An automaton to study imbalances in S-adic words

Application : construction of a C-adic word with infinite imbalance

Mélodie Andrieu
Advisor : Julien Cassaigne
Institut de Mathématiques de Marseille

17^{e} Journées Montoises, Bordeaux, 2018

I - INITIAL QUESTION : imbalance of words associated with Cassaigne-Selmer continued fraction algorithm

II - GENERAL TOOL : an automaton to study imbalance in S-adic words

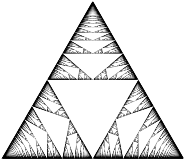
III - RESULT : construction of a C-adic word with infinite imbalance

1D : Euclid's algorithm

Dynamical system	$\begin{array}{rcl} \mathbb{R}^2_+ & \longrightarrow & \mathbb{R}^2_+ \\ (x,y) & \longmapsto & (x-y,y) \text{ if } x \geq y \\ & & (x, y-x) \text{ otherwise.} \end{array}$
Associated words	sturmian words
Substitutions set	$\tau_1 : \begin{array}{rcl} 1 & \mapsto & 1 \\ 2 & \mapsto & 21 \end{array}$ $\tau_2 : \begin{array}{rcl} 1 & \mapsto & 12 \\ 2 & \mapsto & 2 \end{array}$
(Im)balance	1-balanced
Rotation coding	✓

2D : Arnoux-Rauzy algorithm

Dynamical system	$\begin{array}{ccl} \mathcal{G} \subseteq \mathbb{R}^3_+ & \longrightarrow & \mathcal{G} \\ (x,y,z) & \longmapsto & (x-y-z,y,z) \text{ if } x \geq y+z \\ & & (x,y-x-z,z) \text{ if } y \geq x+z \\ & & (x,y,z-x-y) \text{ if } z \geq x+y. \end{array}$
Associated words	Arnoux-Rauzy words
Substitutions set	$\sigma_1 : \begin{array}{ccc} 1 & \mapsto & 1 \\ 2 & \mapsto & 21 \\ 3 & \mapsto & 31 \end{array}$ $\sigma_2 : \begin{array}{ccc} 1 & \mapsto & 12 \\ 2 & \mapsto & 2 \\ 3 & \mapsto & 32 \end{array}$ $\sigma_3 : \begin{array}{ccc} 1 & \mapsto & 13 \\ 2 & \mapsto & 23 \\ 3 & \mapsto & 3 \end{array}$
(Im)balance	There exist AR words with infinite imbalance.
Rotation coding	X



The Rauzy gasket $\mathcal{G}$.

A new candidate for 2D : Cassaigne-Selmer algorithm

Dynamical system	$\begin{array}{ccl} \mathbb{R}^3_+ & \longrightarrow & \mathbb{R}^3_+ \\ (x,y,z) & \longmapsto & (x-z,z,y) \text{ if } x \geq z \\ & & (y,x,z-x) \text{ otherwise.} \end{array}$
Associated words	C-adic words
Substitutions set	$c_1 : \begin{array}{lcl} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array}$ $c_2 : \begin{array}{lcl} 1 & \mapsto & 2 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 3 \end{array}$
(Im)balance	?
Rotation coding	?

$C := \{c_1, c_2\}$.

What can be said about the imbalance of C-adic words ?

A new candidate for 2D : Cassaigne-Selmer algorithm

Dynamical system	$\begin{array}{ccl} \mathbb{R}^3_+ & \longrightarrow & \mathbb{R}^3_+ \\ (x,y,z) & \longmapsto & (x-z,z,y) \text{ if } x \geq z \\ & & (y,x,z-x) \text{ otherwise.} \end{array}$
Associated words	C-adic words
Substitutions set	$c_1 : \begin{array}{ccc} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array}$ $c_2 : \begin{array}{ccc} 1 & \mapsto & 2 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 3 \end{array}$
(Im)balance	There exists C-adic words with infinite imbalance.
Rotation coding	?

$C := \{c_1, c_2\}$.
What can be said about the imbalance of C-adic words ?

Definitions, general problem

Be S a finite set of non-erasing word morphisms over an alphabet $\mathcal{A}$.

