Ergodic theorems and particular trajectories

Brigitte Vallée

Laboratoire GREYC
CNRS et Université de Caen

Journées de l'ANR CODYS, Novembre 2018

Plan of the talk

(I) Main features of the Ergodic Theorem

(II) Another point of view
with focus in the case of a dynamical system of the interval

(III) Probabilistic studies of the truncated generic trajectories
Using the weighted density transformer

(IV) Particular trajectories for the Euclidean system:
- Finite trajectories (rational inputs)
- Periodic trajectories (irrational quadratic inputs)
- Statement of the main result

(V) Study of particular trajectories with three tools
- Dirichlet generating functions
- Singularity analysis
- the weighted transfer operator.

(VI) Return to the notion of size for a periodic trajectory

(I) the Ergodic Theorem and its main features

General ergodic framework.

Definitions. A probability space $(X, \mathcal{X}, \mu)$, and a measurable mapping $T : X \to X$.

– The subset $E \in \mathcal{X}$ is T-invariant iff $T^{-1}E = E$.

– The measure μ is T-invariant [or the map T preserves the measure μ]
iff for every subset $E \in \mathcal{X}$, one has $\mu(T^{-1}E) = \mu(E)$.

– Let $T : X \to X$ be a measure-preserving transformation on $(X, \mathcal{X}, \mu)$.
The map T is ergodic iff
for every T–invariant subset $E \in \mathcal{X}$, one has either $\mu(E) = 0$ or $\mu(E) = 1$.

Particular case of a dynamical system.

– A dynamical system is related to the particular case when:
X (compact) topological space, its Borel set $\mathcal{X}$, $T : X \to X$ continuous.

– A dynamical system (X, T) is said to be uniquely ergodic if
there exists a unique T-invariant Borel probability measure on X.
Such a measure is necessary ergodic for T.

Ergodic Theorem

Definition. Let $T : X \to X$ be a measure-preserving transformation on $(X, \mathcal{X}, \mu)$.
Consider a μ-integrable function c [i.e., $c \in \mathcal{L}^1(\mu)$] (a "cost" or a "weight")
Then, there are the following averages:

– space–average of c

$$\mu[c] = \int_X c(x) d\mu(x) .$$

– time–average of c along a n-th truncated trajectory $(x, Tx, \dots, T^{n-1}x)$

$$\frac{1}{n} C_n(x), \qquad C_n(x) := \sum_{\ell=0}^{n-1} c(T^\ell x) .$$

Ergodic Theorem. If T is μ-ergodic, then one has

$$\lim_{n\to\infty} \left[\frac{1}{n} \sum_{\ell=0}^{n-1} c(T^\ell x)\right] = \mu[c] \qquad \text{for almost } \mu\text{-every } x \in X$$

In the case of the unique ergodicity, when c is moreover continuous,
the previous holds for every $x \in X$.

Main features of the Ergodic Theorem

$$\lim_{n\to\infty} \left[\frac{1}{n} \sum_{\ell=0}^{n-1} c(T^\ell x)\right] = \mu[c] \qquad \text{for almost } \mu\text{-every } x \in X$$

– Almost everywhere?

No information about the exceptional subset $\mathcal{E}$ (of measure 0)
where the Ergodic Theorem "fails"

– Speed of convergence? No information

$\Longrightarrow$ A possible complementary point of view,
notably for dynamical systems of the interval

– Replace almost everywhere by on average

– First on the total set X

– Why not on smaller sets $Y \subset X$?

– Why not on discrete subsets Y .... (of zero measure !)

– Obtain information about the speed of convergence

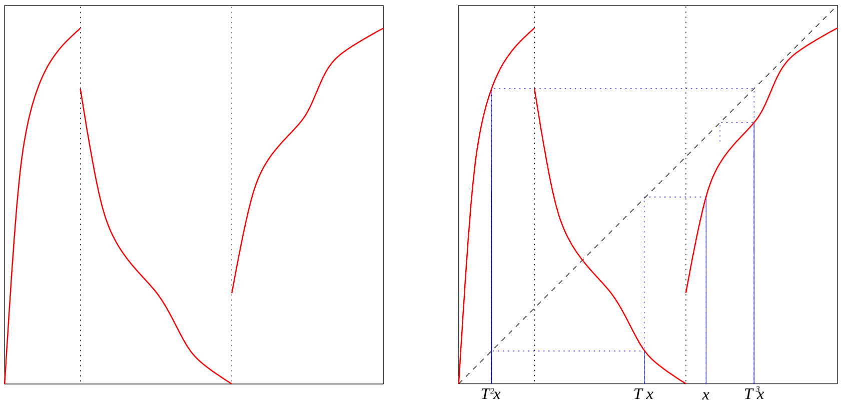
(II) Dynamical system of the interval : generalities

