

Ergodic Theorems for the Transfer Operators of Noisy Dynamical Systems

Eleonora Catsigeras

Instituto de Matemática y Estadística "Rafael Laguardia",
Universidad de la República, Uruguay
Email: eleonora@fing.edu.uy

Abstract. We consider stationary stochastic dynamical systems evolving on a compact metric space, by perturbing a deterministic dynamics with a random noise, added according to an arbitrary probabilistic distribution. We prove the maximal and pointwise ergodic theorems for the transfer operators associated to such systems. The results are extensions to noisy systems of some of the fundamental ergodic theorems for deterministic systems. The proofs are analytic. They follow the rigorous deductive method of the classic proofs in pure mathematics.

Keywords: Ergodic theorems, transfer operator, stochastic systems, noisy dynamics.

1 Introduction

The ergodic theory of deterministic (zero-noise) dynamical systems is based on a series of classical ergodic theorems. The first of those theorems is the pointwise ergodic theorem of Birkhoff-Khinchin [1] that ensures the almost sure convergence of the time averages (see also for instance [2, Chapter II], [3, Theorem 4.1.2], or [4, Chapter 1, p. 11]). Also the ergodic theorem of Kingman [5] ensures the pointwise convergence of the time averages, in a more general formulation, stating the convergence of any subadditive sequence of functions. For a recent more general statement and proof of the subadditive ergodic theorem see [6]. For differentiable dynamical systems, the multiplicative ergodic theorem of Oseledets ([7]) is also a fundamental result in the Ergodic Theory of deterministic dynamical systems, since, among other important consequences, it implies the existence of the Lyapunov exponents and gives a powerful tool for the study of (differentiable) chaos. See also for instance [2, Sections IV.10 and 11], [9, Theorem 4.6.2], or for a relatively short proof and subsequent generalizations of the multiplicative ergodic theorem, see [8].

The proofs of the pointwise ergodic theorems are very different nowadays from the original proofs of Birkhoff, and independent of Khinchin, when they first discovered them in the decade of 1930. Mathematicians now deduce the pointwise ergodic theorem as a particular case of many other more general results: the maximal ergodic theorems [2, Chapter IV], [9, Theorems 2.2.5 and 2.2.6], the subadditive ergodic theorem of Kingman and its generalization for cocycles [6], the operator theoretic ergodic theorems [10], the entangled ergodic theorem [11], the ergodic theorems for group actions [12] and for noncommutative products [13], and the ratio ergodic theorems [14], among others. Besides the ergodic theorems were also generalized for infinite measure spaces [15], [16]. Also the classical multiplicative ergodic theorem of Oseledets is now generalized in several forms; as for instance in Filip's extensions [8], Austin's multiplicative ergodic theorem [17], [18], and González-Tokman-Quass multiplicative ergodic theorem for cocycles [19].

Some of the results cited above apply only to deterministic (zero-noise) dynamical systems. Precisely, the question motivating this paper is: Are those ergodic theorems applicable or extendable also to stochastic or noisy systems? In fact, some of them already have adapted statements that are applicable to stochastic processes, Markovian systems, or random transformations systems (RTS). For instance, very early, the pointwise ergodic theorem for Markovian processes was proved by Kakutani [20] (see also [21, Theorem 6, p. 388] and [23, Corollary 2.2, p. 24]). And much later, the multiplicative ergodic theorem for RTS is stated and proved in [23, Chapter III, p. 88], and also in [24, Chapter 4].

Contributions and novelty of this paper. The purpose of this paper is precisely to state and prove the maximal and pointwise ergodic theorems for stochastic dynamical systems. They are Theorems 1

and 2, and Corollaries 1 and 2. As said above the maximal ergodic theorem and its corollaries, up to now, are known only for deterministic (zero-noise) systems, but in the literature, they are not still stated and proved for stochastic systems.

As a consequence of the corollaries of the maximal ergodic theorem, we also obtain a different proof of Kakutani’s ergodic theorem, when applied to a transfer operator (see Theorem 7). The first part of our pointwise ergodic theorem (initial part of the statement of Theorem 2) is only a reformulation of Kakutani’s ergodic theorem; therefore its statement is not complete new, although its proof is new. Nevertheless, the second part of that theorem, which refers to the convergence to the spatial mean value of the time averages when the probability measure is ergodic, is new.

In Section 6 we provide a concrete example of the abstract results that we state and prove along the paper. Even if the example is known (it was taken from the literature), the application of our new theorems to it, allows us to conclude features of its particular dynamics that, as far as we know, were not previously proved in the literature.

One of the relevance of the results, from the viewpoint of their possible applications, is that they *do not depend on the distribution of the noise*. They work under the hypothesis of two very general conditions of the so called *transfer operator* (see Definition 1, the hypothesis are conditions 1.1 and 1.2). The technique of the transfer operator is well known in the Functional Analysis Theory, and is also known in the theory of Noisy Dynamical Systems of Pure Mathematics. It is very powerful, since, as a main difference with the usual results in the Engineering Scientific literature, it does not depend on particular concrete additional hypothesis neither of which are the formulae of the deterministic dynamics, nor of the particular probabilistic distribution of the noise. It can be for instance, a Gaussian noise, a uniformly distributed noise, a Bernoulli noise, or any locally added noise that has no density and is supported on a Cantor set. Of course, the ergodic stationary measures, and the transfer operators *will depend on the formulae of the concrete deterministic system, and of the distribution of the noise* but the maximal ergodic theorem, the pointwise convergence theorem, and in particular their statements in the case of ergodic probability measures, hold anyhow.

1.1 Setting Up

Along this paper, X is a compact metric space and \mathcal{A} is the Borel sigma-algebra in X . We denote by \mathcal{M} the space of all the probability measures on (X, \mathcal{A}) endowed with the weak*-topology. We consider the functional space $C^0(X, \mathbb{C})$ of all the continuous functions $\varphi : X \mapsto \mathbb{C}$ with the supremum norm. We denote by $C^0(X, \mathbb{R})$ the subspace of real continuous functions, and by $C^0(X, [0, 1])$ its subset of functions whose values belong to the interval $[0, 1]$.

A stochastic dynamical system on X is a stochastic process $x_0, x_1, \dots, x_n, \dots$ with any given initial probability distribution $\mu_0 \in \mathcal{M}$ such that

$$\mu_0(A) = \text{prob}\{x_0 \in A\} \quad \forall A \in \mathcal{A},$$

and a family $\{P(x, \cdot)\}_{x \in X} \subset \mathcal{M}$ of *transition probabilities* $P(x, \cdot)$ (also called *probabilities of noise*) such that

$$P(x, A) = \text{prob}\{x_{n+1} \in A/x_n = x\} \quad \forall n \geq 0, \quad \forall A \in \mathcal{A}, \quad \forall x \in X.$$

The stochastic dynamical system is continuous if the application $x \in X \mapsto P(x, \cdot) \in \mathcal{M}$ is continuous in the weak*-topology.

When studying the properties of continuous stochastic dynamical systems, the following transfer operator $\mathcal{L} : C^0(X, \mathbb{C}) \mapsto C^0(X, \mathbb{C})$ and its dual transfer operator $\mathcal{L}^* : \mathcal{M} \mapsto \mathcal{M}$ are usually considered:

$$(\mathcal{L}\varphi)(x) := \int \varphi(y)P(x, dy) \quad \forall x \in X, \quad \forall \varphi \in C^0(X, \mathbb{C}); \tag{1}$$

$$\int \varphi d(\mathcal{L}^*\mu) := \int (\mathcal{L}\varphi) d\mu \quad \forall \varphi \in C^0(X, \mathbb{C}), \quad \forall \mu \in \mathcal{M}.$$

The ergodic properties of the continuous stochastic dynamical system rely on the convergence μ -a.e. (when it occurs) of the time averages $\frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)$, and also, on the properties of the limits of weak*-convergent subsequences of $\frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^{*j} \mu)$.

We will start by considering any operator \mathcal{L} from $C^0(X, \mathbb{C})$ to itself, that is positive, bounded by 1, and such that $\mathcal{L}(1) = 1$. As said above, along this paper we will study the ergodic properties of the iteration of \mathcal{L} , and of its dual operator \mathcal{L}^* on the space \mathcal{M} of probability measures. A priori, \mathcal{L} is not constructed as the transfer operator of a stochastic dynamical system. Nevertheless, in Section 2-Proposition 3, we will show that there exists such a stochastic system whose transfer operator coincides with the given \mathcal{L} .

1.2 Definitions

Definition 1. (Transfer Operator \mathcal{L} in the space of continuous functions.)

A Transfer Operator \mathcal{L} in the space of continuous functions is a linear operator

$$\mathcal{L} : C^0(X, \mathbb{C}) \mapsto C^0(X, \mathbb{C})$$

such that:

- 1.1 \mathcal{L} is positive; precisely $\mathcal{L}\varphi$ is real and non negative if φ is real and non negative.
- 1.2 $\|\mathcal{L}\| = 1$; precisely $\max_{x \in X} |(\mathcal{L}\varphi)(x)| \leq \max_{x \in X} |\varphi(x)|$ for all $\varphi \in C^0(X, \mathbb{C})$, and $\mathcal{L} \cdot 1 = 1$.

