



HAL
open science

On the Complexity of the Generalized Fibonacci Words

Julien Cassaigne, Idrissa Kaboré

► **To cite this version:**

Julien Cassaigne, Idrissa Kaboré. On the Complexity of the Generalized Fibonacci Words. *RAIRO - Theoretical Informatics and Applications (RAIRO: ITA)*, 2022, 56, pp.5. 10.1051/ita/2022007. hal-03866699

HAL Id: hal-03866699

<https://hal.science/hal-03866699>

Submitted on 22 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

ON THE COMPLEXITY OF THE GENERALIZED FIBONACCI WORDS

JULIEN CASSAIGNE¹ AND IDRISSE KABORE^{2,*} 

Abstract. In this paper we undertake a general study of the complexity function of the generalized Fibonacci words which are generated by the morphism defined by $\sigma_{l,m}(a) = a^l b^m$ and $\sigma_{l,m}(b) = a$.

Mathematics Subject Classification. 68R15.

Received September 4, 2019. Accepted February 25, 2022.

1. INTRODUCTION

The complexity function \mathbf{p} , which counts the number of factors of given length in an infinite word, is a central notion in the field of combinatorics on words. It was introduced in 1975 by Ehrenfeucht *et al.* [8]. It allows one to measure diversity of patterns in an infinite word. It is often used in characterization of some words or families of words; for example eventually periodic words are the only words with bounded complexity function. For more details on this notion we refer the reader to [4, 7].

Let σ be the morphism of the free monoid $\{a, b\}^*$ defined by $\sigma(a) = ab$ and $\sigma(b) = a$. By iterating infinitely many times the morphism σ from a we obtain an infinite word called the Fibonacci word $F = abaababaabaababaab \dots$. This word was widely studied [2, 9, 11, 12] and it is currently very famous for its numerous remarkable properties. The reader may consult [3] for more details on it. Its complexity function is well-known: for any n it admits exactly $n + 1$ factors of length n .

The generalized Fibonacci morphisms of the free monoid $\{a, b\}^*$ are the morphisms $\sigma_{l,m}$ defined by $\sigma_{l,m}(a) = a^l b^m$ and $\sigma_{l,m}(b) = a$, for $l \geq 1$ and $m \geq 2$. By iterating infinitely many times the morphism $\sigma_{l,m}$ from a we obtain an infinite word $F_{l,m}$ called a generalized Fibonacci word (see [1], p. 336). In this paper we are interested in the complexity function of these words.

Precisely, we recall in Section 2 some basic definitions and notations. In Section 3 we describe weak and strong bispecial factors of $F_{l,m}$. These are specific factors which play an important role in the study of the complexity function of an infinite word. Section 4 is devoted to the complexity function of $F_{l,m}$. Then, we study asymptotic behavior of $\frac{\mathbf{p}(n)}{n}$ (Sect. 5). We conclude the paper with some remarks and problems for further work. Note that a preliminary version of this paper was presented at CARI'2012 [6].

Keywords and phrases: Infinite words, special factors, morphisms, complexity.

¹ Inst. de Math. de Marseille, 163 avenue de Luminy, case 907, F-13288 Marseille Cedex 9, France.

² UFR-Sciences Exactes et Appliquées, Université Nazi BONI, 01 BP 1091 Bobo-Dioulasso 01, Burkina Faso.

* Corresponding author: ikaborei@yahoo.fr

2. PRELIMINARIES

We recall here basic notions on words (see for instance [1, 10] for more details).

Let $\mathcal{A} = \{a, b\}$ be a fixed alphabet. \mathcal{A}^* , the set of finite words on \mathcal{A} , is the free monoid generated by \mathcal{A} ; ε the empty word being the neutral element. For any $u \in \mathcal{A}^*$, $|u|$ is called the length of u and represents the number of letters 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 repeated single letter x is simply denoted $u = x^n$, by extension $x^0 = \varepsilon$.

An infinite word is a sequence of letters of \mathcal{A} . The set of infinite words over \mathcal{A} is denoted \mathcal{A}^∞ . A finite word v is a factor of a word u if there exist two words u_1 and u_2 on \mathcal{A} such that $u = u_1vu_2$; we say also that u contains v . The factor v is a prefix (resp. suffix) if u_1 (resp. u_2) is the empty word. We denote by $pref(w)$ (resp. $suf(w)$) the set of prefixes (resp. suffixes) of w .

Let u be an infinite word on \mathcal{A} , w a factor of u and x a letter of \mathcal{A} . The set of factors of u of length n is denoted $\mathcal{L}_n(u)$ and the set of all factors of u , $\mathcal{L}(u)$. The set $\mathcal{L}(u)$ is usually called the language of u . A letter x is a left (resp. right) extension of w in u if xw (resp. wx) is in $\mathcal{L}(u)$. The factor w is a left (resp. right) special factor of u if aw and bw (respectively wa and wb) appear in u . A factor of u which is both left special and right special in u is a bispecial factor.

The complexity function of an infinite word u is the map from \mathbb{N} to \mathbb{N}^* defined by $\mathbf{p}_u(n) = \#\mathcal{L}_n(u)$, where $\#\mathcal{L}_n(u)$ designates the cardinality of the set of factors of u with length n . In all the sequel, the complexity function \mathbf{p}_u of a word u will be simply denoted \mathbf{p} .

We call the function denoted \mathbf{s} , and defined by $\mathbf{s}(n) = \mathbf{p}(n+1) - \mathbf{p}(n)$, the first difference of the complexity function of a word u . So, we have the following formula

$$\mathbf{p}(n) = \mathbf{p}(k_0) + \sum_{k=k_0}^{n-1} \mathbf{s}(k).$$

On a binary alphabet the function \mathbf{s} counts the number of right special factors of a given length in u . It happens that enumeration of some specific bispecial factors allows one to determine the function \mathbf{s} (see [7]). We will come back to this in Sections 3 and 4.

