrm{val}_U(\mathtt{1}) = 1$ and $\mathrm{val}_U(\mathtt{2}) = 2$. As $\mathrm{rep}_{\eta,a}(3) = \mathtt{10}$, we have

$$\begin{aligned} 3 &= \mathrm{val}_U(\mathtt{10}) \\ &= 1 \cdot U_1 + 0 \cdot U_0 \end{aligned}$$

and therefore $U_1 = 3$. Consequently, evaluating the representations of $n \in \{3,4,5\}$ gives $\mathrm{val}_U(\mathtt{10}) = 3$, $\mathrm{val}_U(\mathtt{11}) = 4$, $\mathrm{val}_U(\mathtt{20}) = 6$. This is a contradiction, as $\mathrm{rep}_{\eta,a}(5) = \mathtt{20}$ and

$$\begin{aligned} 5 &\neq \mathrm{val}_U(\mathrm{rep}_{\eta,a}(5)) \\ &= \mathrm{val}_U(\mathtt{20}) \\ &= 2 \cdot U_1 + 0 \cdot U_0 \\ &= 6. \end{aligned}$$

6.1.1 Some extensions of Dumont–Thomas results

In this section, we provide some lemmas [LL23b], which extend the results of Dumont and Thomas. Also, we show that the Dumont–Thomas numeration systems for $\mathbb{N}$ fall into the framework of abstract numeration systems [BR10].

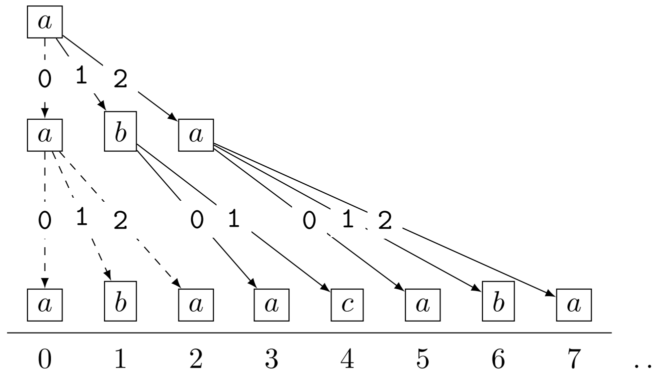


Figure 6.3: Illustration of admissible sequences of length up to 2 associated with the substitution $a \mapsto aba$, $b \mapsto ac$, $c \mapsto a$ and letter a.

First, we recall that we can associate an automaton with every substitution [BR10]. Let $\mathcal{D}$ denote the alphabet $\mathcal{D} = \{0, ..., \max_{c \in A} |\eta(c)| - 1\}$. The deterministic finite automaton with output associated with a substitution $\eta\colon A^* \to A^*$ and a letter $a \in A$ is the 6-tuple $\mathcal{A}_{\eta,a} = (A, \mathcal{D}, \delta, a, B, \xi)$, where the transition function $\delta : A \times \mathcal{D} \to A$ is a partial function such that $\delta(b, i) = c$ if and only if $c = w_i$ and $\eta(b) = w_0 \cdots w_{|\eta(b)|-1}$, B is an alphabet and $\xi : A \to B$ is the output function. We restrict ourselves to the case where $B = A$ and ξ is the identity map. In Figure 6.4, we show the automata associated with a 2-uniform substitution, the Fibonacci substitution φ and the Tribonacci substitution Θ.

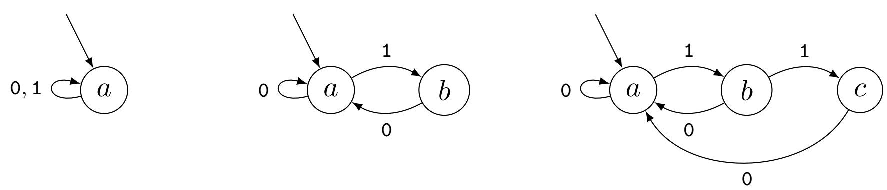


Figure 6.4: The automata $\mathcal{A}_{\omega,a}$, $\mathcal{A}_{\varphi,a}$, $\mathcal{A}_{\Theta,a}$ associated with the 2-uniform substitution $\omega : a \mapsto aa$, the Fibonacci substitution $\varphi\colon a \mapsto ab$, $b \mapsto a$ and the Tribonacci substitution $\Theta : a \mapsto ab$, $b \mapsto ac$, $c \mapsto a$.

We observe that, for every letter $c \in A$, the c-admissible sequences are related to the automaton $\mathcal{A}_{\eta,c}$ in the following way.

Lemma 6.1.9. [LL23b, Lemma 3.6] *Let $\eta : A^* \to A^*$ be a substitution, $k \geq 1$ be an integer and $c \in A$. If $(m_i, a_i)_{i=0,\ldots,k-1}$ is a c-admissible sequence, then, for every $i \in [\![0, k-1]\!]$,*

$$a_i = \mathcal{A}_{\eta,c}(|m_{k-1}| \cdot |m_{k-2}| \cdot \ldots \cdot |m_i|).$$

We can reformulate Lemma 6.1.9 in the following way: a concatenation of the lengths of m_i's of every c-admissible sequence is a finite word accepted by the automaton $\mathcal{A}_{\eta,c}$. The following lemma states that, on the other hand, every finite word accepted by the automaton $\mathcal{A}_{\eta,c}$ is a concatenation of the lengths of m_i's of a c-admissible sequence.

Lemma 6.1.10. [LL23b, Lemma 3.8] *Let $\eta : A^* \to A^*$ be a substitution, $k \geq 1$ be an integer and $c \in A$. If $v_{k-1}v_{k-2} \cdots v_0 \in \mathcal{L}(\mathcal{A}_{\eta,c})$, then there exists a c-admissible sequence $(m_i, a_i)_{i=0,\ldots,k-1}$*

such that, for every $i \in [\![0, k-1]\!]$, we have

$$|m_i| = v_i.$$

We can construct an admissible sequence from a prefix of the image of a letter under the p-th power of a substitution. Consequently, that allows us to define a map which we call the tail map.

Lemma 6.1.11. [LL23b, Lemma 3.9] *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. If $m \in A^*$ and $c \in A$ are such that m is a proper prefix of $\eta^p(c)$, then there exists a unique c-admissible sequence $(m_i, a_i)_{i=0,\dots,p-1}$ such that*

$$|m| = \textstyle\sum_{j=0}^{p-1} |\eta^j(m_j)|. \tag{6.1}$$

Moreover, $m = \eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2}) \cdots \eta^0(m_0)$.

Definition 6.1.12 (tail map)**.** *Let $\eta : A^* \to A^*$ be a substitution and $\mathcal{D} = \{0, ..., \max_{c \in A} |\eta(c)| - 1\}$. Let $p \geq 1$ be an integer. We define a map*

$$\begin{array}{rccl} \mathrm{tail}_{\eta,p,c} : & [\![0, |\eta^p(c)| - 1]\!] & \to & \mathcal{D}^p \\ & n & \mapsto & |m_{p-1}| \cdot |m_{p-2}| \cdot \ldots \cdot |m_0|, \end{array}$$

where $(m_i, a_i)_{i=0,\dots,p-1}$ is the unique c-admissible sequence satisfying Equation (6.1) *with m being the prefix of $\eta^p(c)$ of length n.*

Example 6.1.13 (Continuation of Example 6.1.7)**.** We return to Figure 6.2, where we illustrate the successive images of a under the Tribonacci substitution Θ. The path from the top of the tree to a node of length p reaching a position $n \in \mathbb{N}$ is labeled by $\mathrm{tail}_{\Theta,p,a}(n)$. Their values are shown in the following table.

n	$\mathrm{tail}_{\Theta,1,a}(n)$	$\mathrm{tail}_{\Theta,2,a}(n)$	$\mathrm{tail}_{\Theta,3,a}(n)$
0	0	00	000
1	1	01	001
2		10	010
3		11	011
4			100
5			101
6			110

We observe that considering that the alphabet $\{0, 1\}$ is totally ordered by $0 < 1$, we have that $\mathrm{tail}_{\Theta,3,a}(0) = 000$ is lexicographically smaller than $\mathrm{tail}_{\Theta,3,a}(1) = 001$.

An important property of the tail map is related to the lexicographical order as we observed in Example 6.1.13.

Lemma 6.1.14. [LL23b, Lemma 3.12] *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. Let $c \in A$. Let $n, n' \in [\![0, |\eta^p(c)| - 1]\!]$. Then*

(i) $n = n'$ if and only if $\mathrm{tail}_{\eta,p,c}(n) = \mathrm{tail}_{\eta,p,c}(n')$,

(ii) $n < n'$ if and only if $\mathrm{tail}_{\eta,p,c}(n) <_{lex} \mathrm{tail}_{\eta,p,c}(n')$.

In the next lemma, we describe the relationship between the tail map and the automaton associated with a substitution.

Lemma 6.1.15. [LL23b, Lemma 3.11] *Let $\eta : A^* \to A^*$ be a substitution and $p \geq 1$ be an integer. Let $c \in A$. Then, for every $\ell \in [\![0, |\eta^p(c)| - 1]\!]$, we have*

$$\eta^p(c)[\ell] = \mathcal{A}_{\eta,c}(\mathrm{tail}_{\eta,p,c}(\ell)).$$

We clarify that a Dumont–Thomas numeration system for $\mathbb{N}$ is an abstract numeration system. An abstract numeration system is defined as a triple $\mathcal{S} = (L, A, <)$ where L is an infinite regular language over a totally ordered alphabet $(A, <)$. The map $\mathrm{rep}_{\mathcal{S}} : \mathbb{N} \to L$ is the bijection mapping $n \in \mathbb{N}$ to the nth word in radix order in L. Therefore, for every $m, n \in \mathbb{N}$, it holds that $m < n$ if and only if $\mathrm{rep}_{\mathcal{S}}(m) <_{\mathrm{rad}} \mathrm{rep}_{\mathcal{S}}(n)$ [BR10].

Lemma 6.1.16. *Let $\eta : A^* \to A^*$ be a substitution with a growing letter $a \in A$. Let $\mathcal{D} = \{0, ..., \max_{c\in A} |\eta(c)| - 1\}$. Let $L = \mathcal{L}(\mathcal{A}_{\eta,a}) \setminus \mathtt{0}\mathcal{D}^*$. Then $\mathcal{S} = (L, \mathcal{D}, <)$ is an abstract numeration system and* $\mathrm{rep}_{\eta,a} = \mathrm{rep}_{\mathcal{S}}$.

Proof. Let $\eta : A^* \to A^*$ be a substitution and $a \in A$. Setting $L = \mathcal{L}(\mathcal{A}_{\eta,c}) \setminus \mathtt{0}A^*$, we have that L is an infinite regular language. Therefore $(L, \mathcal{D}, <)$ is an abstract numeration system. We prove that $\mathrm{rep}_{\eta,a} : \mathbb{N} \to L$ is an increasing bijection with respect to the radix order.

Let $n, n' \in \mathbb{N}$ be such that $n < n'$. Applying Theorem 6.1.3 on n (resp., n') we obtain a unique integer k (resp., k') and a unique a-admissible sequence $(m_i, a_i)_{i=0,\dots,k-1}$ (resp., $(m'_i, a'_i)_{i=0,\dots,k'-1}$). From Lemma 6.1.11, we deduce that

$$\mathrm{tail}_{\eta,k,a}(n) = |m_{k-1}| \cdot |m_{k-2}| \cdot \ldots \cdot |m_0|,$$
$$\mathrm{tail}_{\eta,k',a}(n') = |m'_{k'-1}| \cdot |m'_{k'-2}| \cdot \ldots \cdot |m'_0|,$$

and $k \leq k'$. If $k < k'$, then $|\mathrm{tail}_{\eta,k,a}(n)| < |\mathrm{tail}_{\eta,k',a}(n')|$, and thus $|\mathrm{rep}_{\eta,a}(n)| < |\mathrm{rep}_{\eta,a}(n')|$. If $k = k'$, then from Lemma 6.1.14, $\mathrm{tail}_{\eta,k,a}(n) <_{lex} \mathrm{tail}_{\eta,k,a}(n')$ and thus $\mathrm{rep}_{\eta,a}(n) <_{lex} \mathrm{rep}_{\eta,a}(n')$. In both cases, $\mathrm{rep}_{\eta,a}(n) <_{\mathrm{rad}} \mathrm{rep}_{\eta,a}(n')$. Consequently, $\mathrm{rep}_{\eta,a}$ is increasing with respect to the radix order and thus it is injective.

It remains to prove that $\mathrm{rep}_{\eta,a}$ is surjective. Let $v \in \mathcal{L}(\mathcal{A}_{\eta,a}) \setminus \mathtt{0}\mathcal{D}^*$ be of length $k \in \mathbb{N}$. Applying Lemma 6.1.10 on v, we obtain an a-admissible sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that $m_i = v_i$, for every $i \in [\![0, k-1]\!]$. Set $n = \sum_{j=0}^{k-1} |\eta^j(m_j)|$. Using Lemma 6.1.11, we have that $u_0 u_1 \cdots u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2}) \cdots \eta^0(m_0)$. Moreover, $v_{k-1} \neq \mathtt{0}$ and thus $m_{k-1} \neq \varepsilon$. From Theorem 6.1.3, $\mathrm{rep}_{\eta,a}(n) = v$. □

As a consequence, we recover [BR10, Proposition 3.4.12], which we reformulate in the following corollary: the automaton $\mathcal{A}_{\eta,a}$ associated with a substitution η and a growing letter a produces the right-infinite fixed point $\lim_{k\to+\infty} \eta^k(a)$ if it is fed gradually with representations of $n \in \mathbb{N}$ in the Dumont–Thomas numeration system for $\mathbb{N}$. For more details, see the proof of [BR10, Proposition 3.4.12].

Corollary 6.1.17. *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} = \eta(\mathbf{u})$ be a right-infinite fixed point of η with a growing letter $u_0 = a$. Then, for every $n \in \mathbb{N}$,*

$$u_n = \mathcal{A}_{\eta,a}(\mathrm{rep}_{\eta,a}(n)).$$

6.2 Dumont–Thomas numeration systems for $\mathbb{Z}$

In [LL23b], we extended the idea of Dumont–Thomas numeration systems to $\mathbb{Z}$. The following two theorems generalize Theorem 6.1.3 for the right-infinite and left-infinite periodic points of substitutions with a growing letter.

Theorem 6.2.1. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution with a growing letter $a \in A$. Let $\mathbf{u} \in \mathrm{Per}_{\mathbb{N}}(\eta)$ such that $u_0 = a$. Let $p \geq 1$ be a period of $\mathbf{u}$. For every integer $n \geq 1$, there exists a unique integer $k = k(n)$ such that p divides k and a unique sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that*

(i) this sequence is a-admissible and $m_{k-1}m_{k-2}\cdots m_{k-p} \neq \varepsilon$,

(ii) $u_0u_1\cdots u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0)$.

Theorem 6.2.2. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution with a growing letter $b \in A$. Let $\mathbf{u} \in \mathrm{Per}_{\mathbb{Z}_{<0}}(\eta)$ such that $u_{-1} = b$. Let $p \geq 1$ be a period of u. For every integer $n \leq -2$, there exists a unique integer $k = k(n)$ such that p divides k and a unique sequence $(m_i, a_i)_{i=0,\dots,k-1}$ such that*

(i) this sequence is b-admissible and

$$\eta^{p-1}(m_{k-1})\eta^{p-2}(m_{k-2})\cdots\eta^0(m_{k-p})a_{k-p} \neq \eta^p(b), \tag{6.2}$$

(ii) $u_{-|\eta^k(b)|}\cdots u_{n-2}u_{n-1} = \eta^{k-1}(m_{k-1})\eta^{k-2}(m_{k-2})\cdots\eta^0(m_0)$.

With the help of Theorem 6.2.1 and Theorem 6.2.2, it is possible to define Dumont–Thomas numeration systems for $\mathbb{Z}$. Similarly to the Dumont–Thomas numeration systems for $\mathbb{N}$, these numeration systems may not be positional.

Definition 6.2.3 (Dumont–Thomas numeration systems for $\mathbb{Z}$)**.** *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $u_{-1}|u_0$. Let $p \geq 1$ be the period of $\mathbf{u}$. Let $\mathcal{D} = \{0, \dots, \max_{c\in A}|\eta(c)| - 1\}$. We define*

$$\begin{aligned}\mathrm{rep}_u :\ \mathbb{Z} &\to \{\mathtt{0},\mathtt{1}\}\mathcal{D}^*\\ n &\mapsto \begin{cases}\mathtt{0}\cdot|m_{k-1}|\cdot|m_{k-2}|\cdot\ldots\cdot|m_0|, & \text{if } n \geq 1;\\ \mathtt{0}, & \text{if } n = 0;\\ \mathtt{1}, & \text{if } n = -1;\\ \mathtt{1}\cdot|m_{k-1}|\cdot|m_{k-2}|\cdot\ldots\cdot|m_0|, & \text{if } n \leq -2,\end{cases}\end{aligned}$$

where $k = k(n) \geq 0$ is the unique integer and $(m_i, a_i)_{i=0,\dots,k-1}$ is the unique sequence obtained from Theorem 6.2.1 (Theorem 6.2.2) applied on the right-infinite periodic point $\mathbf{u}|_{\mathbb{N}}$ (on the left-infinite periodic point $\mathbf{u}|_{\mathbb{Z}_{<0}}$) if $n \geq 1$ (if $n \leq -2$, respectively) both with period p.

Note that the period $p \in \mathbb{N}$ of $\mathbf{u}$ divides $|\mathrm{rep}_{\mathbf{u}}(n)| - 1$ for every $n \in \mathbb{Z}$. Also, we observe that

$$\mathrm{rep}_{\mathbf{u}}(n) = \begin{cases}\mathtt{0}\cdot\mathrm{tail}_{\eta,k,u_0}(n), & \text{if } n \geq 0;\\ \mathtt{1}\cdot\mathrm{tail}_{\eta,k,u_{-1}}(n), & \text{if } n < 0.\end{cases}$$

We illustrate the Dumont–Thomas numeration systems for $\mathbb{Z}$ based on the Fibonacci and Tribonacci substitution, which are the common thread of this text. We denote $\mathbf{f}'$ the two-sided periodic point of the Fibonacci substitution with the growing seed $b|a$ and we denote $\mathbf{t}'$ the two-sided periodic point of the Tribonacci substitution with the growing seed $c|a$.

