Codes bifixes, combinatoire des mots et systèmes dynamiques symboliques

Bifix Codes, Combinatorics on Words and Symbolic Dynamical Systems

Francesco Dolce

Marne-la-Vallée, 13 septembre 2016

Combinatorics on Words

Combinatorics on Words

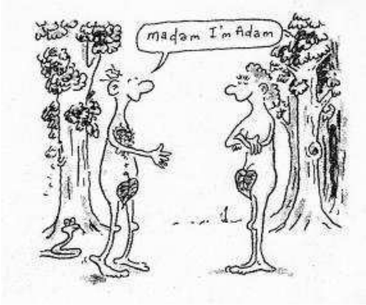
- 011001
- ACGCCTAAT
- ♡♣♡♢♡♠♢
- 1221121221211211221211212211 ⋯

Combinatorics on Words

- 011001
- ACGCCTAAT
- ♡♣♡♢♡♠♢
- 1221121221211211221211212211 $\cdots$

Combinatorics on Words

- 011001
- ACGCCTAAT
- ♡♣♡♢♡♠♢
- 12211212212112112212112122 11 $\cdots$



Combinatorics on Words

Factor complexity

I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be be be be I'mma I'mma be
I'mma be be be be I'mma I'mma be
I'mma be be be be I'mma I'mma be. [1]

Is this the real life ? Is this just fantasy ?
Caught in a landslide, no escape from reality.
Open your eyes, look up to the skies and see.
I'm just a poor boy, I need no sympathy,
Because I'm easy come, easy go, little high,
little low. [2]

[1]. W. Adams, T. Brenneck, M. Deller, F. Duhamel, D. Foder, J.L. Gomez, K. Harris, A.P. Lindo, J. Tankel (2009).

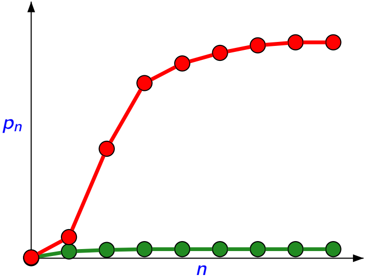
[2]. F. Mercury (1975).

Combinatorics on Words

Factor complexity

I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be, I'mma be, I'mma I'mma I'mma be
I'mma be be be be I'mma I'mma be
I'mma be be be be I'mma I'mma be
I'mma be be be be I'mma I'mma be. [1]

Is this the real life ? Is this just fantasy ?
Caught in a landslide, no escape from reality.
Open your eyes, look up to the skies and see.
I'm just a poor boy, I need no sympathy,
Because I'm easy come, easy go, little high,
little low. [2]



[1]. W. Adams, T. Brenneck, M. Deller, F. Duhamel, D. Foder, J.L. Gomez, K. Harris, A.P. Lindo, J. Tankel (2009).

[2]. F. Mercury (1975).

Codes

$A = \{\bullet, -, ␣\}$

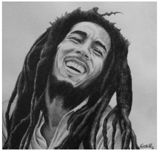
● — ␣	● ● ●␣	— — —␣	● ● ● — ␣
— ● ● ● ␣	● ● ␣	● — — · ␣	● — —␣
— ● — ● ␣	● — — — ␣	— — ● — ␣	— ● ● — ␣
— ● ●␣	— ● —␣	● — ●␣	— ● — — ␣
●␣	● — ● ● ␣	● ● ●␣	— — ● ● ␣
● ● — ● ␣	— — ␣	—␣	
— — ●␣	● — ␣	● ● —␣	

Codes

$A = \{\bullet, -, ␣\}$

● — ␣	● ● ●␣	— — —␣	● ● ● — ␣
— ● ● ● ␣	● ● ␣	● — — · ␣	● — —␣
— ● — ● ␣	● — — — ␣	— — ● — ␣	— ● ● — ␣
— ● ●␣	— ● —␣	● — ●␣	— ● — — ␣
●␣	● — ● ● ␣	● ● ●␣	— — ● ● ␣
● ● — ● ␣	— — ␣	—␣	
— — ●␣	● — ␣	● ● —␣	

— ● — ● ␣ ● ● — ␣ ● — ● ␣ ● ●␣ — — — ␣ ● ● ● ␣ ● ␣ ● — ● ␣



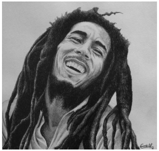
Codes

$A = \{\bullet, -, \sqcup\}$, $B = \{\mathbf{A}, \mathbf{B}, \ldots, \mathbf{Z}\}$, $f : B^* \to A^*$

A ↦ • — ␣	**H** ↦ • • •␣	**O** ↦ — — —␣	**V** ↦ • • • — ␣
B ↦ — • • • ␣	**I** ↦ • • ␣	**P** ↦ • — — · ␣	**W** ↦ • — —␣
C ↦ — • — • ␣	**J** ↦ • — — — ␣	**Q** ↦ — — • — ␣	**X** ↦ — • • — ␣
D ↦ — • •␣	**K** ↦ — • —␣	**R** ↦ • — •␣	**Y** ↦ — • — — ␣
E ↦ •␣	**L** ↦ • — • • ␣	**S** ↦ • • •␣	**Z** ↦ — — • • ␣
F ↦ • • — • ␣	**M** ↦ — — ␣	**T** ↦ —␣	
G ↦ — — •␣	**N** ↦ • — ␣	**U** ↦ • • —␣	

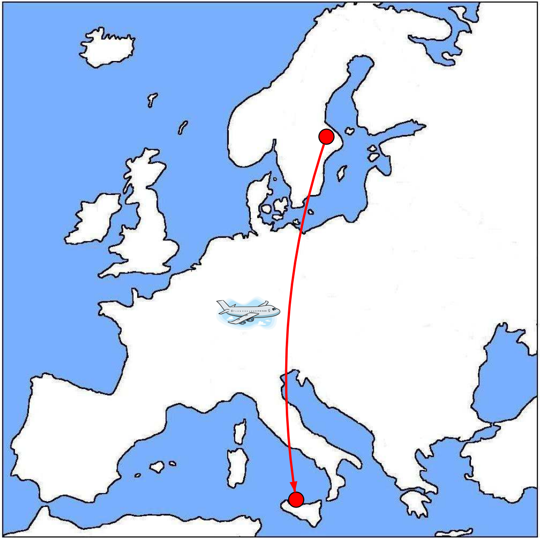
— • — • ␣ • • —␣ • — •␣ • • ␣ — — —␣ • • •␣ •␣ • — •␣

