Palindromic complexity of codings of rotations

Alexandre Blondin Masse, Srečko Brlek, Sébastien Labbé, Laurent Vuillon

Palindromic complexity of codings of rotations

A. Blondin Massé[a,b,*], S. Brlek[a], S. Labbé[a,c], L. Vuillon[b]

[a] *LaCIM, Université du Québec à Montréal,*
C. P. 8888 Succursale "Centre-Ville", Montréal (QC), CANADA H3C 3P8
[b] *Laboratoire de mathématiques, CNRS UMR 5127, Université de Savoie, 73376 Le Bourget-du-lac cedex, France*
[c] *LIRMM, UMR 5506 CNRS, Université Montpellier II, 34392 Montpellier, France*

Abstract

We study the palindromic complexity of infinite words obtained by coding rotations on partitions of the unit circle by inspecting the return words. The main result is that every coding of rotations on two intervals is full, that is, it realizes the maximal palindromic complexity. As a byproduct, a slight improvement about return words in codings of rotations is obtained: every factor of a coding of rotations on two intervals has at most 4 complete return words, where the bound is realized only for a finite number of factors. We also provide a combinatorial proof for the special case of complementary-symmetric Rote sequences by considering both palindromes and antipalindromes occurring in it.

Keywords: Codings of rotations, Sturmian, Rote, return words, full words.

1. Introduction

A coding of rotations is a symbolic sequence obtained from iterative rotations of a point x by an angle α and according to a partition of the unit circle [1]. When the partition consists of two intervals, the resulting coding is a binary sequence. In particular, it yields the famous Sturmian sequences if the size of one interval is exactly α with α irrational [2]. Otherwise, the coding is a Rote sequence if the length of the intervals are rationally independent of α [3] and quasi-Sturmian in the other case [4]. Numerous properties of these sequences have been established regarding subword complexity [1], continued fractions and combinatorics on words [4], or discrepancy and substitutions [5].

The palindromic complexity $|\mathrm{Pal}(w)|$, i.e. the number of distinct palindrome factors, of a finite word w is bounded by $|w| + 1$ (Droubay et al. 2001 [6]) and w is called *full* (Brlek et al. 2004 [7], or *rich* in Glen et al. 2009 [8]) if it realizes that upper bound. Naturally, an infinite word is said to be *full* if all its finite

*Corresponding author.
Email addresses: `alexandre.blondin.masse@gmail.com` (A. Blondin Massé), `Brlek.Srecko@uqam.ca` (S. Brlek), `slabqc@gmail.com` (S. Labbé), `Laurent.Vuillon@univ-savoie.fr` (L. Vuillon)

Preprint submitted to Elsevier *August 24, 2011*

factors are full. The case of periodic full words was completely characterized in [7]. On the other hand, Sturmian words, which are particular cases of coding of rotations, and even episturmian words are full [6]: this result is obtained by showing that the longest palindromic suffix of every nonempty prefix is unioccurrent; rephrasing this property in terms of return words, one has that a word w is full if and only if each complete return word of every palindrome in w is a palindrome [8].

Our main result is Theorem 19 stating that every word generated by codings of rotations is full. To achieve this, we start with a thorough study of partitions of the unit circle into sets I_w according to some trajectories under rotations labeled by w. The proof is based on two cases, whether I_w is an interval or not: Proposition 11 and 16 handle those cases. We use the property that each factor of a coding of rotations on two intervals has at most 4 complete return words, where the bound is realized only for a finite number of factors w, those such that I_w is not an interval and w is some power of a letter.

Moreover, these partitions show some remarkable geometrical symmetries that are useful for handling return words: in particular, if the trajectory of a point is symmetric with respect to some global axis, then the coding of rotations from this point is a palindrome. When the first return function is a bijection then it coincides with some interval exchange transformation, a very useful fact for proving our claim. A direct consequence of our study on return words is that every coding of rotations on two intervals is full.

The paper is divided into four parts. First the basic terminology is introduced, notation and tools relative to combinatorics on words, the unit circle, interval exchange transformations, Poincaré's first return function and codings of rotations. In particular, some conditions for the Poincaré's first return function to be a bijection and consequently an interval exchange transformation are stated. Section 3 contains results about partitions of the unit circle induced by codings of rotations. Section 4 is devoted to the statement and proof of the main result. In Section 5, we provide an alternative proof of the fact that complementary-symmetric words (i.e. words with complexity $f(n) = 2n$ whose language is closed under swapping of letters) are also full by considering both palindromes and antipalindromes occurring in it.

2. Preliminaries

The basic terminology about words is borrowed from M. Lothaire [9]. In what follows, Σ is a finite *alphabet* whose elements are called *letters*. A *word* is a finite sequence of letters $w : [0..n-1] \longrightarrow \Sigma$, where $n \in \mathbb{N}^+ = \mathbb{N} \setminus \{0\}$. The length of w is $|w| = n$ and w_i denote its i-th letter. If k and ℓ are two nonnegative integers, then $w_{[k,\ell]}$ denotes the word $w_k w_{k+1} \cdots w_\ell$. The set of n-length words over Σ is denoted Σ^n, and that of infinite words is Σ^ω. By convention, the *empty* word ε is the unique word of length 0. The free monoid generated by Σ is defined by $\Sigma^* = \bigcup_{n\geq 0} \Sigma^n$, and $\Sigma^\infty = \Sigma^\omega \cup \Sigma^*$. Given a word $w \in \Sigma^\infty$, a *factor* u of w is a word $u \in \Sigma^*$ such that $w = xuy$, for some $x \in \Sigma^*$, and $y \in \Sigma^\infty$. If $x = \varepsilon$ (resp. $y = \varepsilon$) then u is called a *prefix* (resp.

suffix). The set of all factors of w is denoted by $\mathrm{Fact}(w)$, those of length n is $\mathrm{Fact}_n(w) = \mathrm{Fact}(w) \cap \Sigma^n$, and $\mathrm{Pref}(w)$ is the set of all prefixes of w. If $w = pu$, with $|w| = n$ and $|p| = k$, then $p^{-1}w = w_{[k,n-1]} = u$ is the word obtained by erasing the prefix p from w. An *occurrence* of u in w is a position k such that $u = w_{[k,k+|u|-1]}$, the set of all its occurrences is $\mathrm{Occ}(u,w)$. The number of occurrences of u in w is denoted by $|w|_u$. An infinite word is *periodic* if there exists a positive integer p such that $w[i] = w[i+p]$, for all i. An infinite word w is *recurrent* if every factor u of w satisfies $|w|_u = \infty$.