Definition 2. (Transfer Operator \mathcal{L}^* in the space of probability measures.)

For any transfer operator $\mathcal{L} : C^0(X, \mathbb{C}) \mapsto C^0(X, \mathbb{C})$, the Dual Transfer Operator \mathcal{L}^* in the space \mathcal{M} of probabilities is the application $\mathcal{L}^* : \mathcal{M} \mapsto \mathcal{M}$ defined by

$$\int \varphi d\mathcal{L}^*\mu = \int \mathcal{L}\varphi d\mu \quad \forall \varphi \in C^0(X, \mathbb{C}) \quad \forall \mu \in \mathcal{M}. \tag{2}$$

Due to Riesz Theorem, the dual transfer operator \mathcal{L}^* in the space of probability measures exists and is unique for any given transfer operator \mathcal{L} in the space of continuous functions.

We are particularly interested in those probability measures that are fixed by the dual transfer operator \mathcal{L}^* , and more generally, in those probability measures μ that are \mathcal{L}^* -periodic with period $p \geq 1$; i.e. fixed by \mathcal{L}^{*p} for a minimum natural number $p \geq 1$.

Extension of the operator \mathcal{L} to the space L_∞ .

In Section 2-Proposition 6, we prove that the transfer operator \mathcal{L} can be extended to the space L_∞ of bounded measurable functions in such a way that the following equality holds:

$$\int \varphi d\mathcal{L}^*\mu = \int \mathcal{L}\varphi d\mu \quad \forall \varphi \in L_\infty \quad \forall \mu \in \mathcal{M}.$$

Definition 3. (Invariant sets almost everywhere)

Let $\mu \in \mathcal{M}$ and $A \in \mathcal{A}$. We say that A is μ -a.e. \mathcal{L} -invariant if

$$\mathcal{L}\chi_A = \chi_A \quad \mu - a.e.,$$

where χ_A denotes the characteristic function of A . Analogously, the measurable set A is μ -a.e. \mathcal{L}^p -invariant for a natural number $p \geq 1$, if

$$\mathcal{L}^p\chi_A = \chi_A \quad \mu - a.e.$$

Definition 4. (Ergodic measures).

Let $\mu \in \mathcal{M}$. We say that μ is ergodic for \mathcal{L}^* if it is fixed by \mathcal{L}^* and if $\mu(A) \in \{0, 1\}$ for any set $A \subset M$ that is \mathcal{L} -invariant μ -a.e. In other words, if $\chi_A(x) = (\mathcal{L}\chi_A)(x)$ for μ -a.e. $x \in X$, then either $\mu(A) = 0$ or $\mu(A) = 1$.

Analogously, for any natural number $p \geq 1$, we say that μ is ergodic for \mathcal{L}^{*p} if it is fixed by \mathcal{L}^{*p} and if $\mu(A) \in \{0, 1\}$ for any set $A \subset M$ that is \mathcal{L}^p -invariant μ -a.e.

1.3 Statement of the Results.

The purpose of this paper is to prove the following results:

Theorem 1. (Maximal Ergodic Theorem)

Let $\mu \in \mathcal{M}$ such that $\mathcal{L}^* \mu = \mu$. Let $\varphi \in L_\infty$ be a real function. Define

$$\begin{aligned}\varphi_n(x) &:= \max\{\varphi(x), (\varphi + \mathcal{L}\varphi)(x), \dots, (\varphi + \mathcal{L}\varphi + \dots + \mathcal{L}^{n-1}\varphi)(x)\}; \\ E(\varphi) &:= \{x \in X : \sup_{n \geq 1} \varphi_n(x) > 0\}.\end{aligned}$$

Then

$$\int_{E(\varphi)} \varphi d\mu \geq 0. \quad (3)$$

Corollary 1. Let $\mu \in \mathcal{M}$ such that $\mathcal{L}^* \mu = \mu$. Let $\varphi : X \mapsto \mathbb{R}$ be bounded and measurable. For each natural number $n \geq 1$ and each real number α , denote:

$$\begin{aligned}\varphi_n &:= \sum_{j=0}^{n-1} \mathcal{L}^j \varphi. \\ C_\alpha &:= \{x \in X : \sup_{n \geq 1} \frac{\varphi_n(x)}{n} > \alpha\}.\end{aligned}$$

Then, for any measurable set $A \subset C_\alpha$ such that $(\mathcal{L}\chi_A)(x) = \chi_A(x)$ for μ -a.e. $x \in X$, the following inequality holds:

$$\int_A \varphi d\mu \geq \alpha \cdot \mu(A).$$

Corollary 2. Let $\mu \in \mathcal{M}$ such that $\mathcal{L}^* \mu = \mu$. Let $\varphi : X \mapsto \mathbb{R}$ be bounded and measurable. For each natural number $n \geq 1$ and each real number β , denote:

$$\begin{aligned}\varphi_n &:= \sum_{j=0}^{n-1} \mathcal{L}^j \varphi. \\ B_\beta &:= \{x \in X : \inf_{n \geq 1} \frac{\varphi_n(x)}{n} < \beta\}.\end{aligned}$$

Then, for any measurable set $A \subset B_\beta$ such that $(\mathcal{L}\chi_A)(x) = \chi_A(x)$ for μ -a.e. $x \in X$, the following inequality holds:

$$\int_A \varphi d\mu \leq \beta \cdot \mu(A).$$

Theorem 2. (Pointwise Ergodic Theorem for Periodic Measures by the Transfer Operator)

Let μ be a probability measure fixed by the transfer operator \mathcal{L}^{*p} for some natural number $p \geq 1$. Then, for any function $\varphi \in L_\infty$ the following limit exists μ -a.e.:

$$\tilde{\varphi}_p(x) := \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^{jp}\varphi)(x) \quad \mu\text{-a.e. } x \in X. \quad (4)$$

If besides μ is ergodic for \mathcal{L}^{*p} , then

$$\tilde{\varphi}_p(x) = \int \varphi d\mu \quad \mu\text{-a.e. } x \in X. \quad (5)$$

The proofs of Theorems 1 and 2, as well as the proofs of their corollaries and other ergodic theorems, will be developed along Sections 3 and 5. In Section 2, we prove some previous statements.

2 Previous Results

Proposition 3. *Let \mathcal{L} be a transfer operator. Then, there exists a unique family of probability measures $\{P(x, \cdot)\}_{x \in X} \subset \mathcal{M}$ such that:*

$$(\mathcal{L}\varphi)(x) = \int \varphi(y) P(x, dy) \quad \forall x \in X \quad \forall \varphi \in C^0(X, \mathbb{C}).$$

Besides, the probability measure $P(x, \cdot) \in \mathcal{M}$ depends continuously on $x \in X$ in the weak* topology of \mathcal{M} .

Proof. Fix $x \in X$. The transformation $\Lambda_x : C^0(X, \mathbb{R}) \mapsto \mathbb{R}$, defined by $\Lambda_x(\varphi) := (\mathcal{L}\varphi)(x)$ is a linear operator defined on the space of real continuous functions. It is positive, bounded by 1, and $\Lambda_x(1) = 1$. So, applying Riesz Theorem there exists a unique probability measure $P(x, \cdot) \in \mathcal{M}$ such that $\Lambda_x\varphi = \int \varphi(y) P(x, dy)$ for all $\varphi \in C^0(X, \mathbb{C})$. To end the proof of Proposition 3 we must prove that $P(x, \cdot)$ depends continuously on x in the weak* topology of \mathcal{M} . Equivalently, we must prove that if $x_n \rightarrow x \in X$ as $n \rightarrow +\infty$ then, $\lim_{n \rightarrow +\infty} \int \varphi(y) P(x_n, dy) = \int \varphi(y) P(x, dy)$ for any continuous function $\varphi \in C^0(X, \mathbb{C})$. In fact, by construction of the probability measure $P(x_n, \cdot)$, and recalling that $\mathcal{L}\varphi$ is by definition a continuous function if φ is continuous, we have:

$$\lim_{n \rightarrow +\infty} \int \varphi(y) P(x_n, dy) = \lim_{n \rightarrow +\infty} (\mathcal{L}\varphi)(x_n) = (\mathcal{L}\varphi)(x) = \int \varphi(y) P(x, dy),$$

ending the proof. □

Remark 4. As a consequence of Proposition 3, for any measure $\mu \in \mathcal{M}$, we have:

$$A \in \mathcal{A} \text{ is } \mu\text{-a.e. } \mathcal{L}\text{-invariant} \quad \Leftrightarrow \quad \chi_A(x) = (\mathcal{L}\varphi)(x) = P(x, A) \quad \text{for } \mu\text{-a.e. } x \in X. \quad (6)$$

Proposition 5. *The transfer operator $\mathcal{L}^* : \mathcal{M} \mapsto \mathcal{M}$ is continuous in the weak* topology of \mathcal{M} .*

Proof. If $\mu_n \rightarrow \mu$ in the weak* topology, then $\int \mathcal{L}\varphi d\mu_n \rightarrow \int \mathcal{L}\varphi d\mu$ for any continuous function φ (because $\mathcal{L}\varphi$ is also a continuous function). Thus, applying the definition of the measure $\mathcal{L}^*\mu$, we re-write the latter equality as $\int \varphi d\mathcal{L}^*\mu_n \rightarrow \int \varphi d\mathcal{L}^*\mu$ for any continuous function φ . In other words $\mathcal{L}^*\mu_n \rightarrow \mathcal{L}^*\mu$ in the weak* topology of \mathcal{M} (provided that $\mu_n \rightarrow \mu$). We conclude that the transfer operator \mathcal{L}^* is continuous, as wanted. □

Definition 5. We denote by L_∞ the set of bounded functions $\varphi : X \mapsto \mathbb{C}$ such that for any probability measure $\mu \in \mathcal{M}$ there exists a measurable function φ_μ that coincides μ -a.e. with φ .

