Pontifícia Universidade Católica
do Rio de Janeiro

Anselmo de Souza Pontes Junior

Limit laws for dynamical systems with some hyperbolicity

Dissertação de Mestrado

Dissertation presented to the Programa de Pós–graduação em Matemática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Matemática.

Advisor: Prof. Silvius Klein

Rio de Janeiro
April 2024

Pontifícia Universidade Católica
do Rio de Janeiro

Anselmo de Souza Pontes Junior

Limit laws for dynamical systems with some hyperbolicity

Dissertation presented to the Programa de Pós–graduação em Matemática da PUC-Rio in partial fulfillment of the requirements for the degree of Mestre em Matemática. Approved by the undersigned Examination Committee.

Prof. Silvius Klein
Advisor
Departamento de Matemática – PUC-Rio

Prof. Lorenzo Justiniano Díaz Casado
Departamento de Matemática – PUC-Rio

Prof. Isaia Nisoli
Instituto de Matemática – UFRJ

Prof. Manuel Stadlbauer
Instituto de Matemática – UFRJ

Prof. Sergey Borisovich Tikhomirov
Departamento de Matemática – PUC-Rio

Rio de Janeiro, April the 17th, 2024

Anselmo de Souza Pontes Junior

Holds a Bachelor's degree in Mathematics from UFRJ.

Bibliographic data

Pontes Junior, Anselmo de Souza

Limit laws for dynamical systems with some hyperbolicity / Anselmo de Souza Pontes Junior; advisor: Silvius Klein. – Rio de janeiro: PUC-Rio, Departamento de Matemática, 2024.

v., 72 f: il. color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática.

Inclui bibliografia

1. Matemática – Teses. 2. Sistemas de Markov; Operador de Transferência; Mapas Expansores; Grandes Desvios; Teorema Central do Limite.;.
I. Klein, Silvius. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. III. Título.

CDD: 510

To my parents, for their support and encouragement day after day.

Acknowledgments

First and foremost, I am profoundly thankful to God for His unwavering guidance and blessings throughout this academic journey.

To my beloved mother Cristina and my brother Pedro, your endless love, support, and encouragement have been my pillars of strength.

I owe a debt of gratitude to my advisor, Silvius Klein, whose expertise, guidance, and mentorship have been invaluable in shaping this dissertation and my academic growth.

I extend my appreciation to the esteemed members of the committee, Isaia Nisoli, Lorenzo Díaz, Manuel Stadlbauer, and Sergey Tikhomirov, for their insightful feedback.

Special thanks are due to Alexander Arbieto, Isaia Nisoli, Jaqueline Siqueira, and Silvius Klein for their unwavering encouragement, which fueled my determination to pursue mathematics as a career path.

I am indebted to all my professors, particularly Ao Cai (the coolest), Bruno Braga, Lorenzo Díaz, Rangel Baldasso, Sergey Sergeev, and Silvius Klein, for their invaluable teachings.

To my colleagues at PUC Rio, thank you for making our little department a happy second home. Special thanks to Creuza, Carlos, Kátia and Mariana for making this house an organized home.

Last but not least, I am grateful to my "mathbros": Aline, Anthony, Camila Crisprim, Marcelo, Raul, Elias, Fernando, Gabriel (Gaga), Lamartine, Pedro Fidelis, Pedro Bessa, Vinícius Bouça, Vinicius Grijó, and Walter (Dr. Ocelot), for their friendship and shared passion for mathematics.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001. I thank PUC-Rio for the VRAC-I scholarship.

Abstract

Pontes Junior, Anselmo de Souza; Klein, Silvius (Advisor). **Limit laws for dynamical systems with some hyperbolicity**. Rio de Janeiro, 2024. 72p. Dissertação de mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

The study of statistical properties of dynamical systems has been an active research area in recent decades. Its main goal is to investigate when certain deterministic chaotic systems exhibit stochastic behavior when examined through the lens of a relevant invariant measure. Some of the key tools employed in deriving such results are the spectral properties of the transfer operator. However, certain skew product systems, including random and mixed random-quasiperiodic linear cocycles, do not fit this approach. Very recent works have obtained limit laws for these systems by studying the Markov Operator. The purpose of this dissertation is to explain how these operators can be used to derive limit laws, such as Large Deviations Estimates and Central Limit Theorem, for certain skew-product dynamical systems.

Keywords

Markov Systems; Transfer Operator; Expanding Maps; Skew-Products; Arc-sine law.

Resumo

Pontes Junior, Anselmo de Souza; Klein, Silvius. **Leis limite para sistemas dinâmicos com alguma hiperbolicidade**. Rio de Janeiro, 2024. 72p. Dissertação de Mestrado – Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

O estudo das propriedades estatísticas dos sistemas dinâmicos tem sido uma área de pesquisa ativa nas últimas décadas. Seu principal objetivo é investigar quando determinados sistemas caóticos determinísticos exibem comportamento estocástico quando examinados pelas lentes de uma medida invariante relevante. Algumas das principais ferramentas empregadas na obtenção desses resultados são as propriedades espectrais do operador de transferência. No entanto, determinados sistemas do tipo produto torcido, incluindo cociclos lineares aleatórios e cociclos mistos aleatórios-quase periódicos, não se encaixam nessa abordagem. Trabalhos muito recentes obtiveram leis limite para esses sistemas estudando o operador de Markov. O objetivo desta dissertação é explicar como esses operadores podem ser usados para derivar leis limite, como Estimativas de Grandes Desvios e o Teorema do Limite Central, para certos sistemas dinâmicos do tipo produto torcido.

Palavras-chave

Sistemas de Markov; Operador de Transferência; Mapas Expansores; Grandes Desvios; Teorema Central do Limite.;

Table of contents

List of figures

1 Introduction

The field of dynamical systems aims to understand the evolution of the majority of points (in an appropriate sense) $x \in M$ over time through a law $f : M \to M$. It has its origins in Poincaré's studies on Celestial Mechanics and Boltzmann's work on the kinetic theory of gases. Since its inception, the utility of studying such systems from the viewpoint of the extremal measures that are invariant under the system's dynamics, namely the ergodic measures, has become apparent. In the 1930s, the works of Birkhoff and von Neumann provided a theoretical justification for Boltzmann's hypothesis, reformulated as "The time average of an observable converges to its spatial average". Birkhoff's theorem is an analogue of the law of large numbers available for independent and identically distributed random variables, showing that deterministic systems can exhibit statistical properties, and von Neumann's theorem clarifies that the study of invariant measures can be approached through a functional analytical approach.

In the 1960s, revisiting Poincaré's works, Smale and Anosov established the foundations of uniformly hyperbolic systems, a class of systems that exhibit expansion/contraction, leading to chaotic behavior. Since the 1970s, many works have been dedicated to proving more refined statistical properties such as the Central Limit Theorem (CLT) and Large Deviations Type (LDT) estimates for hyperbolic systems, with the study of the so-called Ruelle transfer operator playing a fundamental role. In the late 1990s, Young ([19], [20]) proposed a definition of systems that did not need to be uniformly hyperbolic but only required controlled visit times to a region, which itself is uniformly hyperbolic. This context also proved to be a field where the main ideas used previously could be applied. More recently studied dynamical systems do not fit into the contexts described above. In this case, the central object is also an operator, namely the Markov transition operator. In both cases, the statistical properties of dynamical systems share a functional analytical approach where it is necessary to understand how the powers of the Markov/Ruelle operator evolve in a certain space of observables.

This dissertation contains three classes of examples where we derive LDT and CLT theorems: uniformly expanding systems, that provide a general idea of what happens in the hyperbolic case; piecewise expanding systems, where the influence of singularities on the system's behavior becomes apparent—these examples are studied via the transfer operator; and finally, a more recent example of skew-products that is not predominantly hyperbolic, and where the Markov operator plays a central role.

1.1 Basic Concepts

Let (M, f) be a deterministic dynamical system (DDS), that is, M is a set and $f : M \to M$ is a transformation. Given a state $x \in M$, the law f determines the trajectory

$$x \to f(x) \to f^2(x) \to \cdots \to f^n(x) \to \ldots$$

Definition 1.1 *A measure preserving dynamical system (MPDS) is a tuple $(M, \mathcal{F}, f, \mu)$ where:*

- *$(M, \mathcal{F})$ is a measurable space;*
- *$f : M \to M$ is a measurable map;*
- *μ is an $f-$invariant probability measure, that is:*

$$\mu(A) = \mu(f^{-1}(A)), \quad \forall A \in \mathcal{F}.$$

From now on we assume M to be a compact metric space, $\mathcal{F}$ to be its Borel σ-algebra and the map f to be a continuous function. Note that the invariance of μ with respect to f is equivalent to

$$\mu = f_*\mu = \int_M \delta_{f(x)} d\mu(x).$$

Remark 1.1 *The set $Prob_f(M)$ of $f-$invariant measures is a convex set. Moreover, endowed with the weak*- topology it is compact.*

Definition 1.2 *An MPDS (f, μ) is ergodic if for any set $A \in \mathcal{F}$ which is $f-$invariant (that is, $f^{-1}(A) = A$), we have $\mu(A) = 1$ or $\mu(A) = 0$.*

Remark 1.2 *The ergodic measures are the extremal points of $Prob_f(M)$. From now on, unless otherwise stated, all invariant measures are assumed to be ergodic.*

Given an observable $\varphi \in L^1(M, \mu)$ and $n \in \mathbb{N}$, define the n-th Birkhoff sum of φ by

$$S_n\varphi(x) = \varphi(x) + \varphi(f(x)) + \cdots + \varphi(f^{n-1}(x)) .$$

The Birkhoff ergodic theorem implies that

$$\frac{S_n\varphi(x)}{n} \to \int_M \varphi d\mu, \quad \text{for } \mu\text{-a.e. } x \in M.$$

The result above is an analogue of the law of large numbers for i.i.d. random variables. In this dissertation, we are interested in establishing two other types of statistical properties, namely large deviations type (LDT) estimates, and central limit theorems (CLT) for certain dynamical systems.

The Birkhoff ergodic theorem implies that

$$\mu\left\{x \in M : \left|\frac{S_n\varphi(x)}{n} - \int_M \varphi d\mu\right| > \varepsilon\right\} \to 0, \quad \text{as } n \to \infty.$$

Let $L^\infty(M)$ denote the Banach space of measurable and bounded functions $\varphi : M \to \mathbb{R}$, endowed with the supremum norm and let $\mathcal{B}$ be a Banach subspace of $L^\infty(M)$. The LDT estimate we define below is a quantitative version of the convergence above.

Definition 1.3 *We say that an MPDS (M, f, μ) satisfies large deviations type (LDT) estimates in the space of observables $\mathcal{B}$ if for all $\varphi \in \mathcal{B}$ and $\varepsilon > 0$, there exist $c(\varepsilon) > 0$, $n(\varepsilon) \in \mathbb{N}$ and $C > 0$ such that*

$$\mu\left\{x \in M : \left|\frac{S_n\varphi(x)}{n} - \int_M \varphi d\mu\right| > \varepsilon\right\} \leq Ce^{-c(\varepsilon)n} \quad \textit{for all} \quad n \geq n(\varepsilon).$$

We emphasize that in the definition above the quantities $c(\varepsilon)$ and $n(\varepsilon)$ should depend explicitly on ε and on the other input data (namely the observable φ and the dynamics).

LDT estimates as above correspond to certain types of concentration inequalities in classical probabilities. The simplest such result is Hoeffding's inequality, which states the following. Given $X_1, \ldots, X_n$ independent random variables, denote by $S_n := X_1 + \ldots + X_n$ their sum. If $|X_i| \leq C$ a.s. for $1 \leq i \leq n$, then for all $\epsilon > 0$ we have

$$\mathbb{P}\left(\left|\frac{1}{n}S_n - \frac{1}{n}\mathbb{E}[S_n]\right| > \epsilon\right) \leq 2e^{-c(\epsilon)n},$$

where $c(\epsilon) = (2C)^{-2}\epsilon^2$.

There is a vast literature on large deviations for various classes of dynamical systems and many spaces of observables see [1], [7], [14], [17].

We also want to study the Central Limit Theorem (CLT) for certain dynamical systems.

Definition 1.4 *We say that an ergodic system (f, μ) satisfies the central limit theorem in the space of observables $\mathcal{B}$ if for every $\varphi \in \mathcal{B}$ with zero mean there exists $\sigma(\varphi) > 0$ such that*

$$\frac{S_n \varphi}{\sigma(\varphi)\sqrt{n}} \to \mathcal{N}(0,1), \qquad \textit{in distribution.}$$

In other words, for all $\lambda \in \mathbb{R}$,

$$\mu\left\{x \in M : \frac{S_n\varphi(x)}{\sigma(\varphi)\sqrt{n}} \leq \lambda\right\} \to \int_{\infty}^{\lambda} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}.$$

In order to study such properties we will consider the following stochastic counterpart of deterministic dynamical systems.

Definition 1.5 *A stochastic dynamical system (SDS) is a pair (M, K) where $K : M \to Prob(M)$ is a continuous map and $Prob(M)$ is endowed with the weak* topology. K is called a Markov (or transition) kernel on M.*

For each $x \in M$, $K_x \in Prob(M)$ and given a measurable set $A \subset M$, we may interpret $K_x(A)$ as the probability that the point x transitions to the set A. By induction we define the iterated kernel

$$K_x^{n+1} = \int K_y^n dK_x(y), \quad \forall n \geq 1.$$

We note that $K_x^n(A)$ may be interpreted as the probability that the point x transitions to the set A in n steps.

In this context we have a natural analogue of the definition of invariant measure.

Definition 1.6 *A measure ν is called $K-$stationary if*

$$\nu(A) = \int_M K_x(A) d\nu(x) =: K * \nu$$

for all $A \subset M$ measurable.

Definition 1.7 *A triplet (M, K, ν) with ν being $K-$stationary is called a Markov system.*

Example 1.1 *Given a DDS (M, f), if we put $K_x = \delta_{f(x)}$, we obtain an SDS. Moreover, if μ is f-invariant, then it is K-stationary.*

Remark 1.3 *$K-$stationary measures always exist.*

Given a Markov system (M, K, ν), let $X^+ = M^{\mathbb{N}}$ be the space of sequences on M and let $\mathcal{B}^+$ be the $\sigma-$algebra generated by cylinders on X^+, where

$$C([A_0, A_1, \ldots, A_n]) = \left\{\{x_n\}_{n\geq 0} \in X^+ \colon x_0 \in A_0,\, x_1 \in A_1, \ldots, x_n \in A_n\right\} .$$

By Kolmogorov's extension theorem, given any $\pi \in Prob(M)$, there is a unique probability measure $\mathbb{P}_{K,\pi} \in Prob(X^+)$ such that

$$\mathbb{P}_{K,\pi}(C([A_0, A_1, \ldots, A_n])) = \int_{A_0} \int_{A_n} \cdots \int_{A_1} 1\, dK_{x_0}(x_1) \ldots dK_{x_{n-1}}(x_n)\, d\pi(x_0) .$$

$\mathbb{P}_{K,\pi}$ is called the Markov measure with initial distribution π and transition kernel K. When $\pi = \delta_{x_0}$, for some $x_0 \in M$, we simply write $\mathbb{P}_{x_0}$ for the corresponding Markov measure.

Let us consider a Markov chain $\{Z_n\}_{n\in\mathbb{N}}$ with values in M and with transition kernel K. Given an observable $\varphi : M \to \mathbb{R}$, the corresponding stochastic Birkhoff sums are defined as

$$S_n\varphi = \varphi(Z_0) + \varphi(Z_1) + \cdots + \varphi(Z_{n-1}).$$

In the next section we will describe how to obtain LDT estimates and a CLT for stochastic Birkhoff sums under a mixing hypothesis on the transition kernel.

1.2 Abstract LDT & CLT

In the Markov chain setting we usually do not have independence. A natural substitute for independence is that our system quickly converges to the stationary distribution. This is the concept of mixing. More precisely, we are interested in the convergence $K_x^n \to \nu$ as $n \to \infty$. The most convenient way to study this convergence will be through the action of the kernel on observables. Let $L^\infty(M)$ be the set of bounded measurable functions on M. Endowed with the norm $||\varphi||_\infty := \sup_{x\in M} |\varphi(x)|$ it is a Banach space.

Definition 1.8 *The Markov operator $\mathcal{Q}_K : L^\infty(M) \to L^\infty(M)$ associated to the transition kernel $K\colon M \to Prob(M)$ is defined by the relation*

$$\mathcal{Q}\varphi\,(x) := \int_M \varphi(y)\, dK_x(y) \quad \forall \varphi \in L^\infty(M).$$

The strongest possible type of convergence (and hence of mixing) happens if $\mathcal{Q}^n\varphi \to \displaystyle\int \varphi d\nu$ uniformly for all $\varphi \in L^\infty(M)$. In this setting LDT estimates are already available, see [11].

A weaker form of mixing (which we call spectral strong mixing) is defined as follows: there are constants $C > 0$ and $\sigma \in (0,1)$ such that

$$\|\mathcal{Q}^n\varphi - \int_M \varphi\, d\nu\|_{\mathcal{B}} \leq C\sigma^n \|\varphi\|_{\mathcal{B}}$$

for all $\varphi \in \mathcal{B}$ and $n \in \mathbb{N}$. This form of mixing is equivalent to saying that the operator $\mathcal{Q}$ restricted to $\mathcal{B}$ has the spectral gap property, that is, its spectrum consists of the eigenvalue 1 and a compact set strictly contained in the unit ball. LDT estimates are also available in this setting [14].

The following even weaker notion of mixing is the one which will be the main hypothesis for the abstract theorems of this dissertation.

Definition 1.9 *A Markov system (M, K, ν) is strongly mixing with power rate on a set $\mathcal{B} \subset L^\infty(M)$ of observables if there exist $C > 0$ and $p > 0$ such that*

$$\|\mathcal{Q}^n\varphi - \int \varphi d\nu\|_\infty \leq C\,\frac{1}{n^p}\,\|\varphi\|_{\mathcal{B}} \quad \forall \varphi \in \mathcal{B}.$$

Since we assume that our space $\mathcal{B}$ is a Banach subspace of $L^\infty(M)$, this concept of mixing is much more general then the previously defined spectral strong mixing. Moreover, it is more flexible since it does not require an exponential decay. Now we state the first main theorem of this work, an abstract LDT theorem of Cai et al. for strongly mixing Markov systems, see [6].

Theorem 1.1 *(**Abstract LDT**) Let (M, K, ν) be a strongly mixing Markov system with power rate $r_n = C\frac{1}{n^p}$, $p > 0$ on $\mathcal{B} \subset L^\infty(M)$. Let $\{Z_n\}_{n\geq 0}$ be a Markov chain with transition kernel K. Then for all $\varepsilon > 0$ and $\varphi \in \mathcal{B}$ there are $c(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that for all $x_0 \in M$ and $n \geq n(\varepsilon)$ we have:*

$$\mathbb{P}_{x_0}\left\{\left|\frac{1}{n}S_n(\varphi(Z_i)) - \int \varphi d\nu\right| > \varepsilon\right\} \leq 8\, e^{-c(\varepsilon)n}$$

where $c(\varepsilon)$, $n(\varepsilon)$, depend explicitly on the input data, namely on the norm $||\varphi||_{\mathcal{B}}$ of the observable and the mixing parameters (C, p).

If instead of the strong mixing condition we just ask that $\mathcal{Q}^n\varphi \to \displaystyle\int_M \varphi d\nu$ uniformly, then an LDT estimate holds but not in an effective way, that is, the parameters $c(\varepsilon)$ and $n(\varepsilon)$ cannot be determined explicitly from the input data. Effective LDT estimates have important applications to dynamical systems, see for instance [1], [9]. The next theorem is a CLT for strongly mixing Markov systems.

Theorem 1.2 ***(Abstract CLT)*** *Let $(M, K, \nu, \mathcal{B})$ be a strongly mixing Markov system with mixing rate $r_n = C\frac{1}{n^p}$, where $p > 1$ and $\mathcal{B}$ is a dense subset of $C^0(M)$. Assume that for any open set $U \subset M$ with $\nu(U) > 0$ there exists $\varphi \in \mathcal{B}$ such that $0 \le \varphi \le \mathbf{1}_U$ and $\int_M \varphi d\nu > 0$. For any observable $\varphi \in \mathcal{B}$, if φ is not ν-a.e. constant then the CLT holds.*

The theorem above is a consequence (see [6]) of an abstract central limit theorem due to Gordin-Livšic (see [13], [12]).

