On a generalisation of trapezoidal words

Amy Glen

School of Chemical & Mathematical Sciences
Murdoch University, Perth, Australia

amy.glen@gmail.com
http://amyglen.wordpress.com

Workshop
“Outstanding Challenges in Combinatorics on Words”
BIRS, Banff Centre, Canada
February 23, 2012
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).

Given a finite or infinite word \( w \), let \( C_w(n) \) denote the factor complexity function of \( w \), which counts the number of distinct factors of \( w \) of each length \( n \geq 0 \).
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).

Given a finite or infinite word $w$, let $C_w(n)$ denote the factor complexity function of $w$, which counts the number of distinct factors of $w$ of each length $n \geq 0$.

Amongst many interesting things, de Luca proved the following result.

**Theorem (de Luca 1999)**

If $w$ is a finite Sturmian word, then the graph of its complexity $C_w(n)$ as a function of $n$ (for $0 \leq n \leq |w|$) is that of a regular trapezoid (or possibly an isosceles triangle).
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).

Given a finite or infinite word $w$, let $C_w(n)$ denote the factor complexity function of $w$, which counts the number of distinct factors of $w$ of each length $n \geq 0$.

Amongst many interesting things, de Luca proved the following result.

**Theorem (de Luca 1999)**

If $w$ is a finite Sturmian word, then the graph of its complexity $C_w(n)$ as a function of $n$ (for $0 \leq n \leq |w|$) is that of a regular trapezoid (or possibly an isosceles triangle).

That is:

- $C_w(n)$ increases by 1 with each $n$ on some interval of length $r$.  

Amy Glen (MU, Perth)
On a generalisation of trapezoidal words
February 2012
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).

Given a finite or infinite word $w$, let $C_w(n)$ denote the factor complexity function of $w$, which counts the number of distinct factors of $w$ of each length $n \geq 0$.

Amongst many interesting things, de Luca proved the following result.

**Theorem (de Luca 1999)**

If $w$ is a finite Sturmian word, then the graph of its complexity $C_w(n)$ as a function of $n$ (for $0 \leq n \leq |w|$) is that of a regular trapezoid (or possibly an isosceles triangle).

That is:

- $C_w(n)$ increases by 1 with each $n$ on some interval of length $r$.
- Then $C_w(n)$ is constant on some interval of length $s$. 
Trapezoidal Words

Trapezoidal words were first introduced by de Luca (1999) when studying the behaviour of the factor complexity of finite Sturmian words (i.e., finite factors of infinite Sturmian words).

Given a finite or infinite word $w$, let $C_w(n)$ denote the factor complexity function of $w$, which counts the number of distinct factors of $w$ of each length $n \geq 0$.

Amongst many interesting things, de Luca proved the following result.

**Theorem (de Luca 1999)**

If $w$ is a finite Sturmian word, then the graph of its complexity $C_w(n)$ as a function of $n$ (for $0 \leq n \leq |w|$) is that of a regular trapezoid (or possibly an isosceles triangle).

That is:

- $C_w(n)$ increases by 1 with each $n$ on some interval of length $r$.
- Then $C_w(n)$ is constant on some interval of length $s$.
- Finally $C_w(n)$ decreases by 1 with each $n$ on an interval of length $r$. 
Example

Graph of the factor complexity of the finite Sturmian word \textit{aabaaabab}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{factor_complexity_graph}
\end{figure}
This “trapezoidal property” does not characterise Sturmian words.
This “trapezoidal property” does not characterise Sturmian words. For example, $aaabb$ is trapezoidal ($[1, 2, 3, 3, 2, 1]$), but not Sturmian.
Trapezoidal Words . . .

- This “trapezoidal property” does not characterise Sturmian words. For example, \textit{aaabb} is trapezoidal ([1, 2, 3, 3, 2, 1]), but not Sturmian.

- Note: If $w$ is a \textit{trapezoidal word} (i.e., its “complexity graph” is a regular trapezoid on the interval $[0, |w|]$), then necessarily $C_w(1) = 2$. 

This “trapezoidal property” does not characterise Sturmian words.
For example, $aaabb$ is trapezoidal ($[1, 2, 3, 3, 2, 1]$), but not Sturmian.

Note: If $w$ is a trapezoidal word (i.e., its “complexity graph” is a regular trapezoid on the interval $[0, |w|]$), then necessarily $C_w(1) = 2$. This is because there is 1 factor of length 0, namely the empty word $\varepsilon$. 

