Codings of rotations on two intervals are full

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

[a] *Laboratoire de Combinatoire et d'Informatique Mathématique, Un. du Québec à Montréal, CP 8888 Succ. "Centre-Ville", Montréal (QC), CANADA H3C 3P8*

[b] *Laboratoire de mathématiques, CNRS UMR 5127, Université de Savoie, 73376 Le Bourget-du-lac cedex, France*

EXTENDED ABSTRACT

The coding of rotations is a transformation taking a point x on the unit circle and translating x by an angle α, so that a symbolic sequence is built by coding the iteration of this translation on x according to a partition of the unit circle [2]. If 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 [3]. Otherwise, the coding is a Rote sequence if the length of the intervals are rationally independent of α [11] and quasi-Sturmian in the other case [6]. Many studies show properties of sequences constructed by codings of rotation in terms of their subword complexity [2], continued fractions and combinatorics on words [6] or discrepancy and substitutions [1]. Our goal is to link properties of the sequence given by coding of rotations with the palindromic structure of its subwords. The palindromic complexity $|\mathrm{Pal}(w)|$ of a finite word w is bounded by $|w|+1$, and finite Sturmian (and even episturmian) words realize the upper bound [7]. The palindromic defect of a finite word w is defined in [5] by $D(w) = |w|+1-|\mathrm{Pal}(w)|$, and words for which $D(w) = 0$ are called *full*. Moreover, the case of periodic words is completely characterized in [5]. Our main result is the following.

Theorem 0.1 *Every coding of rotations on two intervals is full.*

Our approach is based on return words of palindromes. Let w be a word, and $u \in \mathrm{Fact}(w)$. Then v is a *return word* of u in w if $u \in \mathrm{Pref}(v)$, $vu \in \mathrm{Fact}(w)$ and $|vu|_{|u|} = 2$. Similarly, v is a *complete return word* of u in w if $v = v'u$, where v' is a return word of u in w. The set of complete return words of u in w is denoted by $\overline{\mathrm{Ret}}_w(u)$. Clearly, the computation of the defect follows from the computation of the palindromic complexity. Indeed, the reader may verify the following characterization of full words (poorly efficient computationnaly speaking) :

$$w \text{ is full } \iff \forall p \in \mathrm{Pal}(w), \overline{\mathrm{Ret}}_w(p) \subseteq \mathrm{Pal}(w). \tag{1}$$

Interval exchange transformations. Let $a, e, \lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ be such that $\lambda_1, \lambda_2, \lambda_3 \geq 0$ and $\lambda_1 + \lambda_2 + \lambda_3 > 0$. Moreover, let $b = a + \lambda_1$, $c = b + \lambda_2$, $d = c + \lambda_3$, $f = e + \lambda_3$, $g = f + \lambda_2$ and $h = g + \lambda_1$. A function $T : [a, d[\to [e, h[$ is called a 3-*intervals exchange transformation* if

$$T(x) = \begin{cases} g + (x - a) & \text{if } a \leq x < b, \\ f + (x - b) & \text{if } b \leq x < c, \\ e + (x - c) & \text{if } c \leq x < d. \end{cases}$$

The subintervals $[a, b[$, $[b, c[$ and $[c, d[$ are said *induced* by T.
Coding of rotations. The notation adopted for studying the dynamical system generated by some partially defined rotations on the circle is from Levitt [8]. The circle is identified with $\mathbb{R}/\mathbb{Z}$, equipped with the natural projection $p : \mathbb{R} \to \mathbb{R}/\mathbb{Z} : x \mapsto x + \mathbb{Z}$. We say that $A \subseteq \mathbb{R}/\mathbb{Z}$ is an *interval of* $\mathbb{R}/\mathbb{Z}$ if there exists an interval $B \subseteq \mathbb{R}$ such that $p(B) = A$.

An interval I of $\mathbb{R}/\mathbb{Z}$ is fully determined by the ordered pair of its end points, $\mathrm{Bord}(I) = \{x, y\}$ where $x \leq y$ or $x \geq y$. The open interval is denoted $]x, y[$, while the closed ones are denoted $[x, y]$. Left-closed and right-open intervals $[x, y[$ as well as left-open and right-closed intervals $]x, y]$ are also considered. The topological closure of I is the closed interval $\bar{I} = I \cup \mathrm{Bord}(I)$, and its interior is the open set $\mathrm{Int}(I) = \bar{I} \setminus \mathrm{Bord}(I)$. Given a real number $\beta \in]0, 1[$, we consider the unit circle $\mathbb{R}/\mathbb{Z}$ partitioned into two nonempty intervals $I_0 = [0, \beta[$ and $I_1 = [\beta, 1[$. The rotation of angle $\alpha \in \mathbb{R}$ of a point $x \in \mathbb{R}/\mathbb{Z}$ is defined by $R_\alpha(x) = x + \alpha \in \mathbb{R}/\mathbb{Z}$, where the addition operation is denoted by the sign $+$ in $\mathbb{R}/\mathbb{Z}$ as in $\mathbb{R}$. As usual this function is extended to sets of points $R_\alpha(X) = \{R_\alpha(x) : x \in X\}$ and in particular to intervals. For convenience and later use, we denote $R_\alpha^y(x) = x + y\alpha \in \mathbb{R}/\mathbb{Z}$, where $y \in \mathbb{Z}$.

