# Discret interval exchange transformations and perfectly clustering words

Mélodie Lapointe

Institut de Recherche en informatique fondamentale

November 20, 2020

# Outline

We study a family of words called

perfectly clustering words

to highlight links between words combinatorics, numbers theory and dynamical system.
We want to

- to understand the structure of perfectly clustering words,
- to generalize the relation between the Raney tree and the Christoffel words,
- to understand their link with the free group.

# Perfectly clustering words and Christoffel words

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

a n a n a s

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

a n a n a s
  n a n a s a

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

```
a n a n a s
  n a n a s a
    a n a s a n
```

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

$$
\begin{array}{l}
\text{a n a n a s} \\
\quad \text{n a n a s a} \\
\qquad \text{a n a s a n} \\
\qquad\quad \text{n a s a n a} \\
\qquad\qquad \text{a s a n a n} \\
\qquad\qquad\quad \text{s a n a n a}
\end{array}
$$

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

a n a n a s
n a n a s a
a n a s a n
n a s a n a
a s a n a n
s a n a n a

a n a n a s

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

```
n a n a s a                 a n a n a s
  a n a s a n               a n a s a n
    n a s a n a
      a s a n a n
        s a n a n a
```

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

$$\begin{array}{cccccc} a & n & a & n & a & s \\ a & n & a & s & a & n \\ a & s & a & n & a & n \\ n & a & n & a & s & a \\ n & a & s & a & n & a \\ s & a & n & a & n & a \end{array}$$

# Perfectly clustering word

The Burrows-Wheeler transform maps a word $w$ to the last column of the matrix containing all conjugates of $w$ in lexicographical order.

Example :

$$\begin{array}{cccccc} a & n & a & n & a & s \\ a & n & a & s & a & n \\ a & s & a & n & a & n \\ n & a & n & a & s & a \\ n & a & s & a & n & a \\ s & a & n & a & n & a \end{array}$$

$$\mathrm{BWT}(\mathrm{ananas}) = \mathrm{snnaaa}$$

# Perfectly clustering word

- $|w|_a$ denote the number of occurrences of the letter $a$ in $w$.
- A word $w$ is $\pi$-clustering if

$$\mathrm{BWT}(w) = a_{\pi(1)}^{|w|_{a_{\pi(1)}}} a_{\pi(2)}^{|w|_{a_{\pi(2)}}} \dots a_{\pi(r)}^{|w|_{a_{\pi(r)}}}$$

and $\pi \neq id$.

- A word $w$ is perfectly clustering if

$$\mathrm{BWT}(w) = a_r^{|w|_{a_r}} a_{r-1}^{|w|_{a_{r-1}}} \dots a_1^{|w|_{a_1}}$$

Example :

| Words | appartement | aluminium | ananas |
|---|---|---|---|
| BWT | tptmeepaanr | mmnauuiil | snnaaa |

# Christoffel words

Geometric definition :
Let $p$ and $q$ be two relatively prime natural number.
The Christoffel path of slope $q/p$ is the path from $(0,0)$ to $(q,p)$ such that

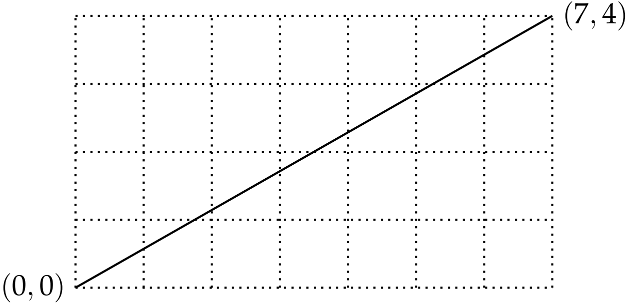
- the path lies below the segment
- the region enclose by the path and the segment contains no integral points besides those on the path.

The Christoffel words of slope $q/p$ is the word on $\{a,b\}^*$ which represent the Christoffel path of slope $q/p$ where

