Palindromes in Sturmian Words

Aldo de Luca[1,2] and Alessandro De Luca[1,3]

[1] Dipartimento di Matematica e Applicazioni "R. Caccioppoli"
Università degli Studi di Napoli Federico II
via Cintia, Monte S. Angelo
I-80126 Napoli, Italy
[2] aldo.deluca@unina.it
[3] aledeluc@studenti.unina.it

Abstract. We study some structural and combinatorial properties of Sturmian palindromes, i.e., palindromic finite factors of Sturmian words. In particular, we give a formula which permits to compute in an exact way the number of Sturmian palindromes of any length. Moreover, an interesting characterization of Sturmian palindromes is obtained.

1 Introduction

Sturmian words have been widely studied for their theoretical importance and their applications to various fields of science. By definition, they are infinite words which are not eventually periodic and have minimal subword complexity. They also enjoy some remarkable characterizations of geometrical nature (*cutting sequences*, *mechanical words*). The reader is referred to [2] and to [1] for general surveys on this subject.

In recent years, some works have investigated Sturmian words by looking at their palindromic factors. A *palindrome* is a finite word which can be read indistinctively from left to right or from right to left.

Palindromes play an essential role in the structure of Sturmian words. In fact, an important theorem of X. Droubay and G. Pirillo [9] shows that an infinite word is Sturmian if and only if it has exactly one palindromic factor of length n for n even, and two for n odd. Moreover, A. de Luca and F. Mignosi [8] proved that the set of palindromic prefixes of all standard Sturmian words is equal to the set PER of *central words*, i.e., words having two periods p and q which are coprime, and length $p+q-2$. Central words are words over a two letter alphabet $\{a, b\}$ and satisfy remarkable structural properties. In particular, a central word w is such that wab and wba can be factorized as a product of two palindromes. Moreover, the set St of factors of all Sturmian words is equal to the set of factors of PER.

There exist, and even they are the majority, Sturmian palindromes which are not central. For instance, there are 14 Sturmian palindromes of length 7, whereas the number of central words of the same length is only 6 (see Table 2).

In this paper we are interested in the combinatorics and the structure of the language of Sturmian palindromes. A main result, proved in Section 4.1, is a simple formula which gives the number of Sturmian palindromes of any length. Moreover, some structural properties of Sturmian palindromes are proved in Section 4.2. A remarkable characterization of Sturmian palindromes is given. In particular, one obtains a new characterization of central words.

2 Preliminaries

Let $\mathcal{A}$ be a 2-letter alphabet $\{a, b\}$. As usual, we denote by $\mathcal{A}^*$ the *free monoid* generated by $\mathcal{A}$, that is the set of all words over $\mathcal{A}$ with the operation of concatenation. The identity element of $\mathcal{A}^*$ is the *empty word* ε. Let $w = a_1 a_2 \cdots a_n$ be a word, with $a_i \in \mathcal{A}$, $1 \leq i \leq n$. The integer n is called the *length* of w, and it is denoted by $|w|$. Conventionally, the length of the empty word is 0.

A word u is a *factor* of $w \in \mathcal{A}^*$ if $w = xuy$ for some words x, y. In the special case $x = \varepsilon$ (resp., $y = \varepsilon$), we call u a *prefix* (resp., *suffix*) of w. A factor u of w is *proper* if $u \neq w$. A factor u of w is *median* if $w = xuy$ with $|x| = |y|$. The set of factors of a word x is denoted by $\mathrm{Fact}\,(x)$. For any $X \subseteq \mathcal{A}^*$, one sets:

$$\mathrm{Fact}\,(X) = \bigcup_{x \in X} \mathrm{Fact}\,(x) \ .$$

Let $w = a_1 a_2 \cdots a_n$ be a word, with $a_i \in \mathcal{A}$, $1 \leq i \leq n$. A positive integer p is a *period* of w if for all i, j, $1 \leq i, j \leq n$, the following condition is satisfied: if $i \equiv j \pmod{p}$, then $a_i = a_j$. The minimal period of w will be denoted by π_w; it is natural to set $\pi_\varepsilon = 1$. For any $w \neq \varepsilon$, the unique positive integer k such that $w = z^k z'$, where $|z| = \pi_w$ and $|z'| < \pi_w$, will be called the *order* of w.

