



HAL
open science

Rigorous computational methods for understanding the statistical behavior of random dynamical systems

Luigi Marangio

► **To cite this version:**

Luigi Marangio. Rigorous computational methods for understanding the statistical behavior of random dynamical systems. Other [cs.OH]. Université Bourgogne Franche-Comté; Università degli studi (Pise, Italie), 2021. English. NNT : 2021UBFCD009 . tel-03346411

HAL Id: tel-03346411

<https://theses.hal.science/tel-03346411>

Submitted on 16 Sep 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT EN CO-TUTELLE
DE L'ÉTABLISSEMENT UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ
PRÉPARÉE À L'UNIVERSITÉ DE FRANCHE-COMTÉ
ET DE L'ÉTABLISSEMENT UNIVERSITÀ DI PISA

École doctorale n°37
Sciences Pour l'Ingénieur et Microtechniques

Doctorat d'Informatique et Mathématique

par

LUIGI MARANGIO

**Rigorous computational methods for understanding the statistical behavior
of random dynamical systems**
**Méthodes de calcul rigoureuses pour comprendre le comportement
statistique des systèmes dynamiques aléatoires**

Thèse présentée et soutenue à Belfort, le 27 avril 2021

Composition du Jury :

VAIENTI SANDRO	Professeur, Université de Toulon, CPT (UMR 7332), Toulon, France	Rapporteur
BAHSOUN WAEL	Professeur, Loughborough University	Rapporteur
NISOLI ISAIA	Professeur, Universidade Federal do Rio de Janeiro	Examineur
GALATOLO STEFANO	Professeur, Università di Pisa	Co-encadrant
GUYEUX CHRISTOPHE	Professeur, Université Bourgogne Franche-Comté	Co-encadrant

ACKNOWLEDGEMENTS

I would like to thank my supervisors Christophe and Stefano. In addition to their academic skills, I met two great people and I believe that now more than ever the world needs people like them.

I'm grateful to both departments where I did research during these years and, of course, to my colleagues. A warm thought goes to Isaia and the UFRJ, a personal and academic experience that will always have a place in my heart.

I thank the Université Franco-Italienne, which has partly financed the mobility of my doctoral program, and which has been very helpful during this last difficult year. I would also like to thank EIPHI for financing my visit to the UFRJ.

A special thank you to the Mesocentre, where part of the experiments reported in this thesis were performed.

Obviously, none of this would have been possible without my colorful and wandering family; I promise that one day I will repay you for all that you give me every day.

Mari ove le lentiggini isolate echeggiano.

CONTENTS

I	Random Dynamical Systems	7
1	Random Dynamical Systems	9
1.1	Random transformation	9
1.1.1	Transfer Operators	11
1.1.2	Stationary measures	13
1.2	Existence and regularity of stationary densities	14
2	Ergodic Theorems	19
2.1	Birkhoff Ergodic Theorem	19
2.1.1	Hopf ergodic maximal lemma and the σ -algebra of the invariant sets	21
2.1.2	Convergence of Birkhoff averages	24
3	Perron-Frobenius Operator	29
3.1	Perron-Frobenius operator	29
3.1.1	Ergodicity and mixing	32
3.1.2	Existence of stationary measure for positive contractions	35
3.1.3	Regularity of Perron-Frobenius iterates	38
II	Rigorous Schemes of Approximation	41
4	Systems with Additive Noise	43
4.1	The Bounded Variation Space	44
4.2	Existence and Regularity of Stationary Densities	50
4.3	Linear Response	52
4.4	Rate of mixing and perturbations	56
4.4.1	Perturbing the map	56

4.4.1.1	Perturbing the noise	57
5	Rigorous Finite Schemes of Approximation	59
5.1	Ulam Approximation	60
5.1.1	Bounding the rate of mixing	62
5.1.2	Stronger bounds and coarse-fine strategy	65
5.2	Fourier Approximation	67
5.2.1	Bounding the rate of mixing	71
5.2.2	Error by using the DFT	75
III	Experiments and Results	77
6	Applications	79
6.1	Associative Memory in the Hippocampus as Random Dynamical System	79
6.2	The Rotation number of the Arnold map	85
6.3	Differentiability of the rotation number for strong nonlinearity	89
6.3.1	Non-monotonic rotation number for strong nonlinearity	94
7	Numerical Aspects	95
7.1	Ulam-based scheme of approximation	98
7.1.1	Building the Ulam matrix	98
7.1.2	Bounding the errors	100
7.2	Fourier-based scheme of approximation	102
7.2.1	Building the Fourier matrix	102
7.2.2	Bounding the errors	103
7.2.3	An introductory example	104