$$a \mapsto [(x,y),(x+1,y)]$$
$$b \mapsto [(x,y),(x,y+1)].$$

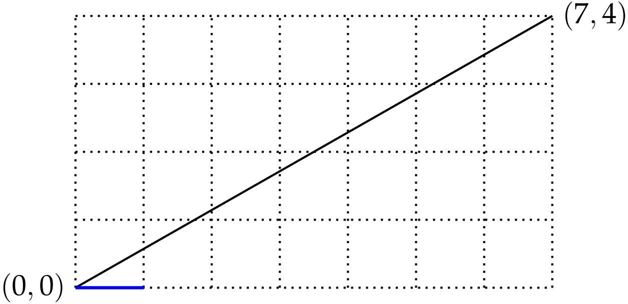
# Christoffel words

Examples : The Christoffel words of slope $7/4$



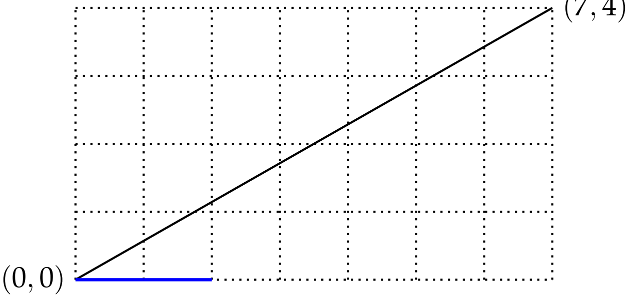
# Christoffel words

Examples : The Christoffel words of slope 7/4



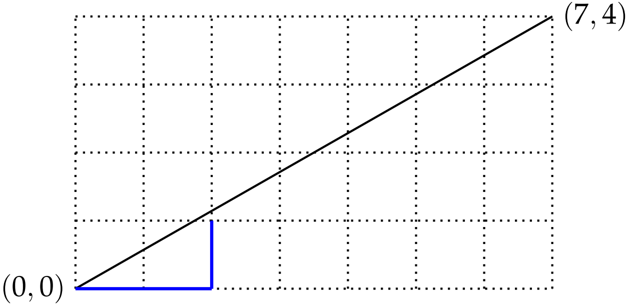
# Christoffel words

Examples : The Christoffel words of slope $7/4$



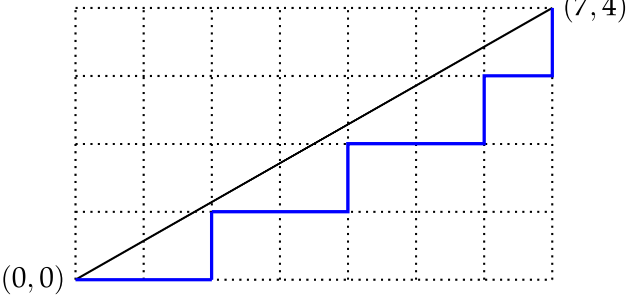
# Christoffel words

Examples : The Christoffel words of slope $7/4$



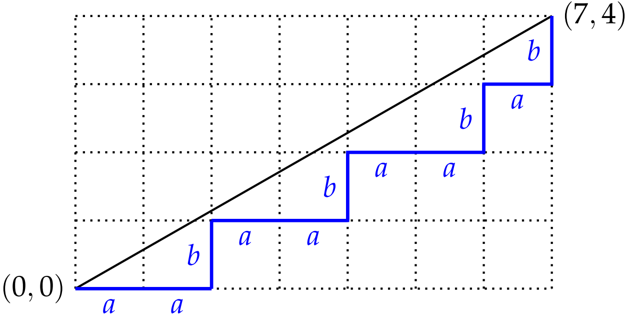
# Christoffel words

Examples : The Christoffel words of slope 7/4



# Christoffel words

Examples : The Christoffel words of slope $7/4$



# Christoffel words

- A power of a word $w$ is

$$w^1 = w \text{ and } w^n = w^{n-1}w \text{ for } n > 1.$$

- A word is primitive if it is not the power of a smaller word.
- A word is a Lyndon word if it is primitive and the smallest of its conjugates.

**Theorem (Mantaci, Restivo et Sciortino, 2003)**

*A binary word $w$ is perfectly clustering if and only if $w$ is a conjugate or a power of a conjugate of a Christoffel word.*

Christoffel words are Lyndon words. Therefore, we will study perfectly clustering words that are Lyndon words.