This “trapezoidal property” does not characterise Sturmian words. For example, $\text{aaabb}$ is trapezoidal ($[1, 2, 3, 3, 2, 1]$), but not Sturmian.

Note: If $w$ is a trapezoidal word (i.e., its “complexity graph” is a regular trapezoid on the interval $[0, |w|]$), then necessarily $C_w(1) = 2$. This is because there is 1 factor of length 0, namely the empty word $\varepsilon$.

So any trapezoidal word is on a binary alphabet and the family of trapezoidal words properly contains all finite Sturmian words.
This “trapezoidal property” does not characterise Sturmian words. For example, $aaabb$ is trapezoidal ($[1, 2, 3, 3, 2, 1]$), but not Sturmian.

Note: If $w$ is a trapezoidal word (i.e., its “complexity graph” is a regular trapezoid on the interval $[0, |w|]$), then necessarily $C_w(1) = 2$. This is because there is 1 factor of length 0, namely the empty word $\varepsilon$.

So any trapezoidal word is on a binary alphabet and the family of trapezoidal words properly contains all finite Sturmian words.

Generalised Trapezoidal Words

We say that a finite word $w$ with $\text{Alph}(w) = A$ ($|A| \geq 2$) is a generalised trapezoidal word (or GT-word for short) if there exist positive integers $m, M$ with $m \leq M$ such that the factor complexity function $C_w(n)$ of $w$ increases by 1 for each $n$ in the interval $[1, m]$, is constant for each $n$ in the interval $[m, M]$, and decreases by 1 for each $n$ in the interval $[M, |w|]$. 
We say that a finite word \( w \) with \( \text{Alph}(w) = \mathcal{A} \) (\(|\mathcal{A}| \geq 2\)) is a **generalised trapezoidal word** (or **GT-word** for short) if there exist positive integers \( m, M \) with \( m \leq M \) such that the factor complexity function \( C_w(n) \) of \( w \) increases by 1 for each \( n \) in the interval \([1, m]\), is constant for each \( n \) in the interval \([m, M]\), and decreases by 1 for each \( n \) in the interval \([M, |w|]\).
Generalised Trapezoidal Words

We say that a finite word $w$ with $\text{Alph}(w) = A$ ($|A| \geq 2$) is a **generalised trapezoidal word** (or **GT-word** for short) if there exist positive integers $m, M$ with $m \leq M$ such that the factor complexity function $C_w(n)$ of $w$ increases by 1 for each $n$ in the interval $[1, m]$, is constant for each $n$ in the interval $[m, M]$, and decreases by 1 for each $n$ in the interval $[M, |w|]$.

So a finite word $w$ consisting of at least two distinct letters is a GT-word if the graph of its factor complexity $C_w(n)$ as a function of $n$ ($0 \leq n \leq |w|$) is either constant or a regular trapezoid (possibly an isosceles triangle when $m = M$) on the interval $[1, |w| - |A| + 1]$. 

Amy Glen (MU, Perth)

On a generalisation of trapezoidal words

February 2012
Generalised Trapezoidal Words

We say that a finite word $w$ with $\text{Alph}(w) = A$ ($|A| \geq 2$) is a **generalised trapezoidal word** (or **GT-word** for short) if there exist positive integers $m, M$ with $m \leq M$ such that the factor complexity function $C_w(n)$ of $w$ increases by 1 for each $n$ in the interval $[1, m]$, is constant for each $n$ in the interval $[m, M]$, and decreases by 1 for each $n$ in the interval $[M, |w|]$.

So a finite word $w$ consisting of at least two distinct letters is a GT-word if the graph of its factor complexity $C_w(n)$ as a function of $n$ ($0 \leq n \leq |w|$) is either constant or a regular trapezoid (possibly an isosceles triangle when $m = M$) on the interval $[1, |w| - |A| + 1]$.
Generalised Trapezoidal Words

We say that a finite word $w$ with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$) is a **generalised trapezoidal word** (or GT-word for short) if there exist positive integers $m$, $M$ with $m \leq M$ such that the factor complexity function $C_w(n)$ of $w$ increases by 1 for each $n$ in the interval $[1, m]$, is constant for each $n$ in the interval $[m, M]$, and decreases by 1 for each $n$ in the interval $[M, |w|]$.

So a finite word $w$ consisting of at least two distinct letters is a GT-word if the graph of its factor complexity $C_w(n)$ as a function of $n$ ($0 \leq n \leq |w|$) is either constant or a regular trapezoid (possibly an isosceles triangle when $m = M$) on the interval $[1, |w| - |\mathcal{A}| + 1]$.