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 $x, f(x), f^2(x), \dots, f^n(x), \dots$ are longer and longer prefixes of u . Then we denote $u = f^\omega(x)$.

Let u be an infinite word on \mathcal{A} and v a factor of u . The Parikh vector of v is $\chi(v) = {}^t(|v|_a, |v|_b)$. We call the following matrix

$$M_\varphi = \begin{pmatrix} |\varphi(a)|_a & |\varphi(b)|_a \\ |\varphi(a)|_b & |\varphi(b)|_b \end{pmatrix},$$

the incidence matrix of a morphism φ . Observe that $\chi(\varphi(v)) = M_\varphi \chi(v)$.

3. NON-ORDINARY BISPECIAL FACTORS OF $F_{l,m}$

Definition 3.1. Let u be an infinite word on \mathcal{A} and v a bispecial factor of u .

- v is called strong bispecial if ava, avb, bva, bvb are factors of u .
- v is called weak bispecial if uniquely ava and bvb , or avb and bva , are factors of u .
- v is called ordinary bispecial if v is neither strong nor weak.

Definition 3.2. A factor of $F_{l,m}$ is said to be short if it does not contain any of the three words a^l, b^m and ba . A factor of $F_{l,m}$ which is not short will be called long.

Lemma 3.3. *Let w be a long factor of $F_{l,m}$. Then, there exists a unique triple of words (p, s, v) verifying $p \in \text{pref}(a^l b^{m-1})$, $s \in \text{suf}(a^{l-1} b^m)$ and $v \in \mathcal{L}(F_{l,m})$ such that $w = s\sigma_{l,m}(v)p$ and $(v \in \mathcal{A}^*b \implies |p| \geq l)$.*

Proof. Existence. Let w be a long factor of $F_{l,m}$. Then, either w is factor of $\sigma_{l,m}(x)$ where $x \in \mathcal{A}$ or $w = s\sigma_{l,m}(v)p$ where s is a proper suffix of $\sigma_{l,m}(x)$, p is a proper prefix of $\sigma_{l,m}(y)$ with $x, y \in \mathcal{A}$ and $xvy \in \mathcal{L}(F_{l,m})$. More precisely, $p \in \text{pref}(a^l b^{m-1})$, $s \in \text{suf}(a^{l-1} b^m)$ if $x = a$ and $s = \varepsilon$ if $x = b$. If $v \notin \mathcal{A}^*b$ it is finished. Suppose $v \in \mathcal{A}^*b$ and $|p| < l$. So, $p = a^{|p|}$. In this case one changes v and p as follows:

$$v \longleftarrow vb^{-1}, \quad p \longleftarrow ap.$$

We still have $w = s\sigma_{l,m}(v)p$ with $p \in \text{pref}(a^l b^{m-1})$ and $|p|$ has increased. We repeat this process until to get $v \notin \mathcal{A}^*b$ or $|p| \geq l$.

Uniqueness. Let w be a long factor of $F_{l,m}$. Suppose $w = s\sigma_{l,m}(v)p = s'\sigma_{l,m}(v')p'$ where

1. $p, p' \in \text{pref}(a^l b^{m-1})$
2. $s, s' \in \text{suf}(a^{l-1} b^m)$
3. $(s, v, p) \neq (s', v', p')$

and verifying:

$$v \in \mathcal{A}^*b \implies |p| \geq l \text{ and } v' \in \mathcal{A}^*b \implies |p'| \geq l. \quad (\star)$$

- Suppose $(v, p) = (v', p')$. Then, we have $s = s'$. That is impossible.
- Suppose $p = p'$ and $v \neq v'$. Then, we have $s\sigma_{l,m}(v) = s'\sigma_{l,m}(v')$.
 - If v and v' are not empty then v and v' must end with the same letter. Then we change v to $\text{pref}_{|v|-1}(v)$ and v' to $\text{pref}_{|v'|-1}(v')$. We repeat the process while $v \neq \varepsilon$ and $v' \neq \varepsilon$.
 - If $v \neq \varepsilon$ and $v' = \varepsilon$ (or conversely) then we have $s\sigma_{l,m}(v) = s'$. Now, we have $|s'| < l + m$. It follows that $0 < |\sigma_{l,m}(v)| < l + m$. Thus, we have $v = b^k$ and $sa^k = s'$. But s' ends with b . That is impossible.
- Suppose $p \neq p'$. Without loss generality let us assume that $|p| > |p'|$. Then p can be written $p = p''p'$ with $p'' \neq \varepsilon$. So, it follows that $s\sigma_{l,m}(v)p'' = s'\sigma_{l,m}(v')$.
 - If v' is empty then $s\sigma_{l,m}(v)p'' = s'$. Now, we have $|s'| < l + m$. So, v takes the form $v = b^k$ and we have $sa^k p'' = s'$. Since $p'' \neq \varepsilon$ then $s' \neq \varepsilon$ and ends with b . Therefore, p'' also ends with b . So, we can write $p'' = a^l b^i$, $i \geq 1$. Thus, s' contains a^l . That is impossible since $s' \in \text{suf}(a^{l-1} b^m)$.
 - If $v' \neq \varepsilon$, let x be the last letter of p'' and of $\sigma_{l,m}(v')$.
 - If $x = a$, then the last letter of v' is b and by (\star) we have $|p'| \geq l$. So, we have $p' = a^l z$ and $p'' = ya$. That implies $p''p' = ya^{l+1}z$. That is impossible.
 - If $x = b$, then the last letter of v' is a and $p'' = a^l b^i$ ($0 < i < m$). Now, the suffix of $\sigma_{l,m}(v')$ of length m is b^m . So, the word $s'\sigma_{l,m}(v') = s\sigma_{l,m}(v)p''$ admits the two words b^m and ab^i as suffix. That is impossible.

