Dendric words and dendric subshifts

Paulina CECCHI B.
(Joint work with **Valérie Berthé**)

Institute de Recherche en Informatique Fondamentale
Université Paris Diderot - Paris 7

Departamento de Matemática y Ciencia de la Computación
Facultad de Ciencia. Universidad de Santiago de Chile

Séminaire CALIN, Paris, July 2018

Let $\mathcal{A}$ be a finite alphabet and consider $x \in \mathcal{A}^{\mathbb{N}}$,

$$x = x_0 x_1 x_2 x_3 \cdots$$

Let $\mathcal{A}$ be a finite alphabet and consider $x \in \mathcal{A}^{\mathbb{N}}$,

$$x = x_0 x_1 x_2 x_3 \cdots$$

The language $\mathcal{L}(x)$ of x is it set of finite subwords or factors.

Let $\mathcal{A}$ be a finite alphabet and consider $x \in \mathcal{A}^{\mathbb{N}}$,

$$x = x_0 x_1 x_2 x_3 \cdots$$

The language $\mathcal{L}(x)$ of x is it set of finite subwords or factors.

The factor complexity of a x is the map $p_x : \mathbb{N} \to \mathbb{N}$ defined by

$$p_x(n) = |\mathcal{L}(x) \cap \mathcal{A}^n|.$$

Let $\mathcal{A}$ be a finite alphabet and consider $x \in \mathcal{A}^{\mathbb{N}}$,

$$x = x_0 x_1 x_2 x_3 \cdots$$

The language $\mathcal{L}(x)$ of x is it set of finite subwords or factors.

The factor complexity of a x is the map $p_x : \mathbb{N} \to \mathbb{N}$ defined by

$$p_x(n) = |\mathcal{L}(x) \cap \mathcal{A}^n|.$$

Given a factor w of x, a right extension of w is a letter $a \in \mathcal{A}$ such that $aw \in \mathcal{L}(x)$. We define analogously a left extension ($wb \in \mathcal{L}(x)$) and a biextension ($awb \in \mathcal{L}(x)$) of w.

Let $\mathcal{A}$ be a finite alphabet and consider $x \in \mathcal{A}^{\mathbb{N}}$,

$$x = x_0 x_1 x_2 x_3 \cdots$$

The language $\mathcal{L}(x)$ of x is it set of finite subwords or factors.

The factor complexity of a x is the map $p_x : \mathbb{N} \to \mathbb{N}$ defined by

$$p_x(n) = |\mathcal{L}(x) \cap \mathcal{A}^n|.$$

Given a factor w of x, a right extension of w is a letter $a \in \mathcal{A}$ such that $aw \in \mathcal{L}(x)$. We define analogously a left extension ($wb \in \mathcal{L}(x)$) and a biextension ($awb \in \mathcal{L}(x)$) of w.

We are interested in a family of words with linear complexity and some restrictions on the possible extensions of their factors.

Outline

- Dendric words.
- Dendric subshifts (symbolic dynamical systems).
- Exploting properties of extension graphs.
    - Balance in dendric words.
    - Invariant measures and orbit equivalence.

Dendric words

Dendric words

- Let $\mathcal{A}$ be a finite non-empty alphabet and $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$. For $w \in \mathcal{L}(x)$, the extensions of w are the following sets,

$$L(w) = \{a \in \mathcal{A} \mid aw \in \mathcal{L}(x)\}$$
$$R(w) = \{a \in \mathcal{A} \mid wa \in \mathcal{L}(x)\}$$
$$B(w) = \{(a, b) \in \mathcal{A} \times \mathcal{A} \mid awb \in \mathcal{L}(x)\}.$$

Dendric words

- Let $\mathcal{A}$ be a finite non-empty alphabet and $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$. For $w \in \mathcal{L}(x)$, the extensions of w are the following sets,

$$L(w) = \{a \in \mathcal{A} \mid aw \in \mathcal{L}(x)\}$$
$$R(w) = \{a \in \mathcal{A} \mid wa \in \mathcal{L}(x)\}$$
$$B(w) = \{(a, b) \in \mathcal{A} \times \mathcal{A} \mid awb \in \mathcal{L}(x)\}.$$

- The extension graph $\mathcal{E}_x(w)$ of w is the undirected bipartite graph whose set of vertices is the disjoint union of $L(w)$ and $R(w)$ and whose edges are the pairs $(a, b) \in B(w)$.

Dendric words

- Example: Consider the Fibonacci word in $\{a, b\}$

$$x = abaababaabaababaababa\cdots$$

produced by the substitution $\varphi : a \mapsto ab, b \mapsto a$ $(x = \varphi^{\omega}(a))$.

Dendric words

- Example: Consider the Fibonacci word in $\{a, b\}$

$$x = abaababaabaababaababa\cdots$$

produced by the substitution $\varphi : a \mapsto ab, b \mapsto a$ $(x = \varphi^{\omega}(a))$.

