<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Prefix Normal Words</td>
</tr>
<tr>
<td>2</td>
<td>Least Number of Palindromes in an Infinite Word</td>
</tr>
<tr>
<td>3</td>
<td>Bispecial Sturmian Words</td>
</tr>
<tr>
<td>4</td>
<td>Trapezoidal Words</td>
</tr>
<tr>
<td>5</td>
<td>The Array of Open and Closed Prefixes</td>
</tr>
<tr>
<td>6</td>
<td>Abelian Repetitions in Sturmian Words</td>
</tr>
<tr>
<td>7</td>
<td>Universal Lyndon Words</td>
</tr>
<tr>
<td>8</td>
<td>Cyclic Complexity</td>
</tr>
</tbody>
</table>
1. Prefix Normal Words
2. Least Number of Palindromes in an Infinite Word
3. Bispecial Sturmian Words
4. Trapezoidal Words
5. The Array of Open and Closed Prefixes
6. Abelian Repetitions in Sturmian Words
7. Universal Lyndon Words
8. Cyclic Complexity
Definition

A prefix normal word with respect to 1 is a binary word with the property that no factor has more 1s than the prefix of the same length.

Example

The words 0000, 1000, 1010 are prefix normal; the word 1011 is not.
Definition

Let $w \in \Sigma^*$, $w \neq \epsilon$. We define, for each $0 \leq k \leq |w|$

$$F_1(w, k) = \max \{|v|_1 \mid v \in \text{Fact}(w) \cap \Sigma^k\},$$

the maximum number of 1s in a factor of $w$ of length $k$. 

Definition

Let $w \in \Sigma^*$. We define, for each $0 \leq k \leq |w|$: 

$$F_1(w, k) = \max\{|v|_1 \mid v \in \text{Fact}(w) \cap \Sigma^k\},$$

the maximum number of 1s in a factor of $w$ of length $k$.

Theorem (Definition of prefix normal form)

Let $w \in \Sigma^*$. Then there exists a unique word $w'$ s.t. $F_1(w') = F_1(w)$ and $w'$ is prefix normal w.r.t. 1.

Example

The prefix normal form (w.r.t. 1) of the word 10100110110001110010 is the word 11101001011001010010.
Prefix normal words and prefix normal forms w.r.t. 0 are defined analogously. Computing prefix normal forms efficiently is important because it allows one to solve the BJPM problem.

**Problem (Binary Jumbled Pattern Matching Problem)**

*Given a binary word $w$ and a Parikh vector $P = (v_1, v_2)$, determine whether $P$ is the Parikh vector of a substring of $w$.***
Prefix normal words and prefix normal forms w.r.t. 0 are defined analogously.
Computing prefix normal forms efficiently is important because it allows one to solve the BJPM problem.

**Problem (Binary Jumbled Pattern Matching Problem)**

Given a binary word $w$ and a Parikh vector $P = (v_1, v_2)$, determine whether $P$ is the Parikh vector of a substring of $w$.

BJPM queries can be answered in constant time if the prefix normal forms of $w$ are known.

The fastest known algorithm to build the pnf of a word of length $n$ runs in $O(n^{1.864})$ time (Chan & Lewenstein, STOC 2015).
Let $L$ be the language of prefix normal words.

**Theorem**

The following holds:

1. $L$ is not context-free;
2. $L$ is (strictly) contained in the language of prefixes of Lyndon words (pre-necklaces);
3. The complexity function $pnw(n)$ of $L$ verifies

$$2^{n-4\sqrt{n\log n}} \leq pnw(n) \leq 2^{n-\log n+1}.$$
Let $L$ be the language of prefix normal words.

**Theorem**

The following holds:

1. $L$ is not context-free;
2. $L$ is (strictly) contained in the language of prefixes of Lyndon words (pre-necklaces);
3. The complexity function $\text{pnw}(n)$ of $L$ verifies

$$2^{n-4\sqrt{n \log n}} \leq \text{pnw}(n) \leq 2^{n-\log n+1}.$$ 

Based on empirical evidence, we conjecture that $\text{pnw}(n) = 2^{n-\Theta((\log n)^2)}$. 

1. Prefix Normal Words
2. Least Number of Palindromes in an Infinite Word
3. Bispecial Sturmian Words
4. Trapezoidal Words
5. The Array of Open and Closed Prefixes
6. Abelian Repetitions in Sturmian Words
7. Universal Lyndon Words
8. Cyclic Complexity

Gabriele Fici
On some Combinatorial and Algorithmic Aspects of Strings
What is the least number of palindromes $\text{MinPal}$ an infinite word must contain?

Without restrictions on the cardinality of the alphabet, one has $\text{MinPal} = 4$. e.g., the fixed point of $a \mapsto \overline{abc}$, $b \mapsto \overline{abc}$, $c \mapsto \overline{abc}$:

$$abcabcabcabcabcabcabcabc \cdots$$

In the aperiodic case, one has $\text{MinPal}(W_{ap}) = 5$. e.g., the fixed point of $a \mapsto \overline{abc}$, $b \mapsto \overline{aab}$, $c \mapsto \overline{ca}$:

$$abcaabcaabcabcaabcaabc \cdots$$

If moreover one requires that the word is closed under reversal, then one still has $\text{MinPal}(W_{cl}) = 5$. (Berstel et al., arXiv, 2009)
What is the least number of palindromes $\text{MinPal}$ an infinite word must contain?

Without restrictions on the cardinality of the alphabet, one has $\text{MinPal} = 4$. e.g., the fixed point of $a \mapsto abc, b \mapsto abc, c \mapsto abc$:

$$abcabcabcabcabcabcabcabc \cdots$$
What is the least number of palindromes $\text{MinPal}$ an infinite word must contain?

Without restrictions on the cardinality of the alphabet, one has $\text{MinPal} = 4$. e.g., the fixed point of $a \mapsto abc, b \mapsto abc, c \mapsto abc$:

$$abcabcabcabcabcabcabcabc \cdots$$

In the aperiodic case, one has $\text{MinPal}(W_{ap}) = 5$. e.g., the fixed point of $a \mapsto abc, b \mapsto aab, c \mapsto ca$:

$$abcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaabcaab
What is the least number of palindromes $MinPal$ an infinite word must contain?

Without restrictions on the cardinality of the alphabet, one has $MinPal = 4$. e.g., the fixed point of $a \mapsto abc, b \mapsto abc, c \mapsto abc$:

$$abcabcabcabcabcabcabcabc \cdots$$

In the aperiodic case, one has $MinPal(W_{ap}) = 5$. e.g., the fixed point of $a \mapsto abc, b \mapsto aab, c \mapsto ca$:

$$abcaabcaabcaabcaabcaabcaabca \cdots$$

If moreover one requires that the word is closed under reversal, then one still has $MinPal(W_{cl}) = 5$. (Berstel et al., arXiv, 2009)
In the case of binary words we have the following:

**Theorem**

Let $A$ be a set with $\#A = 2$. Then:

1. $\text{MinPal}(A_N) = 9$, where $A_N$ denotes the set of all infinite words on $A$.
2. $\text{MinPal}(A_{ap}^N) = 11$, where $A_{ap}^N$ denotes the set of all aperiodic words in $A_N$.
3. $\text{MinPal}(A_{cl}^N) = 13$, where $A_{cl}^N$ denotes the set of all words in $A_N$ closed under reversal.
4. $\text{MinPal}(A_{ap/cl}^N) = 13$, where $A_{ap/cl}^N$ denotes the set of all aperiodic words in $A_N$ closed under reversal.
In some cases, an example can be found among fixed points of primitive substitutions:
In some cases, an example can be found among fixed points of primitive substitutions:

- $MinPal(A^N) = 9$. This is realized by the fixed point of $a \mapsto aababb, b \mapsto aababb$: 

  $aababbaababbaababbaababb\cdots$
In some cases, an example can be found among fixed points of primitive substitutions:

- \( \text{MinPal}(A_N^N) = 9 \). This is realized by the fixed point of 
  \( a \mapsto aababb, b \mapsto aababb \):

  \[
  \text{aababbaababbaababbaababb} \ldots
  \]

- \( \text{MinPal}(A_{ap}^N) = 11 \). This is realized by the fixed point of 
  \( a \mapsto aaababb, b \mapsto aababb \) (M. Müller, private communication):

  \[
  \text{aaababbaaababbaaababbaababbaaababb} \ldots
  \]
In some cases, an example can be found among fixed points of primitive substitutions:

- \( \text{MinPal}(A^N) = 9 \). This is realized by the fixed point of \( a \mapsto aababb, b \mapsto aababb \):
  \[
  aababbaaababbaaababbaababb \cdots
  \]

- \( \text{MinPal}(A^N_{\text{ap}}) = 11 \). This is realized by the fixed point of \( a \mapsto aaababb, b \mapsto aababb \) (M. Müller, private communication):
  \[
  aaababbaaababbaaababbaaababb \cdots
  \]

However, for fixed points of primitive substitutions, one has \( \text{MinPal}(A^N_{\text{cl}}) = \text{MinPal}(A^N_{\text{ap/cl}}) = +\infty \) as a consequence of a theorem of Tan (TCS, 2007).
A word $w$ over an alphabet $\Sigma$ is balanced if for every letter $a \in \Sigma$, any two factors $u, v \in \text{Fact}(w)$ of the same length contain the same number of $a$’s up to 1:

$$\left| |u|_a - |v|_a \right| \leq 1$$

Example

$aaabaa$ and $baaaab$ are balanced, $aabb$ and $aaabab$ are not.
A word $w$ over an alphabet $\Sigma$ is balanced if for every letter $a \in \Sigma$, any two factors $u, v \in \text{Fact}(w)$ of the same length contain the same number of $a$'s up to 1:

$$||u|_a - |v|_a| \leq 1$$

**Example**

$aaabaa$ and $baaaab$ are balanced, $aabb$ and $aaabab$ are not.

A binary balanced word is called a **Sturmian word**. The set of Sturmian words is noted $St$. 
A Sturmian word $w$ is:

- right special (RS) if $wa, wb \in St$
- left special (LS) if $aw, bw \in St$
- bispecial (BS) if $w$ is both right and left special
A Sturmian word $w$ is:

- right special (RS) if $wa, wb \in St$
- left special (LS) if $aw, bw \in St$
- bispecial (BS) if $w$ is both right and left special

Moreover, a bispecial Sturmian word $w$ is

- strictly bispecial (SBS) if $awa, awb, bwa, bwb \in St$
- non-strictly bispecial (NBS), otherwise.
A Sturmian word $w$ is:

- right special (RS) if $wa, wb \in St$
- left special (LS) if $aw, bw \in St$
- bispecial (BS) if $w$ is both right and left special

Moreover, a bispecial Sturmian word $w$ is

- strictly bispecial (SBS) if $awa, awb, bwa, bwb \in St$
- non-strictly bispecial (NBS), otherwise.

**Example**

$aba$ is strictly bispecial,
$ab$ is non-strictly bispecial,
$aab$ is left special but not right special,
$baab$ is neither left special nor right special.
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word \( w_{6,4} = aababaabab \) (left) and the upper Christoffel word \( w'_{6,4} = babaababaa \) (right).
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word $w_{6,4} = aababaabab$ (left) and the upper Christoffel word $w'_{6,4} = babaababaa$ (right).

$$w_{p,q}[i] = \begin{cases} 
  a & \text{if } iq \mod (p + q) > (i - 1)q \mod (p + q), \\
  b & \text{if } iq \mod (p + q) < (i - 1)q \mod (p + q). 
\end{cases}$$
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word $w_{6,4} = aababaabab$ (left) and the upper Christoffel word $w'_{6,4} = babaababaa$ (right).

$$\{i4 \mod(10) \mid i = 0, 1, \ldots, 10\} = \{0, 4, 8, 2, 6, 0, 4, 8, 2, 6, 0\}$$
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word $w_{6,4} = aababaabab$ (left) and the upper Christoffel word $w'_{6,4} = babaababaa$ (right).

Remark

$w'_{p,q}$ is the reversal of $w_{p,q}$. 
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word $w_{6,4} = aababaabab$ (left) and the upper Christoffel word $w'_{6,4} = babaababaa$ (right).

Remark

If $\frac{p}{q} = r \frac{p'}{q'}$, then $w_{p,q} = (w_{p',q'})^r$. 
What are the best grid approximations of a segment with integer coordinates in the Euclidean plane?

Figure: The lower Christoffel word $w_{6,4} = aababaabab$ (left) and the upper Christoffel word $w'_{6,4} = babaababaa$ (right).

Remark

A Christoffel word is **primitive** if and only if $\gcd(p, q) = 1$. 
The maximal internal factor of a word $w = a_1a_2\cdots a_n$, $n \geq 2$, is the factor $a_2a_3\cdots a_{n-1}$.

$BS = \text{bispecial Sturmian words} = SBS \cup NBS$

$SBS = \text{strictly bispecial Sturmian words}$

$NBS = \text{non-strictly bispecial Sturmian words}$
The maximal internal factor of a word $w = a_1a_2\cdots a_n$, $n \geq 2$, is the factor $a_2a_3\cdots a_{n-1}$.

$BS = \text{bispecial Sturmian words} = SBS \cup NBS$

$SBS = \text{strictly bispecial Sturmian words}$

$NBS = \text{non-strictly bispecial Sturmian words}$

**Theorem (Berstel, de Luca, 1997)**

*SBS is the set of maximal internal factors of primitive Christoffel words.*
The maximal internal factor of a word \( w = a_1 a_2 \cdots a_n, \ n \geq 2 \), is the factor \( a_2 a_3 \cdots a_{n-1} \).

\[ BS = \text{bispecial Sturmian words} = SBS \cup NBS \]

\( SBS = \) strictly bispecial Sturmian words
\( NBS = \) non-strictly bispecial Sturmian words

**Theorem (Berstel, de Luca, 1997)**

\( SBS \) is the set of maximal internal factors of primitive Christoffel words.

**Theorem**

\( BS \) is the set of maximal internal factors of Christoffel words.
Let $\phi$ be the Euler totient function. 
The number of strictly bispecial Sturmian words is known to be

$$SBS(n) = \phi(n + 2)$$
Let $\phi$ be the Euler totient function. The number of strictly bispecial Sturmian words is known to be

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

By the theorem we have:

$$NBS(n) = 2(n + 1 - \phi(n + 2))$$

and so

$$BS(n) = 2(n + 1) - \phi(n + 2)$$
Let $L \subseteq \Sigma^*$ be a factorial ($uv \in L \Rightarrow u, v \in L$) language.
Let $L \subseteq \Sigma^*$ be a **factorial** ($uv \in L \Rightarrow u, v \in L$) language.

**Definition**

A word $w \in \Sigma^*$ is a **minimal forbidden word** for the factorial language $L$ if $w \notin L$ but every proper factor of $w$ belongs to $L$.

**Example**

$L = \text{Fact}(aba, aab)$. Its set of m.f.w. is $\{bb, aaa, bab, baa, aaba\}$. 

Gabriele Fici  On some Combinatorial and Algorithmic Aspects of Strings
Let $L \subseteq \Sigma^*$ be a **factorial** ($uv \in L \Rightarrow u, v \in L$) language.

**Definition**

A word $w \in \Sigma^*$ is a **minimal forbidden word** for the factorial language $L$ if $w \notin L$ but every proper factor of $w$ belongs to $L$.

**Example**

$L = \text{Fact}(aba, aab)$. Its set of m.f.w. is \{bb, aaa, bab, baa, aaba\}.

$MFSt = \text{set of m.f.w. for the language } St \text{ of Sturmian words.}$
Let $L \subseteq \Sigma^*$ be a factorial ($uv \in L \Rightarrow u, v \in L$) language.

**Definition**

A word $w \in \Sigma^*$ is a **minimal forbidden word** for the factorial language $L$ if $w \notin L$ but every proper factor of $w$ belongs to $L$.

**Example**

$L = \text{Fact}(aba, aab)$. Its set of m.f.w. is $\{bb, aaa, bab, baa, aaba\}$.

$\text{MFSt} = \text{set of m.f.w. for the language } St \text{ of Sturmian words.}$

**Theorem**

$\text{MFSt} = \{ywx \mid xwy \text{ is a non-primitive Christoffel word, } x, y \in \Sigma\}.$
Corollary

For every $n > 1$, one has $MFSt(n) = 2(n - 1 - \phi(n))$. 
Corollary

For every \( n > 1 \), one has \( MFSt(n) = 2(n - 1 - \phi(n)) \).

It is known that \( St(n) = O(n^3) \), as a consequence of the estimation

\[
\sum_{i=1}^{n} \phi(i) = \frac{3n^2}{\pi^2} + O(n \log n) \tag{1}
\]
Corollary

For every $n > 1$, one has $MFSt(n) = 2(n - 1 - \phi(n))$.

