

# On the number of return words in infinite words with complexity $2n + 1$

Laurent Vuillon\*

**Abstract:** *In this article, we count the number of return words in some infinite words with complexity  $2n + 1$ . We also consider some infinite words given by codings of rotations and interval exchange transformations on  $k$  intervals. We prove that the number of return words over a given word  $w$  for these infinite words is exactly  $k$ .*

*Keywords:* return words, symbolic dynamical systems, Sturmian words, interval exchange transformations, combinatorics on words.

## 1 Introduction

The starting point of this article is the following question: *Does there exist a way to characterize by return words the infinite words with complexity  $2n + 1$  ?* We first give the definition of return words for a recurrent infinite word in a finite alphabet (a recurrent word is an infinite word such that each factor appears infinitely many times). Considering each occurrence of a word  $w$  in a recurrent infinite word  $U$ , we define the set of return words over  $w$  to be the set of all distinct words beginning with an occurrence of  $w$  and ending exactly before the next occurrence of  $w$  in the infinite sequence. This mathematical tool was introduced independently by Durand, Holton and Zamboni in order to study primitive substitutive sequences (see [10, 11, 16]). This notion is quite natural and can be seen as a symbolic version of the first return map for a dynamical system. Recently, many developments of the notion of return words have been given. For example, Allouche, Davison, Queffelec and Zamboni study the transcendence of Sturmian or morphic

---

\*LIAFA, Université Paris 7, case 7014, 2 place Jussieu, F-75251 Paris Cedex 05, France, vuillon@liafa.jussieu.fr

continued fractions and the main tool is to show, using return words, that arbitrary long prefix are “almost squares” (see [2]). We can also use return words to study low complexity infinite words. For example, the author shows that an infinite word is Sturmian if and only if for each word appearing in the infinite word, the cardinality of the set of return words over  $w$  is exactly two (see [21]). Recall that Sturmian infinite words are aperiodic words with complexity  $p(n) = n + 1$  for all  $n$  (the complexity function  $p(n)$  counts the number of distinct factors of length  $n$  in the infinite word) (see [5, 4, 15]). Fagnot and Vuillon show a generalization of the notion of balanced property for Sturmian words and the proof is based on return words and combinatorics on words (see [13]). Cassaigne uses this tool to investigate a conjecture of Rauzy (see [7]).

Our purpose is to compute the set of return words for a class of infinite words with complexity  $p(n) = 2n + 1$  for all  $n$  (see [3, 6, 8, 9, 12, 14, 17]).

The following question is thus natural: is it true that for these infinite words, the cardinality of the set of return words over a factor  $w$  is always three ?

In this direction, Justin and Vuillon [17] show that Arnoux-Rauzy infinite words [3], which are infinite words with complexity  $2n + 1$  (see [3, 8, 9, 12]), have property  $R_3$ . In the sequel, we say that an infinite word has the property  $R_k$  if the cardinality of the set of return words over each factor  $w$  is exactly  $k$ . They give also the structure of return words in the context of Episturmian words [12, 17].

In the opposite, the work of Ferenczi [14] presents a nice substitutive infinite word with complexity  $p(n) = 2n + 1$  given by the Chacon transformation. It is not difficult to show that the cardinality of the set of return words may be larger than three for this infinite word. This is the first example of an infinite word with complexity  $2n+1$  without the property  $R_3$ .

Nevertheless, we find two other classes of infinite words with complexity  $p(n) = 2n+1$  having the property  $R_3$ . The first one is the coding of a rotation  $\alpha$  in the unit circle with three intervals (with rationally independent lengths) and one of these intervals of length  $\alpha$ . The second one is a generalization of this class namely the interval exchange transformation on three intervals.

The structure of the article is the following. First, we recall some basic definitions in combinatorics of words. Secondly, we present the infinite words given by codings of rotation. Then, we show that an infinite word given by a regular interval exchange with three intervals (resp.  $k$  intervals) have the property  $R_3$  (resp.  $R_k$ ). We compute also the length of each return word over a given finite word  $w$ . Lastly, we show that the coding of a rotation by  $\alpha$  in the unit circle with three (resp.  $k$  intervals) intervals with rationally

independent lengths and with one interval of length  $\alpha$  as property  $R_3$  (resp.  $R_k$ )

More generally, we guess that the class of infinite word with complexity  $p(n) = 2n + 1$  and exactly three return words over each factor is the class of infinite words given by a “self-induced discrete dynamical system”. In particular, we may extend these results to bounded remainder sets introduced by Rauzy and Ferenczi.

