Substitutions and Interval Exchange Transformations of Rotation Class

Luis-Miguel Lopez[†] **and Philippe Narbel**[‡]

(†) IGM, Université Marne-La-Vallée. 2, rue de la Butte Verte, 93166 Noisy-le-Grand, France[1]

(‡) LABRI, Université Bordeaux I. 351, Cours de la Libération 33405 Talence, France

e-mails: lopez@univ-mlv.fr, narbel@labri.u-bordeaux.fr

August 1998 - (published in TCS, 255, (2001), pp.323-344)

Abstract. We propose a method for obtaining the symbolic orbits of interval exchange transformations of rotation class over n intervals by composing a finite set of basic substitutions, i.e. by doing simple parallel rewriting. Based on surface theory, this method is shown to be closely related to Rauzy induction. Sturmian objects are known to correspond to interval exchange transformations over 2 intervals. In this respect, our n intervals case is shown to be also related to continued fractions and the obtained words have also linear complexity.

Keywords: *Substitutions, symbolic dynamics, interval exchanges, continued fractions, Sturmian words, Dehn twists.*

Contents

[1]This paper writing has been ended while the first author was invited at the Tokyo Institute of Technology.

1 Introduction

Substitutions are simple parallel rewriting processes on words which are usual endomorphisms when applied to finite words. A substitution θ is defined by the image words of each single involved letter. For instance, over the alphabet $\{a, b\}$, let $\theta(a) = ab$ and $\theta(b) = a$, then $\theta(aba) = \theta(a)\theta(b)\theta(a) = abaab$. Studying iterations of such processes has been extensively done in formal language theory, mostly under the name of D0L-systems (see for instance [RS80, RS97]), as well as in symbolic dynamics [GH55, Qué87, LM95].

Interval exchange transformations are simple piecewise isometric maps acting on an interval of the real line, say $[0, 1)$, which effect is to permute a finite number of semi-open subintervals which makes a partition of it. For instance, let $a \in (0, 1)$, and $T(x) = x + (1 - a)$ if $x \in [0, a)$, $T(x) = x - a$ if $x \in [a, 1)$: this defines an interval exchange transformation over 2 intervals. Such a transformation is fully characterized by the length of the involved subintervals and the permutation which shuffles them. When iterated, interval exchange transformations lead to generic examples of dynamical systems (see for instance [Kea75, Vee78, Vee82, Mas82, Mañ83]).

Iterations of an interval exchange transformation can easily be transformed into symbolic information following a traditional operation in dynamical system theory (see for instance [Kea75]): one assigns a different letter to each of the intervals so that the iterates of the interval exchange transformation, the *orbits*, are translated into words.

The relationship between substitutions and interval exchange transformations over 2 intervals, hence over alphabets of 2 letters has been already extensively deciphered, mostly under the name of Sturmian words and Sturmian substitutions (see [MH40, CH73, Ber95, BS97]).

The main result we prove here is about a set of interval exchange transformations over n intervals. This set consists of the interval exchange of rotation class (see [Vee82, NR97]) which are those which have at most two discontinuities, and for which unique ergodicity is ensured. Of course these include all the interval exchanges over 2 and 3 intervals, but also many over n-intervals where $n \geq 4$. The theorem we prove here is the following:

Theorem: *Consider an irreducible and irrational interval exchange transformation of rotation class over n intervals, $n > 2$ (respect. $n = 2$). The set of its symbolic orbits can be constructively described from compositions of an explicit set of $n + 1$ substitutions (respect. 2 substitutions) over an alphabet of n letters.*

The proof is based on the so-called *Rauzy induction* (see mainly [Rau79, Vee82, Ker85, Zor96, NR97]), which is here geometrically interpreted on surfaces with a structure of non-crossing parallel curves, i.e. singular foliations. It comes out that the considered foliations can be summed up by interval exchange transformations, defined by *first return maps* describing how curves hit some finite transverse segments (see e.g. [Mañ83] p.119). The shorter the segments, the longer the pieces of leaves between successive hits of the transverse segments. In this respect, Rauzy induction can be seen as a mecanism for systematically shrinking a set of transverse segments, and therefore lengthening the pieces of leaves between hits (see e.g. [Zor96]). Using the symbolic translation of the orbits of an interval exchange transformation, the idea of the proof is that this lengthening can be interpreted as the application of substitutions.

Since Sturmian words are the symbolic translation of the case of interval exchange transformations over 2 intervals, we show next, first, how some of the properties of these words are recovered from our framework, and second how these properties hold in the general case over n intervals:

- The complexity of an infinite word w is defined as a map $\mathbb{N}^* \to \mathbb{N}^*$ which gives for each $m \in \mathbb{N}^*$ the number of subwords of length m that occur in w (see e.g. [All94, Lot97]). This is an usual measure for the structure of infinite words. Sturmian words are known to have linear complexity $m+1$ [MH40, CH73, BS97]. Symbolic orbits of interval exchange transformations over n intervals can be proved to have linear complexity $(n-1)m+1$, and therefore:

  Proposition: *All the infinite words generated by the the above theorem have complexity* $(n-1)m+1$.

- Sturmian words are known to be closely related to the classical continued fraction algorithm (see e.g. [Mar82, IY90, Ber95, BS97]). The basic process to generate composition of substitutions in the theorem, i.e. Rauzy induction, can be seen to be this classical algorithm in the 2 intervals case, and gives a multidimensional continued fraction algorithm [Rau77, Ker85, Zor96, NR97] in the general case. We discuss the convergence property of this generalization:

  Proposition: *For every interval exchange transformation of rotation class over n intervals,* $n \geq 2$*, Rauzy induction leads to a weakly convergent multidimensional continued fraction algorithm.*

  We also show how this multidimensional continued fraction algorithm fits into the general framework of Szekeres [Sze70, Bre81].

- Sturmian substitutions are those which leave stable Sturmian words. They can be generated by sets of essentially two substitutions. We show how our framework allows one to recover one of the results about them:

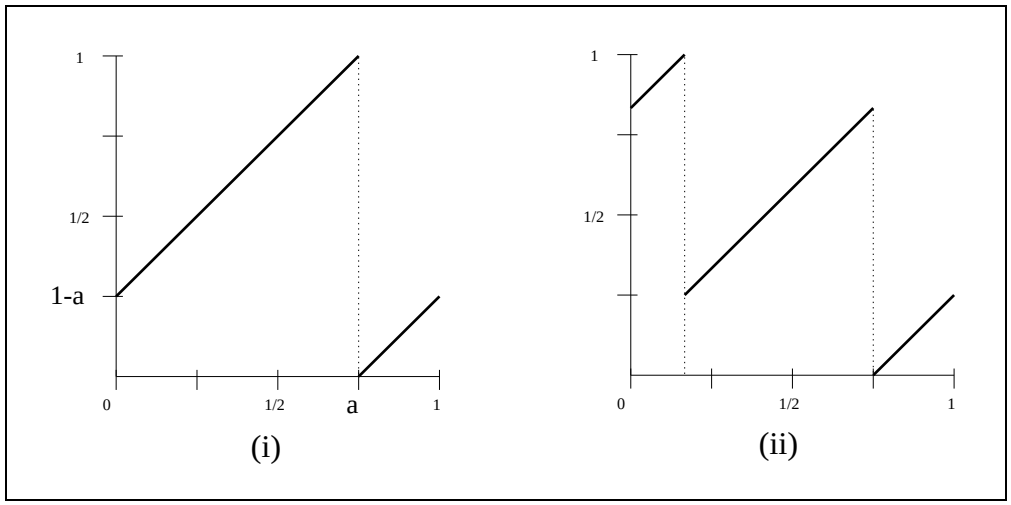
  Proposition: *Consider an interval exchange transformation over 2 intervals. Then the two substitutions given by the above theorem are the Sturmian basic ones used in [IY90] to relate Sturmian words and classical continued fractions.*

2 Interval Exchange Transformations

Let $\lambda = (\lambda_1, \ldots, \lambda_n)$ be an n-dimensional positive vector ($n \geq 2$) such that $\sum_j \lambda_j = 1$, called a **length vector**, and let π be a permutation of $\{1, \cdots, n\}$. An **interval exchange transformation** (see e.g. [Vee82, Mas82, Mañ83]) is a function $T_{\lambda,\pi} : [0,1) \to [0,1)$ whose domain is decomposed according to $b_0 = 0$ and $b_i = \sum_{j=1}^{i} \lambda_j$, for $i = 1, \ldots, n$, i.e. as $\bigsqcup_{i=1}^{n} I_i$ where $I_i = [b_{i-1}, b_i)$, and whose range is decomposed according to the length vector $(\lambda_{\pi^{-1}(1)}, \ldots, \lambda_{\pi^{-1}(n)})$ with $b_0^\pi = 0$ and $b_i^\pi = \sum_{j=1}^{i} \lambda_{\pi^{-1}(j)}$, i.e. as $\bigsqcup_{i=1}^{n} J_i$ where $J_i = [b_{i-1}^\pi, b_i^\pi)$. The complete expression of $T_{\lambda,\pi}$ is then given as $T_{\lambda,\pi}(x) = x - b_{i-1} + b_{\pi(i)-1}^\pi$ for all $x \in I_i$, $i = 1, \ldots, n$. One can see that the intervals J_i's are just the intervals I_i's glued together according to the permutation π, where the interval I_i which is in the i-th place, is sent to the $\pi(i)$-th place .

An interval exchange transformation is said to be **irreducible** when its permutation π does not fix (setwise) any strict subset $\{1, ..., k\} \subset \{1, ..., n\}$. The simplest non-trivial example of an irreducible $T_{\lambda,\pi}$ is given by a decomposition of $[0,1)$ into two intervals (see Figure 2.1(i)): $\lambda = (a, 1-a)$, where $0 < a < 1$, and $\pi(1) = 2$, $\pi(2) = 1$. Hence, $b_0 = 0$, $b_1 = a$, $b_2 = 1$, $I_1 = [0, a)$, $I_2 = [a, 1)$, and $T_{\lambda,\pi}(x) = x + (1-a)$ on I_1, $T_{\lambda,\pi}(x) = x - a$ on I_2. Another interval exchange over 3 intervals is shown in Figure 2.1(ii) where $\pi(1) = 3$, $\pi(2) = 2$, $\pi(3) = 1$, and $\lambda = (1/6, 7/12, 1/4)$. An interval exchange transformation is said to be of **rotation class** iff it has either one or two discontinuities [NR97]. For instance, over 4 intervals, if π is defined as $\pi(1) = 4$, $\pi(2) = 3$, $\pi(3) = 1$ $\pi(4) = 2$, then $T_{\lambda,\pi}(x)$ has only two discontinuities.

