

Characteristic Parameters and Special Trapezoidal Words

Alma D'Aniello Alessandro De Luca

Università di Napoli Federico II

12th International Conference on Words
Loughborough University, September 13th, 2019

Outline

1 Basic Notions

- Characteristic Parameters
- Trapezoidal Words

2 Main Results

- Closed (vs. Open) Prefixes
- Characterizing Special Trapezoidal Words

3 Conclusions



The Parameters H_w , K_w , L_w , R_w

Recall that a factor u of a word w is **left** (resp. **right**) **special** if xu, yu (resp. ux, uy) are factors of w for some letters $x \neq y$.

Notation

Let w be a finite word. Then:

- L_w (resp. R_w) denotes the shortest length for which w has no left (resp. right) special factors;
- H_w (resp. K_w) denotes the length of the shortest unioccurrent prefix (resp. suffix) of w .

The Parameters H_w , K_w , L_w , R_w

Recall that a factor u of a word w is left (resp. right) special if xu, yu (resp. ux, uy) are factors of w for some letters $x \neq y$.

Notation

Let w be a finite word. Then:

- L_w (resp. R_w) denotes the shortest length for which w has no left (resp. right) special factors;
- H_w (resp. K_w) denotes the length of the shortest unioccurrent prefix (resp. suffix) of w .

The Parameters H_w , K_w , L_w , R_w

Recall that a factor u of a word w is left (resp. right) special if xu, yu (resp. ux, uy) are factors of w for some letters $x \neq y$.

Notation

Let w be a finite word. Then:

- L_w (resp. R_w) denotes the shortest length for which w has no left (resp. right) special factors;
- H_w (resp. K_w) denotes the length of the shortest unioccurrent prefix (resp. suffix) of w .

Example

For

$$w = aaaabba$$

we have:

- $L_w = 3$, as bb is a left special factor of maximal length;
- $H_w = 4$, as the prefix $aaaa$ is unioccurrent whereas aaa is not;
- $R_w = 4$, $K_w = 2$.

Example

For

$$w = aaaabbba$$

we have:

- $L_w = 3$, as bb is a left special factor of maximal length;
- $H_w = 4$, as the prefix $aaaa$ is unioccurrent whereas aaa is not;
- $R_w = 4$, $K_w = 2$.

Example

For

$$w = aaaabbba$$

we have:

- $L_w = 3$, as bb is a left special factor of maximal length;
- $H_w = 4$, as the prefix $aaaa$ is unioccurrent whereas aaa is not;
- $R_w = 4$, $K_w = 2$.

Link with Factor Complexity

Theorem (de Luca 1999 etc.)

Let $w \in A^*$ and $\{R_w, K_w\} = \{m, M\}$, with $m \leq M$ and $\text{card } A > 1$. The *factor complexity* $f_w(n)$ of w is

- strictly increasing for $n \leq m$;
- nondecreasing for $m < n \leq M$;
- strictly decreasing (of exactly 1 at each step) when $M < n \leq |w|$.

By symmetry, the same holds when $\{m, M\} = \{L_w, H_w\}$.
In particular, it follows that

$$\max\{R_w, K_w\} = \max\{L_w, H_w\}.$$

Link with Factor Complexity

Theorem (de Luca 1999 etc.)

Let $w \in A^*$ and $\{R_w, K_w\} = \{m, M\}$, with $m \leq M$ and $\text{card } A > 1$. The *factor complexity* $f_w(n)$ of w is

- *strictly increasing for $n \leq m$;*
- *nondecreasing for $m < n \leq M$;*
- *strictly decreasing (of exactly 1 at each step) when $M < n \leq |w|$.*

By symmetry, the same holds when $\{m, M\} = \{L_w, H_w\}$.
In particular, it follows that

$$\max\{R_w, K_w\} = \max\{L_w, H_w\}.$$

Link with Factor Complexity

Theorem (de Luca 1999 etc.)

Let $w \in A^*$ and $\{R_w, K_w\} = \{m, M\}$, with $m \leq M$ and $\text{card } A > 1$. The *factor complexity* $f_w(n)$ of w is

- strictly increasing for $n \leq m$;
- nondecreasing for $m < n \leq M$;
- strictly decreasing (of exactly 1 at each step) when $M < n \leq |w|$.

