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 with 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. [13] showed that a homogeneous sturmian word (i.e., a sturmian word with slope ε and intercept $x_0 = 0$) is invariant under a non-trivial substitution if and only if ε is a Sturm number. In 1998, Allauzen [4] provided a simple characterization of Sturm numbers:

A quadratic irrational number ε with algebraic conjugate ε' is called a Sturm number if

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

For general sturmian words (with arbitrary intercept x_0), the fact that ε is a Sturm number is only a necessary but not a sufficient condition for invariance under a substitution; this is clear since there can only be a countable number of such invariant words, while sturmian words with a given slope are determined by their intercept, hence they are uncountable in number. For a complete characterization, see [21, 7, 9].

In this paper, we study invariance under substitution of infinite words coding non-degenerate exchanges of three intervals with permutation (321).[1] 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 [14]. It is well known that substitutive 3iet words, i.e., 3iet words that are image by a morphism of a fixed point of a substitution correspond to quadratic parameters, see e.g. [1, 3, 11, 14, 19]. Note that, in the present paper, we consider fixed points of substitutions and not substitutive words.

Sturmian words can be equivalently defined as aperiodic words coding a rotation, i.e., an exchange of two intervals with lengths say α, β. The slope of the sturmian word, which we have denoted by ε, is then equal to $\varepsilon = \frac{\alpha}{\alpha+\beta}$. The term 'slope' for parameter ε comes from the fact that the sturmian word with slope ε can be constructed by projection of points of the lattice $\mathbb{Z}^2$ to the straight line $y = \varepsilon x$. It is proved convenient to abuse the language by speaking of the *slope of the rotation*, which 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 being exchanged

Date: March 30, 2008.

[1] An exchange of intervals is non-degenerate if it satisfies i.d.o.c. [18]. For more details, see Section 4.3

but on their ratio, then the lengths are often normalized to satisfy $\alpha + \beta = 1$. In this case α and ε coincide.

The same situation occurs for 3iet words which code exchange of three intervals with lengths, say, α, β, γ. The commonly used normalization of parameters is $\alpha + \beta + \gamma = 1$. However, the normalization $\alpha + 2\beta + \gamma = 1$ seems to be more suitable. Let us mention three arguments in favor, which are results from papers [1, 6, 14, 15]. If u is an infinite word coding exchange of three intervals of lengths α, β, γ, then:

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

In Section 3 and 4, we give 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 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 $\sigma : \{A, B, C\}^* \to \{0, 1\}^*$ and $\sigma' : \{A, B, C\}^* \to \{0, 1\}^*$ be the morphisms defined by*

$$\sigma(A) = 0, \quad \sigma(B) = 01, \quad \sigma(C) = 1$$

$$\sigma'(A) = 0, \quad \sigma'(B) = 10, \quad \sigma'(C) = 1.$$

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*

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

Note that, in this case, the corresponding homogeneous sturmian word is also substitution invariant. A forthcoming paper [8] 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. Preliminary considerations

We work with finite and infinite words over a finite alphabet $\mathcal{A} = \{a_1, \ldots, a_k\}$. The set of all finite words over $\mathcal{A}$ is denoted by $\mathcal{A}^*$. It is a free monoid equipped with the binary operation of concatenation and the empty word. 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 $\mathcal{C}_u : \mathbb{N} \to \mathbb{N}$ 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

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

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

Let $\mathcal{A}$ and $\mathcal{B}$ be two alphabets. A mapping $\varphi : \mathcal{A}^* \to \mathcal{B}^*$ is said to be a *morphism* if $\varphi(w\widehat{w}) = \varphi(w)\varphi(\widehat{w})$ holds for any pair of finite words $w, \widehat{w} \in \mathcal{A}^*$. 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 φ is called a *substitution*.