Fig. 2.1



Let the **positive orbit** (respect. **orbit**) of a point $x \in [0,1)$ be $O^+(x) = \{T^i_{\lambda,\pi}(x),\ i \in \mathbb{N}\}$ (respect. $O(x) = \{T^i_{\lambda,\pi}(x),\ i \in \mathbb{Z}\}$), and let $\mathcal{I}$ be $[0,1) \setminus \bigcup_{i=1}^{n-1} O(b_i)$. Then the pair $(\mathcal{I}, T_{\lambda,\pi})$ is a **dynamical system**, i.e. a pair (X,T) such that X is a metric space and $T : X \to X$ is continuous. Such a system is said to be **minimal** iff for $Y \subset X$, $Closure(T(Y)) = Y$ implies $Y = X$ or $Y = \emptyset$. The system $(\mathcal{I}, T_{\lambda,\pi})$ is minimal iff for each $x \in [0,1)$, the orbit $O^(x)$ is dense in $[0,1)$ [Kea75]. Moreover, if $T_{\lambda,\pi}$ is irreducible and **irrational**, i.e. the only rational relations between the λ_i's are multiples of $\lambda_1 + ... + \lambda_n = 1$, then $T_{\lambda,\pi}$ is minimal [Kea75].

A **topological conjugacy** between two dynamical systems (X_1, T_1) and (X_2, T_2) is a homeomorphism $\phi : X_1 \to X_2$ such that $\phi T_1 = T_2 \phi$, which means that the two systems are essentially equivalent. It is known that, up to a compactification of $\mathcal{I}$ (see e.g. [Kea75]), $(\mathcal{I}, T_{\lambda,\pi})$ has such a conjugacy towards a language of two-way infinite words over an alphabet $A = \{x_1, x_2, \ldots, x_n\}$: Let *cod* be defined as $cod(y) = x_j$ if $y \in I_j$, and be extended by setting $cod(O(y)) = \ldots cod(T^k_{\lambda,\pi}(y)) cod(T^{k+1}_{\lambda,\pi}(y)) \ldots = \ldots x_{i_k} x_{i_{k+1}} \ldots$: the conjugacy is defined by $\phi(y) = cod(O(y))$ towards $(A^{\mathbb{Z}}, \sigma)$ where the **left shift** σ on $A^{\mathbb{Z}}$ is defined by $\sigma(w) = w'$ iff $w'_i = w_{i+1}$ for all $i \in \mathbb{Z}$ and the topology on $A^{\mathbb{Z}}$ is the product topology. We call $\phi(\mathcal{I})$ the **symbolic orbits** of the interval exchange $T_{\lambda,\pi}$. Note that if minimality holds then in each word in $\phi(\mathcal{I})$, every subword of every word in $\phi(\mathcal{I})$ occurs, and every of its subword occurs with bounded gaps (see e.g. [Qué87]). So each distinct orbit contains most of the information about the whole system, and also can be studied locally.

3 Geometric Interpretation of Interval Exchange Transformations

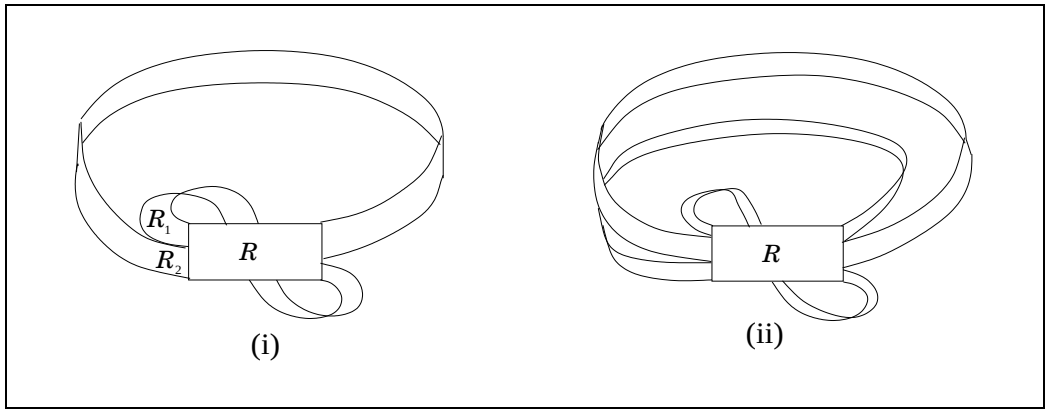
We here recall and develop a geometrical construction from interval exchange transformations (see e.g.[Arn81, Vee82, Ker85]) which comes from the fact that the dynamical system based on iterating $T_{\lambda,\pi}$ can be represented as a foliated surface called a *suspension* (see e.g. [CLN85]). A **foliated box** is a product of two intervals $I \times I'$ for which the individual **leaves** are the sets of the form $I \times \{x\}$, with $x \in I'$. Let $\lambda = (\lambda_1, ..., \lambda_n)$ be the length vector of $T_{\lambda,\pi}$ and let $\mathcal{R} = [0,1] \times [0,1)$ and $\mathcal{R}_i = [0,1] \times [0,\lambda_i)$ be $n+1$ foliated boxes. A foliated surface $\mathcal{M}_{\lambda,\pi}$, henceforth called the **stripped surface** of $T_{\lambda,\pi}$, is obtained by applying the following identification rules for each $i = 1, \ldots, n$:

$$\{1\} \times [0, \lambda_i) \text{ with } \{0\} \times [1 - b_i, 1 - b_{i-1}),$$
$$\{0\} \times [0, \lambda_i) \text{ with } \{1\} \times [1 - b^{\pi}_{\pi(i)}, 1 - b^{\pi}_{\pi(i)-1}).$$

Since individual leaves of the boxes fit together through the identification rules, we get leaves running on $\mathcal{M}_{\lambda,\pi}$. For example, the surface $\mathcal{M}_{\lambda,\pi}$ in Figure 3.1(i) (respect. (ii)) corresponds to the exchange transformation given in Figure 2.1(i) (respect. (ii)).

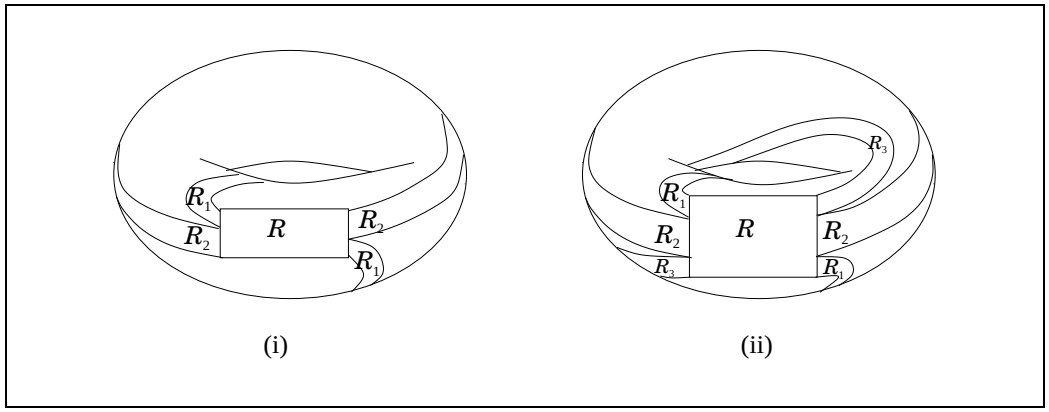
To recover $T_{\lambda,\pi}$ from $\mathcal{M}_{\lambda,\pi}$, consider e.g. the transverse arc $\{0\} \times [0,1)$ of $\mathcal{R}$ and its corresponding first-return map: the two-way infinite leaves of $\mathcal{M}_{\lambda,\pi}$ are bijectively mapped to the orbits of $(\mathcal{I}, T_{\lambda,\pi})$, and therefore also to the symbolic orbits $\phi(\mathcal{I})$.

Fig. 3.1



There is a compact surface $\overline{\mathcal{M}_{\lambda,\pi}}$ containing $\mathcal{M}_{\lambda,\pi}$, obtained by taking the complete rectangles in the above construction, i.e. the rectangles with their four sides. The surface $\mathcal{M}_{\lambda,\pi}$ can now be embedded into an orientable surface of genus g (i.e. g-holed) without boundary: one caps off each boundary component of $\overline{\mathcal{M}_{\lambda,\pi}}$ with a punctured disk. The genus g is given by the Euler-Poincaré characteristic formula: $1 - n + C = 2 - 2g$ where n is the number of intervals and C is equal to the number of boundary components of $\overline{\mathcal{M}_{\lambda,\pi}}$. Figure 3.2 (i) (respect. (ii)) shows a torus of genus 1, in which the irreducible exchange transformation over 2 intervals (respect. 3) of Figure 3.1(i) (respect. (ii)) has been embedded.

