Symbolic dynamical systems and representations

V. Berthé

IRIF-CNRS-Paris-France
http ://www.irif.univ-paris-diderot.fr/~berthe

Journées nationales de calcul formel 2017

Outline

- Symbolic...
- ...dynamics
- Arithmetic dynamics and representations
  - Sturmian words
  - Numeration
  - Continued fractions
- Computational issues

Symbolic dynamics

Words and symbols

An alphabet $\mathcal{A}$ is a finite set

One studies words

- finite words $\mathcal{A}^*$ free monoid
- infinite words $\mathcal{A}^{\mathbb{N}}$, $\mathcal{A}^{\mathbb{Z}}$

from the viewpoint of word combinatorics

Words and symbols

An alphabet $\mathcal{A}$ is a finite set

One studies words

- finite words $\mathcal{A}^*$ free monoid
- infinite words $\mathcal{A}^{\mathbb{N}}$, $\mathcal{A}^{\mathbb{Z}}$

from the viewpoint of word combinatorics

but one can also add more structure

- Topological and measure-theoretic $\rightsquigarrow$ Symbolic dynamics and ergodic theory
- Algebraic $\rightsquigarrow$ Formal languages From free monoids to free groups

A substitution on words : the Fibonacci substitution

Definition A substitution σ is a morphism of the free monoid

Positive morphism of the free group, no cancellations

Example

$$\sigma : 1 \mapsto 12,\ 2 \mapsto 1$$

$$1$$
$$12$$
$$121$$
$$12112$$
$$12112121$$

$$\sigma^{\infty}(1) = 1211212112112 \cdots$$

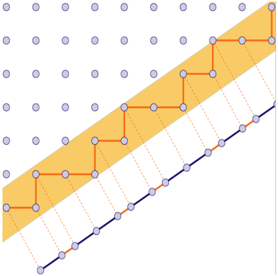
A substitution on words : the Fibonacci substitution

Definition A substitution σ is a morphism of the free monoid

Positive morphism of the free group, no cancellations

Example

$$\sigma : 1 \mapsto 12,\ 2 \mapsto 1 \quad \sigma^{\infty}(1) = 12112121121211212\cdots$$



The Fibonacci word yields a quasicrystal

A substitution on words : the Fibonacci substitution

Definition A substitution σ is a morphism of the free monoid

Positive morphism of the free group, no cancellations

Example

$$\sigma : 1 \mapsto 12, \; 2 \mapsto 1 \quad \sigma^{\infty}(1) = 121121211211212 \cdots$$

Why the terminology Fibonacci word ?

$$\sigma^{n+1}(1) = \sigma^n(12) = \sigma^n(1)\sigma^n(2)$$

$$\sigma^n(2) = \sigma^{n-1}(1)$$

$$\sigma^{n+1}(1) = \sigma^n(1)\sigma^{n-1}(1)$$

The length of the word $\sigma^n(1)$ satisfies the Fibonacci recurrence

Factors and language

Let $u \in \mathcal{A}^{\mathbb{N}}$ be an infinite word

$$u = abaababaabaababaababaab \cdots$$

$$u = abaababaab \underbrace{aa} babaababaab \cdots$$

aa is a factor, bb is not a factor

Let $\mathcal{L}_u$ be the set of factors of u : $\mathcal{L}_u$ is the language of u

Statistical vs. recurrence properties

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

One can consider which factors occur in u and count them for a given length

The factor complexity of u counts the number of factors of a given length

$$p_u(n) = \text{Card}\{\text{factors of } u \text{ of length } n\}$$

But one can also look at these factors from a statistical viewpoint

How often do they occur ?

Word combinatorics vs. symbolic dynamics

Let $u \in \mathcal{A}^{\mathbb{N}}$ be an infinite word.

- Word combinatorics

  Study of the number of factors of a given length (factor complexity), frequencies, repetitions, pattern avoidance, powers

- Symbolic dynamics

  Let $X_u := \overline{\{S^n u \mid n \in \mathbb{N}\}}$ with the shift $S((u_n)_n) = (u_{n+1})_n$

  (X_u, S) is a symbolic dynamical system

  Study of invariant measures, recurrence properties, finding geometric representations, spectral properties

Discrete dynamical system

We are given a dynamical system

$$T\colon X \to X$$

Discrete stands for discrete time
The set X is the set of states
The map T is the law of time evolution

We consider orbits/trajectories of points of X under the action of the map T

$$\{T^n x \mid n \in \mathbb{N}\}$$

Discrete dynamical system

We are given a dynamical system

$$T\colon X \to X$$

Discrete stands for discrete time
The set X is the set of states
The map T is the law of time evolution

We consider orbits/trajectories of points of X under the action of the map T

$$\{T^n x \mid n \in \mathbb{N}\}$$

- Topological dynamics describes the qualitative/topological asymptotic behaviour of trajectories/orbits
  The map T is continuous and the space X is compact
- Ergodicity describes the long term statistical behaviour of orbits
  The space X is endowed with a probability measure and T is measurable $(X, T, \mathcal{B}, \mu)$

Discrete dynamical system

We are given a dynamical system

$$T\colon X \to X$$

Discrete stands for discrete time
The set X is the set of states
The map T is the law of time evolution

We consider orbits/trajectories of points of X under the action of the map T

$$\{T^n x \mid n \in \mathbb{N}\}$$

How well are the orbits distributed ?

According to which measure ?

What are the orbits relevant for computer science ?

Ergodic theorem

We are given a dynamical system $(X, T, \mathcal{B}, \mu)$

$$T \colon X \to X$$

$$\mu(B) = \mu(T^{-1}B) \quad T\text{-invariance}$$

$$T^{-1}B = B \implies \mu(B) = 0 \text{ or } 1 \quad \text{ergodicity}$$

- Average time values : one particle over the long term Orbit
- Average space values : all particles at a particular instant, average over all possible sets

Ergodic theorem space mean= average mean

Theorem $f \in L_1(\mu)$ $\quad \lim_N \frac{1}{N} \sum_{0 \leqslant n < N} f(T^n x) = \int f \, d\mu$ a.e. x

Examples of dynamical systems

- Numeration $T\colon [0,1] \to [0,1],\ x \mapsto 10x - [10x] = \{10x\}$
- Beta-transformation $T\colon [0,1] \to [0,1],\ x \mapsto \{\beta x\}$
- Continued fractions $T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$
- Translation on the torus $R_\alpha\colon x \mapsto \alpha + x \mod 1$
- Symbolic systems $(\mathcal{A}^{\mathbb{N}}, S)$ where S is the shift acting on $\mathcal{A}^{\mathbb{N}}$

$$S((u_n)_n) = (u_{n+1})_n$$

Examples of dynamical systems

- Numeration $T\colon [0,1] \to [0,1],\ x \mapsto 10x - [10x] = \{10x\}$ positive entropy
- Beta-transformation $T\colon [0,1] \to [0,1],\ x \mapsto \{\beta x\}$ positive entropy
- Continued fractions $T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$ positive entropy
- Translation on the torus $R_\alpha\colon x \mapsto \alpha + x \mod 1$ zero entropy
- Symbolic systems $(\mathcal{A}^{\mathbb{N}}, S)$ where S is the shift acting on $\mathcal{A}^{\mathbb{N}}$

$$S((u_n)_n) = (u_{n+1})_n$$

Let $u \in \mathcal{A}^{\mathbb{N}}$ be an infinite word. Let

$$X_u := \overline{\{S^n u \mid n \in \mathbb{N}\}}$$

(X_u, S) is a symbolic dynamical system

Subshifts

- Topology for $u \neq v \in \mathcal{A}^{\mathbb{N}}$, $d(u,v) = 2^{-\min\{n \in \mathbb{N};\ u_n \neq v_n\}}$
- $\mathcal{A}^{\mathbb{N}}$ is complete as a compact metric space.
- $\mathcal{A}^{\mathbb{N}}$ is a Cantor set, that is, a totally disconnected compact set without isolated points.
- The shift map $S((u_n)_{n \in \mathbb{N}}) = (u_{n+1})_{n \in \mathbb{N}}$ is continuous.
- A subshift is a closed shift invariant system included in some $\mathcal{A}^{\mathbb{N}}$.
- Let $X_u := \overline{\mathcal{O}(u)}$ be the orbit closure of the infinite word u under the action of the shift S.

$$\overline{\mathcal{O}(u)} = \{v \in \mathcal{A}^{\mathbb{N}},\ \mathcal{L}_v \subset \mathcal{L}_u\},$$

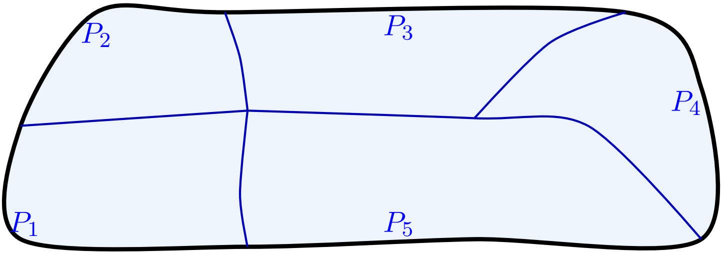
  where $\mathcal{L}_v$ is the set of factors of the sequence v.
- For a word $w = w_0 ... w_r$, the cylinder $[w]$ is the set

$$\{v \in X_u \mid v_0 = w_0, ..., v_r = w_r\}.$$

- Cylinders are clopen (open and closed) sets and form a basis of open sets for the topology of X_u.
- A clopen set is a finite union of cylinders.