By symmetry, the same holds when $\{m, M\} = \{L_w, H_w\}$.
In particular, it follows that

$$\max\{R_w, K_w\} = \max\{L_w, H_w\}.$$

Link with Factor Complexity

Theorem (de Luca 1999 etc.)

Let $w \in A^*$ and $\{R_w, K_w\} = \{m, M\}$, with $m \leq M$ and $\text{card } A > 1$. The *factor complexity* $f_w(n)$ of w is

- strictly increasing for $n \leq m$;
- nondecreasing for $m < n \leq M$;
- strictly decreasing (of exactly 1 at each step) when $M < n \leq |w|$.

By symmetry, the same holds when $\{m, M\} = \{L_w, H_w\}$.
In particular, it follows that

$$\max\{R_w, K_w\} = \max\{L_w, H_w\}.$$

Link with Factor Complexity

Theorem (de Luca 1999 etc.)

Let $w \in A^*$ and $\{R_w, K_w\} = \{m, M\}$, with $m \leq M$ and $\text{card } A > 1$. The factor complexity $f_w(n)$ of w is

- strictly increasing for $n \leq m$;
- nondecreasing for $m < n \leq M$;
- strictly decreasing (of exactly 1 at each step) when $M < n \leq |w|$.

By symmetry, the same holds when $\{m, M\} = \{L_w, H_w\}$.
In particular, it follows that

$$\max\{R_w, K_w\} = \max\{L_w, H_w\}.$$

Motivating a Definition

Theorem (de Luca 1999)

① For all w ,

$$|w| \geq R_w + K_w \quad \text{and} \quad |w| \geq L_w + H_w.$$

② If w is a factor of a Sturmian word, then

$$|w| = R_w + K_w = L_w + H_w.$$

However, some non-Sturmian words such as $aabb$ also verify the equality...

Motivating a Definition

Theorem (de Luca 1999)

① *For all w ,*

$$|w| \geq R_w + K_w \quad \text{and} \quad |w| \geq L_w + H_w.$$

② *If w is a factor of a Sturmian word, then*

$$|w| = R_w + K_w = L_w + H_w.$$

However, some non-Sturmian words such as $aabb$ also verify the equality...

Motivating a Definition

Theorem (de Luca 1999)

① *For all w ,*

$$|w| \geq R_w + K_w \quad \text{and} \quad |w| \geq L_w + H_w.$$

② *If w is a factor of a Sturmian word, then*

$$|w| = R_w + K_w = L_w + H_w.$$

However, some non-Sturmian words such as $aabb$ also verify the equality...

Motivating a Definition

Theorem (de Luca 1999)

① *For all w ,*

$$|w| \geq R_w + K_w \quad \text{and} \quad |w| \geq L_w + H_w.$$

② *If w is a factor of a Sturmian word, then*

$$|w| = R_w + K_w = L_w + H_w.$$

However, some non-Sturmian words such as $aabb$ also verify the equality...

A Finite Analogue of Sturmian Words

Definition

A word w is **trapezoidal** if $|w| = R_w + K_w$
(or *equivalently*, if $|w| = L_w + H_w$).

Theorem (D'Alessandro 2002)

w is trapezoidal $\iff f_w(n) \leq n + 1$ for all $n \geq 0$.

A Finite Analogue of Sturmian Words

Definition

A word w is trapezoidal if $|w| = R_w + K_w$
(or *equivalently*, if $|w| = L_w + H_w$).

Theorem (D'Alessandro 2002)

w is trapezoidal $\iff f_w(n) \leq n + 1$ for all $n \geq 0$.

Central Words

Central words are the palindromic prefixes of standard Sturmian words, and enjoy remarkable characterizations such as:

Theorem (de Luca, Mignosi 1994 etc.)

- ① *A word w is central if and only if it can be written as a^n , b^n , or $uabv = vbau$ for some $n \geq 0$, words u, v , and letters $a \neq b$.*
- ② *w is central if and only if it has two coprime periods p, q such that $|w| = p + q - 2$.*

Example

$abaaba$ is central, whereas $abba$ is not.