Let $\Sigma = \{0, 1\}$ be the alphabet. Given any nonnegative integer n and any $x \in \mathbb{R}/\mathbb{Z}$, we define a finite word C_n by

$$C_n(x) = \begin{cases} \varepsilon & \text{if } n = 0, \\ 0 \cdot C_{n-1}(R_\alpha(x)) & \text{if } n \geq 1 \text{ and } x \in [0, \beta[, \\ 1 \cdot C_{n-1}(R_\alpha(x)) & \text{if } n \geq 1 \text{ and } x \in [\beta, 1[. \end{cases}$$

The *coding of rotations* of x with parameters (α,β) is the infinite word

$$C_\alpha^\beta(x) = \lim_{n\to\infty} C_n(x).$$

For sake of readability, the parameters (α,β) are often omitted when the context is clear. One shows that $C(x)$ is periodic if and only if α is rational. When α is irrational, with $\beta = \alpha$ or $\beta = 1-\alpha$ we get the well known Sturmian words, the case $\beta \notin \mathbb{Z}+\alpha\mathbb{Z}$ yields Rote words, while $\beta \in \mathbb{Z}+\alpha\mathbb{Z}$ the quasi-Sturmian words [1,6].

For each factor w of the infinite word $C(x)$, one defines the nonempty set

$$I_w := \{x \in \mathbb{R}/\mathbb{Z} \mid C_n(x) = w\}.$$

Proposition 0.2 [4] *Let C be the coding of rotation of parameters $\alpha,\beta \in \mathbb{R}/\mathbb{Z}$ and $n \in \mathbb{N}$. Then $P_n = \{I_w \mid w \in \mathrm{Fact}_n(C)\}$ is a partition of $\mathbb{R}/\mathbb{Z}$.*

The set I_w needs not be an interval, but under some constraints, there is a guarantee that I_w is indeed an interval.

Lemma 0.3 *Let I be a finite set. Let $(A_i)_{i\in I}$ be a family of left-closed and right-open intervals $A_i \subseteq \mathbb{R}/\mathbb{Z}$. Let $\ell = \min\{|A_i| : i \in I\}$ and $L = \max\{|A_i| : i \in I\}$. If $\ell + L \leq 1$, then $\bigcap_{i\in I} A_i$ is an interval.*

Lemma 0.4 *Let $C(x)$ be a coding of rotations. If both letters 0 and 1 appear in the word $w \in \mathrm{Fact}(C(x))$, then I_w is an interval.*

Lemma 0.5 *Let C be a coding of rotations of parameters α, β, x. If $\alpha < \beta$ and $\alpha < 1-\beta$, then I_w is an interval for any $w \in \mathrm{Fact}(C)$.*

Definition 1 *We say that a coding of rotations C of parameters α, β, x is* non degenerate *if $\alpha < \beta$ and $\alpha < 1-\beta$. Otherwise, we say that C is* degenerate.

In [4], the authors used the global symmetry axis of the partition 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 $x \mapsto 2y_n - x$.

Lemma 0.6 *Let S_n be the symmetry axis related to $n \in \mathbb{N}$, $x,\alpha \in \mathbb{R}/\mathbb{Z}$ such that $x \notin \mathcal{P}_n$ and $m \in \mathbb{N}$.*

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

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

(iii) *If $S_n(x) = x + m\alpha$, then $C_{n+m}(x)$ is a palindrome.*

Complete return words. this subsection describes the relation between dynamical systems and their associated word C. More precisely, we link complete return words of C to the Poincaré's first return function.

Poincaré's first return function. Let $I, J \subseteq \mathbb{R}/\mathbb{Z}$ be two nonempty left-closed and right-open intervals and $\alpha \in \mathbb{R}$. We define a map $r_\alpha(I,J) : I \to \mathbb{N}^*$ by $r_\alpha(I,J)(x) = \min\{k \in \mathbb{N}^* \mid x + k\alpha \in J\}$ for $x \in I$. The *Poincaré's first return function* $P_\alpha(I,J)$ of R_α on I is the function $P_\alpha(I,J) : I \to J$ given by $P_\alpha(I,J)(x) = x + r_\alpha(I,J)(x) \cdot \alpha$.