Thus, for any $\varphi \in L_\infty$ it is well defined the integral of φ with respect to any measure $\mu \in \mathcal{M}$, by the following equality $\int \varphi d\mu := \int \varphi_\mu d\mu$. In particular, it is well defined the following extension of the transfer operator \mathcal{L} to any real function $\varphi \in L_\infty$:

$$(\mathcal{L}\varphi)(x) := \int \varphi(y) P(x, dy), \quad \forall x \in X, \quad \forall \varphi \in L_\infty, \quad (7)$$

where $P(x, dy)$ is the probability measure constructed by Proposition 3 for each $x \in X$.

Proposition 6. *For any real function $\varphi \in L_\infty$ the function $\mathcal{L}\varphi$ constructed by equality (7) also belongs to L_∞ . Besides*

$$\int \varphi d\mathcal{L}^*\mu = \int \mathcal{L}\varphi d\mu \quad \forall \varphi \in L_\infty \quad \forall \mu \in \mathcal{M}. \quad (8)$$

Proof. Since $|\varphi(y)| \leq k$ for all $y \in X$, we have $(\mathcal{L}\varphi)(x) = \int \varphi(y) P(x, dy) \leq k \quad \forall x \in X$. Therefore $\mathcal{L}\varphi$ is bounded. Let us prove that $\mathcal{L}\varphi$ coincides with a measurable function for μ -a.e. $x \in X$.

1st. step. If $\varphi : X \mapsto \mathbb{C}$ is continuous, then from hypothesis $(\mathcal{L}\varphi)$ is continuous, hence measurable.

2nd. step. Let us prove that for any open set $V \subset X$, the real function $\mathcal{L}\chi_V$ is measurable. Since X is a compact metric space, for any open set $V \subset X$ there exists an increasing sequence $\{K_n\}_{n \geq 1}$ of compact

sets $K_n \subset X$, such that $\bigcup_{n \geq 1} K_n = V$. So $\lim_{n \rightarrow +\infty} \chi_{K_n}(x) = \chi_V(x)$ for all $x \in X$. From Urysohn Lemma, there exists a sequence of continuous functions $\varphi_n : X \mapsto [0, 1]$ such that $\chi_{K_n}(x) \leq \varphi_n(x) \leq \chi_V(x)$ for all $x \in X$. Therefore,

$$\lim_{n \rightarrow +\infty} \varphi_n(x) = \chi_V(x) \quad \forall x \in X. \tag{9}$$

Applying the dominated convergence theorem, we deduce that:

$$\lim_{n \rightarrow +\infty} \int \varphi_n(y) P(x, dy) = \int \chi_V(y) P(x, dy) \quad \forall x \in X.$$

By the definition of the operator \mathcal{L} we deduce that

$$\lim_{n \rightarrow +\infty} (\mathcal{L}\varphi_n)(x) = (\mathcal{L}\chi_V)(x) \quad \forall x \in X. \tag{10}$$

Since φ_n is continuous, $(\mathcal{L}\varphi_n)$ is continuous, hence measurable. Besides, the point-wise limit of measurable functions is measurable. We deduce that $\mathcal{L}\chi_V$ is measurable, as wanted.

3rd. step. Let us prove that for any open set $V \subset X$, the following equality holds:

$$(\mathcal{L}^*\mu)(V) = \int (\mathcal{L}\chi_V) d\mu. \tag{11}$$

In fact, applying equalities (2), (9) and (10), and the dominated convergence theorem, we obtain:

$$\int (\mathcal{L}\chi_V) d\mu = \lim_{n \rightarrow +\infty} \int (\mathcal{L}\varphi_n) d\mu = \lim_{n \rightarrow +\infty} \int \varphi_n d\mathcal{L}^*\mu = \int \chi_V d\mathcal{L}^*\mu = (\mathcal{L}^*\mu)(V).$$

So, equality (11) is proved.

4th. step. For any compact set $K \subset X$, equality (11) holds, with K in the role of V . In fact, $\chi_K = 1 - \chi_V$, where $V = X \setminus K$ is open, hence satisfies equality (11). Besides $\mathcal{L}(1 - \chi_V) = 1 - \mathcal{L}\chi_V$ and $(\mathcal{L}^*\mu)(K) = 1 - (\mathcal{L}^*\mu)(V)$. So equality (11) also holds for K instead of V .

5th. step. Let us prove that for any measurable set $A \subset X$ and any probability measure μ , the function $\mathcal{L}\chi_A$ is measurable μ -a.e. (namely, $\mathcal{L}\chi_A$ coincides with a measurable function up to a set of zero μ -measure).

Since (X, \mathcal{A}) is the measurable space of a compact metric space X with the Borel sigma-algebra \mathcal{A} , any probability measure in (X, \mathcal{A}) is regular. So, for any set $A \in \mathcal{A}$ and any natural number $n \geq 1$, there exists a compact set $K_n \subset A$ and an open set $V_n \supset A$, such that

$$(\mathcal{L}^*\mu)(V_n \setminus K_n) < \frac{1}{n} \quad \forall n \geq 1.$$

It is not restrictive to assume that $K_n \subset K_{n+1}$ and $V_{n+1} \subset V_n$ for all $n \geq 1$. If not, we substitute K_n by $\bigcup_{j=1}^n K_j$, and V_n by $\bigcap_{j=1}^n V_j$.

Since $K_n \subset A \subset V_n$ we have

$$\chi_{K_n} \leq \chi_A \leq \chi_{V_n} \quad \forall n \geq 1.$$

Therefore

$$\lim_{n \rightarrow +\infty} \chi_{K_n}(x) = \chi_{(\bigcup_{n \geq 1} K_n)}(x) \leq \chi_A(x) \leq \chi_{(\bigcap_{n \geq 1} V_n)}(x) = \lim_{n \rightarrow +\infty} \chi_{V_n}(x) \quad \forall x \in X. \tag{12}$$

Thus, applying the dominated convergence theorem, we deduce that:

$$\lim_{n \rightarrow +\infty} \int \chi_{K_n}(y) P(x, dy) \leq \int \chi_A(y) P(x, dy) \leq \lim_{n \rightarrow +\infty} \int \chi_{V_n}(y) P(x, dy) \quad \forall x \in X.$$

Equivalently:

$$\lim_{n \rightarrow +\infty} (\mathcal{L}\chi_{K_n})(x) \leq (\mathcal{L}\chi_A)(x) \leq \lim_{n \rightarrow +\infty} (\mathcal{L}\chi_{V_n})(x) \quad \forall x \in X. \tag{13}$$

Since $V_n \setminus K_n$ is an open set, $(\mathcal{L}\chi_{V_n \setminus K_n})$ is a measurable function, and equality (11) applies:

$$0 \leq \int \mathcal{L}\chi_{V_n \setminus K_n} d\mu = \int \chi_{V_n \setminus K_n} d(\mathcal{L}^*\mu) = (\mathcal{L}^*\mu)(V_n \setminus K_n) < \frac{1}{n}.$$

Thus, applying again the dominated convergence theorem, we obtain:

$$0 \leq \int (\lim_{n \rightarrow +\infty} \mathcal{L}\chi_{V_n} - \lim_{n \rightarrow +\infty} \mathcal{L}\chi_{K_n}) d\mu = \lim_{n \rightarrow +\infty} \int \mathcal{L}\chi_{V_n \setminus K_n} d\mu = 0.$$

But the integrated function is non negative. Thus it must be null μ -a.e. We have proved that

$$\lim_{n \rightarrow +\infty} \mathcal{L}\chi_{V_n}(x) = \lim_{n \rightarrow +\infty} \mathcal{L}\chi_{K_n}(x) \quad \text{for } \mu\text{-a.e. } x \in X.$$

Joining this result with inequalities (13) we conclude that $\chi_A(x)$ coincides, for μ - a.e. $x \in X$ with a measurable function. Precisely

$$(\mathcal{L}\chi_A)(x) = \lim_{n \rightarrow +\infty} (\mathcal{L}\chi_{V_n})(x) = \lim_{n \rightarrow +\infty} (\mathcal{L}\chi_{K_n})(x) \quad \text{for } \mu\text{-a.e. } x \in X.$$

Therefore, taking into account inequalities (12) and that χ_{K_n} and χ_{V_n} satisfy equality (8), we obtain

$$\begin{aligned} \int \mathcal{L}\chi_A d\mu &= \int \lim_{n \rightarrow +\infty} \mathcal{L}\chi_{V_n} d\mu = \\ \lim_{n \rightarrow +\infty} \int \mathcal{L}\chi_{V_n} d\mu &= \lim_{n \rightarrow +\infty} \int \chi_{V_n} d\mathcal{L}^*\mu = \int \lim_{n \rightarrow +\infty} \chi_{V_n} d\mathcal{L}^*\mu \geq \int \chi_A d\mathcal{L}^*\mu, \\ \int \mathcal{L}\chi_A d\mu &= \int \lim_{n \rightarrow +\infty} \mathcal{L}\chi_{K_n} d\mu = \\ \lim_{n \rightarrow +\infty} \int \mathcal{L}\chi_{K_n} d\mu &= \lim_{n \rightarrow +\infty} \int \chi_{K_n} d\mathcal{L}^*\mu = \int \lim_{n \rightarrow +\infty} \chi_{K_n} d\mathcal{L}^*\mu \leq \int \chi_A d\mathcal{L}^*\mu. \end{aligned}$$

We conclude that χ_A also satisfies equality (8).

