Formal intercepts of Sturmian words
and Prefix-Suffix duality for low complexity words

Caius Wojcik - ICJ - Lyon 1

Journées Montoises - LaBRI Bordeaux

10 September 2018
Notations:

- \( \mathcal{A} \) : any set, the alphabet.
- \( \mathcal{A}^+ \) : finite words over \( \mathcal{A} \).
- \( \mathcal{A}^\mathbb{N} \) : infinite words over \( \mathcal{A} \).
- \(|u|\) : length of a finite word \( u \).
- \( \tilde{u} \) : mirror image of a word \( u \).
- \( \mathbb{P}_n(x) \) : prefix of length \( n \geq 1 \) of an infinite word \( x \),
- \( T : \mathcal{A}^\mathbb{N} \longrightarrow \mathcal{A}^\mathbb{N} \) the shift on infinite words
- \( T^n(x) \) : the \( n \)-th suffix of an infinite word \( x \).
1 Introduction

2 Basic properties of Sturmian words

3 Rauzy graphs and repetition function

4 Formal intercepts of Sturmian words

5 Factorisation, extensions and prefix-suffix duality
1 Introduction

2 Basic properties of Sturmian words

3 Rauzy graphs and repetition function

4 Formal intercepts of Sturmian words

5 Factorisation, extensions and prefix-suffix duality
Introduction:

• Thermodynamic has its 1st and 2nd laws,
• Algebra has a fundamental theorem,
• Calculus has a fundamental principle,

Question: What would look like a fundamental principle of combinatorics on words?

Of course, there are an infinite number of answers...
But we introduce and illustrate one of them:

The prefix-suffix duality:

"For any word, the set of its prefixes and the set of its suffixes are in natural bijective correspondence."
Illustration with the Zimin word: The Zimin word $Z$, is defined over the infinite alphabet $A_X = \{x_1, x_2, x_3, \ldots\}$ as

- $Z = \prod_{n \geq 1} x_{\text{val}_2(n)+1}$ where $\text{val}_2$ is the 2-adic valuation.
- $Z = \lim Z_n$ where $Z_1 = x_1$ and $Z_{n+1} = Z_n x_{n+1} Z_n$ for $n \geq 1$.
- $Z = \varphi(Z)$ where $\varphi$ is the morphism defined by $\varphi(x_i) = x_1 x_{i+1}$ for $i \geq 1$.

$Z = x_1 x_2 x_1 x_3 x_1 x_2 x_1 x_4 x_1 x_2 x_1 x_3 x_1 \ldots$
Define the sequences \((U_n)_{n \geq 1}\) and \((V_n)_{n \geq 1}\) with \(U_1 = V_1 = x_1\) and for \(n \geq 1\):

\[
U_{n+1} = x_{n+1} U_1 U_2 \cdots U_n \quad \text{and} \quad V_{n+1} = V_n V_{n-1} \cdots V_1 x_{n+1},
\]

Example: \(U_2 = x_2 x_1\), \(U_3 = x_3 x_1 x_2 x_1\), \(U_4 = x_4 x_1 x_2 x_1 x_3 x_1 x_2 x_1\)

and

\[
Z = \prod_{i \geq 1} U_i
\]

Lemma

Let \(m \geq 1\), written \(m = \sum_{i=0}^{N} b_i 2^i\) in base 2, where the \((b_i)\)'s equal 0 or 1, and equal zero for large \(i \geq 0\). Then:

\[
\tilde{P}_m(Z) = \prod_{i=0}^{N} U_{i+1}^{b_i} \quad \text{and} \quad T^m(Z) = \prod_{i \geq 0} U_{i+1}^{1-b_i}.
\]
Let $\gamma = \sum_{i \geq 0} b_i 2^i$ be a 2-adic number, and define

$$\tilde{P}_\gamma(Z) = \prod_{i=0}^{+\infty} U_{i+1}^{b_i} \quad \text{and} \quad T_\gamma(Z) = \prod_{i=0}^{+\infty} U_{i+1}^{1-b_i}$$

If $\gamma$ is not an integer, then we have the reciprocity formula :

$$T_\gamma(Z) = \tilde{P}_{\bar{\gamma}}(Z)$$

where $\gamma \mapsto \bar{\gamma} = -1 - \gamma$ is an involution on the set of 2-adic numbers that are not integers,

$\bar{\gamma}$ is called the complement of $\gamma$. 
Denote by $\Omega(Z)$ the set of infinite words sharing the same factors as $Z$.

