Combinatorial properties of infinite words associated with cut-and-project sequences

L.-S. Guimond[*] Z. Masáková[†] E. Pelantová[‡]

CRM-2762

October 2001

[*]Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montréal (Québec) H3A 2K6, Canada, guimond@math.mcgill.ca

[†]Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Trojanova 13, 120 00 Praha 2, Czech Republic, masakova@km1.fjfi.cvut.cz

[‡]Department of Mathematics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Trojanova 13, 120 00 Praha 2, Czech Republic, pelantova@km1.fjfi.cvut.cz

Abstract

The aim of this article is to study certain combinatorial properties of infinite binary and ternary words associated to cut-and-project sequences. We consider here the cut-and-project scheme in two dimensions with general orientation of the projecting subspaces. We prove that a cut-and-project sequence arising in such a setting has always either two or three types of distances between adjacent points. A cut-and-project sequence thus determines in a natural way a symbolic sequence (infinite word) in two or three letters. According to the complexity the cut and project construction includes words with complexity $n+1$, $n+const.$ and $2n+1$. The words on two letter alphabet have complexity $n+1$ and thus are Sturmian. The ternary words associated to the cut-and-project sequences have complexity $n+const.$ or $2n+1$. A cut and project scheme has three parameters, two of them specifying the projection subspaces, the third one determining the cutting strip. We classify the triples that correspond to combinatorially equivalent infinite words.

1 Words

A finite alphabet is a set of symbols $\mathcal{A} = \{a_1, \dots, a_k\}$. A concatenation w of letters is called a word. The length of a word w is the number of letters from which w is formed. We denote by $\mathcal{A}^*$ the set of words in the alphabet $\mathcal{A}$. A one-way infinite word u is a sequence $u = (u_n)_{n\in\mathbb{N}} = u_1u_2u_3\dots$ with values in $\mathcal{A}$. One may consider also bidirectional infinite words as sequences $u = (u_n)_{n\in\mathbb{Z}}$. In our article we work with bidirectional infinite words. Two words $(u_n)_{n\in\mathbb{Z}}$ in an alphabet $\mathcal{A}$ and $(v_n)_{n\in\mathbb{Z}}$ in an alphabet $\mathcal{B}$ are combinatorially equivalent if there exists a bijection $h\colon \mathcal{A} \to \mathcal{B}$ and an $n_0 \in \mathbb{Z}$ such that $h(u_n) = v_{n+n_0}$ for all $n \in \mathbb{Z}$.

Let $i \in \mathbb{Z}$, $n \in \mathbb{N}$. A concatenation $u_iu_{i+1}\dots u_{i+n-1}$ is called a factor of u of length n. We define the density of a factor w of u of length n as

$$\varrho_w = \lim_{k\to\infty} \frac{\#\{i \in \mathbb{Z}\cap(-k,k) \mid u_iu_{i+1}\dots u_{i+n-1} = w\}}{2k},$$

if the limit exists.

The function that assigns to a positive integer n the number of different factors of length n in an infinite word u is called the complexity of u, usually denoted by $\mathcal{C}$,

$$\mathcal{C}(n) := \#\{u_iu_{i+1}\dots u_{i+n-1} \mid i \in \mathbb{Z}\}.$$

For the complexity of an infinite word in a finite alphabet $\mathcal{A}$ one has

$$1 \le \mathcal{C}(n) \le (\#\mathcal{A})^n.$$

The complexity function is thus a measure of disorder in the infinite word. For more about the complexity function see [2, 4].

Let us recall some facts about one-way infinite words. We say that an infinite word $u = u_1u_2u_3\dots$ in a finite alphabet $\mathcal{A}$ is eventually periodic, if there exist finite words w_0, w_1 in $\mathcal{A}^*$ such that $w = w_0w_1w_1w_1\dots$. A word which is not eventually periodic is called aperiodic. It is well known that a one-way infinite word is eventually periodic if and only if there exists $n \in \mathbb{N}$, such that $\mathcal{C}(n) \le n$, (see [10]). It follows that the most simple aperiodic words are those of complexity $\mathcal{C}(n) = n + 1$. Such words are called Sturmian. A nice survey of Sturmian and related sequences may be found in [10, 8]. Algebraically, every Sturmian word can be written as a sequence of 0's and 1's given by one of the following formulas

$$\begin{aligned} u_n(\alpha,\beta) &= \lfloor (n+1)\alpha + \beta \rfloor - \lfloor n\alpha + \beta \rfloor, \\ u_n(\alpha,\beta) &= \lceil (n+1)\alpha + \beta \rceil - \lceil n\alpha + \beta \rceil. \end{aligned} \tag{1}$$

J. Cassaigne defines in [5] the quasisturmian sequences as such infinite sequences for which there exist integers k and n_0 such that the complexity function is $\mathcal{C}(n) = n + k$ for $n \ge n_0$. An example of a quasisturmain sequence is a sequenece that arises from a Sturmian word if one substitutes every 0 by a finite word w_0 and every 1 by a finite word w_1. Cassaigne has shown that every quasisturmain sequences is, up to a finite prefix, of this kind.

Using the powerful notion of the Rauzy graphs, the words with complexity $2n + 1$ can be divided into four classes according to the maximal indegree and outdegree in the associated sequence of directed graphs. Arnoux and Rauzy gave geometrical characterization of infinite words of type 3–3. We show that cut-and-project sequences are geometrical reprezentations of infinite words of type 2–2.

In order to explain the classification of sequences with complexity $2n + 1$ into four groups, let us recall the definition of a Rauzy graph. A Rauzy graph for a given infinite word u and given integer n is an oriented graph $\Gamma_n = (V_n, E_n)$, where the vertices are represented by factors of length n and the oriented edges are determined by factors of length $n + 1$. An edge e starts in a vertex v_1 and terminates in a vertex v_2 if the factor v_1 of length n is a prefix of the factor e and v_2 is its suffix. Thus $\#V_n = \mathcal{C}(n)$ and $\#E_n = \mathcal{C}(n + 1)$. The number of edges that start (end) in a vertex v is called the outdegree (indegree) of v. For a bidirectional infinite word, every vertex of every graph has both in- and outdegree at least one. For words of complexity $2n+1$ in every graph Γ_n the number of edges is equal to the number of vertices plus 2. Therefore the maximal outdegree for the graph is either 2 or 3. The same statement is true for maximal indegree. The pair maximal indegree – maximal outdegree for words of complexity $2n+1$ can therefore have only four values. Arnoux and Rauzy in [3] show that if for some n_0 the maximal outdegree in the graph Γ_{n_0} is 2, then it is 2 for every graph Γ_n, $n \ge n_0$. The same statement again holds for indegree. In the article the authors find geometrical representation of words with Rauzy graphs of the type 3–3 for every $n \in \mathbb{N}$. The infinite words considered in this article are of the type 2–2, which means that starting from a certain n_0 both maximal indegree and maximal outdegree are equal to 2.

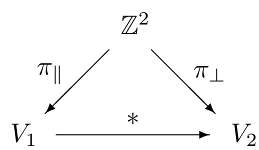
2 Cut-and-project sequences

Generally, cut and project sets are defined as projections of lattices of arbitrary dimensions. In this paper we study the simpliest case. The cut-and-project scheme in $\mathbb{R}^2$ is given by two one-dimensional subspaces V_1 and V_2 and by projections $\pi_{\|}\colon \mathbb{R}^2 \to V_1$ and $\pi_{\perp}\colon \mathbb{R}^2 \to V_2$ which satisfy:

1) $\pi_{\|}$ restricted to $\mathbb{Z}^2$ is a one-to-one mapping.

2) $\pi_{\perp}(\mathbb{Z}^2)$ is dense in V_2.

The scheme is illustrated by the following picture.



In this scheme $\pi_{\|}(\mathbb{Z}^2)$ and $\pi_{\perp}(\mathbb{Z}^2)$ are additive abelian groups. The bijection between them, $\pi_{\perp}^{-1} \circ \pi_{\|}$ is usually denoted by $*$ and called the star map. Its inverse is denoted by $-*$.

Let V_1 be the linear span of a vector $\vec{x_1} = (1, \varepsilon)$ and V_2 the linear span of a vector $\vec{x_2} = (-1, \eta)$. In order to satisfy conditions 1) and 2) we choose ε, η irrational and $\varepsilon \neq -\eta$. Any vector $(p, q) \in \mathbb{Z}^2$ can be written as

$$(p, q) = \frac{1}{\varepsilon + \eta}(q + \eta p)\vec{x}_1 + \frac{1}{\varepsilon + \eta}(q - \varepsilon p)\vec{x}_2 = \pi_{\|}(p, q) + \pi_{\perp}(p, q)\,.$$

For simplicity we omit the common factor $1/(\varepsilon + \eta)$ (which corresponds to different normalization of vectors $\vec{x_1}$ and $\vec{x_2}$) and consider the abelian groups

$$\mathbb{Z}[\eta] = \{p + q\eta \mid p, q \in \mathbb{Z}\} \quad \text{and} \quad \mathbb{Z}[\varepsilon] = \{p + q\varepsilon \mid p, q \in \mathbb{Z}\}\,.$$

The star map $*\colon \mathbb{Z}[\eta] \to \mathbb{Z}[\varepsilon]$, given by

$$x = p + q\eta \quad \mapsto \quad x^* = p - q\varepsilon\,,$$

is an isomorphism of the two groups. Using this formalism we can easily define a cut-and-project sequence as

$$\Sigma_{\varepsilon,\eta}(\Omega) = \{x \in \mathbb{Z}[\eta] \mid x^* \in \Omega\}\,,$$

where Ω is a bounded interval, called the acceptance window. We denote the length of an interval Ω by $|\Omega|$.

It is well known [11] that any cut-and-project set is Delone with finite number of polygons in Voronoï tiling of space. For $\Sigma_{\varepsilon,\eta}(\Omega) \subset \mathbb{R}$ it implies that there exists an increasing sequence $(x_n)_{n\in\mathbb{Z}}$ such that $\Sigma_{\varepsilon,\eta}(\Omega) = \{x_n \mid n \in \mathbb{Z}\}$ and the set of tiles $T = \{t_n = x_{n+1} - x_n \mid n \in \mathbb{Z}\}$ is finite.

If we assign to each tile $t \in T$ a letter $h(t)$ from an alphabet $\mathcal{A}$, the sequence $(h(t_n))_{n\in\mathbb{Z}}$ can thus be viewed as a bidirectional infinite word in the finite alphabet $\mathcal{A}$.

For the description of $\Sigma_{\varepsilon,\eta}(\Omega)$ we can use a different increasing sequence $\tilde{x}_n := x_{n+p}$ for arbitrary fixed $p \in \mathbb{Z}$. The corresponding sequence $(\tilde{t}_n)_{n\in\mathbb{Z}}$ has then the property $\tilde{t}_n = t_{n+p}$ for all $n \in \mathbb{Z}$. The word $(h(\tilde{t}_n))_{n\in\mathbb{Z}}$ is thus a p-shift of the infinite word $(h(t_n))_{n\in\mathbb{Z}}$. In order to avoid specification of the point indexed by 0, we introduce the symbol $u_{\varepsilon,\eta}(\Omega)$ for the entire class of all bidirectional infinite words, that are associated to $\Sigma_{\varepsilon,\eta}(\Omega)$. In [10] the infinite bidirectional words without specific origin are called trajectories.