Definitions

- $w \in \mathcal{A}^{\mathbb{N}}$ is a S-adic word if there exist :
  - a seed $a \in A$
  - a directive sequence $(\sigma_k) \in S^{\mathbb{N}}$ s.t. $w = \lim_{n\to\infty} \sigma_0 \circ \ldots \circ \sigma_{n-1}(a)$.
- $u, v \in \mathcal{A}^*$. The abelianized of u is $ab(u) := (|u|_l)_{l\in\mathcal{A}} \in \mathbb{N}^{\mathcal{A}}$.
  We have : $-ab(u).(1)_{l\in\mathcal{A}} = |u|$
  $-[ab(u) - ab(v)].(1)_{l\in\mathcal{A}} = 0 \iff |u| = |v|$.
  In that case, $ab(u) - ab(v)$ is called the imbalance vector of u and v.
  If $w \in \mathcal{A}^{\mathbb{N}}$, the imbalance of w is :
  $D(w) := \sup_{u,v\in F(w),|u|=|v|} ||ab(u) - ab(v)||_\infty \in \mathbb{N} \cup \{\infty\}$.

General problem : is the imbalance of S-adic words finite ? bounded ?

Ex : Standard sturmians : yes and yes ; standard Arnoux-Rauzy : no and no.

Idea : introduce $\mathcal{F} := \bigcup_{w \text{ S-adic}} \{ab(u) - ab(v)/u, v \in F(w) and |u| = |v|\} \subseteq \mathbb{Z}^{\mathcal{A}}$.

Lemma : The imbalance of S-adic words is bounded if and only if $\mathcal{F}$ is finite.

$\longrightarrow$ We want to explore $\mathcal{F}$.

We need a bigger set $\mathcal{S}$

Where does an imbalance vector come from ? Let's desubstitute !

$$x \in \mathcal{F} \quad \Rightarrow \quad \exists w \text{ S-adic}, u, v \in F(w) \text{ s.t. } ab(u) - ab(v) = x$$

$$\Rightarrow \quad \exists (\sigma_k) \in S^{\mathbb{N}}, a \in \mathcal{A} \quad [\text{s.t. } w = lim \sigma_0 \circ \ldots \circ \sigma_{n-1}(a)]$$
$$\exists n_0 \in \mathbb{N}, \exists u, v\mathcal{A}^* \text{ s.t. } u, v \in F(\sigma_0 \circ \ldots \circ \sigma_{n_0 - 1}(a)) \text{ and } ab(u) - ab(v) = x.$$

$$\Rightarrow \quad \exists n_0 \in \mathbb{N}, \exists (\sigma_k) \in S^{n_0}, \exists a \in \mathcal{A},$$
$$\underline{\exists\, u, v \in F(\sigma_0 \circ \ldots \circ \sigma_{n_0 - 1}(a))} \quad \text{s.t.} \quad ab(u) - ab(v) = x$$

Where does they come ?

$$\sigma_1 \circ \ldots \circ \sigma_{n_0 - 1}(a) = l_0 \ldots\ldots l_i \ldots\ldots l_j \ldots\ldots l_k \quad [\tilde{u} = l_i \ldots l_j]$$
$$\sigma_0 \circ \sigma_1 \circ \ldots \circ \sigma_{n-1}(a) = \sigma_0(l_0) \ldots\ldots \sigma_0(l_i) \ldots\ldots \sigma_0(l_j) \ldots\ldots \sigma_0(l_k) \quad [\mathbf{u}]$$

Problem : $\tilde{u}$ and $\tilde{v}$ may not have the same length !

Solution : we have to consider a bigger set :

$$\mathcal{S} := \bigcup_{w \text{ S-adic}} \{ab(u) - ab(v) / u, v \in F(w)\} \subseteq \mathbb{Z}^{\mathcal{A}}.$$

On this bigger set, we are going to study the converse of the desubstitution

which is NOT the substitution

...but the 'substitute and cut' operation.

The substitute and cut operation on couples of factors

Def : Be $u, \widetilde{u}, v, \widetilde{v} \in \mathcal{A}^*$. Denote by $\alpha(u)$ and $\omega(u)$ the first and last letter of u ; and by $p_k(u)$ [$s_k(u)$] the prefix [suffix] of u of length k.

