S-adic sequences
A bridge between dynamics, arithmetic, and geometry

J. M. Thuswaldner

(joint work with P. Arnoux, V. Berthé, M. Minervino, and W. Steiner)

Marseille, November 2017

PART 1

Sturmian sequences and rotations

Contents

The underlying papers

- Arnoux, P., Sturmian sequences. Substitutions in dynamics, arithmetics and combinatorics, 143–198, Lecture Notes in Math., 1794, Springer, Berlin, 2002.
- Arnoux, P. and Fisher, A. M., The scenery flow for geometric structures on the torus: the linear setting. Chinese Ann. Math. Ser. B 22 (2001), no. 4, 427–470.

Words and sequences

- Alphabet: $\mathcal{A} = \{1, \ldots, d\}$ (set of symbols)
- Set of words: $\mathcal{A}^*$ (finite block of symbols)
- Set of sequences: $\mathcal{A}^{\mathbb{N}}$ (right infinite)

Example

Two symbol alphabet $\mathcal{A} = \{1, 2\}$

$v = 121121$ is a word (over $\mathcal{A}$)

$w = 1212121212\ldots$ is a (right infinite) sequence (over $\mathcal{A}$)

A symbol is also called letter.

Complexity

Let $v = v_0 \dots v_{n-1} \in \mathcal{A}^*$.

$|v| = n$ is the length of v.

Definition (Complexity function)

Let $w = w_0 w_1 \dots \in \mathcal{A}^{\mathbb{N}}$ be a sequence.

The complexity function

$$\begin{aligned} p_w &: \mathbb{N} \to \mathbb{N} \\ & n \mapsto \#\{v \ : \ v \text{ subword of } w, \ |v| = n\} \end{aligned}$$

This function assigns to each integer n the number of subwords of length n occurring in w.

Periodicity

Definition (Ultimate periodicity)

$w \in \mathcal{A}^{\mathbb{N}}$ is ultimately periodic if there exist $k > 0$ and $N > 0$ with $w_n = w_{n+k}$ for each $n \geq N$.

Example

$w = 1121212121212\ldots$ is ultimately periodic

- Subwords of length 1: $\{1, 2\}$
- Subwords of length 2: $\{11, 12, 21\}$
- Subwords of length 3: $\{112, 121, 212\}$
- Subwords of length 4: $\{1121, 1212, 2121\}$
- $\ldots$

Thus $p_w(1) = 2$ and $p_w(n) = 3$ for $n \geq 3$ and p_w is bounded.

Definition of Sturmian sequences

Lemma (Coven and Hedlund, 1973)

$w \in \{1, 2, \ldots, d\}^{\mathbb{N}}$ admits the inequality

$$p_w(n) \leq n$$

for a single choice of n if and only if it is ultimately periodic.

Not ultimately periodic sequences with smallest complexity function:

Definition (Sturmian Sequence)

A sequence $w \in \{1, 2\}^{\mathbb{N}}$ is called a Sturmian sequence if its complexity function satisfies $p_w(n) = n + 1$ for all $n \in \mathbb{N}$.

It is not clear that Sturmian sequences exist!

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 1211212112112121121211212121121 \ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

$\sigma^0(1) = 1$

- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 12112121121121211212112112121121\ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^1(1) = 12$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 1211212112112121121211212121121 \ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^2(1) = 121$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 1211212112112121121211212121121\ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^3(1) = 12112$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 121121211211212112121121211212112 \dots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^4(1) = 12112121$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 12112121121121211212112112121121\ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^5(1) = 1211212112112$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 12112121121121211212112112121121\ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^6(1) = 121121211211212112121$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 1211212112112121121211211212121121\dots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

An example

- Let $\mathcal{A} = \{1, 2\}$ be the alphabet.
- Define a substitution $\sigma : \mathcal{A} \to \mathcal{A}^*$ by

$$\sigma : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 1. \end{array}$$

  This is the Fibonacci substitution.
- We can extend the domain of σ to $\mathcal{A}^*$ and $\mathcal{A}^{\mathbb{N}}$ by concatenation.
- Since $\sigma(1)$ starts with 1 the iteration $\sigma^n(1)$ "converges":

  $\sigma^7(1) = 1211212112112121121211211212112112$
- The sequence

$$\lim_{n\to\infty} \sigma^n(1) = 12112121121121211212112112121121\ldots$$

  is the famous Fibonacci sequence.
- One can prove that this is a Sturmian sequence.

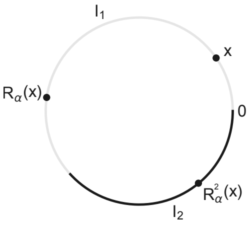
Natural codings of rotations

- Rotation by α: $R_\alpha : \mathbb{T} \to \mathbb{T}$ with $x \mapsto x + \alpha \pmod 1$.
- R_α can be regarded as a two interval exchange of the intervals $I_1 = [0, 1 - \alpha)$ and $I_2 = [1 - \alpha, 1)$.
- $w = w_1 w_2 \ldots \in \{1, 2\}^{\mathbb{N}}$ is a natural coding of R_α if there is $x \in \mathbb{T}$ such that $w_k = i$ if and only if $R_\alpha^k(x) \in I_i$ for each $k \in \mathbb{N}$.