Central Words

Central words are the palindromic prefixes of standard Sturmian words, and enjoy remarkable characterizations such as:

Theorem (de Luca, Mignosi 1994 etc.)

- 1 *A word w is central if and only if it can be written as a^n , b^n , or $uabv = vbau$ for some $n \geq 0$, words u, v , and letters $a \neq b$.*
- 2 *w is central if and only if it has two coprime periods p, q such that $|w| = p + q - 2$.*

Example

abaaba is central, whereas *abba* is not.

Central Words

Central words are the palindromic prefixes of standard Sturmian words, and enjoy remarkable characterizations such as:

Theorem (de Luca, Mignosi 1994 etc.)

- 1 *A word w is central if and only if it can be written as a^n , b^n , or $uabv = vbau$ for some $n \geq 0$, words u, v , and letters $a \neq b$.*
- 2 *w is central if and only if it has two coprime periods p, q such that $|w| = p + q - 2$.*

Example

abaaba is central, whereas *abba* is not.

Central Words

Central words are the palindromic prefixes of standard Sturmian words, and enjoy remarkable characterizations such as:

Theorem (de Luca, Mignosi 1994 etc.)

- 1 *A word w is central if and only if it can be written as a^n , b^n , or $uabv = vbau$ for some $n \geq 0$, words u, v , and letters $a \neq b$.*
- 2 *w is central if and only if it has two coprime periods p, q such that $|w| = p + q - 2$.*

Example

abaaba is central, whereas *abba* is not.

A Known Characterization

Theorem (D'Alessandro 2002)

A word $w \in A^$ is trapezoidal non-Sturmian if and only if*

$$w = pxux \cdot yuyq$$

where u is central, $A = \{x, y\}$, and $p, q \in A^$ are such that $pxux$ (resp. $yuyq$) has the same period as ux (resp. yu).*

Also, in such a case $R_w = |pxux|$ and $K_w = |yuyq|$.

Example

For all $n, m \geq 0$, the word $a^n(ba)^m$ is trapezoidal;
it is not Sturmian if and only if $n \geq 3$ and $m \geq 2$.

A Known Characterization

Theorem (D'Alessandro 2002)

A word $w \in A^$ is trapezoidal non-Sturmian if and only if*

$$w = pxux \cdot yuyq$$

where u is central, $A = \{x, y\}$, and $p, q \in A^$ are such that $pxux$ (resp. $yuyq$) has the same period as ux (resp. yu).*

Also, in such a case $R_w = |pxux|$ and $K_w = |yuyq|$.

Example

For all $n, m \geq 0$, the word $a^n(ba)^m$ is trapezoidal;
it is not Sturmian if and only if $n \geq 3$ and $m \geq 2$.

A Known Characterization

Theorem (D'Alessandro 2002)

A word $w \in A^$ is trapezoidal non-Sturmian if and only if*

$$w = pxux \cdot yuyq$$

where u is central, $A = \{x, y\}$, and $p, q \in A^$ are such that $pxux$ (resp. $yuyq$) has the same period as ux (resp. yu).*

Also, in such a case $R_w = |pxux|$ and $K_w = |yuyq|$.

Example

For all $n, m \geq 0$, the word $a^n(ba)^m$ is trapezoidal;
it is not Sturmian if and only if $n \geq 3$ and $m \geq 2$.

Closed vs. Open

Definition

A word is **closed** (aka periodic-like, complete return) if it has a factor that occurs exactly twice, as a prefix and as a suffix. Otherwise, it is **open**.

Example

aababbaa is closed, but *aababbaaa* is open.

The open/closed duality for prefixes (oc-sequence) has been used to study structural properties of finite and infinite words. For instance...



Closed vs. Open

Definition

A word is closed (aka periodic-like, complete return) if it has a factor that occurs exactly twice, as a prefix and as a suffix. Otherwise, it is open.

Example

aababbaa is closed, but *aababbaaa* is open.

The open/closed duality for prefixes (oc-sequence) has been used to study structural properties of finite and infinite words. For instance...



Closed vs. Open

Definition

A word is closed (aka periodic-like, complete return) if it has a factor that occurs exactly twice, as a prefix and as a suffix. Otherwise, it is open.

Example

aababbaa is closed, but *aababbaaa* is open.