Clearly these words coincide with the (original) trapezoidal words when $|\mathcal{A}| = 2$. 
Some Examples

Length 10 over $\mathcal{A} = \{a, b, c\}$

<table>
<thead>
<tr>
<th>GT-word</th>
<th>$C(n)$ for $n = 0, 1, 2, \ldots, 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaaaaaaaaabc</td>
<td>1, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 1</td>
</tr>
<tr>
<td>abcabcabcabc</td>
<td>1, 3, 4, 4, 4, 4, 4, 4, 3, 2, 1</td>
</tr>
<tr>
<td>abcabcabcabcabc</td>
<td>1, 3, 4, 5, 5, 5, 5, 4, 3, 2, 1</td>
</tr>
<tr>
<td>abcabcabcabcabc</td>
<td>1, 3, 4, 5, 6, 6, 5, 4, 3, 2, 1</td>
</tr>
</tbody>
</table>
Some Examples

**Length 10 over** $A = \{a, b, c\}$

<table>
<thead>
<tr>
<th>GT-word</th>
<th>$C(n)$ for $n = 0, 1, 2, \ldots, 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaaaaaaaabc</td>
<td>1, 3, 3, 3, 3, 3, 3, 3, 3, 3, 2, 1</td>
</tr>
<tr>
<td>abcbcbecbca</td>
<td>1, 3, 4, 4, 4, 4, 4, 4, 4, 3, 2, 1</td>
</tr>
<tr>
<td>abcbcabcba</td>
<td>1, 3, 4, 5, 5, 5, 5, 4, 3, 2, 1</td>
</tr>
<tr>
<td>abcabcabcab</td>
<td>1, 3, 4, 5, 6, 6, 5, 4, 3, 2, 1</td>
</tr>
</tbody>
</table>

**Length 8 over** $A = \{a, b, c, d\}$

<table>
<thead>
<tr>
<th>GT-word</th>
<th>$C(n)$ for $n = 0, 1, 2, \ldots, 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>aaaaaabcd</td>
<td>1, 4, 4, 4, 4, 4, 4, 3, 2, 1</td>
</tr>
<tr>
<td>aaaaabcd</td>
<td>1, 4, 5, 5, 5, 4, 3, 2, 1</td>
</tr>
<tr>
<td>aaabcdab</td>
<td>1, 4, 5, 6, 5, 4, 3, 2, 1</td>
</tr>
</tbody>
</table>
Preliminary Results

Suppose $w$ is a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$).
Preliminary Results

- Suppose \( w \) is a finite word with \( \text{Alph}(w) = A \ (|A| \geq 2) \).

- Let \( R_w (= R) \) denote the smallest positive integer \( p \) such that \( w \) has no right special factor of length \( p \).
Preliminary Results

Suppose $w$ is a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$).

Let $R_w (= R)$ denote the smallest positive integer $p$ such that $w$ has no right special factor of length $p$.

Let $K_w (= K)$ denote the length of the shortest unrepeated suffix of $w$. 
Suppose $w$ is a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$).

Let $R_w (= R)$ denote the smallest positive integer $p$ such that $w$ has no right special factor of length $p$.

Let $K_w (= K)$ denote the length of the shortest unrepeated suffix of $w$.

The graph of the complexity function of any given word $w$ can be described as follows.
Preliminary Results

- Suppose $w$ is a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$).

- Let $R_w$ ($= R$) denote the smallest positive integer $p$ such that $w$ has no right special factor of length $p$.

- Let $K_w$ ($= K$) denote the length of the shortest unrepeated suffix of $w$.

The graph of the complexity function of any given word $w$ can be described as follows.

**Theorem (de Luca 1999)**

Let $w$ be a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$) and let $m = \min\{R_w, K_w\}$, $M = \max\{R_w, K_w\}$.
Preliminary Results

- Suppose \( w \) is a finite word with \( \text{Alph}(w) = A \ (|A| \geq 2) \).
- Let \( R_w (= R) \) denote the smallest positive integer \( p \) such that \( w \) has no right special factor of length \( p \).
- Let \( K_w (= K) \) denote the length of the shortest unrepeated suffix of \( w \).

The graph of the complexity function of any given word \( w \) can be described as follows.

**Theorem (de Luca 1999)**

Let \( w \) be a finite word with \( \text{Alph}(w) = A \ (|A| \geq 2) \) and let \( m = \min\{R_w, K_w\} \), \( M = \max\{R_w, K_w\} \).