□

Let us note that if a factor w is short then it is a factor of $a^{l-1}b^{m-1}$.

Lemma 3.4.

1. $F_{l,m}$ admits exactly one short and weak bispecial factor: b^{m-1} .
2. $F_{l,m}$ admits exactly one short and strong bispecial factor: ε .

Lemma 3.5. *Let w be a factor of $F_{l,m}$. The following assertions are equivalent:*

1. w is a long bispecial factor of $F_{l,m}$.

2. There exists a bispecial factor v of $F_{l,m}$ such that $w = \hat{\sigma}_{l,m}(v)$ where $\hat{\sigma}_{l,m}(v) = \sigma_{l,m}(v)a^l$.
Furthermore, v and w have the same type and $|v| < |w|$.

Proof. Let w be a long bispecial factor. Then wa, wb, aw, bw appear in $F_{l,m}$. Furthermore with the synchronization Lemma there exists a unique triple of words (p, s, v) verifying $p \in \text{pref}(a^l b^{m-1})$, $s \in \text{su}f(a^{l-1} b^m)$ and $v \in \mathcal{L}(F_{l,m})$ such that $w = s\sigma_{l,m}(v)p$ and $(v \in \mathcal{A}^*b \implies |p| \geq l)$. So, the words $s\sigma_{l,m}(v)pa, s\sigma_{l,m}(v)pb, as\sigma_{l,m}(v)p, bs\sigma_{l,m}(v)p$ appear in $F_{l,m}$.

Suppose $\sigma_{l,m}(v)p = \varepsilon$, i.e $w = s = a^i b^m$ with $i < l$. Then, $bw = ba^i b^m$ appears in $F_{l,m}$. That is impossible. Suppose $\sigma_{l,m}(v)p \neq \varepsilon$, $\sigma_{l,m}(v)p$ begin with a . Then, asa and bsa appear in $F_{l,m}$. If $s = b^j$ with $0 < j < m$, then $ab^j a$ appears in $F_{l,m}$, which is impossible. If $s = a^i b^j$ with $0 < i < l$, then $ba^i b^m$ appears in $F_{l,m}$, which is also impossible. Thus $s = \varepsilon$ and we have $w = \sigma_{l,m}(v)p$.

Let us now show that $p = a^l$.

Suppose $|p| > l$. Then, $p = a^l b^i$ with $0 < i < m$. So, $ab^i a$ appears in $F_{l,m}$, which is impossible.

Suppose $|p| < l$. Then $p = a^i$, with $0 \leq i < l$. So, $v \notin \mathcal{A}^*b$. If $v = \varepsilon$ then $w = p = a^i$ and so w is short in $F_{l,m}$. This is impossible because w is assumed long. Otherwise if $v \neq \varepsilon$ then v ends with a . So, the factor $wb = \sigma_{l,m}(v)pb$ of $F_{l,m}$ ends with $a^l b^m a^i b$. Thus, $ba^i b$ appears in $F_{l,m}$ with $0 \leq i < l$, which is again impossible. It follows that $p = a^l$, so $w = \hat{\sigma}_{l,m}(v)$.

The inequality $|v| < |w|$ is obvious.

Conversely, assume that v is a bispecial factor of $F_{l,m}$ and that $w = \hat{\sigma}_{l,m}(v)$. As the words av, bv, va and vb occur in $F_{l,m}$, it follows that $a^l b^m w, aw, wb^m a^l$ and wa occur in $F_{l,m}$. So, w is a bispecial factor of $F_{l,m}$, which is long since it contains a^l .

Finally, if $w = \hat{\sigma}_{l,m}(v)$, then

$$\# \{(x, y) \in \mathcal{A}^2 : xwy \in \mathcal{L}(F_{l,m})\} = \# \{(x', y') \in \mathcal{A}^2 : x'vy' \in \mathcal{L}(F_{l,m})\}.$$

So, v and w have the same type. □

As a consequence, we have:

1. The weak bispecial factors of $F_{l,m}$ are given by the sequence (y_n) defined by $y_1 = b^{m-1}$ and $y_{n+1} = \hat{\sigma}_{l,m}(y_n)$, for $n \geq 1$.
2. The strong bispecial factors of $F_{l,m}$ are given by the sequence (x_n) defined by $x_0 = \varepsilon$ and $x_{n+1} = \hat{\sigma}_{l,m}(x_n)$, for $n \geq 0$.

4. COMPLEXITY OF $F_{l,m}$

In order to understand the complexity function of $F_{l,m}$, we begin this section with a review of some properties of sequences of weak bispecial and strong bispecial factors of $F_{l,m}$.

Definition 4.1. Let $v, w \in \mathcal{A}^*$ and $\chi(v), \chi(w)$ be their Parikh vectors. One says that $\chi(v)$ is less than $\chi(w)$ and one writes $\chi(v) \leq \chi(w)$ when $|v|_x \leq |w|_x$ for all $x \in \mathcal{A}$. Moreover, if $\chi(v) \neq \chi(w)$, one writes $\chi(v) < \chi(w)$.

Note that this define a partial order on words.

Proposition 4.2. Let v, w, v', w' be four finite words such that $v' = \hat{\sigma}_{l,m}(v)$ and $w' = \hat{\sigma}_{l,m}(w)$. Then,

$$\chi(v) < \chi(w) \implies \chi(v') < \chi(w')$$

Proof. On the one hand, we have $|v'|_a = l|v|_a + |v|_b + l$ and $|w'|_a = l|w|_a + |w|_b + l$; so $|v'|_a < |w'|_a$.