INTRODUCTION

In 2019 it has been estimated that the amount of digital data in the world is 40 zettabytes, 40 times the number of observable stars in the universe. Obviously providing an accurate estimate of this amount is impossible, but looking at our society, it is not difficult to imagine that it is so. Managing this amount of data and digital devices is a huge challenge for mankind, involving virtually every area of knowledge, from metallurgy to ethics. Digital data nowadays are produced in every observable context of human experience, and one of their main characteristics is to be in continuous evolution because they represent a changing reality; furthermore, it is clear to all, that human reality is permeated by chaos, that is, the evolution of natural phenomena cannot be exactly predicted. It arises spontaneously then, to wonder if, despite their intrinsic complexity, it is possible to understand, in some sense, the evolution of this type of complex systems. This question is naturally placed in the mathematical field, and one of the techniques currently used to model an observable phenomenon is to describe the evolution of a system through the evolution of a certain stochastic process, a random dynamical system. This allows us to take into account that nature rarely follows laws that are only deterministic or only random, but often everything is generated by the competition of these two tendencies. In any case, recall that chaotic behavior can arise in a purely deterministic system. The purpose of this thesis is to try to answer the following question: given a random dynamical system, are we able to predict (not just simulate) its behavior in the long term? When we say "predict and not simulate," we mean to have mathematical certainty of the long-term behavior of the model under consideration. If I put a drop of black ink in a glass of water, the short term interactions between these two liquids are unpredictable. However, we can all say with certainty that after enough time, the water and ink mix completely, the glass is filled with a grey liquid, the system will have reached its equilibrium state.

The example just exposed, can be formalized in many ways, one of these is to consider a certain operator defined on the appropriate spaces, which maps measures to measures, and prove for example that this operator admits as fixed point the Lebesgue's measure properly normalized (the state of equilibrium). This type of operators, are called Transfer Operators, and are linear operators defined on appropriate Banach spaces, and are the object of study of this thesis.

A transfer operator associated with a dynamical system describes the evolution of the densities according to the dynamics (in the previous example, the drop of ink that we

can model with a delta of Dirac, will evolve in the measure of Lebesgue). The random dynamical systems can be studied substantially from two points of view, annealed and quenched; in the first case, one looks at the average behavior of the evolution of the system with respect to the random parameters, while in the second the focus is on a fixed realization of the random parameters. In this thesis, we will deal only with the annealed case, for results in the quenched case we refer for instance to [18; 1], and [50] for a remarkable application. Indeed, while in the first case the stochastic stability properties necessary to study the system are ensured by the presence of noise, in the second case this is no longer true and the relevant objects of study become the Oseledets-Lyapunov spectrum of some transfer operator cocycle associated to random products of maps.

Therefore we will study the dynamics of a system via annealed transfer operators (see definition 1.5). Under appropriate assumptions, these transfer operators admit fixed points, which are the stationary measures of the system, and our goal is to calculate these fixed points, which represent the statistic behavior of the system under consideration after a long time; a link between stationary measures and statistical behavior of the system, is proved in theorem 2.6, a variant of the Birkhoff's ergodic theorem, which, roughly speaking, show that the time averages of an observable converge to the spatial average with respect to the invariant measure.

Furthermore we can also study how these equilibrium states, the invariant measures, vary with respect to small perturbations of the system, i.e. we can study the **statistical stability** of the system; if this variation is smooth we say that the system admit **Linear Response**.

Unfortunately, in many cases, even simple ones, the calculation of the stationary measure, or better, of its density with respect to the Lebesgue measure, cannot be addressed analytically. The strategy we follow is therefore to rigorously approximate these densities; we remark on the importance of the fact that we are interested in having rigorous quantitative estimates, and not in numerically simulating a system. Numerical simulations are not mathematical proofs and often lead to totally misleading conclusions, see for instance [14; 26; 25; 29; 27], or 7.1 for a simple example. The central core of this interdisciplinary thesis is to use mathematical ideas that allow us to have an explicit estimation of the errors related to these rigorous approximations, combined with an accurate computer science implementation, which leads to have quantitative theorems proved with the aid of a computer.

