Instituto Potosino de Investigación Científica y Tecnológica

División de Control y Sistemas Dinámicos

TÍTULO DE LA TESIS:

Central Limit Theorem for a Class of Contractive Random Dynamical Systems and Critical Behavior in Connected Dynamical Systems

QUE PARA OBTENER EL GRADO DE:

Doctor en Control y Sistemas Dinámicos

PRESENTA:

Hugo Alberto Nieto Loredo

DIRECTOR DE LA TESIS:

Dr. César Maldonado Ahumada

San Luis Potosí, S.L.P.
13 de junio de 2024

Constancia de aprobación de la tesis

Dr. César Octavio Maldonado Ahumada
Director de la tesis

Dr. Haret-Codratian Rosu Barbus
Presidente

Dr. Sergio Iván López Ortega
Jurado en el examen

Dr. Roberto Carlos Álvarez Martínez
Jurado en el examen

Dr. Eric Campos Cantón
Jurado en el examen

Créditos institucionales

Esta tesis fue elaborada en la División de Control y Sistemas Dinámicos del Instituto Potosino de Investigación Científica y Tecnológica, A.C., bajo la dirección del Dr. César Octavio Maldonado Ahumada.

Durante la realización del trabajo el autor recibió una beca académica del Consejo Nacional de Ciencia y Tecnología (862351) y del Instituto Potosino de Investigación Científica y Tecnológica, A. C.

(Página en blanco que se va a utilizar para colocar la copia del acta de examen.)

Resumen

En esta tesis usamos el método espectral para probar un «teorema de límite central» para una clase de sistemas dinámicos aleatorios contractivos. En este método se parte de la existencia de un espacio normado complejo en el que el operador de Perron-Frobenius tiene la propiedad espectral conocida como «quasicompacidad». Posteriormente, se define un nuevo operador mediante una perturbación analítica del operador de Perron-Frobenius y se utiliza teoría de perturbaciones para mostrar que este nuevo operador tiene las mismas propiedades espectrales que el operador de Perron-Frobenius. Finalmente, se expresa la función característica de la variable aleatoria de interés en términos de las iteraciones del operador perturbado y se aplica el teorema de Lèvy para mostrar la convergencia a una distribución normal.

Principalmente, nos basamos en los resultados de la referencia [61] en el que se muestra la quasicompacidad de la medida *pushforward* en una clase de sistemas dinámicos aleatorios que son contractivos. Estos sistemas pueden verse como una transformación en un producto cartesiano en los que la base es un subshift de tipo finito y cuyo operador de Perron-Frobenius correspondiente es quasicompacto. Para la aplicación del método espectral, extendemos este resultado de la quasicompacidad de la medida *pushforward*, de forma que siga siendo válido en un espacio vectorial normado de medidas complejas. Además, obtenemos una cota para la rapidez de convergencia en el Teorema de Límite Central conocida comúnmente como desigualdad de Berry-Esseen. También ilustramos numéricamente el resultado utilizando un mapeo aleatorio contractivo en el intervalo y un sistema de funciones iteradas (SFI).

En la parte final de la tesis, estudiamos también un prototipo de un sistema dinámico formado por una familia de mapeos en el que, en cada iteración, se selecciona uno de estos mapeos mediante una dinámica determinista que depende del mapeo seleccionado en la iteración previa y del valor de la iteración. Nos enfocamos principalmente en el caso de dos mapeos *conectados* mediante huecos en el espacio fase. Más específicamente, la dinámica es la siguiente: se itera el mapeo T_1 hasta que el valor de la iteración caiga en el hueco H_1. Cuando esto suceda, se selecciona el mapeo T_2 y se continúa iterando este mapeo hasta que el valor de la iteración caiga en el hueco H_2. Cuando esto suceda, volvemos a iterar el mapeo T_1 y seguimos aplicando las mismas reglas para cambiar de mapeo. Para este tipo de sistemas, exploramos numéricamente comportamientos críticos como la aparición de ciclos periódicos y orden inducido.

Abstract

In this thesis we use the spectral method to prove a Central Limit Theorem (CLT) for a class of contractive random dynamical systems. In this method, we start from the existence of a complex normed space in which the Perron-Frobenius operator has the spectral property known as “quasicompactness”. Then, we define a new operator as analytical perturbation of the Perron-Frobenius operator and use “Perturbation Theory” to show that this new operator has the same spectral properties than the Perron-Frobenius operator. Finally, we write the characteristic function of the random variable of interest in terms of the iterations of the perturbed operator and we apply Lèvy Theorem to show convergence to a normal distribution.

Mainly, our study is based on the recent results of reference [61] where the authors show quasicompactness of the *pushforward* measure in a class of contractive Random Dynamical Systems. These systems can be seen as a skew product transformation in which the base is a subshift of finite type and whose corresponding Perron-Frobenius operator is quasicompact. For the application of the Spectral Method, we extended this result of quasicompactness of the pushforward measure, so that it remains true on a normed vector space of complex measures. Additionally, we obtain an upper bound for the speed of convergence on the Central Limit Theorem commonly known as Berry-Esseen inequality. Also, we numerically illustrate this result using a contractive Random Dynamical System on the interval and an Iterated Function System (IFS).

In the final part of this thesis, we also study a prototype of dynamical system formed by a family of maps in the interval for which, in each iteration, one of these maps is selected through a deterministic dynamics that depends on the maps that was selected in the previous iteration and the value of the iteration. We focus on the case of two maps *connected* through holes on the space state. More specifically, the dynamics is the following: we iterate map T_1 until the value of the iteration falls in the hole H_1. When this happens, we select the map T_2 and continue iterating this map until the value of the iteration falls in H_2. When this happens, we iterate the map T_1 again and continue applying the same rules to change maps. For these type of systems, we numerically explore critical behaviors such as the appearance of periodic cycles or induced order.

Contents

1. Preliminaries

1.1. Introduction

Broadly speaking, a dynamical system is a *rule* that specifies how a system evolves in time given an initial condition. The collection of all possible states of the system is called the **state space** and is denoted as X. The time in which the systems evolves can be considered either continuous or discrete. For continuous-time dynamical systems, the rule takes the form of a set of differential equations whose solution describes the state of the system at any time. For discrete-time dynamical systems, the rule takes the form of a transformation from the space state to itself $T : X \to X$. In this thesis we restrict ourselves to the study of discrete-time dynamical systems.

Ergodic Theory is the mathematical study of long term behavior of different classes of discrete dynamical systems. Given $x \in X$, the sequence $x, T(x), T^2(x), \ldots$ represents the states of the system and is called the *orbit* of x under T. The state of the system is observed using a function $f : X \to \mathbb{R}$ which is called *observable*. The scenario is depicted in Figure 1.1. Let $S_n f(x) = \sum_{j=0}^{n-1} f(T^j(x))$ be the partial sum of the first n observed values. One of the central questions in Ergodic Theory is to determine under which conditions the following limit exists:

$$f^*(x) = \lim_{n\to\infty} \frac{S_n f(x)}{n}. \tag{1.1}$$

This question can be answered using the well-known Birkhoff ergodic theorem which is formulated in the context of measure theory as follows (see for example Theorem 2.3 from [76]). Suppose that X forms a probability space $(X, \mathcal{B}, \mu)$. When the measure μ satisfies $\mu(A) = \mu(T^{-1}A)$ for all measurable set $A \subseteq X$, T is said to be a **measure preserving transformation**. Under the same condition, μ is said to be a T**-invariant measure**, or in simpler terms, μ is said to be an **invariant measure** of the system. A system having an invariant measure, basically means that the probabilities of events on that system do not change in time (under that probability

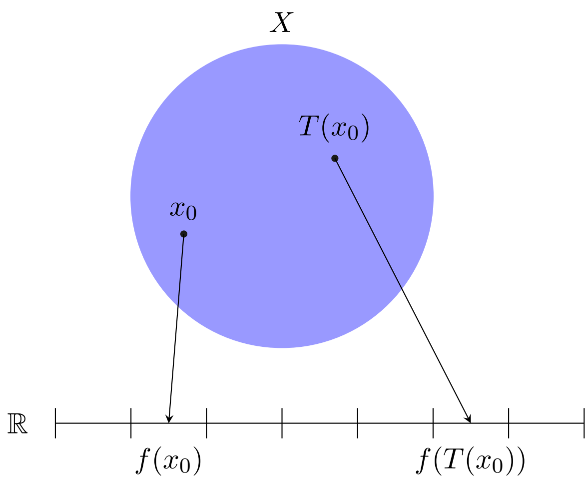


Figure 1.1: The basic setting of a discrete dynamical system with observable $f : X \to \mathbb{R}$.

measure). Such measures are, of course, of great interest from the practical point of view. One can check (see for example Proposition 3.1.1 from [18]) that μ is $T-$invariant if and only if for each continuous observable $\psi : X \to \mathbb{R}$:

$$\int \psi d\mu = \int \psi \circ T d\mu. \tag{1.2}$$

Birkhoff ergodic theorem (see for example Theorem 2.3 from [76]) states that if f is integrable with respect to μ, then the limit:

$$f^*(x) = \lim_{n\to\infty} \frac{S_n f(x)}{n}, \tag{1.3}$$

exists for almost all $x \in X$ (with respect to the measure μ), meaning that the measure of the sets where the limit does not exist, is zero. Moreover, under the assumption that the measure is ergodic, the theorem also gives an explicit value of the limit. Recall that a probability measure is said to be **ergodic** if it is T-invariant and for each $A \in \mathcal{B}$ that satisfies $T^{-1}A = A$, one has that either $\mu(A) = 0$ or $\mu(A) = 1$. Under this assumption, the limit function f^* is simply:

$$f^*(x) = \int f d\mu. \tag{1.4}$$

As a particular case, when f is the indicator function of a measurable set A, i.e.

$$f(x) = \chi_A(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A, \end{cases} \tag{1.5}$$

Birkhoff ergodic theorem reduces to the so called, strong law of large numbers. Indeed, in this case, Equations (1.3) and (1.4) become:

$$\lim_{n\to\infty} \frac{S_n \chi_A(x)}{n} = \chi_A^*(x) = \int \chi_A(x) d\mu = \mu(A) \qquad \mu - a.e.\ x. \tag{1.6}$$

This means that $\frac{1}{n} S_n \chi_A$ converges almost everywhere to the measure of set A, which is exactly what is stated in the strong law of large numbers. [1].

Rather than being interested on the limit behavior of the average only, one can ask about other statistical properties. For example, one can be interested in knowing whether the sequence $\frac{S_n f}{\sqrt{n}}$ satisfies a Central Limit Theorem as in the case of independent and identically distributed (IID for short) random variables. In this thesis we study this specific question for the class of contractive random dynamical systems presented in Chapter 3. In Section 1.4 we also give a brief overview of other statistical properties of general interest. It is important to remark that, given an initial condition x_0, the sequence $x_0, ..., T^n(x_0)$ is not random and therefore, the derived sequence $\frac{S_n f}{\sqrt{n}}$ is not a random variable. However, it turns out that if the transformation is chaotic enough, then the sequence $x_0, ..., T^n(x_0)$ behaves as a random process and $\frac{S_n f}{\sqrt{n}}$ can be treated as a random variable.

Another important remark is that the statistical properties are stated in terms of an invariant measure. That is why invariant measures are one of the most important notions in the study of dynamical systems. When X is a compact metric space and T is a continuous transformation on X, the existence of an invariant measure is guaranteed by the Krylov-Bogolioubov theorem (see for example Theorem 1.2.2 from [18]). However, it can be the case that a dynamical system has multiple invariant measures. In those cases, if there is already a natural measure of interest ν, we are interested in invariant measures μ that satisfy the following: $\mu(A) = 0$ whenever $\nu(A) = 0$. If

[1] In fact, almost everywhere convergence implies convergence in measure, which is the type of convergence stated in the weak law of large numbers. The converse statement is not true in general.

this condition holds, the invariant measure μ is said to be **absolutely continuous** with respect to the reference measure ν and is denoted $\mu \ll \nu$. When $\mu \ll \nu$ the Radon-Nikodym theorem (see for example Theorem 2.2.5 from [18]) guarantees the existence of a unique, non-negative measurable function $\phi : X \to \mathbb{R}$ such that:

$$\mu(A) = \int_A \phi d\nu$$

for every μ-measurable set A. The function ϕ is called the **invariant density**. When ν is the Lebesgue measure and $\mu \ll \nu$, the measure μ is said to be an **acim** (short for absolutely continuous invariant measure) with respect to ν.

It turns out that invariant densities of acim's are fixed points of a certain operator called the **Perron-Frobenius operator**. Moreover, a particularly important spectral property of the Perron-Frobenius operator called **quasicompactness**, is closely related to the existence of statistical properties mentioned above. A formal definition of quasicompactness is given by Definition 2.1.1, but intuitively, the spectrum of a quasicompact operator consists of two disjoint parts: (1) a continuum of points that are contained on a disk of radius r and (2) finitely many points having magnitude is strictly greater than r. Due to the gap between these two disjoint parts of the spectrum, this property is also referred to as **spectral gap** property. A particularly important case found in many applications is the case where the second part of the spectrum contains only the single point $(1, 0)$. The method that uses quasicompactness to obtain statistical properties of a dynamical system is called **Nagaev-Guivarc'h method** or simply the **spectral method**. A detailed explanation of the method and a classical example of its application is presented on Section 2.1.

The spectral method requires the existence of a normed vector space of functions where the Perron-Frobenius operator is quasicompact, but it turns out that proving quasicompactness is not a trivial task. A common way to prove an operator is quasicompact is given by the Ionescu Tulcea-Marinescu theorem (see for example theorem II.5 from [39]). In this thesis, however, we use an existing result where quasicompactness is proved without the use of Ionescu Tulcea-Marinescu theorem. This result is explained in detail in Section 3.2.

1.2. Examples of discrete-time Dynamical Systems

In this section we briefly present examples of discrete Dynamical Systems that have been widely studied in the literature: maps in the interval and subshifts of finite type. We provide conditions and references in which the existence of invariant measures is established. In Section 1.4, when we start the discussion on statistical properties, we make use of these examples and illustrate the known results in each case.

1.2.1 Maps in the interval

A map in the interval is a type of discrete-time Dynamical System which arises when we set $X = [0, 1]$ and $T : X \to X$. An important class of maps in the interval is obtained when the transformation satisfies some regularity and expansive conditions. More specifically, suppose that T satisfies the following:

R1: There exists a partition $0 = a_0 < a_1 < \cdots < a_q = 1$ such that, for $i = 0, ..., q$, $T|_{(a_{i-1},a_i)}$ is a C^2 function and can be extended to $[a_{i-1}, a_i]$ as a C^2 function.

R2: Furthermore, suppose that:

$$\inf_{x \in I} \left| T'(x) \right| > 1. \tag{1.7}$$

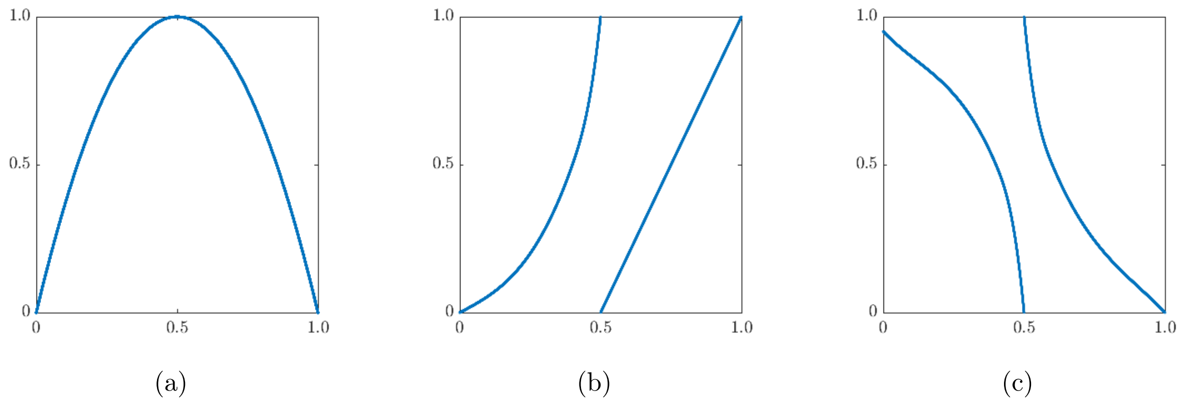


Figure 1.2: Common examples of maps in the interval. (a) Logistic map, used to model population growth subject to constraints. (b) Manneville-Pomeau map, used to model intermittency. (c) Lorenz map, obtained as a 1D projection of the Lorenz flow on a Poincare section.

Such maps are called **expanding maps in the interval** or **Lasota-Yorke maps** and in the classical reference [59], the authors use Functional Analysis techniques to prove that the corresponding Perron-Frobenius operator has a fixed point in the space of functions of bounded variation, and therefore, they prove the existence of an acim. The proof is very constructive and it uses results such as Mazur's Lemma (see Section 1.2, page 6 from [21]), Kakutani-Yosida Theorem (Corollary VIII 5.3 from [19]) and Helly's Lemma (Lemma 2.3.1 from [18]). In that same reference, the authors obtain similar conclusions for piecewise C^2 transformations with a countable number of pieces. In this case, the main condition for the transformation to have an acim is that it satisfies the inequality $\inf_{x\in I}|T'(x)| > 2$.

Although the class of Lasota-Yorke maps is large, many interesting maps in the interval do not satisfy conditions R1 and R2. Three of these maps are depicted in Figure 1.2. Let us now briefly mention some of the results for such maps. On reference [48] the authors prove that a smooth family of one-parameter maps in the interval of the form $T_r(x) = rT(x)$ has an acim on a set of positive Lebesgue measure on the r-axis. This result includes the **Logistic map** $T_4(x) = 4x(1-x)$ shown on Figure 1.2a which does not satisfy Equation (1.7) for $x \in [3/8, 5/8]$.

The **Manneville-Pomeau** map (also known as Liverani-Saussol-Vaienti map) depicted on Figure 1.2b, was proposed as a simple model for intermittency:

$$T(x) = \begin{cases} x(1+2^\alpha x^\alpha) & x < \frac{1}{2}, \\ 2x-1 & x \geq \frac{1}{2}. \end{cases} \tag{1.8}$$

The case where $\alpha = 0$ reduces to the well-known doubling map $T(x) = 2x(\mod 1)$ which satisfies R1 and R2. However, for $\alpha > 0$, the expression (1.7) is violated on $x = 0$ where the map has an indifferent fixed point (meaning that $|T'(0)| = 1$). On this regime, the authors of reference [64] use an interesting probabilistic approach to prove the existence of an acim for $0 < \alpha < 1$.

The case of the **Lorenz map** shown on Figure 1.2c is slightly more complex. The map is given by:

$$T(x) = \begin{cases} \theta|x-0.5|^\alpha & x < 0.5, \\ 1-\theta|x-0.5|^\alpha & x \geq 0.5, \end{cases} \tag{1.9}$$

where $\theta = 109/64$ and $\alpha = 51/64$. It satisfies condition R2 but the derivative on $x = 1/2$ is not defined and therefore $T|_{[0,1/2]}$ is not C^2 (not even C^1). The existence of an acim for a class of maps that includes the Lorenz map was first proved in [53]. More specifically, the author proves quasicompactness of the Perron-Frobenius operator when acting on L_1 under the condition that T is piecewise continuous and monotonous and the inverse of the derivative of the transformation belongs to the space of generalized bounded variation functions. We will revisit the Lorenz map on Chapter 5 where we introduce the notion of a Connected Dynamical System.

1.2.2 Subshifts of finite type

Let $S = \{1, ..., N\}$ and let A be an $N \times N$ matrix whose elements are either 0 or 1. The set [2]:

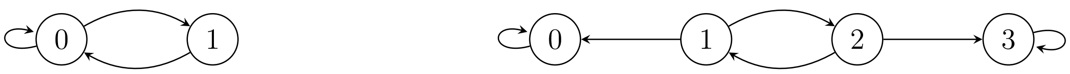
$$\Sigma_A^+ = \left\{ \underline{x} \in S^{\mathbb{N}_0} : A_{x_i, x_{i+1}} = 1, i \geq 0 \right\},$$

together with the map $\sigma : \Sigma_A^+ \to \Sigma_A^+$ given by $(\sigma \underline{x})_i = x_{i+1}$ is called a **subshift of finite type**. The set S is called **alphabet** and the matrix A is called **adjacency matrix**. When the alphabet S is not finite, but countable infinite, the system is called **countable subshift**. In this section (and in most of the thesis) we will mainly focus in subshifts of finite type with finite alphabet. However, in Section 1.4.6 we will talk about countably subshifts with more detail.

A subshift of finite type is also called topological Markov chain. Of course, a Markov chain with transition matrix B is also a subshift of finite type with adjacency matrix given by:

$$A_{i,j} = \begin{cases} 1 & B_{i,j} > 0, \\ 0 & B_{i,j} = 0, \end{cases} \tag{1.10}$$

A graphical representation of two simple (yet interesting) subshifts of finite type, is shown in Figure 1.3.



(a) (b)

Figure 1.3: Graphical representation of two subshifts of finite type (a) with $N = 2$ and $A_{2,2} = 0$ and $A_{i,j} = 1$ in any other case. This system is known as golden subshift of finite type. (b) with $N = 4$ and $A_{1,1} = A_{4,4} = 1$, $A_{k,k-1} = A_{k,k+1} = 1$ for $k = 2, 3$ and $A_{i,j} = 0$ in any other case. This system is called Gambler's ruin (see Section 2.4 and Example 3.10.16 from reference [14]).

Let us now endow Σ_A^+ with some mathematical structure. The set Σ_A^+ can be endowed with the product topology of the discrete topology on S becoming a compact topological space. A base for this topology is formed with **cylinder sets** which are sets of the form:

$$[a_0, ..., a_{n-1}] = \left\{ \underline{x} \in \Sigma_A^+ : x_0 = a_0, ..., x_{n-1} = a_{n-1} \right\}. \tag{1.11}$$

The same topology can be obtained by defining a distance on Σ_A^+ given by:

$$d_\theta(\underline{x}, \underline{y}) = \sum_{i=0}^{\infty} \theta^i \left(1 - \tau(x_i, y_i)\right). \tag{1.12}$$

Where $\theta \in (0, 1)$ and:

$$\tau(x_i, y_i) = \begin{cases} 0 & x_i \neq y_i, \\ 1 & x_i = y_i, \end{cases}$$

Based on this topology, Σ_A^+ can be endowed with the structure of a probability space. For this purpose, consider the Borel sigma-algebra (the sigma-algebra formed with open (closed) sets).

[2]In this thesis, the elements of Σ_A^+ will be denoted with underlined, lower case fonts, for example $\underline{x}, \underline{y}, \underline{z}$, etc.

We say that a probability measure ν on Σ_A^+ is a **Gibbs measure with potential** $w : \Sigma_A^+ \to \mathbb{R}$ if there exist constants $C > 1$ and $P > 0$ such that:

$$C^{-1} \leq \frac{\nu[\underline{y} \in \Sigma_A^+ : y_i = x_i, \forall i \in \{0, ..., m\}]}{\exp\left(-Pm + \sum_{k=0}^{m-1} w(\sigma^k \underline{x})\right)} \leq C,$$

for all $\underline{x} \in \Sigma_A^+$ and $m \geq 1$. Before we continue the discussion of Gibbs measures, let us make some comments about the potential w. Given a sequence $\underline{x} \in \Sigma_A^+$, the real number $w(\underline{x})$ tells us how strong is the interaction between the elements of $\underline{x}$. Roughly speaking, it assigns weights to the occurrences of the elements in the sequence. Consider the case of a Markov chain in which the next state depends only on the current state. In this case the potential is of finite range and it can be seen as the transition matrix of the Markov chain. More generally, we can think that each element of the sequence interacts with every other element and the sum of all interactions results in the real number $w(\underline{x})$. It seems reasonable to assume that the interaction between two elements of the sequence becomes weaker when the elements are far away in the sequence. We will see below that the results of the existence of a Gibbs measure, include a condition of this type for the potential.

From the mathematical point of view, many interesting properties of the subshift of finite type, like the existence of a Gibbs measure, depend on the properties of the matrix A. Let us briefly recall some of these properties. The matrix A is **irreducible** if, for each $i, j \in S$, there exists $m(i,j)$ such that $(A^{m(i,j)})_{i,j} > 0$. The matrix A is **irreducible and aperiodic** if there exists $M \geq 1$ such that $A^M > 0$. It is well known that the dynamical system (Σ_A^+, σ) is topologically mixing (meaning for any two measurable sets A and B there exists N such that $\sigma^n A \cap B \neq \varnothing$ for $n \geq N$) if and only if A is irreducible and aperiodic. Under these circumstances, if $w : \Sigma_A^+ \to \mathbb{R}$ is Lipschitz with respect to d_θ, there exists a unique $\sigma-$invariant probability measure on Σ_A^+ that is also a Gibbs measure (see for example, Theorem 1.4 from [7]).

In fact, the existence of Gibbs measure has been proved for more general settings including potentials of summable variation [89] and countably infinite alphabets [85]. Also, it is important to mention that many expanding maps in the interval are conjugate to a subshift of finite type. This means that there exists a continuous bijective function $h : [0, 1] \to \Sigma_A^+$ such that $\sigma \circ h = h \circ T$. When such a function exists, the expanding map T is said to have the **Markov property**.

The notion of Gibbs measures is connected to (and in fact, it arises from) Statistical Mechanics where the **Gibbs distribution** is used to determine the probability that a physical system is in a certain state as a function of the energy and the temperature. A thorough treatment of the Gibbs measures from both, the mathematical and the physical points of view can be found in references [30] and [7]. The theoretical interest for Gibbs measures arises from the fact that they satisfy a variational principle, i.e. they maximize the difference between the Kolmogorov-Sinai entropy and the expected value of the given potential. For a complete treatment of the variational principle from both, the mathematical and physical points of view, see Section 6.12 and 6.13 from reference [84] and Section 6.9 from [24].

A common exercise that is used to illustrate this point (see for example, Lemma 1.1 from [7] or the claim after definition 1.8 from [5] or Lemma 9.9 from [90]) is the following.

Example 1.2.1. *Let $a(1), ..., a(n) \in \mathbb{R}$ and:*

$$F(p_1, ..., p_n) = \sum_{i=1}^{n} p_i \log(p_i) + \sum_{i=1}^{n} p_i a(i). \tag{1.13}$$

The exercise consists on finding the maximum of F on the set $\{p_1 + ... + p_n = 1, p_1, ..., p_n \geq 0\}$. Note that the first sum on the right side of (1.13) is the Shannon entropy of the probability vector $[p_1, ..., p_n]$ and the second sum is the expected value of the given functions $a(1), ..., a(n)$. So, the *maximum of F will be achieved by a probability vector that represents a Gibbs measure. We can find such a maximum using the Lagrange multipliers as follows. Set*

$$L(p_1, ..., p_n, \eta) = -\sum_{i=1}^{n} p_i \log(p_i) + \sum_{i=1}^{n} p_i a(i) - \eta \left[\sum_{i=1}^{n} p_i - 1\right].$$

The partial derivatives have the form:

$$\frac{\partial L}{\partial p_1} = \log p_1 + 1 + a(1) - \eta = 0.$$
$$\ldots$$
$$\frac{\partial L}{\partial p_n} = \log p_n + 1 + a(n) - \eta = 0.$$
$$\frac{\partial L}{\partial \eta} = -\sum_{i=1}^{n} p_i + 1 = 0.$$

From the derivatives with respect to p_i we see that $\log p_i = \eta - 1 - a(i)$ and therefore using the constraint $\sum_{i=1}^{n} p_i = 1$ we obtain $e^{\eta-1}\sum_{i=1}^{n} e^{a(i)} = 1$ and therefore:

$$p_j = \frac{e^{-a(j)}}{\sum_{i=1}^{n} e^{-a(i)}}. \tag{1.14}$$

The above expression is the Gibbs distribution used in Statistical Mechanics for systems with a finite number of states (a famous example of this type of system is the Ising model, see Section 3.2 from [30]). When the number of states is infinite, Equation (1.14) is not valid anymore, but we can still think of the Gibbs measure as a limit of the Gibbs distribution when the number of states increases. See Section 1.A from [7] for a discussion on this process.

Let us now explain how subshifts of finite type are related with the main chapters on this thesis. In Section 1.5 we will see that a Random Dynamical System can be expressed as a transformation on a product space. The first coordinate on that product space is called **base** and in our results, we are interested in Random Dynamical Systems where the base is a subshift of finite type. All the discussion on Chapters 3 and 4 will refer to these type of systems. Essentially, a subshift of finite type in the base of a Random Dynamical System serves as a process to select the transformation that will be used on the next iteration. On Chapter 5 we present a prototype of a dynamical system where the base of the transformation on the product space is not a subshift of finite type anymore.

1.3. Perron-Frobenius operator

In this section we present the definition of the Perron-Frobenius operator and its characterization. Given a measure μ on X and a transformation $T : X \to X$ one can define in a natural way, the **pushforward measure** $T^*\mu$ by $T^*\mu(A) = \mu(T^{-1}A)$. Also, an integrable function $\psi : X \to \mathbb{C}$ induces a complex measure on X defined by $\psi\mu(A) = \int_A \psi d\mu$. Let us now introduce the **Perron-Frobenius operator** associated to the transformation T and to a reference measure μ. The operator acts on $\varphi \in L_1(\mu)$ as the Radon-Nykodim derivative of $T^*(\varphi\mu)$ with respect to μ:

$$P_T\varphi = \frac{dT^*(\varphi\mu)}{d\mu}. \tag{1.15}$$

Usually, one works with the following characterization of the Perron-Frobenius operator: $P_T\varphi$ is the unique element of $L_1(\mu)$ such that for all $\psi \in L_\infty$:

$$\int \psi \cdot P_T\varphi d\mu = \int (\psi \circ T) \cdot \varphi d\mu. \tag{1.16}$$

Note the above equality characterizes P_T in the sense that, if there exists another operator $P_T^{'}$ such that (1.16) holds for every $\psi \in L_\infty$ then $P_T\varphi = P_T^{'}\varphi$ $\mu-$a.e. Indeed, one can choose:

$$\psi = \operatorname{sign}\left(P_T\varphi - P_T^{'}\varphi\right) = \begin{cases} 1 & P_T\varphi \geq P_T^{'}\varphi, \\ -1 & P_T\varphi < P_T^{'}\varphi, \end{cases}$$

and Equation (1.16) leads to:

$$\begin{aligned} \int \left|P_T\varphi - P_T^{'}\varphi\right| d\mu &= \int \psi \cdot \left(P_T\varphi - P_T^{'}\varphi\right) d\mu = \int \psi \cdot P_T\varphi d\mu - \int \psi \cdot P_T^{'}\varphi d\mu \\ &= \int (\psi \circ T) \cdot \varphi d\mu - \int (\psi \circ T) \cdot \varphi d\mu = 0. \end{aligned}$$

A key property of the Perron-Frobenius operator is that its fixed points correspond to invariant densities of acims (with respect to the measure used to define the operator). Indeed, suppose that ν is an absolutely continuous invariant measure with respect to the reference measure μ and its invariant density is φ. Then, combining equations (1.2) and (1.16) we get $\int \psi \cdot P_T\varphi d\mu = \int \psi \cdot \varphi d\mu$ which implies that $P_T\varphi = \varphi$. This means that, as mentioned above, a measure ν absolutely continuous with respect to μ with density φ, is invariant under T if and only if φ is a fixed point of P_T. In other words, invariant densities of acim's correspond to eigenfunctions of the Perron-Frobenius operator with eigenvalue 1. This spectral description is the basis for the spectral method presented on Section 2.1.

P_T is a **positive operator**, meaning that $\varphi \geq 0$ implies $P_T\varphi \geq 0$ for each $\varphi \in L_1(\mu)$. Also, Equality (1.16) implies that $\int P_T\varphi d\mu = \int \varphi d\mu$. By decomposing φ into its positive and negative parts, $\varphi = \varphi^+ - \varphi^- = \max\{0, \varphi\} - \max\{0, -\varphi\}$, one obtains a weakly contracting property of the Perron-Frobenius operator:

$$\|P_T\varphi\|_1 = \int |P_T\varphi| d\mu \leq \int |\varphi| d\mu = \|\varphi\|_1. \tag{1.17}$$

The inequality is not strict, but as it will be seen on the following sections, under some additional mild conditions, one can find a Banach space where the norm of $P_T^n\varphi$ tends to zero as n tends to infinity. This is basically the content of the famous **Perron-Frobenius Theorem** and it is also an important ingredient on the proof of many statistical properties of the system.

P_T allows to characterize the ergodicity of a transformation. If T is ergodic with respect to μ, then P_T has at most one fixed point. Conversely, if P_T has a unique fixed point that is strictly positive, then T is ergodic with respect to μ (see Theorem 4.4.1 from [18]).

We now use Equality (1.16) to get an explicit expression for the Perron-Frobenius operator for a class of maps in the interval.

Example 1.3.1. *Suppose that T is a piecewise monotonic differentiable map in the interval. Moreover, assume φ is continuous and μ is the Lebesgue measure and set:*

$$s(a,b) = \begin{cases} 1 & a \leq b, \\ -1 & a > b, \end{cases}$$

If we fix ψ in Equality (1.16) to be the indicator function of a measurable set $E = [0, x]$ we get:

$$\int_E P_T\varphi(t)dt = \int_{T^{-1}E} \varphi(t)dt.$$

We will denote the integral with respect to the Lebesgue measure by $\int \varphi(t)dt$. The use of t instead of x is only to avoid confusion with the value of x used to define E. Differentiating the above expression on both sides and using the Fundamental Theorem of Calculus we get:

$$P_T\varphi(x) = \frac{d}{dx}\int_{T^{-1}E} \varphi(t)dt = \sum_{i=1}^{k} \frac{d}{dx}\int_{T_i^{-1}E} \varphi(t)dt = \sum_{i=1}^{k} s\left(T_i^{-1}0, T_i^{-1}x\right) \frac{d}{dx}\int_{T_i^{-1}0}^{T_i^{-1}x} \varphi(t)dt,$$

where $T_i^{-1}0$ and $T_i^{-1}x$ are the preimages of 0 and x under T. As T is piecewise monotonic, $s\left(T_i^{-1}0, T_i^{-1}x\right) = -1$ if and only if $\left(T_i^{-1}x\right)' < 0$. Using the Fundamental Theorem of Calculus on each of the derivatives and then the chain rule for the derivative of a composition of functions, we get [3]:

$$P_T\varphi(x) = \sum_{i=1}^{k}\left|\left(T_i^{-1}x\right)'\right| \cdot \left(\frac{d}{dx}\int_{T_i^{-1}0}^{x}\varphi(t)dt\right)\left(T_i^{-1}(x)\right) \tag{1.18}$$

$$= \sum_{i=1}^{k}\left|\left(T_i^{-1}x\right)'\right| \cdot \varphi\left(T_i^{-1}x\right). \tag{1.19}$$

Finally, using the Inverse Function Theorem, the above expression becomes the usual characterization of the Perron-Frobenius operator for piecewise monotonic maps in the interval:

$$P_T\varphi(x) = \sum_{y\in T^{-1}x}\frac{\varphi(y)}{|T'(y)|}. \tag{1.20}$$

For more general systems, one requires the use of **Jacobian** of the transformation T, associated to the probability measure μ. Assume that $T : X \to X$ is a measurable, non-singular transformation with respect to the probability measure μ. Moreover, suppose that there exists a countable partition of X such that $T|_{X_i}$ is injective for all X_i on the partition. We say that $J_{\mu,T} \in L_1(\mu)$ is the Jacobian of T associated to μ if for each measurable set E for which $T|_E$ is injective, we have:

$$\int_E \varphi \cdot J_{\mu,T}d\mu = \int_{T(E)}\varphi d\mu.$$

We use the notion of Jacobian on Chapter 3.2.1 to obtain an expression for the disintegration of a certain family of probability measures. For a subshift of finite type, the Perron-Frobenius operator can be written as follows:

$$P_\sigma\varphi(\underline{x}) = \sum_{\underline{y}\in\sigma^{-1}\underline{x}} e^{w(\underline{y})}\varphi(\underline{y}), \tag{1.21}$$

where $w : \Sigma_A^+ \to \mathbb{R}$ is the considered potential.

1.4. Statistical properties

In this section we present an overview of some of the statistical properties of interest in Dynamical Systems. A complete survey on the topic of statistical properties is [11]. As mentioned in Section 1.1, all these statistical properties are related to the spectral properties of the associated Perron-Frobenius operator whose definition was presented in Section 1.3. For the setting in this section consider that $(X, \mathcal{B}, \mu)$ is a probability space and $T : X \to X$ is a measure preserving transformation. Let us start with a few words on the motivation of the study of statistical properties of dynamical systems.

1.4.1 Why do we study statistical properties?

Suppose we are given a deterministic Dynamical System that is used to model a phenomenon in physical, chemical or biological sciences. One might be tempted to think that given an initial

[3]Note that the derivative on Equation (1.18) is a function to be evaluated on $T_i^{-1}x$. It is not a multiplication of the two expressions. The derivative is of course obtained with the Fundamental Theorem of Calculus and is simply $\varphi(\cdot)$.

condition, one can run the model and predict the state of the real system at any desired time in the future. It turns out that this will not be possible if the system is **chaotic**. In the classical reference [22], the author presents a simplified model of atmospheric convection that exhibits what was later called **sensitive dependence on initial conditions**. Mathematically, when X is a metric space, this means that there exists $\delta > 0$ such that for every $x \in X$ and $\epsilon > 0$, there exist $y \in X$ with $\mathrm{d}(x, y) < \epsilon$ and $n \in \mathbb{N}$ such that $\mathrm{d}(T^n x, T^n y) > \delta$. This definition formalizes the idea that, no matter how close x and y are, it is impossible to ensure that the orbits of x and y will remain close to each other for an arbitrary long time. This makes the long term behavior of the system impossible to predict accurately.

This limitation is not related to deficiencies of the model to capture the studied phenomenon. It is not related either to poor measuring devices or techniques. Instead, it is an intrinsic property of the dynamics itself. So, the study of statistical properties arises as an approach to understand the long term behavior of these systems. In this approach the focus is to prove statements that do not depend on the initial condition, but instead, are uniform on a large subset of X. The existence of the limit in Equation (1.3) is an example and in this section we present further statistical properties that are frequently studied.

1.4.2 Decay of correlations

Recall that the transformation T is **mixing** if for any two measurable sets A and B, one has that $\lim_{k\to\infty} \mu(A \cap T^{-k}B) = \mu(A)\mu(B)$. This basically means that the events A and $T^{-k}B$ become nearly independent when n tends to infinity. In that sense, mixing is a property that resembles independence of the system at the long-term behavior. A sufficient condition for mixing can be stated as follows. Let $f, g : X \to \mathbb{R}$ two square integrable observables with respect to μ. The **correlation function** $C_{f,g} : \mathbb{Z} \to \mathbb{R}$ is defined by:

$$C_{f,g}(k) = \left| \int (f \circ T^k) g d\mu - \int f d\mu \int g d\mu \right|. \tag{1.22}$$

The transformation T is mixing if for each pair of square integrable functions $f, g : X \to \mathbb{R}$ we have that $\lim_{k\to\infty} C_{f,g}(k) = 0$.[4] When this happens, the system is said to have the property of **decay of correlations**.

In probabilistic terms [5], $C_{f,g}(k)$ is essentially the covariance of the random variables g and $f \circ T^k$. Intuitively, the decay of correlations means that statistics of the long-term behavior of the system does not depend on the initial condition and the rate at which the decay of correlations happens, is a measure of the speed at which the system becomes independent of the initial condition. Such a speed is an indicator of chaotic behavior: the faster the decay of correlations, the more chaotic is the system under consideration. Figure 1.4 shows an estimator of the correlation function for the map $T(x) = 3x(\mod 1)$ and the Manneville-Pomeau map (Equation (1.8)) with $f = \chi_A$ and $g = \chi_B$ where $A = [0, 1/3]$ and $B = [1/2, 1]$. It can be seen from the figure that the rate at which $\widetilde{C}_{f,g}$ tends to zero, is different for each map. The decay of correlations is exponential for the $T(x) = 3x(\mod 1)$ and polynomial for the Manneville-Pomeau map ([64]).

The estimator that we used is given by:

$$\widetilde{C}_{f,g}(k) = \frac{1}{n} \sum_{i=0}^{n-1} f\left(T^i x\right) \cdot g\left(T^{i+k} x\right) - \frac{1}{n^2} \left(\sum_{i=0}^{n-1} f\left(T^i x\right) \right) \left(\sum_{i=0}^{n-1} g\left(T^i x\right) \right), \tag{1.23}$$

where $n + k$ is the size of the sample. As a consequence of Birkhoff ergodic theorem, we can be sure that $\widetilde{C}_{f,f}(k)$ tends to $C_{f,f}(k)$ when n tends to infinity.

[4] In order to prove this assertion, it is enough to set $f = \chi(A)$ and $g = \chi(B)$ so that $0 = \lim_{k\to\infty} C_{f,g}(k) = \lim_{k\to\infty} \mu(A \cap T^{-k}B) - \mu(A)\mu(B)$ implying that T is mixing.

[5] Recall that the covariance of random variables $\mathcal{X}$ and $\mathcal{Y}$ is $\mathrm{Cov}(\mathcal{X}, \mathcal{Y}) = \mathbb{E}(\mathcal{X}\mathcal{Y}) - \mathbb{E}(\mathcal{X})\mathbb{E}(\mathcal{Y})$

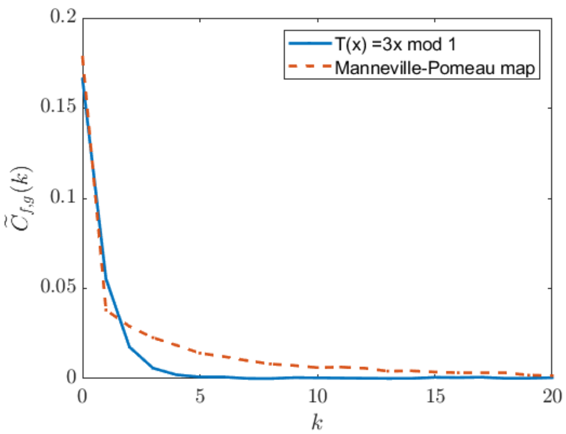


Figure 1.4: Decay of correlations for the map $T(x) = 3x(\mod 1)$ and the Manneville-Pomeau map given by equation (1.8). We used the estimator of the correlation function (1.23) with $f = \chi_A$ and $g = \chi_B$ where $A = [0, 1/3]$ and $B = [1/2, 1]$. It can be seen from the figure that the decay of correlations is slower in the Manneville-Pomeau map.

Decay of correlations has been established for many classes of Dynamical Systems. Two of the most complete compendium on the topic are references [62] and [5]. Let us now mention some of the results that are compiled on this second reference for subshifts of finite type. Locally constant potentials are treated on Proposition 1.1, Lipschitz or equivalently, Hölder potentials are treated on Theorem 1.6 (based on reference [7]). Decay of correlations is exponentially fast in this case. Potentials of summable variation are treated on Theorem 1.11 (based on reference [79]). Decay of correlations in this case is not exponentially fast; instead, it depends on the decay of the variation (see also Chapter 3 of thesis [70] for details on the calculations).

1.4.3 Central Limit Theorem

By elementary Probability Theory, we know that a sequence of IID random variables $\mathcal{X}_n$ with $\mathbb{E}[\mathcal{X}_n] < \infty$ and $0 < \operatorname{Var} \mathcal{X}_n < \infty$, satisfies the Central Limit Theorem. Specifically, if the expected value of $\mathcal{X}_n$ is $\mathbb{E}[\mathcal{X}_n]$ and its variance is $\operatorname{Var}(\mathcal{X}_n)$, then the random variable given by $\frac{S_n(\mathcal{X}_n)-\mathbb{E}[\mathcal{X}_n]}{\sqrt{n}}$ converges in distribution to a normal distribution with mean $\mathbb{E}[\mathcal{X}_n]$ and variance $\operatorname{Var}(\mathcal{X}_n)$.

