

# STURM NUMBERS AND SUBSTITUTION INVARIANCE OF 3IET WORDS

PIERRE ARNOUX, VALÉRIE BERTHÉ, ZUZANA MASÁKOVÁ, AND EDITA PELANTOVÁ

ABSTRACT. In this paper, we give a necessary condition for an infinite word defined by a non-degenerate interval exchange on three intervals (3iet word) to be invariant by a substitution: a natural parameter associated to this word must be a Sturm number. We deduce some algebraic consequences from this condition concerning the incidence matrix of the associated substitution. As a by-product of our proof, we give a combinatorial characterization of 3iet words.

## 1. INTRODUCTION

The original definition of a Sturm number using continued fractions was introduced in 1993 when Crisp et al. [12] showed that a homogeneous sturmian word (i.e., a sturmian word with slope  $\varepsilon$  and intercept  $x_0 = 0$ ) is invariant under a non-trivial substitution if and only if  $\varepsilon$  is a Sturm number. In 1998, Allauzen [4] provided a simple characterization of Sturm numbers:

*A quadratic irrational number  $\varepsilon$  with conjugate  $\varepsilon'$  is called a Sturm number if*

$$(1) \quad \varepsilon \in (0, 1) \quad \text{and} \quad \varepsilon' \notin (0, 1).$$

For general sturmian words (with arbitrary intercept  $x_0$ ), the fact that  $\varepsilon$  is a Sturm number is only a necessary but not a sufficient condition for invariance under a substitution; this is clear since there can be only a countable number of such invariant words, while the sturmian words with a given slope are determined by their intercept, hence they are uncountable in number. For a complete characterization, see [20, 6, 8].

In this paper we study invariance under substitution of infinite words coding non-degenerate exchange of three intervals with permutation (321).<sup>1</sup> These words, which are here called non-degenerate 3iet words, are one of the possible generalizations of sturmian words to a three-letter alphabet. Some combinatorial properties characterizing the language of 3iet words are described in [13]. It is well known that substitutive 3iet words, that is, 3iet words that are image by a morphism of a fixed point of a substitution correspond to quadratic parameters, see e.g. [1, 3, 10, 13, 18]. Let us stress the fact that we consider in the present paper fixed points of substitutions and not substitutive words.

Sturmian words can be equivalently defined as aperiodic words coding a rotation, that is, an exchange of two intervals with lengths say  $\alpha, \beta$ . The slope of the sturmian word, which we have denoted by  $\varepsilon$ , is then equal to  $\varepsilon = \frac{\alpha}{\alpha+\beta}$ . The term ‘slope’ for the parameter  $\varepsilon$  comes from the fact that the sturmian word with slope  $\varepsilon$  can be constructed by projection of points of the lattice  $\mathbb{Z}^2$  to the straight line  $y = \varepsilon x$ ; it will prove convenient to abuse the language by speaking of the *slope of the rotation*, this slope is the complement to 1 of the more usual *angle* of the rotation. Since the sturmian word does not depend on the absolute lengths of the two intervals

---

*Date:* November 2, 2007.

<sup>1</sup>An exchange of intervals is non-degenerate if it satisfies i.d.o.c. [17]. For more details, see Section 4.3

being exchanged but one their ratio, then the lengths are often normalized to satisfy  $\alpha + \beta = 1$ . In this case  $\alpha$  and  $\varepsilon$  coincide.

The same situation appears for 3iet words which code exchange of three intervals with lengths, say,  $\alpha, \beta, \gamma$ . The commonly used normalization of parameters is  $\alpha + \beta + \gamma = 1$ . However, much more suitable appears to be the normalization  $\alpha + 2\beta + \gamma = 1$ . Let us mention three arguments in favor as results of papers [1, 5, 13, 14]. If  $u$  is an infinite word coding exchange of three intervals of lengths  $\alpha, \beta, \gamma$ , then:

- the infinite word  $u$  is aperiodic if and only if  $\frac{\alpha + \beta}{\alpha + 2\beta + \gamma} \notin \mathbb{Q}$ ;
- if  $u$  is assumed to be aperiodic, then  $u$  codes a non-degenerate exchange of three intervals if and only if  $\frac{\alpha + \beta + \gamma}{\alpha + 2\beta + \gamma} \notin \mathbb{Z} + \mathbb{Z}\frac{\alpha + \beta}{\alpha + 2\beta + \gamma}$ ;
- the infinite word  $u$  can be constructed by projection of points of the lattice  $\mathbb{Z}^2$  on the straight line  $y = \frac{\alpha + \beta}{\alpha + 2\beta + \gamma} x$ .

We will give in Section 3 and 4 a short proof of these facts by recalling that such an exchange of three intervals can always be obtained as an induced map of a rotation (exchange of two intervals) on an interval of length  $\alpha + 2\beta + \gamma$ ; in the process, we will give a complete combinatorial characterization of 3iet words, as follows:

**Theorem A.** *Let  $u$  be a sequence on the alphabet  $\{A, B, C\}$  whose letters have positive densities. Let  $\sigma : \{A, B, C\}^* \rightarrow \{0, 1\}^*$  and  $\sigma' : \{A, B, C\}^* \rightarrow \{0, 1\}^*$  be the morphisms defined by*

$$\begin{aligned} \sigma(A) &= 0, & \sigma(B) &= 01, & \sigma(C) &= 1 \\ \sigma'(A) &= 0, & \sigma'(B) &= 10, & \sigma'(C) &= 1. \end{aligned}$$

*The sequence  $u$  is an aperiodic 3iet word if and only if  $\sigma(u)$  and  $\sigma'(u)$  are sturmian words.*

This paper adds yet another argument supporting the normalization  $\alpha + 2\beta + \gamma = 1$ , by the following necessary condition:

**Theorem B.** *If a non-degenerate 3iet word is invariant under a primitive substitution, then*

$$\varepsilon := \frac{\alpha + \beta}{\alpha + 2\beta + \gamma} \quad \text{is a Sturm number.}$$

Remark that, in that case, the corresponding homogeneous sturmian word is also substitution invariant. A forthcoming paper [7] will give a complete characterization of substitution invariant 3iet words. Note that it is a natural question to ask whether, when  $u$  is a substitution invariant 3iet word, one or both of the sturmian words  $\sigma(u)$ ,  $\sigma'(u)$  are also substitution invariant.

This paper is organized as follows. The introductory notation and definitions are given in Section 2. Section 3 and 4 are devoted to the description of a classical exduction process in terms of substitutions and to the proof of Theorem A. Section 5 and Section 6 gather the required material for the proof of Theorem B, namely, properties of translation vectors and balance properties. Theorem B is proved in Section 7.

## 2. PRELIMINARIES

We work with finite and infinite words over a finite alphabet  $\mathcal{A} = \{a_1, \dots, a_k\}$ . The set of all finite words over  $\mathcal{A}$  is denoted by  $\mathcal{A}^*$ . Equipped with the binary operation of concatenation and the empty word, it is a free monoid. The length of a word  $w = w_1 w_2 \cdots w_n$  is denoted by  $|w| = n$ , the number of letters  $a_i$  in the word  $w$  is denoted by  $|w|_{a_i}$ .