It is known that $St(n) = O(n^3)$, as a consequence of the estimation

$$\sum_{i=1}^{n} \phi(i) = \frac{3n^2}{\pi^2} + O(n \log n) \quad (1)$$

From (1) and from the formula of the Corollary above, we have that

$$\sum_{i=1}^{n} MFSt(n) = O(n^2)$$
For every \( n > 1 \), one has \( \text{MFSt}(n) = 2(n - 1 - \phi(n)) \).

It is known that \( \text{St}(n) = O(n^3) \), as a consequence of the estimation

\[
\sum_{i=1}^{n} \phi(i) = \frac{3n^2}{\pi^2} + O(n \log n) \tag{1}
\]

From (1) and from the formula of the Corollary above, we have that

\[
\sum_{i=1}^{n} \text{MFSt}(n) = O(n^2)
\]

Is it possible to give a better estimate?
A Sturmian word has at most \( n + 1 \) distinct factors of each length \( n \).

However, this property does not characterize Sturmian words, e.g. \( w = aaabab \).
A Sturmian word has at most \( n + 1 \) distinct factors of each length \( n \).

However, this property does not characterize Sturmian words, e.g. \( w = aaabab \).

**Definition**

A binary word is **trapezoidal** if it has at most \( n + 1 \) distinct factors of each length \( n \).

So trapezoidal words are a generalization of finite Sturmian words.
A rich word is a word of length \( n \) containing \( n \) non-empty palindromic factors.

For example, \( abaab \) is rich, \( abca \) is not.
A rich word is a word of length $n$ containing $n$ non-empty palindromic factors. For example, $abaab$ is rich, $abca$ is not.

**Proposition**

*Every trapezoidal word is rich.*
A rich word is a word of length $n$ containing $n$ non-empty palindromic factors.

For example, $abaab$ is rich, $abca$ is not.

**Proposition**

Every trapezoidal words is rich.

Moreover, a trapezoidal word is a palindrome if and only if it is a Sturmian palindrome.
An enumerative formula for trapezoidal words can be given.

**Theorem**

*For all* \( n \geq 0 \), *the number of trapezoidal words of length* \( n \) *is*

\[
1 + \sum_{i=1}^{n} (n - i + 1) \phi(i) + \left\lfloor \frac{(n-4)}{2} \right\rfloor \sum_{i=0}^{\left\lfloor \frac{(n-4)}{2} \right\rfloor} 2(n - 2i - 3) \phi(i + 2).
\]

<table>
<thead>
<tr>
<th>( n )</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \text{Trapezoidal Words} \cap \Sigma^n )</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>16</td>
<td>28</td>
<td>46</td>
<td>70</td>
<td>102</td>
<td>140</td>
<td>190</td>
<td>250</td>
<td>318</td>
<td>398</td>
<td>496</td>
</tr>
</tbody>
</table>

This is now OEIS sequence A260881 (thanks to J. Shallit).
1. Prefix Normal Words
2. Least Number of Palindromes in an Infinite Word
3. Bispecial Sturmian Words
4. Trapezoidal Words
5. The Array of Open and Closed Prefixes
6. Abelian Repetitions in Sturmian Words
7. Universal Lyndon Words
8. Cyclic Complexity
Definition

A word is **closed** if it is empty or it contains a factor occurring only as a prefix and as a suffix (with no internal occurrences). Otherwise, it is **open**.
Definition

A word is **closed** if it is empty or it contains a factor occurring only as a prefix and as a suffix (with no internal occurrences). Otherwise, it is **open**.

Example

*abab, ababaaba*, *aaaa* and *a* are closed words.

*ab, abbaba* and *ababc* are open words.
Definition

A word is **closed** if it is empty or it contains a factor occurring only as a prefix and as a suffix (with no internal occurrences). Otherwise, it is **open**.

Example

- *abab*, *abaabaab*, *aaaa* and *a* are closed words.
- *ab*, *abbaba* and *ababc* are open words.

Remark

*Closed words are also known as periodic-like words, or complete (first) returns.*
We can separate the factors of a word in open and closed. We focus here on the prefixes.
We can separate the factors of a word in open and closed. We focus here on the prefixes.

**Definition**

Let $w$ be a finite or infinite word. We define the sequence $oc(w)$ as the sequence whose $n$-th element is:

- 1 if the prefix of length $n$ of $w$ is closed;
- 0 if it is open.
We can separate the factors of a word in open and closed. We focus here on the prefixes.

**Definition**

Let $w$ be a finite or infinite word. We define the sequence $oc(w)$ as the sequence whose $n$-th element is:

- 1 if the prefix of length $n$ of $w$ is closed;
- 0 if it is open.

For example, if $w = abaabab$, then $oc(w) = 1010110$. 
Theorem

The sequence oc characterizes every (finite or infinite) Sturmian word, up to exchanging letters.
The sequence $oc$ characterizes every (finite or infinite) Sturmian word, up to exchanging letters.

That is, if $w_1$ and $w_2$ are Sturmian words, then $oc(w_1) = oc(w_2)$ if and only if $w_1 = w_2$ (up to isomorphisms).
Theorem

The sequence \( oc \) characterizes every (finite or infinite) Sturmian word, up to exchanging letters.

That is, if \( w_1 \) and \( w_2 \) are Sturmian words, then \( oc(w_1) = oc(w_2) \) if and only if \( w_1 = w_2 \) (up to isomorphisms).

The hypothesis that \( w_1 \) and \( w_2 \) be Sturmian words is necessary. For example, \( aaba \) and \( aabb \) have the same sequence \( oc: 1100 \).
Example

Let $F = abaababaabaababaababaabaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababaababa...
Example

Let $F = \text{abaababaabaababaababaababaab} \cdots$ be the Fibonacci word. Then:

$$oc(F) = 10101100111000111100000 \cdots$$

Remarks:

1. the runs of 1’s are followed by equal-length runs of 0’s;
2. the lengths of the runs form the Fibonacci sequence.
Example
Let \( F = abaababaabaababaabaababaabaab \cdots \) be the Fibonacci word. Then:
\[ \text{oc}(F) = 101011001110001111000000 \cdots \]

Remarks:
1. the runs of 1’s are followed by equal-length runs of 0’s;
2. the lengths of the runs form the Fibonacci sequence.

Theorem
Let \( w \) be a prefix of the Fibonacci word \( F \). Then \( w \) is open if and only if there exists \( i \) such that \( F_{i+1} - 1 \leq |w| \leq 2F_i - 2 \).
Example

Let $F = \text{abaababaabaababaababaababaababaababaab} \cdots$ be the Fibonacci word. Then: $\text{oc}(F) = 10101100111000111100000 \cdots$

Remarks:

1. the runs of 1’s are followed by equal-length runs of 0’s;
2. the lengths of the runs form the Fibonacci sequence.

Theorem

Let $w$ be a prefix of the Fibonacci word $F$. Then $w$ is open if and only if there exists $i$ such that $F_{i+1} - 1 \leq |w| \leq 2F_i - 2$.

Question: What can we say about other standard Sturmian words?
Definition

Let $\alpha$ be an irrational number such that $0 < \alpha < 1$, and $[0; d_0 + 1, d_1, \ldots]$ its continued fraction expansion. The sequence of words defined by:

$$s_{-1} = b, \ s_0 = a \text{ and } s_{n+1} = s_n^{d_n} s_{n-1} \text{ for } n \geq 0$$

converges to the infinite **standard Sturmian word** $w_\alpha$. 


Definition

Let $\alpha$ be an irrational number such that $0 < \alpha < 1$, and $[0; d_0 + 1, d_1, \ldots]$ its continued fraction expansion. The sequence of words defined by:

$$s_{-1} = b, \ s_0 = a \text{ and } s_{n+1} = s_n^{d_n} s_{n-1} \text{ for } n \geq 0$$

converges to the infinite standard Sturmian word $w_\alpha$.

1. The sequence of words $s_n$ is called the standard sequence of $w_\alpha$. 
Definition

Let $\alpha$ be an irrational number such that $0 < \alpha < 1$, and $[0; d_0 + 1, d_1, \ldots]$ its continued fraction expansion. The sequence of words defined by:

$$s_{-1} = b, \quad s_0 = a \quad \text{and} \quad s_{n+1} = s_d^n s_{n-1} \quad \text{for} \quad n \geq 0$$

converges to the infinite standard Sturmian word $w_\alpha$.