- Every element $Y$ of $\Omega(Z)$ writes uniquely in the form
  \[ Y = \tilde{P}_\gamma(Z) \]
  where $\gamma$ is a 2-adic number that is not a natural number.

- Moreover, for every $\gamma$ a 2-adic number that is not an integer, the bi-infinite word
  \[ \overline{T\gamma}(Z) \cdot T\gamma(Z) \]
  shares the same factors as $Z$.

This defines a natural bijective correspondence between bi-infinite orbits of $Z$ and the set of 2-adic numbers up to integer equivalence.

This is the so-called Prefix-Suffix duality
1 Introduction

2 Basic properties of Sturmian words

3 Rauzy graphs and repetition function

4 Formal intercepts of Sturmian words

5 Factorisation, extensions and prefix-suffix duality
For an infinite word $x$ and a natural integer $n \geq 1$, we define the complexity function:

$$p(x, n) = \text{number of factors of } x \text{ of length } n.$$ 

**Theorem (Morse-Hedlund)**

Let $x$ be an infinite word. The following statements are equivalent

i) the word $x$ is ultimately periodic,

ii) the complexity function $p(x, \cdot)$ is bounded,

iii) there exists $n \geq 1$ such that $p(x, n) = p(x, n + 1)$,

iv) there exists $n \geq 1$ such that $p(x, n) \leq n$.

The word $x$ is said to be **Sturmian** when

$$\forall n \geq 1, \quad p(x, n) = n + 1$$
For $x \in \{0, 1\}^\mathbb{N}$, the following statements are equivalent:

i) The word $x$ is Sturmian,

ii) $x$ is not ultimately periodic, and for every factors $u, v$ of $x$ with $|u| = |v|$, we have the balanced property: $||u|_1 - |v|_1| \leq 1$,

iii) $x$ is not ultimately periodic, and for every factors $u, v$ of $x$, we have the equivalent balanced property:

$$\left| \frac{|u|_1}{|u|} - \frac{|v|_1}{|v|} \right| < \frac{1}{|u|} + \frac{1}{|v|}.$$

iv) $x$ satisfies the balanced property, and the number

$$\alpha = \lim_{|u| \to +\infty} \frac{|u|_1}{|u|}$$

is an irrational number.

The number $\alpha$ is called the **slope** of $x$. 
The slope is the first parameter that describes a Sturmian word.

**Proposition**

- Two Sturmian words of different slopes only share a finite number of factors,
- Two Sturmian words of same slopes have same set of factors,

A word is Sturmian if and only if for all $n \geq 1$ it has exactly one left special factor, noted $L_n$.

To recover the set of factors from the slope $\alpha$, we get from the balanced property applied to $0L_n$ and $1L_n$:

$$\frac{|L_n|_1}{n} - \alpha < \frac{1}{n} \quad \text{and} \quad \frac{1 + |L_n|_1}{n} - \alpha < \frac{1}{n},$$

so that $\{|L_n|_1\} = \mathbb{Z} \cap [n\alpha - 1, n\alpha[$, determining uniquely the sequence of words $(L_n)_{n \geq 1}$. 
We define the characteristic word $c_\alpha$ of the slope $\alpha$ as:

$$c_\alpha = \lim L_n$$

Amongst Sturmian word of slope $\alpha$, the characteristic word is the only one such that both $0c_\alpha$ and $1c_\alpha$ are Sturmian.