The open/closed duality for prefixes (oc-sequence) has been used to study structural properties of finite and infinite words. For instance...



OC-Sequence of Sturmian Words

Theorem (DL, Fici, Zamboni 2017)

An infinite word w is standard Sturmian if and only if

$$\text{OC}_w = \prod_{n \geq 0} 1^{k_n} 0^{k_n},$$

i.e., if every run of consecutive closed prefixes is followed by an equally long run of open prefixes.

Here the n th symbol of the sequence OC_w is 1 if the prefix $w_{[n]}$ of length n is closed, and 0 otherwise.



Closed vs. Open Trapezoidal

If w is trapezoidal, then $\{L_w, H_w\} = \{R_w, K_w\}$. More precisely,

Theorem (Bucci, DL, Fici 2013)

Let w be trapezoidal. Then:

- $H_w = K_w$ and $L_w = R_w$ if w is closed;
- $H_w = R_w$ and $L_w = K_w$ if w is open.

Our first idea is to refine this further, using prefixes (i.e., the oc-sequence) and their parameters.

Closed vs. Open Trapezoidal

If w is trapezoidal, then $\{L_w, H_w\} = \{R_w, K_w\}$. More precisely,

Theorem (Bucci, DL, Fici 2013)

Let w be trapezoidal. Then:

- $H_w = K_w$ and $L_w = R_w$ if w is closed;
- $H_w = R_w$ and $L_w = K_w$ if w is open.

Our first idea is to refine this further, using prefixes (i.e., the oc-sequence) and their parameters.



Closed vs. Open Trapezoidal

If w is trapezoidal, then $\{L_w, H_w\} = \{R_w, K_w\}$. More precisely,

Theorem (Bucci, DL, Fici 2013)

Let w be trapezoidal. Then:

- $H_w = K_w$ and $L_w = R_w$ if w is closed;
- $H_w = R_w$ and $L_w = K_w$ if w is open.

Our first idea is to refine this further, using prefixes (i.e., the oc-sequence) and their parameters.



Closed vs. Open Trapezoidal

If w is trapezoidal, then $\{L_w, H_w\} = \{R_w, K_w\}$. More precisely,

Theorem (Bucci, DL, Fici 2013)

Let w be trapezoidal. Then:

- $H_w = K_w$ and $L_w = R_w$ if w is closed;
- $H_w = R_w$ and $L_w = K_w$ if w is open.

Our first idea is to refine this further, using prefixes (i.e., the oc-sequence) and their parameters.

Parameters and the OC-Sequence

Lemma

Let w be a word and x a letter. Then:

- ❶ $H_{wx} = H_w + 1$ if wx is closed, and $H_{wx} = H_w$ if wx is open;
- ❷ if wx is trapezoidal, then $L_{wx} = L_w$ whenever wx is closed, and $L_{wx} = L_w + 1$ if wx is open.

In other words, H_w is the number of closed prefixes of w , and if w is trapezoidal, L_w is the number of its open prefixes.

Corollary

Let wx be a trapezoidal word, $x \in A$. Then $K_{wx} = K_w + 1$ and $R_{wx} = R_w$, unless w is open and wx is closed or vice versa, in which case $K_{wx} = R_w + 1$ and $R_{wx} = K_w$ instead.

Parameters and the OC-Sequence

Lemma

Let w be a word and x a letter. Then:

- ❶ $H_{wx} = H_w + 1$ if wx is closed, and $H_{wx} = H_w$ if wx is open;
- ❷ if wx is trapezoidal, then $L_{wx} = L_w$ whenever wx is closed, and $L_{wx} = L_w + 1$ if wx is open.

In other words, H_w is the number of closed prefixes of w , and if w is trapezoidal, L_w is the number of its open prefixes.

Corollary

Let wx be a trapezoidal word, $x \in A$. Then $K_{wx} = K_w + 1$ and $R_{wx} = R_w$, unless w is open and wx is closed or vice versa, in which case $K_{wx} = R_w + 1$ and $R_{wx} = K_w$ instead.

Parameters and the OC-Sequence

Lemma

Let w be a word and x a letter. Then:

- ❶ $H_{wx} = H_w + 1$ if wx is closed, and $H_{wx} = H_w$ if wx is open;
- ❷ if wx is trapezoidal, then $L_{wx} = L_w$ whenever wx is closed, and $L_{wx} = L_w + 1$ if wx is open.