1. The sequence of words $s_n$ is called the standard sequence of $w_\alpha$.
2. A standard word is a finite word belonging to some standard sequence.
Definition

Let \( \alpha \) be an irrational number such that \( 0 < \alpha < 1 \), and \([0; d_0 + 1, d_1, \ldots]\) its continued fraction expansion. The sequence of words defined by:

\[
s_{-1} = b, \quad s_0 = a \quad \text{and} \quad s_{n+1} = s_{dn}^n s_{n-1} \quad \text{for} \quad n \geq 0
\]

converges to the infinite standard Sturmian word \( w_\alpha \).

1. The sequence of words \( s_n \) is called the standard sequence of \( w_\alpha \).
2. A standard word is a finite word belonging to some standard sequence.
3. A central word is a word \( u \) such that \( uxy \) is a standard word, for letters \( x, y \in \Sigma \).
Example

The Fibonacci word $F$ is the standard Sturmian word of slope

$$\alpha = (3 - \sqrt{5})/2 = [0; 2, 1, 1, 1, \ldots]$$

so that $d_n = 1$ for every $n \geq 0$. Therefore, the standard sequence of $F$ is:

$$f_{-1} = b, \quad f_0 = a, \quad f_{n+1} = f_n f_{n-1} \text{ for } n \geq 0.$$  

This sequence is also called the sequence of Fibonacci finite words.
Theorem (Aldo de Luca, 1997)

A word is central if and only if $w$ is the power of a single letter or there exist palindromes $w_1, w_2$ such that $w = w_1 xyw_2 = w_2 yxw_1$, for different letters $x, y$. Moreover, if $|w_1| < |w_2|$, then $w_2$ is the longest palindromic suffix of $w$. 
Theorem (Aldo de Luca, 1997)

A word is central if and only if $w$ is the power of a single letter or there exist palindromes $w_1, w_2$ such that $w = w_1 xyw_2 = w_2 yxw_1$, for different letters $x, y$. Moreover, if $|w_1| < |w_2|$, then $w_2$ is the longest palindromic suffix of $w$.

Every central word is closed.
Theorem (Aldo de Luca, 1997)

A word is central if and only if \( w \) is the power of a single letter or there exist palindromes \( w_1, w_2 \) such that \( w = w_1xyw_2 = w_2yxw_1 \), for different letters \( x, y \). Moreover, if \( |w_1| < |w_2| \), then \( w_2 \) is the longest palindromic suffix of \( w \).

1. Every central word is closed.

2. Every central word is a palindrome.
**Definition**

A *semicentral* word is a word in which the longest repeated prefix, repeated suffix, left special factor and right special factor all coincide.
Definition

A **semicentral** word is a word in which the longest repeated prefix, repeated suffix, left special factor and right special factor all coincide.

Theorem

A word $w$ is semicentral if and only if $w = w_1xyw_1$ for a central word $w_1$ and different letters $x, y$. 
Definition

A **semicentral** word is a word in which the longest repeated prefix, repeated suffix, left special factor and right special factor all coincide.

Theorem

A word $w$ is semicentral if and only if $w = w_1 x y w_1$ for a central word $w_1$ and different letters $x, y$.

In a semicentral word $w_1 x y w_1$, the central word $w_1$ is the longest repeated prefix, repeated suffix, left special factor and right special factor.
Definition

A **semicentral** word is a word in which the longest repeated prefix, repeated suffix, left special factor and right special factor all coincide.

Theorem

A word $w$ is semicentral if and only if $w = w_1 xyw_1$ for a central word $w_1$ and different letters $x, y$.

In a semicentral word $w_1 xyw_1$, the central word $w_1$ is the longest repeated prefix, repeated suffix, left special factor and right special factor.

Semicentral words are therefore open words.
For every $n \geq -1$, one has

$$s_n = u_nxy,$$

(2)

for $x, y$ letters such that $xy = ab$ if $n$ is odd or $ba$ if $n$ is even. Indeed, the sequence $(u_n)_{n \geq -1}$ can be defined by: $u_{-1} = a^{-1}$, $u_0 = b^{-1}$, and, for every $n \geq 1$,

$$u_{n+1} = (u_nxy)^{d_n} u_{n-1},$$

(3)

where $x, y$ are as in (2).
For every $n \geq -1$, one has

$$s_n = u_nxy,$$  \hspace{1cm} (2)

for $x, y$ letters such that $xy = ab$ if $n$ is odd or $ba$ if $n$ is even. Indeed, the sequence $(u_n)_{n \geq -1}$ can be defined by: $u_{-1} = a^{-1}$, $u_0 = b^{-1}$, and, for every $n \geq 1$,

$$u_{n+1} = (u_nxy)^{d_n} u_{n-1},$$  \hspace{1cm} (3)

where $x, y$ are as in (2).

**Theorem**

Let $(u_nxy)$ be the standard sequence of a standard Sturmian word $w$. Let $vx$, $x \in \{a, b\}$, be a prefix of $w$. Then:

1. $v$ is closed and $vx$ is open if and only if there exists $n \geq 0$ such that $v = u_nxyu_{n+1} = u_{n+1}yux_n$ (central prefixes);
2. $v$ is open and $vx$ is closed if and only if there exists $n \geq 1$ such that $v = u_nxyu_n$ (semicentral prefixes).
Table: The structure of the prefixes of a standard Sturmian word $w = \textit{aabaabaaabaabaa} \cdots$ with respect to the $u_n$ prefixes. Here $d_0 = d_1 = 2$ and $d_2 = 1$. 

<table>
<thead>
<tr>
<th>Prefix of $w$</th>
<th>Open/Closed</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_nxyu_n$</td>
<td>open</td>
<td>$aaba$</td>
</tr>
<tr>
<td>$u_nxyu_nx$</td>
<td>closed</td>
<td>$aabaa$</td>
</tr>
<tr>
<td>$u_nxyu_nxy$</td>
<td>closed</td>
<td>$aabaab$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$u_nxyu_{n+1} = u_{n+1}yxu_n$</td>
<td>closed</td>
<td>$aabaabaa$</td>
</tr>
<tr>
<td>$u_{n+1}y xu_ny$</td>
<td>open</td>
<td>$aabaabaa$</td>
</tr>
<tr>
<td>$u_{n+1}yxu_nyx$</td>
<td>open</td>
<td>$aabaabaab$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$u_{n+1}yxu_{n+1}$</td>
<td>open</td>
<td>$aabaabaaabaa$</td>
</tr>
<tr>
<td>$u_{n+1}yxu_{n+1}y$</td>
<td>closed</td>
<td>$aabaabaaabaab$</td>
</tr>
<tr>
<td>prefix of $w$</td>
<td>open/closed</td>
<td>example</td>
</tr>
<tr>
<td>-----------------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>$u_n xy u_n$</td>
<td>open</td>
<td>$aaba$</td>
</tr>
<tr>
<td>$u_n xy u_n x$</td>
<td>closed</td>
<td>$aabaa$</td>
</tr>
<tr>
<td>$u_n xy u_n xy$</td>
<td>closed</td>
<td>$aabaab$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$u_n xy u_{n+1} = u_{n+1} yx u_n$</td>
<td>closed</td>
<td>$aabaabaa$</td>
</tr>
<tr>
<td>$u_{n+1} yx u_n y$</td>
<td>open</td>
<td>$aabaabaabaa$</td>
</tr>
<tr>
<td>$u_{n+1} yx u_n yx$</td>
<td>open</td>
<td>$aabaabaabaaab$</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>$u_{n+1} yx u_{n+1}$</td>
<td>open</td>
<td>$aabaabaaabaaab$</td>
</tr>
<tr>
<td>$u_{n+1} yx u_{n+1} y$</td>
<td>closed</td>
<td>$aabaabaaabaaabab$</td>
</tr>
</tbody>
</table>

**Table:** The structure of the prefixes of a standard Sturmian word $w = aabaabaaabaaabaa \cdots$ with respect to the $u_n$ prefixes. Here $d_0 = d_1 = 2$ and $d_2 = 1$. 
Theorem

Let \( w = w_\alpha \) be a standard Sturmian word, with \( 0 < \alpha < 1/2 \), and let \( \alpha = [0; d_0 + 1, d_1, \ldots] \). The word \( ba^{-1} w \), obtained from \( w \) by complementing the first letter, can be written as an infinite product of squares of reversed standard words in the following way:

\[
ba^{-1} w = \prod_{n \geq 0} (u_{n-1} u_{n+1})^2,
\]

where \( (u_n x y)_{n \geq -1} \) is the standard sequence of \( w \).

