Open and Closed Prefixes of Sturmian Words

Alessandro De Luca[1] and Gabriele Fici[2]

[1] DIETI, Università di Napoli Federico II, Italy
alessandro.deluca@unina.it
[2] Dipartimento di Matematica e Informatica, Università di Palermo, Italy
gabriele.fici@unipa.it

Abstract. A word is closed if it contains a proper factor that occurs both as a prefix and as a suffix but does not have internal occurrences, otherwise it is open. We deal with the sequence of open and closed prefixes of Sturmian words and prove that this sequence characterizes every finite or infinite Sturmian word up to isomorphisms of the alphabet. We then characterize the combinatorial structure of the sequence of open and closed prefixes of standard Sturmian words. We prove that every standard Sturmian word, after swapping its first letter, can be written as an infinite product of squares of reversed standard words.

Keywords: Sturmian word; closed word; standard word; central word; semicentral word.

1 Introduction

In a recent paper with M. Bucci [1], the authors dealt with trapezoidal words, also with respect to the property of being *closed* (also known as periodic-like [2]) or open. Factors of Sturmian words are the most notable example of trapezoidal words, and in fact the last section of [1] showed the sequence of open and closed prefixes of the Fibonacci word, a famous characteristic Sturmian word.

In this paper we build upon such results, investigating the sequence of open and closed prefixes of Sturmian words in general, and in particular in the standard case. More precisely, we prove that the sequence $oc(w)$ of open and closed prefixes of a word w (i.e., the sequence whose n-th element is 1 if the prefix of length n of w is closed, or 0 if it is open) characterizes every (finite or infinite) Sturmian word, up to isomorphisms of the alphabet.

In [1], we investigated the structure of the sequence $oc(F)$ of the Fibonacci word. We proved that the lengths of the runs (maximal subsequences of consecutive equal elements) in $oc(F)$ form the doubled Fibonacci sequence. We prove in this paper that this doubling property holds for every standard Sturmian word, and describe the sequence $oc(w)$ of a standard Sturmian word w in terms of the *semicentral* prefixes of w, which are the prefixes of the form u_nxyu_n, where x, y are letters and u_nxy is an element of the standard sequence of w. As a consequence, we show that the word $ba^{-1}w$, obtained from a standard Sturmian word

w starting with letter a by swapping its first letter, can be written as the infinite product of the words $(u_n^{-1}u_{n+1})^2$, $n \geq 0$. Since the words $u_n^{-1}u_{n+1}$ are reversals of standard words, this induces an infinite factorization of $ba^{-1}w$ in squares of reversed standard words.

Finally, we show how the sequence of open and closed prefixes of a standard Sturmian word of slope α is related to the continued fraction expansion of α.

2   Open and Closed Words

Let us begin with some notation and basic definitions; for those not included below, we refer the reader to [1] and [3].

Let $\Sigma = \{a, b\}$ be a 2-letter alphabet. Let Σ^* and $\widehat{\Sigma}^*$ stand respectively for the free monoid and the free group generated by Σ. Their elements are called *words* over Σ. The *length* of a word w is denoted by $|w|$. The *empty word*, denoted by ε, is the unique word of length zero and is the neutral element of Σ^* and $\widehat{\Sigma}^*$.

A *prefix* (resp. a *suffix*) of a word w is any word u such that $w = uz$ (resp. $w = zu$) for some word z. A *factor* of w is a prefix of a suffix (or, equivalently, a suffix of a prefix) of w. An *occurrence* of a factor u in w is a factorization $w = vuz$. An occurrence of u is *internal* if both v and z are non-empty. The set of prefixes, suffixes and factors of the word w are denoted by $Pref(w)$, $Suff(w)$ and $Fact(w)$, respectively. From the definitions, we have that ε is a prefix, a suffix and a factor of any word. A *border* of a word w is any word in $Pref(w) \cap Suff(w)$ different from w.