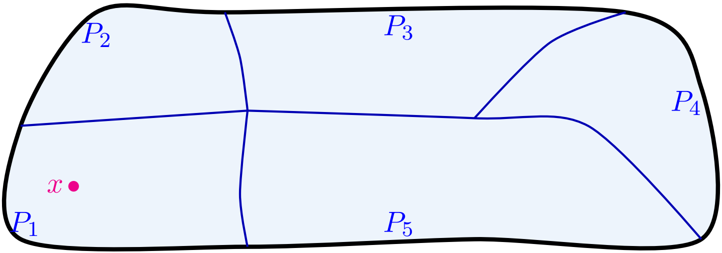
Coding of orbits of $T : X \to X$

Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

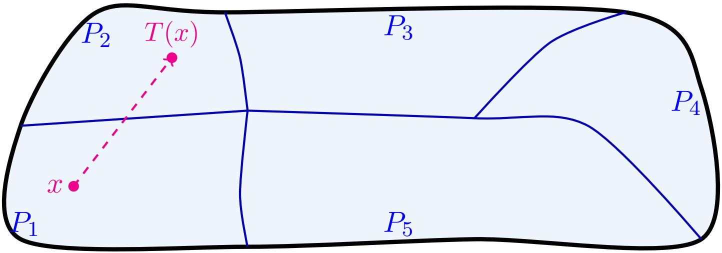
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

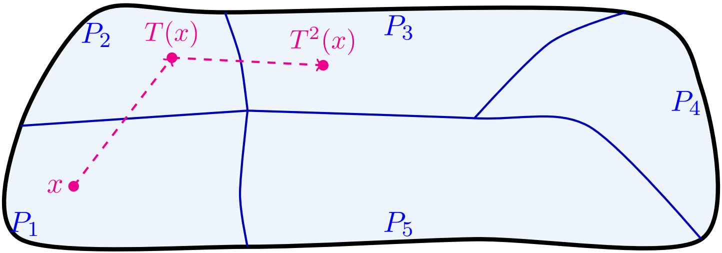
Coding of x $\quad 1$

Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$
Coding of x $\quad 1\ 2$

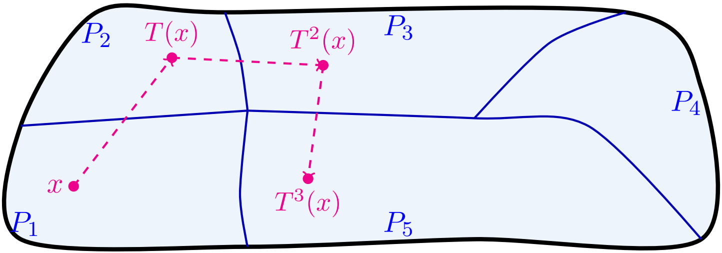
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad 1\ 2\ 3$

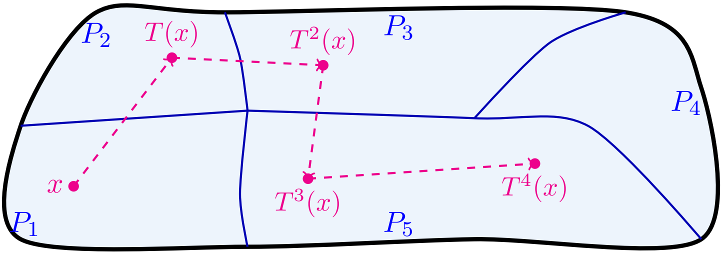
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad$ $1\ 2\ 3\ 5$

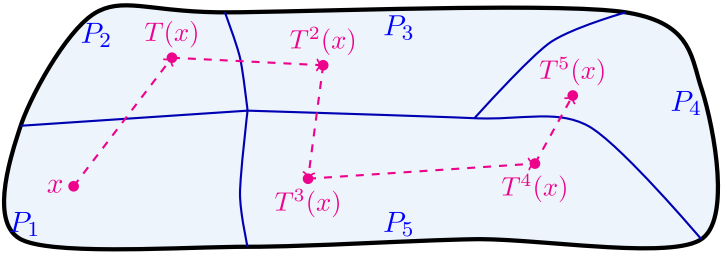
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad$ $1\ 2\ 3\ 5\ 5$

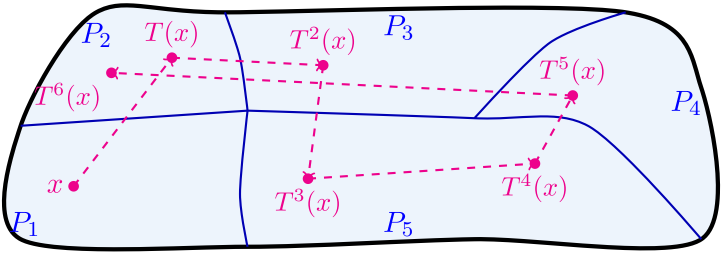
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $1\ 2\ 3\ 5\ 5\ 4$

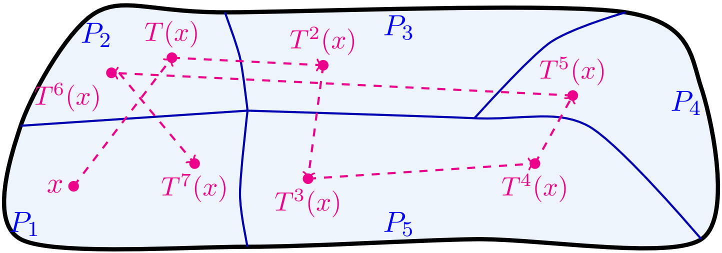
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad 1\ 2\ 3\ 5\ 5\ 4\ 2$

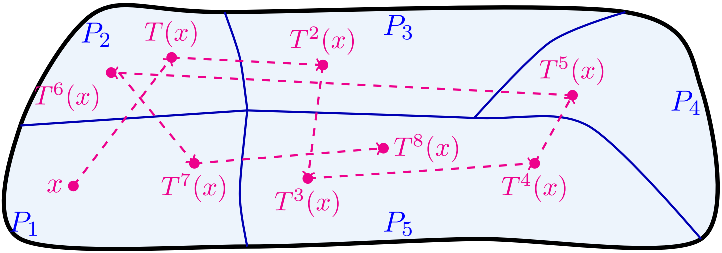
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad 1\ 2\ 3\ 5\ 5\ 4\ 2\ 1$

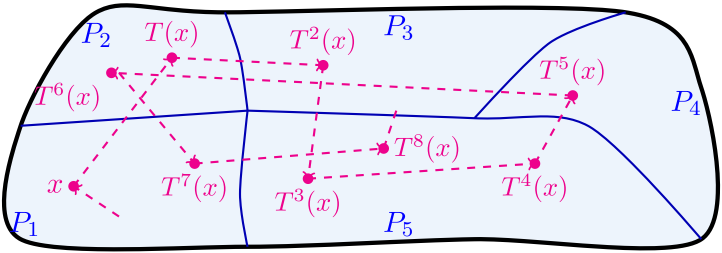
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\quad 1\ 2\ 3\ 5\ 5\ 4\ 2\ 1\ 5$