In other words, one can write

\[
w = a^{d_0} ba^{d_0-1} \prod_{n \geq 1} (u_{n-1} u_{n+1})^2.
\]
Example

Take the Fibonacci word \( F \). Then,

\[
\begin{align*}
  u_1 &= \varepsilon, & u_2 &= a, & u_3 &= aba, & u_4 &= abaaba, & u_5 &= abaababaaba, & \ldots
\end{align*}
\]
Example

Take the Fibonacci word $F$. Then,

$u_1 = \varepsilon$, $u_2 = a$, $u_3 = aba$, $u_4 = abaaba$, $u_5 = abaababaaba$, ...

Indeed, $u_{n-1} u_{n+1}$ is the reversal of the Fibonacci finite word $f_{n-1}$. 

By the previous theorem, we have:

$$F = \mathsf{ab} \coprod_{n \geq 1} (u_{n-1} u_n + 1)$$

$$= \mathsf{ab} \coprod_{n \geq 0} (\bar{f}_n)^2$$

$$= \mathsf{ab} \cdot (\mathsf{a} \cdot \mathsf{a}) \cdot (\mathsf{ba} \cdot \mathsf{ba}) \cdot (\mathsf{aba} \cdot \mathsf{aba}) \cdot (\mathsf{baaba} \cdot \mathsf{baaba}) \cdot ...$$

i.e., $F$ can be obtained by concatenating $\mathsf{ab}$ and the squares of the reversals of the Fibonacci finite words $f_n$ starting from $n = 0$. 

Gabriele Fici
On some Combinatorial and Algorithmic Aspects of Strings
Example

Take the Fibonacci word $F$. Then,

$$u_1 = \varepsilon, \quad u_2 = a, \quad u_3 = aba, \quad u_4 = ababa, \quad u_5 = abababaaba, \ldots$$

Indeed, $u_{n-1}u_n$ is the reversal of the Fibonacci finite word $f_{n-1}$.

By the previous theorem, we have:

$$F = ab \prod_{n \geq 1} (u_{n-1}u_n)^2$$

$$= ab \prod_{n \geq 0} (\tilde{f}_n)^2$$

$$= ab \cdot (a \cdot a)(ba \cdot ba)(aba \cdot aba)(baaba \cdot baaba) \cdots$$

i.e., $F$ can be obtained by concatenating $ab$ and the squares of the reversals of the Fibonacci finite words $f_n$ starting from $n = 0$. 
Question: Is it possible to characterize the sequence of open and closed prefixes of a standard Sturmian word $w_\alpha$ in terms of the continued fraction expansion of $\alpha$?
Question: Is it possible to characterize the sequence of open and closed prefixes of a standard Sturmian word $w_\alpha$ in terms of the continued fraction expansion of $\alpha$?

**Definition**

The *continuants* of an integer sequence $(a_n)_{n \geq 0}$ are defined as $K[\ ] = 1$, $K[a_0] = a_0$, and, for every $n \geq 1$,

$$K[a_0, \ldots, a_n] = a_n K[a_0, \ldots, a_{n-1}] + K[a_0, \ldots, a_{n-2}].$$
Question: Is it possible to characterize the sequence of open and closed prefixes of a standard Sturmian word $w_\alpha$ in terms of the continued fraction expansion of $\alpha$?

Definition

The *continuants* of an integer sequence $(a_n)_{n \geq 0}$ are defined as $K[\ ] = 1$, $K[a_0] = a_0$, and, for every $n \geq 1$,

$$K[a_0, \ldots, a_n] = a_nK[a_0, \ldots, a_{n-1}] + K[a_0, \ldots, a_{n-2}].$$

Let $w$ be a standard Sturmian word and $(s_n)_{n \geq -1}$ its standard sequence. Then:

$$|s_n| = K[1, d_0, \ldots, d_{n-1}].$$
From the previous theorem, we have:

**Corollary**

Let $w = w_{\alpha}$ be a standard Sturmian word, with $0 < \alpha < 1/2$, and $\alpha = [0; d_0 + 1, d_1, \ldots]$. Let, for every $n \geq 0$, $k_n = K \left[1, d_0, \ldots, d_{n-1}, d_n - 1\right]$. Then

$$oc(w) = \prod_{n \geq 0} 1^{k_n} 0^{k_n}.$$
We fix the ordered alphabet $\Sigma = \{a, b\}$.

The Parikh vector of a word $u$ over $\Sigma$ is the vector $(|u|_a, |u|_b)$ counting the number of occurrences of letters of $\Sigma$ in $u$. 
We fix the ordered alphabet $\Sigma = \{a, b\}$.

The Parikh vector of a word $u$ over $\Sigma$ is the vector $(|u|_a, |u|_b)$ counting the number of occurrences of letters of $\Sigma$ in $u$.

**Definition**

An integer $p$ is an abelian period for a word $w$ over $\Sigma$ if $w$ can be written as

$$w = u_0 u_1 \cdots u_{j-1} u_j,$$

where $j \geq 2$ and for every $0 < i < j$ all the $u_i$’s have the same Parikh vector $P$, with sum of components equal to $p$, and the Parikh vectors of $u_0$ and $u_k$ are component-wise smaller than or equal to $P$. 
We fix the ordered alphabet $\Sigma = \{a, b\}$.

The **Parikh vector** of a word $u$ over $\Sigma$ is the vector $(|u|_a, |u|_b)$ counting the number of occurrences of letters of $\Sigma$ in $u$.

**Definition**

An integer $p$ is an **abelian period** for a word $w$ over $\Sigma$ if $w$ can be written as

$$w = u_0 u_1 \cdots u_{j-1} u_j,$$

where $j \geq 2$ and for every $0 < i < j$ all the $u_i$’s have the same Parikh vector $\mathcal{P}$, with sum of components equal to $p$, and the Parikh vectors of $u_0$ and $u_k$ are component-wise smaller than or equal to $\mathcal{P}$.

**Example**

2 is the smallest abelian period of $w = abaab = a \cdot ba \cdot ab \cdot \varepsilon$. 
We fix the ordered alphabet $\Sigma = \{a, b\}$.

The Parikh vector of a word $u$ over $\Sigma$ is the vector $(|u|_a, |u|_b)$ counting the number of occurrences of letters of $\Sigma$ in $u$.

**Definition**

An integer $p$ is an abelian period for a word $w$ over $\Sigma$ if $w$ can be written as

$$w = u_0 u_1 \cdots u_{j-1} u_j,$$

where $j \geq 2$ and for every $0 < i < j$ all the $u_i$’s have the same Parikh vector $P$, with sum of components equal to $p$, and the Parikh vectors of $u_0$ and $u_k$ are component-wise smaller than or equal to $P$.

**Example**

2 is the smallest abelian period of $w = abaab = a \cdot ba \cdot ab \cdot \varepsilon$.

$h = |u_0|$ and $t = |u_k|$ are called resp. head and tail of the abelian period.
Definition

We say that \( w \) is an **abelian repetition** of (abelian) **period** \( m \) and (abelian) **exponent** \( k = |w|/m \) if one can write

\[
w = u_0 u_1 \cdots u_{j-1} u_j
\]

where \( j \geq 3 \) and for every \( 0 < i < j \) all the \( u_i \)'s have the same Parikh vector \( \mathcal{P} \), with sum of components equal to \( m \), and the Parikh vectors of \( u_0 \) and \( u_k \) are component-wise smaller than or equal to \( \mathcal{P} \).
Definition

We say that $w$ is an abelian repetition of (abelian) period $m$ and (abelian) exponent $k = |w|/m$ if one can write

$$w = u_0 u_1 \cdots u_{j-1} u_j$$

where $j \geq 3$ and for every $0 < i < j$ all the $u_i$’s have the same Parikh vector $\mathcal{P}$, with sum of components equal to $m$, and the Parikh vectors of $u_0$ and $u_k$ are component-wise smaller than or equal to $\mathcal{P}$.

An abelian power is an abelian repetition in which $u_0 = u_j = \varepsilon$, i.e., with empty head and empty tail. A word with abelian exponent $k = 1$ is called a degenerated abelian power.
Definition

We say that \( w \) is an **abelian repetition** of (abelian) **period** \( m \) and (abelian) **exponent** \( k = |w|/m \) if one can write

\[
    w = u_0 u_1 \cdots u_{j-1} u_j
\]

where \( j \geq 3 \) and for every \( 0 < i < j \) all the \( u_i \)'s have the same Parikh vector \( P \), with sum of components equal to \( m \), and the Parikh vectors of \( u_0 \) and \( u_k \) are component-wise smaller than or equal to \( P \).

