A Classification of Trapezoidal Words

Gabriele Fici

8th International Conference on Words

Prague – 13 September 2011

- Infinite Sturmian words have exactly $n+1$ factors of length n, for every $n \geq 0$.

Trapezoidal words

- Infinite Sturmian words have exactly $n+1$ factors of length n, for every $n \geq 0$.
- Hence, finite Sturmian words (i.e., balanced words) have at most $n+1$ factors of length n, for every $n \geq 0$.

Trapezoidal words

- Infinite Sturmian words have exactly $n+1$ factors of length n, for every $n \geq 0$.
- Hence, finite Sturmian words (i.e., balanced words) have at most $n+1$ factors of length n, for every $n \geq 0$.
- However, this property does not characterize balanced words, e.g. $w = aaabab$ has $\leq n+1$ factors of length n, $\forall n \geq 0$, but is not balanced.

Trapezoidal words

- Infinite Sturmian words have exactly $n+1$ factors of length n, for every $n \geq 0$.
- Hence, finite Sturmian words (i.e., balanced words) have at most $n+1$ factors of length n, for every $n \geq 0$.
- However, this property does not characterize balanced words, e.g. $w = aaabab$ has $\leq n+1$ factors of length n, $\forall n \geq 0$, but is not balanced.

Definition

A word having at most $n+1$ factors of length n, for every $n \geq 0$, is called a trapezoidal word.

Trapezoidal words

- Infinite Sturmian words have exactly $n+1$ factors of length n, for every $n \geq 0$.
- Hence, finite Sturmian words (i.e., balanced words) have at most $n+1$ factors of length n, for every $n \geq 0$.
- However, this property does not characterize balanced words, e.g. $w = aaabab$ has $\leq n+1$ factors of length n, $\forall n \geq 0$, but is not balanced.

Definition

A word having at most $n+1$ factors of length n, for every $n \geq 0$, is called a trapezoidal word.

Thus, trapezoidal words encompass finite Sturmian words.

Lemma

$$Sturm \subset Trap$$

Trapezoidal words

The name comes from the shape of the factor complexity function of these words.

$$f_w(n) = \text{number of distinct factors of length } n \text{ in the word } w.$$

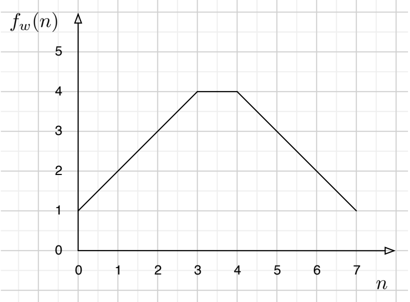


Figure: The factor complexity f_w of the trapezoidal word *aaababa*.

Special Factors

Definition

- v is a left special factor of w if there exist $a \neq b$ such that av and bv are factors of w
- v is a right special factor of w if there exist $a \neq b$ such that va and vb are factors of w
- v is a bispecial factor of w if it is both left and right special

Definition

- v is a left special factor of w if there exist $a \neq b$ such that av and bv are factors of w
- v is a right special factor of w if there exist $a \neq b$ such that va and vb are factors of w
- v is a bispecial factor of w if it is both left and right special

Example

$$w = aaababa$$

ab is left special

Definition

- v is a left special factor of w if there exist $a \neq b$ such that av and bv are factors of w
- v is a right special factor of w if there exist $a \neq b$ such that va and vb are factors of w
- v is a bispecial factor of w if it is both left and right special

Example

$$w = aaababa$$

ab is left special
aa is right special

Special Factors

Definition

- v is a left special factor of w if there exist $a \neq b$ such that av and bv are factors of w
- v is a right special factor of w if there exist $a \neq b$ such that va and vb are factors of w
- v is a bispecial factor of w if it is both left and right special

Example

$$w = aaababa$$

ab is left special
aa is right special
a is bispecial

First Characterization

Lemma

A binary word w is trapezoidal $\Longleftrightarrow$ w has at most one right special factor for each length.

First Characterization

Lemma

A binary word w is trapezoidal $\iff$ w has at most one right special factor for each length.

Analogously,

Lemma

A binary word w is trapezoidal $\iff$ it has at most one left special factor for each length.

First Characterization

Lemma

A binary word w is trapezoidal $\iff$ w has at most one right special factor for each length.

Analogously,

Lemma

A binary word w is trapezoidal $\iff$ it has at most one left special factor for each length.

Example ($w = aaababa$)

The right special factors of w are ε, a, aa.
The left special factors of w are ε, a, ab, aba.