A dynamical system $(\mathcal{I}, T)$ of the unit interval $\mathcal{I}$ is defined by

- a finite or infinite denumerable alphabet Σ,
- a topological partition of $\mathcal{I} :=]0, 1[$ with open intervals $\mathcal{I}_{m, m\in\Sigma}$,
- a shift mapping T
  s.t. $T|_{\mathcal{I}_m}$ is a bijection of class $\mathcal{C}^2$ from $\mathcal{I}_m$ to $\mathcal{J}_m := T(\mathcal{I}_m)$.

Given an input x of $\mathcal{I}$, this gives rise to the trajectory

$$\mathcal{T}(x) := (x, Tx, T^2x, \dots)$$



A dynamical system, with $\Sigma = \{a, b, c\}$ and a word $M(x) = (c, b, a, c \ldots)$.

Correlations between symbols due to

– the geometry of the branches
– the shape of the branches

The geometry of the branches [position of $T(\mathcal{I}_m)$ wrt $\mathcal{I}_\ell$] ;
it describes the set $s(m)$ of possible successors of the symbol m.

Particular cases:

- Complete systems $T(\mathcal{I}_m) = \mathcal{I}$
- Markovian systems $T(\mathcal{I}_m) =$ union of some $\mathcal{I}_\ell$

give rise to a finite characterization of $s(m)$.

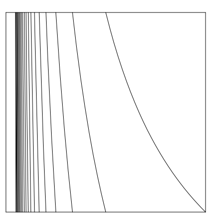
The shape of the branches [derivatives of the branches] also explains how the distribution evolves.

General case of interest.

A complete – or a Markovian – system
– with a possible infinite denumerable alphabet
– topologically mixing – and expansive.

Main instances: – (in this talk) the system defined with the Gauss map

$$T(x) := \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor, \qquad T(0) = 0$$



– (in Frédéric's talk), the system defined with the β-shift (for $\beta > 1$) with
$$T_\beta(x) = \beta x - \lfloor \beta x \rfloor$$

(III) A probabilistic point of view on truncated trajectories.

Focus on dynamical systems (of the interval)

A probabilistic point of view on truncated trajectories.

– Describe the behaviour of the n-th weighted truncated trajectories
– and their limit for $n \to \infty$

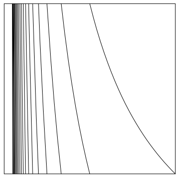
Interest in the random variables $C_n : X \to \mathbb{R}$ associated with a cost c

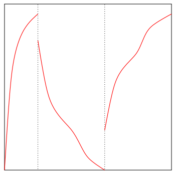
$$x \mapsto C_n(x), \quad \text{with} \quad C_n(x) := \sum_{\ell=0}^{n-1} c(T^\ell x)$$

– What about the sequence of expectations $\mathbb{E}[C_n]$?

– Are there limit distributional results for the random variables C_n?
for instance a limit gaussian law?

The weighted density transformer associated with a complete system.





$\mathcal{H} :=$ {inverse branches}

Density Transformer: for an initial density f on $[0,1]$,
$\mathbf{H}[f]$ is the density on $[0,1]$ after one iteration of the shift

$$\mathbf{H}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|\, f \circ h(x)$$

Weighted density transformer : extension of $\mathbf{H}$ into $\mathbf{H}_{1,w}$.

$$\mathbf{H}_{1,w}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|\, e^{wc(h(x))} f \circ h(x), \qquad \mathbf{H}_{1,0} = \mathbf{H}$$

The n-th iterate generates the weighted n-th truncated trajectories

$$\mathbf{H}^n_{1,w}[f](x) = \sum_{h \in \mathcal{H}^n} |h'(x)|\, e^{wC_n(h(x))}\, f \circ h(x)$$

$$\frac{d}{dw}\Big|_{w=0} \left[w \mapsto \mathbf{H}^n_{1,w}[f](x)\right] = \sum_{h \in \mathcal{H}^n} |h'(x)|\, C_n(h(x))\, f \circ h(x)$$

Remark: If the system is not complete, we must add indicator functions.
For a Markovian sytem, we can replace them by an operator matrix.

