A Classification of Trapezoidal Words

Gabriele Fici

Laboratoire I3S, CNRS & Université de Nice-Sophia Antipolis
2000, Route des Lucioles - 06903 Sophia Antipolis cedex, France
`fici@i3s.unice.fr`

Trapezoidal words are finite words having at most $n + 1$ distinct factors of length n, for every $n \geq 0$. They encompass finite Sturmian words. We distinguish trapezoidal words into two disjoint subsets: open and closed trapezoidal words. A trapezoidal word is *closed* if its longest repeated prefix has exactly two occurrences in the word, the second one being a suffix of the word. Otherwise it is *open*. We show that open trapezoidal words are all primitive and that closed trapezoidal words are all Sturmian. We then show that trapezoidal palindromes are closed (and therefore Sturmian). This allows us to characterize the special factors of Sturmian palindromes. We end with several open problems.

Keywords: trapezoidal words, Sturmian words, special factors, palindromes.

1 Introduction

In combinatorics on words, the most famous class of infinite words is certainly that of Sturmian words. Sturmian words code digital straight line in the discrete plane having irrational slope. They are characterized by the fact that they have exactly $n + 1$ factors of length n, for every $n \geq 0$.

It is well known ([6], Proposition 2.1.17) that a finite word w is a factor of some Sturmian word if and only if w is a binary *balanced* word, that is, there exists a letter a such that for every pair of factors of w of the same length, u and v, one has that u and v contain the same number of a's up to one, i.e.,

$$||u|_a - |v|_a| \leq 1. \tag{1}$$

Finite Sturmian words (finite factors of Sturmian words) have the property that they have *at most* $n + 1$ factors of length n, for every $n \geq 0$. However, this property does not characterize them, as shown by the word $w = aaabab$, which is not Sturmian since the factors aaa and bab do not verify (1).

The set of finite words defined by the property that they have at most $n + 1$ factors of length n, for every $n \geq 0$, is called the set of *trapezoidal words*.

The name comes from the fact that the graph of the complexity function of these words[1] defines a regular trapezium. Trapezoidal words have been defined by de Luca, who observed that Sturmian words are trapezoidal [7]. The non-Sturmian trapezoidal words have been then characterized by D'Alessandro [4].

In this paper, we distinguish trapezoidal words into two distinct classes, accordingly with the definition below.

Definition 1. *Let w be a finite word over an alphabet Σ. We say that w is* closed *if the longest repeated prefix of w has exactly two occurrences in w, the second one being a suffix of w.*

A word which is not closed is called open.

---

[1] The complexity function of a word w is the function that counts the number of distinct factors of each length in w.

For example, the word *aabbaa* is closed, whereas the word *aabbaaa* is open.

Remark 1. *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 does not appear in w followed by different letters.*

The notion of closed word is also related to the concept of complete return *to a factor u in a word w, as considered in [5]. A complete return to u in w is any factor of w having exactly two occurrences of u, one as a prefix and one as a suffix. Therefore, w is a closed word if and only if w is a complete return to its longest repeated prefix.*

In this paper, we distinguish trapezoidal words in open and closed. This allows us to establish some further properties of trapezoidal words. More precisely, we have that open trapezoidal words are all primitive (Lemma 13), while closed trapezoidal words are all Sturmian (Proposition 12). We characterize open trapezoidal words by means of their special factors (Proposition 10) and show that the longest special factor of a closed trapezoidal word is a central word (Lemma 14). We then show that trapezoidal palindromes are closed (Theorem 16). This allows us to characterize the special factors of Sturmian palindromes (Corollary 17).

2   Trapezoidal Words

An *alphabet*, denoted by Σ, is a finite set of symbols. A *word* over Σ is a finite sequence of symbols from Σ. We denote by $alph(w)$ the subset of the alphabet Σ constituted by the letters appearing in w.

The *length* of a word w is denoted by $|w|$ and is the number of its symbols. We denote by w_i the i-th letter of w. For a letter $a \in \Sigma$, we denote by $|w|_a$ the number of a's appearing in w.

The set of all words over Σ is denoted by Σ^*. The set of all words over Σ having length n is denoted by Σ^n. The empty word has length zero and is denoted by ε.

Let $w = a_1 a_2 \cdots a_n$, $n > 0$, be a non-empty word over the alphabet Σ. The word $\tilde{w} = a_n a_{n-1} \cdots a_1$ is called the *reversal* of w. A *palindrome* is a word w such that $\tilde{w} = w$.