An infinite concatenation of letters of  $\mathcal{A}$  forms the infinite word  $u = (u_n)_{n \in \mathbb{N}}$ ,

$$u = u_0 u_1 u_2 \cdots .$$

A word  $w$  is said to be a *factor* of a word  $u = (u_n)_{n \in \mathbb{N}}$  if there is an index  $i \in \mathbb{N}$  such that  $w = u_i u_{i+1} \cdots u_{i+n-1}$ . The set of all factors of  $u$  of length  $n$  is denoted by  $\mathcal{L}_n(u)$ . The *language*  $\mathcal{L}(u)$  of an infinite word  $u$  is the set of all its factors, that is,

$$\mathcal{L}(u) = \bigcup_{n \in \mathbb{N}} \mathcal{L}_n(u).$$

The (*factor*) *complexity*  $\mathcal{C}_u$  of an infinite word  $u$  is the function  $\mathcal{C}_u : \mathbb{N} \rightarrow \mathbb{N}$  defined as

$$\mathcal{C}_u(n) := \#\mathcal{L}_n(u).$$

The density of a letter  $a \in \mathcal{A}$ , representing the frequency of occurrence of the letter  $a$  in an infinite word  $u$ , is defined by

$$\rho(a) := \lim_{n \rightarrow \infty} \frac{\#\{i \mid 0 \leq i < n, u_i = a\}}{n},$$

if the limit exists (this is always the case for 3iet words, as is easy to prove).

Let  $\mathcal{A}$  and  $\mathcal{B}$  be two alphabets. A mapping  $\varphi : \mathcal{A}^* \rightarrow \mathcal{B}^*$  is said to be a *morphism* if  $\varphi(w\hat{w}) = \varphi(w)\varphi(\hat{w})$  holds for any pair of finite words  $w, \hat{w} \in \mathcal{A}^*$ . Obviously, a morphism is uniquely determined by the images  $\varphi(a)$  for all letters  $a \in \mathcal{A}$ . If  $\mathcal{A}$  and  $\mathcal{B}$  coincide and if the images of the letters are never equal to the empty word, then  $\varphi$  is called a *substitution*.

The action of a morphism  $\varphi$  can be naturally extended to infinite words by the prescription

$$\varphi(u) = \varphi(u_0 u_1 u_2 \cdots) := \varphi(u_0) \varphi(u_1) \varphi(u_2) \cdots .$$

A infinite word  $u \in \mathcal{A}^{\mathbb{N}}$  is said to be a *fixed point* of the morphism  $\varphi$  if  $\varphi(u) = u$ .

The incidence matrix of a morphism  $\varphi$  over the alphabet  $\mathcal{A}$  is an important tool which brings a lot of information about the combinatorial properties of the fixed points of the morphism. It is defined by

$$(\mathbf{M}_\varphi)_{ij} = |\varphi(a_i)|_{a_j} = \text{number of letters } a_j \text{ in the word } \varphi(a_i).$$

A morphism  $\varphi$  is called *primitive* if there exists an integer  $k$  such that the matrix  $\mathbf{M}_\varphi^k$  is positive.

Assume that an infinite word  $u$  over the alphabet  $\mathcal{A} = \{a_1, \dots, a_k\}$  is a fixed point of a primitive substitution  $\varphi$ . It is known [19] that in such a case the densities of letters in  $u$  are well defined. The vector

$$\vec{\rho}_u = (\rho(a_1), \dots, \rho(a_k)).$$

is a left eigenvector of the incidence matrix  $\mathbf{M}_\varphi$ , i.e.,  $\vec{\rho}_u \mathbf{M}_\varphi = \Lambda \vec{\rho}_u$ . Since the incidence matrix  $\mathbf{M}_\varphi$  is a non-negative integral matrix, we can use the Perron-Frobenius Theorem stating that  $\Lambda$  is the dominant eigenvalue of  $\mathbf{M}_\varphi$ . Moreover, all eigenvalues of  $\mathbf{M}_\varphi$  are algebraic integers.

### 3. EXCHANGES OF THREE INTERVALS AS INDUCTION OF ROTATIONS

Let  $\alpha, \beta, \gamma > 0$  and denote by

$$I_A := [0, \alpha), \quad I_B := [\alpha, \alpha + \beta), \quad I_C := [\alpha + \beta, \alpha + \beta + \gamma), \quad \text{and } I := I_A \cup I_B \cup I_C,$$

and let  $t_A = \beta + \gamma, t_B = \gamma - \alpha, t_C = -\alpha - \beta \in \mathbb{R}$  be translations vectors; we have:

$$I_A \cup I_B \cup I_C = (I_A + t_A) \cup (I_B + t_B) \cup (I_C + t_C).$$

The map  $T$  defined on  $I$  by  $T(x) = x + t_X$  if  $x \in I_X, X = A, B, C$  is the exchange of three intervals  $I_A, I_B, I_C$  with the permutation (321).

As was already known long time ago (see [16]), this map can be obtained as the induction of a rotation on a suitable interval. We recall the construction; let  $I_D = [\alpha + \beta + \gamma, \alpha + 2\beta + \gamma)$ , and define  $J = I \cup I_D$ . Let  $R$  be the rotation of angle  $\frac{\beta + \gamma}{\alpha + 2\beta + \gamma}$  on  $J$ , defined by  $R(x) = x + \beta + \gamma$  if  $x \in I_A \cup I_B$ , and  $R(x) = x - \alpha - \beta$  if  $x \in I_C \cup I_D$ ; it exchanges the two intervals  $I_A \cup I_B$  and  $I_C \cup I_D$ .

For a subset  $E$  of  $X$ , the *first return time*  $r_E(x)$  of a point  $x \in E$  is defined as  $\min\{n > 0 \mid R^n x \in E\}$ . If the return time is always finite, we define the *induced map* or *first return map* of  $R$  on  $E$  by  $R_E(x) = R^{r_E(x)}(x)$ .

**Lemma 3.1.** *The map  $T$  is the first return map of  $R$  on  $I$ .*

*Proof.* One checks that  $R(I_A) = [\gamma + \beta, \alpha + \gamma + \beta) = T(I_A)$ ,  $R(I_C) = [0, \alpha) = T(I_C)$ ,  $R(I_B) = I_D$ , and  $R^2(I_B) = R(I_D) = T(I_B)$ .  $\square$

Hence  $T$  can be obtained as induction of a rotation on a left interval (for more details, see e.g. [1, 13] or the survey [9]). It can also be obtained as induction on a right interval, and this remark will prove important below: define  $I_E = [-\beta, 0)$ , and  $J' = I_E \cup I$ ; consider the rotation  $R'$  on  $J'$  by the same angle  $\frac{\beta + \gamma}{\alpha + 2\beta + \gamma}$ ; in the same way, one proves that  $T$  is obtained as the first return map of  $R'$  on  $I$ .

The underlying rotation  $R$  turns out to play an important role in the study of  $T$ ; this explains the appearance of the number  $\varepsilon = \frac{\alpha + \beta}{\alpha + 2\beta + \gamma}$  in the introduction: it is the slope of the rotation  $R$ .

**Notation 3.1.** *From now on, we will take the normalization  $\alpha + 2\beta + \gamma = 1$ .*

This amounts to normalize the interval of definition of  $R$  to 1, and will greatly simplify the notation below.