The study of Poincaré's first return function is justified by the following result which establishes a link with complete return words.

Proposition 0.7 *Let $w \in \mathrm{Fact}_n(C)$ and $r = r_\alpha(I_w, I_w)$. Then, $\overline{\mathrm{Ret}}(w) = \{C_{r(x)+n}(x) | x \in I_w\}$.*

Lemma 0.8 *If $I \subseteq \mathbb{R}/\mathbb{Z}$ is a left-closed right-open interval, then $P_\alpha(I,I)$ is a 3-intervals exchange transformation.*

Lemma 0.9 *If $w = a^n$ is a word such that I_w is not an interval, then $P_\alpha(I_{wb}, I_{bw})$, where $b \in \{0,1\}$, is a 3-intervals exchange transformation.*

Properties of complete return words. The next results use the following notation. Let $w \in \mathrm{Fact}_n(C)$. Suppose that $P_\alpha(I_w, I_{\widetilde{w}})$ is an 3-intervals exchange transformation and let J_1, J_2 and J_3 be its induced subintervals. Let $i \in \{1,2,3\}$ and x_i be the middle point of J_i. It follows from the preceding lemmas that $r_\alpha(I_w, I_{\widetilde{w}})(x) = r_\alpha(I_w, I_{\widetilde{w}})(x_i)$ for all $x \in J_i$. Hence, we define $r_i = r_\alpha(I_w, I_{\widetilde{w}})(x_i)$.

Lemma 0.10 *For all $x, y \in J_i$, we have that $C_{r_i}(x) = C_{r_i}(y)$.*

Proposition 0.11 *Assume that I_w is an interval and $r = r_\alpha(I_w, I_w)$. Then, $\overline{\mathrm{Ret}}(w) = \{C_{r_1+n}(x_1), C_{r_2+n}(x_2), C_{r_3+n}(x_3)\}$.*

Corollary 0.12 *Every factor of a non degenerate coding of rotations as well as every factor of any coding of rotations containing both* 0*s and* 1*s has at most* 3 *(complete) return words.*

Proposition 0.13 *Suppose that $I_{w'}$ is not an interval, i.e. $w' = a^{n-1}$, $a \in \{0,1\}$. Let $b \in \{0,1\}$, $b \neq a$, $w = w'b$ and $r = r_\alpha(I_{w'b}, I_{bw'})$. Then,*

$$\overline{\mathrm{Ret}}(w') \subseteq \{a^n\} \cup \{C_{r_1+n}(x_1), C_{r_2+n}(x_2), C_{r_3+n}(x_3)\}.$$

Corollary 0.14 *Every factor w of any coding of rotations has at most* 4 *complete return words. Moreover, this bound is realized only if $w = a^n$.*

The last Corollary is illustrated in the following example.

Example 1 *Let $x = 0.23435636$, $\alpha = 0.422435236$ and $\beta = 0.30234023$. Then $C = C_\alpha^\beta(x) = 0101111010111101111010111101011110101101011111\ldots$. Moreover, $\overline{\mathrm{Ret}}_C(111) = \{111010111, 11101011010111, 1111, 1110111\}$, so that the factor* 111 *has exactly* 4 *complete return words in C, all being palindromes.*

We then have all the necessary tools to prove that every complete return words of a palindrome $w \in C$ is a palindrome and this implies the main result of this paper.

The fact that the number of (complete) return words is bounded by 3 for non degenerate codings of rotations can be found in the work of Keane, Rauzy or Adamczewski [9,10,1] with α irrational. Nevertheless, the proof provided takes into account rational values of α and β. We already know that $|\mathrm{Ret}(w)|$ for a nondegenerate interval exchange on k intervals is k [12], and that $|\mathrm{Ret}(w)|$ for a coding of rotation of the form $C_\alpha^\alpha(x)$ (the Sturmian case) is equal to 2. Here we handled the degenerate case as well.

References

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

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

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

[4] V. Berthé and L. Vuillon, Palindromes and two-dimensional Sturmian sequences, *J. of Automata, Language and Combinatorics*, **6** (2001), 121–138.

[5] S. Brlek, S. Hamel, M. Nivat and C. Reutenauer, On the palindromic complexity of infinite words, *Int. J. Found. Comput. Sci.* **15**: 2 (2004), 293–306

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

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

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

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

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

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

[12] L. Vuillon, On the number of return words in infinite words constructed by interval exchange transformations, *PU.M.A.* Vol. 18 No. 3-4 (2007), 345–355.