STURMIAN WORDS: DYNAMICAL SYSTEMS AND DERIVATED WORDS

Isabel M. Araújo and Véronique Bruyère

Centro de Álgebra da Universidade de Lisboa

Université de Mons–Hainaut

DLT'05 — Palermo, 5th July 2005

Characteristic sturmian words

$\alpha \in]0,1[\setminus \mathbb{Q}$, with continued fraction $[0, 1 + a_1, a_2, \ldots]$

Define $(t_n)_n$

$$t_0 = 0,\ t_1 = 0^{a_1}1,\ t_n = t_{n-1}^{a_n} t_{n-2}\ (n \geqslant 2)$$

Characteristic Sturmian words

$\alpha \in]0,1[\setminus \mathbb{Q}$, with continued fraction $[0, 1 + a_1, a_2, \ldots]$

Define $(t_n)_n$

$$t_0 = 0,\ t_1 = 0^{a_1} 1,\ t_n = t_{n-1}^{a_n} t_{n-2}\ (n \geqslant 2)$$

Def. The infinite word

$$c_\alpha = \lim_{n \to \infty} t_n$$

is called *characteristic Sturmian word* of *slope* α.

CHARACTERISTIC STURMIAN WORDS

$\alpha \in]0,1[\setminus\mathbb{Q}$, with continued fraction $[0, 1+a_1, a_2, \ldots]$

Define $(t_n)_n$

$$t_0 = 0,\ t_1 = 0^{a_1}1,\ t_n = t_{n-1}^{a_n} t_{n-2}\ (n \geqslant 2)$$

Def. The infinite word

$$c_\alpha = \lim_{n\to\infty} t_n$$

is called *characteristic Sturmian word* of *slope* α.

Associated lengths: $q_n = |t_n|$

$$q_0 = 1,\ q_1 = a_1 + 1,\ q_n = a_n q_{n-1} + q_{n-2}\ (n \geqslant 2)$$

Characteristic Sturmian words

$\alpha \in]0,1[\setminus\mathbb{Q}$, with continued fraction $[0, 1+a_1, a_2, \ldots]$

Define $(t_n)_n$

$$t_0 = 0,\ t_1 = 0^{a_1}1,\ t_n = t_{n-1}^{a_n} t_{n-2}\ (n \geqslant 2)$$

Def. The infinite word

$$c_\alpha = \lim_{n\to\infty} t_n$$

is called *characteristic Sturmian word* of *slope* α.

Associated lengths: $q_n = |t_n|$

$$q_0 = 1,\ q_1 = a_1 + 1,\ q_n = a_n q_{n-1} + q_{n-2}\ (n \geqslant 2)$$

- The Fibonacci word is a characteristic Sturmian word.

STURMIAN WORDS

Consider the *dynamical system* generated by c_α

$$S_\alpha = \{x \in \{0,1\}^\omega \mid \mathrm{Fact}(x) = \mathrm{Fact}(c_\alpha)\}$$

STURMIAN WORDS

Consider the *dynamical system* generated by c_α

$$S_\alpha = \{x \in \{0,1\}^\omega \mid \mathrm{Fact}(x) = \mathrm{Fact}(c_\alpha)\}$$

Def. An element of S_α is called *Sturmian word of slope* α.

Sturmian words

Consider the *dynamical system* generated by c_α

$$S_\alpha = \{x \in \{0,1\}^\omega \mid \mathrm{Fact}(x) = \mathrm{Fact}(c_\alpha)\}$$

Def. An element of S_α is called *Sturmian word of slope* α.

- S_α is also the topological closure of

$$R_\alpha = \{\sigma^i(c_\alpha) \mid i \geqslant 0\}$$

where σ is the shift, and the closure of R_α is

$$\{x \in \{0,1\}^\omega \mid \mathrm{Pref}(x) \subseteq \mathrm{Pref}(R_\alpha)\}.$$

EXAMPLE

- c_α the characteristic Sturmian word of slope $\alpha = [0, \overline{3, 2}]$

$$c_\alpha = 0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1 \cdots$$

EXAMPLE

- c_α the characteristic Sturmian word of slope $\alpha = [0, \overline{3, 2}]$

  $$c_\alpha = 0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1 \cdots$$

- The first elements of the characteristic sequence are

  $$t_0 = 0,\ t_1 = 0\,0\,1,\ t_2 = 0\,0\,1\,0\,0\,1\,0$$

EXAMPLE

- c_α the characteristic Sturmian word of slope $\alpha = [0, \overline{3, 2}]$

$c_\alpha = 0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1 \cdots$

- The first elements of the characteristic sequence are

$$t_0 = 0,\ t_1 = 0\,0\,1,\ t_2 = 0\,0\,1\,0\,0\,1\,0$$

- The set R_α is formed by all suffixes of c_α

Example

- c_α the characteristic Sturmian word of slope $\alpha = [0, \overline{3, 2}]$

$$c_\alpha = 0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1 \cdots$$

- The first elements of the characteristic sequence are

$$t_0 = 0,\ t_1 = 0\,0\,1,\ t_2 = 0\,0\,1\,0\,0\,1\,0$$

- The set R_α is formed by all suffixes of c_α
- $1c_\alpha,\ 0c_\alpha \in S_\alpha \backslash R_\alpha$