Lemma

If $w \in \{1, 2\}^{\mathbb{N}}$ is a natural coding of R_α with $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ then w is Sturmian.

An image of the rotation



Natural coding: 112...

Figure: Two iterations of the irrational rotation R_α on the circle $\mathbb{T}$.

The proof of the lemma just follows from the equivalence

$$v_1 \dots v_n \text{ is a factor of a nat'l coding of } R_\alpha \quad \Longleftrightarrow \quad \bigcap_{j=0}^{n} R_\alpha^{-j} I_{v_j} \neq \emptyset.$$

What about the converse of the lemma?

Recurrence

Definition (Recurrence)

A sequence $w \in \{1, 2\}^{\mathbb{N}}$ is called recurrent if each subword of w occurs infinitely often in w.

Lemma

A Sturmian sequence w is recurrent.

Proof.

- Suppose for some v, $|v| = n$ this is wrong.
- Then v doesn't occur in some shift $w' = \Sigma^k w$.
- Then $p_{w'}(n) \leq n$ and w' is ultimately periodic.
- Thus w is ultimately periodic as well, a contradiction.

□

Balance

- $v \in \{1,2\}^*$
- $|v|_i$ is the number of occurrences of the letter i in v.

Definition (Balanced sequence)

A sequence $w \in \{1,2\}^{\mathbb{N}}$ is called balanced if each pair of subwords (v, v') of w with $|v| = |v'|$ satisfies $\big||v|_1 - |v'|_1\big| \leq 1$.

Lemma (Morse and Hedlund, 1940)

Let $w \in \{1,2\}^{\mathbb{N}}$ be given. Then w is a Sturmian sequence if and only if w is not ultimately periodic and balanced.

The proof is combinatorial and a bit tricky.

Related substitutions

Next Goal

Use balance to code a Sturmian word by the Sturmian substitutions

$$\sigma_1 : \begin{array}{l} 1 \mapsto 1, \\ 2 \mapsto 21, \end{array} \qquad \sigma_2 : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 2. \end{array}$$

The domain of these substitutions can naturally be extended from $\{1,2\}$ to $\{1,2\}^*$ and $\{1,2\}^{\mathbb{N}}$, *e.g.*,

$$\sigma_1(1211) = \sigma_1(1)\sigma_1(2)\sigma_1(1)\sigma_1(1) = 12111$$

$$\sigma_2(121\ldots) = \sigma_2(1)\sigma_2(2)\sigma_2(1)\ldots = 12212\ldots$$

Two types of Sturmian sequences

- $w = w_0 w_1 \ldots \in \{1,2\}^{\mathbb{N}}$ a given Sturmian sequence.
- w contains exactly three of the subwords $11, 12, 21, 22$.
- Type 1: w contains $11, 12, 21$
- Type 2: w contains $12, 21, 22$

Desubstitution

Recall: $\sigma_1(1) = 1$, $\sigma_1(2) = 21$
$w = 12112121121121\ldots$ Sturmian sequence of Type 1.

$$w = \underbrace{1}_{\sigma_1(1)}\underbrace{21}_{\sigma_1(2)}\underbrace{1}_{\sigma_1(1)}\underbrace{21}_{\sigma_1(2)}\underbrace{21}_{\sigma_1(2)}\underbrace{1}_{\sigma_1(1)}\underbrace{21}_{\sigma_1(2)}\underbrace{1}_{\sigma_1(1)}\underbrace{21}_{\sigma_1(2)}\ldots$$

$$= \sigma_1(121221212\ldots)$$

Using balance one sees: $121221212\ldots$ is Sturmian again.

S-adic representations of Sturmian sequences

- This desubstitution process is (esssentially) unique.
- Problems can occur at the beginning (in this case an additional shift is needed).

Let w be a Sturmian sequence. Then there is a sequence $(w^{(n)})_{n\geq 0}$ of Sturmian sequences with (modulo shifts)

$$w = w^{(0)} \quad \text{and} \quad w^{(n)} = \sigma_{i_n}(w^{(n+1)}) \text{ for } n \geq 0.$$

Iterating this we see that

$$w = \sigma_{i_0} \circ \cdots \circ \sigma_{i_n}(w^{(n+1)}).$$

The coding sequence $(i_n) \in \{1,2\}^{\mathbb{N}}$ changes its value infinitely often (otherwise w would be ultimately constant).

Variants of S-adic representations

- $w^{(n)}$ begins with the same letter as $w = w^{(0)}$.
- The first letter of $w^{(n)}$ determines a prefix of w whose length tends to infinity with n.

Thus

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

We could also group the blocks of the sequence (i_n). This gives

$$w = \lim_{k\to\infty} \sigma_1^{a_0} \circ \sigma_2^{a_1} \circ \sigma_1^{a_2} \circ \cdots \circ \sigma_1^{a_{2k}}(a).$$

Limit: $\mathcal{A}^{\mathbb{N}}$ carries the product topology of the discrete topology on $\mathcal{A}$.