A factor v of a word w is *left special in w* (resp. *right special in w*) if av and bv are factors of w (resp. va and vb are factors of w). A *bispecial factor* of w is a factor that is both left and right special.

The word $\widetilde{w}$ obtained by reading w from right to left is called the *reversal* (or *mirror image*) of w. A *palindrome* is a word w such that $\widetilde{w} = w$. In particular, the empty word is a palindrome.

We recall the definitions of open and closed word given in [4]:

Definition 1. *A word w is* closed *if and only if it is empty or has a factor $v \neq w$ occurring exactly twice in w, as a prefix and as a suffix of w (with no internal occurrences). A word that is not closed is called* open.

The word aba is closed, since its factor a appears only as a prefix and as a suffix. The word $abaa$, on the contrary, is not closed. Note that for any letter $a \in \Sigma$ and for any $n > 0$, the word a^n is closed, a^{n-1} being a factor occurring only as a prefix and as a suffix in it (this includes the special case of single letters, for which $n = 1$ and $a^{n-1} = \varepsilon$).

More generally, any word that is a power of a shorter word is closed. Indeed, suppose that $w = v^n$ for a non-empty v and $n > 1$. Without loss of generality, we can suppose that v is not a power itself. If v^{n-1} has an internal occurrence in w, then there exists a proper prefix u of v such that $uv = vu$, and it is a basic result in Combinatorics on Words that two words commute if and only if they are powers of a same shorter word, in contradiction with our hypothesis on v.

Remark 2. The notion of closed word is equivalent to that of *periodic-like* word [2]. A word w is periodic-like if its longest repeated prefix is not right special.

The notion of closed word is also closely related to the concept of *complete return* to a factor, as considered in [5]. A complete return to the factor u in a word w is any factor of w having exactly two occurrences of u, one as a prefix and one as a suffix. Hence, w is closed if and only if it is a complete return to one of its factors; such a factor is clearly both the longest repeated prefix and the longest repeated suffix of w (i.e., the longest border of w).

Remark 3. Let w be a non-empty word over Σ. The following characterizations of closed words follow easily from the definition:

1. the longest repeated prefix (resp. suffix) of w does not have internal occurrences in w, i.e., occurs in w only as a prefix and as a suffix;
2. the longest repeated prefix (resp. suffix) of w is not a right (resp. left) special factor of w;
3. w has a border that does not have internal occurrences in w;
4. the longest border of w does not have internal occurrences in w.

Obviously, the negations of the previous properties characterizate open words. In the rest of the paper we will use these characterizations freely and without explicit mention to this remark.

We conclude this section with two lemmas on right extensions.

Lemma 4. *Let w be a non-empty word over Σ. Then there exists at most one letter $x \in \Sigma$ such that wx is closed.*

Proof. Suppose by contradiction that there exist $a, b \in \Sigma$ such that both wa and wb are closed. Let va and $v'b$ be the longest borders of wa and wb, respectively. Since va and $v'b$ are prefixes of w, one has that one is a prefix of the other. Suppose that va is shorter than $v'b$. But then va has an internal occurrence in wa (that appearing as a prefix of the suffix v') against the hypothesis that wa is closed.   □

When w is closed, then exactly one such extension is closed. More precisely, we have the following (see also [2, Prop. 4]):

Lemma 5. *Let w be a closed word. Then wx, $x \in \Sigma$, is closed if and only if wx has the same period of w.*

Proof. Let w be a closed word and v its longest border; in particular, v is the longest repeated prefix of w. Let x be the letter following the occurrence of v as a prefix of w. Clearly, wx is has the same period as w, and it is closed as its border vx cannot have internal occurrences. Conversely, if $y \neq x$ is a letter, then wy has a different period and it is open as its longest repeated prefix v is right special.   □

For more details on open and closed words and related results see [1,2,4,6,7].

