INFINITE ERGODIC THEORY

CLAUDIO BONANNO

Contents

1. Introduction to Ergodic Theory

Let $(X, \mathcal{B})$ be a measurable space with $\mathcal{B}$ a σ-algebra, and let $T : X \to X$ be a measurable transformation (in general not assumed to be invertible). In many cases it is also natural to consider a *reference measure* m on $(X, \mathcal{B})$, and to look to statistical behaviour of the orbits of T with respect to m. A transformation T is called *non-singular* when $m(T^{-1}C) = 0$ iff $m(C) = 0$ for all $C \in \mathcal{B}$.

The existence of "stronger structures" on X leads to more natural choices. For example, if X is a complete metric separable space, it is natural to consider the Borel σ-algebra $\mathcal{B}$ and to restrict to positive Radon measures $\mathcal{M}$, the dual space of $C^0(X)$. If moreover X is compact, then $\mathcal{M}_1$, the set of probability Radon measures, is a compact metric space with respect to the weak* topology. Standard examples are $X = [0, 1]$ or $\mathbb{R}^k$, and m the Lebesgue measure.

When acting on a measurable space with a reference measure $(X, \mathcal{B}, m)$, we always consider non-singular transformations.

Definition 1.1. A measure μ is *T-invariant* if $\mu(T^{-1}C) = \mu(C)$ for all $C \in \mathcal{B}$.

If μ is T-invariant, then T is obviously non-singular with respect to μ. The notion of invariant measure is the analogous of an *invariant distribution* for a stochastic process.

Definition 1.2. A measure μ is *ergodic* if $T^{-1}C = C \pmod{\mu}$ for $C \in \mathcal{B}$ implies $\mu(C) = 0$ or $\mu(X \setminus C) = 0$.

An equivalent definition of ergodicity states that T-invariant measurable functions are μ-a.e. constant.

Remark 1.3. Points of view: μ is invariant/ergodic for (X, T) or T is measure preserving/ergodic for (X, μ).

Example 1.4 (Symbolic dynamics)**.** Let $\mathcal{A} = \{0, \dots, N-1\}$ be an alphabet with $N \in \mathbb{N} \cup \{\infty\}$, and $\Omega = \mathcal{A}^{\mathbb{N}_0} = \{\omega = (\omega_i)_{i \geq 0} \,:\, \omega_i \in \mathcal{A}\}$ be the space of infinite strings. The product topology is induced by the metric

$$d(\omega, \omega') = \min\{k \geq 0 \,:\, \omega_k \neq \omega'_k\}$$

and the Borel σ-algebra is generated by the *cylinders*

$$\mathcal{C}(\omega, k, n) := \left\{\omega' \in \Omega \,:\, \omega'_{k+i} = \omega_{k+i} \ \forall\, i = 0, \dots, n-1\right\}$$

The *(right) shift map* is $\sigma : \Omega \to \Omega$ given by

$$(\sigma\omega)_i = \omega_{i+1} \qquad i \geq 0$$

A *product measure* $\mu_{\mathbb{P}}$ is given by $\mu_{\mathbb{P}}(\mathcal{C}(\omega, k, n)) = \prod_{i=0}^{n-1} p_{\omega_{k+i}}$, where $\mathbb{P} = \{p_0, \dots, p_{N-1}\}$ is a probability vector. The product measures are σ-invariant and ergodic.

$\diamondsuit$

Example 1.5. Let $a > 1$ and $T_a : [0, 1] \to [0, 1]$ given by

$$T(x) = \begin{cases} a\,x\,, & x \in [0, \frac{1}{a}] \\ \frac{a}{a-1}\left(x - \frac{1}{a}\right)\,, & x \in [\frac{1}{a}, 1] \end{cases}$$

then the Lebesgue measure m is T_a-invariant and ergodic. T_2 is the *doubling map* and is conjugated to the shift map with $N = 2$, the conjugation map being the dyadic expansion of a real number.

$\diamondsuit$

Example 1.6. Let $\theta \in \mathbb{R}$ and $T_\theta : [0, 1] \to [0, 1]$ given by $T_\theta(x) = \{x + \theta\}$ is called *rotation of angle* θ. The Lebesgue measure is T-invariant, and it is ergodic if and only if $\theta \in \mathbb{R} \setminus \mathbb{Q}$. If $\theta \in \mathbb{Q}$ all points are periodic.

$\diamondsuit$

Example 1.7 ([5])**.** Let $r \in [0, 1]$ and $T_r : [0, 1] \to [0, 1]$ given by

$$T_r(x) = \begin{cases} \frac{(2-r)\,x}{1-rx}\,, & x \in [0, \frac{1}{2}] \\ \frac{(2-r)(1-x)}{1-r+rx}\,, & x \in [\frac{1}{2}, 1] \end{cases}$$

then the measure $d\mu_r(x) = \frac{r}{-\log(1-r)}\,\frac{1}{1-r+rx}\,dx$, for $r \in [0, 1)$, and $d\mu_1(x) = \frac{1}{x}\,dx$, for $r = 1$, is T_r-invariant and ergodic. The measures μ_r are equivalent to the reference measure m. T_0 is the *tent map*. T_1 is the *Farey (or Parry-Daniels) map*. Notice that $\mu_r([0, 1]) = 1$ for $r \in [0, 1)$, and $\mu_1([0, 1]) = \infty$.

$\diamondsuit$

Example 1.8 ([11])**.** Let $\alpha > 0$ and $T_\alpha : [0, 1] \to [0, 1]$ given by $T_\alpha(x) = \{x + x^{1+\alpha}\}$. These maps are called *Pomeau-Manneville (or Liverani-Saussol-Vaienti)* maps. For each α, there exists only one T_α-invariant measure μ_α equivalent to the reference measure m, which is also ergodic. The density $h_\alpha(x)$ of μ_α is unbounded at 0, and $h_\alpha(x) \sim x^{-\alpha}$ as $x \to 0^+$. Hence $\mu_\alpha([0, 1])$ is finite for $\alpha < 1$, and infinite for $\alpha \geq 1$. For $\alpha = 1$ the density of the invariant measure is given by $h_1(x) = \frac{1}{x} + \frac{1}{x+1}$ ([12]).

$\diamondsuit$

Example 1.9 ([3])**.** Let $T : \mathbb{R} \to \mathbb{R}$ given by $T(x) = x - \frac{1}{x}$ is the *Boole map*. The Lebesgue measure is T-invariant and ergodic (see [1]).

$\diamondsuit$

Other examples of maps on $(\mathbb{R}, m)$ are the *tangent map* $T(x) = \tan x$ which preserves the measure $d\mu(x) = \frac{1}{x^2}\,dx$ and is ergodic, and the *translation map* $T(x) = x + 1$ which preserves the Lebesgue measure m but is not ergodic.

1.1. **Recurrence.** Let T be a non-singular transformation of a space $(X, \mathcal{B}, m)$.

Definition 1.10. A measurable set $W \in \mathcal{B}$ is called a *wandering set* for T if the sets $\{T^{-n}W\}_{n=0}^{\infty}$ are disjoint.

Definition 1.11. The *dissipative part* $D(T)$ of T is the union of the wandering sets. The *conservative part* $C(T)$ of T is $X \setminus D(T)$. The transformation T is called *dissipative* if $D(T) = X \pmod{m}$, and it is called *conservative* if $C(T) = X \pmod{m}$.

It is immediate to verify that the translation map on $\mathbb{R}$ is dissipative. Conservativity of a transformation is equivalent to the validity of *Poincaré Recurrence Theorem*.

Theorem 1.12. *Let T be a transformation of a measurable space $(X, \mathcal{B})$ and let μ be a σ-finite T-invariant measure. Then the following are equivalent:*
(i) T is conservative on $(X, \mathcal{B}, \mu)$;
(ii) for all $C \in \mathcal{B}$ it holds $\sum_{k\ge 1} \chi_C \circ T^k \ge 1$ μ-a.e. on C;
(iii) for all $C \in \mathcal{B}$ it holds $\sum_{k\ge 1} \chi_C \circ T^k = \infty$ μ-a.e. on C.

If T is also ergodic and $\mu(C) > 0$, then conditions (ii) and (iii) hold μ-a.e. on X. Indeed the set where they hold is T-invariant, and it contains C, then it is all $X \pmod{\mu}$.

