International Mathematical Forum, Vol. 7, 2012, no. 44, 2167 - 2177

Study of an Extension of Sturmian Words over a Binary Alphabet

Idrissa Kaboré

Institut des Sciences Exactes et Appliquées
Université polytechnique de Bobo-Dioulasso
01 BP 1091 Bobo-Dioulasso 01, Burkina Faso
ikaborei@yahoo.fr

Abstract

In this paper, we define $(k,\, l)$-Sturmian words. Then, we study their complexity. Finally, we establish a characterization of these words via the action of particular morphisms on Sturmian words.

Mathematics Subject Classification: 68R15

Keywords: Sturmian words, special factors, complexity, morphisms

1 Introduction

A Sturmian word is an infinite binary word which possesses for any integer n, exactly $n+1$ factors of length n. The Sturmian words have been intensively studied over the past three decades (see Berstel's surveys [1, 2]). The numerous investigations established various remarkable characterizations of these words [2, 15, 13]. Considerable works have also been done on their generalizations. We have, for instance, quasi-sturmian words [5, 14, 3, 10] and episturmian words (see [9, 8, 7, 1]).

Our paper deals also with a natural extension of Sturmian words over a binary alphabet: $(k,\, l)$-Sturmian words.

After preliminaries, we recall some basic results of Sturmian words (Section 3). In Section 4, we introduce first of all, $(k,\, l)$-Sturmians words. Then, we study their complexity. Lastly, we establish a characterization of these words with the aid of the action of a family of particular morphisms on Sturmian words.

2 Preliminaries

In all the sequel, except express mention, the alphabet $\mathcal{A}$ considered is the binary alphabet $\{a, b\}$. The set of finite words over $\mathcal{A}$ is denoted by $\mathcal{A}^*$ and ε is the empty word. For any $u \in \mathcal{A}^*$, $|u|$ denotes the length of u ($|\varepsilon| = 0$) and for each $x \in \mathcal{A}$, $|u|_x$ is the number of occurrences of the letter x in u.

A word u of length n written with a single letter x is simply denoted by $u = x^n$. The n-th power of a finite word w denoted by w^n is the word corresponding to the concatenation $(www \dots w)$ n times of w . By extension, $w^0 = \varepsilon$.

An infinite word is a sequence of letters of $\mathcal{A}$. The set of infinite words over $\mathcal{A}$ is denoted by $\mathcal{A}^\omega$.

A finite word v is a factor of u if there exist two words u_1 and u_2 over $\mathcal{A}$ such that $u = u_1vu_2$; we also say that u contains v. The factor v is said to be a prefix (resp. a suffix) if u_1 (resp. u_2) is the empty word.

An infinite word w is ultimately periodic if there exist two words u and v such that $w = uv^\omega$, where v^ω is the infinite concatenation of the word v. It is periodic if u is the empty word. If the infinite word w does not have any of the previous forms we say that it is aperiodic.

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

A non empty factor v of an infinite word u is said to be a right (resp. a left) special factor of u if va and vb (resp. av and bv) are factors of u. The set of right special factors of u of length n will be denoted by $\mathcal{RS}(n)$. We say that a is a right (resp. a left) extension of v in u if va (resp. av) is in u.

An infinite word u is recurrent if each of its factors appears infinitely many times.

The complexity of an infinite word u is the map of $\mathbb{N}$ to $\mathbb{N}^*$ defined by $\mathbf{p}_u(n) = \#\mathcal{L}_n(u)$, where $\#\mathcal{L}_n(u)$ is the cardinality of $\mathcal{L}_n(u)$. The complexity of a word u is increasing and the ultimately periodic words are the only ones whose complexity is bounded. For more details on the complexity, we refer the reader to [4]. In whatever follows, the complexity $\mathbf{p}_u$ of u will be simply denoted by $\mathbf{p}$.

On a binary alphabet, the function $\mathbf{s}$ computes the number of right special factors of a given length in u. It is the same for left special factors if the word u is recurrent. It is used to determine the complexity through the following formula:

$$\mathbf{p}(n) = \mathbf{p}(n_0) + \sum_{m=n_0}^{n-1} \mathbf{s}(m). \tag{1}$$

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

It is said that an infinite word u is generated by a morphism f if there exists a letter $x \in \mathcal{A}$ such that the words $x, f(x), f^2(x), \cdots, f^n(x), \cdots$ are longer and longer prefixes of u. We denote $u = f^\omega(x)$.