An **abelian power** is an abelian repetition in which \( u_0 = u_j = \epsilon \), i.e., with empty head and empty tail. A word with abelian exponent \( k = 1 \) is called a **degenerated abelian power**.

Example

\( w = \text{abaab} \) is an abelian repetition of period 2 and exponent 5/2.
Theorem

Let $s_\alpha$ be a Sturmian word with rotation angle $\alpha$, and $m$ a positive integer. Then $s_\alpha$ contains an abelian power of period $m$ and exponent $k \geq 2$ if and only if $\|m\alpha\| < \frac{1}{k}$, where $\|m\alpha\| = \min(\{m\alpha\}, \{-m\alpha\})$ is the distance of $\alpha$ form the closest integer.
Theorem

Let \( s_\alpha \) be a Sturmian word with rotation angle \( \alpha \), and \( m \) a positive integer. Then \( s_\alpha \) contains an abelian power of period \( m \) and exponent \( k \geq 2 \) if and only if \( \| m\alpha \| < \frac{1}{k} \), where \( \| m\alpha \| = \min(\{m\alpha\}, \{-m\alpha\}) \) is the distance of \( \alpha \) from the closest integer.

Corollary

Let \( s_\alpha \) be a Sturmian word with rotation angle \( \alpha \), and \( m \) a positive integer. Then the maximal exponent \( k_m \) of an abelian power of period \( m \) in \( s_\alpha \) is the largest \( k \) such that \( \| m\alpha \| < \frac{1}{k} \), i.e.,

\[
k_m = \left\lfloor \frac{1}{\| m\alpha \|} \right\rfloor.
\]
Theorem

Let $s_\alpha$ be a Sturmian word with rotation angle $\alpha$, and $m$ a positive integer. Then $s_\alpha$ contains an abelian power of period $m$ and exponent $k \geq 2$ if and only if $\|m\alpha\| < \frac{1}{k}$, where $\|m\alpha\| = \min(\{m\alpha\}, \{-m\alpha\})$ is the distance of $\alpha$ form the closest integer.

Corollary

Let $s_\alpha$ be a Sturmian word with rotation angle $\alpha$, and $m$ a positive integer. Then the maximal exponent $k_m$ of an abelian power of period $m$ in $s_\alpha$ is the largest $k$ such that $\|m\alpha\| < \frac{1}{k}$, i.e.,

$$k_m = \left\lfloor \frac{1}{\|m\alpha\|} \right\rfloor.$$

<table>
<thead>
<tr>
<th>$m$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
<th>17</th>
<th>18</th>
<th>19</th>
<th>20</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_m$</td>
<td>2</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td>11</td>
<td>3</td>
<td>3</td>
<td>17</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>29</td>
<td>2</td>
<td>3</td>
<td>8</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>2</td>
<td>46</td>
</tr>
</tbody>
</table>

Table: The first values of the maximal exponent $k_m$ of an abelian power of period $m$ in the Fibonacci word $f$. The values corresponding to the Fibonacci numbers are in bold.
Definition

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. For every integer $m > 1$, let $k_m$ (resp. $k'_m$) be the maximal exponent of an abelian power (resp. abelian repetition) of period $m$ in $s_\alpha$. The abelian critical exponent of $s_\alpha$ is defined as

$$A(s_\alpha) = \limsup_{m \to \infty} \frac{k_m}{m} = \limsup_{m \to \infty} \frac{k'_m}{m}.$$

The two superior limits coincide since by definition $k_m \leq k'_m < k_m + 2$ for every $m \geq 1$. 

Gabriele Fici  
On some Combinatorial and Algorithmic Aspects of Strings
Definition

Let \( s_\alpha \) be a Sturmian word of angle \( \alpha \). For every integer \( m > 1 \), let \( k_m \) (resp. \( k'_m \)) be the maximal exponent of an abelian power (resp. abelian repetition) of period \( m \) in \( s_\alpha \). The abelian critical exponent of \( s_\alpha \) is defined as

\[
A(s_\alpha) = \limsup_{m \to \infty} \frac{k_m}{m} = \limsup_{m \to \infty} \frac{k'_m}{m}.
\] (4)

The two superior limits coincide since by definition \( k_m \leq k'_m < k_m + 2 \) for every \( m \geq 1 \).

Before studying abelian critical exponent further, we explore its connection to a number-theoretical concept know as the Markov constant.

Definition

Let \( \alpha \) be a real number. The Markov constant of \( \alpha \) is defined as

\[
\mu(\alpha) = \liminf_{k \to \infty} k \| k\alpha \|.
\]
From the previous results, we have:

**Theorem**

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. Then $A(s_\alpha) = \mu(\alpha)^{-1}$ (if $\mu(\alpha) = 0$, then $A(\alpha) = \infty$). In other words, the abelian critical exponent of a Sturmian word is the inverse of the Markov constant of its angle.
From the previous results, we have:

**Theorem**

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. Then $A(s_\alpha) = \mu(\alpha)^{-1}$ (if $\mu(\alpha) = 0$, then $A(\alpha) = \infty$). In other words, the abelian critical exponent of a Sturmian word is the inverse of the Markov constant of its angle.

We can therefore give a formula for the abelian critical exponent:

**Proposition**

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. Then

$$A(s_\alpha) = \limsup_{n \to \infty} \left( [a_{n+1}; a_{n+2}, \ldots] + [0; a_n, a_{n-1}, \ldots, a_1] \right).$$
Furthermore, we have:

**Theorem**

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. The following are equivalent:

(i) $A(s_\alpha)$ is finite,

(ii) $s_\alpha$ is $\beta$-power free for some $\beta \geq 2$,

(iii) $\alpha$ has bounded partial quotients.
Furthermore, we have:

**Theorem**

Let $s_\alpha$ be a Sturmian word of angle $\alpha$. The following are equivalent:

(i) $A(s_\alpha)$ is finite,
(ii) $s_\alpha$ is $\beta$-power free for some $\beta \geq 2$,
(iii) $\alpha$ has bounded partial quotients.

**Theorem**

For every Sturmian word $s_\alpha$ of angle $\alpha$, we have that $A(s_\alpha) \geq \sqrt{5}$. Moreover, $A(s_\alpha) = \sqrt{5}$ if and only if $\alpha$ is equivalent to $\phi - 1$. In particular, the abelian critical exponent of the Fibonacci word is $\sqrt{5}$. 

Gabriele Fici
On some Combinatorial and Algorithmic Aspects of Strings
Concerning the Fibonacci infinite word, we proved the following result, which generalizes to the abelian setting a known result for classical periods.

**Theorem**

*The abelian period of any factor of the Fibonacci word is a Fibonacci number.*
Concerning the Fibonacci infinite word, we proved the following result, which generalizes to the abelian setting a known result for classical periods.

**Theorem**

The abelian period of any factor of the Fibonacci word is a Fibonacci number.

What can be said in general about the abelian period of a factor of a Sturmian word of slope $\alpha$?
Prefix Normal Words
Least Number of Palindromes in an Infinite Word
Bispecial Sturmian Words
Trapezoidal Words
The Array of Open and Closed Prefixes
Abelian Repetitions in Sturmian Words
Universal Lyndon Words
Cyclic Complexity
Given an order $\triangleleft$ on the letters of alphabet $\Sigma$, and a primitive word $w$ over $\Sigma$, we can sort the conjugates of $w$ in lexicographic order. The least element is called a Lyndon word for the order $\triangleleft$. 
Given an order $\prec$ on the letters of alphabet $\Sigma$, and a primitive word $w$ over $\Sigma$, we can sort the conjugates of $w$ in lexicographic order.

The least element is called a Lyndon word for the order $\prec$.

**Example**

Let $w = abcabb$. For the order $a < b < c$, the word $w$ has a smaller conjugate, $abbabc$, which is in fact Lyndon.

Nevertheless, for the order $a < c < b$, the word $w$ is the smallest word in its conjugacy class.
From now on, our alphabet will be $\Sigma_n = \{1, 2, \ldots, n\}$.

Problem

*Given $w$ over $\Sigma_n$, is it possible that for every order on $\Sigma_n$ there is exactly one conjugate of $w$ that is Lyndon for that order?*

*In other words, is it always possible to build a word over $\Sigma_n$ of length $n!$ having $n!$ many Lyndon conjugates?*
From now on, our alphabet will be $\Sigma_n = \{1, 2, \ldots, n\}$.