Trapezoidal words

For a non-empty word w one can define:

- R_w the minimal length for which there are not right special factors in w
- K_w the minimal length of an unrepeated suffix of w

Lemma

A binary word w is trapezoidal $\iff |w| = R_w + K_w$

Trapezoidal words

For a non-empty word w one can define:

- R_w the minimal length for which there are not right special factors in w
- K_w the minimal length of an unrepeated suffix of w

Lemma

A binary word w is trapezoidal $\iff |w| = R_w + K_w$

Analogously, one can define:

- L_w the minimal length for which there are not left special factors in w
- H_w the minimal length of an unrepeated prefix of w

Lemma

A binary word w is trapezoidal $\iff |w| = L_w + H_w$

For a non-empty word w one can define:

- R_w the minimal length for which there are not right special factors in w
- K_w the minimal length of an unrepeated suffix of w

Lemma

A binary word w is trapezoidal $\iff |w| = R_w + K_w$

Analogously, one can define:

- L_w the minimal length for which there are not left special factors in w
- H_w the minimal length of an unrepeated prefix of w

Lemma

A binary word w is trapezoidal $\iff |w| = L_w + H_w$

Example ($w = aaababa$)

One has $K_w = 4$ and $R_w = 3$; $H_w = 3$ and $L_w = 4$. Hence w is trapezoidal.

Lemma

A finite binary word w is non-Sturmian if and only if one can write

$$w = x_1 \cdot aua \cdot x_2 \cdot bub \cdot x_3$$

with $x_1, x_2, x_3 \in \Sigma^*$, $\{a, b\} = \Sigma$ and u a Sturmian palindrome.

The pair $f = aua$, $g = bub$ is the pathological pair of minimal length of w.

Non-Sturmian Words

Lemma

A finite binary word w is non-Sturmian if and only if one can write

$$w = x_1 \cdot aua \cdot x_2 \cdot bub \cdot x_3$$

with $x_1, x_2, x_3 \in \Sigma^*$, $\{a, b\} = \Sigma$ and u a Sturmian palindrome.

The pair $f = aua$, $g = bub$ is the pathological pair of minimal length of w.

Theorem (D'Alessandro, 02)

Let w be a binary non-Sturmian word and z_f, z_g the roots of f and g resp. The word w is trapezoidal if and only if one can write

$$w = pq$$

with $p \in \mathrm{Suff}(\tilde{z}_f{}^*)$ and $q \in \mathrm{Pref}(z_g^*)$.

Lemma

A finite binary word w is non-Sturmian if and only if one can write

$$w = x_1 \cdot aua \cdot x_2 \cdot bub \cdot x_3$$

with $x_1, x_2, x_3 \in \Sigma^*$, $\{a, b\} = \Sigma$ and u a Sturmian palindrome.

The pair $f = aua$, $g = bub$ is the pathological pair of minimal length of w.

Theorem (D'Alessandro, 02)

Let w be a binary non-Sturmian word and z_f, z_g the roots of f and g resp. The word w is trapezoidal if and only if one can write

$$w = pq$$

with $p \in \mathrm{Suff}(\tilde{z}_f{}^*)$ and $q \in \mathrm{Pref}(z_g^*)$.

Example ($w = aaababa$)

One has $f = aaa$, $g = bab$, $\tilde{z}_f = a$, $z_g = ba$. So w is trapezoidal.

Theorem (D'Alessandro, 02)

Let w be a binary non-Sturmian word and z_f, z_g the roots of f and g resp. The word w is trapezoidal if and only if one can write

$$w = pq$$

with $p \in \mathrm{Suff}(\tilde{z}_f{}^*)$ and $q \in \mathrm{Pref}(z_g^*)$.

Theorem (D'Alessandro, 02)

Let w be a binary non-Sturmian word and z_f, z_g the roots of f and g resp. The word w is trapezoidal if and only if one can write

$$w = pq$$

with $p \in \mathrm{Suff}(\tilde{z}_f{}^*)$ and $q \in \mathrm{Pref}(z_g^*)$.

Lemma (Fici, 11)

In the factorization above, the words p and q are Sturmian words.

Theorem (D'Alessandro, 02)

Let w be a binary non-Sturmian word and z_f, z_g the roots of f and g resp. The word w is trapezoidal if and only if one can write

$$w = pq$$

with $p \in \mathrm{Suff}(\tilde{z}_f{}^*)$ and $q \in \mathrm{Pref}(z_g^*)$.

Lemma (Fici, 11)

In the factorization above, the words p and q are Sturmian words.

As a consequence:

Theorem (de Luca, Glen, Zamboni, 08 – Fici, 11)

The following conditions are equivalent:

- *w is a Sturmian palindrome*
- *w is a trapezoidal palindrome*

Theorem (Mignosi, 91)

The number of Sturmian words of length n is

$$1 + \sum_{i=1}^{n} (n - i + 1)\phi(i)$$

where $\phi(i)$ is the totient Euler function.

Counting Trapezoidal Words

Theorem (Mignosi, 91)

The number of Sturmian words of length n is

$$1 + \sum_{i=1}^{n} (n - i + 1)\phi(i)$$

where $\phi(i)$ is the totient Euler function.

Theorem (Bucci, De Luca, Fici, 11)

The number of non-Sturmian trapezoidal words of length $n \geq 4$ is

$$\sum_{i=0}^{\lfloor (n-4)/2 \rfloor} 2(n - 2i - 3)\phi(i + 2)$$

Counting Trapezoidal Words

Theorem (Mignosi, 91)

The number of Sturmian words of length n is

$$1 + \sum_{i=1}^{n} (n - i + 1)\phi(i)$$

where $\phi(i)$ is the totient Euler function.