A *prefix* of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_1 a_2 \cdots a_i$, with $1 \le i \le n$. A *suffix* of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_i a_{i+1} \cdots a_n$, with $1 \le i \le n$. A *factor* of w is a prefix of a suffix of w (or, equivalently, a suffix of a prefix of w). Therefore, a factor of w is any word v such that $v = \varepsilon$ or v is of the form $v = a_i a_{i+1} \cdots a_j$, with $1 \le i \le j \le n$. A factor of a word w is *internal* if it is not a prefix nor a suffix of w.

We denote by $Pref(w)$, $Suff(w)$ and $Fact(w)$, respectively, the set of prefixes, suffixes and factors of the word w.

The *factor complexity* of a word w is the function defined by $f_w(n) = |Fact(w) \cap \Sigma^n|$, for every $n \ge 0$. Notice that $f_w(1)$ is the number of distinct letters occurring in w. A *binary word* is a word w such that $f_w(1) = |alph(w)| = 2$.

A factor u of w is *left special* if there exist $a, b \in \Sigma$, $a \ne b$, such that $au, bu \in Fact(w)$. A factor u of w is *right special* if there exist $a, b \in \Sigma$, $a \ne b$, such that $ua, ub \in Fact(w)$. A factor u of w is *bispecial* if it is both left and right special.

For example, let $w = aabbb$. The left special factors of w are ε, b and bb. The right special factors of w are ε and a. Therefore, the only bispecial factor of w is ε.

A *period* for the word w is a positive integer p, with $0 < p \le |w|$, such that $w_i = w_{i+p}$ for every $i = 1, \ldots, |w| - p$. Since $|w|$ is always a period for w, we have that every non-empty word has at least one period. We can unambiguously define *the* period of the word w as the smallest of its periods. For example the period of $w = aabaaba$ is 3.

The *fractional root* z_w of a word w is its prefix whose length is equal to the period of w. So for example the fractional root of $w = aabaaba$ is $z_w = aab$.

A word w is *a power* if there exists a non-empty word u and an integer $n > 1$ such that $w = u^n$. A word which is not a power is called *primitive*.

The following parameters have been introduced by de Luca [7]:

Definition 2. *Let w be a word over Σ. We denote by H_w the minimal length of a prefix of w which occurs only once in w. We denote by K_w the minimal length of a suffix of w which occurs only once in w.*

Definition 3. *Let w be a word over Σ. We denote by L_w the minimal length for which there are no left special factors of that length in w. Analogously, we denote by R_w the minimal length for which there are no right special factors of that length in w.*

Example 1. *Let $w = aaababa$. The longest left special factor of w is aba, and it is also the longest repeated suffix of w; the longest right special factor of w is aa, and it is also the longest repeated prefix of w. Thus, we have $L_w = 4$, $K_w = 4$, $R_w = 3$ and $H_w = 3$.*

Notice that for every word w such that $|alph(w)| > 1$, the values H_w, K_w, L_w and R_w are positive integers. Moreover, one has $f_w(R_w) = f_w(L_w)$ and $\max\{R_w, K_w\} = \max\{L_w, H_w\}$ ([7], Corollary 4.1).

The following proposition is from de Luca ([7], Proposition 4.2).

Proposition 1. *Let w be a word of length $|w|$ such that $|alph(w)| > 1$ and set $m_w = \min\{R_w, K_w\}$ and $M_w = \max\{R_w, K_w\}$. The factor complexity f_w is strictly increasing in the interval $[0, m_w]$, is nondecreasing in the interval $[m_w, M_w]$ and strictly decreasing in the interval $[M_w, |w|]$. Moreover, for i in the interval $[M_w, |w|]$, one has $f_w(i+1) = f_w(i) - 1$. If $R_w < K_w$, then f_w is constant in the interval $[m_w, M_w]$.*

Proposition 1 allows one to give the following definition.

Definition 4. *A non-empty word w is* trapezoidal *if*

- $f_w(i) = i + 1$ *for* $0 \le i \le m_w$,
- $f_w(i+1) = f_w(i)$ *for* $m_w \le i \le M_w - 1$,
- $f_w(i+1) = f_w(i) - 1$ *for* $M_w \le i \le |w|$.

