Convergence and Factor Complexity for the Arnoux-Rauzy-Poincaré Algorithm

Sébastien Labbé

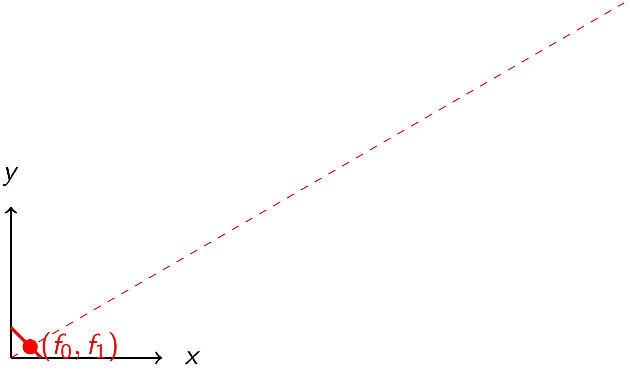
Laboratoire d'Informatique Algorithmique : Fondements et Applications
Université Paris Diderot Paris 7

WORDS 2013
Turku
September 16-20th 2013

Joint work with Valérie Berthé

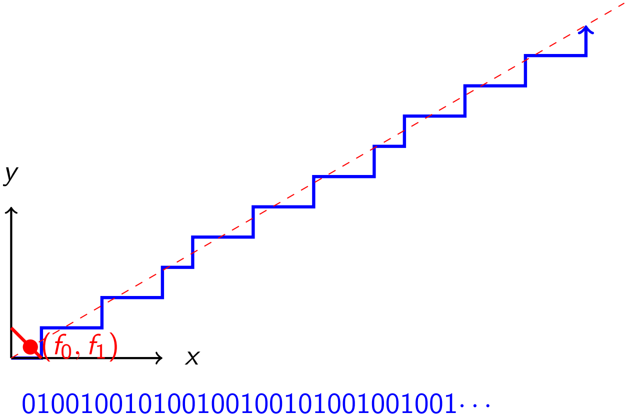
Question (from our talk at WORDS 2011)

Given $(f_0, f_1) \in \mathbb{R}^2$ such that $f_0 + f_1 = 1$, can we construct an infinite sequence $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ on the alphabet $\mathcal{A} = \{0, 1\}$ such that the frequency of digit i is f_i for all $i \in \mathcal{A}$?



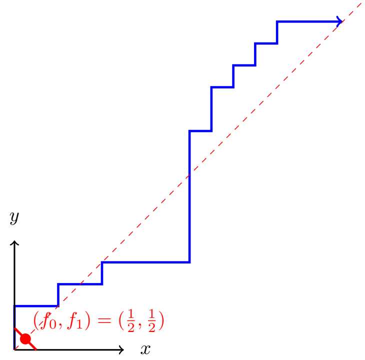
Question (from our talk at WORDS 2011)

Given $(f_0, f_1) \in \mathbb{R}^2$ such that $f_0 + f_1 = 1$, can we construct an infinite sequence $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ on the alphabet $\mathcal{A} = \{0, 1\}$ such that the frequency of digit i is f_i for all $i \in \mathcal{A}$?



One answer is the expansion of π in base 2 :

$$\pi = 11.001001000011111101101010100 0\cdots$$



It is not a good answer because :

- works only for (conjectured) $f_0 = f_1 = \frac{1}{2}$;
- Really different factors appear in the sequence (ex : 0000 and 1111) ;
- All factors (2^k factors of length k) appear in the sequence.

Outline

1. Define the question
2. Why Arnoux-Rauzy-Poincaré MCF algorithm ?
3. Result on factor complexity
4. Conclusion

Plan

1. Define the question
2. Why Arnoux-Rauzy-Poincaré MCF algorithm ?
3. Result on factor complexity
4. Conclusion

Factors, frequencies, vectors

Factor : finite string of consecutive digits. Let

$$w = 010010010100100100101001001001.$$

Then 00100 and 1001 are factors of w. 1001 is a suffix of w.

$|w|$: the length of the factor w. $|w| = 30$.