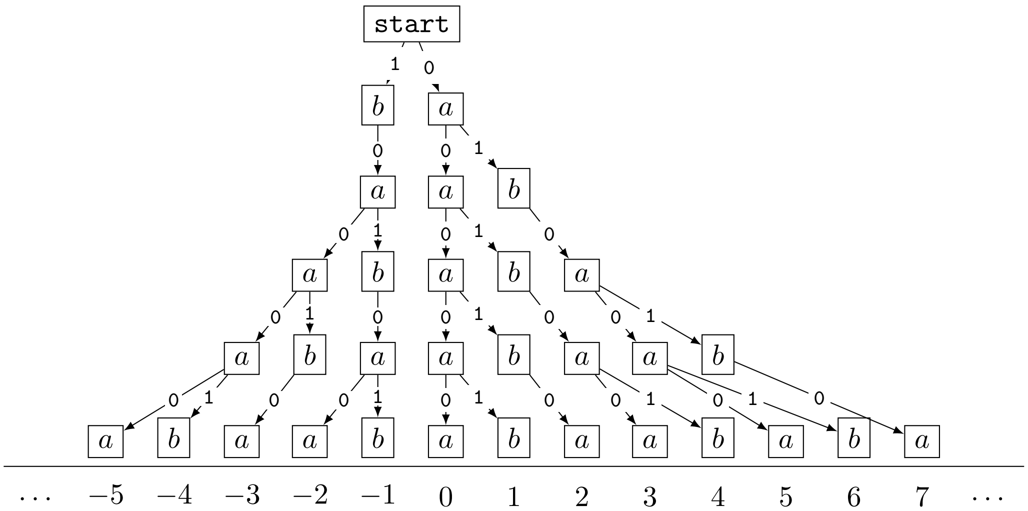


Figure 6.5: Illustration of representations of small integers in the Dumont–Thomas numeration system associated with the two-sided periodic point of the Fibonacci substitution φ and the growing seed $b|a$.

Example 6.2.4. Let $\varphi : a \mapsto ab$, $b \mapsto a$ be the Fibonacci substitution and let

$$\mathbf{f}' = \cdots abaab|abaababaab \cdots$$

be the two-sided periodic point of φ with the growing seed $b|a$. The representations of small integers based on $\mathbf{f}'$ are in Table 6.1. In Figure 6.5, a representation $\operatorname{rep}_{\mathbf{f}'}(n)$ of $n \in \mathbb{Z}$ labels the shortest path from the root of the tree to a node at position $n \in \mathbb{Z}$ such that $|\operatorname{rep}_{\mathbf{f}'}(n)| \bmod 2 = 1$.

n	$\operatorname{rep}_{\mathbf{f}'}(n)$	n	$\operatorname{rep}_{\mathbf{f}'}(n)$	n	$\operatorname{rep}_{\mathbf{f}'}(n)$
−10	1000100	0	0	10	0010010
−9	1000101	1	001	11	0010100
−8	1001000	2	010	12	0010101
−7	1001001	3	00100	13	0100000
−6	1001010	4	00101	14	0100001
−5	10000	5	01000	15	0100010
−4	10001	6	01001	16	0100100
−3	10010	7	01010	17	0100101
−2	100	8	0010000	18	0101000
−1	1	9	0010001	19	0101001

Table 6.1: The Dumont–Thomas numeration system for $\mathbb{Z}$ based on the two-sided point $\mathbf{f}'$ of the Fibonacci substitution φ with the growing seed $b|a$.

Example 6.2.5. The successive images of the growing seed $c|a$ under the Tribonacci substitution $\Theta : a \mapsto ab$, $b \mapsto ac$, $c \mapsto a$ are illustrated in Figure 6.6. We denote $\mathbf{t}' = \cdots abac|abacaba \cdots$ the two-sided periodic point of Θ of period 3 with the growing seed $c|a$. The representations of small integers based on the periodic point $\mathbf{t}'$ are in the following table.

n	−4	−3	−2	−1	0	1	2	3	4	5	6
$\operatorname{rep}_{\mathbf{t}'}(n)$	1000	1001	1010	1	0	0001	0010	0011	0100	0101	0110

In Figure 6.6, a representation $\operatorname{rep}_{\mathbf{t}'}(n)$ of $n \in \mathbb{Z}$ labels the shortest path from the root of the tree to a node at x-position $n \in \mathbb{Z}$ such that $|\operatorname{rep}_{\mathbf{t}'}(n)| \bmod 3 = 1$.

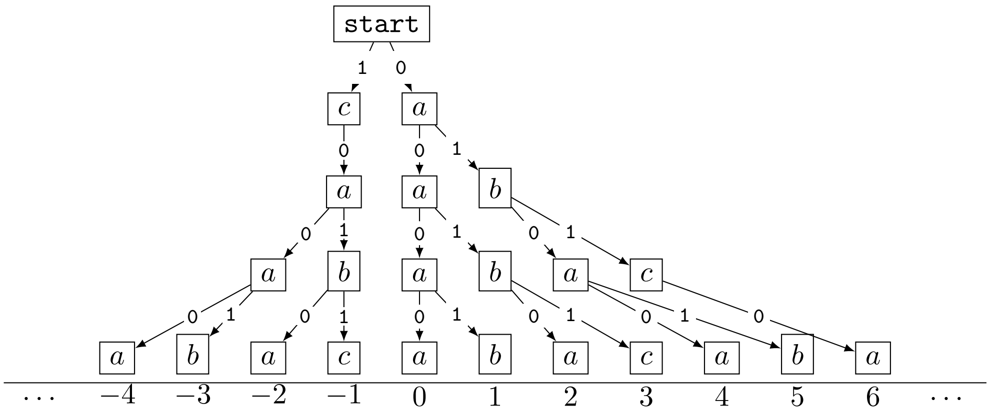


Figure 6.6: Illustration of representations of small integers in the Dumont–Thomas numeration system associated with the two-sided periodic point of the Tribonacci substitution Θ and the growing seed $c|a$.

Definition 6.2.6 (quotient, remainder)**.** *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Let $p \geq 1$ be the period of $\mathbf{u}$. Let $n \in \mathbb{Z} \setminus \{-1, 0\}$ be an integer and $k = k(n)$ be the unique integer and $(m_i, a_i)_{i=0,\ldots,k-1}$ be the unique sequence obtained from Theorem 6.2.1 (Theorem 6.2.2) applied on $\mathbf{u}|_{\mathbb{N}}$ ($\mathbf{u}|_{\mathbb{Z}_{<0}}$) if $n \geq 1$ (if $n \leq -2$, respectively) both with period p. We define the* $\mathbf{u}$-quotient *of n as*

$$q = \begin{cases} |\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2}) \cdots \eta^0(m_p)|, & \text{if } n \geq 1; \\ |\eta^{k-p-1}(m_{k-1})\eta^{k-p-2}(m_{k-2}) \cdots \eta^0(m_p)| - |\eta^{k-p}(u_{-1})|, & \text{if } n \leq -2; \end{cases}$$

and the $\mathbf{u}$-remainder *of n as $r = |\eta^{p-1}(m_{p-1})\eta^{p-2}(m_{p-2}) \cdots \eta^0(m_0)|$.*

The $\mathbf{u}$-quotient q and the $\mathbf{u}$-remainder r of an integer $n \in \mathbb{Z} \setminus \{-1, 0\}$ fulfill the condition that if $n \geq 1$ then $0 \leq q < n$ and if $n \leq -2$ then $-1 \geq q > n$. Consequently, $|q| < |n|$. The choice of the names stems from the fact that if η is K-uniform, then the $\mathbf{u}$-quotient and $\mathbf{u}$-remainder of n correspond to the quotient and remainder of the division of n by K^p. The $\mathbf{u}$-quotient and $\mathbf{u}$-remainder have the following particular role in the description of the substitutive structure of a two-sided periodic point $\mathbf{u}$.

Lemma 6.2.7. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed. Let $p \geq 1$ be the period of $\mathbf{u}$. Let $n \in \mathbb{Z} \setminus \{-1, 0\}$ be an integer. If $q \in \mathbb{Z}$ is the $\mathbf{u}$-quotient and $r \in \mathbb{N}$ is the $\mathbf{u}$-remainder of n, then*

$$u_n = \eta^p(u_q)[r] \quad \text{and} \quad \mathrm{rep}_{\mathbf{u}}(n) = \mathrm{rep}_{\mathbf{u}}(q) \cdot \mathrm{tail}_{\eta,p,u_q}(r).$$

Remark 6.2.8. *Let us note that Theorem 6.2.1 and Theorem 6.2.2 apply for every period $p \geq 1$ of the corresponding periodic point and not just the smallest one. Therefore we could define the Dumont–Thomas representations for $\mathbb{Z}$ using any period $p \geq 1$ of the corresponding periodic point. See Table 6.2 where we illustrate this in the case of the Fibonacci substitution and a period $p = 4$.*

n	$\mathrm{rep}_{\mathbf{f}',4}(n)$	n	$\mathrm{rep}_{\mathbf{f}',4}(n)$	n	$\mathrm{rep}_{\mathbf{f}',4}(n)$
-10	101000100	0	0	10	000010010
-9	101000101	1	00001	11	000010100
-8	101001000	2	00010	12	000010101
-7	101001001	3	00100	13	000100000
-6	101001010	4	00101	14	000100001
-5	10000	5	01000	15	000100010
-4	10001	6	01001	16	000100100
-3	10010	7	01010	17	000100101
-2	10100	8	000010000	18	000101000
-1	1	9	000010001	19	000101001

Table 6.2: The representations $\mathrm{rep}_{\mathbf{f}',4}(n)$ based on Theorem 6.2.1 (resp., Theorem 6.2.2) with the Fibonacci substitution φ, the growing letter a (resp., b) and a period $p = 4$.

6.3 Periodic points as automatic sequences

We recalled in Corollary 6.1.17 that the automaton $\mathcal{A}_{\eta,a}$ associated with a substitution η and a letter a enables us to describe the right-infinite fixed point $\lim_{k\to+\infty} \eta^k(a)$ with the growing letter $u_0 = a$. In this section, we extend the automaton $\mathcal{A}_{\eta,a}$ so that it describes a two-sided periodic point of η. Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Let $\mathcal{D} = \{0, ..., \max_{c\in A} |\eta(c)| - 1\}$. We associate an automaton $\mathcal{A}_{\eta,s}$ with the pair (η, s) by adding a new initial state `start` and two additional edges $\mathtt{start} \xrightarrow{\mathtt{0}} u_0$ and $\mathtt{start} \xrightarrow{\mathtt{1}} u_{-1}$ to the automaton $\mathcal{A}_{\eta,u_0}$; compare Figure 6.4 and Figure 6.7.

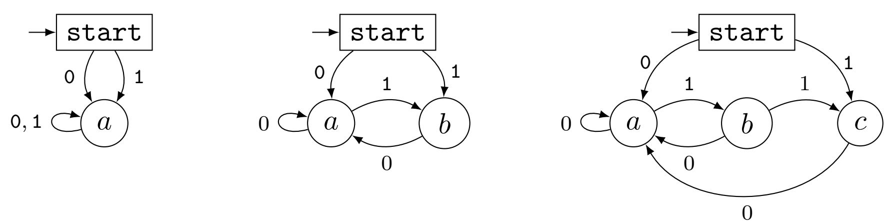


Figure 6.7: The automata $\mathcal{A}_{\omega,s}$, $\mathcal{A}_{\varphi,s}$, $\mathcal{A}_{\Theta,s}$ associated with the 2-uniform substitution $\omega : a \mapsto aa$ and the seed $s = a|a$, the Fibonacci substitution φ: $a \mapsto ab$, $b \mapsto a$ and the seed $s = b|a$, and the Tribonacci substitution $\Theta : a \mapsto ab$, $b \mapsto ac$, $c \mapsto a$ and the seed $s = c|a$.

Denoting the seed $s = b|a$ for some letters $a, b \in A$, the automaton $\mathcal{A}_{\eta,s}$ is related to the usual automata $\mathcal{A}_{\eta,a}$ and $\mathcal{A}_{\eta,b}$ by the following equalities, for every $w \in \mathcal{D}^*$:

$$\mathcal{A}_{\eta,s}(\mathtt{0}w) = \mathcal{A}_{\eta,a}(w) \qquad \text{and} \qquad \mathcal{A}_{\eta,s}(\mathtt{1}w) = \mathcal{A}_{\eta,b}(w). \tag{6.3}$$

Also, if $\mathcal{A}_{\eta,s}(w) = a$ for a nonempty word $w \in \mathcal{D}^+$, then we have, for every $v \in \mathcal{D}^*$,

$$\mathcal{A}_{\eta,a}(v) = \mathcal{A}_{\eta,s}(wv). \tag{6.4}$$

As a consequence, we observe that $\mathcal{L}(\mathcal{A}_{\eta,s}) = \mathtt{0}\mathcal{L}(\mathcal{A}_{\eta,a}) \cup \mathtt{1}\mathcal{L}(\mathcal{A}_{\eta,b})$. The automaton $\mathcal{A}_{\eta,s}$ together with the Dumont–Thomas numeration system $\mathrm{rep}_{\mathbf{u}}$ enables us to describe the periodic point $\mathbf{u}$ in the following way.

Theorem 6.3.1. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Then, for every $n \in \mathbb{Z}$, we have*

$$u_n = \mathcal{A}_{\eta,s}(\mathrm{rep}_{\mathbf{u}}(n)).$$

We denote $\mathtt{W}^{\mathbf{u}}_{\min}$ and $\mathtt{W}^{\mathbf{u}}_{\max}$ the following minimal and maximal element under the tail map:

$$\mathtt{W}^{\mathbf{u}}_{\min} = \mathrm{tail}_{\eta,p,u_0}(0) = \mathtt{0}^p, \qquad \mathtt{W}^{\mathbf{u}}_{\max} = \mathrm{tail}_{\eta,p,u_{-1}}(|\eta^p(u_{-1})| - 1), \tag{6.5}$$

and we observe that the words $\mathtt{W}^{\mathbf{u}}_{\min}$ and $\mathtt{W}^{\mathbf{u}}_{\max}$ play the following neutral role.

Lemma 6.3.2. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Let $w \in \mathcal{L}(\mathcal{A}_{\eta,s})$. Then*

$$\mathcal{A}_{\eta,s}(w) = \begin{cases} \mathcal{A}_{\eta,s}(\mathtt{0}(\mathtt{W}^{\mathbf{u}}_{\min})^* v), & \text{if } w = \mathtt{0}v; \\ \mathcal{A}_{\eta,s}(\mathtt{1}(\mathtt{W}^{\mathbf{u}}_{\max})^* v), & \text{if } w = \mathtt{1}v. \end{cases}$$

The words $\mathtt{W}^{\mathbf{u}}_{\min}$ and $\mathtt{W}^{\mathbf{u}}_{\max}$, which we call the neutral words, enable us to make a link between the complement numeration systems and the Dumont–Thomas numeration systems in Section 6.5 and to extend the Dumont–Thomas numeration systems to higher dimensions in Section 6.6.

Remark 6.3.3. *Let η be a substitution and let $\mathbf{u} \in \mathrm{Per}(\eta)$ with a growing seed $s = u_{-1}|u_0$ and the period $p \geq 1$. We observe that, by Lemma 6.1.9, $\mathtt{W}^{\mathbf{u}}_{\min} \in \mathcal{L}(\mathcal{A}_{\eta,u_0})$ and $\mathtt{W}^{\mathbf{u}}_{\max} \in \mathcal{L}(\mathcal{A}_{\eta,u_{-1}})$. Moreover, combining Lemma 6.1.14 and Equation* (6.5)*, the word $\mathtt{W}^{\mathbf{u}}_{\min}$ (resp., $\mathtt{W}^{\mathbf{u}}_{\max}$) can be obtained as the label of the path of length p in the automaton $\mathcal{A}_{\eta,s}$ starting at the state u_0 (resp., u_{-1}) following the edges with minimal (resp., maximal) lexicographical value.*

6.4 Properties with respect to the total order $\prec$

Dumont–Thomas numeration systems have interesting properties with respect to order. We observed that the Dumont–Thomas numeration systems for $\mathbb{N}$ belong to the class of abstract numeration systems [BR10, §3], which are increasing with respect to the radix order; see Lemma 6.1.16. In this context, it is not surprising that the Dumont–Thomas numeration systems for $\mathbb{Z}$ are increasing with respect to the total order $\prec$, which extends the radix order; see Definition 3.4.1.

Proposition 6.4.1. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed s. Let $p \geq 1$ be the period of $\mathbf{u}$. The map $\mathrm{rep}_{\mathbf{u}} : \mathbb{Z} \to \mathrm{rep}_{\mathbf{u}}(\mathbb{Z})$ is an increasing bijection with respect to the order $\prec$.*

It follows from Theorem 6.3.1 that the representations $\mathrm{rep}_{\mathbf{u}}$ associated with a substitution $\eta : A^* \to A^*$ and a two-sided periodic point $\mathbf{u} \in \mathrm{Per}(\eta)$ with a growing seed $s = u_{-1}|u_0$ form a subset of the language $\mathcal{L}(\mathcal{A}_{\eta,s})$. In the following lemma, this subset is specified.

Lemma 6.4.2. [LL23b] *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. Let $p \geq 1$ be the period of $\mathbf{u}$. Then*

$$\mathrm{rep}_{\mathbf{u}}(\mathbb{Z}) = \bigcup_{\ell \in \mathbb{N}} \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s}) \setminus \{\mathtt{0}\mathtt{W}^{\mathbf{u}}_{\min}, \mathtt{1}\mathtt{W}^{\mathbf{u}}_{\max}\}\mathcal{D}^*.$$

Consequently, we show that a Dumont–Thomas numeration system based on every 2-uniform substitution, which has a two-sided fixed point, is the two's complement numeration system.

Example 6.4.3. Let $\psi : A \to A^*$ be a 2-uniform substitution and let $\mathbf{u} \in \mathrm{Per}(\psi)$ be a two-sided periodic point with the period 1. By Corollary 3.4.4, the map $\mathrm{rep}_{2c} : \mathbb{Z} \to \Sigma^+ \setminus (\mathtt{00}\Sigma^* \cup \mathtt{11}\Sigma^*)$ is the unique increasing bijection with respect to the order $\prec$ such that $0 \mapsto \mathtt{0}$. By Proposition 6.4.1, the map $\mathrm{rep}_{\mathbf{u}} : \mathbb{Z} \to \mathrm{rep}_{\mathbf{u}}(\mathbb{Z})$ is an increasing bijection with respect to the order $\prec$. From

Remark 6.3.3, we have $\mathtt{W}^{\mathbf{u}}_{\min} = \mathtt{0}$ and $\mathtt{W}^{\mathbf{u}}_{\max} = \mathtt{1}$. Also, $\mathrm{rep}_{\mathbf{u}}(0) = \mathtt{0}$ by definition. Finally, we use Lemma 6.4.2 to obtain

$$\begin{aligned}\mathrm{rep}_{\mathbf{u}}(\mathbb{Z}) &= \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{\ell+1}(\mathcal{A}_{\psi,s}) \setminus \{\mathtt{0}\mathtt{W}^{\mathbf{u}}_{\min}, \mathtt{1}\mathtt{W}^{\mathbf{u}}_{\max}\}\{\mathtt{0},\mathtt{1}\}^* \\ &= \mathcal{L}_{\geq 1}(\mathcal{A}_{\psi,s}) \setminus \{\mathtt{0}\mathtt{W}^{\mathbf{u}}_{\min}, \mathtt{1}\mathtt{W}^{\mathbf{u}}_{\max}\}\{\mathtt{0},\mathtt{1}\}^* \\ &= \Sigma^+ \setminus (\mathtt{00}\Sigma^* \cup \mathtt{11}\Sigma^*)\,.\end{aligned}$$

As there is a unique increasing bijection $\mathbb{Z} \to \Sigma^+ \setminus (\mathtt{00}\Sigma^* \cup \mathtt{11}\Sigma^*)$ such that $0 \mapsto \mathtt{0}$, we conclude $\mathrm{rep}_{\mathbf{u}} = \mathrm{rep}_{2c}$.

