ON THE WORDS BY k TO k INSERTION OF A LETTER IN STURMIAN WORDS

Moussa Barro[1], Idrissa Kaboré[2][§], Théodore Tapsoba[3]

[1,2,3]Departement of Mathematics and Informatics
Université Nazi Boni
Bobo-Dioulasso, 01 BP 1091, BURKINA FASO

Abstract: We study the classical complexity of k to k insertion words of a letter in Sturmian words. Then, we determine the Abelian complexity and palindromic complexity of these words. Finally, we show that the k to k insertion of a letter x in Sturmian words preserves the palindromic richness of Sturmian words if and only if $k = 1$.

AMS Subject Classification: 68R15, 11B85
Key Words: Sturmian words, k to k insertion, palindromes, balance, complexity

1. Introduction

The classical complexity function, which counts the number of distinct factors of given length in an infinite word, is often used in characterization of some families of words [1]. For instance, Sturmian words are the infinite words non eventually periodic with minimal complexity. During the last thirty years, Sturmian words were the subject of intensive study. These investigations led to the discovery of numerous characterizations and various properties [3, 6, 7, 11, 14, 15] on these words. During the last two decades, palindromes were used extensively in the literature of infinite words combinatorial study (see citealbaca, droupi, kab). The notion of k to k insertion of a letter in infinite words was introduced in [13]. It consists to insert a letter x in an infinite word $\mathbf{u}$ by steps of k letters.

[§]Correspondence author

Thus, the new word obtained is called "k to k insertion word" of x in $\mathbf{u}$. The study of classical complexity and palindromic properties of k to k insertion word of a foreign letter in Sturmian words was treated in [12] and [13]. In this paper, we propose to complete this study with k to k insertion of an internal letter in Sturmian words.

The paper is organized as follow. In Section 2, we give useful definitions and notations, and we recall some properties of Sturmian words. Then, we determine the classical complexity of k to k insertion word of an internal letter in Sturmian words (Section 3). Finally, in Section 4, we establish the Abelian complexity and palindromic complexity of these words, and we show that they are not rich in palindromes except those for which $k = 1$.

2. Preliminaries

2.1. Definitions and Notations

An alphabet $\mathcal{A}$, is a non empty finite set whose the elements are called letters. A word is a finite or infinite sequence of elements of $\mathcal{A}$. The set of finite words over $\mathcal{A}$ denotes $\mathcal{A}^*$ and ε is the empty word. For any $\mathbf{u} \in \mathcal{A}^*$, the number of letters of $\mathbf{u}$ is called length of $\mathbf{u}$ and it is denoted $|\mathbf{u}|$. Moreover, for any letter x of $\mathcal{A}$, $|\mathbf{u}|_x$ is the number of occurrences of x in $\mathbf{u}$. A word $\mathbf{u}$ of length n written with a unique letter x is simply denoted $\mathbf{u} = x^n$.

Let $\mathbf{u} = u_1 u_2 \cdots u_n$ be a word such that $u_i \in \mathcal{A}$, for all $i \in \{1, 2, \cdots, n\}$. The image of $\mathbf{u}$ by the reversal map is the word denoted $\overline{\mathbf{u}}$ and defined by $\overline{\mathbf{u}} = u_n \cdots u_2 u_1$. The word $\overline{\mathbf{u}}$ is simply called reversal image of $\mathbf{u}$. A finite word $\mathbf{u}$ is called palindrome if $\overline{\mathbf{u}} = \mathbf{u}$. If $\mathbf{u}$ and $\mathbf{v}$ are two finite words over $\mathcal{A}$, we have $\overline{\mathbf{u}\mathbf{v}} = \overline{\mathbf{v}}\,\overline{\mathbf{u}}$.

The set of infinite words over $\mathcal{A}$ is denoted $\mathcal{A}^\omega$ and we write $\mathcal{A}^\infty = \mathcal{A}^* \cup \mathcal{A}^\omega$. The set of letters which appear in a word $\mathbf{u}$, is designated by $alph(\mathbf{u})$. An infinite word $\mathbf{u}$ is said to be eventually periodic if there exist two words $\mathbf{v} \in \mathcal{A}^*$ and $\mathbf{w} \in \mathcal{A}^+$ such that $\mathbf{u} = \mathbf{v}\mathbf{w}^\omega$. If $\mathbf{v} = \varepsilon$, then $\mathbf{u}$ is periodic with period $|\mathbf{w}|$. The n-th power of a finite word $\mathbf{w}$ denoted by $\mathbf{w}^n$ is the word corresponding to the concatenation $(\mathbf{w}\mathbf{w}\cdots\mathbf{w})$ n times of $\mathbf{w}$. By extension, $\mathbf{w}^0 = \varepsilon$.

Let $\mathbf{u} \in \mathcal{A}^\infty$ and $\mathbf{w} \in \mathcal{A}^*$. The word $\mathbf{w}$ is a factor of $\mathbf{u}$ if there exist $\mathbf{u}_1 \in \mathcal{A}^*$ and $\mathbf{u}_2 \in \mathcal{A}^\infty$ such that $\mathbf{u} = \mathbf{u}_1 \mathbf{w} \mathbf{u}_2$. The factor $\mathbf{w}$ is said to be a prefix (resp. a suffix) if $\mathbf{u}_1$ (resp. $\mathbf{u}_2$) is the empty word.

Let $\mathbf{u}$ be an infinite word over $\mathcal{A}$. The set of factors of $\mathbf{u}$ of length n, is written $L_n(\mathbf{u})$ and the set of all factors of $\mathbf{u}$ is denoted by $L(\mathbf{u})$.

For any infinite word $\mathbf{u}$ over $\mathcal{A}$, we shall write $\mathbf{u} = u_0 u_1 u_2 \cdots$, where $u_i \in \mathcal{A}$, $i \geq 0$.

A factor $\mathbf{w}$ of length n of an infinite word $\mathbf{u} = u_0 u_1 u_2 \cdots$ appears in $\mathbf{u}$ at the position l if $\mathbf{w} = u_l u_{l+1} \cdots u_{l+n-1}$.