The factor complexity function \( C_w \) of \( w \) is strictly increasing on the interval \([0, m]\), is non-decreasing on the interval \([m, M]\), and strictly decreasing on the interval \([M, |w|]\).
Preliminary Results

- Suppose \( w \) is a finite word with \( \text{Alph}(w) = A \ (|A| \geq 2) \).
- Let \( R_w \ (= R) \) denote the smallest positive integer \( p \) such that \( w \) has no right special factor of length \( p \).
- Let \( K_w \ (= K) \) denote the length of the shortest unrepeated suffix of \( w \).

The graph of the complexity function of any given word \( w \) can be described as follows.

Theorem (de Luca 1999)

Let \( w \) be a finite word with \( \text{Alph}(w) = A \ (|A| \geq 2) \) and let \( m = \min\{R_w, K_w\} \), \( M = \max\{R_w, K_w\} \).

The factor complexity function \( C_w \) of \( w \) is strictly increasing on the interval \([0, m]\), is non-decreasing on the interval \([m, M]\), and strictly decreasing on the interval \([M, |w|]\).

Moreover, for \( n \in [M, |w|] \), one has \( C_w(n + 1) = C_w(n) - 1 \).
Preliminary Results

- Suppose $w$ is a finite word with $\text{Alph}(w) = A$ ($|A| \geq 2$).
- Let $R_w (= R)$ denote the smallest positive integer $p$ such that $w$ has no right special factor of length $p$.
- Let $K_w (= K)$ denote the length of the shortest unrepeated suffix of $w$.

The graph of the complexity function of any given word $w$ can be described as follows.

**Theorem (de Luca 1999)**

Let $w$ be a finite word with $\text{Alph}(w) = A$ ($|A| \geq 2$) and let $m = \min\{R_w, K_w\}$, $M = \max\{R_w, K_w\}$.

The factor complexity function $C_w$ of $w$ is strictly increasing on the interval $[0, m]$, is non-decreasing on the interval $[m, M]$, and strictly decreasing on the interval $[M, |w|]$.

Moreover, for $n \in [M, |w|]$, one has $C_w(n + 1) = C_w(n) - 1$.

If $R_w < K_w$, then $C_w$ is constant on the interval $[m, M]$.
de Luca (1999) proved that a finite word $w$ is a (binary) trapezoidal word if and only if $|w| = R_w + K_w$. 
de Luca (1999) proved that a finite word $w$ is a (binary) trapezoidal word if and only if $|w| = R_w + K_w$.

**Example**

The binary word $w = aaabb$ is trapezoidal with “complexity sequence” $[1, 2, 3, 3, 2, 1]$ and we see that $R_w = 3$, $K_w = 2$, and $|w| = 5 = R_w + K_w$. 
de Luca (1999) proved that a finite word $w$ is a (binary) trapezoidal word if and only if $|w| = R_w + K_w$.

**Example**

The binary word $w = aaabb$ is trapezoidal with “complexity sequence” $[1, 2, 3, 3, 2, 1]$ and we see that $R_w = 3$, $K_w = 2$, and $|w| = 5 = R_w + K_w$.

Is there a similar combinatorial characterisation for GT-words?
de Luca (1999) proved that a finite word \( w \) is a (binary) trapezoidal word if and only if \(|w| = R_w + K_w\).

**Example**

The binary word \( w = aaabb \) is trapezoidal with “complexity sequence” \([1, 2, 3, 3, 2, 1]\) and we see that \( R_w = 3, K_w = 2, \) and \(|w| = 5 = R_w + K_w\).

Is there a similar combinatorial characterisation for GT-words?

You might guess, for instance, that GT-words are precisely those words \( w \) that satisfy the condition \(|w| = R_w + K_w + |A| - 2\).
Preliminary Results . . .

de Luca (1999) proved that a finite word $w$ is a (binary) trapezoidal word if and only if $|w| = R_w + K_w$.

Example

The binary word $w = aaabb$ is trapezoidal with “complexity sequence” $[1, 2, 3, 3, 2, 1]$ and we see that $R_w = 3$, $K_w = 2$, and $|w| = 5 = R_w + K_w$.

Is there a similar combinatorial characterisation for GT-words?

You might guess, for instance, that GT-words are precisely those words $w$ that satisfy the condition $|w| = R_w + K_w + |A| - 2$.

Whilst it is true that any word satisfying this “$RK$-condition” is a GT-word, the converse does not hold.
de Luca (1999) proved that a finite word $w$ is a (binary) trapezoidal word if and only if $|w| = R_w + K_w$.