#### 4. CHARACTERIZATION OF NON-DEGENERATE 3IET WORDS

**4.1. From 3iet words to sturmian words.** With an initial point  $x_0 \in I$  we associate an infinite word which codes the orbit of  $x_0$  under  $T$  with respect to the natural partition in three intervals (see Definition 4.1 below). It turns out to be useful to shift the interval of definition, so that the free choice of the initial point  $x_0$  for the orbit is replaced by the choice of a parameter that we call  $c$  as the position of the interval. The initial point for the orbit thus can always be chosen as the origin. For this we introduce the new parameters

$$\varepsilon := \alpha + \beta, \quad l := \alpha + \beta + \gamma, \quad c := -x_0,$$

The number  $\varepsilon$  is the slope of the underlying rotation  $R$ , and  $l$  determines the length of the induction interval  $J$ . It is obvious that the above parameters satisfy

$$(2) \quad \varepsilon \in (0, 1), \quad \max(\varepsilon, 1 - \varepsilon) < l < 1, \quad -l < c \leq 0.$$

We redefine in this setting five intervals

$$I_A = [c, c + \alpha), \quad I_B = [c + \alpha, c + \varepsilon), \quad I_C = [c + \varepsilon, c + l), \quad I_D = [c + l, c + 1), \quad I_E = [c - \beta, c).$$

We define  $I = I_A \cup I_B \cup I_C$ ; the map  $T$  (introduced above in Section 3) is defined on  $I$  as the exchange of three intervals  $I_A, I_B, I_C$  according to the permutation (321).

We also define

$$\begin{aligned} J_0 &= I_A \cup I_B, \quad J_1 = I_C \cup I_D \text{ and } J = J_0 \cup J_1 \\ J'_0 &= I_E \cup I_A, \quad J'_1 = I_B \cup I_C \text{ and } J' = J'_0 \cup J'_1. \end{aligned}$$

The rotation  $R$  (resp.  $R'$ ) is then defined on  $J$  (resp.  $J'$ ) by the exchange of  $J_0$  and  $J_1$  (resp.  $J'_0$  and  $J'_1$ ); it has angle  $1 - \varepsilon$ ,  $J_0$  and  $J'_0$  have length  $\varepsilon$ , whereas  $J_1$  and  $J'_1$  have length  $1 - \varepsilon$ .

Let us formulate the definition of 3iet words with the use of these new parameters.

**Definition 4.1.** Let  $\varepsilon, l, c \in \mathbb{R}$  satisfy (2). The infinite word  $(u_n)_{n \in \mathbb{N}}$  defined by

$$(3) \quad u_n = \begin{cases} A & \text{if } T^n(0) \in I_A, \\ B & \text{if } T^n(0) \in I_B, \\ C & \text{if } T^n(0) \in I_C \end{cases}$$

is called the 3iet word with parameters  $\varepsilon, l, c$ .

There is classical and simple way to give a combinatorial interpretation of the induction process of Lemma 3.1 in terms of substitutions. Consider indeed the orbit  $(T^n(0))_{n \in \mathbb{N}}$  of 0 under  $T$ ; it is clear by Lemma 3.1 that it is a subset of the orbit  $(R^n(0))_{n \in \mathbb{N}}$  under  $R$ ; the points of the second orbit which are not in the first are exactly the points in  $I_D$ , and their preimages are exactly the points in  $I_B$ ; the return time of these points to  $I$  is 2. Let  $u$  be the coding of the orbit of 0 under  $T$  with respect to the partition in three intervals  $I_A, I_B, I_C$ ; to obtain the coding of the orbit of the same point under  $R$ , with respect to the partition  $I_A, I_B, I_C, I_D$ , this argument shows that it is enough to introduce a letter  $D$  after each  $B$ , that is to replace  $B$  by  $BD$ ; to obtain the natural sturmian coding with respect to the partition  $J_0, J_1$ , we then project letters  $A, B$  to 0 and  $C, D$  to 1.

**Definition 4.2.** We denote by  $\sigma$  (resp.  $\sigma'$ ) the morphism from  $\{A, B, C\}^*$  to  $\{0, 1\}^*$  defined by  $\sigma(A) = 0, \sigma(B) = 01, \sigma(C) = 1$  (resp.  $\sigma'(A) = 0, \sigma'(B) = 10, \sigma'(C) = 1$ ).

We thus have proved the following:

**Lemma 4.3.** Let  $u$  be the coding of the orbit of 0 under  $T$ , with respect to the partition  $I_A, I_B, I_C$ , and let  $v$  (resp.  $v'$ ) be the coding of the orbit of 0 under  $R$  (resp.  $R'$ ) with respect to the partition  $J_0, J_1$  (resp.  $J'_0, J'_1$ ). Then  $v = \sigma(u), v' = \sigma(u')$ .

This implies that, if  $\varepsilon$  is irrational, then  $\sigma(u)$  and  $\sigma(u')$  are sturmian sequences whose density of 0 equals  $\varepsilon$ .

**4.2. Characterization theorem.** We will now prove the reciprocal (Theorem A below); we need some properties of sturmian sequences.

Let  $v$  be a sturmian sequence that codes the orbit of a rotation of angle  $1 - \varepsilon$  modulo 1 with density of 0 equal to  $\varepsilon$ , and let  $V_n$  be the prefix of  $v$  of length  $n$ , i.e.,  $V_n = v_0 \cdots v_{n-1}$ . Define a map

$$f : \{0, 1\}^* \rightarrow \mathbb{R} \text{ by } f(V) = |V|_0(1 - \varepsilon) - |V|_1\varepsilon.$$

From the definition, we see that

$$\forall n, f(V_n) = |V_n|_0 - n\varepsilon,$$

hence the sequence  $(f(V_n))_{n \in \mathbb{N}}$  is the orbit of 0 under a rotation defined on an interval  $[c, c + 1)$ , with  $c = \inf\{f(V_n) | n \in \mathbb{N}\}$ ; in particular, we have, for all integers  $i, j$ ,  $|f(V_i) - f(V_j)| < 1$ .

We have the following lemma:

**Lemma 4.4.** Let  $v$  be a sturmian sequence, and let  $(n_k)_{k \in \mathbb{N}}$  be a strictly increasing sequence of integers that satisfies  $v_{n_k-1} = 0, v_{n_k} = 1$ . Define a new sequence  $v'$  by:  $v'_{n_k-1} = 1, v'_{n_k} = 0, v'_i = v_i$  otherwise. The sequence  $v'$  is sturmian if and only if for every  $i$  which is not in the sequence  $(n_k)_{k \in \mathbb{N}}$ , and for all  $j$ , we have  $f(V_{n_j}) > f(V_i)$ .

*Proof.* For all  $n$ , let  $V'_n$  stand for the prefix of  $v'$  of length  $n$ . We have  $f(V_i) = f(V'_i)$ , except if  $i = n_k$ , in which case one checks that  $f(V'_{n_k}) = f(V_{n_k}) - 1$ .

