International Journal of Applied Mathematics

**Volume 32 No. 2 2019, 189-204**
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)
**doi:** http://dx.doi.org/10.12732/ijam.v32i2.3

# ABOUT THE WORDS BY $k$ TO $k$ ERASURE OF LETTER AND THE WORDS OF ERASED LETTERS: STURMIAN CASE

Moussa Barro[1], Idrissa Kaboré[2 §], Théodore Tapsoba[3]
[1,2,3]Department of Mathematics and Computer Science
Université Nazi oni
Bobo-Dioulasso 01 BP 1091, BURKINA FASO

**Abstract:** In this paper, we define from an infinite word $u$, the word by $k$ to $k$ erasure of a letter and the word of the erased letters. Then, we study the classical complexity and the palindromic complexity of these words in the case of modulo-recurrent words and more specifically in the Sturmian case.

**AMS Subject Classification:** 68R15, 11B85
**Key Words:** sturmian word, modulo-recurrent, $k$ to $k$ erasure, complexity, palindrome

## 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 [6], [7]. During the last thirty years, Sturmian words are intensively studied. These investigations led to the getting of numerous characterizations and various properties [4], [5], [10], [11] on these words. Over the last two decades, palindromes are used abundantly in the literature of combinatorial study of infinite words (see [2], [8], [9]).

The notion of $k$ to $k$ insertion of a letter in infinite words was introduced in [10], and widely studied in [3], [9].

We introduce the concept of $k$ to $k$ erasure of letter in infinite words. It

Received: September 16, 2018

§Correspondence author

consists to erase a letter steadily with step of length $k$ in an infinite word $u$. Thus, the new word obtained is called *$k$ to $k$ erasure word* of letter in $u$. The erased letters form a word which is called *$k$ to $k$ erased word* of $u$. The paper is devoted to the study of some combinatorial properties of these two types of words.

The paper is organized as follow. In Section 2 we give useful definitions and notations, and recall some properties of Sturmian words and modulo-recurrent words. We determine the classical complexity of the $k$ to $k$ erasure words of letter in non-trivial modulo-recurrent words and the corresponding erased words (Section 3). In Section 4, we study some palindromic proprieties and establish the palindromic complexities of the $k$ to $k$ erasure words and the associated erased words obtained from Sturmian words.

## 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 $\mathcal{A}$ is denoted $\mathcal{A}^*$ and $\varepsilon$, the empty word. For any $u \in \mathcal{A}^*$, the number of letters of $u$ is called length of $u$ and it is denoted $|u|$. Moreover, for any letter $x$ of $\mathcal{A}$, $|u|_x$ is the number of occurrences of $x$ in $u$. A word $u$ of length $n$ written with a unique letter $x$ is simply denoted $u = x^n$.

Let $u = u_1u_2\cdots u_n$ be a word such that $u_i \in \mathcal{A}$, for all $i \in \{1, 2, \cdots, n\}$. The image of $u$ by the reversal map is the word denoted $\overline{u}$ and defined by $\overline{u} = u_n \cdots\ u_2u_1$. The word $\overline{u}$ is simply called reversal image of $u$. A finite word $u$ is called palindrome if $\overline{u} = u$. If $u$ and $v$ are two finite words over $\mathcal{A}$, we have $\overline{uv}=\overline{v}\,\overline{u}$.

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 apprear in a word $u$, is designated by $alph(u)$. An infinite word $u$ is said to be eventually periodic if there exist two words $v \in \mathcal{A}^*$ and $w \in \mathcal{A}^+$ such that $u = vw^\omega$. If $v = \varepsilon$, then $u$ is periodic The *$n$-th* power of a finite word $w$ denoted by $w^n$ is the word corresponding to the concatenation $(ww\cdots w)$ $n$ times of $w$. By extension, $w^0 = \varepsilon$.

Let $u \in \mathcal{A}^\infty$ and $w \in \mathcal{A}^*$. The word $w$ is a factor of $u$ if there exist $u_1 \in \mathcal{A}^*$ and $u_2 \in \mathcal{A}^\infty$ such that $u = u_1wu_2$. The factor $w$ is said to be a prefix (resp. a suffix) if $u_1$ (resp. $u_2$) is the empty word.

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

For any infinite word $u$ over $\mathcal{A}$, we shall write $u = u_0u_1u_2\cdots$, where $u_i \in \mathcal{A}$, $i \geq 0$.

A factor $w$ of length $n$ of an infinite word $u = u_0u_1u_2\cdots$ appears in $u$ in the position $l$ if $w = u_lu_{l+1}\cdots u_{l+n-1}$.

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

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

Let $u = u_0u_1u_2u_3\cdots$ be an infinite word. The window complexity of $u$ is the map, $\mathbf{P}_u^f$ of $\mathbb{N}$ into $\mathbb{N}^*$, defined by

$$\mathbf{P}_u^f(n) = \#\left\{u_{kn}u_{kn+1}\cdots u_{n(k+1)-1} : k \geq 0\right\}.$$

The shift, is the application $S$ of $\mathcal{A}^\omega$ to $\mathcal{A}^\omega$ which erases the first letter of the word.

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

### 2.2. Sturmian words and modulo-recurrent words

In this subsection, we recall some properties of Sturmian words and modulo-recurrent words that will be used in the following.