Trapezoidal words have been considered for the first time by de Luca [7]. The name *trapezoidal* has been given by D'Alessandro [4]. The choice of the name is motivated by the fact that for these words the graph of the complexity function defines a regular trapezium (possibly degenerated in a triangle).

Notice that by definition a trapezoidal word is a binary word.

Example 2. *The word $w = aaababa$ considered in Example 1 is trapezoidal. See Fig. 1.*

In the following proposition we gather some characterizations of trapezoidal words.

Proposition 2. *Let w be a binary word. The following conditions are equivalent:*

1. *w is trapezoidal;*
2. *$|w| = L_w + H_w$;*
3. *$|w| = R_w + K_w$;*
4. *w has at most one left special factor of length n for every $n \ge 0$;*
5. *w has at most one right special factor of length n for every $n \ge 0$;*
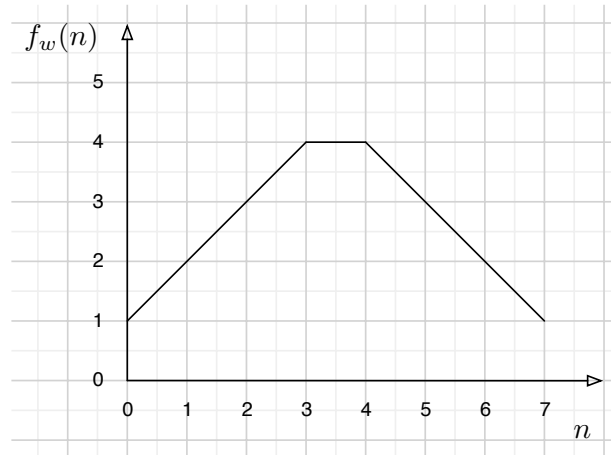6. *w has at most $n + 1$ distinct factors of length n for every $n \ge 0$;*

Figure 1: The graph of the complexity function of the trapezoidal word $w = aaababa$. One has $m_w = \min\{R_w, K_w\} = 3$ and $M_w = \max\{R_w, K_w\} = 4$.

    7. $|f_w(n+1) - f_w(n)| \leq 1$ *for every $n \geq 0$.*

Proof. The equivalence of (1), (2), (3), (4) and (5) is in [7] and [4].

The equivalence of (5) and (6) follows from elementary considerations on the factorial complexity of binary words. Indeed, it is easy to see the number of distinct factors of length $n + 1$ of a binary word w is at most equal to the number of distinct factors of length n plus the number of right special factors of length n.

The equivalence of (7) and (1) comes directly from the definitions and from Proposition 1.

The proof is therefore complete.   □

Recall that a finite word is *Sturmian* if and only if it is balanced, i.e., verifies (1). The following proposition is from de Luca ([7], Proposition 7.1).

Proposition 3. *Let w be a binary word. If w is Sturmian, then w is trapezoidal.*

The inclusion in Proposition 3 is strict, since there exist trapezoidal words that are not Sturmian, e.g. the word $w = aaababa$ considered in Example 1.

Recall that a binary word w is *rich* (or *full*) [5] if it contains $|w| + 1$ distinct palindromic factors, that is the maximum number of distinct palindromic factors a word can contain.

The following proposition is from de Luca, Glen and Zamboni ([11], Proposition 2).

Proposition 4. *Let w be a binary word. If w is trapezoidal, then w is rich.*

Again, the inclusion in Proposition 4 is strict, since there exist rich words that are not trapezoidal, e.g. the word $w = aabbaa$.

D'Alessandro [4] characterized the non-Sturmian trapezoidal words. We report below his characterization.

First, recall that a word w is unbalanced (i.e., w is not balanced) if and only if it contains a pair of *pathological factors* (f, g), that is, f and g are factors of w of the same length but they do not verify (1). Moreover, if f and g are chosen of minimal length, there exists a palindrome u such that $f = aua$ and $g = bub$, for two different letters a and b (see [6], Proposition 2.1.3).

We can also state that f and g are Sturmian words, since otherwise they would contain a pair of pathological factors shorter than $|f| = |g|$ and hence w would contain such a pair of pathological factors, against the minimality of f and g. So the word u is a palindrome such that *aua* and *bub* are Sturmian words, i.e., u is a *central* word [12].

The following lemma is attributed to Aldo de Luca in [4].