The action of a morphism φ can be naturally extended to infinite words by the prescription

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

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

The incidence matrix of a morphism φ 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

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

A morphism φ is called primitive if there exists an integer k such that the matrix $\boldsymbol{M}_\varphi^k$ 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 φ. It is known [20] that in such a case the densities of letters in u are well defined. The vector

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

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

3. Exchanges of three intervals as an induction of rotations

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

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

and let $t_A = \beta + \gamma, t_B = \gamma - \alpha, t_C = -\alpha - \beta \in \mathbb{R}$ be translation 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 a long time ago (see [17]), 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 $R^2(I_B) = R(I_D) = T(I_B)$. □

Hence T can be obtained as the induction of a rotation on a left interval (for more details, see e.g. [1, 14] or the survey [10]). It can also be obtained as an 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 normalizing 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 can thus always be chosen as the origin. For this, we introduce the new parameters

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

The number ε 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

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

We redefine five intervals in this setting

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

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

$$J_0 = I_A \cup I_B,\ J_1 = I_C \cup I_D \text{ and } J = J_0 \cup J_1$$
$$J_0' = I_E \cup I_A,\ J_1' = I_B \cup I_C \text{ and } J' = J_0' \cup J_1'.$$

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 ε, 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*

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

is called the 3iet word with parameters ε, l, c.

There is a simple classical 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 orbit 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, i.e., 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 σ (resp. σ') 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 ε is irrational, then $\sigma(u)$ and $\sigma(u')$ are sturmian sequences whose density of 0 *equals ε.*

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 ε, 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

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

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 $i = n_k$, and then 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 the 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 both belong to $(n_k)_{k\in\mathbb{N}}$, or else if none of them belong 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. Since the latter is sturmian, these densities are irrational, hence v' is an aperiodic balanced sequence and is thus a sturmian sequence according to [16]. □

We are now in position the 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. Assume $v = \sigma(u), v' = \sigma'(u)$ are two sturmian words; by construction, they have the same slope ε, and 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

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

with the choice of 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. Since the image of u by the substitution σ is aperiodic by hypothesis, the sequence u itself is aperiodic. □

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 $(0, 1)$. Note that this stepped line is contained in a "corridor" of width less than 1; the two sturmian lines associated with v and v', obtained by enlarging the corridor on the right or the left to the width of the unit square, are shown in dashed lines.

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.

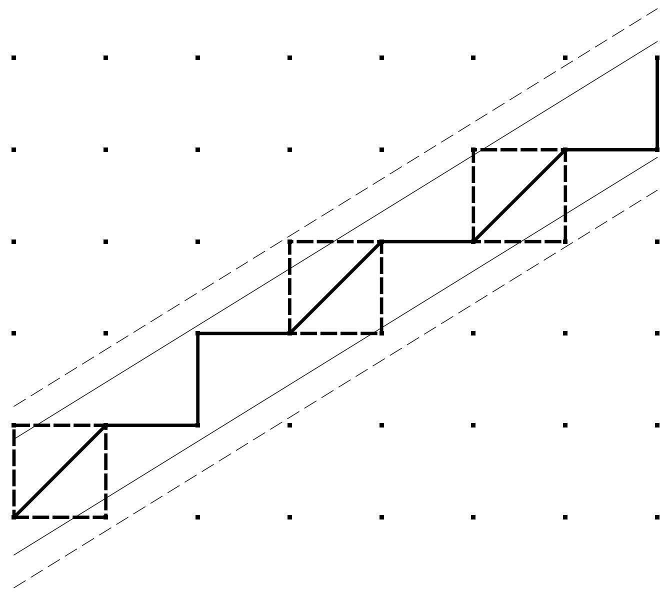

FIGURE 1. The stepped line associated with a 3-iet word and its two sturmian extensions

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.

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 was introduced by Keane [18] 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

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

see [1, 15].

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

If ε is rational, the sequence u is periodic, hence the densities exist in a trivial way.

5. Translation vectors

Let u be a 3iet word 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 furthermore assume that u is fixed by some substitution φ. We will now obtain a more subtle equation using the substitution φ. Let us define a function g (in flavour of the map f defined in Section 4.2) on prefixes of the infinite word u, the fixed point of φ. 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

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

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

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

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

$$E_{\varphi(X)} = \Big\{ (|w|_A, |w|_B, |w|_C)\boldsymbol{M}_\varphi \vec{t} \;\Big|\; wX \text{ is a prefix of } u \Big\}\,. \tag{5}$$

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

- $u_ku_{k+1}u_{k+2}\cdots$ has the prefix $\varphi(A)$;
- $u_0u_1\cdots u_{k-1} = \varphi(u_0u_1\cdots u_{i-1})$ for some $i \in \mathbb{N}$;
- $u_i = A$.