The shift function is the map δ from $\mathcal{A}^\omega$ to $\mathcal{A}^\omega$ which consists in erasing the first letter of the word, for instance $\delta(abaabaaaa\cdots) = baabaaaa\cdots$. By extension, $\delta^0(u) = u$.

3 Sturmian words

In this section, we recall basic results on Stumian words which will be useful in Section 4.

Definition 3.1. *An infinite word u over $\mathcal{A}$ is Sturmian if for all $n \geq 0$, $\mathbf{p}(n) = n + 1$.*

The most known Sturmian word is the famous Fibonacci word generated by the morphism Φ defined by $\Phi(a) = ab$ and $\Phi(b) = a$. Its first few terms are:

$$F = abaababaabaababaababaabaababaaba\cdots \ .$$

Every Sturmian word is recurrent. An infinite word u over $\mathcal{A}$ is Sturmian if and only if, for all $n \geq 0$, u admits a unique right (resp. left) special factor of length n.

Every Sturmian word contains the words ab, ba and one of the two words a^2 and b^2.

A Sturmian word will be said to be a-Sturmian (resp. b-Sturmian) when it contains a^2 (resp. b^2).

We say that a word u over $\mathcal{A}$ is balanced if, for all $n \geq 0$ and all $v, w \in \mathcal{L}_n(u)$, we have $\Big| |v|_x - |w|_x \Big| \leq 1$ for all $x \in \mathcal{A}$. In the opposite case, we say that u is unbalanced.

The following lemma is useful.

Lemma 3.2. *[6] An infinite word u over $\mathcal{A}$ is unbalanced if and only if, there exists a word t with minimal length such that u contains ata and btb.*

The following characterization is well known.

Theorem 3.3. *[6] An infinite word u is Sturmian if and only if, it is aperiodic and balanced.*

Theorem 3.4. *[12] Let u be a recurrent word over $\mathcal{A}$. Then, u is Sturmian if and only if, up to a permutation of letters, it takes the following form:*

$$a^{m_0}ba^{m+\epsilon_1}ba^{m+\epsilon_2}ba^{m+\epsilon_3}b\cdots \tag{2}$$

where $(\epsilon_i)_i$ is a Sturmian sequence in $\{0, 1\}$, m_0 and m are two integers satisfying $m_0 \leq m+1$.

This property will inspire us the introduction of (k, l)-Sturmian words which will be studied in the next section.

4 (k, l)-Sturmian words

4.1 Definition

In all the sequel, (k, l) will designate a pair of positive integers such that $kl \geq 2$.

Definition 4.1. *Let u be a recurrent word over $\mathcal{A}$. We say that u is a (k, l)-Sturmian word if u takes the following form, up to a permutation of letters:*

$$u = \delta^{j_1}\left(a^{m_0}b^l a^{m+\epsilon_1}b^l a^{m+\epsilon_2}b^l a^{m+\epsilon_3}b^l\cdots\right) \tag{3}$$

where $(\epsilon_i)_{i\geq 1}$ is a Sturmian sequence over the alphabet $\{0, k\}$ and j_1, m_0, m are integers satisfying $m \geq 1$, $m_0 \leq m+k$ and $j_1 \leq m+k+l$.

This definition provides an extension of Sturmian words which correspond henceforth to $(1, 1)$-Sturmian words.

The following result is shown in [11].

Theorem 4.2. *Let u be a recurrent word over $\mathcal{A}$. Then, u is a word with complexity ultimately $n+2$ if and only if, it takes one of the two following forms, up to a permutation of letters:*

$$a^{m_0}ba^{m+\epsilon_1}ba^{m+\epsilon_2}ba^{m+\epsilon_3}b\cdots \tag{4}$$

where $(\epsilon_i)_i$ is a Sturmian sequence in $\{0, 2\}$, m_0 and m are two integers such that $m \geq 1$ and $m_0 \leq m+2$.

$$\delta^{j_1}\left(a^{m_0}b^2a^{m+\epsilon_1}b^2a^{m+\epsilon_2}b^2a^{m+\epsilon_3}b^2\cdots\right) \tag{5}$$

where $(\epsilon_i)_i$ is a Sturmian sequence in $\{0, 1\}$ and j_1, m_0 and m are integers such that $m \geq 1$, $m_0 \leq m+1$ and $j_1 \leq m+1$.