$|w|_u$: the number of occurences of the factor u in w.

$$|w|_0 = 19, \quad |w|_1 = 11$$

$$|w|_{00} = 8, \quad |w|_{01} = 11, \quad |w|_{10} = 10, \quad |w|_{11} = 0.$$

$\vec{u} = (|u|_0, |u|_1)$: the abelian vector of the factor u.

$$\overrightarrow{00100} = (4, 1), \quad \overrightarrow{1001} = (2, 2).$$

Factor complexity

Let $w \in \mathcal{A}^{\mathbb{N}}$. The factor complexity is a function $p_w(n) : \mathbb{N} \rightarrow \mathbb{N}$ counting the number of factors of length n, noted $\mathrm{L}_w(n)$, in the sequence w.

$$w = 000100\boxed{0100}0100100010001000100100010001001$$

$$\mathrm{L}_w(4) = \{0001, 0010, 0100, 1000, 1001\}$$

n	$p_w(n)$
0	1
1	2
2	3
3	4
4	5

Upper bound : $p_w(n) \leq |\mathcal{A}|^n$.

Balanced sequences

Definition

An infinite word $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ is said to be finitely balanced or C-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,

$$||\vec{u} - \vec{v}||_\infty \leq C.$$

Base 2 development of π is not 1-balanced because

$$||\overrightarrow{0000} - \overrightarrow{1111}||_\infty = ||(4, -4)||_\infty = 4.$$

If π was proven normal, then 0^k and 1^k would also appear for all k, thus it would not be balanced.

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

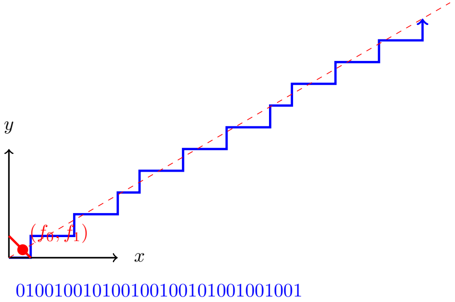
Question (updated)

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 $i \in \mathcal{A}$ is equal to f_i, $\mathbf{w}$ is balanced and has a linear factor complexity?

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

Question (updated)

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



Fact (Answer for $d = 2$, Morse, Hedlund, 1940)

Sturmian words are 1-balanced and satisfy $p(n) = n + 1$.

Definition (Arnoux, Rauzy, 1991)

An infinite word $\mathbf{w} \in \{1, 2, \ldots, d\}^{\mathbb{N}}$ *is an Arnoux-Rauzy word if all its factors occur infinitely often, and if* $p(n) = (d-1)n + 1$ *for all n, with exactly one left special and one right special factor of length n.*

Theorem (Delecroix, Hejda, Steiner, WORDS 2013)

For μ*-almost every* $\mathbf{f}$ *in the Rauzy gasket, the Arnoux-Rauzy word* $w_{AR}(\mathbf{f})$ *is finitely balanced.*

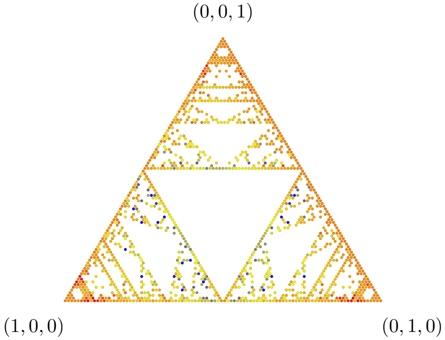
Definition (Arnoux, Rauzy, 1991)

An infinite word $\mathbf{w} \in \{1, 2, \ldots, d\}^{\mathbb{N}}$ *is an Arnoux-Rauzy word if all its factors occur infinitely often, and if* $p(n) = (d-1)n + 1$ *for all* n*, with exactly one left special and one right special factor of length* n*.*

Theorem (Delecroix, Hejda, Steiner, WORDS 2013)

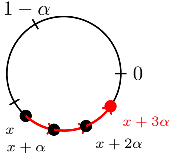
For μ*-almost every* $\mathbf{f}$ *in the Rauzy gasket, the Arnoux-Rauzy word* $w_{AR}(\mathbf{f})$ *is finitely balanced.*

Pierre Arnoux and Štěpán Starosta. The Rauzy Gasket. In Julien Barral and Stéphane Seuret, editors, *Further Developments in Fractals and Related Fields*, Trends in Mathematics, pages 1–23. Birkhäuser Boston, 2013.

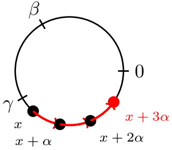


Sturmian words seen as IET

Sturmian word are obtained
from coding of rotations :

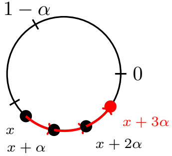


This can be generalized to larger alphabet with coding of rotations on more intervals and more generally to interval exchange transformations (IET) :

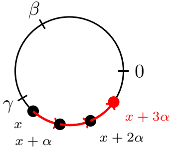


Sturmian words seen as IET

Sturmian word are obtained
from coding of rotations :



This can be generalized to larger alphabet with coding of rotations on more intervals and more generally to interval exchange transformations (IET) :

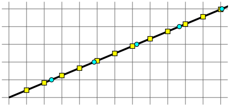


...but such sequences are not balanced.

