Systèmes dynamiques symboliques exceptionnels engendrés par des algorithmes de fractions continues multidimensionnelles

Mélodie Andrieu

Université Bar-Ilan

I3S, Sophia Antipolis, juin 2022.

I - Motivations et questions

$\longrightarrow$ à la recherche d'un bon algorithme de FC multiD

II - Résultats

$\longrightarrow$ une sous-classe de systèmes symboliques étranges

III - Définitions formelles

IV - Un outil général derrière ces résultats

$\longrightarrow$ exploration patiente et ordonnée d'un graphe infini

I. Motivations et questions

Les fractions continues, qu'est-ce que c'est ?

"Fraction continue" = façon de représenter les réels par des suites infinies d'entiers.

$\rightsquigarrow$ Contrairement à l'écriture en base 2, 10, n, ..., les entiers ne sont *pas bornés* a priori.

Exemples.

	écriture décimale	développement en FC
$\frac{19}{12}$	1.5833333...	$[1; 1, 1, 2, 2]$
$\sqrt{13}$	3.605551275...	$[3; 1, 1, 1, 6, 1, 1, 1, 1, 6, 1, 1, 1, 1, 6...]$
π	3.1415926535...	$[3; 5, 15, 1, 292, 1, ...]$
φ	1.618033988...	$[1; 1, 1, 1, 1...]$

Comment retrouver x à partir de son dvp en FC ?

Soit $x \in \mathbb{R}_{\geq 0}$

- **écriture décimale :**

$$...0d_{-m}\dots d_{-1}d_0.d_1d_2d_3... \qquad d_i \in \{0,\dots,9\} \qquad \rightsquigarrow \qquad x = \sum_{j=-m}^{+\infty} d_j 10^{-j}$$

- **développement en FC :**

$$[a_0; a_1, a_2, a_3, \dots] \;\; a_i \in \mathbb{N} \qquad \rightsquigarrow \qquad x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\dots}}}}$$

Exemple :

$$\frac{19}{12} = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2}}}}$$

Comment calculer le dvp en FC d'un réel ?

Le dvp en FC est fondé sur l'algorithme d'Euclide [soustractif].

$$F: \quad \begin{array}{ccl} \mathbb{N}\times\mathbb{N} & \longrightarrow & \mathbb{N}\times\mathbb{N} \\ (y,z) & \longmapsto & \begin{cases} (y-z,z) & \text{si } y \geq z \quad \text{(cas 1)} \\ (y,z-y) & \text{sinon} \quad \text{(cas 2)} \end{cases} \end{array}$$

Exemple :

$$(19,12) \xrightarrow{1} (7,12) \xrightarrow{2} (7,5) \xrightarrow{1} (2,5) \xrightarrow{2} (2,3) \xrightarrow{2} (2,1) \xrightarrow{1} (1,1) \xrightarrow{1} \underset{\substack{\uparrow \\ \mathrm{pgcd}(\mathbf{19},\mathbf{12})}}{(0,1)} \circlearrowright^{2}$$

Def : La **trajectoire symbolique** de $(19,12)$ pour l'itération de F [relativement à sa définition par morceau] est : 1212211222222...

Comment calculer le dvp en FC d'un réel ? [suite]

$$F: \begin{array}{ccc} (\mathbb{R}_{\geq 0})^2 & \longrightarrow & (\mathbb{R}_{\geq 0})^2 \\ (y,z) & \longmapsto & \begin{cases} (y-z, z) & \text{si } y \geq z \quad \text{(cas 1)} \\ (y, z-y) & \text{sinon} \quad \text{(cas 2)} \end{cases} \end{array}$$

Observation. Soit $\lambda \in \mathbb{R}_{>0}$. Les trajectoires symboliques de (y,z) et $\lambda(y,z)$ sont identiques.

Conséquence 1. $\mathrm{traj.\ symb}(y,z) = \mathrm{traj.\ symb}(\frac{y}{z}, 1)$. On peut considérer que F est définie sur le cône projectif $(\mathbb{R}_{\geq 0})^2/\mathbb{R}_{>0}$.