A non empty factor $\mathbf{w}$ of an infinite word $\mathbf{u}$, is said to be right (resp. left) prolongable by a letter x in $\mathbf{u}$ if $\mathbf{w}x$ (resp. $x\mathbf{w}$) appears in $\mathbf{u}$. The number of right (resp. left) extensions of $\mathbf{w}$, is denoted $\partial^+\mathbf{w}$ (resp. $\partial^-\mathbf{w}$). The factor $\mathbf{w}$ is said to be right (resp. left) special factor in $\mathbf{u}$ if $\partial^+\mathbf{w} > 1$ (resp. $\partial^-\mathbf{w} > 1$). An infinite word $\mathbf{u}$ is recurrent if each of its factors appears infinitely times in $\mathbf{u}$. An infinite word $\mathbf{u}$ is said to be uniformly recurrent if any factor of $\mathbf{u}$ appears in $\mathbf{u}$ with bounded gaps, i.e, for any $n \in \mathbb{N}$, there exists an integer N such that any factor of $\mathbf{u}$ of length N contains all factors of $\mathbf{u}$ of length n.

The classical complexity function of an infinite word $\mathbf{u}$ is the map of $\mathbb{N}$ to $\mathbb{N}^*$ defined by $p_{\mathbf{u}}(n) = \#L_n(\mathbf{u})$, where $\#L_n(\mathbf{u})$ designates the cardinal of $L_n(\mathbf{u})$.

The set of palindromes of $\mathbf{u}$ of length n is denoted $PAL_n(\mathbf{u})$, and the set of all palindromes of $\mathbf{u}$, is $PAL(\mathbf{u})$. The palindromic complexity function of an infinite word $\mathbf{u}$ is the map of $\mathbb{N}$ to $\mathbb{N}$, defined by $Pal_{\mathbf{u}}(n) = \#PAL_n(\mathbf{u})$.

Let $\mathbf{u}$ be an infinite word over an alphabet $\mathcal{A} = \{a_1, a_2, \cdots, a_q\}$. The Abelian complexity function of $\mathbf{u}$, is the map of $\mathbb{N}$ to $\mathbb{N}^*$ defined by $\rho_{\mathbf{u}}^{ab}(n) = \#\left\{ (|\mathbf{w}|_{a_1}, |\mathbf{w}|_{a_2}, \cdots, |\mathbf{w}|_{a_q}) \ : \ \mathbf{w} \in L_n(\mathbf{u}) \right\}$.

The Shift, is the application S of $\mathcal{A}^\omega$ to $\mathcal{A}^\omega$ which erases the first letter of the word; for example, $S(abaababaa \cdots) = baababaa \cdots$.

A morphism f is a map of $\mathcal{A}^*$ into itself such that $f(\mathbf{uv}) = f(\mathbf{u})f(\mathbf{v})$ for any $\mathbf{u}, \mathbf{v} \in \mathcal{A}^*$.

2.2. Sturmian Words

In this section we consider the binary alphabet $\mathcal{A} = \{a, b\}$.

Definition 1. An infinite word $\mathbf{u}$ over $\mathcal{A}$ is a Sturmian word if for any natural n, $p_{\mathbf{u}}(n) = n + 1$.

The most well-known Sturmian word is the famous Fibonacci word. It is generated by the morphism φ defined by $\varphi(a) = ab$ and $\varphi(b) = a$.

Definition 2. A Sturmian word $\mathbf{u}$ is said to be a-Sturmian (resp. b-Sturmian) when it contains a^2 (resp. b^2).

Definition 3. An infinite word $\mathbf{u} = u_0 u_1 u_2 \cdots$ is said to be modulo-

recurrent if for all $i \geq 1$, any factor $\mathbf{w}$ of $\mathbf{u}$, appears in $\mathbf{u}$ at all positions modulo i, i.e.

$$\forall i \in \{0, 1, 2, \cdots, l-1\}, \exists g_i \in \mathbb{N} : \mathbf{w} = u_{lg_i+i}u_{lg_i+i+1} \cdots u_{lg_i+i+|w|-1}.$$

Factors of length n occurring in $\mathbf{u}$ at a position which is multiple of n, as above, are called n-window factors of $\mathbf{u}$.

Proposition 1. ([7]) *Let $\mathbf{u}$ be a modulo-recurrent word. Then, for all n, the set of n-window factors of $\mathbf{u}$ is equal to $L_n(u)$.*

Definition 4. A word $\mathbf{u}$ is said to be α-balanced if α is the smallest natural such that for any pair $(\mathbf{v}, \mathbf{w})$ of factors of $\mathbf{u}$ of same length and for any letter x:

$$||\mathbf{v}|_x - |\mathbf{w}|_x| \leq \alpha.$$

If $\alpha = 1$, $\mathbf{u}$ is said to be balanced.

The following theorem presents some classical properties of Sturmian words.

Theorem 2. *Let $\mathbf{u}$ be a Sturmian word. Then, we have the following properties:*

(1) ([8]) $\mathbf{u}$ is not eventually periodic and balanced;

(2) ([8]) For all $n \in \mathbb{N}^*$, $\rho_{\mathbf{u}}^{ab}(n) = 2$;

(3) ([9]) For all $n \in \mathbb{N}$,

$$pal_{\mathbf{u}}(n) = \begin{cases} 1 & \text{if } n \text{ is even} \\ 2 & \text{otherwise} \end{cases};$$

((4) ([13]) $\mathbf{u}$ is modulo-recurrent.

3. Classical Complexity

In the following, k designates a positive integer.

Definition 5. Let $\mathbf{u}$ be an infinite word over $\mathcal{A}$. Let us decompose $\mathbf{u}$ by words of constant length. We have:
$$\mathbf{u} = \mathbf{w}_0 \mathbf{w}_1 \mathbf{w}_2 \mathbf{w}_3 \cdots \mathbf{w}_i \cdots, \ \mathbf{w}_i \in L_k(\mathbf{u}), \ i \in \mathbb{N}.$$

Now, insert in $\mathbf{u}$, between two consecutive factors $\mathbf{w}_i$ and $\mathbf{w}_{i+1}$ a fixed letter x. We get the word $\mathbf{v} = x\mathbf{w}_0 x\mathbf{w}_1 x\mathbf{w}_2 x\mathbf{w}_3 x \cdots x\mathbf{w}_i x \cdots$. This word is called "$k$ to k insertion word" of x in $\mathbf{u}$ and it is denoted $I_k^x(\mathbf{u})$.