From the above result, every recurrent binary word with complexity $n+2$ is either $(1, 2)$-Sturmian or $(2, 1)$-Sturmian.

Remark 4.3. *Every (k, l)-Sturmian word u is aperiodic.*

This remark stems from the fact that the underlying sequence (ϵ_i) is aperiodic since it is Sturmian.

4.2 Complexity

In this section, we show that the complexity of any (k, l)-Sturmian word is ultimately $n+k+l-1$.

Lemma 4.4. *Let u be a recurrent word over $\mathcal{A}$ having the form*

$$u = a^{m_0} b^l a^{m+\epsilon_1} b^l a^{n_0+\epsilon_2} b^l a^{n_0+\epsilon_3} b^l \cdots$$

where $(\epsilon_i)_i$ is a Sturmian sequence in $\{0, 1\}$, l, m_0 and m are integers such that $l \geq 2$, $m \geq 1$ and $m_0 \leq m+1$. Then, $\mathbf{p}(l) = 2l$.

Proof. Observe that for all $n \in \{1, \cdots, l-1\}$, u possesses exactly two right special factors of length n. More precisely, we have:

- If $l \leq m+1$ then, for all $n < l$, $\mathcal{RS}(n) = \{a^n, b^n\}$.
- If $l > m+1$ then, $\mathcal{RS}(n) = \begin{cases} \{a^n, b^n\} & \text{if } n < m+1 \\ \{b^n, b^{n-m}a^m\} & \text{if } m+1 \leq n < l \end{cases}$.

Thus, by applying the formula 1, we have: $\mathbf{p}(l) = \mathbf{p}(1) + 2(l-1) = 2l$. □

Lemma 4.5. *Let u be a recurrent word over $\mathcal{A}$ having the form*

$$u = a^{m_0} b^l a^{m+\epsilon_1} b^l a^{m+\epsilon_2} b^l a^{m+\epsilon_3} b^l \cdots$$

where $(\epsilon_i)_i$ is a Sturmian sequence in $\{0, k\}$ and, k, l, m_0 and m are integers such that $k \geq 2$, $l \geq 1$, $m \geq 1$ and $m_0 \leq m+1$. Then, $\mathbf{p}(n_0) = n_0 + k + l - 1$ for $n_0 = \max(m+k, l)$.

Proof. We have six cases to distinguish according to the values of l. In each of these cases, we shall present the table of values of $\mathcal{RS}(n)$ and $\mathbf{s}(n)$ according to n with $n \leq \max(m+k, l)$, in order to deduce $\mathbf{p}(n_0)$.

Case 1. $l = 1$

n	$[1, m+1[$	$[m+1, m+k[$
$\mathcal{RS}(n)$	$\{a^n\}$	$\{a^n, a^{n-1-m}ba^m\}$
$\mathbf{s}(n)$	1	2

We get

$$\begin{aligned} \mathbf{p}(m+k) &= \mathbf{p}(1) + (m+1-1) + 2(m+k-m-1) \\ &= (m+k) + k + 1 - 1. \end{aligned}$$

Case 2. $1 < l < m+1$

n	$[1, l[$	$[l, m+1[$	$[m+1, m+k[$
$\mathcal{RS}(n)$	$\{a^n, b^n\}$	$\{a^n\}$	$\{a^n, X_n a^m\}$
$\mathbf{s}(n)$	2	1	2

In this table, $X_n a^m$ represents the suffix of length n of $a^{m+k}Va^m$ where V is the shortest factor of u separating two consecutive occurrences of a^{m+k} in u.

$$\begin{aligned}\mathbf{p}(m+k) &= \mathbf{p}(1) + 2(l-1) + (m+1-l) + 2(m+k-m-1)\\ &= (m+k) + k + l - 1.\end{aligned}$$

Case 3. $l = m+1$

n	$[1, m+1[$	$[m+1, m+k[$
$\mathcal{RS}(n)$	$\{a^n, b^n\}$	$\{a^n, X_n a^m\}$
$\mathbf{s}(n)$	2	2