A similar statement is true for 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

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

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

$$\vec{t_\varphi} = \boldsymbol{M}_\varphi \vec{t}. \tag{6}$$

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*

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

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 [16]. The balance properties of the considered generalization of sturmian words, i.e., 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 the 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 φ with incidence matrix $\boldsymbol{M}_\varphi$. Let Λ be the dominant eigenvalue of $\boldsymbol{M}_\varphi$. If $|\lambda| < 1$ for all other eigenvalues λ of $\boldsymbol{M}_\varphi$, 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 ε, l, c satisfying* (2) *and* (4). *Let φ be a primitive substitution such that $\varphi(u) = u$. Then parameter ε is a Sturm number.*

Proof. The density vector of the word u is the vector $\vec{\varrho}_u = \left(1 - \frac{1-\varepsilon}{l}, \frac{1}{l} - 1, 1 - \frac{\varepsilon}{l}\right)$. The vector $\vec{\varrho}_u$ is a left eigenvector corresponding to the Perron-Frobenius eigenvalue Λ. Since $\vec{\varrho}_u$ is not a multiple of any rational vector by Remark 4.5 and $\boldsymbol{M}_\varphi$ is an integral matrix, Λ is either a cubic or a quadratic irrational number. If Λ is cubic, then the eigenvalues of $\boldsymbol{M}_\varphi$ are Λ and its algebraic conjugates. If Λ is quadratic, then the eigenvalues are Λ, its algebraic conjugate, and a rational number. In both cases, $\boldsymbol{M}_\varphi$ has 3 different eigenvalues. Denote the other eigenvalues of $\boldsymbol{M}_\varphi$ by λ_1, λ_2 and by $\vec{x}_1$, $\vec{x}_2$ the right eigenvectors of the matrix $\boldsymbol{M}_\varphi$ corresponding to λ_1 and λ_2, respectively, i.e.,

$$\boldsymbol{M}_\varphi \vec{x}_1 = \lambda_1 \vec{x}_1 \qquad \text{and} \qquad \boldsymbol{M}_\varphi \vec{x}_2 = \lambda_2 \vec{x}_2 \,. \tag{7}$$

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 Λ. Since the vector $\vec{t} = (1-\varepsilon, 1-2\varepsilon, -\varepsilon)^T$ is orthogonal to $\vec{\varrho}_u$, we can write

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

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 $\boldsymbol{M}_\varphi$.

Recall that the translation vector $\vec{t}_\varphi$ satisfies (6). Since this holds for any substitution which has $u_0u_1u_2\cdots$ for its fixed point, one can write

$$\vec{t}_{\varphi^n} = \boldsymbol{M}_{\varphi^n}\ \vec{t} = \boldsymbol{M}_\varphi^n\ \vec{t}. \tag{9}$$

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

$$\boldsymbol{M}_\varphi^n\,\vec{t} \;=\; \mu\lambda_1^n\vec{x}_1 \;+\; \nu\lambda_2^n\vec{x}_2 \tag{10}$$

is bounded.

We shall now distinguish two cases. Recall that the Perron eigenvalue Λ must be an algebraic integer of degree three or two.

The cubic case. Suppose that Λ is a cubic number. Then λ_1, λ_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[2] of degree 3 do not exist, we derive that one of the eigenvalues λ_1, λ_2 is in modulus greater than 1, say $|\lambda_2| > 1$. The boundedness of the sequence of vectors $(\boldsymbol{M}_\varphi^n\,\vec{t})_{n\in\mathbb{N}}$ in (10) implies that $\nu = 0$ and thus $\vec{t}$ is a right eigenvector of the matrix $\boldsymbol{M}_\varphi$, so without loss of generality, let $\vec{x}_1 = \vec{t}$.

Then the components of the vector $\vec{x}_1 = \vec{t} = (1-\varepsilon, 1-2\varepsilon, -\varepsilon)^T$ belong to the field $\mathbb{Q}(\lambda_1)$, but the first plus the last components of the vector $\vec{x}_1$ are equal to the middle one. If a matrix $\boldsymbol{M}$ has such an eigenvector corresponding to the eigenvalue λ, then the vector arising as the sum of the first and second columns of the matrix $\boldsymbol{M} - \lambda\boldsymbol{I}$ and the vector arising as the sum of the second and the third columns of the matrix $\boldsymbol{M} - \lambda\boldsymbol{I}$ are colinear. This implies that λ is a root of a monic polynomial of degree 2. This is a contradiction with cubicity of λ_1.