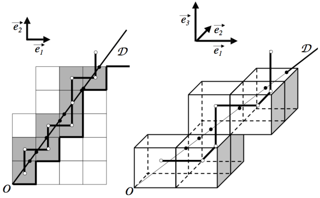
Anton Zorich. Deviation for interval exchange transformations. *Ergodic Theory Dynam. Systems*, 17(6) :1477–1499, 1997.

Sturmian words seen as billiard sequences

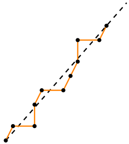
Sturmian word are obtained
from the cutting sequence of a line :



This can be generalized as billiard sequences ($\cong$ E. Andres discrete line) :



Borel (2006)

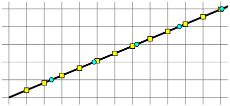


$$-7/2 \leq 2x - 5z < 7/2$$
$$-8/2 \leq 3x - 5y < 8/2$$
$$-5/2 \leq 2y - 3z < 5/2$$

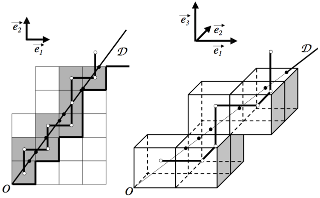
Andres (2003)

Sturmian words seen as billiard sequences

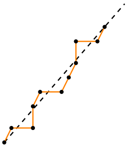
Sturmian word are obtained
from the cutting sequence of a line :



This can be generalized as billiard sequences ($\cong$ E. Andres discrete line) :



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)

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 $p(n) = n^2 + n + 1$.

Plan

Sturmian words seen as iteration of morphisms

From iteration of morphisms $L = \begin{array}{l} 0 \mapsto 0 \\ 1 \mapsto 01 \end{array}$ and $R = \begin{array}{l} 0 \mapsto 10 \\ 1 \mapsto 1 \end{array}$:

$$R^{d_0} L^{d_1} R^{d_2} L^{d_3} R^{d_4} L^{d_5} \cdots (1)$$

where $[d_0; d_1, d_2, d_3, \ldots]$ is the continued fraction expansion of the slope of the word drawn with horizontal and vertical unitary steps.

Sturmian words seen as iteration of morphisms

From iteration of morphisms $L = \begin{array}{l} 0 \mapsto 0 \\ 1 \mapsto 01 \end{array}$ and $R = \begin{array}{l} 0 \mapsto 10 \\ 1 \mapsto 1 \end{array}$:

$$R^{d_0} L^{d_1} R^{d_2} L^{d_3} R^{d_4} L^{d_5} \cdots (1)$$

where $[d_0; d_1, d_2, d_3, \ldots]$ is the continued fraction expansion of the slope of the word drawn with horizontal and vertical unitary steps.

This generalizes to S-adic sequences...

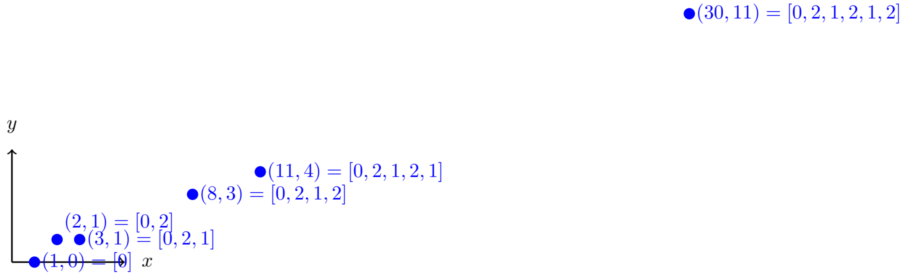
Continued fractions

Let $\alpha = \frac{\sqrt{3}-1}{2} = 0.36602540\cdots$. We have

$$\alpha = 0 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cdots}}}} = [0; 2, 1, 2, 1, 2, 1, \cdots]$$

The convergents p_n/q_n are

$$0, \frac{1}{2}, \frac{1}{3}, \frac{3}{8}, \frac{4}{11}, \frac{11}{30}, \frac{15}{41}, \frac{41}{112}, \frac{56}{153}, \frac{153}{418}, \frac{209}{571}, \frac{571}{1560}, \frac{780}{2131}, \cdots$$