Given $w = a_1 a_2 \cdots a_n$ with all $a_i \in \mathcal{A}$, the *reversal* $w^\sim$ is the word $a_n \cdots a_1$. If $w = \varepsilon$, one sets $\varepsilon^\sim = \varepsilon$. A word $w \in \mathcal{A}^*$ is a *palindrome* if $w = w^\sim$. The set of all palindromes of $\mathcal{A}^*$ is denoted by PAL.

An *infinite word* x is just an infinite sequence of letters:

$$x = c_1 c_2 c_3 \cdots \text{ where } c_i \in \mathcal{A}, \text{ for all } i \geq 1 \ .$$

The product between a finite word and an infinite one is well defined. A (finite) factor of x is either the empty word or any sequence $u = c_i \cdots c_j$ with $i \leq j$, i.e., a finite block of consecutive letters in x. If $i = 1$, then u is a prefix of x. A *suffix* of x is an infinite word y such that $x = uy$ for some $u \in \mathcal{A}^*$. We shall denote by $\mathrm{Fact}\,(x)$ the set of finite factors of x.

An infinite word is *Sturmian* if for each $n \in \mathbb{N}$ it has $n+1$ factors of length n. An equivalent geometrical definition can be given in terms of *cutting sequences*. In fact, a Sturmian word can be defined by considering the sequence of cuts in a squared lattice ($\mathbb{N} \times \mathbb{N}$) made by a ray having a slope which is an irrational number α. A horizontal cut is denoted by the letter b, a vertical by a, and a cut with a corner by ab or ba. A Sturmian word represented by a ray starting from the origin is usually called *standard* or the *characteristic word* associated with the irrational α and it is often denoted by c_α.

Let x be a finite or infinite word over $\mathcal{A}$. A factor u of x is a *right special* factor of x if ua and ub are factors of x. As is well known [2], an infinite word x

is Sturmian if and only if for any $n \geq 0$, there is only one right special factor of x of length n.

3 Sturmian Palindromes

Let St be the set of finite Sturmian words, i.e., factors of infinite Sturmian words over the alphabet $\mathcal{A} = \{a, b\}$. We recall that for any Sturmian word there exists a standard Sturmian word having the same set of factors (cf. [2]). In the sequel we shall be interested in the set $St \cap PAL$, whose elements will be called *Sturmian palindromes*.

As usual, we denote by PER the set of central words, that is, words having two coprime periods p and q and length $p + q - 2$. Conventionally, the empty word ε is central (in this case, $p = q = 1$). It is well known (see [8, 2]) that the set PER coincides with the set of palindromic prefixes of all standard Sturmian words, so that $PER \subseteq St \cap PAL$. However, the previous inclusion is strict since there exist non-central Sturmian palindromes, for instance *abba*. The set PER is particularly important because a finite word is Sturmian if and only if it is a factor of a central word, i.e., $St = \mathrm{Fact}\,(PER)$. We shall prove (Corollary 2) a similar characterization for Sturmian palindromes.

Theorem 1. *Every palindromic factor of a standard Sturmian word c_α is a median factor of a palindromic prefix of c_α.*

The result is attributed to A. de Luca [7] by J.-P. Borel and C. Reutenauer, who gave a geometrical proof in [3]. We shall see later a direct proof which does not use geometrical arguments.

Corollary 2. *A word is a Sturmian palindrome if and only if it is a median factor of some central word.*

Proof. Trivially, every median factor of a palindrome is itself a palindrome. Since $St = \mathrm{Fact}\,(PER)$, it follows that a median factor of an element of PER is a Sturmian palindrome.