Conséquence 2. $\dfrac{y}{z} \in \mathbb{Q} \iff \mathrm{traj.\ symb}(y,z)$ est ultimement constante.

Comment calculer le dvp en FC d'un réel ? [suite et fin]

def : Soit $x \in \mathbb{R}_{\geq 0}$. Le "développement en fraction continue" de x est $[a_0; a_1, a_2, a_3, \ldots]$, où les $a_i \in \mathbb{N}$ sont *déterminés de façon unique* par :

(i) $a_1, a_2, a_3, \ldots \neq 0$,
(ii) $1^{a_0} 2^{a_1} 1^{a_2} 2^{a_3} \ldots$ est la trajectoire symbolique de $(x, 1)$ sous F.

Convention : $[a_0; a_1, a_2, \ldots, a_n, \infty] = [a_0; a_1, a_2, \ldots, a_n]$.

Ex : Si $x = \frac{19}{12}$ alors

$$\text{traj. symb}\left(\frac{19}{12}, 1\right) = \text{traj. symb}(19, 12) = 121221122222\ldots = 1^1 2^1 1^1 2^2 1^2 2^\infty.$$

$\rightsquigarrow$ donc le dvp en FC de $\frac{19}{12}$ est $[1; 1, 1, 1, 2, 2]$.

Intérêt des fractions continues

Le développement en fraction continues :

- caractérise les rationnels : $\boxed{x \in \mathbb{Q} \iff \text{dvp en FC fini}}$
- caractérise les réels quadratiques : :

$$\boxed{x \text{ quadratique} \iff \text{dvp en FC ultimement périodique}}$$

- Fournit les meilleures approximations diophantiennes :

Soit $x = [a_0; a_1, a_2, \ldots] \in \mathbb{R}_{\geq 0}$. Notons $\frac{p_n}{q_n} := [a_0; a_1, \ldots, a_n]$. Alors :

(i) $\left| x - \frac{p_n}{q_n} \right| < \frac{1}{q_n^2}$, $\rightsquigarrow$ "c'est une bonne approximation"

(ii) si $\frac{p}{q}$ est tel que $|qx - p| < |q_n x - p_n|$ alors $q \geq q_n$.

$\rightsquigarrow$ "on ne trouvera pas mieux sans augmenter q"

Des fractions continues en dimension supérieure ?

Depuis Jacobi, on cherche à généraliser les fractions continues en dimension supérieure.

But : - caractériser les réels cubiques (problème d'Hermite);
- détecter les dépendances rationnelles;
- fournir de bonnes approximations diophantiennes simultanées.

Comment ? en proposant un "algorithme d'Euclide" sur les triplets

$$F_{3D}: \begin{array}{ccc} (\mathbb{R}_{\geq 0})^3 & \to & (\mathbb{R}_{\geq 0})^3 \\ (x,y,z) & \mapsto & ??? \end{array}$$

...Mais quelle opération doit être exécutée ??

$\rightsquigarrow$ multiplicité des algorithmes de FC multiD.

...Une question réactivée par la dynamique symbolique

Définitions générales

(i) Un alphabet est un ensemble fini $\mathcal{A}$. ex : $\mathcal{A} = \{1, 2\}$

(ii) Un mot de longueur n est un élément de $\mathcal{A}^n$.

On note $\mathcal{A}^* := \bigcup_{n\in\mathbb{N}} \mathcal{A}^n$ ex : $u = 122$

(iii) Un mot infini est un élément de $\mathcal{A}^{\mathbb{N}}$.

On munit $\mathcal{A}^{\mathbb{N}}$ de la **topologie produit**.

(iv) Une substitution est un élément de $\mathrm{End}\big((\mathcal{A}^*, \cdot)\big)$.