- The extension graphs of a and b are

$\mathcal{E}_x(a)$ $\mathcal{E}_x(b)$

$$\mathcal{L}_3(x) = \{aba, baa, aab, bab\}.$$

Dendric words

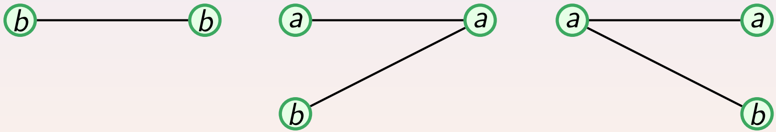
- Example: Consider the Fibonacci word in $\{a, b\}$

$$x = abaababaabaababaababa\cdots$$

produced by the substitution $\varphi : a \mapsto ab, b \mapsto a$.

- The extension graphs of aa, ab, ba are

$\mathcal{E}_x(aa)$ $\mathcal{E}_x(ab)$ $\mathcal{E}_x(ba)$



$$\mathcal{L}_4(x) = \{abaa, baab, aaba, abab, baba\}.$$

Dendric words

- If for all $w \in \mathcal{L}(x)$ the graph $\mathcal{E}_x(w)$ is a tree, x is said to be dendric.

Dendric words

- If for all $w \in \mathcal{L}(x)$ the graph $\mathcal{E}_x(w)$ is a tree, x is said to be dendric.
- The Fibonacci word is a dendric word.

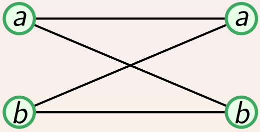
Dendric words

- If for all $w \in \mathcal{L}(x)$ the graph $\mathcal{E}_x(w)$ is a tree, x is said to be dendric.
- The Fibonacci word is a dendric word.
- Consider the Thue-Morse word in $\{a, b\}$ given by

$$y = abbabaabbaababba \cdots$$

produced by the Thue-Morse substitution $\sigma : a \mapsto ab, b \mapsto ba$. This word is not dendric. The extension graph of ϵ is

$$\mathcal{E}_y(\epsilon)$$



$$\mathcal{L}_2(y) = \{aa, ab, ba, bb\}.$$

Dendric words: examples

Dendric words: examples

- Sturmian words: aperiodic (bi)infinite words with factor complexity $p_x(n) = n + 1$.

Dendric words: examples

- Sturmian words: aperiodic (bi)infinite words with factor complexity $p_x(n) = n + 1$.

- Arnoux-Rauzy words. Consider the alphabet $\mathcal{A} = \{1, 2, \cdots, d\}$, $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is an Arnoux-Rauzy word if every factor appears infinitely often; for all n,

$$p_n(x) = (d-1)n + 1,$$

Dendric words: examples

- Sturmian words: aperiodic (bi)infinite words with factor complexity $p_x(n) = n + 1$.
- Arnoux-Rauzy words. Consider the alphabet $\mathcal{A} = \{1, 2, \cdots, d\}$, $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is an Arnoux-Rauzy word if every factor appears infinitely often; for all n,

$$p_n(x) = (d - 1)n + 1,$$

and there exists exactly one left special factor ($L(w) \geq 2$) and one rigth special factor ($R(v) \geq 2$) of each given length.

Dendric words: examples

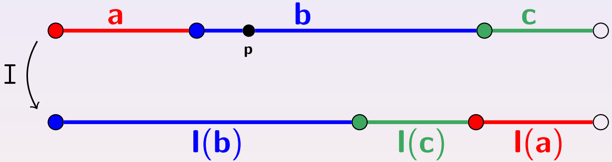
- Sturmian words: aperiodic (bi)infinite words with factor complexity $p_x(n) = n + 1$.
- Arnoux-Rauzy words. Consider the alphabet $\mathcal{A} = \{1, 2, \cdots, d\}$, $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is an Arnoux-Rauzy word if every factor appears infinitely often; for all n,

$$p_n(x) = (d-1)n + 1,$$

  and there exists exactly one left special factor ($L(w) \geq 2$) and one rigth special factor ($R(v) \geq 2$) of each given length.
- Sturmian words are Arnoux-Rauzy words for $d = 2$.

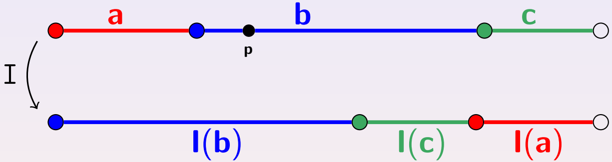
Dendric words: examples

- Codings of regular interval exchanges.



Dendric words: examples

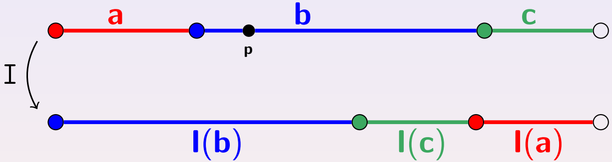
- Codings of regular interval exchanges.



⇝ Regular means that the orbits of nonzero separation points are infinite and disjoint.

Dendric words: examples

- Codings of regular interval exchanges.



⇝ Regular means that the orbits of nonzero separation points are infinite and disjoint.

$$x_p = \cdots abbacb \cdot bacbbacbba \cdots$$

Dendric subshifts

Dendric subshifts

- The set $\mathcal{A}^{\mathbb{Z}}$ equipped with the product topology of the discrete topology is a Cantor space.

Dendric subshifts

- The set $\mathcal{A}^{\mathbb{Z}}$ equipped with the product topology of the discrete topology is a Cantor space.
- The shift map T acts on $\mathcal{A}^{\mathbb{Z}}$ as $T\left((x_n)_{n\in\mathbb{Z}}\right) = (x_{n+1})_{n\in\mathbb{Z}}$.