Characterization of Sturmian sequences

$$\sigma_1 : \begin{array}{l} 1 \mapsto 1, \\ 2 \mapsto 21, \end{array} \qquad \sigma_2 : \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 2. \end{array}$$

Lemma

Let σ_1, σ_2 be the Sturmian substitutions. Then for each Sturmian sequence w there exists a coding sequence $\boldsymbol{\sigma} = (\sigma_{i_n})$, where (i_n) takes each symbol in $\{1, 2\}$ an infinite number of times, such that w has the same language *as*

$$u = \lim_{n\to\infty} \sigma_{i_0} \circ \sigma_{i_1} \circ \cdots \circ \sigma_{i_n}(a).$$

Here $a \in \{1, 2\}$ can be chosen arbitrarily.

Abelianization and incidence matrices

Recall the definition of the Sturmian substitutions

$$\sigma_1: \begin{array}{l} 1 \mapsto 1, \\ 2 \mapsto 21, \end{array} \qquad \sigma_2: \begin{array}{l} 1 \mapsto 12, \\ 2 \mapsto 2. \end{array}$$

For a word $v \in \{1,2\}^*$ define the abelianization

$$\mathbf{l}v = (|v|_1, |v|_2)^t$$

and let $M_i = (|\sigma_i(k)|_j)_{1 \le j,k \le 2}$ be the incidence matrix of σ_i. Then

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

M_i is the abelianized version of σ_i in the sense that

$$\mathbf{l}\sigma_i(v) = M_i \mathbf{l}v.$$

Additive continued fraction algorithm

We start with the well-known additive Euclidean algorithm.

$$F : \mathbb{R}^2_{\geq 0} \setminus \{\mathbf{0}\} \to \mathbb{R}^2_{\geq 0} \setminus \{\mathbf{0}\}$$

$$F(a,b) = \begin{cases} (a-b,b), & \text{if } a > b, \\ (a,b-a), & \text{if } b \geq a. \end{cases}$$

Iterate F on $(a,b) \in \mathbb{R}^2_{>0}$.

- If $a/b \in \mathbb{Q}$ we reach a pair of $(0,c)$ or $(c,0)$ with $c > 0$.
- If $a/b \notin \mathbb{Q}$ we produce an infinite sequence of pairs of different strictly positive numbers.

Additive continued fraction expansion

Recall that

$$M_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \qquad \text{and} \qquad M_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}.$$

We see that

$$F(a,b) = \begin{cases} M_1^{-1}(a,b)^t, & \text{if } a > b, \\ M_2^{-1}(a,b)^t, & \text{if } a \le b. \end{cases}$$

Thus

$$\begin{aligned}(a,b)^t &= M_{i_0} F(a,b) \\ &= M_{i_0} M_{i_1} F^2(a,b) \\ &= M_{i_0} M_{i_1} M_{i_2} F^3(a,b) = \cdots .\end{aligned}$$

This sequence (M_{i_n}) is called the additive continued fraction expansion of (a,b).

Linear additive continued fractions

Let $\mathbb{P}$ be the projective line and

$$X = \{[a : b] \in \mathbb{P} \, : \, a \geq 0, \, b \geq 0\}.$$

Define $M : X \to \{M_1, M_2\}$ by

$$M([a : b]) = \begin{cases} M_1, & \text{if } a > b, \\ M_2, & \text{if } b \geq a. \end{cases}$$

Then the mapping

$$F : X \to X; \quad \mathbf{x} \mapsto M(\mathbf{x})^{-1}\mathbf{x}$$

is called the linear additive continued fraction mapping.

Projective additive continued fractions

Assume $a, b \neq 0$.
Then $[a : b] = [1, b/a]$ if $a > b$ and $[a : b] = [a/b, 1]$ if $a \geq b$.

$$F[1:c] = \begin{cases} [1-c:c] = [\frac{1-c}{c}:1], & \text{if } c > \frac{1}{2}, \\ [1-c:c] = [1:\frac{c}{1-c}] & \text{if } c \leq \frac{1}{2}, \end{cases}$$

$$F[c:1] = \begin{cases} [1:\frac{1-c}{c}], & \text{if } c > \frac{1}{2}, \\ [\frac{c}{1-c}:1] & \text{if } c \leq \frac{1}{2}. \end{cases}$$

Since the coordinate 1 contains no information this defines

$$f : (0,1) \to (0,1)$$

$$x \mapsto \begin{cases} \frac{1-x}{x}, & \text{if } x > \frac{1}{2}, \\ \frac{x}{1-x}, & \text{if } x \leq \frac{1}{2}. \end{cases}$$

called projective additive continued fraction mapping.

A picture

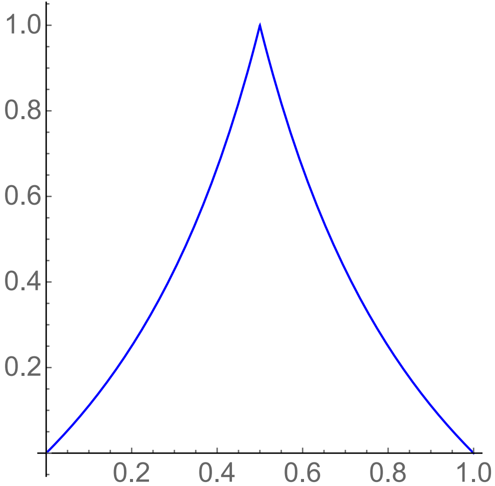


Figure: The projective additive continued fraction mapping.

