Abelian returns in Sturmian words

S. Puzynina
jointly with L. Q. Zamboni
Periodicity

Σ – alphabet
Σ* – finite words over Σ
Σω – (right) infinite words over Σ

A word $w$ is periodic, if there exists $T$ such that $w^n + T = w^n$ for every $n$.

The subword complexity of a word is the function $f(n)$ defined as the number of its factors of length $n$.

Sturmian words are infinite words having the smallest subword complexity among aperiodic words, for Sturmian words $f(n) = n + 1$.

$w \in \Sigma^\omega$ is recurrent if each of its factors occurs infinitely many times in $w$.

$F(w)$: the set of factors of a finite or infinite word $w$.

S. Puzynina jointly with L. Q. Zamboni Abelian returns in Sturmian words
Periodicity

Σ – alphabet
Σ* – finite words over Σ
Σω – (right) infinite words over Σ

- A word $w$ is periodic, if there exists $T$ such that $w_{n+T} = w_n$ for every $n$. 

The subword complexity of a word is the function $f(n)$ defined as the number of its factors of length $n$.

Sturmian words are infinite words having the smallest subword complexity among aperiodic words, for Sturmian words $f(n) = n + 1$.

$w \in \Sigma^\omega$ is recurrent if each of its factors occurs infinitely many times in $w$.

$F(w)$: the set of factors of a finite or infinite word $w$.
Periodicity

Σ – alphabet
Σ* – finite words over Σ
Σω – (right) infinite words over Σ

- A word \( w \) is **periodic**, if there exists \( T \) such that \( w_{n+T} = w_n \) for every \( n \).

- The **subword complexity** of a word is the function \( f(n) \) defined as the number of its factors of length \( n \).

- Sturmian words are infinite words having the smallest subword complexity among aperiodic words, for Sturmian words \( f(n) = n + 1 \).
Periodicity

$\Sigma$ – alphabet
$\Sigma^*$ – finite words over $\Sigma$
$\Sigma^\omega$ – (right) infinite words over $\Sigma$

- A word $w$ is periodic, if there exists $T$ such that $w_{n+T} = w_n$ for every $n$.

- The subword complexity of a word is the function $f(n)$ defined as the number of its factors of length $n$.

- Sturmian words are infinite words having the smallest subword complexity among aperiodic words, for Sturmian words $f(n) = n + 1$.

- $w \in \Sigma^\omega$ is recurrent if each of its factors occurs infinitely many times in $w$. 
Periodicity

$\Sigma$ – alphabet
$\Sigma^*$ – finite words over $\Sigma$
$\Sigma^\omega$ – (right) infinite words over $\Sigma$

- A word $w$ is **periodic**, if there exists $T$ such that $w_{n+T} = w_n$ for every $n$.

- The **subword complexity** of a word is the function $f(n)$ defined as the number of its factors of length $n$.

- Sturmian words are infinite words having the smallest subword complexity among aperiodic words, for Sturmian words $f(n) = n + 1$.

- $w \in \Sigma^\omega$ is **recurrent** if each of its factors occurs infinitely many times in $w$.

- $F(w)$: the set of factors of a finite or infinite word $w$
Definition

\( w = w_1 w_2 \ldots \) a recurrent infinite word, \( u \in F(w) \),

let \( n_1 < n_2 < \ldots \) be all integers \( n_i \) such that \( u = w_{n_i} \ldots w_{n_{i+1}} + |u| - 1 \)

\( w_{n_i} \ldots w_{n_{i+1}} - 1 \) is a return word (or briefly return) of \( u \) in \( w \)
Definition

\[ w = w_1 w_2 \ldots \text{ a recurrent infinite word, } u \in F(w), \]
\[ \text{let } n_1 < n_2 < \ldots \text{ be all integers } n_i \text{ such that } u = w_{n_i} \ldots w_{n_i+|u|-1} \]
\[ w_{n_i} \ldots w_{n_i+1-1} \text{ is a return word (or briefly return) of } u \text{ in } w \]

- introduced independently by F. Durand, C. Holton and L. Q. Zamboni, 1998, and used for a characterization of primitive substitutive sequences
- and then for different problems of combinatorics on words, symbolic dynamical systems and number theory (L. Vuillon, J. Justin, J.-P. Allouche, J. D. Davinson, M. Queffélec, I. Fagnot, J. Cassaigne...)

S. Puzynina jointly with L. Q. Zamboni
Abelian returns in Sturmian words
An infinite word has $k$ returns, if each of its factors has $k$ returns.