Corollary 1.13. *If T is a transformation of a measurable space $(X, \mathcal{B})$ preserving a probability measure μ, then T is conservative on $(X, \mathcal{B}, \mu)$.*

Proof. Let $C \in \mathcal{B}$ with $\mu(C) > 0$ and assume that there exists a measurable $C' \subset C$ with $\mu(C') > 0$ and $\sum_{k\ge1} \chi_C \circ T^k = 0$ on C', hence in particular $C' \cap T^k C' = \emptyset$ for all $k \ge 1$. Let now $C'_n := T^{-n}C'$ for $n \ge 1$, since μ is T-invariant it holds $\mu(C'_n) = \mu(C')$. Moreover we show that $C'_n \cap C'_m = \emptyset$ for $n \neq m$. If there exists $n > m$ with $C'_n \cap C'_m \neq \emptyset$, then $T^n(C'_n \cap C'_m) = C' \cap T^{n-m}C' \neq \emptyset$, which is false.

Hence $\cup_{n\ge1} C'_n$ is a measurable set in X and $1 = \mu(X) \ge \sum_{n\ge1} \mu(C'_n) = \sum_{n\ge1} \mu(C') = \infty$, and we have found a contradiction. ☐

The finiteness of a T-invariant measure is equivalent to finiteness of *mean first return time* to sets of finite measure.

Theorem 1.14 (Kac)**.** *Let μ be a T-invariant measure on $(X, \mathcal{B})$. For $C \in \mathcal{B}$ with $0 < \mu(C) < \infty$ let the first return time of C be defined as*

$$\varphi_C : C \to \mathbb{N}\,, \qquad \varphi_C(x) = \min\{n \ge 1 \,:\, T^n(x) \in C\}\,. \tag{1.1}$$

If T is conservative on $(X, \mathcal{B}, \mu)$ then $\varphi_C(x)$ is finite μ-a.e. on C. If moreover μ is ergodic then

$$\int_C \varphi_C(x)\, d\mu(x) = \mu(X)\,.$$

If μ is a probability measure, starting from a point in C, in mean we have to wait $\frac{1}{\mu(C)}$ iterations of T to return to C.

1.2. **Classical limit theorems.** Let T be a non-singular transformation of a space $(X, \mathcal{B}, m)$. We first get the analogous of the *strong law of large numbers* for the “evolution” of an observable along the orbits of T.

Let $f : X \to \mathbb{C}$ be a measurable observable. The *Birkhoff sums of f* are

$$S_n f(x) := \sum_{k=0}^{n-1} f(T^k(x))$$

Theorem 1.15 (Birkhoff Pointwise Ergodic Theorem)**.** *Let μ be a T-invariant probability on $(X, \mathcal{B})$. For all $f \in L^1(X,\mu)$ there exists $f^* : X \to \mathbb{C}$ such that*

$$\lim_{n\to\infty} \frac{1}{n} S_n f(x) = f^*(x) \qquad \textit{for } \mu\textit{-a.e. } x \in X\,.$$

If moreover μ is ergodic, then $f^*(x) = \int_X f d\mu$.

Remark 1.16. If f is a non-negative non-summable function then $\frac{1}{n} S_n f \to \infty$ μ-a.e., since $f \geq f_N := \min\{f, N\}$ and $\frac{1}{n} S_n f_N \to \int_X f_N \, d\mu$ which is a diverging sequence as $N \to \infty$. Unfortunately does not exist a sequence $\{a_n\}$ such that the limit $\frac{1}{a_n} S_n f$ converges to a real non-zero number for μ-a.e. x. (see [1, Section 2.3]).

If f is non-summable with varying sign, then averaged Birkhoff sums may be undetermined. In figure 1 it is plotted the sequence $\{\frac{1}{n} S_n f(0)\}$ with $f(x) = \frac{\mathrm{sign}(x-\frac{1}{2})}{x(1-x)}$ for the rotation of angle $\theta = \sqrt{2}$, and $n = 1, \ldots, 1100$.

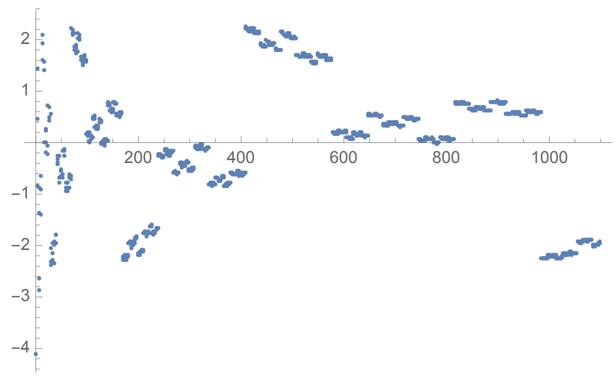

FIGURE 1. An averaged Birkhoff sum for a non-summable observable with varying sign under the rotation of angle $\theta = \sqrt{2}$.

Exercise 1. How many times does a given binary sequence of length ℓ appear in the dyadic representation of a "typical" real number in the first $N >> \ell$ symbols? *Answer:* approximately $\frac{N}{2^\ell}$ for big N.

Exercise 2. Study the asymptotic probability of appearance of $\ell \in \{1, \ldots, b-1\}$ as first digit of the numbers in the sequence $\{2^n\}_{n\geq 0}$ written in base $b \geq 2$. *Answer:* if $\log_b 2$ is irrational then ℓ appears with asymptotic probability $\log_b \frac{\ell+1}{\ell}$; if $\log_b 2$ is rational the first digits of $\{2^n\}_{n\geq 0}$ form a periodic sequence.

Example 1.17 (Continued fractions)**.** Let $F : [0,1] \to [0,1]$ be the Farey map which preserves the measure $d\mu(x) = \frac{1}{x}\, dx$, and consider the following construction. For $x \in [0,1]$ and $C = [\frac{1}{2}, 1]$ define the *hitting time*

$$\tau_C(x) := \min\{n \geq 0 \,:\, F^n(x) \in C\} \tag{1.2}$$

then[1] for all $x \neq 0$, $\tau(x)$ is finite. Then define $G : [0,1] \to [0,1]$ by

$$G(x) := \begin{cases} F^{1+\tau(x)}(x)\,, & \text{if } x \in (0,1] \\ 0\,, & \text{if } x = 0 \end{cases}$$

[1] $\tau(x)$ coincides with $\varphi_C(x)$, the first return time to C, outside C.

The map G is the well-known *Gauss map*, which satisfies $G(x) = \{\frac{1}{x}\}$ for $x \neq 0$, and is related to the continued fractions expansion of a real number. The map G preserves the probability measure $d\nu(x) = \frac{1}{\log 2} \frac{1}{x+1}\, dx$, which is also ergodic.

Let $x = [a_1, a_2, \dots, a_n, \dots]$ be the continued fractions expansion of an irrational number $x \in [0,1]$, and let $e : \mathbb{N} \to \{\pm 1\}$ be the parity function, $e(even) = +1$, $e(odd) = -1$. Then

$$\frac{1}{n} \sum_{k=1}^{n} e(a_k) = \frac{1}{n} \sum_{k=0}^{n-1} f(G^k(x))$$

where $f : [0,1] \to \{\pm 1\}$ is given on irrational numbers by

$$f(x) = \begin{cases} +1\,, & \text{if } x \in \bigcup_{k=1}^{\infty} \left(\frac{1}{2k+1}, \frac{1}{2k}\right) \\ -1\,, & \text{if } x \in \bigcup_{k=1}^{\infty} \left(\frac{1}{2k}, \frac{1}{2k-1}\right) \end{cases}$$

and satisfies

$$\int_0^1 f(x)\, d\mu(x) = \frac{\log 8 - \log \pi^2}{\log 2}.$$

Hence

$$\frac{1}{n} \sum_{k=1}^{n} e(a_k) \to \frac{\log 8 - \log \pi^2}{\log 2} \qquad \text{for } m\text{-a.e. } x \in [0,1]$$

Now show that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} a_k = \infty$$

for m-a.e. $x \in [0,1]$. Moreover (later we will better understand the result)

Theorem 1.18 ([4])**.** *For m-a.e. $x \in [0,1]$ there exists $n_0 = n_0(x)$ such that for all $n \geq n_0$*

$$\sum_{k=1}^{n} a_k = \Big(1 + o(1)\Big)\, n \log_2 n + c(n,x) \max_{1 \leq k \leq n} a_k$$

with $c(n,x) \in [0,1]$.

Let now study for $C = [\frac{1}{2}, 1]$ the average

$$\frac{1}{n} \sum_{k=0}^{n-1} \chi_C(F^k(x)) = \frac{\#\left\{k = 0, \dots, n-1 \,:\, F^k(x) \in C\right\}}{n}$$

Assuming $F^{n-1}(x) \in C$, we have that if $m = \sum_{k=0}^{n-1} \chi_C(F^k(x))$ then $n = \sum_{j=1}^{m} a_j$ where $x = [a_1, a_2, \dots, a_m, \dots]$. In general $n \geq \sum_{j=1}^{m} a_j$. Hence for m-a.e. x we have

$$0 \leq \frac{1}{n} \sum_{k=0}^{n-1} \chi_C(F^k(x)) \leq \frac{m}{\sum_{j=1}^{m} a_j} \to 0$$

as $n \to \infty$.