The *reversal* of $u = u_1u_2\cdots u_n \in \Sigma^n$ is the word $\widetilde{u} = u_nu_{n-1}\cdots u_1$. A *palindrome* is a word p such that $p = \widetilde{p}$. Every word contains palindromes, the letters and ε being necessarily part of them. For a language $L \subseteq \Sigma^\infty$, the set of its palindromic factors is denoted by $\mathrm{Pal}(L)$. Obviously, the palindromic language is closed under reversal, since $\mathrm{Pal}(L) = \mathrm{Pal}(\widetilde{L})$.

Let w be a word, and $u, v \in \mathrm{Fact}(w)$. Then v is a *return word* of u in w if $vu \in \mathrm{Fact}(w)$, $u \in \mathrm{Pref}(vu)$ and $|vu|_u = 2$. Moreover, vu is a *complete return word* of u in w. The set of complete return words of u in w is denoted $\mathrm{CRet}_w(u)$. A natural generalization of complete return words consists in allowing the source word to be different from the target word. Let $u, v \in \mathrm{Fact}(w)$. Then $w_{[i,j+|v|-1]}$ is a *complete return word from u to v in w* if $i \in \mathrm{Occ}(u,w)$ and if j is the first occurrence of v after u, i.e. if j is the minimum of the set of occurrences of v in w strictly greater than i. The set of all complete return words from u to v in w is denoted $\mathrm{CRet}_w(u,v)$. Clearly, $\mathrm{CRet}_w(u,u) = \mathrm{CRet}_w(u)$. From now on, the alphabet is fixed to be $\Sigma = \{0,1\}$.

2.1. Unit circle

The notation adopted for studying the dynamical system generated by some partially defined rotations on the circle is from Levitt [10]. The circle is identified with $\mathbb{R}/\mathbb{Z}$, equipped with the natural projection $p : \mathbb{R} \to \mathbb{R}/\mathbb{Z} : x \mapsto x + \mathbb{Z}$. The set $A \subseteq \mathbb{R}/\mathbb{Z}$ is called an *interval of* $\mathbb{R}/\mathbb{Z}$ if there exists an interval $B \subseteq \mathbb{R}$ such that $p(B) = A$ (see Figure 1).

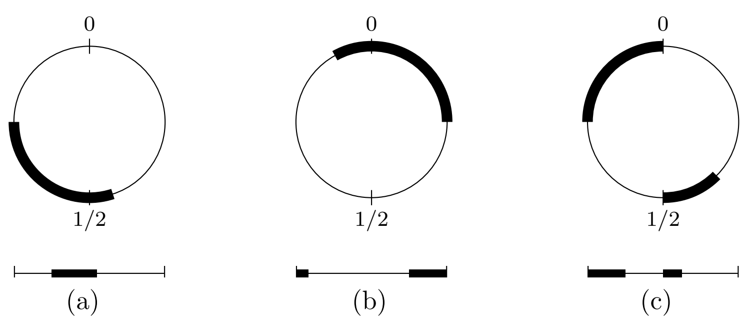

Figure 1: (a) The interval $[0.25, 0.55]$. (b) The interval $[0.75, 0.08]$. (c) Not an interval.

An interval I of $\mathbb{R}/\mathbb{Z}$ is fully determined by the ordered pair of its endpoints, $\partial(I) = \{x, y\}$ where $x \le y$ or $x \ge y$. The topological closure of I is the closed interval $\overline{I} = I \cup \partial(I)$, and its interior is the open set $\mathrm{Int}(I) = \overline{I} \setminus \partial(I)$.

The basic function on $\mathbb{R}/\mathbb{Z}$ considered in our study is the rotation of angle $\alpha \in \mathbb{R}$, defined by $R_\alpha(x) = x + \alpha \in \mathbb{R}/\mathbb{Z}$. Clearly, R_α is a bijection. As usual, this function is extended to sets of points $R_\alpha(X) = \{R_\alpha(x) : x \in X\}$ and in particular to intervals. Conveniently, the iterates of R_α are defined by $R_\alpha^m(x) = x + m\alpha \in \mathbb{R}/\mathbb{Z}$, where $m \in \mathbb{Z}$.

2.2. Interval exchange transformations

An interval exchange transformation is a piecewise affine transformation which maps a partition of the space into intervals to another one according to a permutation. Here, the notation is adapted from Keane and Rauzy (see [11, 12]).

Let $J, K \subseteq \mathbb{R}/\mathbb{Z}$ be two left-closed right-open intervals of the same length λ. Let $q \geq 1$ be an integer and $(\lambda_1, \lambda_2, \ldots, \lambda_q)$ be a vector with values in $\mathbb{R}^+$ such that $\sum_{i=1}^{q} \lambda_i = \lambda$, and let σ be a permutation of the set $\{1, 2, \ldots, q\}$. The intervals J and K are partitioned into q sub-intervals as follows. For $1 \leq i \leq q$, define

$$J_i = \left[\sum_{j<i} \lambda_j, \sum_{j\leq i} \lambda_j \right[\quad \text{and} \quad K_i = \left[\sum_{k<\sigma^{-1}(i)} \lambda_{\sigma(k)}, \sum_{k\leq\sigma^{-1}(i)} \lambda_{\sigma(k)} \right[.$$

The *q-interval exchange transformation* according to σ is a function F such that $F(J_i) = K_i$ and $F|_{J_i}$ is a translation for $i = 1, 2, \ldots, q$.

2.3. Poincaré's first return function

Let $J, K \subseteq \mathbb{R}/\mathbb{Z}$ be two nonempty left-closed and right-open intervals and let $\alpha \in \mathbb{R}$. Define the map $T_\alpha(J, K)$ by

$$\begin{array}{rccl} T_\alpha(J,K): & J & \to & \mathbb{N}^+ \cup \{+\infty\} \\ & x & \mapsto & \inf\{t \in \mathbb{N}^+ \mid x + t\alpha \in K\} \end{array}$$

and the map $P_\alpha(J, K)$ by

$$\begin{array}{rccl} P_\alpha(J,K): & J & \to & K \\ & x & \mapsto & x + T_\alpha(J,K)(x) \cdot \alpha, \end{array}$$

The number $T_\alpha(J, K)(x)$ indicates how many rotations of angle α it takes to move from the point x of J to some target point $P_\alpha(J, K)(x)$ in K. The function $P_\alpha(J, K)$ is a natural generalization of the usual *Poincaré's first return function* (when $J = K$) and so called as well [5].

Note that it is possible that $T_\alpha(J, K) = +\infty$. Indeed, if α is rational, there is no guarantee that the interval K can be reached from the interval J by rotations of angle α. However, if α is irrational, $T_\alpha(J, K)(x)$ is finite (using a density argument) for all $x \in J$. Also, if $J = K$, then $T_\alpha(J, K)(x) \in \mathbb{N}^+$ for all $x \in J$ for all $\alpha \in \mathbb{R}$, even if α is rational.