- A substitute and cut operation from $\widetilde{u}$ to u is a triplet $(\sigma, \beta, \gamma) \in S \times \mathbb{N}^2$ s.t. :
  - $p_\beta(\sigma(\widetilde{u})).u.s_\gamma(\sigma(\widetilde{u})) = \sigma(\widetilde{u})$
  - Cutting conditions : $\begin{cases} \beta = \gamma = 0 \text{ if } \widetilde{u} = \epsilon \text{ (empty word)} \\ \beta + \gamma \leq |\sigma(\widetilde{u})| \text{ and } \beta, \gamma < |\sigma(\widetilde{u})| \text{ if } |u| = 1 \\ \beta < |\sigma(\alpha(\widetilde{u}))| \text{ and } \gamma < |\sigma(\omega(\widetilde{u}))| \text{ otherwise.} \end{cases}$ (✂)
- A substitute and cut operation from $(\widetilde{u}, \widetilde{v})$ to (u, v) is $(\sigma, \beta, \gamma, \delta, \eta) \in S \times \mathbb{N}^4$ s.t. :
  - (σ, β, γ) is a S&C operation from $\widetilde{u}$ to u
  - (σ, δ, η) is a S&C operation from $\widetilde{v}$ to v.

  We denote it ${}^{\beta}_{\delta}\sigma^{\gamma}_{\eta}$.
- A quintuplet $(\sigma, \beta, \gamma, \delta, \eta)$ which satisfy (✂) is said to be allowed.

Ex : $S = C = \left\{ \begin{array}{llll} c1: & 1 \mapsto 1 & \quad c2: & 1 \mapsto 2 \\ & 2 \mapsto 13 & & 2 \mapsto 13 \\ & 3 \mapsto 2 & & 3 \mapsto 3 \end{array} \right\}$ $u = 23$ $v = 33$

There are 4 allowed S&C operations from (u,v) : ${}^0_0c1^0_0$, ${}^1_0c1^0_0$, ${}^0_0c2^0_0$ and ${}^1_0c2^0_0$, which give respectively $(132, 22), (32, 22), (133, 33)$ and $(33, 33)$.

A substitute and cut operation on $\mathcal{S}$?

$\longrightarrow$ We want the S&C operations to be the transitions of our automaton...

Problem : $x \in \mathcal{S}$ can represent in the same time (u, v) and (u', v') for which the sets of allowed S&C operations (and the result they give) are different !

Ex : $x = (0, 0, 0)$ represents $u = v = 132$ as well as $u' = v' = \epsilon$.
${}^0_0c1^1_0$ is allowed for (u, v) but not for (u', v').

Solution : to burst the vectors of $\mathcal{S}$ by adding a 'matrix of extremities' containing the first/last letters of (u, v) - and instead of $\mathcal{S}$, work on :

$$\mathcal{S}' := \bigcup_{w \text{ S-adic}} \{(M_{ext}(u, v), ab(u) - ab(v)), u, v \in F(w)\}.$$

Ex : $\begin{pmatrix} a & b \\ c & c \end{pmatrix}, (1, 1, -1)$ represents (acb, cc), ...
$\begin{pmatrix} a & . \\ c & c \end{pmatrix}, (1, 0, -2)$ represents (a, cc).
$\begin{pmatrix} . & . \\ . & . \end{pmatrix}, (0, 0, 0)$ represents (ϵ, ϵ).

Theorem : The S&C operation is well defined on $\mathcal{S}'$.

The imbalances automaton for S-adic words

We consider the automaton s.t. :

- states : $\mathcal{S}'$
- final states : $\mathcal{F}' = \{(M,x) \in \mathcal{S}' \text{ s.t. } x.(1)_{l\in\mathcal{A}} = 0\}$
- transitions : $X \xrightarrow{{}^{\beta}_{\delta}\sigma^{\gamma}_{\eta}} Y$ whenever ${}^{\beta}_{\delta}\sigma^{\gamma}_{\eta}$ is a S&C operation from X to Y.

Can we construct it ? Does there exist a finite set $\mathcal{I} \subseteq \mathcal{S}'$ s.t. from $\mathcal{I}$ you can reach each elements of $\mathcal{S}'$?