$\diamondsuit$

To obtain more results analogous to those for stochastic processes, we need the system to be "closer" to a sequence of i.i.d. random variables.

Definition 1.19. A probability T-invariant measure μ on $(X, \mathcal{B})$ is *mixing* if for all $A, B \in \mathcal{B}$ it holds

$$\mu(A \cap T^{-n} B) \to \mu(A)\, \mu(B) \tag{1.3}$$

as $n \to \infty$.

If a measure is mixing, then it is ergodic. Moreover condition (1.3) is equivalent to

$$\lim_{n\to\infty} \int_X f(x)\, g(T^n(x))\, d\mu(x) = \left(\int_X f d\mu \right) \left(\int_X g d\mu \right) \qquad \forall\, f \in L^1(X, \mu)\, , g \in L^\infty(X, \mu) \tag{1.4}$$

Condition (1.4) is more useful for the study of the speed of convergence in the limit, what is known as *rate of decay of correlations*. We will consider this problem in Section 3.

For a class of dynamical systems on $([0, 1], \mathcal{B}, m)$, including those in Examples 1.5 and 1.7 for $r \in [0, 1)$, the decay of correlations is exponential for functions of bounded variations. For this class of systems it holds

Theorem 1.20 (Central Limit Theorem)**.** *Let $f : [0, 1] \to \mathbb{R}$ be a function of bounded variation. Then on the space $([0, 1], \mathcal{B}, m)$ the sequence*

$$\frac{S_n f - n \int_X f d\mu}{\sqrt{n}}$$

converges in distribution to the Gaussian $\mathcal{N}(0, \sigma^2)$, where $\sigma = \sigma(f)$ and vanishes if and only if $f = g - g \circ T + const$ for some bounded variation function g.

The same result holds for the Pomeau-Manneville maps of Example 1.8 with $\alpha \in (0, \frac{1}{2})$ and C^1 observables. For $\alpha = \frac{1}{2}$, then the Central Limit Theorem holds with normalization $\sqrt{n \log n}$ if $f(0) \neq \int_X f d\mu_\alpha$, and with $\sqrt{n}$ otherwise. For $\alpha \in (\frac{1}{2}, 1)$, the convergence holds with normalization n^α if $f(0) \neq \int_X f d\mu_\alpha$, but not to a Gaussian law (see [7]). We will come back to these limit laws in Section 4.

Finally we introduce the notion of *exactness*.

Definition 1.21. A non-singular transformation T of a measurable space $(X, \mathcal{B}, m)$ is called *exact* if

$$A \in \bigcap_{n=1}^{\infty} T^{-n}\mathcal{B} \quad \Rightarrow \quad m(A)\, m(A^c) = 0\, .$$

If T preserves a probability measure μ, exactness of T with respect to (X, μ) is equivalent to

$$\lim_{n\to\infty} \mu(T^n A) = 1 \qquad \forall\, A \in \mathcal{B}\, ,\ \mu(A) > 0\, .$$

It can be proved that exactness implies mixing, and mixing implies ergodicity.

2. Inducing, invariant measures and ergodic theorems

A basic construction for the study of ergodic properties of a non-singular conservative transformation T on $(X, \mathcal{B}, m)$ is *inducing*. In this section we use [1, 14].

Let $C \in \mathcal{B}$ and recall the definition φ_{C} of first return time of C given in (1.1). By Theorem 1.12, φ_{C} is m-a.e. finite, hence it is well defined the *induced map* of T on C given by

$$T_C : C \to C\, , \qquad T_C(x) := T^{\varphi_C(x)}(x)$$

The construction of the induced map is similar to those used in Example 1.17 to define the Gauss map, which is called a *jump transformation* of the Farey map.

We introduce the following notations for the level sets of the return time of C:

$$C_k := \{ x \in C \,:\, \varphi_C(x) = k \} \quad \text{and} \quad C_{>k} := \{ x \in C \,:\, \varphi_C(x) > k \} \tag{2.1}$$

By definition the C_k's are disjoint and $C = \cup_{k\geq 1} C_k$, hence

$$T_C^{-1} A = \bigcup_{k\geq 1} C_k \cap T_C^{-1} A = \bigcup_{k\geq 1} C_k \cap T^{-k} A \quad \forall\, A \in \mathcal{B} \tag{2.2}$$

Proposition 2.1. *Let T be a conservative measure preserving and ergodic transformation of $(X, \mathcal{B}, \mu)$. If $C \in \mathcal{B}$ satisfies $0 < \mu(C) < \infty$, then T_C preserves the finite measure $\mu_C = \mu|_C$, hence it is conservative on $(C, \mathcal{B} \cap C, \mu_C)$, and it is ergodic.*

Proof. Let us consider $\mu_C(T_C^{-1} A)$ for $A \in \mathcal{B} \cap C$. By (2.2)

$$\mu_C(T_C^{-1} A) = \sum_{k\geq 1} \mu_C \left(C_k \cap T^{-k} A \right) = \sum_{k\geq 1} \mu_C \left(C \cap \left(T^{-k} A \setminus \cup_{j=1}^{k-1} T^{-j} C \right) \right)$$

Setting $A_0 = A$ and $A_k = T^{-k} A \setminus \cup_{j=0}^{k-1} T^{-j} C$ for $k \geq 1$, we get

$$C \cap \left(T^{-k} A \setminus \cup_{j=1}^{k-1} T^{-j} C \right) = C \cap T^{-1} A_{k-1}$$

and

$$T^{-1} A_k = \left(C \cap T^{-1} A_k \right) \cup \left(C^c \cap T^{-1} A_k \right) \quad \Rightarrow \quad \mu(A_k) = \mu\Big(C \cap T^{-1} A_k \Big) + \mu(A_{k+1})$$

since μ is T-invariant. It follows that

$$\mu_C(T_C^{-1} A) = \sum_{k\geq 1} \Big(\mu(A_{k-1}) - \mu(A_k) \Big) = \mu(A) - \lim_{k\to\infty} \mu(A_k) \leq \mu(A)$$

On the other side, using $T_C^{-1} A = C \setminus T_C^{-1}(C \setminus A)$, and applying the previous argument to $C \setminus A$ to get $\mu_C(T_C^{-1}(C \setminus A)) \leq \mu(C \setminus A)$, we obtain

$$\mu_C(T_C^{-1} A) = \mu_C(C) - \mu_C(T_C^{-1}(C \setminus A)) \geq \mu_C(C) - \mu(C \setminus A) = \mu(A)$$

Hence finally $\mu_C(T_C^{-1} A) = \mu_C(A)$, and we proved that μ_C is T_C-invariant.

We now show that μ_C is ergodic. Let $A \in \mathcal{B} \cap C$ be T_C-invariant and such that $\mu(A) > 0$. If $\mu(C \setminus A) > 0$, then for all $x \in C \setminus A$

$$0 = \sum_{k\geq 1} \chi_C(T_C^k(x)) = \sum_{k\geq 1} \chi_C(T^k(x)) = \infty$$

where in the last equality we have used Theorem 1.12. The proposition is proved. □

Proposition 2.2. *Let $C \in \mathcal{B}$ satisfy $\cup_{n\geq 1} T^{-n} C = X$ and assume that T_C preserves a probability measure ν and is ergodic. Then T has an invariant measure μ given by*

$$\mu(A) := \sum_{n\geq 0} \nu(C_{>n} \cap T^{-n} A)\,, \qquad A \in \mathcal{B}\,,$$

moreover T is conservative on $(X, \mathcal{B}, \mu)$ and ergodic.

Proof. First of all notice that by definition $\mu(C) = \nu(C)$. Moreover

$$\begin{aligned}
\mu(T^{-1}A) &= \sum_{n\geq 0} \nu\Big(C_{>n} \cap T^{-(n+1)}A\Big) = \\
&= \sum_{n\geq 0} \nu\Big(C_{n+1} \cap T^{-(n+1)}A\Big) + \sum_{n\geq 0} \nu\Big(C_{>n+1} \cap T^{-(n+1)}A\Big) = \\
&= \sum_{n\geq 1} \nu\Big(C_n \cap T^{-n}A\Big) + \sum_{n\geq 1} \nu\Big(C_{>n} \cap T^{-n}A\Big) = \\
&= \mu(A) + \sum_{n\geq 1} \nu\Big(C_n \cap T^{-n}A\Big) - \nu(C \cap A) = \\
&= \mu(A) + \nu(T_{C}^{-1}A) - \nu(C \cap A) = \mu(A)
\end{aligned}$$

where in the last equality we have used that ν is T_C-invariant. Hence μ is T-invariant.