**Definition 1.** An infinite word $u$ is a Sturmian word if for any natural $n$, $\mathbf{P}_u(n) = n + 1$.

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

**Definition 2.** An infinite word $u = u_0u_1u_2\cdots$ is said to be modulo-recurrent if for all $i \geq 1$, any factor $w$ of $u$, appears in $u$ at all positions modulo $i$, i.e.

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

**Example 3.** The word $a^\omega$ is modulo-recurrent.

The Thue-Morse word $\mathbf{t}$ is not modulo-recurrent. Indeed, the factorization in factors with length 2 of $\mathbf{t}$

$$\mathbf{t} = ab|ba|ba|ab|ba|ab|ab|ba|ba|ab\cdots$$

shows that the factor $bb$ never appears at an even position.

**Proposition 4.** ([5]) *Let $u$ be an infinite word over $\mathcal{A}$ such that $\boldsymbol{P}_u(n) = (\#\mathcal{A})^n$, for all $n \in \mathbb{N}$. Then, $u$ is modulo-recurrent.*

**Definition 5.** A modulo-recurrent word $u$ over $\mathcal{A}$ is called non-trivial modulo-recurrent word if there exists a positive integer $n_0$ such that for all $n \geq n_0$:

$$\mathbf{P}_u(n) < (\#\mathcal{A})^n.$$

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

**Proposition 6.** ([5]) *Let $u$ be a modulo-recurrent word. Then, for all $n$, the set of $n$-window factors of $u$ is equal to $L_n(u)$.*

**Theorem 7.** ([10]) *Every Sturmian word is modulo-recurrent.*

The following theorem presents a classical characterization of Sturmian words.

**Theorem 8.** ([8]) *Let $u$ be a Sturmian word. Then, for all $n \in \mathbb{N}$, we have:*

$$\boldsymbol{P}_u^{al}(n) = \begin{cases} 1 & \text{if } n \text{ is even} \\ 2 & \text{otherwise} \end{cases}.$$

The modulo-recurrent words can be characterized by their window complexity as follows.

**Theorem 9.** [5] *Let $u$ be a recurrent infinite word. Then, the following assertions are equivalent:*

(1) *$u$ is a modulo-recurrent word,*

(2) $\forall n \geq 1$, $\boldsymbol{P}_u^f(n) = \boldsymbol{P}_u(n)$.

## 3. Classical complexity

We introduce the notion of $k$ to $k$ erasure of letter in the infinite words with $k \geq 1$.

**Definition 10.** Let $u$ be an infinite word over $\mathcal{A}$. Let us decompose $u$ in the form:

$$u = x_0 m_0 x_1 m_1 x_2 m_2 x_3 m_3 \cdots x_i m_i \cdots , \tag{1}$$

where $m_i \in L_k(u)$ and $x_i \in \mathcal{A}$, $i \in \mathbb{N}$.

Now, let us erase the letters $x_i$ in the previous decomposition of $u$. We get the word

$$\mathcal{E}_k(u) = m_0 m_1 m_2 m_3 \cdots m_i \cdots .$$

The word $\mathcal{E}_k(u)$ is called *word by $k$ to $k$ erasure word* of letter in $u$.

The word defined by the sequence of the erased letters $x_i$ in the decomposition of $u$ is:

$$\mathcal{R}_k(u) = x_0 x_1 x_2 x_3 \cdots x_i \cdots .$$

**Example 11.** Consider the Fibonacci word

$$\mathbf{f} = abaababaabaababaababaabababaabaababaa \cdots .$$

Then,

$$\mathcal{E}_3(\mathbf{f}) = baaababaaabababababaaababaaa \cdots$$

and the associated erased word is

$$\mathcal{R}_3(\mathbf{f}) = ababaabab \cdots .$$

**Proposition 12.** *Let $u$ be an infinite word over $\mathcal{A}$ and $v = \mathcal{E}_k(u)$. Then, we have:*

*(i)* $\forall\, n \leq k$, $\boldsymbol{P}_v(n) \leq \boldsymbol{P}_u(n) + (n+1)\boldsymbol{P}_u(n+1)$,

*(ii)* $\forall\, q \geq 1$, $\boldsymbol{P}_v(kq+r) \leq (k-r+1)\boldsymbol{P}_u((k+1)q+r) + r\boldsymbol{P}_u((k+1)q+r+1)$ *with* $0 \leq r < k$.

*Proof.* First, consider $n \leq k$. Then, the factors of $u$ of length $n$ are also factors of $v$. Moreover, the other factors of $v$ of length $n$ are produced from factors of $u$ of length $n+1$ by erasing one letter. Thus, we have the following inequality:

$$\mathbf{P}_v(n) \leq \mathbf{P}_u(n) + (n+1)\mathbf{P}_u(n+1).$$

Now, consider $n > k$. Then, we can write $n = kq + r$ with $q \geq 1$ and $0 \leq r \leq k$. We have two cases to consider.

**Case 1**: The factors of $v$ of length $kq$.

Let us observe that any factor of $v$ of length $kq$ comes from a factor of $u$ of length $(k+1)q-1$, $(k+1)q$ or $(k+1)q+1$. Let $v_1$ be a factor of $v$ of length $kq$ coming from a factor $u_1$ of $u$.

• If $|u_1| = (k+1)q+1$, then $u_1 = x_1m_1x_2m_2\cdots x_qm_qx_{q+1}$ and $v_1 = m_1m_2\cdots m_q$ with $|m_i| = k$, for all $i = 1, \ldots, q$.