Conversely, let u be in $St \cap PAL$. By definition, there exists an infinite (standard) Sturmian word s such that $u \in \mathrm{Fact}\,(s)$. By Theorem 1, u is a median factor of a palindromic prefix of s. Since palindromic prefixes of standard Sturmian words are exactly the elements of PER, the result follows. □

Our proof of Theorem 1, which follows a simple argument suggested by A. Carpi [4], is based on the following results (see [7]):

Proposition 3. *If $w \in \mathrm{Fact}\,(x)$, where x is an infinite Sturmian word, then the reversal $w^\sim$ is a factor of x too. Moreover, if x is standard, then w is a right special factor of x if and only if $w^\sim$ is a prefix of x.*

Corollary 4. *A palindromic factor of an infinite standard Sturmian word x is a right special factor of x if and only if it is a palindromic prefix of x.*

Proof (of Theorem 1). By contradiction, let $c_\alpha = \lambda ux$, where u is a palindrome that is not a median factor of any palindromic prefix of c_α, and $\lambda \in \mathcal{A}^*$ has minimal length for such condition. Since u cannot be a prefix of c_α, we have $|\lambda| \geq 1$. Thus we can assume, without loss of generality, $\lambda = \lambda' a$. Now let z be the first letter of x, so that $x = zx'$. Suppose first $z = a$. The palindrome aua is not a median factor of a palindromic prefix of c_α, otherwise so would be u. But $c_\alpha = \lambda' auax'$ with $|\lambda'| < |\lambda|$, and this contradicts the minimality of $|\lambda|$. Therefore $z = b$, and then aub and $bua = (aub)^\sim$ are factors of c_α. This means in particular that u is a right special factor of c_α. Corollary 4 then implies that u is a prefix of c_α, a contradiction. □

4 Main Results

4.1 Enumeration of Sturmian Palindromes

In this section we shall give an explicit formula for the enumeration function of $St \cap PAL$. We start by recalling some basic facts (see [8, 7]):

Proposition 5. *Let w be a word. The following conditions are equivalent:*

1. *$w \in PER$,*
2. *awb and bwa are Sturmian,*
3. *awa, awb, bwa, and bwb are all Sturmian.*

Proposition 6. *If wa and wb are Sturmian words, then there exists a letter $x \in \mathcal{A}$ such that xwa and xwb are both Sturmian.*

We now prove two easy consequences (see also [7]):

Proposition 7. *Let $w \in \mathcal{A}^*$ be a palindrome. If wa and wb are Sturmian, then w is central.*

Proof. From the previous proposition, there exists a letter $x \in \mathcal{A}$ such that xwa and xwb are both Sturmian. Without loss of generality, we may suppose $x = a$, so that $awb \in St$. Therefore $(awb)^\sim = bwa$ is Sturmian too, thus by Proposition 5 w is central. □

Lemma 8. *Let w be a Sturmian palindrome. If w is not central, then there exists a unique letter $x \in \mathcal{A}$ such that xwx is Sturmian.*

Proof. If awa and bwb are both Sturmian, then w is central by Proposition 7, a contradiction. However, by Corollary 2, w is a (proper) median factor of some central word. □

Now let us introduce the function $g : \mathbb{N} \to \mathbb{N}$ defined for all $n \geq 0$ as

$$g(n) := \mathrm{card}(St \cap PAL \cap \mathcal{A}^n) \ .$$

For any $n \geq 0$, $g(n)$ counts the number of Sturmian palindromes of length n.

Theorem 9. *For any $n \geq 0$, the number $g(n)$ of Sturmian palindromes of length n is given by:*

$$1 + \sum_{i=0}^{\lceil n/2 \rceil - 1} \phi(n-2i) \tag{1}$$

where ϕ is Euler's totient function. Equivalently, for any $n \geq 0$

$$g(2n) = 1 + \sum_{i=1}^{n} \phi(2i) \quad \textit{and} \quad g(2n+1) = 1 + \sum_{i=0}^{n} \phi(2i+1) \ .$$