In this table $X_n a^m$ represents the suffix of length n of $a^{m+k}Va^m$ where V is the shortest factor of u separating two consecutive occurrences of a^{m+k} in u. We get

$$\begin{aligned}\mathbf{p}(m+k) &= \mathbf{p}(1) + 2(m+1-1) + 2(m+k-m-1)\\ &= (m+k) + k + (m+1) - 1.\end{aligned}$$

Case 4. $m+1 < l < m+k$

n	$[1, m+1[$	$[m+1, l[$	$[l, m+k[$
$\mathcal{RS}(n)$	$\{a^n, b^n\}$	$\{a^n, b^n, b^{n-m}a^m\}$	$\{a^n, X_n a^m\}$
$\mathbf{s}(n)$	2	3	2

In this table $X_n a^m$ represents the suffix of length n of $a^{m+k}Va^m$ where V is the shortest factor of u separating two consecutive occurrences of a^{m+k} in u. We get

$$\begin{aligned}\mathbf{p}(m+k) &= \mathbf{p}(1) + 2(m+1-1) + 3(l-m-1) + 2(m+k-l)\\ &= (m+k) + k + l - 1.\end{aligned}$$

Case 5. $l = m+k$

n	$[1, m+1[$	$[m+1, m+k[$
$\mathcal{RS}(n)$	$\{a^n, b^n\}$	$\{a^n, b^n, b^{n-m}a^m\}$
$\mathbf{s}(n)$	2	3

Hence

$$\begin{aligned}\mathbf{p}(m+k) &= \mathbf{p}(1) + 2(m+1-1) + 3(m+k-m-1)\\ &= (m+k) + k + (m+k) - 1.\end{aligned}$$

Case 6. $m + k < l$

n	$[1,\, m+1[$	$[m+1,\, m+k[$	$[m+k,\, l[$
$\mathcal{RS}(n)$	$\{a^n,\, b^n\}$	$\{a^n,\, b^n,\, b^{n-m}a^m\}$	$\{b^n,\, b^{n-m}a^m\}$
$\mathbf{s}(n)$	2	3	2

Hence

$$\begin{aligned}\mathbf{p}(l) &= \mathbf{p}(1) + 2(m+1-1) + 3(m+k-m-1) + 2(l-m-k)\\ &= l + k + l - 1.\end{aligned}$$

The proof of the lemma is complete.

□

Theorem 4.6. *Let u be a (k,l)-Sturmian word. Then, there exists n_0 such that*

$$\forall n \geq n_0,\ \ \mathbf{p}(n) = n + k + l - 1.$$

Proof. *Case 1.* $k = 1$. From Lemma 4.4, for $n = l$ we have $\mathbf{p}(l) = l+k+l-1$. We will show that

$$\forall n \geq l,\ \ \mathbf{p}(n) = n + k + l - 1.$$

It amounts to show that u admits a unique right special factor for any length $n \geq l$.

• For $n \in [l,\, l+m+1[$, observe that u admits a unique right special factor of length n. Indeed, if $l \leq m$ then, for $n \in [l,\, m+1[$ (resp. $n \in [m+1,\, l+m+1[$), the word a^n (resp. $b^{n-m}a^m$) is the unique right special factor of length n of u. Similarly, if $l \geq m+1$ then, for $n \in [l,\, l+m+1[$, the word $b^{n-m}a^m$ is the unique right special factor of length n of u.

• Suppose that there exists a integer $n \geq l+m+1$ such that u possesses two right special factors, D_n and D'_n, of length n. Consider, in this case, n minimal. Then, we can write D_n and D'_n respectively in the form aD and bD. Thus, D is a right special factor of u such that $|D| \geq l+m+1$. Therefore, D ends by $b^l a^m$, the unique right special factor of u of length $l+m$. Similarly, we verify that D begins with $a^m b^l$ in reasoning with left special factors. So, it will be possible to write $D = a^m T a^m$ where T is a factor of u beginning and ending by b^l. Thus, it follows that $b^l a^{m+1} T a^{m+1} b^l$ and $b^l a^m T a^m b^l$ will be in u since aDa and bDb are in u by assumption and a^{m+1} is always preceded or followed by b^l in u. Consequently, the underlying factors of $(\epsilon_i)_i$ in these two factors of u are respectively of the form $1t1$ and $0t0$. It is impossible because the sequence $(\epsilon_i)_i$ is Sturmian.