Multiplicative acceleration

The multiplicative Euclidean algorithm is given by

$$G : \mathbb{R}^2_{\geq 0} \setminus \{\mathbf{0}\} \to \mathbb{R}^2_{\geq 0} \setminus \{\mathbf{0}\}$$

$$(a, b) \mapsto \begin{cases} (a - \lfloor \frac{a}{b} \rfloor b, b), & \text{if } a > b, \\ (a, b - \lfloor \frac{b}{a} \rfloor a), & \text{if } b \geq a. \end{cases}$$

Again this yields a sequence of matrices $M_1^{a_0}, M_2^{a_1}, M_1^{a_2}, \dots$

$$\begin{aligned} (a, b)^t &= M_1^{a_0} G(a, b) \\ &= M_1^{a_0} M_2^{a_1} G^2(a, b) \\ &= M_1^{a_0} M_2^{a_1} M_1^{a_2} G^3(a, b) = \cdots . \end{aligned}$$

Linear multiplicative continued fractions

Set

$$X = \{[a : b] \in \mathbb{P} \; : \; a \geq 0, \; b \geq 0\}.$$

Define $M : X \to \{M_1^c, M_2^c, \; : \; c \geq 1\}$ by

$$M([a : b]) = \begin{cases} M_1^c & \text{if } a > b \text{ and } 0 \leq a - cb < b, \\ M_2^c & \text{if } b \geq a \text{ and } 0 \leq b - ca < b. \end{cases}$$

Then the mapping

$$G : X \to X; \quad \mathbf{x} \mapsto M(\mathbf{x})^{-1}\mathbf{x}$$

is called the linear multiplicative continued fraction mapping.

Projective multiplicative continued fractions

In the same way as before this gives rise to a mapping

$$g : (0,1) \to (0,1), \quad x \mapsto \left\{\frac{1}{x}\right\}.$$

The mapping g is called Gauss map.
It defines the (multiplicative) continued fraction expansion

$$x = \cfrac{1}{a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots}}}, \quad \text{where} \quad a_n = \left\lfloor \frac{1}{g^n(x)} \right\rfloor.$$

Notation: $x = [a_0, a_1, a_2, \ldots]$.

The Gauss map

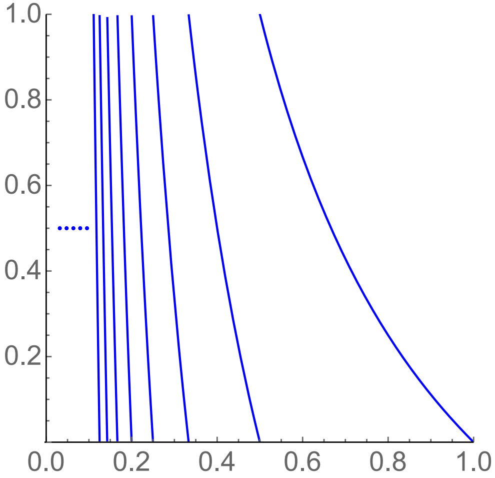


Figure: The Gauss map $x \mapsto \{\frac{1}{x}\}$.

S-adic representation and continued fractions

- Nonabelian: $\sigma_{i_0} \circ \cdots \circ \sigma_{i_n}(a)$ converges to a Sturmian seq.
- Abelian : $M_{i_0} \ldots M_{i_n}$ converges to a vector.

Lemma (Birkhoff 1957 and Furstenberg 1960)

If (i_n) changes its value infinitely often then (M_{i_n}) contains the positive block $M_1 M_2$ infinitely often. This implies that

$$\bigcap_{n\geq 0} M_{i_0} \cdots M_{i_n} \mathbb{R}_+^2 = \mathbb{R}_+ \mathbf{u}.$$

The lemma says that the additive continued fraction algorithm is weakly convergent.

Definition (Generalized right eigenvector)

$\mathbf{u}$ is called the generalized right eigenvector of (M_{i_n}).

The geometric meaning of **u**

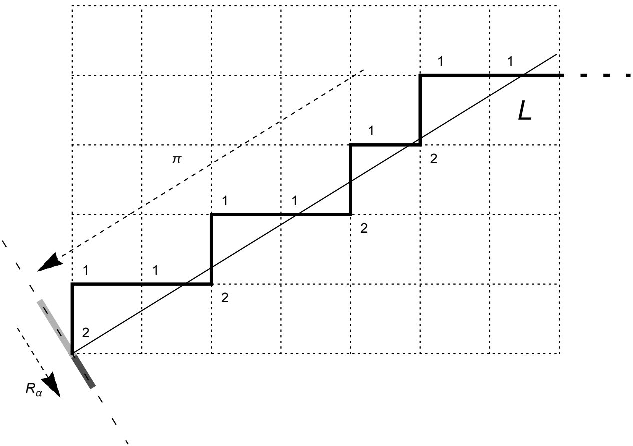


Figure: We can interpret a Sturmian sequence as a broken line. As we will see later, this line approximates the vector **u**.

The combinatorial meaning of **u**

Definition (Uniform letter frequencies)

$w = w_0 w_1 \ldots \in \{1,2\}^{\mathbb{N}}$ has uniform letter frequencies, *i.e.*,

$$f_a(w) = \lim_{\ell \to \infty} \frac{|w_k \ldots w_{\ell-1}|_a}{\ell - k}$$

exists for each $a \in \{1,2\}$ and does not dependent on k.
Frequencies can be defined also for words instead of letters.