We can build the characteristic word with the use of continued fractions: every irrational $\alpha \in ]0, 1[$ writes uniquely as

$$\alpha = [0; a_1, a_2, \ldots] = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \ldots}}}$$

with $a_i \geq 1$ for all $i \geq 1$. 
With $\alpha = [0; a_1, a_2, \ldots]$, we define the standard sequence of finite words:

$$
\begin{align*}
\ s_{-1} &= 1, \quad s_0 = 0, \quad s_1 = s_0^{a_1-1} s_{-1}, \\
\forall n \geq 1, \quad s_{n+1} &= s_n^{a_{n+1}} s_{n-1}
\end{align*}
$$

**Theorem**

$$
c_\alpha = \lim_{n \geq 1} s_n.
$$

- $s_n$ ends with 10 for even $n \geq 2$,
- $s_n$ ends with 01 for odd $n \geq 3$,
- $s_n^{-}$ is a palindrome, where $u^{-}$ denotes a finite word $u$ deprived of its last letter,
Rauzy graph, or factor graph of an infinite word:
Defined as the sequence of directed graphs \((G_m)_{m \geq 1}\) whose:
- vertexes of \(G_m\) are the factors of \(x\) of length \(m\),
- arrows are \(s \rightarrow t\) when there exists a factor \(w\) of \(x\) of length \(m + 1\) and two letters \(a, b\) such that \(w = sb = at\),

For Sturmian words, \(G_m\) has \(m + 1\) vertexes, and is formed as the fusion of two cycles, sharing a common part.

Repetition function:
For an infinite word \(x\) and \(m \geq 1\), define the repetition function \(r(x, \cdot)\) as:

\[
r(x, m) = \max\{k \geq 0 \mid \mathbb{P}_m(x), \mathbb{P}_m(T(x)), \ldots, \mathbb{P}_m(T^{k-1}(x)) \text{ are all distincts} \}
\]
Some results about the repetition function:

- For any infinite word $x$, $r(x, m) \leq p(x, m)$

- Let $x$ be a Sturmian word and $m \geq 2$, then:
  $$r(x, m) = m + 1 \iff r(x, m) \neq r(x, m - 1)$$

- Let $\alpha$ be a slope and $m \geq 2$, then:
  $$P_m(T^{r(c_\alpha, m)}(c_\alpha)) = P_m(c_\alpha) = L_m$$

- Let $z = p01q$ be a palindromic word, with $p$ and $q$ palindromes. Then:
  $$r(z, |p| + 1) = |p| + 2.$$
Define the sequence of continuants \((q_n)_{n \geq -1}\) of \(\alpha\) as the denominators of the irreducible fraction

\[
\frac{p_n}{q_n} = [0; a_1, \ldots, a_n] = \frac{1}{a_1 + \frac{1}{\ldots + \frac{1}{a_n}}}
\]

with \(q_{-1} = 0\) and \(q_0 = 1\). Also : \(q_n = |s_n|\), and the \((q_n)\)'s satisfy the induction relation, for \(n \geq 0\) :

\[
q_{n+1} = a_{n+1}q_n + q_{n-1}
\]

**Theorem**

\[
r(c_\alpha, m) = q_n \quad \text{for} \quad q_n - 1 \leq m \leq q_{n+1} - 2.
\]
Every infinite word $x$ defines a path in the Rauzy graph $G_m$:

$$P_m(x) \rightarrow P_m(T(x)) \rightarrow P_m(T^2(x)) \rightarrow \ldots \rightarrow P_m(T^k(x)) \rightarrow \ldots$$

The value $r(x, m)$ is interpreted in the Rauzy graph $G_m$ as the longest Hamiltonian path in this path, that is the longest path not passing twice on a vertex.

We define the integer intervals for $n \geq 0$, and $1 \leq \ell \leq a_{n+1} - 1$,

$$I_n = [q_n - 1, q_{n+1} - 2] = I_n^0 \cup \bigcup_{\ell=1}^{a_{n+1} - 1} I_n^\ell$$

$$I_n^0 = [q_n - 1, q_n + q_{n-1} - 2]$$

$$I_n^\ell = [\ell q_n + q_{n-1} - 1, (\ell + 1)q_n + q_{n-1} - 2].$$
Proposition

For \( m \in I_n^\ell \), with \( 0 \leq \ell \leq a_{n+1} - 1 \), then

- One cycle in \( G_m \) is of length \( q_n \), called the referent cycle,
- The other one is of length \( \ell q_n + q_{n-1} \)

For any Sturmian word \( x \), \( x \) cannot turn twice around the non-referent cycle.