On the other hand, we have $|v'|_b = m|v|_a$ and $|w'|_b = m|w|_a$; so $|v'|_b \leq |w'|_b$. □

Proposition 4.3. *For all $l \geq 1$ and $m \geq 2$ one has:*

$$\forall n \geq 0, \chi(x_n) < \chi(y_{n+1}) < \chi(x_{n+2})$$

Proof. Since $x_0 = \varepsilon$, $y_1 = b^{m-1}$ and $x_2 = (a^l b^m)^l a^l$ we have $\chi(x_0) < \chi(y_1) < \chi(x_2)$. Suppose these inequalities stay valid until rank n , i.e.,

$$\chi(x_n) < \chi(y_{n+1}) < \chi(x_{n+2}).$$

By Proposition 4.2, we obtain $\chi(x_{n+1}) < \chi(y_{n+2}) < \chi(x_{n+3})$. □

The following Lemma describes the function \mathbf{s} .

Lemma 4.4. *Let $n \in \mathbb{N}$. One has:*

- $\mathbf{s}(n) = 1$ for $n = 0$.
- if $n \in \mathbb{N}^*$, take k the largest integer such that $n > |x_k|$.
 1. If $k = 0$, one has: $\mathbf{s}(n) = \begin{cases} 2 & \text{if } 1 \leq n \leq \min(l, m-1) \\ 1 & \text{if } m \leq n \leq l \end{cases}$.
 2. Otherwise, one has: $\mathbf{s}(n) = \begin{cases} 3 & \text{if } |x_k| < n \leq |y_k| \\ 2 & \text{if } |y_k| < n \leq |y_{k+1}| \\ 1 & \text{if } |y_{k+1}| < n \leq |x_{k+1}| \end{cases}$.

Proof. The function \mathbf{s} is given by the following formula

$$\mathbf{s}(n) = 1 + \# \{w \text{ strong bispecial} : |w| < n\} - \# \{w \text{ weak bispecial} : |w| < n\}.$$

Now $\mathbf{s}(0) = 1$. Also, let us observe that $\mathbf{s}(n) = 2$ if $1 \leq n \leq \min(l, m-1)$ and $\mathbf{s}(n) = 1$ if $m \leq n \leq l$.

Suppose $n \geq |x_1|$. Take k the largest integer such that $n > |x_k|$. Then, it follows:

$$\mathbf{s}(n) = 1 + (k+1) - \begin{cases} k-1 & \text{if } |x_k| < n \leq |y_k| \\ k & \text{if } |y_k| < n \leq |y_{k+1}| \\ k+1 & \text{if } |y_{k+1}| < n \leq |x_{k+1}| \end{cases}.$$

The proof is complete. □

Theorem 4.5. *The complexity function of $F_{l,m}$ satisfies the following inequalities:*

$$n+1 \leq \mathbf{p}(n) \leq 3n+1.$$

Proof. By Lemma 4.4, we have

$$\forall n \geq 0, 1 \leq \mathbf{s}(n) \leq 3.$$

It follows that:

$$1 + \sum_{k=0}^{n-1} 1 \leq \mathbf{p}(n) \leq 1 + \sum_{k=0}^{n-1} 3.$$

□

Lemma 4.6. *We have the following equivalences.*

1. $m \in [2, 2l^2 + 1] \iff \exists k_0 \in \mathbb{N} : \forall k \geq k_0, |x_k| - |y_k| > 0.$
2. $m \in [2l^2 + 2, \infty[\iff \exists k_0 \in \mathbb{N} : \forall k \geq k_0, |x_k| - |y_k| < 0.$

Proof. Consider the sequence $(V_k)_{k \geq 1}$ defined by $V_k = \chi(x_k) - \chi(y_k)$. We have:

$$V_1 = \begin{pmatrix} l \\ -m + 1 \end{pmatrix} \text{ and } V_{k+1} = AV_k$$

where $A = M_\sigma = \begin{pmatrix} l & 1 \\ m & 0 \end{pmatrix}$ is the incidence matrix of $\sigma_{l,m}$. The eigenvalues of the matrix A , being the roots of $X^2 - lX - m$, are

$$\lambda_1 = \frac{l + \sqrt{l^2 + 4m}}{2} \text{ and } \lambda_2 = \frac{l - \sqrt{l^2 + 4m}}{2}.$$

Observe that $\lambda_1 > l \geq 1$ and $-\lambda_1 < \lambda_2 < 0$. Moreover we have

$$\forall k \geq 1, |x_k| - |y_k| = \begin{pmatrix} 1 & 1 \end{pmatrix} V_k.$$

Thus,

$$\begin{aligned} |x_k| - |y_k| &= \begin{pmatrix} 1 & 1 \end{pmatrix} A^{k-1} \begin{pmatrix} l \\ -m + 1 \end{pmatrix} \\ &= \alpha_1 \lambda_1^{k-1} + \alpha_2 \lambda_2^{k-1} \end{aligned}$$

with α_1 and α_2 verifying the following system of equations

$$\begin{cases} \alpha_1 + \alpha_2 &= l - m + 1 \\ \alpha_1 \lambda_1 + \alpha_2 \lambda_2 &= l^2 + lm - m + 1 \end{cases}.$$

We have

$$\alpha_1 = \frac{l^2 + lm - m + 1 - \lambda_2(l - m + 1)}{\lambda_1 - \lambda_2},$$

$$\alpha_2 = \frac{l^2 + lm - m + 1 - \lambda_1(l - m + 1)}{\lambda_2 - \lambda_1}.$$

Since $|\lambda_2| < \lambda_1$, then $|x_k| - |y_k|$ has the same sign as α_1 for k sufficiently large.