Ex :

$$\sigma : \begin{array}{l} 1 \mapsto 1 \\ 2 \mapsto 12 \end{array} \qquad \sigma(122) = 1 \cdot 12 \cdot 12 = 11212.$$

Une classe de mots infinis associée à la fraction continue usuelle

À $F : (y,z) \mapsto \begin{cases} (y-z, z) & \text{si } y \geq z \quad \text{cas 1} \\ (y, z-y) & \text{sinon} \quad \text{cas 2} \end{cases}$ on associe 2 substitutions $\tau_1 : \begin{cases} 1 \mapsto 1 \\ 2 \mapsto 12 \end{cases}$ $\tau_2 : \begin{cases} 1 \mapsto 21 \\ 2 \mapsto 2 \end{cases}$

prop : Soit $(s_n) \in \{\tau_1, \tau_2\}^{\mathbb{N}}$ contenant une infinité de chacune des deux substitutions. Soit $a \in \{1,2\}$. La suite $\big(s_0 \circ s_1 \circ \ldots \circ s_n(a)\big)_{n\in\mathbb{N}}$ converge vers un mot infini $w \in \{1,2\}^{\mathbb{N}}$ qui ne dépend pas de a.

def : - Ces mots sont les **"mots sturmiens [standards]"**.
- (s_n) est appelée **suite directrice**.

Ex :

$$w_{fib} = \lim_{n\to+\infty} \tau_1 \circ \tau_2 \circ \tau_1 \circ \tau_2 \circ \tau_1 \circ \tau_2 \ldots (1) = \underbrace{121121211211212112\ldots}_{\text{"mot de Fibonacci"}}$$

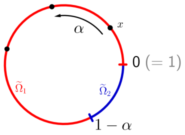
Mots sturmiens

Les mots sturmiens jouissent de nombreuses caractérisations combinatoires, géométriques et dynamiques.

- **[complexité]** w est sturmien ssi pour tout n, w admet exactement $n+1$ facteurs de longueur n [facteur = sous-mot lu avec des lettres consécutives].

- **[équilibre]** w est sturmien ssi w est apériodique et tout couple de facteurs de même longueur admet, à un près, le même nombre de 1 (et donc à un près aussi, le même nombre de 2).

- **[codage de rotation]** Un mot est sturmien ssi il encode la trajectoire d'un point sous l'action d'une rotation irrationnelle, par rapport à une partition bien choisie.



etc.

Classes de mots infinis associées aux algo de FC multiD (prototype)

Idée : Étant donné un algorithme de FC en dimension d noté F_d,

- proposer k substitutions sur l'alphabet $\{1, \ldots, d\}$ associées à la définition en k morceaux de F_d.
- est-ce que les mots
$$w =? \lim_{n \to +\infty} s_0 \circ \ldots \circ s_n(a)$$
où $a \in \{1, \ldots d\}$ et $(s_n)_{n \in \mathbb{N}}$ est une suite à valeur dans ces k substitutions existent ? Sont-ils intéressants ?

Une caractérisation combinatoire [∼1940]

def : Facteur de longueur n de w = sous-mot de w formé par n lettres *consécutives*.

Ex : - 11 est facteur de 21211122212112...
- `art` est facteur de `montmartre`

Propriété d'équilibre :
Les mots sturmiens sont exactement les mots binaires apériodiques dans lesquels toute paire de facteurs de même longueur contient, à une unité près, le même nombre de 1.

Ex :
Un mot commençant par $w = 11211121121112112\ldots$ peut-être sturmien.
Un mot commençant par $w' = 1\underline{22}12221\underline{11}$... ne l'est pas.

Conséquences de l'équilibre

Soit w un mot sturmien.

1. La proportion de 1 parmi les N premières lettres de w tend vers une limite f_1.

2. La trajectoire symbolique de (f_1, f_2) sous l'itération de F est la suite directrice de w.

   Corollaire : $f_1/f_2 \in \mathbb{R}\backslash\mathbb{Q}$

   Ex : Le vecteur fréquence des lettres du mot de Fibonacci est : ...

3. **Mieux :** il existe C tq pour tout N, le nombre de 1 parmi les N premières lettres de w appartient à $[Nf_{\mathbf{1}} - C; Nf_{\mathbf{1}} + C]$.

Interprétation géométrique

La "ligne brisée" constituée des points $P_N := \sum_{n=0}^{N} e_{w[n]}$, où (e_1, e_2) est la base canonique de $\mathbb{R}^2$, reste à distance bornée de sa direction moyenne.