Dendric subshifts

- The set $\mathcal{A}^{\mathbb{Z}}$ equipped with the product topology of the discrete topology is a Cantor space.
- The shift map T acts on $\mathcal{A}^{\mathbb{Z}}$ as $T\left((x_n)_{n\in\mathbb{Z}}\right) = (x_{n+1})_{n\in\mathbb{Z}}$.
- A subshift is a dynamical system (X, T) where X is a closed shift-invariant subset of $\mathcal{A}^{\mathbb{Z}}$. The language $\mathcal{L}_X$ of the subshift is defined as the union of the languaje of its elements.

Dendric subshifts

- The set $\mathcal{A}^{\mathbb{Z}}$ equipped with the product topology of the discrete topology is a Cantor space.
- The shift map T acts on $\mathcal{A}^{\mathbb{Z}}$ as $T\left((x_n)_{n\in\mathbb{Z}}\right) = (x_{n+1})_{n\in\mathbb{Z}}$.
- A subshift is a dynamical system (X, T) where X is a closed shift-invariant subset of $\mathcal{A}^{\mathbb{Z}}$. The language $\mathcal{L}_X$ of the subshift is defined as the union of the languaje of its elements.
- For all $w \in \mathcal{L}_X$, define the cylinder w as

$$[w] = \{x \in X : x_0 \cdots x_{|w|-1} = w\}.$$

The set of all cylinders is a basis of the topology of X.

Dendric subshifts

- (X, T) is minimal if it admits no non-trivial closed and shift-invariant subset: every infinite word $x \in X$ is uniformly recurrent: every word occurring in x occurs infinitely often with bounded gaps, every $x \in X$ has the same language.

Dendric subshifts

- (X, T) is minimal if it admits no non-trivial closed and shift-invariant subset: every infinite word $x \in X$ is uniformly recurrent: every word occurring in x occurs infinitely often with bounded gaps, every $x \in X$ has the same language.

- If for all $w \in \mathcal{L}_X$, the extension graph of w is a tree, X is called a dendric subshift.

Dendric subshifts

- (X, T) is minimal if it admits no non-trivial closed and shift-invariant subset: every infinite word $x \in X$ is uniformly recurrent: every word occurring in x occurs infinitely often with bounded gaps, every $x \in X$ has the same language.
- If for all $w \in \mathcal{L}_X$, the extension graph of w is a tree, X is called a dendric subshift.
- We focus on minimal dendric subshifts. This is the case when there is a dendric word $x = (x_n)_{n\in\mathbb{Z}}$ such that the subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ can be obtained as

$$X = \overline{\{T^k((x_n)_{n\in\mathbb{Z}}) : k \in \mathbb{Z}\}}.$$

Dendric subshifts

- (X, T) is minimal if it admits no non-trivial closed and shift-invariant subset: every infinite word $x \in X$ is uniformly recurrent: every word occurring in x occurs infinitely often with bounded gaps, every $x \in X$ has the same language.

- If for all $w \in \mathcal{L}_X$, the extension graph of w is a tree, X is called a dendric subshift.

- We focus on minimal dendric subshifts. This is the case when there is a dendric word $x = (x_n)_{n\in\mathbb{Z}}$ such that the subshift $X \subseteq \mathcal{A}^{\mathbb{Z}}$ can be obtained as

$$X = \overline{\{T^k((x_n)_{n\in\mathbb{Z}}) : k \in \mathbb{Z}\}}.$$

- Equivalently, $X = \{y \in \mathcal{A}^{\mathbb{Z}} : \mathcal{L}(y) \subseteq \mathcal{L}(x)\}$.

Exploting properties of extension graphs.

Extension graph

Lemma ($\star$)

Let $\mathcal{T}$ be a finite tree, with a bipartition X and Y of its set of vertices, with $|X|, |Y| \geq 2$. Let E be its set of edges. For all $x \in X$, $y \in Y$, define

$$Y_x := \{y \in Y : (x, y) \in E\} \qquad X_y := \{x \in X : (x, y) \in E\}.$$

Let $(G, +)$ be an abelian group and H a subgroup of G. Suppose that there exists a function $g : X \cup Y \cup E \to G$ satisfying the following conditions:

(1) $g(X \cup Y) \subseteq H$;

(2) *for all $x \in X$, $g(x) = \sum_{y \in Y_x} g(x, y)$, and for all $y \in Y$, $g(y) = \sum_{x \in X_y} g(x, y)$.*

Then, for all $(x, y) \in E$, $g(x, y) \in H$.

Extension graph

Proof ideas.

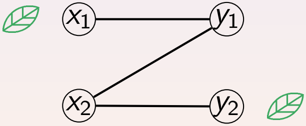
Extension graph
Proof ideas.

- Conditions (1) and (2) imply that the image under g of any edge connected to a leaf belongs to H.

Extension graph
Proof ideas.

- Conditions (1) and (2) imply that the image under g of any edge connected to a leaf belongs to H.

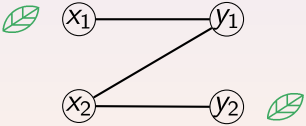
- Let $k := \max\{|X|, |Y|\}$. If $k = 2$, there is only one possibility for $\mathcal{T}$ (modulo relabeling the vertices), since $\mathcal{T}$ is connected and has no cycles, which is



Extension graph
Proof ideas.