We will apply the abstract LDT estimate and the abstract CLT to certain dynamical systems that we describe below. We know that a DDS (f, μ) can be viewed as SDS where $K_x = \delta_{f(x)}$ is the transition Kernel. The associated Markov operator is then $\mathcal{Q}\varphi = \varphi \circ f$, the Koopman operator. Unfortunately, the Koopman operator is not strongly mixing, so the abstract theorems we stated in the previous section do not apply. In what follows we describe two approaches to derive LDT estimates for some classes of dynamical systems. The methods described here are expected to apply to many other classes of examples. Both approaches depend on the behavior of the powers of an operator, either the Markov or the transfer operator.

Given a reference measure μ and a non singular DDS (M, f), for any $h \in L^1(\mu)$, let $\mu_h = h\, d\mu$ and define its evolution by the dynamics as

$$\mathcal{L}h := \frac{df_*\mu_h}{d\mu} \in L^1(\mu)\,.$$

Then the operator $h \to \mathcal{L}h$ is called the transfer operator. The spectral properties of the transfer operator encode information about the absolutely continuous invariant measures (a.c.i.m) of f. The transfer operator does not have rich spectral properties when considered in $L^1(\mu)$. Given a specific DDS (M, f), it is not an easy task to find a Banach space $\mathcal{B}$ in which one can prove that the transfer operator is quasi compact, see [3], [8] and references therein. Once we know that the transfer operator is quasi compact and 1 is a simple eigenvalue on a Banach space $(\mathcal{B}, ||\cdot||_{\mathcal{B}})$ with $||\cdot||_{\mathcal{B}} \ge ||\cdot||_\infty$, we can show that the correlation sequence

$$C_n(\psi, \varphi) := \int \varphi\, \psi \circ f^n d\mu - \int \varphi d\mu \int \psi d\mu,$$

decays exponentially fast when we take $\varphi \in \mathcal{B}$ and $\psi \in L^1(\mu)$. This is a sufficient condition that allows Alves et al. (see [1]) to derive LDT estimates for such system. More precisely we have the following.

Theorem 1.3 (Alves, Freitas, Luzzatto, Vaienti) *Let (M, f, μ) be an ergodic dynamical system and let $\mathcal{B} \subset L^\infty(\mu)$ be a Banach subspace. Let $\varphi \in \mathcal{B}$*

and assume that $\forall \psi \in L^1(\mu)$ *we have*

$$\sum_{n=1}^{\infty} |C_n(\psi, \varphi)| < \infty\,.$$

Then for every $\epsilon > 0$ *there exist* $c(\varepsilon) > 0$ *and* $n(\varepsilon) \in \mathbb{N}$ *such that for all* $n \geq n(\varepsilon)$

$$\mu \left\{ x \in M : \left| \frac{1}{n} S_n\, \varphi(x) - \int_M \varphi d\mu \right| > \varepsilon \right\} \leq 2e^{-c(\varepsilon)n}.$$

The parameters $c(\varepsilon)$, $n(\varepsilon)$ *are explicitly determined by the input data.*

Another way of deriving LDT estimates is to associate to the transfer operator $\mathcal{L}$ a Markov operator $\mathcal{Q}$ by a change of the reference measure (which will be an a.c.i.m) and show that the spectral gap property of $\mathcal{L}$ implies the strong mixing of $\mathcal{Q}$. This will imply LDT estimates for a Markov chain whose transition kernel is defined by assigning to a point x a weighted sum of Dirac masses centered at its pre-images $f^{-1}(\{x\})$.

1.3 Applications to Certain Dynamical Systems

The two aforementioned approaches for establishing LDT estimates will be applied to $C^{1+\theta}$-expanding maps on a compact connected manifold M, as well as to topologically mixing piecewise expanding maps of a compact interval I. Below we state these LDT estimates. For more details and for the statements of the CLT see chapters 4 and 5.

Theorem 1.4 *(LDT) Given* $\theta \in (0, 1)$, *let* $f : M \to M$ *be a* $C^{1+\theta}$ *expanding map and let* $\varphi \in C^\theta$ *be an observable. For any* $\varepsilon > 0$ *there are* $n(\varepsilon) \in \mathbb{N}$ *and* $c(\varepsilon) > 0$ *such that for all* $n \geq n(\varepsilon)$ *we have*

$$\mu \left\{ x \in M\colon \left| \frac{1}{n} S_n\, \varphi(x) - \int_M \varphi d\mu \right| > \varepsilon \right\} \leq Ce^{-c(\varepsilon)n}$$

where $\mu = gdm$ *is the unique a.c.i.m of* f *and the parameters* $n(\varepsilon)$ *and* $c(\varepsilon)$ *only depend (explicitly) on the* C^θ*-norms of* φ, g *and* $\frac{1}{g}$.

Theorem 1.5 *(LDT) Let* $f : I \to I$ *be a topologically mixing piecewise expanding map, and let* $\varphi \in BV(I)$ *be an observable with bounded variation. For any* $\varepsilon > 0$ *there are* $n(\varepsilon) \in \mathbb{N}$ *and* $c(\varepsilon) > 0$ *such that for all* $n \geq n(\varepsilon)$ *we have*

$$\mu \left\{ x \in I\colon \left| \frac{1}{n} S_n \varphi(x) - \int_I \varphi d\mu \right| > \varepsilon \right\} \leq Ce^{-c(\varepsilon)n}$$

where $\mu = gdm$ *is the unique a.c.i.m of* f *and the parameters* $n(\varepsilon)$ *and* $c(\varepsilon)$ *only depend (explicitly) on the* BV*-norms of* φ, g *and* $\frac{1}{g}$.

Another application, this time only of the Markov operator approach is to the study of statistical properties of random torus translations.

Let $\Sigma := \mathbb{T}^d$ be the d-dimensional torus, m its Haar measure and let $\mu \in Prob(\Sigma)$ be another probability measure. Let $X := \Sigma^{\mathbb{Z}}$ and consider $(X, \mu^{\mathbb{Z}})$ as a product probability space of symbols where each symbol is an element of the torus. Let $\sigma : X \to X$ be the bilateral shift map. Finally, define the skew-product map

$$f : X \times \mathbb{T}^d \to X \times \mathbb{T}^d,\ f(\omega, \theta) = (\sigma\omega, \theta + \omega_0).$$

The triple $(X \times \mathbb{T}^d, f, \mu^{\mathbb{Z}} \times m)$ is called a mixed random-quasiperiodic dynamical system. It is in fact a partially hyperbolic dynamical system.

The fiber dynamics encodes the following Markov chain (or random walk on the torus):

$$\theta \to \theta + \omega_0 \to \theta + \omega_0 + \omega_1 \to \dots$$

The corresponding Markov operator is given by

$$\mathcal{Q} : L^\infty(\mathbb{T}^d) \to L^\infty(\mathbb{T}^d),\ \mathcal{Q}\varphi(\theta) = \int \varphi(\theta + \omega_0) d\mu(\omega_0)\,.$$

Definition 1.10 *We say that $\mu \in Prob(T^d)$ satisfies a mixing Diophantine condition (mixing DC) if*

$$|\hat{\mu}(k)| \leq 1 - \frac{\gamma}{|k|^\tau}, \quad \forall k \in \mathbb{Z}^d \setminus \{0\},$$

for some $\gamma, \tau > 0$, where $\hat{\mu}(k)$ are the Fourier coefficients of the measure μ. In this case we write $\mu \in DC(\gamma, \tau)$.

Assuming the mixing Diophantine condition above, we can prove that $\mathcal{Q}$ is strongly mixing and using the abstract theorems we derive LDT estimates and a CLT for this skew-product dynamical system.

Theorem 1.6 *If $\mu \in DC(\gamma, \tau)$ then $\mathcal{Q}$ is strongly mixing with power rate on any space of Hölder continuous functions $C^\alpha(\mathbb{T}^d)$. More precisely, there are constants $C > 0$ and $p > 0$ such that*

$$\|\mathcal{Q}^n\varphi - \int \varphi\, dm\|_{C^0} \leq C\|\varphi\|_\alpha \frac{1}{n^p}, \quad \forall \varphi \in C^\alpha(T^d), \forall n \geq 1.$$

In fact, p can be chosen to be $\frac{\alpha}{\tau} - \iota$, for any $\iota > 0$, in which case C will depend on ι.

Moreover, an effective LDT estimate holds for the Markov chain on the torus $\theta \to \theta + \omega_0 \to \theta + \omega_0 + \omega_1 \to \dots$ starting from any point $\theta \in \mathbb{T}^d$ and with *any observable* $\varphi \in C^\alpha(\mathbb{T}^d)$*. More precisely, for all* $\epsilon > 0$ *there is* $n(\epsilon) \in \mathbb{N}$ *such that for all* $\theta \in \mathbb{T}^d$ *and* $n \geq n(\epsilon)$ *we have*

$$\mu^{\mathbb{N}} \left\{ \left| \frac{1}{n}[\varphi(\theta) + \ldots + \varphi(\theta + \omega_0 + \ldots + \omega_{n-1})] - \int \varphi \, dm \right| > \epsilon \right\} < e^{-c(\epsilon)n}$$

where $c(\epsilon) = c\epsilon^{2+\frac{1}{p}}$, $n(\epsilon) = n\epsilon^{-\frac{1}{p}}$ *for constants* $c > 0$ *and* $n \in \mathbb{N}$ *which depend explicitly and uniformly on the data, namely on* $\|\varphi\|_\alpha, \gamma, \tau$*.*

Theorem 1.7 *Assume that* $\mu \in Prob(\mathbb{T}^d)$ *satisfies a mixing DC with parameters* $\gamma, \tau > 0$ *and let* $\alpha > \tau$*. Then for every* $\varphi \in C^\alpha(\mathbb{T}^d)$ *nonzero with zero mean, there exists* $\sigma = \sigma(\varphi) > 0$ *such that*

$$\frac{S_n\varphi}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

More precisely, for Lebesgue almost every $\theta \in \mathbb{T}^d$ *and for all* $\lambda \in \mathbb{R}$ *we have*

$$\lim_{n\to\infty} \mu^{\mathbb{N}} \left\{ \frac{1}{\sqrt{n}} [\varphi(\theta) + \ldots + \varphi(\theta + \omega_0 + \ldots + \omega_{n-1})] \leq \lambda \right\} = \int_{-\infty}^{\lambda} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \, dx.$$

The rest of this dissertation is organized as follows. In chapter 2 we define the main concepts regarding SDS and prove the abstract LDT estimate of (Cai et al) and the CLT. In chapter 3 we study the basic properties of transfer operators; in particular we show how to use the spectral gap of the operator to derive decay of correlations of observables and then prove the abstract result of (Alves et al) which shows how decay of correlations imply LDT estimates. In chapter 4 we study expanding maps on compact connected manifolds with a focus on proving the spectral gap in the space of Hölder observables and then show how to obtain LDT estimates via decay of correlations and via Markov systems. Chapter 5 follows the same roadmap of chapter 4 for the class of piecewise expanding maps on an interval. In chapter 6 we study the Markov operator of random toral translations, proving that it satisfies the strong mixing property and therefore admits LDT estimates and CLT.

2
Limit Laws for Markov Systems

This chapter introduces Markov systems which represent the main framework to obtain limit theorems in this work. It is mostly based on [6].

2.1
Markov Systems

Let M be a compact metric space. We begin with the concept of stochastic dynamical system.

Definition 2.1 *A stochastic dynamical system (SDS) is a continuous map $K : M \to Prob(M)$, where $Prob(M)$ denotes the set of probability measures on M endowed with the weak* topology.*

Given $x \in M$,

$$K_x(A) = \int \mathbf{1}_A(y) dK_x(y)$$

can be interpreted as the probability that the point x will transition to the set A. Inductively,

$$K_x^n(A) := \int K_y(A) dK_x^{n-1}(y)$$

can be interpreted as the probability that x will transition to A in n steps.

Definition 2.2 *A measure $\nu \in Prob(M)$ is K-stationary if*

$$\nu(A) = \int_M K_x(A)\, d\nu(x),$$

for all Borel sets $A \subset M$.

A stochastic dynamical system (M, K) naturally induces an averaging operator in $L^\infty(M)$, the space of bounded measurable functions $\varphi : M \to \mathbb{R}$. We will consider $L^\infty(M)$ endowed with the supremum norm, so $L^\infty(M)$ is a Banach space.

Definition 2.3 *The Markov operator $\mathcal{Q}_K : L^\infty(M) \to L^\infty(M)$ induced by the stochastic dynamical system K is defined by the relation,*

$$Q\varphi(x) := \int_M \varphi(y) dK_x(y) \quad \forall x \in M, \forall \varphi \in L^\infty(M).$$

Inductively we have,

$$\mathcal{Q}^n\varphi(x) = \int_M \varphi(y)dK_x^n(y)\,.$$

Remark 2.1 *The Markov operator is positive and $\mathcal{Q}\mathbf{1} = \mathbf{1}$. Conversely, if $\mathcal{L}$ is a positive bounded linear operator and $\mathcal{L}\mathbf{1} = \mathbf{1}$, by Riesz-Markov representation theorem there exists an SDS such that its Markov operator is $\mathcal{L}$.*

Now we give an example of SDS.

Example 2.1 *A deterministic dynamical system f can also be naturally viewed as an SDS of the form $K_x = \delta_{f(x)}$. In this case a K-stationary measure μ satisfies*

$$\mu(A) = \int_M \delta_{f(x)}(A)d\mu(x) = \mu(f^{-1}(A)),$$

i.e. μ is f-invariant. The Markov operator associated to this SDS is given by

$$\mathcal{Q}\varphi(x) = \int_M \varphi(y)d\delta_{f(x)}(y) = \varphi(f(x))$$

which is the well known Koopman operator.

Now we describe how to associate to an SDS a Markov process.

On the product space $X^+ = M^{\mathbb{N}}$ consider the sequence of random variables $\{Z_n : X^+ \to M\}_{n\in\mathbb{N}}$, $Z_n(x) := x_n$ where $x = \{x_n\}_n \in X^+$. By Kolmogorov's extension theorem, given $\pi \in Prob(M)$ there exists a unique probability measure $\mathbb{P}_\pi$ on X^+ for which $\{Z_n\}_{n\in\mathbb{N}}$ is a Markov process with transition probability kernel K and initial probability distribution π. If π is a K-stationary measure then P_π is invariant by the shift map

$$\sigma : X^+ \to X^+,\; \sigma(\{x_n\}_{n\in\mathbb{N}}) = \{x_{n+1}\}_{n\in\mathbb{N}}\,.$$

Given an observable $\varphi : M \to \mathbb{R}$ (a priory in $L^\infty(M)$) and $n \in \mathbb{N}$, we define the n-th stochastic Birkhoff sum of φ by $S_n\varphi\colon X^+ \to \mathbb{R}$,

$$\begin{aligned} S_n\varphi(x) &= \varphi(Z_0(x)) + \varphi(Z_1(x)) + \ldots + \varphi(Z_{n-1}(x)) \\ &= \varphi(x_0) + \varphi(x_1) + \ldots + \varphi(x_{n-1})\,. \end{aligned}$$

To establish limit laws for $S_n\varphi$ we will need to understand how the iterates of the Markov operator converge to the integral with respect to the stationary measure. The general idea is that if the convergence $\mathcal{Q}^n\varphi \to \int \varphi d\nu$ is good enough (the convergence can be interpreted in different ways, precise definitions will be given in the next section), then the process $Z_0, Z_1, \ldots, Z_{n-1}$

although not being i.i.d., displays a sort of weak dependence such that it is still possible to recover the classical limit laws available for i.i.d. random variables. The Markov operator $\mathcal{Q}$ defined in $L^\infty(M)$, usually will not provide much information about the weak dependence of the Markov process. So if we want to see richer convergence properties, for instance fast rate of mixing, we must restrict the operator to a Banach subspace $\mathcal{B} \subset L^\infty(M)$. Such Banach space must contain the constant function $\mathbf{1}$ which leads us to the next definition.

Definition 2.4 *A Markov system is a tuple $(M, K, \nu, \mathcal{B})$ where:*

1. *M is a compact metric space,*
2. *K is an SDS,*
3. *ν is a K-stationary measure,*
4. *$(\mathcal{B}, \|.\|_\mathcal{B})$ is a Banach space continuously embedded in $L^\infty(M)$, while $\mathcal{Q}|_\mathcal{B}$ is a continuous operator. In other words, for all $\varphi \in \mathcal{B}$ we have $\|\varphi\|_\infty \leq \|\varphi\|_\mathcal{B}$ and $\|\mathcal{Q}\varphi\|_\mathcal{B} \leq C\|\varphi\|_\mathcal{B}$ for some constant $C \in (0, \infty)$.*

2.2 Strong Mixing and Statistical Properties

In general, mixing refers to the convergence $K_x^n \to \nu$. Such convergence can be studied relative to different norms and spaces. The definition below provides the strongest possible mode of convergence of a Markov system.

Definition 2.5 *We say that $(M, K, \nu, \mathcal{B} = L^\infty(M))$ is uniformly ergodic if there exist $C > 0$ and $0 < \sigma < 1$ such that:*

$$\|\mathcal{Q}^n\varphi - \int \varphi d\nu\|_\infty \leq C\sigma^n\|\varphi\|_\infty, \ \forall \varphi \in L^\infty(M).$$

A weaker form of convergence is the following

Definition 2.6 *A Markov system $(M, K, \nu, \mathcal{B})$ is spectrally strongly mixing if there exist $C > 0$ and $0 < \sigma < 1$ such that:*

$$\|\mathcal{Q}^n\varphi - \int \varphi d\nu\|_\mathcal{B} \leq C\sigma^n\|\varphi\|_\mathcal{B}, \quad \forall \varphi \in \mathcal{B}.$$

The definition above is successfully verified when we can show that the Markov operator $\mathcal{Q}|_\mathcal{B}$ satisfies the spectral gap property, which is the case for many Markov processes such as the ones we will study in chapters 4 and 5.

Definition 2.7 *We say that a bounded linear operator $L : \mathcal{B} \to \mathcal{B}$ has the spectral gap property if we can write it as $L = \lambda P + N$ where:*

1 P is a projection (i.e. $P = P^2$) and $\dim(Im(P)) = 1$;

2 N is a bounded operator with spectral radius $\rho(N) < \lambda$;

3 $PN = NP = 0$.

There is a concept slightly weaker than the spectral gap property which is easier to verify.

Definition 2.8 *A bounded linear operator L on a Banach space $\mathcal{B}$ is called quasi-compact if there exists a direct sum decomposition $\mathcal{B} = F \oplus H$ such that $0 < \rho < \rho(L)$ where*

1. *F and H are closed and L-invariant: $L(F) \subseteq F$, $L(H) \subseteq H$.*
2. *$dim(F) < \infty$ and all eigenvalues of $L|_F : F \to F$ have modulus larger than ρ.*
3. *The spectral radius of $L|_H$ is smaller than ρ.*

Remark 2.2 *The spectral gap property of L immediately implies its quasi compactness. If L is quasi compact and its peripheral spectrum consists of a simple eigenvalue λ, then L has the spectral gap property (see [14]).*

There are, however, interesting examples where it is not possible to have such properties, for instance in the case of the skew-product type dynamics we will study in chapter 6. In [6], the authors motivated by such model introduced the following much weaker concept of mixing.

Definition 2.9 *A Markov system $(M, K, \nu, \mathcal{B})$ is strongly mixing with power rate if there exist $C > 0$ and $p > 0$ such that*

$$||\mathcal{Q}^n \varphi - \int \varphi d\nu||_\infty \leq C \, \frac{1}{n^p} \, ||\varphi||_{\mathcal{B}} \quad \forall \varphi \in \mathcal{B}.$$

Remark 2.3 *This definition is weaker not only because it allows for polynomial instead of exponential decay, but also because it asks for an upper bound on $|| \cdot ||_\infty$ instead of on the stronger norm $|| \cdot ||_{\mathcal{B}}$.*

Now we state an effective LDT estimate for strongly mixing Markov chains as above.