Lemma

Let w be a Sturmian sequence with coding sequence (σ_{i_n}). Then w has uniform letter (and word) frequencies and

$$(f_1(w), f_2(w))^t = \frac{\mathbf{u}}{||\mathbf{u}||_1},$$

u ... *right eigenvector of sequence of incidence matrices (M_{i_n}).*

Sturmian dynamical system

Instead of a single Sturmian sequence w, we study a dynamical system generated by w in a natural way.

- Let w be a Sturmian sequence.
- Let
$$X_w = \overline{\{\Sigma^k w \ : \ k \in \mathbb{N}\}}$$
be the closure of the shift orbit of w.
- Alternatively, X_w can be viewed as the set
$$X_w = \left\{u \in \{1,2\}^{\mathbb{N}} \ : \ L(u) \subseteq L(w)\right\},$$
where $L(u)$ denotes the language of u.
- (X_w, Σ) with the shift Σ defined as
$$\Sigma(u_0 u_1 u_2 \ldots) = u_1 u_2 u_3 \ldots$$
is called Sturmian (dynamical) system.

Properties of sturmian systems

Lemma

Let (X_w, Σ) be a Sturmian system. Then it has the following properties.

(i) *The system (X_w, Σ) is recurrent.*

(ii) *The system (X_w, Σ) is minimal.*

(iii) *The system (X_w, Σ) is uniquely ergodic.*

(iv) *Let $\boldsymbol{\sigma} = (\sigma_n)$ be the coding sequence of w.
Then $X_w = X_{w'} =: X_{\boldsymbol{\sigma}}$ for any Sturmian sequence w' with coding sequence $\boldsymbol{\sigma}$.*

Proof.

(i) follows from recurrence of w, (ii) from primitivity, (iii) is a consequence of the existence of word frequencies, (iv) follows by primitivity and recurrence. □

The main result

Theorem (Morse and Hedlund, 1940)

A sequence $w \in \{1,2\}^{\mathbb{N}}$ is Sturmian if and only if there exists $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ such that w is a natural coding for R_α.

"$\Longleftarrow$" This is easy and was discussed before.

"$\Longrightarrow$" Morse and Hedlund gave a combinatorial proof of this.

In the 1991 Arnoux and Rauzy gave a very beautiful proof of this theorem in which the continued fraction algorithm pops up without being presupposed.

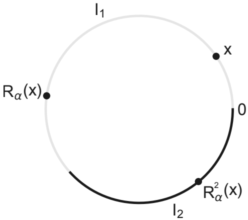
Idea: Show that natural codings of rotations are S-adic as well.

Lemma

For $\alpha \in (0,1)$ *let* u *be the natural coding of the point* $x = 1 - \alpha/(\alpha+1)$ *under an irrational rotation* $R_{\alpha/(\alpha+1)}$*. Then there is a sequence* (σ_{i_n}) *of substitutions such that*

$$u = \lim_{n\to\infty} \sigma_{i_n}(2)$$

The sequence $(i_n) \in \{1,2\}^{\mathbb{N}}$ *is of the form* $1^{a_0}2^{a_1}1^{a_2}2^{a_3}\ldots$ *where* $[a_0, a_1, a_2, \ldots]$ *is the continued fraction expansion of* α*. For* $\alpha > 1$ *a similar result with switched symbols holds.*



Natural coding: 112...

On the proof I

For computational reasons we “stretch” the interval.

- Consider the rotation R by α on the interval $J = [-1, \alpha)$ with the partition $P_1 = [-1, 0)$ and $P_2 = [0, \alpha)$.
- The natural coding u of $1 - \alpha/(\alpha + 1)$ by $R_{\alpha/(\alpha+1)}$ is the natural coding of 0 by R.
- Let R' be the first return map of R to the interval $J' = \left[\alpha \left\lfloor \frac{1}{\alpha} \right\rfloor - 1, \alpha\right)$.
- Let v be a natural coding of the orbit of 0 for R'.

The induction

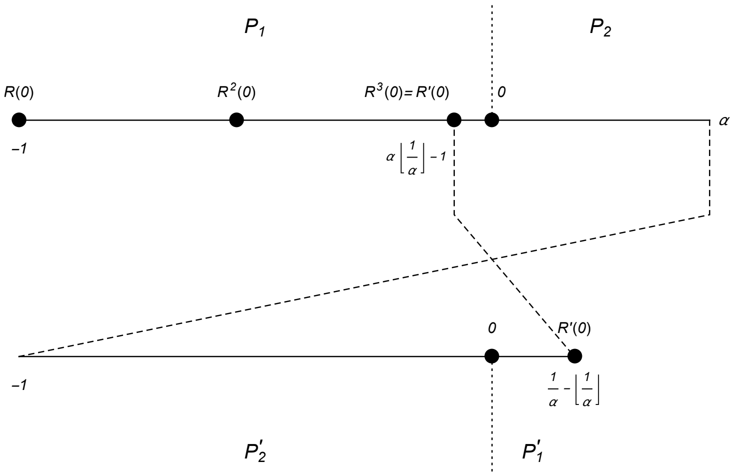


Figure: The rotation R' induced by R.