**Example**

The binary word $w = aaabb$ is trapezoidal with “complexity sequence” [1, 2, 3, 3, 2, 1] and we see that $R_w = 3$, $K_w = 2$, and $|w| = 5 = R_w + K_w$.

Is there a similar combinatorial characterisation for GT-words?

You might guess, for instance, that GT-words are precisely those words $w$ that satisfy the condition $|w| = R_w + K_w + |A| - 2$.

Whilst it is **true** that any word satisfying this “$RK$-condition” is a GT-word, the converse does not hold.

**Example**

The GT-word $ababada$ of length 8 with comp. seq. [1, 4, 5, 5, 5, 4, 3, 2, 1] has $R = 4$ and $K = 1$, but $R + K + 2 \neq 8$, so this GT-word does not satisfy the $RK$-condition.
GT-words satisfying the $RK$-condition

As a first step towards obtaining a combinatorial characterisation of generalised trapezoidal words, we have the following characterisation of finite words $w$ satisfying the condition $|w| = R_w + K_w + |A| - 2$. 
GT-words satisfying the $RK$-condition

As a first step towards obtaining a combinatorial characterisation of generalised trapezoidal words, we have the following characterisation of finite words $w$ satisfying the condition $|w| = R_w + K_w + |A| - 2$.

**Theorem**

A finite word $w$ with $\text{Alph}(w) = A$ ($|A| \geq 2$) satisfies $|w| = R_w + K_w + |A| - 2$ if and only if the factor complexity of $w$ satisfies:

$C_w(0) = 1$,
$C_w(1) = |A|$,
$C_w(i) = C_w(i - 1) + 1$ for $2 \leq i \leq m$,
$C_w(i + 1) = C_w(i)$ for $m \leq i \leq M - 1$,
$C_w(i + 1) = C_w(i) - 1$ for $M \leq i \leq |w|$

where $m = \min\{K_w, R_w\}$ and $M = \max\{K_w, R_w\}$. 
GT-words satisfying the \( R K \)-condition

As a first step towards obtaining a combinatorial characterisation of generalised trapezoidal words, we have the following characterisation of finite words \( w \) satisfying the condition \( |w| = R_w + K_w + |A| - 2 \).

**Theorem**

A finite word \( w \) with \( \text{Alph}(w) = A \) \((|A| \geq 2)\) satisfies \( |w| = R_w + K_w + |A| - 2 \) if and only if the factor complexity of \( w \) satisfies:

\[
\begin{align*}
C_w(0) &= 1, \\
C_w(1) &= |A|, \\
C_w(i) &= C_w(i - 1) + 1 & \text{for } 2 \leq i \leq m, \\
C_w(i + 1) &= C_w(i) & \text{for } m \leq i \leq M - 1, \\
C_w(i + 1) &= C_w(i) - 1 & \text{for } M \leq i \leq |w|
\end{align*}
\]

where \( m = \min\{K_w, R_w\} \) and \( M = \max\{K_w, R_w\} \).

**Corollary**

Let \( w \) be finite word with \( \text{Alph}(w) = A, \ |A| \geq 2. \) If \( |w| = R_w + K_w + |A| - 2 \), then \( w \) is a GT-word.
Proposition

Let \( w \) be a finite word with \( \text{Alph}(w) = A, \ |A| \geq 2 \).
If \( |w| = R_w + K_w + |A| - 2 \), then each factor \( u \) of \( w \) satisfies
\[ |u| = R_u + K_u + |\text{Alph}(u)| - 2. \]
**Proposition**

Let $w$ be a finite word with $\text{Alph}(w) = \mathcal{A}$, $|\mathcal{A}| \geq 2$.
If $|w| = R_w + K_w + |\mathcal{A}| - 2$, then each factor $u$ of $w$ satisfies
$|u| = R_u + K_u + |\text{Alph}(u)| - 2$.

So the language of all words $w$ satisfying the $RK$-condition is closed.
**Proposition**

Let \( w \) be a finite word with \( \text{Alph}(w) = \mathcal{A} \), \( |\mathcal{A}| \geq 2 \).

If \( |w| = R_w + K_w + |\mathcal{A}| - 2 \), then each factor \( u \) of \( w \) satisfies 
\[
|u| = R_u + K_u + |\text{Alph}(u)| - 2.
\]

So the language of all words \( w \) satisfying the \( RK \)-condition is closed.

However, the language of all such words is not closed under reversal.
Preliminary Results

**Proposition**

Let \( w \) be a finite word with \( \text{Alph}(w) = \mathcal{A}, |\mathcal{A}| \geq 2 \).