In [LL23b], we showed that the Fibonacci complement numeration system coincides with a Dumont–Thomas numeration system associated with the Fibonacci substitution. We prove this in a broader framework in the next section.

6.5 Positional Dumont–Thomas numeration systems for $\mathbb{Z}$

We observed that the Dumont–Thomas numeration systems for $\mathbb{Z}$ associated with 2-uniform substitutions are positional numeration systems. We provide a sufficient condition for the Dumont–Thomas numeration systems for $\mathbb{Z}$ to be positional – it suffices to consider Dumont–Thomas numeration systems associated with canonical substitutions of simple Parry numbers.

Let $\beta > 1$ be a simple Parry number with the greedy expansion of unity $d_\beta(1) = t_1 \cdots t_m$ and the quasi-greedy expansion of unity $d^*_\beta(1) = (t_1 t_2 \cdots t_{m-1}(t_m - 1))^\omega$. The canonical substitution η_β over the alphabet $\{1, \ldots, (m-1), m\}$ is defined by [Fab95]

$$\begin{array}{rlcl}\eta_\beta : & 1 & \mapsto & 1^{t_1-1}2 \\ & 2 & \mapsto & 1^{t_2-1}3 \\ & & \vdots & \\ & (m-1) & \mapsto & 1^{t_{m-1}-1}m \\ & m & \mapsto & 1^{t_m-1}.\end{array} \tag{6.6}$$

We recall the notation concerning the numeration systems U_β associated with simple Parry numbers β, which were presented in Section 3.1, such as the language $\mathrm{rep}_\beta(\mathbb{N})$, the canonical alphabet Σ_β and the representation map rep_β. The automaton $\mathcal{H}_{\beta,q_1}$ is the trim minimal automaton accepting $\mathtt{0}^*\,\mathrm{rep}_\beta(\mathbb{N})$. Let $\chi \in [\![1, m]\!]$. Recall from Section 3.3 that we denote $\mathcal{H}_{\beta,q_\chi}$ the automaton which arises from $\mathcal{H}_{\beta,q_1}$ by changing its initial state to the state q_χ and we denote $\mathcal{H}_{\beta,\chi}$ the automaton which arises from the automaton $\mathcal{H}_{\beta,q_1}$ by creating a new initial state $\mathtt{start}$ and adding two additional edges $\mathtt{start} \xrightarrow{\mathtt{0}} q_1$ and $\mathtt{start} \xrightarrow{\mathtt{1}} q_\chi$; see Figure 3.1 and Figure 3.2. We state a simple observation, which makes a link between the automaton $\mathcal{H}_{\beta,\chi}$ associated with a simple Parry number β and χ and the automaton $\mathcal{A}_{\eta_\beta,s}$ associated with the canonical substitution η_β and the growing seed $s = \chi|1$.

Lemma 6.5.1. *Let $\beta > 1$ be a simple Parry number such that $d^*_\beta(1) = (t_1 \cdots t_{m-1}(t_m - 1))^\omega$. Let $\eta_\beta : \{1, 2, \ldots, m)\}^* \to \{1, 2, \ldots, m\}^*$ be the canonical substitution associated with β. Let $\chi \in [\![1, m]\!]$. Let $\mathbf{u} \in \mathrm{Per}(\eta_\beta)$ be a two-sided periodic point with the growing seed $s = \chi|1$. Then the automaton $\mathcal{H}_{\beta,\chi}$ and the automaton $\mathcal{A}_{\eta_\beta,s}$ coincide.*

Proof. It follows from Equation (6.6) that $\mathcal{H}_{\beta,q_1} = \mathcal{A}_{\eta_\beta,1}$; see Remark 3.3.1. As the procedure of modifying the automaton $\mathcal{H}_{\beta,q_1}$ into the automaton $\mathcal{H}_{\beta,\chi}$ is the same as modifying the automaton $\mathcal{A}_{\eta_\beta,u_0}$ into the automaton $\mathcal{A}_{\eta_\beta,s}$, where the growing seed is $s = \chi|1$, we have that the automaton $\mathcal{H}_{\beta,\chi}$ and the automaton $\mathcal{A}_{\eta_\beta,s}$ coincide. □

We show that Dumont–Thomas numeration systems based on canonical substitutions associated with simple Parry numbers coincide with the complement numeration systems associated with simple Parry numbers. In particular, we have that the complement numeration systems associated with simple Parry numbers are Dumont–Thomas. On the other hand, the Dumont–Thomas numeration systems based on simple Parry canonical substitutions are positional.

Proposition 6.5.2. *Let $\beta > 1$ be a simple Parry number such that $d^*_\beta(1) = (t_1 \cdots t_{m-1}(t_m-1))^\omega$. Let $\eta_\beta : \{1, 2, \ldots, m\}^* \to \{1, 2, \ldots, m\}^*$ be the canonical substitution associated with β. Let $\chi \in [\![1, m]\!]$. Let $\mathbf{u} \in \mathrm{Per}(\eta_\beta)$ be a two-sided periodic point with the growing seed $s = \chi|1$. Then*

$$\mathrm{rep}_{\mathbf{u}} = \mathrm{rep}_{\beta,\chi}\,.$$

Proof. Let $\beta > 1$ be a simple Parry number such that $d^*_\beta(1) = (t_1 \cdots t_{m-1}(t_m - 1))^\omega$ and let η_β be the canonical substitution associated with β. Denote $\Sigma = \Sigma_\beta$ the canonical alphabet. Let $\chi \in [\![1, m]\!]$ and let $\mathbf{u} \in \mathrm{Per}(\eta_\beta)$ be a two-sided periodic point with the growing seed $s = \chi|1$. Denote $\mathcal{D} = \{0, \ldots, \max_{c\in A} |\eta_\beta(c)| - 1\}$. From Lemma 6.5.1, we observe that

i) the period of $\mathbf{u}$ is equal to m because of the cycle of maximal edges in the automaton $\mathcal{H}_{\beta,\chi}$,

ii) the alphabets Σ and $\mathcal{D}$ coincide, that is $\Sigma = \mathcal{D}$, and,

iii) using Remark 6.3.3, we have $\mathtt{W}^{\chi}_{\min} = \mathtt{W}^{\mathbf{u}}_{\min}$ and $\mathtt{W}^{\chi}_{\max} = \mathtt{W}^{\mathbf{u}}_{\max}$.

By Proposition 6.4.1, the map $\mathrm{rep}_{\mathbf{u}} : \mathbb{Z} \to \mathrm{rep}_{\mathbf{u}}(\mathbb{Z})$ is an increasing bijection with respect to the order $\prec$. Also, $\mathrm{rep}_{\mathbf{u}}(0) = \mathtt{0}$ by definition. Finally, we use Lemma 6.4.2 and the observations ii) and iii) to obtain

$$\begin{aligned}
\mathrm{rep}_{\mathbf{u}}(\mathbb{Z}) &= \bigcup_{\ell\in\mathbb{N}} \mathcal{L}_{m\ell+1}(\mathcal{A}_{\eta_\beta,s}) \setminus \{\mathtt{0W}^{\mathbf{u}}_{\min}, \mathtt{1W}^{\mathbf{u}}_{\max}\}\mathcal{D}^* \\
&= \mathcal{L}(\mathcal{A}_{\eta_\beta,s}) \cap \Sigma(\Sigma^m) \setminus \{\mathtt{0W}^{\mathbf{u}}_{\min}, \mathtt{1W}^{\mathbf{u}}_{\max}\}\Sigma^* \\
&= \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m) \setminus \{\mathtt{0W}^{\chi}_{\min}, \mathtt{1W}^{\chi}_{\max}\}\Sigma^*.
\end{aligned}$$

By Proposition 3.4.3, the map $\mathrm{rep}_{\beta,\chi} : \mathbb{Z} \to \mathcal{L}(\mathcal{H}_{\beta,\chi}) \cap \Sigma(\Sigma^m)^* \setminus (\mathtt{0W}^{\chi}_{\min}\Sigma^* \cup \mathtt{1W}^{\chi}_{\max}\Sigma^*)$ is the unique increasing bijection with respect to the order $\prec$ such that $0 \mapsto \mathtt{0}$. We conclude $\mathrm{rep}_{\mathbf{u}} = \mathrm{rep}_{\beta,\chi}$. □

Example 6.5.3. Let $\beta = 2$. Then $d_\beta(1) = 2$ and $d^*_\beta(1) = 1^\omega$. The canonical substitution is thus $\eta_\beta : 1 \mapsto 11$, which is 2-uniform. Let $\mathbf{u} \in \mathrm{Per}(\eta_\beta)$ be a two-sided periodic point with the period 1 and the growing seed $1|1$. We observed in Example 6.4.3 that $\mathrm{rep}_{\mathbf{u}} = \mathrm{rep}_{2c}$.

Example 6.5.4. Let $\beta = \tau$ be the golden mean. Then $d_\beta(1) = 11$ and $d^*_\beta(1) = (10)^\omega$. The canonical substitution is thus $\eta_\beta : 1 \mapsto 12$, $2 \mapsto 1$, which is the Fibonacci substitution $\varphi : a \mapsto ab$, $b \mapsto a$. Let $\mathbf{f}' \in \mathrm{Per}(\varphi)$ be a two-sided periodic point with the period 2 and the growing seed $s = b|a$. Thus $\chi = 2$. Observe that the automata $\mathcal{H}_{\tau,\chi}$ in Figure 3.2 and $\mathcal{A}_{\varphi,s}$ in Figure 6.7 coincide. Also, the representations in the Dumont–Thomas numeration system for $\mathbb{Z}$ in Table 6.1 coincide with the representations in the Fibonacci complement numeration system in Table 3.2.

If a Dumont–Thomas numeration system for $\mathbb{Z}$ is positional, we can represent an integer $n \in \mathbb{Z}$ without the knowledge of the other representations in the numeration system.

Example 6.5.5. Let γ be the dominant root of the polynomial x^3-x^2-x-1. Then $d_\gamma(1)=111$ and $d^*_\gamma(1)=(110)^\omega$. The canonical substitution is $\eta_\gamma : 1 \mapsto 12,\ 2 \mapsto 13,\ 3 \mapsto 1$, which is the Tribonacci substitution $\Theta : a \mapsto ab,\ b \mapsto ac,\ c \mapsto a$. Let $\mathbf{t}' \in \mathrm{Per}(\Theta)$ be a two-sided periodic point with the period 3 and the growing seed $s = c|a$. Thus $\chi = 3$.

We construct the value map $\mathrm{val}_{\gamma,\chi}$ defined by Equation (3.15). The sequence $U = (U_n)_{n=0}^{+\infty}$ defined by Equation (3.10) satisfies the property that $U_0 = 1$, $U_1 = 2$, $U_2 = 4$ and $U_n = \sum_{i=1}^{3} U_{n-i}$, for every $n \geq 3$. From Figure 6.4, we have that

$$V_{n,\chi} = \#\mathcal{L}_n(\mathcal{A}_{\Theta,c}) = \#\mathcal{L}_{n-1}(\mathcal{A}_{\Theta,a}) = U_{n-1}.$$

We show the first values of the sequences $(U_n)_{n=0}^{+\infty}$ and $(V_{3n,\chi})_{n=0}^{+\infty}$ in the following table.

$-V_{9,\chi}$	$-V_{6,\chi}$	$-V_{3,\chi}$	$-V_{0,\chi}$	U_0	U_1	U_2	U_3	U_4	U_5	U_6	U_7	U_8	U_9
-149	-24	-4	-1	1	2	4	7	13	24	44	81	149	274

Thus for every $w = w_{k-1}\cdots w_0 \in \{\mathtt{0},\mathtt{1}\}\{\mathtt{0},\mathtt{1},\mathtt{2}\}^{3\ell}$, for some integer $\ell \geq 1$, we have

$$\mathrm{val}_{\gamma,\chi}(w) = \sum_{i=0}^{k-2} U_i - w_{k-1}V_{k-1,\chi} = \sum_{i=0}^{k-2} U_i - w_{k-1}U_{k-2}.$$

We represent $n = -53$ in the Dumont–Thomas numeration system associated with the two-sided periodc point $\mathbf{t}'$. We have

$$\begin{aligned}
-53 &= -149 + 96 \\
&= -149 + 81 + 15 \\
&= -149 + 81 + 13 + 2 \\
&= -V_9 + \mathtt{0}U_8 + \mathtt{1}U_7 + \mathtt{0}U_6 + \mathtt{0}U_5 + \mathtt{1}U_4 + \mathtt{0}U_3 + \mathtt{0}U_2 + \mathtt{1}U_1 + \mathtt{0}U_0
\end{aligned}$$

and therefore $\mathrm{rep}_{\mathbf{t}'}(-53) = \mathtt{1010010010}$.

Example 6.5.6. Let $\beta = 1+\sqrt{2}$ be the silver mean. Then $d_\beta(1) = 21$ and $d^*_\beta(1) = (20)^\omega$. The canonical substitution is thus $\eta_\beta : 1 \mapsto 112,\ 2 \mapsto 1$, which we rename to $a \mapsto aab,\ b \mapsto a$. Let $\mathbf{u} \in \mathrm{Per}(\eta_\beta)$ be a two-sided periodic point with the period 2 and the growing seed $s = b|a$. The representations in the corresponding Dumont–Thomas numeration system for $\mathbb{Z}$ are in Table 6.3.

Another substitution related to the silver mean is the substitution $\rho : a \mapsto abb,\ b \mapsto ab$. This is due to the fact that the silver mean is the Perron–Frobenius eigenvalue of the incidence matrix of ρ. Let $\mathbf{v} \in \mathrm{Per}(\rho)$ be a two-sided periodic point with the period 1 and the growing seed $s = b|a$. We show the representations in the corresponding Dumont–Thomas numeration system for $\mathbb{Z}$ in Table 6.3. We observe that, unlike the Dumont–Thomas numeration system based on the canonical substitution of the silver mean, this numeration system is not positional. Indeed, we assume by contradiction that there exists a value map such that $\mathrm{val}(\mathrm{rep}_{\mathbf{v}}(n)) = n$. As $\mathrm{rep}_{\mathbf{v}}(3) = \mathtt{010}$, we have that $U_0 = 1$ and $U_1 = 3$. Then $\mathrm{val}(\mathrm{rep}_{\mathbf{v}}(5)) = \mathrm{val}(\mathtt{020}) = 6 \neq 5$.

We observed that a Dumont–Thomas numeration system for $\mathbb{Z}$ is positional provided a certain sufficient condition. Moreover, in Example 6.5.6, we showed two substitutions closely related to the silver mean such that the numeration system associated with one of them is positional and the numeration system associated with the other is not. Further research is needed to answer the following open question.

Question 6.5.7. *What is the necessary condition for a Dumont–Thomas numeration system for $\mathbb{Z}$ to be positional?*

n	$\mathrm{rep}_{\mathbf{u}}(n)$
-8	10102
-7	10110
-6	10111
-5	10112
-4	10120
-3	100
-2	101
-1	1

n	$\mathrm{rep}_{\mathbf{u}}(n)$
0	0
1	001
2	002
3	010
4	011
5	012
6	020
7	0100

n	$\mathrm{rep}_{\mathbf{v}}(n)$
-8	1020
-7	1021
-6	1022
-5	100
-4	101
-3	102
-2	10
-1	1

n	$\mathrm{rep}_{\mathbf{v}}(n)$
0	0
1	01
2	02
3	010
4	011
5	020
6	021
7	0100

Table 6.3: Dumont–Thomas numeration systems for $\mathbb{Z}$ related to the silver mean $1+\sqrt{2}$, where $\mathbf{u}$ is the two-sided periodic point of the canonical substitution $a \mapsto aab$, $b \mapsto a$ with the growing seed $b|a$ and $\mathbf{v}$ is the two-sided periodic point of the substitution $\rho : a \mapsto abb$, $b \mapsto ab$ with the growing seed $b|a$.

6.6 Dumont–Thomas numeration systems for $\mathbb{Z}^d$

Dumont–Thomas numeration systems can be extended naturally to higher dimensions. Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ be a two-sided periodic point with a growing seed $s = u_{-1}|u_0$. We use the neutral words $\mathtt{W}^{\mathbf{u}}_{\min}$ and $\mathtt{W}^{\mathbf{u}}_{\max}$ to pad representations to a certain length; see Equation (6.5). Let $\mathcal{A}_{\eta,s}$ be the automaton associated with η and s. Let $w \in \mathcal{L}_{\ell p+1}(\mathcal{A}_{\eta,s})$ for some $\ell \in \mathbb{N}$. Let $t \in \mathbb{N}$ such that $t \geq |w|$ and $t \bmod p = 1$. We define

$$\mathrm{pad}_t(w) = \begin{cases} \mathtt{0}(\mathtt{W}^{\mathbf{u}}_{\min})^m v, & \text{if } w = \mathtt{0}v; \\ \mathtt{1}(\mathtt{W}^{\mathbf{u}}_{\max})^m v, & \text{if } w = \mathtt{1}v, \end{cases}$$

where $m = (t-|w|)/p$. The pad map allows to represent coordinates in $\mathbb{Z}^d$, for every $d \geq 1$.

Definition 6.6.1 (Numeration system for $\mathbb{Z}^d$)**.** *Let $\eta : A^* \to A^*$ be a substitution and $\mathbf{u} \in \mathrm{Per}(\eta)$ with a growing seed. Let $\mathcal{D} = \{0, ..., \max_{c\in A} |\eta(c)| - 1\}$. For every $\boldsymbol{n} = (n_1, n_2, \ldots, n_d) \in \mathbb{Z}^d$, we define*

$$\mathrm{rep}_{\mathbf{u}}(\boldsymbol{n}) = \begin{pmatrix} \mathrm{pad}_t(\mathrm{rep}_{\mathbf{u}}(n_1)) \\ \mathrm{pad}_t(\mathrm{rep}_{\mathbf{u}}(n_2)) \\ \vdots \\ \mathrm{pad}_t(\mathrm{rep}_{\mathbf{u}}(n_d)) \end{pmatrix} \in \{\mathtt{0},\mathtt{1}\}^d(\mathcal{D}^d)^*,$$

where $t = \max\{|\mathrm{rep}_{\mathbf{u}}(n_i)| : 1 \leq i \leq d\}$.