We first assume that  $v'$  is sturmian. Suppose that  $f(V_i) \geq f(V_{n_j})$ , for some  $i, j$ , with  $i$  not in the sequence  $(n_k)_{k \in \mathbb{N}}$ ; then we must have  $f(V'_i) \geq f(V'_{n_j}) + 1$ ; but this is impossible since  $v'$  is a

sturmian sequence with same density of 0's as  $v$ . Hence for every  $i$  which is not in the sequence  $(n_k)_{k \in \mathbb{N}}$ , and for all  $j$ , we have  $f(V_i) < f(V_{n_j})$ .

Conversely, we assume that for every  $i$  which is not in the sequence  $(n_k)_{k \in \mathbb{N}}$ , and for all  $j$ , we have  $f(V_{n_j}) > f(V_i)$ . One checks that for all integers  $i, j$ ,  $|f(V'_i) - f(V'_j)| < 1$ . Indeed, this is immediate if  $i$  and  $j$  belong simultaneously to  $(n_k)_{k \in \mathbb{N}}$ , or else if none of them belongs to this sequence. If  $i$  is not in the sequence  $(n_k)_{k \in \mathbb{N}}$ , then  $|f(V'_i) - f(V'_{n_j})| = |f(V_i) - f(V_{n_j}) + 1| < 1$ . We deduce that the sequence  $v'$  is a balanced sequence. Indeed, take two factors  $W$  and  $W'$  of the same length  $n$  of the sequence  $v'$  that occur respectively at index  $i$  and  $j$ . One has

$$||W|_0 - |W'|_0| = |(f(V'_{i+n}) - f(V'_{j+n})) - (f(V'_i) - f(V'_j))| < 2.$$

We deduce that the densities of letters are well-defined in  $v'$ . By construction, they coincide with the densities of letters for the sequence  $v$ , hence  $v'$  is an aperiodic balanced sequence, it is thus a sturmian sequence, according to [15].  $\square$

We are now in position to prove the first theorem:

**Theorem A.** *Let  $u$  be a sequence on the alphabet  $\{A, B, C\}$  whose letters have positive densities. This sequence is an aperiodic 3iet word if and only if  $\sigma(u)$  and  $\sigma'(u)$  are sturmian words.*

*Proof.* We have proved above (Lemma 4.3) that the condition is necessary. Let us prove it is sufficient. Let  $v = \sigma(u)$ ,  $v' = \sigma'(u)$  be the two sturmian words; by construction, they have the same slope  $\varepsilon$ , and they coincide except on a sequence of pairs of indices  $(n_k - 1, n_k)$ , corresponding to the images of  $B$ , where 0 is replaced by 1 and vice versa.

Define the function  $f$  as above, and define  $c = \inf\{f(V_k) \mid k \in \mathbb{N}\}$ ,  $l = \inf\{f(V_{n_k}) \mid k \in \mathbb{N}\} - c$ . From Lemma 4.4, we deduce that an index  $i$  is of the form  $n_k$  if and only if  $f(V_i) \geq l + c$  if  $\inf\{f(V_{n_k})\} = \min\{f(V_{n_k})\}$  (resp.  $f(V_i) > l + c$  otherwise). Then, one checks that the sequence  $u$  is generated by the exchange  $T$  of the three intervals  $I_A, I_B, I_C$  with either

$$I_A = [c, c + \varepsilon + l - 1), I_B = [c + \varepsilon + l - 1, c + \varepsilon), I_C = [c + \varepsilon, c + l)$$

or

$$I_A = (c, c + \varepsilon + l - 1], I_B = (c + \varepsilon + l - 1, c + \varepsilon], I_C = (c + \varepsilon, c + l],$$

the choice of the intervals being determined by the values of  $u$ , and thus of  $v$  and  $v'$ , at the indices (if any) where the orbit of 0 under  $T$  meets discontinuity points. The interval  $I_B$  corresponds to the times  $n_k - 1$  for the sequence  $v$ , and  $I_A$  and  $I_C$  resp. to value 0 and 1 for the other times. We deduce that  $u$  is aperiodic from the irrationality of  $\varepsilon$ .  $\square$

Figure 1 gives a geometric interpretation of the proof; to the 3iet word  $u$ , we have associated a stepped line (bold line), by associating letter  $A$  to vector  $(1, 0)$ ,  $B$  to  $(1, 1)$ , and  $C$  to  $(0, 1)$ . Remark that this stepped line is contained in a “corridor” of width less than 1; in dashed lines are shown the two sturmian lines associated to  $v$  and  $v'$ , obtained by enlarging the corridor on the right or the left to the width of the unit square.

**4.3. Complexity.** It is known that the factor complexity of the infinite words (3) satisfies  $\mathcal{C}(n) \leq 2n + 1$  for all  $n \in \mathbb{N}$ . A short proof can be given by considering the partition  $\mathcal{P}$  in three intervals; to count the number of factors of length  $n$ , it is enough to count the number of atoms of the partition  $\bigvee_{k=0}^{n-1} T^{-k}\mathcal{P}$ . But it is easy to prove that these atoms are intervals, bounded by reciprocal images of the two discontinuity points. As there can be at most  $2n$  such points between time 0 and  $n - 1$ , there are at most  $2n + 1$  intervals.

The infinite words  $(u_n)_{n \in \mathbb{N}}$  which have full complexity are called *non-degenerate* (or regular) 3iet words; 3iet words for which there exists  $n$  such that  $\mathcal{C}(n) < 2n + 1$  are called degenerate.

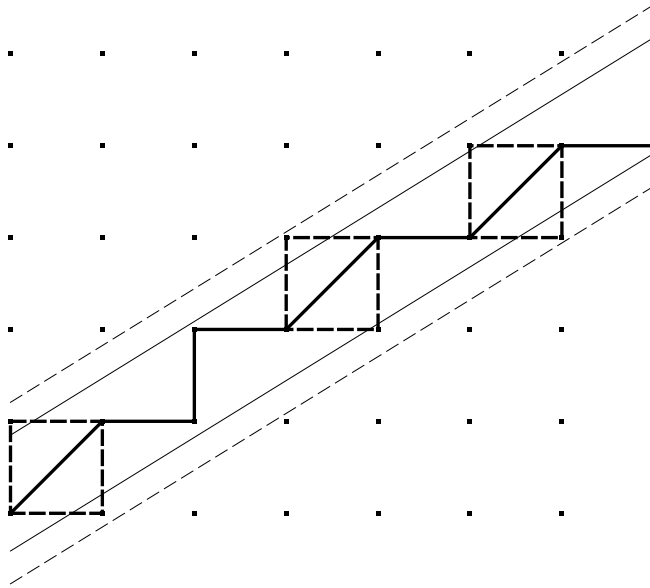


FIGURE 1. The stepped line associated to a 3-iet word and its two sturmian extensions

The necessary and sufficient condition for a word  $(u_n)_{n \in \mathbb{N}}$  coding 3iet to be non-degenerate is the so-called i.d.o.c. (infinite distinct orbit condition). This notion has been introduced by Keane [17] and requires, in this case, that the orbits of the two points of discontinuity of the transformation  $T$  are disjoint, formally  $\{T^n(c + l - 1 + \varepsilon)\}_{n \in \mathbb{N}} \cap \{T^n(c + \varepsilon)\}_{n \in \mathbb{N}} = \emptyset$ . If this condition holds true, then the partition above is limited by exactly  $2n$  points on the interval, hence has  $2n + 1$  atoms. The condition i.d.o.c is equivalent to

$$(4) \quad \varepsilon \notin \mathbb{Q} \quad \text{and} \quad l \notin \mathbb{Z} + \mathbb{Z}\varepsilon =: \mathbb{Z}[\varepsilon],$$

see [1, 14].