Case 2. $k \geq 2$. From Lemma 4.5, for $n = \max(m+k,\, l)$, we have $\mathbf{p}(n) = n+k+l-1$. Let us put $n_0 = \max(m+k,\, l)$. We shall show that

$$\forall n \geq n_0,\ \mathbf{p}(n) = n+k+l-1.$$

As in the previous case, it suffices to show that, for all $n \geq n_0$, u admits a unique right special factor of length n.

Suppose that there exists $n \geq n_0$ such that u possesses two right special factors D_n and D'_n of length n. Consider, in this case, n minimal. Then, D_n and D'_n can be written respectively in the form aD and bD. Hence, aDa and bDb will be in u. This requires that the factor D begins and ends by a^m. Let us check that D contains at least one occurrence of the letter b. If b were not in D then, since bDb is in u and $|D| \geq \max(m+k,\, l)$, D would be necessarily a^{m+k}. Consequently, a^{m+k+2} would be in u since aDa is in u. Therefore, D contains at least one occurrence of b and can be written in the form $D = a^m T a^m$, where T begins and ends by b.

We deduce, as in Case 1, that $b^l a^{m+1} T a^{m+1} b^l$ and $b^l a^m T a^m b^l$ will be in u. Then, we conclude in the same way.

□

In Theorem 4.6, the smallest n_0 verifies:

$$n_0 = \begin{cases} l & \text{if} \quad k = 1 \\ \max(l,\, m+k) & \text{if} \quad k \geq 2 \end{cases}.$$

4.3 Morphic interpretation

In this section, we shall provide a family of morphisms which allow us to state a characterization of $(k,\, l)$-Sturmian words via Sturmian words under their action.

Let k, l and r be positive integers such that $kl \geq 2$, and $0 \leq r < k$. Consider the following morphisms:

$$\varphi_{(k;\,r,\,l)} : \begin{array}{rcl} \{a,\, b\} & \longrightarrow & \{a,\, b\} \\ a & \longmapsto & a^k \\ b & \longmapsto & a^r b^l \end{array}, \quad \overline{\varphi_{(k;\,r,\,l)}} : \begin{array}{rcl} \{a,\, b\} & \longrightarrow & \{a,\, b\} \\ a & \longmapsto & a^k \\ b & \longmapsto & b^l a^r \end{array}$$

where $0 \leq r < k$. For any morphism φ of $\mathcal{A}^*$, let us put

$$< \varphi >= \{\varphi,\, E\varphi,\, \varphi E,\, E\varphi E\}$$

where E (called exchange morphism) is defined by $E(a) = b$ and $E(b) = a$. Consider the set

$$\mathcal{F}_{(k;\,r,\,l)} = \cup_{\varphi \in \left\{\varphi_{(k;\,r,\,l)},\, \overline{\varphi_{(k;\,r,\,l)}}\right\}} < \varphi > .$$

Lemma 4.7. *Let u be a Sturmian word and $\varphi \in \mathcal{F}_{(k;\,r,\,l)}$. Then, $\varphi(u)$ is a $(k,\,l)$-Sturmian word.*

Proof. Let u be a Sturmian word and $\varphi \in \mathcal{F}_{(k;\,r,\,l)}$. Then, we have

$$\varphi \in< \varphi_{(k;\,r,\,l)} > \cup < \overline{\varphi_{(k;\,r,\,l)}} > .$$

Since the reasoning is similar for all $\varphi \in< \varphi_{(k;\,r,\,l)} > \cup < \overline{\varphi_{(k;\,r,\,l)}} >$, we carry on the proof with $\varphi = \varphi_{(k;\,r,\,l)}$.