We are interested in combinatorial properties of words $u_{\varepsilon,\eta}(\Omega)$. In particular, we study

– the cardinality of the alphabet $\{h(t) | t \in T\} \subseteq \mathcal{A}$, in case when different tiles are assigned to different letters.

– the subword complexity of $u_{\varepsilon,\eta}(\Omega)$,

– the triples of parameters ε, η, Ω that give the same bidirectional infinite words, up to choice of the letter assignment h into alphabet $\mathcal{A}$.

3 Distances in $\Sigma_{\varepsilon,\eta}(\Omega)$

In this section we first determine the number of different tiles in $\Sigma_{\varepsilon,\eta}(\Omega)$ and then we derive the lengths of these tiles.

Let us first focus on cut-and-project sequences with symmetric acceptance interval $\Omega = (-d, d)$, with $d > 0$. Let $0 < x_1 < x_2 < \cdots$ denote the positive elements of the set $\Sigma_{\varepsilon,\eta}(-d, d)$. Let x_k be the minimal positive element of $\Sigma_{\varepsilon,\eta}(-d, d)$ with $\operatorname{sign} x_k^* \neq \operatorname{sign} x_1^*$. Then $0 < x_i - x_1 < x_i < x_k$ for all $i = 2, 3, \ldots, k-1$ and $(x_i - x_1)^* \in (-d, d)$. Thus $x_i - x_1$ are positive elements of $\Sigma_{\varepsilon,\eta}(-d, d)$ which is possible only as $x_i - x_1 = x_j$ for some $j = 1, 2, \ldots, i-1$. This implies that

$$x_i = i x_1 , \qquad \text{for } i = 1, 2, \ldots, k-1 . \tag{2}$$

Moreover,

$$|x_k^* - x_1^*| \geq d . \tag{3}$$

Indeed, suppose that $x_k^* - x_1^* = (x_k - x_1)^* \in (-d, d)$. i.e. $x_k - x_1 \in \Sigma_{\varepsilon,\eta}(-d, d)$. Since $x_k > x_k - x_1 > 0$, according to (2) there must exist a $j \in \{0, 1, \ldots, k-1\}$, such that $x_k - x_1 = j x_1$, which gives $x_k^* = (j+1) x_1^*$. The latter is a contradiction with the assumption $\operatorname{sign} x_k^* \neq \operatorname{sign} x_1^*$.

Theorem 3.1. *For any irrational numbers ε, η, $\varepsilon \neq -\eta$, and any bounded interval $\Omega = [c, c+d)$, there exist positive numbers Δ_1 and Δ_2 such that the distances between consecutive points in $\Sigma_{\varepsilon,\eta}(\Omega)$ take at most three values among $\{\Delta_1, \Delta_2, \Delta_1 + \Delta_2\}$.*

Proof. The points of $\Sigma_{\varepsilon,\eta}[c, c+d)$ form an increasing sequence $(x_n)_{n \in \mathbb{Z}}$, i.e. $\Sigma_{\varepsilon,\eta}[c, c+d) = \{x_n \mid n \in \mathbb{Z}\}$. A positive number Δ is a distance between consecutive points of $\Sigma_{\varepsilon,\eta}[c, c+d)$ if there exists $n \in \mathbb{Z}$, such that $\Delta = x_{n+1} - x_n$. For any tile $\Delta > 0$ we have $\Delta \in \Sigma_{\varepsilon,\eta}(-d, d)$. Moreover, let both $\delta > 0$ and $i\delta$, $i \in \mathbb{N}$, $i \geq 2$, belong to $\Sigma_{\varepsilon,\eta}(-d, d)$. Then $i\delta$ is not a tile in $\Sigma_{\varepsilon,\eta}[c, c+d)$, since if x and $x + i\delta$ are elements of $\Sigma_{\varepsilon,\eta}[c, c+d)$ then also $x + \delta$ belongs to $\Sigma_{\varepsilon,\eta}[c, c+d)$, and hence $x + i\delta$ is not the closest neighbour of x.

Denote by Δ_1 and Δ_2 the smallest positive elements of $\Sigma_{\varepsilon,\eta}(-d, d)$ with $\Delta_1^* > 0$ and $\Delta_2^* < 0$. The values Δ_1 and Δ_2 are the two smallest candidates to be lengths of tiles in $\Sigma_{\varepsilon,\eta}[c, c+d)$.

According to (3) we have

$$\Delta_1^* - \Delta_2^* \geq d \tag{4}$$

Let $x_1 \in \Sigma_{\varepsilon,\eta}[c, c+d)$ such that $x_1^* \in [c, c+d-\Delta_1^*)$. Since $x_1^* + \Delta_1^* \in [c, c+d)$ and $x_1^* + \Delta_2^* < c + d - \Delta_1^* + \Delta_2^* \leq c$ (compare (4)), the point $x_1 + \Delta_1$ belongs to $\Sigma_{\varepsilon,\eta}[c, c+d)$ and $x_1 + \Delta_1$ is the closest right neighbour of x_1.

Let $x_2 \in \Sigma_{\varepsilon,\eta}[c, c+d)$ such that $x_2^* \in [c - \Delta_2^*, c+d)$. Similarly, $x_2^* + \Delta_2^* \in [c, c+d)$ and $x_2^* + \Delta_1^* \notin [c, c+d)$, therefore $x_2 + \Delta_2$ is the closest right neighbour of x_2.

Now let $x_3 \in \Sigma_{\varepsilon,\eta}[c, c+d)$ such that $x_3^* \in [c + d - \Delta_1^*, c - \Delta_2^*)$. Then

$$c \leq c + d + \Delta_2^* \leq x_3^* + \Delta_1^* + \Delta_2^* < c + \Delta_1^* < c + d ,$$

and the point $x_3 + \Delta_1 + \Delta_2$ belongs to $\Sigma_{\varepsilon,\eta}[c, c+d)$. We want to show that $x_3 + \Delta_1 + \Delta_2$ is the closest right neighbour of the point x_3.

Suppose that there exists a tile Δ, such that

$$x_3 < x_3 + \Delta < x_3 + \Delta_1 + \Delta_2 , \quad x_3 + \Delta \in \Sigma_{\varepsilon,\eta}[c, c+d) .$$

Combining the inequalities

$$c + d - \Delta_1^* \leq x_3^* < c - \Delta_2^* , \quad \text{and} \quad c \leq x_3^* + \Delta^* < c + d$$

we obtain

$$\Delta_2^* < \Delta^* < \Delta_1^* . \tag{5}$$

Consider the point $\delta := \Delta_1 + \Delta_2 - \Delta > 0$. Since $\Delta > \Delta_1, \Delta_2$ (as Δ_1, Δ_2 are the two smallest candidates for lengths of tiles), it holds that $0 < \delta < \min\{\Delta_1, \Delta_2\}$ but using (5) we get

$$-d < \Delta_2^* < \delta^* = \Delta_1^* + \Delta_2^* - \Delta^* < \Delta_1^* < d .$$

It means that $\delta \in \Sigma_{\varepsilon,\eta}(-d, d)$, which contradicts the definition of Δ_1 and Δ_2 as being the smallest elements of $\Sigma_{\varepsilon,\eta}(-d, d)$. As a result, the closest right neighbour of $x_3 \in \Sigma_{\varepsilon,\eta}[c, c+d)$, with $x_3^* \in [c + d - \Delta_1^*, c - \Delta_2^*)$, is the point $x_3 + \Delta_1 + \Delta_2$.

By that we have determined the right neighbours of all the elements of $\Sigma_{\varepsilon,\eta}[c, c+d)$ and the proof is completed. □

Let us make two comments on the above theorem:

- From the proof of the above theorem it follows that Δ_1 and Δ_2 depend only on the length d of the interval $\Omega = [c, c+d)$ and not on the position of the point c.
- If the acceptance window was an interval $\Omega = (c, c+d]$ or $\Omega = (c, c+d)$ the proof of Theorem 3.1 could be repeated identicaly; for the case $\Omega = [c, c+d]$ a minor amendment is needed. Hence for every acceptance interval Ω with non empty interior the set $\Sigma_{\varepsilon,\eta}(\Omega)$ has at most three types of tiles.

From the proof of Theorem 3.1 we may derive a prescription to find the right neighbour of a given point x in the cut-and-project sequence, according to the position of x^* in the aceptance interval.

Corollary 3.2. *Let $\Omega = [c, c+d)$ be a bounded interval, and let Δ_1, Δ_2, and $\Delta_1 + \Delta_2$ be the tiles in $\Sigma_{\varepsilon,\eta}(\Omega)$ such that $\Delta_1^* > 0 > \Delta_2^*$. The closest right neighbour of a point $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ is*

$$\begin{array}{ll} x + \Delta_1 & \text{if } x^* \in [c, c+d-\Delta_1^*)\,, \\ x + \Delta_1 + \Delta_2 & \text{if } x^* \in [c+d-\Delta_1^*, c-\Delta_2^*)\,, \\ x + \Delta_2 & \text{if } x^* \in [c-\Delta_2^*, c+d)\,. \end{array}$$

In the remaining part of this section we derive the lengths of tiles in $\Sigma_{\varepsilon,\eta}(\Omega)$ for a given interval $\Omega = [c, c+d)$ and positive ε, η. As it will be seen from the study of geometrical similarities of cut-and-project sets, the assumption of ε, η positive does not cause a loss of generality.

Changing continuously the length d of the acceptance interval $\Omega = [c, c+d)$ causes discrete changes of the triplet of distances Δ_1, Δ_2, $\Delta_1 + \Delta_2$. Let $\Sigma[c, c+d)$ be a generic cut-and-project sequence, and let $\Delta_1^* > 0$, $\Delta_2^* < 0$ and $\Delta_1^* + \Delta_2^*$ be the star map images of its tiles.

Denote by d_+ the smallest number such that $d < d_+$ and at least one of the distances Δ_1, Δ_2, $\Delta_1 + \Delta_2$ does not occur in $\Sigma[c, c+d_+)$. Similarly, denote by d_- the largest number such that $d_- < d$ and at least one of the distances Δ_1, Δ_2, $\Delta_1 + \Delta_2$ does not occur in $\Sigma[c, c+d_-)$. Due to Corollary 3.2, we have $d_+ = \Delta_1^* - \Delta_2^*$ and thus $\Sigma[c, c+d_+)$ has only two distances, namely Δ_1 and Δ_2. Hence growing the length of the acceptance interval, the largest distance $\Delta_1 + \Delta_2$ disappears.

Similar situation occurs in $\Sigma[c, c+d_-)$. One of the distances Δ_1, Δ_2, $\Delta_1 + \Delta_2$ disappears, but it happens in such a way that the remaining distances have their star map images of opposite sign. Thus $\Sigma[c, c+d_-)$ has distances

$$\begin{array}{ll} \Delta_1, \Delta_1 + \Delta_2 & \text{if } \Delta_1^* + \Delta_2^* < 0\,, \\ \Delta_2, \Delta_1 + \Delta_2 & \text{if } \Delta_1^* + \Delta_2^* > 0\,. \end{array}$$

From Corollary 3.2 we obtain

$$d_- = \Delta_1^* - (\Delta_1 + \Delta_2)^* = -\Delta_2^*\,, \qquad \text{or} \qquad d_- = \Delta_1^*\,,$$

respectively.