The quadratic case. We have shown that Λ is a quadratic number. In this case, the other eigenvalues of $\boldsymbol{M}_\varphi$ are the conjugate $\lambda_1 = \Lambda'$ of Λ and $\lambda_2 = r \in \mathbb{Z}$. The irrationality of the vector $\vec{t}$ implies that $\mu \neq 0$. Let us also suppose that $\nu \neq 0$. The boundedness of $\boldsymbol{M}_\varphi^n\,\vec{t}$ 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 φ^2 instead of φ. 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

$$\vec{t}_{\varphi^n} = \mu(\Lambda')^n\vec{x}_1 + \nu\vec{x}_2 \qquad \xrightarrow[n\to\infty]{} \qquad \nu\vec{x}_2 \neq \vec{0}\,.$$

We shall make use of the following property of infinite words coding 3iet. For arbitrary factor $w \in \mathcal{L}(u)$, let I_w denote the closure of the set $\{T^n(0) \mid w \text{ is a prefix of } u_nu_{n+1}u_{n+2}\cdots\}$. It is known that I_w is an interval. With growing length of w, the length $|I_w|$ of the interval I_w approaches 0. Since the morphism φ 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 $\lim_{n\to\infty}|I_{\varphi^n(X)}| = 0$.

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

[2]An algebraic integer is called a Salem number, if it is a real number > 1, and all its algebraic conjugates are in modulus ≤ 1 with at least one of them laying on the unit circle. It is known [12] that all Salem numbers are of even degree greater than or equal to 4.

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

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

The proof of Theorem B provides several direct consequences.

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

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

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

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

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 φ, $u_0u_1u_2\cdots = \varphi(u_0)\varphi(u_1)\varphi(u_2)\cdots$, 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) $\displaystyle\bigcup_{x\in\mathcal{A}} \ \bigcup_{\substack{w \text{ is a proper} \\ \text{prefix of } \varphi(X)}} \big(E_{\varphi(X)} + t_w\big) \quad = \quad E_A \cup E_B \cup E_C.$

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

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

Note 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 interval I_A, i.e., the point $\lhd I_A = c$ must coincide with $\triangleleft\big(E_{\varphi(X_1)}+t_{w_1}\big)$, and $\triangleleft\big(E_{\varphi(X_2)}+t_{w_2}\big)$, 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

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

Since $T^n(x) \neq x$ for all $n \neq 0$ and all $x \in I$, we necessarily have $X_1 \neq X_2$. The 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 $\triangleleft\big(E_{\varphi(Y_i)} + t_{v_i}\big)$ 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 Λ' is a unit.

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

8. Conclusions

(1) One might ask wether there are other pairs of morphisms than (σ, σ') that could be used in theorem A. This is indeed the case, and we could in fact compose with any sturmian morphism. Indeed, rotations can be "exduced" in many ways to larger rotations, and this amounts to applying an arbitrary sturmian substitution. In contrast, they can be induced essentially only in one way, which is driven by the continued fraction expansion of the slope.

However, the substitutions σ and σ' are the simplest ones that can be used. A geometric way to understand this is the following: with any interval exchange, we can associate a broken line, as explained in Figure 1. This broken line is given as a cut-and-project scheme, its vertices are exactly the rational points in a slice; the smallest slice containing it which gives a sturmian line is the slice obtained by sliding the unit square, as shown in Figure 1, and this leads to the two substitutions σ and σ'. It is of course possible to take a larger slice, by a change of basis in $\mathbb{Z}^2$, and we obtain in this way other, more complicated, substitutions.

(2) In this paper we have focused on non-degenerate 3iet words. In fact, for the proof of the result stated in Theorem B we need non-vanishing determinant of the substitution matrix, which is ensured only for substitutions fixing non-degenerate 3iet words. For, substitution with degenerate 3iet words as a fixed point may have both singular and non-singular incidence matrix. As an example, consider the transformation $T : [0,2) \mapsto [0,2)$ exchanging intervals of lengths $\alpha = \tau^2$, $\beta = \tau$ and $\gamma = 1$, where $\tau = \frac{1}{2}(\sqrt{5}-1)$. Note that $\tau^2 + \tau = 1$.