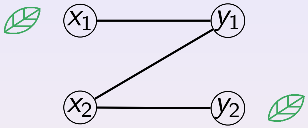
- Conditions (1) and (2) imply that the image under g of any edge connected to a leaf belongs to H.

- Let $k := \max\{|X|, |Y|\}$. If $k = 2$, there is only one possibility for $\mathcal{T}$ (modulo relabeling the vertices), since $\mathcal{T}$ is connected and has no cycles, which is



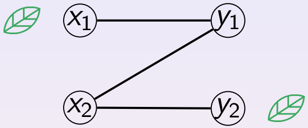
- Both $g(x_1, y_1)$ and $g(x_2, y_2)$ are in H because x_1 and y_2 are leaves.

Extension graph



- By Condition (2), one has $g(x_2) = g(x_2, y_1) + g(x_2, y_2)$, and then $g(x_2, y_1) = g(x_2) - g(x_2, y_2)$.

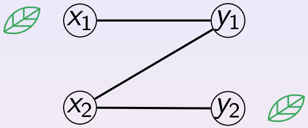
Extension graph



- By Condition (2), one has $g(x_2) = g(x_2, y_1) + g(x_2, y_2)$, and then $g(x_2, y_1) = g(x_2) - g(x_2, y_2)$.

- Since $g(x_2) \in H$ by Condition (1) and H is a group, then $g(x_2, y_1) \in H$.

Extension graph



- By Condition (2), one has $g(x_2) = g(x_2, y_1) + g(x_2, y_2)$, and then $g(x_2, y_1) = g(x_2) - g(x_2, y_2)$.
- Since $g(x_2) \in H$ by Condition (1) and H is a group, then $g(x_2, y_1) \in H$.
- We proceed by induction for $k > 2$. □

An application: Balance.

Balance

- A word $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is balanced on the factor $v \in \mathcal{L}(x)$ if there exists a constant C_v such that for every pair of factors u, w in $\mathcal{L}(x)$ with $|u| = |w|$,

$$||u|_v - |w|_v| \leq C_v.$$

Balance

- A word $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is balanced on the factor $v \in \mathcal{L}(x)$ if there exists a constant C_v such that for every pair of factors u, w in $\mathcal{L}(x)$ with $|u| = |w|$,

$$||u|_v - |w|_v| \leq C_v.$$

- If $x \in X$ and (X, T) is minimal, balance is a property of the language $\mathcal{L}_X$.

Balance

- A word $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is balanced on the factor $v \in \mathcal{L}(x)$ if there exists a constant C_v such that for every pair of factors u, w in $\mathcal{L}(x)$ with $|u| = |w|$,

$$||u|_v - |w|_v| \leq C_v.$$

- If $x \in X$ and (X, T) is minimal, balance is a property of the language $\mathcal{L}_X$.

- Sturmian words are exactly the 1-balanced words on the letters.

Balance

- A word $x \in \mathcal{A}^{\mathbb{N}}$ or $\mathcal{A}^{\mathbb{Z}}$ is balanced on the factor $v \in \mathcal{L}(x)$ if there exists a constant C_v such that for every pair of factors u, w in $\mathcal{L}(x)$ with $|u| = |w|$,

$$||u|_v - |w|_v| \leq C_v.$$

- If $x \in X$ and (X, T) is minimal, balance is a property of the language $\mathcal{L}_X$.
- Sturmian words are exactly the 1-balanced words on the letters.
- They are moreover C-balanced on factors of any length.

Frequencies

- The frequency of a factor $v \in \mathcal{L}(x)$ in $x \in \mathcal{A}^{\mathbb{Z}}$ is defined as the following limit (if it exists),

$$\lim_{n\to\infty} \frac{|x_{-n}\cdots x_n|_v}{2n+1}.$$

Proposition

The language $\mathcal{L}_X$ is balanced in the factor v if and only if v has a frequency μ_v and there exists a constant B_v such that for any factor $w \in \mathcal{L}_X$, we have $||w|_v - \mu_v|w|| \leq B_v$.

- Equivalently, v has a frequency μ_v and there exists B_v such that for all $x \in X$ and for all $n \geq 1$,

$$||x_{[0,n)}|_v - \mu_v n| \leq B_v.$$

Balance

- Balance is a mesure of disorder: convergence speed of $|x_{[0,n)}|_v/n$ towards μ_v.

Balance

- Balance is a mesure of disorder: convergence speed of $|x_{[0,n)}|_v/n$ towards μ_v.
- Balance in factors of length $(n+1)$ implies balance in factors of length n.

Balance

- Balance is a mesure of disorder: convergence speed of $|x_{[0,n)}|_v/n$ towards μ_v.
- Balance in factors of length $(n+1)$ implies balance in factors of length n.
- But balance behaviour can be different in factors of different lengths.

Balance

- Balance is a mesure of disorder: convergence speed of $|x_{[0,n)}|_v/n$ towards μ_v.
- Balance in factors of length $(n+1)$ implies balance in factors of length n.
- But balance behaviour can be different in factors of different lengths.
- Example: the language of the Thue-Morse word is balanced on letters and it is not balanced on the factors of length ℓ for every $\ell \geq 2$.