Weighted density transformer and probabilistic study of generic trajectories

$$\mathbf{H}^n_{1,w}[f](x) = \sum_{h \in \mathcal{H}^n} |h'(x)|\, e^{wC_n(h(x))}\, f \circ h(x)$$

$$\frac{d}{dw}\Big|_{w=0} \left[w \mapsto \mathbf{H}^n_{1,w}[f](x)\right] = \sum_{h \in \mathcal{H}^n} |h'(x)|\, C_n(h(x))\, f \circ h(x)$$

Take the integral over the unit interval + perform a change of variables

$$\mathbb{E}_f[e^{wC_n}] = \int_{\mathcal{I}} \mathbf{H}^n_{1,w}[f](t)dt, \qquad \mathbb{E}_f[C_n] = \frac{d}{dw}\Big|_{w=0} \left[w \mapsto \int_{\mathcal{I}} \mathbf{H}^n_{1,w}[f](t)dt\right]$$

Good dynamical system and cost of moderate growth
$\Longrightarrow$ $\mathbf{H}^n_{1,w}$ behaves as the n-th power of its dominant eigenvalue $\lambda(1,w)$,
uniformy when w is near 0.

$$\mathbb{E}_f[C_n] = n \cdot \mu[c]\,(1 + O(\rho^n)), \qquad \mu[c] = \lambda'_w(1,0)$$

$\mathbb{E}_f[e^{wC_n}]$ behaves as a (uniform) n-th power + Quasi-Powers Theorem
$\Longrightarrow$ Asymptotic Gaussian law for C_n

(IV) Study of particular trajectories of the Euclidean system

Discrete sets $\mathcal{X} \subset \mathcal{I}$ and trajectories $\mathcal{T}(x) := (x, Tx, T^2x, \ldots, T^\ell x, \ldots)$ for $x \in \mathcal{X}$.

Natural discrete sets $\mathcal{X}$:

– The set $\mathcal{Q}$ of x for which the trajectory $\mathcal{T}(x)$ is finite

– The set $\mathcal{P}$ of x for which the trajectory $\mathcal{T}(x)$ is (purely) periodic.

In the particular case of the Euclidean system (Gauss map):

$\mathcal{T}(x)$ finite $\iff x$ rational,

$\mathcal{T}(x)$ (purely) periodic $\iff x$ (reduced) quadratic irrational

In each case, a bijection between $\mathcal{X}$ and a (subset) of $\mathcal{H}^\star = \cup_{p\geq 0}\mathcal{H}^p$

$\mathcal{Q} := \mathbb{Q} \cap [0,1] = \{h(0) \mid h \in \mathcal{H}^\star \times \mathcal{F}\}$

$\mathcal{P} := \mathbb{I} \cap [0,1] = \{h(x_h) \mid h \in \mathcal{H}^\star,\ h \text{ primitive}, x_h \text{ fixed point of } h\}$

There is also a natural notion of size ϵ and a natural definition of cost C

$(\mathcal{Q})$: $\epsilon(x) :=$ denominator of $x = |h'(0)|^{-1/2}$

$(\mathcal{P})$: $\epsilon(x) := |\widehat{h}'(x_{\widehat{h}})|^{-1/2}$, $\widehat{h}$: = the primitive h of even period

$$C(x) := \sum_{\ell=0}^{P(x)-1} c(T^\ell(x)),\quad P(x) := \text{length or period of the trajectory } \mathcal{T}(x)$$

(V) Proof of the result with three tools

– Dirichlet generating functions

– Singularity analysis

– the weighted transfer operator.

Study of the mean cost along particular trajectories: Main result

For $\mathcal{X} \in \{\mathcal{Q}, \mathcal{P}\}$, the set of interest is $\mathcal{X}_N := \{x \in \mathcal{X} \mid \epsilon(x) \leq N\}$

There are two important costs on $\mathcal{X}_N$, the length P, and the cost C.