In [12], the following result is established.

Proposition 3. *Let $\mathbf{u}$ be a Sturmian word over and $\mathbf{v}$ a k to k insertion word of x in $\mathbf{u}$.*

(1) *If $\mathbf{ps}$ is a factor of $\mathbf{u}$ such that $\mathbf{p}$ and $\mathbf{s}$ are words satisfying $|\mathbf{p}| + |\mathbf{s}| \leq k$, then $\mathbf{p}x\mathbf{s}$ is in $\mathbf{v}$.*

(2) *If $\mathbf{pw}_1 \cdots \mathbf{w}_q \mathbf{s}$ is a factor of $\mathbf{u}$ such that $\mathbf{p}$, $\mathbf{s}$ and $\mathbf{w}_i$ are words satisfying $|\mathbf{p}| + |\mathbf{s}| < k$ and $\mathbf{w}_i = k$, $i = 1, \cdots, q$, then, $\mathbf{p}x\mathbf{w}_1 x \cdots \mathbf{w}_q x\mathbf{s}$ is in $\mathbf{v}$.*

Lemma 4. *Let $\mathbf{v}$ be a k to k insertion word of x in a Sturmian word $\mathbf{u}$. Then, any factor of $\mathbf{v}$ is stable by the reversal map.*

Proof. Let $\mathbf{v}_1 \in L(\mathbf{v})$. Then, we can write

$$\mathbf{v}_1 = \mathbf{w}_1 x\mathbf{w}_2 x\mathbf{w}_3 x \cdots x\mathbf{w}_{n-1} x\mathbf{w}_n$$

with $|\mathbf{w}_i| = k$ for all $i \in \{2, 3, \cdots, n-1\}$ and $|\mathbf{w}_1|, |\mathbf{w}_n| \leq k$. Thus, the word $\mathbf{u}_1 = \mathbf{w}_1 \mathbf{w}_2 \cdots \mathbf{w}_{n-1} \mathbf{w}_n$ appears in $\mathbf{u}$. Since $\mathbf{u}$ is a Sturmian word, then $\overline{\mathbf{u}}_1 = \mathbf{w}_n \mathbf{w}_{n-1} \cdots \mathbf{w}_2 \mathbf{w}_1$ also appears in $\mathbf{u}$ [4]. So, $\mathbf{w}_n x\mathbf{w}_{n-1} x \cdots x\mathbf{w}_2 x\mathbf{w}_1$ appears in $\mathbf{v}$ by Proposition 1, since $\mathbf{u}$ is a Sturmian word. In other terms $\overline{\mathbf{v}}_1 \in L(\mathbf{v})$. $\square$

Lemma 5. *Let $\mathbf{u}$ be a Sturmian word over $\{a, b\}$ and $\mathbf{v}$ a k to k insertion word of x in $\mathbf{u}$. Then, two factors of $\mathbf{v}$ with same length come from factors of $\mathbf{u}$ with same or consecutive lengths.*

Proof. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two factors of $\mathbf{v}$ of same length coming from factors $\mathbf{u}_1$ and $\mathbf{u}_2$ of $\mathbf{u}$ respectively. So, we can write

$$\mathbf{v}_1 = \mathbf{m}_0 \underline{x} \mathbf{m}_1 \underline{x} \mathbf{m}_2 \underline{x} \mathbf{m}_3 \underline{x} \cdots \underline{x} \mathbf{m}_h$$

and

$$\mathbf{v}_2 = \mathbf{w}_0 \underline{x} \mathbf{w}_1 \underline{x} \mathbf{w}_2 \underline{x} \mathbf{w}_3 \underline{x} \cdots \underline{x} \mathbf{w}_n.$$

Thus, we have

$$\mathbf{u}_1 = \mathbf{m}_0\mathbf{m}_1\mathbf{m}_2\mathbf{m}_3\cdots\mathbf{m}_h$$

and

$$\mathbf{u}_2 = \mathbf{w}_0\mathbf{w}_1\mathbf{w}_2\mathbf{w}_3\cdots\mathbf{w}_n,$$

where $|\mathbf{m}_i| = |\mathbf{w}_j| = k$, $i = 1, \cdots, h-1$; $j = 1, \cdots, n-1$ and $|\mathbf{m}_0|$, $|\mathbf{m}_h|$, $|\mathbf{w}_0|$, $|\mathbf{w}_n| \leq k$. Therefore, $|\mathbf{v}_1| = |\mathbf{u}_1| + h$ and $|\mathbf{v}_2| = |\mathbf{u}_2| + n$. Let us show that $||\mathbf{u}_1| - |\mathbf{u}_2|| \leq 1$.

If $n = h$, then $|\mathbf{u}_1| = |\mathbf{u}_2|$ because $|\mathbf{v}_1| = |\mathbf{v}_2|$.

Suppose $\alpha = ||\mathbf{u}_1| - |\mathbf{u}_2||$ with $\alpha \geq 1$. Then, we have $|n - h| = \alpha$ since $|\mathbf{v}_1| = |\mathbf{v}_2|$. Without loss of generality, we can assume that $n = h + \alpha$. Thus, we have
$\mathbf{v}_2 = \mathbf{w}_0\underline{x}\mathbf{w}_1\underline{x}\mathbf{w}_2\underline{x}\mathbf{w}_3\underline{x}\cdots\underline{x}\mathbf{w}_h\underline{x}\mathbf{w}_{h+1}\underline{x}\cdots\underline{x}\mathbf{w}_{h+\alpha}$ and $|\mathbf{v}_1| = |\mathbf{m}_0| + k(h-1) + h + |\mathbf{m}_h|$, $|\mathbf{v}_2| = |\mathbf{w}_0| + k(h-1) + h + k\alpha + \alpha + |\mathbf{w}_{h+\alpha}|$.

Since $|\mathbf{v}_1| = |\mathbf{v}_2|$, it follows that $|\mathbf{m}_0| + |\mathbf{m}_h| = |\mathbf{w}_0| + k\alpha + \alpha + |\mathbf{w}_{h+\alpha}|$. So, $k\alpha + \alpha = |\mathbf{m}_0| + |\mathbf{m}_h| - (|\mathbf{w}_0| + |\mathbf{w}_{h+\alpha}|)$. Moreover, we have $|\mathbf{m}_0|$, $|\mathbf{m}_h|$, $|\mathbf{w}_0|$, $|\mathbf{w}_n| \leq k$. Hence, it follows the inequality $k\alpha + \alpha \leq 2k$. Therefore, we obtain $\alpha \leq 1$ and $||\mathbf{u}_1| - |\mathbf{u}_2|| \leq 1$. $\square$

In [13], the two of the co-authors have proved the following result.