If we consider the pointed biinfinite word $u^{(1)}$ coding the orbit of the point $x_0 = 0$, we obtain, by the Rauzy induction on the interval $J = [0, \tau^2)$, the substitution $A \mapsto AC, B \mapsto ACBC, C \mapsto BC$ with the incidence matrix $\left(\begin{smallmatrix} 1 & 0 & 1 \\ 1 & 1 & 2 \\ 0 & 1 & 1 \end{smallmatrix}\right)$, which is singular.

On the other hand, if $x_0 = 1$, we find by induction on the interval $J = [1-\tau^2, 1+\tau^2)$ that the corresponding pointed word $u^{(2)}$ is invariant under the substitution $A \mapsto B, B \mapsto BCB, C \mapsto CAC$ with the incidence matrix $\left(\begin{smallmatrix} 0 & 1 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{smallmatrix}\right)$ of unit determinant.

The pointed biinfinite words $u^{(1)}$ and $u^{(2)}$ differ just by the position of the delimiter, namely

$$u^{(1)} : \qquad \cdots ACBCBC|ACBCACBCBC\cdots$$
$$u^{(2)} : \qquad \cdots ACBCB|CACBCACBCBC\cdots$$

In the same time, it is obvious, that the words are degenerate, since the letter C occurs exactly on the odd (resp. even) positions, and by erasing it, we obtain a sturmian word, namely the Fibonacci word.

(3) Our paper does not study the properties of morphisms with 3iet fixed points. Rather, it describes properties of the corresponding incidence matrices, as summarized in Corollaries 7.1 and 7.2. We can deduce that the eigenvalues of a substitution matrix M_φ are a quadratic unit λ, its algebraic conjugate $\lambda' = \pm\lambda^{-1}$ and $m \in \mathbb{Z}$. The arguments given in this paper do not allow us to derive the value of m. However, using the fact (see [5]) that a substitution fixing a non-degenerate 3iet word is 3iet preserving, we can use the results of [6], to derive that $m = \pm 1$. Note that we call a morphism 3iet preserving, if it preserves the set of 3iet words.

Acknowledgements

We would like to thank the anonymous referee for useful comments and suggestions. The authors are also grateful for financial support from the Czech Science Foundation GA ČR 201/05/0169, through the grant LC06002 of the Ministry of Education, Youth, and Sports of the Czech Republic, and via 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*, J. Math. Soc. Japan **59** (2007), 1199–1234.

[4] C. Allauzen, *Une caractérisation simple des nombres de Sturm*, J. Théor. Nombres Bordeaux **10** (1998), 237–241.

[5] P. Ambrož, E. Pelantová, *3iet preserving morphisms and their fixed points*, in preparation 2008.

[6] P. Ambrož, Z. Masáková, E. Pelantová, *Matrices of 3iet preserving morphisms*, to appear in Theor. Comp. Sci. (2008), 26pp.

[7] P. Baláži, Z. Masáková, E. Pelantová, *Complete characterization of substitution invariant 3iet words*, preprint 2007.

[8] P. Baláži, Z. Masáková, E. Pelantová, *Characterization of substitution invariant Sturmian sequences*, Integers **5** (2005), A14, 23 pp. (electronic)

[9] V. Berthé, H. Ei, S. Ito, H. Rao, *Invertible susbtitutions and Sturmian words: an application of Rauzy fractals*, Theoret. Informatics Appl. **41** (2007), 329–349.

[10] 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.

[11] 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.

[12] D. Boyd, *Small Salem numbers*, Duke Math. J. **44** (1977), 315–328.

[13] D. Crisp, W. Moran, A. Pollington, P. Shiue, *Substitution invariant cutting sequences*, J. Thor. Nombres Bordeaux **5** (1993), 123–137.

[14] 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.

[15] 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.

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

[17] 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.

[18] M. Keane, *Interval exchange transformations*, Math. Z. **141** (1975), 25–31.

[19] G. Poggiaspalla, J. H. Lowenstein, F. Vivaldi, *Geometric representation of interval exchange maps over algebraic number fields*, e-print arXiv:0705.1073 (2007).

[20] M.Queffélec, *Substitution dynamical systems. Spectral analysis*, Lect. Notes in Math. **1294**, Springer-Verlag (1987).

[21] 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 Microlectronique 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`