Fig. 3.2



Punctures of the embedding surface are needed because of the following property:

Remark **3.1** *Any two leaves of $\mathcal{M}_{\lambda,\pi}$ are homotopic iff they correspond to the same symbolic orbit.*

So in respect to homotopy, moves of the leaves on the embedding surface are possible. Now, as announced, we concentrate on interval exchange transformations of rotation class:

Remark **3.2** *([NR97] Cor 3.10). If π is irreducible, then $T_{\lambda,\pi}$ is of* **rotation class** *iff $\mathcal{M}_{\lambda,\pi}$ is embeddable into a torus of genus 1.*

We shall consider specific embeddings of irreducible interval exchange transformations of rotation class over n intervals in the one-holed torus denoted by a pair (α, β): $0 \leq \alpha < n$ is the number of boxes parallely running once along the meridian of the torus, and $0 \leq \beta < n$ is the number of boxes parallely running once along the meridian and once along the longitude

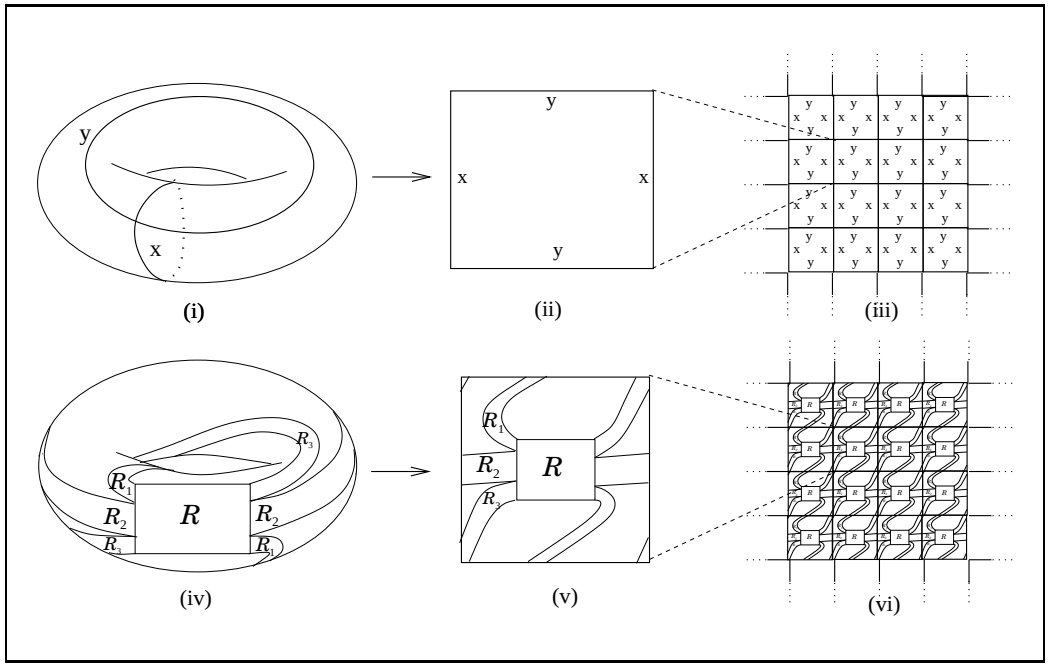
of the torus; the $n-\alpha-\beta$ remaining boxes are constrained to parallely run once along the longitude of the torus. Considering that parallel boxes can be embedded in a global box, this way of embedding $\mathcal{M}_{\lambda,\pi}$ corresponds to embed it as the 3 intervals exchange transformation in Figure 3.2(ii) (see also Figure 3.3(iv)-(vi) and Figure 3.4). Thus for instance, the embeddings shown in Figure 3.2(i) (respect. (ii)) corresponds to $(1,0)$ (respect. $(1,1)$). To such an embedding corresponds a unique permutation π: α is the largest number $0 \leq \alpha \leq n-1$ such that the permutation π restricted to $\{1,2,\ldots,\alpha\}$ sends i to $n-\alpha+i$; and β is such that π restricted to $\{\alpha+1,...,n-\beta\}$ sends $\alpha+i$ to $\beta+i$, and restricted to $\{n-\beta+1,...,n\}$ sends $n-\beta+i$ to i. The converse does not hold:

Remark **3.3** *An interval exchange transformation of rotation class has either a unique embedding of type $(.,.)$ or three different ones with the forms $(\alpha,0)$, $(0,n-\alpha)$, $(\alpha,n-\alpha)$.*

Corollary **3.4** *Among all the embeddings (α,β) for a given permutation π as above, there is only one for which $\alpha>0=\pi^{-1}(n)$ and $\beta>0=n-\pi^{-1}(1)+1$.*

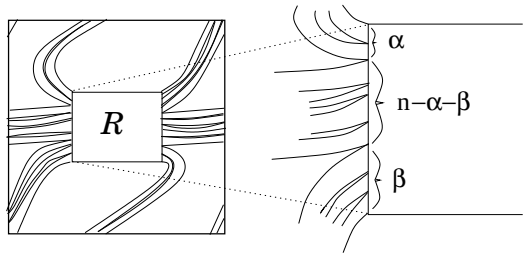
There exists a planar representation of these embeddings. Every surface of genus g can be seen as a polygon whose sides are pairwise identified. Instead of identifying sides of the same polygon, it is also possible to consider infinitely many copies of the polygon with its sides marked, and to glue these copies along their sides so to obtain a tiling of some simply connected infinite surface. This surface together with its tiling is in some sense unique, and is called the *universal covering* (see e.g. [Mas67]). For the torus of genus 1, the polygon is a square with opposite sides identified, called a **flat torus**, and the universal covering is the usual plane $\mathbb{R}^2$. This is shown in Figure 3.3(i)-(iii) where a torus is shown in (i) together with two curves x and y. When cutting along them, one obtains the flat torus shown in (ii) for which x and y are the side pairs, and some of the tiling of the universal covering is shown in (iii). Figure 3.3(iv)-(vi) shows the same but with the embedded stripped surface $\mathcal{M}_{\lambda,\pi}$ of Figure 3.2(ii).

Fig. 3.3



In respect to this representation, Figure 3.4 (i) shows a more general case of an embedding (α,β), i.e. $(2,3)$, of an interval exchange transformation over 9 intervals. Figure 3.4 (ii) zooms in the central box $\mathcal{R}$ indicating the number of respective boxes according to α, β and n.

Fig. 3.4



4 Rauzy Induction

Consider a stripped surface $\mathcal{M}_{\lambda,\pi}$ embedded into a punctured torus of genus 1. According to Remark 3.1, leaves can be moved continuously on this torus without changing their codings. More generally full foliated boxes of $\mathcal{M}_{\lambda,\pi}$ can be continuously moved on the torus without changing their widths λ_i. Let $\overline{\mathcal{R}}, \overline{\mathcal{R}_\jmath}$ be the completed boxes, i.e. the building boxes for $\overline{\mathcal{M}_{\lambda,\pi}}$. A **glueing move** is a continuous move of a box $\overline{R_i}$ of $\overline{\mathcal{M}_{\lambda,\pi}}$, followed by an identification of $\overline{R_i}$'s upper side (resp. lower side) with the union of $\overline{R}$'s lower side (resp. upper side) and one of the boxes' lower sides (resp. upper sides) adjacent to it. Of course a glueing move is not possible for all i, and the following lemma lists all the possibilities for the embeddings of type (α, β) (to better understand its meaning, the reader could first have a look at the examples pictured in Figure 4.1 and Figure 4.2). When a glueing move is possible it yields an embedding of another stripped surface $\mathcal{M}_{\lambda',\pi'}$ for some λ' and π'.

Lemma **4.1** *Let $T_{\lambda,\pi}$ be irreducible and of rotation class over n intervals, $n \geq 2$. Four different kinds of glueing moves are always possible. Their effects on λ and on the embeddings (α, β), and therefore on π, are described as follows, where we denote $n - \beta + 1$ by $\overline{\beta}$:*

- *If $\beta > 0$ and $n - \alpha - \beta > 0$, a* **move under** *$\mathcal{R}_n$ is to glue the box $\mathcal{R}_{\alpha+1}$ under $\mathcal{R}_n$. The effect is:*
  - *$(\lambda_1, \ldots, \lambda_n)$ is sent to $(\lambda_1, \ldots, \lambda_n + \lambda_{\alpha+1})$.*
  - *(α, β) is sent to $(\alpha + 1, \beta)$.*
- *If $\alpha > 0$ and $n - \alpha - \beta > 0$, a* **move under** *$\mathcal{R}_\alpha$ is to glue the box $\mathcal{R}_{\alpha+1}$ under $\mathcal{R}_\alpha$. The effect is:*
  - *$(\lambda_1, \ldots, \lambda_\alpha, \lambda_{\alpha+1}, \ldots, \lambda_n)$ is sent to $(\lambda_1, \ldots, \lambda_\alpha + \lambda_{\alpha+1}, \lambda_{\alpha+2}, \ldots, \lambda_n, \lambda_{\alpha+1})$,*
  - *(α, β) is sent to $(\alpha, \beta + 1)$.*
- *If $\alpha > 0$ and $n - \alpha - \beta > 0$, a* **move over** *$\mathcal{R}_1$ is to glue the box $\mathcal{R}_{\overline{\beta}-1}$ over $\mathcal{R}_1$. The effect is:*
  - *$(\lambda_1, \ldots, \lambda_n)$ is sent to $(\lambda_1 + \lambda_{\overline{\beta}-1}, \lambda_2, \ldots, \lambda_n)$.*
  - *(α, β) is sent to $(\alpha, \beta + 1)$.*
- *If $\beta > 0$ and $n - \alpha - \beta > 0$, a* **move over** *$\mathcal{R}_{\overline{\beta}}$, is to glue the box $\mathcal{R}_{\overline{\beta}-1}$ over $\mathcal{R}_{\overline{\beta}}$. The effect is:*
  - *$(\lambda_1, \ldots, \lambda_{\overline{\beta}-1}, \lambda_{\overline{\beta}}, \ldots, \lambda_n)$ is sent to $(\lambda_{\overline{\beta}-1}, \lambda_1, \ldots, \lambda_{\overline{\beta}} + \lambda_{\overline{\beta}-1}, \lambda_{\overline{\beta}+1}, \ldots, \lambda_n)$.*
  - *(α, β) is sent to $(\alpha + 1, \beta)$.*