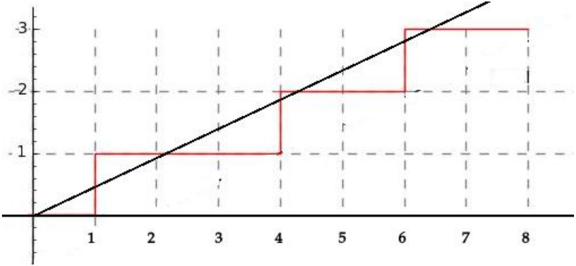


Figure: La ligne brisée de 12111211211...

⟶ Les mots sturmiens approximent les droites de pentes irrationnelles.

[dim 2] L'algorithme d'Arnoux-Rauzy [∼1991]

$$F_{AR} : \begin{array}{ccll} \ldots & \to & \ldots & \\ (x,y,z) & \mapsto & (x-y-z,y,z) & \text{si } x > y+z, \\ & & (x,y-x-z,z) & \text{si } y > x+z, \\ & & (x,y,z-x-y) & \text{si } z > x+y. \end{array}$$

Substitutions : $\sigma_1 : \begin{array}{l} 1 \to 1 \\ 2 \to 12 \\ 3 \to 13 \end{array}$; $\sigma_2 : \begin{array}{l} 1 \to 21 \\ 2 \to 2 \\ 3 \to 23 \end{array}$ et $\sigma_3 : \begin{array}{l} 1 \to 31 \\ 2 \to 32 \\ 3 \to 3. \end{array}$

prop : Soit $(s_n) \in \{\sigma_1, \sigma_2, \sigma_3\}^{\mathbb{N}}$ contenant une infinité de chacune des trois substitutions. Soit $a \in \{1,2,3\}$. La suite $\big(s_0 \circ s_1 \circ \ldots \circ s_n(a)\big)_{n\in\mathbb{N}}$ converge vers un mot infini $w \in \{1,2\}^{\mathbb{N}}$ qui ne dépend pas de a.

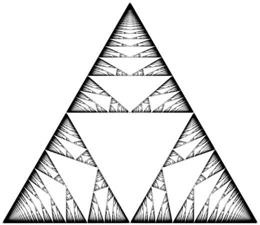
$\rightsquigarrow$ On les appelle mots d'Arnoux-Rauzy [standards].

Ex: $w_{Trib} = 1213121121312121312112131213121312\ldots$ [mot de "Tribonacci", 1982]

[dim 2] L'algorithme d'Arnoux-Rauzy [∼1991]

Contrariété :
F_{AR} n'est pas définie lorsque la somme des petites coordonnées excède la plus grande.

"Baderne de Rauzy" = ensemble des (x, y, z), avec $x + y + z = 1$, pour lesquels F_{AR} peut être itérée indéfiniment.



La baderne de Rauzy

Prop : La baderne de Rauzy est de mesure nulle. [AS13]

Prop : Sa dimension de Hausdorff est <2. [AHS16]

[dim 2] L'algorithme de Cassaigne-Selmer [∼2017]

$$F_C : \begin{array}{ccccl} (\mathbb{R}_{\geq 0})^3 & \rightarrow & (\mathbb{R}_{\geq 0})^3 & & \\ (x,y,z) & \mapsto & (x-z,z,y) & & \text{si } x \geq z, \\ & & (y,x,z-x) & & \text{sinon.} \end{array}$$

$\longrightarrow$ F_C est définie sur tous les triplets !

Cet algorithme engendre la classe des **mots de Cassaigne-Selmer**, à travers les substitutions :

$$c_1 : \begin{array}{ccc} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array} \quad \text{et} \quad c_2 : \begin{array}{ccc} 1 & \mapsto & 2 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 3. \end{array}$$

Ex: $w_{ex} = c_1 \circ c_2 \circ c_1 \circ c_2 \circ c_1 \cdots (1) = 1321213121321312132121321312132 1...$

Questions...

Fait : Chaque mot d'AR et de CS admet un vecteur fréquence des lettres (f_1, f_2, f_3).