**Remark 4.5.** *If  $\varepsilon$  is irrational, it is classical that the rotation  $R$  is uniquely ergodic, which implies that  $T$  is also uniquely ergodic. In that case, the densities of letters in the 3iet aperiodic word are well defined and  $\vec{\rho}_u$  is proportional to the vector of lengths of intervals  $I_A, I_B, I_C$ .*

*If  $\varepsilon$  is rational, the sequence  $u$  is periodic, hence the densities exist in a trivial way.*

## 5. TRANSLATION VECTORS

Let  $u$  be a 3iet word such as defined in Definition 4.1. In our considerations, the column vector of translations will play a crucial role. We denote it by

$$\vec{t} = \begin{pmatrix} t_A \\ t_B \\ t_C \end{pmatrix} = \begin{pmatrix} 1 - \varepsilon \\ 1 - 2\varepsilon \\ -\varepsilon \end{pmatrix}.$$

A first remark is that the vector of translations is orthogonal to the vector of densities; this can be checked directly, and interpreted as the fact that the mean translation is 0, because the orbit under the action of the map  $T$  is bounded.

We assume furthermore that  $u$  is fixed by some substitution  $\varphi$ . We will now obtain a more subtle equation, using the substitution  $\varphi$ . Let us define a function  $g$  (in the flavour of the map

$f$  defined in Section 4.2) on the prefixes of the infinite word  $u$ , the fixed point of  $\varphi$ . For the prefix  $w = u_0u_1 \cdots u_{n-1}$ ,  $n \geq 0$ , we put

$$g(u_0u_1 \cdots u_{n-1}) := T^n(0) = |w|_A t_A + |w|_B t_B + |w|_C t_C.$$

In particular, the image of the empty word equals 0. For  $X \in \{A, B, C\}$ , put

$$E_X := \left\{ g(u_0u_1 \cdots u_{n-1}) \mid u_n = X \right\} = \left\{ (|w|_A, |w|_B, |w|_C) \vec{t} \mid wX \text{ is a prefix of } u \right\}.$$

Clearly, the closure of the set  $E_X$  satisfies  $\overline{E}_X = I_X$ .

The infinite word  $u_0u_1u_2 \cdots = \varphi(u_0)\varphi(u_1)\varphi(u_2) \cdots$  can be imagined as a concatenation of blocks  $\varphi(A)$ ,  $\varphi(B)$ ,  $\varphi(C)$ . Positions, where these blocks start, and the corresponding iterations of  $T$ , are given by the following sets. For  $X \in \{A, B, C\}$ , put

$$E_{\varphi(X)} := \left\{ g(\varphi(u_0u_1 \cdots u_{n-1})) \mid u_n = X \right\} = \left\{ g(\varphi(w)) \mid wX \text{ is a prefix of } u \right\}.$$

From the definition of the matrix  $\mathbf{M}_\varphi$  it follows that

$$(5) \quad E_{\varphi(X)} = \left\{ (|w|_A, |w|_B, |w|_C) \mathbf{M}_\varphi \vec{t} \mid wX \text{ is a prefix of } u \right\}.$$

Obviously,

$$E_{\varphi(A)} \cup E_{\varphi(B)} \cup E_{\varphi(C)} \subset \{T^n(0) \mid n \in \mathbb{N}\} \subset I,$$

and the union is disjoint. The fact that  $T^k(0)$  belongs to  $E_{\varphi(A)}$  is equivalent to

- $u_k u_{k+1} u_{k+2} \cdots$  has the prefix  $\varphi(A)$ ;
- $u_0 u_1 \cdots u_{k-1} = \varphi(u_0 u_1 \cdots u_{i-1})$  for some  $i \in \mathbb{N}$ ;
- $u_i = A$ .

Similar statement is true for the elements of the sets  $E_{\varphi(B)}$  and  $E_{\varphi(C)}$ . Moreover, from the construction of  $E_{\varphi(X)}$  it follows that if  $T^k(0) \in E_{\varphi(X)}$ , then the smallest  $n > k$  for which  $T^n(0) \in E_{\varphi(A)} \cup E_{\varphi(B)} \cup E_{\varphi(C)}$  satisfies  $n - k = |\varphi(X)|$ .

The infinite word  $u$  can therefore be interpreted as a word coding exchange of three sets  $E_{\varphi(A)}$ ,  $E_{\varphi(B)}$ ,  $E_{\varphi(C)}$ , with translations

$$t_{\varphi(X)} := |\varphi(X)|_A t_A + |\varphi(X)|_B t_B + |\varphi(X)|_C t_C.$$

Obviously, one has

$$(E_{\varphi(A)} + t_{\varphi(A)}) \cup (E_{\varphi(B)} + t_{\varphi(B)}) \cup (E_{\varphi(C)} + t_{\varphi(C)}) = E_{\varphi(A)} \cup E_{\varphi(B)} \cup E_{\varphi(C)} \subset I.$$

From the definition of  $t_{\varphi(X)}$ , it follows that the translation vector  $\vec{t}_\varphi = (t_{\varphi(A)}, t_{\varphi(B)}, t_{\varphi(C)})^T$  satisfies

$$(6) \quad \vec{t}_\varphi = \mathbf{M}_\varphi \vec{t}.$$

## 6. BALANCE PROPERTIES OF FIXED POINTS OF SUBSTITUTIONS

**Definition 6.1.** *We say that an infinite word  $u = (u_n)_{n \in \mathbb{N}}$  has bounded balances, if there exists  $0 < K < +\infty$  such that for all  $n \in \mathbb{N}$ , and for all pairs of factors  $w, \hat{w} \in \mathcal{L}_n(u)$ , it holds that*

$$||w|_a - |\hat{w}|_a| \leq K, \quad \text{for all } a \in \mathcal{A}.$$

The above definition is a generalization of the notion of balanced words, which correspond to a constant  $K$  equal to 1. We have used the fact that aperiodic balanced words over a binary alphabet are precisely the sturmian words in the proof of Lemma 4.4 [15]. The balance properties of the considered generalization of sturmian words, the 3iet words, are more complicated. The following is a consequence of results in [1].

**Proposition 6.2.** *Let  $u$  be a 3iet word. Then  $u$  has bounded balances if and only if it is degenerated.*

In this paper we focus on substitution invariant non-degenerate 3iet words. We shall make use of the following result of Adamczewski [2], which describes the balance properties of fixed points of substitutions dependently on the spectrum of the incidence matrix. We mention only that part of his Theorem 13 which will be useful in our considerations.

**Proposition 6.3.** *Let the infinite word  $u$  be invariant under a primitive substitution  $\varphi$  with incidence matrix  $\mathbf{M}_\varphi$ . Let  $\Lambda$  be the dominant eigenvalue of  $\mathbf{M}_\varphi$ . If  $|\lambda| < 1$  for all other eigenvalues  $\lambda$  of  $\mathbf{M}_\varphi$ , then  $u$  has bounded balances.*

## 7. NECESSARY CONDITIONS FOR SUBSTITUTION INVARIANCE OF 3IET WORDS

We now have gathered all the required material for the proof of Theorem B which provides necessary conditions on the parameters of the studied 3iet words to be invariant under substitution.