Return words

$x = x_0 x_1 \cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_i x_{i+1} \cdots x_{i+|w|-1} = w.$$

Return words

$x = x_0 x_1 \cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_i x_{i+1} \cdots x_{i+|w|-1} = w.$$

Return words

$x = x_0x_1\cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_ix_{i+1}\cdots x_{i+|w|-1} = w.$$

Def. [Durand] For i, j, consecutive occurrences of w in x,

$$x_ix_{i+1}\cdots x_{j-1}$$

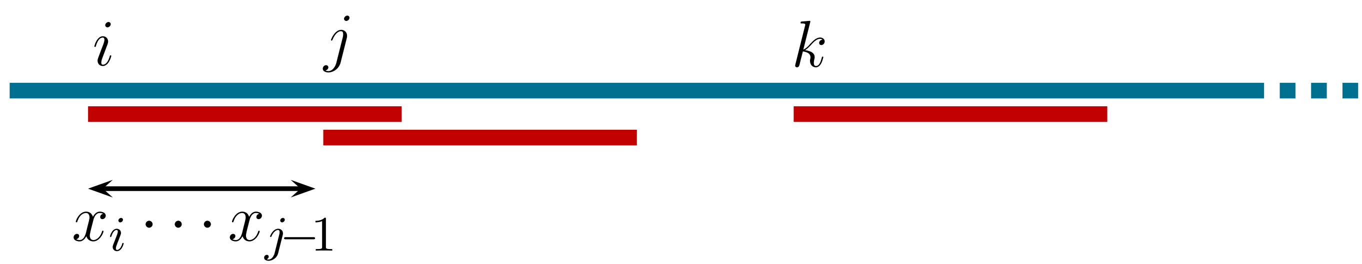
is a *return word* of w in x.

Return words

$x = x_0x_1\cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_ix_{i+1}\cdots x_{i+|w|-1} = w.$$



Def. [Durand] For i, j, consecutive occurrences of w in x,

$$x_ix_{i+1}\cdots x_{j-1}$$

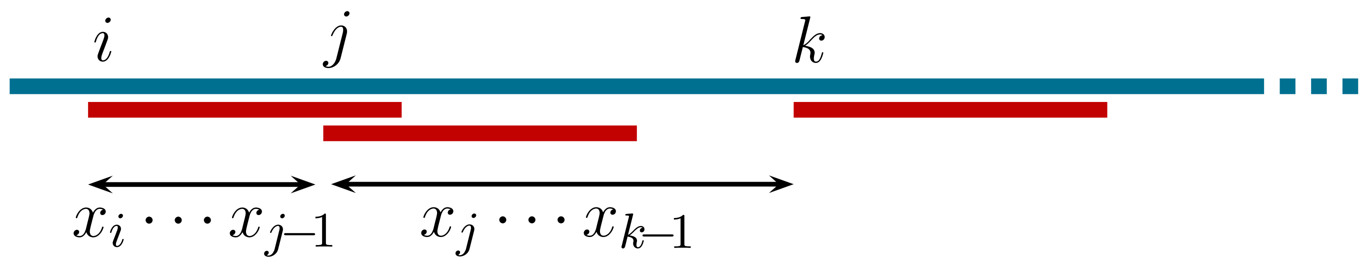
is a *return word* of w in x.