A characterization of Sturmian words via return words:

**Theorem (L. Vuillon, J. Justin, 2000–2001)**

A recurrent infinite word has two returns if and only if it is Sturmian.
Characterization of periodicity via return words:

**Proposition (L. Vuillon, 2001)**

A recurrent infinite word is ultimately periodic if and only if there exists a factor having exactly one return word.
Abelian returns

$u \in \Sigma^*$, $a \in \Sigma$, $|u|_a$ – the number of occurrences of the letter $a$ in $u$

$u, v \in \Sigma^*$ are abelian equivalent if $|u|_a = |v|_a$ for all $a \in \Sigma$
Abelian returns

\(u \in \Sigma^*, \ a \in \Sigma, \ |u|_a - \text{the number of occurrences of the letter } a \text{ in } u\)

\(u, v \in \Sigma^* \text{ are abelian equivalent if } |u|_a = |v|_a \text{ for all } a \in \Sigma\)

**Definition**

\(w \text{ an infinite recurrent word, } u \in F(w), \ n_1 < n_2 < \ldots \text{ all integers } n_i \text{ such that } w_{n_i} \ldots w_{n_i+|u|-1} \approx^{ab} u\)

\(w_{n_i} \ldots w_{n_i+1-1} \) is an abelian return word (or briefly abelian return) of \(u\) in \(w\)
Abelian returns

\( u \in \Sigma^*, \, a \in \Sigma, \, |u|_a \) – the number of occurrences of the letter \( a \) in \( u \)

\( u, v \in \Sigma^* \) are abelian equivalent if \( |u|_a = |v|_a \) for all \( a \in \Sigma \)

**Definition**

\( w \) an infinite recurrent word, \( u \in F(w) \), 
\( n_1 < n_2 < \ldots \) all integers \( n_i \) such that \( w_{n_i} \ldots w_{n_i+|u|-1} \approx^{ab} u \)

\( w_{n_i} \ldots w_{n_i+1-1} \) is an abelian return word (or briefly abelian return) of \( u \) in \( w \)

\( u \) has \( k \) abelian returns in \( w \), if the set of abelian returns of \( u \) consists of \( k \) abelian classes

I. e., we take

- factors up to abelian equivalence
- return words up to abelian equivalence
Example: the Thue-Morse word

the Thue-Morse word

\[ t = 0110100110010110 \ldots \]

Consider abelian returns of its factor 01:
Example: the Thue-Morse word

\[ t = 0110100110010110 \ldots \]

Consider abelian returns of its factor 01:

- 0 ab. ret. 0
- 10 ab. ret. 01
Example: the Thue-Morse word

$t = 0110100110010110 \ldots$

Consider abelian returns of its factor 01: symmetrically

\[
\begin{array}{ccc}
0 & \text{ab. ret. 0} & 0 & \text{ab. ret. 1} \\
01 & \text{ab. ret. 01} & 10 & \text{ab. ret. 10}
\end{array}
\]
Example: the Thue-Morse word

\[ t = 0110100110010110 \ldots \]

Consider abelian returns of its factor 01:

\[
\begin{align*}
0 & \quad \text{ab. ret. 0} & 0 & \quad \text{ab. ret. 1} \\
01 & \quad \text{symmetrically} \quad 10 & \quad 01 & \quad \text{ab. ret. 10} \\
10 & \quad \text{ab. ret. 01} & 01 & \quad \text{ab. ret. 10} \\
\end{align*}
\]

two abelian returns: 0, 1 and 01 $\sim^{ab} 10$. 

S. Puzynina jointly with L. Q. Zamboni

Abelian returns in Sturmian words
Main result

A characterization of Sturmian words via the abelian returns:

Theorem

An aperiodic recurrent infinite word is Sturmian if and only if each of its factors has two or three abelian returns.
A characterization of Sturmian words via the abelian returns:

**Theorem**

An aperiodic recurrent infinite word is Sturmian if and only if each of its factors has two or three abelian returns.

Remind a characterization by Vuillon:
Sturmian $\iff$ each factor has two (normal) returns
Idea of proof

Sturmian $\Rightarrow$ two or three returns


- $w \in \{0, 1\}^q$ balanced word with $|w|_1 = p$, $\gcd(p, q) = 1$. 