**Theorem B.** *Let  $u = (u_n)_{n \in \mathbb{N}}$  be a non-degenerate 3iet word with parameters  $\varepsilon, l, c$  satisfying (2) and (4). Let  $\varphi$  be a primitive substitution such that  $\varphi(u) = u$ . Then the parameter  $\varepsilon$  is a Sturm number.*

*Proof.* The density vector of the word  $u$  is the vector  $\vec{\varrho}_u = (1 - \frac{1-\varepsilon}{l}, \frac{1}{l} - 1, 1 - \frac{\varepsilon}{l})$ . The vector  $\vec{\varrho}_u$  is a left eigenvector corresponding to the Perron-Frobenius eigenvalue  $\Lambda$ . Since  $\vec{\varrho}_u$  is an irrational vector and  $\mathbf{M}_\varphi$  an integral matrix,  $\mathbf{M}_\varphi$  has 3 different eigenvalues. Denote the other eigenvalues of  $\mathbf{M}_\varphi$  by  $\lambda_1, \lambda_2$  and by  $\vec{x}_1, \vec{x}_2$  the right eigenvectors of the matrix  $\mathbf{M}_\varphi$  corresponding to  $\lambda_1$  and  $\lambda_2$ , respectively, i.e.,

$$(7) \quad \mathbf{M}_\varphi \vec{x}_1 = \lambda_1 \vec{x}_1 \quad \text{and} \quad \mathbf{M}_\varphi \vec{x}_2 = \lambda_2 \vec{x}_2.$$

A left eigenvector and a right eigenvector of a matrix corresponding to different eigenvalues are mutually orthogonal. Therefore the vectors  $\vec{x}_1, \vec{x}_2$  form a basis of the orthogonal plane to the left eigenvector corresponding to  $\Lambda$ . Since the vector  $\vec{t} = (1 - \varepsilon, 1 - 2\varepsilon, -\varepsilon)^T$  is orthogonal to  $\vec{\varrho}_u$ , we can write

$$(8) \quad \vec{t} = \mu \vec{x}_1 + \nu \vec{x}_2 \quad \text{for some } \mu, \nu \in \mathbb{C}.$$

Our aim is now to show that either  $\mu = 0$  or  $\nu = 0$ , i.e., that the vector  $\vec{t}$  is a right eigenvector of the matrix  $\mathbf{M}_\varphi$ .

Recall that the translation vector  $\vec{t}_\varphi$  satisfies (6). Since this holds for any substitution which has  $u_0 u_1 u_2 \dots$  for its fixed point, one can write

$$(9) \quad \vec{t}_{\varphi^n} = \mathbf{M}_{\varphi^n} \vec{t} = \mathbf{M}_\varphi^n \vec{t}.$$

Since  $\vec{t}_{\varphi^n}$  represents translations of subsets of a bounded interval  $I$ , the vector  $\vec{t}_{\varphi^n}$  must have bounded components. Combination of (7), (8), and (9) leads to the fact that the sequence of vectors

$$(10) \quad \mathbf{M}_\varphi^n \vec{t} = \mu \lambda_1^n \vec{x}_1 + \nu \lambda_2^n \vec{x}_2$$

is bounded.

We shall now distinguish two cases. Realize that the Perron eigenvalue  $\Lambda$  must be an algebraic integer either of degree three or of degree two.

*The cubic case.* Suppose that  $\Lambda$  is a cubic number. Then  $\lambda_1, \lambda_2$  are its algebraic conjugates. By assumption  $u$  is a non-degenerate 3iet word, and thus using Proposition 6.2 and Proposition 6.3 and the fact that Salem numbers<sup>2</sup> of degree 3 do not exist, we derive that one of the eigenvalues  $\lambda_1, \lambda_2$  is in modulus greater than 1, say  $|\lambda_2| > 1$ . Boundedness of the sequence of vectors  $(\mathbf{M}_\varphi^n \vec{t})_{n \in \mathbb{N}}$  in (10) implies that  $\nu = 0$  and thus  $\vec{t}$  is a right eigenvector of the matrix  $\mathbf{M}_\varphi$ , without loss of generality, we can put  $\vec{x}_1 = \vec{t}$ .

But then the components of the vector  $\vec{x}_1 = \vec{t} = (1 - \varepsilon, 1 - 2\varepsilon, -\varepsilon)^T$  belong to the field  $\mathbb{Q}(\lambda_1)$ , however, the first plus the last components of the vector  $\vec{x}_1$  are equal to the middle one, which gives a quadratic equation for  $\lambda_1$ . This is a contradiction, hence that case is impossible.

*The quadratic case.* We have shown that  $\Lambda$  is a quadratic number. In this case, the other eigenvalues of  $\mathbf{M}_\varphi$  are the conjugate  $\lambda_1 = \Lambda'$  of  $\Lambda$  and  $\lambda_2 = r \in \mathbb{Z}$ . Irrationality of the vector  $\vec{t}$  implies that  $\mu \neq 0$ . Let us suppose that  $\nu \neq 0$ , as well. Boundedness of  $\mathbf{M}_\varphi^n \vec{t}$  in (10) implies that  $|\Lambda'| < 1$  and  $|r| \leq 1$ . By Proposition 6.3, we have  $|r| \geq 1$  and thus  $r = \pm 1$ . Without loss of generality, we can assume that  $r = 1$ , otherwise we consider the morphism  $\varphi^2$  instead of  $\varphi$ . For the vector  $\vec{t}_{\varphi^n}$  of translations of the sets  $E_{\varphi^n(A)}, E_{\varphi^n(B)}, E_{\varphi^n(C)}$ , it holds that

$$\vec{t}_{\varphi^n} = \mu(\Lambda')^n \vec{x}_1 + \nu \vec{x}_2 \quad \xrightarrow{n \rightarrow \infty} \quad \nu \vec{x}_2 \neq \vec{0}.$$

We shall make use of the following property of infinite words coding 3iet. For arbitrary factor  $w \in \mathcal{L}(u)$  denote by  $I_w$  the closure of the set  $\{T^n(0) \mid w \text{ is a prefix of } u_n u_{n+1} u_{n+2} \cdots\}$ . It is known that  $I_w$  is an interval. With growing length of  $w$ , the length  $|I_w|$  of the interval  $I_w$  approaches to 0. Since the morphism  $\varphi$  is primitive, the length  $\varphi^n(X)$  grows to infinity with growing  $n$  for every letter  $X$ . Obviously  $E_{\varphi^n(X)} \subset I_{\varphi^n(X)}$  and  $\lim_{n \rightarrow \infty} |I_{\varphi^n(X)}| = 0$ .