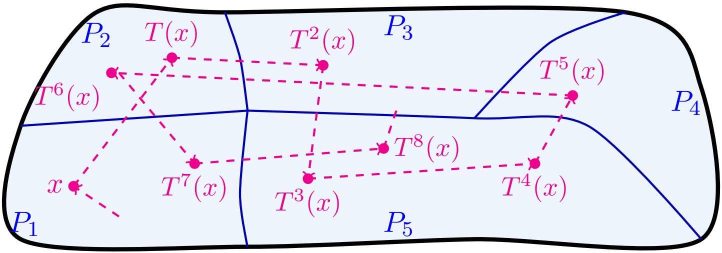
Coding of orbits of $T : X \to X$



Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\cdots 1\ 2\ 3\ 5\ 5\ 4\ 2\ 1\ 5 \cdots$

Coding of orbits of $T : X \to X$



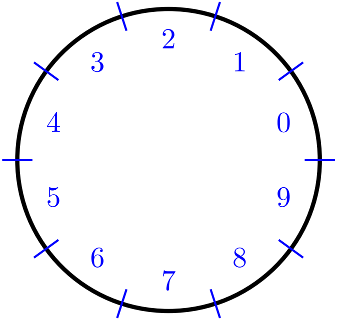
Partition $\mathcal{P} = \{P_1, P_2, P_3, P_4, P_5\}$

Coding of x $\cdots 1\ 2\ 3\ 5\ 5\ 4\ 2\ 1\ 5 \cdots$

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

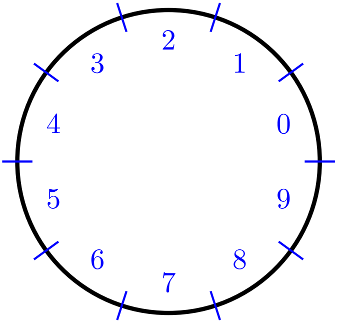
$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



Multiplication by 10 on $[0, 1]$

$X = [0, 1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

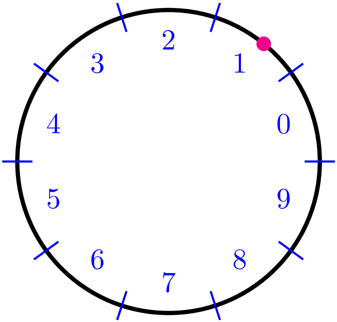


Orbit of $\pi - 3$:

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



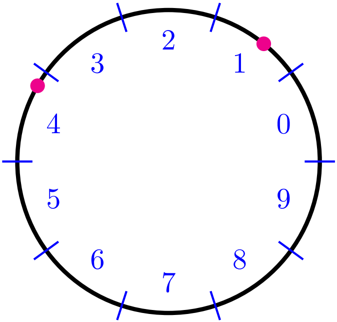
Orbit of $\pi - 3$:

0.1

Multiplication by 10 on $[0,1]$

$$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$$

$$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$$



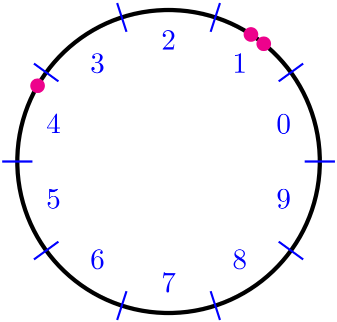
Orbit of $\pi - 3$:

0.14

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

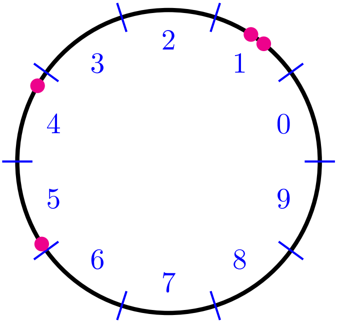


Orbit of $\pi - 3$:
0.141

Multiplication by 10 on $[0, 1]$

$X = [0, 1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

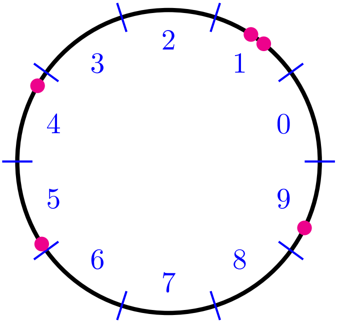


Orbit of $\pi - 3$:
0.1415

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



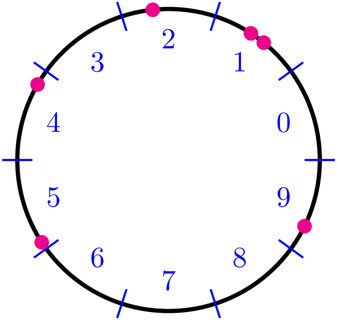
Orbit of $\pi - 3$:

0.14159

Multiplication by 10 on $[0,1]$

$$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$$

$$\mathcal{P} = \left\{ \left[\tfrac{i}{10}, \tfrac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$$

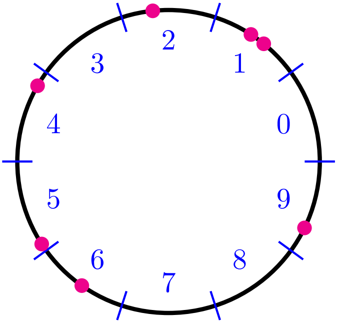


Orbit of $\pi - 3$:
0.141592

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

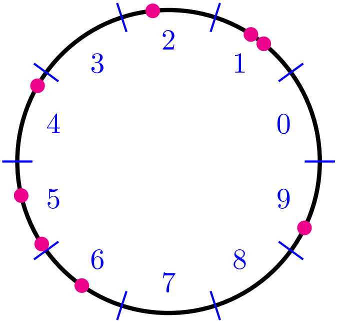


Orbit of $\pi - 3$:
0.1415926

Multiplication by 10 on $[0, 1]$

$X = [0, 1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

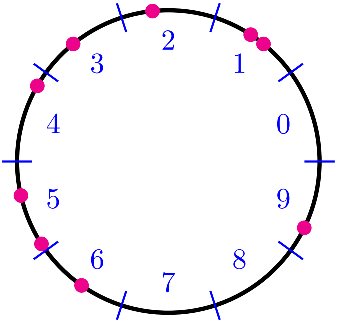


Orbit of $\pi - 3$:
0.14159265

Multiplication by 10 on $[0, 1]$

$X = [0, 1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

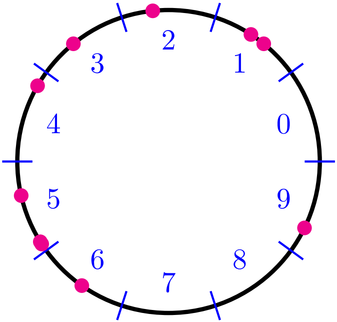


Orbit of $\pi - 3$:
0.141592653

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



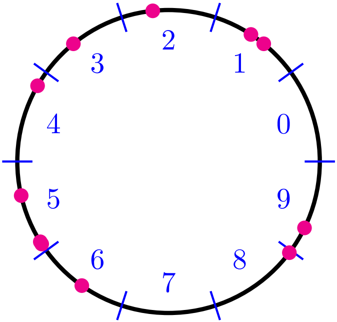
Orbit of $\pi - 3$:

0.1415926535

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

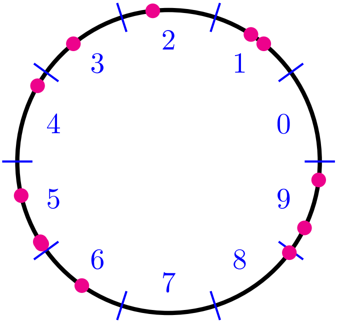


Orbit of $\pi - 3$:
0.14159265358

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

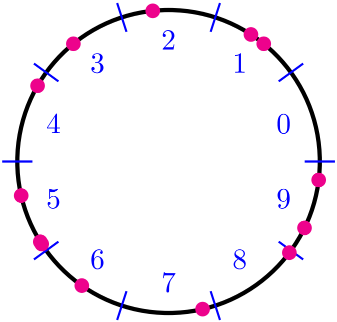


Orbit of $\pi - 3$:
0.141592653589

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

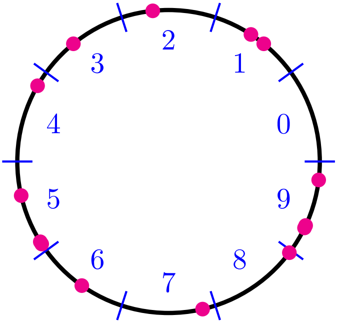


Orbit of $\pi - 3$:
$0.14159265358 97$

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$

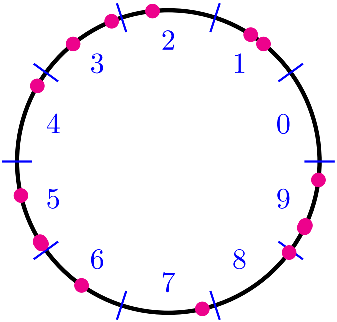


Orbit of $\pi - 3$:
$0.14159265358979312\cdots$

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



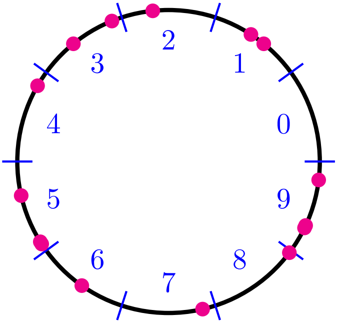
Orbit of $\pi - 3$:
$0.14159265358979312\cdots$

Codings $\Longleftrightarrow$ decimal expansions

Multiplication by 10 on $[0,1]$

$X = [0,1] \qquad T : x \mapsto 10\,x \pmod 1$

$\mathcal{P} = \left\{ \left[\frac{i}{10}, \frac{i+1}{10} \right[: 0 \leqslant i \leqslant 9 \right\}$



Orbit of $\pi - 3$:
$0.14159265358979312\cdots$

Codings $\Longleftrightarrow$ decimal expansions

The coding $\varphi : \{0, \ldots, 9\}^{\mathbb{Z}} \to X$ is not one-to-one
$0.999\cdots = 1.000\cdots$ or $0.46999\cdots = 0.47000\cdots$
(decimal numbers have two preimages)

Symbolic dynamics

- 1898, Hadamard : Geodesic flows on surfaces of negative curvature
- 1912, Thue : Prouhet-Thue-Morse substitution

$$\sigma : a \mapsto ab,\ b \mapsto ba$$

- 1921, Morse : Symbolic representation of geodesics on a surface with negative curvature. Recurrent geodesics

From geometric dynamical systems to
symbolic dynamical systems and backwards

- Given a geometric system, can one find a good partition ?
- And vice-versa ?

Symbolic dynamics and computer algebra

- Sage and word combinatorics
- Sage and interval exchanges etc...
- Computation of densities for invariant measures, Lyapunov exponents etc...
- Roundoffs for numerical simulations,
- Finite state machine simulations
- Computer orbits

Arithmetic dynamics

Arithmetic dynamics

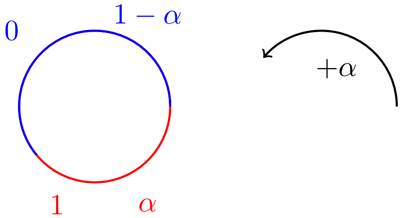
Arithmetic dynamics [Sidorov-Vershik'02] arithmetic codings of dynamical systems that preserve their arithmetic structure

Arithmetic dynamics

Arithmetic dynamics [Sidorov-Vershik'02] arithmetic codings of dynamical systems that preserve their arithmetic structure

Example Let $R_\alpha\colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x+\alpha$ mod 1
One codes trajectories according to the finite partition

$$\{I_0 = [0, 1-\alpha[,\ I_1 = [1-\alpha, 1[\}$$



Sturmian dynamical systems

Sturmian dynamical systems code translations on the one-dimensional torus

Let $R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha$ mod 1

Sturmian dynamical systems

Sturmian dynamical systems code translations on the one-dimensional torus

Let $R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha$ mod 1

Theorem Sturmian words [Morse-Hedlund]
Let $(u_n)_{n\in\mathbb{N}} \in \{0,1\}^{\mathbb{N}}$ be a Sturmian word. There exist $\alpha \in (0,1)$, $\alpha \notin \mathbb{Q}$, $x \in \mathbb{R}$ such that

$$\forall n \in \mathbb{N},\ u_n = i \Longleftrightarrow R_\alpha^n(x) = n\alpha + x \in I_i \pmod 1,$$

with

$$I_0 = [0, 1-\alpha[,\ I_1 = [1-\alpha, 1[$$

or

$$I_0 =]0, 1-\alpha],\ I_1 =]1-\alpha, 1].$$

Sturmian dynamical systems

Sturmian dynamical systems code translations on the one-dimensional torus

Let $R_\alpha\colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha$ mod 1

This yields a measure-theoretic isomorphism

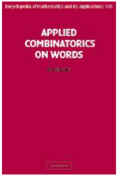
$$\begin{array}{ccc} \mathbb{R}/\mathbb{Z} & \xrightarrow{R_\alpha} & \mathbb{R}/\mathbb{Z} \\ \uparrow & & \uparrow \\ X_\alpha & \xrightarrow[\mathsf{S}]{} & X_\alpha \end{array}$$

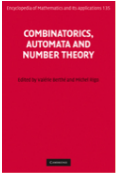
where S is the shift and $X_\alpha \subset \{0,1\}^{\mathbb{N}}$

Sturmian dynamical systems

Sturmian dynamical systems code translations on the one-dimensional torus

Let $R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha$ mod 1





[Lothaire, Algebraic combinatorics on words,
N. Pytheas Fogg, Substitutions in dynamics, arithmetics and combinatorics
CANT Combinatorics, Automata and Number theory]

Sturmian dynamical systems

Sturmian dynamical systems code translations on the one-dimensional torus

Let $R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha$ mod 1

Which trajectories ?

- α real number   generic ones
- α quadratic   substitutive words
- α rational   discrete geometry/Christoffel words

Example In the Fibonacci case

$$\sigma \colon a \mapsto ab, b \mapsto a$$

(X_σ, S) is isomorphic to $(\mathbb{R}/\mathbb{Z}, R_{\frac{1+\sqrt{5}}{2}})$

$$R_{\frac{1+\sqrt{5}}{2}} \colon x \mapsto x + \frac{1+\sqrt{5}}{2} \text{ mod } 1$$

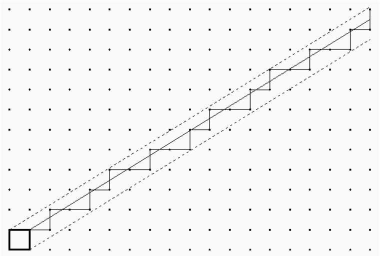
Sturmian words and continued fractions

0110110101101101

Sturmian words and continued fractions

0110110101101101

11 and 00 cannot occur simultaneously



Sturmian words and continued fractions

$$0110110101101101$$

One considers the substitutions

$$\sigma_0\colon 0 \mapsto 0,\ \sigma_0\colon 1 \mapsto 10$$

$$\sigma_1\colon 0 \mapsto 01,\ \sigma_1\colon 1 \mapsto 1$$

One has

$$01\,1\,01\,1\,01\,01\,1\,01\,1\,01 = \sigma_1(0101001010)$$

$$0\,10\,10\,0\,10\,10 = \sigma_0(011011)$$

$$01\,1\,01\,1 = \sigma_1(0101)$$

$$01\,01 = \sigma_1(00)$$

Sturmian words and continued fractions

$$0110110101101101$$

One considers the substitutions

$$\sigma_0 \colon 0 \mapsto 0,\ \sigma_0 \colon 1 \mapsto 10$$

$$\sigma_1 \colon 0 \mapsto 01,\ \sigma_1 \colon 1 \mapsto 1$$

The Sturmian words of slope α are provided by an infinite composition of substitutions

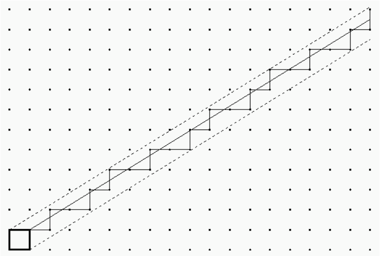
$$\lim_{n \to +\infty} \sigma_0^{a_1} \sigma_1^{a_2} \cdots \sigma_{2n}^{a_{2n}} \sigma_{2n+1}^{a_{2n+1}}(0)$$