3   Open and Closed Prefixes of Sturmian Words

Let Σ^ω be the set of (right) infinite words over Σ, indexed by $\mathbb{N}_0$. An element of Σ^ω is a *Sturmian word* if it contains exactly $n + 1$ distinct factors of length n, for every $n \geq 0$. A famous example of Sturmian word is the Fibonacci word

$$F = abaababaabaabababaababa \cdots$$

If w is a Sturmian word, then aw or bw is also a Sturmian word. A Sturmian word w is *standard* (or *characteristic*) if aw and bw are both Sturmian words. The Fibonacci word is an example of standard Sturmian word. In the next section, we will deal specifically with standard Sturmian words. Here, we focus on finite factors of Sturmian words, called *finite Sturmian words*. Actually, finite Sturmian words are precisely the elements of Σ^* verifying the following balance property: for any $u, v \in Fact(w)$ such that $|u| = |v|$ one has $||u|_a - |v|_a| \leq 1$ (or, equivalently, $||u|_b - |v|_b| \leq 1$).

We let St denote the set of finite Sturmian words. The language St is factorial (i.e., if $w = uv \in St$, then $u, v \in St$) and extendible (i.e., for every $w \in St$ there exist letters $x, y \in \Sigma$ such that $xwy \in St$).

We recall the following definitions given in [8].

Definition 6. *A word $w \in \Sigma^*$ is a left special (resp. right special) Sturmian word if $aw, bw \in St$ (resp. if $wa, wb \in St$). A bispecial Sturmian word is a Sturmian word that is both left special and right special.*

For example, the word $w = ab$ is a bispecial Sturmian word, since aw, bw, wa and wb are all Sturmian. This example also shows that a bispecial Sturmian word is not necessarily a bispecial factor of some Sturmian word (see [9] for more details on bispecial Sturmian words).

Remark 7. It is known that if w is a left special Sturmian word, then w is a prefix of a standard Sturmian word, and the left special factors of w are prefixes of w. Symmetrically, if w is a right special Sturmian word, then the right special factors of w are suffixes of w.

We now define the sequence of open and closed prefixes of a word.

Definition 8. *Let w be a finite or infinite word over Σ. We define the sequence $oc(w)$ as the sequence whose n-th element is 1 if the prefix of length n of w is closed, or 0 otherwise.*

For example, if $w = abaaab$, then $oc(w) = 101001$.

In this section, we prove the following:

Theorem 9. *Every (finite or infinite) Sturmian word w is uniquely determined, up to isomorphisms of the alphabet Σ, by its sequence of open and closed prefixes $oc(w)$.*

We need some intermediate lemmas.

Lemma 10. *Let w be a right special Sturmian word and u its longest repeated prefix. Then u is a suffix of w.*

Proof. If w is closed, the claim follows from the definition of closed word. If w is open, then u is right special in w, and by Remark 7, u is a suffix of w.   □

Lemma 11. *Let w be a right special Sturmian word. Then wa or wb is closed.*

Proof. Let u be the longest repeated prefix of w and x the letter following the occurrence of u as a prefix of w. By Lemma 10, u is a suffix of w. Clearly, the longest repeated prefix of wx is ux, which is also a suffix of wx and cannot have internal occurrences in wx otherwise the longest repeated prefix of w would not be u. Therefore, wx is closed.   □

So, by Lemmas 4 and 11, if w is a right special Sturmian word, then one of wa and wb is closed and the other is open. This implies that the sequence of open and closed prefixes of a (finite or infinite) Sturmian word characterizes it up to exchange of letters. The proof of Theorem 9 is therefore complete.

4   Standard Sturmian Words

In this section, we deal with the sequence of open and closed prefixes of standard Sturmian words. In [1] a characterization of the sequence $oc(F)$ of open and closed prefixes of the Fibonacci word F was given.

Let us begin by recalling some definitions and basic results about standard Sturmian words. For more details, the reader can see [10] or [3].