1. Que dire de ces vecteurs ?

2. A quelle vitesse le nombre d'occurrences moyen de chaque lettre converge vers sa fréquence attendue – c.a.d géométriquement, comment se comporte la ligne brisée 3D associée à un tel mot par rapport à sa direction moyenne [la droite portée par (f_1, f_2, f_3)] ?

II. Résultats

1. Vecteurs fréquence des lettres

Theorème [Cassaigne, Labbé, Leroy 2017]

Le vecteur fréquence des lettres de n'importe quel mot de Cassaigne-Selmer a des coordonnées rationnellement indépendantes.

Theorème [Dynnikov, Hubert, Skripchenko 21; A. 21]

Le vecteur fréquence des lettres de n'importe quel mot d'Arnoux-Rauzy a des coordonnées rationnellement indépendantes.

$\longrightarrow$ conjecturé en 2013 par Arnoux and Starosta.

(Un beau résultat, plus fort)

Théorème [A. 21]

Soit $d \geq d' \geq 1$. Soit w un mot episturmien sur $\{1, ..., d\}$. Notons $f = (f_1, ..., f_d)$ son vecteur fréquence des lettres **[qui existe]**, et par $(s_n)_{n\in\mathbb{N}}$ l'une des suites de substitutions par laquelle il a été obtenu. Les assertions suivantes sont équivalentes.

1. Exactement d' substitutions apparaissent infiniment souvent dans $(s_n)_{n\in\mathbb{N}}$.
2. La dimension de l'espace vectoriel $f_1\mathbb{Q} + ... + f_d\mathbb{Q}$ est d'.

** "mot épisturmien sur $\{1, ..., d\}$" = mot engendré par une *composition quelconque* des d substitutions associées à l'algo d'AR en dimension d, consistant à soustraire la somme des $d-1$ petites composantes à la plus grande...

$\longrightarrow$ **L'algorithme d'Arnoux-Rauzy (généralisé) détecte toutes les dépendances rationnelles !**

2) Propriétés de leurs lignes brisées ?

Fractal de Rauzy = adhérence de la projection de la ligne brisée d'un mot infini, parallèlement à son vecteur fréquence.

Figure: Fractal de Rauzy de w_{trib}

2) Propriétés de leurs lignes brisées / fractals de Rauzy ?

Vieille croyance : "la ligne brisée de tout mot d'AR reste à distance bornée de sa direction moyenne ; autrement-dit, son fractal de Rauzy est borné."

$\longrightarrow$ Vrai pour tous les cas connus...

$\longrightarrow$ contredite par la construction d'un exemple remarquable par [Cassaigne, Ferenczi, Zamboni 2000]

$\longrightarrow$ Toutefois, ces mots "déséquilibrés" sont de mesure nulles. [Delecroix, Hejda, Steiner 2013]

Question : Quid pour les mots de CS ?

2) Propriétés de leurs lignes brisées / fractals de Rauzy ?

Théorème [A. 18]

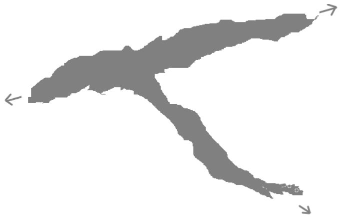
Il existe un mot de Cassaigne-Selmer dont le fractal de Rauzy n'est pas borné.

Aujourd'hui, on ne sait presque rien des propriétés topologiques et géométriques de ces fractals non bornés.

2) Propriétés de leurs lignes brisées / fractals de Rauzy ?

"Conjecture" : La ligne brisée de tout mot d'AR reste à distance bornée d'un plan contenant sa direction moyenne ; autrement-dit, son fractal de Rauzy est contenu dans une bande du plan.

"possible"



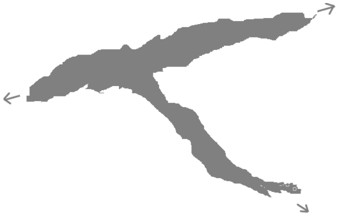
"impossible"

⟶ Intuition suggérée par le théorème d'Oseledets. En effet, si les exposants de Lyapounov du produit infini de matrices associé à w existent, l'un de ces exposants au moins est <0, puisque qu'il y en a déjà un >0, et que leur somme doit être égale à 0.