Moreover, we recall a well-known result which can be deduced from Keane [11]:

Lemma 1. *The induced map $P_\alpha(J,J)$ of the rotation R_α is an exchange transformation of $r \leq 3$ intervals. Moreover, there exists a decomposition*

$$J = J_1 \cup J_2 \cup \ldots \cup J_s, \quad r \leq s \leq 3,$$

into disjoint subintervals and positive integers $t_1, t_2, \ldots, t_s$ such that for $x \in J_i$,

$$P_\alpha(J,J)(x) = R_\alpha^{t_i}(x),$$

with R_α continuous on every interval $R_\alpha^k(J_i)$, $k = 0, 1, \ldots, t_i - 1$. □

To conclude this section, we show that, under some conditions, Poincaré's first return function is an interval exchange transformation. Define the complement of a permutation σ of length n by $\overline{\sigma}(i) = n + 1 - \sigma(i)$, for $i = 1, \ldots, n$.

Proposition 2. *Assume that $|J| = |K| \leq \alpha$. Then $P_\alpha(J,K)$ is a bijection if and only if $P_\alpha(J,K)$ is a q-interval exchange transformation of permutation $\overline{\mathrm{Id}_q}$ for some $q \in \{1,2,3\}$.*

Proof. Let $P = P_\alpha(J,K)$ and $T = T_\alpha(J,K)$.
($\Leftarrow$) By definition of interval-exchange transformation. ($\Rightarrow$) Consider the image $T(J) \subseteq \mathbb{N}^+$. If $|T(J)| = 1$, then P is a 1-interval exchange transformation, so that one may suppose $|T(J)| \geq 2$.

Let t_1 and t_2 be the two smallest values of $T(J)$ with $t_1 < t_2$ and let $J_i = T^{-1}(t_i)$, $K_i = P(J_i)$ for $i = 1, 2$. The set K_1 is an interval (if $1/2 < |K| \leq \alpha$, then $t_1 = 1$ and K_1 cannot overlap both endpoints of K) sharing an endpoint with K. Since $K_1 = K \cap R_\alpha^{t_1}(J)$ and $K_2 = K \cap R_\alpha^{t_2}(J - J_1)$ and since P is a bijection, K_1 and K_2 do not intersect and they both share a distinct endpoint with K. If $|T(J)| = 2$, then $J = J_1 \cup J_2$ and $K = K_1 \cup K_2$, and P is the 2-interval exchange transformation given by J_1 and J_2 of permutation $(2,1)$.

Suppose now $|T(J)| \geq 3$. Let $t_3 = \min(T(J) - \{t_1, t_2\})$, $J_3 = T^{-1}(t_3)$ and $K_3 = P(J_3) = K \cap R_\alpha^{t_3}(J - (J_1 \cup J_2))$. Since K_3 is non empty and because P is a bijection, then $R_\alpha^{t_3}(J - (J_1 \cup J_2))$ must intersect K in the interval $K - (K_1 \cup K_2)$. But the length of this left-closed right-open interval is $|K - (K_1 \cup K_2)| = |J - (J_1 \cup J_2)|$ so that there is just enough space for it. It follows that $R_\alpha^{t_3}(J - (J_1 \cup J_2)) \subset K$ and thus $K_3 = R_\alpha^{t_3}(J - (J_1 \cup J_2))$. Hence, $J_3 = J - (J_1 \cup J_2)$, $|T(J)| = 3$ and P is the 3-interval exchange transformation given by J_1, J_2 and J_3 of permutation $(3,2,1)$. □

2.4. Coding of rotations

Let $x, \alpha, \beta \in \mathbb{R}/\mathbb{Z}$. The unit circle $\mathbb{R}/\mathbb{Z}$ is partitioned into two nonempty intervals $I_1 = [0, \beta[$ and $I_0 = [\beta, 1[$. Then, a sequence of letters $(c_i)_{i \in \mathbb{N}}$ in $\Sigma = \{0,1\}$ is defined by setting

$$c_i = \begin{cases} 1 & \text{if } R_\alpha^i(x) \in [0, \beta[, \\ 0 & \text{if } R_\alpha^i(x) \in [\beta, 1[, \end{cases}$$

and a sequence of words $(\mathbf{C}_n)_{n\in\mathbb{N}}$, by $\mathbf{C}_n(x) = c_0c_1\cdots c_{n-1}$ where $\mathbf{C}_0(x) = \varepsilon$. The *coding of rotations* of x with parameters (α, β) is the infinite word

$$\mathbf{C}(x) = \lim_{n\to\infty} \mathbf{C}_n(x).$$

Example 3. If $x = 0.23435636$, $\alpha = 0.222435236$ and $\beta = 0.30234023$, then

$$\mathbf{W}_{\text{Ex.3}} = \mathbf{C}(x) = 10001000110001000110001000110001000110001000110\ldots$$

Example 4. If $x = 0.23435636$, $\alpha = 0.422435236$ and $\beta = 0.30234023$, then

$$\mathbf{W}_{\text{Ex.4}} = \mathbf{C}(x) = 10100000101000010000101000010100001010010100000 10\ldots$$

It is well-known that $\mathbf{C}(x)$ is periodic if and only if α is rational. When α is irrational, with $\beta = \alpha$ or $\beta = 1-\alpha$, the corresponding coding is a Sturmian word. Otherwise, the case $\beta \notin \mathbb{Z} + \alpha\mathbb{Z}$ yields Rote words [3], while $\beta \in \mathbb{Z} + \alpha\mathbb{Z}$ the quasi-Sturmian words [5, 4].

Poincaré's first return function is linked with complete return words as shown by the next lemma.