On the proof II

- v emerges from u by removing a block of 1s after each letter 2 occurring in u. By the definition of σ_1 this just means that $u = \sigma_1^{\lfloor 1/\alpha \rfloor}(v)$.
- Renormalize: The Gauss map pops up!!!
- Iterate: 1 and 2 are interchanged in the next step.
- This gives a sequence $(u^{(n)})_{n\geq 0}$ of natural codings

$$u = u^{(0)} \quad \text{and} \quad u^{(n)} = \sigma_{i_n}(u^{(n+1)}) \text{ for } n \geq 0$$

  for some sequence (σ_{i_n}) with $(i_n) \in \{1,2\}^{\mathbb{N}}$ having infinitely many changes between the letters 1 and 2.
- Thus

$$u = \lim_{n\to\infty} \sigma_{i_0} \circ \cdots \circ \sigma_{i_n}(a),$$

  where a is the first letter of u.

A consequence for Sturmian systems

Corollary

A Sturmian system (X_σ, Σ) is measurably conjugate to an irrational rotation.

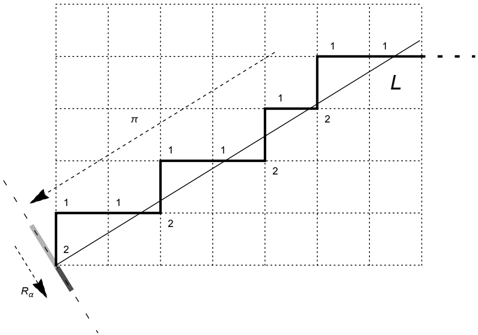


Figure: The rotation R_α is visible on a projection of the broken line

An Example I

- Let $\sigma : \{1,2\} \to \{1,2\}^*$ be given by

$$\sigma = \sigma_1 \circ \sigma_2 : \begin{array}{c} 1 \mapsto 121, \\ 2 \mapsto 21, \end{array}$$

  a reordering of the square of the Fibonacci substitution.
- An associated Sturmian sequence is

$$w = \lim_{n\to\infty} \sigma^n(2) = 211211212112112121121211121\ldots$$

- The asociated S-adic system $(X_{\boldsymbol{\sigma}}, \Sigma)$ is called a substitutive system $(\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_1, \ldots))$.

An Example II

- Let $\varphi = \frac{1+\sqrt{5}}{2}$. By the Perron-Frobenius Theorem its generalized right eigenvector **u** is the eigenvector $(\varphi, 1)^t$ corresponding to the eigenvalue φ^2.
- Let L be the eigenline defined by **u**. Being Sturmian, w is balanced and has uniform letter frequencies

$$(f_1(w), f_2(w))^t = (1, \varphi)^t / \sqrt{1 + \varphi^2}.$$

- This is reflected by the fact that the broken line

$$B = \{\mathbf{l}(p) \ : \ p \text{ is a prefix of } w\}$$

associated with the word w stays at bounded distance from the eigenline L.

An Example III

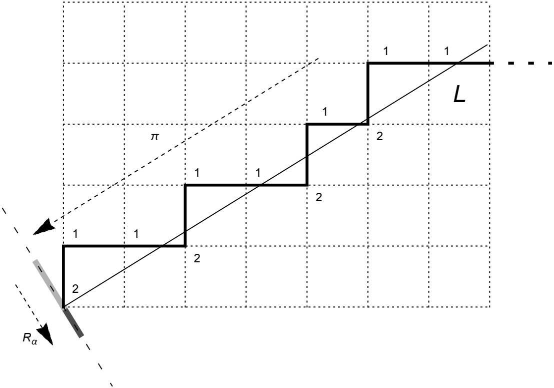


Figure: The broken line and its projection to the Rauzy fractal.

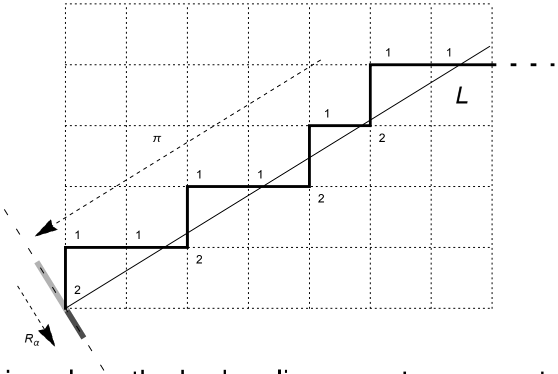
An Example IV

- w is a natural coding of the rotation by φ^{-2} of the point $\varphi^{-1} \in [0,1)$ with respect to the partition $I_1 = [0, \varphi^{-1})$, $I_2 = [\varphi^{-1}, 1)$ of $[0,1)$.
- Let π be the projection along L to the line $L^{\perp}$ orthogonal to L. If we project all points on the broken line and take the closure of the image, due to the irrationality of **u** we obtain the interval

$$\mathcal{R}_{\mathbf{u}} = \overline{\{\pi\mathbf{l}(p) \ : \ p \text{ is a prefix of } w\}}$$

on $L^{\perp}$.
- We color the part of the interval for which we write out 1 at the associated lattice point light grey, the other part dark grey:

$$\mathcal{R}_{\mathbf{u}}(i) = \overline{\{\pi\mathbf{l}(p) \ : \ pi \text{ is a prefix of } w\}} \qquad (i = 1, 2).$$



Thus passing along the broken line one step amounts to exchanging the intervals $\mathcal{R}_{\mathbf{u}}(1)$ and $\mathcal{R}_{\mathbf{u}}(2)$ in the projection. If we identify the end points of $\mathcal{R}_{\mathbf{u}}$ this interval exchange becomes a rotation. This is the rotation which is coded by the Sturmian sequence w. The union $\mathcal{R}_{\mathbf{u}} = \mathcal{R}_{\mathbf{u}}(1) \cup \mathcal{R}_{\mathbf{u}}(2)$ is called the Rauzy fractal associated with the substitution σ.