2) Propriétés de leurs lignes brisées / fractals de Rauzy ?

"Conjecture" : La ligne brisée de tout mot d'AR reste à distance bornée d'un plan contenant sa direction moyenne ; autrement-dit, son fractal de Rauzy est contenu dans une bande du plan.

"possible"



"impossible"

$\longrightarrow$ Intuition suggérée par le théorème d'Oseledets. En effet, si les exposants de Lyapounov du produit infini de matrices associé à w existent, l'un de ces exposants au moins est <0, puisque qu'il y en a déjà un >0, et que leur somme doit être égale à 0.

Cette conjecture est fausse.

Theorème [A. 21]

Il existe des mots d'AR et de CS dont le fractal de Rauzy n'est contenu dans aucune bande du plan.

$\rightarrow$ Nous construisons explicitement des familles de contre exemples.

Méthode

- Réduire le problème dynamique à une question de combinatoire (ici : existe-t-il w de "déséquilibre" infini?)
- Aborder la question de combinatoire avec des outils d'informatique théorique (ici : exploration d'un graphe infini)
- Développer une intuition par l'expérimentation (ici : exploration ordonnée d'une grande partie du graphe).

III. Compléments de définitions

Abélianisation

"L'abélianisé" d'un mot fini u est le vecteur $\mathrm{ab}(u) = (|u|_a)_{a\in A}$, où $|u|_a$ désigne le nombre de a dans u.

Ex: ab(212232)= (1,4,1)

La **"matrice d'incidence"** d'une substitution σ est $M_\sigma := (|\sigma(j)|_i)_{i,j\in A}$.

Ex: $c_1 : \begin{array}{ccc} 1 & \mapsto & 1 \\ 2 & \mapsto & 13 \\ 3 & \mapsto & 2 \end{array}$ donne $M_{c_1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$

Prop : On a toujours $\mathrm{ab}(\sigma(u)) = M_\sigma \cdot \mathrm{ab}(u)$.

Discrépance

Soit $w \in A^{\mathbb{N}}$. Nous souhaitons étudier la fonction :

$$\mathrm{discr}_w : n \longmapsto \max_{a \in A} \left| |\mathrm{pref}_n(w)|_a - Nf_a \right|$$

qui mesure la différence maximale entre le nombre d'occurrences d'une lettre dans le préfixe de longueur n de w, et sa valeur attendue.

$\longrightarrow$ Géométriquement, la discrépance mesure la distance entre la ligne brisée et sa direction moyenne.

Un pendant combinatoire : le déséquilibre

déséquilibre de w : $\quad \mathrm{imb}(w) := \sup_{n\in\mathbb{N}} \sup_{u,v\in\mathcal{F}_n(w)} ||\mathrm{ab}(u)-\mathrm{ab}(v)||_\infty \quad \in \mathbb{N}$ or ∞

** $\mathcal{F}_n(w)$ = ensemble des facteurs de longueur n de w.

$\longrightarrow$ C'est la *plus petite constante* D t.q. pour tout $u, v \in \mathcal{F}_n(w)$ et pour tout $a \in A$:

$$|u|_a - |v|_a \leq D.$$

ex : Le déséquilibre de n'importe quel mot sturmien est 1.

Fait : $1/4\, \mathrm{imb}(w) \leq \sup \mathrm{discr}_w \leq \mathrm{imb}(w)$ [Ada03]

La ligne brisée de w reste à distance bornée de sa direction moyenne si et seulement si le déséquilibre de w est fini.

Mots et sous-shifts S-adiques

Soit S un ensemble *fini* de substitutions sur un même alphabet $\mathcal{A}$.

"mot S-adique": un *mot infini* pouvant s'écrire : $w = \lim_{n\to\infty} s_0 \circ \ldots \circ s_{n-1}(a)$

avec : - $a \in \mathcal{A}$ $\longleftarrow$ "graine"

- $(s_n) \in S^{\mathbb{N}}$ $\longleftarrow$ "suite directrice"

ex : - Les mots sturmiens, d'Arnoux-Rauzy et de Cassaigne-Selmer standards
- en fait tous les mots fournis par des algos de FC...