case 1. Suppose that u can be written in the form:

$$u = \delta^{j_1}\left(a^{m_0}b^l a^{m+\epsilon_1}b^l a^{m+\epsilon_2}b^l a^{m+\epsilon_3}b^l\cdots\right)$$

with $m_0 \le m+k$, $j_1 \le m+k+l$ and $(\epsilon_i)_{i\ge1}$, a Sturmian sequence over $\{0,\,1\}$. We can take without loss of generality $j_1 = 0$. We get

$$\varphi_{(k;\,r,\,l)}(u) = a^{km_0+r}b^l a^{(km+r)+k\epsilon_1}b^l a^{(km+r)+k\epsilon_2}b^l a^{(km+r)+k\epsilon_3}b^l\cdots$$

where the sequence $(k\epsilon_i)_{i\ge1}$ is Sturmian over $\{0,\,k\}$ since the sequence $(\epsilon_i)_{i\ge1}$ is Sturmian over $\{0,\,1\}$. Morever, $\varphi_{(k;\,r,\,l)}(u)$ is recurrent because u is recurrent. Thus, by definition, $\varphi_{(k;\,r,\,l)}(u)$ is a $(k,\,l)$-Sturmian word.

case 2. Suppose that u can be written in the form:

$$u = \delta^{j_1}\left(b^{m_0}a^l b^{m+\epsilon_1}a^l b^{m+\epsilon_2}a^l b^{m+\epsilon_3}a^l\cdots\right)$$

with $m_0 \le m+1$, $j_1 \le m+k+l$ and $(\epsilon_i)_{i\ge1}$, a Sturmian sequence over $\{0,\,1\}$. As in Case 1, we can take $j_1 = 0$. Hence

$$\varphi_{(k;\,r\,l)}(u) = \left(a^r b^l\right)^{m_0} a^k \left(a^r b^l\right)^{m+\epsilon_1} a^k \left(a^r b^l\right)^{m+\epsilon_2} a^k \left(a^r b^l\right)^{m+\epsilon_3} a^k \cdots .$$

Let us notice that $\varphi_{(k;\,r,\,l)}(u)$ is an infinite concatenation of $a^r b^l$ and $a^{r+k}b^l$: $\varphi_{(k;\,r,\,l)}(u) \in \left\{a^r b^l,\, a^{r+k}b^l\right\}^{\omega}$. Furthermore, $\varphi_{(k;\,r,\,l)}(u)$ is modulated after the prefix of length $(r+l)\,m_0$ by the factors in the form

$$a^k(a^r b^l)^{m+\epsilon} = a^{k+r}\quad \underbrace{(b^l a^r b^l)\cdots(b^l a^r b^l)}_{m-1+\epsilon \text{ occurrences of } b^l a^r b^l}$$

which possess, each, exactly $m-1+\epsilon$ occurrences of $b^l a^r b^l$. Thus, $\varphi_{(k;\,r,\,l)}(u)$ will be rewritten in the form

$$\varphi_{(k;\,r,\,l)}(u) = a^r b^l a^{r+\epsilon'_1}b^l a^{r+\epsilon'_2}b^l a^{r+\epsilon'_3}b^l\cdots$$

where $r \le k$ and the sequence $(\epsilon'_i)_i$ is defined over $\{0,\,k\}$ by

$$0^{m_0}k0^{m-1+\epsilon_1}k0^{m-1+\epsilon_2}k0^{m-1+\epsilon_3}k\cdots .$$

It follows that the sequence $(\epsilon'_i)_i$ is Sturmian over $\{0,\,k\}$ since on the one hand it is recurrent and on the other hand the sequence $(\epsilon_i)_i$ is Sturmian over $\{0,\,1\}$.

□

Remark 4.8. *Lemma 4.7 is also valid for all $r \geq k$.*

Theorem 4.9. *Let u be a recurrent word over $\mathcal{A}$. Then, u is a (k, l)-Sturmian word if and only if, there exist $\varphi \in \mathcal{F}_{(k;r,l)}$, a Sturmian word v and $j_0 \in \mathbb{N}$ such that $u = \delta^{j_0}(\varphi(v))$.*

Proof. $\Longrightarrow$) Let u be a recurrent word over $\mathcal{A}$. Assume that u is a (k, l)-Sturmian word. Up to a permutation of letters, u can be written in the following form

$$u = \delta^{j_1}(a^{m_0}b^l a^{m+\epsilon_1}b^l a^{m+\epsilon_2}b^l a^{m+\epsilon_3}b^l \cdots)$$

with $m_0 \leq m+k$, $j_1 \leq m+k+l$ and $(\epsilon_i)_{i\geq 1}$, a Sturmian sequence over $\{0, k\}$. Let us put

$$\hat{u} = \begin{cases} \delta^{j_2}(a^m b^l a^{m+\epsilon_1}b^l a^{m+\epsilon_2}b^l a^{m+\epsilon_3}b^l \cdots) & \text{if } m_0 \leq m \\ \delta^{j_2}(a^{m+k} b^l a^{m+\epsilon_1}b^l a^{m+\epsilon_2}b^l a^{m+\epsilon_3}b^l \cdots) & \text{otherwise} \end{cases}.$$