# Discrete interval exchange transformation

# Discrete interval exchange transformation

- $c = (c_1, \ldots, c_r)$ a length vector
- $\pi$ a permutation
- $|c| = c_1 + \cdots + c_r$
- $\mathcal{A}$ a finite totally ordered alphabet of cardinality $r$

# Discrete interval exchange transformation

A discrete r-interval exchange transformation $T$ with length vector $c$ and permutation $\pi$ is defined on a set of $|c|$ points, $\{1, \ldots, |c|\}$, partionned into $r$ intervals

$$I_{a_i} = \left\{ k : \sum_{j<i} c_j < k \leq \sum_{j \leq i} c_j \right\}$$

by

$$T(x) = x + t_i$$

when $x \in I_{a_i}$ and

$$t_i = \sum_{\pi^{-1}(j) < \pi^{-1}(j)} c_j - \sum_{j<i} c_j.$$

# Discrete interval exchange transformation

- The orbit of a point $z \in \{1, \dots, |c|\}$ is the set
$$\mathcal{O}(z) = \{T^n(z) | n \in \mathbb{Z}\}.$$
- $\mathcal{O}(z)$ is always finite if $T$ is a discrete interval exchange transformation.
- The mapping $T$ is minimal if the orbit of any point in $\{1, \dots, |c|\}$ is dense in $\{1, \dots, |c|\}$.
- Therefore $T$ is minimal if and only if the orbits of all points in $\{1, \dots, |c|\}$ are equals.

# Discrete interval exchange transformation

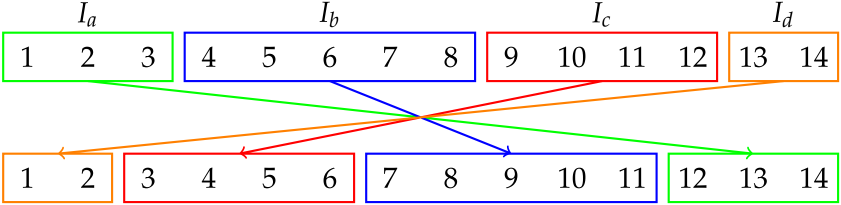
- The trajectory of a point $z$ under $T$ is the infinite sequence $(z_n)_{n \in \mathbb{N}}$ defined by

$$z_n = a_i \text{ if } T^n(z) \in I_{a_i}.$$

- Since $T$ is discrete, the trajectory of a point is a periodic sequence.

# Discrete interval exchange transformation

Example : $T$ with lenght vector $(3,5,4,2)$ and permutation 4321




Orbit : $\mathcal{O}(4) = \{4, 7, 10\}$

Trajectory : $(bbc)^{\omega}$

# Links between perfectly clustering words and DIET

**Theorem (Ferenczi and Zamboni, 2013)**

*Let $w = w_1 \dots w_n$ be a primitive word on an alphabet $\mathcal{A} = \{a_1, \dots, a_r\}$ such that for all $a \in \mathcal{A}$, $|w|_a \geq 1$.*
*The following statement are equivalent :*

- *$w$ is $\pi$-clustering,*
- *$ww$ is a factor in the trajectory of a minimal discret $r$-interval echange transformation with permutation $\pi$,*
- *$ww$ is a factor in the trajectory of a discret $r$-interval echange transformation with permutation $\pi$,*
- *$ww$ is a factor in the trajectory of a continuous $r$-interval echange transformation with permutation $\pi$,*

Remark : From now on, the DIET are always with the symmetric permutation $\pi = r(r-1)\dots 1$.

Some infinite complete binary trees

# Raney tree

Tree construction rule's :

- root : $(1,1)$
- all other nodes :

$$
\begin{array}{ccc}
 & (i,j) & \\
\swarrow & & \searrow \\
(i+j,j) & & (i,i+j)
\end{array}
$$

$$(1,1)$$

$$(2,1) \quad (1,2)$$

$$(3,1) \quad (2,3) \quad (3,2) \quad (1,3)$$

$$(4,1)(3,4)(5,3)(2,5)(5,2)(3,5)(4,3)(1,4)$$

**Theorem (Raney, 1973 and Calkin et Wilf, 2000)**

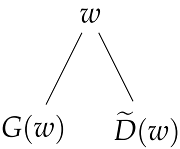
*All pairs of relatively prime integers appears exactly once in the Raney tree.*

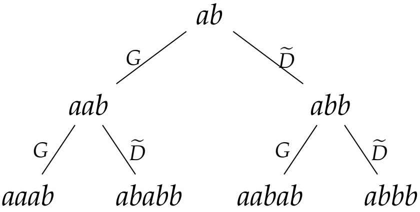
Therefore, all length vectors of minimal discrete 2-interval exchange transformation appears in the Raney tree.