Theorem 2.1 ***(Abstract LDT)*** *Let $(M, K, \nu, \mathcal{B})$ be a strongly mixing Markov system with rate $r_n = C\frac{1}{n^p}$, $p > 0$ and let $\{Z_n\}_{n\geq 0}$ be a K-Markov chain. Then for all $\varepsilon > 0$ and $\varphi \in \mathcal{B}$ there are $c(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that for all $x_0 \in M$ and $n \geq n(\varepsilon)$ we have:*

$$\mathbb{P}_{x_0}\left\{\left|\frac{1}{n}S_n\varphi(Z_i) - \int \varphi d\nu\right| > \varepsilon\right\} \leq 8e^{-c(\varepsilon)n}$$

where $c(\varepsilon) = \bar{c}\varepsilon^{2+\frac{1}{p}}$, $n(\varepsilon) = \bar{n}\varepsilon^{-\frac{1}{p}}$, and $\bar{c}$ and $\bar{n}$ depend explicitly and uniformly on the data. More precisely, $\bar{c} = C(3CL)^{-(2+\frac{1}{p})}$ and $\bar{n} = (3CL)^{\frac{1}{p}}$, where C, p are the parameters in the strong mixing condition and $L := ||\varphi||_{\mathcal{B}}$.

Proof Fix $x_0 \in M$. Without loss of generality, assume that $\int \varphi d\nu = 0$, otherwise consider $\varphi - \int \varphi d\nu$. Moreover, replacing φ by $-\varphi$, it is enough to estimate $\mathbb{P}_{x_0}\{S_n\varphi \geq n\epsilon\}$. Following Bernstein's trick, introducing a parameter $t > 0$ (which will be optimized latter), we have

$$\mathbb{P}_{x_0}\{S_n\varphi \geq n\epsilon\} = \mathbb{P}_{x_0}\{e^{tS_n\varphi} \geq e^{tn\epsilon}\} \leq e^{-tn\epsilon}\, \mathbb{E}_{x_0}(e^{tS_n\varphi}),$$

where we used Chebyshev's inequality.
Thus we need to estimate the exponential moments $\mathbb{E}_{x_0}(e^{tS_n\varphi})$. The next lemma relates these exponential moments with the powers of the Markov operator $\mathcal{Q}$.

Lemma 2.2 *Let $\varphi \in \mathcal{B}$, $\|\varphi\|_{\mathcal{B}} =: L < \infty$. Let $n \geq n_0$ be two integers and denote by $m := \left\lfloor \frac{n}{n_0} \right\rfloor$. Then for all $t > 0$,*

$$\mathbb{E}_{x_0}(e^{tS_n\varphi}) \leq e^{2tn_0L}\, \|\mathcal{Q}^{n_0}(e^{tn_0\varphi})\|_\infty^{m-1}\,.$$

Proof Write $n = mn_0 + r$, with $0 \leq r < n_0$. Fix $t > 0$ and let $f := e^{t\varphi} : M \to \mathbb{R}$, so $0 < f \leq e^{tL}$. Then for all $x = \{x_n\}_{n\geq 0} \in X^+$, we have:

$$\begin{aligned} e^{tS_n\varphi}(x) &= \prod_{j=0}^{n-1} e^{t\varphi(x_j)} = \prod_{j=0}^{n-1} f(x_j) \\ &= f(x_0)\, f(x_{n_0})\, \cdots\, f(x_{(m-1)n_0})\, f(x_1)\, f(x_{n_0+1})\, \cdots\, f(x_{(m-1)n_0+1}) \cdots \\ &\cdots f(x_{n_0-1})\, f(x_{2n_0-1}) \cdots f(x_{mn_0-1})\, f(x_{mn_0})\, f(x_{mn_0+1}) \cdots f(x_{mn_0+r-1}) \\ &=: F_0(x)\, F_1(x) \cdots F_{n_0-1}(x)\, F_{n_0}(x)\,, \end{aligned}$$

where we defined $F_k : X^+ \to \mathbb{R}$ as:

$$F_k(x) := f(x_k)\, f(x_{n_0+k}) \cdots f(x_{(m-1)n_0+k}) \quad \text{for} \quad 0 \leq k \leq n_0 - 1$$

and

$$F_{n_0}(x) := f(x_{mn_0})\, f(x_{mn_0+1}) \cdots f(x_{mn_0+r-1}).$$

By Hölder's inequality,

$$\begin{aligned} \mathbb{E}_{x_0}(e^{tS_n\varphi}) &= \mathbb{E}_{x_0}(F_0 \cdots F_{n_0-1}\, F_{n_0}) \\ &\le \left(\prod_{k=0}^{n_0-1} \mathbb{E}_{x_0}(F_k^{n_0}) \right)^{\frac{1}{n_0}} \|F_{n_0}\|_\infty . \end{aligned}$$

Note that $\|F_{n_0}\|_\infty \le e^{tn_0L}$. We will show that

$$\mathbb{E}_{x_0}(F_k^{n_0}) \le e^{tn_0L}\, \|\mathcal{Q}^{n_0}(e^{tn_0\varphi})\|_\infty^{m-1} \quad \forall\, k = 0, \dots, n_0 - 1$$

which will conclude the proof.

Fix $k \in \{0, \dots, n_0 - 1\}$ and note that

$$F_k^{n_0}(x) = e^{tn_0\varphi(x_k)}\, e^{tn_0\varphi(x_{n_0+k})} \, \cdots \, e^{tn_0\varphi(x_{(m-1)n_0+k})} .$$

To simplify notations, let $G : X^+ \to \mathbb{R}$, $G(x) := F_k^{n_0}(x)$, and $g : M \to \mathbb{R}$, $g(a) := e^{tn_0\varphi(a)}$. Then $0 < g \le e^{tn_0L}$ and

$$G(x) = g(x_k) \cdot g(x_{n_0+k}) \cdot \ldots \cdot g(x_{(m-1)n_0+k}),$$

which is a function that depends on a finite and sparse set of coordinates, arranged in an arithmetic progression of length m with distance n_0 between consecutive terms. We will show that

$$\mathbb{E}_{x_0}(G) \le e^{tn_0L} \|\mathcal{Q}^{n_0} g\|^{m-1} ,$$

where

$$\mathbb{E}_{x_0}(G) = \int_{X^+} G(x)\, d\mathbb{P}_{x_0}(x) = \int_{X^+} G(x) \left(\prod_{i=n}^{1} dK_{x_{i-1}}(x_i) \right) .$$

We split the set of $(m-1)n_0+k$ many indices $I = \{1, 2, \dots, (m-1)n_0+k\}$ into

$$I = \{1, 2, \dots, k\} \cup I_1 \cup \ldots \cup I_{m-1},$$

where for $j = 1, \dots, m-1$, $I_j := \{(j-1)n_0 + k + 1, \dots, jn_0 + k\}$ is a block of length n_0.

Then, since $G(x)$ does not depend on the variables x_j with $j \notin I$, we have

$$
\begin{aligned}
\mathbb{E}_{x_0}(G) &= \int_{X^+} G(x) \left(\prod_{i=n}^{Y-1} dK_{x_{i-1}}(x_i) \right) \\
&= \int g(x_k) \cdots g(x_{(m-2)n_0+k}) \left(\int g(x_{(m-1)n_0+k}) \prod_{i \in I_{m-1}} dK_{x_{i-1}}(x_i) \right) \\
&\quad \times \prod_{i \in I \setminus I_{m-1}} dK_{x_{i-1}}(x_i) \\
&= \int g(x_k) \cdots g(x_{(m-2)n_0+k}) \left(\mathcal{Q}^{n_0} g(x_{(m-2)n_0+k}) \right) \\
&\quad \times \prod_{i \in I \setminus I_{m-1}} dK_{x_{i-1}}(x_i) \\
&\leq \|\mathcal{Q}^{n_0} g\|_\infty \int g(x_k) \cdots g(x_{(m-2)n_0+k}) \prod_{i \in I \setminus I_{m-1}} dK_{x_{i-1}}(x_i).
\end{aligned}
$$

In the last inequality above we used the fact that

$$\mathcal{Q}^{n_0} g(a) \leq \|\mathcal{Q}^{n_0} g\|_\infty$$

which holds for all $a \in M$.

Repeating the argument $m-1$ times, we obtain the desired bound:

$$\mathbb{E}_{x_0}(G) \leq \|\mathcal{Q}^{n_0} g\|^{m-1} \int g(x_k)\, dK_{x_{k-1}}(x_k) \dots dK_{x_0}(x_1) \leq \|\mathcal{Q}^{n_0} g\|^{m-1} e^{tn_0 L},$$

thus finishing the proof of the lemma. ■

We return to the proof of the theorem. Using the strong mixing assumption, for all $n_0 \in \mathbb{N}$ and $\varphi \in \mathcal{B}$, we have that

$$\|\mathcal{Q}^{n_0} \varphi\|_\infty \leq C \|\varphi\|_{\mathcal{B}} \frac{1}{n_0^p} \leq \frac{CL}{n_0^p}.$$

By Lemma 2.2.2, for all $n \geq n_0$,

$$\mathbb{E}_{x_0}(e^{tS_n \varphi}) \leq e^{2tn_0 L} \, \|\mathcal{Q}^{n_0}(e^{tn_0 \varphi})\|_\infty^{\frac{n}{n_0}-1}.$$

However, $\varphi \in \mathcal{B}$ does not necessarily imply that $e^{tn_0\varphi} \in \mathcal{B}$, so the strong mixing condition cannot be directly applied to the observable $e^{tn_0\varphi}$.

The following inequality holds for all $y \in \mathbb{R}$:

$$e^y \leq 1 + y + \frac{y^2}{2} e^{|y|}.$$

Hence we can write

$$e^{tn_0\varphi} = 1 + tn_0\varphi + \frac{1}{2}t^2n_0^2\varphi^2\psi(tn_0\varphi)$$

where the function ψ satisfies the bound $|\psi(tn_0\varphi)| \leq e^{tn_0|\varphi|}$. Then,

$$e^{tn_0\varphi} = 1 + tn_0\varphi + \frac{1}{2}t^2n_0^2\varphi^2\psi(tn_0\varphi)$$

where $\|\psi(tn_0\varphi)\|_\infty \leq e^{tn_0\|\varphi\|_\infty} \leq 2$ if $t \leq \frac{1}{2Ln_0}$. Then we have

$$\mathcal{Q}^{n_0}(e^{tn_0\varphi}) = 1 + tn_0\mathcal{Q}^{n_0}(\varphi) + \frac{1}{2}t^2n_0^2\mathcal{Q}^{n_0}(\varphi^2\psi(tn_0\varphi)).$$

This shows that

$$\|\mathcal{Q}^{n_0}(e^{tn_0\varphi})\|_\infty \leq 1 + 2t^2n_0^2L^2C$$

provided that $tn_0CL^{1/n_0} \leq t^2n_0^2L^2C$ (and that $C \geq 1$, which we may of course assume). Note that we can choose $t \in \mathbb{R}$ satisfying both constraints, namely $\frac{1}{Ln_0^{1+p}} \leq t \leq \frac{1}{2Ln_0}$ if n_0 is large enough that $n_0^p \geq 2$. Using the inequality $(1+y)^{\frac{1}{y}} \leq e$ for $y > 0$, we get

$$\|\mathcal{Q}^{n_0}(e^{tn_0\varphi})\|_{n_0}^{\infty} \leq \Big(1 + 2t^2n_0^2L^2C\Big)^{\frac{2t^2n_0^2L^2C\cdot\frac{n}{n_0}}{2t^2n_0^2L^2C}} \leq e^{2t^2n_0L^2Cn}.$$

Combining this with the estimate given by Lemma 2.1 and recalling that $e^{tn_0L} \leq 2$, we get

$$\mathbb{E}_{x_0}(e^{tS_n\varphi}) \leq e^{2tn_0L}e^{2t^2n_0L^2C/n} \leq 4e^{2t^2n_0L^2C/n}.$$

Fix $\epsilon > 0$. Using Bernstein's trick, we have

$$\mathbb{P}_{x_0}\{S_n\varphi \geq n\epsilon\} \leq e^{-tn\epsilon}\,\mathbb{E}_{x_0}(e^{tS_n\varphi}) \leq 4e^{-tn\epsilon}e^{2t^2n_0L^2C/n} = 4e^{-n(t\epsilon - 2t^2n_0L^2C)}.$$

It remains to maximize $t\epsilon - 2t^2n_0L^2C$ with the proper choice of the free variables. We choose $n_0 = n_0(\epsilon) = \left(\frac{3CL}{\epsilon}\right)^{\frac{1}{p}}$ and $t = t(\epsilon) = \frac{1}{Ln_0^{1+p}}$.

They satisfy the previous constraints provided that $n_0^p = \frac{3CL}{\epsilon} \geq 2$, which is not a restriction since the size ϵ of the deviation cannot exceed $2\|\varphi\|_\infty \leq 2L \leq \frac{3CL}{2}$ as we may, of course, assume that $C \geq \frac{4}{3}$. Then,

$$t\epsilon - 2t^2n_0L^2C = C\left(\frac{\epsilon}{3CL}\right)^{2+\frac{1}{p}} = c(\epsilon),$$

while n must satisfy $n \geq n_0(\epsilon)$ to ensure the applicability of Lemma 2.2.2. ∎

Remark 2.4 *If we just assume that $\mathcal{Q}^n\varphi \to \int_M \varphi\, d\nu$ uniformly for all $\varphi \in \mathcal{B}$, where $\mathcal{B}$ embeds in $C^0(M)$, then one can prove the following non-effective LDT estimate. For any $\epsilon > 0$, there are $n(\epsilon) \in \mathbb{N}$ and $c(\epsilon) > 0$ such that for all $x_0 \in M$, we have*

$$\mathbb{P}_{x_0}\left\{\left|\frac{1}{n}S_n\varphi - \int_M \varphi\, d\nu\right| > \epsilon\right\} \leq 8e^{-c(\epsilon)n}.$$

This can be proven by slightly modifying the previous argument as follows. Given $\varphi \in \mathcal{B}$ with $\|\varphi\|_{\mathcal{B}} \leq L$, and given any $\epsilon > 0$, let $\delta := \frac{\epsilon}{3CL}$ and choose $n(\epsilon) \in \mathbb{N}$ such that

$$\|\mathcal{Q}^{n_0}\varphi - \int_M \varphi\, d\nu\|_{C^0} < \delta \quad \forall n_0 \geq n(\epsilon).$$

Thus, δ will play the role of the mixing rate $r_{n_0} = \frac{1}{n_0^p}$, and the conclusion will hold with $c(\epsilon) = \frac{\epsilon^2}{3CLn(\epsilon)}$. Note that the parameters $n(\epsilon)$ and $c(\epsilon)$ depend in a uniform but not explicit way on the observable φ (in other words, they do not change much as we vary $\varphi \in \mathcal{B}$, but the threshold $n_0(\epsilon)$ for the limiting behavior cannot be determined from the input data).

Now we proceed to establish the Central Limit Theorem (CLT) for Markov systems. In order to do this we will use the following general CLT.

Theorem 2.3 *(Gordin-Livšic) Let (M, K, ν) be an ergodic Markov system, let $\varphi \in L^2(\nu)$ with $\int \varphi\, d\nu = 0$ and assume that*

$$\sum_{n=0}^{\infty} \|\mathcal{Q}^n\varphi\|_2^2 < \infty.$$

Denoting $\psi := \sum_{n=0}^{\infty} \mathcal{Q}^n\varphi$, we have that $\psi \in L^2(\nu)$ and $\varphi = \psi - \mathcal{Q}\psi$. If $\sigma^2(\varphi) := \|\psi\|_2^2 - \|\mathcal{Q}\psi\|_2^2 > 0$, then :

$$\frac{S_n\varphi}{\sigma(\varphi)\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Theorem 2.4 *(**Abstract CLT**) Let $(M, K, \nu, \mathcal{B})$ be a strongly mixing Markov system with mixing rate $r_n = C\frac{1}{n^p}$ with $p > 1$, where $\mathcal{B}$ is a dense subset of $C^0(M)$. Assume that for any open set $U \subset M$ with $\nu(U) > 0$ there exists $\varphi \in \mathcal{B}$ such that $0 \leq \varphi \leq \mathbf{1}_U$ and $\int_M \varphi d\nu > 0$. For any observable $\varphi \in \mathcal{B}$, if φ is not ν-a.e. constant then the CLT holds.*

Proof The strong mixing condition and the density of $\mathcal{B}$ in $C^0(M)$ imply that ν is the unique K-stationary measure, which in turn implies the ergodicity of the Markov system, since the ergodic measures are the extremal points of the set of K-stationary measures. Indeed, if $\tilde{\nu}$ is a K-stationary measure, then for any $\varphi \in C^0(M)$ we have $\int \mathcal{Q}^n \varphi \, d\tilde{\nu} = \int \varphi \, d\tilde{\nu}$ for all $n \in \mathbb{N}$. By strong mixing, for any $\varphi \in \mathcal{B}$ we have that $\mathcal{Q}^n \varphi \to \int \varphi \, d\nu$ uniformly. Integrating with respect to $\tilde{\nu}$, we conclude that $\int \varphi \, d\tilde{\nu} = \int \varphi \, d\nu$ for all $\varphi \in \mathcal{B}$, so for all $\varphi \in C^0(M)$, which shows that $\tilde{\nu} = \nu$.

Let $\varphi \in \mathcal{B}$ be a non ν-a.e. constant observable. We may, of course, assume that $\int \varphi \, d\nu = 0$, otherwise we consider $\varphi - \int \varphi \, d\nu$.

Let $\psi := \sum_{n=0}^{\infty} \mathcal{Q}^n \varphi$. Since $\varphi \in C^0(M)$ and $p > 1$, the strong mixing assumption on $\mathcal{Q}$ implies (via the Weierstrass M-test) that $\psi \in C^0(M)$ as well.

It remains to show that $\sigma^2(\varphi) > 0$, which ensures the applicability of the previous theorem.

Assume by contradiction that $\sigma^2(\varphi) = \|\psi\|_2^2 - \|\mathcal{Q}\psi\|_2^2 = 0$. Then,

$$
\begin{aligned}
0 &\leq \int ((\mathcal{Q}\psi)(x) - \psi(y))^2 \, dK_x(y) \, d\nu(x) \\
&= \int \Big(((\mathcal{Q}\psi)(x))^2 + \psi(y)^2 - 2\psi(y)(\mathcal{Q}\psi)(x) \Big) \, dK_x(y) \, d\nu(x) \\
&= \int \Big(\psi(y)^2 - ((\mathcal{Q}\psi)(x))^2 \Big) \, dK_x(y) \, d\nu(x) \\
&= \int \psi(y)^2 \, dK_x(y) \, d\nu(x) - \int ((\mathcal{Q}\psi)(x))^2 \, d\nu(x) \\
&= \|\psi\|_2^2 - \|\mathcal{Q}\psi\|_2^2 = 0 \quad \text{(since } \nu \text{ is } K\text{-stationary).}
\end{aligned}
$$

Therefore, $\psi(y) = \mathcal{Q}\psi(x)$ for ν-a.e. $x \in M$ and K_x-a.e. $y \in M$. By induction, we obtain that for all $n \geq 1$, $\psi(y) = (\mathcal{Q}^n \psi)(x)$ for ν-a.e. $x \in M$ and for K_x^n-a.e. $y \in M$, which implies that for all $n \geq 1$ and for ν-a.e. $x \in M$, the function ψ is K_x^n-a.e. constant. Let us show that, in fact, ψ is ν-a.e. constant.If ψ is not ν-a.e. constant, then there exist two disjoint open subsets U_1 and U_2 of M such that $\nu(U_1), \nu(U_2) > 0$ and $\psi|_{U_1} < \psi|_{U_2}$. By the assumption, there are two observables $\varphi_1, \varphi_2 \in E$ such that $0 \leq \varphi_i \leq 1_{U_i}$ and $\int \varphi_i \, d\nu > 0$ for $i = 1, 2$. Moreover, for all $x \in M$ and $n \geq 1$,

$$
K_x^n(U_i) = (\mathcal{Q}^n 1_{U_i})(x) \geq (\mathcal{Q}^n \varphi_i)(x) \to \int \varphi_i \, d\nu > 0,
$$

where the above convergence as $n \to \infty$ is uniform in $x \in M$.

Thus, for a large enough integer n and for all $x \in M$, both sets U_1 and U_2 have positive K_x^n measure. However, $\psi|_{U_1} < \psi|_{U_2}$, which contradicts the fact that ψ is K_x^n-a.e. constant for ν-a.e. $x \in M$.