C **U** **R** **I** **O** **S** **E** **R**



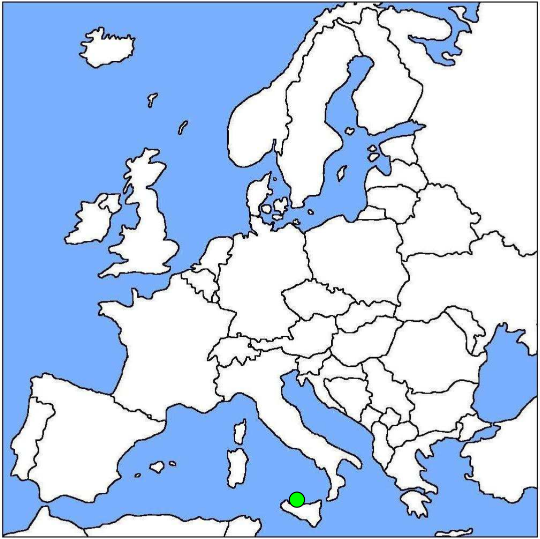
Dynamical Systems

 $: E \to E$



Dynamical Systems

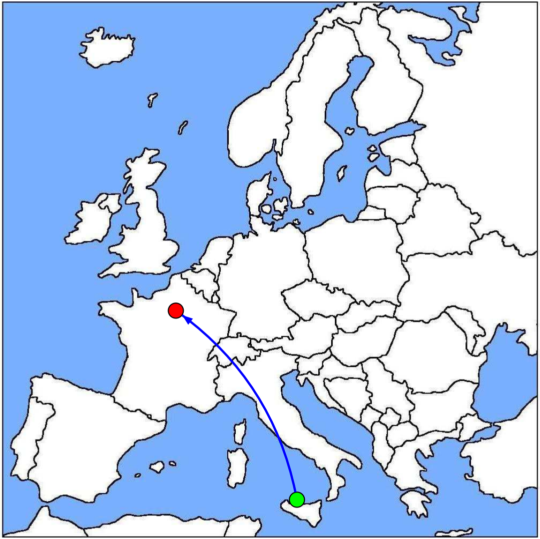
: $E \to E$



Dynamical Systems

 : $E \to E$

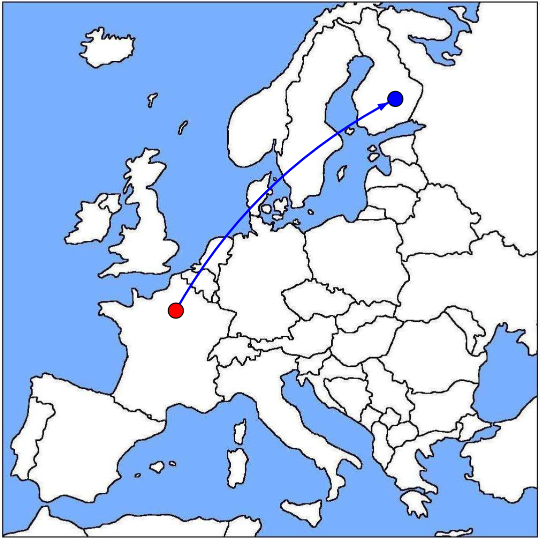
I F



Dynamical Systems

 : $E \to E$

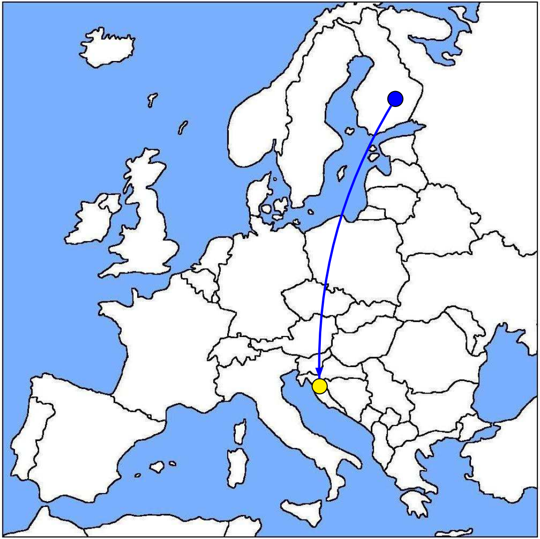
I F S



Dynamical Systems

 : $E \to E$

I F S H



Dynamical Systems

 : $E \to E$

I F S H F

Dynamical Systems

 : $E \to E$

I F S H F F

Dynamical Systems

 : $E \to E$

I F S H F F Č

Dynamical Systems

 : $E \to E$

I F S H F F Č F . . .

SUMMARY

I. Motivation : " *The Fibonacci Menace* "

II. Arnaux-Rauzy sets : " *Attack of the Arnaux-Rauzy* "

III. Interval exchange sets : " *Revenge of the IET* "

IV. Tree and neutral sets : " *Two New Hopes* "

V. Bifix codes : " *Bifix Codes Strike Back* "

VI. Return words : " *Return of the Word* "

VII. To conclude : " *The Audience Awakens* "

Fibonacci

$$x = abaababaabaababa\cdots$$

$$x = \lim_{n\to\infty} \varphi^n(a) \qquad \text{where} \qquad \varphi : \begin{cases} a \mapsto ab \\ b \mapsto a \end{cases}$$

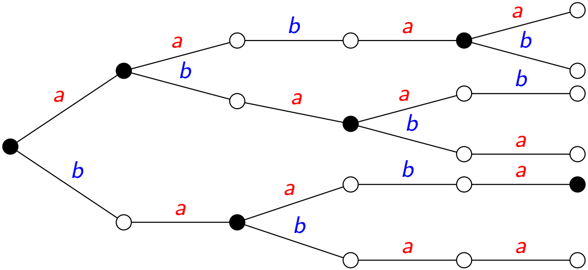
Fibonacci

$$x = abaababaabaababa\cdots$$

The *Fibonacci set* (set of factors of x) is a Sturmian set.

Definition

A *Sturmian* set is a factorial set of factor complexity $p_n = n + 1$.



Fibonacci

$$x = abaababaabaababa \cdots$$

The *Fibonacci set* (set of factors of x) is a Sturmian set. It is a uniformly recurrent set.

Definition

A factorial set S is *recurrent* if for every $u \in S$ there is a $v \in S$ such that uvu is in S.

Fibonacci

$$x = abaababaabaababa\cdots$$

The *Fibonacci set* (set of factors of x) is a Sturmian set. It is a uniformly recurrent set.

Definition

A factorial set S is *recurrent* if for every $u \in S$ there is a $v \in S$ such that uvu is in S. It is *uniformly recurrent* (or *minimal*) if for every $u \in S$ there exists an $n \in \mathbb{N}$ such that u is a factor of every word of length n in S.

Fibonacci

$$x = abaababaabaababa\cdots$$

The *Fibonacci set* (set of factors of x) is a Sturmian set. It is a uniformly recurrent set.