Remarque : la suite directrice peut être vue comme un mot infini sur l'alphabet S...

IV. Un semi-algorithme pour explorer l'ensemble des déséquilibres de tous les mots S-adiques**

** où S est un ensemble fini, arbitraire mais fixé, de substitutions.

Avant-goût

Etant donné S un ensemble fini de substitutions, on souhaite répondre aux questions

1. Est-ce que les déséquilibres des mots S-adiques sont bornés ?
2. Si oui, donner un majorant.
3. Si non, pour n'importe quel $d \in \mathbb{N}$, exhiber un mot S-adique dont le déséquilibre est supérieur à d.

Notre outil : "l'automate des déséquilibres" , un graphe orienté infini tel que :

déséquilibres	$\approx$ état finaux
suites directrices	$\approx$ étiquette des chemins

Formellement...

Theorème [A. 21]

Soit S un ensemble de substitutions non effaçantes définies sur un même alphabet A, et supposons que chaque lettre de A apparaît dans un mot S-adique (pas forcément le même). Si D_S désigne la quantité (éventuellement infinie):

$$D_S = \sup_{w \text{ S-adique}} \operatorname{imb}(w),$$

alors un **parcours en largeur** dans "**l'automate des déséquilibres**", depuis ses états initiaux, fournira, pour chaque $d \leq D_S$, une suite finie de substitutions $(\sigma_i)_{i\in\{1,\dots,n\}}$ telle que tout mot S-adique dont la suite directrice commence par $(\sigma_i)_{i\in\{1,\dots,n\}}$, a un déséquilibre plus grand que d.

$\longrightarrow$ La question "est-ce qu'il existe un mot S-adique de déséquilibre plus grand que d ?" est **semi-décidable**.

$\longrightarrow$ S'il existe un mot S-adique de déséquilibre plus grand que d, le semi-algorithme va le trouver. Sinon, il ne termine pas.

Aperçu de l'automate des déséquilibres sur l'exemple de Cassaigne-Selmer

Sommets : $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

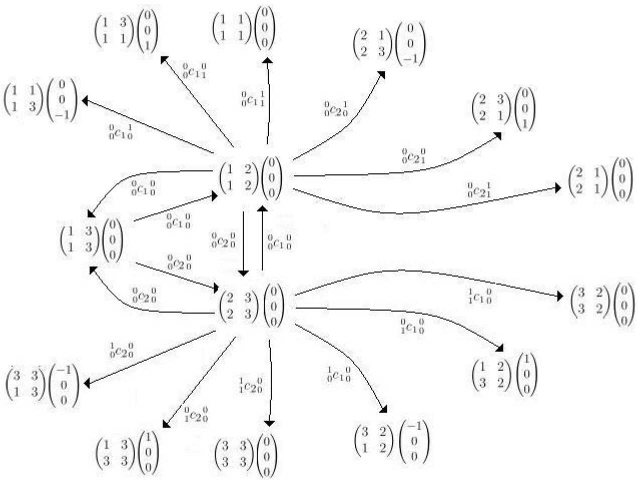
représentant tous les mots finis (u,v) tels que :

- u, v sont facteurs d'un *même* mot S-adique ;
- a et b sont la première et dernière lettre de u ;
- c et d sont la première et dernière lettre de v ;
- $\mathrm{ab}(u) - \mathrm{ab}(v) = (x, y, z)$.

Arcs : $X \longrightarrow Y$ ssi un/tout couple de facteurs représenté par Y est obtenu à partir d'un couple de facteurs (u, v) représenté par X, en substituant *simultanément* u et v par l'une des c_i, puis en rabotant éventuellement les mots obtenus d'une lettre à droite et/ou d'une lettre à gauche.

EXEMPLE

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



Une **portion** de l'automate des déséquilibres pour les substitutions de Cassaigne-Selmer.