In fact, since the transfer operator associated with a random dynamical system is a linear operator, we can expect to approximate it with a finite rank operator; this is done by a suitable composition of the transfer operator with a projection operator on a subspace of finite dimension. At this point, in order to have an invariant density approximation, it will be enough to calculate the eigenvector relative to the maximum eigenvalue of this

matrix. All this procedure, if implemented in floating-point arithmetic, will not give us rigorous information about the system; however, by using the so-called interval arithmetic, we can keep track of all the errors numbers during the computation and if we combine this information with the theoretical estimates that come from the appropriate quantitative stability theorems, we obtain a rigorous approximation of the invariant measure of a system. There are three errors to manage during this process:

- the discretization error, coming from the fact that we are approximating the transfer operator with a finite rank operator; to estimate this error we use theoretical results on the stability of the fixed points of the operators under consideration, see for instance lemma 5.6.
- the approximation error, coming from the fact that we represent this finite rank operator with a matrix that cannot be exactly computed, but will be in fact approximated
- the eigen error, coming from the numerical computation of the eigenvector of this matrix.

These theoretical errors, added to the numerical errors on the computations which we automatically keep track of since we work with the interval arithmetic framework, provide a rigorous approximation of the equilibrium state of a random dynamical system. Our algorithms, return in output an approximation of the invariant density (a vector of coefficients) and a rigorous bound on the norm of the difference between the "true" invariant density, and the one calculated; the goodness of these rigorous approximations, depends on how small is the error computed.

This strategy has been presented and successfully used for instance in [24; 23; 44; 48]. As we said, one way to reduce a linear operator to finite dimension is to compose it with appropriate projections onto spaces of finite size. This procedure is called Galerkin's method, and Ulam's method, which has been used in the above examples, is a special case of it, in which we choose as a finite basis of functions on which to project, the piecewise constant functions. In simple terms, we divide an interval into many small intervals and keep track of how the dynamics of the system moves these intervals. Ulam's method has been studied intensively in the literature [17; 45; 11; 13]. In dynamical system with additive noise of dimension one, the Ulam's method offer a valid and relatively fast method to compute the stationary measure; however if on the one hand, the Ulam basis is easy to implement and does not require assumptions of regularity on the system under consideration, on the other hand, it is not useful to approach multidimensional problems. In fact, its strength, that is not to require regularity, is also its weakness, if there is regularity, the Ulam basis is not able to use it. Of course, considering smaller models, on one hand reduces the information on the system, on the other hand allows to manage

it computationally. For this reason it is important to find the right compromise, and if we want to approach systems of dimension larger than one, Ulam's method may be no longer suitable for rigorous computations, due to its computational cost. In this perspective, in this thesis we present a report of a work in progress [47]: a rigorous approximation strategy, which uses the Fourier basis of functions. To get an idea of the "gain" in terms of approximation that results, we present a first simple, but not trivial, example, where we can obtain errors of the order of 10^{-80} , while with the Ulam method we had errors of the order of 10^{-10} or 10^{-20} ; this comes from the fact that the Fourier basis guarantees us exponential bounds on the discretization and approximation error, if the system is enough smooth. In the case of random system with additive noise, this regularity will be granted by the use of additive noise with Gaussian kernel; we also present theoretical results that allow us to obtain a bound on the approximation error of a system with noise with a different kernel, starting from the bounds for the same system with Gaussian noise. In the author's opinion, the strategy presented represents a promising first step towards the construction of a rigorous approximation algorithm for multidimensional systems.

In literature we can find other examples of rigorous approximation schemes, including Fourier-based ones. Regarding the latter: Pollicott, Jenkinson and others have presented an approach, which involves the computation of sums over periodic orbits, based on the theory of dynamical zeta function [35; 36]. In [56] has been proved that for circular billiards, the convergence of finite-rank approximations using a Fourier basis follows a power law where the power depends on the smoothness of the source distribution driving the system. In [15], using a Fourier-analytic method, it has been shown the first example of a rigorous scheme for approximating the peripheral spectral data of the Perron-Frobenius operator of an Anosov map without mollification. In [63], two spectral Galerkin methods were presented, based on the Fourier exponential bases and the Chebyshev polynomial bases, for a class of systems satisfying certain assumption. The invariant measure for the Lanford map was then computed, leading to a bound on the error of about 10^{-130} . Regarding rigorous Ulam approximations: in [24] a Ulam based approximation strategy was introduced which is the basis of the work presented in this thesis; This strategy has been refined in [23], where it has been proved that in the Matsumoto-Tsuda model of the famous Belosouv-Zabotinsky reaction, an increase in amplitude of the noise causes the Lyapunov exponent to decrease from positive to negative. This phenomenon is known as **Noise Induced Order**. In [48], using a strategy based on linear response theory and Ulam-based rigorous approximation, it was shown that the rotation number associated with a family of Arnold maps is differentiable as a function of one of the parameters over a large range of values, and that in the case of strong nonlinearity, the rotation number is non-monotonic.