**Problem**

*Given* $w$ over $\Sigma_n$, *is it possible that for every order on* $\Sigma_n$ *there is exactly one conjugate of* $w$ *that is Lyndon for that order?*

*In other words, is it always possible to build a word over* $\Sigma_n$ *of length* $n!$ *having* $n!$ *many Lyndon conjugates?*

We call such a word a **Universal Lyndon Word (ULW)** of degree $n$. 

From now on, our alphabet will be $\Sigma_n = \{1, 2, \ldots, n\}$.

**Problem**

*Given* $w$ *over* $\Sigma_n$, *is it possible that for every order on* $\Sigma_n$ *there is exactly one conjugate of* $w$ *that is Lyndon for that order?*

*In other words, is it always possible to build a word over* $\Sigma_n$ *of length* $n!$ *having* $n!$ *many Lyndon conjugates?*

We call such a word a **Universal Lyndon Word (ULW)** of degree $n$.

**Example**

For $n = 2$, $w = 12$ is a ULW.
For $n = 3$, take $w = 123132$. 
From now on, our alphabet will be $\Sigma_n = \{1, 2, \ldots, n\}$.

**Problem**

Given $w$ over $\Sigma_n$, is it possible that for every order on $\Sigma_n$ there is exactly one conjugate of $w$ that is Lyndon for that order?

In other words, is it always possible to build a word over $\Sigma_n$ of length $n!$ having $n!$ many Lyndon conjugates?

We call such a word a **Universal Lyndon Word (ULW)** of degree $n$.

**Example**

For $n = 2$, $w = 12$ is a ULW.
For $n = 3$, take $w = 123132$.

**Question**

*Is it possible to build ULW of any degree?*
Definition (Chung, Diaconis, Graham 1992)

A ucycle (universal cycle) is a circular word containing every object of a particular type exactly once as a cyclic factor.
Definition (Chung, Diaconis, Graham 1992)

A **ucycle (universal cycle)** is a circular word containing every object of a particular type exactly once as a cyclic factor.

An example of ucycles are the binary **de Bruijn words**, which are circular words of length $2^n$ containing every binary word of length $n$ exactly once, e.g. $w = 1111211221212222$ contains all binary words of length 4 once.
Definition (Chung, Diaconis, Graham 1992)

A **ucycle (universal cycle)** is a circular word containing every object of a particular type exactly once as a cyclic factor.

An example of ucycles are the binary **de Bruijn words**, which are circular words of length $2^n$ containing every binary word of length $n$ exactly once, e.g. $w = 1111211221212222$ contains all binary words of length 4 once.

By the way, we have the following remarkable result:

Theorem (Fredricksen and Maiorana, 1978)

*The lexicographically smallest de Bruijn word of degree $n$ is obtained by concatenating, in lexicographic order, the Lyndon words of length dividing $n$.*

\[ w = 1 \cdot 1112 \cdot 1122 \cdot 12 \cdot 1222 \cdot 2 \]
If the objects are the linear orders on $n$ letters or, equivalently, the permutations of $n$ letters, there is no way to build ucycles for $n > 2$.

$$123 \cdot 1 \cdot 2 \cdot 3 \ldots \text{ loop!}$$
If the objects are the linear orders on \( n \) letters or, equivalently, the permutations of \( n \) letters, there is no way to build ucycles for \( n > 2 \).

\[
123 \cdot 1 \cdot 2 \cdot 3 \ldots \text{ loop!}
\]

But if one represents the order \( a_1 < a_2 < \cdots < a_n \) by its “shorthand encoding” \( a_1 a_2 \cdots a_{n-1} \), Jackson in 1993 showed that the corresponding ucycles, called shorthand ucycles for permutations, exist for every \( n \) and are the Eulerian cycles of a particular digraph, called the Jackson Graph.
The Jackson graph of degree \( n \geq 4 \) is a digraph in which:

- the vertices are the words over \( \Sigma_n \) that are permutations of \( n - 2 \) letters;
- there is an edge from vertex \( u \) to vertex \( v \) if and only if the suffix of length \( n - 3 \) of \( u \) is equal to the prefix of length \( n - 3 \) of \( v \) and the first letter of \( u \) is different from the last letter of \( v \).

The label of such an edge is set to the first letter of \( u \).
Proposition

*If* $w$ *is a shorthand ucycle for permutations, i.e., an Eulerian cycle in a Jackson graph, then* $w$ *is a ULW.*

Corollary

*There exist ULW of any degree.*

Shorthand ucycles for permutations $\subseteq$ ULW. Is the inclusion proper?

Yes, e.g. 123431242314132421343214.
Proposition

Let \( w \) be a word over \( \Sigma_n \). Then \( w \) is a ULW if and only if for every cyclic factor \( u \) of \( w \), one has

\[
|w|_u^c = (n - |\text{alph}(u)|)! \tag{5}
\]

where \( |\text{alph}(u)| \) is the number of distinct letters in \( u \) and \( |w|_u^c \) is the number of cyclic occurrences of \( u \) in \( w \).
Proposition

Let $w$ be a word over $\Sigma_n$. Then $w$ is a ULW if and only if for every cyclic factor $u$ of $w$, one has

$$|w|_u^{c} = (n - |\text{alph}(u)|)!$$

(5)

where $|\text{alph}(u)|$ is the number of distinct letters in $u$ and $|w|_u^{c}$ is the number of cyclic occurrences of $u$ in $w$.

Example

$w = 123431242314132421343214$ is a ULW, $|\text{alph}(343)| = 2$, so $343$ must appear in $w$ exactly $(4 - 2)! = 2$ times.
Proposition

Let $w$ be a word over $\Sigma_n$. Then $w$ is a ULW if and only if for every cyclic factor $u$ of $w$, one has

$$|w|_u^c = (n - |\text{alph}(u)|)!$$

(5)

where $|\text{alph}(u)|$ is the number of distinct letters in $u$ and $|w|_u^c$ is the number of cyclic occurrences of $u$ in $w$.

Example

$w = 123431242314132421343214$ is a ULW, $|\text{alph}(343)| = 2$, so 343 must appear in $w$ exactly $(4 - 2)! = 2$ times.

Corollary

The reversal of a ULW is a ULW.
Definition

A set $X \subseteq \Sigma_n^+$ is a **lex-code** of order $n$ if:

1. for every $x \in X$, there exists a unique ordering of $\Sigma_n$ such that $x$ is the lexicographical minimum of $X$ (and, consequently, $|X| = n!$);

2. if $u$ is a proper prefix of some word of $X$, then $u$ is a prefix of at least two distinct words of $X$. 
Definition

A set \( X \subseteq \Sigma^+_n \) is a **lex-code** of order \( n \) if:

1. for every \( x \in X \), there exists a unique ordering of \( \Sigma_n \) such that \( x \) is the lexicographical minimum of \( X \) (and, consequently, \(|X| = n!\));

2. if \( u \) is a proper prefix of some word of \( X \), then \( u \) is a prefix of at least two distinct words of \( X \).

A lex-code \( X \) is a **Hamiltonian lex-code** if the relation

\[
S_X = \{ (x, y) \in X \times X \mid \exists a \in \Sigma, u \in \Sigma^*, xu = ay \} \tag{6}
\]

has a Hamiltonian digraph.
Let $h_w$ denote the shortest unrepeated prefix of $w$. For instance, if $w = 1231121$, then $h_w = 123$. 

Theorem

Let $w$ be a ULW. Then the set $X = \{ h_w^i \mid w^i$ conjugate of $w \}$ (7) is a Hamiltonian lex-code. Conversely, if $X$ is a Hamiltonian lex-code, then there exists a ULW $w$ such that (7) holds true.
Let $h_w$ denote the shortest unrepeated prefix of $w$. For instance, if $w = 1231121$, then $h_w = 123$.

**Theorem**

Let $w$ be a ULW. Then the set

$$X = \{ h_{w_i} \mid w_i \text{ conjugate of } w \}$$  \hspace{1cm} (7)

is a Hamiltonian lex-code.

Conversely, if $X$ is a Hamiltonian lex-code, then there exists a ULW $w$ such that (7) holds true.
Given a word $w$ over $\Sigma_n$, every conjugate of $w$ defines an order on $\Sigma_n$, namely the order of appearance of the letters in the conjugate.
Given a word $w$ over $\Sigma_n$, every conjugate of $w$ defines an order on $\Sigma_n$, namely the order of appearance of the letters in the conjugate.

