On the complexity of a family of S-autom-adic sequences

Sébastien Labbé

Laboratoire d'Informatique Algorithmique : Fondements et Applications
Université Paris Diderot Paris 7

LIAFA
Université Paris-Diderot
Séminaire de l'équipe Automates
February 22, 2013

Joint work with Valérie Berthé

Outline

Plan

Introduction

Let $\Delta_d = \{(f_1, f_2, \ldots, f_d) \in \mathbb{R}_+^d : f_1 + f_2 + \cdots + f_d = 1\}$.

Question

Given a vector $(f_1, f_2, \cdots, f_d) \in \Delta_d$, can we construct an infinite word $\mathbf{w}$ *on the alphabet $\mathcal{A} = \{1, 2, \cdots, d\}$ such that the frequency of each letter $i \in \mathcal{A}$ exists and is equal to f_i ?*

In other words, can we find $g : \Delta_d \to \mathcal{A}^{\mathbb{N}}$ such that $\vec{f} \circ g = Id_{\Delta_d}$?

$$\Delta_d \xrightarrow{\quad g\, =? \quad} \mathcal{A}^{\mathbb{N}} \xrightarrow{\quad \vec{f} \quad} \Delta_d$$

where

$$\begin{array}{rccc} \vec{f}: & \mathcal{A}^{\mathbb{N}} & \to & \Delta_d \\ & \mathbf{w} & \mapsto & (f_1, f_2, \ldots, f_d) \end{array} \qquad \text{and} \qquad f_i = \lim_{p \text{ prefix of } \mathbf{w}} \frac{|p|_i}{|p|}.$$

Periodic words

If $(f_1, f_2, \ldots, f_d) \in \mathbb{Q}_+^d$, then one may think of a periodic word.

For example, if $(f_1, f_2, f_3) = (\frac{1}{6}, \frac{1}{3}, \frac{1}{2})$, then

$$\mathbf{w} = (122333)^\omega = 122333.122333.122333.122333\cdots$$

and

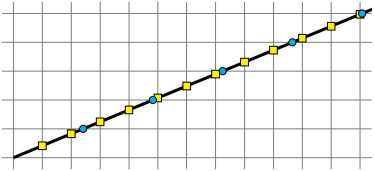
$$\mathbf{w} = (231323)^\omega = 231323.231323.231323.231323\cdots$$

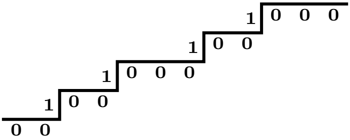
are possible answers. In fact, we want $\mathbf{w}$ to have nice properties, namely

- a linear factor complexity and
- a bounded balance.

Answer for $d = 2$

The Sturmian words of slope $\alpha \in \mathbb{R}_+$ form a well known family of infinite words having a linear factor complexity ($p(n) = n + 1$) and being balanced.





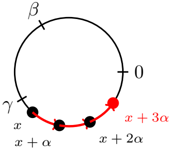
$$c_{1/\sqrt{2}} = \mathbf{0010010001001000100} \cdots$$

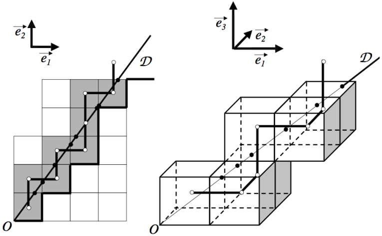
But this is not so easy in higher dimensions.

Possible answers for $d > 2$

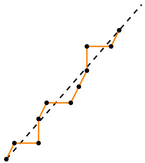
Typical answers to this question are :

	$\forall v \in \mathbb{R}_+^d$	$p(n)$ is linear	Balanced
Coding of Rotations	Yes	Yes	No
Coding of Interval Exchange Tr.	Yes	Yes	No
Billiard words	Yes	No	Yes
The Tribonacci word	No	Yes	Yes
Arnoux Rauzy words	No	Yes	Yes and No





Borel (2006)



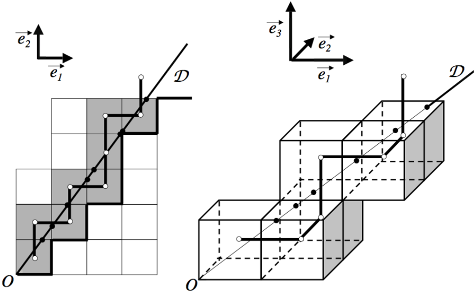
$$-7/2 \leq 2x - 5z < 7/2$$
$$-8/2 \leq 3x - 5y < 8/2$$
$$-5/2 \leq 2y - 3z < 5/2$$

Andres (2003)

Factor Complexity of the Billiard word

Theorem (Baryshnikov, 1995 ; Bédaride, 2003)

If both the direction $(\alpha_1, \alpha_2, \alpha_3)$ and $(\alpha_1^{-1}, \alpha_2^{-1}, \alpha_3^{-1})$ are $\mathbb{Q}$ independent, the number of factors appearing in the Billiard word in a cube is exactly $p(n) = n^2 + n + 1$.



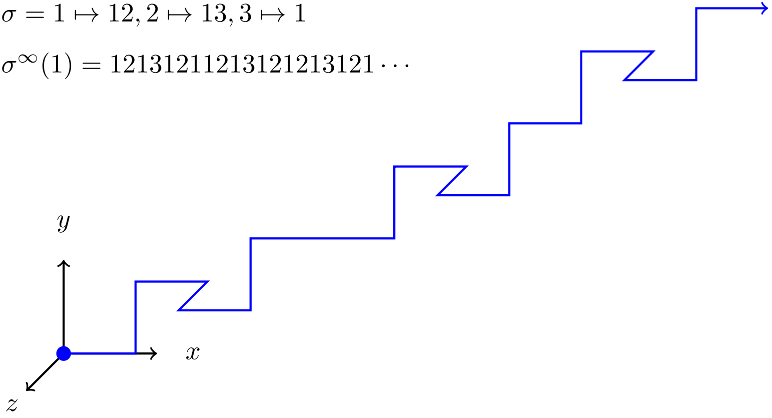
Tribonacci Example from Rauzy (1982)

The Tribonacci word contains $2n+1$ factors of length n and is balanced.

Can we generalize this to any 3D directions ?

$\sigma = 1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 1$

$\sigma^{\infty}(1) = 1213121121312121 3121 \cdots$



The proposed approach

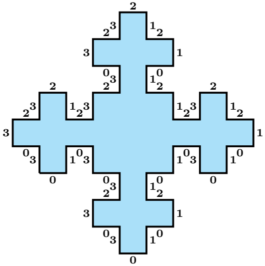
Generalize the Tribonacci word to any $v \in \mathbb{R}^3_+$ using multidimensional continued fractions algorithms and S-adic sequences.

There are other motivations for this approach :

- Study S-adic sequences, in the perspective of the S-adic conjecture concerning factor complexity.
- Study multidimensional continued fractions algorithms from substitutions and combinatorics on words point of view.
- Extend Pisot conjecture and Rauzy fractals (usually defined for fixed point of morphisms) to $\mathcal{S}$-adic sequences.

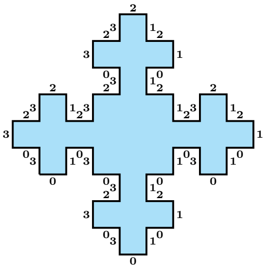
Plan

Factor Complexity

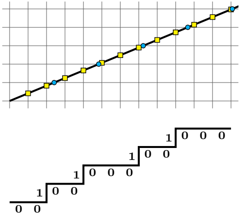


n	$p(n)$	$L(n)$
0	1	$\{\varepsilon\}$
1	4	$\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}\}$
2	8	$\{\mathbf{01}, \mathbf{03}, \mathbf{10}, \mathbf{12}, \mathbf{21}, \mathbf{23}, \mathbf{30}, \mathbf{32}\}$
3	16	$\{\mathbf{030}, \mathbf{212}, \mathbf{301}, \mathbf{032}, \mathbf{103}, \cdots\}$
4	24	$\{\mathbf{1030}, \mathbf{2303}, \mathbf{0121}, \mathbf{1232}, \cdots\}$
5	32	$\{\mathbf{21010}, \mathbf{10303}, \mathbf{32123}, \cdots\}$

Factor Complexity



n	$p(n)$	$L(n)$
0	1	$\{\varepsilon\}$
1	4	$\{\mathbf{0}, \mathbf{1}, \mathbf{2}, \mathbf{3}\}$
2	8	$\{\mathbf{01}, \mathbf{03}, \mathbf{10}, \mathbf{12}, \mathbf{21}, \mathbf{23}, \mathbf{30}, \mathbf{32}\}$
3	16	$\{\mathbf{030}, \mathbf{212}, \mathbf{301}, \mathbf{032}, \mathbf{103}, \cdots\}$
4	24	$\{\mathbf{1030}, \mathbf{2303}, \mathbf{0121}, \mathbf{1232}, \cdots\}$
5	32	$\{\mathbf{21010}, \mathbf{10303}, \mathbf{32123}, \cdots\}$



$c_{1/\sqrt{2}} = \mathbf{0010010001001000100}\cdots$

n	$p(n)$	$L(n)$
0	1	$\{\varepsilon\}$
1	2	$\{\mathbf{0}, \mathbf{1}\}$
2	3	$\{\mathbf{01}, \mathbf{10}, \mathbf{00}\}$
3	4	$\{\mathbf{001}, \mathbf{000}, \mathbf{010}, \mathbf{100}\}$
4	5	$\{\mathbf{1001}, \mathbf{1000}, \mathbf{0100}, \mathbf{0001}, \mathbf{0010}\}$
5	6	$\{\mathbf{00010}, \mathbf{01001}, \mathbf{10010}, \mathbf{00100}, \cdots\}$

Being finitely balanced

Definition

An infinite word $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ is said to be finitely balanced or balanced if there exists a constant $C \in \mathbb{N}$ such that

for all pairs of factors u, v of $\mathbf{w}$ of the same length,

and for all letters $j \in \mathcal{A}$:

$$-C \leq |u|_j - |v|_j \leq C.$$

S-adic construction of a word

Definition

Let $(\mathcal{A}_i)_{i\geq 0}$ be a sequence of alphabets, $(\sigma_i)_{i\geq 0}$ a sequence of morphisms such that $\sigma_i : \mathcal{A}_{i+1}^* \to \mathcal{A}_i^*$, and $(a_i)_{i\geq 0}$ a sequence of letters with $a_i \in \mathcal{A}_i$. Assume that the limit

$$u = \lim_{i\to\infty} (\sigma_0 \circ \sigma_1 \circ \cdots \circ \sigma_{i-1})(a_i)$$

exists and is an infinite word $u \in \mathcal{A}_0^{\mathbb{N}}$. Then $(\sigma_i, a_i)_{i\geq 0}$ is called an s-adic construction of u.