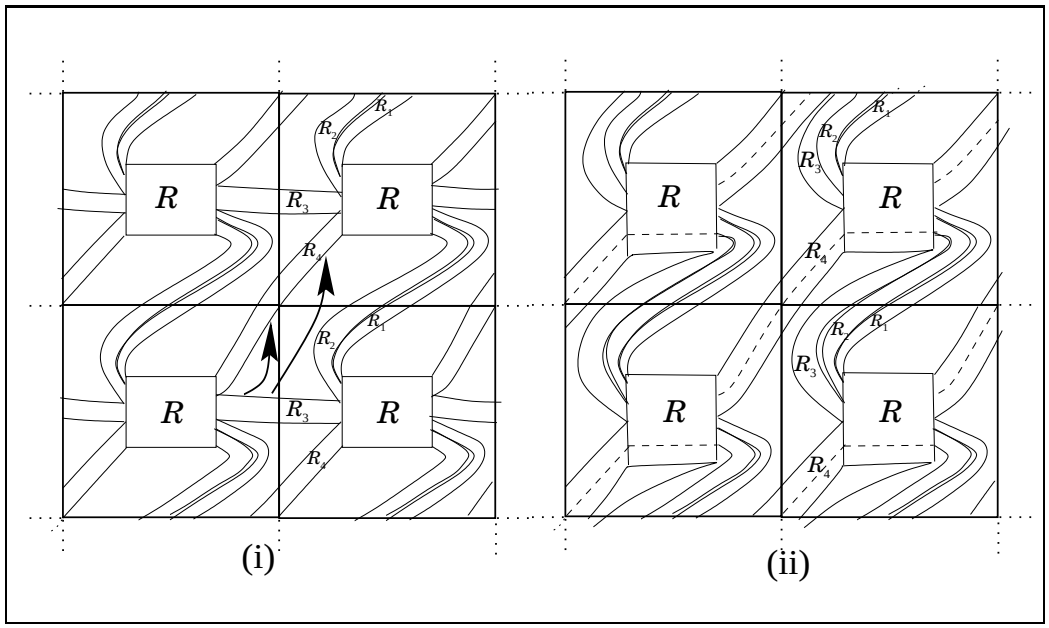
Proof. These effects can be checked directly on the embeddings. To see that all these moves are always possible, note first that if $n - \alpha - \beta = 0$, then the permutation (α, β) can be reembedded as $(0, \beta)$ or as $(\alpha, 0)$. Second if $n - \alpha - \beta > 0$ and $\beta = 0$ then the permutation embedded as $(\alpha, 0)$ can be reembedded as $(0, n - \alpha)$ ($\alpha \neq 0$ since $T_{\lambda,\pi}$ is irreducible), and therefore, moves under $\mathcal{R}_n$ and over $\mathcal{R}_{\overline{\beta}}$ become possible. Otherwise, if $\alpha = 0$ then the permutation embedded as $(0, \beta)$ can be reembedded as $(n - \beta, 0)$ ($\beta \neq 0$ since $T_{\lambda,\pi}$ is irreducible) and therefore moves under $\mathcal{R}_\alpha$ and over $\mathcal{R}_1$ become possible. To see that these moves are the only ones for the embeddings of type (α, β), take in consideration that punctures cannot vanish by the application of a move. $\diamondsuit$

Corollary **4.2** *The set of irreducible interval exchange transformations of rotation class over n intervals, $n \geq 2$ fixed, is setwise invariant under the above four moves.*

Proof. According to the definition of α and β, the only reducible permutation is denoted by (α, β) is $(0, 0)$. Neither moves nor reembeddings as described above can lead to $(0, 0)$. $\diamondsuit$

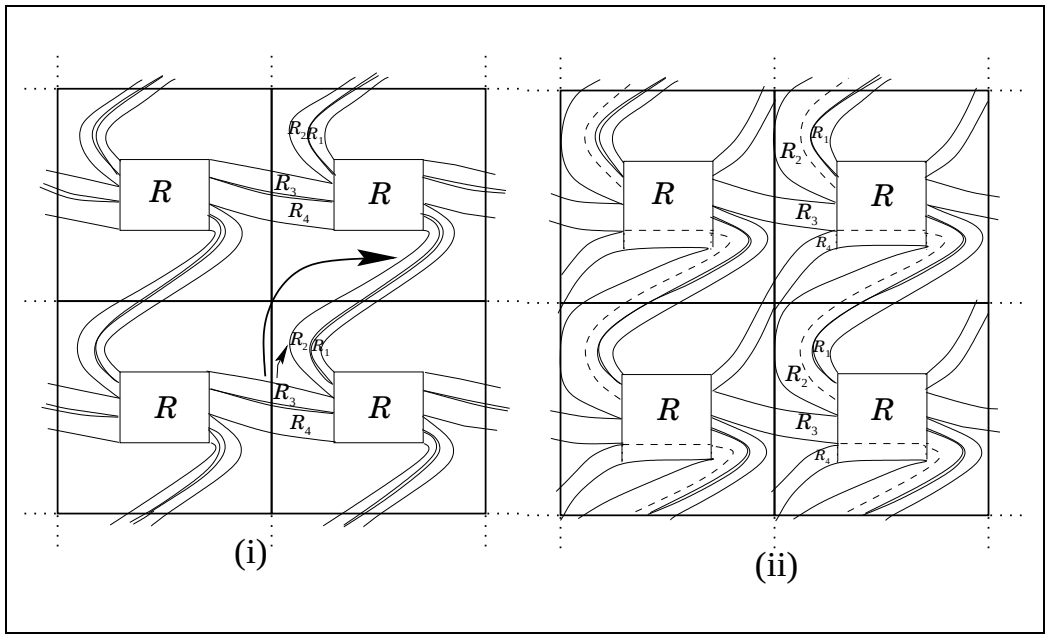
Let us show some examples using the universal covering representations (cf. Figure 3.3). Glueings can be observed by focusing only on four adjacent copies in a square. Consider an interval exchange transformation $T_{\lambda,\pi}$ over 4 intervals, and let $\pi(1) = 3$, $\pi(2) = 4$, $\pi(3) = 2$, $\pi(4) = 1$ embedded as $(2, 1)$. Then a glueing move under $\mathcal{R}_n$ is shown in Figure 4.1: one glues $\mathcal{R}_3$ under $\mathcal{R}_4$. The new stripped surface $\mathcal{M}_{\lambda',\pi'}$ is such that $\pi'(1) = 2$, $\pi'(2) = 3$, $\pi'(3) = 4$, $\pi'(4) = 1$ embedded as $(3, 1)$, and $\lambda' = (\lambda_1, \lambda_2, \lambda_3, \lambda_4 + \lambda_3)$.

Fig. 4.1



Another example is given in Figure 4.2 for which $\pi(1) = 3$, $\pi(2) = 4$, $\pi(3) = 1$, $\pi(4) = 2$ embedded as $(2, 0)$, where a glueing move under $\mathcal{R}_\alpha$ with $\alpha = \pi^{-1}(n)$ is applied: one glues $\mathcal{R}_3$ under $\mathcal{R}_2$. The new stripped surface $\mathcal{M}_{\lambda',\pi'}$ is such that $\pi'(1) = 3$, $\pi'(2) = 4$, $\pi'(3) = 2$, $\pi'(4) = 1$ embedded as $(2, 1)$,and $\lambda' = (\lambda_1, \lambda_2 + \lambda_3, \lambda_4, \lambda_3)$.

Fig. 4.2



Inverting the application of each above glueing move, we get **cutting moves**: beginning with $T_{\lambda^{(j)},\pi^{(j)}}$ and trying to find a $T_{\lambda^{(j+1)},\pi^{(j+1)}}$ such that the former is obtained from the latter by one of the glueing moves. According to Lemma 4.1, this is equivalent to solve the following equation systems:

- If the move is under $\mathcal{R}_n$:

$$\begin{aligned} \lambda_1^{(j)} &= \lambda_1^{(j+1)}, \\ &\cdots \\ \lambda_{n-1}^{(j)} &= \lambda_{n-1}^{(j+1)}, \\ \lambda_n^{(j)} &= \lambda_n^{(j+1)} + \lambda_{\alpha^{(j+1)}+1}^{(j+1)}. \end{aligned} \tag{1}$$

- If the move is under $\mathcal{R}_\alpha$ (here, accordingly to Lemma 4.1, $\alpha^{(j)} = \alpha^{(j+1)}$ is denoted by α):

$$\begin{aligned} \lambda_1^{(j)} &= \lambda_1^{(j+1)}, \\ &\cdots \\ \lambda_{\alpha-1}^{(j)} &= \lambda_{\alpha-1}^{(j+1)}, \\ \lambda_\alpha^{(j)} &= \lambda_\alpha^{(j+1)} + \lambda_{\alpha+1}^{(j+1)}, \\ \lambda_{\alpha+1}^{(j)} &= \lambda_{\alpha+2}^{(j+1)}, \\ &\cdots, \\ \lambda_{n-1}^{(j)} &= \lambda_n^{(j+1)}, \\ \lambda_n^{(j)} &= \lambda_{\alpha+1}^{(j+1)}. \end{aligned} \tag{2}$$

-If the move is over $\mathcal{R}_1$:

$$\begin{aligned} \lambda_1^{(j)} &= \lambda_1^{(j+1)} + \lambda_{\overline{\beta}^{(j+1)}-1}^{(j+1)}, \\ \lambda_2^{(j)} &= \lambda_2^{(j+1)}, \\ &\cdots \\ \lambda_n^{(j)} &= \lambda_n^{(j+1)}. \end{aligned} \tag{3}$$

- If the move is over $\mathcal{R}_{\overline{\beta}}$: (here, accordingly to Lemma 4.1, $\overline{\beta}^{(j)} = \overline{\beta}^{(j+1)}$ denoted by $\overline{\beta}$):

$$\begin{array}{rl} \lambda_1^{(j)} & = \lambda_{\overline{\beta}-1}^{(j+1)}, \\ \lambda_2^{(j)} & = \lambda_1^{(j+1)}, \\ \ldots & \\ \lambda_{\overline{\beta}-1}^{(j)} & = \lambda_{\overline{\beta}-2}^{(j+1)}, \\ \lambda_{\overline{\beta}}^{(j)} & = \lambda_{\overline{\beta}}^{(j+1)} + \lambda_{\overline{\beta}-1}^{(j+1)}, \\ \lambda_{\overline{\beta}+1}^{(j)} & = \lambda_{\overline{\beta}+1}^{(j+1)}, \\ \ldots, & \\ \lambda_n^{(j)} & = \lambda_n^{(j+1)} \end{array} \tag{4}$$

The pair of cutting moves given by Eq. Sys. 1 and 2 (respect. Eq. Sys. 3 and 4) denoted by S_1 (respect. S_2) are two ways of defining **Rauzy induction** ([Rau79] p.322) (see also [Vee82, Zor96]). The difference between the S_1-system and S_2-system is that cutting moves of S_1 are always applied from the last interval, and those for S_2 are always applied from the first interval. The word "induction" is justified since application of cutting moves can be iterated. The separation into two distinct pairs of the above four glueing/cutting moves is justified by the following result:

Remark **4.3** *Let $T_{\lambda,\pi}$ be irreducible, irrational and of rotation class. Rauzy induction can be deterministically iterated to infinity using either S_1 or S_2.*