Lemma 5. *Let w be a non-Sturmian word and (f,g) the pair of pathological factors of w of minimal length. Then f and g do not overlap in w.*

The following is the characterization of trapezoidal non-Sturmian words given by D'Alessandro [4].

Theorem 6. *Let w be a binary non-Sturmian word. Then w is trapezoidal if and only if*

$$w = pq, \quad \text{with } p \in \text{Suff}(\{\tilde{z}_f^*\}), \quad q \in \text{Pref}(\{z_g^*\})$$

where $\tilde{z}_f$ is the mirror image of the fractional root z_f of f, z_g is the fractional root of g, with (f,g) being the pair of pathological factors of w of minimal length.

In particular, $K_w = |q|$ and the longest right special factor of w is the prefix of w of length $R_w - 1$.

Example 3. *Let $w = aaababa$ be the non-Sturmian trapezoidal word considered in Example 1. We have $f = aaa$ and $g = bab$, so that $\tilde{z}_f = a$ and $z_g = ba$. The word w factorizes as $w = pq$, with $p = aaa$ and $q = baba$.*

Hence, trapezoidal words are either Sturmian or of the form described in Theorem 6. The following result of de Luca, Glen and Zamboni states that trapezoidal palindromes are all Sturmian.

Theorem 7 ([11]). *The following conditions are equivalent:*

1. w is a trapezoidal palindrome;

2. w is a Sturmian palindrome.

Let us give a proof of this latter result based on Theorem 6. We first show that the words p and q in the factorization of Theorem 6 are Sturmian words.

Lemma 8. *Let w be a trapezoidal non-Sturmian word and let $w = pq$, with $p \in \text{Suff}(\{\tilde{z}_f^*\})$ and $q \in \text{Pref}(\{z_g^*\})$, be the factorization of Theorem 6. Then p and q are Sturmian words.*

Proof. Recall that any central word u that is not a power of a single letter can be uniquely written as $u = w_1 xy w_2 = w_2 yx w_1$, for two central words w_1, w_2 and different letters x, y (see [3], Proposition 1).

Let u be the central word such that $f = aua$ and $g = bub$. If u is not a power of a single letter, the fractional roots of f and g are xw_2y and yw_1x ([4], Lemma 2). This implies that $\tilde{z}_f$ and z_g both are conjugate to standard Sturmian words[2]. If $u = x^n$, $x \in \Sigma$, $n > 0$, then the fractional roots of f and g are x and yx^n, so even in this case $\tilde{z}_f$ and z_g both are conjugate to standard Sturmian words.

By Theorem 1 in [10], any word whose fractional root is conjugate to a standard Sturmian word is a Sturmian word. This implies that every word belonging to $\tilde{z}_f^*$ or to z_g^* is Sturmian. Since a factor of a Sturmian word is a Sturmian word, the claim follows. $\square$

Now, let w be a trapezoidal palindrome. If w is non-Sturmian we can write, by Theorem 6, $w = pq = v_1 fg v_2 = \tilde{v}_2 gf \tilde{v}_1$, with $v_1, v_2 \in \{a,b\}^*$ such that $p = v_1 f$ and $q = gv_2$. If $|v_1| = |v_2|$, then $f = g$, a contradiction. If $|v_1| \neq |v_2|$ then either f and g overlap (a contradiction with Lemma 5) or p (or q) contains f and g as factors, a contradiction since, by Lemma 8, p and q are Sturmian. So w cannot be a non-Sturmian word.

Hence, we proved that trapezoidal palindromes are Sturmian. Since by Proposition 3, Sturmian words are trapezoidal, the claim of Theorem 7 follows.

---

[2]Standard Sturmian words are words of length one or of the form *uxy*, with *u* central word and *x, y* different letters [12].

3   Open and Closed Trapezoidal Words

In this section we derive some properties of open and closed trapezoidal words.

Proposition 9. *Let w be a trapezoidal word. Then w is open (resp. closed) if and only if $\tilde{w}$ is open (resp. closed).*

Proof. If w is a closed trapezoidal word, then its longest repeated prefix is also its longest repeated suffix and has exactly two occurrences in w. This implies that $\tilde{w}$ has the same property. So $\tilde{w}$ is a closed trapezoidal word. Hence the set of closed trapezoidal words is closed by reversal.