Let α be an irrational number such that $0 < \alpha < 1$, and let $[0; d_0 + 1, d_1, \ldots]$ be the continued fraction expansion of α. The sequence of words defined by $s_{-1} = b$, $s_0 = a$ and $s_{n+1} = s_n^{d_n} s_{n-1}$ for $n \geq 0$, converges to the infinite word w_α, called the *standard Sturmian word of slope* α. The sequence of words s_n is called the *standard sequence* of w_α.

Note that w_α starts with letter b if and only if $\alpha > 1/2$, i.e., if and only if $d_0 = 0$. In this case, $[0; d_1 + 1, d_2, \ldots]$ is the continued fraction expansion of $1 - \alpha$, and $w_{1-\alpha}$ is the word obtained from w_α by exchanging a's and b's. Hence, without loss of generality, we will suppose in the rest of the paper that w starts with letter a, i.e., that $d_0 > 0$.

For every $n \geq -1$, one has

$$s_n = u_n xy, \tag{1}$$

for x, y letters such that $xy = ab$ if n is odd or ba if n is even. Indeed, the sequence $(u_n)_{n \geq -1}$ can be defined by: $u_{-1} = a^{-1}$, $u_0 = b^{-1}$, and, for every $n \geq 1$,

$$u_{n+1} = (u_n xy)^{d_n} u_{n-1}, \tag{2}$$

where x, y are as in (1).

Example 12. The Fibonacci word F is the standard Sturmian word of slope $(3-\sqrt{5})/2$, whose continued fraction expansion is $[0; 2, 1, 1, 1, \ldots]$, so that $d_n = 1$ for every $n \geq 0$. Therefore, the standard sequence of the Fibonacci word F is the sequence defined by: $f_{-1} = b$, $f_0 = a$, $f_{n+1} = f_n f_{n-1}$ for $n \geq 0$. This sequence is also called the sequence of *Fibonacci finite words*.

Definition 13. *A standard word is a finite word belonging to some standard sequence. A central word is a word $u \in \Sigma^*$ such that uxy is a standard word, for letters $x, y \in \Sigma$.*

It is known that every central word is a palindrome. Actually, central words play a central role in the combinatorics of Sturmian words and have several combinatorial characterizations (see [10] for a survey). For example, a word over Σ is central if and only if it is a palindromic bispecial Sturmian word.

Remark 14. Let $(s_n)_{n \geq -1}$ be a standard sequence. It follows by the definition that for every $k \geq 0$ and $n \geq -1$, the word $s_{n+1}^k s_n$ is a standard word. In particular, for every $n \geq -1$, the word $s_{n+1} s_n = u_{n+1} yx u_n xy$ is a standard word. Therefore, for every $n \geq -1$, we have that

$$u_n xy u_{n+1} = u_{n+1} yx u_n \tag{3}$$

is a central word.

The following lemma is a well known result (cf. [11]).

Lemma 15. *Let w be a standard Sturmian word and $(s_n)_{n \geq -1}$ its standard sequence. Then:*

1. *A standard word v is a prefix of w if and only if $v = s_n^k s_{n-1}$, for some $n \geq 0$ and $k \leq d_n$.*
2. *A central word u is a prefix of w if and only if $u = (u_n xy)^k u_{n-1}$, for some $n \geq 0$, $0 < k \leq d_n$, and distinct letters $x, y \in \Sigma$ such that $xy = ab$ if n is odd or ba if n is even.*

Note that $(u_n xy)^{d_n + 1} u_{n-1}$ is a central prefix of w, but this does not contradict the previous lemma since, by (2), $(u_n xy)^{d_n + 1} u_{n-1} = u_{n+1} yx u_n$.

Recall that a *semicentral word* (see [1]) is a word in which the longest repeated prefix, the longest repeated suffix, the longest left special factor and the longest right special factor all coincide. It is known that a word v is semicentral if and only if $v = uxyu$ for a central word u and distinct letters $x, y \in \Sigma$. Moreover, xuy is a factor of $uxyu$ and thus semicentral words are open, while central words are closed.