## 2 Basic definitions and examples

Let  $\mathcal{A} = \{0, 1\}$  be a binary alphabet. We denote by  $\mathcal{A}^*$  the set of finite words on  $\mathcal{A}$  and by  $\mathcal{A}^{\mathbb{N}}$  the one-sided infinite word  $U$ . A word  $w$  is a *factor* of a word  $x \in \mathcal{A}^*$  if there exist some words  $u, v \in \mathcal{A}^*$  such that  $x = u w v$ . An infinite word  $U$  is called *recurrent* if every factor of  $U$  appears infinitely many times in  $U$ . For a finite word  $w = w_1 w_2 \cdots w_n$ , the length of  $w$  is denoted by  $|w|$  and is equal to  $n$ . The set of factors of  $U$  with length  $n$  is denoted by  $L_n(U)$ . The language  $L(U) = \cup_n L_n(U)$  is the set of factors of  $U$ . For two finite words  $w$  and  $u$ , the number of occurrences of  $w$  on  $u$  is denoted by  $|u|_w$  and  $|u|_w = \{i \mid 1 \leq i \leq |u| - |w| + 1 \text{ s.t. } u_i u_{i+1} \cdots u_{i+|w|-1} = w_1 w_2 \cdots w_{|w}|\}$ .

The *position set* of the word  $w$  is the ordered sequence  $i(U, w) = (i_1, i_2, \dots, i_k, \dots)$  where  $i_k$  represents the position of the first letter of the  $k$ -th occurrence of the word  $w$  in the infinite word  $U$ . In a more formal way,  $i_k \in i(U, w)$  if and only if  $U_{i_k} U_{i_k+1} \cdots U_{i_k+|w|-1} = w$  and  $|U_1 \cdots U_{i_k+|w|-1}|_w = k$ . Since the infinite word  $U$  is recurrent the set  $i(U, w)$  is infinite. For a recurrent word  $U$ , the set of *return words over  $w$*  is the set (denoted by  $\mathcal{H}_{U,w}$ ) of all distinct words with the following form:

$$U_{i_k} U_{i_k+1} \cdots U_{i_{k+1}-1}$$

for all  $k \in \mathbb{N}, k > 0$ . This definition is best understood on an example. Let  $U_1 = (0100100001)^\omega$  be an infinite word on the alphabet  $\mathcal{A}$ . By definition, the set of return words over 01 is  $\mathcal{H}_{U_1,01} = \{010, 01000, 01\}$ . Indeed, the infinite word  $U_1$  can be written

$$(\underline{0}1\underline{0}0\underline{1}000\underline{0}1\underline{0}1\underline{0}0\underline{1}000\underline{0}1)^\omega$$

where  $\underline{\quad}$  denotes the position of the first letter for each occurrence of the word 01. We note that between two consecutive occurrences of 01, we have the three possible words 0, 000,  $\epsilon$ . Thus 010, 01000 and 01 are elements of  $\mathcal{H}_{U_1,01}$ . We say that an infinite word have the property  $R_k$  if the cardinality of the

set of return words over each factor  $w$  is exactly  $k$  (i.e.  $\text{Card}(\mathcal{H}_{U,w}) = k \forall w \in L(U)$ ).

The *complexity function*  $p : \mathbb{N} \rightarrow \mathbb{N}$  of an infinite word  $U$  counts the number of distinct factors of  $U$  of given length:

$$p(n) = \text{Card}L_n(U).$$

### 3 A negative result

Let us consider the following substitution extensively studied by Ferenczi (see [14]) which is a recoding of the Chacon substitution:

$$\sigma(1) = 12,$$

$$\sigma(2) = 312,$$

$$\sigma(3) = 3312.$$

The fixed point  $\sigma(x) = x$  of the substitution begins by

$$x = 123123312123123312331212312123123312123123312331233 \dots$$

It is easy to check with few terms of the fixed point  $x$  that the number of return words over  $23$  is upper than  $3$ , indeed:

$$\mathcal{H}_{23} = \{231, 2331, 23121, 233121\}.$$

### 4 Codings of rotation

The aim of this section is to introduce codings of irrational rotation on the unit circle. For  $p \geq 2$ , let  $F = \{\beta_0 < \beta_1 < \dots < \beta_{p-1}\}$  be a set of  $p$  consecutive points of the unit circle (identified in all that follows with  $[0, 1[$  or with the unidimensional torus  $\mathbb{R}/\mathbb{Z}$ ) and let  $\beta_p = \beta_0$ . Let  $\alpha$  be an irrational number in  $]0, 1[$  and let us consider the *positive orbit* of a point  $x$  of the unit circle under the rotation by angle  $\alpha$ , i.e., the set of points  $\{\{\alpha n + x\}, n \in \mathbb{N}\}$ .

The *coding* of the orbit of  $x$  under the rotation by angle  $\alpha$  with respect to the partition  $\{[\beta_0, \beta_1[, [\beta_1, \beta_2[, \dots, [\beta_{p-1}, \beta_p[$  is the infinite word  $u$  defined on the finite alphabet  $\Sigma = \{0, \dots, p-1\}$  as follows:

$$u_n = k \Leftrightarrow \{x + n\alpha\} \in [\beta_k, \beta_{k+1}[, \text{ for } 0 \leq k \leq p-1.$$

A coding of the rotation  $R$  means the coding of the orbit of a point  $x$  of the unit circle under the rotation  $R$  with respect to a finite partition of the unit circle consisting of left-closed and right-open intervals.

#### 4.1 Factors

With the above notation, consider a coding  $u$  of the orbit of a point  $x$  under the rotation by angle  $\alpha$  with respect to the partition  $\{[\beta_0, \beta_1[, [\beta_1, \beta_2[, \dots, [\beta_{p-1}, \beta_p[ \}$ . Let  $I_k = [\beta_k, \beta_{k+1}[$  and let  $R$  denotes the rotation by angle  $\alpha$ .

**Lemma 1**  $w_1 \cdots w_n$  defined on the alphabet  $\Sigma = \{0, 1, \dots, p-1\}$  is a factor of the infinite word  $u$  if and only if  $I(w_1, \dots, w_n) = \bigcap_{j=0}^{n-1} R^{-j}(I_{w_{j+1}}) \neq \emptyset$ .

**Proof** A finite word  $w_1 \cdots w_n$  defined on the alphabet  $\Sigma = \{0, 1, \dots, p-1\}$  is a factor of the infinite word  $u$  if and only if there exists an integer  $k$  such that

$$\{x + k\alpha\} \in I(w_1, \dots, w_n) = \bigcap_{j=0}^{n-1} R^{-j}(I_{w_{j+1}}).$$

As  $\alpha$  is irrational, the sequence  $(\{x + n\alpha\})_{n \in \mathbb{N}}$  is dense in the unit circle, which implies that  $w_1 w_2 \dots w_n$  is a factor of  $u$  if and only if  $I(w_1, \dots, w_n) \neq \emptyset$ . In particular, the set of factors of a coding does not depend on the initial point  $x$  of this coding. ■

## 5 Interval exchange transformations

This section deals with interval exchange transformations. This object is a natural generalization of the codings of rotation  $\alpha$  where  $\alpha$  is the length of one of the intervals. An interval exchange transformation is a piecewise affine transformation which maps a partition into intervals of the space to another partition according to a permutation of the intervals. This transformation could be more complicated than a rotation. Indeed, irrational rotations are uniquely ergodic, whereas a class of interval exchange transformations is non uniquely ergodic. In the sequel, we use the notations of Keane and Rauzy (see [18, 19])

An interval exchange of  $k$  intervals is defined by a vector  $(\lambda_1, \lambda_2, \dots, \lambda_k)$  in  $\mathbb{R}^k$  with strictly positive coordinates and  $\sum_{i=1}^k \lambda_i = 1$ , and a permutation  $\sigma$  of the set  $\{1, 2, \dots, k\}$ . We set the partition of  $[0, 1[$  in  $k$  intervals

$$X_i = \left[ \sum_{j < i} \lambda_j, \sum_{j \leq i} \lambda_j \right] \text{ for } 1 \leq i \leq k$$

( $X_i$  has the length  $\lambda_i$ ).