We now show that T is conservative on $(X, \mathcal{B}, \mu)$. It is enough to show that all wandering sets have vanishing measure. Let $W \in \mathcal{B}$ be a wandering set and fix $n \geq 1$. Then, by the T-invariance of μ we have

$$\mu(C) = \mu(T^{-n}C) \geq \mu\left(T^{-n}C \cap \bigcup_{k=0}^{n} T^{-k}W\right) = \sum_{k=0}^{n} \mu(T^{-n}C \cap T^{-k}W) =$$

$$= \sum_{k=0}^{n} \mu(T^{-(n-k)}C \cap W) = \int_W \left(\sum_{k=0}^{n} \chi_C \circ T^{n-k}\right) d\mu = \int_W \left(\sum_{k=0}^{n} \chi_C \circ T^{k}\right) d\mu$$

Now we recall that $\mu(C) = \nu(C) = 1$, and by the assumption $\cup_{n\geq 1} T^{-n}C = X$ it follows that for all N

$$T^{-N}\left(\bigcup_{n\geq 1} T^{-n}C\right) = \bigcup_{n\geq N+1} T^{-n}C = X\,,$$

hence every orbit visits C infinitely often, that is

$$\sum_{k\geq 0} \chi_C \circ T^k = \infty$$

on X. Hence

$$\mu(C) \geq \int_W \left(\sum_{k=0}^{n} \chi_C \circ T^k\right) d\mu\,, \qquad \forall\, n \geq 1$$

can hold only if $\mu(W) = 0$.

It remains to prove that if ν is ergodic for T_C then μ is ergodic for T. Let $A \in \mathcal{B}$ be a T-invariant set. By the assumption $\cup_{n\geq 1} T^{-n}C = X$ necessarily $A \cap C$ is non-empty and is T_C-invariant, since by (2.2)

$$T_C^{-1}(A \cap C) = \bigcup_{k\geq 1} C_k \cap T^{-k}A = \bigcup_{k\geq 1} C_k \cap A = A \cap C\,.$$

Hence by ergodicity of ν either $\nu(A \cap C) = 0$ or $\nu(A^c \cap C) = 0$. In the first case, we get

$$\mu(T^{-n}C \cap A) = \mu(T^{-n}C \cap T^{-n}A) = \mu(T^{-n}(C \cap A)) = \mu(C \cap A) = 0\,, \qquad \forall\, n \geq 1$$

It follows that

$$\mu(A) = \mu\left(A \cap \bigcup_{n\ge 1} T^{-n}C\right) = 0\,.$$

Analogously in the second case. □

Similar results hold for the *jump transformation* $\tilde{T}_C : X \to X$ defined by $\tilde{T}_C(x) := T^{1+\tau_C(x)}(x)$.

Remark 2.3. If $T|_C$ is bijective onto X, then T is a conjugation between the induced map and the jump transformation. Check it for the Farey and the Gauss maps, and compute the different invariant measures.

2.1. **Ergodic theorem.** We have seen in Example 1.17 how Birkhoff ergodic theorem fails for transformations preserving an infinite measure. We now show that even if for a single observable it is impossible to obtain an almost everywhere convergence rate to a non-null limit, it is possible to obtain meaningful results comparing two different observables.

Theorem 2.4 (Hopf Ratio Ergodic Theorem)**.** *Let T be a conservative ergodic measure preserving transformation on a σ-finite measure space $(X, \mathcal{B}, \mu)$, then for all $f, g \in L^1(X,\mu)$ with g non-negative and $\int_X g\, d\mu > 0$, it holds*

$$\lim_{n\to\infty} \frac{S_n f(x)}{S_n g(x)} = \frac{\int_X f\, d\mu}{\int_X g\, d\mu}$$

μ-a.e. on X.

Corollary 2.5. *Let T be a conservative ergodic measure preserving transformation on a σ-finite measure space $(X, \mathcal{B}, \mu)$ with $\mu(X) = \infty$. Then for all $f \in L^1(X,\mu)$ it holds*

$$\lim_{n\to\infty} \frac{1}{n}\, S_n f \to 0 \qquad \mu - a.e.$$

Proof. For any fixed $N \ge 1$, let $A_N \in \mathcal{B}$ with $\mu(A_N) \ge N$. Then we apply Theorem 2.4 to $f \in L^1(X,\mu)$ and $g = \chi_{A_N}$ to get

$$0 \le \limsup_{n\to\infty} \frac{1}{n}\, S_n f(x) \le \lim_{n\to\infty} \frac{S_n f(x)}{S_n g(x)} = \frac{\int_X f\, d\mu}{N}$$

for μ-a.e. x. Letting $N \to \infty$ we obtain the result. □

Let T be a conservative ergodic measure preserving transformation on a σ-finite measure space $(X, \mathcal{B}, \mu)$, and let $C \in \mathcal{B}$ with $0 < \mu(C) < \infty$. We consider the induced map T_C and its invariant probability measure $\frac{1}{\mu(C)}\, \mu_C$. There is a relation between the Birkhoff sums of the two maps.

Let $f : X \to \mathbb{R}$ be a measurable function, $S_n f$ are its Birkhoff sums for T, and denote by $S^C_n f$ its Birkhoff sums for T_C. As an example, we introduce

$$\varphi_m(x) := S^C_m \varphi_C(x) = \sum_{j=0}^{m-1} \varphi_C(T^j_C(x)) \qquad \text{for } x \in C$$

which is the time of the m-th return of x to C, and the difference $\varphi_k(x) - \varphi_{k-1}(x) = \varphi_C(T^{k-1}_C(x))$ is the length of the k-th excursion outside C.

Lemma 2.6. *Let $f : X \to \mathbb{R}$ be a non-negative measurable function, and let $f^C : C \to \mathbb{R}$ be $f^C = S_{\varphi_C} f$. Then*

$$S_{\varphi_m} f(x) = S^C_m f^C(x)$$

for all $m \ge 1$ and μ-a.e. $x \in C$. Moreover $\int_X f\, d\mu = \int_C f^C\, d\mu$.

Proof. By definition

$$S_{\varphi_2} f(x) = \sum_{k=0}^{\varphi_2(x)-1} f(T^k(x)) = S_{\varphi_1} f(x) + \sum_{k=\varphi_1(x)}^{\varphi_2(x)-1} f(T^k(x)) = S_{\varphi_1} f(x) + \sum_{k=0}^{\varphi_2(x)-\varphi_1(x)-1} f(T^k(T_{C}(x)))$$

hence $S_{\varphi_2} f = S_{\varphi_1} f + S_{\varphi_2 - \varphi_1}(f \circ T_{C})$. In general

$$S_{\varphi_m} f = S_{\varphi_1} f + \sum_{j=1}^{m-1} S_{\varphi_{j+1}-\varphi_j}(f \circ T_{C}^j) = S_{\varphi_1} f + \sum_{j=1}^{m-1} S_{\varphi_{C} \circ T_{C}^j}(f \circ T_{C}^j) =$$

$$= \sum_{j=0}^{m-1} (S_{\varphi_{C}} f) \circ T_{C}^j = S_m^{C}(S_{\varphi_{C}} f) = S_m^{C} f^{C} .$$

Moreover, let first $f = \chi_A$ for $A \in \mathcal{B}$. Then by the relation between μ and the invariant measure for T_{C} of Proposition 2.2, we have

$$\mu(A) = \sum_{n \geq 0} \mu(C_{>n} \cap T^{-n} A)$$

hence

$$\int_X \chi_A \, d\mu = \int_X \left(\sum_{n \geq 0} \chi_{C_{>n}} \chi_A \circ T^n \right) d\mu = \int_C \left(\sum_{k=0}^{\varphi_{C}-1} \chi_A \circ T^k \right) d\mu = \int_C \chi_A^{C} \, d\mu \, .$$

For a measurable function the result follows by standard approximation. □

Let now $f \equiv 1$, hence $f^{C} = \varphi_{C}$, applying Lemma 2.6 we get Kac Theorem 1.14.

Proof of Theorem 2.4. Let us assume that f is non-negative, and recall the function f^{C} on C. By assumption, T_{C} preserves $\frac{1}{\mu(C)} \mu|_{C}$ and is ergodic, hence by Lemma 2.6 for μ-a.e. $x \in C$

$$\lim_{m\to\infty} \frac{1}{m} S_m^{C} f^{C}(x) = \lim_{m\to\infty} \frac{S_{\varphi_m} f(x)}{S_{\varphi_m} \chi_C(x)} = \frac{1}{\mu(C)} \int_C f^{C} \, d\mu = \frac{\int_X f \, d\mu}{\mu(C)} \tag{2.3}$$

Now for any $n \geq 1$ and any fixed $x \in C$, there exists $m = m(n, x)$ such that $\varphi_{m-1}(x) + 1 \leq n \leq \varphi_m(x)$, and since f is non-negative

$$\frac{S_{m-1}^{C} f^{C}(x)}{m} \leq \frac{S_n f(x)}{S_n \chi_C(x)} \leq \frac{S_m^{C} f^{C}(x)}{m-1}$$

From (2.3), it follows that

$$\lim_{n\to\infty} \frac{S_n f(x)}{S_n \chi_C(x)} = \frac{\int_X f \, d\mu}{\mu(C)} \tag{2.4}$$

for μ-a.e. $x \in C$. Moreover, the set on which (2.4) holds is T-invariant and contains C, hence by ergodicity of T it is X (mod μ).