• If $|u_1| = (k+1)q-1$, then $u_1 = m_1x_2m_2\cdots x_qm_q$ and $v_1 = m_1m_2\cdots m_q$ with $|m_i| = k$, for all $i = 1, \ldots, q$.

• If $|u_1| = (k+1)q$, then $u_1 = m_0x_1m_1x_2m_2\cdots x_qm_q$ and $v_1 = m_0m_1m_2\cdots m_q$ with $|m_0| + |m_q| = k$ and $|m_i| = k$, for all $i = 1, \ldots, q-1$.

Consequently, the factors of $u$ of length $(k+1)q-1$ and $(k+1)q+1$ produce the same factors of $v$ of length $kq$. These same factors of $v$ come also from factors of $u$ of length $(k+1)q$ for $|m_0| = 0$.

Furtheremore, any factor of $u$ of length $(k+1)q$ produces at most $k+1$ factors of $v$ of length $kq$ by $k$ to $k$ erasure of letter. In conclusion, we obtain the inequality

$$\mathbf{P}_v(kq) \leq (k+1)\mathbf{P}_u((k+1)q).$$

**Case 2**: The factors of $v$ of length $kq+r$ with $1 \leq r < k$.

Observe that any factor of $v$ of length $kq+r$ comes from a factor of $u$ of length $(k+1)q+r$ or $(k+1)q+r+1$.

Let $v_1$ be a factor of $v$ of $kq+r$ coming from a factor $u_1$ of $u$.

• If $|u_1| = (k+1)q+r$, then $u_1 = m_0x_1m_1x_2m_2\cdots x_qm_q$ with $|m_0|+|m_q| = k+r$ and $|m_0| \leq k$, $|m_q| \leq k$ and $v_1 = m_0m_1m_2\cdots m_q$.
Therefore, $r \leq |m_0| \leq k$ and any factor of $u$ of length $(k+1)q+r$ produces at most $k-r+1$ distinct factors of $v$ of length $kq+r$.

• If $|u_1| = (k+1)q+r+1$, then $u_1 = m_0x_1m_1x_2m_2\cdots x_qm_qx_{q+1}m_{q+1}$ with $|m_0| + |m_{q+1}| = r$ and $v_1 = m_0m_1m_2\cdots m_qm_{q+1}$.

Furthermore, $|m_0| \leq r$. So, any factor of $u$ of length $(k+1)q+r$ produces at most $r+1$ factors of $v$ of length $kq+r$. Moreover, for $|m_0| = 0$ and $|m_0| = r$, the factors of $v$ of length $kq+r$ which come from factors of $u$ of length $(k+1)q+r+1$ are also produced by those of length $(k+1)q+r$. Consequently, any factor of $u$ of length $(k+1)q+r+1$ produces at most $r$ distinct factors of $v$ of length $kq+r$, by $k$ to $k$ erasures of letter.

Finally, we obtain the following inequality:

$$\mathbf{P}_v(kq+r) \leq (k-r+1)\mathbf{P}_u((k+1)q+r) + r\mathbf{P}_u((r+1)q+r+1).$$

□

**Proposition 13.** *Let $u$ be an infinite word and $w = \mathcal{R}_k(u)$. Then, for all positive integer $n$, the classical complexity of $w$ verifies the following inequality:*

$$\boldsymbol{P}_w(n) \leq (k+1)\boldsymbol{P}_u((k+1)n).$$

*Proof.* Let $u_1$ be a sufficiently long factor of $u$. Let $w_1$ be a factor of $w$ such that $w_1$ comes from $u_1$. Then, we can write $u_1 = m_0x_1m_1x_2m_2\cdots x_qm_q$ with $|m_0| \leq k$, $|m_q| \leq k$ and $|m_i| = k$, for $i = 1, \dots, q-1$. Furtheremore,

$$|u_1| \in \{(k+1)q - k, (k+1)q - k + 1, \dots, (k+1)q + k\}.$$

**Case 1**: $|u_1| = (k+1)q - k + r$ with $0 \leq r \leq k$ and $|m_0| + |m_q| = r$. Then, any factor of $u$ of length $(k+1)q - k + r$ produces $r+1$ factors of $w$ of length $q$. Moreover, these factors of $w$ of length $q$ come from the facteur of $u$ of length $(k+1)q$.

**Case 2**: $|u_1| = (k+1)q + k - r$ with $0 \leq r \leq k$, and $|m_0| + |m_q| = 2k - r$. Thus, any factor of $u$ of length $(k+1)q + k - r$ produces $r+1$ factors of $w$ of length $q$. Moreover, these factors of $w$ of length $q$ come from those of $u$ of length $(k+1)q$.

Finally, any factor of $w$ of length $q$ is produced from a factor of $u$ of length $(k+1)q$. In addition, any factor of $u$ of length $(k+1)q$ produces $k+1$ factors of $w$ of length $q$. Thus, we have the following inequality:

$$\mathbf{P}_w(n) \leq (k+1)\mathbf{P}_u((k+1)n).$$

□

**Proposition 14.** *Let $u$ be a modulo-recurrent word over an alphabet $\mathcal{A}$, $k \geq 1$ an integer. Then, the $k$ to $k$ erasure of letter in a factor of $u$ from any possible position produces a factor of $\mathcal{E}_k(u)$.*

*Proof.* Let $u_1$ be a factor of $u$. We have two cases to discuss.