Path of the characteristic word \( c_\alpha \):

For \( m \in I_n^\ell \) with \( 0 \leq \ell \leq a_{n+1} - 1 \):

- The characteristic word \( c_\alpha \) turns exactly \( a_{n+1} - \ell \) times around the referent cycle, that is:

\[
q_n = r(c_\alpha, m) = r(T^{q_n}(c_\alpha), m) = \ldots = r(T^{(a_{n+1}-\ell-1)q_n}(c_\alpha), m) \neq r(T^{(a_{n+1}-\ell)q_n}(c_\alpha), m),
\]
1. Introduction
2. Basic properties of Sturmian words
3. Rauzy graphs and repetition function
4. Formal intercepts of Sturmian words
5. Factorisation, extensions and prefix-suffix duality
Let $\alpha$ be a slope and $(q_n)_{n \geq -1}$ its sequence of continuants. For naturals $(b_i)$, the **Ostrowski conditions** are the equivalent statements:

i) \forall \ell = 1 \ldots k, \quad \sum_{i=0}^{\ell-1} b_{i+1} q_i < q_\ell

ii) 
- $0 \leq b_1 \leq a_1 - 1$
- $\forall i \geq 1, \quad 0 \leq b_i \leq a_i$
- $\forall i \geq 1, \quad b_{i+1} = a_{i+1} \implies b_i = 0$

Let $n \geq 1$. Any $N \in [0, q_n[$ writes uniquely as

$$N = \sum_{i=0}^{n-1} b_{i+1} q_i$$

where the $(b_i)_{i \geq 1}$ satisfy the Ostrowski conditions.
The set $\mathcal{I}_\alpha$ of formal intercepts of the slope $\alpha$ is defined as

$$\mathcal{I}_\alpha = \left\{ (k_n)_{n \geq 1} \in \prod_{n > 0} [0, q_n] \mid \forall n \geq 1, \ k_n = k_{n+1} \mod q_n \right\}$$

If $\rho = (\rho_n)_{n \geq 1}$ is a formal intercept, there exists a unique sequence $(b_i)_{i \geq 1}$ satisfying the Ostrowski conditions, such that

$$\rho = (\rho_n)_{n \geq 1} = \left( \sum_{i=0}^{n-1} b_{i+1} q_i \right)_{n \geq 1}$$

and in this case we directly write

$$\rho = \sum_{i=0}^{+\infty} b_{i+1} q_i$$
For $m \geq n > 0$, set:

$$
\psi_{n+1} : [0, q_{n+1}] \quad \longrightarrow \
[0, q_n]
$$

$$
k \quad \longmapsto \

k \quad [\mod q_n]
$$

$$
\psi_m = \psi_{n+1} \circ \psi_{n+2} \circ \cdots \circ \psi_{m-1} : [0, q_m] \rightarrow [0, q_n]
$$

then

$$
\mathcal{I}_{\alpha} = \lim [0, q_n] = \left\{ (k_n)_{n>0} \in \prod_{n>0} [0, q_n] \mid n \leq m \Rightarrow \psi_n^m(k_m) = k_n \right\}
$$

is the projective limit of the sets $[0, q_n]$ endowed with the functions $\psi_n^m$. Hence are naturally defined the functions $\psi_n : \mathcal{I}_{\alpha} \mapsto [0, q_n] :$

$$
\psi_n \left( \sum_{i=0}^{+\infty} b_{i+1} q_i \right) = \sum_{i=0}^{n-1} b_{i+1} q_i < q_n
$$
Let $\rho = (\rho_n)_{n \geq 1} = \sum_{i \geq 0} b_{i+1} q_i$ be a formal intercept and

$$\lambda_n = q_{n+1} + q_n - 2 - \rho_{n+1},$$

then:

- The words $T^{\rho_n}(c_\alpha)$ and $T^{\rho_{n+1}}(c_\alpha)$ share the same prefixes of length $\lambda_n$. If $b_{n+1} \neq 0$, this is the maximum such length.
- We have the optimal lower bound $\lambda_n \geq q_n - 1$

**Definition**

Let $\rho = (\rho_n)_{n \geq 1}$ be a formal intercept, then we define the infinite word

$$T^{\rho}(c_\alpha) = \lim T^{\rho_n}(c_\alpha)$$

sharing the prefix of length $q_n - 1$ of $T^{\rho_n}(c_\alpha)$, for all $n \geq 1$.

The length of the longest common prefix of $T^{\rho}(c_\alpha)$ and $T^{\rho_n}(c_\alpha)$ is $\lambda_N$, where $N$ is the smallest $N \geq n$ such that $b_{N+1} \neq 0$. 
Theorem

*Every Sturmian word* \( x \) *of slope* \( \alpha \) *writes uniquely in the form*

\[
x = T^\rho(c_\alpha)
\]

*where* \( \rho \) *is a formal intercept of the slope* \( \alpha \).

The reverse bijection is given as follows. For \( x \) a Sturmian word of slope \( \alpha \), the sequence \( (\gamma_n)_{n \geq 1} \) defined as:

\[
\gamma_n = \min\{k \geq 0 \mid \mathbb{P}_{q_{n-1}}(x) = \mathbb{P}_{q_{n-1}}(T^k(c_\alpha))\}
\]

is a formal intercept.

**Example** : The words \( 0c_\alpha \) and \( 1c_\alpha \) have resp. formal intercepts:

\[
\sum_{i \geq 0} a_{2i+2} q_{2i+1} = (q_2 \lfloor \frac{n}{2} \rfloor - 1)_{n \geq 1} \quad \text{and} \\
(a_1 - 1) + \sum_{i \geq 1} a_{2i+1} q_{2i} = (q_2 \lfloor \frac{n}{2} \rfloor + 1 - 1)_{n \geq 1}
\]
Application to the computation of the repetition function:

For $m \in I_n^\ell$, with $n \geq 0$ and $0 \leq \ell \leq a_{n+1} - 1$. We write $m = (\ell + 1)q_n + q_{n-1} - 2 - r$ where $r$ is the length of the common part of $G_m$. Then we have:

- $r(T^{\rho n+1}(c_\alpha), m) = q_n$ for $q_n \leq m \leq q_{n+1} - 2 - \rho_{n+1}$
- $r(T^{\rho n+1}(c_\alpha), m) = q_{n+1} - \rho_{n+1}$ for $q_{n+1} - 1 - \rho_{n+1} \leq m \leq (\ell + 1)q_n + q_{n-1} - 2$
- $r(T^{\rho n+1}(c_\alpha), m) = \ell q_n + q_{n-1}$ for $(\ell + 1)q_n + q_{n-1} - 1 \leq m \leq q_{n+1} - \rho_{n+1} + q_n - 2$
- $r(T^{\rho n+1}(c_\alpha), m) = q_{n+1} - \rho_{n+1} + q_n$ for $q_{n+1} - \rho_{n+1} + q_n - 1 \leq m \leq q_{n+1} - 2$. 
Let $\rho$ be a formal intercept.
For $k \geq 0$, we define $\rho + k$ as the formal intercept of the Sturmian word $T^k(T^\rho(c_{\alpha}))$.
We say that two formal intercepts $\rho$ and $\gamma$ are equivalent when there exists $k, \ell \geq 0$ such that $\rho + k = \gamma + \ell$.