Recall that  $E_{\varphi^n(A)}, E_{\varphi^n(B)}, E_{\varphi^n(C)}$  are disjoint and their union is equal to  $(E_{\varphi^n(A)} + t_{\varphi^n(A)}) \cup (E_{\varphi^n(B)} + t_{\varphi^n(B)}) \cup (E_{\varphi^n(C)} + t_{\varphi^n(C)})$ . Since by assumption  $\lim_{n \rightarrow \infty} \vec{t}_{\varphi^n} = \nu \vec{x}_2 \neq \vec{0}$ , for sufficiently large  $n$ , one of the following is true:

– either there exist  $X, Y \in \{A, B, C\}$ ,  $X \neq Y$  such that

$$E_{\varphi^n(X)} = E_{\varphi^n(Y)} + t_{\varphi^n(Y)},$$

– or for mutually distinct letters  $X, Y, Z$  of the alphabet we have

$$E_{\varphi^n(X)} \cup E_{\varphi^n(Z)} = E_{\varphi^n(Y)} + t_{\varphi^n(Y)}.$$

This would however mean for the densities of letters that  $\varrho(Z) = \varrho(X) = \varrho(Y)$ , or  $\varrho(Y) = \varrho(X) + \varrho(Z)$ , respectively. This contradicts the fact that  $u$  is a non-degenerate 3iet word. Hence the assumption  $\nu \neq 0$  leads to a contradiction.

Thus by (8), the vector  $\vec{t}$  is a right eigenvector of the matrix  $\mathbf{M}_\varphi$  corresponding to the eigenvalue  $\Lambda'$ .

Since  $\Lambda$  is a quadratic number,  $\varepsilon$  is also a quadratic number and  $\Lambda \in \mathbb{Q}(\varepsilon') = \mathbb{Q}(\varepsilon)$ , where  $\varepsilon'$  is the algebraic conjugate of  $\varepsilon$ . Applying the Galois automorphism of the field  $\mathbb{Q}(\varepsilon)$  we obtain that the vector  $\vec{t}' := (1 - \varepsilon', 1 - 2\varepsilon', -\varepsilon')^T$  is a right eigenvector corresponding to  $\Lambda$ , i.e., it has either all components positive or all negative. Therefore we have  $(1 - \varepsilon')\varepsilon' < 0$ , which means that  $\varepsilon$  is a Sturm number.  $\square$

The proof of Theorem B provides several direct consequences.

<sup>2</sup>An algebraic integer is called a Salem number, if all its algebraic conjugates are in modulus  $\leq 1$  and at least one of them lies on the unit circle. It is known [11] that all Salem numbers are of even degree greater than or equal to 4.

**Corollary 7.1.** *Let  $u = (u_n)_{n \in \mathbb{N}}$  be a non-degenerate 3iet word with parameters  $\varepsilon, l, c$  satisfying (2) and (4). Let  $\varphi$  be a primitive substitution such that  $\varphi(u) = u$ . Then*

- *the incidence matrix  $\mathbf{M}_\varphi$  of  $\varphi$  is non-singular;*
- *its Perron-Frobenius eigenvalue is a quadratic number  $\Lambda \in \mathbb{Q}(\varepsilon)$ ;*
- *its right eigenvector corresponding to  $\Lambda$  is equal to  $(1 - \varepsilon', 1 - 2\varepsilon', -\varepsilon')^T$ , where  $\varepsilon'$  is the algebraic conjugate of  $\varepsilon$ .*

Another consequence of the proof of Theorem 7 is that the Perron-Frobenius eigenvalue of the incidence matrix  $\mathbf{M}_\varphi$  of the substitution  $\varphi$  under which a 3iet word is invariant is an algebraic unit. Before stating this result, realize that since  $\vec{t}$  is an eigenvector of  $\mathbf{M}_\varphi$  corresponding to  $\Lambda'$ , the definition of the set  $E_X$  and the equation (5) imply

$$(11) \quad E_{\varphi(X)} = \Lambda' E_X.$$

In accordance with the definition of translations  $t_X$  and  $t_{\varphi(X)}$  for a letter  $X$  in the alphabet  $\mathcal{A} = \{A, B, C\}$  we can more generally introduce the translation  $t_w$  for any finite word  $w \in \mathcal{L}(u)$ , as

$$t_w := |w|_A t_A + |w|_B t_B + |w|_C t_C.$$

With this notation, we can describe several properties of the sets  $E_{\varphi(X)} + t_w$ , where  $w$  is a proper prefix of  $\varphi(X)$ ,  $X \in \mathcal{A}$ . (The number of these sets is  $|\varphi(A)| + |\varphi(B)| + |\varphi(C)|$ .) The substitution invariance of  $u$  under  $\varphi$ ,  $u_0 u_1 u_2 \cdots = \varphi(u_0) \varphi(u_1) \varphi(u_2) \cdots$ , implies the following facts.

- (1)  $E_{\varphi(X)} + t_w = T^{|w|}(E_{\varphi(X)})$ .
- (2) The sets  $E_{\varphi(X)} + t_w$ , where  $w$  is a proper prefix of  $\varphi(X)$ ,  $X \in \mathcal{A}$ , are mutually disjoint.
- (3) For any letter  $X \in \mathcal{A}$  and for every proper prefix  $w$  of  $\varphi(X)$ , there exists a letter  $Y \in \mathcal{A}$  such that  $E_{\varphi(X)} + t_w \subseteq E_Y$ .
- (4) 
$$\bigcup_{x \in \mathcal{A}} \bigcup_{\substack{w \text{ is a proper} \\ \text{prefix of } \varphi(X)}} (E_{\varphi(X)} + t_w) = E_A \cup E_B \cup E_C.$$

**Corollary 7.2.** *Let  $u = (u_n)_{n \in \mathbb{N}}$  be a non-degenerate 3iet word with parameters  $\varepsilon, l, c$  satisfying (2) and (4). Let  $\varphi$  be a primitive substitution such that  $\varphi(u) = u$ . Then the dominant eigenvalue of the incidence matrix  $\mathbf{M}_\varphi$  of  $\varphi$  is a quadratic unit and the parameters  $c, l$  belong to  $\mathbb{Q}(\varepsilon)$ .*

*Proof.* We know already that the Perron-Frobenius eigenvalue  $\Lambda$  of the matrix  $\mathbf{M}_\varphi$  is a quadratic number. For contradiction, assume that  $\Lambda$  is not a unit. Since  $\mathbf{M}_\varphi \vec{t} = \Lambda' \vec{t}$ , we have  $\Lambda' \mathbb{Z}[\varepsilon] \subseteq \mathbb{Z}[\varepsilon]$ . If  $\Lambda$  is not a unit, then  $\Lambda' \mathbb{Z}[\varepsilon]$  is a proper subset of  $\mathbb{Z}[\varepsilon]$  and the quotient abelian group  $\mathbb{Z}[\varepsilon]/\Lambda' \mathbb{Z}[\varepsilon]$  has at least two classes of equivalence. For the purposes of this proof we shall denote by  $\triangleleft J$  the left end-point of a given interval  $J$ .