**Case 1**: $|u_1| \leq k+1$.
Consider $p$ and $s$ two factors of $u$ and $x$ a letter in $\mathcal{A}$ such that $u_1 = pxs$. Thus, since $u$ is modulo-recurrent, then $u_1$ appears in $u$ in some position $h \equiv k+1-|p| \mod (k+1)$. Therefore, $ps$ is a factor of $\mathcal{E}_k(u)$.

**Case 2**: $|u_1| > k+1$.
Consider the words $p$, $s$, $m_1$, $\cdots$, $m_q$ such that
$u_1 = px_1m_1x_2\cdots x_qm_qx_{q+1}s$ with $|p| + |s| \leq k$ and $|m_i| = k$ for $i = 1, \dots, q$. Then, as previously $u_1$ appears in $u$ in some position $h \equiv k+1-|p| \mod (k+1)$. Therefore, it follows that $pm_1\cdots m_qs$ is a factor of $\mathcal{E}_k(u)$. □

The following corollaries are immediate consequences of Proposition 14.

**Corollary 15.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$ and $u_1$ be a factor of $u$. Then, any erased factor of $u_1$ appears in $\mathcal{R}_k(u)$.*

**Corollary 16.** *Let $u$ be a modulo-recurrent word. Then, the words $\mathcal{E}_k(u)$ and $\mathcal{R}_k(u)$ are recurrent.*

**Remark 17.** Let $u$ be a non-trivial modulo-recurrent word. Then, $\mathcal{E}_k(u)$ is not modulo-recurrent.

Indeed, pose $v = \mathcal{E}_k(u)$. Since $u$ is non-trivial modulo-recurrent then, we have
$\mathbf{P}_v^f(k) \leq \mathbf{P}_u^f(k+1) = \mathbf{P}_u(k+1)$ and $\mathbf{P}_v(k) = (k+1)\mathbf{P}_u(k+1)$.
Therefore, $\mathbf{P}_v^f(k) \neq \mathbf{P}_v(k)$. Hence, $v$ is not modulo-recurrent.

**Lemma 18.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$. Then, $u$ is periodic if and only if $u = x^\omega$, $x \in \mathcal{A}$.*

*Proof.* Suppose $u$ a periodic word. So, there exists a factor $m$ of $u$ such that $u = m^\omega$. Thus, $\mathbf{P}_u^f(lq) = 1$ with $|m| = l$ and $q \geq 0$. Since $u$ is modulo-recurrent then, we have $\mathbf{P}_u(lq) = 1$. Consequently, $\mathbf{P}_u(n) = 1$ for all $n$. Hence, $m = x$ and $u = x^\omega$.

The converse is evident. □

**Lemma 19.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$. Then, the following properties are equivalent:*

*(i) $u$ is not eventually periodic;*

*(ii) $\mathcal{E}_k(u)$ is not eventually periodic;*

*(iii) $\mathcal{R}_k(u)$ is not eventually periodic.*

*Proof.* $(i) \Longrightarrow (ii)$. Suppose that $u$ is not eventually periodic. Then, assume that $\mathcal{E}_k(u)$ is eventually periodic. Therefore, $v$ is periodic since it is recurrent. So, there exists a factor $m$ of $\mathcal{E}_k(u)$ such that $\mathcal{E}_k(u) = m^\omega$.

For $k = |m|$, we obtain $u = (x_i m)^\omega$ with $x_i \in \mathcal{A}$. Thus, for $q \geq 1$, we have $\mathbf{P}_u^f((k+1)q) \leq \#\mathcal{A}$ since $u$ is modulo-recurrent. Furthermore, we have

$\mathbf{P}_u((k+1)q) \leq \#\mathcal{A}$. Therefore, $\mathbf{P}_u$ is bounded and $u$ is eventually periodic. This contradicts the fact that $u$ is not eventually periodic. Hence, $\mathcal{E}_k(u)$ is not eventually periodic.

*(ii)* $\Longrightarrow$ *(i)*. Suppose that $\mathcal{E}_k(u)$ is not eventually periodic. By contradiction, assume that $u$ is not eventually periodic. Therefore, $u$ is periodic since it is recurrent. So, there exists a word $t$ such that $u = t^\omega$. Thus, $u$ can be written in the form $u = (xm)^\omega$ with $t = xm$ and $x \in \mathcal{A}$.

For $k = |m|$, we obtain $\mathcal{E}_k(u) = m^\omega$. We have a contradiction. Hence, $u$ is not eventually periodic.

*(iii)* $\Longrightarrow$ *(i)*. Suppose that $\mathcal{R}_k(u)$ is not eventually periodic. By contradiction, assume that $u$ is not eventually periodic. Therefore, $u$ is periodic since it is recurrent. So, there exists a word $t$ such that $u = t^\omega$. Thus, $u$ can be written in the form $u = (xm)^\omega$ with $t = xm$ and $x \in \mathcal{A}$.

For $k = |m|$, we have $\mathcal{R}_k(u) = x^\omega$. This contradicts the fact that $\mathcal{R}_k(u)$ is not evantually periodic. Contradiction. Hence, $u$ is not eventually periodic. □

**Corollary 20.** *Let $u$ be an infinite word. If $u$ is $\tau$-periodic and $k \equiv -1 \mod \tau$, then $\mathcal{E}_k(u)$ is $k$-periodic.*

*Proof.* Consider $u$ a periodic infinite word. Then, we can write $u = t^\omega = (xm)^\omega = xmxmxmxm\cdots$ with $t = xm$, $x \in \mathcal{A}$ and $|t| = \tau$.