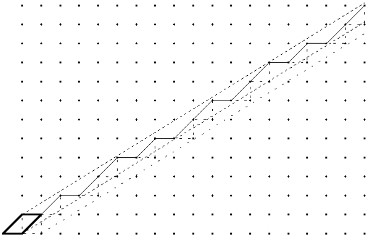
where the a_i are produced by the continued fraction expansion of the slope α

Such a composition of substitutions is called S-adic

Sturmian words and continued fractions

0110110101101101





Euclid algorithm and discrete segments

$$\begin{array}{|ccc|} \hline 11 & = & 2\cdot 4+3 \\ 4 & = & 1\cdot 3+1 \\ 3 & = & 3\cdot 1+0 \\ \hline \end{array}$$

$$\boxed{\frac{4}{11} = \cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{3}}}}$$

$$(11,4) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^2} (3,4) \xleftarrow{\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}} (3,1) \xleftarrow{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^3} (0,1)$$

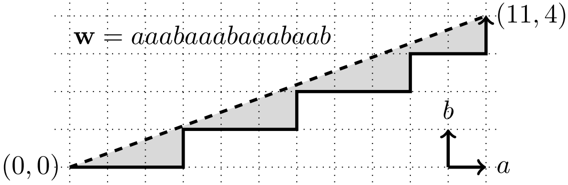
$$\begin{array}{ccccccccc} a & \mapsto & a & a & \mapsto & ab & a & \mapsto & a \\ b & \mapsto & aab & b & \mapsto & b & b & \mapsto & aaab \end{array}$$

$$\mathbf{w} = \mathbf{w}_0 \longleftarrow \mathbf{w}_1 \longleftarrow \mathbf{w}_2 \longleftarrow \mathbf{w}_3 = b$$

Euclid algorithm and discrete segments

$$\begin{array}{ccc} 11 & = & 2\cdot 4+3 \\ 4 & = & 1\cdot 3+1 \\ 3 & = & 3\cdot 1+0 \end{array}$$

$$\frac{4}{11}=\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{3}}}$$



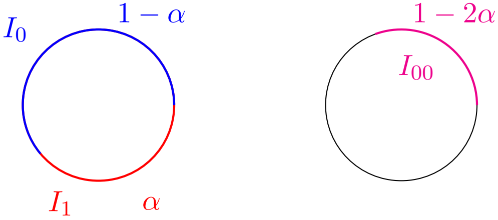
$$(11,4)\xleftarrow{\begin{pmatrix}1&1\\0&1\end{pmatrix}^2}(3,4)\xleftarrow{\begin{pmatrix}1&0\\1&1\end{pmatrix}}(3,1)\xleftarrow{\begin{pmatrix}1&1\\0&1\end{pmatrix}^3}(0,1)$$

$$\begin{array}{ccc} a & \mapsto & a \\ b & \mapsto & aab \end{array}\quad \begin{array}{ccc} a & \mapsto & ab \\ b & \mapsto & b \end{array}\quad \begin{array}{ccc} a & \mapsto & a \\ b & \mapsto & aaab \end{array}$$

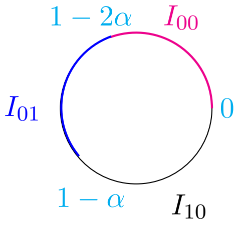
$$\mathbf{w}=\mathbf{w}_0\longleftarrow\mathbf{w}_1\longleftarrow\mathbf{w}_2\longleftarrow\mathbf{w}_3=b$$

From factors to intervals

$$R_\alpha\colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha \text{ mod } 1$$



From factors to intervals



The factors of u of length n are in one-to-one correspondence with the $n+1$ intervals of $\mathbb{T}$ whose end-points are given by

$$-k\alpha \text{ mod } 1, \text{ for } 0 \leqslant k \leqslant n$$

$$w \leadsto I_W = I_{w_1} \cap R_\alpha^{-1} I_{w_2} \cap \cdots R_\alpha^{-n+1} I_{w_n}$$

By uniform distribution of $(k\alpha)_k$ modulo 1, the frequency of a factor w of a Sturmian word is equal to the length of I_w

Balance and frequencies

A word $u \in A^{\mathbb{N}}$ is said to be finitely balanced if there exists a constant $C > 0$ such that for any pair of factors of the same length v, w of u, and for any letter $i \in A$,

$$||v|_i - |w|_i| \leqslant C$$

$|x|_j$ stands for the number of occurrences of the letter j in the factor x

Sturmian words are exactly the 1-balanced words

Fibonacci word $\sigma\colon a \mapsto ab,\ b \mapsto a$

$$\sigma^{\infty}(a) = abaababaabaababaababaabaababaabaababaabab\ldots$$

The factors of length 5 contain 3 or 4 a's

$$abaab, baaba, aabab, ababa, babaa, aabaa$$

Frequencies and unique ergodicity

The frequency f_i of a letter i in u is defined as the following limit, if it exists

$$f_i = \lim_{n\to\infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}$$

Frequencies and unique ergodicity

The frequency f_i of a letter i in u is defined as the following limit, if it exists

$$f_i = \lim_{n\to\infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}$$

One can also consider

$$\lim_{n\to\infty} \frac{|u_k \cdots u_{k+N-1}|_i}{N}$$

If the convergence is uniform with respect to k, one says that u has uniform letter frequencies.

One defines similar notions for factors.

Frequencies and unique ergodicity

The frequency f_i of a letter i in u is defined as the following limit, if it exists

$$f_i = \lim_{n\to\infty} \frac{|u_0 \cdots u_{N-1}|_i}{N}$$

One can also consider

$$\lim_{n\to\infty} \frac{|u_k \cdots u_{k+N-1}|_i}{N}$$

If the convergence is uniform with respect to k, one says that u has uniform letter frequencies.

One defines similar notions for factors.

The symbolic shift (X_u, S) is said to be uniquely ergodic if u has uniform factor frequency for every factor.

Equivalently, there exists a unique shift-invariant probability measure on the symbolic shift (X_u, S).

Theorem Let f be continuous $\lim_N \frac{1}{N} \sum_{0\leqslant n<N} f(T^n x) = \int f \, d\mu$ for a

Symbolic discrepancy

An infinite word $u \in \mathcal{A}^{\mathbb{N}}$ is finitely balanced if and only if

- it has uniform letter frequencies
- there exists a constant B such that for any factor w of u, we have

$$||w|_i - f_i|w|| \leqslant B \quad \text{for all } i$$

Definition The discrepancy of the word u is defined as

$$\Delta_u = \sup_{i \in A,\ n} ||u_0 \cdots u_{n-1}|_i - f_i \cdot n|$$

If u has letter frequencies

bounded discrepancy $\Longleftrightarrow$ finite balance

Particularly good convergence of frequencies

Finite balancedness implies the existence of uniform letter frequencies

Proof Assume that u is C-balanced and fix a letter i

Let N_p be such that for every word of length p of u, the number of occurrences of the letter i belongs to the set

$$\{N_p, N+1, \cdots, N_p + C\}$$

The sequence $(N_p/p)_{p\in\mathbb{N}}$ is a Cauchy sequence. Indeed consider a factor w of length pq

$$pN_q \leqslant |w|_i \leqslant pN_q + pC, \qquad qN_p \leqslant |w|_i \leqslant qN_p + qC$$

$$-C/p \leqslant N_p/p - N_q/q \leqslant C/q$$

Finite balancedness implies the existence of uniform letter frequencies

Proof Assume that u is C-balanced and fix a letter i

Let N_p be such that for every word of length p of u, the number of occurrences of the letter i belongs to the set

$$\{N_p, N+1, \cdots, N_p + C\}$$

The sequence $(N_p/p)_{p\in\mathbb{N}}$ is a Cauchy sequence. Indeed consider a factor w of length pq

$$pN_q \leqslant |w|_i \leqslant pN_q + pC, \qquad qN_p \leqslant |w|_i \leqslant qN_p + qC$$

$$-C/p \leqslant N_p/p - N_q/q \leqslant C/q$$

Let $f_i = \lim N_q/q$

$$-C \leqslant N_p - pf_i \leqslant 0 \quad (q \to \infty)$$

Then, for any factor w

$$||w|_i - f_i|w|| \leqslant C \qquad \rightsquigarrow \text{uniform frequencies}$$

From factors to intervals

$$R_\alpha \colon \mathbb{R}/\mathbb{Z} \to \mathbb{R}/\mathbb{Z},\ x \mapsto x + \alpha \text{ mod } 1$$