---

S. Puzynina jointly with L. Q. Zamboni

Abelian returns in Sturmian words
Idea of proof

Sturmian $\Rightarrow$ two or three returns


- $w \in \{0, 1\}^q$ balanced word with $|w|_1 = p$, $\gcd(p, q) = 1$.
- The shift $\sigma : \{0, 1\}^q \rightarrow \{0, 1\}^q$: $\sigma(w_0 \ldots w_{q-1}) = w_1 \ldots w_{q-1}w_0$. 

### Idea of proof

**Sturmian ⇒ two or three returns**


- \( w \in \{0, 1\}^q \) balanced word with \( |w|_1 = p, \gcd(p, q) = 1 \).
- the *shift* \( \sigma : \{0, 1\}^q \to \{0, 1\}^q \):
  \[
  \sigma(w_0 \ldots w_{q-1}) = w_1 \ldots w_{q-1}w_0.
  \]
- the lexicographic ordering of \( \{\sigma^i(w) : 0 \leq i < q\} \):
  \[
  w(0) <_L w(1) <_L \cdots <_L w(q-1)
  \]

*Lexicographic array* \( A[w] \): \( q \times q \) matrix whose \( i \)th row is \( w(i) \)
Idea of proof

**Sturmian \( \Rightarrow \) two or three returns**

- \( w \in \{0, 1\}^q \) balanced word with \( |w|_1 = p, \gcd(p, q) = 1 \).
- the *shift* \( \sigma : \{0, 1\}^q \rightarrow \{0, 1\}^q \):
  \[
  \sigma(w_0 \ldots w_{q-1}) = w_1 \ldots w_{q-1}w_0.
  \]
- the lexicographic ordering of \( \{ \sigma^i(w) : 0 \leq i < q \} \):
  \[
  w(0) <_L w(1) <_L \cdots <_L w(q-1)
  \]

*Lexicographic array* \( A[w] \): \( q \times q \) matrix whose \( i \)th row is \( w(i) \)

- \( |\text{pref}_j(w(i))|_1 \leq |\text{pref}_j(w(i+1))|_1 \) for all \( 0 \leq i \leq q - 2 \), \( 0 \leq j \leq q - 1 \)
Idea of proof

Sturmian \(\Rightarrow\) two or three returns


- \(w \in \{0, 1\}^q\) balanced word with \(|w|_1 = p, \gcd(p, q) = 1\).
- the shift \(\sigma : \{0, 1\}^q \to \{0, 1\}^q\):
  \(\sigma(w_0 \ldots w_{q-1}) = w_1 \ldots w_{q-1}w_0\).
- the lexicographic ordering of \(\{\sigma^i(w) : 0 \leq i < q\}\):
  \(w(0) <_L w(1) <_L \cdots <_L w(q-1)\)

**Lexicographic array** \(A[w] : q \times q\) matrix whose \(i\)th row is \(w(i)\)

- \(|\text{pref}_j(w(i))|_1 \leq |\text{pref}_j(w(i+1))|_1\) for all \(0 \leq i \leq q - 2, 0 \leq j \leq q - 1\)
- \(j\)-th column of \(A\) is \(\sigma^{jp}u\), where \(u = 0^{q-p}1^p\)
Example

Consider a balanced word $w = 0101001$. 
Idea of proof

Example

Consider a balanced word \( w = 0101001 \).

The lexicographic ordering of \( \{\sigma^i(w) | i = 0, \ldots, 6\} \):

\[
0010101 <_L 0100101 <_L 0101001 <_L 0101010 < \\
<_L 1001010 <_L 1010010 <_L 1010100,
\]
Idea of proof

Example

Consider a balanced word \( w = 0101001 \).

The lexicographic ordering of \( \{\sigma^i(w)|i = 0, \ldots, 6\} \):

\[
0010101 <_L 0100101 <_L 0101001 <_L 0101010 < \]
\[
<_L 1001010 <_L 1010010 <_L 1010100,
\]

The lexicographic array:

\[
A[w] = \begin{pmatrix}
0 & 0 & 1 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0
\end{pmatrix}
\]
Idea of proof

two or three returns $\Rightarrow$ Sturmian

considering abelian returns to factors of special type and restricting
the form of words (quite technical)
Abelian returns of factors of a Sturmian word are either letters or of the form $aBb$, where $a \neq b$ are letters, and $B$ is a bispecial factor.
Abelian returns of factors of a Sturmian word are either letters or of the form $aBb$, where $a \neq b$ are letters, and $B$ is a bispecial factor.

In a Sturmian word for each length $l \geq 2$ there exists at most one abelian return of length $l$. 