Definition

A factorial set S is *recurrent* if for every $u \in S$ there is a $v \in S$ such that uvu is in S. It is *uniformly recurrent* (or *minimal*) if for every $u \in S$ there exists an $n \in \mathbb{N}$ such that u is a factor of every word of length n in S.

Proposition

Uniform recurrence $\Longrightarrow$ recurrence.

Fibonacci

$$x = abaababaabaababa\cdots$$

The *Fibonacci set* (set of factors of x) is a Sturmian set. It is an Arnaux-Rauzy set.

Definition

An *Arnaux-Rauzy* set is a factorial set closed by reversal with $p_n = (\mathrm{Card}\,(A) - 1)n + 1$ having a unique right special factor for each length.

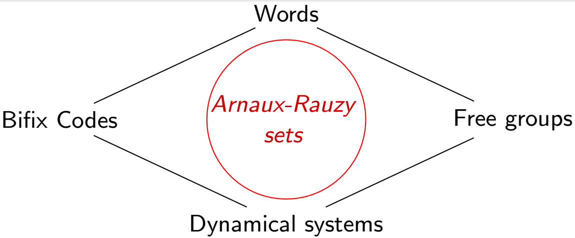
Fibonacci

$x = abaababaabaababa\cdots$

The *Fibonacci set* (set of factors of x) is a Sturmian set. It is an Arnaux-Rauzy set.

Definition

An *Arnaux-Rauzy* set is a factorial set closed by reversal with $p_n = (\mathrm{Card}\,(A) - 1)n + 1$ having a unique right special factor for each length.



[J. Berstel, C. De Felice, D. Perrin, C. Reutenauer, G. Rindone : "**Bifix codes and Sturmian words**" (2012).]

Fibonacci

$$x = ab\ aa\ ba\ ba\ ab\ aa\ ba\ ba\ \cdots$$

Fibonacci

$$x = ab\ aa\ ba\ ba\ ab\ aa\ ba\ ba\ \cdots$$

$$f : \begin{cases} u & \mapsto & aa \\ v & \mapsto & ab \\ w & \mapsto & ba \end{cases}$$

Fibonacci

$$f^{-1}(x) = \; v \; u \; w \; w \; v \; u \; w \; w \; \cdots$$

$$f : \begin{cases} u & \mapsto & aa \\ v & \mapsto & ab \\ w & \mapsto & ba \end{cases}$$

Fibonacci

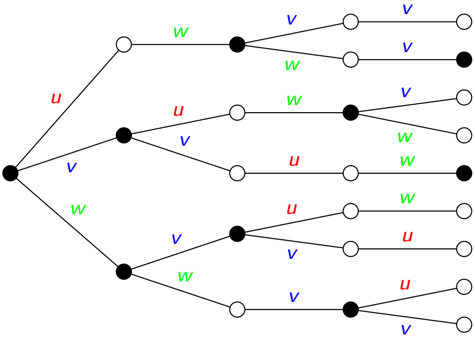
$$f^{-1}(x) = \; v \; u \; w \; w \; v \; u \; w \; w \; \cdots$$

Is the set of factors of $f^{-1}(S)$ an Arnaux-Rauzy set ?

Fibonacci

$$f^{-1}(x) = v\ u\ w\ w\ v\ u\ w\ w\ \cdots$$

Is the set of factors of $f^{-1}(S)$ an Arnaux-Rauzy set ? No !



n	0 ,	1 ,	2 ,	3 ,	4 ,	...
p_n :	1 ,	3 ,	5 ,	7 ,	9 ,	...

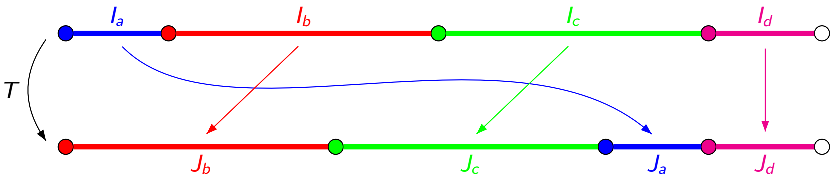
Fibonacci

$$f^{-1}(x) = \; v \; u \; w \; w \; v \; u \; w \; w \; \cdots$$

Interval exchanges

Let $(I_a)_{a\in A}$ and $(J_a)_{a\in A}$ be two partitions of $[0,1[$.
An *interval exchange transformation* (IET) is a map $T : [0,1[\to [0,1[$ defined by

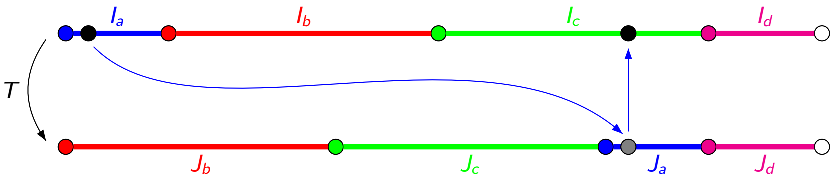
$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in A}$ and $(J_a)_{a \in A}$ be two partitions of $[0, 1[$.
An *interval exchange transformation* (IET) is a map $T : [0, 1[\to [0, 1[$ defined by

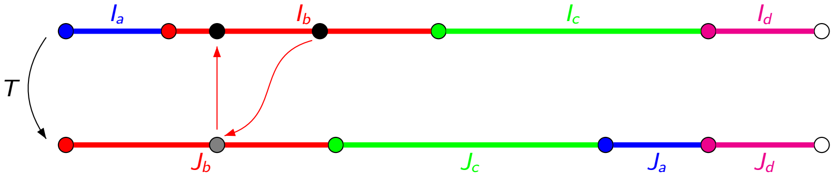
$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



Interval exchanges

Let $(I_a)_{a \in A}$ and $(J_a)_{a \in A}$ be two partitions of $[0,1[$.
An *interval exchange transformation* (IET) is a map $T : [0,1[\to [0,1[$ defined by

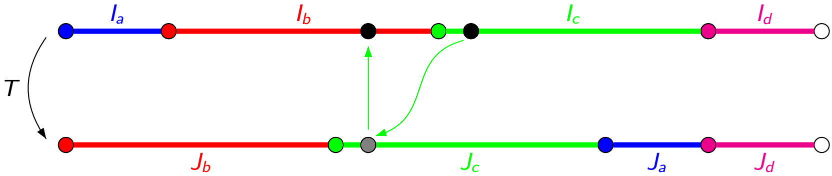
$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



Interval exchanges

Let $(I_a)_{a\in A}$ and $(J_a)_{a\in A}$ be two partitions of $[0,1[$.
An *interval exchange transformation* (IET) is a map $T : [0,1[\to [0,1[$ defined by

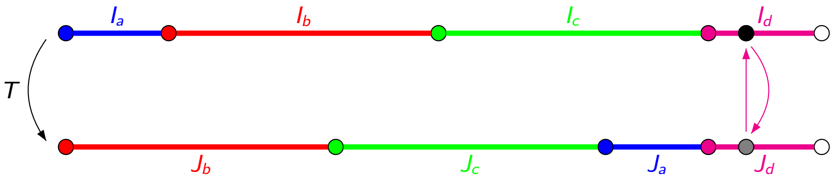
$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



Interval exchanges

Let $(I_a)_{a\in A}$ and $(J_a)_{a\in A}$ be two partitions of $[0,1[$.
An *interval exchange transformation* (IET) is a map $T : [0,1[\to [0,1[$ defined by

$$T(z) = z + \alpha_z \quad \text{if } z \in I_a.$$



Interval exchanges

T is said to be *minimal* if for any point $z \in [0, 1[$ the orbit $\mathcal{O}(z) = \{T^n(z) \mid n \in \mathbb{Z}\}$ is dense in $[0, 1[$.