The interval exchange transformation associated to the ordered pair  $(\sigma, \lambda)$  is the transformation from  $[0, 1[$  to itself, defined as a piecewise affine transformation which maps the partition  $(X_1, X_2, \dots, X_k)$  to the partition  $(X_{\sigma(1)}, X_{\sigma(2)}, \dots, X_{\sigma(k)})$ .

The transformation  $T$  maps the point  $x \in X_i$  to the point

$$T(x) = x + r_i$$

where

$$r_i = \sum_{k < \sigma^{-1}(i)} \lambda_{\sigma(k)} - \sum_{k < i} \lambda_k.$$

Now, we construct an infinite word with values in the alphabet  $\{1, 2, \dots, k\}$  associated with a couple  $(\sigma, \lambda)$  by coding the positive orbit of a point  $x$  under the transformation  $T$  according to the partition  $(X_1, X_2, \dots, X_k)$ . The infinite word  $U(x)$  is given by

$$U(x)_n = \mathcal{I}(T^n(x)), \quad \forall n \in \mathbb{N}$$

where  $\mathcal{I}(y) = i$  if  $y \in X_i$ .

In order to define a transformation such that each orbit is dense in  $[0, 1[$ , we add the following property: an interval exchange transformation is called regular if for all  $0 = a_1 < a_2 < \dots < a_{k+1} = 1$  of the intervals  $X_i = [a_i, a_{i+1}[$  with  $i \in \{1, \dots, k\}$ , we have

$$T^n a_i = a_j, \quad i \text{ and } j \in \{2, 3, \dots, k\}, n \in \mathbb{Z}$$

implies  $n = 0, i = j$ .

M. Keane shows the following result [18]:

**Theorem 1** *An interval exchange transformation  $T$  is regular if and only if for every point  $x$  in  $[0, 1[$  the orbit  $\cup_{n \in \mathbb{Z}} \{T^n x\}$  of  $x$  by  $T$  is dense in  $[0, 1[$ .*

A necessary condition to have a regular interval exchange transformation is to take an irreducible permutation. A permutation is called irreducible if  $\sigma(\{1, 2, \dots, \ell\}) \neq \{1, 2, \dots, \ell\}$  for all  $\ell < k$ .

Now we can state the main theorem of this section:

**Theorem 2** *Let a regular interval exchange transformation  $T$  with  $k$  intervals. The infinite word  $U(x)$  associated with  $T$  has property  $R_k$ .*

In order to show this theorem, we first present a method to construct all the words of length  $n$  factors of the infinite word  $U(x)$  and secondly a way to construct return words by using self-induction.

## 5.1 Factors

The construction of all the factors is not so far from the one used for rotations on the unit circle. The main tool is to consider the negative orbit of all the endpoints  $0 = a_1 < a_2 < \dots < a_{k+1} = 1$  of the intervals  $X_i = [a_i, a_{i+1}[$  with  $i \in \{1, \dots, k\}$ . The set of endpoints  $\{a_i \mid 1 \leq i \leq k+1\}$  is called  $X^{(1)}$ .

A word  $w$  of length  $n$  is a factor of the infinite word  $U(x)$  if and only if there exists an interval  $I_w \subset [0, 1[$  and a point  $y$  in  $I_w$  such that the word  $w$  is the following concatenation of letters:

$$\mathcal{I}(x)\mathcal{I}(T(x)) \cdots \mathcal{I}(T^{n-1}(x)) = w.$$

Using this fact, the number of factors of length one is exactly the number of intervals  $I_1 = X_1, I_2 = X_2, \dots, I_k = X_k$  associated with the letters of the alphabet  $\{1, 2, \dots, k\}$ .

**Proposition 1 (Arnoux, Rauzy [3])** *Let a regular interval exchange transformation  $T$  with  $k$  intervals. The infinite word  $U(x)$  associated with  $T$  has complexity  $p(n) = n(k-1) + 1, \forall n \in \mathbb{N}$ .*

**Proof** The factors  $w = w_1 w_2$  of length 2 are given by the set  $\{\mathcal{I}(x)\mathcal{I}(T(x)), x \in [0, 1[$  or equivalently by the intervals  $I_{w_1 w_2} = I_{w_1} \cap T^{-1}I_{w_2}$ .