**Definition**

A word $w$ over $\Sigma_n$ of length $n!$ such that every conjugate defines a different order is called a **Universal Order Word (UOW)**.
Given a word \( w \) over \( \Sigma_n \), every \textit{conjugate} of \( w \) defines an \textit{order} on \( \Sigma_n \), namely the order of appearance of the letters in the conjugate.

**Definition**

A word \( w \) over \( \Sigma_n \) of length \( n! \) such that every conjugate defines a different order is called a \textit{Universal Order Word (UOW)}.

**Remark**

\textit{Every ULW is a UOW.}
Given a word $w$ over $\Sigma_n$, every conjugate of $w$ defines an order on $\Sigma_n$, namely the order of appearance of the letters in the conjugate.

**Definition**

A word $w$ over $\Sigma_n$ of length $n!$ such that every conjugate defines a different order is called a Universal Order Word (UOW).

**Remark**

*Every ULW is a UOW.*

Is the converse true?
The answer is no!

Example

The word

\[ w = 123421323121424314324134 \]

is a UOW but is not a ULW (e.g. because 323 appears only once).
The answer is no!

**Example**

The word

\[ w = 123421323121424314324134 \]

is a UOW but is not a ULW (e.g. because 323 appears only once).
The answer is no!

Example

The word

$$w = 123421\textbf{323}121424314324134$$

is a UOW but is not a ULW (e.g. because 323 appears only once).

Problem

Provide combinatorial characterizations for UOW.
Prefix Normal Words
Least Number of Palindromes in an Infinite Word
Bispecial Sturmian Words
Trapezoidal Words
The Array of Open and Closed Prefixes
Abelian Repetitions in Sturmian Words
Universal Lyndon Words
Cyclic Complexity

Prefix Normal Words
Least Number of Palindromes in an Infinite Word
Bispecial Sturmian Words
Trapezoidal Words
The Array of Open and Closed Prefixes
Abelian Repetitions in Sturmian Words
Universal Lyndon Words
Cyclic Complexity

1. Prefix Normal Words
2. Least Number of Palindromes in an Infinite Word
3. Bispecial Sturmian Words
4. Trapezoidal Words
5. The Array of Open and Closed Prefixes
6. Abelian Repetitions in Sturmian Words
7. Universal Lyndon Words
8. Cyclic Complexity
Two finite words $u, v$ are **conjugate** if there exist words $w_1, w_2$ such that $u = w_1 w_2$ and $v = w_2 w_1$. The conjugacy relation is an equivalence over $A^*$, which is denoted by $\sim$, whose classes are called **conjugacy classes**.
Two finite words $u, v$ are **conjugate** if there exist words $w_1, w_2$ such that $u = w_1w_2$ and $v = w_2w_1$. The conjugacy relation is an equivalence over $A^*$, which is denoted by $\sim$, whose classes are called **conjugacy classes**.

Note that two words belonging to the same conjugacy class necessarily have the same Parikh vector.
Two finite words $u, v$ are **conjugate** if there exist words $w_1, w_2$ such that $u = w_1w_2$ and $v = w_2w_1$. The conjugacy relation is an equivalence over $A^*$, which is denoted by $\sim$, whose classes are called **conjugacy classes**.

Note that two words belonging to the same conjugacy class necessarily have the same Parikh vector.

**Definition**

The **cyclic complexity** of an infinite word $\omega$ is the function

$$c_\omega(n) = \left| \frac{\text{Fact}(\omega) \cap A^n}{\sim} \right|,$$

i.e., the function that counts the number of distinct conjugacy classes of factors of length $n$ of $\omega$, for every $n \geq 0$. 

Gabriele Fici

On some Combinatorial and Algorithmic Aspects of Strings
Theorem (Morse-Hedlund, 1938)

A word $\omega$ is ultimately periodic if and only if it has bounded factor complexity.

We established the following analogue of the Morse-Hedlund theorem:

Theorem

A word $\omega$ is ultimately periodic if and only if it has bounded cyclic complexity.
The factor complexity does not distinguish between Sturmian words of different slopes. In contrast, for cyclic complexity the situation is quite different:

**Theorem**

Let $x$ be a Sturmian word. If $y$ is an infinite word whose cyclic complexity is equal to that of $x$, then up to renaming letters, $x$ and $y$ have the same set of factors. In particular, $y$ is also Sturmian.
It is easy to prove that a word whose cyclic complexity takes value 1 at some point must be ultimately periodic.

This motivates us to give the following:

**Definition**

We say that an aperiodic word $x$ has **minimal cyclic complexity** if

$$\liminf_{n \to \infty} c_x(n) = 2.$$
Proposition

Let $\Sigma = \{0, 1\}$ and $\mu : 0 \mapsto u0v, 1 \mapsto u1v$, for words $u, v \in \Sigma^*$ such that $|uv| > 0$. Let $x$ be a fixed point of $\mu$. If $x$ is aperiodic, then for every $n \geq 0$ one has $c_x(k^n) = 2$, where $k = |u| + |v| + 1$.

Example

If we take $u = 0$ and $v = \varepsilon$, we obtain the morphism $\mu : 0 \mapsto 00, 1 \mapsto 01$, whose fixed point is the so-called period-doubling word $p = 0100010101000100 \cdots$. By previous proposition, we have $c_p(2^n) = 2$ for every $n \geq 0$. 

Gabriele Fici
On some Combinatorial and Algorithmic Aspects of Strings
A **paperfolding word** is the sequence of ridges and valleys obtained by unfolding a sheet of paper which has been folded infinitely many times.

The regular paperfolding word

$$w_R = 0010011000110110001001110011011010110 \cdots$$

is obtained by folding at each step in the same way.

The regular paperfolding word can also be obtained through the 2-letter substitution rule $\tau$ defined by $\tau(ab) = 0a1b$ for every choice of $a, b \in \{0, 1\}$, beginning from 00. So $\tau(00) = 0010$, $\tau^2(00) = \tau(00)\tau(10) = 00100110$, $\tau^3(00) = \tau(00)\tau(10)\tau(01)\tau(10) = 0010011000110110$, etc. One has $w_R = \lim_{n \to \infty} \tau^n(00)$. 
A **paperfolding word** is the sequence of ridges and valleys obtained by unfolding a sheet of paper which has been folded infinitely many times.

The regular paperfolding word

$$w_R = 00100110001101100010011100110110 \cdots$$

is obtained by folding at each step in the same way.

The regular paperfolding word can also be obtained through the 2-letter substitution rule $\tau$ defined by $\tau(ab) = 0a1b$ for every choice of $a, b \in \{0, 1\}$, beginning from 00. So $\tau(00) = 0010$, $\tau^2(00) = \tau(00)\tau(10) = 00100110$, $\tau^3(00) = \tau(00)\tau(10)\tau(01)\tau(10) = 0010011000110110$, etc. One has $w_R = \lim_{n \to \infty} \tau^n(00)$.

**Proposition**

*Let $w$ be a paperfolding word. Then for every $n \geq 1$ one has $c_w(4 \cdot 2^n) = 4$.***
Let

\[ t = t_0 t_1 t_2 \cdots = 011010011001011010010110 \cdots \]

be the Thue-Morse word, i.e., the fixed point beginning in 0 of the uniform substitution \( \mu : 0 \mapsto 01, 1 \mapsto 10. \)

For the Thue-Morse word, we have the following result:

**Proposition**

Let \( t \) be the Thue-Morse word. Then \( \lim \inf_{n \to \infty} c_t(n) = +\infty. \)
Let
\[ t = t_0 t_1 t_2 \cdots = 0110100110010110010110 \cdots \]
be the Thue-Morse word, i.e., the fixed point beginning in 0 of the uniform substitution \( \mu : 0 \mapsto 01, 1 \mapsto 10 \).

For the Thue-Morse word, we have the following result:

**Proposition**

*Let \( t \) be the Thue-Morse word. Then \( \lim \inf_{n \to \infty} c_t(n) = +\infty \).*

**Problem**

*Find combinatorial characterizations of words with minimal cyclic complexity.*
In the next future, I hope to do the following:

1. Searching for other combinatorial problems about prefix normal words and develop new algorithms;
2. Using the array of open/closed prefixes as a data structure for algorithms on strings;
3. Deepening the study of abelian properties of Sturmian words;
4. Giving combinatorial characterizations, generating algorithms and enumerative formulae for Universal Order Words;
5. Continuing the study of the cyclic complexity also in comparison with other measures of complexity;
6. etc. etc. etc.
Thank You