We conclude that ψ is ν-a.e. constant. Since ν is K-stationary, it follows that $\phi = \psi - \mathcal{Q}\psi = 0$ ν-a.e., which is a contradiction. ∎

3
Transfer Operators

This chapter introduces general facts about the transfer operator and the relationship between its spectral properties and the statistical properties of the underlying dynamical system.

3.1
Basic Properties

Let M be a measurable space and let μ be a reference probability measure on M.

Definition 3.1 *A map $f : M \to M$ is non-singular with respect to μ if given any measurable set $E \subset M$, $\mu(f^{-1}(E)) = 0$ if and only if $\mu(E) = 0$.*

Given $h \in L^1(\mu)$, we define $\mu_h := hd\mu$. If f is a non-singular map then it is easy to see that $f_*\mu_h \ll \mu$. By Radon-Nikodym theorem, there exists the density $\frac{df_*\mu_h}{d\mu} \in L^1(\mu)$. Thus we can define the transfer operator $\mathcal{L} : L^1(\mu) \to L^1(\mu)$,

$$\mathcal{L}h := \frac{df_*\mu_h}{d\mu}\,.$$

Definition 3.2 *Given a map $f : M \to M$, we define its correspondent Koopman operator $U : L^\infty(M) \to L^\infty(M)$ by:*

$$U\varphi = \varphi \circ f, \quad \forall \varphi \in L^\infty(M).$$

The next proposition characterizes the transfer operator as the dual of the Koopman operator.

Proposition 3.1 *$\mathcal{L}h$ is the unique element in $L^1(\mu)$ such that $\forall \varphi \in L^\infty(\mu)$, we have*

$$\int \varphi\,(\mathcal{L}h)\,d\mu = \int (\varphi \circ f)\,h\,d\mu = \int U\varphi\,h\,d\mu\,. \tag{3-1}$$

Proof For every $\varphi \in L^\infty(\mu)$ we have

$$\int \varphi \cdot (\mathcal{L}h)d\mu = \int \varphi\,\frac{df_*\mu_h}{d\mu}d\mu = \int \varphi\,df_*\mu_h = \int (\varphi \circ f)\,h\,d\mu.$$

This identity characterizes $\mathcal{L}h$. Indeed, suppose that there are $h_1, h_2 \in L^1(\mu)$ such that for $i = 1, 2$,

$$\int \varphi h_i \, d\mu = \int (\varphi \circ f) h \, d\mu$$

for all $\varphi \in L^\infty(\mu)$. Choose $\varphi = \operatorname{sgn}(h_1 - h_2)$. Then

$$\begin{aligned}\int |h_1 - h_2| \, d\mu &= \int \varphi \cdot (h_1 - h_2) \, d\mu = \int \varphi h_1 \, d\mu - \int \varphi h_2 \, d\mu \\ &= \int (\varphi \circ f) h \, d\mu - \int (\varphi \circ f) h \, d\mu = 0,\end{aligned}$$

hence $h_1 = h_2$ a.e. ■

Proposition 3.2 *$\mathcal{L}$ is a positive and bounded linear operator, with norm 1 in $L^1(\mu)$.*

Proof Fix $h \geq 0$, and let $\varphi := 1_{\mathcal{L}h<0}$ be the indicator function of the set $\{x \in M : \mathcal{L}h(x) < 0\}$. Then

$$0 \geq \int \varphi \, (\mathcal{L}h) \, d\mu = \int (\varphi \circ f) \, h \, d\mu \geq 0 \, .$$

It follows that $\int_{\{\mathcal{L}h<0\}} (\mathcal{L}h) \, d\mu = 0$. This can only happen if $\mu\{\mathcal{L}h < 0\} = 0$, therefore $\mathcal{L}$ is positive. Now we want to prove that $||\mathcal{L}||_1 = ||h||_1$, let $\varphi := \operatorname{sgn}(\mathcal{L}h)$, then

$$\|\mathcal{L}h\|_1 = \int |(\mathcal{L}h)| \; d\mu = \int \varphi \, (\mathcal{L}h) \, d\mu = \int (\varphi \circ f) \, h \, d\mu \leq \|\varphi \circ f\|_\infty \|h\|_1 = \|h\|_1,$$

hence $\|\mathcal{L}h\|_1 \leq \|h\|_1$.

Finally, if $h > 0$, $\|\mathcal{L}h\|_1 = \int |\mathcal{L}h| \, d\mu = \int \mathcal{L}h \, d\mu = \int (1 \circ f) \, h \, d\mu = \|h\|_1$, so $\|\mathcal{L}\|_1 = 1$. ■

Proposition 3.3 *If $\mathcal{L}h = h$, then $d\mu_h = h d\mu$ is an f-invariant measure absolutely continuous with respect to μ.*

Proof Given any $\varphi \in L^\infty(\mu)$,

$$\int \varphi \; d\mu_h = \int \varphi \; h \; d\mu = \int \varphi \; \mathcal{L}h \; d\mu = \int \varphi \circ f \; d\mu_h$$

which is equivalent to the invariance of μ_h with respect to f. ■

Proposition 3.4 *Suppose μ is an f-invariant measure. Then for all $h \in L^1(\mu)$,*

$$U \, \mathcal{L}h = \mathcal{L}h \circ f = \mathbb{E}_\mu(h | f^{-1}(\mathcal{F})),$$

where U is the Koopman operator, associated to f.

Proof $\mathcal{L}h \circ f$ is clearly $f^{-1}(\mathcal{F})$ measurable. Given a set $A = f^{-1}(B)$, $B \in \mathcal{F}$,

$$\int_A \mathcal{L}h \circ f \, d\mu = \int_{f^{-1}(B)} \mathcal{L}h \circ f \, d\mu = \int_B \mathcal{L}h \, d\mu = \int_A h d\mu,$$

which shows that $U\mathcal{L}$ is the conditional expection $\mathbb{E}_\mu(h|f^{-1}(\mathcal{F}))$. ■

Definition 3.3 *(Correlations) The correlation sequence between two observables φ and ψ is given by:*

$$C_n(\psi, \varphi) := \int \psi \; \varphi \circ f^n d\mu - \int \varphi d\mu \int \psi d\mu \quad \forall n \in \mathbb{N}.$$

Sometimes it is also convenient to consider the normalized version of this sequence.

Definition 3.4 *Let $\mathcal{B}_1$,$\mathcal{B}_2$ be Banach spaces, $\varphi \in \mathcal{B}_1$ and $\psi \in \mathcal{B}_2$, the normalized correlation sequence is given by:*

$$Cor_n(\psi, \varphi) := \frac{C_n(\psi, \varphi)}{||\varphi||_{\mathcal{B}_1} \, ||\psi||_{\mathcal{B}_2}}$$

Definition 3.5 *We say that a dynamical system (f, μ) is mixing if for all measurable sets A, B we have*

$$\lim_{n\to\infty} C_n(\mathbf{1}_A, \mathbf{1}_B) = \lim_{n\to\infty} (\mu(A \cap f^{-1}(B)) - \mu(A)\mu(B)) = 0$$

The next proposition relates the decay of correlations and the convergence of the powers of $\mathcal{L}$.

Proposition 3.5 *If $\psi \in L^\infty(\mu)$ and $\varphi \in L^1(\mu)$, then for all $n \in \mathbb{N}$,*

$$|C_n(\psi, \varphi)| \leq ||\mathcal{L}^n \varphi - \int \varphi d\mu||_{L^1} \; ||\psi||_\infty \, .$$

In particular, the rate of decay of correlations is $O(||\mathcal{L}^n \psi - \int \psi d\mu||_1)$.

Proof We have that

$$\begin{aligned}
|C_n(\psi,\varphi)| :&= \left|\int \varphi\,\psi\circ f^n d\mu - \int \psi d\mu \int \varphi\, d\mu\right| \\
&= \left|\int \mathcal{L}^n\varphi\,\psi d\mu - \int \psi\, d\mu \int \varphi d\mu\right| \\
&= \left|\int (\mathcal{L}^n\varphi - \int \varphi d\mu)\,\psi\, d\mu\right| \\
&\le \int \left|(\mathcal{L}^n\varphi - \int \varphi d\mu)\,\psi\right| d\mu \\
&\le ||\mathcal{L}^n\varphi - \int \varphi d\mu||_1\, ||\psi||_\infty
\end{aligned}$$

which establishes the claim. ■

In view of the above proposition it is important to study the convergence $\mathcal{L}^n\varphi \to \displaystyle\int \varphi d\mu$ in a space of observables $(\mathcal{B}, ||.||_{\mathcal{B}})$ with $||.||_{\mathcal{B}} \ge ||.||_1$, where we can prove quasi-compactness of $\mathcal{L}|_{\mathcal{B}}$. More precisely, we will consider a Banach space $\mathcal{B} \subset L^1(\mu)$ with the following six properties.

1) There exists a semi-norm $|.|_{\mathcal{B}}$, such that $\mathcal{B} = \{\varphi \in L^1(\mu) : |\varphi|_{\mathcal{B}} < \infty\}$ is a Banach space when endowed with the norm $||.||_{\mathcal{B}} = |.|_{\mathcal{B}} + ||.||_1$.
2) The inclusion $\mathcal{B} \to L^1(\mu)$ is compact.
3) $\mathcal{L}(\mathcal{B}) \subset \mathcal{B}$ and $\mathcal{L}|_{\mathcal{B}}$ is bounded with respect to $||\cdot||_{\mathcal{B}}$.
4) (Lasota-Yorke inequality) There are $r \in (0,1)$ and $R > 0$ such that for all $k \ge 1$,
$$||\mathcal{L}^k f||_{\mathcal{B}} \le r^k ||f||_{\mathcal{B}} + R|f|_{\mathcal{B}}\,.$$
5) $\mathcal{B}$ is a Banach algebra with the norm $||.||_{\mathcal{B}}$, that is, there is $C > 0$ such that:
$$||\varphi\,\psi||_{\mathcal{B}} \le C||\varphi||_{\mathcal{B}}\,||\psi||_{\mathcal{B}}, \quad \forall \varphi,\psi \in \mathcal{B}\,.$$
6) $\mathcal{B}$ is continuously embedded in $L^\infty(\mu)$, i.e, there exists $C > 0$ such that:
$$||\varphi||_\infty \le C||\varphi||_{\mathcal{B}}, \quad \forall \varphi \in \mathcal{B}\,.$$

Under conditions 1-4 above, Ionescu-Tulcea and Marinescu theorem (see Theorem 4.3 for its statement) implies that $\mathcal{L}$ is quasicompact on $\mathcal{B}$. If the leading eigenvalue 1 is simple and if $\{z \in \sigma(\mathcal{L}) : |z| = \rho(\mathcal{L})\} = \{1\}$ then $\mathcal{L}$ has the spectral gap property on $\mathcal{B}$.

The iterates of the transfer operator enjoy the following spectral decomposition:

$$\mathcal{L}^n = \Pi + N^n,$$

where Π projects $\varphi \in \mathcal{B}$ onto the fixed points of $\mathcal{L}$,

$$\Pi(\varphi) = h \int \varphi dm$$

and the linear operator N verifies

$$\|N^n(\varphi)\|_{\mathcal{B}} \leq C'' r^n \|\varphi\|_{\mathcal{B}},$$

where $C'' > 0$ and $r \in (0,1)$ are constants depending on f. In particular,

$$\begin{aligned} ||\mathcal{L}^n\varphi - \Pi\varphi||_{\mathcal{B}} &= ||N^n\varphi||_{\mathcal{B}} \\ ||\mathcal{L}^n\varphi||_{\mathcal{B}} &\leq C'' r^n ||\varphi||_{\mathcal{B}}, \end{aligned}$$

for every $\varphi \in \mathcal{B}$ with $\int \varphi\, dm = 0$.

We will apply this argument in chapters 4 and 5 to uniformly expanding maps as well as to piecewise expanding maps to derive the exponential convergence $\mathcal{L}^n\varphi \to \displaystyle\int \varphi d\mu$ for observables φ in an appropriate Banach algebra $\mathcal{B}$. This in turn will imply exponential decay of correlations against $L^1(\mu)$ functions. Furthermore, as we will see next, such decay of correlations implies large deviations estimates.

3.2
Decay of Correlations & Large Deviations

In this section we prove the following theorem due to Alves et al, (see [1]).

Theorem 3.6 (Alves, Freitas, Luzzatto, Vaienti) *Let (M, f, μ) be an ergodic dynamical system and let $\mathcal{B} \subset L^\infty(\mu)$ be a Banach subspace. Let $\varphi \in \mathcal{B}$ and assume that $\forall \psi \in L^1(\mu)$ we have*

$$\sum_{n=1}^{\infty} |C_n(\psi, \varphi)| < \infty\,.$$

Then for every $\epsilon > 0$ there exist $c(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that for all $n \geq n(\varepsilon)$

$$\mu\left\{x \in M : \left|\frac{1}{n} S_n\, \varphi(x) - \int_M \varphi d\mu\right| > \varepsilon\right\} \leq 2e^{-c(\varepsilon)n}$$

The parameters $c(\varepsilon)$, $n(\varepsilon)$ are explicitly determined by the input data.

The approach in [1] is based on a strategy employed by Melbourne [17] to obtain large deviations for certain dynamical systems with slow rate of mixing. The proof of the exponential LDT estimates relies on an application of the Azuma-Hoeffding's inequality for martingale differences which we recall bellow.

Definition 3.6 *Given a filtration of $\sigma-$algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \mathcal{F}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a sequence of random variables $\{X_i\}_{i=1}^n$ is said to be adapted to this filtration if X_i is $\mathcal{F}_i-$ measurable for all $i = 1, \ldots, n$.*

Definition 3.7 *Given a filtration of $\sigma-$algebras $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n \subset \mathcal{F}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, a sequence of adapted random variables $\{X_i\}_{i=1}^n$ is said to be a martingale difference if it satisfies the following conditions:*

1. $\mathbb{E}|X_i| < \infty$,
2. $\mathbb{E}[X_i|\mathcal{F}_{i-1}] = 0$.

Theorem 3.7 (Azuma-Hoeffding) *Let $\{X_i\}_{i=1}^n$ be a sequence of martingale differences. If there is $C > 0$ such that $\|X_i\|_\infty \leq C$, then for all $\varepsilon > 0$ we have*

$$\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^n X_i \geq \varepsilon\right) \leq e^{-\frac{n\varepsilon^2}{2C^2}}.$$

Now we precisely define what is the martingale sequence for which we apply the Azuma-Hoeffding inequality. Recall that (M, f, μ) is an MPDS. Let us start by defining the appropriate filtration. For every $i = 1, \ldots, n$ let

$$\mathcal{F}_i = f^{-(n-i)}(M).$$

Then we clearly have $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \cdots \subset \mathcal{F}_n$.

Fix $\varphi \in \mathcal{B}$ such that $\int \varphi \, d\mu = 0$ and define

$$\chi := \sum_{i=0}^{\infty} \mathcal{L}^i \varphi \quad \text{and} \quad \xi := \varphi + \chi - \chi \circ f.$$

Finally, for all $i = 1, 2, \ldots, n$ consider

$$Z_i = \xi \circ f^{n-i}.$$

It is easy to see that

$$\begin{aligned}
S_n\varphi &= \sum_{i=1}^n \varphi \circ f^{n-i} \\
&= \sum_{i=1}^n (\varphi \circ f^{n-i} + \chi \circ f^{n-i} - \chi \circ f^{n-i+1}) + \sum_{i=1}^n (\chi \circ f^{n-i+1} - \chi \circ f^{n-i}) \\
&= \sum_{i=1}^n Z_i + \chi \circ f^n - \chi.
\end{aligned}$$

Now we show that $\{Z_i\}_{i=0}^n$ is a martingale difference.

Lemma 3.8 $\{Z_i\}_{i=0}^n$ *is a finite martingale difference with respect to the filtration* $\{\mathcal{F}_i\}_{i=0}^n$.

Proof By definition Z_i is $\mathcal{F}_i$ measurable. Note that,

$$\begin{aligned}\mathbb{E}(Z_1) &= \int \varphi \circ f^{n-1} d\mu + \int \chi \circ f^{n-1} d\mu - \int \chi \circ f^n d\mu \\ &= \int \varphi d\mu + \int \chi d\mu - \int \chi d\mu = 0 \quad \text{(by invariance)}.\end{aligned}$$

Moreover,

$$\mathcal{L}\xi = \mathcal{L}\varphi + \mathcal{L}\chi - \mathcal{L}\chi \circ f = \mathcal{L}\varphi + \mathcal{L}\chi - \chi = \mathcal{L}\varphi - \mathcal{L}\varphi = 0,$$

which implies that for all $i = 0, \ldots, n-2$,

$$\mathbb{E}(Z_{n-i} \mid \mathcal{F}_{n-i-1}) = \mathbb{E}(\xi \circ f_i \mid f^{-(i+1)}(M)) = U^{i+1}\mathcal{L}^{i+1}U^i\xi = U^{i+1}\mathcal{L}\xi = 0,$$

where we used the following property that holds when μ is f-invariant, see (Proposition 3.4):

$$\mathbb{E}(\cdot \mid f^{-(i+1)}(M)) = U^{i+1}P^{i+1}\,.$$

We conclude that $\{Z_i\}_{i=0}^n$ is a martingale difference. ■

The next lemma shows how our assumption on the decay of correlations implies that χ is an element of $L^\infty(\mu)$. We emphasize here the importance of having decay against all $L^1(\mu)$ functions.

Lemma 3.9 *Let* $\varphi \in L^\infty(\mu)$ *with* $\int \varphi d\mu = 0$. *If there is a rate function* $r(n)$, $n \geq 0$ *with* $\sum_{n=0}^\infty r(n) < \infty$ *such that* $Cor_n(\psi, \varphi) \leq r(n)$ *for all* $n \geq 0$ *and all* $\psi \in L^1(\mu)$, *then the function*

$$\chi := \sum_{n=0}^\infty \mathcal{L}^n\varphi \ \in L^\infty(\mu)\,.$$

Proof By the Riesz representation theorem, to each $\varphi \in L^\infty(\mu)$ we can associate the linear functional F_φ,

$$F_\varphi(\psi) = \int \varphi\,\psi d\mu, \quad \forall \psi \in L^1(\mu)\,.$$

Moreover we know that $||\varphi||_\infty = ||F_\varphi||$, so for all $n \geq 0$ we have

$$
\begin{aligned}
||\mathcal{L}^n\varphi||_\infty &= \sup \frac{|\int \mathcal{L}^n\varphi\,\psi d\mu|}{||\psi|_{L^1}} \\
&= \sup \frac{\int |\varphi\,\psi \circ f^n| d\mu}{||\psi|_{L^1}} \\
&= \frac{||\varphi||_\infty ||\psi||_{L^1} Cor_n(\psi,\varphi)}{||\psi||_{L^1}} \\
&\le ||\varphi||_\infty r(n),
\end{aligned}
$$

which implies that

$$
||\sum_{n=0}^{\infty} \mathcal{L}^n\varphi||_\infty \le \sum_{n=0}^{\infty} ||\mathcal{L}^n\varphi||_\infty \le ||\varphi||_\infty \sum_{n=0}^{\infty} r(n) \le \infty,
$$

and concludes the proof. ■

Theorem 3.10 *Let $\varphi \in L^\infty(\mu)$ and suppose that*

$$
\chi := \sum_{n=0}^{\infty} \mathcal{L}^n\varphi \in L^\infty(\mu)\,.
$$

Then for every $\epsilon > 0$ there exists $N(\epsilon, \varphi) = \frac{4}{\epsilon||\chi||_\infty}$ such that for $n \ge N$

$$
\mu\left\{x \in M : \left|\frac{1}{n} S_n(\varphi(x)) - \int_M \varphi d\mu\right| > \varepsilon\right\} \le 2e^{-c(\epsilon)\,n},
$$

where $c(\epsilon) = \frac{\epsilon^2}{8}(||\varphi||_\infty + 2(||\sum \mathcal{L}^n\varphi||_\infty)^2$.