Realize that  $E_X \subset \mathbb{Z}[\varepsilon]$ ,  $E_{\varphi(X)} = \Lambda' E_X \subset \Lambda' \mathbb{Z}[\varepsilon]$  and  $E_X \not\subset \Lambda' \mathbb{Z}[\varepsilon]$  for all  $X \in \mathcal{A}$ . Facts (2)–(4) above imply that the left boundary point of the interval  $I_A$ , i.e., the point  $\triangleleft I_A = c$  must coincide with  $\triangleleft(E_{\varphi(X_1)} + t_{w_1})$ , and  $\triangleleft(E_{\varphi(X_2)} + t_{w_2})$ , for some letters  $X_1, X_2 \in \mathcal{A}$  and some prefixes  $w_1, w_2$  of  $\varphi(X_1), \varphi(X_2)$ , respectively. The above property (1) and equation (11) imply

$$\triangleleft(E_{\varphi(X_i)} + t_{w_i}) = T^{|w_i|}(\triangleleft(\Lambda' E_{X_i})).$$

Since  $T^n(x) \neq x$  for all  $n \neq 0$  and all  $x \in I$ , we necessarily have  $X_1 \neq X_2$ . Same reasons imply for the left boundary point of the interval  $I_B$ , that there exist at least two distinct letters  $Y_1 \neq Y_2$ , such that  $\triangleleft(E_{\varphi(Y_i)} + t_{v_i})$  coincide with  $\triangleleft I_B = c + l - (1 - \varepsilon)$  for some proper prefixes  $v_i$  of  $\varphi(Y_i)$ .

Since the distance  $l - 1 + \varepsilon$  between  $\triangleleft I_A$  and  $\triangleleft I_B$  is not an element of  $\mathbb{Z}[\varepsilon]$ , we must have  $Y_i \neq X_j$  for  $i, j = 1, 2$ . This contradicts the fact that the alphabet has only 3 letters. Therefore  $\Lambda'$  is a unit.

The fact that  $\triangleleft I_A = c$ ,  $\triangleleft I_B = c + l - 1 + \varepsilon$ ,  $\triangleleft I_C = c + \varepsilon$  coincide with iterations of points  $\Lambda'c$ ,  $\Lambda'(c + l - 1 + \varepsilon)$ ,  $\Lambda'(c + \varepsilon)$  implies that  $c, l \in \mathbb{Q}(\varepsilon)$ .  $\square$

#### ACKNOWLEDGEMENTS

The authors acknowledge financial support by Czech Science Foundation GA ĀR 201/05/0169, by the grant LC06002 of the Ministry of Education, Youth, and Sports of the Czech Republic, and by the ACINIM NUMERATION.

#### REFERENCES

- [1] B. Adamczewski, *Codages de rotations et phénomènes d'autosimilarité*, J. Théor. Nombres Bordeaux **14** (2002), 351–386.
- [2] B. Adamczewski, *Balances for fixed points of primitive substitutions*, Words. Theoret. Comput. Sci. **307** (2003), 47–75.
- [3] S. Akiyama, M. Shirasaka, *Recursively renewable words and coding of irrational rotations*, Journ. of Math. Soc. Japan.
- [4] C. Allauzen, *Une caractérisation simple des nombres de Sturm*, J. Théor. Nombres Bordeaux **10** (1998), 237–241.
- [5] P. Ambrož, Z. Masáková, E. Pelantová, *Matrices of 3iet preserving morphisms*, submitted to Theor. Comp. Sci. (2007), 26pp.
- [6] P. Baláži, Z. Masáková, E. Pelantová, *Complete characterization of substitution invariant 3iet words*, preprint 2007.
- [7] P. Baláži, Z. Masáková, E. Pelantová, *Characterization of substitution invariant Sturmian sequences*, Integers **5** (2005), A14, 23 pp. (electronic)
- [8] V. Berthé, H. Ei, S. Ito, H. Rao, *Invertible substitutions and Sturmian words: an application of Rauzy fractals*, Theoret. Informatics Appl. **41** (2007), 329–349.
- [9] V. Berthé, S. Ferenczi, L.Q. Zamboni, *Interactions between dynamics, arithmetics, and combinatorics: the good, the bad, and the ugly*, dans *Algebraic and Topological Dynamics*, édité par S. Kolyada, Y. Manin, and T. Ward, Contemporary Mathematics (CONM) **385**, American Mathematical Society, pp. 333–364, 2005.
- [10] M. D. Boshernitzan, C. R. Carroll, *An extension of Lagrange's theorem to interval exchange transformations over quadratic fields*, J. Anal. Math. **72** (1997), 21–44.
- [11] D. Boyd, *Small Salem numbers*, Duke Math. J. **44** (1977), 315–328.
- [12] D. Crisp, W. Moran, A. Pollington, P. Shiue, *Substitution invariant cutting sequences*, J. Thor. Nombres Bordeaux **5** (1993), 123–137.
- [13] S. Ferenczi, C. Holton, L. Zamboni, *Structure of three-interval exchange transformations II. A combinatorial description of the trajectories*, J. Anal. Math. **89** (2003), 239–276.
- [14] L.S. Guimond, Z. Masáková, E. Pelantová, *Combinatorial properties of infinite words associated with cut-and-project sequences*, J. Théor. Nombres Bordeaux **15** (2003), 697–725.
- [15] G. A. Hedlund, M. Morse. Symbolic dynamics II. Sturmian trajectories. *Amer. J. Math.* **62** (1940), 1–42.
- [16] A. B. Katok, A. M. Stepin *Approximations in ergodic theory*, Usp. Math. Nauk. **22** (1967), 81–106 (in Russian), translated in Russian Math. Surveys **22** (1967), 76–102.
- [17] M. Keane, *Interval exchange transformations*, Math. Z. **141** (1975), 25–31.
- [18] G. Poggiaspalla, J. H. Lowenstein, F. Vivaldi, *Geometric representation of interval exchange maps over algebraic number fields*, e-print arXiv:0705.1073 (2007).
- [19] M. Queffélec, *Substitution dynamical systems. Spectral analysis*, Lect. Notes in Math. **1294**, Springer-Verlag (1987).
- [20] S. Yasutomi, *On Sturmian sequences which are invariant under some substitutions*, Number theory and its applications (Kyoto, 1997), Dev. Math. **2**, 347–373, Kluwer Acad. Publ., Dordrecht, 1999.

(Arnoux) INSTITUT DE MATHÉMATIQUES DE LUMINY, CNRS UPR 9016, 163, AVENUE DE LUMINY, CASE 907, 13288 MARSEILLE CEDEX 09 FRANCE

*E-mail address, Arnoux:* `arnoux@iml.univ-mrs.fr`

(Berthé) LABORATOIRE D'INFORMATIQUE, DE ROBOTIQUE ET DE MICROELECTRONIQUE DE MONTPELLIER 161, RUE ADA, 34 392 MONTPELLIER CEDEX 5, FRANCE

*E-mail address, Berthé:* `berthe@lirmm.fr`

(Masáková) DOPPLER INSTITUTE & DEPARTMENT OF MATHEMATICS, FNSPE CZECH TECHNICAL UNIVERSITY, TROJANOVA 13, 120 00 PRAHA 2, CZECH REPUBLIC

*E-mail address, Masáková:* `masakova@km1.fjfi.cvut.cz`

(Pelantová) DOPPLER INSTITUTE & DEPARTMENT OF MATHEMATICS, FNSPE CZECH TECHNICAL UNIVERSITY, TROJANOVA 13, 120 00 PRAHA 2, CZECH REPUBLIC

*E-mail address, Pelantová:* `pelantova@km1.fjfi.cvut.cz`