Balance

Theorem

Let (X, T) be a minimal dendric subshift on a finite alphabet $\mathcal{A}$. Then (X, T) is balanced on the letters if and only if it is balanced on the factors.

Balance

Theorem

Let (X, T) be a minimal dendric subshift on a finite alphabet $\mathcal{A}$. Then (X, T) is balanced on the letters if and only if it is balanced on the factors.

- Proof idea: use the following lemma,

Balance

Theorem

Let (X, T) be a minimal dendric subshift on a finite alphabet $\mathcal{A}$. Then (X, T) is balanced on the letters if and only if it is balanced on the factors.

- Proof idea: use the following lemma,

Lemma

Let (X, T) be a minimal dendric subshift. Let H be the following subset of $C(X, \mathbb{Z})$:

$$H = \left\{ \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^k([a])} : K_a \subseteq \mathbb{Z}, |K_a| < \infty, \alpha(a,k) \in \mathbb{Z} \right\},$$

where χ_A denotes the characteristic function of the set A, for all $A \subseteq X$. Then, for all $v \in \mathcal{L}_X$, $\chi_{[v]}$ belongs to H.

Balance

Proof.

Balance
Proof.

- Given any factor $v \in \mathcal{L}_X$, we want to find, for all $a \in \mathcal{A}$ a finite set $K_a \subseteq \mathbb{Z}$ and for all $k \in K_a$ an integer $\alpha(a, k)$ such that

$$\chi_{[v]} = \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a, k) \chi_{T^k([a])}.$$

Balance

Proof.

- Given any factor $v \in \mathcal{L}_X$, we want to find, for all $a \in \mathcal{A}$ a finite set $K_a \subseteq \mathbb{Z}$ and for all $k \in K_a$ an integer $\alpha(a,k)$ such that

$$\chi_{[v]} = \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^k([a])}.$$

- We procced by induction on $|v|$. If $|v| = 1$, v is a letter and the conclusion follows.

Balance

Proof.

- Given any factor $v \in \mathcal{L}_X$, we want to find, for all $a \in \mathcal{A}$ a finite set $K_a \subseteq \mathbb{Z}$ and for all $k \in K_a$ an integer $\alpha(a, k)$ such that

$$\chi_{[v]} = \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a, k) \chi_{T^k([a])}.$$

- We procced by induction on $|v|$. If $|v| = 1$, v is a letter and the conclusion follows.
- Let $v = v_0 \cdots v_n$,

$$\widetilde{v} = v_1 \cdots v_n,$$

$$v' = v_0 \cdots v_{n-1},$$

$$\bar{v} = v_1 \cdots v_n.$$

Balance

Proof.

Balance

Proof.

- If $\widetilde{v} = v_1 \cdots v_n$ has only one left extension,

Balance

Proof.

- If $\widetilde{v} = v_1 \cdots v_n$ has only one left extension, for all $x \in X$, $\chi_{[v]}(x) = \chi_{[\widetilde{v}]}(Tx)$. $\chi_{[\widetilde{v}]} \in H$ by induction, so for all $x \in X$,

$$\begin{aligned}
\chi_{[v]}(x) &= \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^k([a])}(Tx) \\
&= \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^{k-1}([a])}(x).
\end{aligned}$$

Balance

Proof.

- If $\widetilde{v} = v_1 \cdots v_n$ has only one left extension, for all $x \in X$, $\chi_{[v]}(x) = \chi_{[\widetilde{v}]}(Tx)$. $\chi_{[\widetilde{v}]} \in H$ by induction, so for all $x \in X$,

$$\begin{aligned} \chi_{[v]}(x) &= \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^k([a])}(Tx) \\ &= \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \chi_{T^{k-1}([a])}(x). \end{aligned}$$

Defining $K'_a := \{k-1 : k \in K_a\}$ for all $a \in \mathcal{A}$, and $\beta(a,k) = \alpha(a,k+1)$ for all $k \in K'_a$, we conclude

$$\chi_{[v]}(x) = \sum_{a \in \mathcal{A}} \sum_{k \in K'_a} \beta(a,k) \chi_{T^k([a])},$$

and then $\chi_{[v]}$ belongs to H.

Balance

Proof.

- If $v' = v_0 \cdots v_{n-1}$ has only one right extension,

Balance

Proof.

- If $v' = v_0 \cdots v_{n-1}$ has only one right extension, for all $x \in X$, $\chi_{[v]}(x) = \chi_{[v']}(x)$. We conclude by applying the induction hypothesis.

Balance
Proof.

- If $v' = v_0 \cdots v_{n-1}$ has only one right extension, for all $x \in X$, $\chi_{[v]}(x) = \chi_{[v']}(x)$. We conclude by applying the induction hypothesis.

- If $\widetilde{v}$ has more than one left extension and v' has more than one right extension, let $\mathcal{E}(\bar{v})$ be the extension graph of

$$\bar{v} = v_1 \cdots v_{n-1}.$$

- It is a tree by definition, each of the sets in its bipartition of vertices has cardinality at least two.

Balance

Proof.