A factor of a Sturmian word has two abelian returns if and only if it is singular. A factor of a Sturmian word is called singular if it is the only factor in its abelian class. 

If a factor of a Sturmian word has three abelian returns of lengths $l_1 \leq l_2 < l_3$, then $l_3 = l_1 + l_2$. 

S. Puzynina jointly with L. Q. Zamboni
Abelian returns of factors of a Sturmian word are either letters or of the form $aBb$, where $a \neq b$ are letters, and $B$ is a bispecial factor.

In a Sturmian word for each length $l \geq 2$ there exists at most one abelian return of length $l$.

A factor of a Sturmian word has two abelian returns if and only if it is singular.

A factor of a Sturmian word is called singular if it is the only factor in its abelian class.
Abelian returns of factors of a Sturmian word are either letters or of the form $aBb$, where $a \neq b$ are letters, and $B$ is a bispecial factor.

In a Sturmian word for each length $l \geq 2$ there exists at most one abelian return of length $l$.

A factor of a Sturmian word has two abelian returns if and only if it is singular.

A factor of a Sturmian word is called singular if it is the only factor in its abelian class.

If a factor of a Sturmian word has three abelian returns of lengths $l_1 \leq l_2 < l_3$, then $l_3 = l_1 + l_2$.
A simple sufficient condition for periodicity via abelian returns:

**Lemma**

Let $|\Sigma| = k$. If each factor of a recurrent infinite word over the alphabet $\Sigma$ has at most $k$ abelian returns, then the word is periodic.

**Remark:** not necessary condition for periodicity!
Remind a characterization of periodicity by L. Vuillon, 2001:

periodic $\iff$ there exists a factor having exactly one return word
Remind a characterization of periodicity by L. Vuillon, 2001:

\[ \text{periodic} \iff \text{there exists a factor having exactly one return word} \]

No similar characterization of periodicity in terms of abelian returns exist.

Moreover, in the case of abelian returns it does not hold in both directions.
Remind a characterization of periodicity by L. Vuillon, 2001:

\[
\text{periodic } \iff \text{there exists a factor having exactly one return word}
\]

No similar characterization of periodicity in terms of abelian returns exist.

Moreover, in the case of abelian returns it does not hold in both directions.

\[
\exists \text{ factor having one abelian return } \not\Rightarrow \text{periodicity}
\]

Example: an infinite aperiodic word in \(\{110010, 110100\}\omega\)
the factor 11 has one abelian return 110010 \(\approx_{ab} 110100\)
Remind a characterization of periodicity by L. Vuillon, 2001:

\[
\text{periodic } \iff \text{there exists a factor having exactly one return word}
\]

No similar characterization of periodicity in terms of abelian returns exist.

Moreover, in the case of abelian returns it does not hold in both directions.

\[
\exists \text{ factor having one abelian return } \not\implies \text{periodicity}
\]

Example: an infinite aperiodic word in \(\{110010, 110100\}^\omega\)
the factor 11 has one abelian return \(110010 \approx_{ab} 110100\)

\[
\text{periodicity } \not\implies \exists \text{ factor having one abelian return}
\]

Example: \(w = (001101001011001100110011)^\omega\)
Another version:

**Definition**

$w$ infinite recurrent word, $u \in F(w)$

$w$ has $k$ returns to the abelian class of $u$, if the set of abelian returns of $u$ consists of $k$ different words.
Returns to abelian classes

Another version:

**Definition**

$w$ infinite recurrent word, $u \in F(w)$

$w$ has $k$ returns to the abelian class of $u$, if the set of abelian returns of $u$ consists of $k$ different words.

I. e., now we take

- factors up to abelian equivalence
- return words NOT up to abelian equivalence

S. Puzynina jointly with L. Q. Zamboni

Abelian returns in Sturmian words
Example

The Thue-Morse word

\[ t = 0110100110010110 \ldots \]

4 returns to the abelian class of 01 of \( t \): 0, 1, 01, 10.
The Thue-Morse word

\[ t = 0110100110010110 \ldots \]

4 returns to the abelian class of 01 of \( t \): 0, 1, 01, 10.

Remind: in the sense of previous definition 01 has 3 abelian returns, because \( 01 \approx_{ab} 10 \).
In the sense of the second definition the characterization also holds:

**Theorem**

An aperiodic recurrent infinite word $w$ is Sturmian if and only if for each $u \in F(w)$ the word $w$ has two or three returns to the abelian class of $u$. 
Thank you!