Proof. Let us check it for the S_1-system. According to Lemma 4.1, in Eq. Sys. 1, $\lambda_{\alpha^{(j)}}^{(j)} = \lambda_{\alpha^{(j+1)}+1}^{(j+1)}$, so that its only additive equation becomes $\lambda_n^{(j)} = \lambda_n^{(j+1)} + \lambda_{\alpha^{(j)}}^{(j)}$. In Eq. Sys. 2, $\lambda_n^{(j)} = \lambda_{\alpha+1}^{(j+1)}$, so that its only additive equation becomes $\lambda_{\alpha^{(j)}}^{(j)} = \lambda_{\alpha^{(j)}}^{(j+1)} + \lambda_n^{(j)}$. Since length vectors are in $\mathbb{R}^+$, the choice between the two moves depends on the sign of $\lambda_{\alpha^{(j)}}^{(j)} - \lambda_n^{(j)}$. Indeed, since $T_{\lambda,\pi}$ is irrational, $\lambda_{\alpha^{(j)}}^{(j)} - \lambda_n^{(j)} \neq 0$ for every j. Otherwise this equality would mean that there is a second relation of rational dependance among $\lambda_1^{(j)}, \ldots, \lambda_n^{(j)}$, the first one coming from the normalization relation $\sum_{i=1}^n \lambda_i^{(0)} = 1$ at step 0 which propagates to $\sum_{i=1}^n k_i^{(j)} \lambda_i^{(j)} = 1$ at step j, where $k_i^{(j)} > 0$ for all i, j thus implying independance between this relation and the relation $\lambda_{\alpha^{(j)}}^{(j)} - \lambda_n^{(j)} = 0$. Hence the irrationality rank of $\lambda^{(j)}$, hence of λ, could not be more than $n-2$, which is a contradiction. The same kind of reasoning applies to the S_2-system about the sign of $\lambda_{\overline{\beta}^{(j)}}^{(j)} - \lambda_1^{(j)}$. $\diamondsuit$

So iterating Rauzy induction leads here to a **decomposition algorithm** of an irreducible interval exchange transformation $T_{\lambda,\pi}$ of rotation class. The **expansion** of this iterating process is defined by recording the sequence of cutting moves which are applied. In the S_1-system (respect. S_2-system), putting 0 for Eq. Sys. 1 (respect. Eq. Sys. 3) and 1 for Eq. Sys. 2 (respect. Eq. Sys. 4), it is defined as:

$$\epsilon_{j+1} = \begin{cases} 0 & \text{if } \lambda_n^{(j)} > \lambda_\alpha^{(j)}, \quad (respect.\ \lambda_1^{(j)} > \lambda_{\overline{\beta}}^{(j)}), \\ 1 & \text{if } \lambda_n^{(j)} < \lambda_\alpha^{(j)}, \quad (respect.\ \lambda_1^{(j)} < \lambda_{\overline{\beta}}^{(j)}). \end{cases} \tag{5}$$

We now give a full description of the algorithm in the S_1-system case (it is similar for the S_2-system case). Recall first that cutting moves implies substractions by 1 on (α, β) accordingly to Lemma 4.1, and that $(\alpha, 0)$, $(0, n-\alpha)$ and $(\alpha, n-\alpha)$ are embeddings of the same permutation (cf. Remark 3.3). So whenever at some step $\alpha = 0$ or $\beta = 0$, we systematically reembed the interval exchange transformation using the embedding with no zero values (cf. Corollary 3.4).

Algorithm **4.4** **Input:** $T_{\lambda,\pi}$ *irreducible, irrational and of rotation class, with* $n > 1$ *and embedded as* (α, β)*, where* $\alpha \neq 0$ *and* $\beta \neq 0$*;* **Output:** $\{\epsilon_j\}_{j=1,\ldots}$

1. $j \leftarrow 0$,
   $\lambda^{(0)} \leftarrow \lambda$,
   $(\alpha^{(0)}, \beta^{(0)}) \leftarrow (\alpha, \beta)$.
2. `switch of`
   - $\lambda_n^{(j)} > \lambda_{\alpha^{(j)}}^{(j)}$:
     (a) `solve Eq. Sys.` 1 `to get` $\lambda^{(j+1)}$,
     (b) `if` $\alpha^{(j)} \neq 1$
     `then` $(\alpha^{(j+1)}, \beta^{(j+1)}) \leftarrow (\alpha^{(j)} - 1, \beta^{(j)})$,
     `else` $(\alpha^{(j+1)}, \beta^{(j+1)}) \leftarrow (n - \beta^{(j)}, \beta^{(j)})$. `/* reembedding */`
     (c) $\epsilon_{j+1} \leftarrow 0$.
   - $\lambda_n^{(j)} < \lambda_{\alpha^{(j)}}^{(j)}$:
     (a) `solve Eq. Sys.` 2 `to get` $\lambda^{(j+1)}$,
     (b) `if` $\beta^{(j)} \neq 1$
     `then` $(\alpha^{(j+1)}, \beta^{(j+1)}) \leftarrow (\alpha^{(j)}, \beta^{(j)} - 1)$,
     `else` $(\alpha^{(j+1)}, \beta^{(j+1)}) \leftarrow (\alpha^{(j)}, n - \alpha^{(j)})$. `/* reembedding */`
     (c) $\epsilon_{j+1} \leftarrow 1$.
3. $j \leftarrow j + 1$.
4. `goto` (2).

A full example of the application of this algorithm is given at the end of the next section.

5 The Substitutions

Given an alphabet A, a **substitution** θ over A (see e.g. [Qué87]) is a map which sends every letter x of A to some word over A, and which is extended to any word $w = ...w_i w_{i+1}...$ over A, finite or not, by sending it to $...\theta(w_i)\theta(w_{i+1})....$ Recall that the leaves of the stripped surface $\mathcal{M}_{\lambda,\pi}$ which correspond to the orbits of the interval exchange transformation $T_{\lambda,\pi}$ can be represented into words over $\{x_1, x_2, \ldots, x_n\}$ by marking the boxes as the orbits of $T_{\lambda,\pi}$ visit them. This corresponds to the topological conjugacy ϕ introduced in Section 2, which maps each point of $\mathcal{I} \subset [0, 1)$ to its symbolic orbit. In the previous section, we saw how a glueing move yields an embedding of another stripped surface $\mathcal{M}_{\lambda',\pi'}$ for some λ' and π'. Strictly speaking, just a homotopy has been applied on the leaves, and therefore their original meaning has not been changed. One can however record the change by coding their associated symbolic orbits relatively to the new interval exchange transformation, i.e. $T_{\lambda',\pi'}$. The next result shows indeed that going from the original symbolic orbits to the new one is obtained by applying a substitution:

Proposition **5.1** *Let* $T_{\lambda,\pi}$ *be irreducible and of rotation class over* n *intervals. Let* t *be one of the glueing move of Lemma 4.1. Then there exists a unique substitution* θ_t *such that* $\phi \circ t = \theta_t \circ \phi$.

Proof.
- For a move under $\mathcal{R}_n$: only the leaves coming from $\mathcal{R}_{\alpha+1}$ are modified and stretched: they have to run once through $\mathcal{R}_n$ before going ahead. So $\theta(x_{\alpha+1}) = x_{\alpha+1}x_n$, while the other letters remain unchanged.
- For a move under $\mathcal{R}_\alpha$: only the leaves going into $\mathcal{R}_{\alpha+1}$ are stretched: they have to run once through $\mathcal{R}_\alpha$ and then under $\mathcal{R}_n$ before going their way ahead. So $\theta(x_{\alpha+1}) = x_\alpha x_n$. The letters whose indices start from $\alpha+2$ to n are permuted accordingly.
- For a move over $\mathcal{R}_1$: only the leaves coming from $\mathcal{R}_{\overline{\beta}-1}$ are modified and stretched: they have to run once through $\mathcal{R}_1$ before going ahead. So $\theta(x_{\overline{\beta}-1}) = x_{\overline{\beta}-1}x_1$, while the other letters remain unchanged.
- For a move over $\mathcal{R}_{\overline{\beta}}$: only the leaves going into $\mathcal{R}_{\overline{\beta}-1}$ are stretched: they have to run once over $\mathcal{R}_{\overline{\beta}}$ and then through $\mathcal{R}_1$ before going their way ahead. So $\theta(x_{\overline{\beta}-1}) = x_{\overline{\beta}}x_1$. The letters whose indices start from 1 to $\overline{\beta}-2$ are permuted accordingly. $\diamondsuit$

The substitutions for the S_1-system, i.e. for the pair (under $\mathcal{R}_n$, under $\mathcal{R}_\alpha$), are therefore given as follows, where the third index is the ϵ value of the corresponding equation system:

$$\begin{array}{lll} \theta_{n,\alpha,0}: & x_1 & \mapsto x_1 \\ & & \dots \\ & x_\alpha & \mapsto x_\alpha \\ & x_{\alpha+1} & \mapsto x_{\alpha+1}x_n \\ & x_{\alpha+2} & \mapsto x_{\alpha+2} \\ & & \dots \\ & x_n & \mapsto x_n \end{array} \qquad \begin{array}{lll} \theta_{n,\alpha,1}: & x_1 & \mapsto x_1 \\ & & \dots \\ & x_\alpha & \mapsto x_\alpha \\ & x_{\alpha+1} & \mapsto x_\alpha x_n \\ & x_{\alpha+2} & \mapsto x_{\alpha+1} \\ & & \dots \\ & x_n & \mapsto x_{n-1} \end{array}$$

Taking into account irreducibility, and restrictions given by Lemma 4.1, this gives a set of $2(n-1)$ substitutions: $\theta_{n,\alpha,0}$ with $\alpha \in \{0,..,n-2\}$ since for a move under $\mathcal{R}_n$, then $\beta > 0$ and $n-\alpha-\beta > 0$, so α cannot be greater than $n-2$; and $\theta_{n,\alpha,1}$, where $\alpha \in \{1,..,n-1\}$ since for a move under $\mathcal{R}_\alpha$, then $\alpha > 0$ and $n-\alpha-\beta > 0$.
The substitutions for the S_2-system, i.e. for the pair (over $\mathcal{R}_1$, over $\mathcal{R}_{\overline{\beta}}$) are (recall that $\overline{\beta} = n-\beta+1$):

$$\begin{array}{lll} \theta'_{n,\beta,0}: & x_1 & \mapsto x_1 \\ & & \dots \\ & x_{\overline{\beta}-2} & \mapsto x_{\overline{\beta}-2} \\ & x_{\overline{\beta}-1} & \mapsto x_{\overline{\beta}-1}x_1 \\ & x_{\overline{\beta}} & \mapsto x_{\overline{\beta}} \\ & & \dots \\ & x_n & \mapsto x_n \end{array} \qquad \begin{array}{lll} \theta'_{n,\beta,1}: & x_1 & \mapsto x_2 \\ & & \dots \\ & x_{\overline{\beta}-2} & \mapsto x_{\overline{\beta}-1} \\ & x_{\overline{\beta}-1} & \mapsto x_{\overline{\beta}}x_1 \\ & x_{\overline{\beta}} & \mapsto x_{\overline{\beta}} \\ & & \dots \\ & x_n & \mapsto x_n \end{array}$$