Lemma 5. *Let $u, v \in \text{Fact}(\mathbf{C})$ and $T = T_\alpha(I_u, I_v)$. Then, the set of complete return words from u to v in $\mathbf{C}$ is exactly $\{\mathbf{C}_{T(\gamma)+|v|}(\gamma) \mid \gamma \in I_u\}$.*

Proof. The word w is a complete return word from u to v in $\mathbf{C}$ if and only if $w = \mathbf{C}_{[j,k+|v|-1]}$, j is an occurrence of u and if k is the first occurrence of v in w strictly greater than j, if and only if there exists $\gamma \in I_w$ such that $\gamma \in I_u$, $R_\alpha^{|w|-|v|} \in I_v$ and $R_\alpha^i \notin I_v$ for all $0 < i < |w| - |v|$, if and only if there exists $\gamma \in I_w$ such that $\gamma \in I_u$ and $T_\alpha(I_u, I_v)(\gamma) = |w| - |v|$, if and only if there exists $\gamma \in I_u$ such that $w = \mathbf{C}_{|w|}(\gamma) = \mathbf{C}_{T_\alpha(I_u,I_v)(\gamma)+|v|}(\gamma)$. □

3. Partitions of the unit circle

Let $x, \alpha, \beta \in \mathbb{R}/\mathbb{Z}$. For each word $w \in \Sigma^*$, the set of points I_w from which the word w is read under rotations by α is:

$$I_w = \{\gamma \in \mathbb{R}/\mathbb{Z} \mid \mathbf{C}_{|w|}(\gamma) = w\}.$$

The sets I_w are easily computed from the letters of w and form a partition of the unit circle $\mathbb{R}/\mathbb{Z}$, explicitly:

$$I_w = \bigcap_{0\le i\le n-1} R_\alpha^{-i}(I_{w_i}) \tag{1}$$

$$P_n = \{I_w \mid w \in \text{Fact}_n(\mathbf{C}(\gamma)), \gamma \in \mathbb{R}/\mathbb{Z}\} \tag{2}$$

The set of boundary points $\mathcal{P}_n$ of the partition P_n is

$$\mathcal{P}_n = \bigcup_{I\in P_n} \partial(I).$$

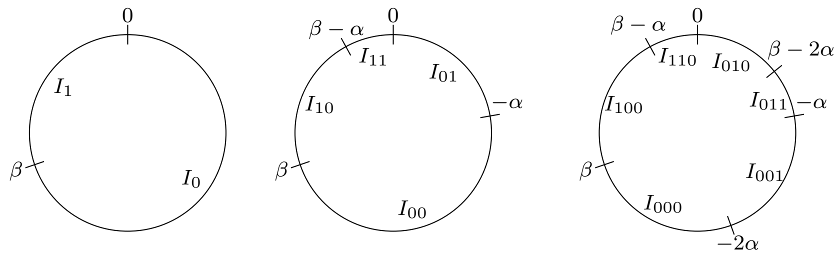

Figure 2: Representation of the partitions P_1, P_2 and P_3 of $\mathbf{W}_{\text{Ex.3}}$

Trivially, when $n = 1$, $\mathcal{P}_1 = \partial(I_0) \cup \partial(I_1) = \{0, \beta\}$ and it can be shown more generally that

$$\mathcal{P}_n = \{-i\alpha \mid i = 0, 1, \cdots, n-1\} \cup \{\beta - i\alpha \mid i = 0, 1, \cdots n-1\}.$$

The partitions P_1, P_2 and P_3 for $\mathbf{W}_{\text{Ex.3}}$ are represented in Figure 2 whereas those of $\mathbf{W}_{\text{Ex.4}}$ are illustrated in Figure 3. Note that in the second case, I_{00} and

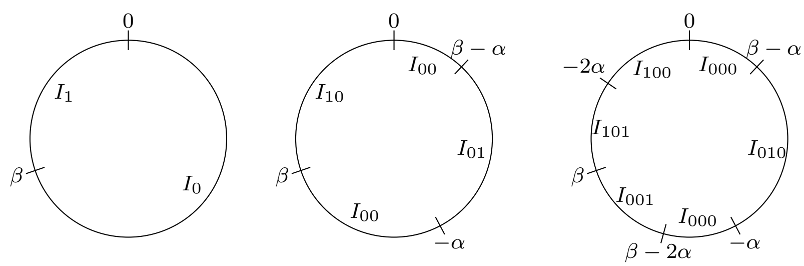

Figure 3: Representation of the partitions P_1, P_2 and P_3 of $\mathbf{W}_{\text{Ex.4}}$

I_{000} are not intervals. The proof of the main result of this paper is precisely based on two cases: whether I_w is an interval (Case 1) or not (Case 2). First, Lemma 1 implies the following result.

Lemma 6. *Let $w \in \Sigma^*$ be such that I_w is an interval. Then*

(i) *$P_\alpha(I_w, I_w)$ is a q-interval exchange transformation, where $q \in \{1, 2, 3\}$.*

(ii) *w has at most 3 complete return words.*

The following lemma is technical and is useful for our goal.

Lemma 7. *Let E be a finite set. Let $(A_i)_{i\in E}$ be a family of left-closed and right-open intervals $A_i \subseteq \mathbb{R}/\mathbb{Z}$. Let $\ell = \min\{|A_i| : i \in E\}$ and $L = \max\{|A_i| : i \in E\}$. If $\ell + L \leq 1$, then $\bigcap_{i\in E} A_i$ is an interval.*

Proof. The proof proceeds by induction on n. If $n = 1$, there is nothing to prove. Otherwise, let $k \in E$ be such that $|A_k| = L = \max\{|A_i| : i \in E\}$ and $\ell = \min\{|A_i : i \in E|\}$. Then

$$\bigcap_{i\in E} A_i = A_k \cap \left(\bigcap_{i\in E\setminus\{k\}} A_i \right).$$

By the induction hypothesis, $\bigcap_{i\in E\setminus\{k\}} A_i$ is an interval and its length is less than ℓ. Since $\ell + L \leq 1$, A_i and $\cap_{i\in E} A_i$ cannot intersect on both of their endpoints at the same time. □

Under some mild condition, it is guaranteed that the set I_w is an interval. This elementary result is provided for sake of completeness.

Lemma 8. *Let $w \in \Sigma^*$, and $\alpha, \beta \in \mathbb{R}/\mathbb{Z}$. The following properties hold:*

(i) *If both letters 0 and 1 appear in w, then I_w is an interval and $|I_w| \leq \alpha$;*

(ii) *If $\alpha < \beta$ and $\alpha < 1 - \beta$, then I_w is an interval.*

Proof. (i) Let $L = \{|R_\alpha^{-i}(I_{w_i})| : 0 \leq i \leq n-1\}$. If both letters 0 and 1 appear in the word w then $L = \{\beta, 1-\beta\}$ so that $\min L + \max L = 1$. Therefore, the intersection of intervals of Equation (1) satisfies the criteria of Lemma 7 and hence I_w is an interval.

The factor 01 or the factor 10 must appear in w. In the first case, the length of I_w is bounded: $|I_w| \leq |I_{01}| = |R_\alpha(I_0) \cup R_\alpha^{-1}(I_1)| \leq \alpha$. A similar inequality is obtained for the factor 10.