The proof is finished by applying (2.4) to two functions f, g and considering the ratio of their Birkhoff sums. □

Exercise 3. Going back to the Farey map, let $C = [\frac{1}{2}, 1]$ and $A = [\frac{1}{3}, \frac{1}{2}]$. The set C corresponds to real numbers with continued fractions expansion with $a_1 = 1$, and the set A to those with $a_1 = 2$. The number of visits of a "typical" orbit $\{F^k(x)\}_{k=0}^{n-1}$ of the Farey map to C corresponds for $x = [a_1, a_2, \dots]$ to $\max\{j \geq 1 \, : \, a_1 + \dots a_j < n\}$, and is equal to the number of visits to C in the first $n-1$ iterations. The number of visits to A corresponds instead to the number of visits to C minus the number of times that '1' appears as a coefficient. Show that the asymptotic ration between the number of visits to C and the number of visits to A in the first $n-1$ iterations of the Farey map is given by $\frac{\log 2}{\log \frac{3}{2}}$.

2.2. **The tail of the first return time.** We have seen in Kac Theorem 1.14, that transformations with an infinite invariant measure have non-summable return times to finite measure sets. Using notations (2.1) we write

$$\mu(X) = \int_C \varphi_C \, d\mu = \sum_{k\geq 1} k\, \mu(C_k) = \sum_{n\geq 0} \mu(C_{>n})$$

hence if $\mu(X) = \infty$, the sequence $\{\mu(C_{>n})\}$ is non-summable.

Proposition 2.7. *Let T be a measure preserving transformation on the σ-finite set $(X, \mathcal{B}, \mu)$. Let $C \in \mathcal{B}$ satisfy $\mu(C) < \infty$ and $\cup_{n\geq 1} T^{-n}C = X$, then*

$$\mu(C_{>n}) = \mu\{x \in X \,:\, \tau_C(x) = n\} \quad \forall\, n \geq 1\,,$$

where τ_C is the hitting time defined in (1.2).

Proof. Let us first notice that

$$T^{-1}C = C_1 \cup \{x \in X \,:\, \tau_C(x) = 1\}$$

with a disjoint union. Hence we can write another disjoint union for

$$T^{-2}C = T^{-1}C_1 \cup C_2 \cup \{x \in X \,:\, \tau_C(x) = 2\}$$

Hence for $n \geq 1$ we write the decomposition

$$T^{-n}C = \bigcup_{k=1}^{n} T^{-(n-k)} C_k \cup \{x \in X \,:\, \tau_C(x) = n\}$$

and the T-invariance of μ yields

$$\mu(C) = \sum_{k=1}^{n} \mu(C_k) + \mu\{x \in X \,:\, \tau_C(x) = n\}$$

from which the thesis follows since $\mu(C_{>n}) = \mu(C) - \sum_{k=1}^{n} \mu(C_k)$. □

Example 2.8. For the Pomeau-Manneville maps of Example 1.8, let $\bar{x}_\alpha \in (0,1)$ satisfy $\bar{x}_\alpha + \bar{x}_\alpha^{1+\alpha} = 1$. Choosing $C = [\bar{x}_\alpha, 1]$, it follows

$$\{x \in X \,:\, \tau_C(x) = n\} = [x_n, x_{n-1}]$$

where $x_0 = \bar{x}_\alpha$ and $T_\alpha^n(x_n) = x_0$. As $n \to \infty$, it holds $x_n \approx n^{-\frac{1}{\alpha}}$, hence

$$\mu_\alpha\{x \in X \,:\, \tau_C(x) = n\} = \int_{x_n}^{x_{n-1}} h_\alpha(x)\, dx \approx \frac{1}{x_n^{\alpha-1}} - \frac{1}{x_{n-1}^{\alpha-1}} \approx n^{-\frac{1}{\alpha}}$$

It follows as already stated above, that

$$\sum_{n\geq 0} \mu_\alpha(C_{>n}) \approx \sum_{n\geq 1} n^{-\frac{1}{\alpha}}$$

converges if and only if $\alpha < 1$. For this reason the case $\alpha = 1$ is called *barely infinite*. It corresponds also to the case of the Farey map for which $x_n = \frac{1}{n+2}$.

♢

Example 2.9. For the Boole map of Example 1.9, let $C = [-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$, and $x_0 = \frac{1}{\sqrt{2}}$. Then

$$\{x \in X \,:\, \tau_C(x) = n\} = [-x_n, -x_{n-1}] \cup [x_{n-1}, x_n]$$

with $T(x_n) = x_{n-1}$. From the definition it follows $x_n \to \infty$ and $x_{n-1} = x_n - \frac{1}{x_n}$, hence

$$x_n^2 - x_{n-1}^2 = 2 - \frac{1}{x_n^2} \to 2$$

and

$$\frac{x_n^2}{n} = \frac{x_0^2}{n} + \frac{1}{n}\sum_{k=1}^{n}(x_k^2 - x_{k-1}^2) \to 2$$

It follows that $x_n \sim \sqrt{2n}$ and $m(C_{>n}) \sim \sqrt{\frac{2}{n}}$.

$\diamondsuit$

3. The transfer operator

We now come back to the problem of getting "stronger" properties as explained in Section 1, in particular *mixing* and *decay of correlations*. The main tool to get a quantitative behaviour is the *transfer operator*.

Let T be a non-singular transformation of $(X, \mathcal{B}, m)$.

Definition 3.1. The *transfer operator* $\mathcal{P}$ associated to T is the linear operator $\mathcal{P} : L^1(X, m) \to L^1(X, m)$ defined as the dual of the composition operator $g \mapsto g \circ T$ on $L^\infty(X, m)$ by

$$\int_X g\,(\mathcal{P}f)\,dm = \int_X (g \circ T)\,f\,dm \qquad \forall\, f \in L^1(X, m)\,,\ g \in L^\infty(X, m) \tag{3.1}$$

The domain of definition of $\mathcal{P}$ can be extended to all non-negative measurable functions by

$$f \mapsto \frac{d(\mathcal{B} \ni A \mapsto \int_{T^{-1}A} f\,dm)}{dm}$$

Proposition 3.2. *(i) The transfer operator $\mathcal{P}$ is a bounded positive linear operator of norm 1.*

(ii) Let $\mu << m$ with density h, then

$$\mu \text{ is } T\text{-invariant} \iff \mathcal{P}h = h$$

(iii) If T is conservative and ergodic on $(X, \mathcal{B}, m)$, then there is at most one σ-finite T-invariant measure $\mu << m$, up to a multiplicative constant.

Proof. (i) By (3.1), positiveness is obvious. Moreover $\|\mathcal{P}f\| = \|f\|$ for non-negative functions, hence in general $\|\mathcal{P}f\| \le \|f\|$, and the norm is clearly 1.

(ii) For $A \in \mathcal{B}$

$$\mu(T^{-1}A) = \int_X (\chi_A \circ T)\,h\,dm = \int_X \chi_A\,(\mathcal{P}h)\,dm = \int_A \mathcal{P}h\,dm$$

Hence $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$ if and only if $\mathcal{P}h = h$ (mod m).

(iii) First, observe that if $\mu << m$ is T-invariant then $\mu \sim m$. Indeed $\mu(T^{-1}[h = 0]) = \mu([h = 0]) = 0$, hence $T^{-1}[h = 0] \subset [h = 0]$ (mod m). If $[h = 0]$ has positive m measure, then $T^{-k}[h = 0] \subset [h = 0]$ for all $k \ge 1$, and $\sum_{k\ge1} \chi_{[h=0]} \circ T^k = 0$ outside $[h = 0]$, which is absurd by conservativeness and ergodicity. Hence, we can assume $\mu = m$.

We first consider the case $m(X) = 1$ and and show that $\mathcal{P}h = h$ (mod m), for $h \in L^1(X, m)$, implies h is constant. Hence we show uniqueness of a T-invariant probability measure absolutely continuous to a reference measure.