Theorem 6.   *Let $\mathbf{u}$ be a Sturmian word over $\mathcal{A} = \{a, b\}$. The classical complexity function of k to k insertion word $\mathbf{v}$ of $c \notin \mathcal{A}$ in $\mathbf{u}$, is given by:*

$$p_{\mathbf{v}}(n) = \left\{ \begin{array}{ll} n^2 + n + 1 & if \quad n \leq k \\ kn + k + 1 & if \quad n > k \end{array} \right. .$$

Here, we propose to extend this result to k to k insertion words of $x \in \{a, b\}$ in Sturmian words.

Lemma 7.   *Let $\mathbf{u}$ be an a-Sturmian word over $\{a, b\}$ and $\mathbf{u}_1$ a factor of $\mathbf{u}$ such that $|\mathbf{u}_1| \leq k$. If the insertion from two different positions of x in $\mathbf{u}_1$ produces the same factor, then the factor which separates these two positions in $\mathbf{u}_1$ is a power of x.*

Proof. Let $\mathbf{u}_1$ be a factor of $\mathbf{u}$ such that $|\mathbf{u}_1| \leq k$. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two words obtained by k to k insertion of x in $\mathbf{u}_1$ from two different positions such that $\mathbf{v}_1 = \mathbf{v}_2$. Then, $\mathbf{u}_1$ can be decomposed in the form $\mathbf{u}_1 = \mathbf{pts}$ with $|\mathbf{t}| > 0$, $|\mathbf{p}|$, $|\mathbf{s}| \geq 0$ and $\mathbf{v}_1 = \mathbf{p}x\mathbf{ts}$, $\mathbf{v}_2 = \mathbf{pt}x\mathbf{s}$. Since $\mathbf{v}_1 = \mathbf{v}_2$, then $\mathbf{p}x\mathbf{ts} = \mathbf{pt}x\mathbf{s}$. So, we have $x\mathbf{t} = \mathbf{t}x$. Thus, $\mathbf{t}$ is a power of x since x is a letter. $\square$

Lemma 8. *Let* **u** *be an a-Sturmian word over* $\{a, b\}$ *and* $\mathbf{u}_1$, *a factor of* **u** *sufficiently long. Then, the k to k insertions of x in $\mathbf{u}_1$ from two different positions produce distinct factors of $I_k^x(\mathbf{u})$.*

Proof. Let $\mathbf{u}_1$ be a factor of **u** such that $\mathbf{u}_1$ contains all the k-window factors of **u**. The word **u** is modulo-recurrent by item (4) of Theorem 2 since it is Sturmian. So, any insertion of x in $\mathbf{u}_1$ produces a factor of $I_k^x(\mathbf{u})$. Consider two factors $\mathbf{v}_1$ and $\mathbf{v}_2$ of $I_k^x(\mathbf{u})$, produced by k to k insertion of x from two different positions in $\mathbf{u}_1$ such that $\mathbf{v}_1 = \mathbf{v}_2$. Then, $\mathbf{v}_1$ and $\mathbf{v}_2$ contain the same number of inserted letters of x.

For $x = a$, by Lemma 7, there exists an integer i_0 such that $\mathbf{u}_1$ can be written in the form $\mathbf{u}_1 = \mathbf{m}_0 a^{i_0} \mathbf{m}_1 a^{i_0} \mathbf{m}_2 a^{i_0} \cdots a^{i_0} \mathbf{m}_n$, where $|\mathbf{m}_i| = |\mathbf{m}_j| = k - i_0$, $i, j \in \{1, \cdots, n-1\}$ and $|\mathbf{m}_o|, |\mathbf{m}_n| \leq k - i_0$. Thus, we obtain $\mathbf{v}_1 = \mathbf{m}_0 \underline{a} a^{i_0} \mathbf{m}_1 \underline{a} a^{i_0} \mathbf{m}_2 \underline{a} a^{i_0} \cdots \underline{a} a^{i_0} \mathbf{m}_n$ and
$\mathbf{v}_2 = \mathbf{m}_0 a^{i_0} \underline{a} \mathbf{m}_1 a^{i_0} \underline{a} \mathbf{m}_2 a^{i_0} \underline{a} \cdots a^{i_0} \underline{a} \mathbf{m}_n$, where $\underline{a}$ is the inserted letter a of $\mathbf{u}_1$. Let **t** be a special factor of $\mathbf{u}_1$ of length $k - i_0$. Since $\mathbf{u}_1$ contains all the k-window factors of **u**, then there exists a word $\mathbf{m}_{j_0}$ with $1 < j_0 < n$, such that $\mathbf{m}_{j_0} = \mathbf{t}$. This contradicts the fact that the words $\mathbf{m}_2, \cdots, \mathbf{m}_{n-1}$ extend to right in $\mathbf{u}_1$ with the letter a.

For $x = b$, by Lemma 7, $\mathbf{u}_1$ can be written in the form

$$\mathbf{u}_1 = \mathbf{m}_0 b \mathbf{m}_1 b \mathbf{m}_2 b \cdots b \mathbf{m}_n$$

with $|\mathbf{m}_i| = |\mathbf{m}_j| = k - 1$, $0 < i, j < n$ and $|\mathbf{m}_o|, |\mathbf{m}_n| \leq k - 1$.