6th. step. Consider a simple measurable function φ ; i.e. φ is a finite linear combination, with real coefficients, of characteristic functions of measurable sets. Then, the function $\mathcal{L}\varphi$ is a finite linear combination of μ -a.e. measurable functions, because the operator \mathcal{L} is linear. Since the finite linear combination of measurable functions is measurable, we conclude that $\mathcal{L}\varphi$ coincides μ -a.e. with a measurable function. Besides, taking into account that the characteristic functions of measurable sets satisfy equality (8), by the linearity of the integrals, we deduce that the simple function φ also satisfies equality (8).

7th. step. Now, consider any bounded measurable function $\varphi : X \mapsto \mathbb{R}$. It is well known that there exists an increasing (in absolute value) sequence $\{\varphi_n\}_{n \geq 1}$ of simple measurable functions such that $\lim_{n \rightarrow +\infty} \varphi_n(x) = \varphi(x)$ for all $x \in X$. Thus, for all $x \in X$ we can apply the dominated convergence theorem of the integrals of the functions φ_n with respect to the probabilities $P(x, \cdot)$. We deduce:

$$(\mathcal{L}\varphi)(x) = \lim_{n \rightarrow +\infty} (\mathcal{L}\varphi_n)(x) \quad \forall x \in X \setminus \Delta.$$

We have already proved that the functions $\mathcal{L}\varphi_n$ coincide with a measurable function for μ -a.e. point in $X \setminus \Delta$, because φ_n is a simple measurable function. So, their point-wise limit also coincides with a measurable function μ -a.e. Besides, since the simple functions satisfy equality (8), by the dominated convergence theorem, the bounded real function φ also satisfies it.

8th. step. Finally, consider any bounded measurable function $\varphi : X \mapsto \mathbb{C}$. By taking real and imaginary parts of φ , and taking into account that \mathcal{L} and the integrals are linear, we conclude that $\mathcal{L}\varphi$ coincides μ -a.e. with a measurable function, and satisfies equality (8). This ends the proof of Proposition 6. \square

3 Proof of the Maximal Ergodic Theorem and Its Corollaries

We will start proving the Maximal Ergodic Theorem 1. To prove it we need some previous Lemmas:

Lemma 1. For any bounded measurable function $\varphi : X \mapsto \mathbb{R}$, consider the positive and negative parts of φ defined by:

$$\varphi^+(x) := \max\{0, \varphi(x)\} \geq 0, \quad \varphi^-(x) := -\min\{0, \varphi(x)\} \geq 0, \quad \varphi = \varphi^+ - \varphi^- \leq \varphi^+.$$

Then

$$(\mathcal{L}(\varphi^+))(x) \geq (\mathcal{L}\varphi)^+(x) \quad \forall x \in X.$$

Proof. On the one hand, applying Proposition 3 we have:

$$(\mathcal{L}(\varphi^+))(x) = \int \varphi^+(y) P(x, dy) \geq \int \varphi(y) P(x, dy) = (\mathcal{L}\varphi)(x) \quad \forall x \in X.$$

On the other hand, since $\varphi^+ \geq 0$, we have $\mathcal{L}(\varphi^+) \geq 0$. So, $\mathcal{L}(\varphi^+) \geq \max\{0, (\mathcal{L}\varphi)\} = (\mathcal{L}\varphi)^+$, as wanted. \square

Lemma 2. Let $\mu \in \mathcal{M}$ such that $\mathcal{L}^*\mu = \mu$. Then, for any bounded measurable function $\varphi : X \mapsto \mathbb{R}$:

$$\int_{\varphi>0} \varphi d\mu \geq \int_{(\mathcal{L}\varphi)>0} (\mathcal{L}\varphi) d\mu.$$

Proof. Applying equality (8), taking into account that $\mathcal{L}^*\mu = \mu$, and applying Lemma 1, we obtain:

$$\int_{\varphi>0} \varphi d\mu = \int \varphi^+ d\mu = \int \varphi^+ d\mathcal{L}^*\mu = \int \mathcal{L}(\varphi^+) d\mu \geq \int (\mathcal{L}\varphi)^+ d\mu = \int_{\mathcal{L}\varphi>0} \mathcal{L}\varphi d\mu.$$

\square

3.1 Proof of Theorem 1.

Proof. The sequence $\{\varphi_n\}_{n \geq 1}$ is non decreasing. Thus, for all $n \geq 1$, the set $E_n := \{x \in X : \varphi_n > 0\}$ is contained in E_{n+1} . Since $E(\varphi) = \bigcup_{n \geq 1} E_n$, we obtain: $\int_{E(\varphi)} \varphi d\mu = \lim_{n \rightarrow +\infty} \int_{E_n} \varphi d\mu$. So, to prove inequality (3), it is enough to prove the following inequality:

$$I_n := \int_{E_n} \varphi d\mu \geq 0 \quad \forall n \geq 1 \quad (\text{to be proved}). \tag{14}$$

We have

$$I_n = \int_{\varphi_n>0} \varphi d\mu = \int_{\varphi_n>0, \mathcal{L}\varphi_n \leq 0} \varphi d\mu + \int_{\varphi_n>0, \mathcal{L}\varphi_n > 0} \varphi d\mu. \tag{15}$$

For all $j \geq 1$, denote $\psi_j := \varphi + \mathcal{L}\varphi + \dots + \mathcal{L}^{j-1}\varphi$. We assert that

$$\mathcal{L}(\max_{1 \leq j \leq n} \psi_j) \geq \max_{1 \leq j \leq n} \mathcal{L}\psi_j. \tag{16}$$

In fact, $\max_{1 \leq j \leq n} \psi_j \geq \psi_i$ for all $1 \leq i \leq n$. Thus

$$(\mathcal{L}(\max_{1 \leq j \leq n} \psi_j))(x) = \int \max_{1 \leq j \leq n} \psi_j(y) P(x, dy) \geq \int \psi_i(y) P(x, dy) = (\mathcal{L}\psi_i)(x) \quad \forall x \in X \quad \forall i = 1, \dots, n.$$

Now, inequality (16) is proved.

Let us compute both integrals at right in equality (15):

$$\mathcal{L}\varphi_n = \mathcal{L}(\max_{1 \leq j \leq n} \psi_j) \geq \max_{1 \leq j \leq n} \mathcal{L}\psi_j = \max\{\mathcal{L}\varphi, \mathcal{L}\varphi + \mathcal{L}^2\varphi, \dots, \mathcal{L}\varphi + \mathcal{L}^2\varphi + \dots + \mathcal{L}^n\varphi\}.$$

Therefore, $\mathcal{L}\varphi_n \leq 0$ implies $\max\{\mathcal{L}\varphi, \mathcal{L}\varphi + \mathcal{L}^2\varphi, \dots, \mathcal{L}\varphi + \mathcal{L}^2 + \dots + \mathcal{L}^n\varphi\} \leq 0$, hence:

$$(\mathcal{L}\varphi_n)(x) \leq 0 \Rightarrow \varphi_n(x) = \max\{\varphi(x), (\varphi + \mathcal{L}\varphi)(x), \dots, (\varphi\mathcal{L}\varphi + \dots + \mathcal{L}^{n-1}\varphi)(x)\} = \varphi(x).$$

Thus, the first integral at right in equality (15) can be written as follows:

$$\int_{\varphi_n > 0, \mathcal{L}\varphi_n \leq 0} \varphi \, d\mu = \int_{\varphi_n > 0, \mathcal{L}\varphi_n \leq 0} \varphi_n \, d\mu. \tag{17}$$

Now, let us compute the second integral at right in equality (15). Applying inequality (16) we obtain:

$$\begin{aligned} \varphi + \mathcal{L}\varphi_n &= \varphi + \mathcal{L}(\max_{1 \leq j \leq n} \psi_j) \geq \varphi + \max_{1 \leq j \leq n} (\mathcal{L}\psi_j) = \\ &= \varphi + \max\{\mathcal{L}\varphi, \mathcal{L}\varphi + \mathcal{L}^2\varphi, \dots, \mathcal{L}\varphi + \mathcal{L}^2\varphi + \dots + \mathcal{L}^n\varphi\} = \\ &= \max\{\varphi + \mathcal{L}\varphi, \varphi + \mathcal{L}\varphi + \mathcal{L}^2\varphi, \dots, \varphi + \mathcal{L}\varphi + \mathcal{L}^2\varphi + \dots + \mathcal{L}^n\varphi\}. \end{aligned} \tag{18}$$