We can now determine by recurrence the lengths d_n, $n \in \mathbb{Z}$, of the acceptance windows for which $\Sigma[c, c+d_n)$ has only two tiles. Let $\Delta_{n1}^* > 0$ and $\Delta_{n2}^* < 0$ be the star images of distances occurring in the sequence $\Sigma[0, d_n)$, i.e. $d_n = \Delta_{n1}^* - \Delta_{n2}^*$.

$$\begin{array}{llll} \text{If } \Delta_{n1}^* + \Delta_{n2}^* > 0 & \text{then } d_{n+1} := \Delta_{n1}^*, & \Delta_{(n+1)1}^* := \Delta_{n1}^* + \Delta_{n2}^*, & \Delta_{(n+1)2}^* := \Delta_{n2}^*. \\ \text{If } \Delta_{n1}^* + \Delta_{n2}^* < 0 & \text{then } d_{n+1} := -\Delta_{n2}^*, & \Delta_{(n+1)1}^* := \Delta_{n1}^*, & \Delta_{(n+1)2}^* := \Delta_{n1}^* + \Delta_{n2}^*. \end{array} \tag{6}$$

Similarly, the algorithm which from the triple $d_n, \Delta_{n1}, \Delta_{n2}$ finds the triple $d_{n-1}, \Delta_{(n-1)1}, \Delta_{(n-1)2}$ has the inverse form

$$\begin{array}{llll} \text{If } \Delta_{n1} > \Delta_{n2} & \text{then } d_{n-1} := \Delta_{n1}^* - 2\Delta_{n2}^*, & \Delta_{(n-1)1}^* := \Delta_{n1}^* - \Delta_{n2}^*, & \Delta_{(n-1)2}^* := \Delta_{n2}^*. \\ \text{If } \Delta_{n1} < \Delta_{n2} & \text{then } d_{n-1} := 2\Delta_{n1}^* - \Delta_{n2}^*, & \Delta_{(n-1)1}^* := \Delta_{n1}^*, & \Delta_{(n-1)2}^* := \Delta_{n2}^* - \Delta_{n1}^*. \end{array} \tag{7}$$

It is now obvious that for the determination of the lengths of tiles in a generic cut and project sequence $\Sigma_{\varepsilon,\eta}[c, c+d)$ with $\varepsilon, \eta, d > 0$, it suffices to find $n \in \mathbb{Z}$ such that $d_n < d < d_{n+1}$. Then numbers Δ_{n1}, Δ_{n2}, $\Delta_{(n+1)1}$, $\Delta_{(n+1)2}$ take three different values that are the lengths of tiles in $\Sigma_{\varepsilon,\eta}[c, c+d)$.

We have yet to determine some initial values for the recurrences (6) and (7). It turns out that it is reasonable to start with a cut-and-project sequence whose acceptance interval is of unit length. Since the distances do not depend on the position of the interval, we can focus on the acceptance interval $[0, 1)$.

Example 1. Let ε, η be fixed positive irrational numbers. Let us study the sequence $\Sigma_{\varepsilon,\eta}[0,1) = \{p + q\eta \mid p, q \in \mathbb{Z}\,,\ 0 \le p - q\varepsilon < 1\}$. Elements of this set have to satisfy $q\varepsilon \le p < 1 + q\varepsilon$. Therefore we can write

$$\Sigma_{\varepsilon,\eta}[0,1) = \{x_q := \lceil q\varepsilon \rceil + q\eta \mid q \in \mathbb{Z}\}\,.$$

From the above formula one may easily observe that the tiles in a $\Sigma_{\varepsilon,\eta}[0,1)$ form a Sturmain word with slope ε. Since $(x_q)_{q\in\mathbb{Z}}$ is an increasing sequence, the lengths of tiles in $\Sigma_{\varepsilon,\eta}[0,1)$ can be computed as

$$x_{q+1} - x_q = \lceil q\varepsilon + \varepsilon \rceil - \lceil q\varepsilon \rceil + \eta = \begin{cases} [\varepsilon] + 1 + \eta =: \Delta_1 \\ [\varepsilon] + \eta =: \Delta_2 \end{cases}$$

The sequence $\Sigma_{\varepsilon,\eta}[0,1)$ has thus two distances ordered as a Sturmian word with slope $\alpha = \varepsilon$ and shift intercept $\beta = 0$, cf. definition of Sturmian words (1). We denote the length of the acceptance interval by $d_0 = 1$. In the notation of the recurrence relations (6) there is

$$\begin{aligned} \Delta^*_{01} &= [\varepsilon] + 1 - \varepsilon &&= 1 - \{\varepsilon\}, \\ \Delta^*_{02} &= [\varepsilon] - \varepsilon &&= -\{\varepsilon\}, \end{aligned} \qquad \text{and} \qquad d_1 = \max(1 - \{\varepsilon\}, \{\varepsilon\})\,.$$

Proposition 3.3. *Let $\varepsilon, \eta > 0$ be irrational numbers and let $\Omega = [c, c + d)$ satisfy $d_1 < d < 1$. The sequence $\Sigma_{\varepsilon,\eta}(\Omega)$ has three types of tiles and the lengths of these tiles are $\{\varepsilon\}^{-*}$, $(1 - \{\varepsilon\})^{-*}$, $(1 - 2\{\varepsilon\})^{-*}$, which are equal to*

$$[\varepsilon] + \eta\,, \qquad [\varepsilon] + 1 + \eta\,, \qquad \textit{and} \qquad 2[\varepsilon] + 1 + 2\eta\,.$$

Let us realize that in this case Δ^*_1 and Δ^*_2 do not depend on $\eta > 0$. Therefore also the algorithm (6), corresponding to shortening the acceptance interval, does not depend on η (unlike the algorithm (7)). The value of η influences only the length of tiles and not their ordering. Therefore we can choose the parameter $\eta > 0$ according to our needs, for example we can set $\eta = \varepsilon^{-1}$ which corresponds to the case when the projection in the cut-and-project scheme is orthogonal.

Proposition 3.4. *Let $\varepsilon, \eta > 0$ and let Ω be an interval of length d, $0 < d \le 1$. Then*

$$u_{\varepsilon,\eta}(\Omega) = u_{\varepsilon,1/\varepsilon}(\Omega)\,.$$

In case that the acceptance window has length $d > 1$, the parameter η plays an important role, as it is shown by the following assertion, which summarizes the results of this section. Its proof is an easy consequence of algorithms (6) and (7).

Proposition 3.5. *Let $\varepsilon, \eta > 0$ be irrational numbers with continued fraction $[a_0, a_1, a_2, \dots]$ and $[b_0, b_1, b_2, \dots]$ respectively. Denote by $\left(\frac{p_k}{q_k}\right)_{k\in\mathbb{N}}$ and $\left(\frac{r_k}{t_k}\right)_{k\in\mathbb{N}}$ the sequences of the convergents associated to ε and η, that is*

$$\frac{p_k}{q_k} = [a_0, a_1, \dots, a_k]\,, \qquad \frac{r_k}{t_k} = [b_0, b_1, \dots, b_k]\,.$$

- *Let $0 < d < 1$.*

$\Sigma_{\varepsilon,\eta}[c, c + d)$ is a cut-and-project sequence with two tiles iff there exist $k \in \mathbb{N}_0$, and $s \in \mathbb{N}$, $1 \le s \le a_{k+1}$, such that $d = |(s - 1)(p_k - \varepsilon q_k) + p_{k-1} + \varepsilon q_{k-1}|$. In this case the lengths of tiles are

$$p_k + \eta q_k \qquad \textit{and} \qquad s(p_k + \eta q_k) + p_{k-1} + \eta q_{k-1}\,.$$

- *Let $1 < d$.*

$\Sigma_{\varepsilon,\eta}[c, c + d)$ is a cut-and-project sequence with two tiles iff there exist $k \in \mathbb{N}_0$, and $s \in \mathbb{N}$, $1 \le s \le b_{k+1}$, such that $d = (s + 1)(r_k + \varepsilon t_k) + r_{k-1} + \varepsilon t_{k-1}$. In this case the lengths of tiles are

$$|r_k - \eta t_k| \qquad \textit{and} \qquad |s(r_k - \eta q_k) + r_{k-1} - \eta t_{k-1}|\,.$$

4 Complexity of cut-and-project sequences

It is useful to introduce a function that allows to determine the neighbour of a point $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ according to the position of x^* in Ω. Its definition is based on Corollary 3.2.

Definition 4.1. Let $\Omega = [c, c+d)$ be a bounded interval, and let Δ_1, Δ_2, and $\Delta_1 + \Delta_2$ be the tiles in $\Sigma_{\varepsilon,\eta}(\Omega)$, such that $\Delta_1^* > 0 > \Delta_2^*$. Let

$$f_{\varepsilon,\Omega}(y) := \begin{cases} y + \Delta_1^* & \text{for } \ y \in [c, c+d-\Delta_1^*)\,, \\ y + \Delta_1^* + \Delta_2^* & \text{for } \ y \in [c+d-\Delta_1^*, c-\Delta_2^*)\,, \\ y + \Delta_2^* & \text{for } \ y \in [c-\Delta_2^*, c+d)\,. \end{cases} \tag{8}$$

The function $f_{\varepsilon,\Omega}$ is called the stepping function of $\Sigma_{\varepsilon,\eta}(\Omega)$. If there is no misunderstanding possible, we shall omit the subscripts.

Note that the right neighbour of $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ is $\big(f_{\varepsilon,\Omega}(x^*)\big)^{-*}$. Similarly, the n-tuple of its right neighbours is given by $(f(x^*))^{-*}$, $(f^{(2)}(x^*))^{-*}$, ..., $(f^{(n)}(x^*))^{-*}$. In particular, for any given $y_0 \in \Omega \cap \mathbb{Z}[\varepsilon]$ we have

$$\Sigma_{\varepsilon,\eta}(\Omega) = \left\{ \left(f^{(n)}_{\varepsilon,\Omega}(y_0) \right)^{-*} \;\middle|\; n \in \mathbb{Z} \right\}.$$

Let us denote the tile $\Delta_1 + \Delta_2$ by letter A, the tile Δ_1 by letter B and the tile Δ_2 by letter C. According to the first n right neighbours, we can associate to every point $x \in \Sigma_{\varepsilon,\eta}(\Omega)$ a word of length n in the alphabet $\mathcal{A} = \{A, B, C\}$. This word will be denoted by $\mathit{word}(x, n)$. If w is a word of length n in the alphabet $\mathcal{A}$, we shall denote by the symbol Ω_w the convex hull of points $x^* \in \Omega \cap \mathbb{Z}[\varepsilon]$ that satisfy $\mathit{word}(x, n) = w$. In particular, we have

$$\Omega_A = [c+d-\Delta_1^*, c-\Delta_2^*)\,, \quad \Omega_B = [c, c+d-\Delta_1^*)\,, \quad \Omega_C = [c-\Delta_2^*, c+d)\,.$$

For $x_1^*, x_2^* \in \Omega \cup \mathbb{Z}[\varepsilon]$ the words $\mathit{word}(x_1, n)$ and $\mathit{word}(x_2, n)$ are different if and only if there exists an $i = 1, 2, \ldots, n$, such that at least one discontinuity point of f lies between $f^{(i-1)}(x_1^*)$ and $f^{(i-1)}(x_2^*)$. Therefore Ω_w is either empty or an interval. It is known that the density of the word w is proportional to the length of interval Ω_w. In particular, we have the following proposition.