- If $v' = v_0 \cdots v_{n-1}$ has only one right extension, for all $x \in X$, $\chi_{[v]}(x) = \chi_{[v']}(x)$. We conclude by applying the induction hypothesis.
- If $\widetilde{v}$ has more than one left extension and v' has more than one right extension, let $\mathcal{E}(\bar{v})$ be the extension graph of

$$\bar{v} = v_1 \cdots v_{n-1}.$$

- It is a tree by definition, each of the sets in its bipartition of vertices has cardinality at least two.
- Note that H is a subgroup of $C(X, \mathbb{Z})$.

Balance

Proof.

- Define $g : L(\bar{v}) \cup R(\bar{v}) \cup E(\bar{v}) \to C(X, \mathbb{Z})$ as follows

Balance

Proof.

- Define $g : L(\bar{v}) \cup R(\bar{v}) \cup E(\bar{v}) \to C(X, \mathbb{Z})$ as follows
  For $a \in L(\bar{v})$, $g(a) = \chi_{[a\bar{v}]}$.
  For $b \in R(\bar{v})$, $g(b) = \chi_{T^{-1}[\bar{v}b]}$.
  For $(a, b) \in E(\bar{v})$, $g(a, b) = \chi_{[a\bar{v}b]}$.

Balance
Proof.

- Define $g : L(\bar{v}) \cup R(\bar{v}) \cup E(\bar{v}) \to C(X, \mathbb{Z})$ as follows
  For $a \in L(\bar{v})$, $g(a) = \chi_{[a\bar{v}]}$.
  For $b \in R(\bar{v})$, $g(b) = \chi_{T^{-1}[\bar{v}b]}$.
  For $(a, b) \in E(\bar{v})$, $g(a, b) = \chi_{[a\bar{v}b]}$.

- Condition (1) of Lemma $(\star)$ holds by induction hypothesis.

Balance
Proof.

- Define $g : L(\bar{v}) \cup R(\bar{v}) \cup E(\bar{v}) \to C(X, \mathbb{Z})$ as follows
  For $a \in L(\bar{v})$, $g(a) = \chi_{[a\bar{v}]}$.
  For $b \in R(\bar{v})$, $g(b) = \chi_{T^{-1}[\bar{v}b]}$.
  For $(a, b) \in E(\bar{v})$, $g(a, b) = \chi_{[a\bar{v}b]}$.

- Condition (1) of Lemma $(\star)$ holds by induction hypothesis.

- For the second condition, let $a \in L(\bar{v})$. One has

$$\chi_{[a\bar{v}]} = \sum_{b \in R(\bar{v}), (a,b) \in E(\bar{v})} \chi_{[a\bar{v}b]}(x)$$

  and thus

$$g(a) = \sum_{b \in R(\bar{v}), (a,b) \in E(\bar{v})} g(a, b).$$

Balance

Proof.

- Similarly, for $b \in R(\bar{v})$ and $x \in X$, one has

$$\chi_{T^{-1}[\bar{v}b]}(x) = \chi_{[\bar{v}b]}(Tx) = \sum_{a \in L(\bar{v}), (a,b) \in E(\bar{v})} \chi_{[a\bar{v}b]}(x).$$

We conclude that for all $b \in R(\bar{v})$,

$$g(b) = \sum_{a \in L(\bar{v}), (a,b) \in E(\bar{v})} g(a, b).$$

Balance

Proof.

- Similarly, for $b \in R(\bar{v})$ and $x \in X$, one has

$$\chi_{T^{-1}[\bar{v}b]}(x) = \chi_{[\bar{v}b]}(Tx) = \sum_{a \in L(\bar{v}), (a,b) \in E(\bar{v})} \chi_{[a\bar{v}b]}(x).$$

We conclude that for all $b \in R(\bar{v})$,

$$g(b) = \sum_{a \in L(\bar{v}), (a,b) \in E(\bar{v})} g(a,b).$$

- Thanks to Lemma ($\star$), every $g(a,b)$ belongs to H. In particular, $g(v_0, v_n) = \chi_{[v]} \in H$.

Balance

Proof of the theorem.

- Suppose (X, T) is balance on every letter. Let C be a constant of balance on the letters.

Balance

Proof of the theorem.

- Suppose (X, T) is balance on every letter. Let C be a constant of balance on the letters.
- Let $v \in \mathcal{L}_X$. Let $n \geq 3$ and let u, w be two factors of $\mathcal{L}_X$ of length $n-1$ with $n-1 > |v|$. Pick a bi-infinite word $x \in X$ such that $u = x_{[i,i+n)}$ and $w = x_{[j,j+n)}$ for some indices $i, j \in \mathbb{Z}$. We have

$$||u|_v - |w|_v| = \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{[v]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{[v]}(T^\ell x) \right|.$$

Balance

Proof of the theorem.

$$||u|_v - |w|_v| = \left| \sum_{a \in \mathcal{A}} \sum_{k \in K_a} \alpha(a,k) \left(\sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right) \right.$$

Balance

Proof of the theorem.

$$
\begin{aligned}
||u|_v - |w|_v| &= \left| \sum_{a\in\mathcal{A}} \sum_{k\in K_a} \alpha(a,k) \left(\sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right) \right| \\
&\leq \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right|
\end{aligned}
$$

Balance

Proof of the theorem.

$$
\begin{aligned}
||u|_v - |w|_v| &= \left| \sum_{a\in\mathcal{A}} \sum_{k\in K_a} \alpha(a,k) \left(\sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right) \right| \\
&\leq \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right|
\end{aligned}
$$

$$
= \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) \right|
$$

Balance

Proof of the theorem.