This also gives a set of $2(n-1)$ substitutions: $\theta'_{n,\beta,0}$ with $\beta \in \{0,..,n-2\}$ and $\theta'_{n,\beta,1}$ with $\beta \in \{1,..,n-1\}$. In case $n > 2$, these two sets can be simplified into smaller sets of $(n+1)$ substitutions. Here is the result for the S_1-system:

Proposition **5.2** *Let $n > 2$. Let ρ be the substitution $x_2 \mapsto x_1x_2$ (the other letters remain unchanged), and τ_i be the substitutions induced by the $n-1$ transpositions $(1\ i)$, for $i = 2, \dots n$, i.e. $x_i \mapsto x_1$ and $x_1 \mapsto x_i$ (the other letters remain unchanged). Then the substitutions $\theta_{n,\alpha,j}$, $j = 0, 1$, can be obtained by a finite composition of ρ, τ_i and $\theta_{n,0,0}$.*

Proof. For the facility of the presentation we set $\tau_1 =$ the identity. A straightforward calculation shows then that for $0 \le \alpha \le n-2$ one has $\theta_{n,\alpha,0} = \tau_{\alpha+1} \circ \theta_{n,0,0} \circ \tau_{\alpha+1}$. Now note that the τ_i's generate the set of all the possible letters permutations induced by a

permutation of $\{1,\ldots,n\}$. For $\alpha = 1,2,\ldots,n-1$, denote by τ'_α the substitutions induced by the transposition $(2 \quad \alpha+1)$ and by r_α the one induced by the circular permutation $(n\ n-1\ \cdots\ \alpha+1)$. For $1 \leq \alpha \leq n-1$ we compute $\tau_\alpha \circ \tau'_\alpha \circ \rho \circ \tau'_\alpha \circ \tau_\alpha$: this is the substitution which sends $x_{\alpha+1}$ to $x_\alpha x_{\alpha+1}$ (equals ρ when $\alpha = 1$). We compose next with r_α on the left and obtain $\theta_{n,\alpha,1}$. $\diamondsuit$

Theorem **5.3** *Let $T_{\lambda,\pi}$ be irreducible, irrational and of rotation class over n intervals, $n > 2$ (respect. $n = 2$). Every subword of its symbolic orbits can be obtained to any prescribed length by applying a composition of $n+1$ substitutions (respect. 2 substitutions) over $\{x_1,...,x_n\}$.*

Proof. We can use either the S_1-system or the S_2-system of Rauzy induction. Let us consider the former and consider the expansion obtained by the decomposition algorithm applied to $T_{\lambda,\pi}$. To each truncation at step N, for some $N > 0$, we get a finite composition of cutting moves. By reversing it, we get a composition of glueing moves which, according to Proposition 5.1, can be translated into a composition of N substitutions $\theta_{i_1} \circ .... \circ \theta_{i_N}$. We therefore obtain a sequence of compositions $\{\theta_{i_1} \circ .... \circ \theta_{i_N}\}_{N\in\mathbb{N}}$. Concentrating on its effect to x_1, this leads to a sequence of words $W = \{\theta_{i_1} \circ .... \circ \theta_{i_N}(x_1)\}_{N\in\mathbb{N}}$ which all are pieces of symbolic orbits of $T_{\lambda,\pi}$. For each substitution in the set corresponding to S_1, $\theta(x_1)$ begins with x_1. Hence, for every $N > 0$, $\theta_{i_1} \circ .... \circ \theta_{i_N}(x_1) = \theta_{i_1} \circ .... \circ \theta_{i_{N-1}}(x_1)v$ for some v over $\{x_1,...,x_n\}$, which means that prefixes are preserved. So indexing the words of W starting from their left ends, there exists a limit word $w_{\lambda,\pi}$ of W, which by construction, represents a positive orbit of $T_{\lambda,\pi}$. The same applies with the S_2-system, but with the substitutions θ'_j and by applying the compositions to x_n.

Now, since $T_{\lambda,\pi}$ is irreducible and irrational, it is minimal (see Section 2). Hence, using the topological conjugacy between $(\mathcal{I}, T_{\lambda,\pi})$ and $(\phi(\mathcal{I}), \sigma)$, the word $w_{\lambda,\pi}$ contains every possible subword occuring in all the symbolic orbits of $T_{\lambda,\pi}$. It is known that minimality means that each subword occurs in $w_{\lambda,\pi}$ with bounded gaps (see e.g. [Qué87]). Therefore, there exists a finite $k > 0$ such that every subword with a prescribed length belongs to $\theta_{i_1} \circ .... \circ \theta_{i_N}(x_1)$ (respect. $\theta'_{i_1} \circ .... \circ \theta'_{i_N}(x_n)$ in the S_2-system) for all $N > k$. $\diamondsuit$

Let us consider a full example. Consider an interval exchange $T_{\lambda,\pi}$ where $\lambda = (1, \sqrt[3]{2}, (\sqrt[3]{2})^2)$ and $\pi(1) = 3$, $\pi(2) = 1$ $\pi(3) = 2$ embedded as $(1,2)$. By applying the decomposition algorithm of the S_1-system of Rauzy induction, one may check that the vector decomposition of the vector goes as follows for the ten first iterations (there, `lambda` indicates the state of the length vector λ and `pi` indicates the new state of the permutation with its embedding):

```
lambda: (1 1.259921 1.5874011)            pi: (1,2)   => Eq. Sys. 1
lambda: (1 1.259921 0.5874011)            pi: (1,1)   => Eq. Sys. 2
lambda: (0.4125989 0.5874011 1.259921)    pi: (2,1)   => Eq. Sys. 1
lambda: (0.4125989 0.5874011 0.8473221)   pi: (1,1)   => Eq. Sys. 1
lambda: (0.4125989 0.5874011 0.259921)    pi: (1,2)   => Eq. Sys. 2
lambda: (0.1526779 0.259921  0.5874011)   pi: (1,2)   => Eq. Sys. 1
lambda: (0.1526779 0.259921  0.4347232)   pi: (1,2)   => Eq. Sys. 1
lambda: (0.1526779 0.259921  0.2820453)   pi: (1,2)   => Eq. Sys. 1
lambda: (0.1526779 0.259921  0.12936734)  pi: (1,1)   => Eq. Sys. 2
lambda: (0.0233105 0.1293674 0.259921)    pi: (2,1)   => Eq. Sys. 1
```

This yields the ten first elements of the expansion:

$$\{\epsilon_i\} = \{0,1,0,0,1,0,0,0,1,0,...\}.$$

In the case of three intervals and in the S_1-system, we have the following four substitutions:

$\theta_{3,0,0}$:			$\theta_{3,1,0}$:			$\theta_{3,1,1}$:			$\theta_{3,2,1}$:		
x_1	$\mapsto$	x_1x_3	x_1	$\mapsto$	x_1	x_1	$\mapsto$	x_1	x_1	$\mapsto$	x_1
x_2	$\mapsto$	x_2	x_2	$\mapsto$	x_2x_3	x_2	$\mapsto$	x_1x_3	x_2	$\mapsto$	x_2
x_3	$\mapsto$	x_3	x_3	$\mapsto$	x_3	x_3	$\mapsto$	x_2	x_3	$\mapsto$	x_2x_3

According to Proposition 5.2, the simplification of the above set of substitutions is the following (in the case $n = 3$, there are also four substitutions):

$\theta_{3,0,0}$:			ρ :			τ_2 :			τ_3 :		
x_1	$\mapsto$	x_1x_3	x_1	$\mapsto$	x_1	x_1	$\mapsto$	x_2	x_1	$\mapsto$	x_3
x_2	$\mapsto$	x_2	x_2	$\mapsto$	x_1x_2	x_2	$\mapsto$	x_1	x_2	$\mapsto$	x_2
x_3	$\mapsto$	x_3	x_3	$\mapsto$	x_3	x_3	$\mapsto$	x_3	x_3	$\mapsto$	x_1

Next, by applying the construction used in Theorem 5.3, one may construct the composition Θ of the corresponding 10 substitutions $\theta_{n,\alpha,i}$, starting from the bottom. For instance, the last line in the above partial expansion indicates that Eq. Sys. 1 is applied with an embedding $(2, 1)$, which means a move under $\mathcal{R}_n$ with $\alpha = 0$, since $(2, 1) = (0, 1)$ (see Remark 3.3), and therefore the application of the substitution $\theta_{3,0,0}$. The above line indicates that Eq. Sys. 2 is applied with an embedding $(1, 1)$, which means a move under $\mathcal{R}_\alpha$ with $\alpha = 1$, and therefore the application of the substitution $\theta_{3,1,1}$. So the whole composition Θ is:

$$\Theta = \theta_{3,0,0} \circ \theta_{3,1,1} \circ \theta_{3,0,0} \circ \theta_{3,1,0} \circ \theta_{3,1,1} \circ \theta_{3,0,0} \circ \theta_{3,0,0} \circ \theta_{3,0,0} \circ \theta_{3,1,1} \circ \theta_{3,0,0}.$$

And,

$$\Theta(x_1) = x_1x_3x_2x_1x_3x_3x_2x_1x_3x_3x_2x_1x_3x_3x_2x_1x_3x_2x_2.$$

This word is a subword of the symbolic orbits of $T_{\lambda,\pi}$.

6 The Sturmian Properties

Sturmian words [MH40] (see the surveys [Ber95, BS97]) are words over two letters with many characterizations and properties. They are known to correspond to symbolic orbits of interval exchange transformations over 2 intervals. The above Theorem 5.3 allows one to get symbolic orbits of interval exchange transformations over n intervals, and therefore symbolic orbits over n letters. In this section, we check in what respect these have or lead to several of the same properties as the Sturmian ones.