We now turn to the description of the chapters that make up this thesis, specifying basic references for each.

In **Chapter 1** In chapter 1 we introduce the main objects of study of this thesis: we will see what is a random transformation and how it induces a linear operator, called transfer operator. We will be interested in the study of the fixed points of this operator, the stationary measures, of which we will demonstrate the existence. In this chapter, as well as in the next two, we will introduce standard notions from the literature, the texts to which we have referred are [2], [8], [40], [41], [20].

We then turn to **Chapter 2**, where we present, in brief, a variant for random dynamical system of the famous Birkhoff's ergodic theorem, one of the classical topics of ergodic theory. Indeed, in **Chapter 3**, we will define a class of operators, the Perron Frobenius operators, induced by non-singular transformations, which satisfy the assumptions of the ergodic theorem; we will introduce the concepts of ergodic transformation and mixing and prove the existence of stationary measures for positive contractions (of which the Perron Frobenius operators are a special case).

In **Chapter 4**, we introduce a particular class of random dynamical systems, dynamical systems with additive noise; we will study the existence and regularity of stationary measures for this type of systems, as well as introduce the necessary function spaces (e.g., the space of bounded-variation functions). We then introduce Linear Response theory, which allows us to analyze how a stationary measure changes in a perturbed random system. Linear Response theory deals with the following problem: if a dynamical system is submitted to some perturbation how does its invariant measure vary? If this change occurs in a smooth way, we say that the system admits Linear Response, and this change can be described by a suitable derivative. We refer to [9] for a survey about linear response for deterministic system (which was first achieved by Ruelle in the uniformly hyperbolic case [55; 54]), meanwhile for the random case, we refer to [6], [5], [4], [22].

After having developed the necessary theoretical framework in the previous chapters, in **Chapter 5** we introduce the finite schemes of rigorous approximation: as we have already said, we are interested in the computation of a fixed point of a linear operator. To do this we project this operator onto a finite basis of functions and study this "discretized" operator. Since it is a finite-dimensional operator, we can calculate a matrix to approximate it and study its properties with a computer. We will also provide estimates of the various errors made during this approximation process, which will eventually lead to what we call a rigorous approximation. In the first section of this chapter, we will describe a strategy based on Ulam's method, i.e. we will choose as a basis of functions, the indicator functions. In the second section, we will present a work in progress, where we will choose as basis of functions the Fourier basis.

In **Chapter 6** we present two applications of the Ulam method described above; in particular we will focus on a family of random dynamical systems, the Arnold maps, and show

how the calculation of the stationary measure combined with Linear Response theory, introduced in [Chapter 4](#), leads to interesting theoretical results, for example, the differentiability of the rotation number. The rotation number, which is a particular function associated with a dynamical system, in the case of the Arnold maps, presents a surprising behavior (see for instance [Fig. 6.5](#)), which seems to go through a smoothing process once the noise has been applied. In [\[48\]](#), a linear response theorem has been applied, where the quantitative estimate required in the assumptions of this statement has been obtained via rigorous computation. As a consequence, the differentiability of the rotation number has been derived; in fact, the latter can be described as the integral of an appropriate observable with respect to the invariant measure of the system. Furthermore, by combining quantitative estimates on the rotation number with stochastic stability theorems, the non-monotonicity of the rotation number for a large set of parameters, in the case of strong nonlinearity, has been deduced.

To conclude in [Chapter 7](#) we describe the numerical details of the work presented and the pseudocodes of the algorithms used, focusing on the Fourier-based algorithm that we have introduced. We conclude by showing an example application of the Fourier-based strategy, which shows the potentiality of this method.



RANDOM DYNAMICAL SYSTEMS

RANDOM DYNAMICAL SYSTEMS

One of the surprises of the '60s and '70s was the discovery that three ordinary differential equations are sufficient to generate an irregular behavior (Lorentz attractor), an example of what today we call deterministic chaos. However, this kind of behavior is not enough to explain (or at least model) everything we observe in nature: for example in the phenomena of ocean-atmosphere interaction, we know that there is a myriad of factors that interact on virtually every spatial and temporal scale, and that actually contribute to the dynamics of the system. Mathematical models are often simplified descriptions that leave out some parameters or details present in reality. One of the ways in which we can integrate a model in a manner that takes into account very different interactions and behaviors is to use random noise. This can be done through different mathematical formalizations, such as stochastic differential equations, but in this thesis, we will use random dynamical systems, whose role in applied sciences has become of fundamental importance in recent years. This first chapter will introduce the basic theory of random dynamical systems, with particular attention to the associated transfer operators and the existence of stationary measures.

