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.
- $\widetilde{u}$: mirror image of a word u.

- $\mathbb{P}_n(x)$: prefix of length $n \geq 1$ of an infinite word x,

- $T : \begin{array}{ccc} \mathcal{A}^{\mathbb{N}} & \longrightarrow & \mathcal{A}^{\mathbb{N}} \\ x_0x_1x_2\ldots & \longmapsto & x_1x_2x_3\ldots \end{array}$ the **shift** on infinite words

- $T^n(x)$: the n-th suffix of an infinite word x.

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 $\mathcal{A}_X = \{x_1, x_2, x_3, \ldots\}$ as

- $Z = \prod_{n \geq 1} x_{val_2(n)+1}$ where 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 φ is the morphism defined by $\varphi(x_i) = x_1 x_{i+1}$ for $i \geq 1$.

$$Z = x_1x_2x_1x_3x_1x_2x_1x_4x_1x_2x_1x_3x_1 \ldots$$

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

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

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 :

$$\widetilde{\mathbb{P}}_m(Z) = \prod_{i=0}^{N}{}^{\uparrow} U_{i+1}^{b_i} \quad \text{and} \quad T^m(Z) = \prod_{i\geq 0}{}^{\uparrow} U_{i+1}^{1-b_i}.$$

Let $\gamma = \sum_{i \geq 0} b_i 2^i$ be a 2-adic number, and define

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

If γ is not an integer, then we have the reciprocity formula :

$$T^\gamma(Z) = \widetilde{\mathbb{P}}_{\overline{\gamma}}(Z)$$

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

$\overline{\gamma}$ is called the **complement** of γ.

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
$$Y = \widetilde{\mathbb{P}}_\gamma(Z)$$
where γ is a 2-adic number that is not a natural number.

- Moreover, for every γ a 2-adic number that is not an integer, the bi-infinite word
$$\widetilde{T^{\overline{\gamma}}(Z)} \cdot T^\gamma(Z)$$
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

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 satifsfies the balanced property, and the number

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

is an irrational number.

The number α 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 α, we get from the balanced property applied to $0L_n$ and $1L_n$:

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

so that $\{|L_n|_1\} = \mathbb{Z} \cap]n\alpha - 1, n\alpha[$, determining uniquely the sequence of words $(L_n)_{n \geq 1}$.

Definition

We define the **characteristic word** c_α *of the slope* α *as :*

$$c_\alpha = \lim L_n$$

Amongst Sturmian word of slope α, 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 irrationnal $\alpha \in]0,1[$ writes uniquely as

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

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 :

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

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 $s \to t$ 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 α be a slope and $m \geq 2$, then :
$$\mathbb{P}_m(T^{r(c_\alpha, m)}(c_\alpha)) = \mathbb{P}_m(c_\alpha) = L_m$$
- 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 α as the denominators of the irreducible fraction

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

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

$$\boxed{q_{n+1} = a_{n+1} q_n + q_{n-1}}$$

Theorem

$$r(c_\alpha, m) = q_n \quad \textit{for} \quad q_n - 1 \leq m \leq q_{n+1} - 2.$$

Every infinite word x defines a path in the Rauzy graph G_m :

$$\mathbb{P}_m(x) \to \mathbb{P}_m(T(x)) \to \mathbb{P}_m(T^2(x)) \to \ldots \to \mathbb{P}_m(T^k(x)) \to \ldots.$$

The value $r(x, m)$ is interpreted in the Rauzy graph G_m as the longuest Hamiltonian path in this path, that is the longuest 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_α : For $m \in I_n^{\ell}$ with $0 \leq \ell \leq a_{n+1} - 1$:

- The characteristic word c_α turns exactly $a_{n+1} - \ell$ times around the referent cycle, that is :

$$\begin{aligned} q_n &= r(c_\alpha, m) = r(T^{q_n}(c_\alpha), m) \\ &= \ldots = r(T^{(a_{n+1}-l-1)q_n}(c_\alpha), m) \neq r(T^{(a_{n+1}-l)q_n}(c_\alpha), m), \end{aligned}$$

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

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

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.

Definition

The set $\mathcal{I}_\alpha$ of **formal intercepts** *of the slope α is defined as*

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

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 :

$$\Psi_n^{n+1} : \begin{array}{ccc} [0, q_{n+1}[& \longmapsto & [0, q_n[\\ k & \longmapsto & k \ [\texttt{mod} \ q_n] \end{array}$$

$$\Psi_n^m = \Psi_n^{n+1} \circ \Psi_{n+1}^{n+2} \circ \cdots \circ \Psi_{m-1}^m \ : \ [0, q_m[\to [0, q_n[$$

then

$$\mathcal{I}_\alpha = \varprojlim [0, q_n[= \left\{ (k_n)_{n>0} \in \prod_{n>0} [0, q_n[\;\middle|\; n \leq m \Rightarrow \Psi_n^m(k_m) = k_n \right\}$$

is the projective limit of the sets $[0, q_n[$ endowed with the functions Ψ_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

$$\lambda_n = q_{n+1} + q_n - 2 - \rho_{n+1}, \qquad \text{then :}$$

- The words $T^{\rho_n}(c_\alpha)$ and $T^{\rho_{n+1}}(c_\alpha)$ share the same prefixes of length λ_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 longuest common prefix of $T^{\rho}(c_\alpha)$ and $T^{\rho_n}(c_\alpha)$ is λ_N, where N is the smallest $N \geq n$ such that $b_{N+1} \neq 0$.