6.1 Linear Complexity

The **complexity function** $P(w, m)$ of an infinite word w is defined as the map from $\mathbb{N}^*$ to $\mathbb{N}^*$ associating to each $m > 0$ the number of distinct subwords of length m that occur in w. Sturmian words are characterized by a complexity function $m + 1$ [MH40, CH73]. The following proposition is acknowledged in word combinatorics (see e.g. [Kea75]), however we could find no written proof in the litterature, so we provide one hereafter:

Proposition **6.1** *Let $T_{\lambda,\pi}$ be irreducible and irrational over n intervals, $n \geq 2$. The complexity function is $P(w, m) = (n - 1)m + 1$ where w is any of its symbolic orbits.*

Proof. The symbolic orbits are words over n letters $A = \{x_1, x_2, \ldots, x_n\}$ and recall that *cod* is defined as $cod(x) = x_j$ if $x \in I_j$, where $I_j = [b_{j-1}, b_j)$ is the jth subinterval involved in $T_{\lambda,\pi}$. Let W_m be the set of the words of length m which appear in the symbolic orbits of T. Let $w \in W_m$ for some m and let $I_w = \{x \in [0, 1) \mid cod(x) \ldots cod(T^{m-1}(x)) = w\}$. This is an interval and clearly, $\{I_w\}_{w \in W_m}$ gives a partition of $[0, 1)$.

For $m = 1$, subwords are the n letters. Assume the property is true for $m \geq 1$. Consider the $(n - 1)m + 1$ subwords and the associated intervals $\{I_w\}_{w \in W_m}$ partitioning $[0, 1)$. Denote by $a_1, ..., a_{(n-1)m+2}$ their extremities. Because of the irreducibility and irrationality of T, the orbits of $B = \{b_0, ...b_n\}$ are infinite and distinct [Kea75]. Moreover, the a_j's come from iterates of B by T. Hence, each b_j, with $j \in \{1, ..., (n - 1)\}$ belongs to the interior of $T^m(I_w)$

for some $w \in W_m$. There are K intervals with $1 \leq K \leq (n-1)$ in $\{T^m(I_w)\}_{w \in W_m}$ which cover B. Each of these, covering h points of B, corresponds to $h+1$ distinct ways of continuing the orbit pieces of length m. Hence, we have a total of $(n-1)m + 1 - K + ((n-1)+K)$ continuations, all longer by one more letter than the previous orbit pieces.

$\diamondsuit$

Corollary **6.2** *Every symbolic positive orbit obtained by composing the substitutions in Proposition 5.1 according to Theorem 5.3 has linear complexity.*

6.2 Continued Fractions

Sturmian words are known for a long time to have a very close relationship with the classical continued fraction algorithm [Mar82]. We can check this within our framework (see also [Ker85, Zor96]), considering e.g. the S_1-system of Rauzy:

Remark **6.3** *Let $T_{\lambda,\pi}$ be an interval exchange transformation over 2 intervals, with $\pi = (1\ 2)$, and $\lambda = (\lambda_1^{(0)}, \lambda_2^{(0)})$. The expansion $\{\epsilon_j\}$ is $(\underbrace{1,...,1}_{a_0}, \underbrace{0,...,0}_{a_1}, \underbrace{1,...,1}_{a_2}, 0, ...)$ iff the expansion given by the classical continued fraction algorithm of λ_1/λ_2 is $[a_0, a_1, a_2, , ...]$.*

Proof. Eq. Sys. 1 and 2 are respectively:

$$\begin{array}{ll} \lambda_1^{(j)} = \lambda_1^{(j+1)}, & \lambda_1^{(j)} = \lambda_1^{(j+1)} + \lambda_2^{(j)} \\ \lambda_2^{(j)} = \lambda_2^{(j+1)} + \lambda_1^{(j)}, & \lambda_2^{(j)} = \lambda_2^{(j+1)}. \end{array}$$

The classical continued fraction algorithm is based on the Euclid algorithm where the expansion corresponds to the multiples of the Euclid's divisions. Assuming that $\lambda_1^{(0)} > \lambda_2^{(0)}$, $\lambda_1^{(0)} = a_0\lambda_2^{(0)} + r_0$, where $a_0 > 0$, this amounts to apply a_0 times Eq. Sys. 2, hence to obtain $\lambda_1^{(a_0)} = r_0$ and $\lambda_2^{(a_0)} = \lambda_2^{(0)}$ with $\lambda_1^{(a_0)} < \lambda_2^{(a_0)}$. So, the second step is given by $\lambda_2^{(a_0)} = a_1\lambda_1^{(a_0)} + r_1$, which is equivalent to apply a_1 times Eq. Sys. 1, hence to obtain $\lambda_1^{(a_0+a_1)} = \lambda_1^{(0)}$ and $\lambda_2^{(a_0+a_1)} = r_1$. Then, carry on the process. $\diamondsuit$

Over $n > 2$ intervals, Rauzy induction implies a multidimensional continued fraction algorithm [Rau77, Ker85, Zor96]. However, properties of this algorithm are not directly checked in the above references. Let us first look at convergence: to each truncation at step N of an expansion, one can reverse the process like in Theorem 5.3 by starting, say from the length vector $c_0 = (1, 0, ...0)$ and thus get a length vector c_N. This defines a sequence $\{c_k\}$ of vectors. We say that such a sequence of approximations is **simply convergent** if $c_k \to \lambda$ when $(k \to \infty)$. It is **weakly convergent** in the sense of the multidimensional continued fraction literature (see e.g. [Bre81]) iff the sequences $\{c_k\}$ simply converge for any starting vector c_0. Convergence of the multidimensional continued fraction in the case of interval exchange transformations of rotation class is a direct consequence of their strong ergodicity:

Proposition **6.4** *Let $T_{\lambda,\pi}$ be irreducible, irrational and of rotation class over $n \geq 2$ intervals. Then the sequence of approximations $\{c_k\}$ obtained by the decomposition algorithm weakly converges to λ.*

Proof. Interval exchange transformation over two and three intervals are known to be uniquely ergodic (see e.g. [Kea75]). So the same strong property holds for the interval exchange transformation with 1 or 2 discontinuities. This means that for any leaf of $\mathcal{M}_{\lambda,\pi}$ and any center point on this leaf, the average number of times the leaf visits each box $\mathcal{R}_i$ in the stripped surface $\mathcal{M}_{\lambda,\pi}$ around the chosen center has limit λ_i. Consider the expansion of the decomposition of $T_{\lambda,\pi}$, and to each truncation step N, $N > 0$, reverse the process and call

h_N the map on the leaves of $\mathcal{M}_{\lambda,\pi}$ corresponding to the composition of the N glueing moves. Accordingly, c_N is the length vector arising at the step N, and T_{c_N,π_N} the interval exchange transformation. Let ℓ be a leaf in $\mathcal{M}_{\lambda,\pi}$ and let ℓ_N in $\mathcal{M}_{c_N,\pi_N}$ be such that $h_N(\ell_N) = \ell$. Put $l_{N,i} = \ell_N \cap R_i$. The number of boxes $\mathcal{R}_j$ visited by $h_N(l_{N,i})$ increases with N, for if not this contradicts minimality. By ergodicity, the number of times $h_N(l_{N,i})$ visits $\mathcal{R}_j$, for any $j \in \{1, 2, \ldots, n\}$, divided by the total number of visited boxes converges to λ_j as N goes to infinity. Since unique ergodicity holds this does not depend on the component chosen for $l_{N,i}$ in $\mathcal{R}_i$, nor on i. $\diamondsuit$

Revisiting the full example given at the end of the precedent section, we can check the convergence by comparing the frequencies of the letters of the subwords obtained by Theorem 5.3 with the length vector of $T_{\lambda,\pi}$. Normalizing $\lambda = (1, \sqrt[3]{2}, (\sqrt[3]{2})^2)$, one gets (0.25992105..., 0.32747998..., 0.41259896...), and for the word $\Theta(x_1)$ given in the example, the normalized frequencies of the three letters are (0.26315789..., 0.31578947..., 0.42105263...). Accordingly, one has just to go further in the expansion to get longer compositions of substitutions, therefore longer words and better approximations of the initial vector. For instance, with the same example and with an expansion prefix length of 50, one gets a word of length 66878 and frequencies (0.25992105..., 0.32747689..., 0.41260205...).

This convergent multidimensional continued fraction can be embedded in the general framework for extending continued fractions due to T.-Sós and Szekeres [Sze70]. One direct consequence is that one can readily extract forward recurrence equations. We quickly show here how the S_1-system of Rauzy induction can be translated into the main formal equations of [Sze70](p.119): the *selected suffix* $\mu(j)$ is the $\alpha^{(j)} > 0$, i.e. $(\pi^{(j)})^{-1}(n)$. Considering the expansion $\{\epsilon_j\}$ (see Eq. 5), we transform Eq. Sys. 1 and Eq. Sys. 2 into ($\alpha^{(j)}$ is denoted by α):

$$\begin{array}{ll} \lambda_n^{(j+1)} & = (1 - 2\epsilon_{j+1})(\lambda_n^{(j)} - \lambda_\alpha^{(j)}), \\ \lambda_\alpha^{(j+1)} & = \epsilon_{j+1}\lambda_n^{(j)} + (1 - \epsilon_{j+1})\lambda_\alpha^{(j)}, \\ \lambda_k^{(j+1)} & = \lambda_k^{(j)} \text{ for } k \neq n, \alpha. \end{array} \tag{6}$$

After having solved system (6) above we apply the permutation $(\alpha\ \alpha+1\ \alpha+2\ \cdots\ n)$ to $\lambda^{(j+1)}$'s coordinates if $\epsilon_{j+1} = 1$; and no permutation if $\epsilon_{j+1} = 0$. Hence, Eq. Sys. 6 is Eq. sys. 1 when $\epsilon_{j+1} = 0$ and Eq. Sys. 2 (after application of the permutation) when $\epsilon_{j+1} = 1$.