(ii) We prove the contrapositive. Assume that there exist a positive integer n and a word $w \in \mathrm{Fact}_n(\mathbf{C})$ such that I_w is an interval while I_{wa} is not, for some letter a. It follows from (i) that $w = a^n$ and $|I_a| > |I_b|$. However, Equation (1) implies $I_w = \bigcap_{i=0}^{n-1} R_\alpha^{-i}(I_a)$. In particular, $I_w \subseteq R_\alpha^{-n+1}(I_a)$ so that

$$R_\alpha^{-n+1}(I_b) \subseteq [0,1[\setminus I_w.$$

Moreover, $I_{wa} = I_w \cap R_\alpha^{-n}(I_a)$ and hence $R_\alpha^{-n}(I_b) \subseteq I_w$. It follows that $R_\alpha^{-n+1}(I_b) \cap R_\alpha^{-n}(I_b) = \emptyset$, $R_\alpha(I_b) \cap I_b = \emptyset$ and $\alpha \geq |I_b| = \min\{\beta, 1-\beta\}$. □

3.1. Symmetry of the partition

In [13], the authors used the global symmetry of the partition $\mathcal{P}_n$, sending the interval I_w on the interval $I_{\widetilde{w}}$. In fact, there are two points y_n and y'_n such that $2 \cdot y_n = 2 \cdot y'_n = \beta - (n-1)\alpha$, and the symmetry S_n of $\mathbb{R}/\mathbb{Z}$ is defined by $x \mapsto 2y_n - x$. This symmetry is useful for describing the structure of return words as illustrated in Figure 4.

Lemma 9. *Let $m \in \mathbb{N}$. The following properties hold.*

(i) *If $S_n(x) = R_\alpha^m(x)$, then $S_n(x+\alpha) = R_\alpha^{m-1}(x)$.*

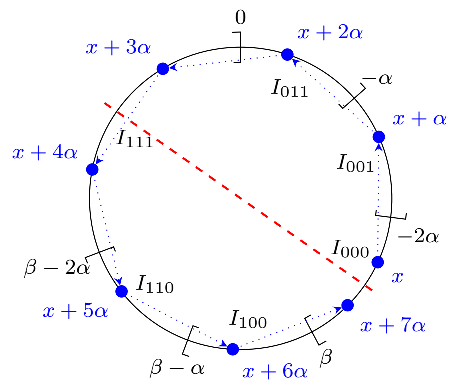

Figure 4: Let $\alpha = 0.135$ and $\beta = 0.578$. The orbit of the point $x = (\beta - 9\alpha)/2 = 0.6815$ is symmetric with respect to $S_3 : x \mapsto \beta - 2\alpha - x$. In fact, one may verify that the points x and $x + 7\alpha$ satisfy $S_3(x) = x + 7\alpha$ and that $\mathbf{C}_{10}(x) = 0001111000$ is a palindrome.

(ii) *If* $x \in \mathrm{Int}(I_w)$, *then* $S_n(x) \in I_{\widetilde{w}}$.

(iii) *Assume that* $x + \ell \cdot \alpha \notin \mathcal{P}_n$ *for all* $0 \le \ell \le m$. *If* $S_n(x) = R_\alpha^m(x)$, *then* $\mathbf{C}_{n+m}(x)$ *is a palindrome.*

Proof. (i) One has

$$\begin{aligned} S_n(x+\alpha) &= 2y_n - x - \alpha = S_n(x) - \alpha \\ &= x + m\alpha - \alpha = x + (m-1)\alpha. \end{aligned}$$

(ii) This follows from the definition of S_n. (iii) In any case, one may suppose that $x \in I_w$ where $|w| = n$. The proof is done by induction on m. If $m = 0$, then $x \in I_{\widetilde{w}}$ from (ii) so that $w = \widetilde{w} = \mathbf{C}_{n+0}(x)$. If $m = 1$, then $x+\alpha \in I_{\widetilde{w}}$ from (ii). Hence, $\mathbf{C}_{n+1}(x) = a\widetilde{w} = wb$ where $a, b \in \{0, 1\}$. Clearly, $a = w_0 = b$ and $\mathbf{C}_{n+1}(x)$ is a palindrome. In general, for $m \ge 2$, $\mathbf{C}_{m+n}(x) = a \cdot \mathbf{C}_{m+n-2}(x+\alpha) \cdot b$ where $a = w_0 = b$ since a is the first letter of w and b is the last letter of $\widetilde{w}$. From (i), we know that $S_n(x+\alpha) = (x+\alpha) + (m-2)\alpha$ so that $\mathbf{C}_{m+n-2}(x+\alpha)$ is a palindrome by the induction hypothesis. □

4. Main results

We describe now the relation between dynamical systems and coding of rotations on two intervals by computing the complete return words of $\mathbf{C}(x)$ using Poincaré's first return function as follows. The key idea is to establish that every complete return word of a palindrome is itself a palindrome. For that purpose, assume that $u \in \mathrm{Pal}(\mathbf{C})$ where $\mathbf{C} = \mathbf{C}(x)$ is such a coding. There are two cases to consider according to whether I_u is an interval or not.

4.1. Case 1: I_u is an interval

Recall from Lemma 6 that $P_\alpha(I_u, I_u)$ is a q-interval exchange transformation with $q \in \{1,2,3\}$. Let $(J_i)_{1\le i\le q\le 3}$ be the q nonempty and maximal sub-intervals of I_u such that $P_\alpha(I_u,I_u)(J_i) = R_\alpha^{t_i}(J_i)$ where $i<j$ implies $t_i<t_j$. Any point of J_i requires the same number t_i of rotations by α to reach the interval I_u. In the general case, two points in J_i may code different words of length t_i under rotations by α. For example, the factor 100 has three complete return words in $\mathbf{W}_{\mathrm{Ex.3}}$ among which two have the same length: 10000100, 1000100, 10001100.

Nevertheless, the next lemma ensures the uniqueness of the return word of length t_i obtained from the interval J_i in the case where I_u is an interval.