Besides

$$\mathcal{L}\varphi_n > 0 \Rightarrow \varphi + \mathcal{L}\varphi_n > \varphi. \tag{19}$$

Joining inequalities (18) and (19), we deduce:

$$\begin{aligned} \mathcal{L}\varphi_n > 0 \Rightarrow \varphi + \mathcal{L}\varphi_n &\geq \max\{\varphi, \varphi + \mathcal{L}\varphi, \dots, \varphi + \mathcal{L}\varphi + \dots + \mathcal{L}^n\varphi\} \geq \\ &= \max\{\varphi, \varphi + \mathcal{L}\varphi, \dots, \varphi + \mathcal{L}\varphi + \dots + \mathcal{L}^{n-1}\varphi\} = \varphi_n. \end{aligned}$$

In brief, we have proved that

$$\mathcal{L}\varphi_n > 0 \Rightarrow \varphi + \mathcal{L}\varphi_n \geq \varphi_n \Rightarrow \varphi \geq \varphi_n - \mathcal{L}\varphi_n.$$

Substituting the latter inequality in the second integral at right of equality (15), we obtain:

$$\int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} \varphi \, d\mu \geq \int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} \varphi_n \, d\mu - \int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} \mathcal{L}\varphi_n \, d\mu. \tag{20}$$

Now, we use equality (17) and inequality (20) to obtain a lower bound of the integral (15):

$$I_n \geq \int_{\varphi_n > 0, \mathcal{L}\varphi_n \leq 0} \varphi_n \, d\mu + \int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} \varphi_n \, d\mu - \int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} (\mathcal{L}\varphi_n) \, d\mu = \int_{\varphi_n > 0} \varphi_n \, d\mu - \int_{\varphi_n > 0, \mathcal{L}\varphi_n > 0} (\mathcal{L}\varphi_n) \, d\mu.$$

Since $\{\varphi_n > 0, \mathcal{L}\varphi_n > 0\} \subset \{\mathcal{L}\varphi_n > 0\}$ and the function $\mathcal{L}\varphi_n$ is positive on those sets, we obtain:

$$I_n \geq \int_{\varphi_n > 0} \varphi_n \, d\mu - \int_{\mathcal{L}\varphi_n > 0} (\mathcal{L}\varphi_n) \, d\mu.$$

Finally, applying Lemma 2, the difference at right in the latter inequality is non negative. We conclude that $I_n \geq 0$, proving assertion (14) as wanted, and ending the proof of Theorem 1. \square

3.2 Proof of Corollary 1.

To prove Corollary 1, we need a previous lemma:

Lemma 3. *Let $\mu \in \mathcal{M}$ (μ is not necessarily fixed by \mathcal{L}^*). Let $A \subset X$ be a measurable set that is μ -a.e. \mathcal{L} -invariant. Then, for any measurable bounded function $\varphi : X \mapsto \mathbb{C}$ the following equality holds:*

$$(\mathcal{L}(\chi_A \cdot \varphi))(x) = \chi_A(x) \cdot (\mathcal{L}\varphi)(x) \quad \mu\text{-a.e. } x \in X.$$

Proof. For μ -a.e. $x \in X$ we have $(\mathcal{L}\chi_A)(x) = P(x, A) = \chi_A(x)$. Therefore

$$P(x, A) = 1 \text{ if } x \in A; \quad P(x, A) = 0 \text{ if } x \notin A, \quad \mu\text{-a.e. } x \in X.$$

$$x \in A \Rightarrow \chi_A(x) \cdot (\mathcal{L}\varphi)(x) = (\mathcal{L}\varphi)(x) = \int \varphi(y) P(x, dy) = \int_A \varphi(y) P(x, dy) + \int_{X \setminus A} \varphi(y) P(x, dy).$$

But for μ -a.e. $x \in A$ we have $P(x, X \setminus A) = 0$. So, the integral at right in the above equality is zero. We obtain:

$$\text{for } \mu\text{-a.e. } x \in A, \quad \chi_A(x) \cdot (\mathcal{L}\varphi)(x) = \int_A \varphi(y) P(x, dy) = \int \chi_A(y) \cdot \varphi(y) P(x, dy) = (\mathcal{L}(\chi_A\varphi))(x).$$

We have proved Lemma 3 for μ -a.e. $x \in A$. Now, let us consider $x \notin A$:

$$x \notin A \Rightarrow \chi_A(x) \cdot (\mathcal{L}\varphi)(x) = 0.$$

Besides, for μ -a.e. $x \notin A$ we have $P(x, A) = 0$. We obtain:

$$\text{for } \mu\text{-a.e. } x \notin A, \quad \chi_A(x) \cdot (\mathcal{L}\varphi)(x) = 0 = \int_A \varphi(y) P(x, dy) = \int \chi_A(y) \cdot \varphi(y) P(x, dy) = (\mathcal{L}(\chi_A\varphi))(x).$$

ending the proof of Lemma 3. □

End of the Proof of Corollary 1.

Proof. First, let us prove Corollary 1 in the particular case $\alpha = 0$. We consider the measurable real function

$$g_n := \chi_A \cdot \varphi_n = \sum_{j=0}^{n-1} \chi_A(\mathcal{L}^j \varphi).$$

Applying Lemma 3, we obtain:

$$g_n = \sum_{j=0}^{n-1} \mathcal{L}^j(\chi_A \cdot \varphi) \quad \mu\text{-a.e.}$$

By construction of g_n , if $x \notin A$ then $g_n(x) = 0$ for all $n \geq 1$. Therefore, if $\sup_{n \geq 1} g_n(x) > 0$ then $x \in A$. Conversely, by hypothesis $A \subset C_0$, hence $\sup_{n \geq 1} g_n(x) > 0$ if $x \in A$. We have proved that

$$A = \{x \in X : \sup_{n \geq 1} g_n(x) > 0\}.$$

Applying Theorem 1 we obtain:

$$\int_A \chi_A \varphi d\mu \geq 0; \quad \int_A \varphi d\mu \geq 0 = \alpha \cdot \mu(A) \quad \text{if } \alpha = 0.$$

We have proved Corollary 1 in the particular case $\alpha = 0$. Now, let us prove it for any real value of α . Consider the function $h_n := \varphi_n - n \cdot \alpha$. Since $\mathcal{L}\alpha = \alpha$ and \mathcal{L} is linear, we obtain:

$$h_n = \sum_{j=0}^{n-1} \mathcal{L}^j(\varphi - \alpha).$$

Besides

$$C_\alpha(\varphi) := \{x \in X : \sup_{n \geq 1} \frac{\varphi_n(x)}{n} > \alpha\} = \{x \in X : \sup_{n \geq 1} \frac{h_n(x)}{n} > 0\} = C_0(h_1).$$

Thus, applying to the measurable function h_1 the result proved in the case $\alpha = 0$, we conclude that

$$\int_A (\varphi - \alpha) d\mu \geq 0; \quad \text{hence } \int_A \varphi \geq \alpha \cdot \mu(A).$$

□

3.3 Proof of Corollary 2

Proof. We apply Corollary 1 to the function $-\varphi$, with $-\beta$ instead of α :

$$\int_A -\varphi d\mu \geq -\beta\mu(A), \quad \text{hence} \quad \int_A \varphi d\mu \leq \beta\mu(A).$$

□

4 Kakutani's Ergodic Theorem

The purpose of this section is to give a proof of the following version of Kakutani's Ergodic Theorem, applied to measures that are stationary (i.e. invariant under the transfer operator \mathcal{L}^*), using the Maximal Ergodic Theorem 1 that we have already proved in Section 3.

Theorem 7. (Kakutani's Ergodic Theorem for \mathcal{L}^* -invariant measures)

If $\mathcal{L}^*\mu = \mu$, then for any $\varphi \in L_\infty$ there exists

$$\tilde{\varphi}(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x) \quad \mu\text{-a.e. } x \in X.$$

Before proving Theorem 7 we will prove a lemmata:

Lemma 4. Let $\varphi : X \mapsto \mathbb{R}$ be a bounded measurable function, and let $\mu \in \mathcal{M}$ such that $\mathcal{L}^*\mu = \mu$. Assume that φ is \mathcal{L} -invariant μ -a.e.; precisely

$$(\mathcal{L}\varphi)(x) = \varphi(x) \quad \mu\text{-a.e. } x \in X.$$

Then, for any real number α the set

$$A_\alpha = \{x \in X : \varphi(x) \geq \alpha\}$$

is \mathcal{L} -invariant μ -a.e.; namely,

$$\chi_{A_\alpha}(x) = (\mathcal{L}\chi_{A_\alpha})(x) = P(x, A_\alpha) \quad \mu\text{-a.e. } x \in X.$$

Proof. By hypothesis $(\mathcal{L}\varphi)(x) = \varphi(x)$ for μ -a.e. $x \in X$. Thus, applying Lemma 1:

$$(\mathcal{L}(\varphi^+))(x) \geq (\mathcal{L}\varphi)^+(x) = \varphi^+(x) \quad \mu\text{-a.e. } x \in X.$$

Besides, applying equality (8) and taking into account that μ is \mathcal{L}^* -invariant, we obtain

$$\int (\mathcal{L}(\varphi^+) - \varphi^+) d\mu = \int \varphi^+ d\mathcal{L}^*\mu - \int \varphi^+ d\mu = \int \varphi^+ d\mu - \int \varphi^+ d\mu = 0.$$