T is said *regular* if the orbits of the separation points $\neq 0$ are infinite and disjoint.

Theorem [M. Keane (1975)]

A regular interval exchange transformation is minimal.

Interval exchanges

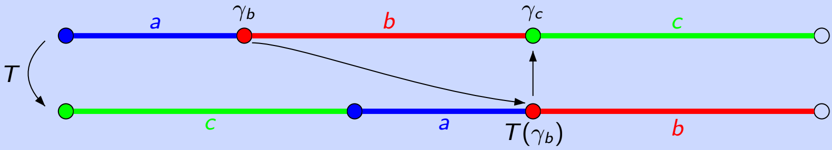
T is said to be *minimal* if for any point $z \in [0,1[$ the orbit $\mathcal{O}(z) = \{T^n(z) \mid n \in \mathbb{Z}\}$ is dense in $[0,1[$.

T is said *regular* if the orbits of the separation points $\neq 0$ are infinite and disjoint.

Theorem [M. Keane (1975)]

A regular interval exchange transformation is minimal.

Example (the converse is not true)

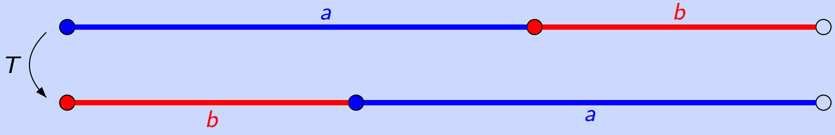


Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)

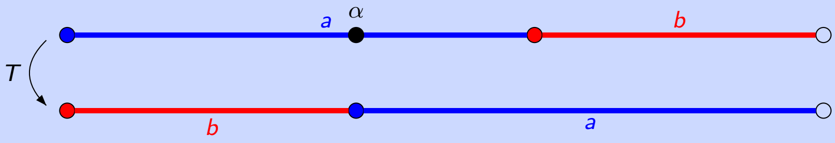


Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)



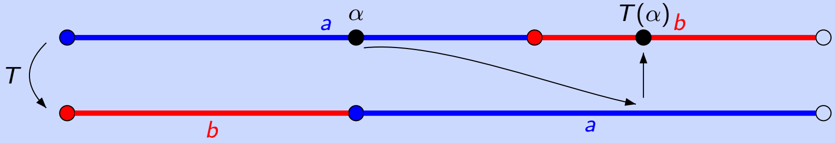
$\Sigma_T(\alpha) = a$

Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)



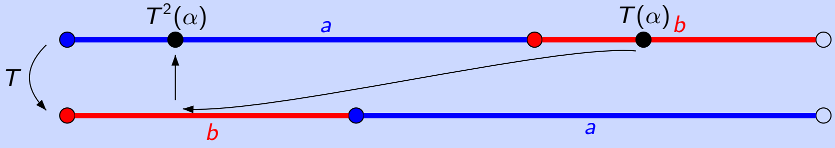
$$\Sigma_T(\alpha) = a\ b$$

Interval exchanges

The *natural coding* of T relative to $z \in [0, 1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3 - \sqrt{5})/2$)



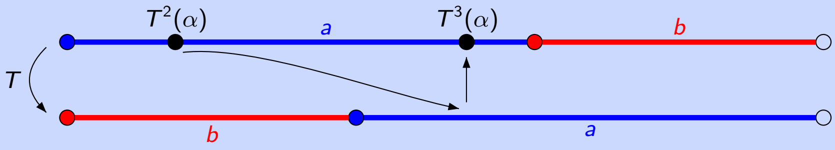
$\Sigma_T(\alpha) = a\, b\, a$

Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)



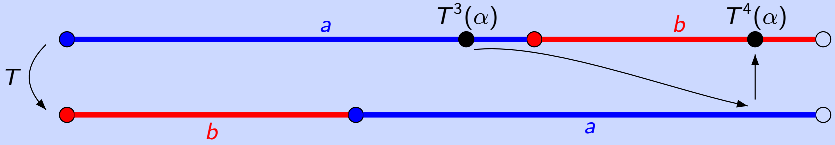
$$\Sigma_T(\alpha) = a\, b\, a\, a$$

Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)



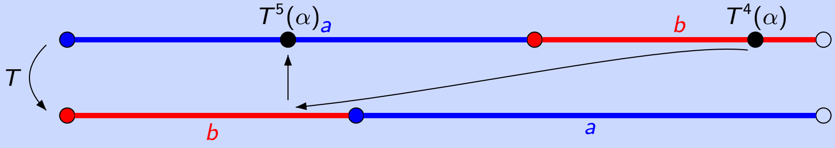
$$\Sigma_T(\alpha) = a\ b\ a\ a\ b$$

Interval exchanges

The *natural coding* of T relative to $z \in [0,1[$ is the infinite word $\Sigma_T(z) = a_0 a_1 \cdots \in A^\omega$ defined by

$$a_n = a \quad \text{if } T^n(z) \in I_a.$$

Example (Fibonacci, $\alpha = (3-\sqrt{5})/2$)



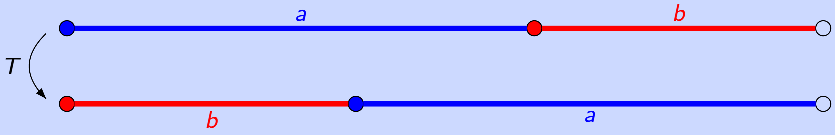
$\Sigma_T(\alpha) = a\,b\,a\,a\,b\,a \cdots$

Interval exchanges

The set $\mathcal{L}(T) = \bigcup_{z\in[0,1[} \mathrm{Fac}(\Sigma_T(z))$ is said a (*minimal*, *regular*) *interval exchange set*.

Remark. If T is minimal, $\mathrm{Fac}(\Sigma_T(z))$ does not depend on the point z.