The mean values of the cost C or the cost P on $\mathcal{X}_N$ are the ratios

$$\mathbb{E}_N[P] := \frac{1}{|\mathcal{X}_N|}\sum_{x\in\mathcal{X}_N} P(x), \qquad \mathbb{E}_N[C] := \frac{1}{|\mathcal{X}_N|}\sum_{x\in\mathcal{X}_N} C(x)$$

Main Theorem. The asymptotic estimates hold on $\mathcal{Q}_N$ or $\mathcal{P}_N$

$$\mathbb{E}_N[P] \sim \frac{2}{h(T)}\log N, \quad \mathbb{E}_N[C] \sim \frac{2}{h(T)}\mu[c]\log N, \qquad (N \to \infty)$$

and involve the entropy $h(T)$ of the Euclidean system

Corollary. When $N \to \infty$, the asymptotic estimate holds: $\dfrac{\mathbb{E}_N[C]}{\mathbb{E}_N[P]} \to \mu[c]$

The particular trajectories (finite or periodic) behave on average in the same way as the generic trajectories behave almost everywhere.

Generating functions and analytic combinatorics

A general tool: A (bivariate) generating function, of Dirichlet type

$$F(s,w) := \sum_{x\in\mathcal{X}} \epsilon(x)^{-s} \cdot e^{wC(x)}, \qquad G(s) = \sum_{x\in\mathcal{X}} C(x)\cdot \epsilon(x)^{-s}$$

Main principles of analytic combinatorics: A good knowledge of

– the dominant singularity of G,

– the behaviour of G at this singularity

leads to the asymptotics of its coefficients, namely the behaviour of

$$\sum_{\substack{x\in\mathcal{X}\\ \epsilon(x)\leq N}} C(x), \quad |\mathcal{X}_N| := \sum_{\substack{x\in\mathcal{X}\\ \epsilon(x)\leq N}} 1, \quad (\text{for } N \to \infty)$$

This solves our question.

But, how to study the function $s \mapsto G(s)$ and "discover" its dominant singularity?

We express $F(s,w)$ – and thus $G(s)$ – in terms of the transfer operator.

Generating functions and transfer operators (I)

The bivariate generating function $F(s, w) := \sum_{x \in \mathcal{X}} \epsilon(x)^{-s} \, e^{wC(x)}$

is first expressed as a sum over the (convenient subset) of $\mathcal{H}^\star$

$$(\mathcal{Q}) \quad F(2s, w) \approx \sum_{h \in \mathcal{H}^\star} |h'(0)|^s \, e^{wC(h(0))} = \sum_{n \geq 0} \sum_{h \in \mathcal{H}^n} |h'(0)|^s \, e^{wC(h(0))}$$

$$(\mathcal{P}) \quad F(2s, w) \approx \sum_{h \in \mathcal{H}^\star} |h'(x_h)|^s \, e^{wC(x_h)} = \sum_{n \geq 0} \sum_{h \in \mathcal{H}^n} |h'(x_h)|^s \, e^{wC(x_h)}$$

Three differences with the previous study (truncated n-th trajectories)

- A sum over all possible "depths"
- The parameter s leads to the weighted transfer operator

$$\mathbf{H}_{s,w}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s \; e^{wc(h(x))} f \circ h(x)$$

- For $(\mathcal{P})$, a variable point x_h for each h and the replacement of $h \mapsto \widehat{h}$.

Some principles for Singularity Analysis

We are interested by the singularities of the operator (as a function of s)

$$\frac{d}{dw}\Big|_{w=0} \left[w \mapsto (I - \mathbf{H}_{s,w})^{-1} \right] = (I - \mathbf{H}_s)^{-1} \circ \mathbf{H}_s^{[c]} \circ (I - \mathbf{H}_s)^{-1}$$

Two main operators:

- The quasi inverse $(I - \mathbf{H}_s)^{-1}$ of the (unweighted) transfer operator

$$\mathbf{H}_s[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s f \circ h(x)$$

It has a dominant singularity at $s = 1$ (in fact a simple pole)

- In the Euclidean case, no other singularities on the line $\Re s = 1$.
  Good news for applying the Tauberian Theorem !!
- Not the case for the β-shift ....! There are an infinite set of poles
  on $\Re s = 1$ located at s for which $1 - \beta^{1-s} = 0$

- The weighted operator $\mathbf{H}_s^{[c]}[f](x) = \sum_{h \in \mathcal{H}} |h'(x)|^s \, c(h(x)) \, f \circ h(x)$

For a cost of moderate growth, it is regular at $s = 1$. It "brings" the factor $\mu[c]$.