In the context of Dynamical Systems, one might ask whether the sequence given by $\frac{S_n f - n\int f d\mu}{\sqrt{n}}$ satisfies a similar statement. Recall from the discussion in Section 1.1 that if the Dynamical System is chaotic enough, then such a sequence behaves as a random variable. So, a Central Limit Theorem in this context can be stated as follows. We say that a transformation $T : X \to X$ with observable f satisfies a Central Limit Theorem with respect to the measure μ if there exists $\rho \geq 0$ such that:

$$\lim_{n\to\infty} \mu\left\{x : \frac{S_n f(x) - n\int f d\mu}{\sqrt{n}} \leq c\right\} = \frac{1}{\rho\sqrt{2\pi}} \int_{-\infty}^{c} \exp\left(-\frac{y^2}{2\rho^2}\right) dy, \tag{1.24}$$

for all $c \in \mathbb{R}$. In words, what a Central Limit Theorem means is that, with high probability, the fluctuations of $S_n f$ around its expected value $n\int f d\mu$, are of order $1/\sqrt{n}$. In analogy with an IID process, ρ^2 is the limiting variance [6] of the process and can be defined as:

$$\rho^2 = \lim_{n\to\infty} \frac{1}{n} \int \left(S_n f - n\int f d\mu\right)^2 d\mu.$$

[6]Recall that the variance of a random variable $\mathcal{X}$ is $\operatorname{Var}(\mathcal{X}) = \mathbb{E}\left(\mathcal{X} - \mathbb{E}(\mathcal{X})\right)^2$

Example 1.4.1. *Consider for example, the case where* $T(x) = 2x(\mod 1)$, $f(x) = z - \frac{1}{2}$ *and* μ *is the Lebesgue measure. Figure 1.5 shows a normalized histogram of* 10000 *realizations of* $\frac{S_n f}{\sqrt{n}}$. *As it can be seen from the figure, the limit distribution can be fitted to a normal distribution with mean* 0 *and variance* $\frac{1}{4}$.

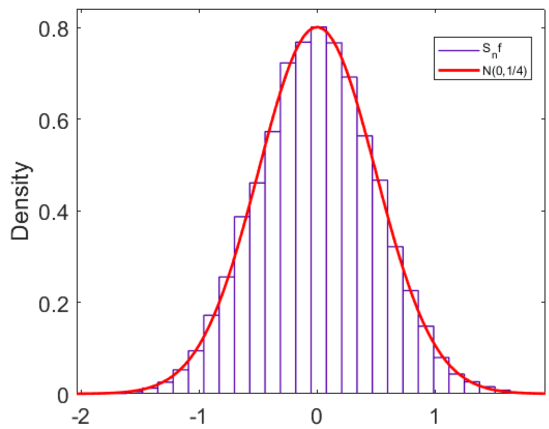


Figure 1.5: Histogram of 10000 realizations of $\frac{S_n f}{\sqrt{n}}$ for transformation $T(x) = 2x \mod 1$ and $f(x) = x - \frac{1}{2}$. The histogram can be fitted to a normal distribution.

Let us now briefly review some results in the literature. One of the first results on the CLT for discrete Dynamical Systems is presented in [82]. In this reference the author proves a CLT for a full-shift with respect to the Gibbs measure with exponential decay of correlations (see Lemma 1.1 in that reference). A similar result with a more general Lebesgue space is presented in [10]. Its drawback is that it only asserts the existence of an observable for which the CLT holds. A more general CLT that uses the spectral decomposition of the Perron-Frobenius operator is proved in [43]. We will actually illustrate a particular case of this result in Section 2.1.2 where we discuss the spectral method for deterministic Dynamical Systems. In a general setting that includes subshifts of finite type with Lipschitz potential, the CLT is proved in reference [39]. For potentials decreasing polinomially, a CLT is proved on reference [79] as a consequence of polynomial decay of correlations.

Our theoretical result in this thesis, presented in Chapter 4, is a Central Limit Theorem for the Random Dynamical Systems studied in Chapter 3. Our result includes an inequality that quantifies the speed of convergence in (1.24). These types of inequalities are known as Berry-Esseen inequalities (see for example [33]).

1.4.4 Berry-Esseen inequalities

Let $\mathcal{X}_n$ be a sequence of IID random variables that satisfy the following moment conditions:

$$\mathbb{E}[\mathcal{X}_i] = 0, \qquad \mathbb{E}[\mathcal{X}_i^2] = \rho^2 > 0, \qquad \mathbb{E}[\mathcal{X}_i^3] < \infty.$$

An inequality for the probability of deviation of the distribution of $\frac{S_n \mathcal{X}_n}{\sqrt{n}}$ with respect to the normal distribution $N(0,1)$ is given in reference [6]. Specifically, the author shows that there exists $D \geq 0$ such that for each $n \geq 0$:

$$\sup_{c \in \mathbb{R}} \left| \mathbb{P}\left\{ \frac{S_n \mathcal{X}_n}{\sqrt{n}} \leq c \right\} - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{c} \exp\left(-\frac{x^2}{2} \right) dx \right| \leq \frac{D}{\sqrt{n}}. \tag{1.25}$$

Inequalities of this type are known as **Berry-Esseen inequalities** and they essentially specify the speed of convergence in the Central Limit Theorem.

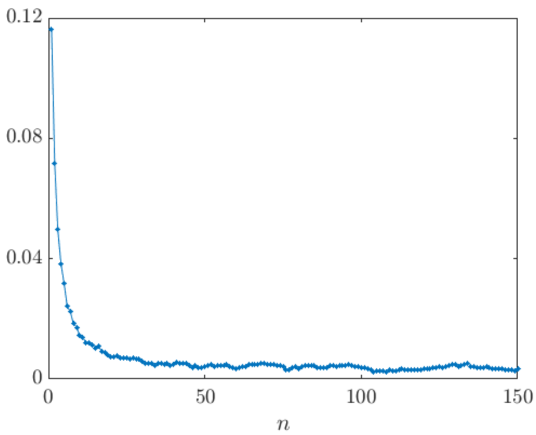


Figure 1.6: Illustration of Berry-Esseen inequality for the map $T(x) = 2x \mod 1$. The absolute value from Inequality (1.25) tends to zero like $\frac{1}{n}$. We used $f(x) = x - \frac{1}{2}$ and $c = -0.4$.

In fact, the author of [6] proves this result for independent variables, not necessarily identically distributed. A similar result for Markov chains was proved in [55] using the spectral gap method. Moreover, the author of reference [69] extends this result to general random processes with weaker dependence conditions that include mixing dynamical systems such as the map $T(x) = 2x \mod 1$ (see Example 3.2 from that reference). For non-uniformly hyperbolic dynamical systems with subexponential decay of correlations, the author of reference [32] proves a similar Berry-Esseen inequality. Figure 1.6 illustrates the Berry-Esseen inequality for the map $T(x) = 2x \mod 1$. The figure shows the absolute value from inequality (1.25) as a function of the number of iterations n.

1.4.5 Large Deviation Principle

We say that a random variable $\mathcal{X}_n$ satisfies a **Large Deviation Principle** with rate function $I : \mathbb{R} \to [0, \infty]$, if the following limit exists:

$$\lim_{\epsilon\to 0}\lim_{n\to\infty} -\frac{1}{n}\ln\left(\mathbb{P}\left\{x \in X : \mathcal{X}_n(x) \in [a-\epsilon, a+\epsilon]\right\}\right) = I(a). \tag{1.26}$$

If Equation (1.26) is satisfied, and $A = [a-\epsilon, a+\epsilon]$ for a small ϵ, we can approximate:

$$\mathbb{P}\left\{x \in X : \mathcal{X}_n(x) \in A\right\} \approx e^{-nI(a)}. \tag{1.27}$$

When the random variables $\mathcal{X}_n$ are independent and identically distributed and the **logarithmic moment generating function** given by:

$$M(t) = \ln \mathbb{E}\left[e^{t\mathcal{X}_i}\right], \tag{1.28}$$

is finite, Cramér Theorem (see for example Theorem 23.3 from [54]) guarantees that a Large Deviation Principle is satisfied.

In a more general setting, when the random variables $\mathcal{X}_n$ are not necessarily IID, a powerful result to obtain a Large Deviation Principle is the Gärtner-Ellis Theorem which, broadly speaking, can be stated as follows. Let:

$$\Lambda(k) = \lim_{n\to\infty}\frac{1}{n}\log \mathbb{E}\left[e^{nk\mathcal{X}_n}\right], \tag{1.29}$$

when the limit exists. Then, if $\Lambda(k)$ is differentiable, $\mathcal{X}_n$ satisfies a Large Deviation Principle with rate function given by:

$$I(a) = \sup_{k\in\mathbb{R}}\left\{ka - \Lambda(k)\right\}.$$

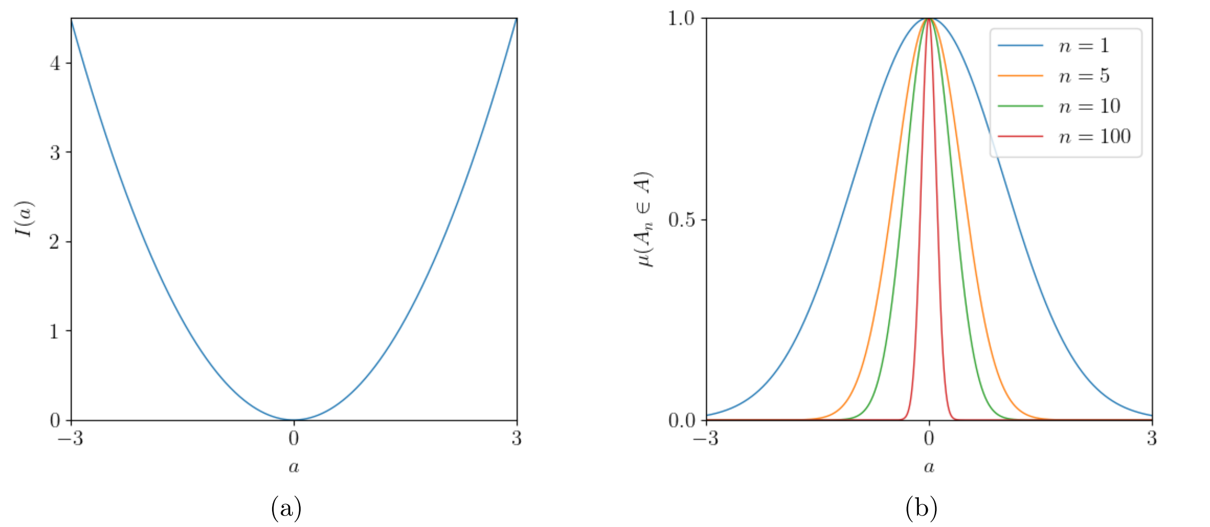


Figure 1.7: Illustration of Large Deviation Principle for the empiric mean of IID random variables. (a) Rate function obtained with the Gärtner-Ellis theorem (b) Probability that $A_n \in A = [a - \epsilon, a + \epsilon]$. For fixed a, the probability decreases as n increases.

Example 1.4.2. *As a first example, consider the empiric mean of IID random variables $\mathcal{X}_n$ taking values in $\{0, 1\}$ and following a Bernoulli distribution with $p = q = \frac{1}{2}$. Using the fact that $\mathcal{X}_n$ is IID, one can show that:*

$$\Lambda(k) = -\log 2 + \log(1 + e^k).$$

The fact that $\Lambda(k)$ is differentiable, allows us to apply the Gärtner-Ellis Theorem and conclude that the empiric mean of $\mathcal{X}_n$ satisfies a Large Deviation Principle with rate function given by:

$$I(a) = \log 2 - a \log a - (1 - a) \log(1 - a).$$

Example 1.4.3. *Suppose now that the random variables are still IID but distributed normally with expected value $\mathbb{E}[\mathcal{X}_n]$ and variance $\mathrm{Var}(\mathcal{X}_n)$. It is easy to show that:*

$$\Lambda(k) = \mathbb{E}[\mathcal{X}_n] \cdot k + \frac{1}{2} \mathrm{Var}(\mathcal{X}_n) \cdot k^2.$$

As $\lambda(k)$ is differentiable, Gärtner-Ellis Theorem allows to conclude that the empiric mean $A_n = \frac{1}{n} S_n(\mathcal{X}_s)$ satisfies a Large Deviation Principle with rate function given by:

$$I(a) = \frac{(a - \mathbb{E}[\mathcal{X}_n])^2}{2 \mathrm{Var}(\mathcal{X}_n)}.$$

Figure 1.7a shows the form of the rate function and Figure 1.7b shows the approximation given by (1.27) for different values of n.

Example 1.4.4. *As commented above, the Gärtner-Ellis Theorem works when the process is not IID. Consider a Markov chain formed with the random variables $\mathcal{X}_1, ..., \mathcal{X}_n$ taking values on a finite set $S = \{1, ..., m\}$. Let A be the associated stochastic matrix and suppose that A is irreducible. We want to use Gärtner-Ellis Theorem to obtain a Large Deviation Principle for the random variable:*

$$A_n f = \frac{1}{n} \sum_{i=1}^{n} f(\mathcal{X}_i), \tag{1.30}$$

where $f : S \to \mathbb{R}$ is an observable. We will treat this example with full detail now.

Let x_0 be a given initial condition. The probability that the Markov chain $\mathcal{X}_1, ..., \mathcal{X}_n$ takes the values $x_1, ..., x_n$ is given by:

$$\mathbb{P}(\mathcal{X}_1 = x_1, ..., \mathcal{X}_n = x_n | \mathcal{X}_0 = x_0) = a_{x_0,x_1} \prod_{i=1}^{n-1} a_{x_i,x_{i+1}}.$$

Let us now obtain the expected value $\mathbb{E}_{x_0}\left[e^{nkA_n}\right]$. Note that:

$$\begin{aligned}
\mathbb{E}_{x_0}\left[e^{nkA_n}\right] = \mathbb{E}_{x_0}\left[e^{k\sum_{i=1}^{n} f(\mathcal{X}_i)}\right] = \mathbb{E}_{x_0}\left[e^{kf(\mathcal{X}_1)} \cdots e^{kf(\mathcal{X}_n)}\right] \\
&= \sum_{x_1,\dots,x_n \in \Sigma} e^{kf(x_1)} \cdots e^{kf(x_n)} a_{x_0,x_1} \prod_{i=1}^{n-1} a_{x_i,x_{i+1}} \\
&= \sum_{x_1,\dots,x_n \in \Sigma} a_{x_0,x_1} e^{kf(x_1)} \cdots a_{x_{n-1},x_n} e^{kf(x_n)} \qquad (1.31) \\
&= \sum_{x_1,\dots,x_n \in \Sigma} \left(\prod_{i=0}^{n-1} a_{x_i,x_{i+1}} e^{kf(x_{i+1})} \right). \qquad (1.32)
\end{aligned}$$

Let $\mathbf{\Pi}_k$ be defined as $(\boldsymbol{A}_k)_{i,j} = a_{i,j} e^{kf(x_j)}$ so we can write the expected value as:

$$\mathbb{E}_{x_0}\left[e^{nkS_n}\right] = \sum_{x_n=1}^{m} (\boldsymbol{A}_k^n)_{x_0,x_n}.$$

Let us now calculate $\Lambda(k)$:

$$\begin{aligned}
\Lambda(k) &= \lim_{n\to\infty} \frac{1}{n} \ln \mathbb{E}_{x_0}\left[e^{nkS_n}\right] \\
&= \lim_{n\to\infty} \frac{1}{n} \ln \left(\sum_{x_n=1}^{m} (\boldsymbol{A}_k^n)_{x_0,x_n} \right).
\end{aligned}$$

$(\boldsymbol{A}_k)_{i,j} = \pi_{i,j} e^{kf(x_j)} \geq 0$, so $\boldsymbol{A}_k$ is an irreducible matrix. Finally, the Perron-Frobenius Theorem for irreducible matrices (see for example Theorem 3.1.1 from [15]) guarantees that:

$$\Lambda(k) = \lim_{n\to\infty} \frac{1}{n} \ln \left(\sum_{x_n=1}^{m} (\boldsymbol{A}_k^n)_{x_0,x_n} \right) = \ln \lambda(\boldsymbol{A}_k).$$

Where $\lambda(\boldsymbol{A}_k)$ is the maximal eigenvalue of $\boldsymbol{A}_k$. According to the Perron-Frobenius Theorem, this expression is valid for all $x_0 \in S$. So, $\Lambda(k)$ is differentiable and we can conclude that $A_n f$ satisfies a Large Deviation Principles with rate function:

$$I(s) = \sup_{k\in\mathbb{R}} \{ks - \ln \lambda(\boldsymbol{A}_k)\}.$$

1.4.6 Concentration inequalities

Consider a sequence of IID random variables $\mathcal{X}_n$. We would like to know how likely is that, for a given $n \geq 0$, $S_n\mathcal{X}_n$ exceeds a prescribed value t. A bound of this type can be obtained using Markov inequality. The result is a bound known as **Chernoff bound** that is valid for all $n \geq 0$ and $t \in \mathbb{R}$:

$$\mathbb{P}\left(\frac{S_n\mathcal{X}_n}{n} \geq t\right) \leq \inf_{a>0} \left\{ \left(\mathbb{E}[e^{a\mathcal{X}_n}]\right)^n \cdot e^{-nta} \right\}.$$

If the distribution of the random variables is known, the expected value of $e^{a\mathcal{X}_n}$ can be obtained explicitly and the above inequality can be optimized by finding the infimum. For example, suppose that the random variables $\mathcal{X}_n$ follow a Bernoulli distribution. Then for any $0 < t < 1$ we have:

$$\mathbb{P}\left(S_n\mathcal{X}_n \geq (1+t)\mathbb{E}[\mathcal{X}_n]\right) \leq \exp\left(-\frac{t^2\mathbb{E}[\mathcal{X}_n]}{3}\right).$$

Similar inequalities can be obtained for independent, not necessarily identically distributed, random variables. Suppose that $0 \leq \mathcal{X}_i \leq 1$. For $0 < t < 1 - \mathbb{E}[\mathcal{X}_n]$ the following bound, called **Hoeffding's inequality**, was first obtained in [42] (see Theorem 1 from that reference):

$$\mathbb{P}\left(\frac{S_n\mathcal{X}_n}{n} - \mathbb{E}\left[\mathcal{X}_n\right] \geq t\right) \leq e^{-2nt^2}. \tag{1.33}$$

This type of inequalities have two important advantages: (1) they are much more sharper than the corresponding Markov or Chebyshev's inequalities and (2) they are valid for all $n \geq 0$, not only in the limit.

In the context of Dynamical Systems, inequalities of the form of (1.33) are called **concentration inequalities**. They are non-asymptotic inequalities for the probabilities of deviation of a general observable $K : X^n \to \mathbb{R}$ from its expected value and they are important because they establish fluctuation bounds for the probability of deviation not only in the limit, but at each $n \geq 0$. Now, we impose a regularity condition on K. We say that $K : X^n \to \mathbb{R}$ is **Lipschitz separately** if for each $i = 0, ..., n-1$, there exists a constant $\mathrm{Lip}_i(K)$ such that:

$$\left|K(x_0, ..., x_i, ..., x_{n-1}) - K(x_0, ..., x_i', ..., x_{n-1})\right| \leq \mathrm{Lip}_i(K)\, \mathrm{d}(x_i, x_i'). \tag{1.34}$$

for all $x_0, ..., x_i, ..., x_{n-1}, x_i' \in X$. We say that T satisfies an **exponential concentration inequality** if there exists $C > 0$ such that, for any separately Lipschitz function K:

$$\int e^{K(x,..,T^{n-1}x) - \int K(y,..,T^{n-1}y)d\mu(y)} d\mu(x) \leq e^{C\sum_{i=0}^{n-1}\mathrm{Lip}_i(K)^2}. \tag{1.35}$$

Note that constant C on (1.35) must not depend on K nor n, it only depends on the dynamic. A standard consequence of the above inequality, is the following estimation for the probability of deviation:

$$\mu\left\{x \in X : K(x, .., T^{n-1}x) - \int K(y, .., T^{n-1}y)d\mu(y) \geq t\right\} \leq e^{-\frac{t^2}{4C\sum_{i=0}^{n-1}\mathrm{Lip}_i(K)^2}}. \tag{1.36}$$

When we set $K(x, .., T^{n-1}x) = S_nf(x)$, inequality (1.36) becomes a bound for the speed of convergence on Birkhoff ergodic theorem (see Chapter 2.4 of [70] for a details on the calculation). Similar bounds can be obtained for the empirical measure or for the estimator of the correlation coefficients.

Let us now give some references on concentration inequalities. A first major reference on the topic is [60]. In that reference, the author compiles many concentration inequalities for different scenarios such as measures on product spaces (Corollary 1.17) and Markov chains (Theorem 3.3). More specific to Dynamical Systems, in reference [73], the authors establish an exponential concentration inequality for piecewise expanding maps in the interval. For subshifts of finite type, in reference [80] a concentration inequality is established for potentials whose variation decreases polinomially. For countably infinite alphabets, a concentration inequality for Hölder continuous potentials is obtained in reference [66]. Finally, for non-uniformly hyperbolic dynamical systems modeled by a Young tower with exponential tails, an exponential concentration inequality was proved in [12].

In reference [66], we prove an exponential concentration inequality for countable subshifts and obtain bounds for the rate of convergence of the Birkhoff ergodic theorem (whose details we present in the next section). The existence of a Gibbs measure for Hölder continuous potentials and the spectral gap of its Perron frobenius operator are proved in [85] and [13] respectively.

1.4.7 Rates of convergence in Birkhoff ergodic theorem

Recall from Section 1.1 that the Birkhoff ergodic theorem establishes sufficient conditions for the existence of the limit $\lim_{n\to\infty} \frac{S_n f}{n}$. From the practical point of view, one might be interested in knowing bounds for the rate of convergence that are valid, not only in the limit when n tends to infinity, but at each $n \geq 0$. Important references where rates of convergence in the Birkhoff ergodic theorem are obtained under different settings, are [51], [50] and [49]. In this section we will talk about the rate of convergence in the specific case of a countable subshift with general observables f that are not necessarily Hölder continuous. This is one of the results we proved in reference [66].

We consider a fairly general class of observables that are bounded and continuous almost everywhere. We denote this class of observables as $\mathcal{BC}$. The set $\mathcal{BC}$ is a much larger class of observables than the Lipschitz or Hölder continuous observables that are usually considered. The use of this type of observables is possible with the application of a general result from [51] that established a rate of convergence of the limit (1.3) for observables that belong to $\mathcal{BC}$. They show that, for each $f \in \mathcal{BC}$ there exist Hölder continuous functions g_1^δ and g_2^δ and there exists also a function $l(\delta)$ with the property that $l(\delta) \to 0$ when $\delta \to 0$, that satisfy the following:

$$g_1^\delta \leq f \leq g_2^\delta.$$

In other words, every $f \in \mathcal{BC}$ is bounded below and above by Hölder continuous functions that can be obtained explicitly in terms of f. The above inequality essentially implies the following (see [51]): if one has a rate of convergence for the Birkhoff ergodic theorem, for Hölder continuous observables, one can use it with g_1^δ and g_2^δ to obtain the corresponding rate for $f \in \mathcal{BC}$.

Returning back to the countably subshift, the concentration inequality proved in reference [66] is then used to obtain a rate of convergence in the Birkhoff ergodic theorem for Hölder continuous observables, which in turn is used to obtain the corresponding rate for observables in $\mathcal{BC}$. It's worth noticing that the distance d under which the functions g_1^δ and g_2^δ are Hölder continuous satisfies the following property: an α-Hölder continuous function with respect to the distance d is Lipschitz continuous with respect to another distance denoted d_α. Let us state the rate of convergence more specifically. Let m be the Gibbs measure associated to a Hölder continuous potential in a countable subshift. Let f be a bounded and continuous m-everywhere observable. Let g_1^δ and g_2^δ be the corresponding Hölder continuous functions that bound f. Then, for each $t > 0$:

$$m\left\{\left|\frac{S_n f}{n} - \int f dm\right| > t + l(\delta)\right\} \leq 4\exp\left\{\frac{nt^2}{4D|g^*|_{d_\alpha}^2}\right\}.$$

In the above inequality, $|g^*|_{d_\alpha} = \max\left\{\left|g_1^\delta\right|_{d_\alpha}, \left|g_2^\delta\right|_{d_\alpha}\right\}$ and $\left|g^\delta\right|_{d_\alpha}$ is the Lipschitz constant with respect to the metric d_α.

If one is interested in a bound for the difference $\left|\frac{s_n f}{n} - \int f dm\right|$ itself, it can also be obtained for $f \in \mathcal{BC}$ as long as the decay of correlations is of the form $\mathcal{O}(n^{-\tau})$ with $\tau > 0$. See Corollary 4.3.1 from [70] for a detailed statement and proof.

1.5. Overview of Random Dynamical Systems

The main theoretical result in this thesis is a Central Limit Theorem for a class of contractive Random Dynamical Systems. We will talk about the specific setting on Chapter 3, but before doing that, on this last section of the first chapter we will give a brief general overview of Random Dynamical Systems (RDS for short) and the associated skew-product transformation. We will see how the notions of invariant measures and the Perron-Frobenius operator can be defined for a Random Dynamical System and we compile some of the existing results that make use of these notions.

When the dynamics of the system is completely determined by its initial conditions, the system is called a **Deterministic Dynamical System**. If the dynamics of the system is affected by one or more random parameters, the system is called a **Random Dynamical System** (RDS for short). The randomness can be seen as a perturbation of the original system and is modelled by an additional abstract dynamical system referred to as **base**. In this thesis we will be interested on the case where the original system forms a compact metric space and the base is a subshift of finite type.

Let $S = \{1, ..., N\}$ and consider N transformations from the unit interval $I = [0, 1]$ to itself $T_s : I \to I$. Suppose that we are given a probability vector $p = [p_1, ..., p_N]$ and an initial condition $x_0 \in I$. The dynamics consists on the following: at each step, we randomly select a transformation according to the probability vector and evaluate the selected transformation on the value of the previous iteration. The random selection of the transformations is represented by a sequence of $S-$valued random variables $\{\xi_n\}_{n=1}^{\infty}$ and the time evolution of the dynamical system, projected on the unit interval, is given by:

$$x_{n+1} = T_{\xi_n}(x_n).$$

The setting described above is a classical prototype of an RDS on the unit interval I that has been extensively studied (see for example [2], [75], [26], [47], [58]; variations of this setting are also studied in [68] and [34]). It is called **product of random mappings** and under this setting, the notion of invariant measure can be extended naturally. We say that a measure μ on I is **stationary** if for every measurable set $E \subseteq I$ we have:

$$\mu(E) = \sum_{s=1}^{N} p_s \cdot \mu\left(T_s^{-1}E\right). \tag{1.37}$$

Similarly, the associated Perron-Frobenius operator can be seen as an averaged version of the individual operators associated to each T_s:

$$P_T\varphi = \sum_{s=1}^{N} p_s P_{T_s}\varphi.$$

In the case that each T_s is a piecewise monotone C^2 map, sufficient conditions for the existence of a stationary probability measure, absolutely continuous with respect to the Lebesgue measure were given in [75]. Here, the main condition is the strict inequality:

$$\sum_{s=1}^{N} \frac{p_s}{|T_s'(x)|} < 1,$$

for all $x \in I$. Of course, this inequality reduces to (1.7) when S is the singleton $S = \{1\}$. There are important results on the statistical properties of RDS's, many of them obtained by properly adapting the spectral method from the deterministic to the random scenario. We will discuss the spectral method for deterministic Dynamical systems on Section 2.1.2 and for Random Dynamical Systems on Section 2.2.1. In the meantime, the rest of this section is devoted to present and discuss a general definition of an RDS stated on [3].

In general, an RDS is not restricted to S being a finite set. Moreover, the process generating the sequence ξ_n is not restricted to be an IID process. For example, it can be a Markov process or another dynamical system. This general setting is considered on the following definition.

Definition 1.5.1. *Let $(\Omega, \mathcal{B}(\Omega), \mathbb{P})$ be a probability space and let $\sigma : \Omega \to \Omega$ be a measure preserving transformation. A random dynamical system on a measurable space $(X, \mathcal{B}(X))$ over $(\Omega, \mathcal{B}(\Omega), \mathbb{P}, \sigma)$ is a map $T : \Omega \times X \to X$ with the following properties:*

1. *Measurability: The map $T : \Omega \times X \to X$ is measurable with respect to the σ-algebras $\mathcal{B}(\Omega) \otimes \mathcal{B}(X)$ and $\mathcal{B}(X)$.*

2. *Cocycle: Let $T_\omega = T(\omega, \cdot) : X \to X$. For all $n, m \in \mathbb{N}$, $\omega \in \Omega$:*

$$T_\omega^{n+m} = T_{\sigma^m \omega}^n \circ T_\omega^m.$$

We recover the setting of product of random mapping by fixing $X = [0, 1]$, $\Omega = S^{\mathbb{N}}$, σ is given by $(\sigma(\omega))_i = \omega_{i+1}$, $\mathbb{P}$ the Markov measure, $\mathcal{B}(X)$ and $\mathcal{B}(\Omega)$ the corresponding Borel σ-algebras and T given by:

$$\begin{aligned} T : \Omega \times X &\to X \\ (\omega, x) &\mapsto T_{(\sigma(\omega))_0}(x). \end{aligned}$$

Note that in this case, Ω is a particular type of subshift of finite type with $A_{i,j} = 1$ for $i, j \in S$ but in general the matrix A can have zero entries as well. Also, the map $\xi : \Omega \to S$ can be seen as a map that retrieves the symbol at position zero of the sequence ω. More generally, (Ω, σ) is not restricted to be a subshift of finite type. For example, the case $\Omega = X = [0, 1]$, $\sigma(\omega) = 2\omega$ mod 1 with a given map $T : [0, 1]^2 \to [0, 1]$ satisfies Definition 1.5.1 with $\xi(\omega) = \omega$.

An RDS can be seen as a transformation on a product space. Indeed, if F is an RDS, then the mapping $F : \Omega \times X \to \Omega \times X$ given by:

$$F(\omega, z) = (\sigma\omega, T(\omega, z)). \tag{1.38}$$

is a measurable transformation on $(\Omega \times X, \mathcal{B}(\Omega) \otimes \mathcal{B}(X))$. Conversely, such a measurable transformation, always defines a cocycle φ and hence, an RDS. The mapping F is called **skew-product transformation**. In Chapter 4 we will study the statistical properties of a skew-product where Ω is a subshift of finite type and T satisfies a contraction property to be specified.

In Chapter 5 we discuss **Connected Dynamical Systems** that can be seen as an extension of Definition 1.5.1 in the sense that these types of systems allow more interdependence between (X, T_ω) and (Ω, σ). This type of dynamical systems can be seen as a skew-product but in certain cases they can also be seen as metastable systems.

1.6. Goals of the research

Now that we have presented an introduction of the topics that we will be discussing in the rest of the thesis, let us state the general and particular goals of this research.

General Goal: Use quasicompactness of the Perron-Frobenius operator and the spectral method to prove the existence of statistical properties of random dynamical systems in which the base is a subshift of finite type.

Particular goals:

- State a Central Limit Theorem for random dynamical systems.
- Once a Central Limit Theorem is established, find an inequality for the speed of convergence.
- Find and study critical behaviors in connected dynamical systems.
- Perform a numerical study of the statistical properties of connected dynamical systems, particularly in places near to the points where critical behaviors occur.

2. Spectral Method and Limit Theorems

2.1. A brief summary of the Spectral Method

The spectral method has been extensively used to prove Central Limit Theorems in various settings of deterministic dynamical systems. Roughly speaking, it consists on finding a Banach space where the transfer operator associated to the dynamics and to the measure of interest, has the property of quasicompactnes. Once quasicompactness is established, one defines a perturbed operator and shows that the characteristic function of the random variable of interest can be expressed in terms of the iterations of the perturbed operator (we will refer to this step as **ST1**). Then one uses perturbation theorems to show that the perturbed operator inherits the spectral properties of the original transfer operator (**ST2**). Then one applies the Lévy continuity theorem to conclude that the random variable converges in distribution to a random variable distributed normally (**ST3**). And finally, one relates the first and second derivatives of the leading eigenvalue of the perturbed operator, to the average and variance of the limit distribution of the random variable of interest (**ST4**). A classic reference on this method is [39] and a more recent survey is [31]. A very complete introduction of the method, in the form of lecture notes, can be found in [27]. In this Section we will present the fundamentals of this method in detail.

2.1.1 Spectral Theory and Quasicompactness

The following discussion is based on references [52] and [39]. Let $\mathcal{B}$ be a normed vector space with norm denoted $\|\cdot\|$. Let $Q : \mathcal{B} \to \mathcal{B}$ be a linear operator acting on $\mathcal{B}$. The operator Q is said to be **bounded** if there exists $C > 0$ such that for all $u \in \mathcal{B}$, one has $\|Qu\| \leq C\|u\|$. The smallest of these constants C, is called **operator norm** and is denoted $\|Q\|$. The use of this name and notation is justified because the set of all bounded linear operators on $\mathcal{B}$, which we will denote $\mathcal{L}_{\mathcal{B}}$, forms a normed vector space with the operator norm. Morevoer, if $\mathcal{B}$ is complete, so is $\mathcal{L}_{\mathcal{B}}$.

A bounded linear operator is continuous with respect to the distance induced by operator norm. Indeed, given $\epsilon > 0$ choose $\delta < \dfrac{\epsilon}{\|Q\|}$. Then, if $u, v \in \mathcal{B}$ and $\|u - v\| < \delta$, then:

$$\|Qu - Qv\| = \|Q(u - v)\| \leq \|Q\|\|u - v\| < \delta\|Q\| < \frac{\epsilon}{\|Q\|}\|Q\| = \epsilon.$$

Also, a continuous operator is **closed**. Suppose u_n converges to $u \in \mathcal{B}$ and Qu_n converges to $v \in \mathcal{B}$. Then by continuity:

$$Qu = Q\left(\lim_{n\to\infty} u_n\right) = \lim_{n\to\infty} Qu_n = v.$$

There are several equivalent expressions for the operator norm; the expressions that we will use are the following:

$$\|Q\| = \inf\{C \geq 0 : \|Qu\| \leq C\|u\|, u \in \mathcal{B}\} = \sup\left\{\frac{\|Qu\|}{\|u\|} : u \neq 0\right\}. \tag{2.1}$$

The **spectrum** of the operator $Q : \mathcal{B} \to \mathcal{B}$ is the set of all complex numbers $z \in \mathbb{C}$ such that $Q - zI$ is not invertible or its inverse is not bounded. The **resolvent set** is the complement of the spectrum:

$$\operatorname{spec}(Q) = \{z \in \mathbb{C} : Q - zI \textit{ is not invertible or its inverse is not bounded.}\}, \tag{2.2}$$
$$\operatorname{res}(Q) = \{\operatorname{spec}(Q)\}^C. \tag{2.3}$$

An element of the spectrum, $\lambda \in \operatorname{spec}(Q)$ is called **eigenvalue of Q** if $Q - \lambda I : \mathcal{B} \to \mathcal{B}$ is not injective.[1] This property guarantees that there exists $u, w \in \mathcal{B}$, $u \neq w$, such that $Qu - \lambda u = Qw - \lambda w$. Therefore, there exists a non-zero vector $v \in \mathcal{B}$ such that $(Q - \lambda I)v = 0$. Such vector is called **eigenvector associated to** λ. For an eigenvalue $\lambda \in \mathbb{C}$, the **geometric multiplicity** is the dimension of the eigenspace $\{u \in \mathcal{B} : (Q - \lambda I)u = 0\}$. The **algebraic multiplicity** is the dimension of the generalized eigenspace $\{u \in \mathcal{B} : \exists m \geq 1 : (Q - \lambda)^m u = 0\}$(see Section 1.3 from [5], after Definition 1.13).

When $\mathcal{B}$ is finite dimensional, the operator Q can always be expressed as a matrix and the spectrum is formed with isolated eigenvalues only, but in the infinite dimensional case, there can be more elements on the spectrum that are not eigenvalues. Moreover, besides isolated points, the spectrum may be formed with a continuum of points in $\mathbb{C}$. In the case of the Perron-Frobenius operator, if one can find a Banach space where the spectrum has a specific configuration, then we can conclude that the system satisfies certain statistical properties. This specific configuration of the spectrum is called **quasicompactness** and it will be described in detail in this section.

The spectral radius of an operator Q, is defined by:

$$\operatorname{spr} Q = \lim_{n\to\infty} \|Q^n\|^{1/n}. \tag{2.4}$$

where $\|Q\|$ is the operator norm. This limit always exists and in fact, one can obtain an equivalent expression as follows. For a fixed $m > 0$, set $n = mq + r$ where q, r are integers such that $q, r \geq 0$ and $0 \leq r < m$. Then $\|Q^n\| = \|Q^{mq+r}\| \leq \|Q^{mq}\|\|Q^r\|$ and therefore:

$$\frac{1}{n}\log\|Q^n\| \leq \frac{q}{n}\log\|Q^m\| + \frac{1}{n}\log\|Q^r\|.$$

This implies:

$$\begin{aligned}
\limsup_{l\to\infty}\left(\frac{1}{l}\log\left\|Q^l\right\|\right) &= \lim_{l\to\infty}\left\{\sup_{n\geq l}\left\{\frac{1}{n}\log\|Q^n\|\right\}\right\} \leq \lim_{l\to\infty}\left\{\sup_{n\geq l}\left\{\frac{q}{n}\log\|Q^m\| + \frac{1}{n}\log\|Q^r\|\right\}\right\} \\
&\leq \lim_{l\to\infty}\left\{\frac{1}{m}\log\|Q^m\| + \frac{1}{l}\max_{0\leq r<m}\{\log\|Q^r\|\}\right\} \\
&= \frac{1}{m}\log\|Q^m\|.
\end{aligned}$$

As this is valid for all $m > 0$, we have:

$$\limsup_{l\to\infty}\left(\frac{1}{l}\log\left\|Q^l\right\|\right) \leq \inf_{l>0}\left\{\frac{1}{l}\log\left\|Q^l\right\|\right\}.$$

On the other hand, $\inf_{l>0}\left\{\frac{1}{l}\log\left\|Q^l\right\|\right\} \leq \liminf_{l\to\infty}\left(\frac{1}{l}\log\left\|Q^l\right\|\right)$, so the following equality holds:

$$\lim_{l\to\infty}\left(\frac{1}{l}\log\left\|Q^l\right\|\right) = \inf_{l>0}\left\{\frac{1}{l}\log\left\|Q^l\right\|\right\}.$$

[1] This means that there exist $u, w \in \mathcal{B}$ such that $Qu - \lambda u = Qw - \lambda w$ and $u \neq w$.

Which finally leads to:

$$\operatorname{spr}(Q) = \inf_{l>0} \left\{ \left\| Q^l \right\|^{1/l} \right\}. \tag{2.5}$$

This means that the spectral radius of Q is bounded by its operator norm. If λ is an eigenvalue of Q, then its magnitude is also bounded by $\|Q\|$. This follows by noticing that $\|Qv\| \leq \|Q\|\|v\|$ and using the fact that $\|Qv\| = |\lambda|\|v\|$. Note that $\operatorname{spr}(Q) < 1$ implies that $\|Q^m\| < 1$ for some $m > 0$. This can be proved by contrapositive. Suppose that $\|Q^m\| \geq 1$ for all $m > 0$. Then $\|Q^m\|^{1/m} \geq 1$ and therefore $\operatorname{spr} Q = \inf_n \|Q^n\|^{1/n} \geq 1$. These observations lead to the conclusion that the spectral radius of Q coincides with the supremum of the magnitudes among all elements of the spectrum (see for example Theorem 1.7.3 from [4]), i.e.

$$\operatorname{spr}(Q) = \sup \{|z| : z \in \operatorname{spec}(Q)\}. \tag{2.6}$$

Before presenting the definition of quasicompactness, let us recall some concepts on Functional Analysis. Let $\mathcal{B}$ be a vector space and M', M'' be subspaces of $\mathcal{B}$. We say that $\mathcal{B}$ is decomposed into the direct sum of M' and M'' and denote it $\mathcal{B} = M' \oplus M''$, if each element of $\mathcal{B}$ can be uniquely expressed as $u = u' + u''$ with $u' \in M'$ and $u'' \in M''$. The following characterization is well known: $\mathcal{B} = M' \oplus M''$, if and only if each element of $\mathcal{B}$ can be expressed as $u = u' + u''$ with $u' \in M'$ and $u'' \in M''$ and $u' + u'' = 0$ implies that $u' = u'' = 0$. Let us now state the formal definition of quasicompactness which will play a crutial role on our result.

Definition 2.1.1. *Let $\mathcal{Q}$ be a bounded linear operator acting on Banach space $(\mathcal{B}, \|\cdot\|)$. We say Q is quasicompact on $\mathcal{B}$, if $\mathcal{B}$ can be decomposed into the direct sum of two invariant subspaces $\mathcal{B} = M' \oplus M''$ with $\operatorname{spr}(Q|_{M''}) < \operatorname{spr}(Q)$, $\dim(M') < \infty$ and each eigenvalue of $Q|_{M'}$ has magnitude $\operatorname{spr}(Q)$.*

In many applications, one finds the particular case in which $\operatorname{spr}(Q) = \dim(M') = 1$. In this case the spectrum of $Q|_{M'}$ consists of a single element λ whose magnitude must be 1. So, there exists $v \in M'$ such that $Qv = \lambda v$. When Q is a Perron-Frobenius operator and $\lambda = 1$, this means that v is density of an absolutely continuous invariant measure. The following provides a description of the special configuration of the spectrum of a quasicompact operator. The same ideas are illustrated in Figure 2.1:

- When an operator is quasicompact, its spectrum can be decomposed into two disjoint parts on the complex plane: Σ' and Σ''. See Lemma 7.1.1 on the Appendix (Chapter 7) for a formal proof of this statement.
- One of this parts, Σ', contains only a finite number of elements.
- The other part, Σ'' may contain an infinite number of elements, but all of them have a magnitude less than the magnitude of the elements of Σ'.

Before providing examples of quasicompact Perron-Frobanius operators, let us make some comments that will be useful later on. The fact that $\mathcal{B}$ is decomposed as a direct sum of two Q-invariant subspaces, provides additional insights for the operator Q. Let $u = u' + u''$ be the unique form in which $u \in \mathcal{B}$ can be expressed as the sum of an element of M' and an element of M''. Then we can set $Pu = u'$ the projection operator on M' which is clearly idempotent ($P^2 = P$). $N = 1 - P$ is the projection operator on M'' and it turns out that the operator Q commutes with P and with N (i.e. $PQ = QP$ and $NQ = QN$). Moreover P also commutes with N and its composition is zero ($PN = NP = 0$). P coincides with $Q|_{M'}$ and N coincides with $Q|_{M''}$. When $\dim(M') = 1$ each element of M' can be expressed as a linear combination of a base element $v \in M'$ which is an eigenvector of Q and whose associated eigenvalue is the unique element of $\operatorname{spec}(Q|_{M'})$, λ. So $Qu' = zQv = \lambda zv = \lambda u'$. Using the linearity of Q and the

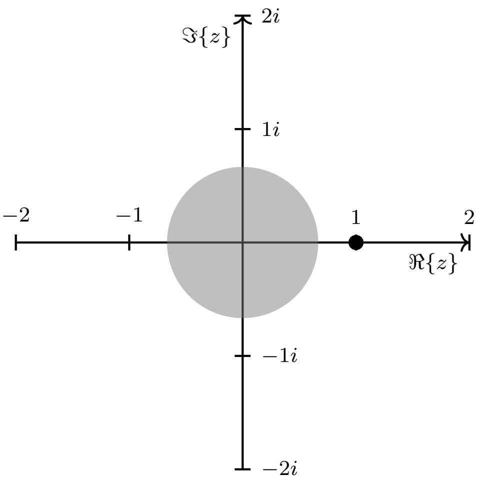


Figure 2.1: Illustration of quasicompactness. The spectrum of Q is decomposed into two disjoint parts on the complex plane: Σ' and Σ''. Σ' only contains a finite number of points ($\Sigma' = \{1\}$ in the figure). The other part, Σ'' may contain a continuum of points, but they all have magnitude less than the magnitude of elements of Σ' (Σ'' is the gray circle).

invariance of M' and M'' under Q, it follows that Q can be expressed as the sum of $\lambda P + NQ$ because for each $u \in \mathcal{B}$:

$$\lambda Pu + NQu = \lambda u' + NQu'' = \lambda u' + (1-P)Qu'' = \lambda u' + Qu'' = Qu' + Qu'' = Qu. \tag{2.7}$$

Let us now see an example where quasicompactness can be established directly using Definition 2.1.1.