Since the whole set of trapezoidal words is closed by reversal ([4], Corollary 7) and since open trapezoidal words form the complement of closed trapezoidal words in the set of trapezoidal words, the set of open trapezoidal words is closed by reversal too.   □

The following proposition gives a characterization of open trapezoidal words.

Proposition 10. *Let w be a trapezoidal word. Then the following conditions are equivalent:*

1. *w is open;*

2. *the longest repeated prefix of w is also the longest right special factor of w;*

3. *the longest repeated suffix of w is also the longest left special factor of w.*

Proof. $(1) \Rightarrow (2)$. Let h be the longest repeated prefix of w and x the letter such that hx is a prefix of w. Since w is not closed, h has a second non-suffix occurrence in w followed by letter y. Since h is the longest repeated prefix of w, we have $y \neq x$. Therefore, h is right special in w.

Suppose that w has a right special factor r longer than h. Since w is trapezoidal, w has at most one right special factor for each length (Proposition 2). Since the suffixes of a right special factor are right special factors, we have that h must be a proper suffix of r. Since r is right special in w, it has at least two occurrences in w followed by different letters. This implies a non-prefix occurrence of hx in w, against the definition of h.

$(2) \Rightarrow (3)$. Let k be the longest repeated suffix of w. We first prove that k is left special in w. Otherwise, k appears in w exactly twice, once as a prefix and once as a suffix of w. This implies that $k = h$, the longest repeated prefix of w – a contradiction, since by hypothesis h is right special in w and therefore it has at least a non-suffix occurrence.

It remains to prove that k is the longest left special factor of w. Since w is trapezoidal, we have, by Proposition 2, $|w| = L_w + H_w = R_w + K_w$. Since by hypothesis the longest repeated prefix of w is also the longest right special factor of w, we have $H_w = R_w$ and therefore $L_w = K_w$. Thus, the longest left special factor of w has length equal to $L_w - 1 = K_w - 1 = |k|$.

$(3) \Rightarrow (1)$. Let k be the longest repeated suffix of w and let ak and bk be factors of w, for different letters a and b. Then we have that $\tilde{k}$ is the longest repeated prefix of $\tilde{w}$, and $\tilde{k}a$ and $\tilde{k}b$ are factors of $\tilde{w}$. This proves that the word $\tilde{w}$ is open. The claim then follows from Proposition 9.   □

Lemma 11. *Let w be a trapezoidal word. If w is open, then $H_w = R_w$ and $K_w = L_w$. If w is closed, then $H_w = K_w$ and $L_w = R_w$.*

Proof. The claim for open trapezoidal words follows directly from Proposition 10.

Suppose that w is a closed trapezoidal word. Then $H_w = K_w$ (since w is closed) and therefore $L_w = R_w$ (since w is trapezoidal).   □

Open trapezoidal words can be Sturmian (e.g. $w = aaabaa$) or not (e.g. $w = aaabab$). Closed trapezoidal words, instead, are always Sturmian, as shown in the following proposition.

Proposition 12. *Let w be a trapezoidal word. If w is closed, then w is Sturmian.*

Proof. Suppose that w is not Sturmian. Then, by Theorem 6, $w = pq$, $p \in \textit{Suff}(\{\tilde{z}_f^*\})$, $q \in \textit{Pref}(\{z_g^*\})$, with (f, g) being the pair of pathological factors of w of minimal length, $K_w = |q|$ (and hence $R_w = |p|$) and the longest right special factor of w is the prefix of w of length $R_w - 1$.

Since w is closed, we have, by Lemma 11, $H_w = K_w$. So the suffix k of q of length $K_w - 1$ is also the longest repeated prefix of w and appears in w only as a prefix and as a suffix.

If $R_w \geq K_w$, then k is a prefix of the longest right special factor of w. This is a contradiction with the fact that k appears in w only as a prefix and as a suffix.

If $R_w < K_w$, then p is a prefix of k and hence p is a factor of q. This implies that f is a factor of q, and therefore q contains both f and g as factors. Hence p would be non-Sturmian, a contradiction with Lemma 8. □

The result stated in Proposition 12 can also be found in a paper of Bucci, de Luca and De Luca ([1], Proposition 3.6).