As the transformation  $T$  is a piecewise affine transformation, it is sufficient to find all the endpoints of the intervals  $I_w$  with  $|w| = 2$ . That is the positions of the endpoints of all intervals with form  $I_{w_1} \cap T^{-1}I_{w_2}$ . These endpoints are given by the ordered set

$$X^{(2)} = \{0 = b_1 < b_2 < \dots < b_{2k} = 1\} = X^{(1)} \cup T^{-1}X^{(1)}.$$

Remark that the number of points in the union is  $2k$  because the points 0 and 1 are elements of the intersection of the two sets  $X^{(1)}$  and  $T^{-1}X^{(1)}$ .

The next step is to prove that each  $I_w$  is connected. By definition, all the intervals associated with words of length 1 are connected. The interval  $I_{m_1 m_2} = I_{m_1} \cap T^{-1}I_{m_2}$  is connected because it is the intersection of two intervals in  $[0, 1[$ .

In other words, for each word  $w$  with  $|w| = 2$  factor of the infinite word  $U(x)$ , there exists an interval  $I_w = [b_\ell, b_{\ell+1}[$  where  $b_\ell$  and  $b_{\ell+1}$  are two consecutive points in the ordered set  $X^{(2)}$ .

By induction, let  $w = w_1 w_2 \cdots w_n$  a factor of length  $n$ , then there exists a point  $x$  such that

$$w = \mathcal{I}(x)\mathcal{I}(T(x)) \cdots \mathcal{I}(T^{n-1}(x))$$

or equivalently

$$I_{w_1 w_2 \cdots w_n} = \bigcap_{i=0 \dots n-1} T^{-i} I_{w_{i+1}}.$$

That is each endpoints of the interval associated with a word of length  $n$  is an element of the ordered set

$$X^{(n)} = \{0 = b_1 < b_2 < \cdots < b_{n(k-1)+2} = 1\} = X^{(1)} \cup T^{-1} X^{(1)} \cup \cdots \cup T^{-n+1} X^{(1)}.$$

Remark that the number of points in the union is  $n(k-1) + 2$  because at each step, we add  $k-1$  new points and for  $n=1$  we have  $k+1$  points in the partition.

Furthermore, the intervals associated with factors of length  $n$  are connected. Indeed, the intervals are the intersection of connected intervals in  $[0, 1[$ .

In other words, for each word  $w$  with  $|w| = n$  factor of the infinite word  $U(x)$ , there exists an interval  $I_w = [b_\ell, b_{\ell+1}[$  where  $b_\ell$  and  $b_{\ell+1}$  are two consecutive points in the ordered set  $X^{(n)}$ .

With this construction, we find  $n(k-1) + 2$  points for the endpoints in the partition  $X^{(n)}$ . We have  $n(k-1) + 1$  intervals in the partition and then the complexity function for a regular interval exchange in  $k$  intervals is equal to  $p(n, k) = n(k-1) + 1, \forall n \in \mathbb{N}$ . ■

Thus an infinite word associated with a regular interval exchange with two intervals is nothing but a Sturmian infinite word (i.e. an infinite word with complexity  $p(n, 2) = n + 1$  for all  $n$ ). And an infinite word associated with a regular interval exchange with three intervals is an infinite word with complexity  $p(n, 3) = 2n + 1$  for all  $n$ .

## 5.2 self-induction

Now, we focus on the construction of return words associated with a word  $w$  factor of the infinite word.

*Theorem 2 Let a regular interval exchange transformation  $T$  with  $k$  intervals. The infinite word  $U(x)$  associated with  $T$  has property  $R_k$ .*

By the previous construction, we find a unique connected interval  $I_w$  associated with the word  $w$ . The main tool is to study the first return map in the adherence of the interval  $I_w$ . This method is used by Rauzy to give a generalized continued fraction algorithm [19, 20].

**Proof** Let  $r(y) = \inf\{t \geq 0 | T^{-t}y \in \overline{I_w}\}$  be the negative first return time in the interval  $\overline{I_w}$ . Keane shows that the points  $\{T^{-r(y)}y | y \in \overline{I_w}, r(y) < \infty\}$  give a partition of the interval  $\overline{I_w}$  in exactly  $k$  intervals  $I_{p_1}, I_{p_2}, \dots, I_{p_k}$  (see [18]). Indeed, the endpoints of these intervals are given by the first time that the negative orbit of the points in  $X^{(1)} = \{0 = a_1 < a_2 < \dots < a_{k+1} = 1\}$  falls in  $\overline{I_w}$ .