In other words, H_w is the number of closed prefixes of w , and if w is trapezoidal, L_w is the number of its open prefixes.

Corollary

Let wx be a trapezoidal word, $x \in A$. Then $K_{wx} = K_w + 1$ and $R_{wx} = R_w$, unless w is open and wx is closed or vice versa, in which case $K_{wx} = R_w + 1$ and $R_{wx} = K_w$ instead.

Parameters and the OC-Sequence

Lemma

Let w be a word and x a letter. Then:

- ❶ $H_{wx} = H_w + 1$ if wx is closed, and $H_{wx} = H_w$ if wx is open;
- ❷ if wx is trapezoidal, then $L_{wx} = L_w$ whenever wx is closed, and $L_{wx} = L_w + 1$ if wx is open.

In other words, H_w is the number of closed prefixes of w , and if w is trapezoidal, L_w is the number of its open prefixes.

Corollary

Let wx be a trapezoidal word, $x \in A$. Then $K_{wx} = K_w + 1$ and $R_{wx} = R_w$, unless w is open and wx is closed or vice versa, in which case $K_{wx} = R_w + 1$ and $R_{wx} = K_w$ instead.

Parameters and the OC-Sequence

Lemma

Let w be a word and x a letter. Then:

- ❶ $H_{wx} = H_w + 1$ if wx is closed, and $H_{wx} = H_w$ if wx is open;
- ❷ if wx is trapezoidal, then $L_{wx} = L_w$ whenever wx is closed, and $L_{wx} = L_w + 1$ if wx is open.

In other words, H_w is the number of closed prefixes of w , and if w is trapezoidal, L_w is the number of its open prefixes.

Corollary

Let wx be a trapezoidal word, $x \in A$. Then $K_{wx} = K_w + 1$ and $R_{wx} = R_w$, unless w is open and wx is closed or vice versa, in which case $K_{wx} = R_w + 1$ and $R_{wx} = K_w$ instead.

More About These Turning Points

Proposition

Let wx be a trapezoidal word, $x \in A$. Then

- ① if w is closed and wx is open, then $L_w < H_w$;
- ② if w is open and wx is closed, then $H_w \leq L_w$.

Example

Let $w = baabaababab$. Then $w_{[n]}$ is closed for $n = 1$ and $4 \leq n \leq 8$, while open otherwise, i.e.,

$$OC_w = 10011111000.$$

As w is trapezoidal, the difference $H - L$ increases (resp. decreases) by 1 at each closed (resp. open) prefix.

More About These Turning Points

Proposition

Let wx be a trapezoidal word, $x \in A$. Then

- ① if w is closed and wx is open, then $L_w < H_w$;
- ② if w is open and wx is closed, then $H_w \leq L_w$.

Example

Let $w = baabaababab$. Then $w_{[n]}$ is closed for $n = 1$ and $4 \leq n \leq 8$, while open otherwise, i.e.,

$$OC_w = 10011111000.$$

As w is trapezoidal, the difference $H - L$ increases (resp. decreases) by 1 at each closed (resp. open) prefix.

More About These Turning Points

Proposition

Let wx be a trapezoidal word, $x \in A$. Then

- ① if w is closed and wx is open, then $L_w < H_w$;
- ② if w is open and wx is closed, then $H_w \leq L_w$.

Example

Let $w = baabaababab$. Then $w_{[n]}$ is closed for $n = 1$ and $4 \leq n \leq 8$, while open otherwise, i.e.,

$$OC_w = 10011111000.$$

As w is trapezoidal, the difference $H - L$ increases (resp. decreases) by 1 at each closed (resp. open) prefix.

More About These Turning Points

Proposition

Let wx be a trapezoidal word, $x \in A$. Then

- ① if w is closed and wx is open, then $L_w < H_w$;
- ② if w is open and wx is closed, then $H_w \leq L_w$.

Example

Let $w = baabaababab$. Then $w_{[n]}$ is closed for $n = 1$ and $4 \leq n \leq 8$, while open otherwise, i.e.,

$$OC_w = 10011111000.$$

As w is trapezoidal, the difference $H - L$ increases (resp. decreases) by 1 at each closed (resp. open) prefix.