Continued fractions : matrices from convergents

With

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

the convergents can be obtained as

$$\begin{aligned}
\begin{pmatrix} q_0 \\ p_0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} &= R^0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\begin{pmatrix} q_1 \\ p_1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} &= R^0 L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} q_2 \\ p_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix} &= R^0 L^2 R^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\begin{pmatrix} q_3 \\ p_3 \end{pmatrix} = \begin{pmatrix} 8 \\ 3 \end{pmatrix} &= R^0 L^2 R^1 L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} q_4 \\ p_4 \end{pmatrix} = \begin{pmatrix} 11 \\ 4 \end{pmatrix} &= R^0 L^2 R^1 L^2 R^1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\
\begin{pmatrix} q_5 \\ p_5 \end{pmatrix} = \begin{pmatrix} 30 \\ 11 \end{pmatrix} &= R^0 L^2 R^1 L^2 R^1 L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

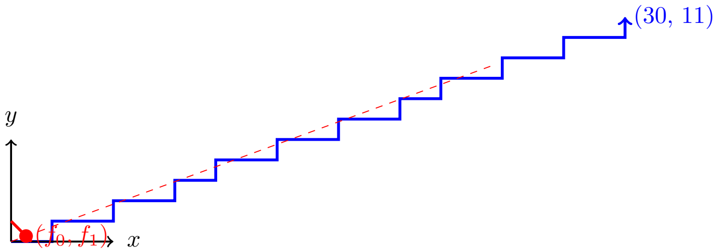
Continued fractions : substitutions from matrices

With

$$L = \begin{array}{l} 0 \mapsto 0 \\ 1 \mapsto 01 \end{array} \qquad \text{and} \qquad R = \begin{array}{l} 0 \mapsto 10 \\ 1 \mapsto 1 \end{array}$$

the convergents can be transformed into finite sequences over $\mathcal{A}$:

$$\begin{array}{lllll}
w_0 & = & R^0(0) & = & 0 \\
w_1 & = & R^0L^2(1) & = & 001 \\
w_2 & = & R^0L^2R^1(0) & = & 0010 \\
w_3 & = & R^0L^2R^1L^2(1) & = & 00100010001 \\
w_4 & = & R^0L^2R^1L^2R^1(0) & = & 001000100010010 \\
w_5 & = & R^0L^2R^1L^2R^1L^2(1) & = & 00100010001001000100010001001000100010001
\end{array}$$



$001000100010010001000100010010001000100 01 \cdots$

Continued fractions : from Euclid Algorithm

With

$$L = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad R = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

the execution of Euclid Algorithm appears as

$$\begin{aligned}
\begin{pmatrix} 30 \\ 11 \end{pmatrix} &= L^2R^1L^2R^1L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 8 \\ 11 \end{pmatrix} &= R^1L^2R^1L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 8 \\ 3 \end{pmatrix} &= L^2R^1L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 2 \\ 3 \end{pmatrix} &= R^1L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\
\begin{pmatrix} 2 \\ 1 \end{pmatrix} &= L^2 \begin{pmatrix} 0 \\ 1 \end{pmatrix}
\end{aligned}$$

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.

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

Experimentations

At WORDS 2011 :

	Min	Mean	Max	Std
Arnoux-Rauzy (AR)	0.6000	0.8922	1.200	0.09953
Fully subtractive	0.5000	6.047	14.21	4.385
Selmer	0.5000	2.151	12.75	2.076
Brun	0.5000	1.100	2.000	0.2625
Poincaré	0.5000	2.476	11.13	2.245
AR-Fully subtractive	0.5000	1.154	4.000	0.3759
AR-Selmer	0.5000	0.9991	1.600	0.1429
AR-Brun	0.5000	0.9169	1.520	0.1170
AR-Poincaré	0.5000	0.9066	1.320	0.1079

TABLE: Statistics for the discrepancy for strictly positive integer vectors (a_1, a_2, a_3) such that $a_1 + a_2 + a_3 = N$ and $N = 100$.

Discrepancy

Definition

The *discrepancy* *of an infinite word* $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ *having frequency* f_i *for each letter* $i \in \mathcal{A}$ *is defined as*

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

Discrepancy

Definition

The discrepancy of an infinite word $\mathbf{w} \in \mathcal{A}^{\mathbb{N}}$ *having frequency* f_i *for each letter* $i \in \mathcal{A}$ *is defined as*

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

$\mathbf{w}$ is balanced $\iff$ $\mathbf{w}$ has finite discrepancy $\iff$ $\mathbf{w}$ stays at bounded distance from the euclidean line of direction (f_0, f_1)

Boris Adamczewski. Balances for fixed points of primitive substitutions. *Theoretical Computer Science*, 307(1) :47 – 75, 2003.