1.1/ RANDOM TRANSFORMATION

Let (S, \mathcal{S}, p) be a probability space and let $(\Omega, \mathcal{A}, \mu)$ be the corresponding symbolic space of the one-sided sequences endowed with the σ -algebra $\mathcal{A} = \mathcal{S}^{\mathbb{N}}$ and the probability measure $\mu = p$; we will fix these notations throughout the text and we will also use the standard notation for the shift map on Ω , $\sigma(\{\omega_i\}_{i \in \mathbb{N}}) = \{\omega_{i+1}\}_{i \in \mathbb{N}}$.

Definition 1.1. Let (M, \mathcal{B}) be a measurable space and let $\mathcal{A} \otimes \mathcal{B}$ be the product σ -algebra on the space $\Omega \times M$. A **random transformation** over σ is a measurable function

$$F : \Omega \times M \rightarrow \Omega \times M, \text{ such that}$$

1. $F(\omega, x) = (\sigma(\omega), F_\omega(x)),$

2. the map $\omega \rightarrow F_\omega$ depends only from the 0-th coordinate of ω ; therefore we shall use, with abuse, the notation F_{ω_0} .

Definition 1.2. Let (M, \mathcal{B}) be a measurable space and let $F : \Omega \times M \rightarrow \Omega \times M$ a random transformation. Given $\omega \in \Omega$, a **random orbit** for F at starting point x is the sequence $\{x_\omega^i\}_{i \in \mathbb{N}}$ defined by

$$x_\omega^0 = x, \quad x_\omega^n := F_{\omega_{n-1}} \circ \cdots \circ F_{\omega_0}(x).$$

We may also consider the canonical projections $\pi_\Omega : \Omega \times M \rightarrow \Omega$ and $\pi_M : \Omega \times M \rightarrow M$; if q is a measure on $\Omega \times M$, we say that q projects to $\pi_\Omega(*q)$ and that $\pi_M(*q)$ is the **marginal** of q .

The following is a standard example of random transformation.

Example 1.3. Let $\{G_1, \dots, G_r\}$ be a finite set of homeomorphism over a compact metric space M , and let $S = \{1, \dots, r\}$, $\Omega = S^{\mathbb{N}}$. The map G defined as

$$G(\sigma(\omega), G_{\omega_0}(x))$$

is a random transformation over σ .

We are going to use the next proposition to define the transfer operator, which is one of the main mathematical objects under study in this thesis.

Proposition 1.4. *The operators $P_x : \Omega \rightarrow \Omega \times M$ and $P^\omega : M \rightarrow \Omega \times M$, defined by $P_x(\omega) = (\omega, x)$ and $P^\omega(x) = (\omega, x)$ are well-defined. Moreover, for any measurable set $E \subset \Omega \times M$*

1. *the slices E_x and E^ω of E are measurable set of \mathcal{A} and \mathcal{B} .*
2. *for any probability measures μ on (Ω, \mathcal{A}) and η on (M, \mathcal{B}) , the maps $x \rightarrow \mu(E_x)$ and $\omega \rightarrow \eta(E^\omega)$ are measurable.*
3. $\int_M \mu(E_x) d\eta(x) = (\mu \times \eta)(E) = \int_\Omega \eta(E^\omega) d\mu(\omega)$.

Proof. P_x is measurable, because $\pi_m \circ P_x(\omega) = x$ and $P_x \circ \pi_\Omega(\omega, x) = (\omega, x)$ are both measurable (the first is a constant function, the second is the identity); similarly P^ω is measurable and $E_x = (P_x)^{-1}(E)$ and $E^\omega = (P^\omega)^{-1}(E)$ are measurable. The proof of the second and third item proceed in standard way [reference]. We will show that the set

$$\mathcal{E} = \{E \subset \Omega \times M : x \mapsto \mu(E_x) \text{ is measurable and } \int_M \mu(E_x) d\eta(x) = (\mu \times \eta)(E)\}$$

is a λ -set that contains the π -system of the sets $A \times B \in \mathcal{A} \times \mathcal{B}$. The proof is analogous for $\omega \mapsto \eta(E^\omega)$.