More About These Turning Points

Proposition

Let wx be a trapezoidal word, $x \in A$. Then

- ① if w is closed and wx is open, then $L_w < H_w$;
- ② if w is open and wx is closed, then $H_w \leq L_w$.

Example

Let $w = baabaababab$. Then $w_{[n]}$ is closed for $n = 1$ and $4 \leq n \leq 8$, while open otherwise, i.e.,

$$OC_w = 10011111000.$$

As w is trapezoidal, the difference $H - L$ increases (resp. decreases) by 1 at each closed (resp. open) prefix.

Sturmian Special Words

Theorem (de Luca, Mignosi 1994)

*A finite Sturmian word w is **strictly bispecial**, i.e., such that awa, awb, bwa, bwb are all Sturmian, if and only if it is central.*

Theorem (de Luca 1997)

w is such that wa, wb (resp. aw, bw) are both Sturmian if and only if it is a suffix (resp. prefix) of a central word.

Sturmian Special Words

Theorem (de Luca, Mignosi 1994)

A finite Sturmian word w is strictly bispecial, i.e., such that awa, awb, bwa, bwb are all Sturmian, if and only if it is central.

Theorem (de Luca 1997)

w is such that wa, wb (resp. aw, bw) are both Sturmian if and only if it is a suffix (resp. prefix) of a central word.

The Simple Bispecial Case

Theorem (Fici 2014)

$w \in A^$ is such that aw, bw, wa, wb are all Sturmian if and only if*

$$w = (uxy)^n u$$

for some central word u , $n \geq 0$, and $A = \{x, y\}$.

In particular, for $n = 1$ we obtain **semicentral words**, which can be equivalently defined by the fact that the longest

- right special factor, left special factor,
- repeated prefix, and repeated suffix

all coincide (Bucci, DL, Fici 2013).



The Simple Bispecial Case

Theorem (Fici 2014)

$w \in A^$ is such that aw, bw, wa, wb are all Sturmian if and only if*

$$w = (uxy)^n u$$

for some central word u , $n \geq 0$, and $A = \{x, y\}$.

In particular, for $n = 1$ we obtain **semicentral words**, which can be equivalently defined by the fact that the longest

- right special factor, left special factor,
- repeated prefix, and repeated suffix

all coincide (Bucci, DL, Fici 2013).



Right (or Left) Special Trapezoidal Words

Theorem

A trapezoidal word w is right special^a if and only if

- w is a suffix of a central word, or*
- $w = pxuxyu$ for a central word u , letters $x \neq y$, and a word p such that $pxux$ has the same period as ux .*

^athat is, such that wa, wb are both trapezoidal

A symmetrical characterization holds for left special trapezoidal words.

Right (or Left) Special Trapezoidal Words

Theorem

A trapezoidal word w is right special^a if and only if

- w is a suffix of a central word, or*
- $w = pxuxyu$ for a central word u , letters $x \neq y$, and a word p such that $pxux$ has the same period as ux .*

^athat is, such that wa, wb are both trapezoidal

A symmetrical characterization holds for left special trapezoidal words.

Strictly Bispecial Ones

Our last main result extends de Luca and Mignosi's characterization.

Theorem

A trapezoidal word w is strictly bispecial if and only if it is central or semicentral.



Further Work

- Characterize the oc-sequences of trapezoidal words, or even just of non-standard Sturmian words;
- Extend Fici's characterization of bispecial Sturmian words to the trapezoidal case.

Thank You



Further Work

- Characterize the oc-sequences of trapezoidal words, or even just of non-standard Sturmian words;
- Extend Fici's characterization of bispecial Sturmian words to the trapezoidal case.

Thank You



Further Work

- Characterize the oc-sequences of trapezoidal words, or even just of non-standard Sturmian words;
- Extend Fici's characterization of bispecial Sturmian words to the trapezoidal case.

Thank You



Further Work

- Characterize the oc-sequences of trapezoidal words, or even just of non-standard Sturmian words;
- Extend Fici's characterization of bispecial Sturmian words to the trapezoidal case.

Thank You