As a corollary of Proposition 12, we have that every trapezoidal word is open or Sturmian. We therefore propose the following

Problem 1. *Give a characterization of open Sturmian words.*

Trapezoidal words, as well as Sturmian words, can be primitive (e.g. $w = aabaaa$) or not (e.g. $w = aabaab$). Open trapezoidal words (and in particular, then, non-Sturmian trapezoidal words) are always primitive.

Lemma 13. *Every open trapezoidal word is primitive.*

Proof. Suppose that w is not primitive. Let $w = u^n$, for a non-empty primitive word u and an integer $n > 1$. The longest repeated prefix of w is therefore u^{n-1}, which is also a suffix of w. Moreover, by elementary combinatorics on words, u^{n-1} cannot have internal occurrences in w. Hence w is closed. □

The converse of Lemma 13 does not hold. Indeed, there exist trapezoidal words that are primitive but not open, e.g. $w = aabaa$.

We now focus on closed trapezoidal words and their special factors.

Lemma 14. *Let w be a closed trapezoidal word and let u be the longest left special factor of w. Then u is also the longest right special factor of w (and thus u is a bispecial factor of w). Moreover, u is a central word.*

Proof. Let u be the longest left special factor of w. Hence there exist different letters $a, b \in \Sigma$ such that au and bu are factors of w.

We claim that both au and bu occur in w followed by some letter. Indeed, suppose the contrary. Then one of the words au and bu, say au, appears in w only as a suffix. Let k be the longest repeated suffix of w. Since u is a repeated suffix of w, we have $|k| \geq |u|$. If $|k| = |u|$, then $k = u$ and since w is closed, u appears in w only as a prefix and as a suffix, against the hypothesis that bu is a factor of w. So $|k| > |u|$ and therefore au must be a suffix of k. This implies an internal occurrence of au in w, a contradiction.

So, there exist letters x, y such that aux and buy are factors of w. Now, we must have $x \neq y$, since otherwise ux would be a left special factor of w longer than u. Thus u is right special in w. Since w is

closed, we have, by Lemma 11, $H_w = K_w$ and $L_w = R_w$, so w cannot contain a right special factor longer than u. Thus u is the longest right special factor of w.

By Proposition 12, w is Sturmian. Since u is a bispecial factor of w, and since a factor of a Sturmian word is a Sturmian word, in order to prove that u is a central word it is sufficient to prove that u is a palindrome. Suppose the contrary. So there exists a prefix z of u and a letter $a \in \Sigma$ such that za is a prefix of u and $b\tilde{z}$ is a suffix of u, for a letter b different from a. Since u is bispecial in w, this implies that aza and $b\tilde{z}b$ are both factors of w. This implies that w is not balanced, a contradiction with Proposition 12. $\qquad\square$

By Theorem 7, trapezoidal palindromes coincide with Sturmian palindromes. Deep and interesting results on Sturmian palindromes can be found in [8] and [9]. In particular, we want to recall the following

Theorem 15 ([9], Theorem 29). *A palindrome $w \in \Sigma^*$ is Sturmian if and only if $\pi_w = R_w + 1$.*

The next theorem shows that Sturmian palindromes are all closed words.

Theorem 16. *Let w be a trapezoidal (Sturmian) palindrome. Then w is closed.*

Proof. By contradiction, suppose that w is open. Then h, the longest repeated prefix of w, is also the longest right special factor of w (Proposition 10). Since w is a palindrome, we have that the longest repeated suffix of w is $\tilde{h}$, the reversal of h. In particular, then, $K_w = H_w$.

By Lemma 11, $H_w = R_w$ and $K_w = L_w$. Thus we have $H_w = R_w = K_w = L_w = |w|/2$, since w is trapezoidal (see Proposition 2). It follows that $w = hxx\tilde{h}$, for a letter $x \in \Sigma$.

By Theorem 15, the period of w is $R_w + 1 = |hxx|$, so we have $\tilde{h} = h$. Therefore, we have $w = hxxh$.

Since h is right special in w, there exists a letter $y \neq x$ such that hy is a factor of w. Now, any occurrence of hy in w cannot be preceded by the letter x, since $h = \tilde{h}$ is the longest repeated suffix of w. Thus w contains the factor yhy.

Hence, w contains both hxx and yhy as factors, and this contradicts the fact that w is Sturmian. $\qquad\square$

Remark 2. *The equivalence between trapezoidal palindromes and Sturmian palindromes (Theorem 7) can also be derived as a consequence of Theorem 16, Proposition 12 and Proposition 3.*

From Theorem 16 and Lemma 14, we derive the following characterization of the special factors of Sturmian palindromes.