Theorem : If $\forall a \in \mathcal{A}$ there exists a S-adic word w s.t. $a \in F(w)$, then there exists a finite set $\mathcal{I} \in \mathcal{S}'$ s.t. :
$\forall X \in \mathcal{S}', \exists X_0 \in \mathcal{I}, \exists (T_i)_{i\in\{0,n_0-1\}} = ({}^{\beta i}_{\delta i}\sigma i^{\gamma i}_{\eta i})_{i\in\{0,n_0-1\}}$ a finite sequence of allowed substitute and cut operations s.t. :

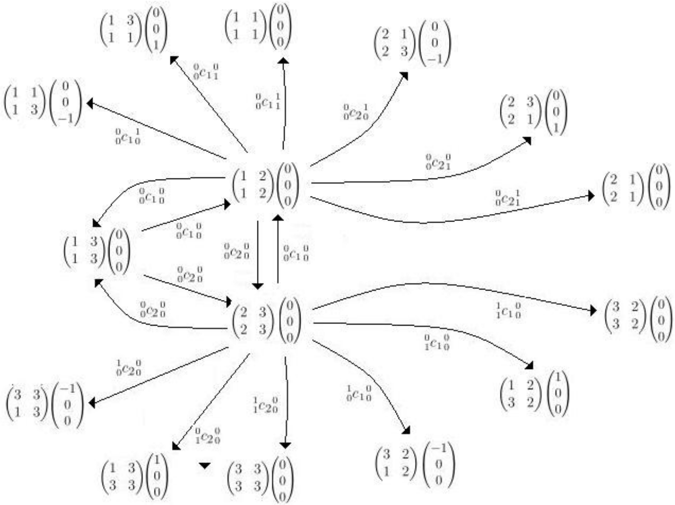
$$X = T_{n_0-1} \circ \ldots \circ T_0(X_0).$$

We can take $\mathcal{I} = \left\{ \begin{pmatrix} a & \cdot \\ a & \cdot \end{pmatrix}, a \in \mathcal{A} \right\}$.

First steps of computation with C

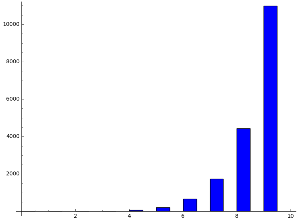
$C = \{c_1, c_2\}$

$$c_1 : \begin{array}{lcl} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array} \qquad c_2 : \begin{array}{lcl} 1 & \mapsto & 2 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 3 \end{array}$$

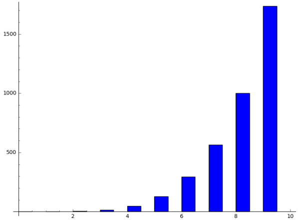


Experimentation (sad) reality

Problem : the tree grows too fast !



Number of vertices in function of depth



Growth after cuttings

Solution : cut branches with no hope to reach new final states...

At depth 9, among 1 000 vertices, we found the first imbalance 3...
At depth 16, among 80 000 vertices, we found the first imbalance 4...

Results for C-adic words

Be w_0 any C-adic word, e.g. $c_1 \circ c_2 \circ c_1 \circ c_2 \circ \ldots (1)$. Consider $w_1 = c_2 \circ c_2 \circ c_2(w_0)$ and for each $n \geq 1$:

$$\begin{cases} w_{n+1} = c_1^{2n+2} \circ c_2(w_n) & \text{if } n \text{ is odd} \\ w_{n+1} = c_2^{2n+2} \circ c_1(w_n) & \text{otherwise.} \end{cases}$$

Theorem 1 : For every n, w_n is a C-adic word satisfying $D(w_n) \geq n$.

$\longrightarrow$ The imbalance of C-adic words is not bounded.

Theorem 2 : There exists a C-adic word with infinite imbalance.

This is a construction from $(w_n)_n$ using the following lemma :

Lemma : If w is a C-adic word s.t. $D(w) \geq 3n$, then $c_1(w)$ (resp. $c_2(w)$) is a C-adic word satisfying $D(w) \geq n$.

Moral & remaining questions

- The imbalance automaton gives intuitions on the nature (bounded/ unbounded) of the set of imbalances of S-adic words - and on rules to construct these imbalances.

- Difficulties :
  - the choice of an initial set $\mathcal{I}$
  - the growing speed.

- Miscellaneous questions :
  - measure of C-adic words with infinite imbalance ?
  - what can be said for imbalances of C-adic words whose the number of consecutive occurencies of a same substitution in the directive sequence is bounded ?
  - Rotation coding ?
  - Does there exist S such that imbalances of S-adic word are finite but not bounded ?
  - Choice of an initial set $\mathcal{I}$ when the condition of the theorem is not satisfied ?

Thank you !