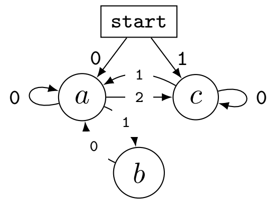
In the following example, we illustrate the procedure to obtain a Dumont–Thomas representation of a position in $\mathbb{Z}^2$ based on a particular periodic point.

Example 6.6.2. Let $\eta : a \mapsto abc$, $b \mapsto a$, $c \mapsto ca$ be a substitution and denote $\mathbf{u} = \mathrm{Per}(\eta)$ with the growing seed $s = c|a$. The corresponding automaton $\mathcal{A}_{\eta,s}$ and the representations $\mathrm{rep}_{\mathbf{u}}(n)$, for $n \in [\![-7, 6]\!]$, are in Figure 6.8. Thus we have that $\mathtt{W}^{\mathbf{u}}_{\min} = \mathtt{00}$ and $\mathtt{W}^{\mathbf{u}}_{\max} = \mathtt{12}$; see Remark 6.3.3. We show how to represent a position $(6, -5) \in \mathbb{Z}^2$. We have that

$$\mathrm{rep}_{\mathbf{u}}(6) = \mathtt{00100} \quad \text{and} \quad \mathrm{rep}_{\mathbf{u}}(-5) = \mathtt{100}.$$

As the word $\mathrm{rep}_{\mathbf{u}}(-5)$ is shorter in length than the word $\mathrm{rep}_{\mathbf{u}}(6)$, we pad the representation $\mathrm{rep}_{\mathbf{u}}(-5)$ with the word $\mathtt{W}^{\mathbf{u}}_{\max} = \mathtt{12}$ so that they have the same length. Hence, we obtain

$$\mathrm{rep}_{\mathbf{u}}((6,-5)) = \begin{pmatrix} \mathtt{00100} \\ \mathtt{100} \end{pmatrix} = \begin{pmatrix} \mathtt{00100} \\ \mathtt{11200} \end{pmatrix}.$$



n	$\mathrm{rep}_{\mathbf{u}}(n)$		n	$\mathrm{rep}_{\mathbf{u}}(n)$
6	00100		-1	1
5	021		-2	111
4	020		-3	110
3	010		-4	101
2	002		-5	100
1	001		-6	11102
0	0		-7	11101

Figure 6.8: The automaton $\mathcal{A}_{\eta,s}$ associated with the substitution $\eta : a \mapsto abc$, $b \mapsto a$, $c \mapsto ca$ and the growing seed $s = c|a$ and the corresponding Dumont–Thomas numeration system.

It is readily possible to generalize Definition 6.6.1 so that a different periodic point is used in each coordinate, provided that all such periodic points have the same period. With Definition 6.2.3 extended to any period $p \geq 1$ of a periodic point (see Remark 6.2.8), Definition 6.6.1 could be generalized even further, using a different periodic point and a different pad map in each coordinate. Our intuition is that the Dumont–Thomas numeration systems for $\mathbb{Z}^d$ using such a generalized definition of the representation map could describe all d-dimensional configurations which are periodic points of d-dimensional substitutions as automatic sequences. For now, we leave this as an open question.

Question 6.6.3. *Is it possible to generalize Theorem 6.3.1 to $\mathbb{Z}^d$, for $d \geq 1$?*

6.6.1 Fibonacci complement numeration system for $\mathbb{Z}^2$

The Fibonacci complement numeration system is a Dumont–Thomas numeration system for $\mathbb{Z}$; see Example 6.5.4. Thus we can extend it to $\mathbb{Z}^2$. In Figure 6.9, we show all representations of the positions $\boldsymbol{n} \in [\![-5, 7]\!] \times [\![-5, 7]\!]$ in the Fibonacci complement numeration system for $\mathbb{Z}^2$. We omit the brackets around representations due to the lack of space. The properties of the Fibonacci complement numeration system extended to $\mathbb{Z}^2$ allow us to find an automatic characterization of a Wang configuration in Chapter 7.

7	10000 01010	10001 01010	10010 01010	10100 01010	10101 01010	00000 01010	00001 01010	00010 01010	00100 01010	00101 01010	01000 01010	01001 01010	01010 01010
6	10000 01001	10001 01001	10010 01001	10100 01001	10101 01001	00000 01001	00001 01001	00010 01001	00100 01001	00101 01001	01000 01001	01001 01001	01010 01001
5	10000 01000	10001 01000	10010 01000	10100 01000	10101 01000	00000 01000	00001 01000	00010 01000	00100 01000	00101 01000	01000 01000	01001 01000	01010 01000
4	10000 00101	10001 00101	10010 00101	10100 00101	10101 00101	00000 00101	00001 00101	00010 00101	00100 00101	00101 00101	01000 00101	01001 00101	01010 00101
3	10000 00100	10001 00100	10010 00100	10100 00100	10101 00100	00000 00100	00001 00100	00010 00100	00100 00100	00101 00100	01000 00100	01001 00100	01010 00100
2	10000 00010	10001 00010	10010 00010	100 010	101 010	000 010	001 010	010 010	00100 00010	00101 00010	01000 00010	01001 00010	01010 00010
1	10000 00001	10001 00001	10010 00001	100 001	101 001	000 001	001 001	010 001	00100 00001	00101 00001	01000 00001	01001 00001	01010 00001
0	10000 00000	10001 00000	10010 00000	100 000	1 0	0 0	001 000	010 000	00100 00000	00101 00000	01000 00000	01001 00000	01010 00000
−1	10000 10101	10001 10101	10010 10101	100 101	1 1	0 1	001 101	010 101	00100 10101	00101 10101	01000 10101	01001 10101	01010 10101
−2	10000 10100	10001 10100	10010 10100	100 100	101 100	000 100	001 100	010 100	00100 10100	00101 10100	01000 10100	01001 10100	01010 10100
−3	10000 10010	10001 10010	10010 10010	10100 10010	10101 10010	00000 10010	00001 10010	00010 10010	00100 10010	00101 10010	01000 10010	01001 10010	01010 10010
−4	10000 10001	10001 10001	10010 10001	10100 10001	10101 10001	00000 10001	00001 10001	00010 10001	00100 10001	00101 10001	01000 10001	01001 10001	01010 10001
−5	10000 10000	10001 10000	10010 10000	10100 10000	10101 10000	00000 10000	00001 10000	00010 10000	00100 10000	00101 10000	01000 10000	01001 10000	01010 10000
	−5	−4	−3	−2	−1	0	1	2	3	4	5	6	7

Figure 6.9: The Fibonacci complement numeration system $\mathcal{F}c$ extended to $\mathbb{Z}^2$.

Chapter 7

Automatic characterization of an aperiodic Wang shift

In this chapter, we present our results concerning a particular Wang shift $\Omega_{\mathcal{Z}}$ [LL21]. More precisely, we characterize it as a 2-dimensional automatic sequence with the help of the Fibonacci complement numeration system $\mathcal{F}c$ extended to $\mathbb{Z}^2$. As a new result not provided in [LL21], we show that $\Omega_{\mathcal{Z}}$ is topologically conjugate to another Wang shift $\Omega_{\mathcal{U}}$ described by Labbé [Lab20], which is minimal, aperiodic and self-similar. As the Wang shift $\Omega_{\mathcal{U}}$ is related to the 11 tile Wang shift discovered by Jeandel and Rao, this creates a link between the Jeandel–Rao Wang shift and $\Omega_{\mathcal{Z}}$.

7.1 Introduction to Wang tiles and Wang shifts

First, we summarize basic notions about Wang tiles. A *Wang tile* is a 4-tuple $(A, B, C, D) \in \mathcal{C}^4$ of letters in an alphabet $\mathcal{C}$ [Wan61]. Geometrically speaking, a Wang tile (A, B, C, D) represents the labeling of edges of a unit square, where A is the east edge label, B is the north edge label, C is the west edge label and D is the south edge label, by convention. We show the tile (A, B, C, D) with an index 0 in Figure 7.1. When clear from context, we refer to tiles as to their indices.

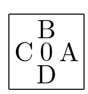


Figure 7.1: A Wang tile (A, B, C, D) with an index 0.

Given a finite set of Wang tiles, we assume that we have infinitely many copies of them, which we can translate but not rotate. We refer to an ordered pair of Wang tiles arranged side by side horizontally (resp., vertically) as to a *horizontal* (resp., *vertical*) *pattern*. A horizontal (or vertical) pattern consisting of two Wang tiles is called *allowed* provided that the labels of the tiles on their common edge match. Otherwise a pattern is called *forbidden*; see Figure 7.2.

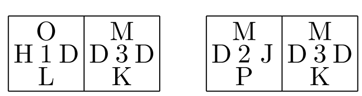


Figure 7.2: An allowed horizontal pattern $(1, 3)$ and a forbidden horizontal pattern $(2, 3)$.

The two main questions that we ask when studying a given set of Wang tiles $\mathcal{T}$ are:

- Is it possible to tile the plane with $\mathcal{T}$, avoiding forbidden patterns?
- If so, is it possible in a periodic way?

To answer the first main question, we need the definition of Wang configurations. A *Wang configuration* over a finite set of Wang tiles $\mathcal{T} = \{t_0, t_1, \dots, t_{k-1}\}$ is a map $f : \mathbb{Z}^2 \to \mathcal{T}$, which assigns a Wang tile from $\mathcal{T}$ to every position in $\mathbb{Z}^2$. A Wang configuration is called *valid* if, for every $\boldsymbol{n} \in \mathbb{Z}^2$, the east label of the tile $f(\boldsymbol{n})$ equals the west label of the tile $f(\boldsymbol{n} + \mathbf{e}_1)$ and the north label of the tile $f(\boldsymbol{n})$ equals the south label of the tile $f(\boldsymbol{n} + \mathbf{e}_2)$, where $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$ denote the canonical generators of $\mathbb{Z}^2$. In other words, a Wang configuration is valid if all of its horizontal and vertical patterns are allowed. A Wang configuration f is called *periodic* if there exists a vector $\boldsymbol{m} \in \mathbb{Z}^2$ such that $f(\boldsymbol{n}+\boldsymbol{m}) = f(\boldsymbol{n})$, for every $\boldsymbol{n} \in \mathbb{Z}^2$; see Example 7.1.1. Therefore, to answer the second main question, the set $\mathcal{T}$ tiles the plane in a periodic way if there exists a periodic Wang configuration $\mathbb{Z}^2 \to \mathcal{T}$.

Example 7.1.1. Let $\mathcal{P} = \{p_0, p_1, p_2, p_3\}$ be the set of Wang tiles shown in Figure 7.3.

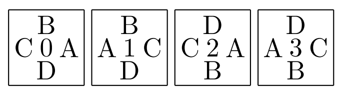


Figure 7.3: A set of Wang tiles.

Let $p\colon \mathbb{Z}^2 \to \{0, 1, 2, 3\}$ be the Wang configuration such that

$$p(\boldsymbol{n}) = \begin{cases} 0, & \text{if } n_1 \bmod 2 = 0 \text{ and } n_2 \bmod 2 = 0; \\ 1, & \text{if } n_1 \bmod 2 = 1 \text{ and } n_2 \bmod 2 = 0; \\ 2, & \text{if } n_1 \bmod 2 = 0 \text{ and } n_2 \bmod 2 = 1; \\ 3, & \text{if } n_1 \bmod 2 = 1 \text{ and } n_2 \bmod 2 = 1. \end{cases}$$

Then we have that $p(\boldsymbol{n}+(2, 2)) = p(\boldsymbol{n})$, for every $\boldsymbol{n} \in \mathbb{Z}^2$, and p is a periodic Wang configuration. We observe that p is formed by the 2×2 pattern shown in Figure 7.4 translated infinitely many times. As by placing the 2×2 pattern side by side does not create forbidden patterns, we have that p is a valid Wang configuration.

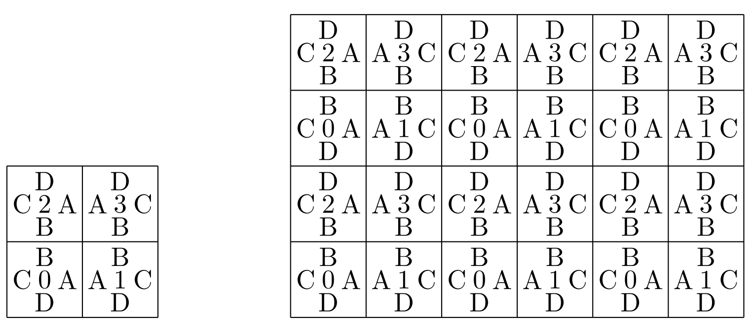


Figure 7.4: A 2×2 pattern created from the tiles $\mathcal{P}$ and a partial valid configuration $p \in \Omega_{\mathcal{P}}$.

A *Wang shift* $\Omega_{\mathcal{T}}$ associated with the set of Wang tiles $\mathcal{T}$ is the set of all valid Wang configurations. A Wang shift $\Omega_{\mathcal{T}}$ is called *aperiodic* if every configuration $f \in \Omega_{\mathcal{T}}$ satisfies the condition that it is not periodic.

Example 7.1.2. We consider a set containing a single Wang tile from Figure 7.1. The horizontal pattern $(0,0)$ is forbidden. Hence, no valid Wang configuration exists and the corresponding Wang shift is empty.

Example 7.1.3 (Continuation of Example 7.1.1)**.** We observe that the Wang shift $\Omega_{\mathcal{P}}$ is nonempty as it contains the valid Wang configuration p. Also, as the Wang shift $\Omega_{\mathcal{P}}$ contains a periodic Wang configuration p, it is not aperiodic.

7.2 Preliminaries on two-dimensional languages and subshifts

We can approach Wang configurations as 2-dimensional words if we only consider the indices of the Wang tiles. In this section, we provide preliminaries on two-dimensional words, languages and subshifts which enables us to study the Wang shifts as symbolic dynamical systems. This subject is broadly studied in the following book [Kur03].

Let A be an alphabet. A *2-dimensional word* over A of *shape* $\boldsymbol{n} = (n_1, n_2) \in \mathbb{N}^2$ is a map $u : [\![0, n_1 - 1]\!] \times [\![0, n_2 - 1]\!] \to A$. We represent it as a matrix with Cartesian coordinates:

$$u = \begin{pmatrix} u_{0,n_2-1} & \dots & u_{n_1-1,n_2-1} \\ \dots & \dots & \dots \\ u_{0,0} & \dots & u_{n_1-1,0} \end{pmatrix}$$

and we call n_1 the *width* of u and n_2 the *height* of u. We denote $\mathrm{width}(u) = n_1$ and $\mathrm{height}(u) = n_2$. The set of all 2-dimensional words over A of shape $\boldsymbol{n}$ is denoted by $A^{\boldsymbol{n}}$ and the set $\bigcup_{\mathbf{n}\in\mathbb{N}^2} A^{\mathbf{n}}$ of all 2-dimensional words over A is denoted by A^{*2}. Let $\boldsymbol{n} = (n_1, n_2)$, $\boldsymbol{m} = (m_1, m_2) \in \mathbb{N}^2$ and let $u \in A^{\boldsymbol{n}}$, $v \in A^{\boldsymbol{m}}$. If $n_2 = m_2$, the concatenation of u and v in direction $\mathbf{e}_1$ is defined as a 2-dimensional word $u \odot^1 v$ of shape $(n_1 + m_1, n_2)$ given as

$$u \odot^1 v = \begin{pmatrix} u_{0,n_2-1} & \dots & u_{n_1-1,n_2-1} & v_{0,n_2-1} & \dots & v_{m_1-1,n_2-1} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ u_{0,0} & \dots & u_{n_1-1,0} & v_{0,0} & \dots & v_{m_1-1,0} \end{pmatrix}.$$

If $n_1 = m_1$, the concatenation of u and v in direction $\mathbf{e}_2$ is defined in an analogous way. A word $v \in A^{*2}$ is a *subword* of a word $u \in A^{*2}$ if there exist words $u_1, u_2, u_3, u_4 \in A^{*2}$ such that

$$u = u_3 \odot^2 (u_1 \odot^1 v \odot^1 u_2) \odot^2 u_4.$$

A subset $L \subset A^{*2}$ is called a *2-dimensional language*. A language $L \subset A^{*2}$ is *factorial* if for every $u \in L$ it holds that $v \in L$ for every subword v of u.

Example 7.2.1. The 2-dimensional words $u_1 = (0)$, $u_2 = \left(\begin{smallmatrix} 2 & 3 \end{smallmatrix}\right)$, $u_3 = \left(\begin{smallmatrix} 1 \\ 3 \end{smallmatrix}\right)$, $u_4 = \left(\begin{smallmatrix} 0 & 1 \\ 2 & 3 \end{smallmatrix}\right)$, have shape $(1,1), (2,1), (1,2), (2,2)$, respectively. We observe that u_1, u_2 and u_3 are subwords of u_4. We can concatenate $u_2 \odot^1 u_1 = \left(\begin{smallmatrix} 2 & 3 & 0 \end{smallmatrix}\right)$ and $u_2 \odot^2 u_4 = \left(\begin{smallmatrix} 0 & 1 \\ 2 & 3 \\ 2 & 3 \end{smallmatrix}\right)$, but $u_1 \odot^1 u_3$ and $u_2 \odot^2 u_3$ are not defined. The language $L = \{u_i : i \in [\![1, 4]\!]\}$ is not factorial, whereas $L \cup \{\varepsilon, (1), (2), (3), \left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \end{smallmatrix}\right)\}$ is factorial.

Let $A^{\mathbb{Z}^2}$ be the set of all maps $\mathbb{Z}^2 \to A$ equipped with the compact product topology. A map $x : \mathbb{Z}^2 \to A$ is called a *configuration*. The compact product topology is compatible with the metric $\mathrm{dist} : A^{\mathbb{Z}^2} \times A^{\mathbb{Z}^2} \to \mathbb{R}$ defined for every $x, y \in A^{\mathbb{Z}^2}$ by

$$\mathrm{dist}(x,y) = \begin{cases} 0, & \text{if } x = y; \\ 2^{-\min\{|n_1|+|n_2| \,:\, x_{\boldsymbol{n}} \neq y_{\boldsymbol{n}}\}}, & \text{otherwise.} \end{cases}$$

The space $A^{\mathbb{Z}^2}$ equipped with the metric dist is a compact metric space. Let $(x_n)_{n=0}^{+\infty} \subset A^{\mathbb{Z}^2}$ be a sequence of configurations and $x \in A^{\mathbb{Z}^2}$. We say that $(x_n)_{n=0}^{+\infty}$ *converges* to x if for all $\varepsilon > 0$ there exists $N \in \mathbb{N}$ such that, for all integers $n \geq N$, $\operatorname{dist}(x_n, x) < \varepsilon$. We denote $\lim_{n\to+\infty} x_n = x$.