Return words

$x = x_0 x_1 \cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_i x_{i+1} \cdots x_{i+|w|-1} = w.$$



Def. [Durand] For i, j, consecutive occurrences of w in x,

$$x_i x_{i+1} \cdots x_{j-1}$$

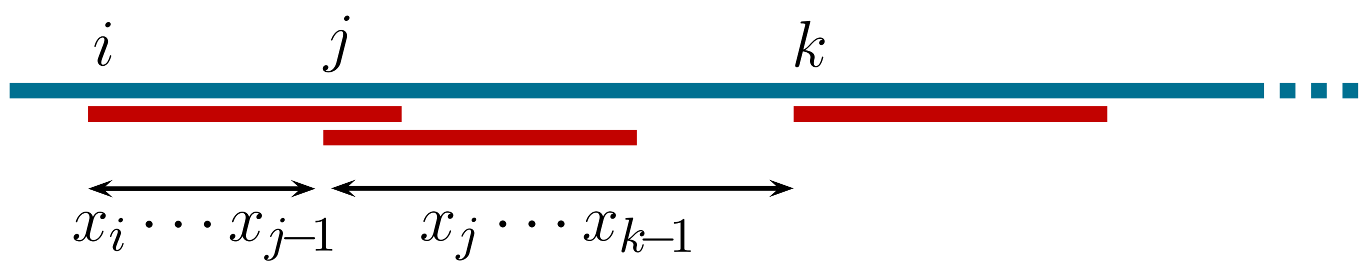
is a *return word* of w in x.

RETURN WORDS

$x = x_0 x_1 \cdots$ a Sturmian word; w a factor of x

An integer i is an *occurrence* of w in x if

$$x_i x_{i+1} \cdots x_{i+|w|-1} = w.$$



Def. [Durand] For i, j, consecutive occurrences of w in x,

$$x_i x_{i+1} \cdots x_{j-1}$$

is a *return word* of w in x.

- [Vuillon] An infinite binary word x is Sturmian *if and only if* each non empty factor of x has exactly two return words.

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$

Example

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\cdots$$

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,\boxed{0\,0\,1}\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0 \cdots$$

- $w = 001$ is non-empty factor of x

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\,0\,0\,1\,0\,0\,0\,1\,0\,0\,1\,0\,0\,1\,0\,0\,0\cdots$$

$\uparrow$ 1 $\uparrow$ 4 $\uparrow$ 8 $\uparrow$ 11 $\uparrow$ 15

- $w = 001$ is non-empty factor of x
- the occurrences of w in x: $1, 4, 8, 11, 15 \ldots$

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,\underline{001}\,\underline{001}\,0\,\underline{001}\,\underline{001}\,0\,\underline{001}\,001\,0\,001\,001\,001\,000\cdots$$

$$\uparrow_{1} \quad \uparrow_{4} \quad \uparrow_{8} \quad \uparrow_{11} \quad \uparrow_{15}$$

- $w = 001$ is non-empty factor of x
- the occurrences of w in x: $1, 4, 8, 11, 15 \ldots$
- the return words of w in x are

001 and 0010.

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,001\,0010\,001\,001\,0\,001\,001\,0001\,001\,001\,000\cdots$$

$$\uparrow 1 \quad \uparrow 4 \quad \uparrow 8 \quad \uparrow 11 \quad \uparrow 15$$

- $w = 001$ is non-empty factor of x
- the occurrences of w in x: $1, 4, 8, 11, 15 \ldots$
- the return words of w in x are

001 and 0010.

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,001\,0010\,001\,001\,0\,001\,001\,0001\,001\,001\,000\cdots$$

$$\uparrow_{1} \quad \uparrow_{4} \quad \uparrow_{8} \quad \uparrow_{11} \quad \uparrow_{15}$$

- $w = 001$ is non-empty factor of x
- the occurrences of w in x: $1, 4, 8, 11, 15\ldots$
- the return words of w in x are

001 and 0010.

EXAMPLE

- c_α the characteristic word of slope $\alpha = [0, \overline{3, 2}]$
- $0c_\alpha \in S_\alpha$

$$0c_\alpha = 0\,001\,0010\,001\,0010\,0010010001001001000\cdots$$

$$\uparrow 1 \quad \uparrow 4 \quad \uparrow 8 \quad \uparrow 11 \quad \uparrow 15$$

- $w = 001$ is non-empty factor of x
- the occurrences of w in x: $1, 4, 8, 11, 15 \ldots$
- the return words of w in x are

001 and 0010.

RETURN WORDS OF STURMIAN WORDS

$x \in S_\alpha$, w a factor of x

- w is a *left special factor* of x if $0w, 1w$ are factors of x

Return words of Sturmian words

$x \in S_\alpha$, w a factor of x

- w is a *left special factor* of x if $0w, 1w$ are factors of x
- $\mathrm{Left}(x)$ the set of left special factors of x

Return words of Sturmian words

$x \in S_\alpha$, w a factor of x

- w is a *left special factor* of x if $0w, 1w$ are factors of x
- $\mathrm{Left}(x)$ the set of left special factors of x

- $\mathrm{Left}(x) = \mathrm{Pref}(c_\alpha)$

Return words of Sturmian words

$x \in S_\alpha$, w a factor of x

- w is a *left special factor* of x if $0w, 1w$ are factors of x
- $\mathrm{Left}(x)$ the set of left special factors of x

- $\mathrm{Left}(x) = \mathrm{Pref}(c_\alpha)$

Prop. Let $w \in \mathrm{Left}(x)$.
The two return words of w are

$$t_n \text{ and } t_n^i t_{n-1},$$

where n and i are determined by the length of w.