6.3 Sturmian Substitutions

The substitutions are the following in the S_1-system:

$$\begin{array}{lcl|lcl} \theta_{2,1,0}: & & & \theta_{2,1,1}: & & \\ x_1 & \mapsto & x_1x_2 & x_1 & \mapsto & x_1 \\ x_2 & \mapsto & x_2 & x_2 & \mapsto & x_1x_2 \end{array}$$

And in the S_2-system:

$$\begin{array}{lcl|lcl} \theta'_{2,0,0}: & & & \theta'_{2,1,1}: & & \\ x_1 & \mapsto & x_1 & x_1 & \mapsto & x_2x_1 \\ x_2 & \mapsto & x_2x_1 & x_2 & \mapsto & x_2 \end{array}$$

These are two sets of basic *Sturmian substitutions* (see the survey [BS97]). The ones of the S_1-system are exactly the same ones as used in [IY90] which are directly related to a theorem of Markov and Venkov [Mar82, Ven70] to describe the sequences over the alphabet $\{0, 1\}$ given as ([.] denotes the bottom function):

$$c(x, y) = \{[nx + y] - [(n-1)x + y]\}_{n=1,2,3\ldots}.$$

These sequences are called *lower mechanical sequences* and these are Sturmian words (see e.g [BS97]). In particular, they run all the finite Sturmian subwords as x runs $[0,1)$. Our framework provides another proof of the main result contained in [IY90](Th. 2.2)[2]:

Proposition **6.5** *Let $\lambda = (\lambda_1, \lambda_2)$ and let $[a_0, a_1, a_2, , ...]$ be the expansion of the traditional continued fraction algorithm of λ_1/λ_2. Let*

$$\Theta = \theta_{2,1,0}^{a_0} \circ \theta_{2,1,1}^{a_1} \circ \theta_{2,1,0}^{a_2} \circ \theta_{2,1,1}^{a_3} \circ ...(x_1)$$

Then the fixed point of Θ is equal to $\tau(c(1 - \frac{\lambda_2}{(\lambda_1+\lambda_2)}, 0))$ where $\tau(0) = x_1$ and $\tau(1) = x_2$.

Proof. First note that $c(x, 0)$ is equal to the orbit starting from 0 of the interval exchange transformation $T_{\lambda,(1,1)}$ where $\lambda = (1 - x, x)$. Hence $(c(\frac{\lambda_2}{\lambda_1+\lambda_2}, 0))$ is equal to the orbit starting from 0 of the exchange interval transformation $T_{\lambda,(1,1)}$ where $\lambda = (\lambda_1, \lambda_2)$. Use Remark 6.3 to conclude. $\diamondsuit$

7 Conclusion

We proposed an effective way of producing the symbolic orbits of any interval exchange transformation of rotation class by using a decomposition algorithm based on Rauzy induction. This algorithm relies on building compositions of substitutions taken among a finite set of generators, following an approximation scheme which happens to be a multidimensional continued fraction. We presented two different systems, i.e. S_1 and S_2, depending on the side (left or right) where the induction process takes place. This yields two different sets of substitutions. We do not know at this time whether these are the only possible systems, i.e. the only possible sets of substitutions which generate the searched for symbolic orbits. Also, we do not know if the minimum number of generators for them is $n + 1$ as given by Proposition 5.2 in case $n > 2$. Another direction of further investigation is the study of multidimensional continued fractions [Sze70, Bre81]. This could lead to a characterization of the periodic expansions, and therefore a characterization of the orbits which can be generated by iterating a single substitution (as it has been done for the 2 intervals case [CMPS93, Ber95]). To this respect, note that the set of possible compositions of substitutions that one can obtain by Rauzy induction can be described by a sofic system (see e.g. [NR97],p.1189). Finally, let us note that Sturmian substitutions can be also related to Dehn twists, i.e. surface homeomorphisms [LN95]. This has been used in a more generalized setting in [LN98] where a conjugacy from Dehn twists towards substitutions has been used. In fact, glueings moves and reembeddings can be related to Dehn twists applications, actually to non-integral ones (as in e.g.[Ker83]).

Acknowledgement

We are greatly endebted to the referee who pointed out to us many remarks to improve and simplify in many aspects the former version of this paper.

[2] To precisely compare our result with the one in [IY90], let us consider only expansions of type $x = [0, a_1, a_2, ...]$. The differences with Proposition 6.5 are that, first the indices of the involved substitutions are inverted, and second, the first term of the composition have power $a_1 - 1$ instead of a_1 (see Theorem 1.1). This can be explained as follows: inverting the indices of the composition of the substitutions generates $c(1 - x, 0)$ instead of $c(x, 0)$ because this amounts to invert the 0's and the 1's (except the first 0) in the limit word. Since $c(x, 0)$ is equal to the orbit starting from 0 of the interval exchange transformation $T_{\lambda,(1,1)}$ where $\lambda = (1 - x, x)$, the inversion leads to the same as the orbit starting from 0 of $T_{\lambda',(1,0)}$ where $\lambda' = (x, 1 - x)$. Now, one can check that if $x = [0, a_1, a_2, ...]$ then $\frac{x}{(1-x)} = [0, a_1 - 1, a_2, ...]$. Summing up, the interval exchange transformation described in [IY90] is the one with $\lambda = (1 - x, x)$, while ours is given by $\lambda = (x, 1)$.

References

[All94] Allouche J.-P. – Sur la complexité des suites infinies. *J. Bull. Belg. Math. Soc*, vol. 1, number 2, 1994, pp. 133–143.

[Arn81] Arnoux P. – Echanges d'intervalles et flots sur les surfaces. *Mono. Enseign. Math.*, vol. 29, 1981, pp. 5–39.

[Ber95] Berstel J. – Recent results on Sturmian words. *Developpements in Language Theory 95*, 1995, pp. 1–12. – World Scientific.

[Bre81] Brentjes A. J. – *Multidimensional continued fraction algorithms.* – Mathematisch Centrum, Amsterdam, 1981, *Mathematical Centre Tracts, 145.*

[BS97] Berstel J. and Séébold P. – Sturmian words. *In : Combinatorics on words. 2nd ed.*, pp. 30–93. – Cambridge University Press, 1997. Lothaire Editor.

[CH73] Coven E.M. and Hedlund G. – Sequences with minimal block growth. *Math. Sys. Th.*, vol. 7, 1973, pp. 138–153.

[CLN85] Camacho C. and Lins Neto A. – *Geometric Theory of Foliations.* – Birkäuser, 1985.

[CMPS93] Crisp D., Moran W., Pollington A. and Shiue P. – Substitution invariant cutting sequences. *Journal de Théorie des Nombres de Bordeaux*, vol. 5, number 1, 1993, pp. 123–138.

[GH55] Gottschalk W.H. and Hedlund G.A. – Topological dynamics. *American Math. Soc. Colloq. Pub*, vol. 36, 1955.

[IY90] Ito S. and Yasutomi S. – On continued fractions, substitutions and characteristic sequences. *Japan. J. Math.*, vol. 16, number 2, 1990, pp. 287–306.

[Kea75] Keane M. – Interval exchange transformations. *Math. Zeit.*, vol. 141, 1975, pp. 25–31.

[Ker83] Kerckhoff, S.P. – The Nielsen realization problem. *Annals of Math. (2)*, vol. 117, number 2, 1983, pp. 235–265.

[Ker85] Kerckhoff S.P. – Simplicial systems for interval exchange maps and measured foliations. *Ergodic Th. and Dynamical Sys.*, vol. 5, 1985, pp. 257–271.

[LM95] Lind D. and Marcus B. – *Symbolic Dynamics and Coding.* – Cambridge Univ Press, 1995.

[LN95] Lopez L-M. and Narbel Ph. – Generalized Sturmian languages. *In : Intl. Conf. in Automata Language and Programming'95.* pp. 336–347. – Szeged, Hungary, 1995.

[LN98] Lopez L.-M. and Narbel Ph. – D0L-systems and surface automorphisms. *In : Math. Foundations in Computer Science.* pp. 522–532. – Springer-Verlag.

[Lot97] Lothaire M. – *Combinatorics on words. 2nd ed.* – Cambridge University Press, 1997, *Encyclopedia of Mathematics and Its Applications.*

[Mañ83] Mañé R. – *Ergodic Theory and Differentiable Dynamics.* – Springer-Verlag, 1983.

[Mar82] Markov A.A. – Sur une question de Jean Bernoulli. *Math. Ann.*, vol. 19, 1882, pp. 27–36.

[Mas67] Massey W. S. – *Algebraic Topology: An Introduction.* – Springer-Verlag, 1967.

[Mas82] Masur H. – Interval exchange transformations and measured foliations. *Ann. of Math.*, vol. 115, 1982, pp. 169–200.

[MH40] Morse M. and Hedlund G.A. – Symbolic dynamics II. Sturmian trajectories. *American Journal of Mathematics*, vol. 62, 1940, pp. 1–42.

[NR97] Nogueira A. and Rudolph D. – Topological weak-mixing of interval exchange maps. *Ergod. Th. and Dynam. Sys*, vol. 17, 1997, pp. 1183–1209.

[Qué87] Quéffelec M. – *Substitution Dynamical Systems. Spectral Analysis.* – Springer-Verlag, 1987, *Lecture Notes in Mathematics.*

[Rau77] Rauzy, G. – Une généralisation du développement en fraction continue. *Semin. Delange-Pisot-Poitou, 18e Année*, vol. 1, 1977, pp. 1501–1515.

[Rau79] Rauzy, G. – Échanges d'intervalles et transformations induites. *Acta Arith.*, vol. 34, 1979, pp. 315–328.

[RS80] Rozenberg G. and Salomaa A. – *The mathematical theory of L systems.* – Academic press, 1980.

[RS97] Rozenberg G. and Salomaa A. – *Handbook of Formal Languages (Vol. 1) Word Languages Grammar.* – Berlin, Springer Verlag, 1997.

[Sze70] Szekeres G. – Multidimensional continued fractions. *Ann. Univ. Sci. Budapest. Rolando Eoetvoes, Sect. Math.*, vol. 13, 1970, pp. 113–140.

[Vee78] Veech W. – Interval exchange transformations. *J. Anal. Math.*, vol. 33, 1978, pp. 222–272.

[Vee82] Veech W. – Gauss measures for transformations on the space of interval exchange maps. *Ann. of Math.*, vol. 115, 1982, pp. 201–242.

[Ven70] Venkov B.A. – *Elementary Number Theory.* – Wolters-Noordhoff, Groningen, 1970.

[Zor96] Zorich A. – Finite Gauss measure on the space of interval exchange transformations and Lyapunov exponents. *Ann. Inst. Fourier*, vol. 46, number 2, 1996, pp. 325–370.