Proposition 4.2. *Let ε, η be irrational numbers, $\varepsilon \neq -\eta$, and let Ω be a bounded interval. Let w be a factor in the word $u_{\varepsilon,\eta}(\Omega)$. Then for the density ϱ_w of the factor w we have*

$$\varrho_w = \frac{|\Omega_w|}{|\Omega|}\,.$$

In particular, for the densities of letters A, B, C we have

$$\varrho_A = \frac{\Delta_1^* - \Delta_2^* - |\Omega|}{|\Omega|}\,, \quad \varrho_B = \frac{|\Omega| - \Delta_1^*}{|\Omega|}\,, \quad \varrho_C = \frac{|\Omega| + \Delta_2^*}{|\Omega|}\,.$$

Recall that the complexity $\mathcal{C}(n)$ of an infinite word u is the number of different factors in u of length n. In our case it is given by the number of all non empty intervals Ω_w that cover Ω. The number of such intervals is determined from the number of left end-points of these intervals. A point z is a left end-point of a non-empty interval Ω_w if and only if it is either a left end-point of Ω itself, i.e. $z = c$, or there exists $i \in \{0, 1, 2 \ldots n-1\}$ such that $f^i(z)$ is a discontinuity point of the function f, i.e. $f^i(z) \in \{c+d-\Delta_1^*, c-\Delta_2^*\}$. Therefore we have the following prescription for the complexity,

$$\mathcal{C}(n) = \#\left\{ c, \alpha, f^{(-1)}(\alpha), \ldots, f^{(-(n-1))}(\alpha), \beta, f^{(-1)}(\beta), \ldots, f^{(-(n-1))}(\beta) \right\} \tag{9}$$

where we denote for simplicity of notation $\alpha = c+d-\Delta_1^*$ and $\beta = c-\Delta_2^*$.

We can now determine the complexity of $u_{\varepsilon,\eta}[c, c+d)$.

Theorem 4.3. *Let $\mathcal{C}$ denote the complexity function of $u_{\varepsilon,\eta}[c, c+d)$.*

- *If $d \notin \mathbb{Z}[\varepsilon]$, then*
$$\mathcal{C}(n) = 2n+1\,, \qquad \textit{for } \ n \in \mathbb{N}\,.$$
- *If $d \in \mathbb{Z}[\varepsilon]$, then there exists a unique non-negative $k \in \mathbb{N}_0$ such that $f^{(k)}(\alpha) = \beta$ or $f^{(k+1)}(\beta) = \alpha$, where α, β are the discontinuity points of the stepping function f. Consequently,*
$$\mathcal{C}(n) = \begin{cases} 2n+1 & \textit{for } \ n \le k\,, \\ n+k+1 & \textit{for } \ n > k\,. \end{cases}$$

Proof. If $f_{\varepsilon,\Omega}$ and $f_{\varepsilon,\Omega+t}$ are stepping functions corresponding to acceptance intervals Ω and $\Omega+t$ respectively, then $f_{\varepsilon,\Omega+t}(y) = f_{\varepsilon,\Omega}(y-t)+t$ for any translation t of the interval Ω. This means that blocs occuring in the word $u_{\varepsilon,\eta}[c,c+d)$ occur also in $u_{\varepsilon,\eta}[c+t,c+d+t)$ and vice versa. Two words with such property are said to belong to the same local isomorphism class. The words $u_{\varepsilon,\eta}[c,c+d)$ and $u_{\varepsilon,\eta}[c+t,c+d+t)$ must have the same complexity. Without loss of generality we can thus consider $\Omega=[0,d)$. Recall that $\Delta_1^*,\Delta_2^*\in\mathbb{Z}[\varepsilon]$, and that the discontinuity points of f for such an interval Ω are $\alpha=d-\Delta_1^*$ and $\beta=-\Delta_2^*$. Our aim is to determine the number of elements in the set

$$M=\Big\{0,\alpha,f^{(-1)}(\alpha),\dots,f^{(-(n-1))}(\alpha),\beta,f^{(-1)}(\beta),\dots,f^{(-(n-1))}(\beta)\Big\}$$

The function $f^{(k)}$ has no fixed point for any k, which implies the following facts:
(i) The points $\alpha,f^{(-1)}(\alpha),\dots,f^{(-(n-1))}(\alpha)$ are mutually distinct. The same holds if we replace α by β.
(ii) One can never have simultaneously $f^k(\alpha)=\beta$ and $f^i(\beta)=\alpha$ for some $k,i\in\mathbb{N}$.
(iii) Since $f(\beta)=0=c$, the point 0 is the closest right neighbour of β^{-*} and we have $f^{-i}(\beta)\neq c$ for all $i\in\mathbb{N}_0$.

- Suppose that $d\notin\mathbb{Z}[\varepsilon]$. Then $\alpha=d-\Delta_1^*\notin\mathbb{Z}[\varepsilon]$ and $\beta=-\Delta_2^*\in\mathbb{Z}[\varepsilon]$. Since $f(y)\in\mathbb{Z}[\varepsilon]$ if and only if $y\in\mathbb{Z}[\varepsilon]$, we have $f^{(-k)}(\alpha)\notin\{0,\beta,f^{(-1)}(\beta),\dots,f^{(-(n-1))}(\beta)\}$. Together with properties (i) and (iii) this means that the set M has $2n+1$ elements.

- Let now $d\in\mathbb{Z}[\varepsilon]$, i.e. the star map images α^{-*}, β^{-*} of both discontinuity points of f and the point 0 belong to $\Sigma_{\varepsilon,\eta}(\Omega)$. Since the point 0 is the closest right neighbour of β^{-*} we have either $\alpha^{-*}\leq\beta^{-*}$ or $0<\alpha^{-*}$. In the first case it means that there exists a non-negative k such that α^{-*} is the k-th left neighbour of β^{-*}, i.e. $f^{(-i)}(\beta)\neq\alpha$ for $i<k$ and $f^{(-k)}(\beta)=\alpha$. In the second case that there exists a non-negative k such that 0 is the k-th left neighbour of α^{-*}, i.e. $f^{(-i)}(\alpha)\notin\{0,\beta\}$ for $i<k$ and $f^{(-k)}(\alpha)=0=f(\beta)$. Starting from th value $n-1=k$ the pairs of elements in the set coincide, which influences the cardinality of this set and consequently also the complexity of the infinite word.

□

Since infinite words with complexity $n+$ *const.* (quasisturmain sequences, see [5]) are well described, we shall now focus on words of complexity $2n+1$. According to the result of Berthé [4], for a given $n\in\mathbb{N}$ the densities of factors of length n can take at most $3\big(\mathcal{C}(n+1)-\mathcal{C}(n)\big)=6$ values. Let us however remark that authors conjecture, upon observation, that the maximal number of different values of densities of factors in infinite words associated to cut-and-project sequences is 5.

Let us determine the maximal in- and outdegree in the Rauzy graphs corresponding to the infinite words in consideration. Let e be a factor of length $n+1$ in the infinite word $u_{\varepsilon,\eta}[c,c+d)$ and let v be the prefix of e of length n. Obviously $\Omega_e\subseteq\Omega_v$. Therefore in the Rauzy graph Γ_n the vertex v is the starting point of the edge e. If $\Omega_e=\Omega_v$, then the outdegree of the vertex v is equal to one. This happens only if the points $f^{(-n)}(\alpha)$, $f^{(-n)}(\beta)$ do not belong to the interval Ω_v. If the interval Ω_v contains exactly one of the points $f^{(-n)}(\alpha)$, $f^{(-n)}(\beta)$, then $\Omega_v=\Omega_{e_1}\cup\Omega_{e_2}$, where e_1,e_2 are the only two factors of length $n+1$ with prefix v and the outdegree of the vertex v is equal to 2. In case that both points $f^{(-n)}(\alpha)$, $f^{(-n)}(\beta)$ happen to belong to the interval Ω_v, the outdegree of the vertex v is equal to 3 and all other vertices in the Rauzy graph Γ_n have outdegree 1. Since the set $\{f^{(-n)}(\alpha),f^{(-n)}(\beta)\mid n\in N\}$ covers the acceptance interval Ω densely and uniformly, the lengths of intervals Ω_v tend to 0 as n tends to infinity. If for every $n\in\mathbb{N}$ both of the points $f^{(-n)}(\alpha)$ and $f^{(-n)}(\beta)$ fall into the same interval Ω_v for some factor v of length n, then $\lim_{n\mapsto\infty}\big(f^{(-n)}(\alpha)-f^{(-n)}(\beta)\big)=0$. Take n large enough so that $\big|f^{(-n)}(\alpha)-f^{(-n)}(\beta)\big|<\delta$. From the properties of the stepping function, as piecewise linear function with slope 1, it can happen that $\big|f^{(-n-1)}(\alpha)-f^{(-n-1)}(\beta)\big|=\big|f^{(-n)}(\alpha)-f^{(-n)}(\beta)\big|$. However, this cannot be true for every n. In the opposite case, we have $\big|f^{(-n-1)}(\alpha)-f^{(-n-1)}(\beta)\big|\gg\big|f^{(-n)}(\alpha)-f^{(-n)}(\beta)\big|$ and thus points $f^{(-n-1)}(\alpha)$, $f^{(-n-1)}(\beta)$ cannot be in the same interval Ω_v. Thus we have proved the following statement.

Proposition 4.4. *Let ε,η be irrational numbers, $\varepsilon\neq-\eta$, and let $\Omega=[c,c+d)$ be a non empty interval. Then there exists an n_0 such that for all $n\geq n_0$ the maximal indegree and the maximal outdegree in the Rauzy graph Γ_n of $u_{\varepsilon,\eta}(\Omega)$ are equal to 2.*

5 Geometrically similar cut-and-project sequences

Two sets Λ and $\hat{\Lambda}\subset\mathbb{R}$ are geometrically similar, if there exist $\varphi,\psi\in\mathbb{R}$, $\varphi\neq0$, such that $\hat{\Lambda}=\phi\Lambda+\psi$. We shall denote this property by $\Lambda\overset{g}{\sim}\hat{\Lambda}$.

Remark 5.1. Note that if two cut-and-project sequences $\Sigma_{\varepsilon,\eta}(\Omega)$ and $\Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega)$ are geometrically similar with a positive similarity factor φ, then the corresponding infinite bidirectional words $u_{\varepsilon,\eta}(\Omega)$ and $u_{\hat\varepsilon,\hat\eta}(\hat\Omega)$ coincide. A converse statement is not true, cf. Proposition 3.4.