- The factors of u of length n are in one-to-one correspondence with the $n+1$ intervals of $\mathbb{T}$ whose end-points are given by $-k\alpha$, for $0 \leqslant k \leqslant n$
- By uniform distribution of $(k\alpha)_k$ modulo 1, the frequency of a factor w of a Sturmian word is equal to the length of I_w
- Sturmian words are 1-balanced
- Intervals I_w have bounded discrepancy Bounded remainder sets
- Kesten's theorem I has bounded discrepancy iff $|I| \in \mathbb{Z} + \alpha\mathbb{Z}$

How to compute frequencies and balances

For primitive substitutions

$\sigma \rightsquigarrow M_\sigma \rightsquigarrow$ Perron-Frobenius eigenvector [Adamczweski]

$M_\sigma[ij]$ counts the number of occurrences of i in $\sigma(j)$

For S-adic words

$\lim \sigma_1 \cdots \sigma_n(a) \rightsquigarrow \cap_n M_1 \cdots M_n \mathbf{e}_a$ Hilbert projective metric [Furstenberg]

For codings of dynamical systems

One uses equidistribution (=unique ergodicity)

Ex : Sturmian words and $(n\alpha)_n \mod 1$

Lyapunov exponents and ergodic deviations

Entropy

Dynamical systems

They can be

- chaotic
- deterministic (zero entropy)

Chaotic systems

- Devaney's definition of chaos A dynamical system is said to be chaotic if
  - it is sensitive to initial conditions
  - its periodic points are dense
  - it is topologically transitive

Chaotic systems

- Devaney's definition of chaos A dynamical system is said to be chaotic if
  - it is sensitive to initial conditions
  - its periodic points are dense
  - it is topologically transitive
- A dynamical system is said to be topologically transitive if there exists a point x such that $\{T^n x\}$ is dense in X
- A map is said to be sensitive to initial conditions if close initial points have divergent orbits, with the separation rate being exponential

- $T_\varphi \colon x \mapsto \varphi \cdot x \mod 1$ is chaotic
- $T_\varphi \colon x \mapsto \varphi + x \mod 1$ is not chaotic

Topological entropy

The factor complexity $p_u(n)$ of an infinite word u counts the number of factors of a given length

Topological entropy

$$\lim_n \frac{\log(p_u(n))}{n}$$

Topological entropy

The factor complexity $p_u(n)$ of an infinite word u counts the number of factors of a given length

Topological entropy

$$\lim_n \frac{\log(p_u(n))}{n}$$

The Fibonacci word $\sigma^\infty(a)$ with $\sigma\colon a \mapsto ab,\ b \mapsto a$ has zero entropy
Substitutive dynamical systems

The golden mean shift (words over $\{0, 1\}$ with no 11) has positive entropy
Subshift of finite type

Topological entropy

The factor complexity $p_u(n)$ of an infinite word u counts the number of factors of a given length

Topological entropy

$$\lim_n \frac{\log(p_u(n))}{n}$$

The measure-theoretic entropy of the shift (X, S, μ) is then defined as

$$H_\mu(X) = \lim_{n \to +\infty} \frac{1}{n} \sum_{w \in \mathcal{L}_X(n)} L(\mu[w])$$

where $L(x) = -x \log_d(x)$ for $x \neq 0$, and $L(0) = 0$ (d stands for the cardinality of the alphabet $\mathcal{A}$)

Lyapounov exponent

It measures the rate of separation of orbits

$$\lambda(x) = \lim_{n\to\infty} \frac{1}{n} \ln\left(|(T^n)'(x)|\right)$$

when this limit exists
with T being defined on the unit interval

$$|T(x) - T(y)| \sim T'(x) \cdot |x - y|$$

$$|T^n(x) - T^n(y)| \sim \prod_{i=0}^{n-1} |T'(T^i x)| \cdot |x - y|$$

$$|T^n(x) - T^n(y)| \sim \exp n\lambda(x) \cdot |x - y|$$

Numeration and representation

Numeration and representation

- Numeration systems
- Continued fractions

Numeration systems

Numeration is inherently dynamical

- How to produce the digits ?
- If one knows how to represent a number, how to represent the next one ?
- The representation of arbitrarily large numbers requires the iteration of a recursive algorithmic process

Base q numeration

How to produce the digits of the expansion of N in base q?

$$N = a_k q^k + \cdots + a_0, \quad \text{for all } i,\ a_i \in \{0, \cdots, q-1\}$$

- Greedy algorithm

$$\text{let } k \text{ s.t. } q^k \leqslant N < q^{k+1},\ a_k := [N/q^k],\ N \mapsto N - a_k q^k$$

$$a_k \to a_{k-1} \cdots \to a_0$$

- Dynamical algorithm

$$T\colon \mathbb{N} \to \mathbb{N},\ n \mapsto \frac{n - (n \bmod q)}{q}$$

$$a_0 \to a_1 \cdots \to a_k$$

Decimal expansions

How to produce the digits of the expansion of x in base 10 ?

$$x = \sum_{i \geqslant 1} a_i 10^{-i}, \quad \text{and for all } i,\ a_i \in \{0, \cdots, 9\}$$

$$T\colon [0,1] \to [0,1],\ x \mapsto 10x - [10x] = \{10x\}$$

Decimal expansions

How to produce the digits of the expansion of x in base 10?

$$x = \sum_{i \geqslant 1} a_i 10^{-i}, \quad \text{and for all } i,\ a_i \in \{0, \cdots, 9\}$$

$$T\colon [0,1] \to [0,1],\ x \mapsto 10x - [10x] = \{10x\}$$

$$x = a_1/10 + \sum_{i \geqslant 2} a_i 10^{-i}$$

$$[10x] = a_1 + \sum_{i \geqslant 1} a_{i+1} 10^{-i}$$

$$T(x) = \{10x\} = \sum_{i \geqslant 1} a_{i+1} 10^{-i}$$

Decimal purely periodic expansions

Which are the real numbers
having a purely periodic decimal expansion ?

Decimal purely periodic expansions

Which are the real numbers
having a purely periodic decimal expansion ?

These are the rational numbers a/b $\quad (\gcd(a,b)=1)$
with b coprime with 10

Decimal expansions of rational numbers

Let

$$T\colon \mathbb{Q}\cap[0,1]\to\mathbb{Q}\cap[0,1],\ x\mapsto 10x-[10x]=\{10x\}$$

Let $a/b\in[0,1]$ with b coprime with 10

$$T(a/b)=\{10\cdot a\}=\frac{10\cdot a-[10\cdot a/b]\cdot\ b}{b}=\frac{10\cdot a \text{ mod } b}{b}$$

- Denominator of $T^k(a/b)=b$
- Numerator of $T^k(a/b)$ belongs to $\{0,1,\cdots,b-1\}$

Decimal expansions of rational numbers

Let

$$T\colon \mathbb{Q}\cap[0,1]\to\mathbb{Q}\cap[0,1],\ x\mapsto 10x-[10x]=\{10x\}$$

Let $a/b\in[0,1]$ with b coprime with 10

$$T(a/b)=\{10\cdot a\}=\frac{10\cdot a-[10\cdot a/b]\cdot\ b}{b}=\frac{10\cdot a \bmod b}{b}$$

- Denominator of $T^k(a/b)=b$
- Numerator of $T^k(a/b)$ belongs to $\{0,1,\cdots,b-1\}$

We thus introduce

$$T_b\colon x\mapsto 10\cdot x \bmod b$$

$$T_b(a)\rightsquigarrow \text{numerator of } T(a/b)$$

We conclude by noticing that T_b is onto and thus one-to-one since we work on a finite set