But the integrated function $\mathcal{L}(\varphi^+) - \varphi^+$ is non negative. So it is zero μ -a.e. We have proved that

$$(\mathcal{L}\varphi^+)(x) = \varphi^+(x) \quad \text{for } \mu\text{-a.e. } x \in X.$$

Since $\varphi^+(x) = \chi_{A_0}(x) \cdot \varphi(x)$ for all $x \in X$, we obtain:

$$\chi_{A_0}(x) \cdot \varphi(x) = \varphi^+(x) = (\mathcal{L}(\varphi^+))(x) = \int \varphi^+(y) P(x, dy) \quad \mu\text{-a.e. } x \in X. \tag{21}$$

For all $x \in A_0$, we have $\chi_{A_0}(x) = 1$. Therefore, from equality (21) we deduce:

$$\text{For } \mu\text{-a.e. } x \in A_0, \quad \varphi(x) = \int \varphi^+(y) P(x, dy) \geq \int \varphi(y) P(x, dy), \tag{22}$$

where the inequality at right is an equality only if $\varphi^+(y) = \varphi(y)$ for $P(x, \cdot)$ -a.e. $y \in X$. This latter assertion occurs only if $P(x, A_0) = 1$. By hypothesis, $\varphi(x) = (\mathcal{L}\varphi)(x) = \int \varphi(y)P(x, dy)$. So, the inequality at right in (22) is indeed an equality. We have proved that $P(x, A_0) = 1$ for μ -a.e. $x \in A_0$. In other words:

$$\chi_{A_0}(x) = 1 \Rightarrow P(x, A_0) = 1 \text{ for } \mu\text{-a.e. } x \in X.$$

Therefore $P(x, A_0) \geq \chi_{A_0}(x)$ for μ -a.e. $x \in X$. But

$$\int (P(x, A_0) - \chi_{A_0}(x)) d\mu(x) = \int (\mathcal{L}\chi_{A_0})(x) d\mu(x) - \int \chi_{A_0}(x) d\mu(x) = \int \chi_{A_0} d\mathcal{L}^*\mu - \int \chi_{A_0} d\mu = 0.$$

So, we deduce that

$$P(x, A_0) = \chi_{A_0}(x) \text{ for } \mu\text{-a.e. } x \in X,$$

ending the proof of Lemma 4 in the case $\alpha = 0$.

Now, let us consider any real value of α . Note that

$$A_\alpha(\varphi) := \{x \in X : \varphi \geq \alpha\} = A_0(\varphi - \alpha).$$

Since φ and the constant α are μ -a.e. \mathcal{L} -invariant, also the function $\varphi - \alpha$ is μ -a.e. \mathcal{L} -invariant. So, applying the case above to the function $\varphi - \alpha$, we deduce that the set $A_0(\varphi - \alpha) = A_\alpha(\varphi)$ is μ -a.e. \mathcal{L} -invariant, as wanted. \square

Lemma 5. *Let $\varphi : X \mapsto \mathbb{R}$ be a bounded measurable function, and let $\mu \in \mathcal{M}$ such that $\mathcal{L}^*\mu = \mu$. Assume that φ is \mathcal{L} -invariant μ -a.e.; precisely*

$$(\mathcal{L}\varphi)(x) = \varphi(x) \text{ } \mu\text{-a.e. } x \in X.$$

Then, for any pair of real numbers α and β , the sets

$$C_\alpha = \{x \in X : \varphi(x) > \alpha\}, \text{ and } B_\beta = \{x \in X : \varphi(x) < \beta\}$$

are \mathcal{L} -invariant μ -a.e.; namely,

$$\chi_{C_\alpha}(x) = (\mathcal{L}\chi_{C_\alpha})(x) = P(x, C_\alpha) \text{ and } \chi_{B_\beta}(x) = (\mathcal{L}\chi_{B_\beta})(x) = P(x, B_\beta) \text{ } \mu\text{-a.e. } x \in X,$$

Proof. On the one hand, applying Lemma 4 we know that the set $\{x \in X : \varphi(x) \geq \beta\}$ is \mathcal{L} -invariant μ -a.e. Hence, its complement B_β is also \mathcal{L} -invariant μ -a.e. In other words

$$\chi_{B_\beta} = P(X, B_\beta) \text{ } \mu\text{-a.e. } x \in X.$$

On the other hand, for all $n \geq 1$ we can apply Lemma 4 to the set

$$E_{\alpha+(1/n)} := \{x \in X : \varphi(x) \geq \alpha + (1/n)\}.$$

We deduce that $E_{\alpha+(1/n)}$ is \mathcal{L} -invariant μ -a.e. Thus,

$$\chi_{E_{\alpha+(1/n)}}(x) = P(x, E_{\alpha+(1/n)}) \text{ } \mu\text{-a.e. } x \in X. \tag{23}$$

Besides, $C_\alpha = \bigcup_{n \geq 1} E_{\alpha+(1/n)}$. Thus, $\chi_{C_\alpha}(x) = \lim_{n \rightarrow +\infty} \chi_{E_{\alpha+(1/n)}}(x)$ for all $x \in X$, and by the dominated convergence theorem, we deduce that $P(x, C_\alpha) = \lim_{n \rightarrow +\infty} P(x, E_{\alpha+(1/n)}) \quad \forall x \in X$. Finally, taking $n \rightarrow +\infty$ in equality (23) with x fixed, we obtain $\chi_{C_\alpha}(x) = \lim_{n \rightarrow +\infty} P(x, E_{\alpha+(1/n)}) \text{ } \mu\text{-a.e. } x \in X$, concluding that $\chi_{C_\alpha}(x) = P(x, C_\alpha)$ for μ -a.e. $x \in X$, as wanted. \square

Lemma 6. *Let $\varphi : X \mapsto \mathbb{R}$ be a bounded measurable function, and let α, β be real numbers. Construct the set*

$$A = \{x \in X : \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x) > \alpha, \quad \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x) < \beta\}.$$

Then, for any measure $\mu \in \mathcal{M}$ that is fixed by the transfer operator \mathcal{L}^* , the set A is \mathcal{L} -invariant μ -a.e.

Proof. It is standard to check that the following real functions

$$\psi_1(x) := \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x), \quad \psi_2(x) := \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x),$$

are \mathcal{L} -invariant. Thus, applying Lemma 5 the sets

$$C_\alpha := \{x \in X : \psi_2 > \alpha\}, \quad B_\beta := \{x \in X : \psi_1 < \beta\},$$

satisfy

$$\chi_{C_\alpha}(x) = P(x, C_\alpha) \quad \mu\text{-a.e. } x \in X, \quad \chi_{B_\beta}(x) = P(x, B_\beta) \quad \mu\text{-a.e. } x \in X.$$

On the one hand, for μ -a.e. $x \in C_\alpha \cap B_\beta$, we have $\chi_{C_\alpha} = P(x, C_\alpha) = 1$ and $\chi_{B_\beta} = P(x, B_\beta) = 1$. Since the intersection of two sets of probability 1 also has probability 1, we deduce that

$$P(x, C_\alpha \cap B_\beta) = 1 \quad \text{for } \mu\text{-a.e. } x \in C_\alpha \cap B_\beta.$$

On the other hand, for μ -a.e. $x \notin C_\alpha \cap B_\beta$, we have $\chi_{C_\alpha} = P(x, C_\alpha) = 0$ or $\chi_{B_\beta} = P(x, B_\beta) = 0$. The intersection of two sets, when at least one of them has zero probability, also has zero probability. We deduce that

$$P(x, C_\alpha \cap B_\beta) = 0 \quad \text{for } \mu\text{-a.e. } x \notin C_\alpha \cap B_\beta.$$

Finally, observe that $A = C_\alpha \cap B_\beta$. We conclude that $\chi_A(x) = P(x, A)$ for μ -a.e. $x \in X$, ending the proof of Lemma 6. \square

4.1 Proof of Theorem 7.

Proof. Due to the linearity of the transfer operator \mathcal{L} , it is enough to prove Theorem 7 for real functions $\varphi \in L_\infty$. Denote $\varphi_n := \sum_{j=0}^{n-1} \mathcal{L}^j \varphi$. For any pair of rational numbers α and β such that $\alpha > \beta$, we construct the set

$$A_{\alpha,\beta} = \{x \in X : \limsup_{n \rightarrow +\infty} \frac{1}{n} \varphi_n(x) > \alpha, \liminf_{n \rightarrow +\infty} \frac{1}{n} \varphi_n(x) < \beta\}.$$

Applying Lemma 6, the set $A_{\alpha,\beta}$ is \mathcal{L} -invariant μ -a.e. Besides, if $x \in A_{\alpha,\beta}$, then $\sup_{n \geq 1} \varphi_n(x)/n > \alpha$ and $\inf_{n \geq 1} \varphi_n(x)/n < \beta$. Thus, applying Corollaries 1 and 2, we obtain:

$$\alpha \cdot \mu(A_{\alpha,\beta}) \leq \int_{A_{\alpha,\beta}} \varphi \leq \beta \cdot \mu(A_{\alpha,\beta}).$$