In order to describe classes of mutually geometrically similar cut-and-project sequences, let us consider the group of symmetries of the lattice $\mathbb{Z}^2$. We have the translations $\mathbb{Z}^2 + \left(\begin{smallmatrix} a \\ b \end{smallmatrix}\right) = \mathbb{Z}^2$, for $a, b \in \mathbb{Z}$, and the group of rotation symmetries of $\mathbb{Z}^2$. They are given by all integer valued matrices $\mathbb{A}$ with determinant ± 1. Then $\mathbb{A}\mathbb{Z}^2 = \mathbb{Z}^2$. These symmetries of $\mathbb{Z}^2$ correspond to the following transformations of cut-and-project sequences.

Proposition 5.2. *Let ε, η be irrational numbers, $\varepsilon \neq -\eta$, and let Ω be a bounded interval.*

- *For $x \in \mathbb{Z}[\eta]$ we have*
$$\Sigma_{\varepsilon,\eta}(\Omega) + x = \Sigma_{\varepsilon,\eta}(\Omega + x^*) .$$
- *Let a, b, c, d be integers, such that $ad - bc = \pm 1$. Then*
$$\Sigma_{\varepsilon,\eta}(\Omega) = (a + c\eta)\Sigma_{\frac{-b+d\varepsilon}{a-c\varepsilon},\frac{b+d\eta}{a+c\eta}}\left(\tfrac{1}{a-c\varepsilon}\Omega\right) . \tag{10}$$

Proof. First, let $x = a + b\eta \in \mathbb{Z}[\eta]$. Then

$$\begin{aligned}
\Sigma_{\varepsilon,\eta}(\Omega) + x &= \big\{(p + a) + (q + b)\eta \;\big|\; p, q \in \mathbb{Z},\, p - q\varepsilon \in \Omega\big\} = \\
&= \big\{p + q\eta \;\big|\; p, q \in \mathbb{Z},\, (p - a) - (q - b)\varepsilon \in \Omega\big\} = \\
&= \big\{p + q\eta \;\big|\; p, q \in \mathbb{Z},\, p - q\varepsilon \in \Omega + a - b\varepsilon\big\} = \Sigma_{\varepsilon,\eta}(\Omega + x^*) .
\end{aligned}$$

Now let $\mathbb{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ be an integer valued matrix with determinant ± 1. Then $\mathbb{A}\mathbb{Z}^2 = \mathbb{Z}^2$ and therefore

$$\begin{aligned}
\Sigma_{\varepsilon,\eta}(\Omega) &= \big\{p + q\eta \;\big|\; p, q \in \mathbb{Z},\, p - q\varepsilon \in \Omega\big\} = \\
&= \big\{(1, \eta)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \;\big|\; p, q \in \mathbb{Z},\, (1, -\varepsilon)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\big\} = \\
&= \big\{(1, \eta)\mathbb{A}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \;\big|\; p, q \in \mathbb{Z},\, (1, -\varepsilon)\mathbb{A}\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\big\} = \\
&= \big\{(a + c\eta, b + d\eta)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \;\big|\; p, q \in \mathbb{Z},\, (a - c\varepsilon, b - d\varepsilon)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \Omega\big\} = \\
&= (a + c\eta)\left\{\left(1, \tfrac{b+d\eta}{a+c\eta}\right)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \;\middle|\; p, q \in \mathbb{Z},\, \left(1, \tfrac{b-d\varepsilon}{a-c\varepsilon}\right)\left(\begin{smallmatrix} p \\ q \end{smallmatrix}\right) \in \tfrac{1}{a-c\varepsilon}\Omega\right\} = \\
&= (a + c\eta)\Sigma_{\frac{-b+d\varepsilon}{a-c\varepsilon},\frac{b+d\eta}{a+c\eta}}\left(\tfrac{1}{a-c\varepsilon}\Omega\right) .
\end{aligned}$$

□

The group of integer valued matrices with determinant ± 1 is generated by matrices $\mathbb{A}_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, $\mathbb{A}_2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, $\mathbb{A}_3 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Elementary transformations of cut-and-project sequences for these matrices are

$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon-1,\ \eta+1}(\Omega) \tag{11}$$
$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{-\varepsilon,-\eta}(-\Omega) \tag{12}$$
$$\Sigma_{\varepsilon,\eta}(\Omega) = \eta\, \Sigma_{\frac{1}{\varepsilon},\frac{1}{\eta}}(-\tfrac{1}{\varepsilon}\Omega) \tag{13}$$

Using the above elementary transformation we show that we may limit our considerations to ε, η and Ω with certain properties, without loosing any infinite bidirectional word. The aim of this section is to prove the following theorem.

Theorem 5.3. *Let ε, η are irrational numbers, $\varepsilon \neq \eta$. Let Ω be a bounded interval. Then there exist irrational numbers $\hat\varepsilon, \hat\eta$ and an interval $\hat\Omega$, such that*

$$\Sigma_{\varepsilon,\eta}(\Omega) \overset{g}{\sim} \Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega) , \qquad 0 < \hat\varepsilon < 1 ,\, \hat\eta > 0 , \qquad \max(1 - \hat\varepsilon, \hat\varepsilon) < |\hat\Omega| \leq 1 .$$

The proof of the theorem will be divided into two lemmas. Before stating the lemmas, let us recall certain properties of continued fractions [7].

If $\xi > 0$ is an irrational number with continued fraction $[a_0, a_1, a_2, \dots]$ and $\left(\frac{p_n}{q_n}\right)_{n\in\mathbb{N}_0}$ the sequence of the convergents assotiated to ξ, then for every $n \in \mathbb{N}_0$ we have

(i) $$\frac{p_{2n}}{q_{2n}} < \xi < \frac{p_{2n+1}}{q_{2n+1}}$$

(ii) $$\frac{1}{q_{n+2}} < |q_n\xi - p_n| < \frac{1}{q_{n+1}}$$

(iii) $$p_{n+1}q_n - p_nq_{n+1} = (-1)^n$$

Lemma 5.4. *For any irrational ε, η, $\varepsilon \neq -\eta$, and bounded interval Ω, there exist irrational numbers $\hat{\varepsilon}, \hat{\eta}$ and a bounded interval $\hat{\Omega}$ such that*

$$\Sigma_{\varepsilon,\eta}(\Omega) \overset{g}{\sim} \Sigma_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega}), \qquad \hat{\varepsilon} > 0, \hat{\eta} > 0, \qquad \textit{and} \qquad |\hat{\Omega}| \leq 1.$$

Proof. Using transformation (12) we may assume without loss of generality that $\eta > 0$. We first find convergents $\frac{p_{2n}}{q_{2n}}$ and $\frac{p_{2n+1}}{q_{2n+1}}$ associated to η, so that $-\varepsilon \notin \left[\frac{p_{2n}}{q_{2n}}, \frac{p_{2n+1}}{q_{2n+1}}\right]$. From the property (i) of continued fractions we have

$$\frac{p_{2n}}{q_{2n}} < \eta < \frac{p_{2n+1}}{q_{2n+1}} \quad \Longleftrightarrow \quad \frac{-p_{2n} + \eta q_{2n}}{p_{2n+1} - \eta q_{2n+1}} > 0 \tag{14}$$

and

$$\left(\frac{p_{2n}}{q_{2n}} + \varepsilon\right)\left(\frac{p_{2n+1}}{q_{2n+1}} + \varepsilon\right) > 0 \quad \Longleftrightarrow \quad \frac{p_{2n} + \varepsilon q_{2n}}{p_{2n+1} + \varepsilon q_{2n+1}} > 0 \tag{15}$$

Let us define the matrix $\mathbb{A} = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) = \left(\begin{smallmatrix} p_{2n+1} & -p_{2n} \\ -q_{2n+1} & q_{2n} \end{smallmatrix}\right)$ for transformation (10). Then from (14) we have

$$\hat{\eta} := \frac{b + d\eta}{a + c\eta} = \frac{-p_{2n} + \eta q_{2n}}{p_{2n+1} - \eta q_{2n+1}} > 0,$$

and from (15) we have

$$\hat{\varepsilon} := \frac{-b + d\varepsilon}{a - c\varepsilon} = \frac{p_{2n} + \varepsilon q_{2n}}{p_{2n+1} + \varepsilon q_{2n+1}} > 0.$$

From the property (iii) of continued fractions we obtain

$$\det \mathbb{A} = p_{2n+1}q_{2n} - p_{2n}q_{2n+1} = 1.$$

In the transformation (10) with the matrix $\mathbb{A}$ we obtain a new acceptance window $\hat{\Omega} := \frac{1}{a - c\varepsilon}\Omega$. Let us find an estimate of the denominator in the fraction,

$$\begin{aligned} |a - c\varepsilon| = |c|\left|\eta + \varepsilon - \left(\frac{a}{c} + \eta\right)\right| \; &\geq \; q_{2n+1}|\eta + \varepsilon| - |q_{2n+1}\eta - p_{2n+1}| \\ &\geq \; q_{2n+1}|\eta + \varepsilon| - \frac{1}{q_{2n+2}} \; \geq \; q_{2n+1}|\eta + \varepsilon| - 1. \end{aligned}$$

Note that in the estimate we have used property (ii) of continued fractions.

Since $\lim_{n\to\infty} q_{2n+1} = +\infty$ and $|\eta + \varepsilon| \neq 0$, it is possible to choose a sufficiently large n, so that the length of the interval $\hat{\Omega}$ is smaller or equal to 1. ☐

Lemma 5.5. *Let $\varepsilon, \eta > 0$, ε, η irrational and let Ω be an interval of length $|\Omega| < 1$. Then there exist irrational numbers $\hat{\varepsilon}, \hat{\eta}$ and an interval $\hat{\Omega}$ satisfying*

$$\Sigma_{\varepsilon,\eta}(\Omega) \overset{g}{\sim} \Sigma_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega}), \qquad 0 < \hat{\varepsilon} < 1, \hat{\eta} > 0, \qquad \max(1 - \hat{\varepsilon}, \hat{\varepsilon}) < |\hat{\Omega}| \leq 1.$$

Proof. Without loss of generality we may assume that $\varepsilon < 1$. Otherwise we would use the transformation (11) to get

$$\Sigma_{\varepsilon,\eta}(\Omega) = \Sigma_{\varepsilon - [\varepsilon], \eta + [\varepsilon]}(\Omega),$$

where $\varepsilon - [\varepsilon] < 1$. We first show that there exist $0 < \check{\varepsilon} < 1$, $\check{\eta} > 0$ and $\check{\Omega}$ such that

$$\Sigma_{\varepsilon,\eta}(\Omega) \overset{g}{\sim} \Sigma_{\check{\varepsilon},\check{\eta}}(\check{\Omega}), \qquad \text{and} \qquad \check{\varepsilon} < |\check{\Omega}| \leq 1.$$

The case of $|\Omega| > \varepsilon$ is trivial, because it suffices to put $\check{\varepsilon} = \varepsilon$, $\check{\eta} = \eta$ and $\check{\Omega} = \Omega$. Assume that for given Ω we have $|\Omega| \leq \varepsilon$. Let $\left(\frac{p_n}{q_n}\right)_{n\in\mathbb{N}_0}$ are the convergents associated to ε. According to property (ii) of continued fractions, the sequence $\left(|p_n - q_n\varepsilon|\right)_{n\in\mathbb{N}_0}$ is decreasing and thus it is possible to find $n \in \mathbb{N}$ so that

$$|p_n - q_n\varepsilon| < |\Omega| \leq |p_{n-1} - q_{n-1}\varepsilon| \leq |p_0 - q_0\varepsilon| = \varepsilon. \tag{16}$$

Let us transform the cut and project set $\Sigma_{\varepsilon,\eta}(\Omega)$ using the transformation (10) with matrix $\mathbb{A} = \left(\begin{smallmatrix} p_{n-1} & p_n \\ q_{n-1} & q_n \end{smallmatrix}\right)$. We obtain $\Sigma_{\varepsilon,\eta}(\Omega) \overset{\mathrm{g}}{\sim} \Sigma_{\check\varepsilon,\check\eta}(\check\Omega)$, where

$$\check\varepsilon = \frac{-p_n + q_n\varepsilon}{p_{n-1} - q_{n-1}\varepsilon}\,, \qquad \check\eta = \frac{p_n + q_n\eta}{p_{n-1} + q_{n-1}\eta}\,, \qquad \check\Omega = \frac{1}{p_{n-1} - q_{n-1}\varepsilon}\Omega\,.$$

According to properties (i) and (ii) of the continued fractions we have $0 < \check\varepsilon < 1$. The number $\check\eta$ is clearly positive. Since for $|\Omega|$ we have (16), for the length of $\check\Omega$ we may write

$$\check\varepsilon = \frac{-p_n + q_n\varepsilon}{p_{n-1} - q_{n-1}\varepsilon} < |\check\Omega| \le 1\,,$$

what was to be shown.