For any $A \in \mathcal{B}$ with $m(A) < \infty$, since $\|\frac{1}{n}\sum_{k=0}^{n-1} \chi_A \circ T^k\|_\infty \le 1$ for all $n \ge 1$, there exists a subsequence $\{n_j\}$ and a function $g_A \in L^\infty(X,m)$ such that

$$\frac{1}{n_j}\sum_{k=0}^{n_j-1} \chi_A \circ T^k \quad \stackrel{\star}{\rightharpoonup} \quad g_A$$

in the weak * topology of L^∞. Moreover g_A is T-invariant, hence m-a.e. constant by ergodicity, and

$$\int_X g_A\, dm = \lim_{j\to\infty} \int_X \left(\frac{1}{n_j}\sum_{k=0}^{n_j-1} \chi_A \circ T^k \right) dm = m(A)$$

by T-invariance of m. It follows that $g_A \equiv m(A)$ m-a.e. Finally

$$\int_A h\, dm = \int_X \chi_A \left(\frac{1}{n_j}\sum_{k=0}^{n_j-1} \mathcal{P}^k h \right) dm = \int_X \left(\frac{1}{n_j}\sum_{k=0}^{n_j-1} \chi_A \circ T^k \right) h\, dm \to m(A) \int_X h\, dm$$

hence h is m-a.e. constant.

In the general case, if $\nu << m$ is a σ-finite T-invariant measure, then chosen $C \in \mathcal{B}$ with $0 < m(C), \nu(C) < \infty$, by what we have shown above, the induced T_C-invariant probability measures m_C and ν_C coincide (up to a multiplicative constant we can assume that $m(C) = \nu(C) = 1$). Hence by Proposition 2.2, also m and ν coincide. □

Remark 3.3. Let T be a piecewise C^1 invertible transformation of a Euclidean set (X,m), that is there exists a partition of open sets $\mathcal{I} = \{I_j\}_{j\in J}$, with J finite or countable, such that $m(I_j \cap I_{\tilde{j}}) = 0$, $X = \cup_j I_j$ (mod m), $T \in C^1(I_j)$ and $T|_{I_j}$ is injective. Let us denote by $\phi_j = (T|_{I_j})^{-1} : T(I_j) \to I_j$.

Then for all measurable $f : X \to \mathbb{R}^+$ and $A \in \mathcal{B}$ it holds

$$\int_{T^{-1}A} f\, dm = \sum_{j\in J} \int_{\phi_j(A\cap T(I_j))} f\, dm = \sum_{j\in J} \int_{A\cap T(I_j)} f \circ \phi_j\, |\det J\phi_j|\, dm$$

It follows that the transfer operator can be written as

$$\mathcal{P}f(x) = \sum_{y\in T^{-1}(x)} \frac{f(y)}{|\det JT(y)|} \tag{3.2}$$

Using (3.2) and Proposition 3.2-(ii), we can verify the existence of invariant measures for the Examples 1.5-1.9.

Example 3.4. Let $X = [0,1]$ and T be a piecewise C^1 invertible transformation with $\mathcal{I} = \{I_0, I_1\}$ and $T(I_j) = [0,1]$ for all $j = 0, 1$. Then

$$\mathcal{P}f(x) = \sum_{j\in J} f(\phi_j(x))\, |\phi_j'(x)|$$

Let h be the density of a T-invariant measure $\mu << m$, and letting $C := I_1$ assume $0 < \mu(C) < \infty$. We consider then the jump transformation T_C, and write the transfer operator $\mathcal{P}_C$ associated to T_C. The transformation T_C is piecewise C^1 invertible on the partition $\{C_n\}_{n\ge 0}$ of level sets of the hitting time τ_C. Moreover letting $\psi_n := (T_C|_{C_n})^{-1} : [0,1] \to C_n$, it holds

$$\psi_n = (\phi_0)^n \circ \phi_1\,,$$

hence

$$\mathcal{P}_C g = \sum_{n\ge 0} g \circ \psi_n\, |\psi_n'| = \sum_{n\ge 0} (g \circ (\phi_0)^n \circ \phi_1)\, |((\phi_0)^n \circ \phi_1)'| = \sum_{n\ge 0} (\mathcal{P}_1 \circ \mathcal{P}_0^n) g$$

where

$$\mathcal{P}_0 f := f \circ \phi_0 \, |\phi_0'| \qquad \mathcal{P}_0 f := f \circ \phi_1 \, |\phi_1'|$$

We obtain the operator equality

$$(1 - \mathcal{P}_C) \circ (1 - \mathcal{P}_0) = 1 - \mathcal{P}$$

from which it follows

Proposition 3.5. *If the measure $d\mu = h\,dm$ is T-invariant then the jump transformation T_C preserves the measure $d\mu_C = k\,dm$ with*

$$k = (1 - \mathcal{P}_0) h\,.$$

Viceversa, if the measure $d\mu_C = k\,dm$ is T_C-invariant then T preserves the measure $d\mu = h\,dm$ with

$$h = \sum_{n \geq 0} \mathcal{P}_0^n k\,.$$

This example applies in particular to the Farey and Gauss maps.

$\diamondsuit$

We now turn to the connections with ergodic properties. Let T be a non-singular transformation of $(X, \mathcal{B}, m)$ and assume that $d\mu = hdm$ is a σ-finite T-invariant measure.

Proposition 3.6. *(i) T is conservative and ergodic if and only if $\sum_{n\geq 0} \mathcal{P}^n f = \infty$ m-a.e., for all non-negative $f \in L^1(X,m)$ with $\int_X f\,dm > 0$.*

(ii) If μ is finite, T is mixing if and only if $\mathcal{P}^n f \rightharpoonup h \int_X f dm$ weakly in L^1 for all $f \in L^1(X,m)$. Moreover, if $\mathcal{P}^n f \to h \int_X f dm$ in L^1, the rate of decay of correlations in (1.4) is measured by $\|\mathcal{P}^n f - h \int_X f dm\|_{L^1}$.

(iii) T is exact if and only if $\|\mathcal{P}^n f\|_{L^1} \to 0$ for all $f \in L^1(X,m)$ with $\int_X f\,dm = 0$.

Proof. (i) Using (3.1) with $f \in L^1(X,m)$ and $g = \chi_A$ for $A \in \mathcal{B}$ with $m(A) > 0$, it holds

$$\int_A \Big(\sum_{k=0}^{n-1} \mathcal{P}^k f \Big)\, dm = \int_X \Big(\sum_{k=0}^{n-1} \chi_A \circ T^k \Big)\, f\, dm$$

and the last integral is infinite by conservativity and ergodicity.

(ii) By (1.4) and (3.1), T is mixing if and only if for all $F \in L^1(X,\mu)$ and $g \in L^\infty(X,\mu)$

$$\Big(\int_X F\,d\mu \Big) \Big(\int_X g\,d\mu \Big) = \lim_{n\to\infty} \int_X F(x)\, g(T^n(x))\, d\mu(x) = \lim_{n\to\infty} \int_X (\mathcal{P}^n(Fh))\, g\, dm$$

and we recall that $h > 0$ (mod m)and $Fh \in L^1(X,m)$. Moreover

$$\left| \int_X F(x)\, g(T^n(x))\, d\mu(x) - \Big(\int_X F\,d\mu \Big) \Big(\int_X g\,d\mu \Big) \right| \leq \int_X \left| \mathcal{P}^n(Fh) - h \int_X F\,d\mu \right| g\, dm$$

(iii) Assume T to be exact and consider a given $f \in L^1(X,m)$ with $\int_X f\,dm = 0$. Since $L^\infty = (L^1)'$, there exists a sequence $\{g_n\}$ in L^∞ with $\|\mathcal{P}^n f\|_{L^1} = \int_X g_n\, \mathcal{P}^n f\, dm = \int_X (g_n \circ T^n)\, f\, dm$. Since $\mathcal{P}$ has norm 1, it follows that $\{g_n \circ T^n\}$ is a bounded sequence in L^∞, whence up to subsequence it converges to some $g \in L^\infty$ in weak* sense, and g is then measurable with respect to the tail σ-algebra $\cap_{n=1}^\infty T^{-n}\mathcal{B}$, hence it is constant μ-a.e. It follows that

$$\lim_{n\to\infty} \|\mathcal{P}^n f\|_{L^1} = \lim_{n\to\infty} \int_X (g_n \circ T^n)\, f\, dm = g \int_X f\, dm = 0$$