Since $\alpha > \beta$, we deduce that $\mu(A_{\alpha,\beta}) = 0$. The set of all the pair of rational numbers α and β such that $\alpha > \beta$ is countably infinite. Thus, $\mu \left(\bigcup_{\alpha,\beta \in \mathbb{Q}, \alpha > \beta} A_{\alpha,\beta} \right) = 0$. Finally, we observe that

$$\bigcup_{\alpha,\beta \in \mathbb{Q}, \alpha > \beta} A_{\alpha,\beta} = \left\{ x \in X : \lim_{n \rightarrow +\infty} \frac{1}{n} \varphi_n(x) \text{ does not exist} \right\}.$$

to conclude that $\lim_{n \rightarrow +\infty} \frac{1}{n} \varphi_n(x)$ exists μ -a.e. $x \in X$, ending the proof of Theorem 7. \square

5 Ergodic Measures

Proposition 8. *A probability measure μ that is fixed by the operator \mathcal{L}^* is ergodic if and only if any function $\varphi \in L_\infty$ that is μ -a.e. \mathcal{L} -invariant is constant μ -a.e.*

Proof. First, assume that any function $\varphi \in L_\infty$ that is μ -a.e. \mathcal{L} -invariant is constant μ -a.e. Let us prove that μ is ergodic according to Definition 4. Consider a μ -a.e. \mathcal{L} -invariant set $A \subset M$, Equivalently, χ_A is a μ -a.e. \mathcal{L} -invariant function. Thus, it is constant μ -a.e. Since χ_A can take only the values 1 or 0, we deduce that either $\chi_A(x) = 0$ for μ -a.e. $x \in X$, or $\chi_A(x) = 1$ for μ -a.e. $x \in X$. In other words, either $\mu(A) = 0$ or $\mu(A) = 1$, proving that μ is ergodic.

Conversely, assume that μ is ergodic according to Definition 4, and consider any function $\varphi \in L_\infty$ that is \mathcal{L} -invariant μ -a.e. Denote

$$A_\alpha := \{x \in X : \varphi(x) \geq \alpha\}.$$

Applying Lemma 4 the set A_α is \mathcal{L} -invariant μ -a.e. So, from Definition 4, we deduce that $\mu(A_\alpha) \in \{0, 1\}$ for all $\alpha \in \mathbb{R}$. By construction $\mu(A_\alpha)$ is non increasing with α , and it is zero for all values of α large enough (because φ is bounded). Thus, there exists

$$k := \sup\{\alpha \in \mathbb{R} : \mu(A_\alpha) = 1\} \in \mathbb{R}.$$

We deduce that $\mu(A_{k+\epsilon}) = 0$ and $\mu(A_{k-\epsilon}) = 1$ for all $\epsilon > 0$. So

$$\mu\left((X \setminus A_{k+\epsilon}) \cap A_{k-\epsilon}\right) = 1 \quad \forall \epsilon > 0.$$

In other words:

$$\mu\left(\left\{x \in X : k - \epsilon \leq \varphi(x) < k + \epsilon\right\}\right) = 1 \quad \forall \epsilon > 0.$$

In particular,

$$\mu\left(\left\{x \in X : k - \frac{1}{n} \leq \varphi(x) < k + \frac{1}{n}\right\}\right) = 1 \quad \forall n \geq 1.$$

Then,

$$\mu\left(\bigcap_{n \geq 1} \left\{x \in X : k - \frac{1}{n} \leq \varphi(x) < k + \frac{1}{n}\right\}\right) = 1,$$

or equivalently $\varphi(x) = k$ for μ -a.e. $x \in X$. We have proved Proposition 8. □

Corollary 3. (Ergodic Theorem for \mathcal{L}^* ergodic measures)

If μ is ergodic for the operator \mathcal{L}^ , then for all $\varphi \in L_\infty$:*

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x) = \int \varphi d\mu \quad \mu\text{-a.e. } x \in X.$$

Proof. Applying Theorem 7 the above limit exists for μ -a.e. $x \in X$. Denote it by $\tilde{\varphi}(x)$. Since the function $\tilde{\varphi}$ is \mathcal{L} -invariant μ -a.e., and the measure μ is ergodic, we apply Proposition 8 to deduce that there exists a constant k such that $\tilde{\varphi}(x) = k$ for μ -a.e. $x \in X$. Now, it is enough to check that $k = \int \varphi d\mu$. In fact, by the dominated convergence theorem, we have

$$\int \tilde{\varphi} d\mu = \int \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j \varphi)(x) d\mu(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \mathcal{L}^j \varphi d\mu.$$

But taking into account that μ is \mathcal{L}^* -invariant, we have

$$\int \mathcal{L}^j \varphi d\mu = \int \varphi d\mathcal{L}^{*j} \mu = \int \varphi d\mu.$$

Therefore,

$$\int \tilde{\varphi} d\mu = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \int \varphi d\mu = \int \varphi d\mu.$$

Finally, since $\tilde{\varphi} = k$ for μ -a.e. $x \in X$, the above equality implies that $k = \int \varphi d\mu$, as wanted. □

5.1 End of the Proof of Theorem 2

Proof. For each fixed natural value of $p \geq 1$, the operator \mathcal{L}^p transforms continuous functions into continuous functions. Besides, it is positive, bounded by 1, and $\mathcal{L}^p(1) = 1$. So, \mathcal{L}^p satisfies Definition 1 and is a transfer operator. Besides, after applying equality (2) p times, we deduce that the dual transfer operator of \mathcal{L}^p is \mathcal{L}^{*p} . So, we can apply all the results proved along this paper to \mathcal{L}^p instead of \mathcal{L} , and to \mathcal{L}^{*p} instead of \mathcal{L}^* . In particular, Theorem 2 follows from Theorem 7 and Corollary 3, using \mathcal{L}^p instead of \mathcal{L} , and \mathcal{L}^{*p} instead of \mathcal{L}^* . \square

6 Example

6.1 One-dimensional Piecewise Contractive Affine System Plus Uniform Noise.

In [22, section 3.1] it is posed the deterministic piecewise contractive affine system defined by infinite iteration of the recurrence equation

$$x_{n+1} = f(x_n), \quad f(x) = \alpha x + \beta \pmod{1},$$

where $0 < \alpha, \beta < 1$ are fixed real parameters, $x \in [0, 1)$ is the state variable, and “(mod. 1)” denotes that $f(x)$ is the fractionary part of $\alpha x + \beta$ at each iteration. This system is one of the simplest nonlinear models that one can take for a mathematical model of a physical phenomenon in dimension one, with a discontinuity. Its graphic increases linearly, with slope $\alpha < 1$, until it reaches the level 1, at which it is reset to zero, from where it starts increasing again linearly, with the same slope, until the independent variable reaches the level 1. Even if in each of both continuous pieces, the graphic is affine (linear), the system is nonlinear due to the discontinuity point. One should be aware that the independent variable of that graphic is not time, but the previous spacial state of the system. Time is the number n of iterations of the recurrent process.

Although the recurrence equation is very simple, and the system is typically non chaotic, it is not possible to solve the equation explicitly by means of a formula, to find the state of the system for any time $n \in \mathbb{N}$ and its asymptotic behaviour as time $n \rightarrow +\infty$. Nevertheless, for certain values of the parameters, it is checked that the rotation number is rational (see [22, Figure 1 and Table 1]); hence the asymptotic attractor is a periodic orbit. Precisely, for $\alpha = 0.8$ and $\beta = 0.5$ the attractor is a periodic orbit of period 5 and its rotation number is $2/5$.

In [22, Section 3.2], it is added a noise with uniform distribution in a neighborhood of radius ϵ (amplitude of the noise) at each iteration of the deterministic (zero-noise) described above. Precisely, the noisy recurrent equation is now

$$x_{n+1} = f(x_n) + \omega_n \pmod{1},$$

where ω_n is a random real variable (the *noise*) with a probability distribution $P(x_n, \cdot)$ that has *uniform density* with respect to Lebesgue, supported on the interval $(f(x_n) - \epsilon, f(x_n) + \epsilon)$. The purpose in the article [22] is to find, by computer experiments an approximation of the density of the stationary ergodic measure for $\alpha = 0.8$, $\beta = 0.5$ and four different values of the noise amplitude $\epsilon = 0.02; 0.07; 0.3$ and 0.8 (see [22, Figures 2, 3, 4 left and 4 right], respectively). We remark that the densities of the stationary ergodic measures so obtained in all the cases studied in [22] are not constant. In other words, even if the probability of the noise is uniformly distributed, the stationary ergodic probability measures are not necessarily uniformly distributed.

We denote by ρ the density function of the stationary ergodic measure μ , whose graphic is approximately found in [22, Figures 2, 3, 4 left and 4 right] in that example, after computer numerical experiments. Therefore, for any bounded function φ , its integration with respect to the stationary ergodic measure μ is computed by

$$\int \varphi d\mu = \int_0^1 \varphi(x) \cdot \rho(x) dx.$$

What is the meaning and relevance of finding the stationary ergodic measures and their densities of the noisy system in that example? To answer this question, we will apply Theorem 2, precisely its equalities (4) and (5) to this example.