DERIVATED WORDS

$x \in S_\alpha$; $w \in \mathrm{Left}(x)$; u, v return words

$x = px'$, where $|\, p \,|$ is the first occurrence of w in x

Derivated words

$x \in S_\alpha$; $w \in \mathrm{Left}(x)$; u, v return words
$x = px'$, where $|\, p \,|$ is the first occurrence of w in x

- x' is written in a unique way as a concatenation of u, v:

$$x = p\, x' = p\ z_1 z_2 \cdots \quad (z_i \in \{u, v\}).$$

EXAMPLE

- $x = 0c_\alpha$ with $\alpha = [0, \overline{3, 2}]$
- $w = 001 \in \mathrm{Left}(0c_\alpha)$

$$x = 0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,\underline{0\,0\,1}\,0\,\underline{0\,0\,1}\cdots$$

[$\longrightarrow$]

EXAMPLE

- $x = 0c_\alpha$ with $\alpha = [0, \overline{3, 2}]$
- $w = 001 \in \mathrm{Left}(0c_\alpha)$, $p = 0$, $x' = c_\alpha$

$$x = 0\,001\,0010\,001\,0010\,001\,0010\,001\,001\,0010\,001\cdots$$

[$\longrightarrow$]

EXAMPLE

- $x = 0c_\alpha$ with $\alpha = [0, \overline{3, 2}]$
- $w = 001 \in \mathrm{Left}(0c_\alpha)$, $p = 0$, $x' = c_\alpha$

$$x = 0\,001\,0010\,001\,0010\,001\,0010\,001\,001\,0010\,001\cdots$$

[⟶]

Example

- $x = 0c_\alpha$ with $\alpha = [0, \overline{3, 2}]$
- $w = 001 \in \mathrm{Left}(0c_\alpha)$, $p = 0$, $x' = c_\alpha$

$$x = 0\,001\,0010\,001\,0010\,001\,0010\,001\,001\,0010\,001\cdots$$

Δ

$$001 \mapsto 0$$

$$0010 \mapsto 1$$

$$D_w(x) = 0\;1\;0\;1\;0\;1\;0\;0\;1\;\cdots$$

[$\longrightarrow$]

EXAMPLE

- $x = 0c_\alpha$ with $\alpha = [0, \overline{3, 2}]$
- $w = 001 \in \mathrm{Left}(0c_\alpha)$, $p = 0$, $x' = c_\alpha$

$$x = 0\,001\,0010\,001\,0010\,001\,0010\,001\,001\,0010\,001\cdots$$

Δ: $001 \mapsto 0$, $0010 \mapsto 1$

Θ: $0 \mapsto 001$, $1 \mapsto 0010$

$$D_w(x) = 0\,1\,0\,1\,0\,1\,0\,0\,1\ \cdots$$

- x' is obtained from this "*derivated word*" by applying the morphism $0 \mapsto 001$, $1 \mapsto 0010$ to it.

[$\longrightarrow$]

DERIVATED WORDS

$x \in S_\alpha$; $w \in \mathrm{Left}(x)$; u, v return words
$x = px'$, where $|\, p \,|$ is the first occurrence of w in x

- x' is written in a unique way as a concatenation of u, v:

$$x = p\, x' = p\ z_1 z_2 \cdots \quad (z_i \in \{u, v\}).$$

DERIVATED WORDS

$x \in S_\alpha$; $w \in \mathrm{Left}(x)$; u, v return words
$x = px'$, where $|\, p \,|$ is the first occurrence of w in x

- x' is written in a unique way as a concatenation of u, v:

$$x = p\,x' = p\; z_1 z_2 \cdots \quad (z_i \in \{u, v\}).$$

Consider a bijection $\Delta : \{u, v\} \longrightarrow \{0, 1\}$.

Derivated words

$x \in S_\alpha$; $w \in \mathrm{Left}(x)$; u, v return words
$x = px'$, where $|\, p \,|$ is the first occurrence of w in x

- x' is written in a unique way as a concatenation of u, v:

$$x = p\, x' = p\; z_1 z_2 \cdots \quad (z_i \in \{u, v\}).$$

Consider a bijection $\Delta : \{u, v\} \longrightarrow \{0, 1\}$.