Example 2.1.2. *Let $X = [0,1]$ and $T : X \to X$ defined by $T(x) = 2x \mod 1$. Set $\mathcal{B}$ the Banach space of Lipschitz functions $u : X \to \mathbb{C}$ with the Lipschitz norm given by:*

$$\|u\|_{\mathrm{Lip}} = \|u\|_\infty + |u|_{\mathrm{Lip}},$$

where:

$$|u|_{\mathrm{Lip}} = \sup_{x \neq y} \frac{|u(x) - u(y)|}{|x-y|}.$$

Let us verify that the definition of quasicompactnes is satisfied for the Perron-Frobenius operator acting on $\mathcal{B}$. Using Equation (1.20), the Perron-Frobenius operator associated to the Lebesgue measure and to the transformation T can be written as:

$$P_T u(x) = \frac{1}{2}\left[u\left(\frac{x}{2}\right) + u\left(\frac{x+1}{2}\right)\right]. \tag{2.8}$$

It is easy to see that P_T is a bounded linear operator on $\mathcal{B}$ (indeed, Inequality (2.9) below implies that P_T is bounded). $\mathcal{B}$ can be decomposed into $M' = \{z \cdot 1_X : z \in \mathbb{C}\}$ and $M'' = \left\{u - \int u d\mu : u \in \mathcal{B}\right\}$ and it is easy to see that, if $u' + u'' = 0$ with $u' \in M'$ and $u'' \in M''$, then $u' = u'' = 0$. Also, $P_T M' \subseteq M'$ and $P_T M'' \subseteq M''$. The fact that $P_T 1_X = 1_X$ guarantees that $1 \in \operatorname{spec}(P_T)$ and together with Equation (2.6) we get that $\operatorname{spr}(P_T) \geq 1$. The reverse inequality can be obtained by noticing that $|P_T u|_{\mathrm{Lip}} \leq \frac{1}{2}|u|_{\mathrm{Lip}}$ and:

$$\|P_T^n u\|_{\mathrm{Lip}} = |P_T^n u|_{\mathrm{Lip}} + \|P_T^n u\|_\infty \leq \frac{1}{2^n}|u|_{\mathrm{Lip}} + \|u\|_\infty \leq \left(\frac{1}{2^n} + 1\right)\|u\|_{\mathrm{Lip}}. \tag{2.9}$$

The above implies that $\|P_T^n\|_{\text{Lip}} \leq \frac{1}{2^n} + 1$ *and using Equation 2.4 we obtain:*

$$\operatorname{spr}(P_T) = \lim_{n\to\infty} \|P_T^n\|_{\text{Lip}}^{1/n} \leq \lim_{n\to\infty} \left(\frac{1}{2^n} + 1\right)^{1/n} = 1. \tag{2.10}$$

Which leads to the equality $\operatorname{spr}(P_T) = 1$. *We now obtain an upper bound for* $\operatorname{spr}(P_T|_{M''})$. *For this, we will show that for* $u \in M''$, *we have* $\|u\|_\infty \leq 2|u|_{\text{Lip}}$. *Indeed, for* $u \in M''$, *there exist* $x_1, x_2 \in \mathbb{C}$ *such that* $\operatorname{Re}(u(x_1)) = \operatorname{Im}(u(x_2)) = 0$, [2] *therefore:*

$$\begin{aligned}
|u(x)| &\leq |u(x) - \operatorname{Re}(u(x_1)) - i\operatorname{Im}(u(x_2))| \\
&\leq |\operatorname{Re}(u(x)) - \operatorname{Re}(u(x_1))| + |\operatorname{Im}(u(x)) - \operatorname{Im}(u(x_2))| \\
&\leq |u(x) - u(x_1)| + |u(x) - u(x_2)| \\
&\leq |u|_{\text{Lip}}|x - x_1| + |u|_{\text{Lip}}|x - x_2| \\
&\leq 2|u|_{\text{Lip}},
\end{aligned}$$

and consequently, $\|u\|_\infty \leq 2|u|_{\text{Lip}}$[3]. *Actually* $\|\cdot\|_\infty$ *and* $|\cdot|_{\text{Lip}}$ *are equivalent norms in* M'' *(because* u *is a bounded Lipschitz function, we have that* $|u|_{\text{Lip}} \leq 2\|u\|_\infty$*). For our purposes, the inequality* $\|u\|_{\text{Lip}} = |u|_{\text{Lip}} + \|u\|_\infty \leq 3|u|_{\text{Lip}}$ *will be enough:*

$$\begin{aligned}
\|(P_T|_{M''})^n u\|_{\text{Lip}} \leq 3|(P_T|_{M''})^n u|_{\text{Lip}} &\leq \frac{3}{2^n}|u|_{\text{Lip}} \leq \frac{3}{2^n}\|u\|_{\text{Lip}}, \\
\operatorname{spr}(P_T|_{M''}) = \lim_{n\to\infty} \|(P_T|_{M''})^n\|_{\text{Lip}}^{1/n} &\leq \lim_{n\to\infty} \left(\frac{3}{2^n}\right)^{1/n} = \frac{1}{2},
\end{aligned}$$

which implies that $\operatorname{spr}(P_T|_{M''}) \leq \frac{1}{2} < 1 = \operatorname{spr}(P_T)$. $1_X \in M'$ *and* 1 *is the only eigenvalue of* $P_T|_{M'}$ *and has magnitude* 1. *Finally,* M' *is a one-dimensional subspace because each* $u' \in M'$ *can be written as a linear combination of the base element* 1_X. *We have verified each requirement of definition 2.1.1 and we can safely conclude that* P_T *is a quasicompact operator for* $T(x) = 2x \pmod 1$.

Proving quasicompactness by directly using the definition as we did in Example 2.1.2 is much more elaborate on more general scenarios. But there exists a functional result that has been used as an important tool for this purpose: this is the called **Ionescu-Tulcea and Marinescu Theorem (I-TM Theorem)**. Many important results in Deterministic and Random Dynamical Systems have been established using quasicompactness of the Perron-Frobenius operator obtained by means of the I-TM Theorem. We briefly discuss an example on how this result is used, however, we remark that the authors in references [61] and [28] which are the base of our work, present a new method to prove quasicompactness without the need to use the I-TM Theorem. We think that this is an outstanding result and that it provides a tool to prove statistical properties in other settings that have not previously considered. A step in this direction is the central limit theorem that we present in Chapter 4.

Let us now provide a description of the I-TM Theorem and a basic example. Such result was established on reference [87] and it was later generalized on references [38] and [39] (Theorem II.5 or more generally, Theorem XIV.3). Suppose that Q is a bounded operator on a Banach space and $|\cdot|$ is a continuous seminorm on $\mathcal{B}$. Moreover, suppose the following conditions hold:

(ITM1): The image of the unit ball of the norm $\|\cdot\|$ under Q [4] is conditionally compact in $(\mathcal{B}, |\cdot|)$.

(ITM2): There exists a constant $C > 0$ such that for all $u \in \mathcal{B}$, $|Qu| \leq C|u|$.

[2] To see why it is enough to observe that $\int u d\mu = 0$ for all $u \in M''$. A Lipschitz function whose integral is zero, must cross the zero axis at some value.

[3] Note that we have used the fact that $|x - y| \leq 1$ for any $x, y \in X$.

[4] i.e. $Q\,(u \in \mathcal{B} : \|u\| \leq 1)$.

(ITM3): There exist $k \in \mathbb{N}$ and $R, r \in \mathbb{R}$ with $r < \operatorname{spr}(Q)$ such that for all $u \in \mathcal{B}$:

$$\left\| Q^k u \right\| \leq R|u| + r^k \|u\|. \tag{2.11}$$

Before providing an example, let us state some comments on the above conditions. Condition ITM1 means that any sequence $\{Qu_n\}_n$ with $\|u_n\| \leq 1$ must contain a subsequence that converges on $\mathcal{B}$ under the seminorm $|\cdot|$: there exists a subsequence Qu_{n_k} and $u \in \mathcal{B}$ such that $\lim_{k\to\infty} |Qu_{n_k} - u| = 0$ [5]. Sometimes Condition ITM1 is stated in a different equivalent form as:

(ITM1a): Q is a compact operator from $(\mathcal{B}, \|\cdot\|)$ to $(\mathcal{B}, |\cdot|)$.

The inequality on Condition ITM3 is called **Lasota-Yorke inequality** and we will see how it is satisfied on the example below. In Section 3.2.2 we also present the Lasota-Yorke inequality obtained in [61] for the class of Random Dynamical Systems that we are interested in.

Example 2.1.3. *Let $X = \Sigma_A^+$ be a subshift of finite type associated to an irreducible and aperiodic matrix A and let $w : X \to \mathbb{R}$ be a Lipschitz potential. As we saw on Section 1.2.2, we can endow X with a distance given by:*

$$d_\theta(\underline{x}, \underline{y}) = \sum_{i=0}^{\infty} \theta^i \left(1 - \tau(x_i, y_i)\right), \tag{2.12}$$

Where $\theta \in (0,1)$ and:

$$\tau(x_i, y_i) = \begin{cases} 0 & x_i \neq y_i, \\ 1 & x_i = y_i. \end{cases}$$

Let $\mathcal{B}$ be the Banach space of Lipschitz functions $\varphi : X \to \mathbb{C}$ endowed with the norm $\|\cdot\|_\theta = \|\cdot\|_\infty + |\cdot|_\theta$ where:

$$|\varphi|_\theta = \sup_{\underline{x}, \underline{y} \in \Sigma_A^+} \left\{ \frac{|\varphi(\underline{x}) - \varphi(\underline{y})|}{d_\theta(\underline{x}, \underline{y})} \right\}. \tag{2.13}$$

The operator $P_\sigma : \mathcal{B} \to \mathcal{B}$ is well defined. Suppose further that P_σ is normalized. It means that the potential $w : \Sigma_A^+ \to \mathbb{R}$ satisfies the following for all $\underline{x} \in X$:

$$\sum_{\underline{y} \in \sigma^{-1}\underline{x}} e^{w(\underline{y})} = 1.$$

Under this scenario, we want to verify the conditions ITM1 to ITM3 in I-TM Theorem. The above normalization condition implies that $P_\sigma 1_X = 1_X$ and $\|P_\sigma \varphi\|_\infty \leq \|\varphi\|_\infty$. So ITM2 is satisfied with seminorm $\|\cdot\|_\infty$ and $C = 1$. The following basic inequality is obtained in reference [74] (Proposition 2.1):

$$|P_\sigma^n \varphi|_\theta \leq \theta^n |\varphi|_\theta + C\|\varphi\|_\infty \qquad \forall n \geq 1. \tag{2.14}$$

A Lasota-Yorke inequality (ITM3) can be obtained from the above as follows:

$$\|P_\sigma^n \varphi\|_\theta = |P_\sigma^n \varphi|_\theta + \|P_\sigma^n \varphi\|_\infty \leq \theta^n |\varphi|_\theta + C\|\varphi\|_\infty + \|\varphi\|_\infty = \theta^n \|\varphi\|_\theta + (C+1)\|\varphi\|_\infty. \tag{2.15}$$

We will use Inequality (2.15) *later in Section 3.2.2 as an ingredient to establish spectral gap for the Perron-Frobenius operator acting on certain space of signed measures. Note that the*

[5] In topological terms, a conditionally compact subset of a topological space is a subset whose closure is compact. Condition 1 is understood considering that $\mathcal{B}$ is endowed with the pseudometric topology defined by the seminorm $|\cdot|$.

seminorm from Inequality (2.11) *corresponds to the supremum norm from inequality* (2.15) [6]*. It is left to show that the image of the unit ball of the norm* $\|\cdot\|_\theta$ *under* P_σ *is conditionally compact in* $(\mathcal{B}, \|\cdot\|_\infty)$ *(see discussion after Equation* (2.11) *for an explanation of a conditionally compact set). For this purpose, let* φ_n *be a sequence of Lipschitz functions such that* $\|\varphi_n\|_\theta \leq 1$*. We will remark the following properties of* $P_\sigma\varphi_n$*. For each* $n \geq 1$*:*

1. $P_\sigma\varphi_n$ *is uniformly bounded by 1:*

$$|P_\sigma\varphi_n(\underline{x})| \leq \|P_\sigma\varphi_n\|_\infty \leq \|\varphi_n\|_\infty \leq \|\varphi_n\|_\theta \leq 1.$$

2. $|P_\sigma\varphi_n|_\theta$ *is bounded by* $\theta + C$ *(using* (2.15)*):*

$$|P_\sigma\varphi_n|_\theta \leq (\theta + C)\,|\varphi_n|_\theta \leq \theta + C.$$

Property 2 implies that $P_\sigma\varphi_n$ *is* ***uniformly equicontinuous****: given* $\epsilon > 0$*, choose* $\delta = \frac{\epsilon}{\theta+C}$ *so that if* $d_\theta(\underline{x}, \underline{y}) < \delta$ *we have:*

$$\left|P_\sigma\varphi_n(\underline{x}) - P_\sigma\varphi_n(\underline{y})\right| < (\theta + C)d_\theta(\underline{x}, \underline{y}) < (\theta + C)\delta = (\theta + C)\frac{\epsilon}{\theta + C} = \epsilon.$$

By the Ascoli Theorem, $P_\sigma\varphi_n$ *contains a subsequence that converges uniformly n* $\mathcal{B}$*. Uniform convergence on* $\mathcal{B}$ *implies convergence in the supremum norm* [7]*, so we can conclude that ITM1 is satisfied and consequently, the Perron-Frobenius operator* $P_\sigma : \mathcal{B} \to \mathcal{B}$ *is quasicompact.*

2.1.2 Perturbed Operator

In this Section we define a perturbed operator and use Perturbation Theory to establish quasi-compactness of such perturbed operator. Throughout this Section let X be a compact metric space, $T : X \to X$ be a mixing transformation with respect to the invariant probability measure μ. We also assume that the Perron-Frobenius operator P_T has spectral gap on a Banach space $\mathcal{B}$ and the observable $f : X \to \mathbb{R}$ belongs to $\mathcal{B}$ and satisfies $\int f d\mu = 0$. Let us start by defining the **perturbed transfer operator** $P_{f,t} : \mathcal{B} \to \mathcal{B}$ by:

$$P_{f,t}\varphi = P_T\left(e^{itf}\varphi\right).$$

$P_{f,t}$ is a bounded linear operator if P_T is. Using Equation (1.16) we obtain $\int P_{f,t}\varphi d\mu = \int 1_X \cdot P_t\left(e^{itf}\varphi\right) d\mu = \int e^{itf}\varphi d\mu$. By induction, we obtain:

$$\int P^n_{f,t}\varphi d\mu = \int e^{itS_nf}\varphi d\mu, \tag{2.16}$$

with $\varphi = 1_X$ the above equality becomes a relation between the iterations of the perturbed transfer operator and the characteristic function of the random variable $S_nf/\sqrt{n}$:

$$\phi_{n,X}(t) = \int e^{itS_nf/\sqrt{n}}d\mu = \int P^n_{f,t/\sqrt{n}}1_X d\mu. \tag{2.17}$$

This accomplishes step ST1 mentioned at the beginning of this section. For the next step we need to establish the regularity of $P_{f,t}$ with respect to the parameter t. Such regularity will enable us to use Perturbation Theory at the end of this section.

[6] In other words, the seminorm $|\cdot|_\theta$ is not the seminorm required on ITM Theorem.

[7] Uniform convergence means that for all $\epsilon_2 > 0$, there exists N_2 such that for all $\underline{x} \in X$ and $n \geq N$, $|P_\sigma\varphi_n(\underline{x}) - P_\sigma\varphi(\underline{x})| < \epsilon_2$. To establish convergence in norm, choose $\epsilon_1 = \frac{\epsilon_2}{2}$.

Let $DP_{f,t} : \mathcal{B} \to \mathcal{B}$ be defined by $DP_{f,t}\varphi = P_T\left(if\varphi e^{itf}\right)$. Our objective is to verify that the operator defined by $\Delta \mapsto \Delta \cdot DP_{f,t}$ for $\Delta \in \mathbb{R}$, is the derivative of $P_{f,t}$ (at any $t \in \mathbb{R}$). For this purpose we note that for all $\varphi \in \mathcal{B}$:

$$\begin{aligned}\lim_{\Delta\to 0}\frac{1}{\Delta}\left(P_{f,t+\Delta}(\varphi) - P_{f,t}(\varphi) - \Delta \cdot DP_{f,t}(\varphi)\right) &= P_T\left(\varphi e^{itf} \cdot \lim_{\Delta\to 0}\frac{1}{\Delta}\left(e^{i\Delta f} - 1 - \Delta i f\right)\right)\\ &= P_T(\varphi e^{itf} \cdot 0) = P_T(0) = 0.\end{aligned}$$

Note that we have used the continuity of P_T to exchange positions with the limit. Also, $P_T(0) = 0$ follows by considering that $\varphi \geq 0$ implies that $P_T\varphi \geq 0$ and using equation (1.16) with $\varphi = 0$ and $\psi = 1_X$. This means that:

$$\lim_{|\Delta|\to 0}\frac{\|P_{f,t+\Delta} - P_{f,t} - \Delta \cdot DP_{f,t}\|}{|\Delta|} = 0. \tag{2.18}$$

According to Section 7.1.2 on the Appendix, the above means that $DP_{f,t}$ is the derivative of $P_{f,t}$ with respect to t. Similarly, one can verify that the higher order derivatives can be expressed as $D^nP_{f,t}\varphi = P_T\left((if)^n\varphi e^{itf}\right)$. Therefore, the map $t \mapsto P_{f,t}$ is C^∞ and can be seen as analytical perturbation of the bounded, linear operator P_T. We now want to use Perturbation Theory in order to conclude spectral gap for $P_{f,t}$. For that purpose we will use the classical reference on this topic [52]. Let us define a function, similar to a distance between two closed operators. Let M, N two subspaces from a Banach space $\mathcal{B}$ and:

$$\delta(M,N) = \sup_{u\in M, \|u\|=1}\left(\inf_{v\in N}\|u-v\|\right), \tag{2.19}$$

$$\hat{\delta}(M,N) = \max\left[\delta(M,N), \delta(N,M)\right]. \tag{2.20}$$

For two closed operators Q, S acting on a Banach space $\mathcal{B}$, let:

$$\delta(Q,S) = \delta(\mathrm{G}(Q), \mathrm{G}(S)), \tag{2.21}$$

$$\hat{\delta}(Q,S) = \max\left[\delta(\mathrm{G}(Q), \mathrm{G}(S)), \delta(\mathrm{G}(S), \mathrm{G}(T))\right]. \tag{2.22}$$

where $G(Q)$ is the **graph of the operator** Q defined by:

$$\mathrm{G}(Q) = \{(u, Qu) : u \in \mathcal{B}\}.$$

The norm from equation (2.21) can be fixed to be the euclidean norm on the product space $\mathcal{B}^2$:

$$\|(\varphi, P_{f,t}\varphi)\| = \left(\|\varphi\|^2 + \|P_{f,t}\varphi\|^2\right)^{1/2}.$$

For our purposes, the following property of $\hat{\delta}(P_T, P_{f,t})$ will be used:

Proposition 2.1.4. $\lim_{t\to 0}\hat{\delta}\left(P_T, P_{f,t}\right) = 0$.

Proof. $P_{f,t}$ is analytic (in particular, continuous) on the real parameter t, so:

$$\lim_{t\to 0}\|P_T\varphi - P_{f,t}\varphi\| = 0.$$

Therefore:

$$\begin{aligned}\lim_{t\to 0}\delta(\mathrm{G}(P_T), \mathrm{G}(P_{f,t})) &= \lim_{t\to 0}\left[\sup_{u\in \mathrm{G}(P_T), \|u\|=1}\left(\inf_{v\in \mathrm{G}(P_{f,t})}\|u-v\|\right)\right]\\ &= \sup_{u\in \mathrm{G}(P_T), \|u\|=1}\left(\lim_{t\to 0}\left(\inf_{v\in \mathrm{G}(P_{f,t})}\|u-v\|\right)\right).\end{aligned} \tag{2.23}$$

For $\varphi \in \mathcal{B}$ such that $\|u\| = \|(\varphi, P_T\varphi)\| = 1$, let $v = (\varphi, P_{f,t}\varphi)$ (i.e. u and v have the same first element, $\pi_1 u = \pi_1 v$). Then:

$$\|u - v\| = \|(\varphi, P_T\varphi) - (\varphi, P_{f,t}\varphi)\| = \|P_T\varphi - P_{f,t}\varphi\|.$$

Therefore, the infimum on (2.23) is bounded by $\|P_T\varphi - P_{f,t}\varphi\|$ and the supremum on that equation is bounded by:

$$\lim_{t\to 0} \delta(\mathrm{G}(P_T), \mathrm{G}(P_{f,t})) \leq \lim_{t\to 0} \|P_T\varphi - P_{f,t}\varphi\| = 0, \tag{2.24}$$

with the same process we obtain:

$$\lim_{t\to 0} \delta(\mathrm{G}(P_{f,t}), \mathrm{G}(P_T)) = 0, \tag{2.25}$$

(2.24) and (2.25) confirm that the proposition holds. □

The above proposition means that, given $\epsilon > 0$, there exists $\beta > 0$ such that if $|t| < \beta$, then $\hat{\delta}\left(P_T, P_{f,t}\right) < \epsilon$. In particular, we can fix $\epsilon = \delta$ where δ is the real number that exists as consequence of Theorem IV 3.16 from reference [52]. More specifically, as the spectrum of $P_{f,0} = P_T$ can be expressed as the disjoint union of $\Sigma'(0) = \{1\}$ and $\Sigma''(0) = \mathrm{spec}(P_T|_{M''})$, one can draw a simple closed curve Γ on the complex plane that contains $\Sigma'(0)$ in its interior and $\Sigma''(0)$ in its exterior. Then, Theorem IV 3.16 guarantees that for small enough $|t|$, the same is valid for $P_{f,t}$ (with the same Γ). The associated decomposition $\mathcal{B} = M'(t) \oplus M''(t)$ is formed with $M'(t) = \Pi_{f,t}\mathcal{B}$ and $M''(t) = (1 - \Pi_{f,t})\mathcal{B}$ where:

$$\Pi_{f,t} = -\frac{1}{2\pi i} \oint_\Gamma \left(P_{f,t} - zI\right)^{-1} dz.$$

Theorem IV 3.16 also guarantees that there exists an isomorphism between $M'(0)$ and $M'(t)$; in particular, $\dim(M'(t)) = \dim(M'(0)) = 1$, thus there exist $\lambda(t) \in \mathbb{R}$ and $v(t) \in M'(t)$ such that $P_{f,t}|_{M'(t)}v(t) = \lambda(t)v(t)$ and:

$$\mathrm{spec}(P_{f,t}|_{M'(t)}) = \{\lambda(t)\}.$$

Actually, $M'(t)$ is the span of $v(t)$ and both are analytical functions of t.

Note that Γ does not cross any point of the spectrum, so the integral is well defined. Despite its complicated appearance, the operator $\Pi_{f,t}$ is just a projection on $M'(t)$ (idempotent operator). Indeed, set $c(t) = \left(\lambda^2(t) - z\lambda(t)\right)^{-1}$ and for $\varphi \in M'(t)$:

$$\begin{aligned} c(t)P_{f,t}|_{M'(t)}\left(P_{f,t} - zI\right)\varphi &= c(t)P_{f,t}|_{M'(t)}\left(P_{f,t}\varphi - z\varphi\right) \\ &= c(t)\left(P_{f,t}^2|_{M'(t)}\varphi - zP_{f,t}|_{M'(t)}\varphi\right) \\ &= c(t)(\lambda(t) - z)\lambda(t)\varphi = \varphi. \end{aligned}$$

Where we have used the fact that $M'(t)$ is unidimensional. Similarly, one can verify that $\left(P_{f,t} - zI\right)c(t)P_{f,t}|_{M'(t)}\varphi = \varphi$. We can conclude that $\left(P_{f,t} - zI\right)^{-1} = c(t)P_{f,t}|_{M'(t)}$. Using the Residue Theorem:

$$\Pi_{f,t}\varphi = -\frac{1}{2\pi i}\oint_\Gamma c(t)P_{f,t}|_{M'(t)}\varphi dz = -\frac{P_{f,t}|_{M'(t)}\varphi}{2\pi i}\oint_\Gamma (\lambda^2(t) - z\lambda(t))^{-1}dz = \frac{P_{f,t}|_{M'(t)}\varphi}{\lambda(t)} = \varphi.$$

In order to establish quasicompactnes of the operator $P_{f,t}$ it is left to show that $\mathrm{spr}(P_{f,t}|_{M''(t)}) < \mathrm{spr}(P_{f,t}) = |\lambda(t)|$. The spectrum of $P_{f,t}$ is **upper semicontinuous** (see Theorem IV 3.1 and Remark IV 3.2 from [52]). This means that for all $\epsilon > 0$ there exists $\delta > 0$ such that if $\|P_{f,t} - P_T\| < \delta$ then:

$$\sup_{z\in\mathrm{spec}(P_{f,t}|_{M''(t)})} \left\{ \inf_{w\in\mathrm{spec}(P_T|_{M''})} |z - w| \right\} < \epsilon.$$

Choose $\epsilon < |\lambda(t)| - \mathrm{spr}(P_T|_{M''})$. If $z \in \mathrm{spec}(P_{f,t}|_{M''(t)})$ and $|z| \geq \mathrm{spr}(P_T|_{M''})$ then:

$$|z| - \mathrm{spr}(P_T|_{M''}) \leq \inf_{w \in \mathrm{spec}(P_T|_{M''})} |z - w| < \epsilon < |\lambda(t)| - \mathrm{spr}(P_T|_{M''}).$$

This implies that $|z| < |\lambda(t)|$ for all $z \in \mathrm{spec}(P_{f,t}|_{M''(t)})$ and therefore:

$$\mathrm{spr}(P_{f,t}) = \max\left\{|\lambda(t)|, \mathrm{spr}(P_{f,t}|_{M''(t)})\right\} = |\lambda(t)|.$$

We have verified all the requirements for quasicompactness, so we can conclude that the perturbed operator $P_{f,t}$ is quasicompact on $\mathcal{B}$ and this completes step ST2 from the beginning of this section.

2.2. Central Limit Theorem for Deterministic Dynamical Systems

Quasicompactness of $P_{f,t}$ provides the main ingredients for the Central Limit Theorem in the deterministic case. Let us define an operator $N_{f,t} = (1 - \Pi_{f,t})P_{f,t}$. It is clear that

$$P_{f,t} = \lambda(t)\Pi_{f,t} + N_{f,t}, \tag{2.26}$$
$$\Pi_{f,t}N_{f,t} = N_{f,t}\Pi_{f,t} = 0, \tag{2.27}$$

The first equality is obtained as consequence of the following:

$$\Pi_{f,t}P_{f,t}\varphi = z\Pi_{f,t}P_{f,t}v(t) = z\lambda(t)\Pi_{f,t}v(t) = \lambda(t)\Pi_{f,t}(z \cdot v(t)) = \lambda(t)\Pi_{f,t}\varphi.$$

The other two equalities are obtained considering that $\Pi_{f,t}$ is idempotent ($\Pi_{f,t}^2 = \Pi_{f,t}$) and the fact that $\Pi_{f,t}P_{f,t} = P_{f,t}\Pi_{f,t}$. We still need to find a bound for the norm of the iterations of the operator $N_{f,t}$. For this purpose, recall that $\mathrm{spr}(P_T|_{M''(0)}) < \mathrm{spr}(P_T) = 1$. Then using Lemma 7.1.3 we obtain that there exists $\eta > 0$ and $c > 0$ such that if $\left\|P_{f,t}|_{M''(t)} - P_T|_{M''(0)}\right\| < \eta$ then:

$$\left\|N_{f,t}^n\right\| \leq c \cdot \xi^n,$$

for $\mathrm{spr}(P_T|_{M''(0)}) < \xi < 1$ and $n \geq 1$.

We can now continue with ST3. Using (2.27) we obtain for the iterations of the perturbed operator:

$$P_{f,t}^n\varphi = (\lambda(t)\Pi_{f,t} + N_{f,t})^n\,\varphi = \lambda^n(t)\Pi_{f,t}\varphi + N_{f,t}^n\varphi.$$

By (2.17) we know that the characteristic function of the random variable S_nf can be expressed in terms of the iterations of the Perron-Frobenius operator.

$$\begin{aligned}
\lim_{n\to\infty} \phi_{n,X}(t) &= \lim_{n\to\infty} \int P_{f,t/\sqrt{n}}^n 1_X\,d\mu = \lim_{n\to\infty}\left(\lambda^n(t/\sqrt{n})\int \Pi_{f,t/\sqrt{n}}1_X\,d\mu + \int N_{f,t/\sqrt{n}}^n 1_X\,d\mu\right)\\
&= \lim_{n\to\infty} \lambda^n(t/\sqrt{n})\Pi_{f,t/\sqrt{n}}1_X\\
&= \lim_{n\to\infty}\left(1 + \lambda'(0)\frac{t}{\sqrt{n}} + \lambda''(0)\frac{t^2}{2n} + \cdots\right)^n \Pi_{f,t/\sqrt{n}}1_X.
\end{aligned} \tag{2.28}$$

A calculation similar to the presented in Section 7.4 in the Appendix, confirms that $\lambda'(0) = 0$ and:

$$\lambda''(0) = -\rho^2 = \int f^2 d\mu + 2\sum_{k=1}^{\infty}\int f \cdot (f \circ T^k)\,d\mu.$$

The limit in Equation (2.28) can be now evaluated as:

$$\lim_{n\to\infty} \phi_{n,X}(t) = e^{-\rho^2 t^2/2}.$$

The above equation means that the characteristic function of the random variable $S_n f/\sqrt{n}$ converges pointwise to the characteristic function of a Gaussian random variable. By Lévy Continuity Theorem (see for example Theorem 14.15 from [25]), the random variable $S_n f/\sqrt{n}$ converges in distribution to a normal distribution with mean zero and variance ρ^2. This completes step ST3. As we have mentioned above, Section 7.4 in the Appendix relates the first derivative of the eigenvalue $\lambda(t)$ with the mean and the second derivative with the variance. This completes step ST4 and completes the discussion on the Spectral Method for the deterministic case.

It is worth mentioning here that we have used reference [52] for all the steps that require elements of Perturbation Theory. However, another shorter approach can be taken using Theorem III.8 from [39] which already summarizes the conditions specifically required to establish a Central Limit Theorem. In Chapter 4 we make use of Theorem III.8 to prove our main result. In Section 7.1.3 in the Appendix, we provide a detailed proof of the theorem. All the ideas come from the original reference [39], but we have added comments and details that hopefully make the proof easier to follow.

2.2.1 Spectral Method for Random Dynamical Systems

The Spectral Method has been successfully adapted to Random Dynamical Systems to prove Central Limit Theorems under several scenarios. In this Section we will briefly talk about these results. In all cases, the selection process of the maps is an IID process. A CLT is proved in [47] in the case where each of the involved transformation is an expanding map in the interval. It's worth noting that in this result, the limit distribution is a convex combination of the individual Normal distributions corresponding to each map.

In [40] the authors establish a CLT when the maps are Lipschitz and satisfy a property of average contraction. In [2] the authors establish a CLT when the averaged Perron-Frobenius operator satisfies the I-TM Theorem including a Lasota-Yorke inequality. In the same reference, the authors obtain other interesting statistical properties for RDS like a Large Deviation Principle and an exponential Concentration Inequality. A quenched CLT is also obtained, but in this case the phase space is restricted to be the unit interval and all the individual maps are required to preserve the Lebesgue measure.

In a recent paper ([36]), the author proved statistical properties for non-uniformly expanding, random dynamical systems. He obtained a Berry-Esseen theorem, a local central limit theorem and large and moderate deviations principles for random non-uniformly expanding dynamical systems with exponential first return times. See also [35, 37] for other related results in the case of expanding maps where the base satisfies a mixing condition.

An interesting variation of random dynamical system is considered in [44]. In this paper, the maps are randomly selected according to probability vectors that depend on the position. Moreover, the system experiences jumps at random times, and these jumps are defined by randomly selecting a map that also depends on the position. In reference [45] the author proves a CLT for this interesting system.

3. Setting of Interest and Spectral Gap

In this chapter we describe the class of contractive RDS that we are interested in, following the lines of reference [61]. In this reference the authors prove that the pushforward measure has spectral gap on a convenient real normed vector space of signed measures. In Section 3.1 we define the setting of interest and then, in Section 3.2 we present an overview of the results of reference [61], which includes the existence of a unique invariant probability measure, decay of correlations and spectral gap of the pushforward measure. We will use these results in Chapter 4 where we state and prove a Central Limit Theorem for such class of contractive random dynamical systems.

3.1. Contractive Random Dynamical Systems

This section will be devoted to describe the setting of interest which is essentially the same as [61]. Let K be a compact metric space whose distance is denoted d. Given an irreducible and aperiodic matrix A, consider Σ_A^+ to be the one-sided subshift of finite type associated to A and endowed with the distance $d_\theta(\underline{x}, \underline{y}) = \sum_{i=0}^{\infty} \theta^i (1 - \tau(x_i, y_i))$, where $\underline{x}$, $\underline{y} \in \Sigma_A^+$, $\theta \in (0, 1)$ and for all i, $\tau(\cdot, \cdot)$ is given by:

$$\tau(x_i, y_i) = \begin{cases} 0 & \text{if } x_i \neq y_i, \\ 1 & \text{if } x_i = y_i. \end{cases}$$

Let $\Sigma = \Sigma_A^+ \times K$ and let μ be a measure on the Borel sigma-algebra of Σ. Consider the dynamics $F : \Sigma \to \Sigma$ given by $F(\underline{x}, z) = (\sigma \underline{x}, G(\underline{x}, z))$, where $\sigma(x_0 x_1 x_2 ...) = x_1 x_2 x_3 ...$ is the shift map and $G : \Sigma \to K$ satisfies the following two conditions:

C1: There exists $0 < \alpha < 1$ such that $d(G(\underline{x}, z_1), G(\underline{x}, z_2)) \leq \alpha d(z_1, z_2)$ for all $\underline{x} \in \Sigma_A^+$ and for all $z_1, z_2 \in K$.

C2: For each $z \in K$ there exists k_z such that $d(G(\underline{x}, z), G(\underline{y}, z)) \leq k_z d_\theta(\underline{x}, \underline{y})$ and $\operatorname{ess\,sup}_{z \in K} k_z < \infty$.

Let $F^*\mu$ be the pushforward measure $F^*\mu(E) = \mu(F^{-1}E)$ for any measurable set E. We define the projections $\pi_1 : \Sigma \to \Sigma_A^+$, given by $\pi_1(\underline{x}, z) = \underline{x}$ and $\pi_2 : \Sigma \to K$ given by $\pi_2(\underline{x}, z) = z$. Let μ be a complex measure and let f be an integrable function with respect to μ. We denote a new complex measure $f\mu$ by $f\mu(A) = \int_A f d\mu$. Let us assume that $f : \Sigma_A^+ \to \mathbb{R}$ and $g : K \to \mathbb{R}$ are Lipschitz functions with respect to d_θ and d respectively. We denote $|f|_\theta$ and $|g|_d$ to their Lipschitz seminorms given by:

$$|f|_\theta = \sup_{\underline{x}, \underline{y} \in \Sigma_A^+} \left\{ \frac{|f(\underline{x}) - f(\underline{y})|}{d_\theta(\underline{x}, \underline{y})} \right\}, \qquad |g|_d = \sup_{z_1, z_2 \in K} \left\{ \frac{|g(z_1) - g(z_2)|}{d(z_1, z_2)} \right\}. \tag{3.1}$$

We endow Σ_A^+ with the Gibbs measure denoted m. Recall that the transfer operator P_σ associated to the map σ and to the measure m, satisfies a Lasota-Yorke inequality when acting on the vector space of Lipschitz functions with norm defined by $\|\cdot\|_\theta = \|\cdot\|_\infty + |\cdot|_\theta$. More specifically, denote:

$$\mathcal{F}_\theta = \left\{ f : \Sigma_A^+ \to \mathbb{R} : |f|_\theta < \infty \right\}.$$

Then, there exists $C > 0$ such that for all $f \in \mathcal{F}_\theta$ and $n \geq 0$:

$$\|P_\sigma^n f\|_\theta \leq \theta^n \|f\|_\theta + C\|f\|_\infty. \tag{3.2}$$

This is in fact the Lasota-Yorke inequality that we obtained in Example 2.1.3. Moreover, as a consequence of I-TM Theorem, the operator P_σ has spectral gap in $\mathcal{F}_\theta$. Namely, $P_\sigma : \mathcal{F}_\theta \to \mathcal{F}_\theta$ can be written as $P_\sigma = Q_\sigma + N_\sigma$ (see Equation (2.7) and the discussion around it), where Q_σ is idempotent, the spectral radius of N_σ is strictly less than one and there exist constants $r < 1$ and $D > 0$ such that for all $f \in \ker Q_\sigma$ and $n \geq 0$:

$$\|P_\sigma^n f\|_\theta \leq D r^n \|f\|_\theta. \tag{3.3}$$

Inequalities (3.2) and (3.3) will be used below on the proof of the spectral properties of operator F^* when acting on a space of signed measures. Given two signed measures μ_1 y μ_2 on a compact metric space, we denote:

$$W_1^0(\mu_1, \mu_2) = \sup\left\{\left|\int g d\mu_1 - \int g d\mu_2\right| : \|g\|_\infty \leq 1, |g|_d \leq 1\right\}.$$

The above expression corresponds to the dual form of the **Wasserstein distance** between the signed measures μ_1 and μ_2. Such distance is connected to the famous **optimal transportation problem** which roughly speaking, consists on finding the optimal way to move a given mass that is originally distributed with μ_1, so that it ends up distributed with μ_2. A complete treatment of the problem can be found in reference [88]. In fact, the Wasserstein distance has been found applications in many areas including Economics [81] and Computer Vision [83].

Let $\mathcal{SB}$ be the space of signed measures on Σ, i.e., $\mu \in \mathcal{SB}$ if μ is a real valued function, defined on the Borel sigma algebra of Σ, $\mu(\varnothing) = 0$ and for every partition $\{E_n\}$ of the measurable set $E \subseteq \Sigma$ we have:

$$\sum_{n=1}^{\infty} \mu(E_n) = \mu\left(\bigcup_{n=1}^{\infty} E_n\right).$$

Given a measure $\mu \in \mathcal{SB}$, we denote the Jordan decomposition of μ by $\mu = \mu^+ - \mu^-$, with:

$$\mu^+ = \frac{1}{2}\left(|\mu| + \mu\right), \qquad \mu^- = \frac{1}{2}\left(|\mu| - \mu\right),$$

where $|\mu|(A) = \sup \sum_{i=1}^{\infty} |\mu(A_i)|$ and the supremum is taken over all partitions of the set A. Now we define the following $[0,1]$-valued measures on Σ:

$$\overline{(\mu^+)}(A) = \begin{cases} \dfrac{\mu^+(A)}{\mu^+(\Sigma)}, & \text{if } \mu^+(\Sigma) \neq 0; \\ 0, & \text{if } \mu^+(\Sigma) = 0. \end{cases} \qquad \overline{(\mu^-)}(A) = \begin{cases} \dfrac{\mu^-(A)}{\mu^-(\Sigma)}, & \text{if } \mu^-(\Sigma) \neq 0; \\ 0, & \text{if } \mu^-(\Sigma) = 0. \end{cases} \tag{3.4}$$

$\mu^\pm(\Sigma) = 0$ if and only if $\mu^\pm = 0$ (see Lemma 4.2.2 in the next chapter), so the above expressions are well defined for all $\mu \in \mathcal{SB}$. We will be interested on measures whose projection on Σ_A^+ is absolutely continuous with respect to the Gibbs measure m. Let:

$$\mathcal{AB} = \left\{\mu \in \mathcal{SB} : \mu^+ \circ \pi_1^{-1} \ll m, \mu^- \circ \pi_1^{-1} \ll m\right\}.$$

For a measure $\mu \in \mathcal{AB}$, we denote ϕ_{μ^+} and ϕ_{μ^-} to the corresponding densities with respect to m, and $\phi_\mu = \phi_{\mu^+} - \phi_{\mu^-}$.

We will use the Rokhlin Disintegration Theorem ([71]) to define a family of signed measures on K. This theorem states that a probability measure μ on a compact metric space can be *desintegrated* with respect to a measurable partition Γ. It means that there exists a family of

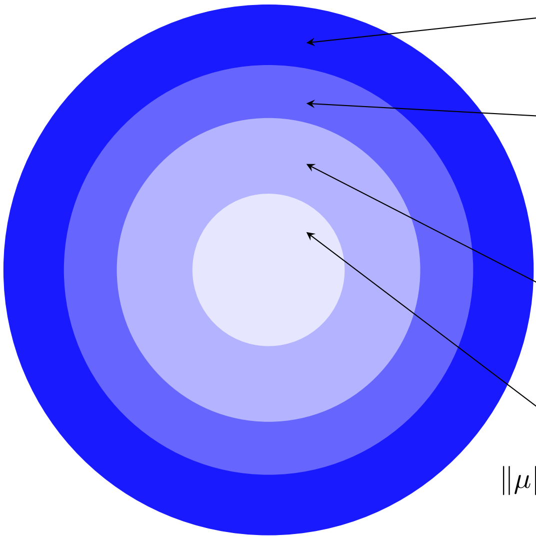


Figure 3.1: Steps on the construction of the space S^∞. We start with the set of all signed measures on Σ, denoted $\mathcal{SB}$. All signed measures can be written as difference of two non-negative measures $\mu = \mu^+ - \mu^-$. Then the space $\mathcal{AB}$ is formed with measures $\mu \in \mathcal{SB}$ such that $\mu^+ \circ \pi_1^{-1} \ll m$ and $\mu^- \circ \pi_1^{-1} \ll m$. Then the set $\mathcal{L}^\infty$ is defined as the set of measures where $\|\mu\|_\infty < \infty$ and finally, the set S^∞ is defined as the set of measures such that $\|\mu\|_{S^\infty} < \infty$.

probability measures $\{\mu_\gamma\}_{\gamma\in\Gamma}$ such that the original measure μ can be obtained by integrating μ_γ (for a precise statement see Theorem 5.1.11 from [71]). Moreover, if the sigma-algebra admits a countable generator, then the disintegration is unique (Proposition 5.1.7 from [71]).