If \( |w| = R_w + K_w + |\mathcal{A}| - 2 \), then each factor \( u \) of \( w \) satisfies
\[
|u| = R_u + K_u + |\text{Alph}(u)| - 2.
\]

So the language of all words \( w \) satisfying the \( RK \)-condition is closed.

However, the language of all such words is not closed under reversal.

For example, \( abbcc \) satisfies the \( RK \)-condition, but its reversal \( ccba \) does not since it has \( R = 2 \) and \( K = 1 \).
Some More Basic Properties

The language of all GT-words is closed . . .

**Theorem**

If $w$ is a GT-word, then each factor of $w$ (containing at least two different letters) is also a GT-word.
Some More Basic Properties

The language of all GT-words is closed . . .

**Theorem**

If $w$ is a GT-word, then each factor of $w$ (containing at least two different letters) is also a GT-word.

Moreover, the language of all GT-words is closed under reversal.

**Theorem**

A finite word $w$ is a GT-word if and only if its reversal is a GT-word.
A Combinatorial Characterisation

Theorem

Let $w$ be a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$) and suppose $p$ is the longest prefix of $w$ such that $K_p \neq 1$. Then $w$ is a GT-word if and only if

$$|w| = R_p + K_p + |\mathcal{A}| - 2.$$
A Combinatorial Characterisation

Theorem

Let \( w \) be a finite word with \( \text{Alph}(w) = \mathcal{A} (|\mathcal{A}| \geq 2) \) and suppose \( p \) is the longest prefix of \( w \) such that \( K_p \neq 1 \). Then \( w \) is a GT-word if and only if

\[
|w| = R_p + K_p + |\mathcal{A}| - 2.
\]

Example

Recall that the GT-word \( w = ababadac \) does not satisfy the \( RK \)-condition.
A Combinatorial Characterisation

Theorem

Let \( w \) be a finite word with \( \text{Alph}(w) = \mathcal{A} \ (|\mathcal{A}| \geq 2) \) and suppose \( p \) is the longest prefix of \( w \) such that \( K_p \neq 1 \). Then \( w \) is a GT-word if and only if

\[
|w| = R_p + K_p + |\mathcal{A}| - 2.
\]

Example

Recall that the GT-word \( w = ababadac \) does not satisfy the \( RK \)-condition, but it does indeed satisfy the condition

\[
|w| = R_p + K_p + |\mathcal{A}| - 2.
\]
**Theorem**

Let $w$ be a finite word with $\text{Alph}(w) = \mathcal{A}$ ($|\mathcal{A}| \geq 2$) and suppose $p$ is the longest prefix of $w$ such that $K_p \neq 1$. Then $w$ is a GT-word if and only if

$$|w| = R_p + K_p + |\mathcal{A}| - 2.$$ 

**Example**

Recall that the GT-word $w = ababadac$ does not satisfy the $RK$-condition, but it does indeed satisfy the condition

$$|w| = R_p + K_p + |\mathcal{A}| - 2$$

since $p = ababada$ with $R_p = 4$, $K_p = 2$, and $|w| = 4 + 2 + 2 = 8$. 
Back to the Binary Case

In the case when $|\mathcal{A}| = 2$, we have proved the following.

**Theorem** (de Luca-G.-Zamboni 2008)

Let $w$ be a binary palindrome. Then $w$ is trapezoidal if and only if $w$ is Sturmian.
Back to the Binary Case

In the case when $|A| = 2$, we have proved the following.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a **binary palindrome**. Then $w$ is trapezoidal if and only if $w$ is Sturmian.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a **binary trapezoidal word**. Then $w$ contains $|w| + 1$ distinct palindromes (including $\varepsilon$).
Back to the Binary Case

In the case when $|\mathcal{A}| = 2$, we have proved the following.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a **binary palindrome**. Then $w$ is trapezoidal if and only if $w$ is Sturmian.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a **binary trapezoidal word**. Then $w$ contains $|w| + 1$ distinct palindromes (including $\varepsilon$).

That is, binary trapezoidal words (and hence finite Sturmian words) are “rich” in palindromes in the sense that they contain the maximum number of distinct palindromic factors, according to the following result.
Back to the Binary Case

In the case when $|\mathcal{A}| = 2$, we have proved the following.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a binary palindrome. Then $w$ is trapezoidal if and only if $w$ is Sturmian.

**Theorem (de Luca-G.-Zamboni 2008)**

Let $w$ be a binary trapezoidal word. Then $w$ contains $|w| + 1$ distinct palindromes (including $\varepsilon$).