S-adic construction of a word in Sage

Available in Sage since December 2009 :

```
  sage: tm = WordMorphism('a->ab,b->ba')
  sage: fib = WordMorphism('a->ab,b->a')
  sage: w = words.s_adic([fib,tm,tm,fib,tm,fib]*3,'a'); w
word: abaaabaababaabaaababaaababaaabaababaabaa...
```

The following examples illustrates an S-adic word defined over an infinite set S of morphisms x_h :

```
  sage: from itertools import repeat
  sage: x = lambda h:WordMorphism({1:[2],2:[3]+[1]*(h+1),3:[3]+[1]*h})
  sage: for h in [0,1,2,3]: print "x_%s : %s" % (h, x(h))
  x_0 : 1->2, 2->31, 3->3
  x_1 : 1->2, 2->311, 3->31
  x_2 : 1->2, 2->3111, 3->311
  x_3 : 1->2, 2->31111, 3->3111
  sage: w = Word(lambda n : valuation(n+1, 2) ); w
  word: 0102010301020104010201030102010501020103...
  sage: s = words.s_adic(w, repeat(3), x); s
  word: 3232232232322322322323223223232232232232...
```

S-adic Conjecture

Conjecture

There exists a condition C such that a sequence has sub-linear complexity if and only if it is an S-adic sequence satisfying Condition C for some finite set S of morphisms.

- Fabien Durand, Julien Leroy, and Gwénaël Richomme. Towards a statement of the s-adic conjecture through examples. *arXiv :1208.6376*, August 2012.
- Sébastien Ferenczi. Rank and symbolic complexity. *Ergodic Theory Dynam. Systems*, 16(4) :663–682, 1996.

Arnoux-Rauzy words : $\mathcal{S}$-adic representation

Let $\mathcal{A} = \{1, 2, 3\}$, $\mathcal{S} = \{\sigma_i : i \in \mathcal{A}\}$ where

$$\sigma_i : i \mapsto i, j \mapsto ji \text{ for } j \notin \mathcal{A}.$$

An Arnoux-Rauzy word can be represented as an $\mathcal{S}$-adic sequence

$$\mathbf{w} = \lim_{n\to\infty} (\sigma_{i_0} \circ \sigma_{i_1} \circ \cdots \circ \sigma_{i_n})(1).$$

where every letter in $\mathcal{A}$ occurs infinitely often in $(i_m)_{m\geq 0}$. The sequence $(i_m)_{m\geq 0} \in \mathcal{A}^{\mathbb{N}}$ is called the $\mathcal{S}$-directive sequence.

Pierre Arnoux and Gérard Rauzy. Représentation géométrique de suites de complexité 2n+1. *Bull. Soc. Math. France*, 119(2) :199–215, 1991.

J. Cassaigne, S. Ferenczi, and L. Q. Zamboni. Imbalances in Arnoux-Rauzy sequences. *Ann. Inst. Fourier (Grenoble)*, 50(4) :1265–1276, 2000.

Tribonacci Example from Rauzy (1982)

The Tribonacci word is the fixed point of

$$1 \mapsto 12, 2 \mapsto 13, 3 \mapsto 1$$

and is also an Arnoux-Rauzy word with $\mathcal{S}$-directive sequence $123123 \cdots$.

It has complexity $2n+1$ and is 2-balanced.

Gwénaël Richomme, Kalle Saari, and Luca Q. Zamboni. Balance and abelian complexity of the Tribonacci word. *Advances in Applied Mathematics*, 45(2) :212–231, August 2010.

Balance properties of Arnoux-Rauzy words

Let **w** be an Arnoux-Rauzy word with $\mathcal{S}$-directive sequence $(i_m)_{m\geq 0}$.

Theorem (Berthé, Cassaigne, Steiner, 2012)

If the weak partial quotients are bounded by h, i.e., if we do not have $i_m = i_{m+1} = \cdots = i_{m+h}$ *for any* $m \geq 0$*, then* **w** *is* $(2h+1)$*-balanced.*

Theorem (Berthé, Cassaigne, Steiner, 2012)

Let X *be the set of words* $\{1121, 1122, 12121, 12122\}$ *together with all the words that are obtained from one of these four words by a permutation of the letters* 1, 2, *and* 3. *If* $(i_m)_{m\geq 0}$ *contains no factor in* X, *then* **w** *is 2-balanced.*

Valérie Berthé, Julien Cassaigne, and Wolfgang Steiner. Balance properties of arnoux-rauzy words. *arXiv :1212.5106*, December 2012.

Plan

2D : Euclid algorithm on $(11, 4)$

$$\begin{array}{|ccc|}\hline 11 & = & 2 \cdot 4 + 3 \\ 4 & = & 1 \cdot 3 + 1 \\ 3 & = & 3 \cdot 1 + 0 \\ \hline \end{array}$$

$$\boxed{\frac{4}{11} = 0 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3}}}}$$

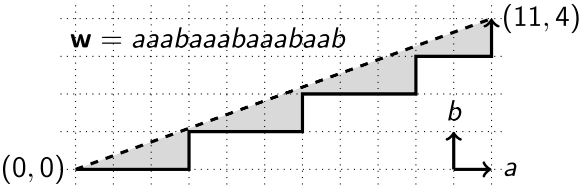
$$(11,4) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2} (3,4) \xleftarrow{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}} (3,1) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^3} (0,1)$$

$$\mathbf{w} = \mathbf{w}_0 \xleftarrow[\begin{array}{ccc} a & \mapsto & a \\ b & \mapsto & aab \end{array}]{} \mathbf{w}_1 \xleftarrow[\begin{array}{ccc} a & \mapsto & ab \\ b & \mapsto & b \end{array}]{} \mathbf{w}_2 \xleftarrow[\begin{array}{ccc} a & \mapsto & a \\ b & \mapsto & aaab \end{array}]{} \mathbf{w}_3 = b$$

2D : Euclid algorithm on $(11, 4)$

$$\begin{aligned} 11 &= 2 \cdot 4 + 3 \\ 4 &= 1 \cdot 3 + 1 \\ 3 &= 3 \cdot 1 + 0 \end{aligned}$$

$$\frac{4}{11} = 0 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3}}}$$



$$(11,4) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2} (3,4) \xleftarrow{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}} (3,1) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^3} (0,1)$$

$$\mathbf{w} = \mathbf{w}_0 \xleftarrow{\begin{array}{l} a \mapsto a \\ b \mapsto aab \end{array}} \mathbf{w}_1 \xleftarrow{\begin{array}{l} a \mapsto ab \\ b \mapsto b \end{array}} \mathbf{w}_2 \xleftarrow{\begin{array}{l} a \mapsto a \\ b \mapsto aaab \end{array}} \mathbf{w}_3 = b$$

3D : Brun's algorithm on $(7,4,6)$

$$
(7,4,6) \xleftarrow[\begin{smallmatrix}1\mapsto 1\\ 2\mapsto 2\\ 3\mapsto 13\end{smallmatrix}]{\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}} (1,4,6) \xleftarrow[\begin{smallmatrix}1\mapsto 1\\ 2\mapsto 23\\ 3\mapsto 3\end{smallmatrix}]{\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}} (1,4,2) \xleftarrow[\begin{smallmatrix}1\mapsto 1\\ 2\mapsto 2\\ 3\mapsto 223\end{smallmatrix}]{\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}} (1,0,2) \xleftarrow[\begin{smallmatrix}1\mapsto 133\\ 2\mapsto 2\\ 3\mapsto 3\end{smallmatrix}]{\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}} (1,0,0)
$$

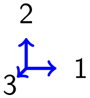
$$
\mathbf{w}_0 \longleftarrow \mathbf{w}_1 \longleftarrow \mathbf{w}_2 \longleftarrow \mathbf{w}_3 \longleftarrow \mathbf{w}_4
$$

Its (Hausdorff) distance to the euclidean line is 1.3680.

$$(7,4,6) \xleftarrow{\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}} (1,4,6) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}} (1,4,2) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}} (1,0,2) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}} (1,0,0)$$

$$\mathbf{w}_0 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&2\\3&\mapsto&13\end{array}} \mathbf{w}_1 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&23\\3&\mapsto&3\end{array}} \mathbf{w}_2 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&2\\3&\mapsto&223\end{array}} \mathbf{w}_3 \xleftarrow{\begin{array}{ccc}1&\mapsto&133\\2&\mapsto&2\\3&\mapsto&3\end{array}} \mathbf{w}_4$$

$$\mathbf{w} = \mathbf{w}_0 = 1213213132132 1313$$



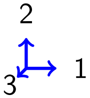
Its (Hausdorff) distance to the euclidean line is 1.3680.