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,

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

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

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

Quadratic complexity for ARP sequences

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,

$$p_s(n) = (1, 3, 5, 8, 11, 15, 19, 23, 27, 31, 35, 38, \cdots)$$

Quadratic complexity for ARP sequences

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,

$$p_s(n) = (1, 3, 5, 8, 11, 15, 19, 23, 27, 31, 35, 38, \cdots)$$

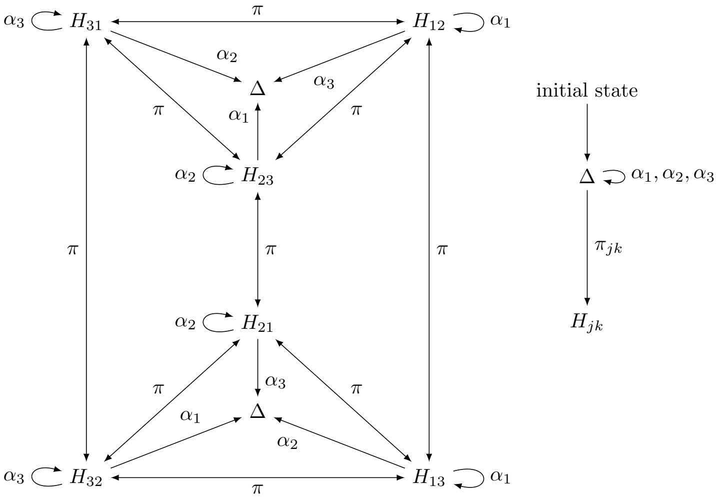
Even worse, the fixed point of

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

starting with letter 1 has a quadratic factor complexity.

Language of Arnoux-Rauzy Poincaré algorithm

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



Let $\mathbf{w} = \lim_{n\to\infty} \sigma_0\sigma_1\cdots\sigma_n(1)$ be an S-adic word generated by the Arnoux-Rauzy-Poincaré algorithm from a totally irrational vector $\mathbf{x} \in \Delta_3$.

Theorem (Factor Complexity)

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

- $p(n) \leq 3n + 1$ *for all* $n \geq 0$ *;*
- $p(n+1) - p(n) \in \{2, 3\}$ *for all* $n \geq 0$ *;*
- $\limsup_{n\to\infty} \frac{p(n)}{n} \leq \frac{5}{2} < 3$ *(not sharp).*

Let $\mathbf{w} = \lim_{n\to\infty} \sigma_0\sigma_1 \cdots \sigma_n(1)$ be an S-adic word generated by the Arnoux-Rauzy-Poincaré algorithm from a totally irrational vector $\mathbf{x} \in \Delta_3$.

Theorem (Factor Complexity)

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

- $p(n) \leq 3n + 1$ *for all* $n \geq 0$ *;*
- $p(n+1) - p(n) \in \{2, 3\}$ *for all* $n \geq 0$ *;*
- $\limsup_{n\to\infty} \frac{p(n)}{n} \leq \frac{5}{2} < 3$ *(not sharp).*

Theorem (Frequencies and Convergence)

The symbolic dynamical system generated by $\mathbf{w}$ *is uniquely ergodic, and the frequencies of letters are proved to exist in* $\mathbf{w}$ *and to be equal to the coordinates of* $\mathbf{x}$*.*

Furthermore, the Arnoux-Rauzy-Poincaré algorithm is a weakly convergent algorithm, that is, for Lebesgue almost every $\mathbf{x} \in \Delta$*, if* $(M_n)_n$ *stands for the sequence of matrices produced by the Arnoux-Rauzy-Poincaré algorithm, then one has* $\cap_n M_0 \cdots M_n(\mathbb{R}^3_+) = \mathbb{R}_+\mathbf{x}$*.*

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 :

$$\begin{aligned} p(n) &= 1,3,5,7,9,11,14,17,20,22,24,26,28, \\ s(n)=p(n+1)-p(n) &= 2,2,2,2,2,3,3,3,2,2,2,2, \\ b(n)=s(n+1)-s(n) &= 0,0,0,0,+1,0,0,-1,0,0,0, \end{aligned}$$