If moreover $|\check\Omega| > 1 - \check\varepsilon$, we may set $\hat\varepsilon := \check\varepsilon$, $\hat\eta := \check\eta$, and $\hat\Omega := \check\Omega$, and the lemma is proved. It thus remains to solve the case

$$\check\varepsilon < \check\Omega \le 1 - \check\varepsilon \tag{17}$$

This, however, is possible only if $\check\varepsilon < \frac{1}{2}$, i.e. its continued fraction has the form $\check\varepsilon = [0, c_1, c_2, \dots]$ with $c_1 \ge 2$. Since

$$1 - \check\varepsilon > 1 - 2\check\varepsilon > \dots > 1 - (c_1 - 1)\check\varepsilon > \check\varepsilon\,, \tag{18}$$

it is possible to find a minimal $s \in \{1, 2, \dots, c_1 - 1\}$ so that

$$|\check\Omega| \le 1 - s\check\varepsilon\,. \tag{19}$$

Now we use the transformation (10) with the matrix $\mathbb{A} = \left(\begin{smallmatrix} 1 & 0 \\ s & 1 \end{smallmatrix}\right)$ to get $\Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega) \overset{\mathrm{g}}{\sim} \Sigma_{\check\varepsilon,\check\eta}(\check\Omega)$, for

$$\hat\varepsilon := \frac{\check\varepsilon}{1 - s\check\varepsilon}\,, \qquad \hat\eta := \frac{\check\eta}{1 + s\check\eta}\,, \qquad \hat\Omega := \frac{1}{1 - s\check\varepsilon}\check\Omega\,.$$

Let us verify that $\hat\varepsilon$, $\hat\eta$ and $\hat\Omega$ satisfy the inequalities required by the lemma. The parameter $\hat\eta$ is positive as it is a ratio of positive numbers. For $\hat\varepsilon$ we use the inequality (18) to obtain $0 < \hat\varepsilon < 1$. For the estimate of $\hat\Omega$ we derive from (17) and (19) that

$$\hat\varepsilon = \frac{\check\varepsilon}{1 - s\check\varepsilon} < |\hat\Omega| \le 1\,. \tag{20}$$

In order to complete the proof of the lemma, it remains to show that $|\hat\Omega| > 1 - \hat\varepsilon$. If the minimal s satisfies $s \le c_1 - 2$, we have $1 - (s+1)\check\varepsilon < |\check\Omega| \le 1 - s\check\varepsilon$ which implies

$$1 \ge |\hat\Omega| > \frac{1 - (s+1)\check\varepsilon}{1 - s\check\varepsilon} = 1 - \frac{\check\varepsilon}{1 - s\check\varepsilon} = 1 - \hat\varepsilon\,.$$

If the minimal s is $s = c_1 - 1$, then automatically $\hat\varepsilon > 1 - \hat\varepsilon$ which, together with (20) implies in both cases $1 \ge |\hat\Omega| > \max(\hat\varepsilon, 1 - \hat\varepsilon)$ and proves the lemma. □

Combination of Lemmas 5.4 and 5.5 constitutes the proof of Theorem 5.3.

Remark 5.6. Note that in Theorem 5.3 we may choose the interval $\hat\Omega$ in such a way that the similarity factor between $\Sigma_{\varepsilon,\eta}(\Omega)$ and $\Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega)$ is positive, thus infinite bidirectional words associated to these cut-and-project sets are identical. If it happens that $\Sigma_{\varepsilon,\eta}(\Omega) = \varphi\Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega)$ for some $\varphi < 0$, we can use the transformation matrix $\mathbb{A} = \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$ to obtain

$$\Sigma_{\hat\varepsilon,\hat\eta}(\hat\Omega) = -\Sigma_{\hat\varepsilon,\hat\eta}(-\hat\Omega)\,.$$

Therefore using the interval $-\hat\Omega$ instead of $\hat\Omega$ gives us a positive similarity factor $-\varphi$.

6 Combinatorial classification of cut-and-project sequences

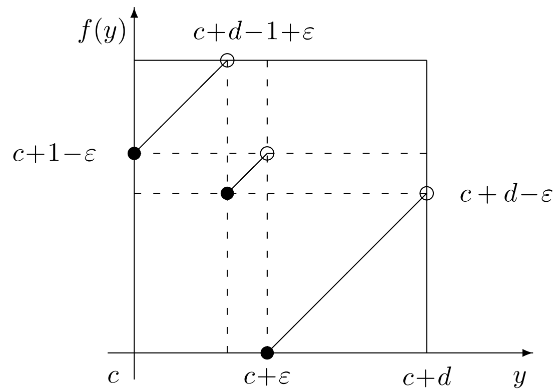
Let us recall that two words $u = (u_n)_{n\in\mathbb{Z}}$ and $\hat u = (\hat u_n)_{n\in\mathbb{Z}}$ in alphabets $\mathcal{A}$ and $\hat{\mathcal{A}}$, respectively, are combinatorially equivalent, if there exists a bijection $h\colon \mathcal{A} \to \hat{\mathcal{A}}$ and an $n_0 \in \mathbb{Z}$ such that $\hat u_n = h(u_{n+n_0})$ for all $n \in \mathbb{Z}$. We denote this by $u \overset{\mathrm{c}}{\sim} \hat u$. Note that if a bijection h should exist, certainly the density of a letter a in the word u must be the same as the density of $h(a)$ in $\hat u$.

Remark 6.1. Recall that two words associated to cut-and-project sets which are geometrically similar with positive similarity factor are combinatorially equivalent, see Remark 5.1. Using Theorem 5.3 and Remark 5.6 we may thus limit our consideration, without loss of generality, to words $u_{\varepsilon,\eta}(\Omega)$ with parameters satisfying

$$0 < \varepsilon < 1\,, \qquad \eta > 0\,, \qquad \max(\varepsilon, 1-\varepsilon) < |\Omega| \leq 1\,.$$

In this case the stepping function has the form

$$f(y) = \begin{cases} y + 1 - \varepsilon, & y \in [c, c+d-1+\varepsilon)\,, \\ y + 1 - 2\varepsilon, & y \in [c+d-1+\varepsilon, c+\varepsilon)\,, \\ y - \varepsilon, & y \in [c+\varepsilon, c+d)\,. \end{cases}$$



and thus we have

$$\Omega_C = \big[c+\varepsilon, c+d\big)\,, \qquad \Omega_B = \big[c, c+d-1+\varepsilon\big)\,, \qquad \Omega_A = \big[c+d-1+\varepsilon, c+\varepsilon\big)\,.$$

For the densities of tiles A, B, and C we have in this case

$$\varrho_A = \frac{1}{d} - 1\,, \qquad \varrho_B = 1 - \frac{1-\varepsilon}{d}\,, \qquad \text{and} \qquad \varrho_C = 1 - \frac{\varepsilon}{d}\,,$$

respectively, cf. Proposition 4.2.

Proposition 3.4 shows that two geometrically non similar cut-and-project sequences may correspond to combinatorially equivalent infinite words. Let us show yet another example of combinatorially equivalent words.

Example 2. Consider Sturmian words u, $\hat{u}$ in the alphabet $\{0,1\}$ with the slope α, and $\hat{\alpha} = 1-\alpha$ respectively, where $0 < \alpha < 1$. Obviously, u and $\hat{u}$ are combinatorially equivalent since $\hat{u}$ arises from u by replacing 0's with 1's and vice versa. Such Sturmian words correspond to cut and project sequences with Ω a semi-closed interval of length 1, cf. Example 1. For $\eta > 0$ we have

$$\Sigma_{\alpha,\eta}[c, c+1) = \{x_n \equiv \lceil c + n\alpha \rceil + n\eta \mid n \in \mathbb{Z}\}$$
$$\Sigma_{1-\alpha,\eta}(-c-1, -c] = \{y_n \equiv \lfloor -c + (1-\alpha)n \rfloor + n\eta \mid n \in \mathbb{Z}\}\,.$$

Both of the sets have two tiles of lengths η and $\eta + 1$. More precisely, the distances are

$$x_{n+1} - x_n = \eta + \lceil c + n\alpha + \alpha \rceil - \lceil c + n\alpha \rceil\,,$$
$$y_{n+1} - y_n = \eta + 1 - \lceil c + n\alpha + \alpha \rceil + \lceil c + n\alpha \rceil\,.$$

We can see that

$$\begin{aligned} x_{n+1} - x_n &= \eta & &\Longrightarrow & y_{n+1} - y_n &= \eta + 1\,, \\ x_{n+1} - x_n &= \eta + 1 & &\Longrightarrow & y_{n+1} - y_n &= \eta\,. \end{aligned}$$

The above example illustrates that for acceptance window of length 1 the interchange $\varepsilon \leftrightarrow 1-\varepsilon$, $\Omega \leftrightarrow -\Omega$ gives combinatorially equivalent words. The following lemma states that the same property holds for general acceptance window as well.