Example (Fibonacci)



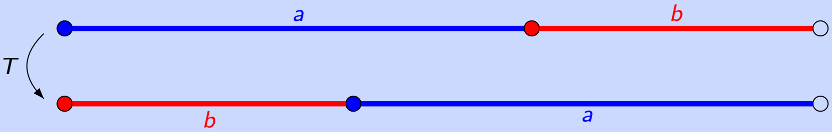
$$\mathcal{L}(T) = \{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, \dots\}$$

Interval exchanges

The set $\mathcal{L}(T) = \bigcup_{z \in [0,1[} \mathrm{Fac}\,(\Sigma_T(z))$ is said a (*minimal*, *regular*) *interval exchange set*.

Remark. If T is minimal, $\mathrm{Fac}\,(\Sigma_T(z))$ does not depend on the point z.

Example (Fibonacci)

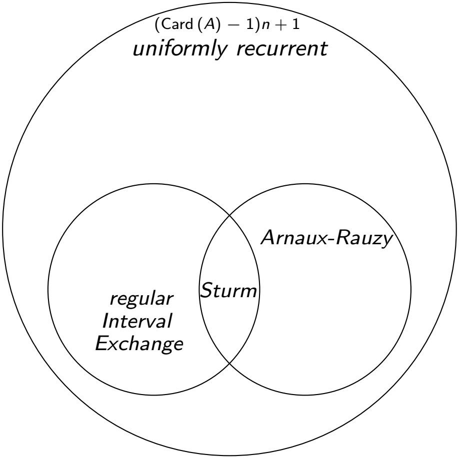


$$\mathcal{L}(T) = \left\{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, \ldots\right\}$$

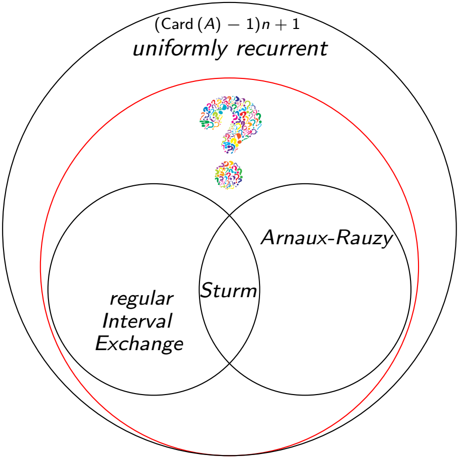
Proposition

Regular interval exchange sets have factor complexity $p_n = (\mathrm{Card}\,(A) - 1)n + 1$.

Arnaux-Rauzy and Interval Exchanges



Arnaux-Rauzy and Interval Exchanges



SUMMARY

IV. Tree and neutral sets : *" Two New Hopes "*

- Extension graphs and multiplicity
- Tree and neutral sets

V. Bifix codes : *" Bifix Codes Strike Back "*

- Maximal bifix decoding
- Cardinality of bifix codes
- Finite index basis property

VI. Return words : *" Return of the Word "*

- Cardinality of return words
- Return theorem
- Derivation

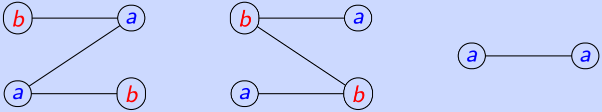
Extension graphs

The *extension graph* of a word $w \in S$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{a \in A \,|\, aw \in S\}, \\ R(w) &= \{a \in A \,|\, wa \in S\}, \\ B(w) &= \{(a,b) \in A \,|\, awb \in S.\} \end{aligned}$$

Example (Fibonacci, $S = \{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, bab, \ldots\}$)

$\mathcal{E}(\varepsilon)$ $\quad$ $\mathcal{E}(a)$ $\quad$ $\mathcal{E}(b)$



Extension graphs

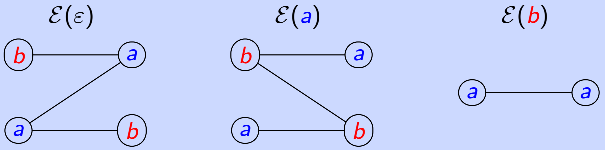
The *extension graph* of a word $w \in S$ is the undirected bipartite graph $\mathcal{E}(w)$ with vertices $L(w) \sqcup R(w)$ and edges $B(w)$, where

$$\begin{aligned} L(w) &= \{a \in A \,|\, aw \in S\}, \\ R(w) &= \{a \in A \,|\, wa \in S\}, \\ B(w) &= \{(a, b) \in A \,|\, awb \in S.\} \end{aligned}$$

The *multiplicity* of a word w is the quantity

$$m(w) = \text{Card}\,(B(w)) - \text{Card}\,(L(w)) - \text{Card}\,(R(w)) + 1.$$

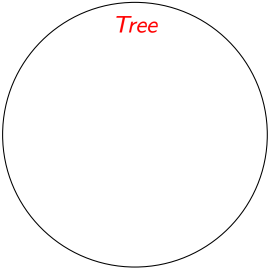
Example (Fibonacci, $S = \{\varepsilon, a, b, aa, ab, ba, aab, aba, baa, bab, \ldots\}$)



Tree and neutral sets

Definition

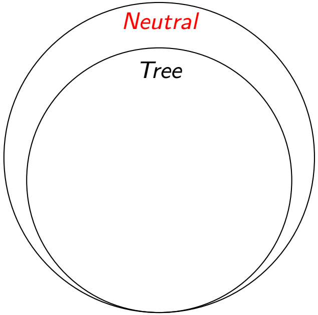
A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$.



Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$. It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

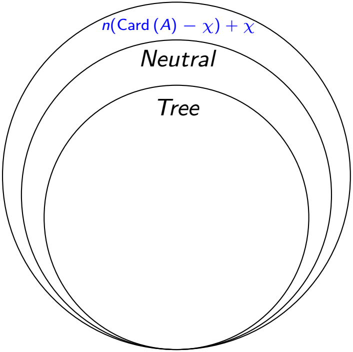


Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$. It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

The *characteristic* of a neutral/tree set S is the quantity $\chi(S) = 1 - m(\varepsilon)$.



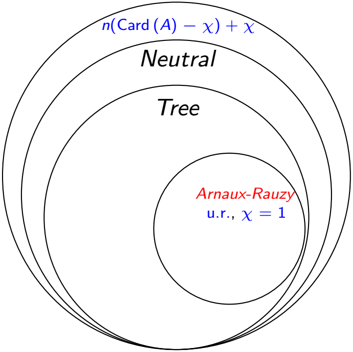
[J. Cassaigne : "**Complexité et facteurs spéciaux**" (1997).]

Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$.
It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

The *characteristic* of a neutral/tree set S is the quantity $\chi(S) = 1 - m(\varepsilon)$.