Case 1. $l - m + 1 \geq 0$. Then, we have $-\lambda_2(l - m + 1) \geq 0$ since $\lambda_2 < 0$. So, $\alpha_1 > 0$ since $\lambda_1 - \lambda_2 > 0$ and $l^2 + lm - m + 1 = l^2 + l(m - 1) + 1 > 0$.

Case 2. $l - m + 1 < 0$. We have:

$$\begin{aligned} \alpha_1 > 0 &\iff \lambda_2 > \frac{l^2 + lm - m + 1}{l - m + 1} \\ &\iff \frac{l - \sqrt{l^2 + 4m}}{2} > \frac{l^2 + lm - m + 1}{l - m + 1} \end{aligned}$$

$$\iff l^2 + 4m < \frac{(-l^2 - 3lm + l + 2m - 2)^2}{(l - m + 1)^2}$$

since $-l^2 - 3lm + l + 2m - 2 = m(2 - 3l) - l(l - 1) - 2 < 0$ and $l - m + 1 < 0$. The last inequality can be turned into the following one:

$$P_l(m) = m^3 + m^2(-2l^2 + l - 3) + m(-2l^3 + 3l^2 - 2l + 3) + l^3 - l^2 + l - 1 < 0.$$

If $l = 1$ then $P_l(m) = m(m^2 - 4m + 2)$. So,

$$P_l(m) < 0 \iff m \in \{2, 3\}.$$

Suppose from now on that $l \geq 2$. The derivative

$$P'_l(m) = 3m^2 + 2m(-2l^2 + l - 3) - 2l^3 + 3l^2 - 2l + 3$$

admits two roots of opposite signs

$$\beta_1 = \frac{2l^2 - l + 3 - \sqrt{4l^4 + 2l^3 + 4l^2}}{3} < 0,$$

$$\beta_2 = \frac{2l^2 - l + 3 + \sqrt{4l^4 + 2l^3 + 4l^2}}{3} > 0$$

and is negative between these two roots. So, P_l is decreasing on $[\beta_1, \beta_2]$ which contains 0, and is increasing on $]-\infty, \beta_1] \cup [\beta_2, +\infty[$. Furthermore, one verifies that $P_l(0) > 0$, $P_l(1) < 0$, $P_l(2l^2 + 1) < 0$ and $P_l(2l^2 + 2) > 0$. It results that P_l admits two positive roots m_1, m_2 and one negative root m_3 with $0 < m_1 < 1$, $2l^2 + 1 < m_2 < 2l^2 + 2$ and $m_1 \leq \beta_2 \leq m_2$. Thus, P_l is negative on $]m_1, m_2[$ and positive on $]m_2, +\infty[$. Then, $\alpha_1 > 0$ if $m \in [2, 2l^2 + 1]$ and $\alpha_1 < 0$ if $m \geq 2l^2 + 2$.

So, we have:

1. $m \in [2, 2l^2 + 1] \iff \exists k_0 \in \mathbb{N} : \forall k \geq k_0, |x_k| - |y_k| > 0$.
2. $m \in [2l^2 + 2, \infty[\iff \exists k_0 \in \mathbb{N} : \forall k \geq k_0, |x_k| - |y_k| < 0$.

□

Theorem 4.7. 1. If $m \in [2, 2l^2 + 1]$, then there exists a constant δ and an integer n_0 such that for all $n > n_0$

$$n + 1 \leq \mathbf{p}(n) \leq 2n + \delta.$$

2. If $m \in [2l^2 + 2, +\infty[$, then there exists a constant δ and an integer n_0 such that for all $n > n_0$

$$2n + \delta \leq \mathbf{p}(n) \leq 3n + 1.$$

Proof. Suppose $m \in [2, 2l^2 + 1]$. Then, there exists k_0 such that for all $k \geq k_0$, $|x_k| - |y_k| > 0$. In this case, for $n \in]|x_k|, |x_{k+1}|]$ we have by Lemma 4.4

$$\mathbf{s}(n) = \begin{cases} 2 & \text{if } |y_k| < n \leq |y_{k+1}| \\ 1 & \text{if } |y_{k+1}| < n \leq |x_{k+1}| \end{cases}$$

as the case $|x_k| < n \leq |y_k|$ is empty. Thus,

$$\forall n \geq |x_{k_0}|, 1 \leq \mathbf{s}(n) \leq 2.$$

By summation, it follows that

$$\sum_{k=|x_{k_0}|}^{n-1} 1 \leq \sum_{k=|x_{k_0}|}^{n-1} \mathbf{s}(k) \leq \sum_{k=|x_{k_0}|}^{n-1} 2.$$

Therefore

$$\mathbf{p}(|x_{k_0}|) + n - |x_{k_0}| \leq \mathbf{p}(n) \leq \mathbf{p}(|x_{k_0}|) + 2(n - |x_{k_0}|).$$

It follows that

$$n + 1 \leq \mathbf{p}(n) \leq 2n + \delta, \text{ with } \delta \in \mathbb{Z}.$$

Suppose $m \geq 2l^2 + 2$. Then, by Lemma 4.6, there exists k_0 such that for all $k \geq k_0$, $|x_k| - |y_k| < 0$. In this case, for $n \in]|x_k|, |x_{k+1}]$ we have

$$\mathbf{s}(n) = \begin{cases} 3 & \text{if } |x_k| < n \leq |y_k| \\ 2 & \text{if } |y_k| < n \leq |x_{k+1}| \end{cases}.$$

It follows that,

$$\forall n \geq |x_{k_0}|, 2 \leq \mathbf{s}(n) \leq 3.$$

Thus, in the similar way as previously, we get

$$2n + \delta \leq \mathbf{p}(n) \leq 3n + 1, \text{ with } \delta \in \mathbb{Z}.$$

□

5. ASYMPTOTIC BEHAVIOR OF $\frac{p(n)}{n}$

Before the statement of the main result we need some technical lemmas.