That is, binary trapezoidal words (and hence finite Sturmian words) are “rich” in palindromes in the sense that they contain the maximum number of distinct palindromic factors, according to the following result.

**Theorem (Droubay-Justin-Pirillo 2001)**

A finite word $w$ contains at most $|w| + 1$ distinct palindromes (including $\varepsilon$).
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)
A finite word $w$ is said to be *rich* if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:
- $abac$ is rich, whereas $abca$ is *not* rich.
- The word *rich* is rich ... and *poor* is rich too!
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
- The word $rich$ is rich . . . and $poor$ is rich too!
- Any binary trapezoidal word is rich, but not conversely.
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)
A finite word $w$ is said to be \textit{rich} if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:
- $abac$ is rich, whereas $abca$ is \textbf{not} rich.
- The word \textit{rich} is rich . . . and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.
  
  E.g., $aabbbaa$ is rich, but not trapezoidal
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
- The word $rich$ is rich . . . and $poor$ is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., $aabbaa$ is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
- The word $rich$ is rich . . . and $poor$ is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., $aabbaa$ is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be *rich* if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
- The word *rich* is rich . . . and *poor* is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., $aabbbaa$ is rich, but not trapezoidal ($C(1) = 2, C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: $abaabaaabaaaabaaaab \cdots$
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word \( w \) is said to be \textit{rich} if \( w \) contains exactly \(|w| + 1\) distinct palindromes (including \( \varepsilon \)).

Examples:

- \( abac \) is rich, whereas \( abca \) is \textbf{not} rich.
- The word \textit{rich} is rich . . . and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.

  E.g., \( aabbaa \) is rich, but not trapezoidal \((C(1) = 2, C(2) = 4)\)

Roughly speaking, a finite or infinite word is rich if and only if a \textbf{new} palindrome is introduced at each new position.

Example: \( abaabaabaaaabaaaaab \cdots \)
Rich Words

**Definition (G.-Justin-Widmer-Zamboni 2009)**

A finite word \( w \) is said to be *rich* if \( w \) contains exactly \( |w| + 1 \) distinct palindromes (including \( \varepsilon \)).

**Examples:**

- \( abac \) is rich, whereas \( abca \) is **not** rich.
- The word *rich* is rich . . . and *poor* is rich too!
- Any binary trapezoidal word is rich, but not conversely.

  E.g., \( aabbaa \) is rich, but not trapezoidal \( (C(1) = 2, \ C(2) = 4) \)

Roughly speaking, a finite or infinite word is rich if and only if a **new** palindrome is introduced at each new position.

**Example:** \( ababaaabaaabaaabaaabab \cdots \)
Rich Words

**Definition (G.-Justin-Widmer-Zamboni 2009)**

A finite word $w$ is said to be *rich* if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

**Examples:**

- $abac$ is rich, whereas $abca$ is not rich.
- The word *rich* is rich ... and *poor* is rich too!
- Any binary trapezoidal word is rich, but not conversely.

  E.g., $aabbaa$ is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

**Example:** $abaabaaabaaabaaabaaabaaabaaabab \cdots$
Rich Words

**Definition (G.-Justin-Widmer-Zamboni 2009)**

A finite word $w$ is said to be *rich* if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

**Examples:**

- *abac* is rich, whereas *abca* is not rich.
- The word *rich* is rich . . . and *poor* is rich too!
- Any binary trapezoidal word is rich, but not conversely.
  
  E.g., *aabbaa* is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

**Example:** *abaaabaaabaaabaaaaab* · · ·
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word \( w \) is said to be \textit{rich} if \( w \) contains exactly \( |w| + 1 \) distinct palindromes (including \( \varepsilon \)).

Examples:

- \( abac \) is rich, whereas \( abca \) is not rich.
- The word \textit{rich} is rich . . . and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., \( aabbaa \) is rich, but not trapezoidal (\( C(1) = 2, \ C(2) = 4 \))

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: \( \underline{abaab}aaabaaabaaaaaab \cdots \)
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word \( w \) is said to be \textit{rich} if \( w \) contains exactly \(|w| + 1\) distinct palindromes (including \( \varepsilon \)).

Examples:

- \( abac \) is rich, whereas \( abca \) is \textbf{not} rich.
- The word \textit{rich} is rich . . . and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.