Let $x \in A^{\mathbb{Z}^2}$ be a configuration and let $\boldsymbol{n} \in \mathbb{Z}^2$. The *shift action* is the map $\sigma : \boldsymbol{n} \mapsto \sigma^{\boldsymbol{n}}$ of the additive group $\mathbb{Z}^2$ on $A^{\mathbb{Z}^2}$ defined for every $\boldsymbol{m} \in \mathbb{Z}^2$ by the rule $(\sigma^{\boldsymbol{n}}(x))_{\boldsymbol{m}} = x_{\boldsymbol{m}+\boldsymbol{n}}$. For every $X \subset A^{\mathbb{Z}^2}$, the shift-closure of X is defined as

$$\overline{X}^{\sigma} := \{\sigma^{\boldsymbol{n}}(x) \mid x \in X,\ \boldsymbol{n} \in \mathbb{Z}^2\}.$$

A subset $X \subset A^{\mathbb{Z}^2}$ closed with respect to the compact product topology such that $X = \overline{X}^{\sigma}$ is called a *subshift*. Consequently, X is a compact metric space equipped with the metric dist.

Let $x \in A^{\mathbb{Z}^2}$ be a configuration. We call x *periodic* if there is a vector $\boldsymbol{n} \in \mathbb{Z}^2 \setminus \{(0,0)\}$ such that $x = \sigma^{\boldsymbol{n}}(x)$. Otherwise we call it *non-periodic*. We denote $\mathcal{L}(x)$ the language containing all subwords of x. We observe that $\mathcal{L}(x)$ is factorial. If $X \subset A^{\mathbb{Z}^2}$ is a subshift, we call the factorial language $\mathcal{L}(X) = \bigcup_{x\in X} \mathcal{L}(x)$ the *language of* X. Conversely, if $L \subset A^{*^2}$ is a factorial language, we let $\mathcal{X}_L = \{x \in A^{\mathbb{Z}^2} \mid \mathcal{L}(x) \subset L\}$ denote the subshift generated by L.

Let $X \subset A^{\mathbb{Z}^2}$ and $Y \subset B^{\mathbb{Z}^2}$ be subshifts. A map $\theta : X \to Y$ is *continuous* if for all sequences $(x_n)_{n=0}^{+\infty} \subset X$ and all $x \in X$, it holds that $\lim_{n\to+\infty} \theta(x_n) = \theta(x)$ if and only if $\lim_{n\to+\infty} x_n = x$. A *homeomorphism* is a map $X \to Y$ which fulfills the condition that

- $\theta : X \to Y$ is a bijection,
- $\theta : X \to Y$ is continuous,
- the inverse map $\theta^{-1} : Y \to X$ is continuous.

We say that X and Y are *topologically conjugate* if there exists a homeomorphism $\theta : X \to Y$ such that $\theta \circ \sigma^{\boldsymbol{n}} = \sigma^{\boldsymbol{n}} \circ \theta$, for every $\boldsymbol{n} \in \mathbb{Z}^2$. We refer to the last property as to the *commuting property* of θ.

Let A and B be alphabets and let $L \subset A^{*^2}$ be a factorial language. A *2-dimensional morphism* is a map $\phi : L \to B^{*^2}$ fulfilling the condition that for every $i \in \{1,2\}$ and every $u, v \in L$ such that $u \odot^i v$ is defined and $u \odot^i v \in L$, we have that the concatenation $\phi(u) \odot^i \phi(v)$ in direction $\mathbf{e}_i$ is defined and

$$\phi(u \odot^i v) = \phi(u) \odot^i \phi(v).$$

Thus, a 2-dimensional morphism ϕ is completely defined from images of the letters $a \in A$ and it can be extended to a continuous map $\phi : \mathcal{X}_L \to B^{\mathbb{Z}^2}$ in such a way that the origin of $\phi(x)$ is at zero position in the word $\phi(x_{(0,0)})$, for all $x \in \mathcal{X}_L$. A 2-dimensional morphism ϕ is called *expansive* if, for every letter $a \in A$, the width and height of $\phi^k(a)$ goes to $+\infty$. The *language* $\mathcal{L}_\phi$ of an expansive 2-dimensional morphism ϕ is the set

$$\mathcal{L}_\phi = \{u \in A^{*^2} : u \text{ is a subword of } \phi^n(a) \text{ for some } a \in A \text{ and } n \in \mathbb{N}\}.$$

Let $X \subset A^{\mathbb{Z}^2}$ be a subshift. We say that X is

- *minimal* if for every subshift $Y \subset X$ it holds that $Y = \emptyset$ or $Y = X$;
- *self-similar* if there exists an expansive morphism $\phi : \mathcal{L}_\phi \to A^{*^2}$ such that $X = \overline{\phi(X)}^{\sigma}$;
- *aperiodic* if every configuration $x \in X$ is non-periodic.

7.3 Wang shifts related to the Jeandel–Rao Wang shift

In 2021, Jeandel and Rao constructed a set of 11 Wang tiles, which generates an aperiodic Wang shift and they proved that this size of an aperiodic Wang set is minimal [JR21]. In other words, every set of less than 11 Wang tiles admits a periodic configuration. Based on the set of 11 Wang tiles described by Jeandel and Rao, a set of 19 Wang tiles $\mathcal{U}$ with some remarkable properties was constructed [Lab21]; see Figure 7.5. The Wang shift $\Omega_{\mathcal{U}}$ was proved to be self-similar, minimal and aperiodic [Lab19]. Moreover, it has three characterizations: as a self-similar subshift, as a Wang shift and as the coding of a toral $\mathbb{Z}^2$-rotation [Lab20].

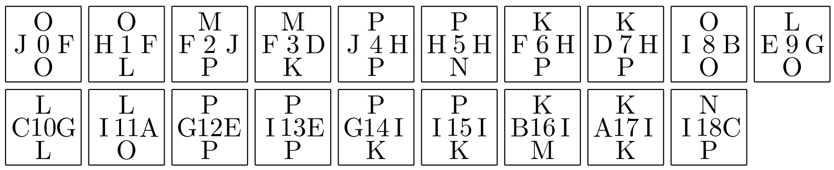


Figure 7.5: The set $\mathcal{U} = \{u_0, \ldots, u_{18}\}$ of 19 Wang tiles.

Theorem 7.3.1. [Lab21] *The Wang shift $\Omega_{\mathcal{U}}$ is minimal, aperiodic and self-similar. The 2-dimensional morphism Ψ such that $\Omega_{\mathcal{U}} = \overline{\Psi(\Omega_{\mathcal{U}})}^{\sigma}$ is defined by*

$$\Psi : \quad [\![0,18]\!] \to [\![0,18]\!]^{*^2}$$

$$\begin{cases} 0 \mapsto (17), & 1 \mapsto (16), \qquad 2 \mapsto (15,\ 11), \qquad 3 \mapsto (13,\ 9), \\ 4 \mapsto (17,\ 8), & 5 \mapsto (16,\ 8), \qquad 6 \mapsto (15,\ 8), \qquad 7 \mapsto (14,\ 8), \\ 8 \mapsto \begin{pmatrix} 6 \\ 14 \end{pmatrix}, & 9 \mapsto \begin{pmatrix} 3 \\ 17 \end{pmatrix}, \qquad 10 \mapsto \begin{pmatrix} 3 \\ 16 \end{pmatrix}, \qquad 11 \mapsto \begin{pmatrix} 2 \\ 14 \end{pmatrix}, \\ 12 \mapsto \begin{pmatrix} 7 & 1 \\ 15 & 11 \end{pmatrix}, & 13 \mapsto \begin{pmatrix} 6 & 1 \\ 14 & 11 \end{pmatrix}, \quad 14 \mapsto \begin{pmatrix} 7 & 1 \\ 13 & 9 \end{pmatrix}, \quad 15 \mapsto \begin{pmatrix} 6 & 1 \\ 12 & 9 \end{pmatrix}, \\ 16 \mapsto \begin{pmatrix} 5 & 1 \\ 18 & 10 \end{pmatrix}, & 17 \mapsto \begin{pmatrix} 4 & 1 \\ 13 & 9 \end{pmatrix}, \quad 18 \mapsto \begin{pmatrix} 2 & 0 \\ 14 & 8 \end{pmatrix}. \end{cases} \tag{7.1}$$

We introduce a set of Wang tiles $\mathcal{Z}$, which is a simplification of the set $\mathcal{U}$ after the unification of some labels; see Figure 7.6. We prove that the Wang shift $\Omega_{\mathcal{Z}}$ has the same properties as $\Omega_{\mathcal{U}}$.

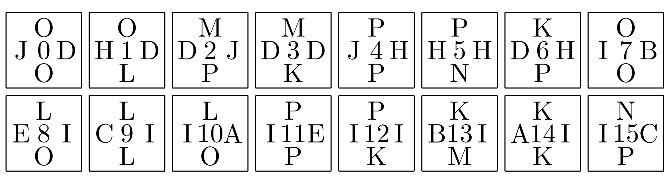


Figure 7.6: The set $\mathcal{Z} = \{z_0, \ldots, z_{15}\}$ of 16 Wang tiles.

Theorem 7.3.2. *The Wang shift $\Omega_{\mathcal{Z}}$ is minimal, aperiodic and self-similar. The 2-dimensional morphism ϕ such that $\Omega_{\mathcal{Z}} = \overline{\phi(\Omega_{\mathcal{Z}})}^{\sigma}$ is defined by*

$$\phi : \quad [\![0,15]\!] \to [\![0,15]\!]^{*^2}$$

$$\begin{cases} 0 \mapsto (14), & 1 \mapsto (13), \qquad 2 \mapsto (12,\ 10), \qquad 3 \mapsto (11,\ 8), \\ 4 \mapsto (14,\ 7), & 5 \mapsto (13,\ 7), \qquad 6 \mapsto (12,\ 7), \qquad 7 \mapsto \begin{pmatrix} 6 \\ 12 \end{pmatrix}, \\ 8 \mapsto \begin{pmatrix} 3 \\ 14 \end{pmatrix}, & 9 \mapsto \begin{pmatrix} 3 \\ 13 \end{pmatrix}, \qquad 10 \mapsto \begin{pmatrix} 2 \\ 12 \end{pmatrix}, \qquad 11 \mapsto \begin{pmatrix} 6 & 1 \\ 12 & 10 \end{pmatrix}, \\ 12 \mapsto \begin{pmatrix} 6 & 1 \\ 11 & 8 \end{pmatrix}, & 13 \mapsto \begin{pmatrix} 5 & 1 \\ 15 & 9 \end{pmatrix}, \quad 14 \mapsto \begin{pmatrix} 4 & 1 \\ 11 & 8 \end{pmatrix}, \quad 15 \mapsto \begin{pmatrix} 2 & 0 \\ 12 & 7 \end{pmatrix}. \end{cases} \tag{7.2}$$

7.3.1 Proof of Theorem 7.3.2

We approach the Wang shifts $\Omega_{\mathcal{U}}$ and $\Omega_{\mathcal{Z}}$ as subshifts; see Section 7.2. The subshifts $\Omega_{\mathcal{U}}$ and $\Omega_{\mathcal{Z}}$ are equipped with the metric dist. A map $\theta : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$ is continuous if for all sequences $(x_n)_{n=0}^{+\infty} \subset \Omega_{\mathcal{U}}$ and all $x \in \Omega_{\mathcal{U}}$, it holds that $\lim_{n\to+\infty} \theta(x_n) = \theta(x)$ in $\Omega_{\mathcal{Z}}$ if and only if $\lim_{n\to+\infty} x_n = x$ in $\Omega_{\mathcal{U}}$. A homeomorphism is a map $\theta : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$ which fulfills the condition that θ is a continuous bijection, and the inverse map $\theta^{-1} : \Omega_{\mathcal{Z}} \to \Omega_{\mathcal{U}}$ is continuous. We show that $\Omega_{\mathcal{U}}$ and $\Omega_{\mathcal{Z}}$ are topologically conjugate by discovering a homeomorphism $\Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$ which has the commuting property. As $\Omega_{\mathcal{Z}}$ and $\Omega_{\mathcal{U}}$ are compact metric spaces, we will use the following observation during the proof. It reformulates Theorem 4.17 from [Rud76].

Lemma 7.3.3. [Rud76] *Suppose θ is a continuous bijection of a compact metric space X onto a metric space Y. Then the inverse map θ^{-1} defined on Y by $\theta^{-1}(\theta(x)) = x$, for $x \in X$, is a continuous map of Y onto X.*

We have a simple observation concerning the language $\mathcal{L}(\Omega_{\mathcal{Z}})$ of the subshift $\Omega(\mathcal{Z})$. If a horizontal pattern is in the language $\mathcal{L}(\Omega_{\mathcal{Z}})$, then it is allowed. The reverse, however, is not true. In the following lemma, we show some horizontal patterns, which are allowed, but which do not occur in the language $\mathcal{L}(\Omega_{\mathcal{Z}})$.

Lemma 7.3.4. *The horizontal patterns (z_3, z_2), (z_3, z_3), (z_8, z_7), (z_8, z_{10}), (z_8, z_{15}), (z_9, z_7), (z_9, z_{10}), (z_9, z_{15}), (z_{12}, z_{12}) do not occur in $\mathcal{L}(\Omega_{\mathcal{Z}})$.*

Proof. We show the statement case by case based on Figure 7.6.

- (z_3, z_2): the north label of both tiles is M and the only tile with the south label M is z_{13}. The pattern (z_{13}, z_{13}) is forbidden, hence $(z_3, z_2) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$.

- (z_3, z_3): the same argument as for the pattern (z_3, z_2).

- (z_8, z_7): the south label of both tiles is O. The set of tiles having O as a north label is $\{z_0, z_1, z_7\}$. For all $i, j \in \{0, 1, 7\}$, the pattern (z_i, z_j) is forbidden. Hence $(z_8, z_7) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$.

- (z_8, z_{10}): the same argument as for the pattern (z_8, z_7).

- (z_8, z_{15}): the north labels of this pattern read LN. The set of tiles having L as a south label is $\{z_1, z_9\}$ and the only tile with the south label N is z_5. The patterns (z_1, z_5) and (z_9, z_5) are forbidden, Hence, $(z_8, z_{15}) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$.

- (z_9, z_{15}): the same argument as for the pattern (z_8, z_{15}).

- (z_9, z_7): assume by contradiction that $(z_9, z_7) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. The south labels of this pattern read LO. The tiles having L as the north label are $\{z_8, z_9, z_{10}\}$ and the tiles having O as the north label are $\{z_0, z_1, z_7\}$. The only horizontal patterns from the set $\{z_8, z_9, z_{10}\} \times \{z_0, z_1, z_7\}$, which are allowed, are (z_8, z_7) and (z_9, z_7). We showed in a previous part that $(z_8, z_7) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$. Thus $\left(\begin{smallmatrix} z_9 & z_7 \\ z_9 & z_7 \end{smallmatrix}\right) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. The west label of z_9 is C and the only tile with east label C is z_{15}. Thus $\left(\begin{smallmatrix} z_{15} \\ z_{15} \end{smallmatrix}\right) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. As the pattern $\left(\begin{smallmatrix} z_{15} \\ z_{15} \end{smallmatrix}\right)$ is forbidden, we have a contradiction.

- (z_9, z_{10}): an analogous argument as for the pattern (z_9, z_7).

- (z_{12}, z_{12}): the south labels of this pattern read KK. The set of tiles having K as a north label is $\{z_6, z_{13}, z_{14}\}$. For all $i, j \in \{6, 13, 14\}$, the pattern (z_i, z_j) is forbidden. Hence, $(z_{12}, z_{12}) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$. $\square$

Let $m_{\mathcal{UZ}} : \mathcal{U} \to \mathcal{Z}$ be a map, which assigns to a tile $u \in \mathcal{U}$ such a tile $z \in \mathcal{Z}$ that z is created from u by unifying the labels $F \mapsto D$ and $G \mapsto I$. We show the map $m_{\mathcal{UZ}}$ in Figure 7.7 and we observe some of its properties in the following lemma.

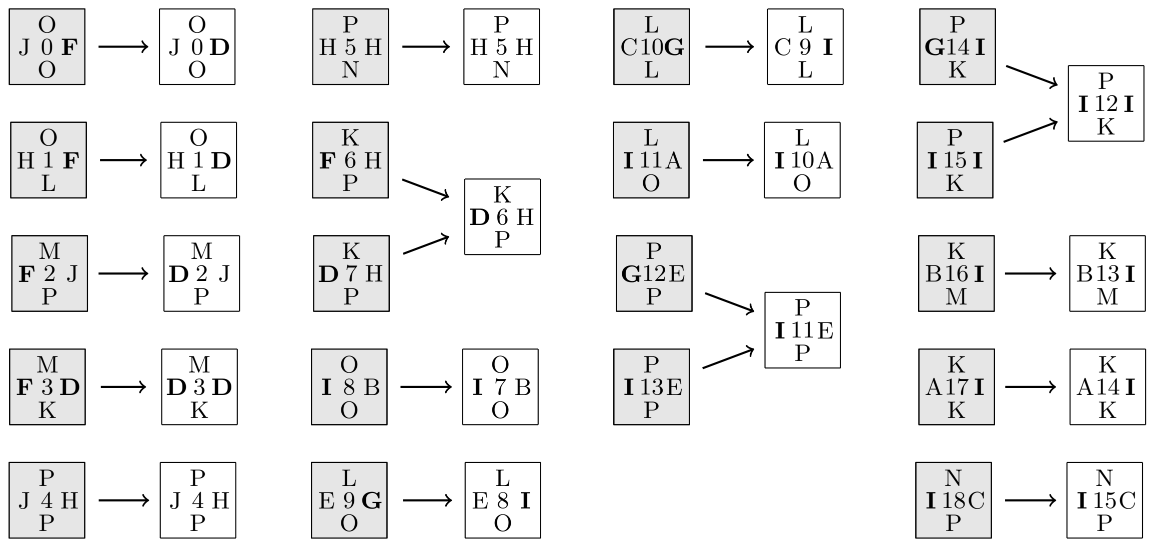


Figure 7.7: The map $m_{\mathcal{UZ}} : \mathcal{U} \to \mathcal{Z}$ that maps a Wang tile $u \in \mathcal{U}$ to a Wang tile $z \in \mathcal{Z}$ so that the tile z is created from u by replacing $F \mapsto D$ and $G \mapsto I$. We draw the labels F, D, G, I in bold. The tiles $u \in \mathcal{U}$ are shown in gray color to avoid confusion with the tiles $z \in \mathcal{Z}$.

Lemma 7.3.5. *Let $z, \tilde{z} \in \mathcal{Z}$ and $u, v \in \mathcal{U}$. The map $m_{\mathcal{UZ}} : \mathcal{U} \to \mathcal{Z}$ has the following properties.*

i) The preimage $m_{\mathcal{UZ}}^{-1}(z)$ contains one or two elements. It contains two elements if and only if $z \in \{z_6, z_{11}, z_{12}\}$.

ii) If $m_{\mathcal{UZ}}(u) = m_{\mathcal{UZ}}(v)$ and $u \neq v$, then $\{u, v\} \in \{\{u_6, u_7\}, \{u_{12}, u_{13}\}, \{u_{14}, u_{15}\}\}$, and u and v differ exactly on their west label.

iii) If (u, v) is allowed, then $(m_{\mathcal{UZ}}(u), m_{\mathcal{UZ}}(v))$ is allowed.

iv) If $z, \tilde{z} \notin \{z_6, z_{11}, z_{12}\}$ are such that $(z, \tilde{z}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$, then $(m_{\mathcal{UZ}}^{-1}(z), m_{\mathcal{UZ}}^{-1}(\tilde{z}))$ is allowed.