Proof. Given $w \in St \cap PAL$, at least one of its "extensions" awa and bwb is Sturmian. Indeed, according to Lemma 8, if $w \notin PER$, then exactly one of these extensions is in St. If $w \in PER$, then from Proposition 5, both awa and bwb are Sturmian palindromes. Since the number of central words of length n is $\phi(n+2)$ (see [8]), one gets:

$$g(n+2) = g(n) + \phi(n+2)$$

and this implies the desired formula, because $g(0) = 1$ and $g(1) = 2$. □

We define a function $f : \mathbb{N} \to \mathbb{N}$ by setting, for $n \geq 0$:

$$f(2n) = 1 + \frac{n(n+1)}{2} \text{ and } f(2n+1) = 2 + n(n+1) \ .$$

It is easy to verify that $g(n) \leq f(n)$ for all $n \geq 0$. Moreover, for any $n \geq 0$ we set $h(n) = \text{card}(PER \cap \mathcal{A}^n) = \phi(n+2)$. In Table 1 we list the values of the functions g, f, and h for $0 \leq n \leq 17$.

As an example, in Table 2 we report the set of all 14 Sturmian palindromes of length 7. The six central words in it are underlined.

Table 1. The functions g, f, and h

n	$g(n)$	$f(n)$	$h(n)$	n	$g(n)$	$f(n)$	$h(n)$
0	1	1	1	9	20	22	10
1	2	2	2	10	14	16	4
2	2	2	2	11	30	32	12
3	4	4	4	12	18	22	6
4	4	4	2	13	42	44	8
5	8	8	6	14	24	29	8
6	6	7	4	15	50	58	16
7	14	14	6	16	32	37	6
8	10	11	4	17	66	74	18

Table 2. Sturmian palindromes of length 7 (central words are underlined)

$\underline{aaaaaaa}$	$\underline{bbbbbbb}$
$\underline{aaabaaa}$	$\underline{bbbabbb}$
$aababaa$	$bbababb$
$abaaaba$	$babbbab$
$\underline{abababa}$	$\underline{bababab}$
$abbabba$	$baabaab$
$abbbbba$	$baaaaab$

4.2 Structural properties

We have seen in Section 3 that a Sturmian palindrome is a median factor of a central word. In this section we shall give some further results concerning the structure of Sturmian palindromes.

Proposition 10. *A palindrome $w \in \mathcal{A}^*$ with minimal period $\pi_w > 1$ can be uniquely represented as*

$$w = w_1 xy w_2 = w_2 yx \widetilde{w_1}$$

with $x, y \in \mathcal{A}$, w_2 the longest proper palindromic suffix of w, and $|w_1 xy| = \pi_w$. The word w is not central if and only if either $w_1 \notin PAL$ or $w = (w_1 xx)^k w_1$ where $k \geq 1$ is the order of w.

Proof. As is well known (cf. [7]), a palindrome w has a period $p < |w|$ if and only if it has a palindromic suffix (prefix) of length $|w| - p$. Since the minimal period π_w of a palindrome is less than $|w|$ and $\pi_w > 1$, it follows that w can be uniquely factorized as $w = w_1 xy w_2$ where w_2 is the longest proper palindromic suffix of w and $|w_1 xy| = \pi_w$. Since w is a palindrome, we can write

$$w = w_1 xy w_2 = w_2 yx \widetilde{w_1} \ .$$

From a classic result on central words (see [7]), when $\pi_w > 1$, w is central if and only if $w_1 \in PAL$ and $x \neq y$. Therefore, in the case $w_1 \in PAL$, w is not central if and only if $w = w_1 xx w_2 = w_2 xx w_1$. The word w has the two periods

$$\pi_w = |w_1 xx| \text{ and } q = |w_2 xx| \tag{2}$$

and length $\pi_w + q - 2$. Thus $w \notin PER$ if and only if $d = \gcd(\pi_w, q) > 1$. Since $|w| \geq \pi_w + q - d$, by Fine and Wilf's theorem (cf. [10]) w has the period $d \geq \pi_w$. This occurs if and only if $q = k\pi_w$ with $k \geq 1$. From (2) this condition is equivalent to the statement $w_2 xx = (w_1 xx)^k$, i.e., $w = (w_1 xx)^k w_1$. $\square$

Example 11. Let $w = aaabaaaaaabaaa \in St \cap PAL$, with $\pi_w = 7$. The word w can be factorized as $(aaaba)aa(aaabaaa)$, where $aaabaaa$ is the longest proper palindromic suffix of w, $|aaaba| = \pi_w - 2 = 5$. The prefix $aaaba$ is not a palindrome, thus w is not central.