Proposition 16. *The semicentral prefixes of w are precisely the words of the form $u_n xy u_n$, $n \geq 1$, where x, y and u_n are as in (1).*

Proof. Since u_n is a central word, the word $u_n xy u_n$ is a semicentral word by definition, and it is a prefix of $u_n xy u_{n+1} = u_{n+1} yx u_n$, which in turn is a prefix of w by Lemma 15.

Conversely, assume that w has a prefix of the form $u\xi\eta u$ for a central word u and distinct letters $\xi, \eta \in \Sigma$. From Lemma 15 and (1), we have that

$$u\xi\eta u = (u_n xy)^k u_{n-1} \cdot \xi\eta \cdot (u_n xy)^k u_{n-1},$$

for some $n \geq 1$, $k \leq d_n$, and distinct letters $x, y \in \Sigma$ such that $xy = ab$ if n is odd or ba if n is even. In particular, this implies that $\xi\eta = yx$.

If $k = d_n$, then $u = u_{n+1} yx u_{n+1}$, and we are done. So, suppose by contradiction that $k < d_n$. Now, on the one hand we have that $(u_n xy)^{k+1} u_{n-1} yx$ is a prefix of w by Lemma 15, and so $(u_n xy)^{k+1} u_{n-1}$ is followed by yx as a prefix of w; on the other hand we have

$$\begin{aligned}
u\xi\eta u &= (u_n xy)^k u_{n-1} \cdot yx \cdot (u_n xy)^k u_{n-1} \\
&= (u_n xy)^k \cdot u_{n-1} yx u_n xy \cdot (u_n xy)^{k-1} u_{n-1} \\
&= (u_n xy)^k \cdot u_n xy u_{n-1} xy \cdot (u_n xy)^{k-1} u_{n-1} \\
&= (u_n xy)^{k+1} \cdot u_{n-1} xy \cdot (u_n xy)^{k-1} u_{n-1},
\end{aligned}$$

so that $(u_n xy)^{k+1} u_{n-1}$ is followed by xy as a prefix of w, a contradiction. $\qquad\square$

The next theorem shows the behavior of the runs of open and closed prefixes in w by determining the structure of the last elements of the runs.

Theorem 17. *Let vx, $x \in \Sigma$, be a prefix of w. Then:*

1. *v is open and vx is closed if and only if there exists $n \geq 1$ such that $v = u_n xy u_n$;*
2. *v is closed and vx is open if and only if there exists $n \geq 0$ such that $v = u_n xy u_{n+1} = u_{n+1} yx u_n$.*

Proof. 1. If $v = u_n xy u_{n+1} = u_{n+1} yx u_n$, then v is semicentral and therefore open. The word vx is closed since its longest repeated prefix $u_n x$ occurs only as a prefix and as a suffix in it.

Conversely, let vx be a closed prefix of w such that v is open, and let ux be the longest repeated suffix of vx. Since vx is closed, ux does not have internal occurrences in vx. Since u is the longest repeated prefix of v (suppose the longest repeated prefix of v is a z longer than u, then vx, which is a prefix of z, would be repeated in v and hence in vx, contradiction) and v is open, u must have an internal occurrence in v followed by a letter $y \neq x$. Symmetrically, if ξ is the letter preceding the occurrence of u as a suffix of v, since u is the longest repeated suffix of v one has that u has an internal occurrence in v preceded by a letter $\eta \neq \xi$. Thus u is left and right special in w. Moreover, u is the longest special factor in v. Indeed, if u' is a left special factor of v, then u must be a prefix of u'. But ux cannot appear in v since vx is closed, and if uy was a left special factor of v, it would be a prefix of v. Symmetrically, u is the longest right

special factor in v. Thus v is semicentral, and the claim follows from Proposition 16.