Proof By the definition of Z_i it is clear that for $i = 1, 2, \ldots, n$ we have,

$$
||Z_i||_\infty \le ||\varphi||_\infty + 2||\chi||_\infty \tag{3-2}
$$

By lemma 3.1.2 we know that $\{Z_i\}_{i=0}^n$ is a martingale difference. Then Azuma-Hoeffding inequality is applicable and we get.

$$
\mu\left(\frac{1}{n}\left|\sum Z_i\right| > \frac{\varepsilon}{2}\right) \le 2\exp\left\{-\frac{\varepsilon^2 n}{8(||\varphi||_\infty + 2||\chi||_\infty)^2}\right\}
$$

for all $n \in \mathbb{N}$. In particular for $n \ge N$ where $\frac{2}{N||\chi||_\infty} \le \frac{\varepsilon}{2}$ we obtain,

$$
\begin{aligned}
\mu\left(\frac{1}{n}|S_n| > \varepsilon\right) &\le \mu\left(\frac{1}{n}\left|\sum Z_i\right| + \frac{2}{N}||\chi||_\infty > \varepsilon\right) \\
&\le \mu\left(\frac{1}{n}\left|\sum Z_i\right| > \frac{\varepsilon}{2}\right) \le 2\exp\left\{-\frac{\varepsilon^2 n}{8(||\varphi||_\infty + 2||\chi||_\infty)^2}\right\},
\end{aligned}
$$

which establishes the result. ■

Together with the previous lemma, this theorem establishes the main result of this section, Theorem 3.6.

3.3 An Associated Markov System

The previous section provided a criterion for establishing LDT estimates for non singular maps via decay of correlations, which in turn can be derived from the spectral properties of the transfer operator.

In this section, we explain how to derive limit laws (LDT estimates and CLT) for non singular maps via the abstract results in section 2.2, assuming that the transfer operator satisfies the spectral gap property. The transfer operator is a Markov operator if and only if $\mathcal{L}\mathbf{1} = \mathbf{1}$, i.e, the reference measure is the absolutely continuous invariant measure. That is not typically the case, so the abstract theorems cannot be directly applied to transfer operators. To bypass this technical problem we will associate to our original transfer operator a new one, which will be Markov.

Suppose that (M, f, μ) is non singular and that the corresponding transfer operator $\mathcal{L}$ has the spectral gap property on a Banach algebra $\mathcal{B} \subset L^\infty(\mu)$. Let h be such that $\mathcal{L}h = h$. Let us consider $d\mu = hdm$ as our new reference measure, set $\mu_\varphi = \varphi d\mu$ and define $\mathcal{Q}\varphi(x) = \frac{df_*\mu_\varphi}{d\mu}$, the transfer operator of f with reference measure μ. This operator has the following properties.

Proposition 3.11

1 *$\mathcal{Q}$ is a Markov operator and μ is a stationary measure with respect to $\mathcal{Q}$.*

2 $(\mathcal{Q}\varphi)\, h = \mathcal{L}(\varphi h)$.

Proof Since $\mathcal{Q}$ is a transfer operator we already know that it is positive. Note that $Q\mathbf{1} = \mathbf{1}$ since for all $\varphi \in L^\infty(M)$ we have,

$$\int \varphi\, \mathcal{Q}\mathbf{1}\, d\mu = \int \varphi \circ f\, \mathbf{1}\, d\mu = \int \varphi \circ f\, d\mu = \int \varphi\, d\mu.$$

Thus $\mathcal{Q}$ is a Markov operator.

Let $\psi \in L^\infty(M)$. Using the duality characterization of $\mathcal{Q}$ and $\mathcal{L}$,

$$\begin{aligned}\int \psi \mathcal{Q}\varphi h dm &= \int \psi \mathcal{Q}\varphi d\mu \\ &= \int \psi \circ f \varphi d\mu \\ &= \int \psi \circ f \varphi h dm \\ &= \int \psi \mathcal{L}(\varphi h) dm.\end{aligned}$$

Since this holds for any $\psi \in L^\infty(M)$ we conclude that $(\mathcal{Q}\varphi)\, h = \mathcal{L}(\varphi h)$. ■

Proposition 3.12 *The Markov system $(M, K, \mu, \mathcal{B})$ is strongly mixing with exponential rate.*

Proof By the previous proposition it follows by induction that for all $n \in \mathbb{N}$,

$$(\mathcal{Q}^n\varphi)\, h = \mathcal{L}^n(\varphi h)\,.$$

Then

$$\mathcal{Q}^n\varphi - \int \varphi d\mu = \frac{1}{h}\left(\mathcal{L}^n(\varphi h) - \int \varphi h\, d\mu\right)\,.$$

Applying the mixing property of $\mathcal{L}$ and the fact that $\mathcal{B}$ is a Banach algebra we get:

$$\begin{aligned}||\mathcal{Q}^n\varphi - \int \varphi d\mu||_\infty &\leq ||\frac{1}{h}(\mathcal{L}^n(\varphi h) - \int \varphi h d\mu)||_{\mathcal{B}} \\ &\leq C(h)||(\mathcal{L}^n(\varphi h) - \int \varphi h d\mu)||_{\mathcal{B}} \\ &\leq C(h)\sigma^n||\varphi h||_{\mathcal{B}} \\ &\leq C'(h)\sigma^n||\varphi||_{\mathcal{B}}\,,\end{aligned}$$

where $C(h)$ and $C'(h)$ depend only on the $\mathcal{B}$ norms of h and $\frac{1}{h}$. ■

In light of the proposition above, the abstract results from the previous chapter imply LDT estimates and CLT for a Markov chain with transition kernel K corresponding to the Markov operator Q defined above. This transition kernel K assigns to a point x a weighted average of Dirac masses supported on the pre-images $f^{-1}(x)$. These LDT estimates and CLT for such a Markov chain translate into similar properties for the dynamical system (M, f, μ). We will show this (with full details) in concrete examples in chapters 4 and 5.

In summary, we have two approaches to establish LDT estimates for systems whose transfer operator $\mathcal{L}$ have a spectral gap. One is through decays of correlations (i.e. the result of Alves et al). The other stems from the abstract LDT theorem for strongly mixing Markov systems of Cai et al. In the same setting the abstract CLT theorem is also applicable.

4
Expanding Maps

The main goal of this chapter is to introduce expanding maps on compact, connected manifolds. These maps were extensively studied in thermodynamic formalism and represent a typical model when one is looking to understand the statistical properties of hyperbolic systems. We mostly follow [3] and plan to:

i Define expanding maps and list their basic properties.

ii Define the transfer operator and list its basic properties.

iii Study the action of the transfer operator on the space of Hölder densities. The main goal here is to prove its quasi-compactness.

iv Use the quasi-compactness results to derive statistical properties for Birkhoff sums. This will be done by applying the abstract results of the previous chapter.

4.1
Basic Properties

Let M be a compact, connected, finite dimensional Riemannian manifold and let d denote its metric.

Definition 4.1 *Given $r > 1$, a C^r map $f : M \to M$ is uniformly expanding if there exists $\lambda > 1$ such that for all $x \in M$ and $v \in T_xM$ we have,*

$$||D_x f v|| \geq \lambda ||v||. \tag{4-1}$$

Remark 4.1 *By the inverse function theorem and the compactness of M, $f^{-1}(\{x\})$ is finite for every $x \in M$. Moreover, since M is connected it follows that $\#f^{-1}(\{x\})$ is constant. We denote this constant by $deg(f)$ and call it the degree of the map f.*

Example 4.1 *If $M = \mathbb{S}^1$, $f(z) = z^2$ is called the angle doubling map, arguably the most simple example of an expanding map.*

Definition 4.2 *A dynamical system (M, f) is topologically mixing if given two open sets $A, B \subset M$ there is an integer number N such that $n \geq N$ implies $A \cap f^n(B) \neq \emptyset$.*

Remark 4.2 *Since M is compact and f is an expanding map, the change of variables formula implies for every open set $A \subset M$ there exist $n := n(A)$ such that $f^n(A) = M$. Therefore, every expanding map is topologically mixing.*

An interesting dynamical property of expanding maps that will be used latter is the strong backwards shadowing. Let f be a λ-expanding map that is, a map satisfying (4-1) for some $\lambda > 1$. For any $x, y \in M$ there exists a bijection p between the sets $f^{-1}(x)$ and $f^{-1}(y)$ such that for all $x' \in f^{-1}(x)$,

$$d(x', p(x')) \leq \frac{d(x,y)}{\lambda}.$$

To see this first note that by the inverse function theorem, for each $x' \in f^{-1}(x)$ there exists a neighborhood B of x' and $g : B \to M$ such that $f \circ g = id$ and $g(x') = x$. Since f is expanding, g must be contracting, that is:

$$d(g(x'), g(x'')) \leq \lambda^{-1} d(x', x''), \quad \forall x', x'' \in B.$$

By compactness, we may choose B as an open ball with the same radius $\rho > 0$, for all $x \in M$, which then shows that the property is satisfied.

Remark 4.3 *Another simple fact is that if f is λ- expanding then* $\inf_{x \in M} \det |D_x f| \geq \lambda^n$.

4.2 Transfer Operators

In this section we define the transfer operator of a uniformly expanding map f on M and recall some of its basic properties. We fix m the Lebesgue measure on the manifold M as our reference measure. Given $h \in L^1(m)$, recall that the transfer operator is given by the Radon-Nikodym derivative $\mathcal{L}h = \frac{df_* m_h}{dm}$, where $dm_h := hdm$.

Example 4.2 *Let $f : \mathbb{S}^1 \to \mathbb{S}^1$ be the angle doubling map. Then*

$$\mathcal{L}h(x) = \frac{h(\frac{x}{2}) + h(\frac{x+1}{2})}{2}.$$

In fact, the transfer operator admits a simple closed form formula which generalizes the example above.

Proposition 4.1 *Let $f : M \to M$ be an expanding map. Then*

$$\mathcal{L}h(x) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{\det |D_{x'} f|}. \tag{4-2}$$

Proof (See [18]) ■

Remark 4.4 *In what follows we will use the notation* $g(x) := \frac{1}{\det|D_x f|}$, *therefore the transfer operator is written as*

$$\mathcal{L}h(x) = \sum_{x' \in f^{-1}(x)} g(x')h(x').$$

Remark 4.5 *By the change of variables formula we know that* $\mathcal{L}^* m = m$, *so* 1 *is an eigenvalue of* $\mathcal{L}$.

4.3
Quasi-Compactness

In this section we prove the quasi compactness of the transfer operator of $C^{1+\theta}$ expanding maps in the space of C^θ densities. First, we define what we mean by a $C^{1+\theta}$ map.

Definition 4.3 *We say that* $f \in C^{1+\theta}$ *for* $0 < \theta < 1$ *if* $f \in C^1$ *and it has a* θ*-Hölder continuous derivative, meaning that there exists* $C > 0$ *such that:*

$$||D_{x'} f - D_{y'} f|| \le C d(x,y)^\theta \tag{4-3}$$

for all $x \in M$, y *in the same coordinate chart of* x *and* x', y' *being the respective representations of* x, y *in Euclidean coordinates.*

Now we define precisely the Banach space in which we prove the quasi compactness of $\mathcal{L}$.

Definition 4.4 *Let*

$$C^\theta(M) = \left\{ \varphi : M \to \mathbb{R} : \exists C > 0 \;\; s.t \;\; |\varphi(x) - \varphi(y)| \le C d(x,y)^\theta \quad \forall x, y \in M \right\}.$$

Let $||\varphi||_\theta = ||\varphi||_\infty + v_\theta(\varphi)$, *where*

$$v_\theta(\varphi) := \sup_{x \ne y} \frac{|\varphi(x) - \varphi(y)|}{d(x,y)^\theta}$$

is the θ*-Hölder seminorm.*

Then $||\cdot||_\theta$ *is a norm and* $(C^\theta(M), ||\cdot||_\theta)$ *is a Banach space.*

The main result of this section is the following.

Theorem 4.2 *Let* $\lambda > 1$, $0 < \theta < 1$, *and* $f : M \to M$ *be a* $C^{1+\theta}$ *and* λ*-expanding map of a compact, connected, Riemannian manifold. Then the correspondent transfer operator has a spectral gap on the space of* θ*-Hölder densities.*

Remark 4.6 *Since f is topologically mixing and we already know that $1 \in \sigma(\mathcal{L})$, it suffices to prove that the transfer operator is quasi-compact. For more details (see [8] page 52).*

Remark 4.7 *We point out that our approach to prove quasi compactness differs from the one in [3] where the spectral gap property is proved via projective cones.*

In order to prove the above result we use the following abstract quasi-compactness result of Ionescu-Tulcea and Marinescu [15]. This version of the theorem can be found in [14].

Theorem 4.3 *(Ionescu-Tulcea and Marinescu) Suppose $(X, ||.||)$ is a Banach space and $L : X \to X$ is a bounded linear operator with spectral radius $\rho(L)$. Assume that there exists a semi-norm $||.||'$ with the following properties:*

1. ***Continuity**: $||.||'$ is continuous in $(X, ||\cdot||)$.*
2. ***Pre-compactness**: for any sequence $\{f_n\}_n \subset X$, if $\sup ||f_n|| < \infty$ then there exist a subsequence $\{n_k\}_k$ and $g \in X$ such that $||Lf_{n_k} - g||' \to 0$.*
3. ***Boundness**: There exists $M > 0$ such that $||Lf||' \leq M||f||'$ for all $f \in X$.*
4. ***Lasota-Yorke (or Doeblin-Fortet)**: there are $k \geq 1$, $0 < r < \rho(L)$, and $R > 0$ such that*

$$||L^k f|| \leq r^k ||f|| + R||f||'. \tag{4-4}$$

Then $L : X \to X$ is quasi-compact.

We will check the conditions in Ionescu-Tulcea and Marinescu theorem for the transfer operator where $(X, ||\cdot||) = C^\theta((M), ||\cdot||_\theta)$ and $||\cdot||' = ||\cdot||_\infty$.

In the next lemmas we will repeatedly use the following two facts:

1 If f is a λ-expanding map on a n-dimensional manifold M, then $|\det D_x f| \geq \lambda^n$(see Remark 4.3).

2 By the strong backwards shadowing property, $\forall x, y \in M$, there exists a bijection $p : f^{-1}(x) \to f^{-1}(y)$ such that $d(x', p(x')) \leq \frac{d(x,y)}{\lambda}$.

Lemma 4.4 *(Invariance) $\mathcal{L}(C^\theta(M)) \subset C^\theta(M)$.*

Proof Let $\varphi \in C^\theta(M)$. Then

$$\begin{aligned}\frac{|\mathcal{L}\varphi(x) - \mathcal{L}\varphi(y)|}{d(x,y)^\theta} &\le \frac{\sum_{x' \in f^{-1}(x)} |g(x')\varphi(x') - g(p(x'))\varphi(p(x'))|}{d(x,y)^\theta} \\ &\le \frac{\lambda^\theta}{\lambda^n} \sum_{x' \in f^{-1}(x)} \frac{|\varphi(x) - \varphi(p(x'))|}{d(x,p(x'))^\theta} \\ &\le \lambda^{(\theta-n)} deg(f)|\varphi|_\theta \\ &\le \lambda^\theta |\varphi|_\theta,\end{aligned}$$

which shows that $\mathcal{L}\varphi$ is θ-Hölder continuous. ■

Lemma 4.5 *(Continuity) The norm $||.||_\infty$ is continuous in $C^\theta(M)$ endowed with the Hölder norm $||\cdot||_\theta$.*
Proof *If $||\varphi_n - \varphi||_\theta \to 0$ by definition we must have $||\varphi_n - \varphi||_\infty \to 0$.* ■

Lemma 4.6 *(Pre-compactness) For any sequence $\{f_n\}_n \subset C^\theta(M)$, if $\sup ||f_n|| < \infty$ then there exist a subsequence $\{n_k\}_k$ and $g \in C^\theta(M)$ such that $||\mathcal{L}f_{n_k} - g||_\infty \to 0$.*

Proof If $\sup ||f_n||_\theta < \infty$ then lemma 4.4 implies that $\{\mathcal{L}f_n\}_n$ is an equicontinuous family and by Arzelà-Ascoli theorem there exists a subsequence $\{f_{n_k}\}$ and $g \in C^0(M)$ such that $||\mathcal{L}f_{n_k} - g||_\infty \to 0$.

Furthermore, $||g||_\theta \le \sup ||\mathcal{L}f_n||_\theta < \infty$. ■

Lemma 4.7 *(Boundness) $||\mathcal{L}f||_\infty \le M||f||_\infty$ for all $\varphi \in C^\theta(M)$.*

Proof Let $\varphi \in C^\theta(M)$. Then

$$\begin{aligned}||\mathcal{L}\varphi(x)||_\infty &\le \sum_{x' \in f^{-1}(x)} |\varphi(x')| g(x') \\ &\le \deg(f)\lambda^{-n} ||\varphi||_\infty \le ||\varphi||_\infty\,.\end{aligned}$$

■

Lemma 4.8 *(Lasota-Yorke Inequality) There are $k \ge 1$, $0 < r < \rho(L)$, and $R > 0$ such that*

$$||\mathcal{L}^k\varphi||_\theta \le r^k ||\varphi||_\theta + R||\varphi||_\infty \quad \forall \varphi \in C^\theta(M). \tag{4-5}$$

Proof We first prove that there exists $0 < r < 1$ and $R_1 > 0$ such that

$$|\mathcal{L}\varphi|_\theta \le r|\varphi|_\theta + R_1||\varphi||_\infty. \tag{4-6}$$

We have

$$\begin{aligned}
\frac{|\mathcal{L}\varphi(x)-\mathcal{L}\varphi(y)|}{d(x,y)^\theta} &= \frac{1}{d(x,y)^\theta}|\sum_{x'\in f^{-1}(x)} g(x')\varphi(x') - \sum_{y'\in f^{-1}(y)} g(y')\varphi(y')| \\
&\le \frac{1}{d(x,y)^\theta}\sum_{x'\in f^{-1}(x)} |(g(x')\varphi(x') - g(p(x'))\varphi(p(x')))| \\
&\le \sum_{x'\in f^{-1}(x)} |g(x')|\frac{|\varphi(x')-\varphi(p(x')|}{d(x,y)^\theta} + \\
&\quad \frac{|g(x')-g(p(x'))|}{d(x,y)^\theta}\varphi(p(x')) \\
&\le \sum_{x'\in f^{-1}(x)} |g(x')|\lambda^{-\theta}\frac{|\varphi(x')-\varphi(p(x')|}{d(x,p(x'))^\theta} + \\
&\quad \lambda^{-\theta}\frac{|g(x')-g(p(x'))|}{d(x,p(x'))^\theta}\varphi(p(x')) \\
&\le \deg(f)\lambda^{-(\theta+n)}|\varphi|_\theta + \deg(f)\lambda^{-(\theta+2n)}|\det Df|_\theta|\varphi|_\infty\,.
\end{aligned}$$

Therefore by the definition of the Hölder norm,

$$|\mathcal{L}\varphi|_\theta \le \lambda^{-\theta}|\varphi|_\theta + |\det Df|_\theta|\varphi|_\infty\,.$$

By induction we conclude that

$$\begin{aligned}
|\mathcal{L}^n\varphi|_\theta &\le \lambda^{-\theta}|\mathcal{L}^{n-1}\varphi|_\theta + R|\varphi|_\infty \\
&\le \lambda^{-n\theta}|\varphi|_\theta + \sum_{j=1}^{n-1}\lambda^{-\theta j}R|\varphi|_\infty.
\end{aligned}$$

Defining $R_1 = \frac{R}{1-\lambda^{-\theta}}$ finishes the proof. ■

By Theorem 4.3 we conclude that $\mathcal{L}$ is quasi-compact and its essential spectral radius is smaller than $\lambda^{-\theta} < 1$. Since $||\mathcal{L}||_\theta \le 1$, there exists a finite set $\Theta \subset [0, 2\pi)$ such that

$$\mathcal{L} = \sum_{\theta\in\Theta} e^{i\theta}\Pi_\theta + N,$$

where Π_θ are finite rank operators such that $\Pi_\theta\Pi_{\theta'} = \delta_{\theta,\theta'}$, $\Pi_\theta N = N\Pi_\theta = 0$ and $\rho(N) < \lambda^{-\theta}$. Furthermore, $1 \in \sigma(\mathcal{L}^*)$ which implies that $0 \in \Theta$. It follows that,

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} e^{-ki\theta}\mathcal{L}^k = \sum_{\theta'\in\Theta}\lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} e^{ki(\theta'-\theta)}\Pi_\theta + e^{-ik\theta}N^k = \Pi_\theta\,.$$

Let $h_* = \Pi_0 1$. Then $\mathcal{L}h_* = \mathcal{L}\Pi_0 1 = \Pi_0 1 = h_*$. Applying the equation above we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \mathcal{L}^k 1 = \Pi_0 1 = h_*.$$

Since $\mathcal{L}$ is a positive operator we have conclude that $h_* \geq 0$. In fact, h_* is bounded away from zero by the next lemma.