Then, if (q, r) is the pair of integers such that $m = qk + r$ with $0 \leq r < k$, the word $\hat{u}$ can be rewritten

$$\hat{u} = \delta^{j_2}\left(a^{k\left(q+\epsilon'_0\right)+r}b^l a^{k\left(q+\epsilon'_1\right)+r}b^l a^{k\left(q+\epsilon'_2\right)+r}b^l a^{k\left(q+\epsilon'_3\right)+r}b^l \cdots\right)$$

where $(\epsilon'_i)_{i\geq 0}$ verifies $(\epsilon_i)_{i\geq 0} = (k\epsilon'_i)_{i\geq 0}$ and

$$j_2 = \begin{cases} m - m_0 & \text{if } m_0 \leq m \\ m + k - m_0 & \text{otherwise} \end{cases}.$$

Thus, the sequence $(\epsilon'_i)_{i\geq 1}$ is Sturmian over $\{0, 1\}$ since the sequence $(\epsilon_i)_{i\geq 1}$ is Sturmian over $\{0, k\}$. Furthermore, $\hat{u}$ can be written $\hat{u} = \delta^{j_2}(\varphi(v))$ where

$$v = a^{(q+r)+\epsilon'_1}ba^{(q+r)+\epsilon'_1}ba^{(q+r)+\epsilon'_2}b^l a^{(q+r)+\epsilon'_3}b^l \cdots$$

with $(\epsilon'_i)_{i\geq 1}$, a Sturmian sequence over $\{0, 1\}$. Therefore, v is Sturmian. Now, $u = \delta^{j_1}(\hat{u})$ and $\hat{u} = \delta^{j_2}(\varphi(v))$. So, $u = \delta^{j_0}(\varphi(v))$ with $j_0 = j_1 + j_2$.

$\Longleftarrow$) The converse is due to Lemma 4.7. □

References

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

[2] Berstel, J., Recent results on Sturmian words, in *Developments in Language Theory II*, eds. J. Dassow and A. Salomaa, World Scientific, 1996, pp. 13–24.

[3] Cassaigne, J., Sequences with grouped factors, in *Developments in Language Theory III (DLT'97)* pp. 211–222, Aristotle University of Thessaloniki, 1998.

[4] Cassaigne, J., Nicolas, F., Factors complexity, *Combinatorics, Automata and Number Theory*, V. Berth, M. Rigo (Eds), Encyclopedia of Mathematics and its Applications 135, Cambridge University Press (2010).

[5] Coven, E. M., Sequences with minimal block growth, II, *Math. Systems Theory* 8 (1975), 376–382.

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

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

[8] Glen, A., Justin, J., Episturmian words: a survey, *RAIRO–Theoretical Informatics and Applications* 43 (2009), 402–433.

[9] Glen, A., Levé, F., Richomme, G., Directive words of episturmian words: equivalences and normalization, *RAIRO–Theoretical Informatics and Applications* 43 (2009), 299–319.

[10] Heinis, A., Arithmetics and combinatorics of words of low complexity, *PhD Thesis*, University of Leiden, 2001.

[11] Kaboré, I., Tapsoba, T., Mots binaires récurrents de complexité ultimement $n+2$, *Annales Mathématiques Africaines Serie 1*, 2 (2011) pp. 47–57.

[12] Kaboré, I., Tapsoba, T., Combinatoire de mots récurrents de complexité $n+2$, *RAIRO-Theoret. Inform. Appl.* 41 (2007), 425–446.

[13] Lothaire, M., *Algebraic combinatorics on words*, Cambridge University Press, 2002.

[14] Paul, M. E., Minimal symbolic flows having minimal block growth, *Math. Systems Theory* 8 (1975), 309–315.

[15] Pytheas Fogg, N., *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Mathematics vol. 1794, Springer-Verlag, Berlin, 2002.

Received: April, 2012