Theorem

Every Sturmian word x of slope α writes uniquely in the form

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

where ρ is a formal intercept of the slope α.

The reverse bijection is given as follows. For x a Sturmian word of slope α, 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 ρ 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 ρ and γ are **equivalent** when there exists $k, \ell \geq 0$ such that $\rho + k = \gamma + \ell$.

Lemma

Let ρ 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 $\rho = \sum_{i \geq 0} b_{i+1} q_i$ be a formal intercept. The following are equivalent :

i) ρ is equivalent to zero,

ii) one of the two sequences $(\Psi_n(\rho))_{n \geq 0}$ or $(q_n - \Psi_n(\rho))_{n \geq 0}$ 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 $\rho = \sum_{i \geq 0} b_{i+1} q_i$ and $\gamma = \sum_{i \geq 0} c_{i+1} q_i$ be two formal intercept not equivalent to zero. Then ρ and γ 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 $\mathrm{Supp}(\rho)$ of ρ as

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

and the function Λ_ρ as :

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

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

- $n \in \mathrm{Supp}(\rho)$ 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 $n \in \mathrm{Supp}(\rho)$, 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 :

$$\mathbb{P}_m(c_\alpha) = \prod_{i=0}^{N}{}^{\downarrow} s_i^{b_{i+1}} = s_N^{b_{N+1}} s_{N-1}^{b_N} \dots s_0^{b_1}.$$

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 :

$$s_{N+1}^{--} = \prod_{i=0}^{N}{}^{\downarrow} s_i^{b_{i+1}} \cdot \prod_{i=0}^{N}{}^{\uparrow} \widetilde{s_i}^{c_{i+1}}.$$

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* $\widetilde{\mathbb{P}_\rho}(c_\alpha)$ *as the infinite product*

$$\widetilde{\mathbb{P}_\rho}(c_\alpha) = \prod_{i=0}^{+\infty} \widetilde{s_i}^{b_{i+1}} = \lim_{N \to +\infty} \prod_{i=0}^{N\uparrow} \widetilde{s_i}^{b_{i+1}} = \lim_{N \to +\infty} \widetilde{\mathbb{P}_{\Psi_N(\rho)}}(c_\alpha).$$

Theorem

Let $\rho = \sum_{i \geq 0} b_{i+1} q_i$ *be a formal intercept not a natural number. The formal intercept of* $\widetilde{\mathbb{P}_\rho}(c_\alpha)$ *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 preceeding computation to the two formal intercepts of the words $01c_\alpha$ and $10c_\alpha$, we obtain the following factorisation for the characteristic word c_α :

- If $a_1 \geq 2$, then :

$$c_\alpha = \widetilde{s_0}^{a_1-2} \prod_{i\geq 1} \widetilde{s_{2i}}^{a_{2i+1}} \quad \text{and} \quad c_\alpha = \widetilde{s_0}^{a_1-1} \widetilde{s_1}^{a_2-1} \prod_{i\geq 1} \widetilde{s_{2i+1}}^{a_{2i+2}}.$$

- If $a_1 = 1$ and $a_2 \geq 2$, then :

$$c_\alpha = \widetilde{s_1}^{a_2-2} \widetilde{s_2}^{a_3-1} \prod_{i\geq 2} \widetilde{s_{2i}}^{a_{2i+1}} \quad \text{and} \quad c_\alpha = \widetilde{s_1}^{a_2-2} \prod_{i\geq 1} \widetilde{s_{2i+1}}^{a_{2i+2}}$$

- If $a_1 = 1$ and $a_2 = 1$, then :

$$c_\alpha = \widetilde{s_2}^{a_3-1} \prod_{i\geq 2} \widetilde{s_{2i}}^{a_{2i+1}} \quad \text{and} \quad c_\alpha = \prod_{i\geq 1} \widetilde{s_{2i+1}}^{a_{2i+2}}.$$

For ρ 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 ρ.

<u>technical lemma :</u> If ρ is not equivalent to zero, then $\overline{\rho}$ is not equivalent to zero either.

Also we have the formula :

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

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

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

Theorem

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

$$T^{\rho}(c_\alpha) = \widetilde{\mathbb{P}_{\overline{\rho}}}(c_\alpha)$$

Moreover, the bi-infinite word

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

is a Sturmian orbit.

Note : **This is the prefix-suffix duality for Sturmian words... !**
As a consequence, the set of non-trivial dynamical orbit orbit of the characteristic Sturmian word of slope α is in a natural correspondence with the set of non-zero equivalence classes of formal intercepts of the slope α.

Application :

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

- Let x be a Sturmian word of slope α. 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 :
  $$x = \widetilde{\mathbb{P}_\rho}(c_\alpha)$$
  for a formal intercept ρ.

− − − ====== THE END====== − − −

Thank you for your attention !