Lemma 4.9 *The invariant density h_* is bounded away from zero.*

Proof We already know that $h_* \geq 0$ and that h_* is continuous. Suppose by contradiction that there exists $\bar{x} \in M$ such that $h_*(\bar{x}) = 0$. It follows that

$$0 = h_*(\bar{x}) = \mathcal{L}^n h_*(\bar{x}) = \sum_{y \in f^{-n}(\bar{x})} \frac{1}{|\det D_y f^n|} h(y),$$

therefore $h(y) = 0$ for all $y \in f^{-n}(x), n \in \mathbb{N}$. Since f is expanding, we have that the preimages of f are dense in M. By the continuity of h_* we conclude that $h_* \equiv 0$, which is a contradiction. ■

4.4
LDT via Decay of Correlations

In this section we use the information on the spectral gap to derive statistical properties for expanding maps. We remark that some of these properties were already stated in the past chapter but here we collect the precise constants for the class of maps we have at hand.

Theorem 4.10 *(Exponential Decay of Correlations) For all $\varphi \in C^\theta$ with $\int \varphi \, d\mu = 0$ and $\psi \in L^1$, we have*

$$C_n(\psi, \varphi) \leq C \, \lambda^{-n\theta} \, ||\psi||_{L^1} \, .$$

Proof We estimate the decay of correlations as follows:

$$\begin{aligned} |C_n(\psi, \varphi)| &:= \left| \int \varphi \psi \circ f^n d\mu - \int \varphi d\mu \int \psi d\mu \right| \\ &= \left| \int \mathcal{L}^n \varphi \psi d\mu \right| \\ &= \left| \int \mathcal{L}^n \varphi \, \psi \, d\mu \right| \\ &\leq \int |\mathcal{L}^n \varphi \, \psi| \, d\mu \\ &\leq ||\mathcal{L}^n \varphi||_\theta \, ||\psi||_{L^1} \leq C \, \lambda^{-n\theta} \, ||\psi||_{L^1}, \end{aligned}$$

which establishes the result. ■

Proposition 4.11 *(Mixing) The absolutely continuous invariant measure μ of a $C^{1+\theta}$, λ-expanding map is mixing.*

Proof Let $\varphi, \psi \in L^2(\mu)$. We show that $C_n(\varphi, \psi) \to 0$ when $n \to \infty$, which is well known to be equivalent to the system (f, μ) being mixing(see [18]). Without loss of generality we can suppose that $\int \varphi d\mu = 0$. The space of θ-Hölder functions is dense in L^2, so we can consider sequences of Hölder functions $\{\varphi_k\}_k$ and $\{\psi_k\}_k$ such that $||\varphi_k - \varphi||_2$ and $||\psi_k - \psi||_2$ approach zero when k goes to infinity. Then

$$\begin{aligned}
|C_n(\varphi, \psi)| &:= |\int \psi\varphi \circ f^n d\mu| \\
&\leq |\int \psi_k\varphi_k \circ f^n d\mu| + |\int (\varphi_k \circ f^n)(\psi - \psi_k) d\mu| \\
&+ |\int ((\varphi - \varphi_k) \circ f^n)\psi d\mu| \\
&\leq C\lambda^{-n\theta}|\psi_k|_\theta \int |\varphi_k| d\mu + \int \varphi_k d\mu \int \psi_k d\mu + ||\varphi_k||_2||\psi - \psi_k||_2 \\
&+ ||\psi_k||_2||\varphi - \varphi_k||_2 \,.
\end{aligned}$$

Since $||\varphi_k - \varphi||_2 \to 0$ and $\int \varphi d\mu = 0$ we have that $\int \varphi_k \to 0$ as $k \to \infty$. Letting n and k go to ∞, we have that $C_n(\varphi, \psi)) \to 0$. ■

Theorem 4.12 *Let (f, μ) be a λ-expanding map equipped with its a.c.i.m. μ. Let $\mathcal{B} \subset L^\infty(\mu)$ be the Banach space $C^\theta(M)$and $\varphi \in \mathcal{B}$. Then for every $\epsilon > 0$ there exist $\tau(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that for all $n \geq n(\varepsilon)$*

$$\mu\left\{x \in M : \left|\frac{1}{n}S_n(\varphi(x)) - \int_M \varphi d\mu\right| > \varepsilon\right\} \leq 2e^{-c(\varepsilon)n} \,,$$

where $c(\varepsilon)$ and $n(\varepsilon)$ are explicitly determined by the input data.

Proof By Theorem 4.10, we know that the sequence $\xi(n) = C\lambda^{-n\theta}$ satisfies that $C_n(\varphi, \psi) \leq \xi(n)||\psi||_{L^1}$. Since $\sum_{n=0}^{\infty} \xi(n) < \infty$, then theorem 3.6 is applicable and it implies the LDT estimates. ■

4.5 Statistical Properties via Markov Systems

In this section we apply the abstract theorems of chapter 2 to derive statistical properties for Birkhoff sums of expanding maps. The transfer operator is Markov if and only if $\mathcal{L}\mathbf{1} = \mathbf{1}$, i.e, the Lebesgue measure is the absolutely continuous invariant measure of f. That is not typically the case so the abstract theorems cannot be directly applied to transfer operators as explained before. To bypass this technical problem we will associate to our original transfer operator a new one which will be Markov. Let h be the unique absolutely continuous invariant measure of f, let us consider $d\mu = hdm$ as our new reference measure and define $\mathcal{Q}\varphi(x) = \frac{df_*\mu_\varphi}{d\mu}$, the transfer operator of f with reference measure μ. This operator has the following properties, as shown already in the general context of the previous chapter (Proposition 3.11).

Proposition 4.13 *The operator $\mathcal{Q}$ defined above has the following properties.*

1. *$\mathcal{Q}$ is Markov,*
2. *μ is a stationary measure with respect to $\mathcal{Q}$,*
3. *$(\mathcal{Q}\varphi)h = \mathcal{L}(\varphi h)$.*

Remark 4.8 *The formula in the above proposition implies that $\mathcal{Q}$ has the following closed form:*

$$\mathcal{Q}\varphi(x) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{h(x)} \frac{1}{|\det D_{x'} f|} \varphi(x') \tag{4-7}$$

so the associated Markov kernel is

$$K_x(y) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{h(x)} \frac{1}{|\det D_{x'} f|} \delta_{x'} \tag{4-8}$$

so this Markov chain sends x to one of its pre-images.

Applying proposition 3.12 to this Markov system we conclude the following.

Corollary 1 *The Markov system (M, K, μ, C^θ) is strongly mixing with exponential rate.*

Theorem 4.14 *(LDT) Let $f : M \to M$ be a $C^{1+\theta}(M)$ and expanding map, and let $\varphi \in C^\theta$. Given any $\varepsilon > 0$ there are $n(\varepsilon) \in \mathbb{N}$ and $c(\varepsilon) > 0$ such that for all $n \geq n(\varepsilon)$ we have*

$$\mu \left\{ x : M \left| \frac{1}{n} S_n(\varphi(x)) - \int_M \varphi d\mu \right| > \varepsilon \right\} \leq C e^{-c(\varepsilon)n},$$

where the parameters $n(\varepsilon)$ and $c(\varepsilon)$ depend (explicitly) only on the C^θ-norms of φ, g and $\frac{1}{g}$.

Proof Let $X^+ = M^{\mathbb{N}}$ and let $\mathbb{P}_\mu$ be the Markov measure with initial distribution μ and transition kernel

$$K_x = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{h(x)} \frac{1}{|det D_{x'} f|} \delta_{x'} .$$

Since the Markov system (M, K, μ, C^θ) was shown to be strongly mixing, it follows from the Abstract LDT Theorem 2.1 that the stochastic Birkhoff sums

$$S_n\varphi(\omega) = \varphi(\omega_0) + \varphi(\omega_1) + \ldots \varphi(\omega_{n-1})$$

satisfy an LDT estimate w.r.t $\mathbb{P}_\mu$ for all observables $\varphi \in C^\theta$. More precisely,

$$\mathbb{P}_\mu \left\{ \left| \frac{1}{n} S_n \varphi - \int_M \varphi d\mu \right| > \varepsilon \right\} \leq C e^{-c(\varepsilon) n} .$$

Now we will show that the above estimate can actually be transferred to the deterministic Birkhoff sums of our dynamical system (f, μ). First thing we notice is that although the set of possible symbols is uncountable(these are the points of M) the allowed transitions are finite (recall that K_x is supported on the pre-images of x. So the space of allowed transitions is described by

$$\Omega = \underset{j\in\mathbb{N}}{\cap} \{\omega \in X^+ : f(\omega_{j+1}) = \omega_j\} = \underset{j\in\mathbb{N}}{\cap} \sigma^{-j}\{\omega \in X^+ : f(\omega_1) = \omega_0\}.$$

Notice that

$$\mathbb{P}\{\omega \in X^+ : f(\omega_1) = \omega_0\} = \int K_{\omega_0}(\omega_1) d\mu(\omega_0).$$

By stationarity, we have that $\sigma^{-j}\{\omega \in X^+ : f(\omega_1) = \omega_0\} = 1$ for all $j \in \mathbb{N}$ and from here it follows that $\mathbb{P}(\Omega) = 1$. Consider the deviation set for the deterministic Birkhoff sum, that is, let

$$E := \left\{ x \in M : \left| \frac{1}{n} S_n(\varphi(x)) - \int_M \varphi d\mu \right| > \varepsilon \right\} .$$

By stationarity of the measure μ,

$$\begin{aligned} \mu(E) &= \int K_x^n(E) d\mu(x) = \int \mathbb{P}_x(\omega \in X^+ : \omega_n \in E) d\mu(x) \\ &= \mathbb{P}_x(\omega \in X^+ : \omega_n \in E) = \mathbb{P}_x(\omega \in \Omega : \omega_n \in E) \\ &= \mathbb{P}_\mu \left\{ x \in \Omega \left| \frac{1}{n} S_n(\varphi(\omega)) - \int_M \varphi d\mu \right| > \varepsilon \right\} \leq C e^{-c(\varepsilon) n} \end{aligned}$$

which concludes the proof. ■

In the same context we also have the following central limit theorem.

Theorem 4.15 *(CLT) If $\int_M \varphi = 0$ and φ is not a coboundary, i.e, there is no $\psi \in C^0(M)$ such that $\varphi(x) = \psi(x) - \psi(f(x))$, then there exists $\sigma(\varphi) > 0$,* *such that for any $x \in M$*

$$\frac{S_n(\varphi(x))}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

Proof Assume that φ has zero μ mean and it is not a coboundary. Let us verify the hypothesis of the Abstract CLT for the Markov system (M, K, μ, C^θ). Define $\psi := \sum_{n=0}^{\infty} \mathcal{Q}^n \varphi$ and suppose by contradiction that $\sigma^2(\varphi) := ||\psi||_2^2 - ||\mathcal{Q}\psi||_2^2 = 0$. This would imply that $\mathcal{Q}\psi(x) = \psi(y)$ for μ and K_x a.e point. By the definition of K this implies that $\psi(y) = \mathcal{Q}\psi(f(y))$ for μ a.e. But since ψ , $\mathcal{Q}\psi$ are continuous and $d\mu = hdm$ is continuous and bounded away from zero, we must have that $\psi(y) = \mathcal{Q}\psi(f(y))$ for all $x \in M$. Then

$$\varphi = \psi - \mathcal{Q}\psi = \mathcal{Q}\psi \circ f - \mathcal{Q}\psi$$

which shows that φ is a coboundary a contradiction. Applying the abstract CLT theorem we have that

$$\mathbb{P}\left\{\omega \in X^+ : \frac{S_n\varphi(\omega)}{\sigma\sqrt{n}} \leq \lambda\right\} \to \int_{\infty}^{\lambda} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}. \quad \text{as}\, n \to \infty.$$

By the same argument of the previous theorem we transfer the CLT from the stochastic to the deterministic system thus obtaining

$$\frac{S_n(\varphi(x))}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

■

In summary, the hyperbolicity and regularity of f allow us to show that its transfer operator acts in a smoothing way on the space of θ-Hölder observables. Since we have a spectral gap, we can apply the results developed in the previous chapters to obtain LTD estimates. One path is through decays of correlations (i.e. using Theorem 3.6). The other path, shown above, is via the abstract LDT estimates for strongly mixing Markov systems, namely Theorem 2.1. The two paths lead to similar types of estimates, although the parameters in the latter may be more explicit. Moreover, in the same setting the abstract CLT, Theorem 2.4, is also applicable.

5
Piecewise Expanding Maps

The main goal of this chapter is to introduce piecewise expanding maps on the interval. This is the simplest case where statistical properties are studied in the presence of discontinuities. We plan to:

1 Define expanding maps and list their basic properties;

2 Study some examples of transfer operators for this class of dynamics;

3 Study the action of the transfer operator on the space of densities of bounded variation;

4 Use the quasi-compactness results to derive statistical properties for Birkhoff sums. This will be done in a similar fashion to the previous chapter.

5.1
Basic Properties

Definition 5.1 *(Piecewise monotone map) Let* $I = [a, b]$ *be a compact interval. A map* $f : I \to I$ *is called piecewise monotonic if there is a partition* $\mathcal{P}$ *of* I*,* $a = a_0 < a_1 < \cdots < a_q = b$*, and a number* $n \geq 1$ *such that:*

1 $f|_{(a_{i-1},a_i)}$ *is a* C^r *function, For all* $i = 1, \ldots, q$ *which admits a extension to a* C^r *function on the closed interval* $[a_{i-1}, a_i]$*.*

2 $|f'(x)| > 0$ *on* (a_{i-1}, a_i)*, for* $i = 1, \ldots, q$*.*

Example 5.1 *(tent map) Let* $I = [0, 1]$ *and* $u \in I$ *the tent map with height* u *is given by:*

$$f(x) = \begin{cases} 2ux & \text{if } 0 < x < \frac{1}{2} \\ 2u(1-x) & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases} \tag{5-1}$$

Example 5.2 *Let* $I = [0, 1]]$*,* $f(x) = 2x \mod 1$ *is called the doubling map:*

$$f(x) = \begin{cases} 2x & \text{if } 0 < x < \frac{1}{2} \\ 2x - 1 & \text{if } \frac{1}{2} \leq x \leq 1 \end{cases} \tag{5-2}$$

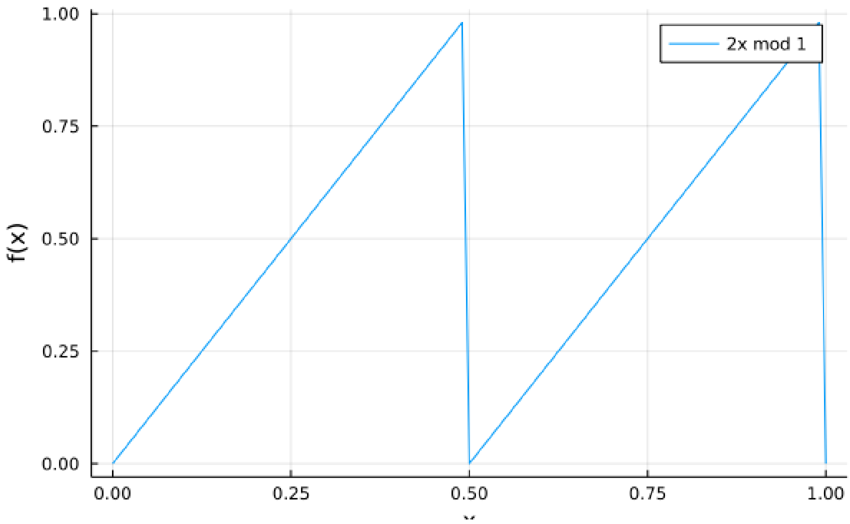

Figure 5.1: Doubling Map

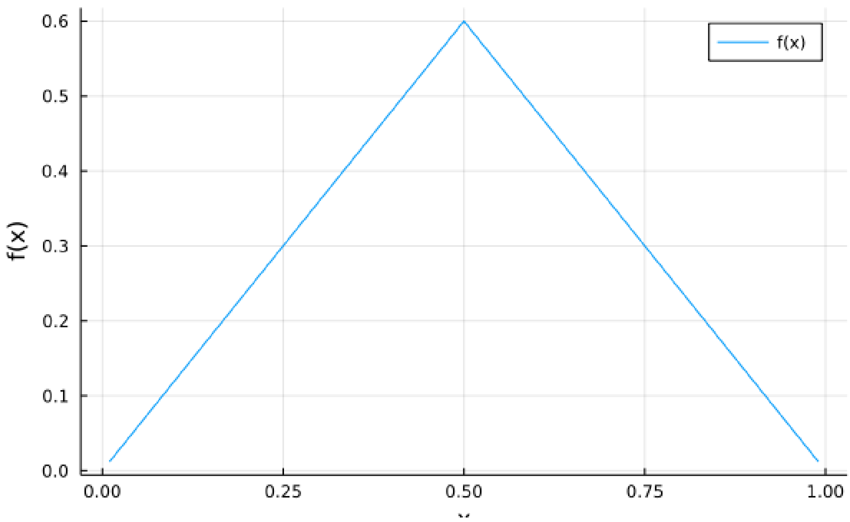

Figure 5.2: Tent Map

5.2 Transfer Operators

Piecewise expanding maps are non-singular so the construction of the transfer operator is the same.

Example 5.3 *(Doubling Map) Let $f : I \to I$, then $\mathcal{L}h(x) = \frac{h(\frac{x}{2})+h(\frac{x+1}{2})}{2}$.*

In fact, the transfer operator admits a nice closed formula which generalizes the example above.

Proposition 5.1 *Let $f : I \to I$ be a piecewise monotone map. Then*

$$\mathcal{L}h(x) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{|f'(x)|}\,. \tag{5-3}$$

Proof See ([4] page 86) ■

Example 5.4 *(Transfer Operator for the Tent Map)*

$$\mathcal{L}h(x) = \frac{h(\frac{x}{2u})}{2u} + \frac{h(\frac{2u-x}{2u})}{2u}\,. \tag{5-4}$$

Unlike the case of expanding maps, the space of Hölder continuous functions is not necessarily preserved by the transfer operator for any monotonic function. Let us define the specific class of dynamics we will study.

Definition 5.2 *(Piecewise Expanding) A map $f : I \to I$ is called piecewise expanding if there exists a partition $\mathcal{P}$ of I, $a = a_0 < a_1 < \cdots < a_q = b$, and a number $n \geq 1$ such that:*

1) The restriction $f|_{(a_{i-1},a_i)}$ to each (a_{i-1}, a_i) is C^2 , and $|f'(x)| \geq \lambda > 2$ for all $i = 1, \ldots, q$ and for all $x \in (a_{i-1}, a_i)$.

2) $\displaystyle\int_I \frac{|f''|}{f'^2} dm < \infty$.

Remark 5.1 *The Tent and Doubling maps trivially satisfy condition 2) above. In fact, these examples already satisfy $\|\frac{f''}{f'^2}\|_\infty < \infty$. The next example, which appears in the study of the Lorenz attractor (see [2]), does not have a uniform bound on the distortion $\dfrac{|f''|}{f'^2}$, but it admits a bounded L^1 norm. Therefore, the results we derive in this chapter also apply to this map.*

Example 5.5 *(One dimensional Lorenz map) Let $I = [0, 1]$. The one dimensional Lorenz map with parameters $\theta \geq 0$ and $0 \leq \alpha < 1$ is given by*

$$f(x) = \begin{cases} \theta|x-2|^\alpha & \text{if } 0 \leq x < \frac{1}{2} \\ (1-\theta)|x-2|^\alpha & \text{if } \frac{1}{2} < x \leq 1 \end{cases}$$

The action of transfer operators corresponding to piecewise expanding maps improves the regularity of the functions of bounded variation.

5.3 Bounded Variation Functions

In this section we collect some relevant facts about bounded variation (BV) functions.

Definition 5.3 *Let $\varphi : I \to \mathbb{R}$. The **variation** var φ is defined by:*

$$var\varphi = \sup \sum_{i=1}^{n} |\varphi(x_{i-1}) - \varphi(x_i)| \tag{5-5}$$

where the supremum is taken over all the finite partitions of the interval I. Given a specific partition $\mathcal{P}$, we denote by $var_{\mathcal{P}}\varphi$ the variation over the partition $\mathcal{P}$.