Let $v = abaabababaaba \in St \cap PAL$. We factorize v as

$$v = (abaabab)ab(abaaba)$$

where $abaaba$ is the longest proper palindromic suffix of v. Also in this case $abaabab$ is not a palindrome, so that $w \notin PER$.

Let $u = abbabbabba \in St \cap PAL$. We factorize u as $(a)bb(abbabba)$, where $abbabba$ is the longest palindromic suffix of u. In this case, the prefix a is a palindrome, and $u = (abb)^3a$. Hence u is not central.

Lemma 12. *If $w = w_1xyw_2 = w_2yx\tilde{w_1}$, where w_2 is the longest proper palindromic suffix of w and $x, y \in \mathcal{A}$, then $w' = ywy$ has the minimal period $\pi_{w'} = \pi_w$.*

Proof. Since w is a factor of w', one has $\pi_{w'} \geq \pi_w$. The word yw_2y is a palindromic proper suffix of $w' = yw_1xyw_2y$, so that w' has the period $|yw_1x|$. Hence, $\pi_{w'} \leq |yw_1x| = |w_1xy| = \pi_w$. Thus $\pi_w = \pi_{w'}$. □

The next lemma is essentially a restatement of Lemma 2 in [5]. Note that its first part is an obvious consequence of Lemma 12.

Lemma 13. *Let $w = w_1xyw_2 = w_2yxw_1 \in PER$, with $|w_2| > |w_1|$, $\{x, y\} = \mathcal{A}$. The word $v = ywy$ has minimal period $\pi_v = \pi_w$, whereas $v' = xwx = xw_1xyw_2x$ has minimal period $\pi_{v'} = |w_2| + 2 = |w| - \pi_w + 2$.*

Let $w \in (St \cap PAL) \setminus PER$. We denote by u the (unique) shortest median extension of w in PER, and by v the longest central median factor of w. Note that also v is unique.

Theorem 14. *Let $w \in (St \cap PAL) \setminus PER$. With the preceding notation, one has $\pi_u = \pi_w$. Moreover, either $\pi_w = \pi_v$ or $\pi_w = |v| - \pi_v + 2$.*

Proof. We consider first the case that $\pi_v = 1$, so that $v = x^n$ with $x \in \mathcal{A}$. In such a case w has also the median palindromic factor $v_1 = yx^ny$, where $\{x, y\} = \mathcal{A}$ (recall that v is the longest central median factor of w). Moreover, $n = |v|$ is at least 2, otherwise v_1 would be equal to $yxy \in PER$. One has $\pi_{v_1} = |yx^n| = n + 1 = |v| - \pi_v + 2$. Now we define, for $2 \leq i \leq n$:

$$v_i = xv_{i-1}x = x^{i-1}yx^nyx^{i-1} = (x^{i-1}yx^{n-i+1})(x^{i-1}yx^{i-1}) \ . \tag{3}$$

The word $v_n = x^{n-1}yx^nyx^{n-1}$ is central, whereas for $i < n$ one has $v_i \notin PER$. From Lemma 8 it follows that the words v_i are the *only* Sturmian extensions of v_1 which are median factors of v_n. Since for $i < n$ one has $v_i \notin PER$, one derives that $w = v_k$ for some $1 \leq k < n$, and $u = v_n$. As shown in (3), by Lemma 12 all the v_i's have the same period, for $1 \leq i \leq n$. The result in this case follows: $\pi_w = \pi_u = |v| - \pi_v + 2$.

Now let us assume $\pi_v > 1$. In this case $v = w_1xyw_2 = w_2yxw_1$, with $w_1, w_2 \in PAL$, $x \neq y$. We suppose $|w_1| < |w_2|$, so that $\pi_v = |w_1| + 2$. From the definition of v, it follows that there exists a letter $z \in \mathcal{A}$ such that $v_1 = zvz$ is a median factor of w which is not central. By Lemma 13, we have $\pi_{v_1} = \pi_v$ if $z = y$, or else $\pi_{v_1} = |v| - \pi_v + 2$ if $z = x$.