Lemma 10. *If I_u is an interval, then for all $x, y \in J_i$ and $1 \le i \le q$ one has $\mathbf{C}_{t_i}(x) = \mathbf{C}_{t_i}(y)$.*

Proof. Without loss of generality one may assume that $x < y$. By contradiction, suppose that there exists k, $0 \le k \le t_i$, such that $R_\alpha^k(x)$ lies in $I_0 = [0, \beta[$ and $R_\alpha^k(y)$ lies in $I_1 = [\beta, 0[$ (the proof is the same for the other case). Then,

$$\beta \in]R_\alpha^k(x), R_\alpha^k(y)] \subset R_\alpha^k(\mathrm{Int}(J_i)).$$

If $k < n$, then $\beta - k\alpha \in \mathrm{Int}(J_i)$ which is a contradiction because it is a point of the set $\mathcal{P}_n$. If $k \ge n$, then $\beta - \ell\alpha \in R_\alpha^{k-\ell}(\mathrm{Int}(J_i))$ for all $0 \le \ell < n$. This is a contradiction as well because at least one of the boundary points of I_u is of the form $\beta - \ell\alpha$ which contradicts the minimality of t_i. □

A well-chosen representative allows one to compute the word coded from the interval J_i. It appears that the middle point m_i of J_i is convenient for being symmetric: indeed, it follows from Lemma 5 and Lemma 10 that if $u \in \mathrm{Fact}_n(\mathbf{C})$ is a palindrome such that I_u is an interval, then

$$\mathrm{CRet}_{\mathbf{C}}(u,u) = \{\mathbf{C}_{t_i+n}(m_i) \mid 1 \le i \le q\}. \tag{3}$$

Proposition 11. *If I_u is an interval then every complete return word of u is a palindrome.*

Proof. Let $n = |u|$ and $w \in \mathrm{CRet}_{\mathbf{C}}(u)$. Since I_u is an interval, there exists $i \in \{1,2,3\}$ such that $w = \mathbf{C}_{T(m_i)+n}(m_i)$ where $T = T_\alpha(I_u, I_u)$ and m_i is the middle point of the interval J_i, for $i = 1,2,3$. Let $\sigma_i : \gamma \mapsto 2m_i - \gamma$ denote the reflection with respect to the middle point m_i of the sub-interval J_i. Moreover, since $P_\alpha(I_u, I_u)$ is an interval exchange transformation, the following equalities hold

$$m_i + T(m_i)\alpha = P_\alpha(I_u,I_u)(m_i) = (S_n \circ \sigma_i)(m_i) = S_n(m_i).$$

From Lemma 10, we know that none of the points $m_i + \ell \cdot \alpha$ are in $\mathcal{P}_n$ so that Lemma 9 (iii) can be applied, and w is a palindrome. □

4.2. Case 2: I_u is not an interval

In this case Lemma 8 implies that $u = a^n$ is some power of a single letter. Then every complete return word of u is either (i) of the form a^{n+1} or (ii) belongs to the set $a^n b\Sigma^* \cap \Sigma^* ba^n$, with $a \neq b$. The first case is trivial because a^{n+1} is clearly a palindrome, so only the second case is described in detail.

Proposition 12. *If $u' = a^n b$ and $v' = ba^n$ then $P_\alpha(I_{u'}, I_{v'})$ is a bijection.*

Proof. It suffices to show that $P_{-\alpha}(I_{v'}, I_{u'})$ is the inverse of $P_\alpha(I_{u'}, I_{v'})$. By contradiction, assume that it is not the case. Then there exist $x \in I_{u'}$ and $y \in I_{v'}$ such that $y = P_\alpha(I_{u'}, I_{v'})(x)$ and $T_{-\alpha}(I_{v'}, I_{u'})(y) < T_\alpha(I_{u'}, I_{v'})(x)$, i.e. the orbit of y falls within $I_{u'}$ before reaching x when making rotations of $-\alpha$. Therefore, by Lemma 5, there exists a complete return word w from $u' = a^n b$ to $v' = ba^n$ such that u' occurs twice in w. But this implies that v' occurs twice in w as well, which contradicts the definition of complete return word. $\square$

Corollary 13. *If $u' = a^n b$ and $v' = ba^n$ then $P_\alpha(I_{u'}, I_{v'})$ is a q-interval exchange transformation, where $q \in \{1, 2, 3\}$.* $\square$

Let $(J_i)_{1 \leq i \leq q}$ be the q nonempty sub-intervals of $I_{u'}$ as defined in the proof of Proposition 2. It follows from the preceding lemmas that any point of J_i requires the same number of rotations by α to reach the interval $I_{v'}$, i.e. $T_\alpha(I_{u'}, I_{v'})(x) = T_\alpha(I_{u'}, I_{v'})(y)$ for all $x, y \in J_i$. Hence, for $1 \leq i \leq q$, let

$$t_i = T_\alpha(I_{u'}, I_{v'})(J_i). \tag{4}$$

As pointed out above, two points in the interval J_i might code different words of length t_i under rotations by α. Nevertheless, the uniqueness of the return word obtained from the interval J_i is ensured in our case.

Lemma 14. *If $u' = a^n b$ and $v' = ba^n$, then for all $x, y \in J_i$ and $1 \leq i \leq q$ one has $\mathbf{C}_{t_i}(x) = \mathbf{C}_{t_i}(y)$ where t_i is defined by Equation (4).*

Proof. Without loss of generality one may assume that $x < y$. By contradiction, suppose that there exists k, $0 \leq k \leq t_i$, such that $R_\alpha^k(x)$ lies in $I_0 = [0, \beta[$ and $R_\alpha^k(y)$ lies in $I_1 = [\beta, 0[$ (the proof is the same for the other case). Then,

$$\beta \in]R_\alpha^k(x), R_\alpha^k(y)] \subset R_\alpha^k(\mathrm{Int}(J_i)).$$

If $k \leq n$, then $\beta - k\alpha \in \mathrm{Int}(J_i)$ which is a contradiction because it is a point of the set $\mathcal{P}_{n+1}$. If $k > n$, then $\beta - \ell\alpha \in R_\alpha^{k-\ell}(\mathrm{Int}(J_i))$ for all $0 \leq \ell \leq n$. This is a contradiction as well because at least one of the boundary points of $I_{u'}$ or $I_{v'}$ is of the form $\beta - \ell\alpha$ which contradicts the fact that $P_\alpha(I_{u'}, I_{v'})$ is a bijection. $\square$

Once again, we choose the middle point m_i as representative. It follows from Lemma 5 and Lemma 14 that if $u' = a^n b$ and $v' = ba^n$, then

$$\mathrm{CRet}_{\mathbf{C}}(u', v') = \{\mathbf{C}_{t_i+n+1}(m_i) \mid 1 \leq i \leq q\}. \tag{5}$$

Proposition 15. *Let* $T = T_\alpha(I_{a^n b}, I_{ba^n})$. *Then,*

$$\mathrm{CRet}(a^n) \subseteq \{a^{n+1}\} \cup \{\mathbf{C}_{T(m_i)+n+1}(m_i) \mid 1 \leq i \leq q\}.$$

Proof. Let $us \in \mathrm{CRet}(u)$ and let b be the first letter of s. If $b = a$, then $us = ua = a^{n+1}$. If $b \neq a$, then $us = ubtu = a^n bta^n$ where the last letter of bt must be b: these are the complete return words from $a^n b$ to ba^n as given by Equation (5). □