Definition 5.4 *A function φ is said to be of **bounded variation** if $var\varphi < \infty$. We define the space of functions of bounded variation as*

$$BV = \{\varphi \in L^1 : \varphi < \infty\}.$$

Proposition 5.2 *The following simple properties of the variation hold.*

1. $var(\varphi_1 + \varphi_2) \leq var\varphi_1 + var\varphi_2$.
2. $var(\varphi_1\varphi_2) \leq var\varphi_1 sup\varphi_2 + sup\varphi_1 var\varphi_2$.
3. $var(\varphi_1\psi) \leq var\varphi_1 sup\varphi_2 + sup\psi' \int |\varphi| dm$, *if ψ is C^1.*
4. $var|\varphi| \leq var\varphi$.

Proof see [3] ■

5.4 Quasi-Compactness

In this section we consider the space BV endowed with the norm $||\varphi||_{BV} := var\varphi + ||\varphi||_{L^1}$. Instead of studying the space of densities directly we will actually study the transfer operator on measures. First we define the action of the following functionals on the set of measures on I.

$$|\mu| = \sup_{\varphi \in C^0, \|\varphi\|_\infty = 1} |\mu(\varphi)| \tag{5-6}$$

$$\|\mu\| = \sup_{\varphi \in C^1, \|\varphi\|_\infty = 1} |\mu(\varphi'), | \tag{5-7}$$

where we denote by $\mu(\varphi)$

$$\mu(\varphi) = \int_I \varphi d\mu.$$

These functionals are in fact norms on the space of signed measures M(I).

Note that, for each $\varphi \in C^0(I;\mathbb{R})$ and $\epsilon > 0$, one can find $\varphi_\epsilon \in C^1(I;\mathbb{R})$ such that

$$|\varphi - \varphi_\epsilon| \leq \epsilon |\varphi|_\infty .$$

Writing $\mu(\varphi) = \mu(\varphi - \varphi_\epsilon) + \mu(\varphi_\epsilon)$ we obtain

$$\mu(\varphi) \leq |\mu|\epsilon|\varphi|_\infty + \mu(\varphi_\epsilon) = |\mu|\,\epsilon|\varphi|_\infty + \mu(\frac{d}{dt}\int_0^t \varphi_\epsilon) = (|\mu|\epsilon + ||\mu||(1+\epsilon))|\varphi|_\infty.$$

Since ϵ is arbitrary we get

$$|\mu| \leq |||\mu|| .$$

Lemma 5.3 *Let $\mathcal{B} := \{\mu \in M(I) : ||\mu|| < \infty\}$. If $\mu \in \mathcal{B}$ then μ is absolutely continuous with respect to the Lebesgue measure. Moreover,*

$$\frac{d\mu}{dm} \in L^p, \quad \textit{for all } p < \infty.$$

Proof Let $\varphi \in C^0(I;\mathbb{R})$. Then for each $\epsilon \in (0,1)$ there exists $\varphi_\epsilon \in C^1(\mathbb{R};\mathbb{R})$, supported in $[-\epsilon, 1+\epsilon]$, such that $|\varphi - \varphi_\epsilon|_{C^0(I;\mathbb{R})} \leq \epsilon$, $|\varphi_\epsilon|_\infty \leq 1$, and $|\varphi_\epsilon - \varphi|_\infty \leq |\varphi|_\infty(1+\epsilon)$. Let,

$$F(\psi) := -\frac{1}{2}||\psi||$$

and define the following convolution

$$w_\epsilon(x) = \int F(x-z)\varphi_\epsilon(z)dz.$$

Note that $w''_\epsilon = \varphi_\epsilon$, Therefore,

$$\mu(\varphi) \leq C(|\mu| + ||\mu||)|\varphi|_{L^p}$$

so the linear functional $\mu : C^0 \to \mathbb{R}$ can be extended to a linear functional on L^p. Since the dual of L^p is L^q where q is the dual of p, that is, $\frac{1}{p} + \frac{1}{q} = 1$. It follows that there exists $h \in L^q$ such that $\mu(\varphi) = \int h(x)\varphi(x)dx$. ■

Remark 5.2 *For all $\mu \in \mathcal{B}$ setting $h := \frac{d\mu}{dm}$, we have that $|h|_{L^1} = |\mu|$ and $|h|_{BV} = ||\mu||$. See [3] chapter 3.*

Theorem 5.4 *Let $f : I \to I$ be a piecewise expanding map. Then the corresponding transfer operator is quasi-compact on the space of BV densities.*

As in chapter 4, the proof of the quasicompactness theorem comes from the verification of the hypothesis in Ionescu-Tulcea and Marinescu's theorem. We do that in the following series of lemmas.

Lemma 5.5 *(Continuity) $||.||_{L^1}$ is continuous in $BV(I)$.*
Proof *If $||\varphi_n - \varphi||_{BV} \to 0$ by definition we must have $||\varphi_n - \varphi||_{L^1} \to 0$. By the dominated convergence theorem it follows that $||\varphi_n||_{L^1} \to ||\varphi||_{L^1}$* ■

Lemma 5.6 *(Pre-compactness) The ball $B = \{\mu \in \mathcal{B} : \|\mu\| \leq 1\}$ is relatively compact in $(\mathcal{M}(X), |.|)$*
Proof *For each $t \in \mathbb{N}$, let us consider a partition $\{A_j\}$ of $[0,1]$ into intervals of size t^{-1}. Define*

$$P'_t(x) = t\sum_j 1_{A_j}(x)\int_{A_j}\varphi(x)dx.$$

Firstly of all, note that

$$P'_t\mu(\varphi) := \mu(P_t\varphi) = \int hP_t\varphi = \int P_th \cdot \varphi$$

and

$$P'_t\mu(\varphi') = \int hP_t\varphi' = ||\mu||\,|\int_0^x P_t\varphi' dz| \leq 4||\mu||$$

In addition,

$$\mu(P_t\varphi - \varphi) = ||\mu||\,|\int_0^x P_t\varphi - \varphi dz|\,.$$

If $x_k \in A_j = [j/t, (j+1)/t)$, *then*

$$|\int_0^x (P_t\varphi - \varphi)dz| = \int_{jt^{-1}}^{x_k} \varphi \leq |\varphi|_\infty t^{-1}\,.$$

Thus

$$||P'_t\mu|| \leq 4||\mu|| \quad \textit{and} \quad |P'_t\mu - \mu| \leq 4^2\, t^{-1}\,.$$

Furthermore, $P'_t\mu = t\sum_{i=1}\mu(A_i)m_{A_i}$. *So the range of* P'_t *is a finite dimensional space. This implies that if* $\{\mu_{n_j}\} \subset B$, *then* $\{P'_t\mu\}$ *is in a finite dimensional bounded set, so it is compact. Therefore, there exists* μ_t *and* n_j *such that* $\lim_{j\to\infty}||P'_t\mu_{n_j} - \mu_t|| = 0$. *In addition, for* $t' \geq t$,

$$|\mu_t - \mu_{t'}| \leq |\mu_t - P'_t\mu_{n_j}| + |\mu_t - P'_{t'}\mu_{n_j}| + |P'_t\mu_{n_j} - P'_{t'}\mu_{n_j}| \leq Ct^{-1}$$

It follows that there exists t_j *and a measure* μ *such that*

$$\lim_{j\to\infty} |\mu - P_{t_j}\mu_{n_j}| = 0\,,$$

which proves the pre-compactness of $\mathcal{B}$ ∎

Remark 5.3 *We are working with measures in* $\mathcal{B}$, *which by lemma 5.3 must be absolutely continuous. We recall that the transfer operator and the pushforward* f_* *are related by* $d(f_*\mu) = (\mathcal{L}h)dx$, *if* $d\mu = hdx$.

Lemma 5.7 *(Boundness) For each* $\mu \in \mathcal{B}$, *we have* $|f_*\mu| \leq |\mu|$.

Proof Given $\mu \in \mathcal{B}$,

$$|f_*\mu| = \sup_{\varphi\in C^0, |\varphi|_\infty = 1} |\mu(\varphi \circ f)| \leq |\mu|,$$

which concludes the proof. ∎

The next lemma is a Lasota-Yorke inequality. We closely follow the proof in [10].

Lemma 5.8 *(Lasota-Yorke Inequality) Let f be a piecewise expanding map and $\mu \in \mathcal{B}$. Then*

$$\|f_*\mu\| \leq \frac{2}{\inf f'}\|\mu\| + \frac{2}{\min(d_i - d_{i+1})}\mu(1) + 2\mu\left(\frac{f''}{(f')^2}\right) \tag{5-8}$$

Proof Note that

$$f_*\mu(\varphi') = \sum_{Z\in\{(d_i,d_{i+1})|i\in(1,\ldots,n-1)\}} f_*\mu(\varphi'\chi_Z),$$

since $f_*\mu$ gives zero weight to the points d_i ($f_*\mu$ is absolutely continuous).For each such Z define φ_Z to be linear and such that $\varphi_Z = \varphi$ on ∂Z, then define $\psi_Z = \varphi - \varphi_Z$, on Z, and extend it to $[0,1]$ by setting it to zero outside Z. This is a continuous function. Moreover, for each $x \in Z$, $|\varphi_Z'|_\infty \leq \frac{2|\varphi|_\infty}{\min(d_i - d_{i+1})}$.Thus,

$$|f_*\mu(\varphi')| = \left|\sum_Z \mu(\psi_Z' \circ f\chi_{f^{-1}(Z)}) + \mu(\varphi_Z' \circ f\chi_{f^{-1}(Z)})\right|$$

Note that on each Z we have $\psi' \circ Z \circ f = (\frac{\psi_Z' \circ f}{f'})' + \frac{(\psi_Z \circ f)f''}{(f')^2}$,

$$\begin{aligned}|f_*\mu(\varphi')| &\leq \sum_Z \mu((\psi_Z \circ f(f')^{-1})'\chi_{f^{-1}(Z)}) \\ &\quad + \sum_Z \mu((\psi_Z \circ f)f''(f')^{-2}\chi_{f^{-1}(Z)}) \\ &\quad + \frac{2|\varphi|_\infty\mu(1)}{\min(d_i - d_{i+1})} \\ &\leq |\mu((\psi_Z \circ f(f')^{-1})')| \\ &\quad + 2|\varphi|_\infty\mu(|f''(f')^{-2}|) + \frac{2|\varphi|_\infty\mu(1)}{\min(d_i - d_{i+1})}.\end{aligned} \tag{5-9}$$

Note that

$$|\mu((\frac{\psi_Z \circ f}{f'})')| \leq |\mu|\,|\frac{\psi_Z \circ f}{f'}|_\infty \leq ||\mu||\frac{1}{\inf f'}|\varphi|_\infty$$

From which we obtain

$$|f_*\mu| \leq \frac{2|\varphi|_\infty}{\inf f'}\|\mu\| + \frac{2\varphi_\infty}{\min(d_i - d_{i+1})}\mu(1) + 2\,\varphi_\infty\mu\left(\frac{|f''|}{(f')^2}\right)$$

Taking the supremum over all $\varphi \in C^1(I)$, with $|\varphi|_\infty = 1$ we get,

$$\|f_*\mu\| \leq \frac{2}{\inf f'}\|\mu\| + \frac{2}{\min(d_i - d_{i+1})}\mu(1) + 2\mu\left(\frac{|f''|}{(f')^2}\right).$$

This concludes the proof of the Lasota-Yorke inequality. ∎

5.5 LDT via Decay of Correlations

All the hypotheses of Ionescu-Tulcea and Marinescu theorem have been verified. We then conclude that the transfer operator $\mathcal{L}$ associated to piecewise expanding maps is quasi compact on the space of functions of bounded variation. We may have multiple peripheral eigenvalues. In fact, the peripheral spectrum $\sigma_{BV}(\mathcal{L}) \cap \{|z| = 1\}$ is a finite union of cyclic groups (see [8] Lemma 1.58). The next lemma relates the topological transitivity (mixing) of f to the peripheral spectrum of $\mathcal{L}$.

Lemma 5.9 *If the map f is topologically transitive, then 1 is a simple eigenvalue for f. If all the powers of f are topologically transitive, then $\{1\}$ is the entire peripheral spectrum.*

Proof If 1 is not a simple eigenvalue, then there exists an invariant set A such that $\mu(A) \notin \{0, 1\}$. However, 1_A is in BV, which implies that A contains an open set, and the same applies to A^c (this is true only for $d = 1$) . By topological transitivity, there is an orbit that visits both of these open sets; hence, the sets are not invariant. The same argument applied to f^n establishes the lemma. ∎

In the following we will restrict ourselves to the case where f is topologically mixing so by the lemma above it has a unique absolutely continuous invariant measure.

Corollary 2 *Let f be a piecewise expanding map. Then f has exponential decay of correlations of the BV functions against the L^1 functions.*

Proof This follows the same reasoning as that of proposition 4.10. ∎

From the corollary above and Theorem 3.6, we conclude the following.

Theorem 5.10 *Let $f : M \to M$ be a topologically mixing piecewise expanding map. Let $\mathcal{B} \subset L^\infty(\mu)$ be the set of functions of bounded variation, $\varphi \in \mathcal{B}$. Then for every $\epsilon > 0$ there exist $c(\varepsilon) > 0$ and $n(\varepsilon) \in \mathbb{N}$ such that for all $n \geq n(\varepsilon)$*

$$\mu\left\{x \in M : \left|\frac{1}{n}S_n(\varphi(x)) - \int_M \varphi d\mu\right| > \varepsilon\right\} \leq 2e^{-c(\varepsilon)n},$$

We note that $\tau(\varepsilon)$ and $n(\varepsilon)$ are explicitly determined by the input data.

5.6
Statistical Properties via Markov Systems

This section follows the same routine as the last section of the previous chapter. We will associate to our original transfer operator a new one which will be Markov. Let h be the unique absolutely continuous invariant measure of f, let us consider $d\mu = hdm$ as our new reference measure and define $\mathcal{Q}\varphi(x) = \frac{df_*\mu_\varphi}{d\mu}$, the transfer operator of f with reference measure μ. We recall some properties of this operator already proved in chapter 3.

Proposition 5.11 *The operator $\mathcal{Q}$ defined above satisfies the following properties.*

1. *$\mathcal{Q}$ is Markov,*
2. *μ is a stationary measure with respect to $\mathcal{Q}$,*
3. *$(\mathcal{Q}\varphi)\, h = \mathcal{L}(\varphi h)$.*

Remark 5.4 *The formula in the above proposition implies that $\mathcal{Q}$ has the following closed form:*

$$\mathcal{Q}\varphi(x) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{h(x)} \frac{1}{|detD_{x'}f|}\varphi(x')\,. \tag{5-10}$$

It follows that the associated Markov kernel is

$$K_x(y) = \sum_{x' \in f^{-1}(x)} \frac{h(x')}{h(x)} \frac{1}{|detD_{x'}f|}\delta_{x'}\,, \tag{5-11}$$

so this Markov chain sends x to one of its pre-images.

Proposition 5.12 *The Markov system (M, K, μ, BV) is strongly mixing with exponential rate.*
Proof *It follows from Proposition 3.12 .* ∎

Theorem 5.13 *(LDT) Let $f : M \to M$ a topologically mixing piecewise expanding map, consider $\varphi \in BV$. Given any $\varepsilon > 0$ there are $n(\varepsilon) \in \mathbb{N}$ and $c(\varepsilon) > 0$ such that for all $n \geq n(\varepsilon)$ we have*

$$\mu\left\{\left|\frac{1}{n}S_n(\varphi(x)) - \int_M \varphi d\mu\right| > \varepsilon\right\} \leq Ce^{-c(\varepsilon)n}\,.$$

The parameters $n(\varepsilon)$ and $c(\varepsilon)$ only depend (explicitly) on the BV-norms of φ, g and $\frac{1}{g}$.

Proof This follows the same reasoning as that of theorem 2.2. ∎

In the same context we also have a CLT.

Theorem 5.14 *(CLT) If $\int_M \varphi = 0$ and φ is not a coboundary, i.e, there is no $\psi \in C^0(M)$ such that $\varphi(x) = \psi(x) - \psi(f(x))$, then there exists $\sigma(\varphi) > 0$, such that*

$$\frac{S_n(\varphi(x))}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1).$$

6
Limit Laws for Random Toral Translations

In this chapter we study a partially hyperbolic dynamical system which is not predominantly hyperbolic, namely the skew-product encoding a random toral translation.

Let $\Sigma := \mathbb{T}^d$, where $\mathbb{T} = {}^{\mathbb{R}}/_{\mathbb{Z}}$ is the one-dimensional torus, m the Haar measure on the d-dimensional torus $\mathbb{T}^d$ and let $\mu \in Prob(\Sigma)$ be another probability measure.

Let $X := \Sigma^{\mathbb{Z}}$ and consider $(X, \mu^{\mathbb{Z}})$ as a product probability space of symbols where each symbol is an element of the torus. Let $\sigma : X \to X$ be the bilateral shift map. Finally, define the skew-product map

$$f : X \times \mathbb{T}^d \to X \times \mathbb{T}^d,\ f(\omega, \theta) = (\sigma\omega, \theta + \omega_0).$$

The triple $(X \times \mathbb{T}^d, f, \mu^{\mathbb{Z}} \times m)$ is called a mixed random-quasiperiodic dynamical system. This dynamical system was studied by (Cai et al) in [5, 6] and this chapter closely follows their work. The fiber dynamics encodes the following Markov chain:

$$\theta \to \theta + \omega_0 \to \theta + \omega_0 + \omega_1 \to \dots$$

Its transition kernel is given by $K\colon \mathbb{T}^d \to Prob(\mathbb{T}^d)$,

$$K_\theta = \int \delta_{\theta+\omega_0} d\mu(\omega_0)\,.$$

The corresponding Markov operator is

$$\mathcal{Q} : L^\infty(\mathbb{T}^d) \to L^\infty(\mathbb{T}^d),\ \mathcal{Q}\varphi(\theta) = \int \varphi(\theta + \omega_0) d\mu(\omega_0).$$

The Lebesgue measure m is K-stationary. Therefore $(\mathbb{T}^d, K, m)$ is a Markov system. Now we recall a basic definition in Fourier analysis.

Definition 6.1 *Given $\mu \in Prob(\mathbb{T}^d)$, the Fourier coefficients of μ are*

$$\hat{\mu}(k) = \int_{\mathbb{T}^d} e_k(x) d\mu(x)\,,$$

where $e_k(x) = e^{2\pi i<k,x>}$ *for* $k \in \mathbb{Z}^d$ *and* $x \in \mathbb{T}^d$ *are the characters of the multiplicative group* $\mathbb{T}^d$.

Lemma 6.1 *The characters* $\{e_k : k \in \mathbb{Z}^d\}$ *form a complete basis of eigenvectors for the Markov operator* $\mathcal{Q} : L^2(\mathbb{T}^d) \to L^2(\mathbb{T}^d)$. *That is,* $\mathcal{Q}e_k = \hat{\mu}(k)e_k$, $\forall k \in \mathbb{Z}^d$, *and if* $\varphi = \sum_{k\in\mathbb{Z}^d} \hat{\varphi}(k)e_k$ *in* $L^2(\mathbb{T}^d, m)$, *then*

$$\mathcal{Q}\varphi = \sum_{k\in\mathbb{Z}^d} \hat{\mu}(k)\hat{\varphi}(k)e_k \quad \text{in } L^2(\mathbb{T}^d, m).$$

Proof By the linearity and boundedness of $\mathcal{Q}$, it is enough to prove the first equality. For any $\theta \in T^d$ and any $k \in \mathbb{Z}^d$, we have

$$\begin{aligned}
\mathcal{Q}e_k(\theta) &= \int_{T^d} e_k(\theta + \omega_0)\, d\mu(\omega_0) \\
&= \int_{T^d} e_k(\theta)e_k(\omega_0)\, d\mu(\omega_0) \\
&= e_k(\theta) \int_{T^d} e_k \, d\mu \\
&= e_k(\theta)\hat{\mu}(k).
\end{aligned}$$

Thus, the result follows. ■

We now recall the definition of mixing for the Markov operator and characterize it in terms of the Fourier coefficients of the μ measure.