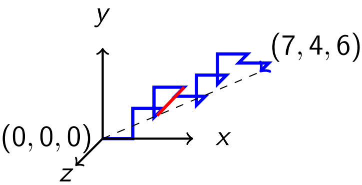
3D : Brun's algorithm on $(7,4,6)$

$$(7,4,6) \xleftarrow{\begin{pmatrix}1&0&1\\0&1&0\\0&0&1\end{pmatrix}} (1,4,6) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&0\\0&1&1\end{pmatrix}} (1,4,2) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&2\\0&0&1\end{pmatrix}} (1,0,2) \xleftarrow{\begin{pmatrix}1&0&0\\0&1&0\\2&0&1\end{pmatrix}} (1,0,0)$$

$$\mathbf{w}_0 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&2\\3&\mapsto&13\end{array}} \mathbf{w}_1 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&23\\3&\mapsto&3\end{array}} \mathbf{w}_2 \xleftarrow{\begin{array}{ccc}1&\mapsto&1\\2&\mapsto&2\\3&\mapsto&223\end{array}} \mathbf{w}_3 \xleftarrow{\begin{array}{ccc}1&\mapsto&133\\2&\mapsto&2\\3&\mapsto&3\end{array}} \mathbf{w}_4$$

$$\mathbf{w} = \mathbf{w}_0 = 1213213132132131 3$$





Its (Hausdorff) distance to the euclidean line is 1.3680.

3D Continued fraction algorithms

Brun's Algorithm : Subtract the second largest to the largest.

$(7,4,6) \to (1,4,6) \to (1,4,2) \to (1,2,2) \to (1,0,2) \to (1,0,1) \to (0,0,1).$

Selmer's Algorithm : Subtract the smallest to the largest.

$(7,4,6) \to (3,4,6) \to (3,4,3) \to (3,1,3) \to (2,1,3) \to (2,1,2) \to (1,1,2)$
$\to (1,1,1) \to (0,1,1) \to (0,0,1)$

Poincaré's Algorithm : Subtract the smallest to the mid and the mid to the largest.

$$(7,4,6) \to (1,4,2) \to (1,2,1) \to (1,1,0) \to (1,0,0)$$

Arnoux-Rauzy's Algorithm : Subtract the sum of the two smallest to the largest (not always possible).

$$(7,4,6) \to \text{Impossible}$$

Fully subtractive's Algorithm : Subtract the smallest to the other two.

$$(7,4,6) \to (3,4,2) \to (1,2,2) \to (1,1,1) \to (1,0,0)$$

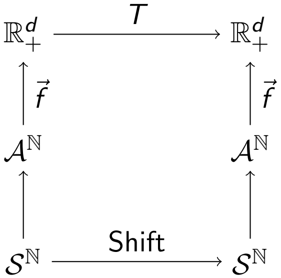
A multidimensional continued fraction algorithm can be given as

$$T : \mathbb{R}_+^d \to \mathbb{R}_+^d : x \mapsto A_x^{-1} x$$

where $A_x \in SL_d(\mathbb{N})$ is a matrix depending on x.
Let $\mathcal{S}$ be a set of substitutions

$$\mathcal{S} = \{s_x : \text{incidence matrix of } s_x = A_x\}.$$

Is there a language $\mathcal{L} \subset \mathcal{S}^{\mathbb{N}}$ such that the following diagram commutes ?



$$\text{Shift} : (\sigma_i)_{i \geq 0} \mapsto (\sigma_i)_{i \geq 1}$$

Plan

Discrepancy

Definition

The discrepancy of an infinite word $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ *is defined as*

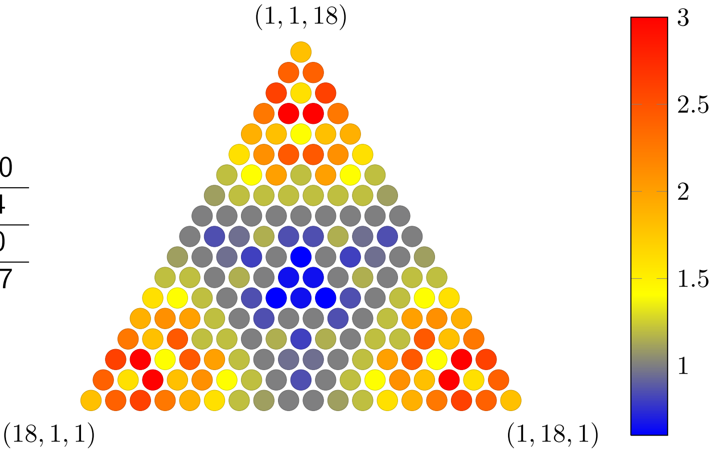
$$\limsup_{i \in \mathcal{A}, p \text{ prefix of } \mathbf{w}} \left| f_i \cdot |p| - |p|_i \right|.$$

where f_i *is the frequency of the letter* $i \in \mathcal{A}$, *if it exists :*

$$f_i = \lim_{p \text{ prefix of } \mathbf{w}} \frac{|p|_i}{|p|}$$

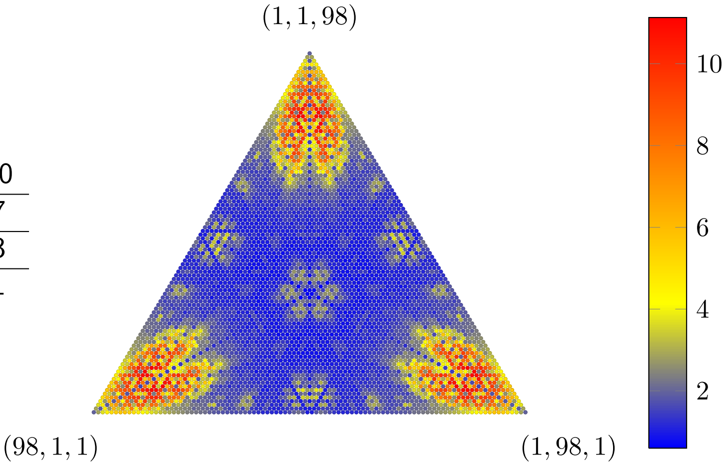
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 20$ pour l'algorithme **Poincaré**.

min	0.6000
moy.	1.484
max	3.000
E.T.	0.6137



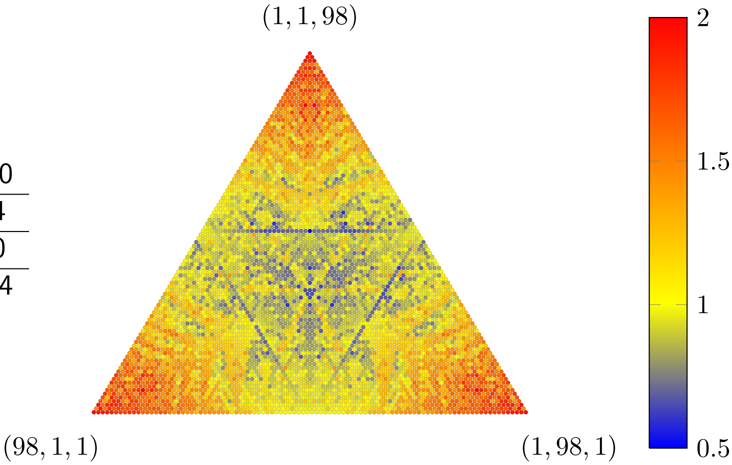
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Poincaré**.

min	0.6000
moy.	2.527
max	11.13
E.T.	2.261



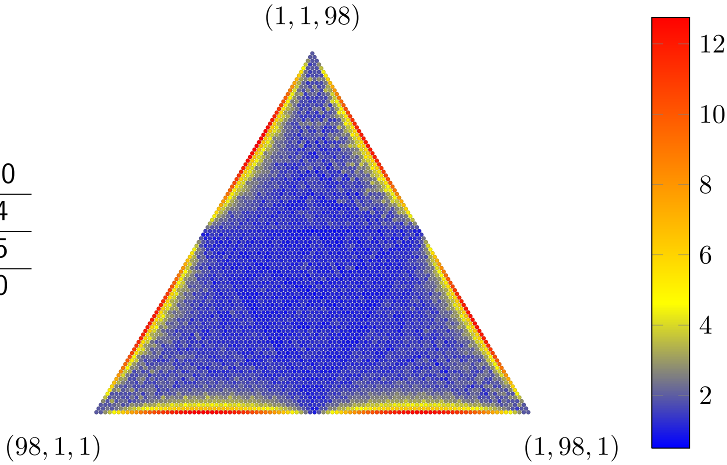
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Brun**.

min	0.5000
moy.	1.114
max	2.000
E.T.	0.2664



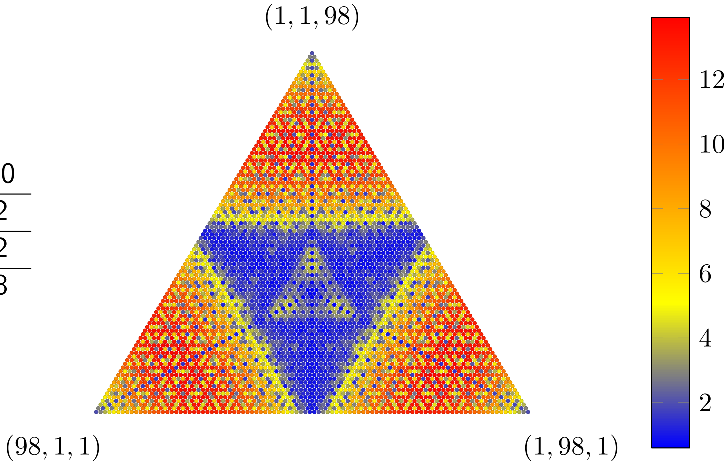
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Selmer**.