For $\mu \in \mathcal{AB}$ and $\gamma \in \Sigma_A^+$, the disintegration theorem 7.3.1 induces a family of signed measures on K defined by:

$$\begin{aligned} \overline{(\mu^+)}|_\gamma &= \phi_{\mu^+}(\gamma) \cdot \overline{(\mu^+)}_\gamma \circ \pi_2^{-1}, \\ \overline{(\mu^-)}|_\gamma &= \phi_{\mu^-}(\gamma) \cdot \overline{(\mu^-)}_\gamma \circ \pi_2^{-1}, \\ \mu|_\gamma &= \overline{(\mu^+)}|_\gamma - \overline{(\mu^-)}|_\gamma. \end{aligned} \tag{3.5}$$

The set of measures $\mu|_\gamma$ with $\mu \in \mathcal{AB}$ and $\gamma \in \Sigma_A^+$ defines a vector space in which we now define a norm given by:

$$\|\mu|_\gamma\|_W = W_1^0(\mu|_\gamma, 0) = \sup\left\{\left|\int gd\,(\mu|_\gamma)\right| : \|g\|_\infty \le 1, |g|_d \le 1\right\}.$$

We define $\mathcal{L}^\infty$ and $\|\cdot\|_\infty$ by:

$$\mathcal{L}^\infty = \left\{\mu \in \mathcal{AB} : \operatorname*{ess\,sup}_\gamma \{\|\mu|_\gamma\|_W\} < \infty\right\}, \qquad \|\mu\|_\infty = \operatorname*{ess\,sup}_\gamma \{\|\mu|_\gamma\|_W\}.$$

Finally, we define the vector space (see section 7.5) S^∞ and the norm $\|\cdot\|_{S^\infty}$:

$$S^\infty = \{\mu \in \mathcal{AB} : \|\mu\|_{S^\infty} < \infty\}, \qquad \|\mu\|_{S^\infty} = \|\phi_\mu\|_\theta + \|\mu\|_\infty.$$

where $\|\cdot\|_\theta = \|\cdot\|_\infty + |\cdot|_\theta$. Figure 3.1 shows the construction of the space S^∞ graphically.

3.2. Spectral Gap on Signed Measures

Under the setting above, the authors of [28] (see also [61]) show that the operator F^* has spectral gap in S^∞ in the sense of the following theorem:

Theorem 3.2.1 ([28]). *On the setting described above, we have the following:*

1. *There exists a unique $F-$invariant probability measure on S^∞, denoted μ_0.*
2. *Define $P, N : S^\infty \to S^\infty$ by $P\mu = \mu(\Sigma)\mu_0$ and $N\mu = F^*\mu - F^*P\mu$. The operator $F^* : S^\infty \to S^\infty$ can be expressed as $F^* = P + N$.*
3. *The operator P is idempotent (i.e. $P^2 = P$), $PN = NP = 0$ and $\dim(\mathrm{Im}(P)) = 1$.*
4. *There exist constants $0 < \xi < 1$ and $C > 0$ such that, for all $\mu \in S^\infty$:*

$$\|N^n \mu\|_{S^\infty} \leq \|\mu\|_{S^\infty} \xi^n C. \tag{3.6}$$

The rest of this section, is devoted to discuss the main parts of the proof presented in [28].

3.2.1 Disintegration for $F^*\mu$ and induced family of measures on K

The first step of the proof of Theorem 3.2.1 is to find the corresponding expressions for the disintegration and for the family of measures in K associated to the pushforward measure $F^*\mu$. First, note that when $\mu \in \mathcal{AB}$ we also have that $F^*\mu \in \mathcal{AB}$. Moreover:

$$\phi_{F^*\mu} = P_\sigma \phi_\mu. \tag{3.7}$$

Indeed, using the definition of the transfer operator:

$$P_\sigma(\phi_{\mu^+}) = \frac{d(\phi_{\mu^+} m \circ \sigma^{-1})}{dm},$$

where $\phi_{\mu^+} m(E) = \int_E \phi_{\mu^+} dm = \mu^+ \circ \pi_1^{-1}(E)$, one gets for each measurable set $E \subseteq \Sigma_A^+$:

$$\int_E P_\sigma \phi_{\mu^+} dm = \int_E d(\phi_{\mu^+} m \circ \sigma^{-1}) = \mu^+ \circ \pi_1^{-1} \circ \sigma^{-1}(E) = \mu^+(\sigma^{-1}(E) \times K). \tag{3.8}$$

Also:

$$(F^*\mu)^+ \circ \pi_1^{-1}(E) = \mu^+ \circ F^{-1} \circ \pi_1^{-1}(E) = \mu^+(\sigma^{-1}(E) \times K). \tag{3.9}$$

Similar results are obtained for μ^- and as a consequence, (3.7) is established. The fact that $(F^*\mu)^+ = F^*(\mu^+)$ arises by noting that, if A_i forms a partition of A, then $F^{-1}(A_i)$ forms a partition of $F^{-1}A$. Conversely, if B_i forms a partition of $F^{-1}A$, then $F(B_i)$ forms a partition of A.

Let us denote $\nu = \overline{(F^*\mu)^+}$ just to simplify notation. The next step is to find an expression for the disintegration associated to ν. Such disintegration is the unique family of probability measures on Σ, denoted ν_γ, with $\gamma \in \Sigma_A^+$ that satisfy:

$$\nu(A) = \int_{\Sigma_A^+} \nu_\gamma(A) d(\nu \circ \pi_1^{-1})(\gamma). \tag{3.10}$$

The above is obtained using Equation (7.8) from Disintegration Theorem 7.3.1 with $g = \chi_A$. Recall that Disintegration Theorem 7.3.1 guarantees the existence of a family of probability measures in Σ denoted ν_γ, that satisfies Equation (3.10). Let $B_2 = \left\{\gamma \in \Sigma_A^+ : P_\sigma(\phi_{\mu^+})(\gamma) = 0\right\}$ and $B_3 = B_2^c$ so the above equation can be written as:

$$\nu(A) = \int_{B_3} \nu_\gamma(A) d(\nu \circ \pi_1^{-1})(\gamma) + \int_{B_2} \nu_\gamma(A) d(\nu \circ \pi_1^{-1})(\gamma).$$

Note that the second term is always zero, regardless of the expression for ν_γ. Indeed, with the change of variables $\gamma = \sigma\beta$ and using (3.8) and (3.9) one gets:

$$\begin{aligned}\int_{B_2} \nu_\gamma(A) d(\nu \circ \pi_1^{-1})(\gamma) &= \frac{1}{\mu^+(\Sigma)} \int_{B_2} \nu_\gamma(A) d\left(\mu^+ \circ F^{-1} \circ \pi_1^{-1}\right)(\gamma) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{B_2} \nu_\gamma(A) P_\sigma \phi_{\mu^+}(\gamma) dm(\gamma) = 0 \qquad (3.11) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{B_2} \nu_\gamma(A) d\left(\mu^+ \circ \pi_1^{-1}\sigma^{-1}\right)(\gamma) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{\sigma^{-1}B_2} \nu_{\sigma\beta}(A) d\left(\mu^+ \circ \pi_1^{-1}\right)(\beta) \\ &= \int_{\sigma^{-1}B_2} \nu_{\sigma\beta}(A) d\overline{(\mu^+)} \circ \pi_1^{-1}(\beta). \qquad (3.12)\end{aligned}$$

So, we only need to find ν_γ for $\gamma \in B_3$. It will be convenient to define the map σ restricted to cylinder sets of size one. Let $i \in S$ and define $[i] = \left\{\gamma \in \Sigma_A^+ : \gamma_0 = i\right\}$ and $\sigma_i = \sigma|_{[i]}$. We will now show that for $\gamma \in B_3$:

$$\nu_\gamma = \frac{1}{P_\sigma \phi_{\mu^+}(\gamma)} \sum_{i=1}^{N} \frac{\phi_{\mu^+}}{J_{m,\sigma_i}} \circ \left(\sigma_i^{-1}(\gamma)\right) \cdot \chi_{\sigma[i]}(\gamma) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma} \circ F^{-1},$$

where J_{m,σ_i} is the Jacobian of the transformation σ_i associated to the measure m. See the discussion at the end of Section 1.3 for a precise definition of the Jacobian. It is enough to show that (3.10) is satisfied:

$$\begin{aligned}&\int_{B_3} \frac{1}{P_\sigma \phi_{\mu^+}(\gamma)} \sum_{i=1}^{N} \frac{\phi_{\mu^+}}{J_{m,\sigma_i}} \circ \left(\sigma_i^{-1}(\gamma)\right) \cdot \chi_{\sigma[i]}(\gamma) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma} \circ F^{-1}(A) d\left(\nu \circ \pi_1^{-1}\right)(\gamma) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{B_3} \sum_{i=1}^{N} \frac{\phi_{\mu^+}}{J_{m,\sigma_i}} \circ \left(\sigma_i^{-1}(\gamma)\right) \cdot \chi_{\sigma[i]}(\gamma) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma}(F^{-1}A) dm(\gamma) \\ &= \frac{1}{\mu^+(\Sigma)} \sum_{i=1}^{N} \int_{\sigma[i]\cap B_3} \frac{\phi_{\mu^+}}{J_{m,\sigma_i}} \circ \left(\sigma_i^{-1}(\gamma)\right) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma}(F^{-1}A) dm(\gamma) \\ &= \frac{1}{\mu^+(\Sigma)} \sum_{i=1}^{N} \int_{[i]\cap\sigma_i^{-1}B_3} \phi_{\mu^+}(\beta) \cdot \overline{(\mu^+)}_\beta(F^{-1}A) dm(\beta) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{\sigma^{-1}B_3} \phi_{\mu^+}(\beta) \cdot \overline{(\mu^+)}_\beta(F^{-1}A) dm(\beta) \\ &= \frac{1}{\mu^+(\Sigma)} \int_{\sigma^{-1}B_3} \overline{(\mu^+)}_\beta(F^{-1}A) d\left(\phi_{\mu^+} m\right)(\beta) \\ &= \int_{\sigma^{-1}B_3} \overline{(\mu^+)}_\beta(F^{-1}A) d\overline{(\mu^+)} \circ \pi_1^{-1}(\beta) = \overline{(\mu^+)}\left(F^{-1}A\right) = \overline{(F^*\mu)^+}(A). \qquad (3.13)\end{aligned}$$

Note that we are able to use the Disintegration Theorem again on equation (3.13) because (3.12) guarantees that:

$$\int_{\sigma^{-1}B_2} \overline{(\mu^+)}_\beta(F^{-1}A) d\overline{(\mu^+)} \circ \pi_1^{-1}(\beta) = 0.$$

Now we want to obtain an expression for the family of probability measures induced on K by the pushforward measure and its corresponding density (i.e. an expression for $\nu|_\gamma$). For any measurable set $J \subseteq K$ we have:

$$\begin{aligned}
\overline{((F^*\mu)^+)}|_\gamma(J) &= P_\sigma\phi_{\mu^+}(\gamma) \cdot \overline{((F^*\mu)^+)}_\gamma \cdot \pi_{2,\gamma}^{-1}(J) \\
&= \sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \cdot \phi_{\mu^+}\left(\sigma_i^{-1}(\gamma)\right) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma} \circ F^{-1} \circ \pi_{2,\gamma}^{-1}(J) \\
&= \sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \cdot \phi_{\mu^+}\left(\sigma_i^{-1}(\gamma)\right) \cdot \overline{(\mu^+)}_{\sigma_i^{-1}\gamma} \circ \pi_{2,\sigma_i^{-1}\gamma}^{-1}(J') \\
&= \sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \cdot \overline{(\mu^+)}|_{\sigma_i^{-1}\gamma}(J'),
\end{aligned}$$

where $J' = \pi_{2,\sigma_i^{-1}\gamma} \circ F^{-1} \circ \pi_{2,\sigma_i^{-1}\gamma}^{-1}(J)$. It will be convenient to define $F_\gamma : K \to K$ by $F_\gamma(z) = G(\gamma, z)$ so that:

$$\overline{((F^*\mu)^+)}|_\gamma(J) = \sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \cdot \overline{(\mu^+)}|_{\sigma_i^{-1}\gamma} \circ F_{\sigma_i^{-1}\gamma}^{-1}(J). \tag{3.14}$$

3.2.2 Lasota-Yorke inequality

In this section we will derive inequalities for the norms of the pushforward measure. These inequalities will be used at the end of the section to establish a Lasota-Yorke inequality and the existence of a unique $F-$invariant probability measure on Σ. Let us start by establishing the following inequality:

$$\left\|\mu|_\gamma \circ F_\gamma^{-1}\right\|_W \leq \|\mu|_\gamma\|_W. \tag{3.15}$$

Recall that d is the distance in K. If $g : K \to \mathbb{R}$ satisfies $\|g\|_\infty \leq 1$ and $|g|_d \leq 1$, this is also true for $g \circ F_\gamma$:

$$1 \geq |g|_d \geq \frac{|g(G(\gamma, z_1)) - g(G(\gamma, z_2))|}{d\left(G(\gamma, z_1), G(\gamma, z_2)\right)}.$$

which leads to:

$$|g(G(\gamma, z_1)), g(G(\gamma - z_2))| \leq d\left(G(\gamma, z_1), G(\gamma, z_2)\right) \leq \alpha d\left(z_1, z_2\right).$$

So we use the following property of the integral with respect to a pushforward measure:

$$\int gd\left(\overline{(\mu^+)}|_\gamma \circ F_\gamma^{-1}\right) = \int g \circ F_\gamma d\overline{(\mu^+)}|_\gamma. \tag{3.16}$$

Together with the corresponding one for μ^-, it is clear that:

$$\left|\int gd\left(\mu|_\gamma \circ F_\gamma^{-1}\right)\right| = \left|\int g \circ F_\gamma d\left(\mu|_\gamma\right)\right|. \tag{3.17}$$

By taking the supremum on (3.17) over g, we obtain (3.15). Equations (3.14) and (3.15) imply that the infinity norm can not be expanded by the pushforward measure in the sense that $\|F^*\mu\|_\infty \leq \|\mu\|_\infty$. Indeed, if we set $c(\gamma) = \|\mu|_\gamma\|_W$, we have the following:

$$\begin{aligned}
\|F^*\mu\|_\infty &= \operatorname*{ess\,sup}_\gamma \left\{\|(F^*\mu)\,|_\gamma\|_W\right\} = \operatorname*{ess\,sup}_\gamma \left\{\left\|\sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \cdot \mu|_{\sigma_i^{-1}\gamma} \circ F_{\sigma_i^{-1}\gamma}^{-1}\right\|_W\right\} \\
&\leq \operatorname*{ess\,sup} \left\{\sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \left\|\mu|_{\sigma_i^{-1}\gamma} \circ F_{\sigma_i^{-1}\gamma}^{-1}\right\|_W\right\} \\
&\leq \operatorname*{ess\,sup} \left\{\sum_{i=1}^{N} \frac{\chi_{\sigma[i]}(\gamma)}{J_{m,\sigma_i}\left(\sigma_i^{-1}(\gamma)\right)} \left\|\mu|_{\sigma_i^{-1}\gamma}\right\|_W\right\} \\
&\leq \|P_\sigma(c)\|_\infty \leq \|c\|_\infty = \|\mu\|_\infty.
\end{aligned} \tag{3.18}$$

Inequalities similar to (3.15) and (3.18) will be needed to obtain a Lasota-Yorke inequality. In order to obtain an inequality similar to (3.15). Let μ be a signed measure in K and $g : K \to \mathbb{R}$ a function satisfying $\|g\|_\infty \leq 1$ and $|g|_d \leq 1$. Let $z_1 \in K$ such that $g \circ F_\gamma(z_1) \leq 1$ and set $u = g \circ F_\gamma(z_1)$. Then, for any $z \in K$:

$$|g \circ F_\gamma(z) - u| \leq \alpha d(z, z_1) \leq \alpha D(K),$$

where $D(K)$ is the diameter of K. This implies that $\left\| \frac{g \circ F_\gamma - u}{\alpha D(K)} \right\|_\infty \leq 1$. Also, for $z_1, z_2 \in K$:

$$\left| \frac{g(F_\gamma(z_1)) - u}{D(K)} - \frac{g(F_\gamma(z_2)) - u}{D(K)} \right| \leq \frac{\alpha d(z_1, z_2)}{D(K)} \leq 1.$$

Therefore $\left| \frac{g \circ F_\gamma - u}{\alpha D(K)} \right|_d \leq 1$. Finally:

$$\begin{aligned} \left| \int gd\left(\mu \circ F_\gamma\right) \right| &= \left| \int g \circ F_\gamma d\mu \right| \leq \left| \int (g \circ F_\gamma - u) d\mu \right| + \left| \int u d\mu \right| \\ &= \alpha D(K) \left| \int \frac{g \circ F_\gamma - u}{\alpha D(K)} d\mu \right| + u|\mu(K)|. \end{aligned}$$

By taking the supremum over g we obtain $\|\mu \circ F_\gamma\|_W \leq \alpha D(K) \|\mu\|_W + |\mu(K)|$. Now we can use exactly the same process that was used to get inequality (3.18) in order to obtain:

$$\|F^* \mu\|_\infty \leq \alpha D(K) \|\mu\|_\infty + \|\phi_\mu\|_\infty.$$

If $\alpha D(K) < 1$, the norm of the iterations is bounded by:

$$\|F^{*n} \mu\|_\infty \leq \alpha^n (D(K))^n \|\mu\|_\infty + \frac{1}{1 - \alpha D(K)} \|\phi_\mu\|_\infty,$$

then a Lasota-Yorke inequality follows directly:

$$\begin{aligned} \|F^{*n} \mu\|_{S^\infty} &= \|F^{*n} \mu\|_\infty + \|P_\sigma^n \phi_\mu\|_\theta \\ &\leq \alpha^n (D(K))^n \|\mu\|_\infty + \frac{1}{1 - \alpha D(K)} \|\phi_\mu\|_\infty + \theta^n \|\phi_\mu\|_\theta + C_2 \|\phi_\mu\|_\infty \qquad (3.19) \\ &\leq 2\alpha_1^n \|\mu\|_{S^\infty} + B_4 \|\mu\|_\infty, \qquad (3.20) \end{aligned}$$

with $\alpha_1 = \max\{\alpha D(K), \theta\}$ and $B_4 = \frac{1}{1 - \alpha D(K)} + C_2$. The last ingredient to establish the uniqueness of the $F-$invariant measure is the convergence to zero of $\|F^{*n}\|_\infty$ for measures that belong to $\mathcal{V}$, where:

$$\mathcal{V} = \{\mu \in S^\infty : Q_\sigma \phi_\mu = 0\}.$$

This is proved as follows. As $\phi_\mu \in \ker(Q_\sigma)$, we have that $\|P_\sigma \phi_\mu\|_\theta \leq Dr^n \|\phi_\mu\|_\theta \leq Dr^n \|\mu\|_{S^\infty}$. For a given $n \geq 1$, let $b \in \{0, 1\}$ and $l \geq 0$ such that $n = 2l + b$. Then:

$$\begin{aligned} \|F^{*n} \mu\|_\infty &= \left\| F^{*2l+b} \mu \right\|_\infty \leq (\alpha D(K))^l \left\| F^{*l+b} \mu \right\|_\infty + (B_4 - C_2) \left\| P_\sigma^l \phi_\mu \right\|_\infty \\ &\leq (\alpha D(K))^l \|\mu\|_\infty + (B_4 - C_2) \left\| P_\sigma^l \phi_\mu \right\|_\theta \\ &\leq (\alpha D(K))^l \|\mu\|_\infty + (B_4 - C_2) Dr^l \|\phi_\mu\|_{S^\infty} \\ &\leq (1 + (B_4 - C_2) D) \beta_1^{-b} \beta_1^n \|\mu\|_{S^\infty} \leq D_2 \beta_1^n \|\mu\|_{S^\infty}, \qquad (3.21) \end{aligned}$$

where $D_2 = \dfrac{1 + (B_4 - C_2) D}{\beta_1}$ and $\beta_1 = \max\left\{(\alpha D(K))^{1/2}, r^{1/2}\right\}$.

3.2.3 Existence and uniqueness of the invariant measure

The existence of an $F-$invariant probability measure on Σ follows from arguments on lifting measures. Specifically, the proof relies on the fact that, for each continuous function $\psi : \Sigma \to \mathbb{R}$, the following limits exist and are equal:

$$\lim_{n\to\infty} \int (\psi \circ F^n)_+ dm = \lim_{n\to\infty} \int (\psi \circ F^n)_- dm,$$

where $\psi_+, \psi_- : \Sigma_A^+ \to \mathbb{R}$ are defined as follows:

$$\psi_+(\underline{x}) = \sup_{\{\underline{x}\}\times K} \psi(\underline{x}, z), \qquad \psi_-(\underline{x}) = \inf_{\{\underline{x}\}\times K} \psi(\underline{x}, z).$$

The functional $\mu : C^0(\Sigma) \to \mathbb{R}$ defined by $\mu(\psi) = \lim_{n\to\infty} \int (\psi \circ F^n)_+ dm$ is a continuous linear functional satisfying $\mu(1) = 1$, $\mu(\psi) \geq 0$ whenever $\psi \geq 0$, $|\mu(\psi)| \leq 1$ for ψ on the unit ball $\|\psi\|_\infty \leq 1$. Consequently, the Riesz Representation Theorem allows to conclude the existence of a probability measure on Σ, denoted μ_0 such that:

$$\mu(\psi) = \int \psi d\mu_0.$$

The measure μ_0 is clearly invariant under F because for each continuous function $\psi : \Sigma \to \mathbb{R}$:

$$\int \psi \circ F d\mu_0 = \mu(\psi \circ F) = \lim_{n\to\infty} \int (\psi \circ F \circ F^n)_+ dm = \mu(\psi) = \int \psi d\mu_0.$$

Suppose now that there exists another probability measure μ_1 on S^∞ that is $F-$invariant. Then $\mu_0(\Sigma) = \mu_1(\Sigma) = 1$ and therefore, $\mu_0 - \mu_1 \in \mathcal{V}$. Equation (3.21) guarantees that $\|F^{*n}(\mu_0 - \mu_1)\|_\infty \leq D_2\beta_1^n \|\mu_0 - \mu_1\|_\infty$. Using the invariance of μ_0 and μ_1 we get:

$$\lim_{n\to\infty} \|F^{*n}\mu_0 - F^{*n}\mu_1\|_\infty = 0$$
$$\lim_{n\to\infty} \|\mu_0 - \mu_1\|_\infty = 0.$$

3.2.4 Spectral gap

Item 1 of Theorem 3.2.1 was proved on the previous section. Assertions 2 and 3 are obtained below. First note that F^* and P commute (i.e. $F^*P = PF^*$):

$$F^*P\mu = F^*(\mu(\Sigma)\mu_0) = \mu(\Sigma)F^*\mu_0 = \mu(\Sigma)\mu_0 = P\mu.$$
$$PF^*\mu = P(\mu \circ F^{-1}) = \mu \circ F^{-1}(\Sigma)\mu_0 = \mu(\Sigma)\mu_0 = P\mu.$$

So $N = F^* - F^*P = F^* - P$. The operator P is idempotent:

$$P^2\mu = P(P\mu) = P(\mu(\Sigma)\mu_0) = \mu(\Sigma)P\mu_0 = \mu(\Sigma)\mu_0(\Sigma)\mu_0 = P\mu.$$

The operators P and N commute (i.e. $PN = NP$) and its composition is zero:

$$NP\mu = N(\mu(\Sigma)\mu_0) = \mu(\Sigma)\left(F^*\mu_0 - P\mu_0\right) = 0,$$
$$PN\mu = N\mu(\Sigma)\mu_0 = \left(F^*\mu(\Sigma) - P\mu(\Sigma)\right)\mu_0 = \mu(\Sigma)\mu_0 - \mu(\Sigma)\mu_0 = 0.$$

It is easy to prove that $\dim(\mathrm{Im}(P)) = 1$, it is enough to show that μ_0 is a base for $\mathrm{Im}(P)$. For each $\mu \in \mathrm{Im}(P)$, there exists $c(\mu) \in \mathbb{R}$ such that $\mu = c(\mu) \cdot \mu_0$. As $\mu \in \mathrm{Im}(P)$, there exists $\nu \in S^\infty$ such that:

$$\mu = P\nu = \nu(\Sigma)\mu_0.$$

So, we can set $c(\mu) = \nu(\Sigma)$. Finally, in order to prove (3.6), let us first consider a measure $\mu \in \mathcal{V}$. For $n \geq 1$, let $b \in \{0, 1\}$ and l such that $n = 2l + b$. Using Lasota-Yorke inequality (3.20) and convergence to zero inequality (3.21) we obtain:

$$\begin{aligned}
\|F^{*n}\mu\|_{S^\infty} &\leq 2\alpha_1^l \left\|F^{*l+b}\mu\right\|_{S^\infty} + B_4 \left\|F^{*l+b}\mu\right\|_\infty \leq 2\alpha_1^l \left(2 + B_4\right) \|\mu\|_{S^\infty} + B_4 \left\|F^{*l}\mu\right\|_\infty \\
&\leq 2\alpha_1^l \left(2 + B_4\right) \|\mu\|_{S^\infty} + B_4 D_2 \beta_1^l \|\mu\|_{S^\infty} \\
&\leq \left[2\left(2 + B_4\right) + B_4 D_2\right] \xi^{-b} \xi^n \|\mu\|_{S^\infty},
\end{aligned}$$

where $\xi = \max\left\{\sqrt{\alpha_1}, \sqrt{\beta_1}\right\}$. Note that $\mu - P\mu \in \mathcal{V}$ and $N^n\mu = F^{*n}\left(\mu - P\mu\right)$. So, by the above inequality we get $\|N^{*n}\mu\|_{S^\infty} \leq K\xi^n \|\mu\|_{S^\infty}$ with $K = \left[2\left(2 + B_4\right) + B_4 D_2\right] \xi^{-b}$.

3.2.5 Decay of correlations

We have already mentioned about decay of correlations in the standard setting of expanding maps in the interval (see Section 1.4.2). Let us now present and discuss the proof for decay of correlations for the Contractive Random Dynamical Systems introduced in this section. It can be shown that the pushforward measure satisfies the following relation for $g \in L_1(\mu_0)$ and $h \in L_\infty$:

$$\int g \cdot (h \circ F) d\mu_0 = \int h dF^*(g\mu_0). \tag{3.22}$$

Also, $P(f\mu_0) = (f\mu_0)(\Sigma)\mu_0 = \mu_0 \cdot \int f d\mu_0$. Having this in mind, suppose that $g : \Sigma_A^+ \to \mathbb{R}$ is Lipschitz and $f : \Sigma_A^+ \to \mathbb{R}$ is such that $f\mu_0 \in S^\infty$. Then:

$$\begin{aligned}
C_{f,g}(n) = \left| \int f \cdot (g \circ F^n) d\mu_0 - \int g d\mu_0 \int f d\mu_0 \right| &= \left| \int g dF^{*^n}(f\mu_0) - \int g dP(f\mu_0) \right| \\
&\leq \max(|g|_\theta, \|g\|_\infty) \left\|F^{*^n}(f\mu_0) - P(f\mu_0)\right\|_W \\
&\leq \|N^n(f\mu_0)\|_{S^\infty} \max(|g|_\theta, \|g\|_\infty) \\
&\leq \max(|g|_\theta, \|g\|_\infty) K\xi^n \|\mu\|_{S^\infty},
\end{aligned}$$

where we have used the fact that $\frac{g}{\max(|g|_\theta, \|g\|_\infty)}$ is a Lipschitz function whose uniform norm and Lipschitz seminorm are both bounded by 1. Recall that $\xi = \max\left\{\sqrt{\alpha_1}, \sqrt{\beta_1}\right\} < 1$, so, $C_{f,g}(n)$ decays exponentially with n.

4. Central Limit Theorem for Contractive RDS

4.1. Statement of the Central Limit Theorem

In this section we state the main result of this thesis, which is a Central Limit Theorem for the class of contractive random dynamical system introduced in Chapter 3. Before presenting this result, let us recall the usual setting of the Spectral Method for Hyperbolic Dynamical Systems and point out some differences with the setting described in the previous chapter.

In the Spectral Method for deterministic dynamical systems, the transfer operator (associated to the dynamics $T : X \to X$ and the measure of interest μ) has spectral gap on a Banach space $\mathcal{B}$ of functions $f : X \to \mathbb{C}$. Then one defines a family of operators $P_{f,t}$ whose iterations are related to the characteristic function φ of the random variable of interest, which is $\frac{S_n f}{\sqrt{n}}$. Then the use of perturbation theory allows to conclude spectral gap of the family of operators and then the convergence of φ to the characteristic function of a normally distributed random variable follows by applying Lévy theorem. In the context of random dynamical systems, one can define an **averaged Perron-Frobenius operator** acting on the same space as $P_{f,t}$ and apply essentially the same method (see for example [2]).

As a reminder, we consider a transformation in the product space Σ given by $F(\underline{x}, z) = (\sigma\underline{x}, G(\underline{x}, z))$ where $G : \Sigma \to K$ satisfies the following conditions:

C1: There exists $0 < \alpha < 1$ such that $d(G(\underline{x}, z_1), G(\underline{x}, z_2)) \leq \alpha d(z_1, z_2)$ for all $\underline{x} \in \Sigma_A^+$ and for all $z_1, z_2 \in K$.

C2: For each $z \in K$ there exists k_z such that $d(G(\underline{x}, z), G(\underline{y}, z)) \leq k_z d_\theta(\underline{x}, \underline{y})$ and $\operatorname{ess\,sup}_{z \in K} k_z < \infty$.

The setting we are considering differs from the standard setting in several ways. First, the operator that we are considering is actually the pushforward measure F^* acting on a normed vector space of signed measures. This pushforward measure can be easily related to the transfer operator P_F associated to the dynamic F and to the unique F-invariant probablity measure on Σ, denoted μ_0. So, the operator that we are considering is not an averaged version of the individual transfer operators, but instead, is the transfer operator acting on the product space Σ.

Before presenting our main result, we shall also specify the class of observables for which our result is satisfied. Recall from Section 3.1 that the signed measure $f\mu_R$ is defined by $f\mu_R(A) = \int_A f d\mu_R$ and, when $f\mu_R \in \mathcal{AB}$ we define:

$$\phi_{f\mu_R} = \phi_{(f\mu_R)^+} - \phi_{(f\mu_R)^-},$$

where $\phi_{(f\mu_R)^\pm}$ are the densities of $(f\mu_R)^\pm \circ \pi_1^{-1}$ with respect to the Gibbs measure m. We will be interested in observables $f : \Sigma \to \mathbb{R}$ that satisfy the following properties:

P1: f is Lipschitz with respect to the distance $d + d_\theta$.

P2: $\phi_{f\mu_R}$ is Lipschitz with respect to d_θ and $\int f d\mu_0 = 0$.

We are now ready to present our main results.

Theorem 4.1.1. *Let (K, d) be a compact metric space and Σ_A^+ the one-sided subshift of finite type associated to the aperiodic matrix A. Denote $\Sigma = \Sigma_A^+ \times K$ and $F : \Sigma \to \Sigma$ given by $F(\underline{x}, z) = (\sigma\underline{x}, G(\underline{x}, z))$, where $G : \Sigma \to K$ satisfies conditions C1 and C2. Let $f : \Sigma \to \mathbb{R}$ a function that satisfies conditions P1, P2 and P3. Denote $S_n f = f + f \circ F + \cdots + f \circ F^{n-1}$. Then there exists $\rho \geq 0$ such that for all $c \in \mathbb{R}$:*

$$\lim_{n\to\infty} \mu_0 \left\{ (\underline{x}, z) : \frac{S_n f(\underline{x}, z)}{\sqrt{n}} \leq c \right\} = N(0, \rho^2)\Big|_0^c. \tag{4.1}$$

Remark 4.1.2. *We remark that this result is obtained following a line of argumentation which, although requires a Lasota-Yorke inequality to hold, it does not require the unit ball of the norm to be relatively compact on the seminorm. In our case, the Lasota-Yorke inequality holds with two norms but the conditions of the Ionescu-Tulcea and Marinescu's theorem are not satisfied.*

Remark 4.1.3. *Although we focus on the case where the base of the random dynamical system is a subshift of finite type Σ_A^+, let us remark that the spectral gap result in [28] only requires that the base is a dynamical system whose the transfer operator has spectral gap itself. So, our result is still valid, provided that the dynamic on the base satisfies (C1) and (C2). In this case the resulting skew-product is not a random dynamical system anymore. See Chapter 5 for a discussion on the non-random selection process with different settings.*

Our next result is a Berry-Esseen inequality which gives us an upper bound for the speed of convergence for the limit given by Equation (4.1) and valid for every n. This bound is obtained using the so-called Berry-Esseen lemma (see for example Lemma XVI.3.2 from [23] or Lemma VI.3 from [39]). The inequality says essentially that the speed of convergence goes like $\frac{1}{\sqrt{n}}$.

Theorem 4.1.4. *Under the same hypothesis of Theorem 4.1.1, if $\rho > 0$, then there exists a constant $D > 0$ such that the following bound for the speed of convergence holds:*

$$\left| \mu_0 \left\{ (\underline{x}, z) : \frac{S_n f(\underline{x}, z)}{\sqrt{n}} \leq c \right\} - \frac{1}{\rho\sqrt{2\pi}} \int_{-\infty}^{c} \exp\left(-\frac{x^2}{2\rho^2}\right) dx \right| \leq \frac{D}{\sqrt{n}}. \tag{4.2}$$

In the rest of this chapter we provide a detailed proof of theorems 4.1.1 and 4.1.4 and at the end of the chapter we also illustrate the results with simple examples.

4.2. Auxiliary propositions and lemmas

We will divide the proof of the main result into several auxiliary propositions and lemmas. First, in Lemma 4.2.2 we clarify how expressions (3.5) are defined when $\mu^{\pm} = 0$. This is needed in order to ensure the consistency of the norm $\|\cdot\|_W$. In Section 4.3 we recall how quasicompactness is implied by Theorem 3.2.1. The quasicompactness of the transfer operator will allow us to apply Perturbation Theorem 7.1.2, which is one of the main tools we use.

The steps that we use to prove Theorem 4.1.1 are as follows. First we extend the setting to a normed space of complex measures and show quasicompactness of the transfer operator on this complex setting (Corollary 4.4.3). Then on section 4.5 we present the properties of the observables we are interested in. Lemmas 4.5.1 and 4.5.2 show that these properties imply a set of inequalities that allow us to conclude that if $\mu \in S^\infty$ then $\|f\mu\|_{S^\infty}$ is bounded in terms of $\|\mu\|_{S^\infty}$. Propositions 4.6.1 and 4.6.2 use this result to conclude that the family of operators presented in Section 4.6 is bounded and well defined. Moreover, Proposition 4.7.1 establishes a relation between the characteristic function of the random variable of interest and the iterations of the transfer operator (in the form of a pushforward measure). Then the proof of the limit in Equation (4.1) follows by applying the standard techniques of [39], in particular, we use Theorem 7.1.2 presented on that reference.

Remark 4.2.1. *According to Equation* (3.4), *when* $\mu^\pm(\Sigma) = 0$, *we have that* $\overline{(\mu^\pm)}$ *is not a probability measure. Therefore, the Desintegration Theorem 7.3.1 cannot be applied and expressions* (3.5) *might seem undefined. The following lemma will allow us to handle that case.*

Lemma 4.2.2. $\mu^+(\Sigma) = 0$ *if and only if* $\mu^+ = 0$ *and* $\mu^-(\Sigma) = 0$ *if and only if* $\mu^- = 0$.

Proof. The *only if* part is obvious and the *if* part follows straightforward from the definitions: suppose $\mu^+(\Sigma) = 0$ and that there exists a measurable set B such that $\mu^+(B) = b > 0$. Given the fact that $\mu^+(\Sigma) = \mu^+(B) + \mu^+(B^c)$, we have $\mu^+(B^c) = -b$, which is a contradiction because μ^+ is a non-negative measure. The same happens with μ^-. □

4.3. Quasicompactness of $F^* : S^\infty \to S^\infty$

Quasicompactness of the operator $F^* : S^\infty \to S^\infty$ will play a crucial role on our result. In this section we recall how Theorem 3.2.1 implies quasicompactness in the sense of Definition 2.1.1. The proof is analogous to Example 2.1.2 that we presented in Section 2.1.1. As a reminder, recall that the operators $F^*, P, N : S^\infty \to S^\infty$ are defined by $F^*\mu = \mu \circ F^{-1}$, $P\mu = \mu(\Sigma) \cdot \mu_0$ and $N\mu = F^*(1-P)\mu$

Remark 4.3.1. *The following statements are a direct consequence of the definitions and will be used in Proposition 4.3.2:*

1. $F^*P = PF^*$ *and* $F^*N = NF^*$.
2. $\mu \in \mathrm{Im}(P) \Rightarrow F^*\mu = P\mu$.
3. $\mu \in \mathrm{Im}(N) \Rightarrow F^*\mu = N\mu$.
4. $\mu_0 \in \mathrm{Im}(P)$ *and* $0 \in \mathrm{Im}(P) \cap \mathrm{Im}(N)$.

Proposition 4.3.2. *The operator* F^* *is quasicompact in* S^∞.

Proof. Define $M' = \mathrm{Im}(P)$ y $M'' = \mathrm{Im}(1-P)$. It is straightforward to show that M' and M'' are closed subspaces of S^∞ and that $S^\infty = M' \oplus M''$. Using Remark 4.3.1 and the fact that P satisfies $P^2 = P$, it follows that M' and M'' are invariant under F^*.

Now, we will show that $\mathrm{spr}(F^*) = 1$. Using Proposition 5.7 from [61], one can see that the operator F^{*n} is bounded with the operator norm given by $\|F^{*n}\|_{S^\infty} \leq 2\alpha^n + B_4$. Indeed:

$$\|F^{*n}\mu\|_{S^\infty} \leq 2\alpha^n\|\mu\|_{S^\infty} + B_4\|\mu\|_\infty \leq 2\alpha^n\|\mu\|_{S^\infty} + B_4\|\mu\|_{S^\infty} \leq (2\alpha^n + B_4)\,\|\mu\|_{S^\infty}. \tag{4.3}$$

By definition of spectral radius, one has that,

$$\mathrm{spr}(F^*) = \lim_{n\to\infty} \|F^{*n}\|_{S^\infty}^{1/n} \leq \lim_{n\to\infty} (2\alpha^n + B_4)^{1/n} = 1,$$

since $\alpha < 1$. By the first item of Theorem 3.2.1, we have that $1 \in \mathrm{spec}(F^*)$ and therefore $\mathrm{spr}(F^*) = 1$. Next, in order to show that $\mathrm{spr}(F^*|_{M''}) < 1$, note that the spectral radius of the operator $F^*|_{M''} = 1 - P = N$ can be written as:

$$\mathrm{spr}(N) = \inf_n \|N^n\|_{S^\infty}^{1/n}.$$

Item 4 of Theorem 3.2.1 implies that $\|N^n\|_{S^\infty} \leq C\xi^n$, since $C > 0$ and $\xi < 1$, one has that:

$$\mathrm{spr}(N) = \inf_n \|N^n\|_{S^\infty}^{1/n} \leq \inf_n (C\xi^n)^{1/n} = \inf_n C^{1/n}\xi = \xi \inf_n C^{1/n} = \xi \lim_{n\to\infty} C^{1/n} = \xi < 1.$$

Item 3 of Theorem 3.2.1 guarantees the dimension of the subspace M' is finite, so the only thing left to show is that each eigenvalue of the operator $F^*|_{M'}$ has magnitude $\mathrm{spr}(F^*) = 1$. Suppose λ is an eigenvalue of P. Then $P\mu = \lambda\mu$ which is equivalent to $P^2\mu = \lambda P\mu$. As P is idempotent, we have $P\mu = \lambda P\mu$ which means $\lambda = 1$. This shows, not only that each eigenvalue of P has magnitude 1, but also that 1 is the only eigenvalue of P. This completes the proof of Proposition 4.3.2. □

Remark 4.3.3. *The quasicompactness of the operator F^* implies that the spectrum of F^* can be expressed as* $\mathrm{spec}(F^*) = \mathrm{spec}(P) \cup \mathrm{spec}(N)$*, or equivalently* $\mathrm{spec}(F^*) = Z \cup \{1\}$ *where Z is contained in the interior of a disc of radius less than* 1.

4.4. Extension to complex measures

In this section we show that Theorem 3.2.1 can be extended to a subset of the space of complex measures in Σ. This extension is needed in order to apply the Spectral Method to obtain limit theorems. The reason is the following. The characteristic function of the random variable of interest has to be expressed in terms of the iterations of the perturbed operator (that will be introduced on the next section). In order to achieve this, the perturbed operator must be defined as the action of the original operator acting on a vector that is perturbed in a complex-exponential manner. This means the original operator must be defined on a complex vector space. In our case, the vector space is formed with measures, so, as stated above, we need to extend the quasicompact action of the operator F^* to include complex measures.

Let us start by recalling the definition of a complex measure:

Definition 4.4.1. *A complex measure μ on a measure space (X, Σ) is a function $\mu : \Sigma \to \mathbb{C}$ that satisfies the following: for every partition $\{E_i\}$ of the measurable set $E \subseteq \Sigma$:*

$$\mu\,(E) = \sum_{i=1}^{\infty} \mu(E_i).$$

Given two signed measures on Σ, μ_R and μ_I, we can define a complex measure μ in the following way:

$$\mu(E) = \mu_R(E) + i\mu_I(E).$$

for every measurable set E. Let us denote:

$$\mathcal{S}^\infty = \{\mu = \mu_R + i\mu_I : \mu_R \in S^\infty,\ \mu_I \in S^\infty\}\,.$$

$\mathcal{S}^\infty$ is a complex vector space that can be converted into a normed vector space using the following norm:

$$\|\mu\|_{\mathcal{S}^\infty} = \sup_{\kappa\in[-2\pi,2\pi]} \|\cos(\kappa)\cdot\mu_R + \sin(\kappa)\cdot\mu_I\|_{S^\infty}.$$

It is straightforward to verify that the following inequality holds: $\|\mu_R\|_{S^\infty} + \|\mu_I\|_{S^\infty} \leq 2\|\mu\|_{\mathcal{S}^\infty}$. This will be used below to prove Proposition 4.6.2. Let us now define complex extensions of the operators F^*, P y N

$$\begin{aligned}
\mathcal{F}^*\mu &= (\mu_R + i\mu_I)\circ F^{-1} = F^*\mu_R + iF^*\mu_I,\\
\mathcal{P}\mu &= P\mu_R + iP\mu_I,\\
\mathcal{N}\mu &= N\mu_R + iN\mu_I.
\end{aligned}$$

The following corollaries show properties of the operator $\mathcal{F}^*$, that are basically inherited from the original operator F^*. These properties will be useful later.

Corollary 4.4.2. $\mathcal{F}^* : \mathcal{S}^\infty \to \mathcal{S}^\infty$ *is a bounded linear operator on* $\mathcal{S}^\infty$.