If $k = q\tau - 1$, then $\mathcal{E}_k(u) = (m(xm)^{q-1})^\omega = (mt^{q-1})^\omega$. Thus, $v$ is periodic and $|mt^{q-1}| = q\tau - 1 = k$. □

**Lemma 21.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$. Let $u_1$ be a factor of $u$ such that $|u_1| \leq k+1$. If the $k$ to $k$ erasures of letter in $u_1$ from two different positions produce a same factor then,*

*(i) the erased letters are identical;*

*(ii) the factor which separates these two positions in $u_1$ is a power of the erased letter.*

*Proof.* Let $u_1$ be a factor of $u$ such that $|u_1| \leq k+1$. Let $v_1$ and $v_2$ be two words obtained by $k$ to $k$ erasures of letter from two different positions in $u_1$ such that $v_1 = v_2$. Then, $u_1$ can be written in the form $u_1 = px_1tx_2s$ with $|p|, |t|, |s| \geq 0$ and $v_1 = ptx_2s$, $v_2 = px_1ts$. Since $v_1 = v_2$, then we get $ptx_2s = px_1ts$. It follows that $x_1 = x_2$ *(i)*.

In addition, $ptx_2s = px_1ts$ implies that $tx_2 = x_1t$. Thus, by *(i)* we have $tx = xt$ with $x = x_1 = x_2$. Therefore, $t$ is a power of $x$ because $x$ is a letter *(ii)*. □

For the sequel, we need the following definitions.

**Definition 22.** Let $u$ be a modulo-recurrent word over $\mathcal{A}$ and $k \geq 1$.

(1) A factor $u_1$ of $u$ is said to be sufficiently long if $u_1$ contains all the $(k+1)$-window factors of $u$.

(2) A factor $v_1$ of $\mathcal{E}_k(u)$ is said to be sufficiently long if $v_1$ comes from a sufficiently long factor of $u$.

**Lemma 23.** *Let $u$ be a non-trivial modulo-recurrent word over $\mathcal{A}$. Let $u_1$ be a sufficiently long factor of $u$. Then, the $k$ to $k$ erasures of letter in $u_1$ from two different positions produce distinct factors of $\mathcal{E}_k(u)$.*

*Proof.* Let $u_1$ be a sufficiently long factor of $u$. The word $u$ is modulo-recurrent and non-trivial since it is Sturmian. So, any $k$ to $k$ erasure of letter in $u_1$ produces a factor of $\mathcal{E}_k(u)$. Consider two factors $v_1$ and $v_2$ of $\mathcal{E}_k(u)$ obtained by $k$ to $k$ erasures of letter from two different positions in $u_1$ such that $v_1 = v_2$. Then, $v_1$ and $v_2$ contain the same number of erased letters. So, $u_1$ can be written in two forms $u_1 = m_0x_1m_1x_2m_2 \cdots x_qm_q$ and $u_1 = t_0y_1t_1y_2t_2 \cdots y_qt_q$ with $x_i, y_i \in \mathcal{A}$. Therefore, we obtain $v_1 = m_0m_1m_2 \cdots m_q$ and $v_2 = t_0t_1t_2 \cdots t_q$.

Since $v_1 = v_2$, we have three cases to discuss.

**Case 1**: $|m_0| = |t_0|$. So $m_0 = t_0$ since $v_1 = v_2$. Then, $m_i = t_i$ et $x_i = y_i$, for all $i = 0, 1, \ldots, q$. Thus, the two erasures are identical.

**Case 2**: $|m_0| < |t_0|$. We can write $t_0 = m_0x_1t'_1$ since $v_1 = v_2$. Thus, $u_1$ can be written in the form

$$u_1 = m_0x_1t'_1y_1t''_1x_2t'_2y_2t''_2 \cdots x_qt'_qy_qt_q$$

with $t_i = t''_ix_{i+1}t'_{i+1}$, $i = 1, \ldots, q-1$. According to Lemma 21, we have $x_i = y_i$ and $t'_i = x_i^l$ with $l \geq 1$, for $i = 1, \ldots, q-1$. Therefore, we obtain $u_1 = m_0x_1^{l+2}t''_1x_2^{l+2}t''_2 \cdots x_q^{l+2}t_q$. Now, we know that $|x_i^{l+2}t''_i| = k+1$, for all $i = 1, \ldots, q-1$. So, all the $(k+1)$-window factors of $S^{|m_0|}(u_1)$ begin with $x_i^2$. Thus, $\mathbf{P}_u^f(k+1) \leq \mathbf{P}_u(k+1) - (\#\mathcal{A})$ since $L_{k+1}(u)$ contains all the factors of $u$ of length $k+1$ begining with $x_ix_j$ where $x_i, x_j \in \mathcal{A}$. This contradicts the fact that $\mathbf{P}_u^f(k+1) = \mathbf{P}_u(k+1)$ by Theorem 9 since $u$ is modulo-recurrent. Thus, $|m_0| = |t_0|$ and we come back to Case 1.