**Lemma**

Let $\rho$ be a formal intercept and $k \geq 0$ such that $\rho + k$ is not a natural number. Then there exists $N \geq 0$ such that for all $n \geq N$,

$$\Psi_n(\rho + k) = \Psi_n(\rho) + k.$$

Let $\rho = \sum_{i\geq0} b_{i+1}q_i$ be a formal intercept. The following are equivalent:

i) $\rho$ is equivalent to zero,

ii) one of the two sequences $(\Psi_n(\rho))_{n\geq0}$ or $(q_n - \Psi_n(\rho))_{n\geq0}$ converges,

iii) One of the following holds:

- $b_i = 0$ for $i \geq 1$ large enough,
- $b_{2i} = a_{2i}$ and $b_{2i+1} = 0$ for $i \geq 1$ large enough,
- $b_{2i} = 0$ and $b_{2i+1} = a_{2i+1}$ for $i \geq 1$ large enough.

Let $\rho = \sum_{i\geq0} b_{i+1}q_i$ and $\gamma = \sum_{i\geq0} c_{i+1}q_i$ be two formal intercepts not equivalent to zero. Then $\rho$ and $\gamma$ are equivalent if and only if

$$b_i = c_i \quad \text{for } i \geq 1 \text{ large enough}.$$
Let $\rho = \sum_{i \geq 0} b_{i+1} q_i$ be a formal intercept not a natural number. We define the support $\text{Supp}(\rho)$ of $\rho$ as

$$\text{Supp}(\rho) = \{n \geq 0 \mid b_{n+1} \neq 0\}$$

and the function $\Lambda_\rho$ as :

$$\Lambda_\rho(n) = \min(\text{Supp}(\rho) \cap [n, +\infty[).$$

Let $n \geq 0$, and $\rho$ a formal intercept. Then

- $n \in \text{Supp}(\rho)$ if and only if $\Psi_{n+1}(\rho) \geq q_n$, and we have :
  $$\Psi_{\Lambda_\rho(n)+1}(\rho) \geq q_{\Lambda_\rho(n)} \geq q_n$$
- If $n \in \text{Supp}(\rho)$, then $b_{n+2} \neq a_{n+2}$, and we have :
  $$\Psi_{\Lambda_\rho(n)+2}(\rho) < q_{\Lambda_\rho(n)+2} - q_{\Lambda_\rho(n)} = a_{\Lambda_\rho(n)+2} q_{\Lambda_\rho(n)+1}$$
- $\Psi_{\Lambda_\rho(n)+1}(\rho) = b_{\Lambda_\rho(n)+1} q_{\Lambda_\rho(n)} + \Psi_n(\rho)$
Let $m = \sum_{i=0}^{N} b_{i+1} q_i$, where the $(b_i)$'s satisfy the Ostrowski conditions. Then:

$$\mathbb{P}_m(c_\alpha) = \prod_{i=0}^{N} s_{i}^{b_{i+1}} = s_{N}^{b_{N+1}} s_{N-1}^{b_{N}} \cdots s_{0}^{b_{1}}.$$

Let $m, p \geq 1$ be such that $m + p = q_{N+1} - 2$, with:

$$m = \sum_{i=0}^{N} b_{i+1} q_i \quad \text{et} \quad p = \sum_{i=0}^{N} c_{i+1} q_i$$

where the $(b_i)_{i=1}^{N+1}$ et $(c_i)_{i=1}^{N+1}$ satisfy the Ostrowski conditions. Then we have the product formula:

$$s_{N+1}^{-} = \prod_{i=0}^{N} s_{i}^{b_{i+1}} \cdot \prod_{i=0}^{N} \tilde{s}_{i}^{c_{i+1}}.$$
Definition

Let \( \rho = \sum_{i \geq 0} b_{i+1}q_i \) be a formal intercept not a natural number. We define the Sturmian word \( \tilde{P}_\rho(c_\alpha) \) as the infinite product

\[
\tilde{P}_\rho(c_\alpha) = \prod_{i=0}^{+\infty} \tilde{s}_i b_{i+1} = \lim_{N \to +\infty} \prod_{i=0}^{N} \tilde{s}_{i}^{(b_{i+1})} = \lim_{N \to +\infty} \tilde{P}_\Psi N(\rho)(c_\alpha).
\]