For any state x of the noisy system, let us take as *observable function* $\varphi(x)$ the state x itself. In other N dimensional examples, for which the state x of the system is vectorial and has N real coordinates, one can take a single coordinate of x as the observable function $\varphi(x)$, or any other real function of x that takes values in a (bounded) interval of the real line. From the construction of the transfer operator by equality (1), $(\mathcal{L}^j\varphi)(x)$ is the expected value (E.V.) at time j of the observable state of the system $\varphi(x_j)$ along the noisy orbits x_0, x_1, \dots, x_j with initial state $x_0 = x$. Since in our example $\varphi(x_j) = x_j$, we have

$$(\mathcal{L}^j\varphi)(x_0) = \text{E.V.}(x_j).$$

Therefore, the following equality gives the E.V. of the *time average* along the noisy orbits of the system up to time $n \geq 1$:

$$\text{E.V. (time average up to time } n) = \text{E.V.}\left(\frac{x_0 + x_1 + \dots + x_{n-1}}{n}\right) = \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j\varphi)(x_0).$$

Now, applying equality (4) of Theorem 2, there exists the limit as time $n \rightarrow +\infty$ of the E.V. of the time averages, which is called E.V. of *asymptotic time average*:

$$\text{E.V. asymptotic time average} = \lim_{n \rightarrow +\infty} \text{E.V.}\left(\frac{x_0 + x_1 + \dots + x_{n-1}}{n}\right) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} (\mathcal{L}^j\varphi)(x_0) \quad \mu - \text{a.e.}$$

The relevant conclusion is given by applying equality (5) of Theorem 2:

$$\text{E.V. asymptotic time average} = \int \varphi(x) d\mu = \int_0^1 x \cdot \rho(x) dx,$$

meaning that the E.V of the asymptotic time average, hence the E.V. of the *time averages for time large enough, equals the spacial average with respect to the ergodic stationary probability* μ . This is a version like Boltzmann hypothesis, applied to the example of noisy system that was posed at the beginning.

7 Conclusions and Further Research

We have proved the Maximal Ergodic Theorem 1 and the Pointwise Ergodic Theorem 2 of Periodic Measures, for the transfer operator that is associated to a Markovian stochastic dynamical system, obtained by adding noise with any probability distribution at each iterate of a deterministic continuous system. As a consequence, we have also proved Corollaries 1 and 2, which provide different statements of the Maximal Ergodic Theorem for noisy systems. Besides, we have proved Kakutani's Ergodic Theorem 7, also as a consequence of them.

The relevance of the ergodic theorems for the transfer operators that are proved along this paper, is the extension they provide to stochastic markovian processes, of the classical pointwise ergodic theorems for deterministic systems. They hold in particular for noisy systems, i.e. for dynamical systems which are randomly perturbed by noise, independently on the probabilistic distribution of the noise. The interpretation of their meaning allows their application to other sciences and engineering, although the results are proven by pure mathematics. In fact, a mathematical model of the dynamics of certain physical phenomenon or human-made technology, may be not purely deterministic. "The real world" which is modeled, for instance by a differential equation, usually behaves (or is perceived by the observer or is constructed by the engineer), as noisy, exhibiting random perturbations, more or less near a supposed deterministic model. This noise may be due to multiple causes. For instance, the physical phenomenon may need much more complexity to be completely described than the variables and parameters that are considered in the simplified mathematical equations. Instead of taking its full complexity as it is, it may be convenient to add certain type of random perturbations to a simplified mathematical model. Also the unavoidable inexactitude of the experimental data and of the observations, from which the mathematical deterministic model was designed, may require the consideration of a noisy dynamics. Finally, the noisy dynamical systems, and the application of the ergodic theorems that were proved along this paper, may

mathematically explain better some physical phenomena, just because epistemologically, the intrinsic nature of them may be not deterministic, but have predominant random components. We cite Kifer in [23, Introduction, p. 1]:

“Mathematicians often face the question to which extent mathematical models describe processes of the real world. These models are derived from experimental data, hence they describe real phenomena only approximately. ... Global stability in the presence of noise ... can be described as recovering parameters of dynamical systems from the study of their random perturbations. ... In this way (they) can be considered ... having physical sense.”

Finally, we propose some related subjects for further research:

- a) Estimates for the rates of convergence of the time averages for the pointwise convergence ergodic theorems: the abstract tools used in the proof of the ergodic theorems for the transfer operator of noisy systems, may be also used to extend to stochastic dynamical systems the estimates of the rate of convergence and deviation, already obtained for deterministic systems for instance in [25] and [26].
- b) Relationships between the ergodic components of stationary or periodic measures that satisfy Theorems 2 and 7, and the spectral properties of the invariant measures supported on the attractor: we propose the research of the possible extensions to stochastic dynamical systems of the mixing properties, and of the almost periodicity or asymptotic periodicity of the transfer operators, as for instance is proved in [27] and [28] in some particular cases.

Acknowledgments. The author thanks CSIC of Universidad de la República (Uruguay) for the partial financial support of Project of the group "Sistemas Dinámicos".

References

1. G.D. Birkhoff, "Proof of the Ergodic Theorem", *Proc. Nat. Acad. Sciences*, vol. 17, pp. 656–660, 1931.
2. R. Mañé, *Ergodic Theory and Differentiable Dynamics*, Springer-Verlag, Berlin, 1987.
3. A. Katok, B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge Univ. Press, Cambridge, 1999.
4. I.P. Cornfeld, S.V. Fomin, Ya. G. Sinai, *Ergodic Theory*, Springer-Verlag, Berlin, Re-edition 2012.
5. J.F. Kingman, "Subadditive Ergodic Theory", *Annals Prob.*, vol. 1. pp. 889–909, 1973.
6. S. Gouëzel, A. Karlsson, "Subadditive and Multiplicative Ergodic Theorems", *ArXiv*: 1509.07733 [math.DS], 2015.
7. V.I. Oseledets, "A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems", *Trans. Moscow Math. Soc.*, vol. 19, pp. 197–231, 1968.
8. S. Filip, "Notes on the multiplicative ergodic theorem", *Ergodic Th. & Dyn. Sys.*, Published on line, 2017.
9. E. Catsigeras, *Teoría Ergódica de los Sistemas Dinámicos Discretos* (in Spanish), Univ. de la República, Montevideo, 2013.
10. T. Eisner, B. Farkas, M. Haase, R. Nagel, *Operator Theoretic Aspects of Ergodic Theory*, Springer, Heidelberg-New York, 2015.
11. T. Eisner, D. Kunszenti-Kovács, "On the entangled ergodic theorem", *Ann. Scuola Norm. Super. Pisa Cl. Sci.*, vol XII, pp. 141–156, 2013.
12. A. Tempelman, *Ergodic theorems for group actions: Informational and Thermodynamical Aspects*, Springer, Dordrecht, Re-Edition 2013.
13. A. Karlsson, F. Ledrappier, "Noncommutative Ergodic Theorems", *ArXiv* 1110.6847 [math.DS], 2011.
14. H.H. Rugh, D. Thomine, "Ratio ergodic theorems: from Hopf to Birkhoff and Kingman", *ArXiv* 1802.08439, 2018.
15. J. Aaronson, *An Introduction to Infinite Ergodic Theory*, Math. Surveys and Monographs Vol. 50, Amer. Math. Soc., Providence, 1997.
16. E. Roy, "Poisson suspensions and infinite ergodic theory", *Erg. Th. & Dyn. Sys.*, vol. 29, pp. 667–683, 2009.
17. T. Austin, "On the norm convergence of non-conventional ergodic averages", *Erg. Th. & Dyn. Sys.*, vol. 30, pp. 321–338, 2010.
18. T. de la Rue, "Notes on Austin's multiple ergodic theorem", *ArXiv* 0907.0538, 2009.
19. C. González-Tokman, A. Quass, "A semi-invertible operator Oseledets theorem", *Erg. Th. & Dyn. Sys.*, vol. 34, pp. 1230–1272, 2014.

20. S. Kakutani, "Random ergodic theorems and Markoff processes with a stable distribution", . *Proc 2nd. Berkeley Symp.*, pp. 247–261, 1951.
21. K. Yosida, *Functional Analysis*, Springer-Verlag, Heidelberg, 1978.
22. D. Faranda, J. Milhazes Freitas, P. Guiraud, S. Vaienti, "Statistical properties of random dynamical systems with contracting direction", *Journ. of Physics A: Math. Theor.* , vol. 49, No. 204001, 2016.
23. Y. Kifer: *Ergodic theory of random transformations*, Birkhäuser, Boston, 1986.
24. L. Arnold: *Random dynamical systems*. Springer-Verlag, Heidelberg–Berlin, 1998.
25. A.G. Kachurovskii, V.V. Sedalishchev, "Constants in estimates for the rates of convergence in von Neumann's and Birkhoff's ergodic theorems", *Sbornik: Mathematics*, vol. 202, 2011.
26. L. Rey-Bellet, L.S. Young, "Large deviations in non-uniformly hyperbolic dynamical systems", *Erg. Th. & Dyn. Sys.*, vol. 28, pp. 587–612, 2008.
27. I. Mezic, "Spectral Properties of Dynamical Systems, Model Reduction and Decompositions", *Nonlinear Dynamics*, vol. 41, pp. 309–325, 2005.
28. W. Bahsoun, H. Hu, S. Vaienti, "Pseudo-Orbits, Stationary Measures and Metastability", *Dynamical Systems*, vol. 29, pp. 322–336, 2014.