# Tree of dual Christoffel words

Tree construction rule's :

- root : $ab$
- all other nodes :




**Theorem**

*All Christoffel words appears exactly once in the tree of dual Christoffel words.*

Remarks : $G : a \mapsto a$ and $b \mapsto ab$
$\widetilde{D} : a \mapsto ab$ and $b \mapsto a$

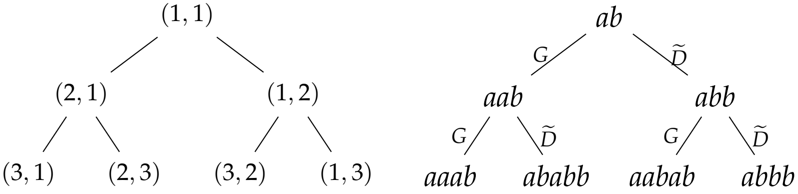
# Tree of dual Christoffel words

**Theorem**

*The Raney tree and the dual of the Christoffel tree are isomorphic.*

Proof outline :

- $w \mapsto (|w|_a, |w|_b)$
- Construct the dual of Christoffel words of slope $j/i$



# Perfectly clustering word

- Can we generalize both trees ?
- What are the relationship between the generalization of those trees ?

# Generalization of Raney tree

# Circular composition

Goal : Labelling an infinite tree with lenght associated to minimal DIET.
What is known :

- $c = (c_1, c_2)$, $T$ is minimal iff $gcd(c) = 1$ (Raney tree, Stern-Brocot tree) ;
- $c = (c_1, c_2, c_3)$, $T$ is minimal iff $gcd(c_1 + c_2, c_2 + c_3) = 1$ (Redlich and Pak, 2009) ;
- $c = (c_1, \ldots, c_r)$ and $r > 3$, there do not exist homogeneous polynomials $f_1$ and $f_2$ of arbitrary degree with interger coefficients such that

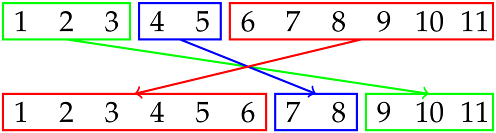
  $$gcd(f_1(c), f_2(c)) = 1$$

  iff $T$ is minimal (Karnauhova and Liebscher, 2015).

# Circular composition

- A discrete $r$-interval exchange transformation $T$ can be viewed as a permutation.
- $T$ is minimal iff its permutations is circular i.e. has 1 cycle.
- A composition $C$ is circular if the symmetric DIET $T$ is minimal

Example : $T$ with lenght vector $(3, 2, 6)$




Permutation : $\mathcal{P}(T) = (1, 9, 4, 7, 2, 10, 5, 8, 3, 11, 6)$

# Circular composition

$$\psi_i(x) = \begin{cases} x & \text{if } x \in \{1, \dots, m\}, \\ x \cdot (x + |t_i|) & \text{if } x \in \{m+1, \dots, m + |t_i|\} \text{ and } t_i \geq 0, \\ (x + |t_i|) \cdot x & \text{if } x \in \{m+1, \dots, m + |t_i|\} \text{ and } t_i < 0, \\ x + |t_i| & \text{otherwise} \end{cases}$$

where $x \in \{1, \dots, |c|\}$ is a letter, $\cdot$ is word concatenation and

$$m = \begin{cases} \sum_{j \leq i} c_j & \text{if } t_i \geq 0, \\ \sum_{j > i} c_j & \text{if } t_i < 0. \end{cases}$$

Recall : $t_i = \sum_{j>i} c_j - \sum_{j<i} c_j$

# Circular composition

$$\psi_i(x) = \begin{cases} x & \text{if } x \in \{1, \dots, m\}, \\ x \cdot (x + |t_i|) & \text{if } x \in \{m+1, \dots, m+|t_i|\} \text{ and } t_i \geq 0, \\ (x + |t_i|) \cdot x & \text{if } x \in \{m+1, \dots, m+|t_i|\} \text{ and } t_i < 0, \\ x + |t_i| & \text{otherwise} \end{cases}$$

**Proposition**

*Let $i$ be in $\{1, \dots, r\}$ and $c$ a lenght vector. Then,*

$$\psi_i(\mathcal{P}(T_c)) \equiv \mathcal{P}(T_{c'})$$

*where $c' = (c_1, \dots, c_i + |t_i|, \dots, c_r)$.*

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | | | | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | | | | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6,7,8\} \\ x + 3 & \text{otherwise} \end{cases}$

$$m = \begin{cases} \sum_{j \le i} c_j & \text{if } t_i \ge 0, \\ \sum_{j > i} c_j & \text{if } t_i < 0. \end{cases}$$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | | | | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \dots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6,7,8\} \\ x + 3 & \text{otherwise} \end{cases}$