min	0.5000
moy.	2.184
max	12.75
E.T.	2.070



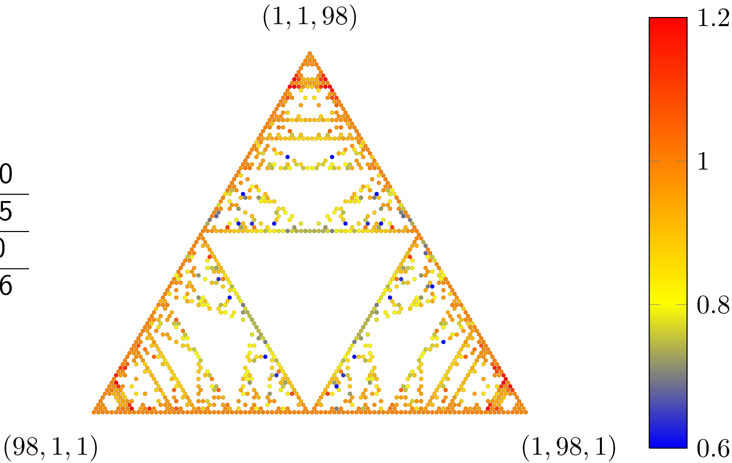
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Fully subtractive**.

min	0.6000
moy.	5.982
max	13.92
E.T.	4.388



Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Arnoux-Rauzy**.

min	0.6000
moy.	0.9055
max	1.200
E.T.	0.1006



3D Continued fraction algorithms : fusions

Arnoux-Rauzy and Selmer Do Arnoux-Rauzy if possible, otherwise Selmer.

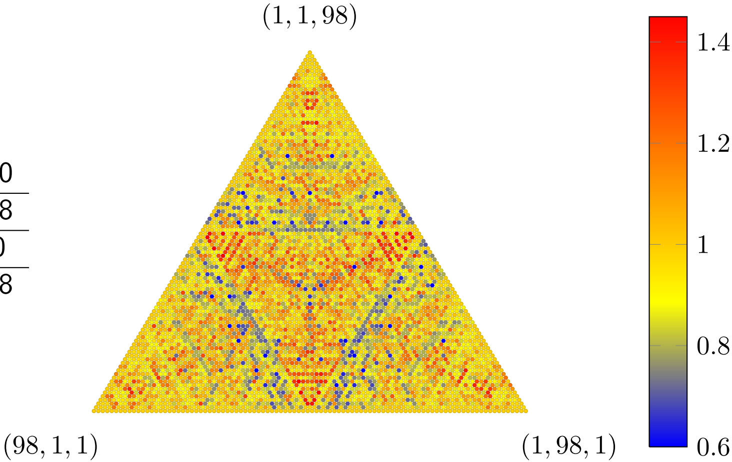
Arnoux-Rauzy and Fully Do Arnoux-Rauzy if possible, otherwise Fully subtractive.

Arnoux-Rauzy and Brun Do Arnoux-Rauzy if possible, otherwise Brun.

Arnoux-Rauzy and Poincaré Do Arnoux-Rauzy if possible, otherwise Poincaré.

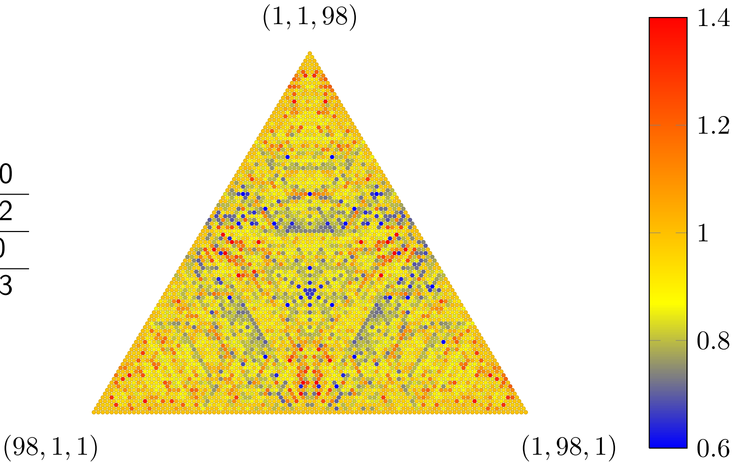
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Arnoux-Rauzy-Selmer**.

min	0.6000
moy.	0.9678
max	1.450
E.T.	0.1438



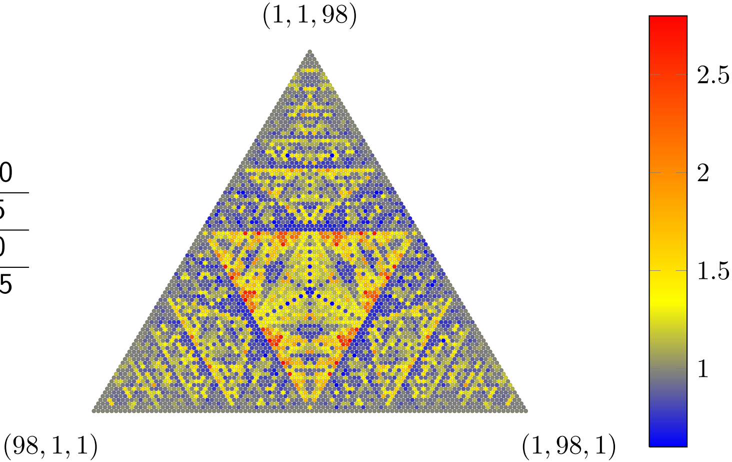
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Arnoux-Rauzy-Brun**.

min	0.6000
moy.	0.9132
max	1.400
E.T.	0.1143



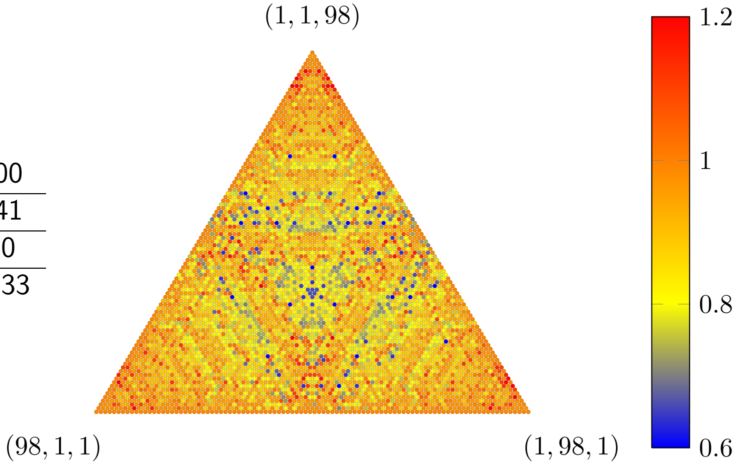
Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Arnoux-Rauzy-Fully subtractive**.

min	0.6000
moy.	1.095
max	2.800
E.T.	0.3105



Discrépance pour les triplets d'entiers strictement positifs (a_1, a_2, a_3) tels que $a_1 + a_2 + a_3 = N$ et $N = 100$ pour l'algorithme **Arnoux-Rauzy-Poincaré**.

min	0.6000
moy.	0.8941
max	1.200
E.T.	0.09733



Plan

Arnoux-Rauzy and Poincaré substitutions

For all $\{i, j, k\} = \{1, 2, 3\}$, we consider

$$\begin{array}{ll} \pi_{jk} : i \mapsto ijk, j \mapsto jk, k \mapsto k & \text{(Poincaré substitutions)} \\ \alpha_k : i \mapsto ik, j \mapsto jk, k \mapsto k & \text{(Arnoux-Rauzy substitutions)} \end{array}$$

Namely,

$$\pi_{23} = \begin{cases} 1 \mapsto 123 \\ 2 \mapsto 23 \\ 3 \mapsto 3 \end{cases}, \quad \pi_{13} = \begin{cases} 1 \mapsto 13 \\ 2 \mapsto 213 \\ 3 \mapsto 3 \end{cases}, \quad \alpha_3 = \begin{cases} 1 \mapsto 13 \\ 2 \mapsto 23 \\ 3 \mapsto 3 \end{cases},$$

$$\pi_{12} = \begin{cases} 1 \mapsto 12 \\ 2 \mapsto 2 \\ 3 \mapsto 312 \end{cases}, \quad \pi_{32} = \begin{cases} 1 \mapsto 132 \\ 2 \mapsto 2 \\ 3 \mapsto 32 \end{cases}, \quad \alpha_2 = \begin{cases} 1 \mapsto 12 \\ 2 \mapsto 2 \\ 3 \mapsto 32 \end{cases},$$

$$\pi_{31} = \begin{cases} 1 \mapsto 1 \\ 2 \mapsto 231 \\ 3 \mapsto 31 \end{cases}, \quad \pi_{21} = \begin{cases} 1 \mapsto 1 \\ 2 \mapsto 21 \\ 3 \mapsto 321 \end{cases}, \quad \alpha_1 = \begin{cases} 1 \mapsto 1 \\ 2 \mapsto 21 \\ 3 \mapsto 31 \end{cases}.$$

Complexity examples

In general, it is possible that $p(n+1) - p(n) > 3$ for some values of n. Let

$$s = \pi_{23}\pi_{23}\pi_{13}\pi_{23}\pi_{23}\alpha_1\alpha_3\alpha_2(1).$$

Indeed,

$$p(7) = 23 > 22 = 3 \cdot 7 + 1 \quad \text{and} \quad p(6) - p(5) = 19 - 15 = 4.$$