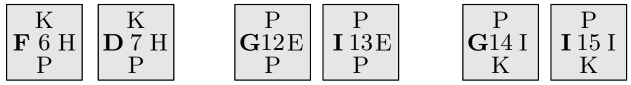
v) If $z \in \mathcal{Z}$ and $\tilde{z} \in \{z_6, z_{12}\}$ are such that $(z, \tilde{z}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$, then $m_{\mathcal{UZ}}^{-1}(z)$ contains one element and there exists a unique $v \in m_{\mathcal{UZ}}^{-1}(\tilde{z})$ such that $(m_{\mathcal{UZ}}^{-1}(z), v)$ is allowed.

vi) If $z \in \mathcal{Z} \setminus \{z_{12}\}$ is such that $(z, z_{11}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$, then $m_{\mathcal{UZ}}^{-1}(z)$ contains one element and there exists a unique $v \in m_{\mathcal{UZ}}^{-1}(z_{11})$ such that $(m_{\mathcal{UZ}}^{-1}(z), v)$ is allowed.

vii) If $z \in \mathcal{Z}$ is such that $(z, z_{12}, z_{11}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$, then $m_{\mathcal{UZ}}^{-1}(z)$ contains one element and there exists a unique pair $u \in m_{\mathcal{UZ}}^{-1}(z_{12})$ and $v \in m_{\mathcal{UZ}}^{-1}(z_{11})$ such that $(m_{\mathcal{UZ}}^{-1}(z), u, v)$ is allowed.

Proof. i): This observation is trivial from Figure 7.7.

ii): From Figure 7.5, the unordered pairs of tiles $u, v \in \mathcal{U}$ such that $u \neq v$ and $m_{\mathcal{UZ}}(u) = m_{\mathcal{UZ}}(v)$ are the following:



We observe that they differ exactly on their west label.

iii): If (u, v) is an allowed horizontal pattern with respect to $\mathcal{U}$, then the east label of u equals to the west label of v. As $m_{\mathcal{U}\mathcal{Z}}$ is a map, we have that the east label of $m_{\mathcal{U}\mathcal{Z}}(u)$ equals to the west label of $m_{\mathcal{U}\mathcal{Z}}(v)$. Thus the horizontal pattern $(m_{\mathcal{U}\mathcal{Z}}(u), m_{\mathcal{U}\mathcal{Z}}(v))$ is allowed.

iv): From part i), we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ and $m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z})$ contain one element. As $(z, \tilde{z})$ is an allowed horizontal pattern, we have that the east label of z equals to the west label of $\tilde{z}$. Denote this label R. If $R \notin \{D, I\}$, then $(m_{\mathcal{U}\mathcal{Z}}^{-1}(z), m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z}))$ is allowed as the map $m_{\mathcal{U}\mathcal{Z}}$ acts as the identity on R. Assume $R \in \{D, I\}$.

- $R = D$: By inspection of the tiles $\mathcal{Z}$ in Figure 7.6, we have that $(z, \tilde{z}) \in \{z_0, z_1, z_3\} \times \{z_2, z_3\}$ (recall that we do not consider z_6 by assumption). By Lemma 7.3.4, $(z, \tilde{z}) \in \{z_0, z_1\} \times \{z_2, z_3\}$. Then the east label of $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ equals to the west label of $m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z})$; see Figure 7.7.

- $R = I$: By inspection of the tiles $\mathcal{Z}$ in Figure 7.6, we have that $(z, \tilde{z}) \in \{z_8, z_9, z_{13}, z_{14}\} \times \{z_7, z_{10}, z_{15}\}$ (recall that we do not consider z_{11} and z_{12} by assumption). By Lemma 7.3.4, $(z, \tilde{z}) \in \{z_{13}, z_{14}\} \times \{z_7, z_{10}, z_{15}\}$. Then the east label of $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ equals to the west label of $m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z})$; see Figure 7.7.

v): We assume that $z \in \mathcal{Z}$ and $\tilde{z} \in \{6, 12\}$ are such that $(z, \tilde{z}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$.

- Assume $\tilde{z} = z_6$. From Figure 7.7, we have $m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z}) = \{u_6, u_7\}$ and we observe that the tile u_6 has the west label F and the tile u_7 has the west label D. As $(z, \tilde{z})$ is allowed, we have from Figure 7.6 that $z \in \{z_0, z_1, z_3\}$. By Lemma 7.3.5 i), $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ contains one element. If $z \in \{z_0, z_1\}$, then $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label F and we choose $v = u_6$. If $z = z_3$, then $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label D and we choose $v = u_7$.

- Assume $\tilde{z} = z_{12}$. From Figure 7.7, we have $m_{\mathcal{U}\mathcal{Z}}^{-1}(\tilde{z}) = \{u_{14}, u_{15}\}$ and we observe that the tile u_{14} has the west label G and the tile u_{15} has the west label I. From Figure 7.6, we have $z \in \{z_8, z_9, z_{12}, z_{13}, z_{14}\}$. By Lemma 7.3.4, the horizontal pattern $(z_{12}, z_{12}) \notin \mathcal{L}(\Omega_{\mathcal{Z}})$. Thus $z \in \{z_8, z_9, z_{13}, z_{14}\}$ and by Lemma 7.3.5 i), $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ contains one element. If $z = \{z_8, z_9\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label G and we choose $v = u_{14}$. If $z = \{z_{13}, z_{14}\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label I and we choose $v = u_{15}$.

vi): We assume that $z \in \mathcal{Z} \setminus \{z_{12}\}$ is such that $(z, z_{11}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. From Figure 7.7, we have $m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{11}) = \{u_{12}, u_{13}\}$. By inspection of the tiles $\mathcal{Z}$ in Figure 7.6, we have $z \in \{z_8, z_9, z_{13}, z_{14}\}$, because $z \neq z_{12}$ by assumption. By part i), $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ contains one element. If $z \in \{z_8, z_9\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label G and we choose $v = u_{12}$. If $z \in \{z_{13}, z_{14}\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label I and we choose $v = u_{13}$.

vii): We assume that $z \in \mathcal{Z}$ is such that $(z, z_{12}, z_{11}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. From Figure 7.7, we have $m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{11}) = \{u_{12}, u_{13}\}$. By inspection of the tiles $\mathcal{Z}$ in Figure 7.6, we have $z \in \{z_8, z_9, z_{12}, z_{13}, z_{14}\}$ and, using Lemma 7.3.4, we have $z \in \{z_8, z_9, z_{13}, z_{14}\}$. Consequently, $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ contains one element. From Figure 7.7, we observe that both elements in $m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{12}) = \{u_{14}, u_{15}\}$ have the same east label I, whereas the tile $u_{12} \in m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{11})$ has the west label G and the tile $u_{13} \in m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{11})$ has the west label I. Therefore we have a unique choice $v = u_{13}$ so that (u, v) is allowed, for every $u \in m_{\mathcal{U}\mathcal{Z}}^{-1}(z_{12})$. If $z \in \{z_8, z_9\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label G and we choose $u = u_{12}$. If $z \in \{z_{13}, z_{14}\}$, then we have that $m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$ has the east label I and we choose $u = u_{13}$. $\square$

We define a map, which maps a valid Wang configuration over $\mathcal{U}$ to a Wang configuration over $\mathcal{Z}$ by merging labels $F \mapsto D$ and $G \mapsto I$:

$$\begin{aligned} \psi : \Omega_{\mathcal{U}} &\to \mathcal{Z}^{\mathbb{Z}^2} \\ y &\mapsto x \quad \text{ such that } x_{\boldsymbol{n}} = m_{\mathcal{U}\mathcal{Z}}(y_{\boldsymbol{n}}), \text{ for every } \boldsymbol{n} \in \mathbb{Z}^2. \end{aligned}$$

We observe in the following lemma that the resulting Wang configuration is valid.

Lemma 7.3.6. *The map ψ fulfills the condition that $\psi : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$.*

Proof. It follows directly from Lemma 7.3.5 iii). □

As a counterpart to the map ψ, we define a map ρ, which maps a valid Wang configuration $x \in \Omega_{\mathcal{Z}}$ to a Wang configuration over $\mathcal{U}$ in the following way:

$$\rho : \Omega_{\mathcal{Z}} \to \mathcal{U}^{\mathbb{Z}^2}$$

$$x \mapsto y \quad \text{ such that } y_{\boldsymbol{n}} = \begin{cases} m_{\mathcal{U}\mathcal{Z}}^{-1}(x_{\boldsymbol{n}}), & \text{if } x_{\boldsymbol{n}} \notin \{6, 11, 12\}; \\ \theta_1(x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}}), & \text{if } x_{\boldsymbol{n}} \in \{6, 12\}; \\ \theta_2(x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}}), & \text{if } x_{\boldsymbol{n}} = 11 \text{ and } x_{\boldsymbol{n}-\mathbf{e}_1} \neq 12; \\ \theta_3(x_{\boldsymbol{n}-2\mathbf{e}_1}, x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}}), & \text{if } x_{\boldsymbol{n}} = 11 \text{ and } x_{\boldsymbol{n}-\mathbf{e}_1} = 12; \end{cases}$$

where $\theta_1 : \mathcal{Z} \times \{6, 11, 12\} \to \mathcal{U}$ maps $(x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}})$ to the unique tile v from Lemma 7.3.5 v), $\theta_2 : \mathcal{Z} \setminus \{12\} \times \{11\} \to \mathcal{U}$ maps $(x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}})$ to the unique tile v from Lemma 7.3.5 vi), and $\theta_3 : \mathcal{Z} \times \{12\} \times \{11\} \to \mathcal{U}$ maps $(x_{\boldsymbol{n}-2\mathbf{e}_1}, x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}})$ to the unique tile v from Lemma 7.3.5 vii).

Remark 7.3.7. *The map ρ fulfills the condition that, for every $\boldsymbol{n} \in \mathbb{Z}^2$, we have $y_{\boldsymbol{n}} \in m_{\mathcal{U}\mathcal{Z}}^{-1}(x_{\boldsymbol{n}})$.*

Lemma 7.3.8. *The map ρ fulfills the condition that $\rho : \Omega_{\mathcal{Z}} \to \Omega_{\mathcal{U}}$.*

Proof. We prove that $\rho(\Omega_{\mathcal{Z}}) \subset \Omega_{\mathcal{U}}$, i.e., all Wang configurations in $\rho(\Omega_{\mathcal{Z}})$ are valid. Let $x \in \Omega_{\mathcal{Z}}$ and let $y = \rho(x)$. Let $\boldsymbol{n} \in \mathbb{Z}^2$. Then $(x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}}) \in \mathcal{L}(\Omega_{\mathcal{Z}})$. We prove that $(y_{\boldsymbol{n}-\mathbf{e}_1}, y_{\boldsymbol{n}})$ is allowed.

- If $x_{\boldsymbol{n}-\mathbf{e}_1}, x_{\boldsymbol{n}} \notin \{z_6, z_{11}, z_{12}\}$, then $(y_{\boldsymbol{n}-\mathbf{e}_1}, y_{\boldsymbol{n}})$ is allowed by Lemma 7.3.5 iv).
- If $x_{\boldsymbol{n}} \in \{z_6, z_{12}\}$, then $(y_{\boldsymbol{n}-\mathbf{e}_1}, y_{\boldsymbol{n}})$ is allowed by Lemma 7.3.5 v).
- If $x_{\boldsymbol{n}} = z_{11}$ and $x_{\boldsymbol{n}-\mathbf{e}_1} \neq z_{12}$, then $(y_{\boldsymbol{n}-\mathbf{e}_1}, y_{\boldsymbol{n}})$ is allowed by Lemma 7.3.5 vi).
- If $x_{\boldsymbol{n}} = z_{11}$ and $x_{\boldsymbol{n}-\mathbf{e}_1} = z_{12}$, then $(y_{\boldsymbol{n}-\mathbf{e}_1}, y_{\boldsymbol{n}})$ is allowed by Lemma 7.3.5 vii). □

Theorem 7.3.9. *The Wang shifts $\Omega_{\mathcal{Z}}$ and $\Omega_{\mathcal{U}}$ are topologically conjugate.*

Proof. We show that $\psi : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$ is continuous, injective and onto $\Omega_{\mathcal{Z}}$, that ψ commutes with the shift action σ and that there exists an inverse map ψ^{-1}. This will finish the proof as $\Omega_{\mathcal{U}}$ and $\Omega_{\mathcal{Z}}$ are compact metric spaces and the inverse map of a continuous bijection between two compact metric spaces is continuous; see Lemma 7.3.3.

We show that ψ *is onto* $\Omega_{\mathcal{Z}}$ by showing that ρ is the right inverse of ψ. Assume $x \in \Omega_{\mathcal{Z}}$. From Lemma 7.3.8 and Lemma 7.3.6, we have $\psi(\rho(x)) \in \Omega_{\mathcal{Z}}$. For every tile $z \in \mathcal{Z}$, we have that $m_{\mathcal{U}\mathcal{Z}}(v) = z$, for every $v \in m_{\mathcal{U}\mathcal{Z}}^{-1}(z)$. Thus, by Remark 7.3.7, we have $\psi(\rho(x)) = x$.

We show that ψ *is injective.* Assume by contradiction that there exist $y, y' \in \Omega_{\mathcal{U}}$ such that $y \neq y'$ and $\psi(y) = \psi(y')$. Denote $x = \psi(y)$ and $x' = \psi(y')$. Let $\boldsymbol{n} \in \mathbb{Z}^2$ such that $y_{\boldsymbol{n}} \neq y'_{\boldsymbol{n}}$. Thus $x_{\boldsymbol{n}} = x'_{\boldsymbol{n}}$ and from the definition of the map ψ, we have $m_{\mathcal{U}\mathcal{Z}}(y_{\boldsymbol{n}}) = m_{\mathcal{U}\mathcal{Z}}(y'_{\boldsymbol{n}})$. From Lemma 7.3.5 ii), we have that $\{y_{\boldsymbol{n}}, y'_{\boldsymbol{n}}\} \in \{\{u_6, u_7\}, \{u_{12}, u_{13}\}, \{u_{14}, u_{15}\}\}$. This implies one of the following three cases, where we assume without loss of generality that the pair $\{y_{\boldsymbol{n}}, y'_{\boldsymbol{n}}\}$ is ordered.

- $y_{\boldsymbol{n}} = u_6, y'_{\boldsymbol{n}} = u_7$: From Figure 7.5, we have $y_{\boldsymbol{n}-\mathbf{e}_1} \in \{u_0, u_1\}$ and $y'_{\boldsymbol{n}-\mathbf{e}_1} = u_3$. From Figure 7.7, we have that $m_{\mathcal{UZ}}(y_{\boldsymbol{n}-\mathbf{e}_1}) \in \{z_0, z_1\}$ and $m_{\mathcal{UZ}}(y'_{\boldsymbol{n}-\mathbf{e}_1}) = z_3$. Thus $m_{\mathcal{UZ}}(y_{\boldsymbol{n}-\mathbf{e}_1}) \neq m_{\mathcal{UZ}}(y'_{\boldsymbol{n}-\mathbf{e}_1})$ and $\psi(y) \neq \psi(y')$, which is a contradiction.

- $y_{\boldsymbol{n}} = u_{12}, y'_{\boldsymbol{n}} = u_{13}$: From Figure 7.5, we have that $y_{\boldsymbol{n}-\mathbf{e}_1} \in \{u_9, u_{10}\}$ and $y'_{\boldsymbol{n}-\mathbf{e}_1} \in \{u_{14}, u_{15}, u_{16}, u_{17}\}$. From Figure 7.7, we have $m_{\mathcal{UZ}}(y_{\boldsymbol{n}-\mathbf{e}_1}) \in \{z_8, z_9\}$ and $m_{\mathcal{UZ}}(y'_{\boldsymbol{n}-\mathbf{e}_1}) \in \{z_{12}, z_{13}, z_{14}\}$. Thus $m_{\mathcal{UZ}}(y_{\boldsymbol{n}-\mathbf{e}_1}) \neq m_{\mathcal{UZ}}(y'_{\boldsymbol{n}-\mathbf{e}_1})$ and $\psi(y) \neq \psi(y')$, a contradiction.

- $y_{\boldsymbol{n}} = u_{14}, y'_{\boldsymbol{n}} = u_{15}$: The east label of u_{12} and u_{14} coincide and the east label of u_{13} and u_{15} coincide; see Figure 7.5. Therefore, the same argument applies as in the previous case.

We show that ρ *is the inverse of* ψ. As the map ψ is injective and onto $\Omega_{\mathcal{Z}}$, it has an inverse map $\psi^{-1} : \Omega_{\mathcal{Z}} \to \Omega_{\mathcal{U}}$ equal to its right inverse map, which is ρ. Thus $\psi^{-1} = \rho$.

We show that ψ *commutes with* σ. Let $x \in \Omega_{\mathcal{U}}$ and denote $y = \psi(x) \in \Omega_{\mathcal{Z}}$. Let $\boldsymbol{n} \in \mathbb{Z}^2$. Then, for every $\boldsymbol{m} \in \mathbb{Z}^2$, we have

$$\begin{aligned}(\psi(\sigma^{\boldsymbol{n}}(x)))_{\boldsymbol{m}} &= m_{\mathcal{UZ}}((\sigma^{\boldsymbol{n}}(x))_{\boldsymbol{m}}) \\ &= m_{\mathcal{UZ}}(x_{\boldsymbol{m}+\boldsymbol{n}}) \\ &= y_{\boldsymbol{m}+\boldsymbol{n}} \\ &= (\sigma^{\boldsymbol{n}}(y))_{\boldsymbol{m}} \\ &= (\sigma^{\boldsymbol{n}}(\psi(x)))_{\boldsymbol{m}}.\end{aligned}$$

Thus $\psi \circ \sigma^{\boldsymbol{n}} = \sigma^{\boldsymbol{n}} \circ \psi$, for every $\boldsymbol{n} \in \mathbb{Z}^2$.

We show that ψ *is continuous.* If x, $y \in \Omega_{\mathcal{U}}$, then

$$\operatorname{dist}(\psi(x), \psi(y)) \leq \operatorname{dist}(x, y).$$

Thus if $(x_n)_{n\in\mathbb{N}}$ is a sequence of tilings $x_n \in \Omega_{\mathcal{U}}$ such that $x = \lim_{n\to\infty} x_n$ exists, then $\psi(x) = \lim_{n\to\infty} \psi(x_n)$ and ψ is continuous. □

From Theorem 7.3.1 and Theorem 7.3.9, we deduce the proof of Theorem 7.3.2.