Proposition 16. *If* I_u *is not an interval then every complete return word of* u *is a palindrome.*

Proof. Let $n = |u|$ and let $w \in \mathrm{CRet}_{\mathbf{C}}(u)$. Since I_u is not an interval, $u = a^n$ with $a \in \{0, 1\}$ by Lemma 8. Proposition 15 implies that $w = a^{n+1}$ which is a palindrome or $w = \mathbf{C}_{T(m_i)+n+1}(m_i)$ where $T = T_\alpha(I_{a^n b}, I_{ba^n})$. Again, let $\sigma_i : \gamma \mapsto 2m_i - \gamma$ denote the reflection with respect to the middle point m_i of the sub-interval J_i. Then

$$m_i + T(m_i)\alpha = P_\alpha(I_{a^n b}, I_{ba^n})(m_i) = (S_{n+1} \circ \sigma_i)(m_i) = S_{n+1}(m_i).$$

so that w is a palindrome by Lemma 9 (iii). □

4.3. Concluding statements

First, it is worth mentioning that a slight technical refinement of Katok's results is obtained from Lemma 10 and Proposition 15.

Proposition 17. *The set of words* $w \in \mathrm{Fact}(\mathbf{C})$ *having* 4 *complete return words is finite. Moreover,* I_w *is not an interval and* w *is the power of a single letter.*

This result is illustrated in the following example.

Example 18. The factor 000 has four complete return words in $\mathbf{W}_{\mathrm{Ex.4}}$ namely

$$0000, \quad 0001000, \quad 000101000 \quad \text{and} \quad 00010100101000$$

all being palindromes.

This example illustrates that complete return words of palindromes are palindromes, a consequence of Propositions 11 and 16. In other words:

Theorem 19. *Every coding of rotations on two intervals is full.* □

In [14], the authors showed that an infinite word w whose set of factors is closed under reversal is full if and only if

$$\mathrm{Pal}_n(w) + \mathrm{Pal}_{n+1}(w) = \mathrm{Fact}_{n+1}(w) - \mathrm{Fact}_n(w) + 2 \quad \text{for all } n \in \mathbb{N}. \tag{6}$$

Since Equation (6) is verified for the case where the number of return words is at most 3, it is likely that Theorem 19 could be deduced from results in [14]. However, it is not clear if this also holds for the more involved case where the bound of 4 return words is realized. Here all possible cases were handled, including the periodic one.

Note. The fact that the number of (complete) return words is bounded by 3 when $\alpha < \min\{\beta, 1-\beta\}$ can be found in the work of Keane [11], Rauzy [12] or Adamczewski [5] with α irrational. Moreover, it was already known that $|\mathrm{CRet}(w)| = k$ for a non degenerate interval exchange on k intervals [15] and that $|\mathrm{CRet}(w)| = 2$ for a coding of rotations with $\alpha = \beta$ (the Sturmian case). Nevertheless, the proofs provided here hold for any values of α and β, taking into account rational values as well.

5. Complementary-symmetric Rote words

The special case when $\beta = 1/2$ deserves some attention. These words built on the alphabet $\Sigma = \{0, 1\}$ are called *complementary-symmetric Rote words* [3]. We provide here a combinatorial proof for the case where α is irrational and $\beta = 1/2$, based on the peculiar structure of antipalindromes, a generalization of palindromes.

An *antipalindrome* q is a word such that $\overline{q} = \widetilde{q}$ where $\overline{\cdot}$ is the non trivial involution on Σ^* (swapping of letters) defined by $0 \mapsto 1, 1 \mapsto 0$. Given two palindromes p and q, one writes $p \prec q$ if there exists a word x such that $x^{-1}q = p\widetilde{x}$ or equivalently $q = xp\widetilde{x}$. The *difference of* w, denoted by $\Delta(w)$, is the word $v = v_1 v_2 \cdots v_{|w|-1}$ defined by

$$v_i = (w_{i+1} - w_i) \bmod 2, \quad \text{for } i = 1, 2, \ldots, |w| - 2.$$

Complementary-symmetric words are connected to Sturmian words by a structural theorem.

Theorem 20 (Rote [3])**.** *An infinite word* $\mathbf{w}$ *is a complementary-symmetric Rote word if and only if the infinite word* $\Delta(\mathbf{w})$ *is a Sturmian word.*

For instance, the following word is complementary-symmetric

$$x = 11000110001100111001110011100011000\cdots$$

and its associated Sturmian word is

$$\begin{aligned} y &= \Delta(1110000001111000)\cdots \\ &= 0100101001010100101001010010101001\cdots \end{aligned}$$

The key idea is to exploit the link with Sturmian words and to use both palindromes and antipalindromes. First, we state without proof some elementary properties of the operator Δ.

Lemma 21. *Let* $u, v \in \Sigma^*$*, where* $|u|, |v| \geq 2$*. Then*

(i) $\Delta(u) = \Delta(v)$ *if and only if* $v = u$ *or* $v = \overline{u}$*,*

(ii) u *is either a palindrome or an antipalindrome if and only if* $\Delta(u)$ *is a palindrome and*

(iii) *u is an antipalindrome if and only if $\Delta(u)$ is an odd palindrome with central letter* 1.

The following fact, established in [16, 17], is useful.

Theorem 22. *A binary word $\mathbf{w}$ is Sturmian if and only if every nonempty factor u of $\mathbf{w}$ satisfies $|\mathrm{CRet}_{\mathbf{w}}(u)| = 2$.*

The lattice of palindrome factors of Sturmian words has the following factorial closure property.

Lemma 23. *Let $\mathbf{s}$ be a Sturmian word and $p, q \in \mathrm{Pal}(\mathbf{s})$, where $|p| \geq |q|$. Assume that there exists a nonempty word r such that $r \prec p$ and $r \prec q$. Then $q \prec p$.*

Proof. The proof proceeds by contradiction. Assume that $q \not\prec p$ and let r' be the longest palindrome such that $r' \prec p, q$. Clearly, $r' \neq \varepsilon$ since $r \neq \varepsilon$. Moreover, $r' \neq p, q$ since $q \not\prec p$. Therefore, there exist two distinct letters a and b such that $ar'a \prec p$ and $br'b \prec q$, i.e. $ar'a$ and $br'b$ are both factors of $\mathbf{s}$. This is a contradiction with the balance property of Sturmian words, since $|ar'a|_a - |br'b|_a = 2 > 1$. □

A last lemma is useful to prove Theorem 25.