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 :

$$\begin{aligned} p(n) &= 1,3,5,7,9,11,14,17,20,22,24,26,28, \\ s(n) = p(n+1) - p(n) &= 2,2,2,2,2,3,3,3,2,2,2,2, \\ b(n) = s(n+1) - s(n) &= 0,0,0,0,+1,0,0,-1,0,0,0, \end{aligned}$$

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

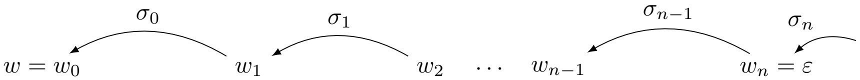
Theorem (Cassaigne, 1997 ; Cassaigne, Nicolas, 2010)

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

$$s(n) = \sum_{w \in RS_n(\mathbf{u})} (d^+(w) - 1) \qquad \text{and} \qquad b(n) = \sum_{w \in BS_n(\mathbf{u})} m(w)$$

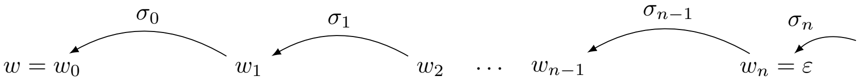
Life of a bispecial factor under ARP substitutions

- Some synchronization lemmas allows to define uniquely the antecedent w_{k+1} of a (bispecial) factor $w_k = s \cdot \sigma_k(w_{k+1}) \cdot p$.
- w_k is an extended image of w_{k+1}



Life of a bispecial factor under ARP substitutions

- Some synchronization lemmas allows to define uniquely the antecedent w_{k+1} of a (bispecial) factor $w_k = s \cdot \sigma_k(w_{k+1}) \cdot p$.
- w_k is an extended image of w_{k+1}



- We always have $|w_k| > |w_{k+1}|$.
- The history of w is $\sigma_0\sigma_1\cdots\sigma_n$.
- The life of w is $(w_k)_{0\leq k\leq n}$.
- w_{k+1} has only one bispecial extended image w_k under an Arnoux-Rauzy substitution.
- w_{k+1} has one or two extended images under Poincaré substitution.

Extension type of bispecial words

The extension type of a factor w of $\mathbf{u}$ is

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

If $1w2$, $2w3$, $3w1$, $3w2$ and $3w3$ are such factors, then the extension type is represented as :

$E(w) =$

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

Extension type of bispecial words

The extension type of a factor w of $\mathbf{u}$ is

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

If $1w2$, $2w3$, $3w1$, $3w2$ and $3w3$ are such factors, then the extension type is represented as :

$$E(w) = \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & & \times & \\ 2 & & & \times \\ 3 & \times & \times & \times \end{array}$$

The bilateral multiplicity of a factor w is

$$\begin{aligned} m(w) &= \mathrm{Card}\, E(w) - d^-(w) - d^+(w) + 1 \\ &= 5 - 3 - 3 + 1 = 0. \end{aligned}$$

A bispecial factor w is said

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

A bispecial factor w is ordinary if

$$\text{row} \cap \text{column} \quad \subseteq \quad E(w) \quad \subseteq \quad \text{row} \cup \text{column}.$$

Ordinary bispecial factor from an ordinary

Let

$$\begin{aligned} \mathcal{S}_\alpha &= \{\alpha_1, \alpha_2, \alpha_3\}, \\ \mathcal{S}_\pi &= \{\pi_{12}, \pi_{13}, \pi_{23}, \pi_{21}, \pi_{31}, \pi_{32}\} \\ \mathcal{S} &= \mathcal{S}_\alpha \cup \mathcal{S}_\pi . \end{aligned}$$

CASE 1 : If $\mathrm{history}(w_0) \in \mathcal{S}^* \, \pi_{jk} \, \mathcal{S}_\alpha^* \{\alpha_k\}$, then $m(w_0) = 0$:

$$w_0 \xleftarrow{\mathcal{S}^*} w_\ell \xleftarrow{\pi_{jk}} w_{\ell+1} \xleftarrow{\mathcal{S}_\alpha^*} w_n = \varepsilon \quad (\alpha_k)$$

	i	j	k
i	×	×	×
j			
k			×

ordinary
$m(w_0) = 0$

	i	j	k
i			
j			×
k	×	×	×

	i	j	k
i			×
j			×
k	×	×	×

	i	j	k
i			×
j			×
k	×	×	×

ordinary
$m(w_n) = 0$

Two ordinary bispecial factors from an ordinary

CASE 2 : If history$(w_0) \in \mathcal{S}^* \pi_{jk} \mathcal{S}_\alpha^* \{\alpha_i, \alpha_j\}$, then $m(w_0) = 0$:

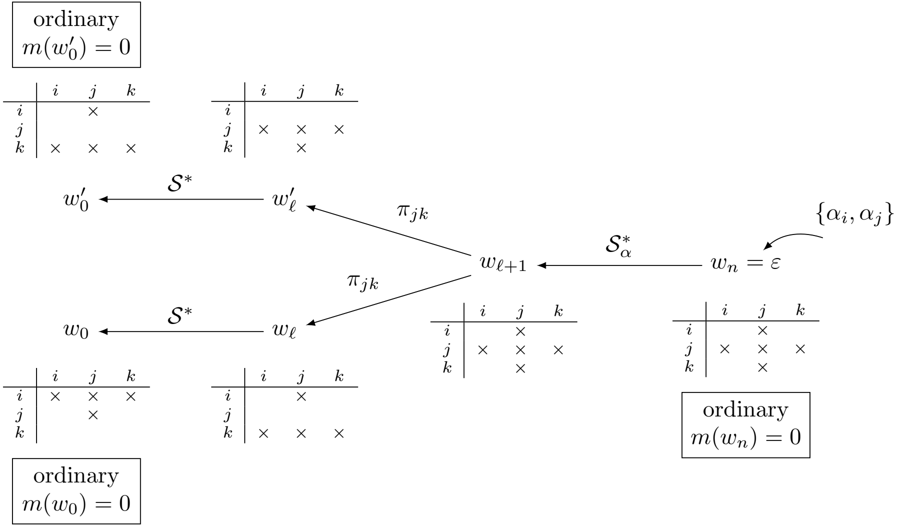


Table for w_0':

	i	j	k
i		×	
j			
k	×	×	×

Table for w_ℓ':

	i	j	k
i			
j	×	×	×
k		×	

Table for $w_{\ell+1}$:

	i	j	k
i		×	
j	×	×	×
k		×	

Table for w_n:

	i	j	k
i		×	
j	×	×	×
k		×	

Table for w_0:

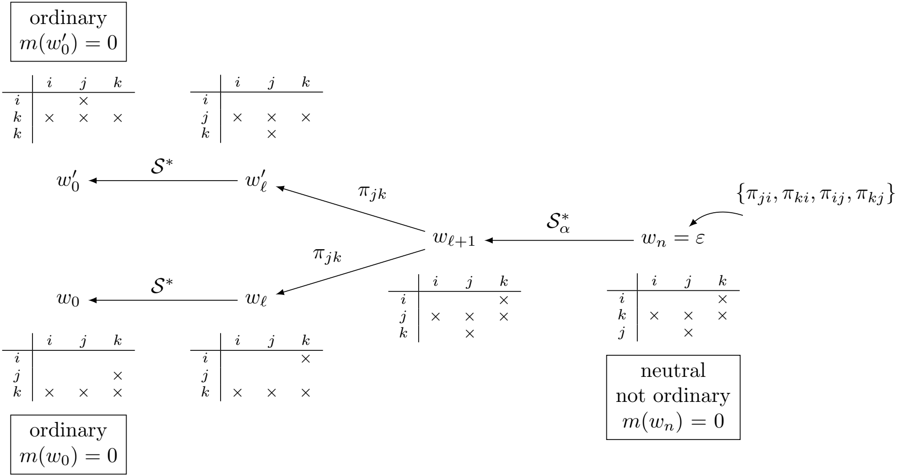
	i	j	k
i	×	×	×
j		×	
k			

Table for w_ℓ:

	i	j	k
i		×	
j			
k	×	×	×

Two ordinary bispecial factors from an non ordinary

CASE 3 : If history$(w_0) \in \mathcal{S}^* \pi_{jk} \mathcal{S}_\alpha^* \{\pi_{ji}, \pi_{ki}, \pi_{ij}, \pi_{kj}\}$, then $m(w_0) = 0$:



	i	j	k
i		×	
k	×	×	×
k			

	i	j	k
i			
j	×	×	×
k		×	

	i	j	k
i			×
j	×	×	×
k		×	

	i	j	k
i			×
k	×	×	×
j		×	

	i	j	k
i			
j			×
k	×	×	×

	i	j	k
i			×
j			
k	×	×	×

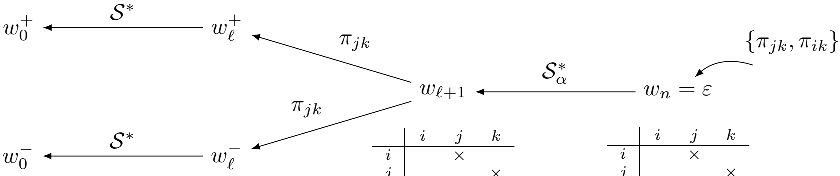
Strong and weak bispecial factors from an non ordinary