Lemma 6.2. *Let $0 < \varepsilon < 1$, $\eta > 0$ be irrational numbers, and let Ω be an interval of length $\max(\varepsilon, 1-\varepsilon) < |\Omega| \leq 1$. Then*

$$u_{\varepsilon,\eta}(-\Omega) \overset{c}{\sim} u_{1-\varepsilon,\eta}(\Omega)\,.$$

Proof. Although $\mathbb{Z}[\varepsilon] = \mathbb{Z}[1-\varepsilon]$, the star maps from $\mathbb{Z}[\eta]$ to $\mathbb{Z}[\varepsilon] = \mathbb{Z}[1-\varepsilon]$ are different for the cut-and-project sequences $\Sigma_{\varepsilon,\eta}(-\Omega)$ and $\Sigma_{1-\varepsilon,\eta}(\Omega)$. We shall therefore distinguish them by indices,

$$\begin{aligned} *_1 \,:\quad & p + q\eta \ \mapsto\ p - q\varepsilon & &\text{for } \Sigma_{\varepsilon,\eta}(-\Omega)\,, \\ *_2 \,:\quad & p + q\eta \ \mapsto\ p - q(1-\varepsilon) & &\text{for } \Sigma_{1-\varepsilon,\eta}(\Omega)\,. \end{aligned}$$

According to Proposition 3.3 in both cut-and-project sets the three distances are η, $1+\eta$ and $1+2\eta$. Their images under the star map are however different for $*_1$ and $*_2$, namely,

$$\begin{array}{cccl} -\varepsilon, & 1-\varepsilon, & 1-2\varepsilon, & \text{for } *_1\,, \\ \varepsilon-1, & \varepsilon, & 2\varepsilon-1, & \text{for } *_2\,. \end{array}$$

For the proof we use two facts:

$$(F1)\quad f^{-1}_{\varepsilon,\Omega}(x) = f_{1-\varepsilon,\Omega}(x) \ \text{ for all } \ x\in\Omega\,, \qquad\qquad (F2)\quad \Sigma_{\varepsilon,\eta}(\Omega) = -\Sigma_{\varepsilon,\eta}(-\Omega)\,.$$

Let $x_0 \in \Omega\cap\mathbb{Z}[\varepsilon]$. Then the fact (F2) says that the sequence of distances to the left from the point $x_0^{-*_1}$ in the set $\Sigma_{\varepsilon,\eta}(\Omega)$ is the same as the sequence of distances to the right from the point $-x_0^{-*_1}$ in the set $\Sigma_{\varepsilon,\eta}(-\Omega)$.

The fact (F1) says that the sequence of steps to the right from the point $x_0^{-*_2}$ in the set $\Sigma_{1-\varepsilon,\eta}(\Omega)$ is the same as the sequence of steps to the left from the point $x_0^{-*_1}$ in the set $\Sigma_{\varepsilon,\eta}(\Omega)$, only the length of steps η and $1+\eta$ are interchanged. □

Note that the density of the shortest tile η in the set $\Sigma_{\varepsilon,\eta}(-\Omega)$ is equal to $1-\frac{\varepsilon}{d}$, whereas the density of the shortest tile η in the set $\Sigma_{1-\varepsilon,\eta}(\Omega)$ is equal to $1-\frac{1-\varepsilon}{d}$. It is therefore obvious that the cut-and-project sets $\Sigma_{\varepsilon,\eta}(-\Omega)$ and $\Sigma_{1-\varepsilon,\eta}(\Omega)$ are not geometrically similar. Nevertheless, the associated infinite bidirectional words are combinatorially equivalent.

From every Sturmian sequence in the alphabet $\{A,B\}$ we can form a special word in the alphabet $\{A,B,C\}$ in the way that in between every two letters of the Sturmian sequence we insert k-times the letter C. Such a sequence we call a k-padded Sturmain sequence.

Definition 6.3. Let u be a Sturmain word in the alphabet $\{A,B\}$. Let $k\in\mathbb{N}$ and h be a morphism $h\colon \{A,B\}\to\{A,B,C\}^*$, given by $h(A)=AC^k$, $h(B)=BC^k$. The word $h(u)$ is called a k-padded Sturmian word.

The following proposition shows that k-padded Sturmain sequences can also be cut and project sequences.

Proposition 6.4. *Every k-padded Sturmian word is a word associated to a cut and project sequence.*

Proof. Let u be a Sturmain word with slope α, $0<\alpha<1$. Similarly as in Examples 1 and 2 we can consider u to be the word $u_{\alpha,\mu}[c,c+1)$, for some $c\in\mathbb{R}$ and $\mu>0$. For a given $k\in\mathbb{N}$ we form a k-padded Sturmian sequence and want to find parameters ε, η, Ω, so that the k-padded Sturmain word is a word $u_{\varepsilon,\eta}(\Omega)$. We put

$$\varepsilon = \frac{1}{k+2-\alpha}\,,\qquad \eta=\mu\,,\qquad \Omega=[\varepsilon c,\varepsilon(c+k+1))\,.$$

Note that Ω can be written as $\Omega=[\tilde c,\tilde c+\tilde d)$ for $\tilde c=\varepsilon c$ and $\tilde d=(k+1)\varepsilon$. Since $\frac{1}{k+2}<\varepsilon<\frac{1}{k+1}$, we have $\max\{1-\varepsilon,\varepsilon\}<\tilde d<1$ and therefore star images of tiles in $\Sigma_{\varepsilon,\eta}(\Omega)$ are ε, $1-\varepsilon$, and $1-2\varepsilon$.

For such parameters the stepping function $f_{\varepsilon,\Omega}$ is given by

$$f_{\varepsilon,\Omega}(y) := \begin{cases} y+1-\varepsilon & \text{for } \ y\in\big[\varepsilon c,(c+k+2)\varepsilon-1\big) & =\Omega_A\,,\\ y+1-2\varepsilon & \text{for } \ y\in\big[(c+k+2)\varepsilon-1,\varepsilon c+\varepsilon\big) & =\Omega_B\,,\\ y-\varepsilon & \text{for } \ y\in\big[\varepsilon c+\varepsilon,(c+k+1)\varepsilon\big) & =\Omega_C\,. \end{cases}$$

We can verify easily that

$$\begin{aligned} f^{(i)}_{\varepsilon,\Omega}(\Omega_A\cup\Omega_B) &\subseteq \Omega_C\,, \qquad \text{for } i=1,2,\ldots,k\,,\\ f^{(k+1)}_{\varepsilon,\Omega}(\Omega_A\cup\Omega_B) &= (\Omega_A\cup\Omega_B)\,. \end{aligned}$$

Therefore every letter A or B is followed by the string of k letters C, whereas the $(k+1)$-th letter is different from C. Moreover, the restriction of $f^{(k+1)}_{\varepsilon,\Omega}$ to $\Omega_A\cup\Omega_B$ is a scaled stepping function $f_{\alpha,[c,c+1)}$. □

Theorem 6.5.

- *For any irrational numbers ε, η, $\varepsilon\neq-\eta$, and a bounded interval Ω there exist irrational $\hat\varepsilon$ and an interval $\hat\Omega$ satisfying*

$$0<\hat\varepsilon<\frac12\,,\qquad 1-\hat\varepsilon<|\hat\Omega|\le 1\,,$$

such that $u_{\varepsilon,\eta}(\Omega)\overset{c}{\sim}u_{\hat\varepsilon,\hat\eta}(\hat\Omega)$ for any irrational $\hat\eta>0$.

- *Let ε, $\hat{\varepsilon}$, η, $\hat{\eta}$ be positive irrational numbers such that $0 < \varepsilon, \hat{\varepsilon} < \frac{1}{2}$. Let $\Omega, \hat{\Omega}$ be intervals satisfying $1-\varepsilon < |\Omega| \le 1$, $1-\hat{\varepsilon} < |\hat{\Omega}| \le 1$, $u_{\varepsilon,\eta}(\Omega) \overset{c}{\sim} u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$. Then*
  - *either $\varepsilon = \hat{\varepsilon}$ and there exists $x \in \mathbb{Z}[\varepsilon]$ such that $\Omega \cap \mathbb{Z}[\varepsilon] = (x + \hat{\Omega}) \cap \mathbb{Z}[\varepsilon]$, i.e. $u_{\varepsilon,\eta}(\Omega) = u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$.*
  - *or there exists $k \in \mathbb{N}$, and an irrational α, $0 < \alpha < 1$, such that $u_{\varepsilon,\eta}(\Omega)$ is k-padded Sturmian word with slope α and $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ is k-padded Sturmian word with slope $1-\alpha$.*

The proof of the first statement of the theorem is a direct consequence of Lemma 6.2 and Remark 6.1. In the remaining part of the section we are going to prove the second statement.

Let us determine using the stepping function the possible lengths of blocks of one letter in the word $u_{\varepsilon,\eta}(\Omega)$. Let us denote by $\min(A)$, $\max(A)$, (resp. $\min(B)$, $\max(B)$) the minimal and the maximal number of letters C that follow a letter A (resp. B) in the word $u_{\varepsilon,\eta}(\Omega)$.

Lemma 6.6. *Let $0 < \varepsilon < \frac{1}{2}$, $\eta > 0$ be irrational numbers, and let Ω be an interval of length $1-\varepsilon < |\Omega| \le 1$. Then*

$$\min(A) = \lfloor d\varepsilon^{-1} \rfloor - 1\,, \quad \max(A) = \lfloor \varepsilon^{-1} \rfloor - 1\,, \quad \min(B) = \lfloor \varepsilon^{-1} \rfloor - 1\,, \quad \max(B) = \lceil d\varepsilon^{-1} \rceil - 1\,.$$

Proof. Let us derive the value of $\min(A)$. From the properties of the stepping function, the shortest block of letters C following a letter A is found after the point $x = c + d - 1 + \varepsilon$. Therefore

$$\min(A) = \max\{k \mid c + d - k\varepsilon \ge c + \varepsilon\} = \lfloor d\varepsilon^{-1} \rfloor - 1\,.$$

Similarly the lowest number of letters C following a B is found after the point $x = c$, hence

$$\min(B) = \max\{k \mid c + \varepsilon \le c + 1 - k\varepsilon\} = \lfloor \varepsilon^{-1} \rfloor - 1\,.$$

We derive the values of $\max(A)$, $\max(B)$ from the right end-points of intervals Ω_A, Ω_B. We have

$$\begin{aligned} \max(A) &= \max\{k \mid c + 1 - k\varepsilon > c + \varepsilon\} = \lfloor \varepsilon^{-1} \rfloor - 1\,, \\ \max(B) &= \max\{k \mid c + d - (k-1)\varepsilon > c + \varepsilon\} = \lceil d\varepsilon^{-1} \rceil - 1\,. \end{aligned}$$

□

Using the stepping function we can determine some relations for the densities of factors in the word $u_{\varepsilon,\eta}(\Omega)$, that will help us in proving the main theorem.