Def. [Durand] The *derivated word* of x with respect to w is

$$D_w(x) = \Delta(z_1)\Delta(z_2) \cdots \; .$$

[$\longleftarrow$]

Derivated words (2)

Consider also the inverse bijection $\Theta : \{0, 1\} \to \{u, v\}$.

DERIVATED WORDS (2)

Consider also the inverse bijection $\Theta : \{0,1\} \rightarrow \{u,v\}$.

- Θ can be extended to a morphism

Derivated words (2)

Consider also the inverse bijection $\Theta : \{0, 1\} \to \{u, v\}$.

- Θ can be extended to a morphism
- Θ is the “decoding” morphism that recovers x' from $D_w(x)$

Derivated words (2)

Consider also the inverse bijection $\Theta : \{0, 1\} \to \{u, v\}$.

- Θ can be extended to a morphism
- Θ is the "decoding" morphism that recovers x' from $D_w(x)$

Proposition. The derivated word $D_w(c_\alpha)$ of a characteristic Sturmian word c_α is a characteristic Sturmian word c_δ. The slope δ can be determined from α and w.

THEOREMS

An infinite word w is *substitutive* if it is the image by a literal morphism of a *morphic word*.
A morphism is *Sturmian* if the image of a Sturmian word is a Sturmian word.

Theorem. [Crisp *et al.*] [Araújo et Bruyère]
The following are equivalent

- α is a Sturm number;
- $D_w(c_\alpha) = c_\alpha$, for some $w \in \mathrm{Pref}(c_\alpha)$;
- c_α is a fixed point of a morphism.

THEOREMS

An infinite word w is *substitutive* if it is the image by a literal morphism of a *morphic word*.
A morphism is *Sturmian* if the image of a Sturmian word is a Sturmian word.

Theorem. [Araújo et Bruyère]
The following are equivalent

- α is a Sturm number;
- $D_w(x)$ has slope α, for some $w \in \mathrm{Left}(x)$;
- $\varphi(x)$ has slope α, for some Sturmian morphism φ;

for $x \in S_\alpha$.

THEOREMS (2)

Theorem. The following are equivalent

- the continued fraction of α is ultimately periodic;
- the set of all derivated words of c_α is finite;
- c_α is substitutive.

THEOREMS (2)

Theorem. The following are equivalent

- the continued fraction of α is ultimately periodic;
- there exist a Sturmian word y, and words x, x', such that $D_w(x)$, y and $D_{w'}(y)$ have the same slope;
- there exist a Sturmian word y, and two Sturmian morphisms ψ and φ, such that

  x and $\psi(y)$ have the same slope,

  y and $\varphi(y)$ have the same slope;

for $x \in S_\alpha$.

QUESTION

Theorem. For $x \in R_\alpha$, the following are equivalent

- the continued fraction of α is ultimately periodic;
- the set of all derivated words of x is finite.

QUESTION

Theorem. For $x \in R_\alpha$, the following are equivalent

- the continued fraction of α is ultimately periodic;
- the set of all derivated words of x is finite.

This equivalence is no longer valid if $x \in S_\alpha \setminus R_\alpha$ (counter-example built using [Bernardi, Guerziz and Koskas]).

QUESTION

Theorem. For $x \in R_\alpha$, the following are equivalent

- the continued fraction of α is ultimately periodic;
- the set of all derivated words of x is finite.

This equivalence is no longer valid if $x \in S_\alpha \backslash R_\alpha$ (counter-example built using [Bernardi, Guerziz and Koskas]).

Question. Under which conditions is the set of derivated words of a Sturmian word finite?

Bibliography

- I.M. Araújo and V. Bruyère, Sturmian words and a criterium by Michaux–Villemaire. *Theoret. Comput. Sci.* 339:88–102, 2005.
- I.M. Araújo and V. Bruyère, Words derivated from Sturmian words. *Theoret. Comput. Sci.*, to appear.
- D. Bernardi, A. Guerziz and M. Koskas, Sturmian words: description and orbits. *Preprint*, 2004.
- F. Durand, A characterisation of substitutive sequences using return words. *Discrete Mathematics*, 179:89–101, 1998.
- L. Vuillon A characterisation of Sturmian words by return words. *European J. Combin.*, 22:263–275, 2001.

Dynamical systems and Sturmian words

- c_α the characteristic word
- $R_\alpha = \{\sigma^i(c_\alpha) \mid i \geqslant 1\}$ the suffixes of c_α
- S_α the closure of R_α

Words in S_α:

- same set of factors
- same return words with respect to a given factor w
- $\mathrm{Left}(x) = \mathrm{Pref}(c_\alpha)$