Therefore, we have

$$\mathbf{v}_1 = \mathbf{m}_0 \underline{b} b \mathbf{m}_1 \underline{b} b \mathbf{m}_2 \underline{b} b \cdots \underline{b} b \mathbf{m}_n$$

and

$$\mathbf{v}_2 = \mathbf{m}_0 b \underline{b} \mathbf{m}_1 b \underline{b} \mathbf{m}_2 b \underline{b} \cdots b \underline{b} \mathbf{m}_n.$$

Observe that the words $\mathbf{m}_2, \cdots, \mathbf{m}_{n-1}$ extend to right and to left in $\mathbf{u}_1$ with the letter b only. So, they begin and end by a letter a. Thus, for all $i = 2, \cdots, n-1$, we have $S(\mathbf{m}_i) b \in L_{k-1}(\mathbf{u})$. Since **u** is balanced, there exist two factors $\mathbf{t}_1$ and $\mathbf{t}_2$ of $\mathbf{u}_1$ of length $k - 1$ such that $|\mathbf{t}_1|_b - |\mathbf{t}_2|_b = 1$. In addition, since $\mathbf{u}_1$ contains all the k-window factors of **u**, then there exist two words $\mathbf{m}_{i_0}$ and $\mathbf{m}_{j_0}$, $1 < i_0, j_0 < n$ such that $\mathbf{m}_{i_0} = \mathbf{t}_1$ and $\mathbf{m}_{j_0} = \mathbf{t}_2$. It follows that $|S(\mathbf{m}_{i_0}) b|_b - |\mathbf{m}_{j_0}|_b = 2$. This contradicts the fact that **u** is balanced. Hence, we deduce that $\mathbf{v}_1$ and $\mathbf{v}_2$ are distinct factors. $\square$

Lemma 9.   *Let* $\mathbf{u}$ *be an a-Sturmian word and* $\mathbf{u}_1, \mathbf{u}_2$ *two distinct factors of* $\mathbf{u}$ *sufficiently long. Then, the k to k insertion words of* x *in the words* $\mathbf{u}_1$ *and* $\mathbf{u}_2$ *are distinct factors.*

Proof.   Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two distinct factors of $\mathbf{u}$ such that $\mathbf{u}_1$ and $\mathbf{u}_2$ contain all the k-window factors of $\mathbf{u}$. Decompose $\mathbf{u}_1 = \mathbf{m}_0 \mathbf{m}_1 \mathbf{m}_2 \cdots \mathbf{m}_h$ and $\mathbf{u}_2 = \mathbf{w}_0 \mathbf{w}_1 \mathbf{w}_2 \cdots \mathbf{w}_n$, where $|\mathbf{w}_i| = |\mathbf{m}_j| = k$, $0 < i < h$, $0 < j < n$ and $|\mathbf{w}_o|, |\mathbf{m}_o|, |\mathbf{w}_n|, |\mathbf{m}_h| \le k$. Consider $\mathbf{v}_1 = \mathbf{m}_0 \underline{x} \mathbf{m}_1 \underline{x} \mathbf{m}_2 \underline{x} \cdots \underline{x} \mathbf{m}_n$ and $\mathbf{v}_2 = \mathbf{w}_0 \underline{x} \mathbf{w}_1 \underline{x} \mathbf{w}_2 \underline{x} \cdots \underline{x} \mathbf{w}_h$ obtained by k to k insertion of x in the factors $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively. Suppose $\mathbf{v}_1 = \mathbf{v}_2$. By Lemma 5, we have $||\mathbf{u}_1| - |\mathbf{u}_2|| \le 1$ since $|\mathbf{v}_1| = |\mathbf{v}_2|$.

If $|\mathbf{m}_o| = |\mathbf{w}_o|$, then we have $h = n$ since $|\mathbf{v}_1| = |\mathbf{v}_2|$. Thus, we obtain $\mathbf{m}_0 = \mathbf{w}_0, \mathbf{m}_1 = \mathbf{w}_1, \cdots, \mathbf{m}_h = \mathbf{w}_n$, i.e. $\mathbf{u}_1 = \mathbf{u}_2$. This is impossible because $\mathbf{u}_1$ and $\mathbf{u}_2$ are distinct factors.

Now, consider $|\mathbf{w}_o| < |\mathbf{m}_o|$. We can write

$$\mathbf{m}_0 = \mathbf{w}_0 x \mathbf{m}_0',$$

$$\mathbf{v}_1 = \mathbf{w}_0 x \mathbf{m}_0' \underline{x} \mathbf{m}_1 \underline{x} \mathbf{m}_2 \underline{x} \cdots \underline{x} \mathbf{m}_n,$$

and

$$\mathbf{v}_2 = \mathbf{w}_0 \underline{x} \mathbf{w}_1 \underline{x} \mathbf{w}_2 \underline{x} \cdots \underline{x} \mathbf{w}_h.$$

Since $\mathbf{v}_1 = \mathbf{v}_2$, we have $\mathbf{w}_i = \mathbf{w}_i' x \mathbf{w}_i''$, where the letter x coincides with the inserted $\underline{x}$ in $\mathbf{m}_{i-1} \underline{x} \mathbf{m}_i$, for $i = 1, \cdots, n - 1$.

Similarly, $\mathbf{m}_j = \mathbf{m}_j' x \mathbf{m}_j''$ for $j = 1, \cdots, h - 1$ with $x \in \{a, b\}$. So, $\mathbf{m}_i' = \mathbf{w}_i''$ and $\mathbf{m}_i'' = \mathbf{w}_{i+1}'$, for all $i = 1, \cdots, n - 1$. Thus, we can write

$$\mathbf{u}_1 = \mathbf{w}_0 x \mathbf{w}_1' \mathbf{w}_1'' x \mathbf{w}_2' \mathbf{w}_2'' x \mathbf{w}_3' \mathbf{w}_3'' x \cdots x \mathbf{w}_n' \mathbf{w}_n''$$

and

$$\mathbf{u}_2 = \mathbf{w}_0 \mathbf{w}_1' x \mathbf{w}_1'' \mathbf{w}_2' x \mathbf{w}_2'' \mathbf{w}_3' x \mathbf{w}_3'' \cdots \mathbf{w}_n' x \mathbf{w}_n''.$$

Moreover, $|\mathbf{w}_i'| = |\mathbf{w}_j'|$ and $|\mathbf{w}_i''| = |\mathbf{w}_j''|$ for all $i, j = 1, \cdots, n - 1$.

If $x = b$, then b is the unique right extension of $\mathbf{w}_i'$ in $\mathbf{u}_1$. So, $\mathbf{w}_i''$ begins with b in $\mathbf{u}_2$. This is not possible, otherwise $\mathbf{u}_2$ will contain b^2 because b is the unique left extension of $\mathbf{w}_i''$ in $\mathbf{u}_1$.