Lemma 5.1. *Let $(r_k), (s_k)$ be two strictly increasing sequences of integers and l be a real number such that:*

$$\lim_{k \rightarrow \infty} \frac{s_{k+1} - s_k}{r_{k+1} - r_k} = l. \text{ Then, } \lim_{k \rightarrow \infty} \frac{s_k}{r_k} = l.$$

Proof. Let $(r_k), (s_k)$ be two strictly increasing sequences of integers and l be a real number such that

$$\lim_{k \rightarrow \infty} \frac{s_{k+1} - s_k}{r_{k+1} - r_k} = l. \text{ Write } m_k = \frac{s_{k+1} - s_k}{r_{k+1} - r_k}. \text{ Let } \varepsilon > 0. \text{ There exists } k_0 \in \mathbb{N} \text{ such that:}$$

$$\forall k \geq k_0, l - \varepsilon \leq m_k \leq l + \varepsilon. \tag{5.1}$$

So, for all $k > k_0$ we have

$$\begin{aligned}
s_k &= s_{k_0} + \sum_{j=k_0}^{k-1} (s_{j+1} - s_j) \\
&= s_{k_0} + \sum_{j=k_0}^{k-1} m_j (r_{j+1} - r_j) \\
&= s_{k_0} + \sum_{j=k_0}^{k-1} l (r_{j+1} - r_j) + \sum_{j=k_0}^{k-1} (m_j - l) (r_{j+1} - r_j) \\
&= s_{k_0} + l (r_k - r_{k_0}) + R_k
\end{aligned} \tag{5.2}$$

where $R_k = \sum_{j=k_0}^{k-1} (m_j - l) (r_{j+1} - r_j)$. Thereby $\frac{s_k}{r_k} = l + \frac{s_{k_0} - l r_{k_0}}{r_k} + \frac{R_k}{r_k}$. From (5.1) and since (r_k) is increasing, it follows that $|R_k| \leq \varepsilon (r_k - r_{k_0})$ and so $R_k \leq \varepsilon r_k$.

Thus, $\lim_{k \rightarrow \infty} \frac{s_k}{r_k} = l$. \square

Lemma 5.2. For all $k \geq 1$, $|x_k| = \beta_1 \lambda_1^k + \beta_2 \lambda_2^k + \beta_3$ and $|y_k| = \gamma_1 \lambda_1^k + \gamma_2 \lambda_2^k + \gamma_3$ where λ_1 and λ_2 are the eigenvalues of matrix A , and $(\beta_1, \beta_2, \beta_3)$, $(\gamma_1, \gamma_2, \gamma_3)$ are some triples of real numbers.

Proof. Let w, w' be two words such that $w' = \hat{\sigma}_{l,m}(w) = \sigma_{l,m}(w) a^l$. Put $W = \begin{pmatrix} |w|_a \\ |w|_b \\ 1 \end{pmatrix}$. Observe that

$|w| = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} W$ and $W' = BW$ where $B = \begin{pmatrix} l & 1 & l \\ m & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. With these relations we are able to determine

the two sequences $(x_k)_{k \geq 0}$, $(y_k)_{k \geq 1}$, and the length of x_k and y_k for all k .

Namely, since $x_{k+1} = \hat{\sigma}_{l,k}(x_k)$, $y_{k+1} = \hat{\sigma}_{l,k}(y_k)$, $x_0 = \varepsilon$ and $y_1 = b^{m-1}$ we have $X_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, $Y_1 = \begin{pmatrix} 0 \\ m-1 \\ 1 \end{pmatrix}$,

$X_{k+1} = BX_k$ and $Y_{k+1} = BY_k$, and so $X_k = B^k X_0$ and $Y_k = B^{k-1} Y_1$. The eigenvalues of the matrix B are λ_1 , λ_2 and 1. Then, it follows that

$$|x_k| = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} B^k X_0 \tag{5.3}$$

$$= \beta_1 \lambda_1^k + \beta_2 \lambda_2^k + \beta_3 \tag{5.4}$$

where $(\beta_1, \beta_2, \beta_3)$ is the solution of the following system of equations

$$\begin{cases} \beta_1 + \beta_2 + \beta_3 &= 0 \\ \beta_1 \lambda_1 + \beta_2 \lambda_2 + \beta_3 &= l \\ \beta_1 \lambda_1^2 + \beta_2 \lambda_2^2 + \beta_3 &= l^2 + lm + l \end{cases}$$

and

$$|y_k| = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix} B^{k-1} Y_1 \tag{5.5}$$

$$= \gamma_1 \lambda_1^k + \gamma_2 \lambda_2^k + \gamma_3 \tag{5.6}$$

where $(\gamma_1, \gamma_2, \gamma_3)$ is given by the following system of equations

$$\begin{cases} \gamma_1 \lambda_1 + \gamma_2 \lambda_2 + \gamma_3 & = m - 1 \\ \gamma_1 \lambda_1^2 + \gamma_2 \lambda_2^2 + \gamma_3 & = m - 1 + l \\ \gamma_1 \lambda_1^3 + \gamma_2 \lambda_2^3 + \gamma_3 & = (m + l)^2 - m \end{cases} .$$