$$
\begin{aligned}
||u|_v - |w|_v| &= \left| \sum_{a\in\mathcal{A}} \sum_{k\in K_a} \alpha(a,k) \left(\sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right) \right| \\
&\leq \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right|
\end{aligned}
$$

$$
= \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) \right|
$$

$$
= \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \cdot \underbrace{\left||(T^{-k}x)_{[i,i+n-|v|)}|_a - |(T^{-k}x)_{[j,j+n-|v|)}|_a\right|}_{\leq C}.
$$

Balance

Proof of the theorem.

$$
\begin{aligned}
||u|_v - |w|_v| &= \left| \sum_{a\in\mathcal{A}} \sum_{k\in K_a} \alpha(a,k) \left(\sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right) \right| \\
&\leq \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{T^k[a]}(T^\ell x) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{T^k[a]}(T^\ell x) \right|
\end{aligned}
$$

$$
= \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \left| \sum_{\ell=i}^{i+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) - \sum_{\ell=j}^{j+n-1-|v|} \chi_{[a]}(T^\ell(T^{-k}x)) \right|
$$

$$
= \sum_{a\in\mathcal{A}} \sum_{k\in K_a} |\alpha(a,k)| \cdot \underbrace{\left| |(T^{-k}x)_{[i,i+n-|v|)}|_a - |(T^{-k}x)_{[j,j+n-|v|)}|_a \right|}_{\leq C}.
$$

Then, $||u|_v - |w|_v| \leq |\mathcal{A}|KC$, $K = \max_{a\in\mathcal{A}} \left\{ \sum_{k\in K_a} |\alpha(a,k)| \right\}$.

Balance

- Let $\mathcal{A} = \{1, \cdots, d\}$. The set of Elementary Arnoux-Rauzy substitutions defined on $\mathcal{A}$ is $\{\sigma_i : i \in \mathcal{A}\}$ given by

$$\sigma_i : \quad i \mapsto i; \quad j \mapsto ji \quad \mathit{for} \quad i \neq j.$$

Balance

- Let $\mathcal{A} = \{1, \cdots, d\}$. The set of Elementary Arnoux-Rauzy substitutions defined on $\mathcal{A}$ is $\{\sigma_i : i \in \mathcal{A}\}$ given by

$$\sigma_i : \quad i \mapsto i; \quad j \mapsto ji \quad \text{for} \quad i \neq j.$$

- An Arnoux-Rauzy substitution is a finite product of σ_i's.

Balance

- Let $\mathcal{A} = \{1, \cdots, d\}$. The set of Elementary Arnoux-Rauzy substitutions defined on $\mathcal{A}$ is $\{\sigma_i : i \in \mathcal{A}\}$ given by

$$\sigma_i : \quad i \mapsto i; \quad j \mapsto ji \quad \textit{for} \quad i \neq j.$$

- An Arnoux-Rauzy substitution is a finite product of σ_i's.
- A substitution σ on $\mathcal{A}$ is said to be primitive if there exists a positive integer N such that for every $a, b \in \mathcal{A}$, b appears in $\sigma^N(a)$.
- Words produced by primitive Arnoux-Rauzy substitutions are known to be balanced on the letters (they are Pisot substitutions).

Balance

- Let $\mathcal{A} = \{1, \cdots, d\}$. The set of Elementary Arnoux-Rauzy substitutions defined on $\mathcal{A}$ is $\{\sigma_i : i \in \mathcal{A}\}$ given by

$$\sigma_i : \quad i \mapsto i; \quad j \mapsto ji \quad \textit{for} \quad i \neq j.$$

- An Arnoux-Rauzy substitution is a finite product of σ_i's.
- A substitution σ on $\mathcal{A}$ is said to be primitive if there exists a positive integer N such that for every $a, b \in \mathcal{A}$, b appears in $\sigma^N(a)$.
- Words produced by primitive Arnoux-Rauzy substitutions are known to be balanced on the letters (they are Pisot substitutions).
- We know automatically that they are balanced on all factors.

Balance

- Let $\mathcal{A} = \{1, \cdots, d\}$. The set of Elementary Arnoux-Rauzy substitutions defined on $\mathcal{A}$ is $\{\sigma_i : i \in \mathcal{A}\}$ given by

$$\sigma_i : \quad i \mapsto i; \quad j \mapsto ji \quad \textit{for} \quad i \neq j.$$

- An Arnoux-Rauzy substitution is a finite product of σ_i's.
- A substitution σ on $\mathcal{A}$ is said to be primitive if there exists a positive integer N such that for every $a, b \in \mathcal{A}$, b appears in $\sigma^N(a)$.
- Words produced by primitive Arnoux-Rauzy substitutions are known to be balanced on the letters (they are Pisot substitutions).
- We know automatically that they are balanced on all factors.
- Another sufficient conditions exist to guarrantee balance on letters for Arnoux-Rauzy words.

Another application: Invariant measures and Orbit equivalence.

Invariant measures

- A probability measure μ on the compact metric space X is T-invariant if for all Borel subset B of X,

$$\mu(T^{-1}(B)) = \mu(B).$$

Invariant measures

- A probability measure μ on the compact metric space X is T-invariant if for all Borel subset B of X,

$$\mu(T^{-1}(B)) = \mu(B).$$

- Every dynamical system has invariant measures (Krylov-Bogolyubov's theorem).