Lemma 6.7. *Let $0 < \varepsilon < \frac{1}{2}$, $\eta > 0$ be irrational numbers, and let Ω be an interval of length $1-\varepsilon < |\Omega| \le 1$. Then*

$$\varrho_C > \varrho_B\,, \qquad \varrho_{BC} = \varrho_{CB} = \varrho_B\,, \qquad \text{and} \qquad \varrho_{AB} = \varrho_{BA} = 0\,.$$

Proof. The fact $\varrho_C > \varrho_B$ is obvious since $\varepsilon < \frac{1}{2}$. Let us determine the set Ω_{CB}. We have $x \in \Omega_{CB}$ if $x \in \Omega_C$ and $f(x) \in \Omega_B$, hence $\Omega_{CB} = \Omega_C \cap f^{-1}(\Omega_B)$. From the properties of the given stepping function it is clear that $f^{-1}(\Omega_B) \subset \Omega_C$ and therefore $\Omega_{CB} = f^{-1}(\Omega_B)$. The function f is piecewise linear with slope 1 and thus $|\Omega_{CB}| = |f^{-1}(\Omega_B)| = |\Omega_B|$. This implies $\varrho_{CB} = \varrho_B$. All the other relations can be derived in a similar way. □

Proof of Theorem 6.5. Let $0 < \varepsilon, \hat{\varepsilon} < \frac{1}{2}$, $\eta, \hat{\eta} > 0$, $1-\varepsilon < |\Omega| \le 1$, $1-\hat{\varepsilon} < |\hat{\Omega}| \le 1$. The sequences $u_{\varepsilon,\eta}(\Omega)$ and $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ are bidirectional infinite words in the alphabet $\{A, B, C\}$. If $u_{\varepsilon,\eta}(\Omega) \overset{c}{\sim} u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$, then the word $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ arises from $u_{\varepsilon,\eta}(\Omega)$ by a simple permutation of letters in the alphabet. For the purposes of the proof we shall distinguish the letters A, B, C in the word $u_{\varepsilon,\eta}(\Omega)$ and letters $\hat{A}, \hat{B}, \hat{C}$ in the word $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$. We have 6 permutations on three letters, given by

$$\begin{aligned} h_1: &\quad A \leftrightarrow \hat{A}, \quad B \leftrightarrow \hat{B}, \quad C \leftrightarrow \hat{C} \\ h_2: &\quad A \leftrightarrow \hat{B}, \quad B \leftrightarrow \hat{A}, \quad C \leftrightarrow \hat{C} \\ h_3: &\quad A \leftrightarrow \hat{C}, \quad B \leftrightarrow \hat{B}, \quad C \leftrightarrow \hat{A} \\ h_4: &\quad A \leftrightarrow \hat{C}, \quad B \leftrightarrow \hat{A}, \quad C \leftrightarrow \hat{B} \\ h_5: &\quad A \leftrightarrow \hat{B}, \quad B \leftrightarrow \hat{C}, \quad C \leftrightarrow \hat{A} \\ h_6: &\quad A \leftrightarrow \hat{A}, \quad B \leftrightarrow \hat{C}, \quad C \leftrightarrow \hat{B} \end{aligned}$$

Obviously, the densities of factors of finite length must correspond in $u_{\varepsilon,\eta}(\Omega)$ and $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$. According to Lemma 6.7, we have $\varrho_C > \varrho_B$ and $\varrho_{\hat{C}} > \varrho_{\hat{B}}$, thus we may exclude the permutation h_6.

Let us exclude the permutation h_5. We use Lemma 6.7 to find that $\varrho_B = \varrho_{CB} = \varrho_{\hat{B}\hat{A}} = 0$, but this implies that $1 - \varepsilon = d$ which is a contradiction. In the same way we exclude h_4, since h_4 and h_5 differ only by interchange of letters with and without hats. For the permutation h_3 one would have $\varrho_B = \varrho_{CB} = \varrho_{\hat{A}\hat{B}} = 0$, which is again a contradiction.

The remaining permutations are h_1 or h_2. The permutation h_1 implies $\varrho_A = \varrho_{\hat{A}}$ and thus $d = \hat{d}$. Then necessarily also $\varrho_C = \varrho_{\hat{C}}$ which implies $\varepsilon = \hat{\varepsilon}$. The combinatorial equivalence of $u_{\varepsilon,\eta}(\Omega)$ and $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ means that for $\hat{\eta} = \eta$ the sets $\Sigma_{\varepsilon,\eta}(\Omega)$ and $\Sigma_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ are geometrically similar. The only way this can happen is that $\hat{\Omega} \cap \mathbb{Z}[\varepsilon] = (x + \Omega) \cap \mathbb{Z}[\varepsilon]$ for some $x \in \mathbb{Z}[\varepsilon]$, cf. Proposition 5.2.

Permutation h_2 provides

$$\begin{aligned} \varrho_C = \varrho_{\hat{C}} \quad &\Longrightarrow \quad \frac{\varepsilon}{d} = \frac{\hat{\varepsilon}}{\hat{d}}\,, \\ \max(A) = \max(\hat{B}) \quad &\Longrightarrow \quad \left\lfloor \varepsilon^{-1} \right\rfloor = \left\lceil \hat{d}\hat{\varepsilon}^{-1} \right\rceil\,, \\ \min(B) = \min(\hat{A}) \quad &\Longrightarrow \quad \left\lceil \varepsilon^{-1} \right\rceil = \left\lfloor \hat{d}\hat{\varepsilon}^{-1} \right\rfloor\,. \end{aligned}$$

This is possible only if $\frac{\hat{d}}{\hat{\varepsilon}} = \frac{d}{\varepsilon} = k+1$ for some $k \in \mathbb{N}$. From the relation $\min(A) = \min(\hat{B})$ we then have $\frac{d}{\varepsilon} = \left\lfloor \frac{1}{\hat{\varepsilon}} \right\rfloor$ and therefore $\left\lfloor \frac{1}{\varepsilon} \right\rfloor = \left\lfloor \frac{1}{\hat{\varepsilon}} \right\rfloor = k+1$. The latter gives us finally

$$\frac{1}{k+2} < \frac{1}{\varepsilon}, \frac{1}{\hat{\varepsilon}} < \frac{1}{k+1}\,, \qquad d = (k+1)\varepsilon\,, \quad \hat{d} = (k+1)\hat{\varepsilon}\,.$$

Therefore both $u_{\varepsilon,\eta}(\Omega)$ and $u_{\hat{\varepsilon},\hat{\eta}}(\hat{\Omega})$ are k-stuffed Sturmian words derived from Sturmian sequences with slopes $\alpha = \frac{(k+2)\varepsilon - 1}{\varepsilon}$ and $1 - \alpha = \frac{(k+2)\hat{\varepsilon} - 1}{\hat{\varepsilon}}$. □

7 Conclusions

Infinite ternary words associated with cut-and-project sequences are natural generalizations of Sturmian sequences. Every cut-and-project sequence is characterized by a triple ε, η, Ω, where ε, η are irrational numbers such that $\varepsilon \neq -\eta$, and Ω is a bounded interval. We have introduced an equivalence on the set of such triples $(\varepsilon, \eta, \Omega)$ using the combinatorial equivalence of corresponding infinite words and we have characterized the equivalence classes, cf. Theorem 6.5.

According to the complexity we may distinguish between quasisturmian sequences with complexity $n + const.$, and words of complexity $2n+1$. We have shown that a cut-and-project sequence has the complexity $n + const.$, if and only if the length of the acceptance interval Ω belongs to $\mathbb{Z}[\varepsilon]$. Rauzy graphs $(\Gamma_n)_{n\in\mathbb{N}}$ associated to an infinite word of complexity $2n+1$ are starting from an n_0 of the type 2–2, and therefore the cut-and-project sequences are different from those studied in [3]. On the other hand, construction of words of complexity $\leq 2n+1$ using cut-and-project sequences can be substituted with construction by 3-interval exchange [6].

If we impose a requirement that the set $\Sigma_{\varepsilon,\eta}(\Omega)$ is self-similar, i.e. there exists a $\gamma > 1$ such that $\gamma\Sigma_{\varepsilon,\eta}(\Omega) \subset \Sigma_{\varepsilon,\eta}(\Omega)$, then necessarily ε is a quadratic integer and $\eta = -\varepsilon'$. Moreover, γ must be a quadratic Pisot unit in the same algebraic field, $\gamma \in \mathbb{Q}[\varepsilon]$. Aperiodic sets with selfsimilarity play an important role in mathematical modeling of quasicrystals, i.e. those diffracting materials whose diffraction pattern reveals crystallographically forbidden 5-, 8-, or 12-fold symmetries. These symmetries imply that the model sets are described as subsets of rings $\mathbb{Z}[\varepsilon]$, where ε is a quadratic Pisot unit. In [9] we thus study when a selfimilar cut-and-project sequence can be generated by a substitution rule for a special case when ε is a solution of equation $x^2 = mx \pm 1$, $m \in \mathbb{N}$. We give conditions on Ω, in order that there exists an infinite word v in a finite alphabet $\mathcal{A}$ and a morphism $h : \mathcal{A} \to \{A, B, C\}$ such that v is invariant under a nontrivial substitution φ and $h(v) = u_{\varepsilon,-\varepsilon'}(\Omega)$. Mathematical tools used for the demonstration of the above statement can be extended for arbitrary quadratic irrational number ε. The invariance of words constructed by a 3-interval exchange under a morphism is studied in [1]. For construction of substitution rules for cut-and-project sequences one can use a program by Jan Patera available at [12]. A question that remains open is how one characterizes the class of matrices corresponding to these substitutions.

Acknowledgements

E.P. and Z.M. are grateful for the hospitality of CRM, Université de Montréal, and L.S.G. is grateful for the hospitality of the Department of Mathematics, FNSPE, Czech Technical University, where parts of the work were done. The authors acknowledge the financial support of the Grant Agency of the Czech Republic, GA 201/01/0130. L.S.G. acknowledges financial support by NSERC of Canada.

References

[1] B. Adamczewski, *Codages de rotations et phénomènes d'autosimilarité,* preprint Institut de Mathématiques de Luminy 2001.

[2] J. P. Allouche, *Sur la complexité des suites infinies,* Bull. Belg. Math. Soc. **1** (1994), 133–143.

[3] P. Arnoux, G. Rauzy, *Représentation géométrique de suites de complexité 2n+1,* Bull. Soc. math. France **119** (1991), 199–215.

[4] V. Berthé, *Sequences of low complexity: automatic and sturmian sequences,* in *Topics in Symbolic Dynamics and Applications*, Eds. F. Blanchard, A. Maass, A. Nogueira, Cambridge Univ. Press (2000), pp. 1–34.

[5] J. Cassaigne, *Sequences with grouped factors,* in *DLT'97, Developments in Language Theory III, Thessaloniki, Aristotle University of Thessaloniky*, (1998) 211–222, available at `ftp://iml.univ-mrs.fr/pub/cassaigne/publis/grouped.ps.gz`.

[6] J. Cassaigne, S. Ferenczi, L.Q. Zamboni, *Imbalances in Arnoux-Rauzy sequences,* Ann. Inst. Fourier (Grenoble) **50** (2000), 1265–1276.

[7] R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete mathematics. A foundation for computer science.* Second edition, Addison Wesley, Reading MA, 1994.

[8] W. F. Lunnon, P. A. B. Pleasants, *Characterization of two-distance sequences,* J. Austral. Math. Soc. (Series A) **53** (1992), 198–218.

[9] Z. Masáková, J. Patera, E. Pelantová, *Substitution rules for aperiodic sequences of the cut and project type,* J. Phys. A: Math. Gen. **33** (2000), 8867-8886.

[10] M. Morse, G. A. Hedlund, *Symbolic dynamics II: Sturmian trajectories,* Amer. J. Math. **62** (1940), 1–42.

[11] R.V. Moody, *Meyer sets and their duals,* in *Mathematics of Long Range Aperiodic Order*, Proc. NATO ASI, Waterloo, 1996, ed. R. V. Moody, Kluwer (1996) 403–441.

[12] Jan Patera, `http://kmlinux.fjfi.cvut.cz/~patera/makerule.cgi`.