2. If $v = u_n x y u_{n+1} = u_{n+1} y x u_n$, then v is a central word and therefore it is closed. Its longest repeated prefix is u_{n+1}. The longest repeated prefix of vx is either a^{d_0-1} (if $n = 0$) or $u_n x$ (if $n > 0$); in both cases, it has an internal occurrence as a prefix of the suffix $u_{n+1} x$. Therefore, vx is open.

Conversely, suppose that vx is any open prefix of w such that v is closed. If $vx = a^{d_0} b$, then $v = u_0 x y u_1 = u_1 y x u_0$ and we are done. Otherwise, by 1), there exists $n \geq 1$ such that $|u_n \xi y u_n| < |v| < |u_{n+1} y \xi u_{n+1}|$, where $\{\xi, y\} = \{a, b\}$. We know that $u_n \xi y u_{n+1}$ is closed and $u_n \xi y u_{n+1} \xi$ is open; it follows $v = u_n \xi y u_{n+1} = u_n x y u_{n+1}$, as otherwise there should be in w a semicentral prefix strictly between $u_n x y u_n$ and $u_{n+1} y x u_{n+1}$. $\qquad\square$

Note that, for every $n \geq 1$, one has:

$$\begin{aligned} u_{n+1} y x u_{n+1} &= u_{n+1} y x u_n (u_n^{-1} u_{n+1}) \\ &= u_n x y u_{n+1} (u_n^{-1} u_{n+1}) \\ &= u_n x y u_n (u_n^{-1} u_{n+1})^2. \end{aligned}$$

Therefore, starting from an (open) semi-central prefix $u_n x y u_n$, one has a run of closed prefixes, up to the prefix $u_n x y u_{n+1} = u_{n+1} y x u_n = u_n x y u_n (u_n^{-1} u_{n+1})$, followed by a run of the same length of open prefixes, up to the prefix $u_{n+1} y x u_{n+1} = u_{n+1} y x u_n (u_n^{-1} u_{n+1}) = u_n x y u_n (u_n^{-1} u_{n+1})^2$. See Table 1 for an illustration.

prefix of w	open/closed	example
$u_n x y u_n$	open	$aaba$
$u_n x y u_n x$	closed	$aabaa$
$u_n x y u_n x y$	closed	$aabaab$
...	...	...
$u_n x y u_{n+1} = u_{n+1} y x u_n$	closed	$aabaabaa$
$u_{n+1} y x u_n y$	open	$aabaabaaa$
$u_{n+1} y x u_n y x$	open	$aabaabaaab$
...	...	...
$u_{n+1} y x u_{n+1}$	open	$aabaabaaabaa$
$u_{n+1} y x u_{n+1} y$	closed	$aabaabaaabaab$

Table 1. The structure of the prefixes of a standard Sturmian word $w = aabaabaaabaabaa\cdots$ with respect to the u_n prefixes. Here $d_0 = d_1 = 2$ and $d_2 = 1$.

In Table 2, we show the first elements of the sequence $oc(w)$ for a standard Sturmian word $w = aabaabaaabaabaa\cdots$ of slope $\alpha = [0; 3, 2, 1, \ldots]$, i.e., with $d_0 = d_1 = 2$ and $d_2 = 1$. One can notice that the runs of closed prefixes are followed by runs of the same length of open prefixes.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
w	a	a	b	a	a	b	a	a	a	b	a	a	b	a	a
$oc(w)$	1	1	0	0	1	1	1	1	0	0	0	0	1	1	1

Table 2. The sequence $oc(w)$ of open and closed prefixes for the word $w = aabaabaaabaabaa\cdots$