$$\psi_i(x) = \begin{cases} x & \text{if } x \in \{1, \dots, m\}, \\ x \cdot (x + |t_i|) & \text{if } x \in \{m+1, \dots, m+|t_i|\} \text{ and } t_i \geq 0, \\ (x + |t_i|) \cdot x & \text{if } x \in \{m+1, \dots, m+|t_i|\} \text{ and } t_i < 0, \\ x + |t_i| & \text{otherwise} \end{cases}$$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | | | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \dots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \dots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | 5, | | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | 5, | 8, 11, | | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | 5, | 8, 11, | 3, | | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \dots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | 5, | 8, 11, | 3, | 14, | |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \dots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Circular composition

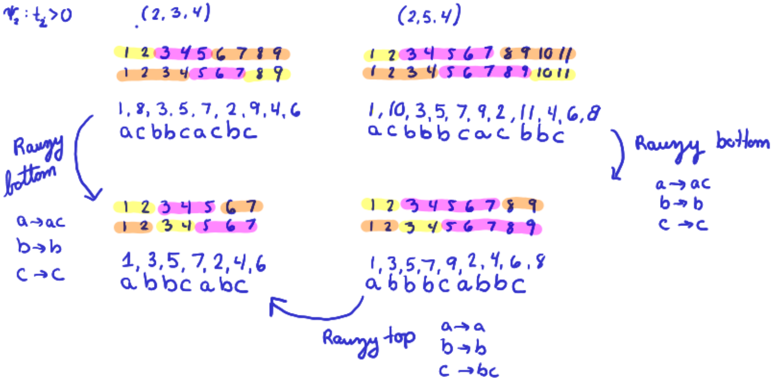
Example : $\psi_2(\mathcal{P}(T_{(3,2,6)}))$

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\mathcal{P}(T_{(3,2,6)})$ | 1, | 9, | 4, | 7, | 2, | 10, | 5, | 8, | 3, | 11, | 6 |
| $\psi_2(\mathcal{P}(T_{(3,2,6)}))$ | 1, | 12, | 4, | 7, 10, | 2, | 13, | 5, | 8, 11, | 3, | 14, | 6, 9 |

- $t_2 = 6 - 2 = 3 > 0$
- $m = 3 + 2 = 5$
- $\psi_2(x) = \begin{cases} x & \text{if } x \in \{1, \ldots, 5\} \\ x \cdot (x+3) & \text{if } x \in \{6, 7, 8\} \\ x + 3 & \text{otherwise} \end{cases}$

# Rauzy Induction

How can you induce on the middle interval ?



# Rauzy Induction

How can you induce on the middle interval ?

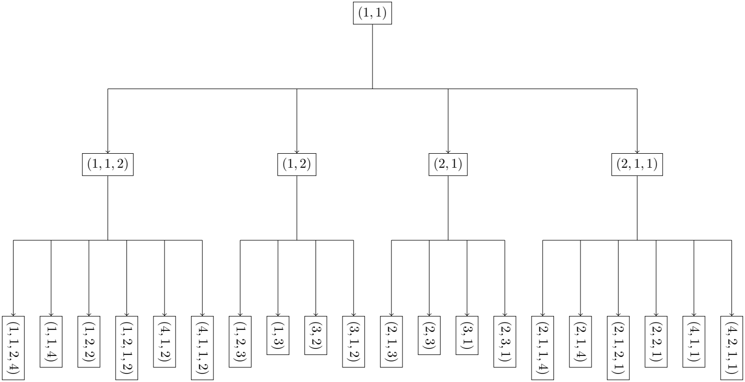
| $\psi_i$ | Rauzy | Morphismes | $t_i > 0$ |
|---|---|---|---|
| $\psi_1$ | Rauzy à gauche bottom 2x | $\lambda_a = (a, ab, ac)$ | Vrai |
| $\psi_2$ | Rauzy à gauche (bottom + top + inverse bottom) | $\lambda_b = (ab^{-1}, b, bc)$ | Vrai |
| $\psi_2$ | Rauzy à droite (bottom + top + inverse bottom) | $\rho_b = (ab, b, b^{-1}c)$ | Faux |
| $\psi_3$ | Rauzy à droite bottom 2x | $\rho_c = (ac, bc, c)$ | Faux |

# Tree of circular composition

- Roots : (1,1)
- Each node $c = (c_1, \ldots, c_r)$ has children :
  - $(c_1, \ldots, c_i + |t_i|, \ldots, c_r)$ for all $i \in \{1, \ldots, r\}$
  - Let $\delta_i = |\sum_{j>i} c_j - \sum_{j \leq j} c_j|$, if $\delta_i > c_{i+1}$ or $-\delta_i > c_i$, then $(c_1, \ldots, c_i, \delta_i, c_{i+1}, \ldots, c_r)$ for all $i \in \{0, \ldots, r\}$

Remark : The criterion on $\delta_i$ check that $\delta_i \neq 0$ and is necessary for the unicity.