□

Theorem 5.3. 1. If $m \in [1, 2l^2 + 1]$, we have

$$\liminf \frac{p(n)}{n} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\beta_1 (\lambda_1 - 1)}, \text{ and } \limsup \frac{p(n)}{n} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\gamma_1 (\lambda_1 - 1)}.$$

2. If $m \geq 2l^2 + 2$ we have

$$\liminf \frac{p(n)}{n} = 2 + \frac{\gamma_1 - \beta_1}{\beta_1 (\lambda_1 - 1)} \text{ and } \limsup \frac{p(n)}{n} = 2 + \frac{\lambda_1 (\gamma_1 - \beta_1)}{\gamma_1 (\lambda_1 - 1)}.$$

Proof. Case 1. $m \in [1, 2l^2 + 1]$. From Lemma 4.6, there exists k_0 such that for all $k \geq k_0$,

$$|y_k| < |x_k| < |y_{k+1}| < |x_{k+1}|.$$

In this case, we have

$$\forall n \in]|x_k|, |x_{k+1}|], \mathbf{s}(n) = \begin{cases} 2 & \text{if } |x_k| < n \leq |y_{k+1}| \\ 1 & \text{if } |y_{k+1}| < n \leq |x_{k+1}| \end{cases}.$$

So, $\liminf \frac{\mathbf{p}(n)}{n} = \liminf \frac{\mathbf{p}(|x_k|+1)}{|x_k|+1}$ and $\limsup \frac{\mathbf{p}(n)}{n} = \limsup \frac{\mathbf{p}(|y_k|+1)}{|y_k|+1}$. Furthermore, with Lemma 5.2 we have

$$\begin{aligned} \mathbf{p}(|x_{k+1}| + 1) - \mathbf{p}(|x_k| + 1) &= \sum_{n=|x_k|+1}^{|x_{k+1}|} \mathbf{s}(n) \\ &= 2(|y_{k+1}| - |x_k|) + (|x_{k+1}| - |y_{k+1}|) \\ &= 2(\gamma_1 \lambda_1^{k+1} - \beta_1 \lambda_1^k) + (\beta_1 \lambda_1^{k+1} - \gamma_1 \lambda_1^{k+1}) + o(\lambda_1^k) \\ &= \gamma_1 \lambda_1^{k+1} + \beta_1 \lambda_1^{k+1} - 2\beta_1 \lambda_1^k + o(\lambda_1^k). \end{aligned} \tag{5.7}$$

Moreover

$$|x_{k+1}| - |x_k| = \beta_1 \lambda_1^{k+1} - \beta_1 \lambda_1^k + o(\lambda_1^k).$$

So

$$\frac{\mathbf{p}(|x_{k+1}| + 1) - \mathbf{p}(|x_k| + 1)}{|x_{k+1}| - |x_k|} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\beta_1 (\lambda_1 - 1)} + o(1).$$

By Lemma 5.1 it follows that $\lim_{k \rightarrow \infty} \frac{\mathbf{p}(|x_k|+1)}{|x_k|+1} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\beta_1 (\lambda_1 - 1)}$. In a similar way, we have

$$\begin{aligned} \mathbf{p}(|y_{k+1}|+1) - \mathbf{p}(|y_k|+1) &= \sum_{n=|y_k|+1}^{|y_{k+1}|} \mathbf{s}(n) \\ &= (|x_k| - |y_k|) + 2(|y_{k+1}| - |x_k|) \\ &= 2\gamma_1 \lambda_1^{k+1} - \gamma_1 \lambda_1^k - \beta_1 \lambda_1^k + o(\lambda_1^k) \end{aligned} \quad (5.8)$$

and

$$|y_{k+1}| - |y_k| = \gamma_1 \lambda_1^{k+1} - \gamma_1 \lambda_1^k + o(\lambda_1^k).$$

So, $\lim_{k \rightarrow \infty} \frac{\mathbf{p}(|y_k|+1)}{|y_k|+1} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\gamma_1 (\lambda_1 - 1)}$.

Case 2. $m > 2l^2 + 1$. From Lemma 4.6, there exists k_0 such that for all $k \geq k_0$,

$$|x_k| < |y_k| < |x_{k+1}| < |y_{k+1}|.$$

In this case, we have

$$\forall n \in]|x_k|, |x_{k+1}][, \mathbf{s}(n) = \begin{cases} 3 & \text{if } |x_k| < n \leq |y_k| \\ 2 & \text{if } |y_k| < n \leq |x_{k+1}| \end{cases}.$$

So, $\liminf \frac{\mathbf{p}(n)}{n} = \liminf \frac{\mathbf{p}(|x_k|+1)}{|x_k|+1}$ and $\limsup \frac{\mathbf{p}(n)}{n} = \limsup \frac{\mathbf{p}(|y_k|+1)}{|y_k|+1}$. Furthermore

$$\begin{aligned} \mathbf{p}(|x_{k+1}|+1) - \mathbf{p}(|x_k|+1) &= \sum_{n=|x_k|+1}^{|x_{k+1}|} \mathbf{s}(n) \\ &= 3(|y_k| - |x_k|) + 2(|x_{k+1}| - |y_k|) \\ &= 2\beta_1 \lambda_1^{k+1} - 3\beta_1 \lambda_1^k + \gamma_1 \lambda_1^k + o(\lambda_1^k) \end{aligned} \quad (5.9)$$

and

$$|x_{k+1}| - |x_k| = \beta_1 \lambda_1^{k+1} - \beta_1 \lambda_1^k + o(\lambda_1^k).$$

So

$$\frac{\mathbf{p}(|x_{k+1}|+1) - \mathbf{p}(|x_k|+1)}{|x_{k+1}| - |x_k|} = 2 + \frac{\gamma_1 - \beta_1}{\beta_1 (\lambda_1 - 1)} + o(1).$$