Computations in the software Sage are shown below.

```
sage: p23 = WordMorphism({3:[3],2:[2,3],1:[1,2,3]})
sage: p13 = WordMorphism({3:[3],1:[1,3],2:[2,1,3]})
sage: a1 = WordMorphism({1:[1],2:[2,1],3:[3,1]})
sage: a2 = WordMorphism({1:[1,2],2:[2],3:[3,2]})
sage: a3 = WordMorphism({1:[1,3],2:[2,3],3:[3]})
sage: s = words.s_adic([p23,p23,p13,p23,p23,a1,a3,a2],[1]); s
word: 1232333233123233332331232333333123233323...
sage: map(s.number_of_factors, range(12))
[1, 3, 5, 8, 11, 15, 19, 23, 27, 31, 35, 38]
sage: [3*n+1 for n in range(12)]
[1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34]
```

Quadratic complexity example

In fact the fixed point of

$$\pi_{13}\pi_{23} : \begin{cases} 1 \mapsto 132133 \\ 2 \mapsto 2133 \\ 3 \mapsto 3 \end{cases}$$

starting with letter 1 has a quadratic factor complexity. This follows from Pansiot theorem.

```
sage: p23 = WordMorphism({3:[3],2:[2,3],1:[1,2,3]})
sage: p13 = WordMorphism({3:[3],1:[1,3],2:[2,1,3]})
sage: m = p13 * p23
sage: print m
WordMorphism: 1->132133, 2->2133, 3->3
sage: w = m.fixed_point(1); w
word: 1321332133132133332133132133313213332...
sage: p = w[:30000]
sage: map(p.number_of_factors, range(12))
[1, 3, 5, 8, 11, 15, 20, 25, 31, 38, 46, 54, 63]
```

Arnoux-Rauzy Poincaré algorithm

Now in order to partition Δ, we consider the following nine matrices :

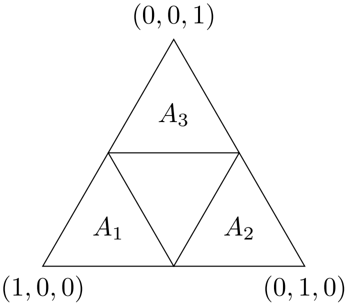
$$P_{23} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad P_{13} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

$$P_{12} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad P_{32} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$P_{31} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad P_{21} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

For each $\{i, j, k\} = \{1, 2, 3\}$, P_{jk} is the incidence matrix of the substitution π_{jk} and A_k is the incidence matrix of α_k.

Arnoux-Rauzy Poincaré algorithm



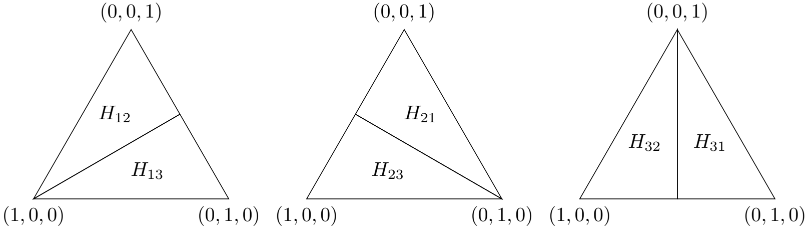
Arnoux-Rauzy Poincaré algorithm

We also define six other matrices :

$$H_{21} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad H_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \quad H_{13} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

$$H_{31} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad H_{32} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad H_{23} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

whose columns describe half triangles in Δ.

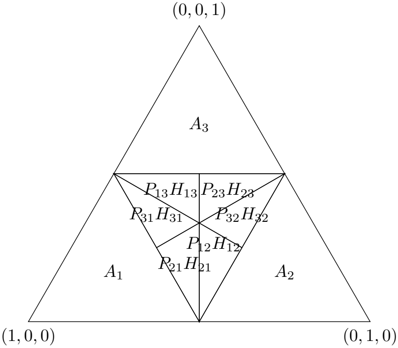


Arnoux-Rauzy Poincaré algorithm

Then, the column vectors of

$$A_1, A_2, A_3, P_{31}H_{31}, P_{13}H_{13}, P_{23}H_{23}, P_{32}H_{32}, P_{12}H_{12} \text{ and } P_{21}H_{21}$$

describe a disjoint partition of Δ.



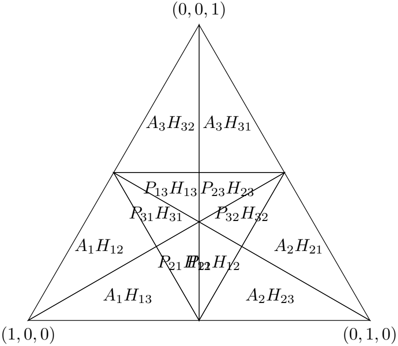
$$
\begin{array}{rcl}
T : \Delta & \to & \mathbb{R}^3_+ \\
\mathbf{x} & \mapsto & \begin{cases} A_k^{-1}\mathbf{x}, & \text{if } \mathbf{x} \in A_k\Delta \\ P_{jk}^{-1}\mathbf{x}, & \text{if } \mathbf{x} \in P_{jk}H_{jk}\Delta. \end{cases}
\end{array}
$$

We defined the **Arnoux-Rauzy-Poincaré** multidimensionnal continued fractions algorithm as the iteration of the function

$$
\begin{array}{rcl}
\overline{T} : \quad \Delta & \to & \Delta \\
\mathbf{x} & \mapsto & \dfrac{T(\mathbf{x})}{||T(\mathbf{x})||}
\end{array}
$$

and $||\mathbf{x}|| = x_1 + x_2 + x_3$.

$$\mathcal{P} = \{A_j H_{jk} : \{i,j,k\} = \{1,2,3\}\} \cup \{P_{jk} H_{jk} : \{i,j,k\} = \{1,2,3\}\}$$



Proposition

The transformation $\overline{T}$ is a Markov transformation for the partition $\mathcal{P}$.

Démonstration.

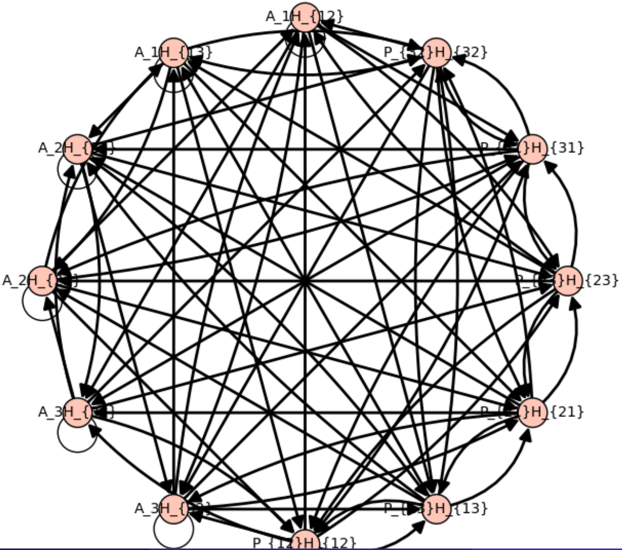
The image of $A_j H_{jk}$ and of $P_{jk} H_{jk}$ under $\overline{T}$ are :

$$\overline{T}(A_j H_{jk}) = \overline{T}(P_{jk} H_{jk}) = H_{jk}.$$

But this half triangle H_{jk} is an union of elements of $\mathcal{P}$. □

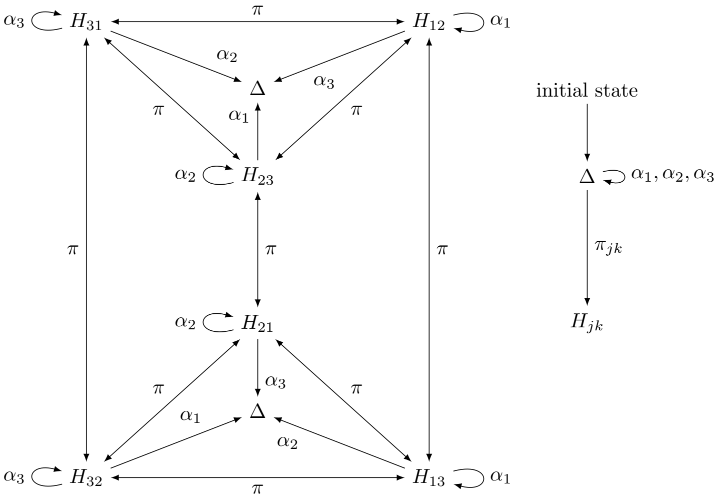
Language of Arnoux-Rauzy Poincaré algorithm

This automaton describe the language $\mathcal{L} \subset \mathcal{S}^{\mathbb{N}}$ of ARP algorithm.



Language of Arnoux-Rauzy Poincaré algorithm

Deterministic and minimized automaton recognizing the language $\mathcal{L} \subset \mathcal{S}^{\mathbb{N}}$ of ARP algorithm :



Result on complexity

Let

$$\mathbf{w} = \lim_{n\to\infty} \sigma_1 \circ \sigma_2 \circ \cdots \circ \sigma_n(a),$$

where $a \in \{1, 2, 3\}$ and $(\sigma_i)_{i\geq 0} \in \mathcal{L} \subset \mathcal{S}^{\mathbb{N}}$.

Theorem (Berthé, L., 2013)

The factor complexity of $\mathbf{w}$ *is such that*

- $p(n+1) - p(n) \in \{2, 3\}$ *and*
- $\limsup_{n\to\infty} \frac{p(n)}{n} < 3$.

Arnoux-Rauzy Poincaré algorithm

Theorem (Boshernitzan, 1984)

A minimal symbolic system (X, S) *such that* $\limsup_{n\to\infty} \frac{p(n)}{n} < 3$ *is uniquely ergodic.*

Theorem (see CANT, Prop. 7.2.10)

A symbolic system (X_x, S) *is uniquely ergodic if, and only if,* x *has uniform frequencies.*

Sébastien Ferenczi and Thierry Monteil. Infinite words with uniform frequencies, and invariant measures. In *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.*, pages 373–409. Cambridge Univ. Press, Cambridge, 2010.