Invariant measures

- A probability measure μ on the compact metric space X is T-invariant if for all Borel subset B of X,

$$\mu(T^{-1}(B)) = \mu(B).$$

- Every dynamical system has invariant measures (Krylov-Bogolyubov's theorem).
- (X, T) is said to be uniquely ergodic if there exists only one invariant measure.

Invariant measures

- A probability measure μ on the compact metric space X is T-invariant if for all Borel subset B of X,

$$\mu(T^{-1}(B)) = \mu(B).$$

- Every dynamical system has invariant measures (Krylov-Bogolyubov's theorem).
- (X, T) is said to be uniquely ergodic if there exists only one invariant measure.
- For minimal systems, unique ergodicity is equivalent to having frequencies.

Invariant measures

- Let $\mathcal{M}(X, T)$ the set of $T-$invariant measures on X. The Image subgroup of (X, T) is the following subgroup of $\mathbb{R}$.

$$I(X, T) = \bigcap_{\mu \in \mathcal{M}(X,T)} \left\{ \int f d\mu : f \in C(X, \mathbb{Z}) \right\}.$$

Invariant measures

- Let $\mathcal{M}(X,T)$ the set of $T-$invariant measures on X. The Image subgroup of (X,T) is the following subgroup of $\mathbb{R}$.

$$I(X,T) = \bigcap_{\mu\in\mathcal{M}(X,T)} \left\{ \int f d\mu : f \in C(X,\mathbb{Z}) \right\}.$$

- One can show that

$$I(X,T) = \bigcap_{\mu\in\mathcal{M}(X,T)} \langle \mu([w]) : w \in \mathcal{L}_X \rangle .$$

Invariant measures

- Let $\mathcal{M}(X,T)$ the set of $T-$invariant measures on X. The Image subgroup of (X,T) is the following subgroup of $\mathbb{R}$.

$$I(X,T) = \bigcap_{\mu\in\mathcal{M}(X,T)} \left\{ \int f d\mu : f \in C(X,\mathbb{Z}) \right\}.$$

- One can show that

$$I(X,T) = \bigcap_{\mu\in\mathcal{M}(X,T)} \langle \mu([w]) : w \in \mathcal{L}_X \rangle .$$

- If (X,T) is uniquely ergodic with measure μ,

$$I(X,T) = \langle \mu([w]) : w \in \mathcal{L}_X \rangle .$$

Invariant measures

- Let $\mathcal{M}(X, T)$ the set of $T-$invariant measures on X. The Image subgroup of (X, T) is the following subgroup of $\mathbb{R}$.

$$I(X, T) = \bigcap_{\mu \in \mathcal{M}(X,T)} \left\{ \int f d\mu : f \in C(X, \mathbb{Z}) \right\}.$$

- One can show that

$$I(X, T) = \bigcap_{\mu \in \mathcal{M}(X,T)} \langle \mu([w]) : w \in \mathcal{L}_X \rangle .$$

- If (X, T) is uniquely ergodic with measure μ,

$$I(X, T) = \langle \mu([w]) : w \in \mathcal{L}_X \rangle .$$

Invariant measures

- The triple $(I(X,T), I(X,T)\cap\mathbb{R}^+, 1)$ is an order group with unit.

Invariant measures

- The triple $(I(X,T), I(X,T)\cap\mathbb{R}^+, 1)$ is an order group with unit.
- $(I(X,T), I(X,T)\cap\mathbb{R}^+, 1)$ is total invariant for Orbit equivalence [Giordano-Putnam-Skau95].
- (X,T) and (Y,S) are orbit equivalent if there is a homeomorphism $h : X \to Y$ such that for all $x \in X$

$$h(\{T^n(x) : n \in \mathbb{Z}\}) = \{S^n(h(x)) : n \in \mathbb{Z}\},$$

that is, h sends orbits onto orbits.

Invariant measures

Theorem

Let (X, T) a uniquely ergodic dendric subshift over the alphabet $\mathcal{A}$ and μ its unique invariant measure. Then, the image subgroup of (X, T) is

$$I(X, T) = \sum_{a \in \mathcal{A}} \mathbb{Z}\mu([a]).$$

Invariant measures

Theorem

Let (X, T) a uniquely ergodic dendric subshift over the alphabet $\mathcal{A}$ and μ its unique invariant measure. Then, the image subgroup of (X, T) is

$$I(X, T) = \sum_{a \in \mathcal{A}} \mathbb{Z}\mu([a]).$$

Corollary

Two minimal uniquely ergodic dendric subshifts are orbit equivalent if and only if they have the same additive group of letter frequencies.

Invariant measures

Theorem

Let (X, T) a uniquely ergodic dendric subshift over the alphabet $\mathcal{A}$ and μ its unique invariant measure. Then, the image subgroup of (X, T) is

$$I(X, T) = \sum_{a \in \mathcal{A}} \mathbb{Z}\mu([a]).$$

Corollary

Two minimal uniquely ergodic dendric subshifts are orbit equivalent if and only if they have the same additive group of letter frequencies.

Corollary

All dendric subshifts over a three-letter alphabet with the same letter frequencies are orbit equivalent.