Lack of injectivity

- Gauss map: $x = [a_0, a_1, a_2, \ldots] \in (0,1)$ yields

$$g(x) = [a_1, a_2, a_3, \ldots].$$

  The partial quotient a_0 cannot be reconstructed from $g(x)$.

- Sturmian Recoding: Let $w = \lim_{n\to\infty} \sigma_{i_0} \circ \cdots \circ \sigma_{i_n}(a)$ be a Sturmian word. Then there is a recoded Sturmian word u with $w = \sigma_{i_0}(u)$, *viz.*

$$u = \lim_{n\to\infty} \sigma_{i_1} \circ \cdots \circ \sigma_{i_n}(a).$$

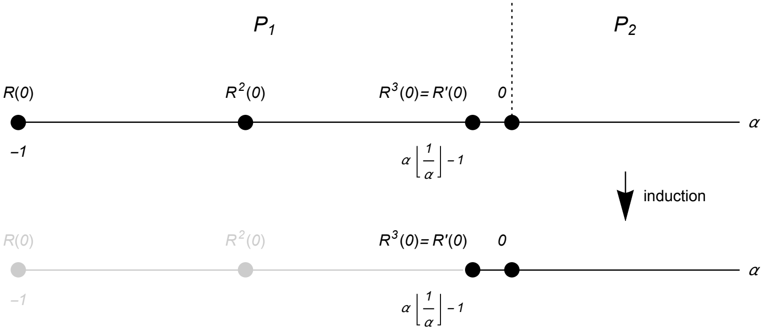
  The substitution σ_{i_0} cannot be reconstructed form u.

Rokhlin's Natural extension

- Let (X, T) be a dynamical system. T a surjection.
- Consider
$$Y = \{(x_i)_{i \in \mathbb{N}} \; : \; T(x_{i+1}) = x_i))\},$$
which is an inverse limit.
- Then $(Y, \hat{T})$ with $\hat{T} : Y \to Y$ given by
$$(x_1, x_2, \ldots) \mapsto (T(x_1), T(x_2), \ldots) = (T(x_1), x_1, x_2)$$
is the natural extension of (X, T).
- This goes back to Rokhlin.

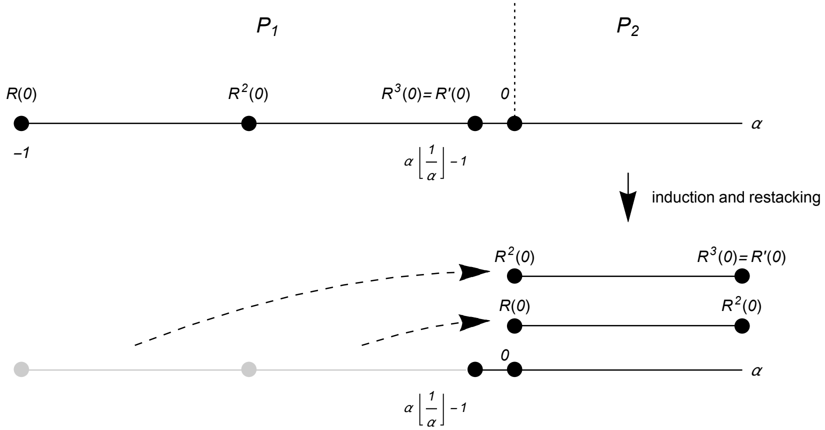
This is an abstract way of recording the past. We want it more concrete.

Inducing without restacking



We loose some part of the information: the intervals $[R(0), R^2(0))$ and $[R^2(0), R^3(0))$ depicted in light gray are no longer present in the induced rotation.

Inducing with restacking



The intervals $[R(0), R^2(0))$ and $[R^2(0), R^3(0))$ are stacked on one interval of the induced rotation. No information lost!

Boxes

Restacking the intervals has the disadvantage that we can go "back" only finitely many steps. Here is a better way of doing it (Arnoux and Fisher, 2001):

Build rectangular boxes above intervals

- One rectangle has width 1, the other one has width α.
- The sum of the areas of the rectangles is 1.
- Common lower vertex is 0.
- Now we can restack the rectangles.