Proof. Building on Proposition 5.7 from [61], we can see that $\|F^*\mu\|_{S^\infty} \leq (2\alpha + B_4)\|\mu\|_{S^\infty}$ Therefore:

$$\begin{aligned}\|\mathcal{F}^*\mu\|_{\mathcal{S}^\infty} &= \sup_\kappa \|\cos(\kappa)\cdot F^*\mu_R + \sin(\kappa)\cdot F^*\mu_I\|_{S^\infty} \\ &= \sup_\kappa \|F^*\left(\cos(\kappa)\cdot\mu_R + \sin(\kappa)\cdot\mu_I\right)\|_{S^\infty} \\ &\leq (2\alpha + B_4)\sup_\kappa \|\cos(\kappa)\cdot\mu_R + \sin(\kappa)\cdot\mu_I\|_{S^\infty} \\ &= (2\alpha + B_4)\|\mu\|_{\mathcal{S}^\infty}.\end{aligned}$$

□

Corollary 4.4.3. *The operator $\mathcal{F}^*$ is quasicompact in $\mathcal{S}^\infty$.*

Proof. It is enough to show that $\mathcal{F}^*$, $\mathcal{P}$ and $\mathcal{N}$ satisfy properties analogous to those of the original operators F^*, P and N. As μ_0 is the only probability measure invariant under F, and since all probability measures on $\mathcal{S}^\infty$ are elements of S^∞, we have that μ_0 is also the only probability measure on $\mathcal{S}^\infty$ invariant under F. It is straightforward to verify that $\mathcal{F}^* = \mathcal{P} + \mathcal{N}$, $\mathcal{P}^2 = \mathcal{P}$ and $\mathcal{P}\mathcal{N} = \mathcal{N}\mathcal{P} = 0$:

$$\begin{aligned}\mathcal{P}\mu + \mathcal{N}\mu &= P\mu_R + iP\mu_I + N\mu_R + iN\mu_I = (P+N)\mu_R + i(P+N)\mu_I = F^*\mu_R + iF^*\mu_I = \mathcal{F}^*\mu \\ \mathcal{P}^2\mu &= \mathcal{P}\mathcal{P}\mu = \mathcal{P}(P\mu_R + iP\mu_I) = PP\mu_R + PP\mu_I = P^2\mu_R + P^2\mu_I = P\mu_R + p\mu_I = \mathcal{P}\mu \\ \mathcal{P}\mathcal{N}\mu &= \mathcal{P}(N\mu_R + iN\mu_I) = PN\mu_R + iPN\mu_I = 0 \\ \mathcal{N}\mathcal{P}\mu &= \mathcal{N}\left(P\mu_R + iP\mu_I\right) = NP\mu_R + iNP\mu_I = 0.\end{aligned}$$

In order to prove that $\dim(\mathrm{Im}(\mathcal{P})) = 1$ it is enough to show that, for each $\mu \in \mathrm{Im}(\mathcal{P})$, there exist $z \in \mathbb{C}$ and $\nu \in \mathrm{Im}(\mathcal{P})$ (ν independent of μ), such that $\mu = z\nu$. Indeed, as consequence of the operator $\mathcal{P}$ being idempotent, $\mathcal{P}\mu = \mu$ and therefore:

$$\mu = \mathcal{P}\mu = P\mu_R + iP\mu_I = \mu_R(\Sigma)\mu_0 + i\mu_I(\Sigma)\mu_0 = (\mu_R(\Sigma) + i\mu_I(\Sigma))\mu_0.$$

Finally, in order to show the norm $\|\cdot\|_{\mathcal{S}^\infty}$ is contracted by the operator $\mathcal{N}$, we note that:

$$\begin{aligned}\|\mathcal{N}^n\mu\|_{\mathcal{S}^\infty} &= \sup_\kappa \|\cos(\kappa)\cdot N^n\mu_R + \sin(\kappa)\cdot N^n\mu_I\|_{S^\infty} \\ &= \sup_\kappa \|N^n\left(\cos(\kappa)\cdot\mu_R + \sin(\kappa)\cdot\mu_I\right)\|_{S^\infty} \\ &\leq \xi^n K \sup_\kappa \|\cos(\kappa)\cdot\mu_R + \sin(\kappa)\cdot\mu_I\|_{S^\infty} \\ &= \xi^n K \|\mu\|_{\mathcal{S}^\infty}.\end{aligned}$$

□

4.5. Properties of the observable

Recall from Section 3.1 that the signed measure $f\mu_R$ is defined by $f\mu_R(A) = \int_A f d\mu_R$ and, when $f\mu_R \in \mathcal{AB}$, we define:

$$\phi_{f\mu_R} = \phi_{(f\mu_R)^+} - \phi_{(f\mu_R)^-},$$

where $\phi_{(f\mu_R)^\pm}$ are the densities of $(f\mu_R)^\pm \circ \pi_1^{-1}$ with respect to the Gibbs measure m. From now on, we will be interested in observables $f : \Sigma \to \mathbb{R}$ that satisfy properties P1: and P2: defined in Section 4.1.

The following lemma shows that for any bounded measurable observable f and for any $\mu_R \in \mathcal{AB}$, the measure $f\mu_R$ belongs to $\mathcal{AB}$. Moreover, an expression for the density $\phi_{f\mu_R}$ is obtained using Disintegration Theorem 7.3.1.

Lemma 4.5.1. *Let $\mu_R \in \mathcal{AB}$ and $f : \Sigma \to \mathbb{R}$ be a bounded measurable function. Denote:*

$$\psi_f^+(\gamma) = \int_K f(\gamma, z) d\overline{(\mu_R^+)}|_\gamma(z), \qquad \psi_f^-(\gamma) = \int_K f(\gamma, z) d\overline{(\mu_R^-)}|_\gamma(z).$$

Then $f\mu_R \in \mathcal{AB}$ and:

$$\phi_{f\mu_R}(\gamma) = \psi_f^+(\gamma) - \psi_f^-(\gamma), \qquad m\text{-a.e.}$$

Proof. Let $E \subset \Sigma_A^+$ and define $A = \pi_1^{-1}E$. It is enough to prove:

$$f\mu_R(A) = \int_E \Big(\psi_f^+(\gamma) - \psi_f^-(\gamma)\Big) dm(\gamma).$$

Below we use Equation (7.8) from the Disintegration Theorem 7.3.1:

$$\begin{aligned}
\int_E \psi_f^+(\gamma) dm(\gamma) &= \int_E \int_K f(\gamma, z) d\overline{(\mu_R^+)}|_\gamma(z) dm(\gamma) \\
&= \mu_R^+(\Sigma) \int_E \int_K f(\gamma, z) d(\overline{\mu_R^+})_\gamma \circ \pi_2^{-1}(z) d(\overline{\mu_R^+}) \circ \pi_1^{-1}(\gamma) \qquad (4.4)\\
&= \mu_R^+(\Sigma) \int_E \int_A f(\gamma, z) d(\overline{\mu_R^+})_\gamma d(\overline{\mu_R^+}) \circ \pi_1^{-1}(\gamma) = \mu_R^+(\Sigma) \int_A f(\gamma, z) d(\overline{\mu_R^+})(\gamma, z) \\
&= \int_A f(\gamma, z) d\mu_R^+(\gamma, z). \qquad (4.5)
\end{aligned}$$

Similarly we get:

$$\int_E \psi_f^-(\gamma) dm(\gamma) = \int_A f(\gamma, z) d\mu_R^-(\gamma, z). \tag{4.6}$$

Therefore:

$$\begin{aligned}
\int_E \Big(\psi_f^+(\gamma) - \psi_f^-(\gamma)\Big) dm(\gamma) &= \int_A f(\gamma, z) d\mu_R^+(\gamma, z) - \int_A f(\gamma, z) d\mu_R^-(\gamma, z) \\
&= \int_A f(\gamma, z) d\mu_R(\gamma, z) = f\mu_R(A).
\end{aligned}$$

This completes the proof. $\square$

The following lemma shows that for an observable satisfying property P1, both, $\|f\mu_R\|_\infty$ and $\|\phi_{f\mu_R}\|_\infty$ are finite and bounded by the same expression.

Lemma 4.5.2. *Let $\mu_R \in \mathcal{AB}$ and $f : \Sigma \to \mathbb{R}$ be a measurable function satisfying property P1. Then there exists $D_2(f) \geq 0$ such that:*

$$\|f\mu_R\|_\infty \leq D_2(f) \|\mu_R\|_\infty, \tag{4.7}$$
$$\|\phi_{f\mu_R}\|_\infty \leq D_2(f) \|\mu_R\|_\infty. \tag{4.8}$$

Moreover, if f satisfies P2 then the following inequality also holds:

$$\|f\mu_R\|_{S^\infty} \leq 4D_2(f) \|\mu_R\|_{S^\infty}. \tag{4.9}$$

Proof. Let $h : K \to \mathbb{R}$ with $\|h\|_\infty \leq 1$ and $|h|_d \leq 1$ and set $g : \Sigma \to \mathbb{R}$ by:

$$g(\gamma, z) = \frac{h(z) f(\gamma, z)}{2 \max(\|f\|_\infty, |f|_{d+d_\theta})}.$$

According to item ii) of Proposition 7.5.2 from the Appendix, if $\mu \in \mathcal{AB}$, then for all $c \in \mathbb{R}$, $c\mu \in \mathcal{AB}$ and $\phi_{c\mu} = c\phi_\mu$. We can use the above with $\mu = hf\mu_R$ and $c = \frac{1}{2\max(\|f\|_\infty, |f|_{d+d_\theta})}$ and obtain:

$$\phi_{hf\mu_R}(\gamma) = D_2(f)\phi_{g\mu_R}(\gamma), \tag{4.10}$$

where $D_2(f) = 2\max(\|f\|_\infty, |f|_{d+d_\theta})$. Note that for each $\gamma \in \Sigma_A^+$, $g_\gamma : K \to \mathbb{R}$ defined by $g_\gamma(z) = g(\gamma, z)$ satisfies $\|g_\gamma\|_\infty \leq 1$ and $|g_\gamma|_d \leq 1$. Using Lemma 4.5.1, we can write equation (4.10) as:

$$\begin{aligned}\left|\int_K hd\overline{(f\mu_R)^+}|_\gamma - \int_K hd\overline{(f\mu_R)^-}|_\gamma\right| &= D_2(f)\left|\int_K g_\gamma d\overline{(\mu_R^+)}|_\gamma - \int_K g_\gamma d\overline{(\mu_R^-)}|_\gamma\right| \\ &\leq D_2(f)\|\mu_R|_\gamma\|_W \leq D_2(f)\|\mu_R\|_\infty. \end{aligned}\tag{4.11}$$

Inequality (4.7) follows by taking supremums on inequality (4.11) (we first take the supremum over all h with $\|h\|_\infty \leq 1$ and $|h|_d \leq 1$ and then we take the supremum over all $\gamma \in \Sigma_A^+$). Inequality (4.8) follows from (4.11) with $h = 1$.

Now suppose that f satisfies P2. To prove Inequality (4.9) note that in this case, $|\phi_{f\mu_R}|_\theta \leq 2\|\phi_{f\mu_R}\|_\infty$ and Inequalities (4.7) and (4.8) imply that:

$$\begin{aligned}\|f\mu_R\|_{S^\infty} = \|f\mu_R\|_\infty + \|\phi_{f\mu_R}\|_\infty + |\phi_{f\mu_R}|_\theta &\leq 2D_2(f)\|\mu_R\|_\infty + 2\|\phi_{f\mu_R}\|_\infty \\ &\leq 2D_2(f)\|\mu_R\|_\infty + 2D_2(f)\|\mu_R\|_\infty \\ &\leq 4D_2(f)\|\mu_R\|_{S^\infty}.\end{aligned}$$

□

Remark 4.5.3. *If $\mu_R \in S^\infty$, then the observables of the form $f(\gamma, z) = f_1(\gamma)$ satisfying P1, will also satisfy P2. In this case:*

$$\begin{aligned}\phi_{f\mu_R}(\gamma) &= \int_K f_1(\gamma)d\overline{(\mu_R^+)}|_\gamma(z) - \int_K f_1(\gamma)d\overline{(\mu_R^-)}|_\gamma(z) \\ &= f_1(\gamma)\int_K d\overline{(\mu_R^+)}|_\gamma(z) - f_1(\gamma)\int_K d\overline{(\mu_R^-)}|_\gamma(z) \\ &= f_1(\gamma)\phi_{\mu_R}(\gamma).\end{aligned}$$

Therefore, $\phi_{f\mu_R}$ is Lipschitz with respect to d_θ. Moreover, we can also obtain a bound for its Lipschitz norm in terms of the Lipschitz norm of ϕ_{μ_R}:

$$\begin{aligned}\|\phi_{f\mu_R}\|_\theta = \|\phi_{f\mu_R}\|_\infty + |\phi_{f\mu_R}|_\theta &\leq \|f_1\|_\infty\|\phi_{\mu_R}\|_\infty + \|f_1\|_\infty|\phi_{\mu_R}|_\theta + \|\phi_{\mu_R}\|_\infty|f_1|_\theta \\ &\leq \|f_1\|_\infty\left(\|\phi_{\mu_R}\|_\infty + |\phi_{\mu_R}|_\theta\right) + 2\|f_1\|_\infty\|\phi_{\mu_R}\|_\infty \\ &\leq 3\|f_1\|_\infty\|\phi_{\mu_R}\|_\theta.\end{aligned}$$

Remark 4.5.4. *If μ_R has Lipschitz regularity in the sense that:*

$$|\mu_R|_\theta = \sup_{\gamma_1 \neq \gamma_2}\left\{\frac{\|\mu_R|_{\gamma_1} - \mu_R|_{\gamma_2}\|_W}{d_\theta(\gamma_1, \gamma_2)}\right\} < \infty.$$

Then $\phi_{f\mu_R}$ is automatically Lipschitz because:

$$\begin{aligned}|\phi_{f\mu_R}(\gamma_1) - \phi_{f\mu_R}(\gamma_2)| &= \left|\int fd\overline{\mu_R^+}|_{\gamma_1} - \int fd\overline{\mu_R^-}|_{\gamma_1} - \int fd\overline{\mu_R^+}|_{\gamma_2} + \int fd\overline{\mu_R^-}|_{\gamma_2}\right| \\ &= \left|\int fd\left(\overline{\mu_R^+}|_{\gamma_1} - \overline{\mu_R^-}|_{\gamma_1} - \overline{\mu_R^+}|_{\gamma_2} + \overline{\mu_R^-}|_{\gamma_2}\right)\right| \\ &= \left|\int fd\left(\mu_R|_{\gamma_1} - \mu_R|_{\gamma_2}\right)\right| \\ &\leq \max\left\{\|f\|_\infty, |f|_{d+d_\theta}\right\}\|\mu_R|_{\gamma_1} - \mu_R|_{\gamma_2}\|_W.\end{aligned}$$

Therefore:

$$\sup_{\gamma_1 \neq \gamma_2} \left\{ \frac{|\phi_{f\mu_R}(\gamma_1) - \phi_{f\mu_R}(\gamma_2)|}{d_\theta(\gamma_1, \gamma_2)} \right\} \leq \max \left\{ \|f\|_\infty, |f|_{d+d_\theta} \right\} |\mu_R|_\theta < \infty.$$

Lipschitz regularity of the invariant measure is also proved on reference [61] (see Theorem 7.10), so if $\mu_R \in span\{\mu_0\}$ *condition P2 is automatically satisfied.*

4.6. Family of operators $\mathcal{F}^*_{f,t}$

Let us define a family of operators dependent of a real parameter t and an observable $f : \Sigma \to \mathbb{R}$ satisfying P1-P2. We denote this family $\mathcal{F}^*_{f,t}$ and we define it by:

$$\mathcal{F}^*_{f,t}(\mu) = \mathcal{F}^* \left(e^{itf} \mu \right),$$

where $e^{itf}\mu$ is a complex measure defined by:

$$e^{itf}\mu(A) = \int_A e^{itf} d\mu,$$

for all measurable set $A \subseteq \Sigma$.

Proposition 4.6.1. *If* $\mu \in \mathcal{S}^\infty$ *then* $e^{itf}\mu \in \mathcal{S}^\infty$ *and hence* $\mathcal{F}^*_{f,t}\mu \in \mathcal{S}^\infty$

Proof. Note that $e^{itf}\mu = \cos(tf)\mu_R - \sin(tf)\mu_I + i\cos(tf)\mu_I + i\sin(tf)\mu_R$, so, it is enough to show that the signed measures: $\cos(tf)\mu_R$, $\sin(tf)\mu_I$, $\cos(tf)\mu_I$ and $\sin(tf)\mu_R$ are elements of S^∞. We will show that $\cos(tf)\mu_R \in S^\infty$. The proof for the other signed measures is analogous. Note that Inequality (4.9) implies a similar inequality for $\cos(tf)\mu_R$. Indeed, for each $(\gamma, z) \in \Sigma$, the following sum converges:

$$\sum_{n=0}^{\infty} \left| \frac{(-1)^n t^{2n}}{(2n)!} f^{2n}(\gamma, z) \right| \leq \sum_{n=0}^{\infty} \frac{t^{2n}}{(2n)!} \|f\|_\infty^{2n} = \cosh(t \cdot \|f\|_\infty) < \infty.$$

So, using Taylor expansion, and Fubini's theorem we have:

$$(\cos(tf)\mu_R)(A) = \int_A \cos(tf) d\mu_R = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \int_A f^{2n} d\mu_R = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} f^{2n} \mu_R(A).$$

Using inequality $\|f\mu_R\|_{S^\infty} \leq D_2(f)\|\mu_R\|_{S^\infty}$ we get:

$$\begin{aligned} \|\cos(tf)\mu_R\|_{S^\infty} &= \left\| \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} f^{2n} \mu_R \right\|_{S^\infty} \leq \sum_{n=0}^{\infty} \left| \frac{t^{2n}}{(2n)!} \right| \|f^{2n}\mu_R\|_{S^\infty} \\ &\leq \sum_{n=0}^{\infty} \left| \frac{t^{2n}}{(2n)!} \right| D_2^{2n}(f) \|\mu_R\|_{S^\infty} \\ &\leq \cosh(t \cdot D_2(f)) \|\mu_R\|_{S^\infty}. \end{aligned}$$

This implies that $e^{itf}\mu \in \mathcal{S}^\infty$, which in view of Corollary 4.4.2, it follows that $\mathcal{F}^*_{f,t}\mu \in \mathcal{S}^\infty$ □

Proposition 4.6.2. $\mathcal{F}^*_{f,t}$ *is a family of bounded linear operators in* $\mathcal{S}^\infty$.

Proof. As we have already shown that $\mathcal{F}^*$ is a bounded linear operator on $\mathcal{S}^\infty$ and $e^{itf}\mu \in \mathcal{S}^\infty$ when $\mu \in \mathcal{S}^\infty$, to show this proposition it is enough to show that there exists a constant $C > 0$, independent of μ, such that $\left\|e^{itf}\mu\right\|_{\mathcal{S}^\infty} \le C\|\mu\|_{\mathcal{S}^\infty}$. This is a direct consequence of Proposition 4.6.1:

$$\begin{aligned}
\left\|e^{itf}\mu\right\|_{\mathcal{S}^\infty} &\le \|\cos(tf)\mu_R - \sin(tf)\mu_I\|_{S^\infty} + \|\cos(tf)\mu_I + \sin(tf)\mu_R\|_{S^\infty} \\
&\le \|\cos(tf)\mu_R\|_{S^\infty} + \|\sin(tf)\mu_I\|_{S^\infty} + \|\cos(tf)\mu_I\|_{S^\infty} + \|\sin(tf)\mu_R\|_{S^\infty} \\
&\le \cosh(t \cdot D_2(f)) \left(\|\mu_R\|_{S^\infty} + \|\mu_I\|_{S^\infty}\right) \\
&+ \sinh(t \cdot D_2(f)) \left(\|\mu_R\|_{S^\infty} + \|\mu_I\|_{S^\infty}\right) \\
&\le 2\left(\cosh(t \cdot D_2(f)) + \sinh(t \cdot D_2(f))\right) \|\mu\|_{\mathcal{S}^\infty} = 2\exp\left[t \cdot D_2(f)\right] \|\mu\|_{\mathcal{S}^\infty}.
\end{aligned}$$

□

The inequality in Proposition 4.6.2 is enough to prove that each operator $\mathcal{F}^*_{f,t}$ is a bounded linear operator in $\mathcal{S}^\infty$ because the expression $2\exp\left[t \cdot D_2(f)\right]$ does not depend on μ.

4.7. Characteristic function and Perron-Frobenius operator

Now we stablish the relationship between the Perron-Frobenius operator (written as a pushforward measure) and the characteristic function of the random variable of interest. The Perron-Frobenius operator associated to the dynamics $F : \Sigma \to \Sigma$ and to the invariant measure μ_0 is defined in the same way as in (1.15). For a function $g : \Sigma \to \mathbb{C}$, integrable with respect to μ_0, the Perron-Frobenius operator P_F is defined as the Radon-Nykodim derivative of the complex measure $F^*(g\mu_0)$ with respect to μ_0:

$$P_F g = \frac{dF^*(g\mu_0)}{d\mu_0}.$$

Using (1.16), it can be shown the transfer operator satisfies the following equation for $g \in L_1(\mu_0)$ and $h \in L_\infty$:

$$\int g \cdot (h \circ F) d\mu_0 = \int h dF^*(g\mu_0). \tag{4.12}$$

The following lemma relates the iterations of the perturbed operator with the characteristic function of the random variable $S_n f$. The result is an equality that is analogous to (2.16).

Proposition 4.7.1. $\left(\mathcal{F}^*_{f,t}\right)^n \mu_0(\Sigma) = \int e^{itS_n f} d\mu_0$ *where* $S_n f = f + f \circ F + \cdots + f \circ F^{n-1}$.

Proof. By definition, we have $\mathcal{F}^*_{f,t}\mu(\Sigma) = \int e^{itf} d\mu$, therefore:

$$\begin{aligned}
\left(\mathcal{F}^*_{f,t}\right)^2 \mu_0(\Sigma) = \mathcal{F}^*_{f,t}\left(\mathcal{F}^*_{f,t}\mu_0(\Sigma)\right) &= \int e^{itf} d(\mathcal{F}^*(e^{itf}\mu_0)) \\
&= \int e^{itf} d(F^*(e^{itf}\mu_0)_R + iF^*(e^{itf}\mu_0)_I) \\
&= \int e^{itf} dF^*(\cos(tf)\mu_0) + i\int e^{itf} dF^*(\sin(tf)\mu_0) \\
&= \int \cos(tf) \cdot e^{itf} \circ F d\mu_0 + i\int \sin(tf) \cdot e^{itf} \circ F d\mu_0 \\
&= \int e^{itf} \circ F \cdot \left[\cos(tf) + i\sin(tf)\right] d\mu_0 \\
&= \int e^{itf} \cdot e^{itf} \circ F d\mu_0 \\
&= \int e^{itf} \cdot e^{itf\circ F} d\mu_0 = \int e^{it(f + f\circ F)} d\mu_0.
\end{aligned}$$

So, by induction the result follows. □

Proposition 4.7.1 shows that for each t, the iterations of operator $\mathcal{F}^*_{f,t}$ can be expressed in terms of the characteristic function of the random variable $S_n f$ with respect to the probability measure μ_0.

4.8. Proof of the main results

We are now ready to show the Central Limit Theorem with respect to μ_0, which is the unique probability measure invariant under the skew-product $F : \Sigma \to \Sigma$.

Proof of Theorem 4.1.1. We will show the characteristic function of the random variable $\frac{S_n f(x,z)}{\sqrt{n}}$ converges to the characteristic function of a normally distributed random variable. The main tool here is the Perturbation Theorem 7.1.2 from reference [39]. We provide an extended proof of this theorem in Section 7.1.3 in the appendix. Let us now verify that the hypothesis of that theorem holds true. We have already verified that $\mathcal{F}^*_{f,t}$ is a bounded linear operator on $\mathcal{S}^\infty$ (Proposition 4.6.2). Following the same argument as in Section 2.1.2, we obtain that $\mathcal{F}^*_{f,t}$ is of class C^∞. Moreover, $\mathcal{F}^*_{f,0} = \mathcal{F}^*$ has one simple leading eigenvalue (this follows from Lemma 7.1.4) and $\operatorname{spr}(\mathcal{F}^*) = 1$ (Corollary 4.4.3).

We are now able to apply Theorem 7.1.2. Let $t \in I_0$, $\lambda(t)$ be the eigenvalue of $\mathcal{F}^*_{f,t}$ associated to the eigenvector $\mu(t)$ and $\mu^*(t)$ be the corresponding eigenvector on the dual space of $\mathcal{S}^\infty$ (item i). For $\nu \in \mathcal{S}^\infty$, denote $\Pi_t(\nu) = \langle \mu^*(t), \nu \rangle \mu(t)$ and note that $\Pi_0 \mu_0 = \mu_0$. Using item iii) of Theorem 7.1.2 and Proposition 4.7.1 we get:

$$\left(\mathcal{F}^*_{f,t/\sqrt{n}}\right)^n = \left(\lambda\left(t/\sqrt{n}\right)\right)^n \cdot \Pi_{t/\sqrt{n}} + \mathcal{N}^n_{t/\sqrt{n}}, \tag{4.13}$$

with $\mathcal{N}_t = \mathcal{F}^*_{f,t} - \lambda(t)\Pi_t$. As $\lambda(t)$ depends analytically on t, it can be expanded in its Taylor series:

$$\begin{aligned}\left(\mathcal{F}^*_{f,t/\sqrt{n}}\right)^n &= \left(1 + \lambda'(0)\frac{t}{\sqrt{n}} + \lambda''(0)\frac{t^2}{2n} + \cdots\right)^n \cdot \Pi_{t/\sqrt{n}} + \mathcal{N}^n_{t/\sqrt{n}} &(4.14)\\ &= \left(1 - \rho^2\frac{t^2}{2n} + \cdots\right)^n \cdot \Pi_{t/\sqrt{n}} + \mathcal{N}^n_{t/\sqrt{n}}, &(4.15)\end{aligned}$$

where we have used the fact that $\lambda'(0) = 0$ and:

$$\rho^2 = -\lambda''(0) = \int f^2 d\mu_0 + 2\sum_{k=1}^{\infty} \int f \cdot (f \circ F^k) d\mu_0.$$

See Theorem 7.4.1 in the appendix for the standard proof of this statement. Taking limits in Equation (4.15):

$$\lim_{n\to\infty} \left(\mathcal{F}^*_{f,t/\sqrt{n}}\right)^n = \exp\left(-\frac{\rho^2 t^2}{2}\right)\Pi_0.$$

Now we use Proposition (4.7.1) and the fact that $\Pi_0\mu_0 = \mu_0$:

$$\lim_{n\to\infty}\left(\int e^{i\frac{t}{\sqrt{n}}S_n f} d\mu_0\right) = \exp\left(-\frac{\rho^2 t^2}{2}\right)\mu_0(\Sigma) = \exp\left(-\frac{\rho^2 t^2}{2}\right). \tag{4.16}$$

Note the left-hand side of equation (4.16) is the limit of the characteristic function of the random variable $\frac{S_n f}{\sqrt{n}}$, and the right-hand side of the same equation is the characteristic function of a random variable distributed normally with mean 0 and variance ρ^2. By Lévy theorem (see for example Theorem 14.15 from reference [25]) $\frac{S_n f}{\sqrt{n}}$ converges in distribution to a normal distribution with mean 0 and variance ρ^2. □

Let us now get the speed of convergence (4.2) as an application of Lemma 7.2.1. These bounds are commonly known as Berry-Esseen inequalities.

Proof of Theorem 4.1.4. Let $n \geq 1$ and p_1 and p_2 be the probabilities defined by the distribution functions:

$$F_1(u) = \mu_0 \left\{ (\underline{x}, z) \in \Sigma : \frac{S_n f(\underline{x}, z)}{\sqrt{n}} \leq u \right\}, \qquad F_2(u) = \frac{1}{\rho\sqrt{2\pi}} \int_{-\infty}^{u} \exp\left(-\frac{x^2}{2\rho^2} \right) dx. \tag{4.17}$$

Note that F_2 satisfies $\sup_{u\in\mathbb{R}} F_2'(u) = \dfrac{1}{\sqrt{2\pi}}$ and:

$$\tilde{p}_1(t) = \int e^{it\frac{S_n f}{\sqrt{n}}} d\mu_0 = \left(\lambda\left(t/\sqrt{n}\right)\right)^n \cdot \Pi_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma) + \mathcal{N}^n_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma), \qquad \tilde{p}_2(t) = e^{-\rho^2 t^2/2}.$$

Applying Lemma 7.2.1 with $L = \rho\epsilon\sqrt{n}$ we obtain:

$$\begin{aligned}
&\left|\mu_0 \left\{ (\underline{x}, z) \in \Sigma : \frac{S_n f(\underline{x}, z)}{\sqrt{n}} \leq u \right\} - \frac{1}{\rho\sqrt{2\pi}} \int_{-\infty}^{u} \exp\left(-\frac{x^2}{2\rho^2} \right) dx\right| \\
&\leq \frac{1}{\pi} \int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} \frac{1}{|t|} \left| \left(\lambda\left(t/\sqrt{n}\right)\right)^n \cdot \Pi_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma) + \mathcal{N}^n_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma) - e^{-\rho^2 t^2/2} \right| dt + \frac{24}{\rho\pi\epsilon\sqrt{2\pi n}} \\
&\leq \frac{1}{\pi} \int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} \left| \frac{\mathcal{N}^n_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma)}{t} \right| dt + \frac{1}{\pi} \int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} \left| \left(\lambda\left(t/\sqrt{n}\right)\right)^n \right| \left| \frac{\Pi_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma) - 1}{t} \right| dt \\
&+ \frac{1}{\pi} \int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} \left| \frac{\left(\lambda\left(t/\sqrt{n}\right)\right)^n - e^{-\rho^2 t^2/2}}{t} \right| dt + \frac{24}{\rho\pi\epsilon\sqrt{2\pi n}}.
\end{aligned}$$

Let us now find upper bounds for each of the three integrals above. As $t \mapsto \mathcal{N}_t$ is in particular C^1, it is also locally Lipschitz, and therefore, there exists a constant $L_1 \geq 0$ such that for all t in a neighborhood of 0:

$$\left| \mathcal{N}_{t/\sqrt{n}}(\mu_0)(\Sigma) - \mathcal{N}_0(\mu_0)(\Sigma) \right| \leq L_1 \left| \frac{t}{\sqrt{n}} \right|. \tag{4.18}$$

With $t = 0$, Equation (4.13) becomes:

$$1 = \int e^{i0S_n f} d\mu_0 = \left(\mathcal{F}^*_{f,0}\right)^n \mu_0(\Sigma) = \left(\lambda\left(0\right)\right)^n \cdot \Pi_0\left(\mu_0\right)(\Sigma) + \mathcal{N}^n_0\left(\mu_0\right)(\Sigma). \tag{4.19}$$

Recall that $\Pi_0\mu_0(\Sigma) = 1$ which combined with (4.19), results in $\mathcal{N}^n_0(\mu_0)(\Sigma) = 0$ for all n. Therefore:

$$|\mathcal{N}^n_t(\mu_0)(\Sigma)| = |\mathcal{N}^n_t(\mu_0)(\Sigma) - \mathcal{N}^n_0(\mu_0)(\Sigma)| = |(\mathcal{N}_t\mu_0 - \mathcal{N}_0\mu_0)(\Sigma)| \sum_{k=0}^{n-1} \left| \left(\mathcal{N}_0^{n-k-1}\mathcal{N}_t^k \right)(\mu_0)(\Sigma) \right|.$$

So, for small enough t and ϵ, there exists a constant $M_1 \geq 0$ such that [1]:

$$\int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} \left| \frac{\mathcal{N}^n_{t/\sqrt{n}}\left(\mu_0\right)(\Sigma)}{t} \right| dt \leq \frac{2ncr^{n-1}L_1\|\mu_0\|_{S^\infty}}{\sqrt{n}} \int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}} dt \leq 4n\epsilon\rho cr^{n-1}L_1 \leq \frac{M_1}{\sqrt{n}}, \tag{4.20}$$

[1]For small enough ϵ, the inequality $n\epsilon r^{n-1} \leq 1/\sqrt{n}$ holds for all $n \geq 1$.

where we have used the following bound valid for all $\mu \in \mathcal{S}^\infty$ (in particular, valid for $\mathcal{N}_0^n\mu_0(\Sigma)$ and $\mathcal{N}_t^n\mu_0(\Sigma)$):

$$\begin{aligned}
\|\mu\|_{\mathcal{S}^\infty} &\geq \frac{1}{2}\|\mu_R\|_{S^\infty} + \frac{1}{2}\|\mu_I\|_{S^\infty} \geq \frac{1}{2}|\phi_{\mu_R}(\gamma)| + \frac{1}{2}|\phi_{\mu_I}(\gamma)| \\
\|\mu\|_{\mathcal{S}^\infty} &\geq \frac{1}{2}\int |\phi_{\mu_R}(\gamma)|dm(\gamma) + \frac{1}{2}\int |\phi_{\mu_I}(\gamma)|dm(\gamma) \\
&\geq \frac{1}{2}\left|\int \phi_{\mu_R}(\gamma)dm(\gamma)\right| + \frac{1}{2}\left|\int \phi_{\mu_R}(\gamma)dm(\gamma)\right| \\
&= \frac{1}{2}|\mu_R(\Sigma)| + \frac{1}{2}|\mu_I(\Sigma)| \geq \frac{1}{2}|\mu(\Sigma)|.
\end{aligned}$$

For small enough t, we have the inequality $|\lambda(t)| \leq e^{-\rho^2t^2/4}$. This can be seen by the following:

$$\begin{aligned}
|\lambda(t)| &= \left|1 - \frac{\rho^2t^2}{2} + \sum_{k=3}^{\infty}\frac{\lambda^{(k)}(0)}{k!}t^k\right| \leq \left|1 - \frac{\rho^2t^2}{2}\right| + \left|\sum_{k=3}^{\infty}\frac{\lambda^{(k)}(0)}{k!}t^k\right| \\
&\leq 1 - \frac{\rho^2t^2}{2} + \sum_{k=3}^{\infty}\left|t^k\right|\left|\frac{\lambda^{(k)}(0)}{k!}\right| \leq 1 - \frac{\rho^2t^2}{2} + t^2\sum_{k=3}^{\infty}\left|t^{k-2}\right|\left|\frac{\lambda^{(k)}(0)}{k!}\right| \\
&\leq 1 - \frac{\rho^2t^2}{2} + \frac{\rho^2t^2}{4} = 1 - \frac{\rho^2t^2}{4} \leq e^{-\rho^2t^2/4},
\end{aligned}$$

where we have used a range of t such that: (1) $\frac{\rho^2t^2}{2} \leq 1$ and (2) $\sum_{k=3}^{\infty}\left|t^{k-2}\right|\left|\frac{\lambda^k(0)}{k!}\right| \leq \frac{\rho^2}{4}$. The elementary inequality $1+x \leq e^x$ was also used. Also, Π_t satisfies an inequality similar to (4.18): $\left|\Pi_{t/\sqrt{n}}(\mu_0)(\Sigma) - \Pi_0(\mu_0)(\Sigma)\right| \leq L_2\left|\frac{t}{\sqrt{n}}\right|$, therefore:

$$\int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}}\left|\left(\lambda\left(t/\sqrt{n}\right)\right)^n\right|\left|\frac{\Pi_{t/\sqrt{n}}\left(\mu_0\right)\left(\Sigma\right) - 1}{t}\right|dt \leq \frac{M_2}{\sqrt{n}}\int_{-\infty}^{\infty} e^{-\rho^2t^2/4}dt. \tag{4.21}$$

Note that:

$$\begin{aligned}
\left|\lambda(t) - e^{-\rho^2t^2/2}\right| &= \left|1 - \frac{\rho^2t^2}{2} + \mathcal{O}(t^3) - \left(1 - \frac{\rho^2t^2}{2} + \mathcal{O}(t^3)\right)\right| \\
&\leq L_3|t|^3.
\end{aligned}$$

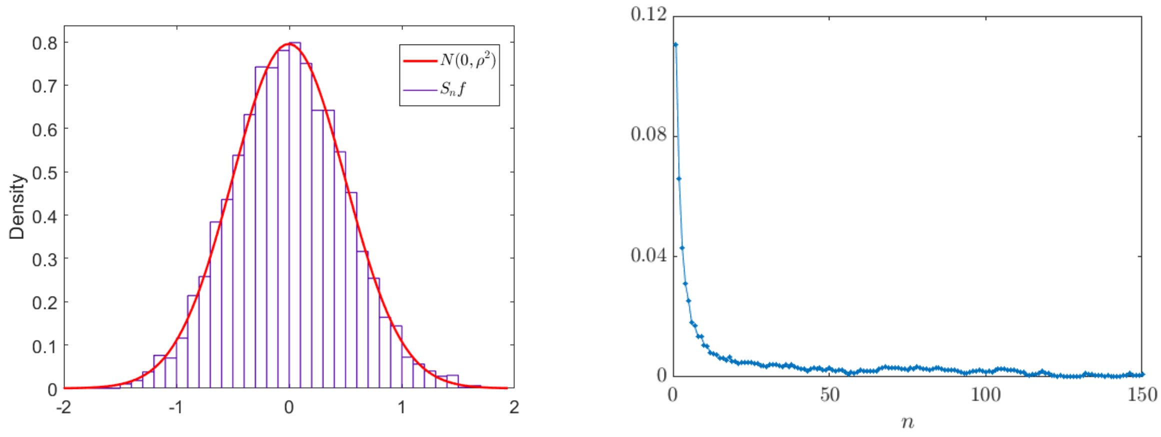
Finally, rescaling t:

$$\begin{aligned}
\left|\left(\lambda\left(t/\sqrt{n}\right)\right)^n - e^{-\rho^2t^2/2}\right| &\leq \left|\sum_{k=0}^{n-1}\left(\lambda(t/\sqrt{n})\right)^{n-k-1}\left(e^{-\rho^2t^2/2n}\right)^k\left(\lambda(t/\sqrt{n}) - e^{-\rho^2t^2/2n}\right)\right| \\
&\leq n\left(e^{-\rho^2t^2/4n}\right)^{n-1}L_3\left|\frac{t}{\sqrt{n}}\right|^3.
\end{aligned}$$

Which implies:

$$\int_{-\rho\epsilon\sqrt{n}}^{\rho\epsilon\sqrt{n}}\left|\frac{\left(\lambda\left(t/\sqrt{n}\right)\right)^n - e^{-\rho^2t^2/2}}{t}\right|dt \leq \frac{M_3}{\sqrt{n}}\int_{-\infty}^{\infty} t^2e^{-\rho^2t^2/4}dt. \tag{4.22}$$

Inequalities (4.20), (4.21) and (4.22) imply the speed of convergence (4.2). □



(a) (b)

Figure 4.1: (a) Histogram of 10000 realizations of $\frac{S_n f}{\sqrt{n}}$ from Example 4.9.1. The histogram shows consistency with the central limit theorem. (b) Illustration of Berry-Esseen inequality. 4.9.1. The absolute value from inequality (4.2) goes to zero like $\frac{1}{\sqrt{n}}$.

4.9. Examples

Let us illustrate an application of the above results by using two simple examples:

Example 4.9.1. *Let $K = [0,1]$ and $d(z_1, z_2) = |z_1 - z_2|$ for $z_1, z_2 \in K$. Let Σ_A^+ be the full shift associated to the alphabet $\{1,2\}$ and the aperiodic matrix A with $A_{i,j} = 1$ for $i,j \in \{1,2\}$. Let us endow Σ_A^+ with a Markov measure using the probability vector $p = [\frac{1}{2}, \frac{1}{2}]$. Consider the map $G : \Sigma \to K$ given by:*

$$G(\underline{x}, z) = \begin{cases} T_1(z) & x_0 = 1 \\ T_2(z) & x_0 = 2 \end{cases},$$

with $T_1(z) = \frac{z}{2}$ and $T_2(z) = \frac{z+1}{2}$. It is straightforward to see that G statisfies C1 and C2 with $\alpha = \frac{1}{2}$ and $\operatorname{ess\,sup}_{z\in K} k_z = \frac{1}{2}$. Denote $\mathbb{P}$ to the product measure of the Lebesgue measure L and the Markov measure m. $\mathbb{P}$ is invariant under the skew-product $F(\underline{x}, z) = (\sigma\underline{x}, G(\underline{x}, z))$ as it satisfies:

$$L(A) = \sum_{i=1}^{2} p_i \cdot L\left(T_i^{-1} A\right).$$

Clearly, $\mathbb{P} \circ \pi_1^{-1} = m$ and therefore $\mathbb{P} \in S^\infty$. As a consequence of the uniqueness of μ_0, we can conclude that $\mathbb{P} = \mu_0$. Let us now set an observable $f(\gamma, z) = z - \frac{1}{2}$. This is a Lipschitz observable that satisfies P1-P2 enabling us to apply Theorem 4.1.1 and conclude that the sequence $\frac{S_n f}{\sqrt{n}}$ satisfies a central limit theorem with speed of convergence given by Inequality (4.2). Figure 4.1 shows a numerical simulation that illustrates our theorem.

Example 4.9.2. *Let $K \subset \mathbb{R}^2$ be the triangular region delimited with vertices on $v_1 = (0,0)$, $v_2 = (1,0)$ and $v_3 = (0.5, \sin(\pi/3))$. We endow K with the euclidean distance:*

$$d(z_1, z_2) = \sqrt{[(z_1)_x - (z_2)_x]^2 + [(z_1)_y - (z_2)_y]^2} \tag{4.23}$$

Let Σ_A^+ be the full shift associated to the alphabet $\{1,2,3\}$ and the aperiodic matrix A with $A_{i,j} = 1$ for $i,j \in \{1,2,3\}$. Let us endow Σ_A^+ with a Markov measure using the probability vector

$p = [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}]$. *Consider the map* $G : \Sigma \to K$ *given by:*

$$G(\underline{x}, z) = \frac{1}{2}\left((z)_x + (v_{x_0})_x, (z)_y + (v_{x_0})_y\right).$$

It is well known that iterations of the map G *produce the Sierpiński triangle depicted on figure 4.2a. Let us now show that* G *satisfies C1 and C2. For any* $\underline{x} \in \Sigma_A^+$ *and* $z_1, z_2 \in K$*:*

$$d(G(\underline{x}, z_1), G(\underline{x}, z_2)) = \sqrt{[(G(\underline{x}, z_1))_x - (G(\underline{x}, z_2))_x]^2 + [(G(\underline{x}, z_1))_y - (G(\underline{x}, z_2))_y]^2} \tag{4.24}$$

$$= \sqrt{\frac{1}{4}[(z_1)_x - (z_2)_x]^2 + \frac{1}{4}[(z_1)_y - (z_2)_y]^2} = \frac{1}{2}d(z_1, z_2). \tag{4.25}$$

So, condition C1 is satisfied with $\alpha = \frac{1}{2}$*. Also, for any* $z \in K$ *and* $\underline{x}, \underline{y} \in \Sigma_A^+$*, we have the following: if* $\underline{x}$ *and* $\underline{y}$ *belong to the same cylinder, then* $d(G(\underline{x}, z), G(\underline{y}, z)) = 0$*, otherwise we get:*

$$d(G(\underline{x}, z), G(\underline{y}, z)) < 1 \leq d_\theta(\underline{x}, \underline{y}). \tag{4.26}$$

And condition C2 is also satisfied with $k_z = 1$*. In reference [46] the author shows that for these type of systems, there exists a probability measure* μ_H *satisfying:*

$$\mu_H(A) = \sum_{i=1}^{3} p_i \cdot \mu_H\left(T_i^{-1}A\right), \tag{4.27}$$

where $T_i(z) = G(\underline{x}, z)|_{\underline{x}:x_0=i}$*. Let* $\mathbb{P}$ *be the product measure between* μ_H *and* m*. Then we can argue as in the previous example, and conclude that* $\mu_0 = \mathbb{P}$*. We will use the observable* $f(\underline{x}, z) = d(z, 0) - \mathfrak{b}$ *where* $\mathfrak{b}$ *is the average euclidean distance from the vertices to the origin. Let us roughly estimate* $\mathfrak{b}$*. The iterative construction of the Sierpiński gasket gives the bottom left vertices of the triangles of iteration* m *to be:*

$$l_s = \sum_{k=1}^{m} \frac{v_{s_k}}{2^k},$$

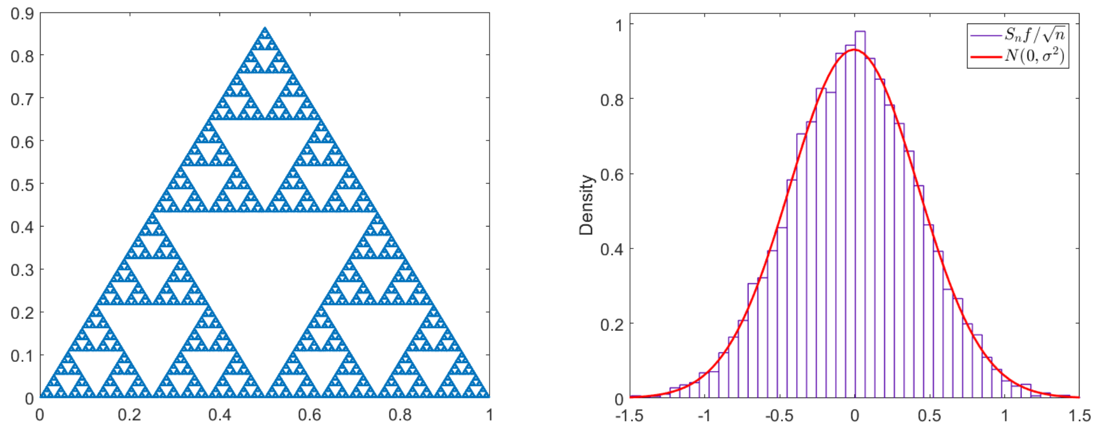
where s_k *is the element in position* k *the sequence* $s \in \{1, 2, 3\}^m$*. Then we can approximate* $\mathfrak{b}$ *by (see [41]):*

$$\mathfrak{b}_m = \frac{1}{|\{1, 2, 3\}^m|} \sum_s \sqrt{(l_s)_x^2 + (l_s)_y^2} \approx 0.618.$$

So, our observable is $f(\underline{x}, z) = d(z, 0) - 0.618$*. Using the inequality* $\left|\sqrt{a^2 + b^2} - \sqrt{c^2 + d^2}\right| \leq \sqrt{(a - c)^2 + (b - d)^2}$ *we get:*

$$\begin{aligned} |f(\gamma_1, z_1) - f(\gamma_2, z_2)| &= |d(z_1, 0) - d(z_2, 0)| = \left|\sqrt{(z_1)_x^2 + (z_1)_y^2} - \sqrt{(z_2)_x^2 + (z_2)_y^2}\right| \\ &\leq \left|\sqrt{((z_1)_x - (z_2)_x)^2 + ((z_1)_y - (z_2)_y)^2}\right| = d(z_1, z_2) \\ &\leq d(z_1, z_2) + d_\theta(\gamma_1, \gamma_2). \end{aligned}$$

So, f *is Lipschitz with respect to the distance* $d + d_\theta$*. Also, using the fact that* $d(z, 0) \leq 1$*,* f *satisfies properties P1-P2, so we can conclude that the random variable* $\frac{S_n f}{\sqrt{n}}$ *converges in distribution to a normally distributed random variable. Figure 4.2b shows a numerical simulation of this result.*



(a) (b)

Figure 4.2: (a) Sierpiński Triangle generated with iterations of the random dynamical system of Example 4.9.2. (b) Histogram of 10000 realizations of $\frac{S_n f}{\sqrt{n}}$ illustrating the central limit theorem.