Using Lemma 12, we shall now define a sequence of palindromes with the same minimal period as v_1. Let us first suppose that $z = y$, so that $v_1 = yw_1xyw_2y$. We set $v_2 = xv_1x = (xyw_1)(xyw_2yx)$. Moreover, if $w_1 = p_1p_2 \cdots p_k$ with $p_j \in \mathcal{A}$ for $1 \leq j \leq k$, we set $v_i = p_{k-i+3}v_{i-1}p_{k-i+3}$ for $i \geq 3$, so that

$$v_3 = p_k v_2 p_k = (p_k xyp_1 \cdots p_{k-1})(p_k xyw_2yxp_k) \ ,$$
$$\vdots$$
$$v_{k+2} = p_1 v_{k+1} p_1 = p_1 \cdots p_k xyw_1xyw_2yxp_k \cdots p_1 = w_1xyw_1xyw_2yxw_1^{\sim} \ .$$

Since $w_1 = w_1^{\sim}$, the last equation can be written as

$$v_{k+2} = (w_1)xy(w_1xyw_2yxw_1) = (w_1xyw_2yxw_1)yx(w_1)$$

showing, by Proposition 10, that the word v_{k+2} is central, so that for any $i \leq \pi_v = k+2$, $v_i \in St \cap PAL$. Let $s \leq k+2$ be the minimal integer such that $v_s \in PER$. Since for $i < s$ one has $v_i \notin PER$, and from Lemma 8, one derives that $u = v_s$ and $w = v_r$ for some integer $r < s$. One has $\pi_w = \pi_{v_s} = \pi_u$, and in this case $\pi_w = \pi_v$.

The case $z = x$ is similarly dealt with, but interchanging the roles of w_1 and w_2. Thus one assumes $w_2 = q_1 \cdots q_k$, and defines v_i as $q_{k-i+3}v_{i-1}q_{k-i+3}$ for $i \geq 3$, starting from $v_2 = yv_1y = (yxw_2)(yxw_1xy)$ and ending with

$$v_{k+2} = w_2yxw_2yxw_1xyw_2 \in PER \ .$$

Therefore there exist integers r, s such that $1 \leq r < s \leq k+2 = |v| - \pi_v + 2$, $w = v_r$, and $u = v_s$, so that $\pi_w = \pi_u$ and $\pi_w = \pi_{v_1} = |v| - \pi_v + 2$. □

Example 15. Let $w = baaabaaab \in (St \cap PAL)$. Following the notations of Theorem 14, one has $v = aaabaaa$, $v_1 = w$, and $u = v_3 = aabaaabaaabaa$. Thus $\pi_w = \pi_u = \pi_v = 4$.

Let $w = babbbbab$. In this case we have $v = bbbb$, $w = v_2$, and $u = v_4 = bbbabbbbabbb$, so that $\pi_w = \pi_u = 5 = |v| + 1 = |v| - \pi_v + 2$.

For any word $w \in \mathcal{A}^*$, we denote by R_w the minimal nonnegative integer such that there is no right special factor of w of length R_w, and by K_w the length of the shortest unrepeated suffix of w. Conventionally, one assumes $R_\varepsilon = K_\varepsilon = 0$. The following theorem gives a further criterion, different from Proposition 10, to discriminate whether a (Sturmian) palindrome over $\mathcal{A}$ is central or not.

Theorem 16. *Let $w \in \mathcal{A}^*$ be a palindrome, with $\pi_w > 1$. Then w is central if and only if its prefix of length $\pi_w - 2$ is a right special factor of w.*

Proof. From Proposition 10, we can write

$$w = w_1xyw_2 = w_2yxw_1^{\sim} \qquad (4)$$

where $x, y \in \mathcal{A}$, w_2 is the longest proper palindromic suffix of w, $|w_1| = \pi_w - 2$, and w is central if and only if $w_1 \in PAL$ and $x \neq y$. Therefore we have to prove that w_1 is a right special factor of w if and only if $w_1 = w_1^{\sim}$ and $x \neq y$.