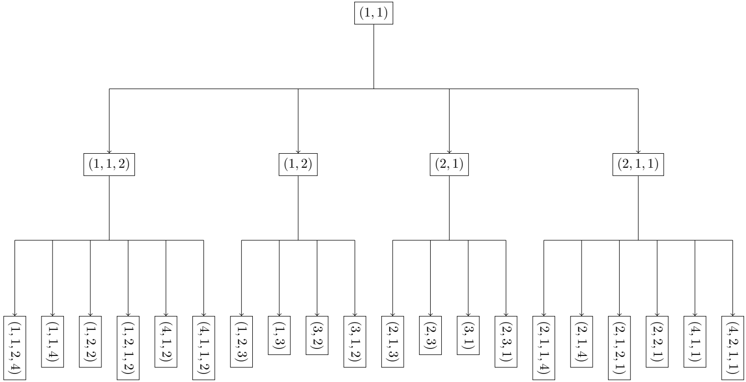
# Tree of circular composition



# Tree of circular composition

**Theorem**

*All circular compositions appear exactly once in the tree of circular composition.*



## Generalization of the tree of dual Christoffel words

# Free group morphisms

The first idea is to work on the free group $F(\mathcal{A})$ instead of the free monoid $\mathcal{A}$. Recall that

- The inverse of an element $l \in \mathcal{F}(\mathcal{A})$ is denoted by $l^{-1}$.
- Each element of the free group may be represent by a reduced word, which is a product of the letters or their inverses, without the factors $xx^{-1}$ or $x^{-1}x$ for $x \in \mathcal{A}$.
- A element $w$ of the free group is called positive if $w \in \mathcal{A}^*$.

# Free group morphisms

**Morphism**

$$\rho_\ell : F(\mathcal{A}) \to F(\mathcal{A})$$
$$\rho_\ell(x) = \begin{cases} x & \text{si } x = \ell \\ x\ell & \text{si } x < \ell \\ \ell^{-1}x & \text{si } x > \ell \end{cases}$$

**Example**

$$\rho_b : F(\{a,b,c\}) \to F(\{a,b,c\})$$
$$\rho_b(x) = \begin{cases} ab & \text{si } x = a \\ b & \text{si } x = b \\ b^{-1}c & \text{si } x = c \end{cases}$$

Remarks : $\rho_b = \widetilde{D}$

# Free group morphisms

**Morphism**

$$\lambda_\ell : F(\mathcal{A}) \to F(\mathcal{A})$$
$$\lambda_\ell(x) = \begin{cases} x & \text{si } x = \ell \\ x\ell^{-1} & \text{si } x < \ell \\ \ell x & \text{si } x > \ell \end{cases}$$

**Example**

$$\lambda_b : F(\{a, b, c\}) \to F(\{a, b, c\})$$
$$\lambda_b(x) = \begin{cases} ab^{-1} & \text{si } x = a \\ b & \text{si } x = b \\ bc & \text{si } x = c \end{cases}$$

Remarks : $\lambda_a = G$ $\qquad \rho_\ell^{-1} = \lambda_\ell$

# Free group morphisms

Morphism

$$\Delta_\ell : \mathcal{A}^* \to (\mathcal{A} \cup \{b\})^*$$

$$\Delta_\ell(x) = \begin{cases} x & \text{si } x < \ell \\ N(x) & \text{si } x \geq \ell \end{cases}$$

Example

$$\Delta_b : \{a, b, c\}^* \to \{a, b, c, d\}^*$$

$$\Delta_b(x) = \begin{cases} a & \text{si } x = a \\ c & \text{si } x = b \\ d & \text{si } x = c \end{cases}$$

Remarks : $N(x)$ maps a letter $x$ to the next letter in the lexicographical order.

# Free group morphisms

Let $w$ be a positive word of $F(\mathcal{A})$.
The morphism $\lambda_\ell(w)$ and $\rho_\ell(w)$ are necessarily positive but their exists a condition on the parikh vector of perfectly clustering word.