Definition 6.2 *The system* $(\mathcal{Q}, m)$ *is called mixing if* $\forall\varphi \in C^0(\mathbb{T}^d)$,

$$\mathcal{Q}^n\varphi(\theta) \to \int_{\mathbb{T}^d} \varphi\, dm \quad \text{as} \quad n \to \infty, \quad \forall\theta \in \mathbb{T}^d.$$

Theorem 6.2 *The following statements are equivalent:*

1. $(\mathcal{Q}, m)$ *is mixing.*
2. $|\hat{\mu}(k)| < 1$, $\forall k \in \mathbb{Z}^d \setminus \{0\}$.
3. $\forall k \in \mathbb{Z}^d \setminus \{0\}$, $\forall E \subset \mathbb{T}^d$ *with* $\mu(E) = 1$, $\exists \alpha \neq \beta \in E$ *such that* $\langle k, \alpha - \beta\rangle \notin \mathbb{Z}$.
4. *The semigroup generated by the set* $S := \{\alpha - \beta : \alpha, \beta \in supp(\mu)\}$ *is dense in* $\mathbb{T}^d$.

We say that μ *is mixing if any of the statements above is true.*

Proof

1. (1) $\Rightarrow$ (2): If there is $k \in \mathbb{Z}^d \setminus \{0\}$ such that $|\hat{\mu}(k)| = 1$, then since $\mathcal{Q}^n e_k = \hat{\mu}(k)^n e_k$ for all $n \in \mathbb{N}$, we have $|\mathcal{Q}^n e_k| = |\hat{\mu}(k)^n e_k| = 1 \nrightarrow 0 = \int e_k \, dm$ as $n \to \infty$. This contradicts the mixing condition.

2. (2) $\Rightarrow$ (1): We first establish the convergence in Definition 6.2 for trigonometric polynomials, then proceed by approximation. Let $p = \sum_{|k| \leq N} c_k e_k$ be a trigonometric polynomial. Note that $\int p \, dm = c_0$ and $\hat{\mu}(0) = 1$, so we have

$$\mathcal{Q}^n p - \int p \, dm = \sum_{0 < |k| \leq N} c_k \hat{\mu}(k)^n e_k.$$

Hence,

$$\left\| \mathcal{Q}^n p - \int p \, dm \right\|_{C_0} \leq \sum_{0 < |k| \leq N} |c_k| |\hat{\mu}(k)|^n.$$

Let $\sigma := \max\{|\hat{\mu}(k)| : 0 < |k| \leq N\} < 1$. Then,

$$\left\| \mathcal{Q}^n p - \int p \, dm \right\|_{C_0} \leq \sum_{0 < |k| \leq N} |c_k| \sigma^n \to 0 \quad \text{as} \quad n \to \infty.$$

Given any observable $\varphi \in C_0(\mathbb{T}^d)$ and given $\epsilon > 0$, by the Weierstrass approximation theorem, there exists a trigonometric polynomial p such that $\|\varphi - p\|_{C_0} < \epsilon$. Moreover, by the previous argument, there is $n(\epsilon) \in \mathbb{N}$ such that $\|\mathcal{Q}^n p - \int p \, dm\|_{C_0} < \epsilon$ for all $n \geq n(\epsilon)$. Writing $\varphi = p + \varphi - p$, we have

$$\mathcal{Q}^n \varphi = \mathcal{Q}^n p + \mathcal{Q}^n (\varphi - p)$$

and

$$\int \varphi \, dm = \int p \, dm + \int (\varphi - p) \, dm.$$

Then, for all $n \geq n(\epsilon)$,

$$\left\| \mathcal{Q}^n \varphi - \int \varphi \, dm \right\|_{C_0} \leq \left\| \mathcal{Q}^n p - \int p \, dm \right\|_{C_0} + \|\varphi - p\|_{C_0} + \|\mathcal{Q}^n (\varphi - p)\|_{C_0} \leq 3\epsilon,$$

which proves the mixing of $(\mathcal{Q}, m)$, and it also shows that the convergence in Definition 3.1 must be uniform.

3. (2) $\Leftrightarrow$ (3): Let $k \in \mathbb{Z}^d \setminus \{0\}$. Then $|\hat{\mu}(k)| = 1$ if and only if

$$\int e^{2\pi i \langle k, \alpha \rangle} \, d\mu(\alpha) = 1,$$

which, by the lemma below, is equivalent to $e^{2\pi i \langle k, \alpha \rangle}$ being constant for

μ-a.e. $\alpha \in \Sigma$. This holds if and only if there is $E \subset \Sigma$ with $\mu(E) = 1$ such that for all $\alpha, \beta \in E$, $e^{2\pi i\langle k,\alpha\rangle} = e^{2\pi i\langle k,\beta\rangle}$. This is equivalent to $e^{2\pi i\langle k,\alpha-\beta\rangle} = 1$ for all $\alpha, \beta \in E$, establishing the claim.

4. (3) $\Rightarrow$ (4): The closed semigroup H generated by S can be written as $H = \bigcup_{n\geq 1} S_n$, where $S_n := S + S_{n-1}$. By the Poincaré recurrence theorem, H is also a group. Assuming by contradiction that $H \neq \mathbb{T}^d$, by Pontryagin's duality for locally compact abelian groups, there exists a nontrivial character $e_k : \mathbb{T}^d \to \mathbb{C}$ containing H in its kernel. In particular, this implies that there exists $k \in \mathbb{Z}^d \setminus \{0\}$ such that $\langle k, \theta\rangle \in \mathbb{Z}$ for all $\theta \in S$, which is a contradiction.

5. (4) $\Rightarrow$ (3): Assume by contradiction that for some $k \in \mathbb{Z}^d \setminus \{0\}$ and $E \subset \mathbb{T}^d$, with full μ-measure, we have $\langle k, \beta - \alpha\rangle \in \mathbb{Z}$ or equivalently $e^{2\pi i\langle k,\beta-\alpha\rangle} = 1$, for all $\alpha, \beta \in E$. Because E is dense in $\mathrm{supp}(\mu)$, this implies by continuity that $e^{2\pi i\langle k,\theta\rangle} = 1$ for all $\theta \in S$. Then e_k is a nontrivial character of $\mathbb{T}^d$, and $H := \{\theta \in \mathbb{T}^d : e_k(\theta) = 1\}$ is a proper compact subgroup of $\mathbb{T}^d$. The assumption implies that $S \subset H$, hence the closed semigroup generated by S is contained in H, a contradiction with (4).

■

Lemma 6.3 *Let (Ω, ρ) be a probability space. Assume that $f : \Omega \to \mathbb{C}$ is Lebesgue integrable. If*

$$\int_\Omega f\, d\rho = \int_\Omega |f|\, d\rho,$$

then $\arg f$ *is constant ρ-a.e. That is, $\exists \theta_0 \in \mathbb{R}$ such that $f(x) = e^{i\theta_0}|f(x)|$ for ρ-a.e. $x \in \Omega$.*

Proof Let $\theta_0 := \arg\left(\int_\Omega f\, d\rho\right)$, so we can write

$$\int_\Omega f\, d\rho = e^{i\theta_0} \int_\Omega |f|\, d\rho.$$

Then

$$\begin{aligned}
0 &= \int_\Omega f\,d\rho - \int_\Omega |f|\,d\rho \\
&= e^{-i\theta_0}\int_\Omega f\,d\rho - \int_\Omega |f|\,d\rho \\
&= \int_\Omega e^{-i\theta_0} f - |f|\,d\rho \\
&= \mathrm{Re}\left(\int_\Omega e^{-i\theta_0} f - |f|\,d\rho\right) \\
&= \int_\Omega \mathrm{Re}(e^{-i\theta_0} f) - |f|\,d\rho.
\end{aligned}$$

Since $\mathrm{Re}(e^{-i\theta_0}f) \le |e^{-i\theta_0}f|$, it follows that $\mathrm{Re}(e^{-i\theta_0}f) = |e^{-i\theta_0}f| \ge 0$, ρ-a.e. In particular, $\mathrm{Im}(e^{-i\theta_0}f) = 0$ ρ-a.e. Therefore,

$$e^{-i\theta_0} f = \mathrm{Re}(e^{-i\theta_0}f) = |f|\ \rho\text{-a.e.}$$

which implies $f = e^{i\theta_0}|f|$ ρ-a.e. ■

Now we aim to obtain a strong mixing condition as in Definition 2.9. In order to do that we must consider a quantitative version of mixing for the measure μ.

Definition 6.3 *We say that $\mu \in Prob(\mathbb{T}^d)$ satisfies a mixing Diophantine condition (mixing DC) if*

$$|\hat{\mu}(k)| \le 1 - \frac{\gamma}{|k|^\tau}, \quad \forall k \in \mathbb{Z}^d \setminus \{0\},$$

for some $\gamma, \tau > 0$. In this case, we write $\mu \in DC(\gamma, \tau)$.

Remark 6.1 *The mixing DC is inspired by the concept of the Diophantine condition (DC) for points on the torus. We say that $\alpha \in \mathbb{T}^d$ satisfies the Diophantine condition $DC(\gamma,\tau)$ if*

$$\inf_{j\in\mathbb{Z}} |\langle k, \alpha\rangle - j| \ge \frac{\gamma}{|k|^\tau}, \quad \forall k \in \mathbb{Z}^d \setminus \{0\}.$$

It is usually assumed when talking about a DC for points on the torus that $\gamma > 0$ and $\tau > d$. This is because when $\tau < d$, the set of points satisfying $DC(\gamma,\tau)$ is empty; when $\tau = d$, it has Lebesgue measure zero on $\mathbb{T}^d$; while when $\tau > d$, the set $\bigcup_{\gamma>0} DC(\gamma,\tau)$ has full Lebesgue measure.

We will also need the following theorem on approximation by trigonometric polynomials.

Theorem 6.4 *(Jackson) If $u \in C^\alpha(\mathbb{T}^d)$, that is, if u is α-Hölder continuous, then for all $n \geq 1$ there exists a trigonometric polynomial p_n of degree $\leq n$ satisfying:*

$$||u - p_n||_\infty \leq Cn^{-\alpha}$$

for some universal constant $C > 0$.

Proof See p.49 in [16] ■

The following result shows that the mixing DC of a measure μ ensures the strong mixing of the corresponding Markov system $(\mathcal{Q}, m)$.

Proposition 6.5 *If $\mu \in DC(\gamma, \tau)$, then $\mathcal{Q}$ is strongly mixing with a power rate on any space of α-Hölder continuous functions $C^\alpha(\mathbb{T}^d)$. More precisely, there exist $C < \infty$ and $p > 0$ such that*

$$\left\| \mathcal{Q}^n \varphi - \int \varphi \, dm \right\|_{C^0} \leq C \|\varphi\|_\alpha \frac{1}{n^p} \quad \forall \varphi \in C^\alpha(\mathbb{T}^d), \quad n \geq 1.$$

In fact, p can be chosen as $\frac{\alpha}{\tau} - \iota$ for any $\iota > 0$, in which case C will depend on ι.

Proof Fix an observable $\varphi \in C^\alpha(\mathbb{T}^d)$. The trick for obtaining a sharp rate of convergence is to approximate φ by trigonometric polynomials, with an error bound (and algebraic complexity) correlated to the number of iterates of the Markov operator. Fix n to be this number of iterates. Let N be the degree of approximation, which will be chosen later. Since $\varphi \in C^\alpha(\mathbb{T}^d)$, by Jackson's approximation theorem, there exists a trigonometric polynomial p_N, with $\deg p_N \leq N$, such that for some universal constant $C_0 < \infty$,

$$\|\varphi - p_N\|_{C^0} \leq C_0 \|\varphi\|_\alpha \frac{1}{N^\alpha}.$$

In fact, p_N is the convolution of φ with the Jackson kernel, so

$$p_N = \sum_{|k| \leq N} c_k e_k,$$

where the coefficients c_k satisfy $|c_k| \leq |\hat{\varphi}(k)| \leq \|\varphi\|_\alpha$. We can then write

$$\varphi = p_N + (\varphi - p_N) =: p_N + r_N.$$

By linearity, we have

$$\mathcal{Q}^n \varphi = \mathcal{Q}^n p_N + \mathcal{Q}^n r_N$$

and

$$\int \varphi \, dm = \int p_N \, dm + \int r_N \, dm.$$

Thus,

$$\mathcal{Q}^n\varphi - \int \varphi\, dm = \mathcal{Q}^n p_N - \int p_N\, dm + \mathcal{Q}^n r_N - \int r_N\, dm,$$

which shows that

$$\left\| \mathcal{Q}^n\varphi - \int \varphi\, dm \right\|_{C^0} \le \left\| \mathcal{Q}^n p_N - \int p_N\, dm \right\|_{C^0} + \|\mathcal{Q}^n r_N\|_{C^0} + \int |r_N|\, dm.$$

Due to the bound on $r_N = \varphi - p_N$ and the fact that $\mathcal{Q}$ is a bounded operator with norm 1 on $C^0(\mathbb{T}^d)$, the second and third terms on the right-hand side above are smaller than $C_0\|\varphi\|_\alpha \frac{1}{N^\alpha}$. It remains to estimate the first term. Since $p_N = \sum_{|k|\le N} c_k e_k$,

$$\mathcal{Q}^n p_N = \sum_{|k|\le N} c_k \mathcal{Q}^n e_k = \sum_{|k|\le N} c_k \hat{\mu}(k)^n e_k.$$

This implies

$$\begin{aligned}\left\| \mathcal{Q}^n p_N - \int p_N\, dm \right\|_{C^0} &\le \sum_{0<|k|\le N} |c_k||\hat{\mu}(k)|^n \\ &\le \|\varphi\|_\alpha \sum_{0<|k|\le N} (1-\gamma|k|^{-\tau})^n \\ &\le |\varphi\|_\alpha (2N)^d (1-\gamma N^\tau)^n.\end{aligned}$$

Using the inequality $(1-x)^{\frac{1}{x}} \le e^{-1}$ for $x \in (0,1)$, we have (for N large enough)

$$(1-\gamma N^\tau)^n \le e^{-n\gamma N^\tau}.$$

Combining the above estimates, we obtain

$$\left\| \mathcal{Q}^n\varphi - \int \varphi\, dm \right\|_{C^0} \le 2d\|\varphi\|_\alpha (2N)^d e^{-n\gamma N^\tau} + 2C_0\|\varphi\|_\alpha \frac{1}{N^\alpha}.$$

Fix any $\epsilon > 0$ and choose $N := (n\gamma)^{\frac{1}{1-\epsilon}\tau}$. Then

$$\frac{1}{N^\alpha} = \frac{1}{\gamma^{\frac{\alpha}{\tau}}(1-\epsilon)^{\frac{1}{n}(\frac{\alpha}{\tau}-1)}} = C_1\frac{1}{n^p},$$

where $p := \frac{\alpha}{\tau}(1-\epsilon) = \frac{\alpha}{\tau} - o(1)$, while

$$(2N)^d e^{-n\gamma N^\tau} = (n\gamma)^{\frac{d}{1-\epsilon}} e^{-(n\gamma)^\epsilon} \ll \frac{1}{n^p}$$

for n large enough. This completes the proof provided the constant C is chosen large enough depending on α, γ, τ, d, and ϵ. ■

As a consequence of the mixing of this operator and the abstract CLT & LDT theorems, we conclude the following.

Theorem 6.6 *Assume that $\mu \in DC(\gamma, \tau)$. Then for all $\varphi \in C^\alpha(\mathbb{T}^d)$, for all $\theta \in \mathbb{T}^d$, and for all $\epsilon > 0$, there exists $n(\epsilon) \in \mathbb{N}$ such that for all $n \geq n(\epsilon)$, we have*

$$\mu^{\mathbb{N}} \left\{ \left| \frac{1}{n} S_n\varphi - \int \varphi \, dm \right| > \epsilon \right\} < e^{-c(\epsilon)n},$$

where $c(\epsilon) = c\epsilon^{2+\frac{1}{p}}$, $n(\epsilon) = n\epsilon^{-\frac{1}{p}}$ for constants $c > 0$ and $n \in \mathbb{N}$, which depend explicitly and uniformly on the data.

Theorem 6.7 *Assume that $\mu \in DC(\gamma, \tau)$ and let $\alpha > \tau$. Then for every $\varphi \in C^\alpha(\mathbb{T}^d)$ nonzero with zero mean, there exists $\sigma = \sigma(\varphi) > 0$ such that*

$$\frac{S_n\varphi}{\sigma\sqrt{n}} \xrightarrow{d} \mathcal{N}(0, 1).$$

Bibliography

[1] ALVES, J. F.; FREITAS, J. M.; LUZZATTO, S. ; VAIENTI, S. From rates of mixing to recurrence times via large deviations, **Adv. Math.**, v.228, n.2, p. 1203–1236, 2011.

[2] ARAÚJO, V.; GALATOLO, S. ; PACIFICO, M. J. Decay of correlations for maps with uniformly contracting fibers and logarithm law for singular hyperbolic attractors, **Math. Z.**, v.276, n.3-4, p. 1001–1048, 2014.

[3] BALADI, V. **Positive transfer operators and decay of correlations**, volume 16 of **Advanced Series in Nonlinear Dynamics**. World Scientific Publishing Co., Inc., River Edge, NJ, 2000. x+314p.

[4] BOYARSKY, A.; GORA, P. **Laws of chaos: invariant measures and dynamical systems in one dimension**. Springer Science & Business Media, 2012.

[5] CAI, A.; DUARTE, P. ; KLEIN, S. Mixed random-quasiperiodic cocycles, **Bull. Braz. Math. Soc. (N.S.)**, v.53, n.4, p. 1469–1497, 2022.

[6] CAI, A.; DUARTE, P. ; KLEIN, S. **Statistical properties for mixing markov chains with applications to dynamical systems**, 2024.

[7] CHAZOTTES, J.-R.; GOUËZEL, S. Optimal concentration inequalities for dynamical systems, **Comm. Math. Phys.**, v.316, n.3, p. 843–889, 2012.

[8] DEMERS, M. F.; KIAMARI, N. ; LIVERANI, C. **Transfer operators in hyperbolic dynamics—an introduction**. 33 º Colóquio Brasileiro de Matemática. Instituto Nacional de Matemática Pura e Aplicada (IMPA), Rio de Janeiro, 2021. 238p.

[9] DUARTE, P.; KLEIN, S. **Lyapunov exponents of linear cocycles**, volume 3 of **Atlantis Studies in Dynamical Systems**. Atlantis Press, Paris, 2016. xiii+263p. Continuity via large deviations.

[10] GALATOLO, S.; NISOLI, I. An elementary approach to rigorous approximation of invariant measures, **SIAM Journal on Applied Dynamical Systems**, v.13, n.2, p. 958–985, Jan. 2014.

[11] GLYNN, P. W.; ORMONEIT, D. Hoeffding's inequality for uniformly ergodic Markov chains, **Statist. Probab. Lett.**, v.56, n.2, p. 143–146, 2002.

[12] GORDIN, M.; HOLZMANN, H. The central limit theorem for stationary Markov chains under invariant splittings, **Stoch. Dyn.**, v.4, n.1, p. 15–30, 2004.

[13] GORDIN, M. I.; LIFŠIC, B. A. Central limit theorem for stationary Markov processes, **Dokl. Akad. Nauk SSSR**, v.239, n.4, p. 766–767, 1978.

[14] HENNION, H.; HERVÉ, L. **Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness**, volume 1766 of **Lecture Notes in Mathematics**. Springer-Verlag, Berlin, 2001. viii+145p.

[15] IONESCU TULCEA, C. T.; MARINESCU, G. Théorie ergodique pour des classes d'opérations non completement continues, **Ann. Math. (2)**, v.52, p. 140–147, 1950.

[16] KATZNELSON, Y. **An introduction to harmonic analysis**. Camb. Math. Libr. 3rd ed.. ed., Cambridge: Cambridge University Press, 2004.

[17] MELBOURNE, I. Large and moderate deviations for slowly mixing dynamical systems, **Proc. Amer. Math. Soc.**, v.137, n.5, p. 1735–1741, 2009.

[18] OLIVEIRA, K.; VIANA, M. Fundamentos da teoria ergódica, **IMPA, Brazil**, p. 3–12, 2014.

[19] YOUNG, L.-S. Statistical properties of dynamical systems with some hyperbolicity, **Annals of Mathematics**, v.147, n.3, p. 585–650, 1998.

[20] YOUNG, L.-S. Recurrence times and rates of mixing, **Israel Journal of Mathematics**, v.110, n.1, p. 153–188, 1999.