Lemma 24. *Let $\mathbf{r}$ be a complementary-symmetric Rote word and $u \in \mathrm{Pal}(\mathbf{r})$. Then there exist a palindrome p and an antipalindrome q such that $u \in \mathrm{Pref}(p) \cap \mathrm{Pref}(q)$ and*

$$\mathrm{CRet}_{\mathbf{r}}(\Delta(u)) = \{\Delta(p), \Delta(q)\}.$$

Proof. From Theorem 20, we know that $\Delta(\mathbf{r})$ is Sturmian and from Lemma 21 that $\Delta(u)$ is a palindrome. Therefore, it follows from Theorem 22 that $\Delta(u)$ has two complete return words. Moreover, since $\Delta(\mathbf{r})$ is full, these two complete return words are palindromes.

Let p and q be the two words such that $u \in \mathrm{Pref}(p) \cap \mathrm{Pref}(q)$ and

$$\mathrm{CRet}_{\mathbf{r}}(\Delta(u)) = \{\Delta(p), \Delta(q)\}.$$

By Lemma 21(i), p and q are indeed unique and it follows from Lemma 21(ii) that p and q are either palindromes or antipalindromes.

First, one shows that p and q cannot be both antipalindromes. Arguing by contradiction, assume that the contrary holds. Then $\Delta(p)$ and $\Delta(q)$ are both palindromes of odd length having 1 as a central factor. By Lemma 23, one concludes that $\Delta(p) \prec \Delta(q)$ or $\Delta(q) \prec \Delta(p)$. The former case implies $|\Delta(q)|_{\Delta(u)} \geq 4$ while the latter implies $|\Delta(p)|_{\Delta(u)} \geq 4$, contradicting the fact that $\Delta(p)$ and $\Delta(q)$ are complete return words.

It remains to show that p and q cannot be both palindromes. Since $\mathbf{r}$ is recurrent for being Sturmian, there exists $v \in \mathrm{Fact}(\mathbf{r})$ such that $u \in \mathrm{Pref}(v)$, $\overline{u} \in \mathrm{Suff}(v)$ and $|v|_u = |v|_{\overline{u}} = 1$, i.e. $\Delta(v)$ is a complete return word of $\Delta(u)$ in $\Delta(\mathbf{r})$. But v is not a palindrome since $u \in \mathrm{Pref}(v)$ and $\overline{u} \in \mathrm{Suff}(v)$, so that it must be an antipalindrome. □

As a consequence the fullness property holds.

Theorem 25. *Rote words with $\beta = 1/2$ are full.*

Proof. Let $\mathbf{r}$ be a Rote sequence, $u \in \mathrm{Pal}(\mathbf{r})$ and v a complete return word of u in $\mathbf{r}$. It suffices to show that v is a palindrome.

First, note that $|v|_u = 2$ but it is possible to have $|v|_{\overline{u}} > 0$. Let $n = |v|_{\overline{u}}$. By Lemma 24, there exist a palindrome p and an antipalindrome q such that $\Delta(p)$ and $\Delta(q)$ are the two complete return words of $\Delta(u)$ in $\mathbf{r}$, where $u \in \mathrm{Pref}(p) \cap \mathrm{Pref}(q)$. If $n = 0$, then $v = p$ is a palindrome, as desired. Otherwise,

$$v = (qu^{-1})(\overline{p}u^{-1})^n\overline{q}.$$

Therefore,

$$\begin{aligned} \widetilde{v} &= \widetilde{\overline{q}}(\widetilde{u^{-1}\widetilde{\overline{p}}})^n(\widetilde{u^{-1}\widetilde{q}}) \\ &= q(u^{-1}\overline{p})^n(u^{-1}\overline{q}) = v, \end{aligned}$$

so that v is a palindrome. Hence, $\mathbf{r}$ is full. □

Acknowledgements

This paper is an extended version of a communication that appeared in [18]. The results presented in this paper were discovered by computer exploration using the open-source mathematical software `Sage` [19]. We are also grateful to the anonymous referees for pointing out crucial points that improved the presentation.

This research was supported by NSERC (Canada).

[1] P. Alessandri, V. Berthé, Three distance theorems and combinatorics on words, Enseig. Math. 44 (1998) 103–132.

[2] J. Berstel, P. Séébold, Sturmian words, in: Lothaire, Combinatorics on Words, Cambridge University Press, 2002.

[3] G. Rote, Sequences with subword complexity $2n$, J. Number Theory 46 (1994) 196–213.

[4] G. Didier, Codages de rotations et fractions continues, J. Number Theory 71 (1998) 275–306.

[5] B. Adamczewski, Codages de rotations et phénomènes d'autosimilarité, J. Théorie des nombres de Bordeaux 14 (2002) 351–386.

[6] X. Droubay, J. Justin, G. Pirillo, Episturmian words and some constructions of de Luca and Rauzy, Theoret. Comput. Sci. 255 (2001) 539–553.

[7] S. Brlek, S. Hamel, M. Nivat, C. Reutenauer, On the palindromic complexity of infinite words, Int. J. Found. Comput. Sci. 15 (2004) 293–306.

[8] A. Glen, J. Justin, S. Widmer, L. Q. Zamboni, Palindromic richness, Eur. J. Comb. 30 (2009) 510–531.

[9] M. Lothaire, Combinatorics on Words, Addison-Wesley, 1983.

[10] G. Levitt, La dynamique des pseudogroupes de rotations, Inventiones mathematicae 113 (1993) 633–670.

[11] M. Keane, Interval exchange transformations, Math. Zeit. 141 (1975) 25–31.

[12] G. Rauzy, Echange d'intervalles et transformations induites, Acta. Arith. 34 (1979) 315–328.

[13] V. Berthé, L. Vuillon, Palindromes and two-dimensional sturmian sequences, J. of Automata, Languages and Combinatorics 6 (2001) 121–138.

[14] M. Bucci, A. D. Luca, A. Glen, L. Zamboni, A connection between palindromic and factor complexity using return words, Adv. in Appl. Math. 42 (2009) 60–74.

[15] L. Vuillon, On the number of return words in infinite words constructed by interval exchange transformations, PU.M.A. 18 (2007) 345–355.

[16] J. Justin, L. Vuillon, Return words in sturmian and episturmian words, RAIRO Theoretical Informatics and Applications 34 (2000) 343–356.

[17] L. Vuillon, A characterization of sturmian words by return words, European Journal of Combinatorics 22 (2001) 263–275.

[18] A. Blondin Massé, S. Brlek, S. Labbé, L. Vuillon, Codings of rotations on two intervals are full, in: J. Nešetřil, A. Raspaud (Eds.), EuroComb 2009, 7th International Conference on Words, volume 34, pp. 289–293.

[19] W. Stein, et al., Sage Mathematics Software (Version 4.3.3), The Sage Development Team, 2010. http://www.sagemath.org.