5. Connected Dynamical Systems

In the previous two chapters we discussed a class of contractive random dynamical systems and its statistical properties. In this chapter we will be interested in a different setting. Roughly speaking, in this chapter we will discuss a class of deterministic dynamical systems in which, at every iteration, a map is selected based on the map that was selected on the previous iteration and the value of such iteration. We will focus on the case of two maps in the interval through holes in the space state. We refer to such systems as **Connected Dynamical Systems**. We will present some interesting phenomena that we found numerically and discuss a process that allows us to obtain a single map on the interval with essentially the same dynamics as the Connected Dynamical System. We use this single map to provide a heuristic explanation of the phenomena that we see on the connected dynamical system.

5.1. From random to non-random selection process

In a random dynamical system, each iteration of the dynamics consists on the random selection of one map among a family of maps. In this chapter we discuss a deterministic variation of this setting where the map is not selected randomly. Instead, the selection of the map becomes a process that depends on the last map that was iterated and the value of that iteration. Throughout this chapter we will refer to this process as the **selection process**. We will use this terminology for both, the random and the deterministic cases.

The selection process in a random dynamical system can be an IID process, meaning that the probability of selecting a map for the i-th iteration does not depend on the maps that were selected on any of the previous iterations. A first step for a non-IID selection process is the Markov case: the probability of selecting a map for the i-th iteration depends on the map that was selected on the $(i-1)$-th iteration but it does not depend on the maps that were selected on any of the iterations previous to $i-1$.

In the setting that we discuss this chapter, the selection process is allowed to have a dynamics on its own. Moreover, such dynamics depends on both, the map that was selected on the previous iteration and the value of that iteration. As explained above, we refer to such systems as connected dynamical systems. We explain how the setting of a Connected Dynamical System can be regarded as a selection process that depends on the value of the previous iteration and also depends on the map that was iterated. The result is a transformation on a product space that is similar to the transformation F that was studied in the case of a Random Dynamical System (see Equation (1.38) in Section 1.5).

5.2. Overview of Dynamical Systems with holes

In this section we will talk about dynamical systems with holes, which have some similarities with the connected dynamical systems we are interested. The notion of dynamical systems with holes was introduced by Pianigiani and Yorke in [78]. They studied a class of C^2 expanding maps defined on sets that are not invariant under the dynamics ($T : A \to \mathbb{R}^n$ with $A \subset T(A)$ strictly). This means the orbit might eventually escape from the state space and terminate. Suppose that we choose $x \in A$ according to a given initial measure ν_0. Let $E \subset A$ and $\nu_m(E)$ be the probability that $T^m(x) \in E$ given the fact that $x, ..., T^m(x) \in A$. If ν_m converges to a measure ν_∞ that is independent of ν_0, then ν_∞ is said to be a **conditionally invariant measure**. The authors of reference [78] show that this measure exists for expanding dynamical systems. Building on that work, Collet, Martínez and Schmitt [72] show that the conditionally invariant measure converges to a T-invariant probability measure on a limit Cantor set. Under the same setting, one might also be interested in characterizing the elements of A whose orbits do not escape from A; the so-called, survivor set. In reference [63] the authors provide conditions that allow them to obtain the Hausdorff dimension of the survivor set. Let H be the set such that $T(H) = T(A) \setminus A$. The set H is called hole and one might be interested in knowing how the position of H affects the escape rate. References [9] and [1] are outstanding results in this direction. In reference [9] the authors consider escape rates through holes that are preimages of a Markov partition and they show that the larger the Poincaré recurrence time, the larger the escape rate through the hole. In reference [1] the authors find upper and lower bounds for the probability that an orbit escapes through a hole in the $i-th$ iteration. A particularly transparent survey of results for dynamical systems with holes is [17]. Other related results for non uniformly hyperbolic systems are found in [8] and [16].

Note, however, the dynamics on these types of systems terminates when the orbit falls in the holes. For the systems we have in mind, the orbit does not terminate but instead, it continues with a different transformation. This can be seen as a situation in which a system T_1 that is in a state of equilibrium, is placed in contact with another system T_2, and as a result, the new system is out of equilibrium. In that sense, our setting is related to the metastable systems presented on reference [86]. In that reference, the authors obtain an approximation of the invariant density of the metastable system as a convex combination of the invariant densities of the stable systems. We further discuss this relation in Section 5.7.1.

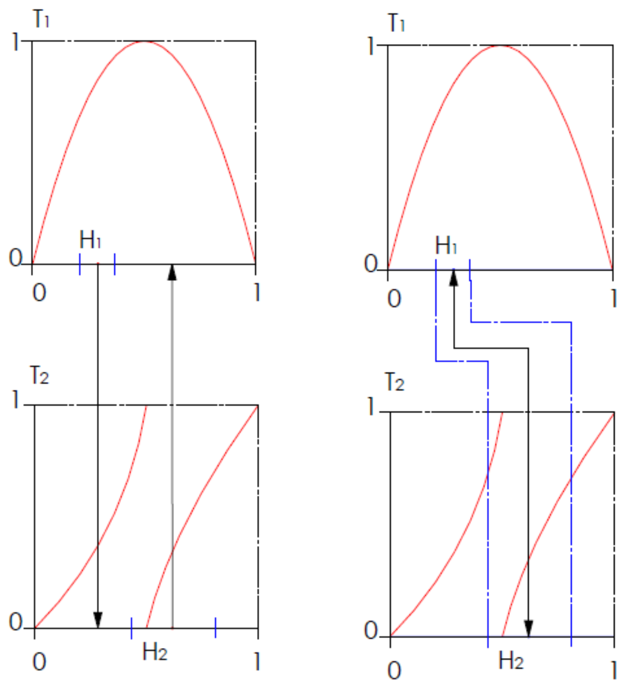
5.3. Connected Dynamical Systems

In this section we study a prototype of dynamical system consisting on two transformations on the unit interval, T_1, T_2 that are *connected* through holes on the space state. More specifically, we are interested in the dynamics generated in this way: There are two dynamics T_1 and T_2 that are connected through a hole from a region H_1 in the space state of T_1 to another region H_2 in the space state of T_2. We numerically explore some interesting critical behaviors arising on these type of maps. We also point out some similarities and differences with other previously studied settings such as position dependent random dynamical systems and metastable systems.

We will discuss two types of connected dynamical systems. The specific setting setting is as follows. Let $I = [0, 1]$ and consider the following prototype of dynamical system. We are given two maps $T_1, T_2 : I \to I$. Associated to each map, we are given two closed intervals $H_1 = H_{T_1} \subseteq I$ and $H_2 = H_{T_2} \subseteq I$ which we call *holes*. We are interested in the dynamics generated as follows: we iterate the map T_1 with a random, uniformly distributed initial condition on I until the value of the iteration falls on H_1. When this happens, we change maps and we will continue the orbit with iterations of T_2 until the value of the iteration falls on H_2. We continue this way, changing maps when the value of the iteration falls in the corresponding hole. We studied two different ways in which we can continue the dynamics when falling in a hole:

1. **Type 1 Connected Dynamical System (CDS-1)**: When the orbit falls in the hole of T_1, the dynamics continue on the exact same point, but with T_2. Similarly, when the orbit falls in the hole of T_2, the dynamics continue on the exact same point but with T_1. Figure 5.1a depicts this setting.

2. **Type 2 Connected Dynamical System (CDS-2)**: when the orbit falls in the hole of



(a) CDS-1 (b) CDS-2

Figure 5.1: Graphical representation of the studied settings. When the orbit falls in H_1: (CDS-1) we take that value and start iterating T_2. (CDS-2) we use τ_1 to take that value to the other hole and then we start iterating T_2 with the resulting value.

T_1, the value is taken to H_2 with the application of the linear map $\tau_1 : H_1 \to H_2$ given by:

$$\tau_1(x) = \min(H_2) + \frac{(\max(H_2) - \min(H_2))(x - \min(H_1))}{\max(H_1) - \min(H_1)},$$

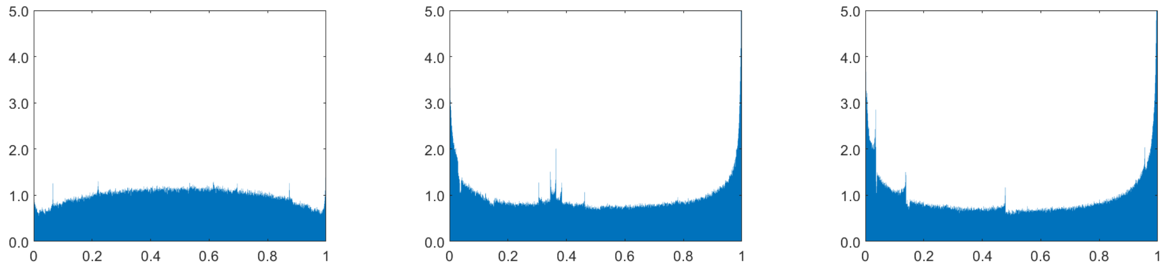
and we iterate now the map T_2 with initial condition $\tau_1(x)$ until the value of the iteration falls in H_2. When this happens, we take that value and apply the linear map $\tau_2 : H_2 \to H_1$ given by:

$$\tau_2(x) = \min(H_1) + \frac{(\max(H_1) - \min(H_1))(x - \min(H_2))}{\max(H_2) - \min(H_2)}.$$

Figure 5.1b depicts this second setting.

Remark 5.3.1. *In general, both settings are different because they represent classes of dynamical systems that are not contained one on the other. The only circumstance where the orbits will coincide, is when $H_1 = H_2$, because in this case $\tau_1(x) = \tau_2(x) = x$.*

We will now describe some of the critical phenomena that we found on specific Connected Dynamical Systems. These phenomena are described in terms of the qualitative behavior of the **empirical density**, presented as a histogram and obtained using 5×10^5 iterations and 1×10^3 bins.



(a) $\beta = 0.97$ (b) $\beta = 0.79$ (c) $\beta = 0.51$

Figure 5.2: Empirical density of the CDS-2 given by (5.1) with $H_1 = [0, 0.1]$ and $H_2 = [\beta, 1]$ for different values of β. For values of β that are close to 1, the empirical density resembles the acim of the Lorenz map. As β decreases, it resembles the one of the logistic map.

5.4. From Chaos to Periodic Orbits

Let us fix a CDS-2 with the logistic map with parameter $r = 4$ and the one-dimensional Lorenz map with the parameters $\theta = 109/64$ and $\alpha = 51/64$, namely:

$$\begin{aligned} T_1(x) &= 4x(1-x), \\ T_2(x) &= \begin{cases} \theta|x-0.5|^{\alpha}, & x < 0.5, \\ 1-\theta|x-0.5|^{\alpha}, & x \geq 0.5. \end{cases} \end{aligned} \tag{5.1}$$

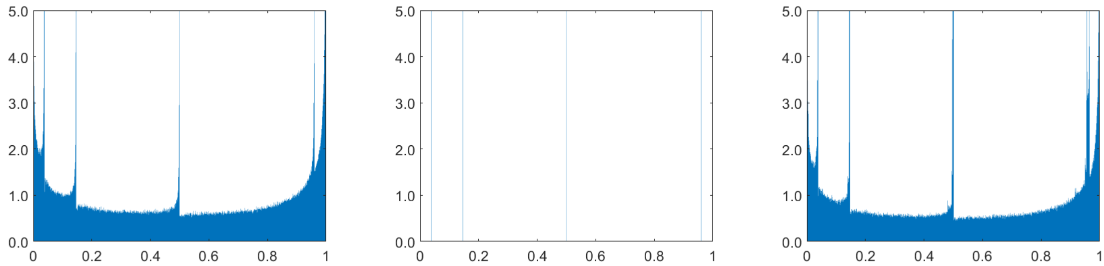
Let $H_1 = [0, 0.01]$ and $H_2 = [\beta, 1]$, where β is a parameter. The results of numerical simulations with three selected values of β are shown on Figure 5.2.

As expected, the empirical density of the connected system resembles the invariant density of the Lorenz map when β is close to 1 because it results in a small hole H_2, which causes the dynamics to spend more time iterating T_2, which is the Lorenz map. As β decreases, H_2 becomes a larger hole and the dynamics spends more time iterating the logistic map, and therefore, the empirical density, in this case, resembles the invariant density of the logistic map.

One would expect the empirical density of the connected dynamical system will look more and more similar to the one of the logistic map as we decrease the value of β. However, this is not exactly the case. Numerical simulations show that if we continue decreasing the value of β, the empirical density exhibits high peaks on specific values and eventually, at $\beta \approx 0.0125$, the values on these peaks form a periodic orbit causing the empirical density to vanish. Even more, if we continue decreasing the values of β, we see that the empirical density returns at $\beta \approx 0.0095$. This phenomenon, shown on Figure 5.3, can be seen as a **phase transition**, in the sense of transition from existence to non-existence of invariant measure. The figure shows that with $\beta = 0.0126$, there exists an empirical density, but with $\beta = 0.0125$ the orbit is periodic with period 6, and yet when β falls to 0.0095, there is again an empirical density supported on the whole unit interval.

Further numerical simulations show that if $\beta \in [0.0096, 0.0125]$, then the orbits have a periodic behavior and it is observed that the period of the orbit is given in multiples of 6 and it changes at specific critical values of β that can be localized empirically.

In order to support this claim, we have estimated the Lyapunov exponent for the connected map. Figure 5.4 shows how the Lyapunov exponent behaves as a function of the parameter β. Showing that the orbits of the system enter into a different regime (of an ordered system) when $\beta \in [0.0095, 0.0125]$. Outside of this range, the exponent is positive and remains approximately constant. But once it is in that range, the exponent abruptly decreases and becomes negative because of the appearance of periodic orbits.



(a) $\beta = 0.0126$ (b) $\beta = 0.0125$ (c) $\beta = 0.0095$

Figure 5.3: Phase transition for the empirical densities occurring when $0.0095 < \beta < 0.0125$, for the connected map defined by (5.1). For $H_1 = [0, 0.01]$ and $H_2 = [\beta, 1]$, with $\beta = 0.0125$, the numerical simulation results in a periodic orbit of period 6.

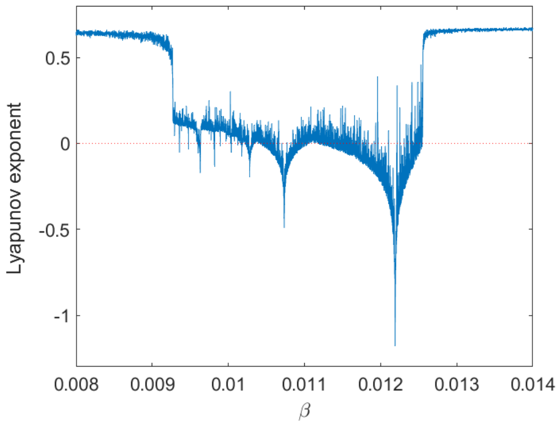


Figure 5.4: Lyapunov exponent for the CDS-2 formed with maps (5.1) and holes $H_1 = [0, 0.01]$ and $H_2 = [\beta, 1]$. The system exhibits an ordered regime for values of the parameter $\beta \in [0.0095, 0.0126]$.

We will revisit this example in Section 5.6 where we provide an argument for the existence of the periodic cycle shown on Figure 5.3b. The technique that we use is a rescaling process that allows us to see the connected maps as a single map in the interval. The resulting map still exhibits the transitions from existence of empirical measure to a periodic cycle of period six, and then again from the periodic cycle to existence of empirical measure. To the best of our knowledge, a simliar transition has not been reported in the literature before, so it becomes an interesting phenomenon to study.

5.5. Inducing order by connecting maps

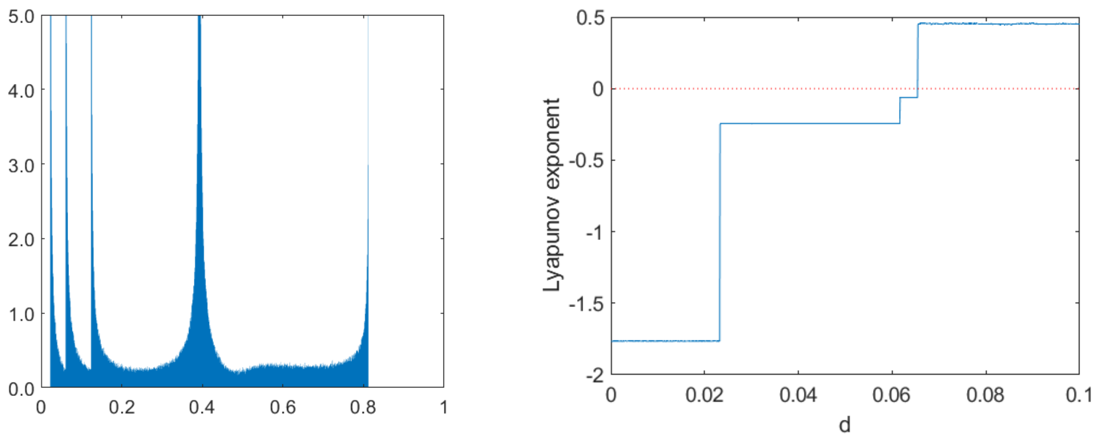
The **Belousov-Zhabotinsky** map is an interval map that was proposed to model the chemical reaction of the same name. It is composed of three regions and a typical orbit will oscillate around several values but it will never stabilize on any periodic orbit. Matsumoto and Tsuda [67] reported a noise-induced order phenomenon on this map that occurrs when additive noise is injected to the system. When this happens, a transition occurs and the systems moves from a chaotic regime to an ordered regime. Recently, Galatolo S., Monge M. and Nisoli I. [29] have shown this transition analytically, although they still use numerical simulations to obtain certain parameters of the model.

In this section we will see how the Connected Dynamical System composed of the logistic map and the Beloúsov-Zhabotinski map exhibits an *induced order* property. That is similar to the phenomenon reported in [67] or to the experimental phenomenon reported in [77], with the difference that in those cases the order is induced by noise and here, by the fact of connecting the maps.

Let us consider the connected map formed with the logistic map and the Beloúsov-Zhabotinski map, the later is given by:

$$T_2(x) = \begin{cases} \left(-(0.125-x)^{1/3}+a\right)e^{-x}+b & x<0.125, \\ \left((x-0.125)^{1/3}+a\right)e^{-x}+b & x<0.3, \\ c\left(10xe^{-10x/3}\right)^{19}+b & x\geq 0.3, \end{cases} \tag{5.2}$$

where the parameters are given by the classical values: $a = 0.50607357, b = 0.0232885279$ and $c = 0.121205692$ (see [67]).



(a) (b)

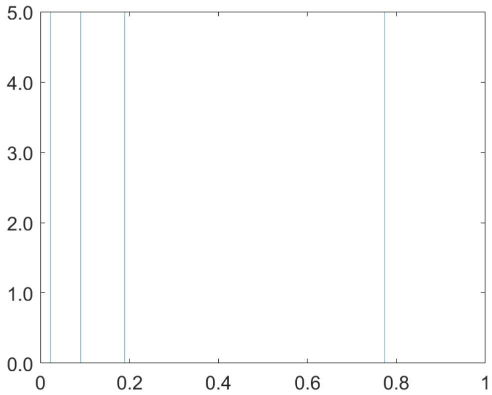
Figure 5.5: (a) Empirical density for the Beloúsov-Zhabotinsky map given by (5.2). (b) Lyapunov exponent for the CDS-2 formed with the logistic map and the Beloúsov-Zhabotinsky map, with holes $H_1 = H_2 = [d, d+0.1]$. The Lyapunov exponent is negative for $d \leq 0.0654$ where the periodic orbits of period 13, 10 and 4, appear.

Here, we considered the connected map using the logistic map and the Beloúsov-Zhabotinsky map with holes given by $H_1 = H_2 = [d, d+0.1]$ (since the holes are equal, this system can be seen as either CDS-1 or CDS-2).

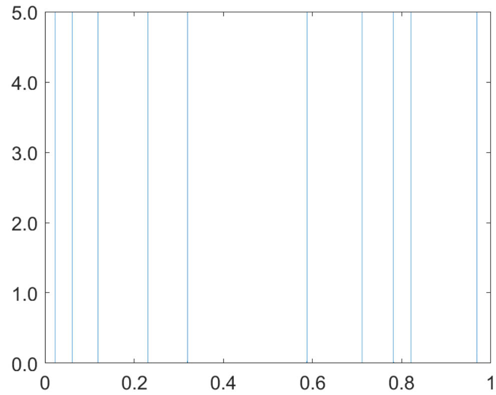
Similar to the previous section, we identified a region of parameter values d, for which the connected system have orbits ending up in periodic orbits. When $d \in [0, 0.0654]$ our simulations indicate that the dynamics end up in periodic orbits, displaying a negative Lyapunov exponent (see Figure 5.5b) and a positive Lyapunov exponent (around a fixed value) for all $d \geq 0.0654$. A 4-period cycle appears when $d \in [0, 0.0231]$, a cycle of period 10 appears for $d \in (0.0231, 0.0615]$, and a 13-period cycle shows up for $d \in (0.0615, 0.0654]$. For larger values of d the system stays at a chaotic regime, having empirical densities, see Figure 5.6. The existence of the periodic region, can be seen as a stabilization phenomenon of the Beloúsov-Zhabotinski map by means of connecting it with a chaotic dynamical systems, similar to a noise induced order phenomenon obtained experimentally in reference [77] by means of a feedback algorithm.

5.6. Rescaling a connected dynamical system into a map in the interval

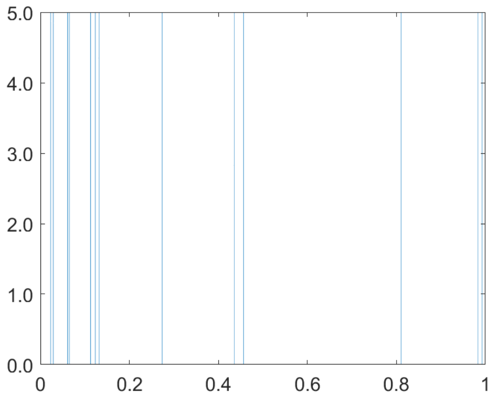
The dynamics of a Connected Dynamical System consists in the iteration of either T_1 or T_2 depending on the value of the previous iteration: if the value of the previous iteration has fallen into the hole of the map under iteration, then we use the other map for the next iteration,



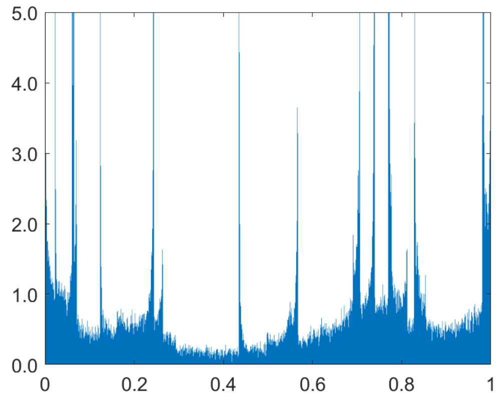
(a) $d = 0.0200$



(b) $d = 0.0400$



(c) $d = 0.0630$



(d) $d = 0.0800$

Figure 5.6: Histograms of the Connected Dynamical System formed with the logistic map and the Beloúsov-Zhabotinski map. The holes are set as $H_1 = H_2 = [d, d + 0.1]$. The numerical simulations of the system show 4 different regimes depending on the value of d. For $d \in [0, 0.0231]$, the system reaches a periodic orbit of period four. While if $d \in (0.0231, 0.0615]$, the period increases to ten and when $d \in (0.0615, 0.0654]$, it further increases to 13. Finally, when $d > 0.0654$ the system enters into a chaotic regime.

otherwise we continue. In this section we will see that a Connected Dynamical System can be seen as a single scaled map acting in the unit interval. We will use this strategy to give a plausibility argument for the existence of the periodic cycle of period 6 for the CDS-2 presented in Section 5.4.

Let us start by setting the following notation. For given closed subintervals of positive Lebesgue measure, $A, B \subseteq [0, 1]$ and $x \in A$, let us define the map $\phi : A \to B$ by:

$$\phi_{A,B}(x) = \min(B) + \frac{(\max(B) - \min(B))(x - \min(A))}{\max(A) - \min(A)}. \tag{5.3}$$

Also, for $x \in I$ let:

$$\psi(x) = \begin{cases} \phi_{I_1,I}(x) & x \in I_1, \\ \phi_{I_2,I}(x) & x \in I_2. \end{cases}$$

Equation (5.3) is similar to the maps τ_1 and τ_2 whose purpose is to move the values from one of the holes to the other on a CDS-2. The map $\phi_{A,B}$ will work similar, but A and B will not necessarily be holes.

Let $T_1, T_2 : I \to I$ and H_1, H_2 be closed subintervals of I. The CDS-1 and the CDS-2 formed with (T_1, T_2, H_1, H_2) can be treated as a single map on the interval by properly scaling the involved intervals. More precisely, we will find the maps $\widetilde{T}_{CDS-1}, \widetilde{T}_{CDS-2} : I \to I$ such that:

1. The orbit of the CDS-1 formed with (T_1, T_2, H_1, H_2) can be obtained by applying ψ to the orbit of $\widetilde{T}_{CDS-1}$.
2. Similarly, the orbit of the CDS-2 formed with (T_1, T_2, H_1, H_2) can be obtained by applying ψ to the orbit of $\widetilde{T}_{CDS-2}$.

Let $\widetilde{T}_1 : I_1 \to I_1$ and $\widetilde{T}_2 : I_2 \to I_2$ be defined by the following commutative diagrams:

$$\begin{tikzcd} I_1 \arrow[r, "\widetilde{T}_1"] \arrow[d, "\phi_{I_1,I}"] & I_1 \arrow[d, "\phi_{I_1,I}"] \\ I \arrow[r, "T_1"] & I \end{tikzcd} \qquad \begin{tikzcd} I_2 \arrow[r, "\widetilde{T}_2"] \arrow[d, "\phi_{I_2,I}"] & I_2 \arrow[d, "\phi_{I_2,I}"] \\ I \arrow[r, "T_2"] & I \end{tikzcd}$$

Then set new holes $\widetilde{H}_1 \subseteq I_1$ and $\widetilde{H}_2 \subseteq I_2$ as follows:

$$\begin{aligned} \widetilde{H}_1 &= \phi_{I,I_1}(H_1), \\ \widetilde{H}_2 &= \phi_{I,I_2}(H_2). \end{aligned} \tag{5.4}$$

Finally, the maps $\widetilde{T}_{CDS-1}$ and $\widetilde{T}_{CDS-2}$ are defined respectively as follows:

$$\widetilde{T}_{CDS-1}(x) = \begin{cases} \widetilde{T}_1(x) & x \in I_1, x \notin \widetilde{H}_1, \\ \widetilde{T}_2 \circ \phi_{I_1,I_2}(x) & x \in I_1, x \in \widetilde{H}_1, \\ \widetilde{T}_2(x) & x \in I_2, x \notin \widetilde{H}_2, \\ \widetilde{T}_1 \circ \phi_{I_2,I_1}(x) & x \in I_2, x \in \widetilde{H}_2, \end{cases} \tag{5.5}$$

$$\widetilde{T}_{CDS-2}(x) = \begin{cases} \widetilde{T}_1(x) & x \in I_1, x \notin \widetilde{H}_1, \\ \widetilde{T}_2 \circ \phi_{\widetilde{H}_1,\widetilde{H}_2}(x) & x \in I_1, x \in \widetilde{H}_1, \\ \widetilde{T}_2(x) & x \in I_2, x \notin \widetilde{H}_2, \\ \widetilde{T}_1 \circ \phi_{\widetilde{H}_2,\widetilde{H}_1}(x) & x \in I_2, x \in \widetilde{H}_2, \end{cases} \tag{5.6}$$

Note that Equations (5.5) and (5.6) are maps in the interval that encapsulate the dynamics of the original Connected Dynamical System. These maps in the interval can be studied using standard techniques and they will provide information of the original Connected Dynamical System they come from. We illustrate this fact on Section 5.6.1 where we present a plausibility argument for the existence of the periodic cycle of period six that was numerically found on Section 5.4.

Let us now present an example that shows how the maps (5.5) and (5.6) can be obtained from the corresponding Connected Dynamical System.

Example 5.6.1. *Let T be the CDS-1 given by:*

$$T_1(x) = \begin{cases} 2x & x \in \left[0, \frac{1}{2}\right), \\ 2x - 1 & x \in \left[\frac{1}{2}, 1\right]. \end{cases} \qquad T_2(x) = \begin{cases} 3x & x \in \left[0, \frac{1}{3}\right), \\ -x + 4/3 & x \in \left[\frac{1}{3}, \frac{2}{3}\right), \\ -2x + 2 & x \in \left[\frac{2}{3}, 1\right]. \end{cases} \tag{5.7}$$

These maps are depicted in Figure 5.7. Let the holes be holes $H_1 = [0, 1/10]$ and $H_2 = [19/20, 1]$. Let us define scaled versions of T_1 and T_2 as follows:

$$\widetilde{T}_1(x) = \begin{cases} 2x & x \in \left[0, \frac{1}{4}\right), \\ 2x - \frac{1}{2} & x \in \left[\frac{1}{4}, \frac{1}{2}\right]. \end{cases} \qquad \widetilde{T}_2(x) = \begin{cases} 3x - 1 & x \in \left[\frac{1}{2}, \frac{2}{3}\right), \\ -x + \frac{5}{3} & x \in \left[\frac{2}{3}, \frac{5}{6}\right), \\ -2x + \frac{5}{2} & x \in \left[\frac{5}{6}, 1\right]. \end{cases} \tag{5.8}$$

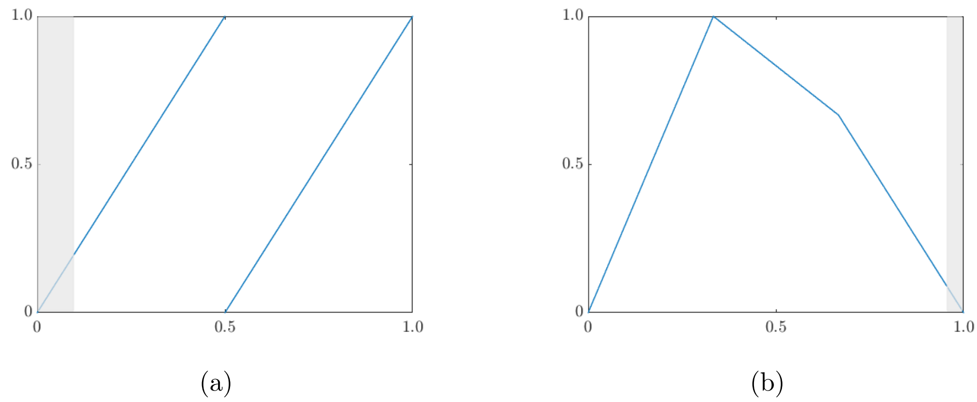


Figure 5.7: The maps T_1 and T_2 from Example 5.6.1. The holes of the connected dynamical system are shown in gray.

The holes will also be scaled accordingly:

$$\widetilde{H}_1 = \left[0, \frac{1}{20}\right], \qquad \widetilde{H}_2 = \left[\frac{39}{40}, 1\right]. \tag{5.9}$$

These are the sets that will trigger the change between the two almost invariant sets $[0, 0.5)$ *and* $[0.5, 1]$*. The dynamics of the single map of the interval* $\widetilde{T}_{CDS-1}$ *is given given by:*

$$\widetilde{T}_{CDS-1}(x) = \begin{cases} 3x + \frac{1}{2} & x \in \left[0, \frac{1}{20}\right), \\ 2x & x \in \left[\frac{1}{20}, \frac{1}{4}\right), \\ 2x - \frac{1}{2} & x \in \left[\frac{1}{4}, \frac{1}{2}\right), \\ 3x - 1 & x \in \left[\frac{1}{2}, \frac{2}{3}\right), \\ -x + \frac{5}{3} & x \in \left[\frac{2}{3}, \frac{5}{6}\right), \\ -2x + \frac{5}{2} & x \in \left[\frac{5}{6}, \frac{39}{40}\right), \\ 2x - \frac{3}{2} & x \in \left[\frac{39}{40}, 1\right). \end{cases} \tag{5.10}$$

Figure 5.8 shows the resulting dynamical system and a typical orbit. Let us now transform the orbit of $\widetilde{T}$ *to obtain the orbit of the original CDS-1. In order to do that, we can simply rescale again the values to cover the whole interval* $[0, 1]$*:*

$$T^n(x) = \begin{cases} 2\widetilde{T}^n_{CDS-1}(x) & \widetilde{T}^n_{CDS-1}(x) \in \left[0, \frac{1}{2}\right), \\ 2\left[\widetilde{T}^n_{CDS-1}(x) - \frac{1}{2}\right] & \widetilde{T}^n_{CDS-1}(x) \in \left[\frac{1}{2}, 1\right). \end{cases} \tag{5.11}$$

Figure 5.9 shows the empirical density of: (a) the CDS-1 of Example 5.6.1 and (b) the corresponding map $\widetilde{T}_{CDS-1}$*. As it can be seen, the densities match consistently on the whole interval.*

Example 5.6.2. *Let us suppose that we have the same maps and holes but now it is a CDS-2, which means we will apply* τ_1 *and* τ_2 *when the orbit falls into* H_1 *and* H_2 *respectively. This system can also be seen as a single map in the interval, but now the definition of* $\widetilde{T}_{CDS-2}$ *will change to reflect the map* τ_1 *in the first linear part and* τ_2 *on the last linear part:*

$$\widetilde{T}_{\text{CDS-2}}(x) = \begin{cases} 4x - \frac{39}{10} & x \in \left[0, \frac{1}{20}\right), \\ 2x & x \in \left[\frac{1}{20}, \frac{1}{4}\right), \\ 2x - \frac{1}{2} & x \in \left[\frac{1}{4}, \frac{1}{2}\right), \\ 3x - 1 & x \in \left[\frac{1}{2}, \frac{2}{3}\right), \\ -x + \frac{5}{3} & x \in \left[\frac{2}{3}, \frac{5}{6}\right), \\ -2x + \frac{5}{2} & x \in \left[\frac{5}{6}, \frac{39}{40}\right), \\ \frac{x}{2} + \frac{39}{40} & x \in \left[\frac{39}{40}, 1\right). \end{cases} \tag{5.12}$$

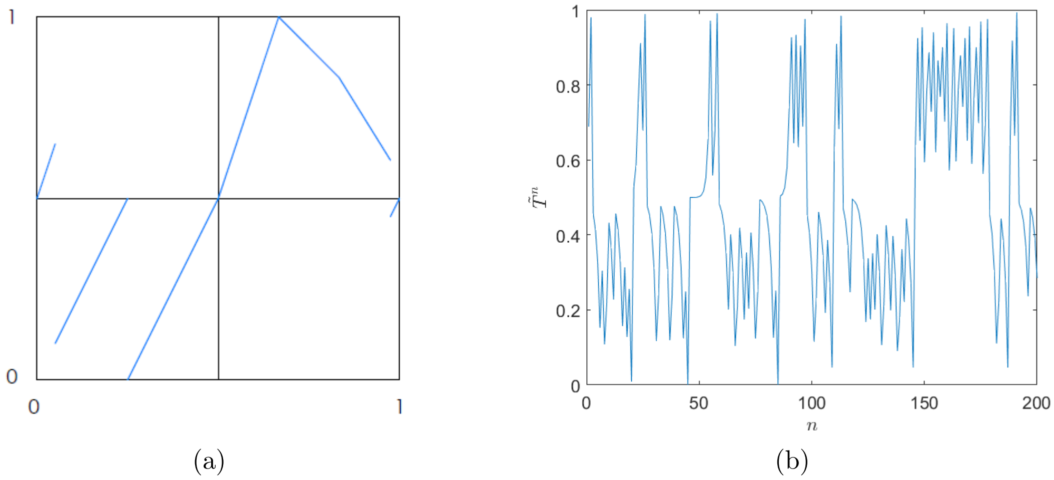


Figure 5.8: Resulting $\widetilde{T}_{CDS-1}$ from example 5.6.1.

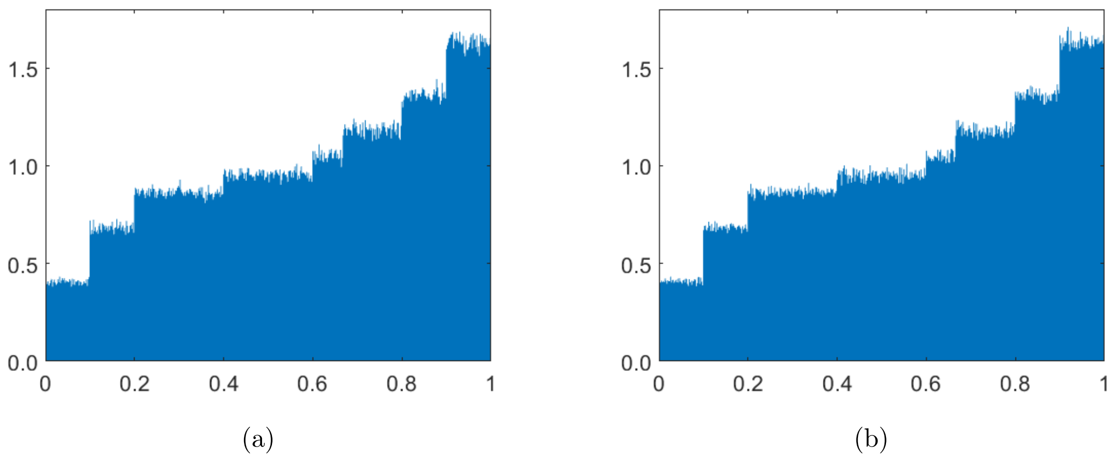


Figure 5.9: Empirical densities of: (a) the original CDS-1 (5.7) and (b) the map on the interval $\widetilde{T}_{CDS-1}$ given by (5.11).

5.6.1 Attracting period 6: a plausibility argument

Let us now revisit the CDS-2 presented in Section 5.4. A plausibility argument for the existence of an attracting orbit of period 6 when $\beta = 0.0125$ can be given as follows. We first rescale the CDS-2 into a single map in the interval using the process explained on the previous section. We obtain $\widetilde{H}_1 = [0, 0.005]$ and $\widetilde{H}_2 = [0.50625, 1]$ for the scaled holes. We will now follow the iterations of the scaled system in the interval $A = [0.245, 0.255]$. The first and second plots of the resulting map are given in Figure 5.10a. The plot of the second iteration shows that $\widetilde{T}^2(A) = \widetilde{T}_1^2(A) \subseteq \widetilde{H}_1$ and hence:

$$\widetilde{T}^3(A) = \widetilde{T}_2 \circ \phi_{\widetilde{H}_1, \widetilde{H}_2}(\widetilde{T}_1^2(A)).$$

Similarly, Figure 5.10b shows that $\widetilde{T}^3(A) \subseteq \widetilde{H}_2$ and hence:

$$\widetilde{T}^4(A) = \widetilde{T}_1 \circ \phi_{\widetilde{H}_2, \widetilde{H}_1}(\widetilde{T}^3(A)).$$

Finally, Figure 5.10c shows that $\widetilde{T}^i(A) \notin \widetilde{H}_1$ for $i = 4, 5$. Therefore:

$$\widetilde{T}^6(A) = \widetilde{T}_1^3 \circ \phi_{\widetilde{H}_2, \widetilde{H}_1}(\widetilde{T}^3(A)).$$

Figure 5.10d shows the plot of $\widetilde{T}^6(A)$ along with the identity map. The horizontal line indicates the point above which $|(\widetilde{T}^6(A))'| < 1$, where we have evaluated the derivative numerically. The figure shows that $\widetilde{T}^6$ and the identity map intersect at a point well above the dashed line, around $r \approx 0.249851...$. This means that the point of intersection r represents an attractive fixed point of $\widetilde{T}^6$ implying that it belongs to a periodic cycle of period 6 of $\widetilde{T}$.

5.7. Relation with Other Types of Dynamical Systems

In this section we illustrate how Connected Dynamical Systems CDS-1 and CDS-2 are related to other settings previously studied in the literature. The first obvious relation is with a standard closed dynamical system: A CDS-2 becomes a closed dynamical system when $T_1 = T_2$ and $H_1 = H_2$. For a CDS-1, it is enough to set $T_1 = T_2$ as the holes will not produce any effect on this case. For both settings, a degenerated case in which $H_1 = [0, 0]$ will also result on a closed dynamical system. So, the setting we propose can be seen as a deterministic generalization of a closed dynamical system.

Another relation can be found on position dependent random dynamical systems. Such dynamical systems are studied in [34]. In their setting, one of the transformations is randomly selected based on probabilities $p_1(x)$ and $p_2(x)$ that depend on the position of the previous iteration. The following example shows how a specific class of position dependent random dynamical system can be seen as a CDS-1.

Example 5.7.1. *Let $T_1, T_2 : I \to I$, $a \in (0, 1)$, $A = [0, a]$ and $B = (a, 1]$. Set $p_1(x) = 1_A(x)$ and $p_2(x) = 1 - p_1(x)$. The resulting position dependent random dynamical system can be seen as a CDS-1 if we set $H_1 = B$ and $H_2 = A$.*

The indicator functions $p_1(x)$ and $p_2(x)$ as defined above, have the same effect as the holes on a CDS-1. But when $p_1(x)$ and $p_2(x)$ are not indicator functions, they will not act as holes anymore. So, the system presented on the above example is the only class of position dependent random dynamical systems that can be seen as connected dynamical system.