CASE 4 : If history$(w_0) \in \mathcal{S}^* \pi_{jk} \mathcal{S}_\alpha^* \{\pi_{jk}, \pi_{ik}\}$, then $m(w_0) = \pm 1$:

strong
$m(w_0^+) = +1$

	i	j	k
i		×	×
j			
k	×	×	×

	i	j	k
i			
j		×	×
k	×	×	×



	i	j	k
i		×	
j			×
k	×	×	×

	i	j	k
i		×	
j			×
k	×	×	×

neutral
not ordinary
$m(w_n) = 0$

	i	j	k
i		×	
j			
k			×

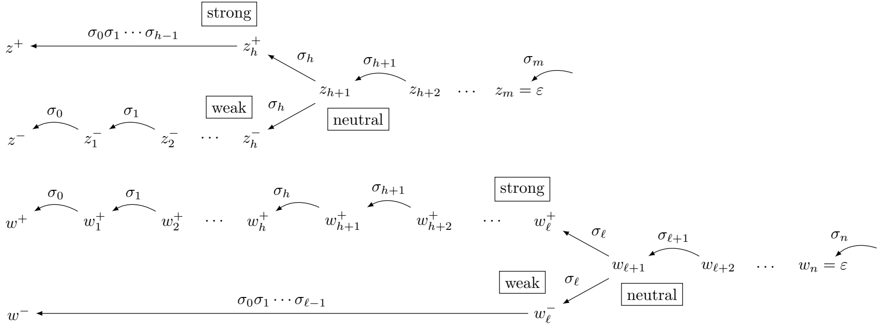
	i	j	k
i		×	
j			
k			×

weak
$m(w_0^-) = -1$

Idea of the proof on complexity

Show that the lifes of two pairs of strong and weak bispecial factors do not intersect, i.e. the following equality is preserved :

$$|z^+| < |z^-| < |w^+| < |w^-|.$$



Plan

Conclusion

	$\forall v \in \mathbb{R}^3_+$	$p(n)$ linear	Balanced
Arnoux Rauzy words	No	Yes	Almost always
Billiard, Andres discrete line	Yes	No	Yes
Coding of rotations and of IET	Yes	Yes	No
Brun S-adic sequences	Yes	$\approx 5n$?	Almost always
ARP S-adic sequences	Yes	Yes : $\frac{5}{2}n$	?
Other S-adic sequences	?	?	?

Conclusion

	$\forall v \in \mathbb{R}^3_+$	$p(n)$ linear	Balanced
Arnoux Rauzy words	No	Yes	Almost always
Billiard, Andres discrete line	Yes	No	Yes
Coding of rotations and of IET	Yes	Yes	No
Brun S-adic sequences	Yes	$\approx 5n$?	Almost always
ARP S-adic sequences	Yes	Yes : $\frac{5}{2}n$	?
Other S-adic sequences	?	?	?

Other motivations :

- 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.

Complexity of Brun sequences

An example using Sage shows that $p(n+1) - p(n) = 5$ is possible for Brun algorithm :

```
sage: Bij = WordMorphism('i->ij,j->j,k->k')
sage: Bjk = WordMorphism('i->i,j->jk,k->k')
sage: Bki = WordMorphism('i->i,j->j,k->ki')
sage: from itertools import cycle, repeat
sage: w = words.s_adic(cycle([Bij, Bjk, Bki]), repeat('i'))
sage: w
word: ijjkjkkijjkkijkijijjkjkkijkijijjkkijijjk...
sage: prefix1000 = w[:1000]
sage: map(prefix1000.number_of_factors, range(20))
[1, 3, 7, 11, 15, 20, 25, 30, 35, 40, 45, 50,
              55, 60, 65, 70, 75, 80, 85, 90]
```

About $\limsup_{n\to\infty} \frac{p(n)}{n}$ for ARP

I think the bound $\frac{5}{2}$ may be improved :

```
sage: p23 = WordMorphism('1->123,2->23,3->3')
sage: a1 = WordMorphism('1->1,2->21,3->31')
sage: m = p23 * a1
sage: x = m.fixed_point('1')
sage: p = x[:100000]
sage: L = [(n,p.number_of_factors(n)/float(n)) for n in range(10, 2000,10)]
sage: point(L, figsize=3)
```