If $x = a$, we have $\mathbf{u}_1 = \mathbf{w}_0 |a \mathbf{w}_1' \mathbf{w}_1''| a \mathbf{w}_2' \mathbf{w}_2''| a \mathbf{w}_3' \mathbf{w}_3''| a \cdots |a \mathbf{w}_n' \mathbf{w}_n''$. So, the factors $a \mathbf{w}_i' \mathbf{w}_i''$ begin with a in $\mathbf{u}_1$, for all $i = 1, \cdots, n$. Thus, from Proposition 1, we obtain a contradiction since $\mathbf{u}_2$ contains all the k-window factors of $\mathbf{u}$. In conclusion, $\mathbf{v}_1$ and $\mathbf{v}_2$ are distinct factors of $\mathbf{v}$. $\square$

Theorem 10. *Let* **u** *be a Sturmian word. Then, the classical complexity of k to k insertion word of x in* **u**, *is eventually $kn + k + 1$.*

Proof. By Lemma 9, there exists an integer n_0 such that two factors of **u** of length greater than n_0, produce by insertion of x distinct factors of $I_k^x(\mathbf{u})$. Let l_0 be the minimal length of factors produced by those with length n_0. Consequently, using Theorem 6, we deduce that the classical complexity of $I_k^x(\mathbf{u})$ satisfies $kn + k + 1$ for all $n \geq l_0$. □

4. Abelian Complexity and Palindromic Complexity

This section is devoted to balance, Abelian complexity and palindromicity in words obtained by k to k insertion of a letter in Sturmian words.

4.1. Balance and Abelian Complexity

In this paragraph, we consider the alphabet $\{a, b, c\}$.

Theorem 11. *Let* **u** *be a Sturmian word over $\{a, b\}$, and the k to k insertion word* **v** *of x in* **u**. *Then,* **v** *is 2-balanced.*

Proof. Let $\mathbf{u}_1$ and $\mathbf{u}_2$ be two factors of **u**. Let $\mathbf{v}_1$ and $\mathbf{v}_2$ be two k to k insertion words of x in $\mathbf{u}_1$ and $\mathbf{u}_2$ respectively such that $|\mathbf{v}_1| = |\mathbf{v}_2|$. So, by Lemma 5, we have $||\mathbf{u}_1| - |\mathbf{u}_2|| \leq 1$ and the factors $\mathbf{v}_1$ and $\mathbf{v}_2$ contain q or $q + 1$ inserted letters x. First, let us show that $||\mathbf{u}_1| - |\mathbf{u}_2|| \leq 1$ implies $||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 2$ for any $y \in \{a, b\}$. Without loss of generality, we can assume that $|\mathbf{u}_2| \geq |\mathbf{u}_1|$. Thus, we can write $\mathbf{u}_2 = \mathbf{u}_2' z$ with $z \in \{\varepsilon, x, y\}$. It follows that $||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| = ||\mathbf{u}_1|_y - |\mathbf{u}_2' z|_y| = ||\mathbf{u}_1|_y - (|\mathbf{u}_2'|_y + |z|_y)| \leq ||\mathbf{u}_1|_y - |\mathbf{u}_2'|_y| + |z|_y \leq 2$.

Consider $x = c$. Then, we have $||\mathbf{v}_1|_c - |\mathbf{v}_2|_c| \leq 1$, since $||\mathbf{u}_1| - |\mathbf{u}_2|| \leq 1$ and $|\mathbf{v}_1| = |\mathbf{v}_2|$. Let $y \in \{a, b\}$, then $|\mathbf{v}_1|_y = |\mathbf{u}_1|_y$ and $|\mathbf{v}_2|_y = |\mathbf{u}_2|_y$ because y is not an inserted letter in **v**. So, $||\mathbf{v}_1|_y - |\mathbf{v}_2|_y| = ||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 2$.

Now, consider $x \in \{a, b\}$. We have two cases to discuss.

Case 1: $|\mathbf{u}_1| = |\mathbf{u}_2|$. Let $y \in \{a, b\}$.

• Suppose $y \neq x$. Then, $|\mathbf{v}_1|_y = |\mathbf{u}_1|_y$ and $|\mathbf{v}_2|_y = |\mathbf{u}_2|_y$. Thus, we have $||\mathbf{v}_1|_y - |\mathbf{v}_2|_y| = ||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 1$ because **u** is balanced.

• Suppose $y = x$. Then, $|\mathbf{v}_1|_y = |\mathbf{u}_1|_y + q$ and $|\mathbf{v}_2|_y = |\mathbf{u}_2|_y + q$, where q is the number of insertions of x.

So, $||\mathbf{v}_1|_y - |\mathbf{v}_2|_y| = ||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 1$, since **u** is balanced.

Case 2: $||\mathbf{u}_1| - |\mathbf{u}_2|| = 1$. Let $y \in \{a, b\}$. By Lemma 5, we have $||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 2$.

• Suppose $y \neq x$, then $||\mathbf{v}_1|_y - |\mathbf{v}_2|_y| = ||\mathbf{u}_1|_y - |\mathbf{u}_2|_y| \leq 2$, since $|\mathbf{v}_1|_y = |\mathbf{u}_1|_y$ and $|\mathbf{v}_2|_y = |\mathbf{u}_2|_y$.

• Suppose $y = x$. Then, it follows:

$$
\begin{aligned}
||\mathbf{v}_1|_y - |\mathbf{v}_2|_y| &= |(|\mathbf{u}_1|_y + q + 1) - (|\mathbf{u}_2|_y + q)| \quad \text{with } |\mathbf{u}_2| = |\mathbf{u}_1| + 1 \\
&= ||\mathbf{u}_1|_y + 1 - |\mathbf{u}_2|_y| \\
&= ||\mathbf{u}_1|_y + 1 - |\mathbf{u}_2'z|_y| \text{ with } \mathbf{u}_2'z = \mathbf{u}_2 \text{ and } z \in \{a, b\} \\
&= ||\mathbf{u}_1|_y - |\mathbf{u}_2'|_y + (1 - |z|_y)| \\
&\leq ||\mathbf{u}_1|_y - |\mathbf{u}_2'|_y| + |1 - |z|_y| \\
&\leq 1 + 1 = 2.
\end{aligned}
$$