Arnoux-Rauzy Poincaré algorithm

Corollary

The frequencies of factors and letters in $\mathbf{w}$ *exist.*

Proposition (Berthé, L, 2013)

Arnoux-Rauzy Poincaré algorithm provides a function

$$\begin{array}{rccc} g: & \Delta & \rightarrow & \mathcal{A}^{\mathbb{N}} \\ & v & \mapsto & \mathbf{w} \end{array}$$

for $\mathcal{A} = \{1, 2, 3\}$ *such that* $\vec{f} \circ g(v) = v$ *and* $\mathbf{w}$ *has linear factor complexity.*

Idea of the proof on complexity

Let $p(n)$ be the factor complexity function of $\mathbf{w}$. Let $s(n)$ and $b(n)$ be its sequences of finite differences of order 1 and 2 :

$$\begin{aligned} s(n) &= p(n+1) - p(n), \\ b(n) &= s(n+1) - s(n). \end{aligned}$$

Lemma

$p(n+1) - p(n) \in \{2, 3\}$ *if and only if* $\sum_{\ell=0}^{n-1} b(\ell) \in \{0, 1\}$.

Proof. Of course, we have

$$p(n+1) - p(n) = s(n) = s(0) + \sum_{\ell=0}^{n-1} b(\ell)$$

But, $s(0) = p(1) - p(0) = 3 - 1 = 2$. The result follows.

Idea of the proof on complexity

Let $p(n)$ be the factor complexity function of $\mathbf{w}$. Let $s(n)$ and $b(n)$ be its sequences of finite differences of order 1 and 2 :

$$s(n) = p(n+1) - p(n),$$
$$b(n) = s(n+1) - s(n).$$

Lemma

$p(n+1) - p(n) \in \{2, 3\}$ *if and only if* $\sum_{\ell=0}^{n-1} b(\ell) \in \{0, 1\}$.

Lemma

If the sequence of finite differences of order 2 *is such that*

$$(b(n))_n = 0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0, 1, 0, \ldots, 0, -1, \ldots$$

then $\sum_{\ell=0}^{n-1} b(\ell) \in \{0, 1\}$.

Idea of the proof on complexity

Functions s and b are related to special and bispecial factors of $\mathbf{w}$.

Theorem (Cassaigne, Nicolas, 2010)

Let $\mathbf{u} \in A^{\mathbb{N}}$ be a infinite [recurrent] word. Then, for all $n \in \mathbb{N}$:

1. $$s(n) = \sum_{w \in RS_n(\mathbf{u})} (d^+(w) - 1)$$
2. $$s(n) = \sum_{w \in LS_n(\mathbf{u})} (d^-(w) - 1)$$
3. $$b(n) = \sum_{w \in BS_n(\mathbf{u})} m(w)$$

Julien Cassaigne and François Nicolas. Factor complexity. In *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.*, pages 163–247. Cambridge Univ. Press, Cambridge, 2010.

Special and bispecial words

A language is a subset of the free monoid $\mathcal{A}^*$. A language L is factorial if

$$w \in L \quad \text{and} \quad u \text{ factor of } w \implies u \in L$$

The language of factors an infinite word $\mathbf{u} \in \mathcal{A}^{\mathbb{N}}$ is

$$L(\mathbf{u}) = \{w : w \text{ is factor of } \mathbf{u}\}.$$

Right extensions and right valence :

$$E^+(w) = \{x \in \mathcal{A} | wx \in L(\mathbf{u})\} \qquad d^+(w) = \mathrm{Card} E^+(w).$$

Left extensions and left valence :

$$E^-(w) = \{x \in \mathcal{A} | xw \in L(\mathbf{u})\} \qquad d^-(w) = \mathrm{Card} E^-(w).$$

A factor w is

- right special if $d^+(w) \geq 2$,
- left special if $d^-(w) \geq 2$,
- bispecial if it is left and right special.

Special and bispecial words

The set of right special, left special and bispecial factors of length n are identified respectively by $RS_n(\mathbf{u})$, $LS_n(\mathbf{u})$ and $BS_n(\mathbf{u})$.
The extension type of a factor w of $\mathbf{u}$ is the set of pairs (a, b) of $\mathcal{A} \times \mathcal{A}$ such that w can be extended in both directions as awb :

$$E(w) = \{(a, b) \in \mathcal{A} \times \mathcal{A} | awb \in L(\mathbf{u})\}.$$

The bilateral multiplicity of a factor w is

$$m(w) = \mathrm{Card}\, E(w) - d^-(w) - d^+(w) + 1.$$

(Assuming u is recurrent, $m(w) \neq 0$ implies that w is bispecial.)
A bispecial factor is said

- strong if $m(w) > 0$,
- weak if $m(w) < 0$,
- neutral if $m(w) = 0$.

Examples of extension types $E(w)$ of bispecial factors w

$$m(w) = \mathrm{Card}\, E(w) - d^-(w) - d^+(w) + 1.$$

	1	2
1		$\times$
2	$\times$	

$m(w) = -1$
weak

	1	2
1	$\times$	$\times$
2	$\times$	

$m(w) = 0$
neutral and ordinary

	1	2
1	$\times$	$\times$
2	$\times$	$\times$

$m(w) = 1$
strong

Examples of extension types $E(w)$ of bispecial factors w

$$m(w) = \mathrm{Card}\, E(w) - d^-(w) - d^+(w) + 1.$$

	1	2
1		$\times$
2	$\times$	

$m(w) = -1$
weak

	1	2
1	$\times$	$\times$
2	$\times$	

$m(w) = 0$
neutral and ordinary

	1	2
1	$\times$	$\times$
2	$\times$	$\times$

$m(w) = 1$
strong

A bispecial factor w is ordinary if

$$E(w) \subseteq (\{a\} \times A) \cup (A \times \{b\}) \quad \text{for a pair of letters} \quad (a, b) \in E(w).$$

Lemma

If a bispecial factor is ordinary, then it is neutral.

On a binary alphabet, the reciprocal also holds.

Examples of extension types $E(w)$ of bispecial factors w

	1	2	3
1		$\times$	
2		$\times$	
3	$\times$	$\times$	$\times$

$m(w) = 0$
neutral and ordinary

	1	2	3
1		$\times$	
2			
3	$\times$	$\times$	$\times$

$m(w) = 0$
neutral and ordinary

	1	2	3
1		$\times$	
2			$\times$
3	$\times$	$\times$	$\times$

$m(w) = 0$
neutral but not ordinary

	1	2	3
1		$\times$	
2		$\times$	
3	$\times$		$\times$

$m(w) = -1$
weak

	1	2	3
1		$\times$	
2			
3			$\times$

$m(w) = -1$
weak

	1	2	3
1			
2		$\times$	$\times$
3	$\times$	$\times$	$\times$

$m(w) = 1$
strong

Idea of the proof on complexity

Functions s and b are related to special and bispecial factors of $\mathbf{w}$.

Theorem (Cassaigne, Nicolas, 2010)

Let $\mathbf{u} \in A^{\mathbb{N}}$ be a infinite [recurrent] word. Then, for all $n \in \mathbb{N}$:

1. $$s(n) = \sum_{w \in RS_n(\mathbf{u})} (d^+(w) - 1)$$
2. $$s(n) = \sum_{w \in LS_n(\mathbf{u})} (d^-(w) - 1)$$
3. $$b(n) = \sum_{w \in BS_n(\mathbf{u})} m(w)$$

Julien Cassaigne and François Nicolas. Factor complexity. In *Combinatorics, automata and number theory*, volume 135 of *Encyclopedia Math. Appl.*, pages 163–247. Cambridge Univ. Press, Cambridge, 2010.

Idea of the proof on complexity

In our case, we show that

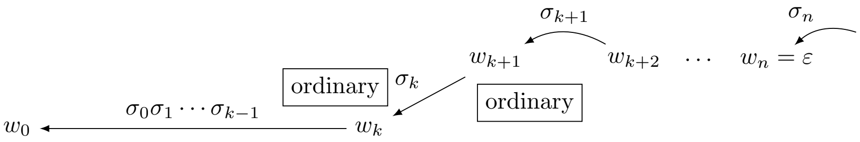
$$(b(n))_n = 0, \ldots, 0, 1, 0, \ldots, 0, -1, 0, \ldots, 0, 1, 0, \ldots, 0, -1, \ldots$$

but we also have

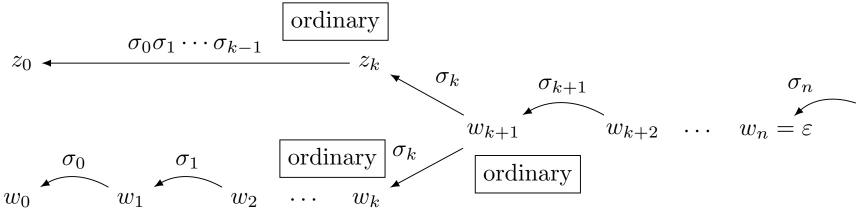
- $|BS_\ell| = 1$,
- $m(w) \in \{-1, 0, 1\}$.