Indeed, assume that these two latter conditions are satisfied. Since $w_1^{\sim} = w_1$ and w_2 is the longest proper palindromic suffix (and prefix) of w, one has that w_1 is a proper prefix and suffix of w_2. This implies, from (4), that w_1 is a right special factor of w.

Conversely, suppose w_1 is a right special factor of w. Let us first prove that $w_1 \in PAL$. By hypothesis, we have $\pi_w - 2 = |w_1| \le R_w - 1$, that is $R_w \ge \pi_w - 1$. Since in general one has $\pi_w \ge R_w + 1$ (see [6, Corollary 5.3]), it follows $\pi_w = R_w + 1$. This implies $|w| = R_w + K_w$, again by [6, Corollary 5.3]. The suffix $w_1^{\sim}$ of w is repeated, because w_1 is a right special factor of w, which is a palindrome. This leads to

$$\pi_w - 2 = |w_1^{\sim}| \le K_w - 1$$

and thus to $|w| = R_w + K_w \ge 2\pi_w - 2$. If $|w| = 2\pi_w - 2$, then $|w_1| = |w_2|$ so that one derives $w_1 = w_2 \in PAL$. If $|w| \ge 2\pi_w - 1$, then w has the prefix w_1xyw_1x, so that $yw_1x \in \mathrm{Fact}\,(w)$. Let z be the letter such that $\mathcal{A} = \{x, z\}$. The word w_1z is a factor of w because w_1 is right special. Moreover, since w_1z is not a prefix, there exists a letter y' such that $y'w_1z \in \mathrm{Fact}\,(w)$. One has $y \neq y'$, for otherwise yw_1 would be a right special factor of w of length $\pi_w - 1 = R_w$, which is a contradiction. As w is a palindrome, the words $xw_1^{\sim}y$ and $zw_1^{\sim}y'$ are factors of w too, so that $w_1^{\sim}$ is a right special factor of w. By [6, Proposition 4.7], this implies $w_1 = w_1^{\sim}$. Therefore we get $w_1 \in PAL$ again.

We shall now prove that $x \neq y$. By contradiction, suppose w has the factorization

$$w = (w_1xx)^k w_1\,, \text{ with } k \ge 1$$

as granted by Proposition 10. As before, assume $\mathcal{A} = \{x, z\}$. Since w_1 is a right special factor of w, one has $w_1z \in \mathrm{Fact}\,(w)$. Thus we have either $w_1z = xw_1$ or $w_1z = v_2xxv_1z$, where v_1z is a prefix of w_1 and v_2 is a suffix of w_1. Since $|w_1| = |w_1z| - 1$, we can write $w_1 = v_1z\alpha v_2$, with $\alpha \in \mathcal{A}$. The first case is impossible since w_1 is a palindrome and $x \neq z$. In the latter case, one obtains:

$$v_1z\alpha v_2 = w_1 = w_1^{\sim} = v_1^{\sim}xxv_2^{\sim}$$

which is absurd again, because $x \neq z$. □

Example 17. The word $w = baab$ is a Sturmian palindrome of minimal period $\pi_w = 3$. Its prefix of length 1 is not a right special factor, hence $w \notin PER$. The word $v = ababab bababa$ is a Sturmian palindrome having minimal period 7, and its prefix *ababa* of length 5 is not right special. Therefore $v \notin PER$. On the contrary, the word $u = aabaabaa$ has minimal period 3, and its prefix of length 1 is a right special factor, so that u is central.

Proposition 18. *A palindrome $w \in \mathcal{A}^*$ is Sturmian if and only if $\pi_w = R_w + 1$.*

Proof. The result is trivially true if $\pi_w = 1$. Since for any $w \in \mathcal{A}^*$ one has $\pi_w \ge R_w + 1$ (cf. [6]), in the case $\pi_w > 1$ the condition $\pi_w = R_w + 1$ is equivalent to the existence of a right special factor s of w of length $|s| = \pi_w - 2$.