Continued fractions

Euclid algorithm

We start with two nonnegative integers u_0 and u_1

$$u_0 = u_1 \left[\frac{u_0}{u_1} \right] + u_2$$

$$u_1 = u_2 \left[\frac{u_1}{u_2} \right] + u_3$$

$$\vdots$$

$$u_{m-1} = u_m \left[\frac{u_{m-1}}{u_m} \right] + u_{m+1}$$

$$u_{m+1} = \mathsf{gcd}(u_0, u_1)$$

$$u_{m+2} = 0$$

Euclid algorithm

We start with two nonnegative integers u_0 and u_1

$$u_0 = u_1 \left[\frac{u_0}{u_1}\right] + u_2$$

$$u_1 = u_2 \left[\frac{u_1}{u_2}\right] + u_3$$

$$\vdots$$

$$u_{m-1} = u_m \left[\frac{u_{m-1}}{u_m}\right] + u_{m+1}$$

$$u_{m+1} = \mathsf{gcd}(u_0, u_1)$$

$$u_{m+2} = 0$$

One subtracts the smallest number to the largest as much as we can

Euclid algorithm and continued fractions

We start with two coprime integers u_0 and u_1

$$u_0 = u_1 a_1 + u_2$$

$$\vdots$$

$$u_{m-1} = u_m a_m + u_{m+1}$$

$$u_m = u_{m+1} a_{m+1} + 0$$

$$u_{m+1} = 1 = \mathsf{gcd}(u_0, u_1)$$

Euclid algorithm and continued fractions

We start with two coprime integers u_0 and u_1

$$u_0 = u_1 a_1 + u_2$$

$$\vdots$$

$$u_{m-1} = u_m a_m + u_{m+1}$$

$$u_m = u_{m+1} a_{m+1} + 0$$

$$u_{m+1} = 1 = \mathsf{gcd}(u_0, u_1)$$

$$\frac{u_1}{u_0} = \frac{1}{a_1 + \frac{u_2}{u_1}}$$

$$u_1/u_0 = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots + \frac{1}{a_m + 1/a_{m+1}}}}}$$

Continued fractions

We represent real numbers in $(0, 1)$ as

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cdots}}}}$$

with partial quotients (digits) $a_i \in \mathbb{N}^*$

Continued fractions

One represents α as

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

in order to find good rational approximations of α

Continued fractions

One represents α as

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

in order to find good rational approximations of α

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots + \cfrac{1}{a_n}}}}$$

Continued fractions

One represents α as

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}}$$

in order to find good rational approximations of α

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots + \cfrac{1}{a_n}}}}$$

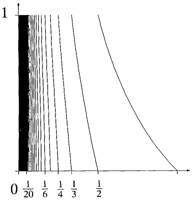
$$|\alpha - p_n/q_n] \leqslant 1/q_n^2$$

[http ://images.math.cnrs.fr/Nombres-et-representations.html]

Continued fractions and dynamical systems

Consider the Gauss map

$$T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$$



Let $x \in (0,1)$

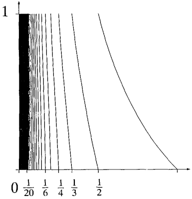
$$x_1 = T(x) = \{1/x\} = \frac{1}{x} - \left[\frac{1}{x}\right] = \frac{1}{x} - a_1$$

$$x = \frac{1}{a_1 + x_1}$$

Continued fractions and measure-theoeretic dynamical systems

Consider the Gauss map

$$T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$$



A measure is said to be T-invariant if $\mu(B) = \mu(T^{-1}B), \forall B \in \mathcal{B}$
The Gauss measure is defined as

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{1}{1+x} \mathrm{dx}$$

The Gauss measure is T invariant

Continued fractions and ergodicity

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{1}{1+x} \mathsf{dx}, \;\; \mu(B) = \mu(T^{-1}B) \quad T\text{-invariance}$$

Continued fractions and ergodicity

$$\mu(B) = \frac{1}{\log 2} \int_B \frac{1}{1+x} \mathsf{dx}, \;\; \mu(B) = \mu(T^{-1}B) \quad T\text{-invariance}$$

Theorem The Gauss map is ergodic with respect to the Gauss measure

Definition of ergodicity $T^{-1}B = B \implies \mu(B) = 0$ or 1

Continued fractions and ergodicity

$$\mu(B) = \frac{1}{\log 2}\int_B \frac{1}{1+x}\mathsf{dx},\ \mu(B) = \mu(T^{-1}B)\quad T\text{-invariance}$$

Theorem The Gauss map is ergodic with respect to the Gauss measure

Definition of ergodicity $T^{-1}B = B \implies \mu(B) = 0$ or 1

Ergodic theorem For a.e. x (=on a set of measure 1)

$$\lim_n \frac{1}{n}\sum_{j=0}^{n-1} f(T^j x) = \int f d\mu,\ \forall f \in L_1(\mu)$$

Take $f = \mathbf{1}_B$ for some measurable set B

Time mean= Mean value along an orbit =
=mean value of f w.r.t. μ = Spatial mean

Measure-theoretic results

Sets of zero measure for the Gauss measure= sets of zero measure for the Lebesgue measure

Almost everywhere (a.e.) = on a set of measure 1

- For a.e. $x \in [0,1]$

$$\lim \frac{\log q_n}{n} = \frac{\pi^2}{12 \log 2}$$

- For a.e. x and for $a \geqslant 1$

$$\lim_{N \to \infty} \frac{1}{N} \{k \leqslant N;\ a_k = a\} = \frac{1}{\log 2} \log \frac{(a+1)^2}{a(a+2)}$$

- Gauss measure

$$\mu(A) = \frac{1}{\log 2} \int_A \frac{dx}{1+x}$$

Continued fractions vs. decimal expansions

Let x_n, y_n with $x_n < x < y_n$ be the two consecutive n-th decimal approximations of x

We fix n Let $k_n(x)$ be the largest integer $k \geqslant 0$ such that

$$x_n = [a_0; a_1, \cdots, a_k, \cdots]$$

$$y_n = [a_0; a_1, \cdots, a_k, \cdots]$$

Theorem [Lochs'64] For almost every irrational number x (with respect to the Lebesgue measure)

$$\lim \frac{k_n(x)}{n} = \frac{6 \log 10 \log 2}{\pi^2} \sim 0.9702 = \frac{\text{Entropy base 10}}{\text{Entropy Gauss}}$$

Continued fractions vs. decimal expansions

Let x_n, y_n with $x_n < x < y_n$ be the two consecutive n-th decimal approximations of x

We fix n Let $k_n(x)$ be the largest integer $k \geqslant 0$ such that

$$x_n = [a_0; a_1, \cdots, a_k, \cdots]$$

$$y_n = [a_0; a_1, \cdots, a_k, \cdots]$$

Theorem [Lochs'64] For almost every irrational number x (with respect to the Lebesgue measure)

$$\lim \frac{k_n(x)}{n} = \frac{6 \log 10 \log 2}{\pi^2} \sim 0.9702 = \frac{\text{Entropy base 10}}{\text{Entropy Gauss}}$$

- "The n first decimals determine the n first partial quotients"
- The first 1000 decimals of π give the first 968 partial quotients
- The continued fraction is only slightly more efficient at representing real numbers than the decimal expansion

Formal power series with coefficients in $\mathbb{F}_q$

Formal power series

Let q be a power of a prime number p

We have the correspondence

- $\mathbb{Z} \sim \mathbb{F}_q[X]$
- $\mathbb{Q} \sim \mathbb{F}_q(X)$
- $\mathbb{R} \sim \mathbb{F}_q((X^{-1}))$

$$f = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0 + a_{-1} X^{-1} + \cdots$$

Laurent formal power series

Formal power series

Let $f \in \mathbb{F}_q((X^{-1}))$ $\qquad f \neq 0$

$$f = a_n X^n + a_{n-1} X^{n-1} + \cdots \qquad a_n \neq 0$$

- Degree $\qquad \deg f = n$
- Distance $\qquad |f| = q^{\deg f}$

Ultrametric space

$$|f + g| \leqslant \max(|f|, |gl)$$

No carry propagation !

Continued fractions

One can expand series f into continued fractions

$$f \;=\; a_0(X) + \cfrac{1}{a_1(X) + \cfrac{1}{a_2(X) + \ddots}} := [a_0(X); a_1(X), a_2(X), \cdots]$$

The digits $a_i(X)$ are polynomials of positive degree

$$a_k \geqslant 1 \rightsquigarrow \deg a_k(X) \geqslant 1$$

- Unique expansion even if f does not belong to $\mathbb{F}_q(X)$
- Finite expansion iff $f \in \mathbb{F}_q(X)$
- But there exist explicit examples of algebraic series with bounded partial quotients [Baum-Sweet]
- Roth's theorem does not hold for algebraic series (see e.g. [Lasjaunias-de Mathan])

[B.-Nakada, Expositiones Mathematicae]

Why is everything simpler ?

Ultrametric space !

- Digits are equidistributed : the Haar measure is invariant

Why is everything simpler ?

Ultrametric space !