Furthermore, by Theorem 10, observe that $\mathbf{v}$ is neither eventually periodic nor a Sturmian word. So, there exists a pair $(\mathbf{v}_1, \mathbf{v}_2)$ of factors of $\mathbf{v}$ and a letter x such that $||\mathbf{v}_1|_x - |\mathbf{v}_2|_x| \geq 2$. Thus, the value 2 is reached. $\qquad \square$

Lemma 12. *Let $\mathbf{u}$ be a Sturmian word and $\mathbf{v}$, the k to k insertion word of x in $\mathbf{u}$. Then, any factor of $\mathbf{v}$ of length $(k+1)q$, is produced by a factor of $\mathbf{u}$ of length kq, $q \geq 0$.*

Proof. Let $\mathbf{u}_1$ be a factor of $\mathbf{u}$ and $\mathbf{v}_1$, a k to k insertion word of x in $\mathbf{u}_1$ such that $|\mathbf{v}_1| = (k+1)q$, $q \geq 0$. First, let us observe that any factor of $\mathbf{v}$ of length $k + 1$ contains only one insertion of x and comes from a factor of $\mathbf{u}$ of length k. Since $|\mathbf{v}_1| = (k+1)q$, $\mathbf{v}_1$ contains exactly q inserted letters. In addition, we have $|\mathbf{v}_1| = (k+1)q = kq + q$. Thus, it results that $|\mathbf{u}_1| = kq$. $\quad \square$

Theorem 13. *Let $\mathbf{u}$ be a Sturmian word over $\{a, b\}$ and $\mathbf{v}$, the k to k insertion word of x in $\mathbf{u}$. Then, the Abelian complexity of $\mathbf{v}$ verifies:*

• *For $x \in \{a, b\}$, we have $\rho_{\mathbf{v}}^{ab}(n) \in \{2, 3\}$, for all $n \geq 1$.*

• *For $x = c$, we have :*

$$
\rho_{\mathbf{v}}^{ab}(n) = \left\{
\begin{array}{ll}
3 & \text{if } n = 1 \\
2 & \text{if } n \text{ is multiple of } k + 1 \\
4 & \text{otherwise}
\end{array}
\right. .
$$

Proof. The word $\mathbf{v}$ being non eventually periodic, its Abelian complexity verifies $\rho_{\mathbf{v}}^{ab}(n) \geq 2$, for all $n \geq 1$. Consider $n \geq 1$ such that $n = (k+1)q$, where $q \geq 1$. By Lemma 12, we have

$$
L_n(\mathbf{v}) = \{I_k^x(\mathbf{u}_1) : \mathbf{u}_1 \in L_{kq}(\mathbf{u})\}
$$

for any $x \in \{a, b, c\}$. Thus, it follows $\rho_{\mathbf{v}}^{ab}(n) = \rho_{\mathbf{u}}^{ab}(kq)$. So, by item (2) of Theorem 2, we deduce $\rho_{\mathbf{v}}^{ab}(n) = 2$, for any $x \in \{a, b, c\}$.

For $x \in \{a, b\}$, we deduce that $\rho_{\mathbf{v}}^{ab}(n) \leq 3$, since $\mathbf{v}$ is a 2-balanced binary word.

Now, let us take $x = c$.

If $n = 1$, then $L_1(\mathbf{v}) = \{a, b, c\}$ and so, $\rho_{\mathbf{v}}^{ab}(1) = 3$.

If $1 < n \leq k$, then $L_n(\mathbf{v}) = L_n(\mathbf{u}) \cup \{\mathbf{u}_1 c \mathbf{u}_2 \; : \; \mathbf{u}_1 \mathbf{u}_2 \in L_{n-1}(\mathbf{u})\}$. So, $\rho_{\mathbf{v}}^{ab}(n)) = \rho_{\mathbf{n}}^{ab}(n) + \rho_{\mathbf{u}}^{ab}(n-1)$. Hence, $\rho_{\mathbf{v}}^{ab}(n) = 2 + 2 = 4$, by item (2) of Theorem 2.

If $n > k + 1$ and $k + 1$ does not divide n, then we can write $n = (k + 1)q + r + 1$ where $q \geq 1$ and $0 \leq r < k$. So, by Lemma 5, it follows that $L_n(\mathbf{v}) = \{I_k^c(\mathbf{u}_1) : \mathbf{u}_1 \in L_{kq+r}(\mathbf{u})\} \cup \{I_k^c(\mathbf{u}_2) : \mathbf{u}_2 \in L_{kq+r+1}(\mathbf{u})\}$. Therefore, $\rho_{\mathbf{v}}^{ab}(n) = \rho_{\mathbf{u}}^{ab}(kq + r) + \rho_{\mathbf{u}}^{ab}(kq + r + 1)$. From item (2) of Theorem 2, we deduce $\rho_{\mathbf{v}}^{ab}(n) = 4$. $\square$

4.2. Palindromic Properties

In this subsection, we show that the palindromes of a word obtained by k to k insertion of some letter in a Sturmian word are produced by some palindromes of the related Sturmian word. After, we determine the palindromic complexity of these words and prove that they are not rich in palindromes in general.

The following theorem is shown in [12].

Theorem 14. *Let $\mathbf{u}$ be a Sturmian word over $\mathcal{A}$ and $c \notin \mathcal{A}$.*

(1) If k is even, the palindrome complexity of $I_k^c(\mathbf{u})$, Pal, is given by:

$$\forall n \in \mathbb{N} \; : \; Pal(n) = \begin{cases} 1 & if \;\; n \;\; is \;\; even, \;\; n \leq k + 1 \\ 3 & if \;\; n \;\;\; is \;\; odd, \;\; n \leq k + 1 \\ 1 & if \;\; n > k + 1 \end{cases}.$$

(2) If k is odd, the palindrome complexity of $I_k^c(\mathbf{u})$, Pal, is given by:

$$\forall n \in \mathbb{N} \; : \; Pal(n) = \begin{cases} 1 & if \;\; n \;\; is \;\; even, \;\; n \leq k + 1 \\ 0 & if \;\; n \;\; is \;\; even, \;\; n > k + 1 \\ 3 & if \;\; n \;\; is \;\; odd \end{cases}.$$