There are  $k+1$  points in  $X^{(1)}$ . This shows that for a general interval  $[\alpha, \beta]$ , the number of induced points in  $] \alpha, \beta[$  is  $k+1$  and that the number of induced intervals is  $k+2$ . As the endpoints of  $\overline{I_w}$  are both in the negative orbit of two different points in  $X^{(1)}$  and as the interval exchange transformation is regular, we have

$$T^n a_i = a_j, \quad i \text{ and } j \in \{2, 3, \dots, k\}, n \in \mathbb{Z}$$

implies  $n = 0, i = j$ . Thus

$$\alpha = T^{n_{a_0}} a_0, \beta = T^{n_{a_{k+1}}} a_{k+1}.$$

That is the number of induced points in the interior of  $\overline{I_w}$  is  $k-1$ . Thus the number of induced intervals is  $k$ . Such intervals with endpoints on the negative orbit of  $a_0$  and  $a_{k+1}$  are called acceptable intervals by Rauzy (see [19]). Indeed, for interval exchange transformation with  $k$  intervals, the induced transformation on an acceptable intervals is also an interval exchange transformation with  $k$  intervals.

By construction, for each  $p_i$  there exist  $y \in I_w$  and  $t$  such that the word  $p_i = \mathcal{I}(y)\mathcal{I}(T(y)) \dots \mathcal{I}(T^t(y)) \dots \mathcal{I}(T^{t+|w|-1}(y))$ . Furthermore, the prefix of length  $|w|$  of  $p_i$  is exactly  $w$  (indeed  $y \in I_w$ ) and the suffix of length  $|w|$  of  $p_i$  is exactly  $w$  (indeed  $t$  is the positive first return time of  $y$  in the interval  $I_w$  defined by  $r^+(y) = \inf\{t \geq 0 | T^t y \in \overline{I_w}\}$ ). In other words, the number of return words over  $w$  is exactly the number of induced intervals. By construction  $I_w$  is an acceptable interval. Thus the number of return words over  $w$  is exactly  $k$ . That is, the infinite word associated with regular interval exchange transformation on  $k$  intervals has property  $R_k$ . ■

**Corollary 1** *The length of the  $i$ -th return word of  $w$  associated with the  $i$ -th induced interval  $p_i$  is  $k_i + k'_i$  where  $k_i$  is the given by the smallest  $k_i \geq 0$  such that  $T^{-k_i} \ell_i \in I_w$ , where  $k_i$  is the given by the smallest  $k'_i > 0$  such that  $T^{k'_i} \ell_i \in I_w$  and where  $\ell_i$  is the left endpoint of  $p_i$ .*

**Proof** By construction, the left endpoint of  $p_1$  is  $\alpha$ . Each orbit of the points in  $p_1$  according to the induce transformation is a translation of the

orbit of  $\alpha$ . Then the length of the return word associated with  $p_1$  is  $k'_1$  where  $k'_1$  is the smallest  $k'_1 > 0$  such that  $T^{k'_1}\alpha \in I_w$ . In other words, the length of the return word associated with the interval  $p_1$  is exactly the sum of the positive first return time  $k'_1$  of the point  $\alpha$  in  $I_w$  and the length of  $w$ .

By construction, each interval  $p_i$  for  $i = 1, \dots, k$  has  $T^{-k_i}\ell_i$  for left endpoint. Furthermore, each orbit of the points in  $p_i$  according to the induce transformation is a translation of the orbit of  $T^{-k_i}\ell_i$ . Then the length of the return word associated with  $p_i$  is  $k_i + k'_i$  where  $k_i$  is the smallest  $k_i > 0$  such that  $T^{-k_i}\ell_i \in I_w$  and where  $k'_i$  is the smallest  $k'_i > 0$  such that  $T^{k'_i}\ell_i \in I_w$ . In other words, the length of the return word associated with the interval  $p_i$  is exactly given by the sum of the first time that  $T^{-k_i}\ell_i$  is mapped on the point  $\ell_i$  ( $\ell_i$  is element of the partition  $X^{(1)}$ ) and the positive first return time  $k'_i$  of the point  $\ell_i$  in  $I_w$  and the length of  $w$ . ■