Let us first prove that every Sturmian palindrome w such that $\pi_w \ge 2$ has such a factor. If w is central, the result follows directly from Theorem 16. Thus we suppose $w \notin PER$, and as in Theorem 14 we denote by v the central median factor of w of maximal length. If $\pi_v = 1$, then there exists a letter $x \in \mathcal{A}$ and an integer $n \geq 1$ such that $v = x^n$. From the maximality condition, one has $n > 1$. In this case, by Theorem 14 one derives $\pi_w = |v| + 1 = n + 1$ and $yx^n y \in \mathrm{Fact}\,(w)$, where $\{x, y\} = \mathcal{A}$; therefore x^{n-1} is the desired right special factor of w, of length $n - 1 = \pi_w - 2$. If $\pi_v > 1$, using Proposition 10 we write v as v_1xyv_2, with $\pi_v = |v_1xy|$. By Theorem 14, one has either $\pi_w = \pi_v$ or $\pi_w = |v| - \pi_v + 2$. In the first case, the result is a consequence of Theorem 16. Indeed, the prefix v_1 of the central word v, whose length is $\pi_v - 2 = \pi_w - 2$, is a right special factor of v, and then of w. In the latter case, one derives that the word $xvx = xv_1xyv_2x = xv_2yxv_1x$ is a factor of w, so that v_2 is a right special factor of w, of length $|v| - \pi_v = \pi_w - 2$.

Conversely, let us assume that w has a right special factor s with $|s| = \pi_w - 2$. By Proposition 10, one can write

$$w = w_1xyw_2 = w_2yxw_1^\sim$$

with $x, y \in \mathcal{A}$, $|w_1xy| = \pi_w$. Moreover, let k be the order of w. By definition, k is the maximal integer such that $(w_1xy)^k$ is a prefix of w. This implies, by periodicity, that w is a prefix of $(w_1xy)^{k+1}$. By applying repeatedly Lemma 12, one obtains that the word $u = (w_1xy)^{k+1}w(yxw_1^\sim)^{k+1}$ has minimal period π_w. Since $s \in \mathrm{Fact}\,((w_1xy)^{k+1})$, there exists a median factor t of u which starts with s and has w as a factor, i.e., $t = s\delta w\delta^\sim s^\sim$ for some word δ. We have $\pi_w \leq \pi_t \leq \pi_u = \pi_w$, therefore from Theorem 16 it follows that $t \in PER$, whence $w \in St$. □

Example 19. Let $w = abba \in St \cap PAL$. One has $\pi_w = 3 = R_w + 1$. The Sturmian word $u = ababaa$ is not a palindrome, but $\pi_u = 5 = R_u + 1$. However, the word $v = aabab \in St$ has $\pi_v = 5 > 3 = R_w + 1$.

The palindrome word $s = aabbaa$ is not Sturmian. One has $\pi_s = 4 > 3 = R_s + 1$.

References

1. Allouche, J.P., Shallit, J.: Chapters 9–10. In: Automatic Sequences. Cambridge University Press, Cambridge UK (2003)
2. Berstel, J., Séébold, P.: Sturmian words. In Lothaire, M., ed.: Algebraic Combinatorics on Words, Chapter 2. Cambridge University Press, Cambridge UK (2001)
3. Borel, J.P., Reutenauer, C.: Palindromic factors of billiard words. Theoretical Computer Science, to appear (2005)
4. Carpi, A.: Private communication (2005)
5. D'Alessandro, F.: A combinatorial problem on Trapezoidal words. Theoretical Computer Science **273** (2002) 11–33
6. de Luca, A.: On the combinatorics of finite words. Theoretical Computer Science **218** (1999) 13–39
7. de Luca, A.: Sturmian words: structure, combinatorics, and their arithmetics. Theoretical Computer Science **183** (1997) 45–82

8. de Luca, A., Mignosi, F.: Some combinatorial properties of Sturmian words. Theoretical Computer Science **136** (1994) 361–385
9. Droubay, X., Pirillo, G.: Palindromes and Sturmian words. Theoretical Computer Science **223** (1999) 73–85
10. Lothaire, M.: Combinatorics on Words. Addison-Wesley, Reading MA (1983)