Typical life of ordinary bispecial factors

If $\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{S}_\alpha^*\{\alpha_k\} \cup \mathcal{S}^*\{\pi_{ij},\pi_{kj},\pi_{ji},\pi_{ki}\}\mathcal{S}_\alpha^*\{\alpha_k\}$:

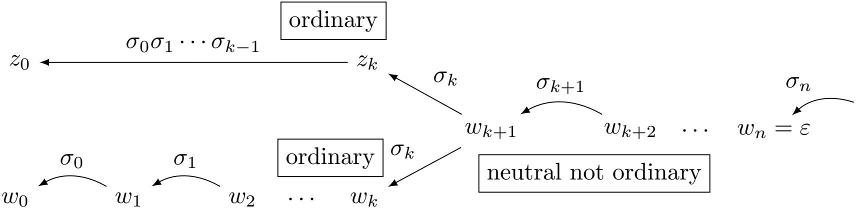


If $\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{S}^*\{\pi_{ik},\pi_{jk}\}\mathcal{S}_\alpha^*\{\alpha_k\}$:

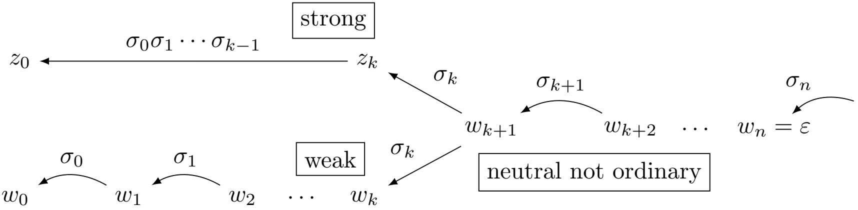


Typical life of neutral not ordinary bispecial factors

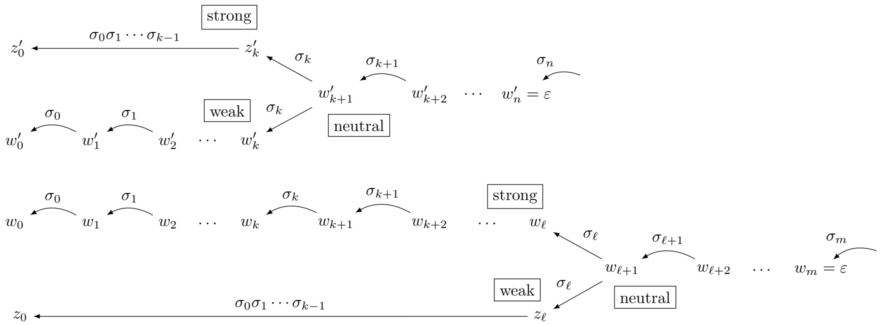
If $\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{S}^*\{\pi_{ij},\pi_{kj},\pi_{ji},\pi_{ki}\}\mathcal{S}_\alpha^*\{\pi_{jk}\}$:



If $\sigma_0\sigma_1\cdots\sigma_n \in \mathcal{S}^*\{\pi_{ik},\pi_{jk}\}\mathcal{S}_\alpha^*\{\pi_{jk}\}$:



Idea of the proof on complexity



Plan

Evaluate the invariant measure

From a starting point

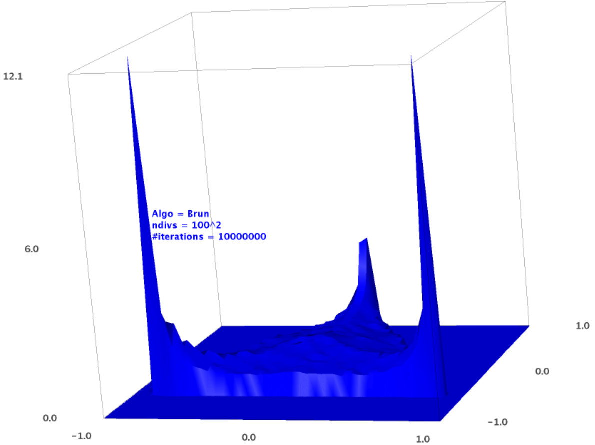
$$p_0 \in \Delta = \{(x_1, x_2, x_3) \in \mathbb{R}^3_+ : x_1 + x_2 + x_3 = 1\},$$

compute the sequence $(p_n)_{n\in\mathbb{N}}$ such that

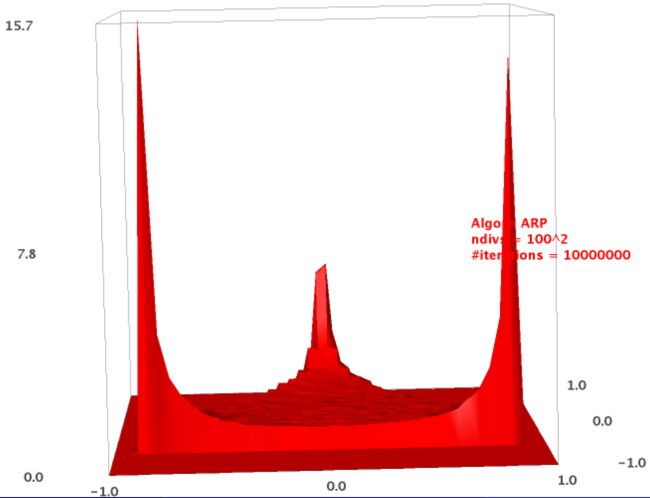
$$p_{n+1} = T(p_n)$$

where T is a Multidimensional continued fraction algorithm.

Brun, 10M iterations



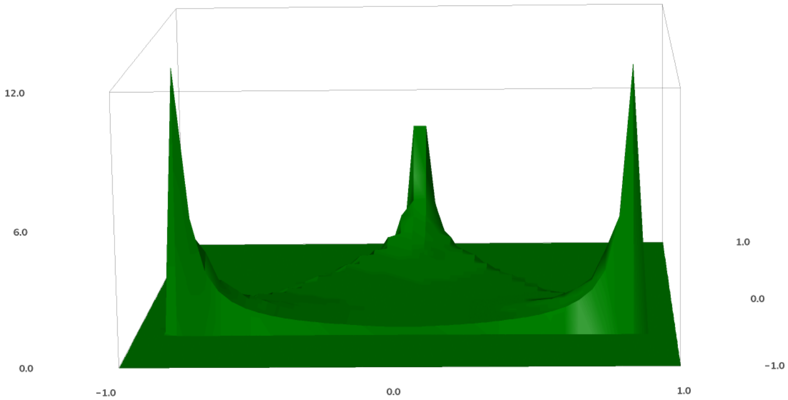
Arnoux-Rauzy-Poincaré, 10M iterations



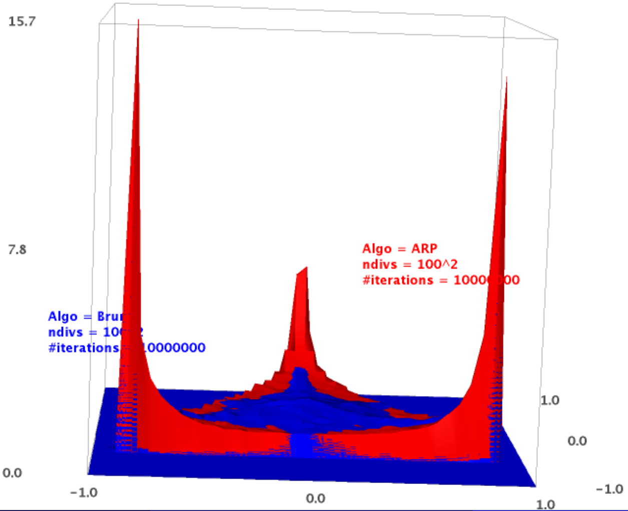
Brun theoretical

On $\{(\alpha, \beta) : 0 < \alpha < \beta < 1\}$, the absolutely continuous invariant measure with respect to Lebesgue is $\frac{12}{\pi}\frac{1}{\beta(1+\alpha)}$. Below, we show it for the simplex

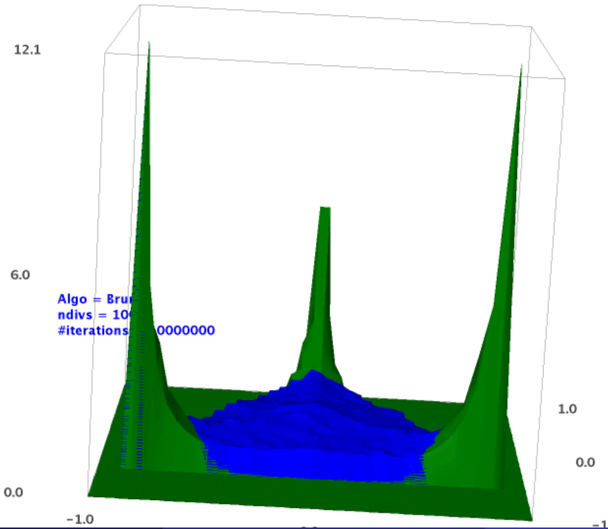
$$\Delta = \{(x_1, x_2, x_3) \in \mathbb{R}_+^3 : x_1 + x_2 + x_3 = 1\}.$$