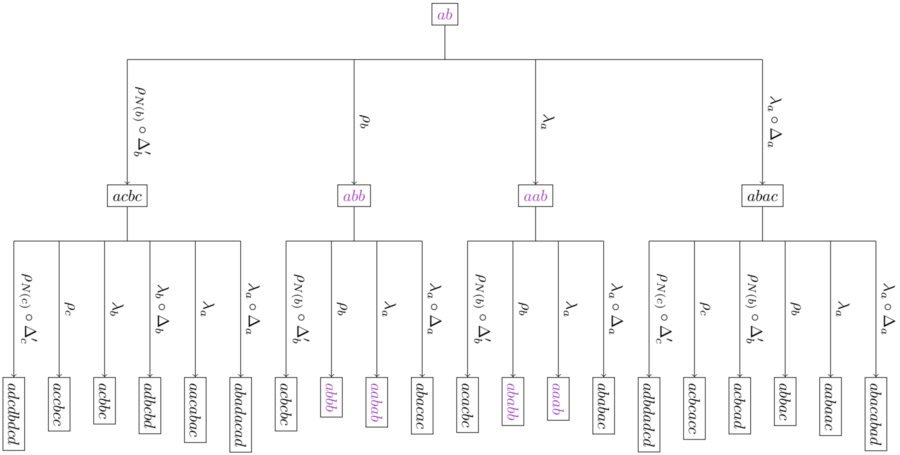
**Proposition**

*Let $w$ be a perfectly clustering Lyndon word.*

- *If $\sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a$, then $\lambda_\ell(w)$ is positive.*
- *If $\sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a$, then $\rho_\ell(w)$ is positive.*

Remark : In the construction of the tree, $\sum_{a>\ell} |w|_a$ and $\sum_{a<\ell} |w|_a$ are never equal.

# Tree of perfectly clustering Lyndon words



# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

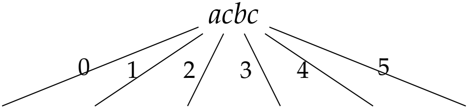
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :



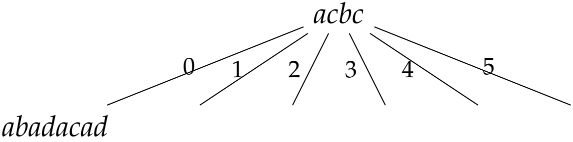
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :



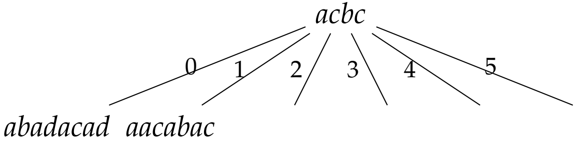
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :



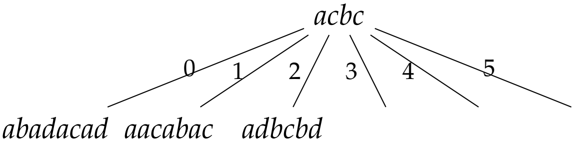
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :

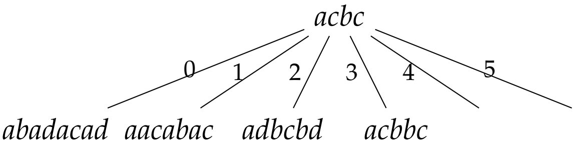


# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.
Example :



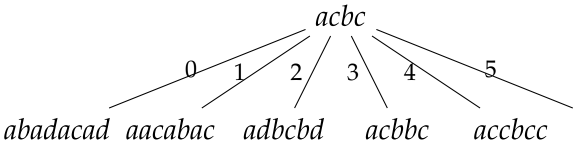
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :



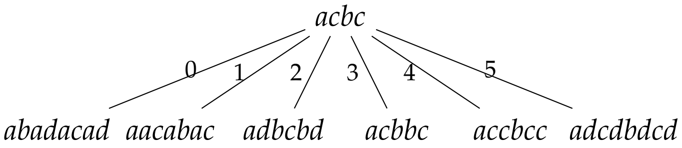
# Tree of perfectly clustering Lyndon words

Let $w$ be a perfectly clustering word appearing in the tree. Then,

$$S_i(w) = \begin{cases} \lambda_\ell \circ \Delta_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \lambda_\ell(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a > \sum_{a<\ell} |w|_a; \\ \rho_\ell(w), & \text{if } i \text{ is even and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a; \\ \rho_{N(\ell)} \circ \Delta_{N(\ell)}(w), & \text{if } i \text{ is odd and } \sum_{a>\ell} |w|_a < \sum_{a<\ell} |w|_a, \end{cases}$$

where $\ell = a_{\lfloor i/2 \rfloor + 1}$ and $i \in \{0, \ldots, 2|Alph(w)| - 1\}$.