Lemma 15. *Let $\mathbf{u}$ be a Sturmian word and $\mathbf{v}$, the k to k insertion word of x in $\mathbf{u}$. Then, any palindrome of $\mathbf{v}$ sufficiently long, is produced by a palindrome of $\mathbf{u}$.*

Proof. Let $\mathbf{v}_1$ be a palindrome of $\mathbf{v}$ sufficiently long. Then, $\mathbf{v}_1$ comes from a factor $\mathbf{u}_1$ of $\mathbf{u}$. Since $\mathbf{u}$ is a Sturmian word, then $\overline{\mathbf{u}}_1$ is a factor of $\mathbf{u}$. Hence, $\overline{\mathbf{v}}_1$ comes from $\overline{\mathbf{u}}_1$. So, as $\mathbf{v}_1$ is a palindrome, then $\mathbf{v}_1$ comes from $\overline{\mathbf{u}}_1$. Since $\mathbf{v}_1$ is sufficiently long, by Lemma 9, $\mathbf{v}_1$ comes from an only one factor of $\mathbf{u}$ and so, $\overline{\mathbf{u}}_1 = \mathbf{u}_1$.   □

Theorem 16. *Let $\mathbf{u}$ be a Sturmian word and $\mathbf{v}$, the k to k insertion word of x in $\mathbf{u}$. Then, there exists $n_0 \geq 1$ such that for $n \geq n_0$, we have:*

$$Pal_{\mathbf{v}}(n) = \begin{cases} 1 & if \quad k \quad is \quad even \\ 3 & if \quad k \quad and \quad n \quad are \quad odd \\ 0 & if \quad k \quad is \quad odd \quad and \quad n \quad is \quad even \end{cases}.$$

Proof. Let $\mathbf{u}$ be a Sturmian word and $\mathbf{v}$ the k to k insertion word of the letter x in $\mathbf{u}$. According to Lemma 9, any factor of $\mathbf{v}$ sufficiently long comes from a unique factor of $\mathbf{u}$. Thus, by Theorem 14, we obtain the result.   □

Let $\mathbf{w}$ be a finite word. We denote by $P(\mathbf{w})$, the number of distinct palindromes contained in $\mathbf{w}$. We have $P(\mathbf{w}) = \sum_{n=0}^{|\mathbf{w}|} Pal_{\mathbf{w}}(n)$.

Definition 6. A word $\mathbf{u}$ (finite or infinite) is said to be rich if for any factor $\mathbf{w}$ of $\mathbf{u}$, we have $P(\mathbf{w}) = |\mathbf{w}| + 1$.

The following result is established in [5].

Theorem 17. *Let $\mathbf{u}$ be an infinite word such that the set of factors is stable by the reversal map. Then, $\mathbf{u}$ is rich if and only if*

$$\forall n \in \mathbb{N}, Pal_{\mathbf{u}}(n) + Pal_{\mathbf{u}}(n+1) = p_{\mathbf{u}}(n+1) - p_{\mathbf{u}}(n) + 2.$$

Theorem 18. *Let $\mathbf{u}$ be a Sturmian word. Then, the word $I_k^x(\mathbf{u})$ is rich if and only if $k = 1$.*

Proof. Let $\mathbf{u}$ be a Sturmian word. From Lemma 9, the word $I_k^x(\mathbf{u})$ is stable by the reversal map.

• For k even, using Theorems 10, 14 and 16, $\mathbf{v}$ is rich if and only if we have the following equivalences:

$$1 + 1 = k(n+1) + k + 1 - (kn + k + 1) + 2 \iff 2 = k + 2$$

$$\Longleftrightarrow \quad k = 0.$$

This means that $\mathbf{v}$ is not rich.

• For k odd, using Theorems 10, 14 and 16, $\mathbf{v}$ is rich if and only if we have the following equivalences:

$$3 = k(n+1) + k + 1 - (kn + k + 1) + 2 \quad \Longleftrightarrow \quad 3 = k + 2$$
$$\Longleftrightarrow \quad k = 1.$$

□

References

[1] J.-P. Allouche, Sur la complexité des suites infinies, *Bull. Belg. Math. Soc. Simon Stevin*, **1** (1994), 133-143.

[2] J.-P. Allouche, M. Baake, J. Cassaigne, D. Damanik, Palindrome complexity, *Theoret. Comput. Sci.*, **292** (2003), 9-31.

[3] J. Berstel, Sturmian and episturmian words (a survery of some recent results), In: *Proceedings of CAI 2007*, Vol. 4728 of *Lecture Notes in Computer Science*, Springer-Verlag (2007).

[4] V. Berthée, Fréquences des facteurs des suites sturmiennes, *Theoret. Comput. Sci.*, **165** (1996), 295-309.

[5] M. Bucci, A. De Luca, A. Glen, L.Q. Zamboni, A connection between palindromic and factor complexity using retun words, *Adv. Appl. Math.*, **42**, No 1 (2009), 60-74.

[6] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc. Simon Stevin*, **4** (1997), 67-88.

[7] J. Cassaigne, I. Kaboré, T. Tapsoba, On a new notion of complexity on infinite words, *Acta Univ. Sapentiae, Mathematica*, **2** (2010), 127-136.

[8] E.M. Coven, G.A. Hedlund, Sequences with minimal block growth, *Math. Syst. Theory*, **7** (1973), 138-153.

[9] X. Droubay, G. Pirillo, Palindromes and Sturmian words, *Theoret. Comput. Sci.*, **223**, No 1-2 (1999), 73-85.

[10] S. Dulucq, D. Gouyou-Beauchamps, Sur les facteurs des suites de Sturm, *Theoret. Comput. Sci.*, **71** (1990), 381-400.

[11] A. Glen, J. Justin, Episturmian words: A survery, *RAIRO-Theo. Inf. Appl.*, **43** (2009), 402-433.

[12] I. Kaboré, About k to k insertion words of Sturmian words, *Intern. J. Pure Appl. Math.*, **79** No 4 (2012), 561-572.

[13] I. Kaboré, T. Tapsoba, Combinatoire des mots récurrents de complexité $n + 2$, *RAIRO-Theo. Inf. Appl.*, **41** (2007), 425-446.

[14] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press (2002).

[15] F. Mignosi, P. Sébold, Morphismes sturmiens et règles de Rauzy, *J. Théor. Nombres Bordeaux*, **5** (1993), 221-233.