  E.g., \( aabbaa \) is rich, but not trapezoidal \((C(1) = 2, \ C(2) = 4)\)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: \textcolor{green}{abaaba}aabaaaabaaaaaab\ldots
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- \textit{abac} is rich, whereas \textit{abca} is not rich.
- The word \textit{rich} is rich \ldots and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., \textit{aabbaa} is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: \textit{ababaaabaaaaaab} \ldots
Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word \( w \) is said to be \textit{rich} if \( w \) contains exactly \(|w| + 1 \) distinct palindromes (including \( \varepsilon \)).

Examples:

- \( abac \) is rich, whereas \( abca \) is \textbf{not} rich.
- The word \textit{rich} is rich . . . and \textit{poor} is rich too!
- Any binary trapezoidal word is rich, but not conversely.
  
  E.g., \( aabbaa \) is rich, but not trapezoidal \((C(1) = 2, C(2) = 4)\)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: \( abaabaaaabaaaabaaaab \cdots \)
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be rich if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- $abac$ is rich, whereas $abca$ is not rich.
- The word $rich$ is rich . . . and $poor$ is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., $aabbaa$ is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: $abaa_{\text{baaab}}aaaabaaaabab\cdots$
Rich Words

Definition (G.-Justin-Widmer-Zamboni 2009)

A finite word $w$ is said to be *rich* if $w$ contains exactly $|w| + 1$ distinct palindromes (including $\varepsilon$).

Examples:

- *abac* is rich, whereas *abca* is not rich.
- The word *rich* is rich . . . and *poor* is rich too!
- Any binary trapezoidal word is rich, but not conversely.

E.g., *aabbaa* is rich, but not trapezoidal ($C(1) = 2$, $C(2) = 4$)

Roughly speaking, a finite or infinite word is rich if and only if a new palindrome is introduced at each new position.

Example: *abaabaaabaaaaaabaaaaaab* · · ·
Here are some other characteristic properties of rich words that were previously established by Droubay-Justin-Pirillo (2001) and G.-Justin-Widmer-Zamboni (2009).

**Characteristic Properties of Rich Words**

For any finite or infinite word $w$, the following conditions are equivalent:

i) $w$ is rich;

ii) every prefix of $w$ has a unioccurrent palindromic suffix (and equivalently, when $w$ is finite, every suffix of $w$ has a unioccurrent palindromic prefix);

iii) for each factor $u$ of $w$, every prefix (resp. suffix) of $u$ has a unioccurrent palindromic suffix (resp. prefix);

iv) for each palindromic factor $p$ of $w$, every complete return to $p$ in $w$ is a palindrome.
Richness & GT-words when $|\mathcal{A}| \geq 3$

Unlike in the binary case ($|\mathcal{A}| = 2$), not all GT-words are palindromic-rich.
Richness & GT-words when $|\mathcal{A}| \geq 3$

Unlike in the binary case ($|\mathcal{A}| = 2$), not all GT-words are palindromic-rich.

**Example**

The GT-word $ababadbc$ is not rich since it contains a non-palindromic complete return to $b$, namely $badb$. 
Richness & GT-words when $|A| \geq 3$

Unlike in the binary case ($|A| = 2$), **not** all GT-words are palindromic-rich.

**Example**

The GT-word $ababadbc$ is not rich since it contains a non-palindromic complete return to $b$, namely $badb$.

However, **all palindromic GT-words are rich** by the following more general result.
Richness & GT-words when $|\mathcal{A}| \geq 3$

Unlike in the binary case ($|\mathcal{A}| = 2$), **not** all GT-words are palindromic-rich.

**Example**

The GT-word $abavadbc$ is not rich since it contains a non-palindromic complete return to $b$, namely $badb$.

However, **all palindromic GT-words are rich** by the following more general result.

**Theorem**

Suppose $w$ is a GT-word and let $v$ denote the unique factor of $w$ such that $w = bve$ where $b$ is the longest (possibly empty) prefix of $w$ such that $|w|_x = 1$ for each $x \in \text{Alph}(b)$ and $e$ is the longest (possibly empty) suffix of $w$ such that $|w|_x = 1$ for each $x \in \text{Alph}(e)$.

If $v$ is a palindrome, then $w$ is rich.
Examples

- The GT-word $w = abacabade$ has $v = abacaba$ (a palindrome) and $w$ is indeed rich.
Examples

- The GT-word \( w = abacabade \) has \( v = abacaba \) (a palindrome) and \( w \) is indeed rich.

- The converse of the theorem does not hold.
Examples

- The GT-word $w = abacabade$ has $v = abacaba$ (a palindrome) and $w$ is indeed rich.

- The converse of the theorem does not hold. For example, the GT-word $ababadac$ is rich, but the corresponding $v$ is $ababada$ (non-palindromic).
Thank You!