The words $u_n^{-1}u_{n+1}$ are reversals of standard words, for every $n \geq 1$. Indeed, let $r_n = \widetilde{s_n}$ for every $n \geq -1$, so that $r_{-1} = b$, $r_0 = a$, and $r_{n+1} = r_{n-1}r_n^{d_n}$ for $n \geq 0$. Since by (1) $s_n = u_n xy$ and $s_{n+1} = u_{n+1}yx$, one has $r_n = yxu_n$ and $r_{n+1} = xyu_{n+1}$, and therefore, by (3),

$$u_n r_{n+1} = u_{n+1} r_n. \tag{4}$$

Multiplying (4) on the left by u_n^{-1} and on the right by r_n^{-1}, one obtains

$$r_{n+1}r_n^{-1} = u_n^{-1}u_{n+1}. \tag{5}$$

Since $r_{n+1} = r_{n-1}r_n^{d_n}$, one has that $r_{n+1}r_n^{-1} = r_{n-1}r_n^{d_n-1}$, and therefore $r_{n+1}r_n^{-1}$ is the reversal of a standard word. By (5), $u_n^{-1}u_{n+1}$ is the reversal of a standard word.

Now, note that for $n = 0$, one has $u_0 xy u_1 = u_1 yx u_0 = a^{d_0}$ and $(u_0^{-1}u_1) = ba^{d_0-1}$. Thus, we have the following:

Theorem 18. *Let w be the standard Sturmian word of slope α, with $0 < \alpha < 1/2$, and let $[0; d_0 + 1, d_1, \ldots]$, with $d_0 > 0$, be the continued fraction expansion of α. The word $ba^{-1}w$ obtained from w by swapping the first letter can be written as an infinite product of squares of reversed standard words in the following way:*

$$ba^{-1}w = \prod_{n \geq 0} (u_n^{-1}u_{n+1})^2,$$

where $(u_n)_{n \geq -1}$ is the sequence defined in (1).

In other words, one can write

$$w = a^{d_0}ba^{d_0-1}\prod_{n \geq 1}(u_n^{-1}u_{n+1})^2.$$

Example 19. Take the Fibonacci word. Then, $u_1 = \varepsilon$, $u_2 = a$, $u_3 = aba$, $u_4 = abaaba$, $u_5 = abaababaaba$, etc. So, $u_1^{-1}u_2 = a$, $u_2^{-1}u_3 = ba$, $u_3^{-1}u_4 = aba$, $u_4^{-1}u_5 = baaba$, etc. Indeed, $u_n^{-1}u_{n+1}$ is the reversal of the Fibonacci finite word

f_{n-1}. By Theorem 18, we have:

$$F = ab \prod_{n \geq 1} (u_n^{-1} u_{n+1})^2$$

$$= ab \prod_{n \geq 0} (\widetilde{f_n})^2$$

$$= ab \cdot (a \cdot a)(ba \cdot ba)(aba \cdot aba)(baaba \cdot baaba) \cdots$$

i.e., F can be obtained by concatenating ab and the squares of the reversals of the Fibonacci finite words f_n starting from $n = 0$.

Note that F can also be obtained by concatenating the reversals of the Fibonacci finite words f_n starting from $n = 0$:

$$F = \prod_{n \geq 0} \widetilde{f_n}$$

$$= a \cdot ba \cdot aba \cdot baaba \cdot ababaaba \cdots$$

and also by concatenating ab and the Fibonacci finite words f_n starting from $n = 0$:

$$F = ab \prod_{n \geq 0} f_n$$

$$= ab \cdot a \cdot ab \cdot aba \cdot abaab \cdot abaababa \cdots$$

One can also characterize the sequence of open and closed prefixes of a standard Sturmian word w in terms of the directive sequence of w.

Recall that the *continuants* of an integer sequence $(a_n)_{n \geq 0}$ are defined as $K[\ ] = 1$, $K[a_0] = a_0$, and, for every $n \geq 1$,

$$K[a_0, \ldots, a_n] = a_n K[a_0, \ldots, a_{n-1}] + K[a_0, \ldots, a_{n-2}].$$

Continuants are related to continued fractions, as the n-th convergent of $[a_0; a_1, a_2, \ldots]$ is equal to $K[a_0, \ldots, a_n] / K[a_1, \ldots, a_n]$.