Theorem

Let \( \rho = \sum_{i \geq 0} b_{i+1}q_i \) be a formal intercept not a natural number. The formal intercept of \( \tilde{P}_\rho(c_\alpha) \) is given by the sequence

\[
(\psi_n(q_{\Lambda_\rho(n)}+1 - 2 - \rho_{\Lambda_\rho(n)}+1))_{n \geq 0}.
\]
Application: By applying the preceding computation to the two formal intercepts of the words $01c_\alpha$ and $10c_\alpha$, we obtain the following factorisation for the characteristic word $c_\alpha$:

- **If $a_1 \geq 2$, then:**
  \[ c_\alpha = \tilde{s}_0^{a_1-2} \prod_{i \geq 1} \tilde{s}_2^{a_2_i+1} \quad \text{and} \quad c_\alpha = \tilde{s}_0^{a_1-1} \tilde{s}_1^{a_2-1} \prod_{i \geq 1} \tilde{s}_2^{a_2_i+2}. \]

- **If $a_1 = 1$ and $a_2 \geq 2$, then:**
  \[ c_\alpha = \tilde{s}_1^{a_2-2} \tilde{s}_2^{a_3-1} \prod_{i \geq 2} \tilde{s}_2^{a_2_i+1} \quad \text{and} \quad c_\alpha = \tilde{s}_1^{a_2-2} \prod_{i \geq 1} \tilde{s}_2^{a_2_i+2}. \]

- **If $a_1 = 1$ and $a_2 = 1$, then:**
  \[ c_\alpha = \tilde{s}_2^{a_3-1} \prod_{i \geq 2} \tilde{s}_2^{a_2_i+1} \quad \text{and} \quad c_\alpha = \prod_{i \geq 1} \tilde{s}_2^{a_2_i+2}. \]
For $\rho$ a formal intercept not a natural number, we define the formal intercept $\overline{\rho}$ as the formal intercept of $\mathbb{P}_\rho(c_\alpha)$, given by the sequence:

$$\psi_n(\overline{\rho}) = \psi_n(q_{\Lambda_\rho(n)+1} - 2 - \psi_{\Lambda_\rho(n)+1}(\rho)),$$

and call it the **complement** of $\rho$.

**technical lemma**: If $\rho$ is not equivalent to zero, then $\overline{\rho}$ is not equivalent to zero either.

Also we have the formula:

$$T(\overline{\mathbb{P}_{\rho+1}(c_\alpha)}) = \overline{\mathbb{P}_\rho(c_\alpha)}.$$

And as a consequence, for all $k \geq 0$, we have:

$$\rho + k + k = \overline{\rho}.$$
Theorem

The map $\rho \mapsto \overrightarrow{\rho}$, from the set of formal intercept not equivalent to zero to itself is an involution. Hence the reciprocity formula:

$$T^\rho(c_\alpha) = \overrightarrow{P_{\overrightarrow{\rho}}}(c_\alpha)$$

Moreover, the bi-infinite word

$$\overrightarrow{T^\rho(c_\alpha)} \cdot T^\rho(c_\alpha)$$

is a Sturmian orbit.

Note: This is the prefix-suffix duality for Sturmian words... !

As a consequence, the set of non-trivial dynamical orbit of the characteristic Sturmian word of slope $\alpha$ is in a natural correspondence with the set of non-zero equivalence classes of formal intercepts of the slope $\alpha$. 

Application:

- Let $x$ be a Sturmian word of slope $\alpha$. The following are equivalent:
  
  i) $x$ is a suffix of $0c_\alpha$ or $1c_\alpha$,
  
  ii) There exists two distinct formal intercepts $\rho$ and $\gamma$ such that
      \[ x = \widetilde{P}_\rho(c_\alpha) = \widetilde{P}_\gamma(c_\alpha) \]

- Let $x$ be a Sturmian word of slope $\alpha$. The following are equivalent:
  
  i) One of the two words $01c_\alpha$ and $10c_\alpha$ is a suffix of $x$.
  
  ii) The word $x$ has no factorisation of the form:
      \[ x = \widetilde{P}_\rho(c_\alpha) \]
      for a formal intercept $\rho$. 

Thank you for your attention!