Example :



# Tree of perfectly clustering word

**Theorem**

*Every complete perfectly clustering Lyndon word appears exactly once in the tree of perfectly clustering word.*

# Tree of perfectly clustering word

**Theorem**

*Every complete perfectly clustering Lyndon word appears exactly once in the tree of perfectly clustering word.*

**Theorem**

*The tree of circular composition and the tree of complete perfectly clustering Lyndon words are isomorphic.*

# Tree of perfectly clustering word

**Theorem**

*Every complete perfectly clustering Lyndon word appears exactly once in the tree of perfectly clustering word.*

**Theorem**

*The tree of circular composition and the tree of complete perfectly clustering Lyndon words are isomorphic.*

**Corollary**

*If $w$ is a complete perfectly clustering word, then their exists a sequence of group morphism*

$$\xi_1 \circ \xi_2 \circ \cdots \circ \xi_k(ab) = w$$

*such that $\xi_i \in \{\lambda_\ell, \rho_\ell, \lambda_\ell \circ \Delta_\ell, \rho_{N(\ell)\circ\Delta_{N(\ell)}} \mid \ell \in \mathcal{A}\}$.*

Primitive element of $F(\mathcal{A})$

# Primitive elements

- A basis of the free group $F(\mathcal{A})$ is a $|\mathcal{A}|$-uplet $u = (u_1, \dots u_{|\mathcal{A}|})$ such that the function $a_i \mapsto u_i$ extends to an automorphims of $F(\mathcal{A})$.
- An element $g$ is primitive if there exists elements $h_1, \dots, h_{|\mathcal{A}|-1}$ such that $(g, h_1, \dots, h_{|\mathcal{A}|-1})$ is a basis of $F(\mathcal{A})$.
- A primitive element of $F(\mathcal{A})$ is primitive on $\mathcal{A}^*$ but a primitive word on $\mathcal{A}^*$ is not necessarily a primitive element of $F(\mathcal{A})$.

# Primitive elements

**Theorem**

*If $w$ is a perfectly clustering words on $\mathcal{A}^*$, then $w$ is a positive primitive element of $F(\mathcal{A})$.*

Idea of the proof :

- $\lambda_\ell$ and $\rho_\ell$ are automorphisms.
- If $\mathcal{A} \subseteq \mathcal{B}$, then a primitive element of $F(\mathcal{A})$ is a primitive element of $F(\mathcal{B})$.
- If $w$ is a primitive element of $F(\mathcal{A})$, then $\lambda_\ell \circ \Delta_\ell(w)$ and $\rho_{N(\ell)} \circ \Delta_{N(\ell)}(w)$ are primitive element of $F(\mathcal{A} \cup b)$.

# Primitive elements

**Theorem**

*If $w$ is a perfectly clustering words on $\mathcal{A}^*$, then $w$ is a positive primitive element of $F(\mathcal{A})$.*

**Theorem (Osbone, Zieschang (1981) and Kassel, Reutenauer (2007))**

*The word $\{a, b\}^*$ that are conjugates of Christoffel words are exactly the positive primitive elements of the free group $F(\{a, b\})$.*

For the free group $F(\mathcal{A})$ the converse is not true :
$(cabaaba, caba, cababa)$ is a basis of $F(\{a, b, c\})$, but the word $cabaaba$ is not perfectly clustering since $\mathrm{BWT}(cabaaba) = bcabaaa$.

# Summary

- Perfectly clustering word are a generalization of Christoffel words.
- Perfectly clustering words are codage of discrete interval exchange transformation.
- We construct a a generalization of the Raney tree.
- We construct a generalization of the tree of dual Christoffel words.
- They can be enumerate using free group automorphism.
- Perfeclty clustering words are positive primitive elements of the free group.

# Summary

The discrete $r$-interval exchange transformation with $r \in \{2, 3\}$ are special, since their exists *gcd* criterion to compute the minimality.

Do you know examples of substitutions on the 4-letters alphabets corresponding to 4-interval exchange transformation?