Proof of Theorem 7.3.2. We prove that $\Omega_{\mathcal{Z}}$ is minimal, aperiodic and self-similar. We use the following observations in the proof. We denote $\theta : \Omega_{\mathcal{U}} \to \Omega_{\mathcal{Z}}$ the homeomorphism derived in Theorem 7.3.9 and we denote $\theta^{-1} : \Omega_{\mathcal{Z}} \to \Omega_{\mathcal{U}}$ its inverse map. Thus we have $\theta(\Omega_{\mathcal{U}}) = \Omega_{\mathcal{Z}}$ and $\theta^{-1}(\Omega_{\mathcal{Z}}) = \Omega_{\mathcal{U}}$, which implies the observations (1), (2), (3) and (4). The observation (5) is the commuting property of θ. From Theorem 7.3.1, we have a 2-dimensional morphism Ψ such that $\Omega_{\mathcal{U}} = \overline{\Psi(\Omega_{\mathcal{U}})}^{\sigma}$ and thus $\Omega_{\mathcal{U}} = \{\sigma^{\boldsymbol{n}}(\Psi(y)) : \boldsymbol{n} \in \mathbb{Z}^2, y \in \Omega_{\mathcal{U}}\}$. This implies observation (6).

(1) $x \in \Omega_{\mathcal{Z}}$ if and only if $\theta^{-1}(x) \in \Omega_{\mathcal{U}}$,

(2) $y \in \Omega_{\mathcal{U}}$ if and only if there exists $x' \in \Omega_{\mathcal{Z}}$ so that $y = \theta^{-1}(x')$,

(3) for every $y, y' \in \Omega_{\mathcal{U}}$, $y = y'$ if and only if $\theta(y) = \theta(y')$,

(4) for every $x, x' \in \Omega_{\mathcal{Z}}$, $x = x'$ if and only if $\theta^{-1}(x) = \theta^{-1}(x')$,

(5) $\theta^{-1} \circ \sigma^{\boldsymbol{n}} = \sigma^{\boldsymbol{n}} \circ \theta^{-1}$, for every $\boldsymbol{n} \in \mathbb{Z}^2$,

(6) $y \in \Omega_{\mathcal{U}}$ if and only if there exists $y' \in \Omega_{\mathcal{U}}$ and $\boldsymbol{n} \in \mathbb{Z}^2$ so that $y = \sigma^{\boldsymbol{n}}(\Psi(y'))$.

(Minimality): Let $Y \subset \Omega_\mathcal{Z}$ be a subshift. Then $\theta^{-1}(Y) \subset \theta^{-1}(\Omega_\mathcal{Z}) = \Omega_\mathcal{U}$. As θ^{-1} is a continuous map, we have that $\theta^{-1}(Y)$ is a closed shift-invariant set – a subshift. By Theorem 7.3.1, $\Omega_\mathcal{U}$ is minimal and thus it contains no nonempty proper subshift. Therefore $\theta^{-1}(Y)$ is either empty or $\theta^{-1}(Y) = \Omega_\mathcal{U}$. If $\theta^{-1}(Y) = \emptyset$, then $Y = \emptyset$. If $\theta^{-1}(Y) = \Omega_\mathcal{U}$, then $Y = \Omega_\mathcal{Z}$. Hence $\Omega_\mathcal{Z}$ is minimal.

(Aperiodicity): Assume by contradiction that $x \in \Omega_\mathcal{Z}$ is a periodic configuration, i.e., there exists $\boldsymbol{n} \in \mathbb{Z}^2$ such that $x = \sigma^{\boldsymbol{n}}(x)$. Applying θ^{-1} and using (4) and (5), we have

$$\theta^{-1}(x) = \theta^{-1}(\sigma^{\boldsymbol{n}}(x)) = \sigma^{\boldsymbol{n}}(\theta^{-1}(x)).$$

Together with (1), we have that $\theta^{-1}(x) \in \Omega_\mathcal{U}$ is a periodic configuration. This is a contradiction as $\Omega_\mathcal{U}$ is an aperiodic Wang shift; see Theorem 7.3.1.

(Self-similarity): We aim to prove that there exists a 2-dimensional morphism $\phi : \Omega_\mathcal{Z} \to \Omega_\mathcal{Z}$ such that $\Omega_\mathcal{Z} = \overline{\phi(\Omega_\mathcal{Z})}^\sigma$, where the shift-closure is defined as $\overline{\phi(\Omega_\mathcal{Z})}^\sigma = \{\sigma^{\boldsymbol{n}}(\phi(x')) : \boldsymbol{n} \in \mathbb{Z}^2, x' \in \Omega_\mathcal{Z}\}$. We obtain a series of equivalent propositions:

$$\begin{aligned}
x \in \Omega_\mathcal{Z} &\overset{(1)}{\longleftrightarrow} \theta^{-1}(x) \in \Omega_\mathcal{U} \\
&\overset{(6)}{\longleftrightarrow} \text{ there exist } \boldsymbol{n} \in \mathbb{Z}^2 \text{ and } y' \in \Omega_\mathcal{U} \text{ such that } \theta^{-1}(x) = \sigma^{\boldsymbol{n}}(\Psi(y')) \\
&\overset{(3)}{\longleftrightarrow} \text{ there exist } \boldsymbol{n} \in \mathbb{Z}^2 \text{ and } y' \in \Omega_\mathcal{U} \text{ such that } x = \theta(\sigma^{\boldsymbol{n}}(\Psi(y'))) \\
&\overset{(5)}{\longleftrightarrow} \text{ there exist } \boldsymbol{n} \in \mathbb{Z}^2 \text{ and } y' \in \Omega_\mathcal{U} \text{ such that } x = \sigma^{\boldsymbol{n}}(\theta(\Psi(y'))) \\
&\overset{(2)}{\longleftrightarrow} \text{ there exist } \boldsymbol{n} \in \mathbb{Z}^2 \text{ and } x' \in \Omega_\mathcal{Z} \text{ such that } x = \sigma^{\boldsymbol{n}}(\theta(\Psi(\theta^{-1}(x')))) \\
&\leftrightarrow x \in \overline{\phi(\Omega_\mathcal{Z})}^\sigma,
\end{aligned}$$

where $\phi : \Omega_\mathcal{Z} \to \Omega_\mathcal{Z}$ is the continuous map defined as $\phi = \theta \circ \Psi \circ \theta^{-1}$. Thus $\Omega_\mathcal{Z}$ is self-similar. Moreover, ϕ is well-defined from letters in the alphabet $\mathcal{Z}$ as, for every $y \in \Omega_\mathcal{U}$ and every $\boldsymbol{n} \in \mathbb{Z}^2$, we have $(\theta(y))_{\boldsymbol{n}} = m_{\mathcal{U}\mathcal{Z}}(y_{\boldsymbol{n}})$. Consequently, for every $z \in \mathcal{Z}$, we have $\phi(z) = m_{\mathcal{U}\mathcal{Z}} \circ \Psi \circ m^{-1}_{\mathcal{U}\mathcal{Z}}(z)$ which is well-defined as we show in Figure 7.8. See Equation (7.1) and Figure7.7.

Observe that, indeed, ϕ is the map in Equation (7.2). □

7.4 Automatic characterization of $\Omega_\mathcal{Z}$

The Wang shift $\Omega_\mathcal{Z}$ is a minimal, aperiodic and self-similar Wang shift related to the Jeandel–Rao Wang shift. In this part, we characterize a configuration $x \in \Omega_\mathcal{Z}$ as a 2-dimensional automatic sequence. As the Wang shift $\Omega_\mathcal{Z}$ is minimal, in a certain sense this result characterizes $\Omega_\mathcal{Z}$ as a whole. Note that it is the Fibonacci complement numeration system for $\mathbb{Z}^2$ which is used in the automatic characterization.

Theorem 7.4.1. [LL21] *There exists a deterministic finite automaton with output $\mathcal{A}$ such that the configuration*

$$\begin{array}{rccl} x: & \mathbb{Z}^2 & \to & \{0, 1, \ldots, 15\} \\ & \boldsymbol{n} & \mapsto & \mathcal{A}(\mathrm{rep}_{\mathcal{F}c}(\boldsymbol{n})) \end{array}$$

satisfies the condition that $x \in \Omega_\mathcal{Z}$.

We do not repeat the proof of Theorem 7.4.1 which is present in [LL21]. We show, however, how the configuration x and the automaton $\mathcal{A}$ are constructed. The inspiration for these steps

$$z \xrightarrow{m^{-1}_{\mathcal{UZ}}} u \xrightarrow{\Psi} \Psi(u) \xrightarrow{m_{\mathcal{UZ}}} \phi(z)$$

z	u	$\Psi(u)$	$\phi(z)$
0	0	17	14
1	1	16	13
2	2	$(15, 11)$	$(12, 10)$
3	3	$(13, 9)$	$(11, 8)$
4	4	$(17, 8)$	$(14, 7)$
5	5	$(16, 8)$	$(13, 7)$
6	6	$(15, 8)$	$(12, 7)$
	7	$(14, 8)$	
7	8	$\left(\begin{smallmatrix}6\\14\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}6\\12\end{smallmatrix}\right)$

z	u	$\Psi(u)$	$\phi(z)$
8	9	$\left(\begin{smallmatrix}3\\17\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}3\\14\end{smallmatrix}\right)$
9	10	$\left(\begin{smallmatrix}3\\16\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}3\\13\end{smallmatrix}\right)$
10	11	$\left(\begin{smallmatrix}2\\14\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}2\\12\end{smallmatrix}\right)$
11	12	$\left(\begin{smallmatrix}7&1\\15&11\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}6&1\\12&10\end{smallmatrix}\right)$
	13	$\left(\begin{smallmatrix}6&1\\14&11\end{smallmatrix}\right)$	
12	14	$\left(\begin{smallmatrix}7&1\\13&9\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}6&1\\11&8\end{smallmatrix}\right)$
	15	$\left(\begin{smallmatrix}6&1\\12&9\end{smallmatrix}\right)$	
13	16	$\left(\begin{smallmatrix}5&1\\18&10\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}5&1\\15&9\end{smallmatrix}\right)$
14	17	$\left(\begin{smallmatrix}4&1\\13&9\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}4&1\\11&8\end{smallmatrix}\right)$
15	18	$\left(\begin{smallmatrix}2&0\\14&8\end{smallmatrix}\right)$	$\left(\begin{smallmatrix}2&0\\12&7\end{smallmatrix}\right)$

Figure 7.8: The map $\phi : \Omega_{\mathcal{Z}} \to \Omega_{\mathcal{Z}}$ defined from letters.

comes from [AA20] where a similar procedure was done for a fixed point $\mathbb{N}^2 \to A$ over an alphabet A. Recall that the 2-dimensional morphism ϕ which satisfies the condition that $\Omega_{\mathcal{Z}} = \overline{\phi(\Omega_{\mathcal{Z}})}^{\sigma}$ is in Equation (7.2). Let

$$s = \left(\begin{smallmatrix}8 & 12\\ 1 & 6\end{smallmatrix}\right).$$

We observe that $\phi^2(s)$ prolongates s at the origin.

$$\begin{array}{r|l} 8 & 12 \\ \hline 1 & 6 \end{array} \xrightarrow{\phi} \begin{array}{r|ll} 3 & 6 & 1 \\ 14 & 11 & 8 \\ \hline 13 & 12 & 7 \end{array} \xrightarrow{\phi} \begin{array}{rr|lll} 11 & 8 & 12 & 7 & 13 \\ 4 & 1 & 6 & 1 & 3 \\ 11 & 8 & 12 & 10 & 14 \\ \hline 5 & 1 & 6 & 1 & 6 \\ 15 & 9 & 11 & 8 & 12 \end{array} \xrightarrow{\phi} \cdots$$

In other words, there exists a configuration $x : \mathbb{Z}^2 \to A$ such that $x = \phi^2(x)$ and $x|_{\{-1,0\}\times\{-1,0\}} = s$. As $x = \phi^2(x)$, we have that $x \in \Omega_{\mathcal{Z}}$, i.e., it is a valid Wang configuration. A finite part of the configuration x is shown in Figure 7.9. The tiles $x_{\boldsymbol{n}}$, for every $\boldsymbol{n} \in \{-1, 0\} \times \{-1, 0\}$, are drawn in yellow color. It is the tiles contained in s. The tiles that emerge from s by applying ϕ^2 are drawn in green color. The tiles that emerge from s by applying ϕ^4 are drawn in blue color.

In analogy with the 1-dimensional substitutions, it is possible to associate with ϕ and $a \in A$ a deterministic finite automaton with output $\mathcal{A}_{\phi,a}$ [BR10]. Let $\mathcal{D} = \{\mathtt{0}, \mathtt{1}\} \times \{\mathtt{0}, \mathtt{1}\}$ and $A = \{0, 1, \ldots, 15\}$. The automaton associated with ϕ and $a \in A$ is the 6-tuple $\mathcal{A}_{\phi,a} = (A, \mathcal{D}, \delta_\phi, a, A, \xi)$, where the transition function $\delta_\phi : A \times \mathcal{D} \to A$ is a partial function such that $\delta_\phi(c, e) = b$ for any $c, d \in A$ and $e \in \mathcal{D}$ if and only if d is in $\phi(c)$ at position e and the output

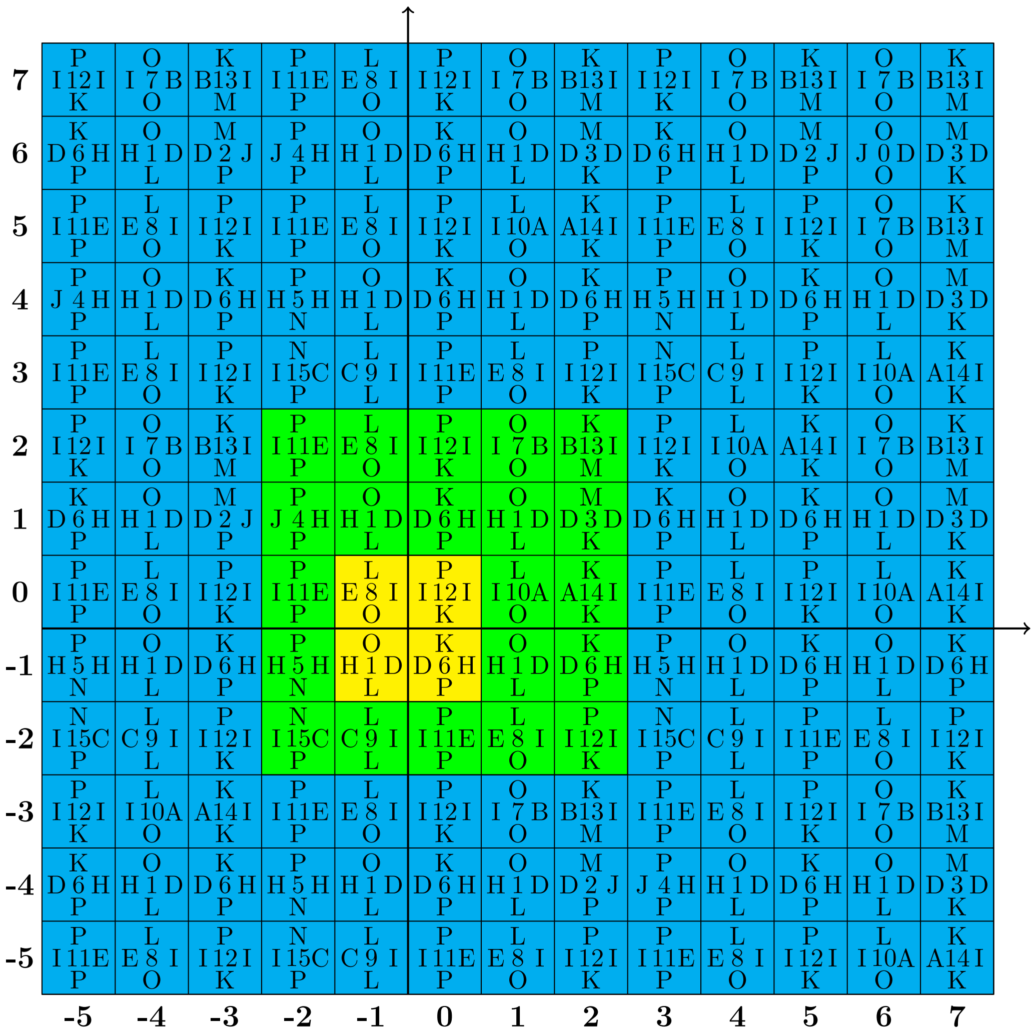


Figure 7.9: A finite part of the valid configuration $x \in \Omega_{\mathcal{Z}}$ from Theorem 7.4.1.

function ξ is the identity map. In analogy with the procedure described in Section 6.3, we modify the automaton $\mathcal{A}_{\phi,a}$ into an automaton $\mathcal{A}_{\phi,s}$ by adding a new initial state `start` and additional edges connecting the new initial state to the letters contained in s. The deterministic finite automaton with output

$$\mathcal{A}_{\phi,s} = (A \cup \{\texttt{start}\}, \mathcal{D}, \delta_{\phi,s}, \texttt{start}, A, \xi)$$

has the transition function $\delta_{\phi,s}$ such that

- $\delta_{\phi,s}(\texttt{start}, \binom{0}{0}) = 12$, $\delta_{\phi,s}(\texttt{start}, \binom{0}{1}) = 8$,
- $\delta_{\phi,s}(\texttt{start}, \binom{1}{0}) = 6$, $\delta_{\phi,s}(\texttt{start}, \binom{1}{1}) = 1$,
- $\delta_{\phi,s}|_{A \times \mathcal{D}} = \delta_{\phi}$.

We illustrate in the following example that feeding an $\mathcal{F}c$-representation of a position $\boldsymbol{n} \in \mathbb{Z}^2$ gives the index of the tile in the configuration x at position $\boldsymbol{n}$.