## 6 Return words for codings of rotation

As the codings of rotation are a special case of interval exchange transformations, we have the following corollary:

**Corollary 2** *Let  $T$  be a coding of rotations on  $k$  intervals with rationally independent lengths and by a rotation of length  $\alpha$  equal to the length of one of the intervals. The infinite word  $U(x)$  associated with  $T$  has property  $R_k$ .*

**Proof** As the rotation is defined on the circle, we can translate all the intervals in order to have the first interval with length equal to  $\alpha$ . The coding of rotation of angle  $\alpha$  on  $k$  intervals with the first interval of length  $\alpha$  is equivalent to an interval exchange transformation with either the permutation  $(23 \dots k1)$  if  $\alpha > 0$  or the permutation  $(k123 \dots k-1)$  if  $\alpha < 0$ . As the lengths of the intervals are rationally independent in the coding of rotation then the associated interval exchange transformation is regular. Consequently, the infinite word  $U(x)$  associated with  $T$  has property  $R_k$ . ■

## References

- [1] P. Alessandri, V. Berthé *Three distance theorems and combinatorics on words*, Enseig. Math. **44** (1998), 103–132.
- [2] J.-P. Allouche, J. L. Davison, M. Queffélec and L. Q. Zamboni *Transcendence of Sturmian or morphic continued fractions*, preprint 1999.

- [3] P. Arnoux et G. Rauzy, *Représentation géométrique de suites de complexité  $2n+1$* , Bull. Soc. Math. France. **119** (1991), 199–215.
- [4] J. Berstel and P. Séébold , *Sturmian words*, In M. Lothaire, Algebraic combinatorics on Words. (2000). To appear.
- [5] V. Berthé, *Fréquences des facteurs des suites sturmiennes*, Theoret. Comput. Sci. **165** (1996), 295–309.
- [6] V. Berthé and L. Vuillon, *A two-dimensional generalization of Sturmian sequences : tilings and rotations*, to appear in Discr. Math..
- [7] J. Cassaigne, *Ideas for a proof of Rauzy’s conjecture on the recurrence functions of infinite words*, talk in Rouen “words 1999”.
- [8] J. Cassaigne, S. Ferenczi and L.Q. Zamboni *Imbalances in Arnoux-Rauzy Sequences*, Preprint IML 1999.
- [9] M.G. Castelli, F. Mignosi and A. Restivo, *Fine and Wilf’s theorem for three periods and a generalization of Sturmian words*, Theoret. Comp. Sci., to appear.
- [10] F. Durand, *A characterization of substitutive sequences using return words*, Discrete Math. **179** (1998), 89–101.
- [11] F. Durand, *Contributions à l’étude des suites et systèmes dynamiques substitutifs*, Ph.D. Thesis, Université de la Méditerranée (Aix-Marseille II) 1996.
- [12] X. Droubay, J. Justin and G. Pirillo, *Episturmian words and some constructions of de Luca and Rauzy*, Theoret. Comp. Sci., to appear.
- [13] I. Fagnot and L. Vuillon *Generalized balances in Sturmian words* Preprint 02/2000 LIAFA.
- [14] S. Ferenczi *Les transformations de Chacon: combinatoire, structure géométrique, lien avec les systèmes de complexité  $2n+1$* , Bull. Soc. Math. France, **123** n0. 2 (1995), 271–292.
- [15] G. A. Hedlund, Morse *Symbolic dynamics II: Sturmian trajectories*, Amer. J. Math. **62** (1940), 1–42.
- [16] C. Holton, L. Q. Zamboni *Descendants of primitive substitutions*, Theory Comput. Systems **32**, (1999) 133–157.

- [17] J. Justin and L. Vuillon *Return words in Sturmian and episturmian words*, Preprint 10/2000 LIAFA.
- [18] M. Keane *Interval exchange transformations* Math. Zeit. **141** (1975), 25–31.
- [19] G. Rauzy, *Echange d'intervalles et transformations induites*, Acta. Arith. **34** (1979), 315–328.
- [20] G. Rauzy, *Une généralisation du développement en fraction continue*, Semin. Delange-Pisot-Poitou, 18e Année **1** (1977), 1501–1515.
- [21] L. Vuillon *A characterization of Sturmian words by return words*, to appear in European. J. of Combin..