Theorem (Bucci, De Luca, Fici, 11)

The number of non-Sturmian trapezoidal words of length $n \geq 4$ is

$$\sum_{i=0}^{\lfloor (n-4)/2 \rfloor} 2(n - 2i - 3)\phi(i + 2)$$

hence we have an enumerative formula for trapezoidal words.

Definition

A word *w* is closed if its longest repeated prefix has exactly two occurrences in the word, the second one being a suffix of the word.
Otherwise *w* is open.

Open and Closed Words

Definition

A word w is closed if its longest repeated prefix has exactly two occurrences in the word, the second one being a suffix of the word.
Otherwise w is open.

Example

$w = aabbaa$ is closed; $w = aabbaaa$ is open.

Open and Closed Words

Definition

A word w is closed if its longest repeated prefix has exactly two occurrences in the word, the second one being a suffix of the word.
Otherwise w is open.

Example

$w = aabbaa$ is closed; $w = aabbaaa$ is open.

Remark

Closed words are also called periodic-like words or complete returns.

Open and Closed Words

Definition

A word w is closed if its longest repeated prefix has exactly two occurrences in the word, the second one being a suffix of the word.
Otherwise w is open.

Example

$w = aabbaa$ is closed; $w = aabbaaa$ is open.

Remark

Closed words are also called periodic-like words or complete returns.

We want to study open and closed trapezoidal words.

Open and Closed Words

Recall that:

R_w is the minimal length for which there are not right special factors in w

K_w is the minimal length of an unrepeated suffix of w

L_w is the minimal length for which there are not left special factors in w

H_w is the minimal length of an unrepeated prefix of w

Lemma

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 $L_w = R_w$.

Open and Closed Words

Recall that:

R_w is the minimal length for which there are not right special factors in w

K_w is the minimal length of an unrepeated suffix of w

L_w is the minimal length for which there are not left special factors in w

H_w is the minimal length of an unrepeated prefix of w

Lemma

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 $L_w = R_w$.

Remark

One can have $R_w = K_w = L_w = H_w$ as for example in $w = abba$ (closed) or $w = aaba$ (open).

Proposition

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

- *w is open;*
- *the longest repeated prefix of w is also the longest right special factor of w;*
- *the longest repeated suffix of w is also the longest left special factor of w.*

Open Trapezoidal Words

Proposition

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

- *w is open;*
- *the longest repeated prefix of w is also the longest right special factor of w;*
- *the longest repeated suffix of w is also the longest left special factor of w.*

Lemma

Every open trapezoidal word is primitive.

Proposition

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

- *w is open;*
- *the longest repeated prefix of w is also the longest right special factor of w;*
- *the longest repeated suffix of w is also the longest left special factor of w.*

Lemma

Every open trapezoidal word is primitive.

Problem

Give a characterization of open Sturmian words.

Closed Trapezoidal Words

Proposition (Bucci, de Luca, De Luca, 09)

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

Proposition (Bucci, de Luca, De Luca, 09)

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

Lemma

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. Moreover, u is a central Sturmian word.

Proposition (Bucci, de Luca, De Luca, 09)

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

Lemma

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. Moreover, u is a central Sturmian word.

Example

Let $w = aababaaba$.

The longest repeated prefix is *aaba*, which is also the longest repeated suffix and does not have internal occurrences.

The longest left special factor of w is *aba*, which is also its longest right special factor and it is a central word.

Closed Trapezoidal Words

Proposition (Bucci, de Luca, De Luca, 09)

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

Lemma

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. Moreover, u is a central Sturmian word.

Example

Let $w = aababaaba$.

The longest repeated prefix is *aaba*, which is also the longest repeated suffix and does not have internal occurrences.

The longest left special factor of w is *aba*, which is also its longest right special factor and it is a central word.

Theorem

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

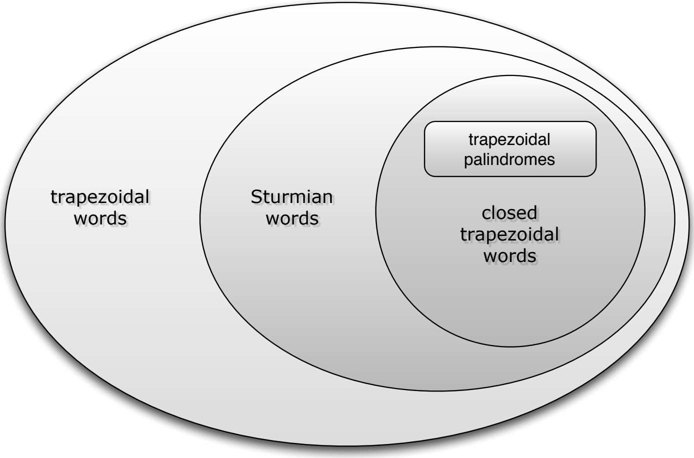
Theorem

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

Corollary

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 Sturmian word.

Venn Diagram for Trapezoidal Words



Open Problems

Some open problems:

- Characterize open Sturmian words
- Give an enumerative formula for open and closed trapezoidal words
- Exploit the dichotomy open/closed to study other classes of words

Thank you!