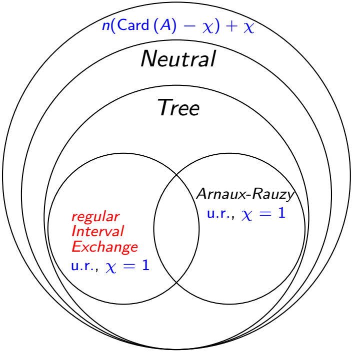
[Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone : "**Acyclic, connected and tree sets**" (2014).]

Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$. It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

The *characteristic* of a neutral/tree set S is the quantity $\chi(S) = 1 - m(\varepsilon)$.



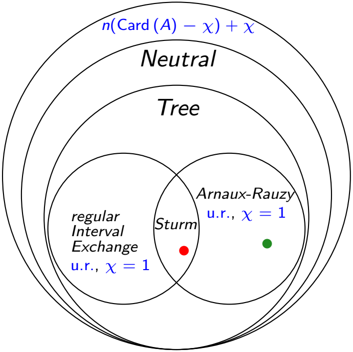
[Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone : "**Bifix codes and interval exchanges**" (2015).]

Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$. It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

The *characteristic* of a neutral/tree set S is the quantity $\chi(S) = 1 - m(\varepsilon)$.



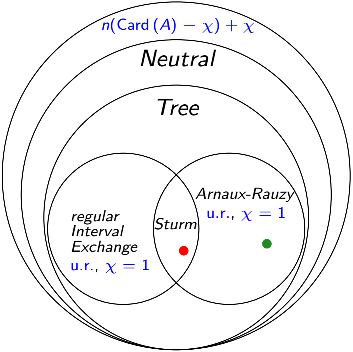
- Fibonacci
- Tribonacci

Tree and neutral sets

Definition

A factorial set S is called a *tree set* if the graph $\mathcal{E}(w)$ is a tree for any nonempty $w \in S$. It is called *neutral* if every nonempty word has multiplicity $m(w) = 0$.

The *characteristic* of a neutral/tree set S is the quantity $\chi(S) = 1 - m(\varepsilon)$.



- ● Fibonacci
- ? 2-coded Fibonacci
- ● Tribonacci
- ? 2-coded Tribonacci

Bifix codes

Definition

A *bifix code* is a set $X \subset A^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

Example

- $\{aa, ab, ba\}$
- $\{aa, ab, bba, bbb\}$
- $\{ac, bcc, bcbca\}$

Bifix codes

Definition

A *bifix code* is a set $X \subset A^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

A bifix code $X \subset S$ is S*-maximal* if it is not properly contained in a bifix code $Y \subset S$.

Example (Fibonacci)

The set $X = \{aa, ab, ba\}$ is an S-maximal bifix code.
It is not an A^*-maximal bifix code, since $X \subset X \cup \{bb\}$.

Bifix codes

Definition

A *bifix code* is a set $X \subset A^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

A bifix code $X \subset S$ is *S-maximal* if it is not properly contained in a bifix code $Y \subset S$.

A *coding morphism* for a bifix code $X \subset A^+$ is a morphism $f : B^* \to A^*$ which maps bijectively B onto X.

Example

The map $f : \{u, v, w\}^* \to \{a, b\}^*$ is a coding morphism for $X = \{aa, ab, ba\}$.

$$f : \begin{cases} u \mapsto aa \\ v \mapsto ab \\ w \mapsto ba \end{cases}$$

Bifix codes

Definition

A *bifix code* is a set $X \subset A^+$ of nonempty words that does not contain any proper prefix or suffix of its elements.

A bifix code $X \subset S$ is *S-maximal* if it is not properly contained in a bifix code $Y \subset S$.

A *coding morphism* for a bifix code $X \subset A^+$ is a morphism $f : B^* \to A^*$ which maps bijectively B onto X.

Example

The map $f : \{u, v, w\}^* \to \{a, b\}^*$ is a coding morphism for $X = \{aa, ab, ba\}$.

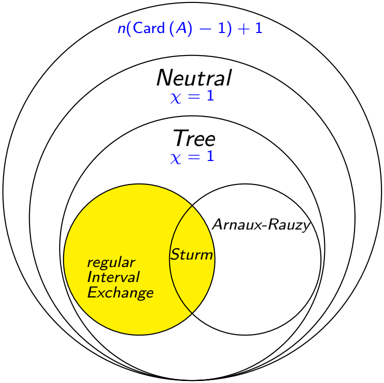
$$f : \begin{cases} u \mapsto aa \\ v \mapsto ab \\ w \mapsto ba \end{cases}$$

When S is factorial and X is an S-maximal bifix code, the set $f^{-1}(S)$ is called a *maximal bifix decoding* of S.

Maximal bifix decoding

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014)]

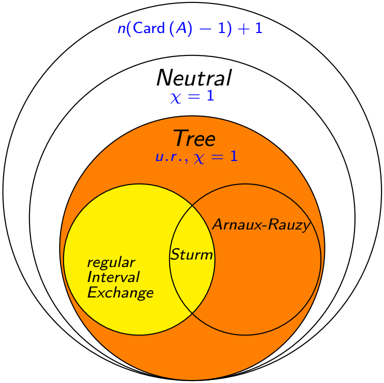
The family of regular interval exchange sets is closed under maximal bifix decoding.



Maximal bifix decoding

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

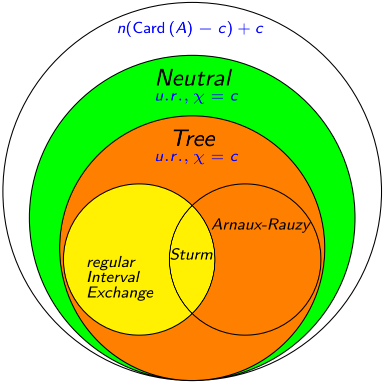
The family of *(uniformly)* recurrent tree sets of characteristic 1 is closed under maximal bifix decoding.



Maximal bifix decoding

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015) ; Dolce, Perrin (2016)]

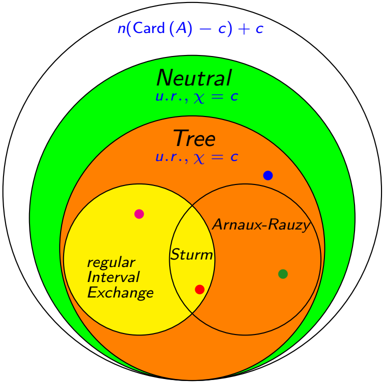
The family of *(uniformly)* recurrent neutral sets (resp. tree sets) of characteristic c is closed under maximal bifix decoding.



Maximal bifix decoding

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015) ; Dolce, Perrin (2016)]

The family of *(uniformly)* recurrent neutral sets (resp. tree sets) of characteristic c is closed under maximal bifix decoding.