**Case 3**: $|m_0| > |t_0|$. We carry on the reasoning similarly in Case 2. ☐

**Lemma 24.** *Let $u$ be a non-trivial modulo-recurrent word over $\mathcal{A}$. Let $u_1$ and $u_2$ be two distinct and sufficiently long factors of $u$. Then, the words* *obtained by $k$ to $k$ erasures of letter in the words $u_1$ and $u_2$ are distinct factors of $\mathcal{E}_k(u)$.*

*Proof.* Let $u_1$ and $u_2$ be two distinct and sufficiently long factors of $u$. We decompose $u_1$ and $u_2$ in the following forms $u_1 = m_0x_1m_1x_2m_2 \cdots x_qm_q$ and $u_2 = t_0y_1t_1y_2t_2 \cdots y_qt_q$ with $x_i, y_i \in \mathcal{A}$, and $m_i, t_i \in L_k(u)$, $1 \leq i \leq q-1$. Thus, we obtain $v_1 = m_0m_1m_2 \cdots m_q$ and $v_2 = t_0t_1t_2 \cdots t_q$.

Suppose $v_1 = v_2$.

**Case 1**: $|m_0| = |t_0|$. Since $v_1 = v_2$, then $m_0 = t_0$. So, we have $m_i = t_i$ and $x_i = y_i$, for all $i = 0, 1, \cdots, q$. Thus, it results that $u_1 = u_2$. This is impossible because $u_1$ and $u_2$ are distinct factors. Therefore, $v_1$ and $v_2$ are distinct factors of $\mathcal{E}_k(u)$.

For the cases $|m_0| < |t_0|$ and $|m_0| > |t_0|$, we carry on the reasoning similarly to Lemma 23. In conclusion, $v_1$ and $v_2$ are distinct factors of $\mathcal{E}_k(u)$. ☐

**Corollary 25.** *Let $u$ be a non-trivial modulo-recurrent word over $\mathcal{A}$. Let $u_1$ and $u_2$ be two distinct and sufficiently long factors of $u$. Let $v_1$ and $v_2$ (resp. $w_1$ and $w_2$) be two words obtained by $k$ to $k$ erasures of letters (resp. two erased words) in the words $u_1$ and $u_2$ respectively. If $v_1 = v_2$, then we have $w_1 = w_2$.*

*Proof.* It suffices to apply successively Lemma 24 and Lemma 23. ☐

We show that the inequality of the Proposition 12 becomes an equality for $n$ sufficiently large and $u$ is a non-trivial modulo-recurrent word.

**Theorem 26.** *Let $u$ be a non-trivial modulo-recurrent word over $\mathcal{A}$, $v = \mathcal{E}_k(u)$ and $w = \mathcal{R}_k(u)$. Then, for $q$ sufficiently large, we have:*

(*i*) $\boldsymbol{P}_v(kq + r) = (k - r + 1)\boldsymbol{P}_u((k + 1)q + r) + r\boldsymbol{P}_u((k + 1)q + r + 1)$,

(*ii*) $\boldsymbol{P}_w(q) = (k+1)\boldsymbol{P}_u((k+1)q)$.

*Proof.* (*i*) According to Lemma 23, the $k$ to $k$ erasures from two different positions in a sufficiently long factor of $u$ produce distinct factors of $v$. Moreover, by Lemma 24, the $k$ to $k$ erasures of letter in two distinct and sufficiently long factors of $u$ produce distinct factors of $v$. Therefore, by Proposition 12, any factor of $u$ of length $(k + 1)q + r$ (resp. $(k + 1)q + r + 1$) produce after erasure $(k - r + 1)$ (resp. $r$) distinct factors of $v$ since $u$ is modulo-recurrent and non-trivial. Hence, the inequality in Proposition 12 becomes an equality.

(*ii*) By Corollary 25, any sufficiently long factor of $w$ comes from an only one factor of $u$. So, any factor of $u$ of length $(k+1)q$ produces $(k+1)$ factors of length $q$ of $u$. Thus, the inequality of the Proposition 13 becomes an equality. □

Observe that the Theorem 26 is not true for the trivial modulo-recurrent words.

Indeed, if $\mathbf{P}_u(n) = 1$, for all $n \in \mathbb{N}$, then the word $u$ is periodic. From Lemma 18, it follows that $u = x^\omega$ and $v = u$. Consequently, for all $n \in \mathbb{N}$, $\mathbf{P}_v(n) = 1$.

If $\mathbf{P}_u(n) = (\#\mathcal{A})^n$, for all $n \in \mathbb{N}$, then the classical complexity of $u$ is maximal. Since $u$ is modulo-recurrent, so $\mathbf{P}_v(n) \geq \mathbf{P}_u(n)$, for all $n \in \mathbb{N}$. Thus, we have $\mathbf{P}_v(n) = \mathbf{P}_u(n)$, for all $n \in \mathbb{N}$.

Taking $r = 0$, in Theorem 26, we have the following:

**Remark 27.** Let $u$ be a non-trivial modulo-recurrent word over $\mathcal{A}$, $v = \mathcal{E}_k(u)$ and $w = \mathcal{R}_k(u)$. Then, for $q$ sufficiently large we have

$$\mathbf{P}_w(q) = \mathbf{P}_v(kq).$$

**Corollary 28.** *Let $u$ be a Sturmian words over $\mathcal{A}$, $v = \mathcal{E}_k(u)$ and $w = \mathcal{R}_k(u)$. Then, for all sufficiently large $n$, we have:*

(*i*) $$\boldsymbol{P}_v(n) = (k+2)n+k+\lfloor\frac{n}{k}\rfloor+1,$$

*where $\lfloor . \rfloor$ denotes the floor function.*

(*ii*) $$\boldsymbol{P}_w(n) = (k+1)^2n+k+1.$$

*Proof.* Since the word $u$ is Sturmian, then it is a non-trivial modulo-recurrent word. By applying Theorem 26 for $n = kq + r$, we have:

(*i*)
$$\begin{aligned}\mathbf{P}_v(kq+r) &= (k-r+1)\mathbf{P}_u((k+1)q+r) + r\mathbf{P}_u((k+1)q+r+1)\\ &= (k-r+1)((k+1)q+r+1) + r((k+1)q+r+2)\\ &= (k+2)(kq+r)+k+q+1\\ &= (k+2)n+k+\lfloor\frac{n}{k}\rfloor+1 \text{ since } q = \frac{n-r}{k} = \lfloor\frac{n}{k}\rfloor .\end{aligned}$$

(*ii*)
$$\begin{aligned}\mathbf{P}_w(n) &= (k+1)\mathbf{P}_u((k+1)n)\\ &= (k+1)((k+1)n+1)\\ &= (k+1)^2n+k+1.\end{aligned}$$
□

## 4. Palindromic complexity

In this section, we determine the palindromic complexities of $\mathcal{E}_k(u)$ and $\mathcal{R}_k(u)$.

**Lemma 29.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$. Then, the following assertions are equivalent:*
*(i) $L(u)$ is stable by reversal map;*
*(ii) $L(\mathcal{E}_k(u))$ is stable by reversal map;*
*(iii) $L(\mathcal{R}_k(u))$ is stable by reversal map.*

*Proof.* Let $u_1$ be a factor of $u$ such that $u_1 = m_0x_1m_1x_2m_2 \cdots x_qm_q$. Then, we have $v_1 = m_0m_1m_2 \cdots m_q$ and $w_1 = x_1x_2 \cdots x_q$.

$(i) \Longrightarrow (ii)$. Suppose that $L(u)$ is stable by reversal map. Then we have $\overline{u}_1 = \overline{m}_qx_q \cdots \overline{m}_1x_1\overline{m}_0 \in L(u)$. So, by Proposition 14, we get $\overline{v}_1 = \overline{m}_q \cdots \overline{m}_1\overline{m}_0 \in L(\mathcal{E}_k(u))$ since $u$ is modulo-recurrent.

$(ii) \Longrightarrow (i)$. Suppose that $L(\mathcal{E}_k(u))$ is stable by reversal map. Then, we have $\overline{v}_1 \in L(\mathcal{E}_k(u))$. According to Proposition 14, $\overline{v}_1$ comes from $\overline{u}_1$ since $u$ is modulo-recurrent. Thus, we have $\overline{u}_1 = \overline{m}_qx_q \cdots \overline{m}_1x_1\overline{m}_0 \in L(u)$.

$(iii) \Longrightarrow (i)$. Suppose that $L(\mathcal{R}_k(u))$ is stable by reversal map. Then, we have $\overline{w}_1 \in L(\mathcal{R}_k(u))$. In addition, $\overline{w}_1$ comes from $\overline{u}_1$ because $u$ is modulo-recurrent. Hence, we have:

$$\overline{u}_1 = \overline{m}_qx_q \cdots \overline{m}_1x_1\overline{m}_0 \in L(u).$$

□

**Theorem 30.** *Let $u$ be a modulo-recurrent word over $\mathcal{A}$. Let $v_1$ be a palindrome of $\mathcal{E}_k(u)$ coming from a sufficiently long factor $u_1$ of $u$ and $w_1$ the associated erased factor of $u_1$. Then, $u_1$ and $w_1$ are palindromes.*

*Proof.* Let $u_1$, $v_1$ and $w_1$ be words satisfying the hypothesis of the Theorem. Then, by Lemma 29, we have $\overline{u}_1 \in L(u)$. Since $v_1$ comes from $u_1$, by Proposition 14, $\overline{v}_1$ comes from $\overline{u}_1$. Furthermore, $v_1$ is palindrome, i.e, $\overline{v}_1 = v_1$. So, by Lemma 23, $v_1$ comes from $\overline{u}_1$ since $v_1$ is sufficiently long. According to Lemma 24, it results that $\overline{u}_1 = u_1$.

Moreover, we have $\overline{w}_1 \in L(\mathcal{R}_k(u))$ because $w_1 \in L(\mathcal{R}_k(u))$. Therefore, $u_1$ being sufficiently long, by Proposition 14, $\overline{w}_1$ comes from $\overline{u}_1$. In addition, $u_1$ being a palindrome, by Corollary 25, we deduce that $\overline{w}_1 = w_1$. □

**Theorem 31.** *Let $u$ be an infinite word over $\mathcal{A}$, admitting palindromes and $v = \mathcal{E}_k(u)$. We have:*

• *For $n \leq k$,*

$$\boldsymbol{P}_v^{al}(n) \leq \boldsymbol{P}_u^{al}(n) + \boldsymbol{P}_u^{al}(n+1).$$

• *For $n = kq + r$ with $0 \leq r < k$,*

$$\boldsymbol{P}_v^{al}(kq+r) \leq \begin{cases} \boldsymbol{P}_u^{al}((k+1)q+r) + \boldsymbol{P}_u^{al}((k+1)q+r+1) \; if\, k, r\, are\, even \\ \boldsymbol{P}_u^{al}((k+1)q+r) \quad if\ k, r\ are\ odd \\ \boldsymbol{P}_u^{al}((k+1)q+r+1) \quad if\ k\ is\ odd, r\ is\ even \\ 0 \quad if\ k\ is\ even, r\ is\ odd \end{cases}.$$

*Proof.* Let us observe that any palindrome of $v$ comes from a palindrome of $u$ by Theorem 30.