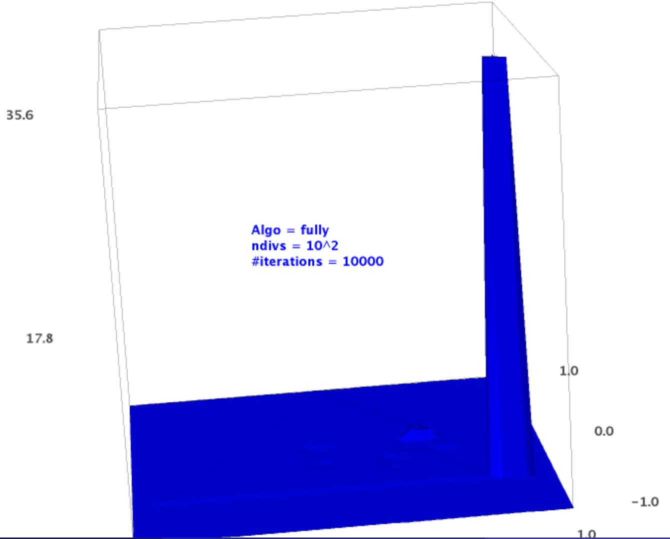
Brun vs ARP, 10M iterations



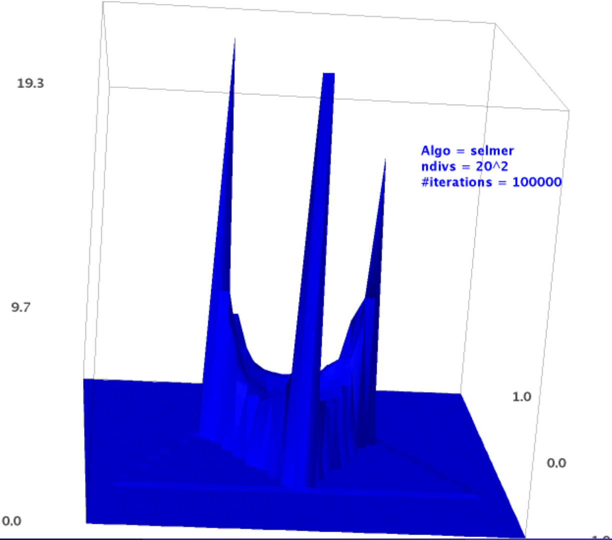
Brun 10M iterations vs Brun theretical



Fully substractive, 10K iterations



Selmer, 100K iterations



Natural extension

From a pair of starting points

$$p_0 \in \Delta, \quad q_0 \in \Delta$$

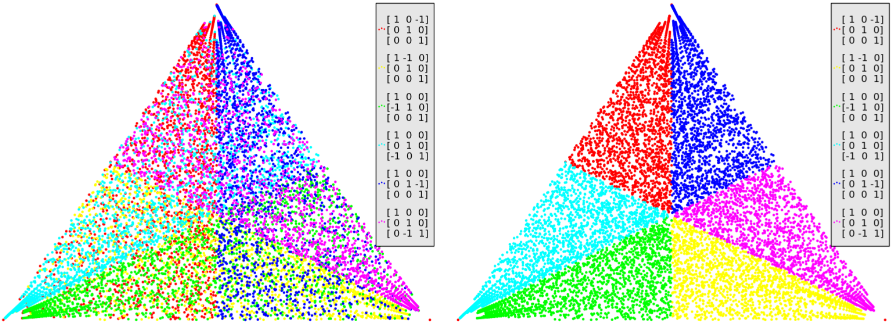
compute the sequences $(p_n)_{n\in\mathbb{N}}$ and $(q_n)_{n\in\mathbb{N}}$ such that

$$\begin{array}{rcccl} p_{n+1} & = & T(p_n) & = & A^{-1} p_n \\ q_{n+1} & = & & = & A^T q_n \end{array}$$

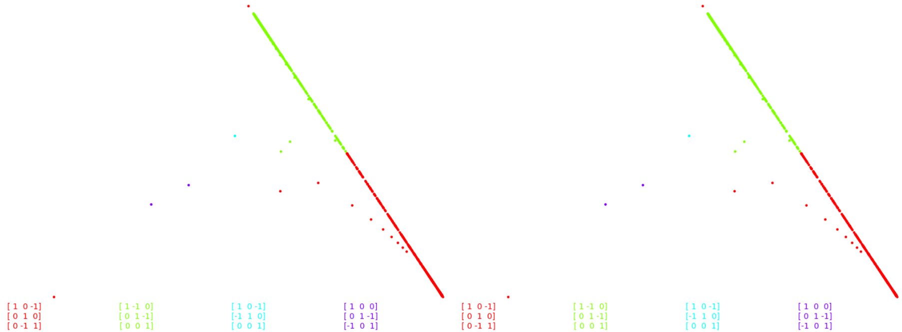
where T is a Multidimensional continued fraction algorithm.

Pierre Arnoux and Arnaldo Nogueira. Mesures de gauss pour des algorithmes de fractions continues multidimensionnelles. *Annales Scientifiques de l'École Normale Supérieure. Quatrième Série*, 26(6) :645–664, 1993.

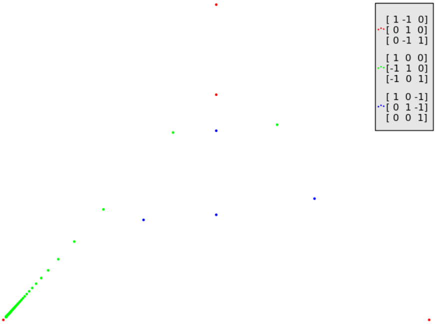
Natural extension of Brun, 10K iterations

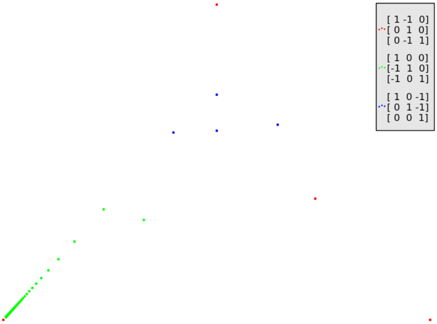


Natural extension of Poincaré, 1K iterations

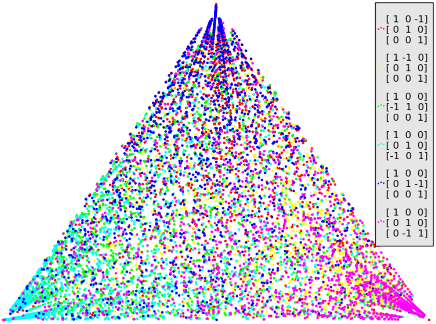


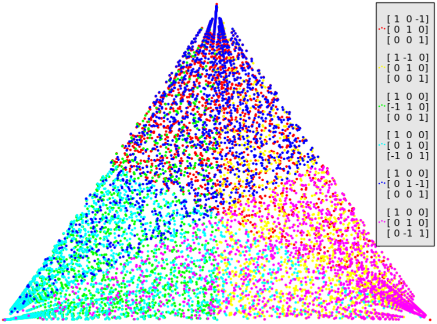
Natural extension of Fully sub., 100 iterations



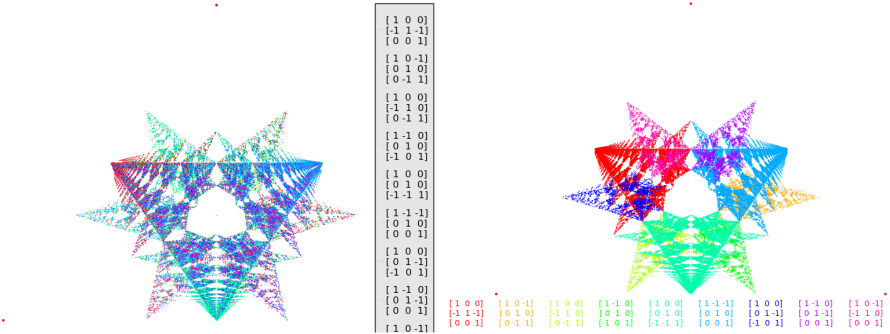


Natural extension of Selmer, 10K iterations

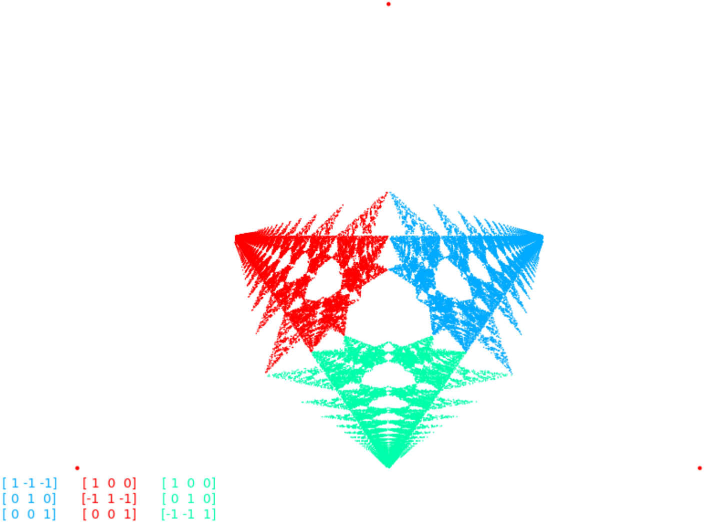




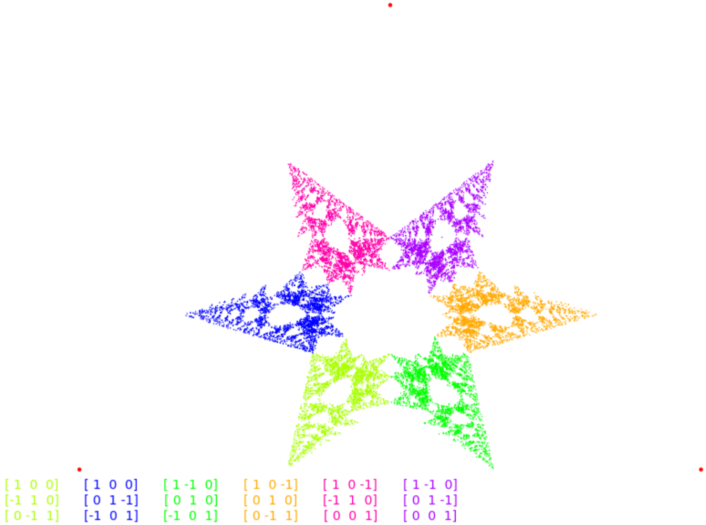
Natural extension of ARP, 100K iterations



Arnoux-Rauzy part in last picture



Poincaré part in last picture



Plan

Questions

- Compute, if it exists, the invariant measure associated to the **Arnoux-Rauzy-Poincaré** algorithm.
- Find under which conditions the **Brun** and **Arnoux-Rauzy-Poincaré** algorithms lead to bounded balance sequences.
- Compute the factor complexity of S-adic sequences obtain using **Brun** algorithm.
- Study ergodic properties of those previous fusion algorithms.