Corollary 17. *Let w be a trapezoidal (Sturmian) palindrome. Then the longest left special factor of w is also the longest right special factor of w and it is a central word.*

Example 4. *Let $w = aababaa$. The longest left special factor of w is aba, which is also its longest right special factor and is a central word.*

4   Conclusions and Open Problems

In this paper we distinguished trapezoidal words into two disjoint classes: open and closed. We derived some combinatorial and structural properties of these two classes of words.

Many further development directions can arise. For example, a challenging problem could be that of finding a characterization of open Sturmian words, that is, of Sturmian words for which the longest repeated prefix is also the longest right special factor (Problem 1).

Another interesting problem concerns enumeration. Enumeration formulae for Sturmian words [13] and for primitive Sturmian words [10] are known. To the best of our knowledge, an enumeration formula for trapezoidal words is not yet known. A possible direction for finding it could be enumerating open and closed trapezoidal words separately.

5   Acknowledgments

References

[1] M. Bucci, A. de Luca & A. De Luca (2009): *Rich and Periodic-Like Words*. In V. Diekert & D. Nowotka, editors: *Developments in Language Theory, 13th International Conference, DLT 2009, Stuttgart, Germany, June 30–July 3, 2009, Proceedings, Lecture Notes in Comput. Sci.* 5583, Springer, pp. 145–155, doi:10.1007/978-3-642-02737-6_11.

[2] A. Carpi & A. de Luca (2001): *Periodic-like words, periodicity and boxes. Acta Informatica* 37, pp. 597–618, doi:10.1007/PL00013314.

[3] A. Carpi & A. de Luca (2005): *Central Sturmian Words: Recent Developments*. In C. De Felice & A. Restivo, editors: *Developments in Language Theory, 9th International Conference, DLT 2005, Palermo, Italy, July 4-8, 2005, Proceedings, Lecture Notes in Comput. Sci.* 3572, Springer, pp. 36–56, doi:10.1007/11505877_4.

[4] F. D'Alessandro (2002): *A combinatorial problem on Trapezoidal words. Theoret. Comput. Sci.* 273(1-2), pp. 11–33, doi:10.1016/S0304-3975(00)00431-X.

[5] A. Glen, J. Justin, S. Widmer & L. Q. Zamboni (2009): *Palindromic richness. European J. Combin.* 30, pp. 510–531, doi:10.1016/j.ejc.2008.04.006.

[6] M. Lothaire (2002): *Algebraic Combinatorics on Words*. Encyclopedia of Mathematics and its Applications, Cambridge Univ. Press, New York, NY, USA.

[7] A. de Luca (1999): *On the combinatorics of finite words. Theoret. Comput. Sci.* 218, pp. 13–39, doi:10.1016/S0304-3975(98)00248-5.

[8] A. de Luca & A. De Luca (2005): *Palindromes in Sturmian Words*. In C. De Felice & A. Restivo, editors: *Developments in Language Theory, 9th International Conference, DLT 2005, Palermo, Italy, July 4-8, 2005, Proceedings, Lecture Notes in Comput. Sci.* 3572, Springer, pp. 199–208, doi:10.1007/11505877_18.

[9] A. de Luca & A. De Luca (2006): *Combinatorial properties of Sturmian palindromes. Internat. J. Found. Comput. Sci.* 17(3), pp. 557–573, doi:10.1142/S0129054106003991.

[10] A. de Luca & A. De Luca (2006): *Some characterizations of finite Sturmian words. Theoret. Comput. Sci.* 356, pp. 118–125, doi:10.1016/j.tcs.2006.01.036.

[11] A. de Luca, A. Glen & L. Q. Zamboni (2008): *Rich, Sturmian, and Trapezoidal words. Theoret. Comput. Sci.* 407, pp. 569–573, doi:10.1016/j.tcs.2008.06.009.

[12] A. de Luca & F. Mignosi (1994): *Some combinatorial properties of Sturmian words. Theoret. Comput. Sci.* 136(2), pp. 361–385, doi:10.1016/0304-3975(94)00035-H.

[13] F. Mignosi (1991): *On the number of factors of Sturmian words. Theoret. Comput. Sci.* 82, pp. 71–84, doi:10.1016/0304-3975(91)90172-X.