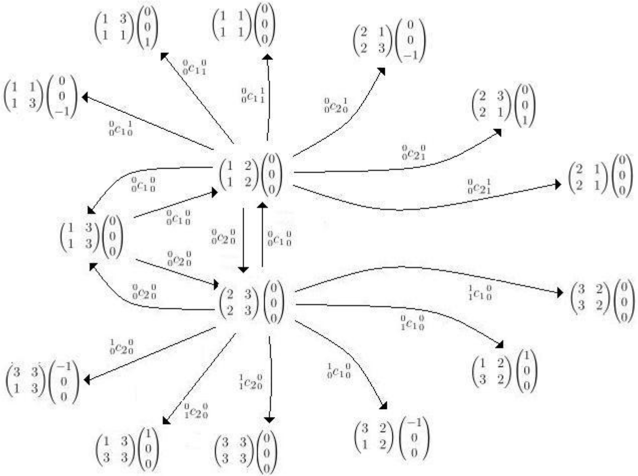
Propriétés de l'automate des déséquilibres

- On peut calculer Y *directement* à partir de X...
- Ce graphe n'est pas connexe.
- Toutefois, tous les sommets sont accessibles à partir d'un ensemble fini, connu, de sommets "initiaux".
- Tout sommet admet un nombre fini (et même borné) d'arêtes sortantes.

$\longrightarrow$ Bien qu'il soit infini, on peut [commencer à] parcourir ce graphe en largeur...

Intérêt : Il existe un mot S-adique de déséquilibre plus grand que d ssi on peut trouver un sommet (M, x) t.q. :
- la somme des coordonnées de x est égale à zéro;
- $||x||_\infty \geq d$.

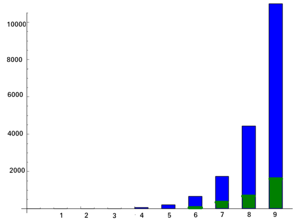
Exemple : donner un minorant du déséquilibre suprémal de la classe des mots de Cassaigne-Selmer...



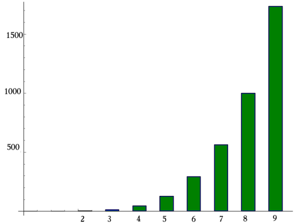
Implantation du semi-algorithme pour les substitutions de Cassaigne-Selmer et d'Arnoux-Rauzy

(Triste) réalité : exemple de Cassaigne-Selmer

Problème : l'arbre croît trop vite !



Nombre de sommets en fonction de la profondeur



Croissance après élagage

Solution : élaguer les branches qui ne conduiront jamais à un nouvel état final

A profondeur 9, parmi 1 500 sommets, apparaît le premier déséquilibre 3...
A profondeur 16, parmi 80 000 sommets, apparaît le premier déséquilibre 4...

Et pourtant, ça a marché...

And yet, results!

1) I managed to find back the families of words constructed in [CFZ00]!
2) I spotted families of words with growing imbalances for Cassaigne-Selmer words as well:

From any CS word w_0, construct:

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

Lemma 2: For all n, w_n is a CS word satisfying $\mathrm{imb}(w_n) \geq n$.

Theorem (A. 18)

There exists a Cassaigne-Selmer word w_∞ with infinite imbalance.

$\longrightarrow$ w_∞ is contructed from $(w_n)_n$ by a pumping method which relies on:

Lemma 3: If w is a CS word s.t. $\mathrm{imb}(w) \geq 3n$, then $w' := c_1(w)$ (resp. $c_2(w)$) is a CS satisfying $\mathrm{imb}(w') \geq n$.

And unexpected results!

Lemma 4: For any $(a, b, c) \in \mathbb{Z}^3$, there exists $s \in (S_{AR})^*$ and there exist $u, v \in \mathcal{F}(s(1))$ that satisfy $\mathrm{ab}(u) - \mathrm{ab}(v) = (a, b, c)$.

Rk: s can be explicitly described.

Theorem (A. 20)

There exists an Arnoux-Rauzy word $w_{\infty\infty}$ whose stepped line is not trapped between two parallel planes - or, equivalently, whose Rauzy fractal is not trapped between two parallel lines.

$\longrightarrow$ $w_{\infty\infty}$ is contructed from Lemma 4 by a pumping method which relies on the invertibility of incidences matrices of AR substitutions.

Merci !