5.7.1 Relation with Metastable Systems

A close examination of the process followed in Section 5.6 reveals that the resulting equations (5.4), will always have the following distinctive properties (let T below denote either $\widetilde{T}_{CDS-1}$ or $\widetilde{T}_{CDS-2}$):

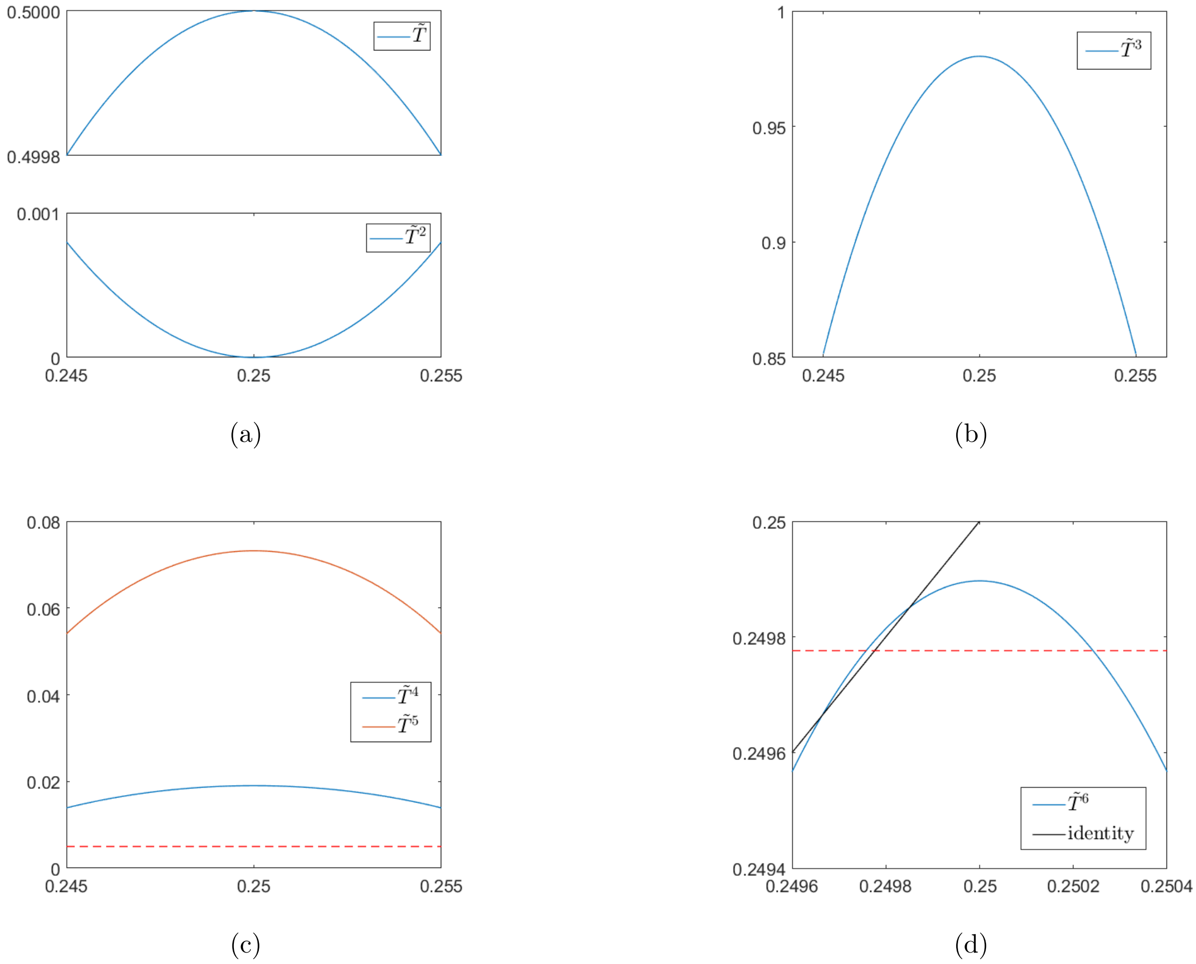


Figure 5.10: Plausibility argument for the existence of a periodic orbit of period 6. We follow the orbit of $A = [0.245, 0.255]$ under $\widetilde{T}$. (a) $\widetilde{T}^2(A) \subseteq \widetilde{H}_1$, so $\widetilde{T}_2$ will be applied on the next iteration. (b) $\widetilde{T}^3(A) \subseteq \widetilde{H}_2$ so $\widetilde{T}_1$ will be applied in the next iteration. (c) $\widetilde{T}^i(A) \notin \widetilde{H}_1$ so $\widetilde{T}_1$ will be applied on iteration 6. (d) Finally, $\widetilde{T}^6$ intersects the identity map on a point where the absolute value of its derivative is less than 1 (dashed line).

M1: $T(H_1) \subseteq I_2$ and $T(H_2) \subseteq I_1$.

M2: $T(I_1 \setminus H_1) \subseteq I_1$ and $T(I_2 \setminus H_2) \subseteq I_2$.

If the holes are small, the system will stay on I_1 for a long time but it will eventually switch to I_2 where it will also stay for a long time until it switches again to I_1. Moreover, the degenerated case when $\text{Leb}(H_1) = \text{Leb}(H_2) = 0$ produces a system, in which I_1 and I_2 become invariant under T.

There is another class of dynamical systems that follows the same qualitative behavior: metastable systems. In reference [86], the authors study a family of transformations $T_\epsilon : I \to I$ where the index ϵ is considered a smooth perturbation of an original system T_0. The sets I_1 and I_2 are invariant under T_0 but the perturbation destroys this invariance by creating holes through which the dynamics can move from I_1 to I_2 and viceversa. The authors of [86] show that for small ϵ, the invariant density of the system T_ϵ can be approximated as a convex combination of the densities of T_1 and T_2.

Despite having the same qualitative behaviour, the mechanism that destroys the invariance is completely different in a metastable system compared to a Connected Dynamical System. In a metastable system, the holes on the phase space arise as a consequence of a smooth perturbation of the stable system. On the other hand, in a Connected Dynamical System the given holes cause an abrupt change with respect to the stable system.

5.7.2 Relation with *skew-product* dynamics

In this section we present a general setting where the selection process is allowed to depend on variables related to the dynamics. Let $I = [0,1]$, S be an index set, not necessarily finite and suppose that we are given a family of transformations $T_s : I \to I$ with $s \in S$. On each iteration of the dynamic, we will select one of the maps T_s and we will also select an $x \in I$ where the map will be evaluated. For this purpose, let $\widetilde{\Omega}$ be an arbitrary non empty set and denote $\Omega = \widetilde{\Omega} \times I$. The selection process will be carried out by the maps $\sigma : \Omega \to \Omega$, $\xi_1 : \Omega \to S$ and $\xi_2 : \Omega \to I$. The dynamics can be seen as a skew-product transformation $\Theta : \Omega \to \Omega$ given by:

$$\Theta(\omega) = (\pi_1(\sigma(\omega)), T_{\xi_1(\omega)}(\xi_2(\omega))). \tag{5.13}$$

Under this setting, we are interested on the projection of Θ seen from the unit interval, i.e. given an initial condition $\omega \in \Omega$, we are interested on the limit behaviour of:

$$x_n = T_{\xi_1(\sigma^{n-1}\omega)}\left(\xi_2(\sigma^{n-1}\omega)\right). \tag{5.14}$$

Let us now present some examples of the types of Dynamical Systems that can be obtained with the above setting.

Example 5.7.2. *(Connected Dynamical Systems formed with 2 maps) Let $\widetilde{\Omega} = S = \{1,2\}$ and $k : S \to S$ defined by $k(s) \neq s$ (in other words, $k(1) = 2$ and $k(2) = 1$) and:*

$$\tau_s = \phi_{H_s, H_{k(s)}}.$$

Then we can set:

$$\sigma(s,x) = \begin{cases} (1,x) & s = 1, x \notin H_1, \\ (2,\tau_1(x)) & s = 1, x \in H_1, \\ (1,\tau_2(x)) & s = 2, x \in H_2, \\ (2,x) & s = 2, x \notin H_2, \end{cases} \tag{5.15}$$

and ξ_1 and ξ_2 as the projections on S and I respectively and we will recover the setting of a CDS-2 from the previous sections. A CDS-1 can also be recovered if we modify the definition of σ so that its projection on the second coordinate is the identity map (i.e. it always returns the same second coordinate).

Example 5.7.3. *(Connected Dynamical Systems formed with n maps) If $S = \{1,2,...,n\}$ we can use partitions on I to define the action of σ. Let $\mathcal{P}_i = \mathcal{P}_{T_i}$, $i \in S$ be partitions of I in intervals and let $k(P)$ be an element of S associated to each $P \in \mathcal{P}_i$. Moreover, let $P(i,x)$ be the element of partition $\mathcal{P}_i$ such that $x \in P(i,x)$. For a CDS-2 note we will need to add the following conditions so that our setting can be extended to n maps. For each fixed i and for each $P \in \mathcal{P}_i$ with $k(P) \neq i$ we will require the following:*

1. *$k(P) \neq k(Q)$ for all $Q \in \mathcal{P}_i$.*
2. *There exists $Q \in \mathcal{P}_{k(P)}$ such that $k(Q) = i$.*

The first condition guarantees that two elements of the same partition do not have the same element of S associated. The second condition guarantees that the holes are set consistently (i.e. if T_i has a hole to T_j, then T_j must have a hole to T_i). The action of σ is then defined as:

$$\sigma(i,x) = (k(P(i,x)), \tau_i(x)),$$

where:

$$\tau_i(x) = \begin{cases} \phi_{P,Q}, & k(P) \neq i, \\ x, & k(P) = i, \end{cases}$$

where $P = P(i, x)$ *and* $Q \in \mathcal{P}_{k(P)}$ *is such that* $k(Q) = i$ *(which exists as consequence of item 2 above).*

Example 5.7.4. *(An infinite family of maps) As stated above, the setting allows* S *to be infinite. For example, let* $S = [\gamma, 2]$*, where* γ *is a parameter and let* T_s *defined by:*

$$T_s(x) = \begin{cases} sx & x \leq \frac{1}{2}, \\ 2x - 1 & x > \frac{1}{2}, \end{cases}$$

Let $\widetilde{\Omega} = I$ *and* $\sigma(x, y) = (2x \mod 1, y)$ *for* $(x, y) \in \Omega$, $\xi_1(x, y) = \phi_{I,S}(x)$ *and* $\xi_2(x, y) = y$*. The resulting dynamical system is essentially the second family of maps studied on reference [65] (see Section IV in that reference). The only difference resides on the nature of the selection process. While in that paper, the selection process is an IID random process, uniformly distributed on* S*, here the selection process is the dyadic transformation* $2x \mod 1$*.*

Under these conditions, the question of obtaining the critical value γ_c *such that for all* $\gamma > \gamma_c$*, the system has an acim, can be still answered with the results of [68]. Specifically, one needs to solve the following expression for* γ_c*:*

$$-\int_0^1 \log\left((2 - \gamma_c)\omega + \gamma_c\right) d\omega = 0. \tag{5.16}$$

The solution $\gamma_c \approx 0.262583$ *coincides with the one obtained on [65] for an IID selection process.*

5.8. Approximation of the invariant density

If we restrict ourselves to the case where $S = \{1, 2\}$, then we can obtain an approximation of the invariant measure as follows. The maps $\widetilde{T}_{CDS-1}$ and $\widetilde{T}_{CDS-2}$, obtained in Section 5.6 are maps on the interval whose Perron-Frobenius operator, that we will denote $P_{\widetilde{T}}$, is well defined. Therefore, we can use Ulam's method (see for example Chapter 6.1 from [18]) to obtain an approximation of the fixed point of $P_{\widetilde{T}}$ (and hence, an approximation of the acim). Let $\widetilde{u}$ be such approximation. Then $U : \Omega \to \mathbb{R}$ defined by:

$$U(i, x) = \widetilde{u}(\phi_{I,I_i}(x)), \tag{5.17}$$

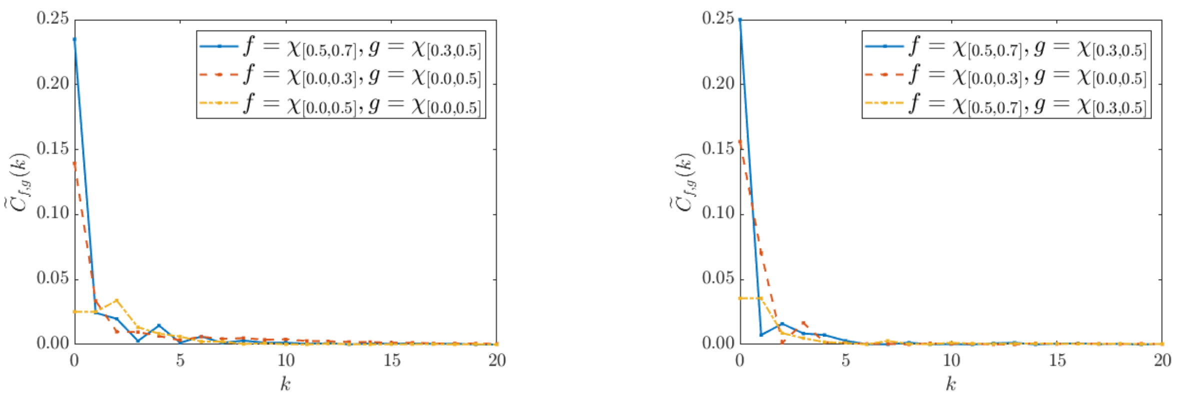
is an approximation of the fixed point of P_Θ (and hence, an approximation of the acim). Recall from Section 5.6 that the definition of ϕ is given by:

$$\phi_{I,I_i}(x) = \min(I_i) + \frac{(\max(I_i) - \min(I_i))(x - \min(I))}{\max(I) - \min(I)}. \tag{5.18}$$

The method above can be extended to the case of n maps, but it cannot be further extended to the case of infinite maps.

5.9. Numerical explorations on statistical properties

Due to the fact that we still do not have analytical results that tell us how the spectrum of the Perron-Frobenius operator behaves on the complex plane, we cannot apply the spectral method to obtain statistical properties. Therefore, we studied some of the statistical properties numerically. For this purpose, we used two systems:



(a) (b)

Figure 5.11: Estimator of correlation coefficients given by $\widetilde{C}_{f,g}(k)$. f and g are indicator functions of different intervals. (a) The system given by (5.19). (b) The system given by (5.20) where $H_2 = [0.97, 1]$. In both cases $\widetilde{C}_{f,g}(k)$ tends to zero when k tends to infinity.

1. The system studied in Example 5.6.1, namely, the CDS-1 given by:

$$T_1(x) = \begin{cases} 2x & x \in \left[0, \frac{1}{2}\right), \\ 2x - 1 & x \in \left[\frac{1}{2}, 1\right]. \end{cases} \qquad T_2(x) = \begin{cases} 3x & x \in \left[0, \frac{1}{3}\right), \\ -x + 4/3 & x \in \left[\frac{1}{3}, \frac{2}{3}\right), \\ -2x + 2 & x \in \left[\frac{2}{3}, 1\right]. \end{cases} \tag{5.19}$$

   and holes $H_1 = [0, 1/10]$ and $H_2 = [19/20, 1]$.

2. The system presented in Section 5.4, namely, the CDS-2 given by:

$$\begin{aligned} T_1(x) &= 4x(1-x), \\ T_2(x) &= \begin{cases} \theta|x - 0.5|^{\alpha}, & x < 0.5, \\ 1 - \theta|x - 0.5|^{\alpha}, & x \geq 0.5, \end{cases} \end{aligned} \tag{5.20}$$

   where $\theta = 109/64$ and $\alpha = 51/64$ are fixed parameters and $H_1 = [0, 0.01]$ and $H_2 = [\beta, 1]$. We used several values for β in order to explore the transition between the chaotic and the ordered regimes. We used $\beta = 0.97$ to illustrate the chaotic regime (see Figure 5.2) and we used $\beta \in \{0.01254, 0.01255, 0.01256, 0.01257\}$ to illustrate the ordered regime. As shown in Figure 5.3, the periodic cycle of period six appears in the system around these values.

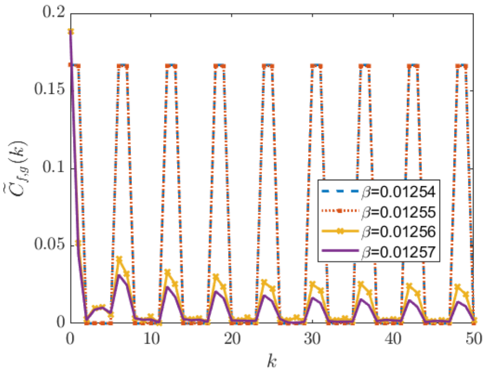
5.9.1 Decay of Correlations

We used the same estimator for the coefficients of correlation given by Equation (1.23) in Section 1.4.2. The estimator is given by:

$$\widetilde{C}_{f,g}(k) = \frac{1}{n}\sum_{i=0}^{n-1} f\left(T^i x\right) \cdot g\left(T^{i+k} x\right) - \frac{1}{n^2}\left(\sum_{i=0}^{n-1} f\left(T^i x\right)\right)\left(\sum_{i=0}^{n-1} g\left(T^i x\right)\right), \tag{5.21}$$

where $n + k$ is the size of the sample. Figure 5.11 shows how this estimator behaves with indicator functions of different intervals. The figure shows the numerical simulations with three combinations of f and g: (1) $f = \chi_{[0.5,0.7]}$ and $g = \chi_{[0.3,0.5]}$ (2) $f = \chi_{[0,0.3]}$ and $g = \chi_{[0,0.5]}$ (3) $f = \chi_{[0,0.5]}$ and $g = \chi_{[0,0.5]}$. In all cases, we numerically observed that $\widetilde{C}_{f,g}(k)$ tends to 0 when k tends to infinity.

We will now explore the transition from chaos to order in the second system (5.20). As commented in Section 5.4, this transition occurs around $\beta = 0.0125$. Figure 5.12 shows how the correlations behave with different values of β around that value. The transition seems to occur abruptly after $\beta = 0.01255$.

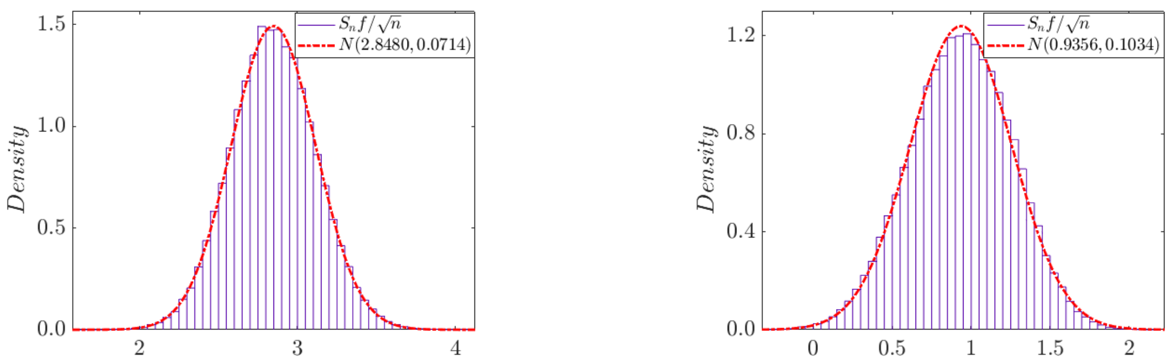


(a)

Figure 5.12: Correlation function for the system given by (5.20). Correlations do not decay for $\beta = 0.01254$ nor $\beta = 0.01255$. There seems to be an abrupt change between $\beta = 0.01255$ and $\beta = 0.01256$.

5.9.2 Central Limit

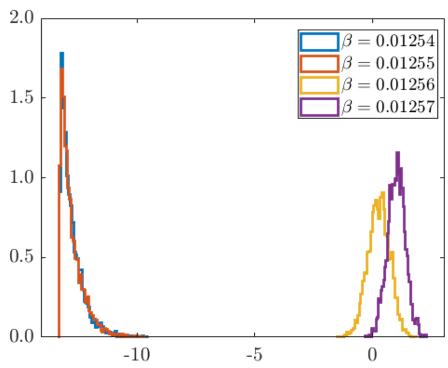
In this section, we use an observable given by $f(x) = x - \frac{1}{2}$. As we did in Section 1.4.3, we plotted a normalized histogram of 10000 realizations of $\frac{S_n f}{\sqrt{n}}$. The limit distribution for systems (5.19) and (5.20) with $H_2 = [0.97, 1]$ can be fitted to a normal distribution. This scenario corresponds to the chaotic regime (see Figure 5.2).



(a) (b)

Figure 5.13: Fitting to a normal distribution. (a) The system given by (5.19). (b) The system given by (5.20) with $H_2 = [0.97, 1]$. In both cases, a normal distribution fits the data adequately.

The corresponding histograms for the ordered regime are shown in Figure 5.14. The figure shows the histograms of the system with different values of β. A transition seems to occur between $\beta = 0.0125$ and $\beta = 0.0126$. When $\beta \geq 0.01256$, the histogram can be fitted to a normal distribution. But when $\beta \leq 0.01255$, the histogram becomes highly non-symmetric and skewed. Note the transition is quite fast: a change in 0.00001 in the parameter β results in a completely different histogram.

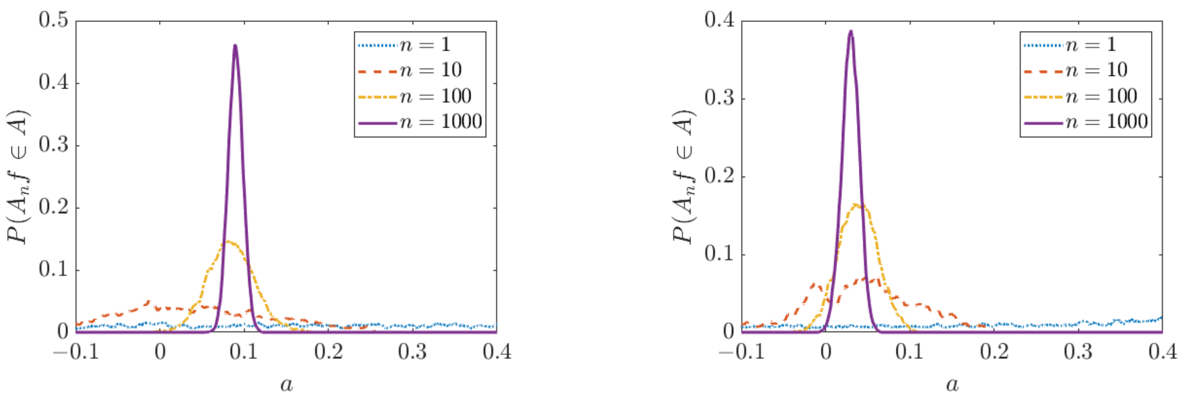


(a)

Figure 5.14: Distribution of 10000 realizations of $\frac{S_n f}{\sqrt{n}}$ with $f(x) = x - \frac{1}{2}$. The system is given by Equations (5.20) with different values of β. Similar to the correlations shown in Figure 5.12 , this figure shows that there is an abrupt change in the form of the histogram between $\beta = 0.01255$ and $\beta = 0.01256$.

5.9.3 Large Deviation Principle

When a Large Deviation Principle holds, the rate function $I(a)$ allows us to estimate the probability that $A_n f \in A$ for each $n \geq 0$ and $A = [a - \epsilon, a + \epsilon]$ with small values of ϵ. In figure 5.15 we obtained these probabilities empirically for different values of n and a with $f(x) = x - \frac{1}{2}$ and $\epsilon = 0.005$. The figure shows that, as n increases, the probability concentrates around its expected value. These probabilities can be used to obtain an approximation of the rate function $I(a)$ using expression (1.27).



(a) (b)

Figure 5.15: Numerical estimation of the probability $\mathbb{P}(A_n f \in A)$ where $A = [a - \epsilon, a + \epsilon]$. The figure shows different values of n and different values of a. For large values of n, the probability is concentrated around its expected value. (a) The system given by (5.19) and (b) the system given by (5.20). In both cases, we used $f(x) = x - \frac{1}{2}$, $\epsilon = 0.005$, .

6. Contributions and possible future work

6.1. Contributions

- Published paper where we prove a Gaussian concentration inequality and a rate of convergence for the Birkhoff ergodic theorem for subshifts with countable alphabet and general observables. The result is used to obtain a rate of convergence of the ergodic theorem for this type of systems.

  Cesar Maldonado, Humberto Muñiz y Hugo Nieto Loredo (2021) *Concentration inequalities and rates of convergence of the ergodic theorem for countable shifts with Gibbs measures*, Journal of Difference Equations and Applications,
  27:11, 1594-1607, DOI: 10.1080/10236198.2021.2000970.

- Manuscript ready to be submitted where we introduce connected dynamical systems and report interesting phenomena that motivate their study such as induced order, phase transitions, etc.

  Cesar Maldonado, Hugo Nieto Loredo y Ricardo A. Pérez Otero (2024) *Critical behavior in connected dynamical systems: a numerical approach*,

- Submitted for publication where we prove a Central Limit Theorem and a Berry-Esseen inequality for contractive random dynamical systems. Part of this work was also presented in the 55th Congress of the Mexican Mathematical Society in October 2022.

  Cesar Maldonado y Hugo Nieto Loredo (2024) *Central Limit Theorem for a Class of Contractive Random Dynamical Systems*,
  Journal of Difference Equations and Applications,

6.2. Possible future work

6.2.1 Other Statistical Properties

As commented in Section 1.3, the spectral gap property of the Perron-Frobenius operator is a key ingredient to prove many statistical properties of the dynamics. An option for future work is to prove other statistical properties using the same spectral gap result from [28]. For example, in order to prove a concentration inequality for this type of systems, we can try to adapt the technique presented in [73]. Similarly, we can try to adapt the technique from [2] to prove a large deviation principle.

6.2.2 Improve spectral gap result

Instead of focusing on proving other statistical properties, another possible future work would be the improvement of the spectral gap result from [28]. Condition C2 presented in section 3.1 means that the dynamics needs to be uniformly contractive because the same value α must work for all $\underline{x} \in \Sigma_A^+$ and for all $z_1, z_2 \in K$. This condition is quite restrictive and a possible future work can be focused on finding more relaxed conditions under which the spectral gap result continues to be true.

6.2.3 Perron-Frobenius operator for connected dynamical systems

Regarding the connected dynamical systems presented in Chapter 5, a possible direction for future work is to carry out a rigorous study of the invariant density. Such an invariant density should be obtained as a fixed point of the Perron-Frobenius operator associated to the skew-product (5.13). A starting point in this direction would be to consider the single map in the interval that we presented in Section 5.6. The idea here is to use the fact that the Perron-Frobenius operator for the single map can be obtained explicitly.

6.2.4 Theoretical results in connected dynamical systems

The numerical explorations of the statistical properties of connected dynamical systems that were presented in Section 5.9 can be studied from a rigorous point of view. The first step would be to show the existence of a stationary measure for the skew product (5.13). We believe that a technique similar to the one that was used in reference [68] can be used for this purpose.

7. Appendix

7.1. Statement and proof of perturbation theorem

Perturbation theorem 7.1.2 is a result that allows to obtain key properties of a perturbed quasi-compact operator that are used on the proof of a central limit theorem. Roughly speaking, the theorem establishes regularity conditions for a perturbed operator to inherit quasicompactness from the unperturbed operator when the perturbation is small. This important result was stated in [39] and in this section we organize the proof on several preliminary lemmas and provide details and comments not stated in the original reference. Before presenting the theorem, let us introduce some ideas that will be needed.

7.1.1 Spectral theory

Lemma 7.1.1. *Suppose that $Q : \mathcal{B} \to \mathcal{B}$ is a bounded linear operator and there exists closed Q-invariant subspaces F and H such that $\mathcal{B} = F \oplus H$. Then* $\mathrm{spec}(Q) = \mathrm{spec}(Q|_F) \cup \mathrm{spec}(Q|_H)$.

Proof. We will use contrapositive for both implications.

($\Leftarrow$) Suppose that $\lambda \notin \mathrm{spec}(Q)$. Then $Q - \lambda$ is invertible and there exists a bounded linear operator $T : \mathcal{B} \to \mathcal{B}$ that satisfies the following:

$$T(Q - \lambda)u = (Q - \lambda)Tu = u \qquad \forall u \in \mathcal{B} \tag{7.1}$$

If $u' \in F$ then $(Q - \lambda)u' \in F$. Due to the invertibility of $Q - \lambda$, we obtain that $Tu' \in F$ and therefore $T|_F : F \to F$ is a well defined, bounded linear operator that satisfies:

$$T|_F(Q|_F - \lambda)u' = (Q|_F - \lambda)T|_F u' = u' \qquad \forall u' \in F \tag{7.2}$$

Hence, $Q|_F - \lambda$ is invertible and $\lambda \notin \mathrm{spec}(Q|_F)$. Replacing F by H on the above, we can also conclude that $\lambda \notin \mathrm{spec}(Q|_H)$.

($\Rightarrow$) Suppose that $\lambda \notin \mathrm{spec}(Q|_F)$ and $\lambda \notin \mathrm{spec}(Q|_H)$. Then $(Q|_F - \lambda)^{-1}$ and $(Q|_H - \lambda)^{-1}$ are well defined on F and H respectively. Define $T : \mathcal{B} \to \mathcal{B}$ by:

$$Tu = (Q|_F - \lambda)^{-1}u' + (Q|_H - \lambda)^{-1}u''$$

where $u = u' + u''$ with $u' \in F$ and $u'' \in H$. Using the fact that $(Q - \lambda)u = (Q|_F - \lambda)u' + (Q|_H - \lambda)u''$ we obtain that T is a bounded linear operator that satisfies $T(Q - \lambda)u = (Q - \lambda)Tu = u$ for all $u \in \mathcal{B}$. Hence, $Q - \lambda$ is invertible and $\lambda \notin \mathrm{spec}(Q)$. $\square$

We say that $Q : \mathcal{B} \to \mathcal{B}$ **has s simple leading eigenvalues** if there exist closed subspaces, F and H such that

i) $\mathcal{B} = F \oplus H$,

ii) $QF \subseteq F$, $QH \subseteq H$,

iii) $\dim(F) = s$ and $Q|_F$ has s simple eigenvalues λ_k, $k = 1, ..., s$,

iv) $\operatorname{spr}(Q|_H) < \min\{|\lambda_k| : k = 1, ..., s\}$,

When Q has 1 simple leading eigenvalue, the operator is quasicompact. Conversely, when an operator Q is quasicompact with $\dim F = 1$, the operator also has one simple leading eigenvalue. Indeed, recall that the spectrum of an n-dimensional operator consists of, at most, n isolated eigenvalues. It follows that $\operatorname{spec}(Q|_F) = \{\lambda\}$. Using the fact that $\operatorname{spec}(Q) = \operatorname{spec}(Q|_F) \cup \operatorname{spec}(Q|_H)$ (see Lemma 7.1.1) we get:

$$\operatorname{spr}(Q) = \sup\{\operatorname{spec}(Q)\} = \max\{\operatorname{spr}(Q|_F), \operatorname{spr}(Q|_H)\} = |\lambda| > \operatorname{spr}(Q|_H).$$

The above shows that the operator Q satisfies the definition of quasicompactness.

7.1.2 Notions of derivatives on normed spaces

Let $\mathcal{B}_1$ and $\mathcal{B}_2$ two normed spaces and $U \subseteq \mathcal{B}_1$ be an open subset of $\mathcal{B}_1$. Let $\mathcal{B}(U, \mathcal{B}_2)$ denote the collection of all bounded operators from U to $\mathcal{B}_2$. An operator $Q \in \mathcal{B}(U, \mathcal{B}_2)$ is said to be **differentiable at** $u \in U$ if there exists a bounded linear operator $DQ_u : \mathcal{B}_1 \to \mathcal{B}_2$ such that:

$$\lim_{\|\Delta\| \to 0} \frac{\|Q(u + \Delta) - Q(u) - DQ_u(\Delta)\|}{\|\Delta\|} = 0. \tag{7.3}$$

DQ_u is called the **derivative of** Q **at** u. If Q is differentiable at every $u \in U$ and the map $u \mapsto DQ_u$ is continuous (with respect to the corresponding norms) [1], the operator Q is said to be of class C^1.

Higher order derivatives are defined similar: Q is said to be **twice differentiable at** $u \in U$ if the map $u \mapsto DQ_u$ is differentiable at $u \in U$, i.e. if there exists a bounded linear operator $D^2Q_u : \mathcal{B}_1 \to \mathcal{B}(\mathcal{B}_1, \mathcal{B}(\mathcal{B}_1, \mathcal{B}_2))$ such that:

$$\lim_{\|\Delta\| \to 0} \frac{\left\|DQ_{u+\Delta} - DQ_u - D^2Q_u(\Delta)\right\|}{\|\Delta\|} = 0. \tag{7.4}$$

The operator Q is said to be of class C^2, if Q is twice differentiable at every $u \in U$ and the map $u \mapsto D^2Q_u$ is continuous. Q is said to be **analytic** if Q is of class C^∞. If Q is a bounded linear operator itself, then it is also of class C^∞, its first derivative is the operator itself and the higher order derivatives are all 0 (see Propositions 3.6 and 3.8 from [56]).

If the operator Q depends on several variables, the partial derivatives with respect to each of them is defined by fixing the other variables separately. For example, let U and V be open subsets of normed spaces and $Q : U \times V \to \mathcal{B}_2$. The **partial derivative of** Q **at** $(t_0, h_0) \in U \times V$ **with respect to** t is simply the derivative of the operator $Q_{h_0} : U \to \mathcal{B}_2$ defined by $Q_{h_0}(t) = Q(t, h_0)$. Q is of class C^k if and only if each partial derivative exists and is of class C^{k-1} (Proposition 3.5 from [56]).

We denote $\mathcal{B}^*$ the dual space of $\mathcal{B}$. Recall that the dual operator of Q, $Q^* : \mathcal{B}^* \to \mathcal{B}^*$ is defined by $Q^*\varphi(u) = \langle \varphi, Qu \rangle$.

7.1.3 Statement of perturbation theorem

We state and prove the theorem in the case in which the operator has one simple leading eigenvalue. As stated above, a quasicompact operator with $\dim F = 1$ has one simple leading eigenvalue. This is suitable for our purposes as the pushforward measure that we are interested in this thesis, has this property.

Theorem 7.1.2 ([39])**.** *Let $\mathcal{L}_\mathcal{B}$ denote the family of bounded linear operators acting on a Banach space $\mathcal{B}$. Let I be an open interval centered at* 0 *and $\{Q(t)\}_{t \in I}$ be a family of operators that belong to $\mathcal{L}_\mathcal{B}$ such that:*

[1] Note that this is different than requesting that $DQ_u : \mathcal{B}_1 \to \mathcal{B}_2$ is continuous.

C1) $t \mapsto Q(t)$ *is of class* C^k *with* $k > 0$.

C2) $Q(0)$ *has one simple leading eigenvalue and* $\operatorname{spr}(Q(0)) = 1$.

Then there exist an open interval centered at 0, $I_0 \subseteq I$ *and functions of class* C^k: $\lambda(t)$, $v(t)$, $\varphi(t)$, $N(t)$ *taking values respectively on* $\mathbb{C}$, $\mathcal{B}$, $\mathcal{B}^*$ *and* $\mathcal{L}_{\mathcal{B}}$ *such that, for* $t \in I_0$:

i) $Q(t)v(t) = \lambda(t)v(t)$ *and* $Q^*(t)\varphi(t) = \lambda(t)\varphi(t)$.

ii) $\langle\varphi(t), v(t)\rangle = 1$.

iii) $Q^n(t) = \lambda^n(t)\langle\varphi(t), \cdot\rangle v(t) + N^n(t)$.

iv) $\|N^n(t)\| \leq c|\lambda|^n$.

7.1.4 Preliminary lemmas

As commented above, we divide the proof presented in [39] into several lemmas and provide further comments not stated on the original reference.

Lemma 7.1.3. *Let* V_0 *be a bounded linear operator and suppose there exists* r *such that* $\operatorname{spr}(V_0) < r$. *Then there exists* $\eta > 0$ *and* $c > 0$ *such that if* $\|V - V_0\| < \eta$, *then* $\|V^n\| \leq cr^n$ *for all* $n \geq 1$.

Proof. Let us first show that $\|V - V_0\| \to 0$ implies $\|V^n - V_0^n\| \to 0$. Let $\eta > 0$ and suppose $\|V - V_0\| < \eta$. Then:

$$\begin{aligned}\left\|V^2 - V_0^2\right\| &= \|(V - V_0)V_0 + V_0(V - V_0) + (V - V_0)(V - V_0)\| \\ &\leq \|V - V_0\|\|V_0\| + \|V_0\|\|V - V_0\| + \|V - V_0\|^2 \\ &< \eta\,(2\|V_0\| + \eta)\,.\end{aligned}$$

By induction, one obtains that for each $n, \eta > 0$ there exists $\delta(n, \eta) > 0$ such that if $\|V - V_0\| < \eta$ then $\|V^n - V_0^n\| < \delta(n, \eta)$, with $\delta(n, \eta) \to 0$ as $\eta \to 0$. Choose n_0 such that the strict inequality $r > \|V_0^{n_0}\|^{1/n_0}$ holds [2] and set $\eta > 0$ small enough so that $0 < \delta(n_0, \eta) < r^{n_0} - \|V_0^{n_0}\|$. If $\|V - V_0\| < \eta$ then we have by the triangle inequality that $\|V\| < \|V_0\| + \eta$, $\|V^{n_0}\| < \|V_0^{n_0}\| + \delta(n_0, \eta)$ and:

$$\|V^{n_0}\| < \|V_0^{n_0}\| + \delta(n_0, \eta) < r^{n_0}.$$

For $n \geq 1$, set $n = kn_0 + l$ with $0 \leq l < n_0$:

$$\|V^n\| \leq \|V^{n_0}\|^k\|V\|^l < r^{kn_0}\|V\|^l = r^n\left(\frac{\|V\|}{r}\right)^l < r^n\left(\frac{\|V_0\| + \eta}{r}\right)^l \leq cr^n.$$

where $c = \max\left\{\left(\dfrac{\|V_0\| + \eta}{r}\right)^l : l = 0, ..., n_0 - 1\right\}$. □

The following lemma provides a characterization for the operator Q having one simple leading eigenvalue.

Lemma 7.1.4. *Let* $\mathcal{B}$ *be a normed vector space and* $Q : \mathcal{B} \to \mathcal{B}$ *be a bounded, linear operator,* $\lambda \in \mathbb{C}$, $v \in \mathcal{B}$ *and* $\varphi \in \mathcal{B}^*$ *such that:*

[2] It is always possible to choose such an n_0 because $r > \operatorname{spr}(V_0) = \inf_n\|V_0^n\|^{1/n}$. By contradiction: if such an n_0 does not exist, then $r \leq \|V_0^n\|^{1/n}$ for all $n \geq 1$, this means that r is a lower bound of $\|V_0^n\|^{1/n}$, but $\operatorname{spr}(V_0)$ is the greatest of all lower bounds of $\|V_0^n\|^{1/n}$, so it follows that $\operatorname{spr}(V_0) \geq r$, contradicting the hypothesis of the lemma.

A1. $Qv = \lambda v$ *and* $Q^*\varphi = \lambda\varphi$

A2. $\langle\varphi, v\rangle = 1$

A3. If $N = Q - \lambda\langle\varphi, \cdot\rangle v$ *then* $\operatorname{spr}(N) < |\lambda|$.

Then the operator Q *has* 1 *simple leading eigenvalue, i.e. there exist closed* Q*-invariant subspaces* F *and* H *such that* $\mathcal{B} = F \oplus H$*, with* $\dim F = 1$*,* $Q|_F(v) = \lambda v$ *and* $\operatorname{spr}(Q|_H) < |\lambda|$*. Conversely, if the operator* Q *has 1 simple leading eigenvalue, then there exist* $\lambda \in \mathbb{C}$*,* $v \in \mathcal{B}$ *and* $\varphi \in \mathcal{B}^*$ *that satisfy assertions 1-3 above.*

Proof. ($\Rightarrow$) Let $F = \{u \in \mathcal{B} : (Q - \lambda I)u = 0\}$ and $H = \{h \in \mathcal{B} : \langle\varphi, h\rangle = 0\}$. Let us explain why $\mathcal{B} = F \oplus H$. F is the eigenspace associated to the eigenvalue λ, so F is formed with elements of the form $u = z \cdot v$ with $z \in \mathbb{C}$. So, each $u \in \mathcal{B}$ can be expressed as $u = \langle\varphi, u\rangle v + (u - \langle\varphi, u\rangle v)$ where $\langle\varphi, u\rangle v \in F$ and $(u - \langle\varphi, u\rangle v) \in H$:

$$\langle\varphi, u - \langle\varphi, u\rangle v\rangle = \langle\varphi, u\rangle - \langle\varphi, u\rangle\langle\varphi, v\rangle = \langle\varphi, u\rangle - \langle\varphi, u\rangle = 0.$$

If $u \in F$, then u can be written as $u = z \cdot v$ with $v \in F$ and it is clear that $Qu \in F$. Also, if $h \in H$ then $\langle\varphi, h\rangle = 0$ and:

$$\langle\varphi, Qh\rangle = Q^*\varphi(h) = \lambda\varphi(h) = \lambda\langle\varphi, h\rangle = 0,$$

and therefore $Qh \in H$. Due to the fact that $Qv = \lambda v$, we have that $v \in F$, so $Q|_F(v) = \lambda v$. Let $n \geq 1$. For all $h \in H$ we have: $\|Q|_H^n h\| \leq \|Q|_H^n\|\|h\|$ where $\|Q|_H^n\|$ is the smallest constant that satisfies $\|Q|_H^n h\| \leq C\|h\|$ for all $h \in H$. Also, for all $u \in \mathcal{B}$ $\|N^n u\| \leq \|N^n\|\|u\|$. But $H \subseteq \mathcal{B}$ so $\|N^n h\| \leq \|N^n\|\|h\|$ for all $h \in H$. Using the fact that $Qh = Nh$ we get $\|Q|_H^n h\| \leq \|N^n\|\|h\|$ for all $h \in H$. But $\|Q|_H^n\|$ is the smallest of all constants that satisfy that inequality on H, so, we must have that $\|Q|_H^n\| \leq \|N^n\|$. This implies that $\|Q|_H^n\|^{1/n} \leq \|N^n\|^{1/n}$ for all $n \geq 1$, which in turn, implies that $\operatorname{spr}(Q|_H) \leq \operatorname{spr}(N) < |\lambda|$.

($\Leftarrow$) Suppose that Q has 1 simple leading eigenvalue. It is clear that $Qv = \lambda v$. As $\mathcal{B} = F \oplus H$, each $u \in \mathcal{B}$ can be written in a unique way as $u = u' + u''$ with $u' \in F$ and $u'' \in H$. Also, as $\dim F = 1$, for each $u' \in F$, there exists $z \in \mathbb{C}$ such that $u' = zv$. Let $\varphi \in \mathcal{B}^*$ be the linear functional defined by $\langle\varphi, u\rangle = z$. By definition, $\langle\varphi, v\rangle = 1$ and:

$$\langle Q^*\varphi, u\rangle = \langle\varphi, Qu' + Qu''\rangle = \langle\varphi, z\lambda v + Qu''\rangle = \lambda z = \lambda\langle\varphi, u\rangle,$$

and we get $Q^*\varphi = \lambda\varphi$. Note that $u' \in F \Rightarrow Nu' = 0$ and $u'' \in H \Rightarrow \langle\varphi, u''\rangle v = 0$ so $N = Q\Pi_H$ where $\Pi_H : \mathcal{B} \to H$ is defined by $\Pi_H u = u''$. Therefore $\|N^n\| \leq \|Q|_H^n\|\|\Pi_H\|$ and $\|N^n\|^{1/n} \leq \|Q|_H^n\|^{1/n}\|\Pi_H\|^{1/n}$ (recall that Π_H is idempotent, i.e. $\Pi_H^2 = \Pi_H$). Then $\operatorname{spr}(N) \leq \operatorname{spr}(Q|_H) < |\lambda|$. $\square$

Lemma 7.1.5. *Under the assumptions A1-A2 of Lemma 7.1.4, suppose the function* $t \mapsto Q(t)$ *is continuous at* 0 *and let* $0 < \epsilon < |\lambda|$*. If* $h \in H$ *satisfies:*

$$\|h\| < \frac{\epsilon\|v\|}{\epsilon + 2\|\varphi\|\|v\|\|Q(0)\|}. \tag{7.5}$$

Then there exists $\delta(\epsilon) > 0$ *such that if* $|t| < \delta(\epsilon)$ *then* $\langle\varphi, Q(t)(v + h)\rangle \neq 0$.