- Fibonacci
- 2-coded Fibonacci
- Tribonacci
- 2-coded Tribonacci

Parse and degree

Definition

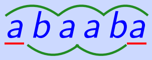
A *parse* of a word w with respect to a bifix code X is a triple (q, x, p) such that :

- $w = qxp$,
- q has no suffix in X,
- $x \in X^*$ and
- p has no prefix in X.

Example

Let $X = \{aa, ab, ba\}$ and $w = abaaba$. The two possible parses of w are :

- $(\varepsilon, ab\,aa\,ba, \varepsilon)$,
- $(a, ba\,ab, a)$.



Parse and degree

Definition

A *parse* of a word w with respect to a bifix code X is a triple (q, x, p) such that :

- $w = qxp$,
- q has no suffix in X,
- $x \in X^*$ and
- p has no prefix in X.

The S-*degree* of X is the maximal number of parses with respect to X of a word of S.

Example (Fibonacci)

- The set $X = \{aa, ab, ba\}$ has S-degree 2.
- The set $X = S \cap A^n$ has S-degree n.

Cardinality of bifix codes

Theorem [Dolce, Perrin (2016)]

Let S be a neutral set of characteristic c.
For any finite S-maximal bifix code X of S-degree n, one has

$$\mathrm{Card}\,(X) = n\big(\mathrm{Card}\,(A) - c\big) + c.$$

Example (Fibonacci)

The S-maximal bifix codes $X = \{aa, ab, ba\}$ and $Y = \{a, bab, baab\}$ of S-degree 2 satisfy

$$\mathrm{Card}\,(X) = \mathrm{Card}\,(Y) = 2(2-1) + 1 = 3.$$

Cardinality of bifix codes

Theorem [Dolce, Perrin (2016)]

Let S be a neutral set of characteristic c.
For any finite S-maximal bifix code X of S-degree n, one has

$$\mathrm{Card}\,(X) = n\big(\mathrm{Card}\,(A) - c\big) + c.$$

Theorem [Dolce, Perrin (2016)]

Let S be a uniformly recurrent set.
If every finite S-maximal bifix code of S-degree n has $n\big(\mathrm{Card}\,(A) - c\big) + c$ elements, then S is neutral of characteristic c.

Finite index basis property

Example (Fibonacci)

The S-maximal bifix code $X = \{aa, ab, ba\}$ of S-degree 2 is a basis of $\langle A^2 \rangle$. Indeed

$$bb = ba\,(aa)^{-1}\,ab$$

Finite index basis property

Example (Fibonacci)

The S-maximal bifix code $X = \{aa, ab, ba\}$ of S-degree 2 is a basis of $\langle A^2 \rangle$. Indeed

$$bb = ba\,(aa)^{-1}\,ab$$

Also $S \cap A^3 = \{aab, aba, baa, bab\}$ is a basis of $\langle A^3 \rangle$:

$$\begin{array}{rcl} aaa & = & aab(bab)^{-1}baa \\ abb & = & aba(baa)^{-1}bab \\ bba & = & bab(aab)^{-1}aba \\ bbb & = & bba(aba)^{-1}aab \end{array}$$

Finite index basis property

Definition

A set $S \subset A^+$ satisfies the *finite index basis property* if for any finite bifix code $X \subset S$:

X is an S-maximal bifix code of S-degree d if and only if it is a basis of a subgroup of index d of the free group on A.

Finite index basis property

Definition

A set $S \subset A^+$ satisfies the *finite index basis property* if for any finite bifix code $X \subset S$:

X is an S-maximal bifix code of S-degree d if and only if it is a basis of a subgroup of index d of the free group on A.

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014)]

A regular interval exchange set satisfies the finite index basis property.

Finite index basis property

Definition

A set $S \subset A^+$ satisfies the *finite index basis property* if for any finite bifix code $X \subset S$:

X is an S-maximal bifix code of S-degree d if and only if it is a basis of a subgroup of index d of the free group on A.

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

A *(uniformly)* recurrent tree set of characteristic 1 satisfies the finite index basis property.

Finite index basis property

Definition

A set $S \subset A^+$ satisfies the *finite index basis property* if for any finite bifix code $X \subset S$:

X is an S-maximal bifix code of S-degree d if and only if it is a basis of a subgroup of index d of the free group on A.

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

A *(uniformly)* recurrent tree set of characteristic 1 satisfies the finite index basis property.

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2015)]

A uniformly recurrent set satisfying the finite index basis property is a tree sets of characteristic 1.

Return words

A (*right*) *return word* to w in S is a nonempty word u such that $wu \in S$ starts and ends with w but has no w as an internal factor. Formally,

$$\mathcal{R}_S(w) = \{u \in A^+ \mid wu \in S \cap (A^+w \setminus A^+wA^+)\}$$

Example (Fibonacci)

$\mathcal{R}_S(b) = \{a\underline{b}, aa\underline{b}\}$

$$\varphi(a)^\omega = abaabab\underline{b}aabaababaabab aa\underline{b}aababaabaab\cdots$$

Return words

A (*right*) *return word* to w in S is a nonempty word u such that $wu \in S$ starts and ends with w but has no w as an internal factor. Formally,

$$\mathcal{R}_S(w) = \{u \in A^+ \mid wu \in S \cap (A^+ w \setminus A^+ w A^+)\}$$

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{b\underline{aa}, bab\underline{aa}\}$

$$\varphi(a)^\omega = abaabab\underline{aa}baababaababaab\underline{aa}babaabaab\cdots$$

Cardinality of return words

Theorem [L. Vuillon (2001)]

Let S be a Sturmian set. For every $w \in S$, one has

$$\mathrm{Card}\left(\mathcal{R}_S(w)\right) = 2.$$

Cardinality of return words

Theorem [L. Vuillon (2001) ; L. Balková, E. Pelantová, W. Steiner (2008)]

Let S be a recurrent neutral set of characteristic 1. For every $w \in S$, one has

$$\mathrm{Card}\,(\mathcal{R}_S(w)) = \mathrm{Card}\,(A)\,.$$

Cardinality of return words

Theorem [L. Vuillon (2001) ; L. Balková, E. Pelantová, W. Steiner (2008) ; Dolce, Perrin (2016)]

Let S be a recurrent neutral set. For every $w \in S$, one has

$$\text{Card}\,(\mathcal{R}_S(w)) = \text{Card}\,(A) - \chi(S) + 1.$$

Cardinality of return words

Theorem [L. Vuillon (2001) ; L. Balková, E. Pelantová, W. Steiner (2008) ; Dolce, Perrin (2016)]

Let S be a recurrent neutral set. For every $w \in S$, one has

$$\mathrm{Card}\,(\mathcal{R}_S(w)) = \mathrm{Card}\,(A) - \chi(S) + 1.$$

Corollary

A recurrent neutral set is uniformly recurrent.

Proof. A recurrent set S is uniformly recurrent if and only if $\mathcal{R}_S(w)$ is finite for all $w \in S$.