- Digits are equidistributed : the Haar measure is invariant
- Hence, understanding the polynomial case can help the understanding of the integer case

Dynamical analysis

Rational vs. irrational parameters

Euclid algorithm $\rightsquigarrow$ gcd $\rightsquigarrow$ rational parameters

Continued fractions $\rightsquigarrow$ irrational parameters

Is it relevant to compare generic orbits
and orbits for integer parameters ?

Rational vs. irrational parameters

- When computing a gcd, we work with integer/rational parameters
- This set has zero measure
- Ergodic methods produce results that hold only almost everywhere

Average-case analysis vs. a.e. results

Fact Orbits of rational points tend to behave like generic orbits

And their probabilistic bevaviour can be captured thanks to the methods of dynamical analysis of algorithms

Number of steps for the Euclid algorithm

Consider

$$\Omega_m := \{(u_1, u_2) \in \mathbb{N}^2,\ 0 \leqslant u_1, u_2 \leqslant m\}$$

endowed with the uniform distribution

- Theorem The mean value $\mathbb{E}_m[L]$ of the number of steps satisfies

$$\mathbb{E}_m[L] \sim \frac{2}{\pi^2/(6\log 2)} \log m = \frac{1}{\lambda_1} \log m$$

$\pi^2/(6\log 2)$ is the entropy

[Heilbronn'69,Dixon'70,Hensley'94,Baladi-Vallée'03...]

Number of steps for a generalized Euclid algorithm

Consider parameters $(u_1, \cdots, u_d)$ with $0 \leqslant u_1, \cdots, u_d \leqslant m$

To be expected

$$\mathbb{E}_m[L] \sim \frac{\text{dimension}}{\text{Entropy}} \times \log m$$

- Formal power series with coefficients in a finite field and ploynomials with degree less than m

$$\frac{2}{2\frac{q}{q-1}}m = \frac{q-1}{q}m$$

- Brun [B.-Lhote-Vallée]

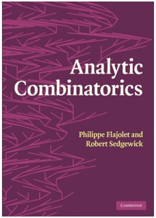
Dynamical analysis of algorithms [Vallée]

It belongs to the area of

- Analysis of algorithms [Knuth'63]

probabilistic, combinatorial, and analytic methods

- Analytic combinatorics [Flajolet-Sedgewick]



generating functions and complex analysis,
analytic functions, analysis of the singularities

Dynamical analysis of algorithms [Vallée]

It mixes tools from

• dynamical systems (transfer operators, density transformers, Ruelle-Perron-Frobenius operators)

• analytic combinatorics (generating functions of Dirichlet type)

the singularities of (Dirichlet) generating functions
are expressed in terms of transfer operators

Average analysis of algorithms

- [mean value] Computation of the asymptotic mean

$$\mathbb{E}_n[X] \underset{n\to\infty}{\sim} a_n$$

ex : what is the average bit complexity of the algorithm when the input size n is large ? Is it linear in n ? Quadratic in n ?...

- [variance] $\mathbb{V}_n[X] \underset{n\to\infty}{\sim} b_n$

ex : what is asymptotically the probability to be far from the mean value ?

- [limit law] What is the limit law of X

$$\mathbb{P}_n\left[\frac{X - a_n}{\sqrt{b_n}} \in [x, x + dx]\right] \underset{n\to\infty}{\sim} f(x)$$

ex : what is asymptotically the probability that X is in the interval $[a, b]$?

Distributional dynamical analysis

$\gcd(u_0, u_1) = 1 \quad N \geqslant u_0 > u_1 > \cdots \quad u_{k-1} = a_k u_k + u_{k+1}$

Cost of moderate growth $c(a) = O(\log a)$

- Number of steps in Euclid algorithm $c \equiv 1$
- Number of occurrences of a quotient $c = \mathbf{1}_a$
- Binary length of a quotient $c(a) = \log_2(a)$

Distributional dynamical analysis

$$\gcd(u_0, u_1) = 1 \quad N \geqslant u_0 > u_1 > \cdots \quad u_{k-1} = a_k u_k + u_{k+1}$$

Cost of moderate growth $c(a) = O(\log a)$

- Number of steps in Euclid algorithm $c \equiv 1$
- Number of occurrences of a quotient $c = \mathbf{1}_a$
- Binary length of a quotient $c(a) = \log_2(a)$

Theorem [Baladi-Vallée'05]

$$\mathbb{E}_N[\text{Cost}] = \frac{12 \log 2}{\pi^2} \cdot \hat{\mu}(c) \cdot \log N + O(1)$$

The distribution is asymptotically Gaussian (CLT)

$$\hat{\mu}(c) = \int_0^1 c([1/x]) \cdot \frac{1}{\log 2} \frac{1}{1+x} dx$$

$$C_n(x) = \sum_{i=1}^{n} c(a_i(x)) \qquad a_i = \left[\frac{1}{T^{i-1}(x)} \right]$$

Discrete case/Euclid Continuous case/truncated trajectories

Finite state simulation

By finite state machine simulation of the dynamical system (X, T), we mean the following : we consider

- a finite set $\hat{X}$, which is a set of finite sequences, this is a discretization of the space X,
- a coding map $\varphi \colon X \to \hat{X}$, i.e., a projection onto the discretized space $\hat{X}$,
- and a map $\hat{T}$ that acts on $\hat{X}$ with $\hat{T}(\hat{X}) \subset \hat{X}$, whose action is defined as a finite state machine
- we also want the behavior of $\hat{T} \circ \varphi$ to be close to $\varphi \circ T$

Dynamics and computation

One can consider uniform or nonuniform (floating point) discretizations

Consider a finite state machine simulation of a dynamical system

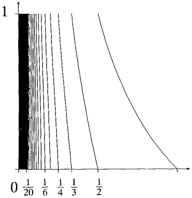
- all the orbits are ultimately periodic

- Are there generic orbits among computable orbits ?
- How far are computed orbits from exact ones ?
- How far are computed orbits from generic orbits ?
- How far are periodic orbits from generic ones ?
- Round-off errors
- Which invariants can be computed numerically (entropy, Lyapounov exponents) ?

The floating-point Gauss map

Consider the Gauß map

$$T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$$



The Gauss map has a singularity at point 0

The floating-point Gauss map

Consider the Gauß map

$$T\colon [0,1] \to [0,1],\ x \mapsto \{1/x\}$$

Floating-point Gauss map

$$\widehat{T}(0) = 0,\ \widehat{T}(x) = 1/x \bmod 1 \text{ otherwise}$$

- Are there orbits which do not go to 0 ? How do the orbits behave nearby 0 ?
- How far are calculated orbits from exact orbits ?

Theorem Orbits under the floating-point Gauss map are close to corresponding exact orbits

[R. M. Corless, Continued fractions and Chaos
[P. Góra, A. Boyarsky, Why do computers like Lebesgue measure]
[P.-A. Guihéneuf, Dynamical properties of spatial discretizations of a generic homeomorphism]

Random mappings on finite sets [Knuth,Flajolet-Odlyzko'89]

We consider random maps defined on a finite set with N elements
Orbits are ultimately periodic
In average...

- The purely periodic part has length $\sqrt{\pi N/8}$
- The preperiod has length $\sqrt{\pi N/8}$
- A connected component has size $2N/3$
- The number of components is $1/2 \log N$
- The number of cyclic nodes is $\sqrt{\pi N/2}$

Random mappings on finite sets [Knuth,Flajolet-Odlyzko'89]

We consider random maps defined on a finite set with N elements
Orbits are ultimately periodic
In average...

- The purely periodic part has length $\sqrt{\pi N/8}$
- The preperiod has length $\sqrt{\pi N/8}$
- A connected component has size $2N/3$
- The number of components is $1/2 \log N$
- The number of cyclic nodes is $\sqrt{\pi N/2}$

In summary, one has a one giant component and few large trees

Methods come from

analysis of algorithms/ combinatorial analysis
singularities of (exponential) generating functions

Bridges between Automatic Sequences and Algebra and Number Theory

April 24 - 28, 2017 (Spring school), May 1 - 5, 2017 (Workshop), CRM, Montréal, Canada

Speakers for the school

B. Adamczewski, Y. Bugeaud, C. Reutenauer, R. Yassawi

Organizing Committee

J. Bell, V. Berthé, Y. Bugeaud, S. Labbé

Part of the Winter 2017 thematic session Algebra and Words in Combinatorics at CRM