For $n \leq k$, according to Proposition 12, we have:

- $PAL_n(v) = \{P, Qx\overline{Q} : P \in PAL_n(u), Qxx\overline{Q} \in PAL_{n+1}(u)\}$ if $n$ is odd.
- $PAL_n(v) = \{P, Q\overline{Q} : P \in PAL_n(u), Qx\overline{Q} \in PAL_{n+1}(u)\}$ if $n$ is even.

For $n > k$, we can write $n = kq+r$ with $0 \leq r < k$. According to Proposition 12, we have:

$$\begin{aligned} PAL_n(v) = \{&m_0 m_1 m_2 \cdots m_q : m_0 x_1 m_1 x_2 \cdots x_q m_q \in PAL_{(k+1)q+r}(u), \\ &|m_0| = |m_q| = \frac{k+r}{2}\Big\} \\ \cup \{&m_0 m_1 \cdots m_q m_{q+1} : m_0 x_1 m_1 \cdots x_{q+1} m_{q+1} \in PAL_{(k+1)q+r+1}(u), \\ &|m_0| = |m_{q+1}| = \frac{r}{2}\Big\}. \end{aligned}$$

Thus, it follows that:

• If $k$ and $r$ have the same parity, any palindrome of $u$ of length $(k+1)q+r$ produces a palindrome of $v$ of length $kq+r$.

• If $r$ is even, any palindrome of $u$ of length $(k+1)q+r+1$ produces a palindrome of $v$ of length $kq+r$. □

**Remark 32.** For $n$ sufficiently large the inequality of the Theorem 31 becomes an equality, if $u$ is a non-trivial modulo-recurrent word.

**Corollary 33.** *Let $u$ be a Sturmian word over $\mathcal{A}$ and $v = \mathcal{E}_k(u)$. Then, for any $n$ sufficiently large, we have:*

$$\boldsymbol{P}_v^{al}(n) = \begin{cases} 3 & if\ k, n\ are\ even \\ 0 & if\ k\ is\ even, n\ is\ odd \\ 2 & if\ k\ is\ odd \end{cases}.$$

*Proof.* Consider $n$ sufficiently large and pose $n = kq + r$ with $0 \leq r < k$. The word $u$ being Sturmian, it is non-trivial modulo-recurrent word. From Theorem 26, any palindrome of $u$ of length $n$ produces a unique palindrome of $v$. Consequently, the inequality of the Theorem 31 becomes an equality.

• For $k$ even, then $r$ and $n$ have the same parity. So, according to Theorem 8, $v$ admits three palindromes if $n$ is even and neither otherwise.

• For $k$ odd and $n$ even, then $r$ and $q$ have the same parity. Thus, if $r$ is even (resp. odd), then $(k+1)q + r + 1 = n + q + 1$ (resp. $(k+1)q + r = n + q$) is odd. Hence, $v$ admits two palindromes of length $n$.

• For $k$ and $n$ odd, then $r$ and $q$ have the different parities. If $r$ is even (resp. odd), then $(k+1)q + r = n + q$ (resp. $(k+1)q + r + 1 = n + q + 1$) is odd. Thus, $v$ admits also two palindromes of length $n$. □

**Corollary 34.** *Let $u$ be a Sturmian word over $\mathcal{A}$, $v = \mathcal{E}_k(u)$ and $w = \mathcal{R}_k(u)$. Then, for $n$ sufficiently large, we have:*

$$\boldsymbol{P}_w^{al}(n) = \begin{cases} 1 & if\ k, n\ are\ even \\ 2 & if\ k\ is\ even, n\ is\ odd \\ 2 & if\ k\ is\ odd \end{cases}$$

*Proof.* Suppose $n$ sufficiently large. So, from Proposition 13, we have:

$$PAL_n(w) = \{x_1x_2\cdots x_n \ : \ m_0x_1m_1x_2\cdots x_nm_n \in PAL(u), \\ |m_0| = |m_n|\}.$$

Thus, any palindrome of $w$ of length $n$ comes from a palindrome of $u$ of length $(k+1)n - k$ i.e., $|m_0| = |m_n| = 0$.

• If $k$ is odd, then $(k+1)n - k$ is odd. So, according to Theorem 8, we deduce that $w$ admits two palindromes of length $n$.

• If $k$ is even, then $(k+1)n - k$ and $n$ have the same parity. So, $w$ admits a unique palindrome if $n$ is even and two palindromes otherwise. □

## 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] M. Barro, I. Kaboré, T. Tapsoba, On the words by $k$ to $k$ insertion of a letter in Sturmian words, *International Journal of Applied Mathematics* **30**, No 5 (2017), 387-400; doi: 10.12732/ijam.v30i5.3.

[4] J. Berstel, Sturmian and episturmian words (a survery of some recent resultats), In: *Proc. of CAI 2007*, *LNCS* **4728**, Springer-Verlag, 2007.

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

[6] E.M. Coven, Sequences with minimal block growth, *Math. Syst. Theory*, **8** (1975), 376-382.

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

[8] X. Droubay, G. Pirillo, Palindromes and Sturmian words, *Theoret. Comput. Sci.*, **223**, (1999), 73-85.

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

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

[11] M. Lothaire, *Algebraic Combinatorics on Words*, Cambridge University Press, 2002.