Return Theorem

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014)]

Let S be a recurrent tree set of characteristic 1.
Then, for any $w \in S$, the set $\mathcal{R}_S(w)$ is a basis of the free group on A.

Return Theorem

Theorem [Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014)]

Let S be a recurrent tree set of characteristic 1.
Then, for any $w \in S$, the set $\mathcal{R}_S(w)$ is a basis of the free group on A.

Example (Fibonacci)

The set $\mathcal{R}_S(aa) = \{baa, babaa\}$ is a basis of the free group. Indeed,

$$\begin{aligned} a &= baa\,(babaa)^{-1}\,baa \\ b &= baa\,a^{-1}\,a^{-1} \end{aligned}$$

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$. Let $f : B^* \to A^*$ defined by $f(u) = baa$, $f(v) = babaa$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$
$$f^{-1}(a^{-1}x) = u\ v\ u\ v\ v\ u\ v\ u\ \cdots$$

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$. Let $f : B^* \to A^*$ defined by $f(u) = baa$, $f(v) = babaa$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$
$$f^{-1}(a^{-1}x) = u\ v\ u\ v\ v\ u\ v\ u\cdots$$

Definition

Let S be a recurrent, $w \in S$ and $f : B^* \to A^*$ a morphism such that $B \overset{1:1}{\longleftrightarrow} \mathcal{R}_S(w)$. The set $\mathcal{D}_f(S) = f^{-1}\left(w^{-1}S\right)$ is called the *derived set* of S with respect to f.

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$. Let $f : B^* \to A^*$ defined by $f(u) = baa$, $f(v) = babaa$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$
$$f^{-1}(a^{-1}x) = u\ v\ u\ v\ v\ u\ v\ u \cdots$$

Definition

Let S be a recurrent, $w \in S$ and $f : B^* \to A^*$ a morphism such that $B \overset{1:1}{\longleftrightarrow} \mathcal{R}_S(w)$.
The set $\mathcal{D}_f(S) = f^{-1}\left(w^{-1}S\right)$ is called the *derived set* of S with respect to f.

Theorem [J. Justin, L. Vuillon (2000)]

The family of Sturmian sets is closed under derivation.

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$. Let $f : B^* \to A^*$ defined by $f(u) = baa$, $f(v) = babaa$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$
$$f^{-1}(a^{-1}x) = u\ v\ u\ v\ v\ u\ v\ u\cdots$$

Definition

Let S be a recurrent, $w \in S$ and $f : B^* \to A^*$ a morphism such that $B \xleftrightarrow{1:1} \mathcal{R}_S(w)$. The set $\mathcal{D}_f(S) = f^{-1}\left(w^{-1}S\right)$ is called the *derived set* of S with respect to f.

Theorem [J. Justin, L. Vuillon (2000) ; Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014)]

The family of regular interval exchange sets is closed under derivation.

Derivation

Example (Fibonacci)

$\mathcal{R}_S(aa) = \{baa, babaa\}$. Let $f : B^* \to A^*$ defined by $f(u) = baa$, $f(v) = babaa$.

$$x = abaababaabaababaababaabaababaabaa\cdots$$
$$f^{-1}(a^{-1}x) = u\ v\ u\ v\ v\ u\ v\ u\ \cdots$$

Definition

Let S be a recurrent, $w \in S$ and $f : B^* \to A^*$ a morphism such that $B \xleftrightarrow{1:1} \mathcal{R}_S(w)$.
The set $\mathcal{D}_f(S) = f^{-1}\left(w^{-1}S\right)$ is called the *derived set* of S with respect to f.

Theorem [J. Justin, L. Vuillon (2000) ; Berthé, De Felice, Dolce, Leroy, Perrin, Reutenauer, Rindone (2014, 2015)]

The family of recurrent tree sets of characteristic 1 is closed under derivation.

SUMMARY

VII. To conclude : *" The Audience Awakens "*

- Specular sets, linear involutions and a few other stuff
- Further research directions

Specular sets

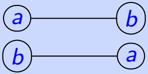
A *specular set* with respect to a linear involution θ on an alphabet A is a :

- biextendable
- tree set of characteristic 2
- θ-reduced (no factor of the form $a\theta(a)$ for $a \in A$)
- θ-symmetric (if $w \in S$ then $\theta(w) \in S$)

Example

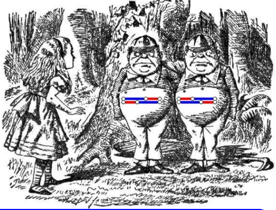
Let $A = \{a, b\}$ and θ be the identity on A. The set of factors of $(ab)^\omega$ is a specular set.

$\mathcal{E}(\varepsilon)$



$\mathcal{E}(bababa)$

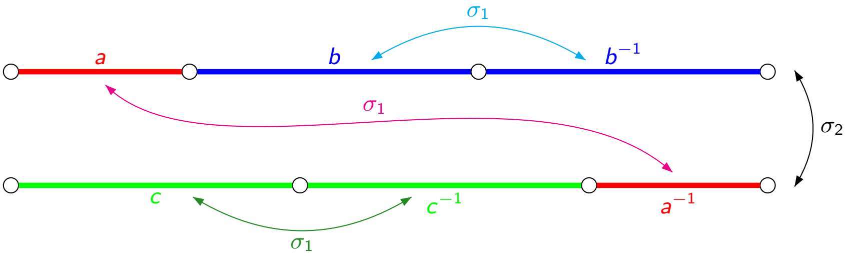
Linear involutions



Theorem [Berthé, De Felice, Delecroix, Dolce, Leroy, Perrin, Reutenauer, Rindone (2015)]

The natural coding of a linear involution without connections is a specular set.

$$T = \sigma_2 \circ \sigma_1$$



And a few other stuff

"Oh dear! Oh dear! I shall be too late!"

- Weak/strong and acyclic/connected sets;
- Palindromes in tree and specular sets;
- Branching Rauzy induction;
- Interval exchanges over a quadratic field;
- ...

Further research directions

- Decidability of the tree condition

  [Work in progress with Revekka Kyriakoglou and Julien Leroy]

- Converse of the Return Theorem

  [$\mathcal{R}_S(w)$ basis of F_A for all $w \in S \overset{?}{\Longrightarrow} S$ is a tree set of $\chi = 1$]

- Extension graphs and palindromic defect

  [$D(u) > 0$, $\mathcal{L}(u)$ closed under reversal $\Longrightarrow \mathcal{E}(w)$ has a cycle]

- *Quasi*-tree sets

  [Sets with a finite number of non-neutral elements.]

- Connections with profinite monoids/groups.

  [S tree of $\chi = 1 \Longrightarrow$ its Schützenberger group is a free profinite group]

Thank you
for attending
my ~~Tea Party~~!
PhD defense