Assume now that $\|\mathcal{P}^n f\|_{L^1} \to 0$ for all $f \in L^1(X,m)$ with $\int_X f\,dm = 0$, and argue by contradiction. If T is not exact, there exists $A \in \cap_{n=1}^{\infty} T^{-n}\mathcal{B}$ with $m(A) > 0$ and $m(A^c) > 0$. Hence we can find $f \in L^1(X,m)$ with $\int_X f\,dm = 0$ and $\int_A f\,dm > 0$. For the sequence $\{A_n\} \subset \mathcal{B}$ satisfying $A = T^{-n}A_n$ it holds

$$\|\mathcal{P}^n f\|_{L^1} \geq \int_{A_n} \mathcal{P}^n f\,dm = \int_X \chi_{A_n}\mathcal{P}^n f\,dm = \int_X \chi_A f\,dm > 0$$

since $\chi_{A_n} \circ T^n = \chi_A$. We get the contradiction. $\square$

Corollary 3.7. *If $\mu(X) = 1$ and T is exact, then it is mixing.*

Proof. If T is exact then for $f \in L^1(X,m)$ we write $f = f - h\int_X f\,dm + h\int_X f\,dm$, and notice that $\int_X (f - h\int_X f\,dm)\,dm = 0$. Hence

$$\mathcal{P}^n f = \mathcal{P}^n(f - h\int_X f\,dm) + \mathcal{P}^n(h\int_X f\,dm) = \mathcal{P}^n(f - h\int_X f\,dm) + h\int_X f\,dm \to h\int_X f\,dm$$

in L^1. $\square$

Proposition 3.8. *If μ is infinite and T is exact, then $\int_A \mathcal{P}^n f\,dm \to 0$ for all $A \in \mathcal{B}$ with $\mu(A) < \infty$ and all non-negative $f \in L^1(X,m)$.*

Proof. Fix $N \in \mathbb{N}$ and $C \in \mathcal{B}$ with $N \leq \mu(C) \leq \infty$. For $f \in L^1(X,m)$ let us write $f = f - h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm + h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm$. Since $\int_X \left(f - h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm\right)dm = 0$, it follows

$$0 \leq \int_A \mathcal{P}^n f\,dm \leq \left\|\mathcal{P}^n\left(f - h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm\right)\right\|_{L^1} + \int_X \chi_A \left|\mathcal{P}^n\left(h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm\right)\right| dm$$

Since T is exact, the first term on the right vanishes as $n \to \infty$. Morevoer

$$\int_X \chi_A \left|\mathcal{P}^n\left(h\,\frac{\chi_C}{\mu(C)}\int_X f\,dm\right)\right| dm = \frac{\mu(C \cap T^{-n}A)}{\mu(C)}\int_X f\,dm \leq \frac{\mu(A)}{N}\int_X f\,dm$$

Hence $\int_A \mathcal{P}^n f\,dm \to 0$. $\square$

Remark 3.9. For relations between the transfer operator and the central limit theorem, and stronger probabilistic results for dynamical systems, we refer to [7].

4. When and how to recover some probabilistic results for infinite measures

We have seen that asking to a dynamical systems to be "close" to a sequence of i.i.d random variables improves the result that we can obtain for the behaviour of Birkhoff sums of observables. Unfortunately, this approach breaks down if the system preserves an infinite measure (see Corollary 2.5 and Proposition 3.8). In this last section we give a glimpse of how to try to circumvent these problems by asking for stronger properties or by changing definitions.

4.1. **Distributional limit.** Let T be a conservative ergodic measure preserving transformation on $(X,\mathcal{B},\mu)$. We have shown in Corollary 2.5 that the strong law of large numbers for $\{f \circ T^n\}$, where f is a summable observable, fails if $\mu(X) = \infty$. Weaker convergence laws can be obtained under additional assumptions on the systems. For more detail we refer to [1, Chapter 3].

It turns out that, contrarily to the Birkhoff sums of an observable, it is possible in some cases to control the sums of iterates of the transfer operator. This is due to the "regularizing" nature of the transfer operator occurring for example in systems which locally expand the state space (see Figure 2).

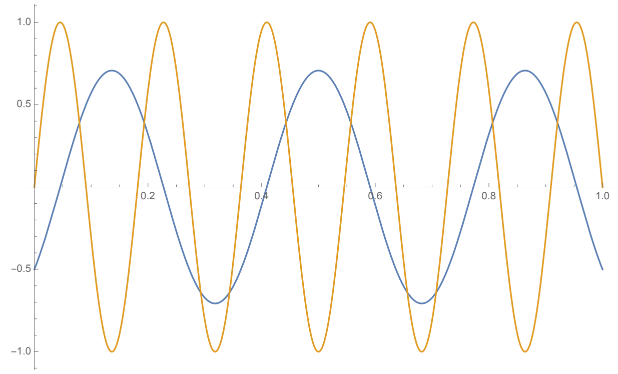

FIGURE 2. Let $f(x) = \sin(11\pi x)$. We plot f in orange and $\mathcal{P}f$ in blue, where $\mathcal{P}$ is the transfer operator of the doubling map of Example 1.5.

Definition 4.1. The transformation T is *pointwise dual ergodic* if there exists a sequence $\{a_n(T)\}$ such that

$$\frac{1}{a_n(T)} \sum_{k=0}^{n-1} \mathcal{P}^n f \to \int_X f\, d\mu\,, \qquad \mu - a.e.$$

for all $f \in L^1(X, \mu)$.

We now introduce a method to prove pointwise dual ergodicity, which is again based on inducing. We first need a definition, which is a strengthening of the notion of mixing.

Definition 4.2. Let S be a measure preserving transformation of the probability space $(Y, \mathcal{B}, \nu)$, and $\mathcal{C} \subset \mathcal{B}$ be a countable measurable partition which is generating for S. Let us denote by $\mathcal{C}^k$ the iterated partitions, that is $\mathcal{C}^k = \vee_{j=0}^{k-1} S^{-j}\mathcal{C}$. The system $(Y, \mathcal{B}, S, \nu, \mathcal{C})$ is said to be ψ*-mixing* (or continued-fraction mixing) if the numbers

$$\psi_n := \sup_{C \in \mathcal{C}^k,\, A \in \mathcal{B},\, \nu(A)>0} \frac{\left|\nu(C \cap T^{-(k+n)}A) - \nu(C)\nu(A)\right|}{\nu(C)\nu(A)}$$

satisfy $\psi_n \to 0$ as $n \to \infty$.

Proposition 4.3. *Let T be a conservative, ergodic measure preserving transformation of the space $(X, \mathcal{B}, \mu)$. Let $C \in \mathcal{B}$ with $0 < \mu(C) < \infty$ and let T_C be the induced map. Let $\mathcal{C} \subset \mathcal{B} \cap A$ be a countable measurable partition which is generating $\mathcal{B}$ under T_C, such that φ_C is $\mathcal{C}$-measurable. If $(C, \mathcal{B} \cap C, \frac{\mu|_C}{\mu(C)}, T_C, \mathcal{C})$ is ψ-mixing, then T is pointwise dual ergodic.*

We say (with a slight abuse of notation) that a subset $C \in \mathcal{B}$ satisfying the assumptions of Proposition 4.3, is a *Darling-Kac set*.

Example 4.4 (Continued fractions again...)**.** Proposition 4.3 can be applied to the Farey map and the Gauss map, its jump transformation on $C = [\frac{1}{2}, 1]$ (which is related to the induced map T_C, see Remark 2.3). In particular the Gauss map is ψ-mixing (see [8, Chapter 5]), hence the Farey map is pointwise dual ergodic.

$\diamondsuit$

The existence of a Darling-Kac set is useful also to find the asymptotic behaviour of the sequence $\{a_n(T)\}$. We recall that a sequence $\{b_n\}$ is called *regularly varying of index* $\rho \in \mathbb{R}$ if $b_n = n^\rho\, \ell(n)$, where $\{\ell(n)\}$ satisfies $\frac{\ell(cn)}{\ell(n)} \to 1$ as $n \to \infty$ for all $c > 0$.

Proposition 4.5. *Let T be a conservative, ergodic measure preserving transformation of the space $(X,\mathcal{B},\mu)$ and let $C\in\mathcal{B}$ be a Darling-Kac set. If the sequence $w_n(C) := \mu(C)+\sum_{k=1}^{n-1}\mu(C_{>k})$, called the* wandering rate of C*, is regularly varying with index $1-\alpha$, for some $\alpha\in[0,1]$, then*

$$a_n(T) \sim \frac{1}{\Gamma(2-\alpha)\Gamma(1+\alpha)}\,\frac{n}{w_n(C)}$$

We can finally state the main result.

Theorem 4.6. *Let T be a conservative, ergodic measure preserving transformation of the space $(X,\mathcal{B},\mu)$ and let $C\in\mathcal{B}$ be a Darling-Kac set. If the wandering rate $w_n(C)$ is regularly varying with index $1-\alpha$, for some $\alpha\in[0,1]$, then for all $f\in L^1(X,\mu)$ and all $t>0$ it holds*

$$\frac{1}{\mu(C)}\,\mu|_C\left[\frac{1}{a_n(T)}\,S_nf\le t\right]\to\mathbb{P}\left[\mathcal{M}_\alpha\le\frac{t}{\int_X f\,d\mu}\right]$$

as $n\to\infty$, where $\mathcal{M}_\alpha$ is a non-negative real random variable distributed according to the normalized Mittag-Leffler distribution of order α.