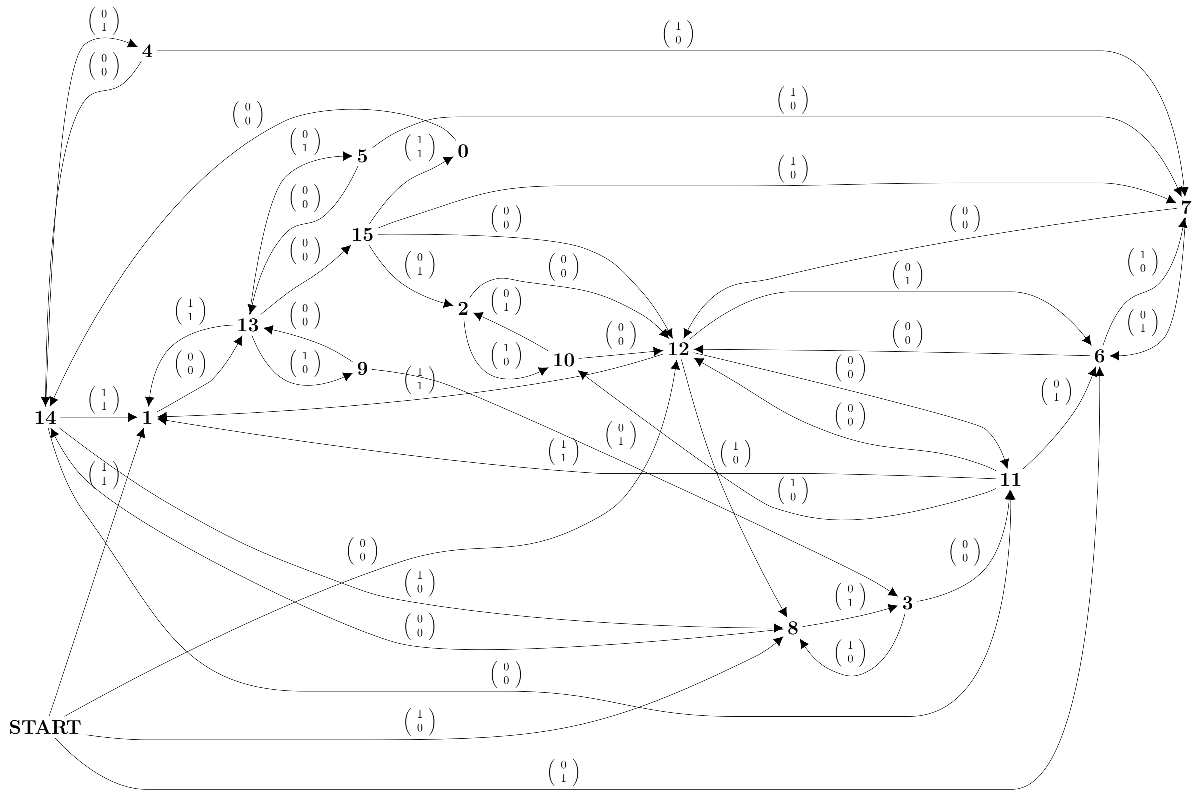


Figure 7.10: The automaton $\mathcal{A}_{\phi,s}$ associated with the 2-dimensional morphism ϕ and $s = \left(\begin{smallmatrix} 8 & 12 \\ 1 & 6 \end{smallmatrix}\right)$.

Example 7.4.2. Let $\boldsymbol{n} = (-1, 6) \in \mathbb{Z}^2$. Then, $\mathrm{rep}_{\mathcal{F}c}(\boldsymbol{n}) = \left(\begin{smallmatrix} \mathtt{10101} \\ \mathtt{01001} \end{smallmatrix}\right)$ and $\mathcal{A}_{\phi,s}$ gives

$$\mathtt{start} \xrightarrow{\binom{1}{0}} 8 \xrightarrow{\binom{0}{1}} 3 \xrightarrow{\binom{1}{0}} 8 \xrightarrow{\binom{0}{0}} 14 \xrightarrow{\binom{1}{1}} 1.$$

The tile at position $\boldsymbol{n}$ in the tiling x in Figure 7.9 is indeed $x_{\boldsymbol{n}} = 1$.

Other Wang shifts with an automatic characterization might exist. For now, we leave this as an open question.

Question 7.4.3. *Are there other minimal aperiodic Wang shifts which have an automatic characterization?*

Chapter 8

Conclusion and open problems

This work embodies three main subjects of interest – combinatorics on words, number theory and discrete geometry. There are remarkable links between these areas. For instance, the Fibonacci word may be defined purely from the point of view of combinatorics on words as the fixed point of the morphism $a \mapsto ab$, $b \mapsto a$. However, it may be described also as a cutting sequence with the slope of the golden mean, which links the combinatorics on words to discrete geometry and number theory. Moreover, the Fibonacci word is an automatic sequence created with the use of the Fibonacci numeration system. In this work, we continued developing these links, but open problems remain to be tackled.

In Chapter 3, we described a class of positional numeration systems for $\mathbb{Z}$ associated with simple Parry numbers. As the two's complement numeration system for $\mathbb{Z}$ belongs to this class, we called the class the complement numeration systems. It would be interesting to generalize this construction to non-simple Parry numbers. Also, an analogue for $\mathbb{Z}$ of the Fibonacci numeration system belongs to this class. We constructed a Mealy machine which performs addition in this numeration system. The question, whether there exists a Mealy machine performing addition in a given complement numeration system for $\mathbb{Z}$, remains open.

Question 3.3.10. *Can we construct complement numeration systems for $\mathbb{Z}$ associated with non-simple Parry numbers?*

Question 3.5.9. *Can we construct a Mealy machine, which performs addition in a given complement numeration system associated with a simple Parry number?*

In Chapter 4, we studied the repetition rate of regular Arnoux–Rauzy words which are a subclass of episturmian words. These words generalize Sturmian words to d-ary alphabets, for $d \geq 2$. There is a monoid of morphisms which map episturmian words to episturmian words. In the case of a binary alphabet, this monoid is called the monoid of Sturm and its submonoid called the special Sturmian monoid $\mathcal{M}$ is of particular interest to us. In Chapter 5, we faithfully represented $\mathcal{M}$ by 3×3 matrices with nonnegative entries. We ask whether the faithful representation may be found in the case of the monoid of episturmian morphisms.

Question 5.3.7. *Can we faithfully represent the monoid of morphisms which map episturmian words to episturmian words?*

In Chapter 6, we approached numeration systems in a broader framework as the regular languages which describe two-sided automatic sequences. We provided a sufficient condition for these numeration systems, which we call Dumont–Thomas, to be positional and we ask naturally what are the necessary conditions. Also, we showed that the Dumont–Thomas numeration systems can be extended naturally to $\mathbb{Z}^d$, for $d \geq 1$. We believe that the Dumont–Thomas numeration systems extended to $\mathbb{Z}^d$ may be used to describe all d-dimensional words $\mathbb{Z}^d \to A$

which are periodic points of d-dimensional substitutions.

Question 6.5.7. *What is the necessary condition for a Dumont–Thomas numeration system for $\mathbb{Z}$ to be positional?*

Question 6.6.3. *Is it possible to generalize Theorem 6.3.1 to $\mathbb{Z}^d$, for $d \geq 1$?*

In Chapter 7, we observed a remarkable link between a particular Wang shift, a numeration system for $\mathbb{Z}^2$ and a 2-dimensional morphism. More precisely, we described a particular Wang configuration as an automatic sequence. Naturally, we ask the question: are there other Wang shifts sharing this property? Our knowledge of Dumont–Thomas numeration systems for $\mathbb{Z}^2$ might be a cornerstone for the future research in this area.

Question7.4.3. *Are there other minimal aperiodic Wang shifts which have an automatic characterization?*

Bibliography

[AA20] Shigeki Akiyama and Pierre Arnoux, editors. *Substitution and Tiling Dynamics: Introduction to Self-inducing Structures*. Springer International Publishing, 2020.

[AR91] Pierre Arnoux and Gérard Rauzy. Représentation géométrique de suites de complexité 2n+1. *Bulletin de la Société mathématique de France*, 119:101–117, 01 1991.

[AS03] Jean-Paul Allouche and Jeffrey Shallit. *Automatic sequences*. Cambridge University Press, Cambridge, 2003. Theory, applications, generalizations.

[Bas02] Frédérique Bassino. *Beta-Expansions for Cubic Pisot Numbers*. January 2002. Pages: 152.

[Ber66] Robert Berger. The undecidability of the domino problem. *Mem. Amer. Math. Soc. No.*, 66:72, 1966.

[Ber86] Jean Berstel. Fibonacci words - a survey. In *The book of L*, pages 13–27. 1986. dedic. A. Lindenmayer Occas. 60th Birthday.

[BG13] Michael Baake and Uwe Grimm. *Aperiodic Order. Vol. 1*, volume 149 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2013.

[BH97] Véronique Bruyère and Georges Hansel. Bertrand numeration systems and recognizability. *Theoretical Computer Science*, 181(1):17–43, July 1997.

[BHMV94] Véronique Bruyère, Georges Hansel, Christian Michaux, and Roger Villemaire. Logic and p-recognizable sets of integers. volume 1, pages 191–238. 1994. Journées Montoises (Mons, 1992).

[BM89] Anne Bertrand-Mathis. Comment ecrire les nombres entiers dans une base qui n'est pas entiere. *Acta Mathematica Hungarica*, 54(3):237–241, September 1989.

[BR10] Valérie Berthé and Michel Rigo, editors. *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2010.

[Car68] L. Carlitz. Fibonacci representations. *Fibonacci Quart.*, 6(4):193–220, 1968.

[Cas08] Julien Cassaigne. On extremal properties of the Fibonacci word. *RAIRO - Theoretical Informatics and Applications - Informatique Théorique et Applications*, 42(4):701–715, 2008.

[CCS22] Émilie Charlier, Célia Cisternino, and Manon Stipulanti. A full characterization of Bertrand numeration systems, February 2022. arXiv:2202.04938 [math].

[CdL00] Arturo Carpi and Aldo de Luca. Special factors, periodicity, and an application to sturmian words. *Acta Inf.*, 36:983–1006, 07 2000.

[CKR10] Emilie Charlier, Tomi Kärki, and Michel Rigo. Multidimensional generalized automatic sequences and shape-symmetric morphic words. *Discrete Math.*, 310(6-7):1238–1252, 2010.

[Cob72] Alan Cobham. Uniform tag sequences. *Math. Systems Theory*, 6:164–192, 1972.

[Cul96] Karel Culik, II. An aperiodic set of 13 Wang tiles. *Discrete Math.*, 160(1-3):245–251, 1996.

[Day60] D. E. Daykin. Representation of natural numbers as sums of generalised Fibonacci numbers. *J. London Math. Soc.*, 35:143–160, 1960.

[DD14] Gregory Dresden and Zhaohui Du. Binet-type formulas for r-generalized Fibonacci numbers. *Journal of Integer Sequences*, 17(4), 2014.

[DDP23] Francesco Dolce, L'ubomíra Dvořáková, and Edita Pelantová. On balanced sequences and their critical exponent. *Theoretical Computer Science*, 939:18–47, 2023.

[DJP01] Xavier Droubay, Jacques Justin, and Giuseppe Pirillo. Episturmian words and some constructions of de luca and rauzy. *Theoretical Computer Science*, 255:539–553, 03 2001.

[DL02] David Damanik and Daniel Lenz. The index of sturmian sequences. *European Journal of Combinatorics*, 23(1):23–29, 2002.

[DL23] L'ubomíra Dvořáková and Jana Lepšová. Critical exponents of regular Arnoux-Rauzy sequences. In *Combinatorics on Words*, volume 13899 of *Lecture Notes in Computer Science*, pages 130–142. Springer, 2023. https://doi.org/10.1007/978-3-031-33180-0_10.

[DP23] L'ubomíra Dvořáková and Edita Pelantová. The repetition threshold of episturmian sequences, 2023. `https://arxiv.org/abs/2309.00988`.

[DPOS22] L'ubomíra Dvořáková, Edita Pelantová, Daniela Opočenská, and Arseny M. Shur. On minimal critical exponent of balanced sequences. *Theoretical Computer Science*, 922:158–169, 2022.

[DT89] Jean-Marie Dumont and Alain Thomas. Systemes de numeration et fonctions fractales relatifs aux substitutions. *Theoret. Comput. Sci.*, 65(2):153–169, 1989.

[Fab95] Stéphane Fabre. Substitutions et β-systèmes de numération. *Theoretical Computer Science*, 137(2):219–236, 1995.

[Fie08] Miroslav Fiedler. *Special matrices and their applications in numerical mathematics*. Dover Publications, Inc., Mineola, NY, second edition, 2008. Translated from the Czech by Petr Přikryl and Karel Segeth.

[Fog02] N. Pytheas Fogg. *Substitutions in Dynamics, Arithmetics and Combinatorics*, volume 1794 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2002. Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel.

[Fro91] Christiane Frougny. Fibonacci representations and finite automata. *IEEE Trans. Inform. Theory*, 37(2):393–399, 1991.

[GJ09] Amy Glen and Jacques Justin. Episturmian words: a survey. *RAIRO - Theoretical Informatics and Applications*, 43(3):403–442, mar 2009.

[Gle07] Amy Glen. Powers in a class of a-strict standard episturmian words. *Theoretical Computer Science*, 380(3):330–354, 2007. Combinatorics on Words.

[GS87] Branko Grünbaum and G. C. Shephard. *Tilings and patterns*. W. H. Freeman and Company, New York, 1987.

[Hol98] M. Hollander. Greedy Numeration Systems and Regularity. *Theory of Computing Systems*, 31(2):111–133, April 1998.

[JP02] Jacques Justin and Giuseppe Pirillo. Episturmian words and episturmian morphisms. *Theor. Comput. Sci.*, 276:281–313, 04 2002.

[JR21] Emmanuel Jeandel and Michaël Rao. An aperiodic set of 11 Wang tiles. *Advances in Combinatorics*, January 2021.

[JV00] J. Justin and L. Vuillon. Return words in sturmian and episturmian words. *RAIRO-Theoret. Inf. Appl.*, 34:343–356, 2000.

[Kar96] Jarkko Kari. A small aperiodic set of Wang tiles. *Discrete Math.*, 160(1-3):259–264, 1996.

[KMPS18] Karel Klouda, Kateřina Medková, Edita Pelantová, and Štěpán Starosta. Fixed points of sturmian morphisms and their derivated words. *Theoretical Computer Science*, 743:23–37, 2018.

[Knu69] Donald E. Knuth. *The art of computer programming. Vol. 2: Seminumerical algorithms*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont, 1969.

[Knu98] Donald E. Knuth. *The art of computer programming. Vol. 2*. Addison-Wesley, Reading, MA, 1998. Seminumerical algorithms, Third edition.

[Knu11] Donald E. Knuth. *The art of computer programming. Vol. 4A. Combinatorial algorithms. Part 1*. Addison-Wesley, Upper Saddle River, NJ, 2011.

[Kur03] Petr Kurka. *Topological and symbolic dynamics*, volume 11 of *Cours Spécialisés [Specialized Courses]*. Société Mathématique de France, Paris, 2003.

[Lab19] Sébastien Labbé. A self-similar aperiodic set of 19 Wang tiles. *Geom. Dedicata*, 201:81–109, 2019.

[Lab20] Sébastien Labbé. Three characterizations of a self-similar aperiodic 2-dimensional subshift. December 2020. `http://arxiv.org/abs/2012.03892`.

[Lab21] Sébastien Labbé. Substitutive Structure of Jeandel-Rao Aperiodic Tilings. *Discrete Comput. Geom.*, 65(3):800–855, 2021.

[Lek52] C. G. Lekkerkerker. Voorstelling van natuurlijke getallen door een som van getallen van Fibonacci. *Simon Stevin*, 29:190–195, 1952.

[Lin12] Peter Linz. *An introduction to formal languages and automata.* Sudbury, MA: Jones & Bartlett Learning, 5th ed. edition, 2012.

[LL21] Sébastien Labbé and Jana Lepšová. A numeration system for Fibonacci-like Wang shifts. In *Combinatorics on words*, volume 12847 of *Lecture Notes in Comput. Sci.*, pages 104–116. Springer, Cham, 2021.

[LL23a] Sébastien Labbé and Jana Lepšová. A Fibonacci analogue of the two's complement numeration system. *RAIRO-Theor. Inf. Appl.*, 57:12, 2023.

[LL23b] Sébastien Labbé and Jana Lepšová. Dumont-Thomas numeration systems for $\mathbb{Z}$, 2023. `https://arxiv.org/abs/2302.14481`.

[Lot02] M. Lothaire. *Algebraic Combinatorics on Words*, volume 90 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 2002.

[LPS23] Jana Lepšová, Edita Pelantová, and Štěpán Starosta. On a faithful representation of sturmian morphisms. *European Journal of Combinatorics*, 110:103707, may 2023.

[Med19] Kateřina Medková. Derived sequences of Arnoux–Rauzy sequences. In Robert Mercaş and Daniel Reidenbach, editors, *Combinatorics on Words*, pages 251–263, Cham, 2019. Springer International Publishing.

[MH40] Marston Morse and Gustav A. Hedlund. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.*, 62:1–42, 1940.

[MPR19] Adeline Massuir, Jarkko Peltomäki, and Michel Rigo. Automatic sequences based on Parry or Bertrand numeration systems. *Adv. in Appl. Math.*, 108:11–30, 2019.

[Pel] Edita Pelantová. personal communication.

[Pel21] Jarkko Peltomäki. Initial nonrepetitive complexity of regular episturmian words and their diophantine exponents. *CoRR*, abs/2103.08351, 2021.

[Pen79] Roger Penrose. Pentaplexity a class of non-periodic tilings of the plane. *The Mathematical Intelligencer*, 2:32–37, 1979.

[PS21] Edita Pelantová and Štěpán Starosta. On sturmian substitutions closed under derivation. *Theoretical Computer Science*, 867:128–139, 2021.

[PW17] Jarkko Peltomäki and Markus A. Whiteland. A square root map on Sturmian words. *Electron. J. Combin.*, 24(1):Paper No. 1.54, 50, 2017.

[Rau82] Gérard Rauzy. Nombres algébriques et substitutions. *Bulletin de la Société Mathématique de France*, 110:147–178, 1982.

[Rig14] Michel Rigo. *Formal Languages, Automata and Numeration Systems.* Wiley-ISTE, 2014.

[RM02] Michel Rigo and Arnaud Maes. More on generalized automatic sequences. *J. Autom. Lang. Comb.*, 7(3):351–376, 2002.

[Rud76] W. Rudin. *Principles of Mathematical Analysis*. International series in pure and applied mathematics. McGraw-Hill, 1976.

[Saa10] Kalle Saari. Everywhere α-repetitive sequences and Sturmian words. *European J. Combin.*, 31(1):177–192, 2010.

[Sak87] Jacques Sakarovitch. Easy multiplications. I. The realm of Kleene's theorem. *Inform. and Comput.*, 74(3):173–197, 1987.

[Sen96] Marjorie Senechal. *Quasicrystals and Geometry*. Encyclopedia of Mathematics and its Applications. Cambridge University Press, Cambridge, 1996.

[SMKGS23] David Smith, Joseph Samuel Myers, Craig S. Kaplan, and Chaim Goodman-Strauss. An aperiodic monotile, 2023.

[TW07] Bo Tan and Zhi-Ying Wen. Some properties of the tribonacci sequence. *Eur. J. Comb.*, 28:1703–1719, 08 2007.

[Wan61] Hao Wang. Proving theorems by pattern recognition – II. *Bell System Technical Journal*, 40(1):1–41, January 1961.

[Zec72] E. Zeckendorf. Représentation des nombres naturels par une somme de nombres de Fibonacci ou de nombres de Lucas. *Bull. Soc. Roy. Sci. Liège*, 41:179–182, 1972.