Let w be a standard Sturmian word and $(s_n)_{n \geq -1}$ its standard sequence. Since $|s_{-1}| = |s_0| = 1$ and, for every $n \geq 1$, $|s_{n+1}| = d_n|s_n| + |s_{n-1}|$, then one has, by definition, that for every $n \geq 0$

$$|s_n| = K[1, d_0, \ldots, d_{n-1}].$$

For more details on the relationships between continuants and Sturmian words see [12].

By Theorems 17 and 18, all prefixes up to a^{d_0} are closed; then all prefixes from $a^{d_0}b$ till $a^{d_0}ba^{d_0-1}$ are open, then closed up to $a^{d_0}ba^{d_0-1} \cdot u_1^{-1}u_2$, open again up to $a^{d_0}ba_{d_0-1} \cdot (u_1^{-1}u_2)^2$, and so on. Thus, the lengths of the successive runs of closed and open prefixes are: d_0, d_0, $|u_2| - |u_1|$, $|u_2| - |u_1|$, $|u_3| - |u_2|$, $|u_3| - |u_2|$, etc. Since $d_0 = K[1, d_0 - 1]$ and, for every $n \geq 1$,

$$|u_{n+1}| - |u_n| = |s_{n+1}| - |s_n| = (d_n - 1)|s_n| + |s_{n-1}|$$
$$= K[1, d_0, \ldots, d_{n-1}, d_n - 1],$$

we have the following:

Corollary 20. *Let w and α be as in the previous theorem and let, for every $n \geq 0$, $k_n = K[1, d_0, \ldots, d_{n-1}, d_n - 1]$. Then*

$$oc(w) = \prod_{n \geq 0} 1^{k_n} 0^{k_n}.$$

Acknowledgments

References

1. Bucci, M., De Luca, A., Fici, G.: Enumeration and Structure of Trapezoidal Words. Theoretical Computer Science **468** (2013) 12–22
2. Carpi, A., de Luca, A.: Periodic-like words, periodicity and boxes. Acta Informatica **37** (2001) 597–618
3. Lothaire, M.: Algebraic Combinatorics on Words. Encyclopedia of Mathematics and its Applications. Cambridge Univ. Press, New York, NY, USA (2002)
4. Fici, G.: A Classification of Trapezoidal Words. In: WORDS 2011, 8th International Conference on Words. Volume 63 of Electronic Proceedings in Theoretical Computer Science. (2011) 129–137
5. Glen, A., Justin, J., Widmer, S., Zamboni, L.Q.: Palindromic richness. European J. Combin. **30** (2009) 510–531
6. Bucci, M., de Luca, A., De Luca, A.: Rich and Periodic-Like Words. In: DLT 2009, 13th International Conference on Developments in Language Theory. Volume 5583 of Lecture Notes in Comput. Sci. Springer (2009) 145–155
7. Fici, G., Lipták, Zs.: Words with the Smallest Number of Closed Factors. In: 14th Mons Days of Theoretical Computer Science. (2012)
8. de Luca, A., Mignosi, F.: Some combinatorial properties of Sturmian words. Theoret. Comput. Sci. **136** (1994) 361–385
9. Fici, G.: A Characterization of Bispecial Sturmian Words. In: MFCS 2012, 37th International Symposium on Mathematical Foundations of Computer Science. Volume 7464 of Lecture Notes in Comput. Sci., Springer Berlin Heidelberg (2012) 383–394
10. Berstel, J.: Sturmian and episturmian words. In: CAI 2007, Second International Conference on Algebraic Informatics. Volume 4728 of Lecture Notes in Computer Science., Springer (2007) 23–47
11. Fischler, S.: Palindromic prefixes and episturmian words. J. Combin. Theory Ser. A **113** (2006) 1281–1304
12. de Luca, A.: Some extremal properties of the Fibonacci word. Internat. J. Algebra Comput. (to appear)