Similar results can be obtained under weaker assumptions on T, as for example in [15].

Remark 4.7. Mittag-Leffler distributions are ubiquitous in Infinite Ergodic Theory. We refer to [13] for their appearance in a Central Limit Theorem for a class of infinite measure preserving transformations.

Example 4.8 (...and continued fractions again)**.** We continue Example 4.4, applying Proposition 4.5 and Theorem 4.6 to the Farey map F on $([0,1],\mathcal{B},\mu)$, with $d\mu(x)=\frac{1}{\log 2}\frac{1}{x}dx$ and its Darling-Kac set $C=[\frac{1}{2},1]$ (the normalization of μ gives $\mu(C)=1$).

First of all, by Proposition 2.7

$$\mu(C_{>n})=\mu\left\{x\in[0,1]\,:\,\tau_C=n\right\}=\mu\left(\left(\frac{1}{n+2},\frac{1}{n+1}\right)\right)=\log_2\left(1+\frac{1}{n+1}\right)$$

(see Example 2.8), hence for the wandering rate we find

$$w_n(C)=1+\sum_{k=1}^{n-1}\log_2\left(1+\frac{1}{k+1}\right)=\log_2(n+1)$$

which is regularly varying of index 0. Hence we apply Proposition 4.5 and obtain that the Farey map is pointwise dual ergodic with sequence

$$a_n(F)\sim\frac{n}{\log_2 n}$$

Let us now come to the distributional behaviour of Theorem 4.6. We can apply it to the Farey map with $\alpha=1$, and since $\mathcal{M}_1$ turns out to be a constant random variable, convergence in distribution implies convergence in probability. Hence we obtain that for all $f\in L^1(X,\mu)$ and all $\varepsilon>0$ it holds

$$\nu\left(\left|\frac{\log_2 n}{n}\,S_nf-\int_X f\,d\mu\right|\ge\varepsilon\right)\to 0$$

as $n\to\infty$, for any probability measure $\nu<<m$. We have thus shown that in the barely infinite case, we can obtain a weak law of large numbers.

Applying the previous results to χ_C give information about the return times to C, and in duality information about the coefficients of the continued fractions expansion of a real number.

Corollary 4.9 (Khintchin's weak law). *The coefficients $\{a_n(x)\}$ of the continued fractions expansion of a number $x \in [0,1]$ satisfy*

$$\nu\left(\left|\frac{1}{n\,\log_2 n}\sum_{k=0}^{n-1} a_k(x) - 1\right| \geq \varepsilon\right) \to 0$$

as $n \to \infty$, for any probability measure $\nu << m$.

Recall Theorem 1.18 to have a strengthening of this result.

$\diamondsuit$

4.2. **Scaling rates and infinite mixing.** We have seen that to study the distributional behaviour of Birkhoff sums, we have considered convergence of a normalized mean of the iterations of the transfer operator. If we want to talk about a "mixing" property we need to consider instead the convergence of the iterations of the transfer operator.

It turns out that for transformations T conservative and ergodic on $(X, \mathcal{B}, m)$, and preserving a σ-finite infinite measure $\mu << m$, under strong assumptions on the behaviour of T on a Darling-Kac set (essentially, we need to have regularly varying wandering rates and nice spectral properties for the transfer operator of the induced map), it is possible to rescale $\{\mathcal{P}^n f\}$ for nice $f \in L^1(X,\mu)$ to obtain a finite limit (recall Proposition 3.8). In particular, this applies for the Farey map and the Pomeau-Manneville maps with $\alpha \geq 1$.

We now state a result from [6, 10].

Theorem 4.10. *Let T be a map of the family of the Pomeau-Manneville maps with $\alpha \geq 1$ or the Farey map. There exists a constant c, depending only on T, such that for all Lipschitz functions f supported on $(0,1]$ it holds*

$$n^{1-\frac{1}{\alpha}}\,\mathcal{P}^n f \to c\int_X f\,d\mu \qquad \textit{uniformly on compact subsets of } (0,1]$$

for the Pomeau-Manneville maps with $\alpha > 1$; and

$$\log n\,\mathcal{P}^n f \to c\int_X f\,d\mu \qquad \textit{uniformly on compact subsets of } (0,1]$$

for the Pomeau-Manneville maps with $\alpha = 1$ and for the Farey map.

Notice that the normalizing sequence is nothing but the wandering rate of the Darling-Kac set we found in Example 2.8. The asymptotic behaviour of this sequence is also called the *scaling rate* of T, and it is one way to recover a "mixing" notion in Infinite Ergodic Theory.

Unfortunately with Theorem 4.10 we can't get information for the behaviour of $\int_X (g \circ T^n)\, f\, dm$ for a general $g \in L^\infty$. Of course, if $\mu(X) = \infty$, also $\int_X g\,d\mu$ might be infinite (so we have to change definition (1.4) anyway). We now introduce the approach to the problem initially studied in [9], and applied in [2] to some maps of the interval. In particular it applies to the Farey map and to maps of the "Pomeau-Manneville type".

Let T be a conservative ergodic transformation of $(X, \mathcal{B}, m)$, with $X = [0,1]$, and let $\mu << m$ be an infinite T-invariant measure with density unbounded at 0. Let us introduce the space

$$\mathcal{G} := \left\{ g \in L^\infty(X,m) \,:\, \bar{\mu}(g) := \lim_{a\to 0^+} \frac{1}{\mu((a,1))}\int_a^1 g\,d\mu \text{ exists finite}\right\}$$

Functions in $\mathcal{G}$ are called *global observables.*

Theorem 4.11 ([2]). *If T is exact and there exists $\bar{f} \in L^1(X,\mu)$ such that $\mathcal{P}^n \bar{f}$ is positive and increasing for all $n \geq 0$, then*

$$\int_X (g\circ T^n)\, f\, d\mu \to \bar{\mu}(g)\int_X f\,d\mu \tag{4.1}$$

for all $f \in L^1(X,\mu)$ *and all* $g \in \mathcal{G}$*.*

Systems satisfying (4.1) are called in [9] *global-local mixing*, and analogous definitions are introduced by using global observables or L^1 functions in place of f or g.

References

[1] J. Aaronson, "An Introduction to Infinite Ergodic Theory", American Mathematical Society, 1997

[2] C. Bonanno, P. Giulietti, M. Lenci, *Infinite mixing for one-dimensional maps with an indifferent fixed point*, Nonlinearity, to appear (look at arXiv)

[3] G. Boole, *On the comparison of transcendence, with certain applications to the theory of definite integrals*, Philos. Trans. R. Soc. Lond **147** (1857), 745–803

[4] G.H. Diamond, J.D. Vaaler, *Estimates for partial sums of continued fraction partial quotients*, Pacific J. Math. **122** (1986), 73–82

[5] M. Giampieri, S. Isola, *A one parameter family of analytic Markov maps with an intermittency transition*, Discrete Cont. Dyn. Syst. **12** (2005), 115–136

[6] S. Gouëzel, *Correlation asymptotics from large deviations in dynamical systems with infinite measure*, Colloq. Math. **125** (2011), 193–212

[7] S. Gouëzel, *Limit theorems in dynamical systems using the spectral method*, in Proceedings of Symposia in Pure Mathematics, vol. **89**, 2015, pp. 161–193

[8] M. Kesseböhmer, S. Munday, B.O. Stratmann, "Infinite Ergodic Theory of Numbers", De Gruyter, 2016

[9] M. Lenci, *On infinite-volume mixing*, Commun. Math. Phys. **298** (2010), 485–514

[10] I. Melbourne, D. Terhesiu, *Operator renewal theory and mixing rates for dynamical systems with infinite measure*, Invent. Math. **189** (2012), 61–110

[11] Y. Pomeau, P. Manneville, *Intermittent transition to turbulence in dissipative dynamical systems*, Comm. Math. Phys. **74** (1980), 189–197

[12] M. Thaler, *Infinite ergodic theory. Examples: One-dimensional maps with indifferent fixed points*, course notes, Marseille, 2001

[13] D. Thomine, *A generalized central limit theorem in infinite ergodic theory*, Probab. Theory Relat. Fields **158** (2014), 597–636

[14] R. Zweimuller, *Surrey notes on infinite ergodic theory*, available at `http://homepage.univie.ac.at/roland.zweimueller`

[15] R. Zweimuller, *Infinite measure preserving transformations with compact first regeneration*, J. Anal. Math. **103** (2007), 93–131

Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56127 Pisa
E-mail address: `claudio.bonanno@unipi.it`