By Lemma 5.1 again it follows that

$$\lim_{k \rightarrow \infty} \frac{\mathbf{p}(|x_k|+1)}{|x_k|+1} = 1 + \frac{\gamma_1 \lambda_1 - \beta_1}{\beta_1 (\lambda_1 - 1)}.$$

In a similar way, we have

$$\begin{aligned}
\mathbf{p}(|y_{k+1}| + 1) - \mathbf{p}(|y_k| + 1) &= \sum_{n=|y_k|+1}^{|y_{k+1}|} \mathbf{s}(n) \\
&= 2(|x_{k+1}| - |y_k|) + 3(|y_{k+1}| - |x_{k+1}|) \\
&= 3\gamma_1\lambda_1^{k+1} - 2\gamma_1\lambda_1^k - \beta_1\lambda_1^{k+1} + o(\lambda_1^k)
\end{aligned} \tag{5.10}$$

and

$$|y_{k+1}| - |y_k| = \gamma_1\lambda_1^{k+1} - \gamma_1\lambda_1^k + o(\lambda_1^k)$$

$$\text{So, } \lim_{k \rightarrow \infty} \frac{\mathbf{p}(|y_k|+1)}{|y_k|+1} = 2 + \frac{\lambda_1(\gamma_1 - \beta_1)}{\gamma_1(\lambda_1 - 1)}.$$

□

6. CONCLUDING REMARKS AND FURTHER WORK

It results from Theorem 4.7 that:

$$\text{if } m \in [2, 2l^2 + 1], \text{ then } 1 \leq \liminf \frac{\mathbf{p}(n)}{n} \leq \limsup \frac{\mathbf{p}(n)}{n} \leq 2; \tag{6.1}$$

$$\text{if } m > 2l^2 + 1, \text{ then } 2 \leq \liminf \frac{\mathbf{p}(n)}{n} \leq \limsup \frac{\mathbf{p}(n)}{n} \leq 3. \tag{6.2}$$

By Theorem 5.3 one observes that for $F_{l,m}$, the values of $\liminf \frac{\mathbf{p}(n)}{n}$ and $\limsup \frac{\mathbf{p}(n)}{n}$ are strictly dependent with those of the parameters l and m . Indeed, we check that (6.1) and (6.2) become

$$\text{if } m \in [2, 2l^2 + 1], \text{ then } 1 < \liminf \frac{\mathbf{p}(n)}{n} < \limsup \frac{\mathbf{p}(n)}{n} < 2; \tag{6.3}$$

$$\text{if } m > 2l^2 + 1, \text{ then } 2 < \liminf \frac{\mathbf{p}(n)}{n} < \limsup \frac{\mathbf{p}(n)}{n} < 3. \tag{6.4}$$

For $l \geq 1$, $m \geq 2$, let us write $\alpha_{l,m} = \liminf \frac{\mathbf{p}(n)}{n}$ and $\beta_{l,m} = \limsup \frac{\mathbf{p}(n)}{n}$. In (6.1) the value 1 is reached when $m = 1$. In this case $F_{l,m}$ is Sturmian and we have excluded it by taking $m \geq 2$. The values 2 and 3 are never reached, but we can prove that they are accumulation points for $\alpha_{l,m}$ and $\beta_{l,m}$.

In further work, it will be interesting to describe the region covered by the cloud of points $(\alpha_{l,m}, \beta_{l,m})$ in the first quadrant of the plane.

Another problem is to undertake a similar study in the case of S -adic words where morphisms are all generalized Fibonacci morphisms.

REFERENCES

- [1] J.-P. Allouche and J. Shallit, *Automatic Sequences: Theory, Applications, Generalizations*. Cambridge University Press, UK (2003).
- [2] J. Berstel, *Mots de Fibonacci*, *L.I.T.P. Séminaire d'Informatique Théorique*, Paris (1980–1981), 57–78.
- [3] J. Cassaigne, On extremal properties of the Fibonacci word. *RAIRO-Theor. Inf. Appl.* **42** (2008) 701–715.

- [4] J. Cassaigne, Complexité et facteurs spéciaux, *Bull. Belg. Math. Soc.* **4** (1997) 67–88.
- [5] J. Cassaigne, An algorithm to test if a given circular HD0L-language avoids a pattern, in *IFIP World Computer Congress'94, North-Holland* (1994) 459–464.
- [6] J. Cassaigne and I. Kaboré, Etude de la complexité du mot de Fibonacci généralisé, in: *Proceedings of 11th African Conference on Research in Computer Science and Applied Mathematics, CARI'12* (2012) 62–69.
- [7] J. Cassaigne and F. Nicolas, Complexity, in: *Combinatorics, Automata and Number Theory*, edited by V. Berthé and M. Rigo. Vol. 135 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press (2010).
- [8] A. Ehrenfeucht, K.P. Lee and G. Rozenberg, Subword complexities of various classes of deterministic developmental languages without interaction. *Theoret. Comput. Sci.* **1** (1975) 59–75.
- [9] A. de Luca, A combinatorial property of the Fibonacci words. *Inform. Process. Lett.* **12** (1981) 193–195.
- [10] M. Lothaire, *Algebraic combinatorics on words*. Cambridge University Press (2002).
- [11] F. Mignosi and G. Pirillo, Repetitions in the Fibonacci infinite word. *RAIRO-Theor. Inf. Appl.* **26** (1992) 199–204.
- [12] G. Pirillo, From the Fibonacci word to Sturmian words. *Publ. Math. Debrecen* **54** (1999) 961–971.

Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/math-s2o-programme>