Generating functions and transfer operators (II)

In case $(\mathcal{Q})$ (finite trajectories)

$$F(2s, w) \approx \sum_{h \in \mathcal{H}^\star} |h'(0)|^s e^{wC(h(0))} = \sum_{n \geq 0} \sum_{h \in \mathcal{H}^n} |h'(0)|^s e^{wC(h(0))}$$

$$F(2s, w) \approx (I - \mathbf{H}_{s,w})^{-1}[1](0)$$

In case $(\mathcal{P})$ (periodic trajectories)

$$F(2s, w) \approx \sum_{n \geq 0} \sum_{h \in \mathcal{H}^n} |h'(x_h)|^s e^{wC(x_h)}$$

the presence of the fixed point x_h of h leads to the trace of the operator.
Each component of the transfer operator is a composition operator,

$$f \mapsto |h'|^s e^{wC \circ h} f \circ h$$

On a good functional space, the set of its eigenvalues is

$$\{ |h'(x_h)|^s e^{wC(x_h)} \cdot (-1)^n h'(x_h)^n \mid n \in \mathbb{N} \}$$

Finally:

$$F(2s, w) \approx \mathrm{Tr} \left[\mathbf{H}_{s,w}^2 (I - \mathbf{H}_{s,w}^2)^{-1} \right]$$

(VI) Return to the notion of size for a periodic trajectory
Case of the Euclidean system

A natural notion of the size of a (reduced) quadratic irrational (rqi) ?
The arithmetical point of view.

A real x for which the trajectory $\mathcal{T}(x)$ is purely periodic of period $P(x)$.
Then x is defined by the relation $h(x) = x$,
with a (primitive) inverse branch $h \in \mathcal{H}^{P(x)}$

$$h(x) = x \Longrightarrow ax^2 + bx + c = 0, \qquad \text{with } \gcd(a, b, c) = 1$$

$$\Longrightarrow x \in \mathbb{Q}(\sqrt{\Delta}), \qquad \text{with } \Delta = b^2 - 4ac > 0$$

There is a fundamental unit in the quadratic (real) number field $\mathbb{Q}(\sqrt{\Delta})$
When chosen > 1 it is denoted as $\epsilon(x)$ and is chosen as the size of x.

Its computation involves the smallest inverse branch $\widehat{h}$ of even length.

$$\epsilon(x) = |\widehat{h}'(x_{\widehat{h}})|^{-1/2}$$

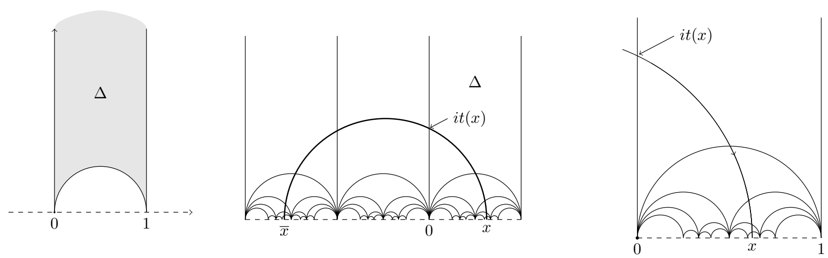
A natural notion of the size of a (reduced) quadratic irrational (rqi) ?
The dynamical point of view.

The hyperbolic plane $\mathbb{H} := \{z = x + iy \in \mathbb{C} \mid y > 0\}$.
When endowed with the metric $ds = |dz|/y$, its geodesics are
the semi-circles centered on the real axis and the vertical lines.

The set $\{h(\Delta) \mid h \in SL_2(\mathbb{Z})\}$ where Δ is the triangle with cusps $i\infty$, 0 and 1
defines the Farey tessellation of $\mathbb{H}$.

With a rqi number, $x \in [0, 1]$ one associates
- its conjugate $\bar{x}$ that satisfies $\bar{x} < -1$.
- its minimal even period q
- the sum of the digits along the even period $M := \sum_{i \le q} m_i$.

A natural notion of the size of a (reduced) quadratic irrational (rqi) ?
The dynamical point of view : Some geodesic pictures



The oriented geodesic $\gamma(x)$ that links $\bar{x}$ to x
- intersects the imaginary axis at $it(x)$.
- defines the oriented curve $\gamma_+(x)$ that links $it(x)$ to x
- this curve "crosses" domains $h(\Delta)$,

The primitive part of the geodesic is defined by the portion of $\gamma_+(x)$
that corresponds to the first M domains it crosses.
Its length $\rho(x)$ is related to the size $\epsilon(x)$ via $\rho(x) = \log \epsilon(x)$