Restacking and renormalizing boxes

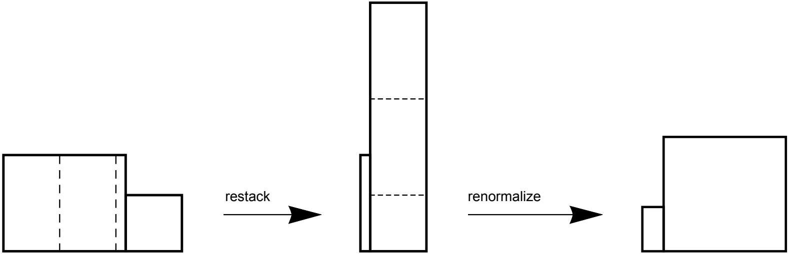


Figure: Step 1: Restack the boxes. Step 2: Renormalize in a way that the larger box has length 1 again.

$$\begin{array}{rclrcl} a & = & \text{length of large } \square, & b & = & \text{length of small } \square, \\ d & = & \text{width of large } \square, & c & = & \text{width of small } \square. \end{array}$$

Mapping in two variables since $\sup\{a, b\} = 1$ and $ad + bc = 1$.

The associated mapping

- Δ_m: Set of pairs of rectangles $(a \times d, b \times c)$ as above such that $a > b$ is equivalent to $d > c$ (the one with larger height has also larger width) with $\sup\{a, b\} = 1$ and $ad + bc = 1$.
- $\Delta_m = \Delta_{m,0} \cup \Delta_{m,1}$, where $a = 1$ in $\Delta_{m,0}$ and $b = 1$ in $\Delta_{m,1}$.

Definition

The map Ψ is defined on $\Delta_{m,1}$ by

$$(a, d) \mapsto \left(\left\{\frac{1}{a}\right\}, a - d^2 a\right),$$

and analogously on $\Delta_{m,0}$. This is the natural extension of the Gauss map.

Ψ preserves the Lebesgue measure and can be used to determine the invariant measure of the Gauss map.

Boxes and Sturmian words

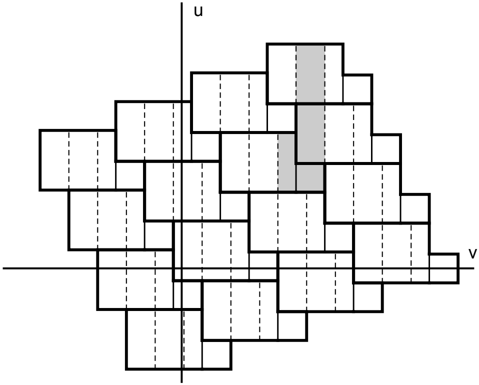


Figure: The vertical line is coded by a Sturmian word *u*, the horizontal line by a Sturmian word *v*. The restacking procedure desubstitutes *u* and substitutes *v*.

Lattices

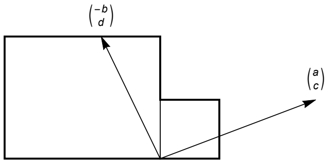


Figure: A pair of boxes is a fundamental domain of a lattice

- A pair of boxes is a fundamental domain of the lattice $\left\langle \begin{pmatrix} a \\ c \end{pmatrix}, \begin{pmatrix} -b \\ d \end{pmatrix} \right\rangle_{\mathbb{Z}} \in SL_2(\mathbb{Z}) \setminus SL_2(\mathbb{R})$.
- Renormalization can be done by multiplying the lattice by $\begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}$ from the right. This is the geodesic flow on $SL_2(\mathbb{Z}) \setminus SL_2(\mathbb{R})$.

Poincaré section

- Hit the pair of rectangles with the geodesic flow until the shorter rectangle has length 1.
- Restack; then the shorter rectangle has length 1. Restacking doesn't affect the lattice or the flow, only the basis.
- Repeat the procedure.

Poincaré section

The procedure above can be used to show that the natural extension of the Gauss map is a Poincaré section of the geodesic flow on $SL_2(\mathbb{Z}) \setminus SL_2(\mathbb{R})$.

Results of that type go back to Artin and were studied *e.g.* by Series, Arnoux, and Fisher.

The scenery flow (Arnoux and Fisher, 2001)

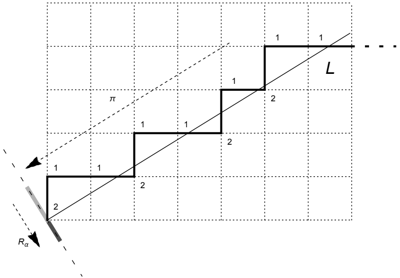
- Mark a point in the fundamental domain.
- This gives a torus fiber on the space of lattices.
- The geodesic flow acting on this extended space is called scenery flow.
- On each fiber we have a vertical and a horizontal flow. These flows are rotations that code Sturmian words as seen above.

Parametrization of Sturmian sequences & systems

Remark

- Set of points in a given pair of rectangles parametrizes the pairs of Sturmian words inside a given Strumian system (offsets with given slope).
- Set of pairs of rectangles parametrizes the set of natural extensions of sturmian Systems (slopes).

Pictures at the end



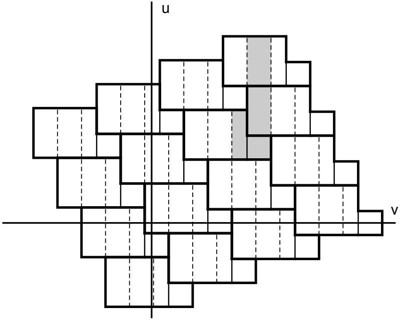


Figure: Sturmian sequences and their natural extensions