Proof. First note that for all $h \in H$ we have $\langle\varphi, v + h\rangle = \langle\varphi, v\rangle = 1$ and therefore:

$$\langle\varphi, Q(0)(v + h)\rangle = \langle Q^*(0)\varphi, v + h\rangle = \lambda\langle\varphi, v + h\rangle = \lambda.$$

The operator $Q(t)$ is continuous with respect to t, it means, for all $\epsilon_1 > 0$ there exists $\delta_1(\epsilon_1) > 0$ such that $|t| < \delta_1(\epsilon_1) \Rightarrow \|Q(t) - Q(0)\| < \epsilon_1$. So, given $\epsilon > 0$, choose $\epsilon_1 < \frac{\epsilon}{2\|\varphi\|\|v\|}$ and $|t| < \delta_1(\epsilon_1)$. It follows that $\|Q(t) - Q(0)\| < \frac{\epsilon}{2\|\varphi\|\|v\|}$ and:

$$\begin{aligned}
|\langle \varphi, Q(t)(v+h)\rangle - \lambda| &= |\langle \varphi, Q(t)v\rangle + \langle \varphi, Q(t)h\rangle - \langle \varphi, \lambda v\rangle| = |\langle \varphi, Q(t)v - \lambda v\rangle + \langle \varphi, Q(t)h\rangle| \\
&\leq |\langle \varphi, Q(t)v - Q(0)v\rangle| + |\langle \varphi, Q(t)h\rangle| \\
&\leq \|\varphi\|\|Q(t) - Q(0)\|\|v\| + \|\varphi\|\|Q(t)\|\|h\| \\
&\leq \|\varphi\|\|Q(t) - Q(0)\|\|v\| + \|\varphi\|\|h\| \left(\|Q(t) - Q(0)\| + \|Q(0)\|\right) \\
&< \frac{\epsilon}{2} + \|h\| \left(\frac{\epsilon}{2\|v\|} + \|\varphi\|\|Q(0)\| \right) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,
\end{aligned}$$

where we have used the fact that $|\langle \varphi, u\rangle| \leq \|\varphi\|\|u\|$ and $\langle \varphi, v\rangle = 1$. Set $\delta(\epsilon) = \delta_1(\epsilon_1)$. If $|\langle \varphi, Q(t)(v+h)\rangle| \geq |\lambda|$ the lemma is true directly. If $|\langle \varphi, Q(t)(v+h)\rangle| < |\lambda|$ we have $|\lambda| - |\langle \varphi, Q(t)(v+h)\rangle| < \epsilon$ which implies $|\langle \varphi, Q(t)(v+h)\rangle| > |\lambda| - \epsilon$. Therefore any $\epsilon < |\lambda|$ will guarantee that $\langle \varphi, Q(t)(v+h)\rangle \neq 0$. □

In what follows, fix an ϵ for the application of lemma 7.1.5 and assume that $|t| < \delta(\epsilon)$ and (7.5) is satisfied. Let us denote:

$$\begin{aligned}
t_\epsilon &= \{|t| < \delta(\epsilon)\}. \\
h_\epsilon &= \{h \in H : (7.5) \textit{ is satisfied}\}.
\end{aligned}$$

Lemma 7.1.6. *Under the assumptions of Lemma 7.1.5, assume further that $Q(t)$ is a family of bounded linear operators (i.e. $Q(t)(u+v) = Q(t)u + Q(t)v$) for all $u, v \in \mathcal{B}$) and let $\mathcal{R} : t_\epsilon \times h_\epsilon \to H$ be defined by:*

$$\mathcal{R}(t,h) = \frac{Q(t)(v+h)}{\langle \varphi, Q(t)(v+h)\rangle} - (v+h). \tag{7.6}$$

If the map $t \mapsto Q(t)$ is of class C^k with $k > 0$ then $\mathcal{R}$ is also of class C^k.

Proof. It is enough to show that the numerator and denominator of the above expression are both of class C^k (with respect to both, $t \in t_\epsilon$ and $h \in h_\epsilon$). Let us show that the numerator is of class C^k. The same arguments can be used to show that the denominator is also of class C^k.

- $Q(t)(v + \cdot)$ **is of class** C^k **with respect to** h: As $Q(t)$ is a bounded linear operator on $\mathcal{B}$, the derivative of the map $h \mapsto Q(t)(v+h)$ (i.e. the partial derivative of $Q(t)(v+h)$ with respect to h) is $Q(t)$ and the higher order derivatives are zero. All these derivatives are continuous (this can be seen from the fact that the maps $h \mapsto Q(t)$ and $h \mapsto 0$ do not depend on h), so $Q(t)(v + \cdot)$ is of class C^∞ with respect to h.

- $Q(\cdot)(v+h)$ **is of class** C^k **with respect to** t: Let $DQ(t), D^2Q(t), ..., D^kQ(t)$ be the first k derivatives of the map $t \mapsto Q(t)$ [3]. By hypothesis, these derivatives are continuous. Then $\lim_{|\Delta| \to 0} \frac{\|Q(t+\Delta) - Q(t) - DQ(\Delta)\|}{|\Delta|} = 0$ and:

$$\begin{aligned}
&\lim_{|\Delta| \to 0} \frac{\|Q(t+\Delta)(v+h) - Q(t)(v+h) - DQ(\Delta)(v+h)\|}{|\Delta|} \\
&\qquad \leq \|v+h\| \lim_{|\Delta| \to 0} \frac{\|Q(t+\Delta) - Q(t) - DQ(\Delta)\|}{|\Delta|} = 0.
\end{aligned}$$

  And therefore, the derivative with respect to t of the map $t \mapsto Q(t)(v+h)$ is the map $DQ(t)(v+h)$. The continuity of this map follows from the continuity of $t \mapsto DQ(t)$.

[3] Note all the derivatives are bounded and linear on $\mathcal{B}$, i.e. the normed space $\mathcal{B}_2$ from the definition of the derivative on Section 7.1.2 is the set of all bounded linear operators in $\mathcal{B}$.

Indeed, for a given $\varepsilon > 0$ set $\varepsilon_1 < \frac{\varepsilon}{\|v+h\|}$ and let $\delta(\varepsilon_1)$ such that $|t-t_0| < \delta(\varepsilon_1) \Rightarrow \|DQ(t) - DQ(t_0)\| < \varepsilon_1$. Then we get:

$$\|DQ(t)(v+h) - DQ(t_0)(v+h)\| \leq \|DQ(t) - DQ(t_0)\|\|v+h\| \leq \varepsilon_1\|v+h\| < \varepsilon.$$

We can do the same for the other derivatives $D^2Q(t), ..., D^kQ(t)$ and conclude that $Q(\cdot)(v+h)$ is of class C^k with respect to t.

□

Lemma 7.1.7. *Under the assumptions of Lemma 7.1.6, assume further that the map $t \mapsto Q(t)$ is of class C^k with $k > 0$. Then there exists an open neighborhood of 0, denoted $t'_\epsilon \subseteq t_\epsilon$ and functions of class C^k, $\lambda(t) \in \mathbb{C}$, $v(t) \in \mathcal{B}$ and $h(t) \in h_\epsilon$ such that $h(0) = 0$ for $t \in t'_\epsilon$ we have that $\mathcal{R}(t, h(t)) = 0$ and $Q(t)v(t) = \lambda(t)v(t)$.*

Proof. The partial derivative of $\mathcal{R}$ with respect to h at point $(0,0)$ must satisfy:

$$\lim_{\|\Delta\|\to 0} \frac{\|\mathcal{R}(0,\Delta) - \mathcal{R}(0,0) - D\mathcal{R}(0,\Delta)\|}{\|\Delta\|} = 0$$

Note that $\mathcal{R}(0,0) = 0$ and for $h \in H$:

$$\mathcal{R}(0,h) = \frac{Q(0)(v+h)}{\langle\varphi, Q(0)(v+h)\rangle} - (v+h) = \frac{\lambda v + Q(0)h}{\lambda} - v - h = \frac{Q(0)h}{\lambda} - h$$

So, the partial derivative of $\mathcal{R}$ at $(0,0)$ is the operator defined by $D\mathcal{R}(0,h) = \frac{Q|_H h - \lambda h}{\lambda}$. Using the fact that $\mathrm{spr}(Q|_H) < |\lambda|$, it is clear that $\lambda \notin \mathrm{spec}(Q|_H)$, therefore $D\mathcal{R}(0,\cdot)$ is an invertible operator with bounded inverse. The implicit function theorem (see for example Theorem 2.1 from [57]) guarantees the existence of an open neighborhood of 0 denoted $t'_\epsilon \subseteq t_\epsilon$ and a function $h : t'_\epsilon \to h_\epsilon$ which is of class C^k, such that $h(0) = 0$ and $\mathcal{R}(t, h(t)) = 0$.

In what follows, we suppose that $t \in t'_\epsilon$ and we fix $h(t)$ to be the function found above. Set $v(t) = v + h(t)$ and $\lambda(t) = \langle\varphi, Q(t)v(t)\rangle$. Using the fact that $\mathcal{R}(t, h(t)) = 0$, it is clear that $Q(t)v(t) = \lambda(t)v(t)$. □

Lemma 7.1.8. *Under the assumptions of Lemma 7.1.7, there exists an open neighborhood of 0, denoted $t''_\epsilon \subseteq t_\epsilon$ and functions of class C^k, $\varphi(t) \in \mathcal{B}^*$ and $N(t) \in \mathcal{L}_\mathcal{B}$ such that for $t \in t''_\epsilon$, $Q^*(t)\varphi(t) = \lambda(t)\varphi(t)$, $Q^n(t) = \lambda^n(t)\langle\varphi(t),\cdot\rangle v(t) + N^n(t)$ and $\|N^n(t)\| \leq c|\lambda|^n$ for all $n \geq 1$.*

Proof. Let $F' = \{u^* \in \mathcal{B}^* : (Q^* - \lambda)u^* = 0\}$ and $H'(t) = \{u^* \in \mathcal{B}^* : \langle u^*, v(t)\rangle = 0\}$. Each $u^* \in \mathcal{B}^*$ can be written in a unique way as $u^* = \langle u^*, v\rangle\varphi + (u^* - \langle u^*, v\rangle\varphi)$ where $\langle u^*, v\rangle\varphi \in F'$ and $u^* - \langle u^*, v\rangle\varphi \in H'(0)$. So, $B^* = F' \oplus H'(0)$. Set the function $\mathcal{G}_1 : t'_\epsilon \times H'(0) \to H'(t)$ by:

$$\mathcal{G}_1(t, u^*) = Q^*(t)(\varphi + u^*) - \lambda(t)(\varphi + u^*).$$

and let $\pi : H'(t) \to H'(0)$ be the projection onto $H'(0)$ parallel to F'. In other words, for $u^* \in H'(t) \subseteq B^*$, let $u^* = u^*_F + u^*_H$ be the unique way in which u^* can be decomposed into the sum of an element of F' plus an element of $H'(0)$, then $\pi(u^*) = u^*_H$. Note that π is injective: Suppose that $u^*_1, u^*_2 \in H'(t)$ so that $\langle u^*_1, v(t)\rangle = \langle u^*_2, v(t)\rangle = 0$. Suppose further that $\pi(u^*_1) = \pi(u^*_2)$. Then, necessarily $(u^*_1)_H = (u^*_2)_H$ and therefore $\langle (u^*_1)_F, v(t)\rangle = \langle (u^*_2)_F, v(t)\rangle$. Let $(u^*_1)_F = z_1\varphi$ and $(u^*_2)_F = z_2\varphi$. Then $z_1\langle\varphi, v(t)\rangle = z_2\langle\varphi, v(t)\rangle$ implying that $z_1 = z_2$ and therefore $(u^*_1)_F = (u^*_2)_F$ and $u^*_1 = u^*_2$. Finally, let $\mathcal{G} = \pi \circ \mathcal{G}_1$. We use the same steps as in Lemma 7.1.7. The partial derivative of $\mathcal{G}$ with respect to u^* at point $(0,0)$ must satisfy:

$$\lim_{\|\Delta\|\to 0} \frac{\|\mathcal{G}(0,\Delta) - \mathcal{G}(0,0) - D\mathcal{G}(0,\Delta)\|}{\|\Delta\|} = 0.$$

Note that $\mathcal{G}(0,0) = 0$ and for $u^* \in H'(0)$:

$$\mathcal{G}(0,u^*) = \pi\left(Q^*(0)(\varphi+u^*) - \lambda(0)(\varphi+u^*)\right) = \pi\left(Q^*(0)u^* - \lambda(0)u^*\right) = \left(Q^*(0) - \lambda(0)\right)u^*.$$

So, the partial derivative of $\mathcal{G}$ with respect to u^* at $(0,0)$ is the invertible map $D\mathcal{G}(0,u^*) = Q^*|_{H'(0)}u^* - \lambda u^*$. [4] The fact that $\mathcal{G}$ is a function of class C^k with respect to both, u^* and t follows from the same arguments as in lemma 7.1.7. We can now apply the implicit function theorem as before and conclude there exists an open neighborhood of 0, $t''_\epsilon \subseteq t'_\epsilon$ and a function $\psi : t''_\epsilon \to H'(0)$ of class C^k such that $\psi(0) = 0$ and for $t \in t''_\epsilon$, $\mathcal{G}(t,\psi(t)) = 0$ which in view of the injective action of π, implies that $\mathcal{G}_1(t,\psi(t)) = 0$. Using the continuity of $v(t)$ and $\psi(t)$ and the fact that $\langle\varphi, v\rangle = 1$, we can assume that $\langle v + \psi(t), v(t)\rangle \neq 0$ and set $\varphi(t) = \frac{\varphi+\psi(t)}{\langle\varphi+\psi(t),v(t)\rangle}$. Using the fact that $\mathcal{G}_1(t,\psi(t)) = 0$, it is clear that $Q^*(t)\varphi(t) = \lambda(t)\varphi(t)$ and $\langle\varphi(t), v(t)\rangle = 1$. □

7.1.5 Proof of perturbation theorem

Proof of theorem 7.1.2. $Q(0)$ has one simple leading eigenvalue, so the converse of Lemma 7.1.4 guarantees that assertions A1-A3 on that lemma, hold true. Using continuity of $t \mapsto Q(t)$, Lemma 7.1.5 implies that the operator $\mathcal{R}$ from Equation (7.6) is well defined on $t_\epsilon \times h_\epsilon$. Furthermore, under the assumption that $t \mapsto Q(t)$ is of class C^k, Lemma 7.1.6 guarantees that $\mathcal{R}$ is also of class C^k. With the application of the implicit function theorem on Lemmas 7.1.7 and 7.1.8, assertions i) and ii) of Theorem 7.1.2 are satisfied. For assertion iii) it is enough to note that $\lambda(t)\langle\varphi(t), \lambda(t)\langle\varphi(t),u\rangle v(t)\rangle v(t) = \lambda^2(t)\langle\varphi(t),u\rangle v(t)$ and:

$$N(t)\left(\lambda(t)\langle\varphi(t),u\rangle v(t)\right) = Q(t)\left(\lambda(t)\langle\varphi(t),u\rangle v(t)\right) - \lambda^2(t)\langle\varphi(t),u\rangle v(t) = 0.$$
$$\lambda(t)\langle\varphi(t), N(t)u\rangle v(t) = \lambda(t)\langle\varphi(t), Q(t)u\rangle v(t) - \lambda^2(t)\langle\varphi(t),u\rangle v(t) = 0.$$

Finally, by Lemma 7.1.3 with $V = N(t)$ and $V_0 = N(0)$ and using assertion A3 from lemma 7.1.4, for small enough $|t|$ we have that $\|N(t) - N(0)\| \leq \eta$ and there exists $c > 0$ such that $\|N^n(t)\| \leq c|\lambda|^n$ for all $n \geq 1$. □

7.2. Berry-Esseen lemma

The so-called Berry-Esseen inequalities, provide an upper bound for the speed of convergence in the Central Limit Theorem. This type of inequalities were first introduced for IID random variables, but they are now commonly used in the context of dynamical systems with the same purpose. The following lemma is used in the proof of Theorem 4.1.4. It is stated and proved in [20] (see Lemmas 3.4.18 and 3.4.19 from that reference).

Lemma 7.2.1 ([20]). *Let p_1 and p_2 probability measures on $\mathbb{R}$. Denote F_1 and F_2 their respective distribution functions and $\tilde{p}_1$ and $\tilde{p}_2$ their characteristic functions. We assume the following:*

1. $\int |u|dp_1(u) < \infty$ *and* $\int |u|dp_2(u) < \infty$.
2. $\sup_{u\in\mathbb{R}} F_2'(u) = a < \infty$.

Then, for any $u \in \mathbb{R}$ and $L > 0$:

$$|F_1(u) - F_2(u)| \leq \frac{1}{\pi}\int_{-L}^{L} \frac{|\tilde{p}_1(t) - \tilde{p}_2(t)|}{|t|}dt + \frac{24a}{\pi L}. \tag{7.7}$$

[4]The fact that F' is the dual space of F follows from arguments on linear algebra: Consider the base of F formed with the eigenvector v alone. The dual basis is formed with the eigenvector $\varphi \in F^*$. The span of the dual basis must coincide with the whole dual space of F and as a result, F^* is the span of φ, i.e. $F^* = F'$. F' is the annihilator of H and therefore, it is the dual space of F. Similarly, $H'(0)$ is the annihilator of F and therefore, it is the dual of H (see the discussions on I-3.6 and III-3.4 of [52]). This means that $\mathrm{spec}(Q|_F) = \mathrm{spec}(Q^*|_{F'}) = \{\lambda\}$ and $\mathrm{spec}(Q|_H) = \mathrm{spec}(Q^*|_{H'(0)})$ which results in $\mathrm{spr}(Q^*|_{H'(0)}) < |\lambda|$ implying that $\lambda \notin \mathrm{spec}(Q^*|_{H'(0)})$ and $Q^*|_{H'(0)} - \lambda$ is invertible with bounded inverse.

7.3. Rokhlin Disintegration theorem

The following theorem states that a probability measure μ on a compact metric space can be 'disintegrated' with respect to a measurable partition Γ. This means that there exists a family of probability measures $\{\mu_\gamma\}_{\gamma\in\Gamma}$ such that the original measure μ can be obtained by integrating μ_γ (see Theorem 5.1.11 from [71]). The second part of the theorem states sufficient conditions for the disintegration to be unique (Proposition 5.1.7 from [71]). Specifically, this family of probability measures is used to define expressions (3.5) and Equation (7.8) is used in the proof of Lemma 4.5.1.

Theorem 7.3.1 ([71])**.** *Let $(\Sigma, \mathcal{B}, \mu)$ a probability space, (Σ, d) a complete separable metric space and Γ a measurable partition of Σ. Let $\pi : \Sigma \to \Gamma$ be the function that takes an element $x \in \Sigma$ and returns the element of Γ that contains x (i.e. $\pi(x) = \gamma$, where $\gamma \in \Gamma$ is such that $x \in \gamma$). Then μ admits a disintegration with respect to Γ. It means that there exists a family of probability measures on Σ, denoted $\{\mu_\gamma\}_{\gamma\in\Gamma}$ and a measure on Γ, $\hat{\mu} = \mu \circ \pi$ such that:*

A) $\mu_\gamma(\gamma) = 1$ $\hat{\mu}$-a.e. $\gamma \in \Gamma$.

B) The function $\Gamma \to \mathbb{R}$ defined by $\gamma \mapsto \mu_\gamma(E)$ is measurable.

C) For all bounded measurable $g : \Sigma \to \mathbb{R}$:

$$\int_\Sigma g(x) d\mu(x) = \int_{\pi^{-1}\Sigma} \int_\Sigma g(x) d\mu_\gamma(x) d(\mu \circ \pi^{-1}(\gamma)). \tag{7.8}$$

Even more, if $\mathcal{B}$ is countably generated, then the disintegration is unique in the sense that, if $(\{\mu_\gamma\}_{\gamma\in\Gamma}, \hat{\mu})$ y $(\{\mu'_\gamma\}_{\gamma\in\Gamma}, \hat{\mu}')$ satisfy the three properties above, then $\mu_\gamma = \mu'_\gamma$, $\hat{\mu}$-a.e. $\gamma \in \Gamma$.

7.4. Derivatives of $\lambda(t)$

The following is a standard calculation used to obtain the first and second derivatives of the eigenvalue $\lambda(t)$ under the condition that $\int f d\mu_0 = 0$.

Theorem 7.4.1. *If $\int f d\mu_0 = 0$, then $\lambda'(0) = 0$ and $\lambda''(0) = -\int f^2 d\mu_0 - 2\sum_{k=1}^{\infty} \int f \cdot (f \circ F^k) d\mu_0$*

Proof. Let μ_t be the eigenmeasure associated to the eigenvalue $\lambda(t)$. Then:

$$\mathcal{F}^*_{f,t}\mu_t = \lambda(t)\mu_t,$$

with $h \in L_\infty$ and using Equation (3.22) we get:

$$\int e^{itf} \cdot (h \circ F)\, d\mu_t = \int hd\left(\mathcal{F}^*\left(e^{itf}\mu_t\right)\right) = \int hd\left(\mathcal{F}^*_{f,t}\mu_t\right) = \lambda(t)\int hd\mu_t.$$

Taking the extremes of the last equation and deriving with respect to t:

$$\int (if) \cdot e^{itf} \cdot (h \circ F) d\mu_t + \int e^{itf} \cdot (h \circ F) d\mu'_t = \lambda'(t) \int hd\mu_t + \lambda(t) \int hd\mu'_t, \tag{7.9}$$

with $t = 0$, $h = 1$ and recalling that $\lambda(0) = 1$:

$$i\int f d\mu_0 + \int d\mu'_0 = \lambda'(0)\int d\mu_0 + \int d\mu'_0,$$

which implies $\lambda'(0) = i\int f d\mu_0 = 0$. Note Equation (7.9) allows us to find an expression for $\int h d\mu_0'$ knowing that $\lambda'(0) = 0$.

$$\begin{aligned} i\int f\cdot(h\circ F)d\mu_0 + \int(h\circ F)d\mu_0' &= \int h d\mu_0' \\ i\int f\cdot(h\circ F^2)d\mu_0 + \int(h\circ F^2)d\mu_0' &= \int(h\circ F)d\mu_0' \\ i\int f\cdot(h\circ F^3)d\mu_0 + \int(h\circ F^3)d\mu_0' &= \int(h\circ F^2)d\mu_0' \\ &\vdots \end{aligned}$$

Therefore:

$$i\sum_{k=1}^{\infty}\int f\cdot(h\circ F^k)d\mu_0 = \int h d\mu_0'. \tag{7.10}$$

Now we derive again Equation (7.9) with respect to t and fix $h = 1$:

$$-\int f^2\cdot e^{itf}d\mu_t + i\int f\cdot e^{itf}d\mu_t' + i\int f\cdot e^{itf}d\mu_t' + \int e^{itf}d\mu_t'' =$$
$$\lambda''(t)\mu_t(\Sigma) + \lambda'(t)\mu_t'(\Sigma) + \lambda'(t)\mu_t'(\Sigma) + \lambda(t)\mu_t''(\Sigma),$$

with $t = 0$ and using (7.10), this expression becomes:

$$\lambda''(0) = -\int f^2 d\mu_0 + 2i\int f d\mu_0' = -\int f^2 d\mu_0 - 2\sum_{k=1}^{\infty}\int f\cdot(f\circ F^k)d\mu_0.$$

Note the sum converges as consequence of the decay of correlation shown in [61]. □

7.5. Properties of S^∞

Lemma 7.5.1. *Let $E\subset\Sigma_A^+$ a measurable set and $A = \pi_1^{-1}E$. If $\mu\in\mathcal{AB}$, then $\mu(A) = \int_E \phi_\mu dm$*

Proof. This is a straightforward consequence of the definitions:

$$\begin{aligned} \mu(A) = \mu^+(A) - \mu^-(A) = \mu^+(\pi_1^{-1}E) - \mu^-(\pi_1^{-1}E) &= \int_E \phi_{\mu^+}dm - \int_E \phi_{\mu^-}dm \\ = \int_E \left(\phi_{\mu^+} - \phi_{\mu^-}\right)dm = \int_E \phi_\mu dm. \end{aligned}$$

□

Proposition 7.5.2. *S^∞ is a vector space. Also, if $c\in\mathbb{R}$, $\mu,\mu_1,\mu_2\in S^\infty$, then $\phi_{c\mu} = c\cdot\phi_\mu$ and $\phi_{\mu_1+\mu_2} = \phi_{\mu_1} + \phi_{\mu_2}$.*

Proof. i) We will show that $\mu_1+\mu_2$ defined by $(\mu_1+\mu_2)(A) = \mu_1(A) + \mu_2(A)$ belongs to S^∞ and $\phi_{\mu_1+\mu_2} = \phi_{\mu_1} + \phi_{\mu_2}$. Suppose that $E\subset\Sigma_A^+$ is such that $m(E) = 0$ and let $A = \pi_1^{-1}E$. In order to show that $\mu_1+\mu_2\in\mathcal{AB}$ it is enough to show that $(\mu_1+\mu_2)^+(A) = (\mu_1+\mu_2)^-(A) = 0$. Note that:

$$\begin{aligned} (\mu_1+\mu_2)^+(A) &= \frac{1}{2}\left[|\mu_1+\mu_2|(A) + (\mu_1+\mu_2)(A)\right] \\ &= \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_1(A_i)+\mu_2(A_i)| + \mu_1(A) + \mu_2(A)\right] \\ &\le \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_1(A_i)| + \mu_1(A)\right] + \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_2(A_i)| + \mu_2(A)\right] \\ &\le \mu_1^+(A) + \mu_2^+(A) = 0. \end{aligned}$$

In the last inequality we have used the fact that μ_1 and $\mu_2 \in \mathcal{AB}$ and hence $\mu_1^+(A) = \mu_2^+(A) = 0$. Similarly:

$$\begin{aligned}(\mu_1+\mu_2)^-(A) &= \frac{1}{2}\left[|\mu_1+\mu_2|(A) - (\mu_1+\mu_2)(A)\right] \\ &= \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_1(A_i)+\mu_2(A_i)| - \mu_1(A) - \mu_2(A)\right] \\ &\leq \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_1(A_i)| - \mu_1(A)\right] + \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|\mu_2(A_i)| - \mu_2(A)\right] \\ &\leq \mu_1^-(A) + \mu_2^-(A) = 0.\end{aligned}$$

This shows that $\mu_1+\mu_2 \in \mathcal{AB}$. Using Lemma 7.5.1 with $\mu_1+\mu_2$ we obtain

$$(\mu_1+\mu_2)(A) = \int_E \phi_{\mu_1+\mu_2} dm. \tag{7.11}$$

But also

$$(\mu_1+\mu_2)(A) = \mu_1(A)+\mu_2(A) = \int_E \phi_{\mu_1} dm + \int_E \phi_{\mu_2} dm = \int_E (\phi_{\mu_1}+\phi_{\mu_2})\, dm, \tag{7.12}$$

Equations (7.11) and (7.12) hold for every measurable set E and this means that $\phi_{\mu_1+\mu_2} = \phi_{\mu_1}+\phi_{\mu_2}$. Therefore, $\mu_1+\mu_2 \in S^\infty$.

ii) We will show that $c\mu$ defined by $c\mu(A) = c\cdot\mu(A)$ belongs to S^∞ and $\phi_{c\mu} = c\cdot\phi_\mu$. First, note that if $c=0$, then $c\mu = 0$, $\phi_{c\mu} = \phi_0 = 0$ and the claim is trivially true. Otherwise we have the following for $(c\mu)^+$ and $(c\mu)^-$:

$$\begin{aligned}(c\mu)^\pm(A) &= \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|(c\mu)(A_i)|\right] \pm \frac{1}{2}(c\mu)(A) = \frac{1}{2}\left[\sup\sum_{i=1}^{\infty}|c||\mu(A_i)|\right] \pm \frac{c}{2}\mu(A) \\ &= \frac{|c|}{2}\sup\sum_{i=1}^{\infty}|\mu(A_i)| \pm \frac{c}{2}\mu(A).\end{aligned}$$

If $c>0$, the above means $(c\mu)^+ = c\cdot\mu^+$ and $(c\mu)^- = c\cdot\mu^-$, so we can conclude $c\mu \in \mathcal{AB}$[5] and:

$$\int_E \phi_{(c\mu)^+} dm = (c\mu)^+ \circ \pi_1^{-1}E = c\cdot\mu^+(\pi_1^{-1}E) = c\int_E \phi_{\mu^+} dm. \tag{7.13}$$

Equation (7.13) holds for every measurable set E which means $\phi_{(c\mu)^+} = c\cdot\phi_{\mu^+}$. Exactly the same process with $(c\mu)^-(A)$ allows us to show $\phi_{(c\mu)^-} = c\cdot\phi_{\mu^-}$ so:

$$\phi_{c\mu} = \phi_{(c\mu)^+} - \phi_{(c\mu)^-} = c\cdot\phi_{\mu^+} - c\cdot\phi_{\mu^-} = c\cdot\phi_\mu. \tag{7.14}$$

Similarly, when $c<0$ we have that $(c\mu)^+ = -c\cdot\mu^-$ and $(c\mu)^- = -c\cdot\mu^+$, so $c\mu \in \mathcal{AB}$ and:

$$\int_E \phi_{(c\mu)^+} dm = (c\mu)^+ \circ \pi_1^{-1}E = -c\cdot\mu^-(\pi_1^{-1}E) = -c\int_E \phi_{\mu_-} dm. \tag{7.15}$$

Again, Equation (7.15) holds for every measurable set E and that means $\phi_{(c\mu)^+} = -c\phi_{\mu^-}$. Exactly the same process with $(c\mu)^-(A)$ allows us to show $\phi_{(c\mu)^-} = -c\phi_{\mu^+}$, so:

$$\phi_{c\mu} = \phi_{(c\mu)^+} - \phi_{(c\mu)^-} = -c\cdot\phi_{\mu^-} + c\cdot\phi_{\mu^+} = c\cdot\phi_\mu \tag{7.16}$$

Equations (7.14) and (7.16) imply that $c\mu \in S^\infty$ with any $c\in\mathbb{R}$.

□

[5] To see why, suppose $E \subset \Sigma_A^+$ is such that $m(E)=0$. Then $\mu^+ = \mu^- = 0$ and therefore $(c\mu)^+ = (c\mu)^- = 0$

Bibliography

[1] V. S. Afraimovich and L. A. Bunimovich. Which hole is leaking the most: a topological approach to study open systems. Nonlinearity, 23(3):643, feb 2010.

[2] R. Aimino, M. Nicol, and S. Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. Probability Theory and Related Fields, 162:233–274, 2015.

[3] L. Arnold. Random Dynamical Systems. Springer Monographs in Mathematics. Springer Berlin Heidelberg, 2013.

[4] W. Arveson. Spectral Theory and Banach Algebras, pages 1–38. Springer New York, New York, NY, 2002.

[5] V. Baladi. Positive Transfer Operators and Decay of Correlations. Advanced series in nonlinear dynamics. World Scientific Publishing Company, 2000.

[6] A. C. Berry. The accuracy of the gaussian approximation to the sum of independent variates. Transactions of the American Mathematical Society, 49:122–136, 1941.

[7] R. Bowen and D. Ruelle. Equilibrium state and the ergodic theory of anosov diffeomorphisms. Lecture Notes in Mathematics, 470:1–83, 01 2008.

[8] H. Bruin, M. Demers, and I. Melbourne. Existence and convergence properties of physical measures for certain dynamical systems with holes. Ergodic Theory and Dynamical Systems, 30(3):687–728, 2010.

[9] L. A. Bunimovich and A. Yurchenko. Where to place a hole to achieve a maximal escape rate. Israel Journal of Mathematics, 182(1):229–252, Mar 2011.

[10] R. Burton and M. Denker. On the central limit theorem for dynamical systems. Transactions of the American Mathematical Society, 302(2):715–726, 1987.

[11] R. Chazottes. Fluctuations of Observables in Dynamical Systems: From Limit Theorems to Concentration Inequalities, pages 47–85. Springer International Publishing, Cham, 2015.

[12] R. Chazottes and Sébastien Gouëzel. Optimal concentration inequalities for dynamical systems. Communications in Mathematical Physics, 316(3):843–889, Dec 2012.

[13] V. Cyr and O. Sarig. Spectral gap and transience for ruelle operators on countable markov shifts. Communications in Mathematical Physics, 292(3):637–666, Dec 2009.

[14] M. H. DeGroot and M. J. Schervish. Probability and Statistics. Addison-Wesley, 2012.

[15] A. Dembo and O. Zeitouni. Large Deviations Techniques and Applications. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 2009.

[16] M. Demers. Markov extensions and conditionally invariant measures for certain logistic maps with small holes. Ergodic Theory and Dynamical Systems, 25(4):1139–1171, 2005.

[17] M. Demers and L.S. Young. Escape rates and conditionally invariant measures. Nonlinearity, 19(2):377, Dec 2005.

[18] J. Ding and A. Zhou. Statistical Properties of Deterministic Systems. Springer Berlin, Heidelberg, 01 2009.

[19] N. Dunford and J.T. Schwartz. Linear Operators: General theory. Linear Operators. Interscience Publishers, 1958.

[20] R. Durrett. Probability: Theory and Examples. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

[21] I. Ekeland and R. Temam. Convex Analysis and Variational Problems. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 1999.

[22] Edward N. E.Lorenz. Deterministic Nonperiodic Flow. Journal of Atmospheric Sciences, 20(2):130–148, March 1963.

[23] W. Feller. An Introduction to Probability Theory and Its Applications. Number v. 2 in An Introduction to Probability Theory and Its Applications. Wiley, 1968.

[24] S. Friedli and Y. Velenik. Statistical Mechanics of Lattice Systems: A Concrete Mathematical Introduction. Cambridge University Press, 2017.

[25] B. E. Fristedt and L.F. Gray. A Modern Approach to Probability Theory. Probability and Its Applications. Birkhäuser Boston, 1996.

[26] G. Froyland. Ulam's method for random interval maps. Nonlinearity, 12(4):1029–1051, jan 1999.

[27] S. Galatolo. Statistical properties of dynamics. introduction to the functional analytic approach. arXiv: Dynamical Systems, 2015.

[28] S. Galatolo and R. Lucena. Spectral gap and quantitative statistical stability for systems with contracting fibers and lorenz-like maps. Discrete and Continuous Dynamical Systems, 40(3):1309–1360, 2020.

[29] S. Galatolo, M. Monge, and I. Nisoli. Existence of noise induced order, a computer aided proof. Nonlinearity, 33(9):4237, jul 2020.

[30] H. O. Georgii. Gibbs Measures and Phase Transitions. De Gruyter studies in mathematics. De Gruyter, 2011.

[31] S. Gouëzel. Limit theorems in dynamical systems using the spectral method. Hyperbolic dynamics, fluctuations and large deviations, 89:161–193, 2015.

[32] S. Gouëzel. Berry-esseen theorem and local limit theorem for non uniformly expanding maps. Annales de l'I.H.P. Probabilités et statistiques, 41(6):pp. 997–1024, 2005.

[33] S. Gouëzel. Berry–esseen theorem and local limit theorem for non uniformly expanding maps. Annales de l'Institut Henri Poincare (B) Probability and Statistics, 41(6):997–1024, 2005.

[34] P. Góra and A. Boyarsky. Absolutely continuous invariant measures for random maps with position dependent probabilities. Journal of Mathematical Analysis and Applications, 278(1):225–242, 2003.

[35] Y. Hafouta. Limit theorems for some skew products with mixing base maps. Ergodic Theory and Dynamical Systems, 41(1):241–271, 2021.

[36] Y. Hafouta. Limit theorems for random non-uniformly expanding or hyperbolic maps with exponential tails. Annales Henri Poincaré, 23(1):293–332, Jan 2022.

[37] Y. Hafouta. Large deviations, moment estimates and almost sure invariance principles for skew products with mixing base maps and expanding-on-average fibers. Ergodic Theory and Dynamical Systems, 44(1):118–158, 2024.

[38] H. Hennion. Sur un théorème spectral et son application aux noyaux lipchitziens. Proceedings of the American Mathematical Society, 118(2):627–634, 1993.

[39] H. Hennion and L. Hervé. Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness. Springer, 2001.

[40] H. Hennion and L. Hervé. Central limit theorems for iterated random lipschitz mappings. The Annals of Probability, 32(3):1934–1984, 2004.

[41] A. M. Hinz and A. Schief. The average distance on the sierpiński gasket. Probability Theory and Related Fields, 87:129–138, 1990.

[42] W. Hoeffding. Probability inequalities for sums of bounded random variables. Journal of the American Statistical Association, 58(301):13–30, 1963.

[43] F. Hofbauer and G. Keller. Ergodic properties of invariant measures for piecewise monotonic transformations. The Theory of Chaotic Attractors, pages 120–141, 2004.

[44] K. Horbacz. Random dynamical systems with jumps. Journal of Applied Probability, 41(3):890–910, 2004.

[45] K. Horbacz. The central limit theorem for random dynamical systems. Journal of Statistical Physics, 164(6):1261–1291, Sep 2016.

[46] J. Hutchinson. Fractals and self similarity. Indiana University Mathematics Journal, 30(5):713–747, 1981.

[47] H. Ishitani. Central limit theorems for the random iterations of 1-dimensional transformations. Dynamics of Complex Systems, 1404:21–31, 2004.

[48] M. V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. Communications in Mathematical Physics, 81(1):39 – 88, 1981.

[49] A. G. Kachurovskii and I. V. Podvigin. Large deviations and the rate of convergence in the birkhoff ergodic theorem. Mathematical Notes, 94(3):524–531, Sep 2013.

[50] A. G. Kachurovskii and I. V. Podvigin. Correlations, large deviations, and rates of convergence in ergodic theorems for characteristic functions. Doklady Mathematics, 91(2):204–207, Mar 2015.

[51] A. G. Kachurovskii and I. V. Podvigin. Large deviations and rates of convergence in the birkhoff ergodic theorem: From holder continuity to continuity. Doklady Mathematics, 93(1):6–8, 2016.

[52] T. Kato. Perturbation Theory for Linear Operators. Classics in Mathematics. Springer Berlin Heidelberg, 1995.

[53] G. Keller. Generalized bounded variation and applications to piecewise monotonic transformations. Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 69(3):461–478, Sep 1985.

[54] A. Klenke. Probability Theory: A Comprehensive Course. Universitext. Springer London, 2007.

[55] Benoît Kloeckner. Effective berry–esseen and concentration bounds for markov chains with a spectral gap. The Annals of Applied Probability, 29(3):pp. 1778–1807, 2019.

[56] S. Lang. Fundamentals of Differential Geometry. Graduate Texts in Mathematics. Springer New York, 2001.

[57] S. Lang. Real and Functional Analysis. Graduate Texts in Mathematics. Springer New York, 2012.

[58] A. Larkin. Quenched decay of correlations for one-dimensional random lorenz maps. Journal of Dynamical and Control Systems, Jan 2022.

[59] A. Lasota and J. Yorke. On the existence of invariant measures for piecewise monotonic transformations. Transactions of the American Mathematical Society, 186:481–488, 1973.

[60] M. Ledoux. The Concentration of Measure Phenomenon. Mathematical surveys and monographs. American Mathematical Society, 2001.

[61] D. Lima and R. Lucena. Lipschitz regularity of the invariant measure and statistical properties for a class of random dynamical systems. arXiv preprint arXiv:2001.08265, 2020.

[62] C. Liverani. Decay of correlations. Annals of Mathematics, 142(2):239–301, 1995.

[63] C. Liverani and V. Maume-Deschamps. Lasota–yorke maps with holes: conditionally invariant probability measures and invariant probability measures on the survivor set. Annales de l'Institut Henri Poincare (B) Probability and Statistics, 39(3):385–412, 2003.

[64] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. Ergodic Theory and Dynamical Systems, 19(3):671–685, 1999.

[65] C. Maldonado and R. Pérez Otero. Phase transition in piecewise linear random maps in the interval. Chaos: An Interdisciplinary Journal of Nonlinear Science, 31:093–112, 09 2021.

[66] Cesar Maldonado, Humberto Muñiz, and Hugo Nieto Loredo. Concentration inequalities and rates of convergence of the ergodic theorem for countable shifts with gibbs measures. Journal of Difference Equations and Applications, 27(11):1594–1607, 2021.

[67] K. Matsumoto and I. Tsuda. Noise-induced order. Journal of Statistical Physics, 31(1):87–106, Apr 1983.

[68] T. Morita. Asymptotic behavior of one-dimensional random dynamical systems. Journal of the Mathematical Society of Japan, 37(4):651 – 663, 1985.

[69] J. Moritz. Berry–esseen theorems under weak dependence. The Annals of Probability, 44(3):2024–2063, 2016.

[70] H. Nieto. Cotas en la rapidez de convergencia en el teorema ergódico de birkhoff en sistemas dinámicos simbólicos con medida de gibbs. Master's thesis, División de Control y Sistemas Dinámicos, 2019.

[71] K. Oliveira and M. Viana. Fundamentos da teoria ergódica. In Colecao Fronteiras da Matematica, 2014.

[72] S. Martinez P. Collet and B. Schmitt. The yorke-pianigiani measure and the asymptotic law on the limit cantor set of expanding systems. Nonlinearity, 7(5):1437, sep 1994.

[73] S. Martinez P. Collet and B. Schmitt. Exponential inequalities for dynamical measures of expanding maps of the interval. Probability Theory and Related Fields, 123(3):301–322, Jul 2002.

[74] W. Parry and M. Pollicott. Zeta Functions and the Periodic Orbit Structure of Hyperbolic Dynamics. Asterisque; 187-188. Société mathématique de France, 1990.

[75] S. Pelikan. Invariant densities for random maps of the interval. Transactions of the American Mathematical Society, 281(2):813–825, 1984.

[76] K. Petersen. Ergodic Theory. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1983.

[77] V. Petrov, V. Gáspár, J. Masere, and K. Showalter. Controlling chaos in the belousov—zhabotinsky reaction. Nature, 361(6409):240–243, Jan 1993.

[78] A. Pianigiani and J. Yorke. Expanding maps on sets which are almost invariant: Decay and chaos. Transactions of the American Mathematical Society, 252:351–366, 1979.

[79] M. Pollicott. Rates of mixing for potentials of summable variation. Transactions of the American Mathematical Society, 352(2):843–853, 1999.

[80] J. Moles R. Chazottes and E. Ugalde. Gaussian concentration bound for potentials satisfying walters condition with subexponential continuity rates. Nonlinearity, 33(3):1094, jan 2020.

[81] S. Rachev, S. Stoyanov, and F. Fabozzi. A probability metrics approach to financial risk measures. John Wiley & Sons, 2011.

[82] M. Ratner. The central limit theorem for geodesic flows onn-dimensional manifolds of negative curvature. Israel Journal of Mathematics, 16(2):181–197, Jun 1973.

[83] Y. Rubner, C. Tomasi, and L. J. Guibas. The earth mover's distance as a metric for image retrieval. International journal of computer vision, 40:99–121, 2000.

[84] D. Ruelle. Thermodynamic Formalism: The Mathematical Structure of Equilibrium Statistical Mechanics. Cambridge Mathematical Library. Cambridge University Press, 2004.

[85] O. Sarig. Existence of gibbs measures for countable markov shifts. Proceedings of the American Mathematical Society, 131(6):1751–1758, 2003.

[86] C. González Tokman, B. R. Hunt, and P. Wright. Approximating invariant densities of metastable systems. Ergodic Theory and Dynamical Systems, 31:1345 – 1361, 2009.

[87] C. T. Ionescu Tulcea and G. Marinescu. Theorie ergodique pour des classes d'operations non completement continues. Annals of Mathematics, 52(1):140–147, 1950.

[88] C. Villani. Topics in Optimal Transportation. Graduate studies in mathematics. American Mathematical Society, 2003.

[89] P. Walters. Ruelle's operator theorem and g-measures. Transactions of the American Mathematical Society, 214:375–387, 1975.

[90] P. Walters. An Introduction to Ergodic Theory. Graduate Texts in Mathematics. Springer New York, 2000.