We first observe that $A \times B \in \mathcal{A} \times \mathcal{B}$ is a π -system, because $(A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B')$; moreover a set of the form $A \times B$ belongs to \mathcal{E} since

$$\mu((A \times B)_x) = \begin{cases} \mu(A) & \text{if } x \in B, \\ 0 & \text{if } x \in M \setminus B. \end{cases}$$

and

$$\int_M \mu((A \times B)_x) d\eta(x) = \mu(A)\eta(B) = (\mu \times \eta)(A \times B).$$

Finally for \mathcal{E} being a λ -system it must hold that $\Omega \times M \in \mathcal{E}$. The maps $x \mapsto \mu(\Omega \times M)_x = \mu(\Omega)$ and $x \mapsto \mu((E^c)_x) = \mu(\Omega) - \mu(E_x)$ are measurable and

$$\begin{aligned} \int_M \mu(\Omega \times M)_x d\eta(x) &= \mu(\Omega)\eta(M) = (\mu \times \eta)(\Omega \times M), \\ \int_M \mu(E^c)_x d\eta(x) &= \mu(\Omega)\eta(M) - \int_M \mu(E_x) d\eta(x) = (\mu \times \eta)(E^c). \end{aligned}$$

Let $\{E(i)\}_{i \in \mathbb{N}}$ a sequence of pairwise disjoint sets in \mathcal{E} ; observe that $(\cup_{i \in \mathbb{N}} E(i))_x = \cup_{i \in \mathbb{N}} (E(i)_x)$. Thus the map $x \mapsto \mu((\cup_{i \in \mathbb{N}} E(i))_x) = \sum_{i \in \mathbb{N}} \mu(E(i)_x)$ is given by a series of positive measurable functions, hence it is measurable and

$$\begin{aligned} \int_M \mu((\cup_{i \in \mathbb{N}} E(i))_x) d\eta(x) &= \sum_{i \in \mathbb{N}} \int_M \mu(E(i)_x) d\eta(x) = \sum_{i \in \mathbb{N}} (\mu \times \eta)(E(i)_x) = \\ &= (\mu \times \eta)((\cup_{i \in \mathbb{N}} E(i))_x). \end{aligned}$$

□

1.1.1/ TRANSFER OPERATORS

Suppose to have a random orbit $\{x_i\}_{i \in \mathbb{N}}$, of a random transformation $F : \Omega \times M \rightarrow \Omega \times M$, for which we know the starting point $x_0 = x \in M$, it is a natural question whenever or not we can compute the expected value $\mathbb{E}[\phi(x_1) | x_0 = x]$, where $\phi \in L^p(\eta)$, $p \in [1, \infty]$ is an **observable** that quantifies some properties of M . Since the iterate depends on an outcome $\omega \in \Omega$ which is distributed according to μ , this latter quantity can be formally computed in the following way

$$U\phi(x) = \int_M \phi(F_\omega(x)) d\mu(\omega) = \int_M \phi(F_{\omega_0}(x)) dp(\omega_0). \quad (1.1)$$

Moreover $U\phi(\cdot)$ defines a bounded, measurable function on M according to proposition 1.4, in fact if $\phi = \chi_B$, for some $B \in \mathcal{B}$, the integral

$$U\phi(x) = \mu(F^{-1}(\Omega \times B)_x), \quad (1.2)$$

is well defined.

By standard measure theory argument we can extend this to every bounded measurable function ϕ .

Definition 1.5. The operator U , as defined in 1.1, is called the **transfer operator** associated to F , which maps bounded (or nonnegative) measurable functions into bounded (or nonnegative) measurable functions. The **adjoint transfer operator** acts in the space of probability measures η on M and it is defined by

$$U^*\eta(B) = \int \eta(F^{-1}(\omega)(B))d\mu(\omega) = \int \eta(F^{-1}(\Omega \times B)^\omega)d\mu(\omega); \quad (1.3)$$

this integral is well defined by proposition 1.4 and $U^*\eta$ is a well-defined probability, by proposition 1.11.

The following lemma explains in which sense U and U^* are dual operators.

Lemma 1.6. *Let $\phi : M \rightarrow \mathbb{R}$ be a bounded or nonnegative measurable function. Then*

$$\int \phi d(U^*\eta) = \int U\phi d\eta.$$

To prove this lemma we consider the positive linear functionals defined by

$$f(\phi) = \int \phi d(U^*\eta), \quad g(\phi) = \int U\phi d\eta,$$

and we want to show that $f(\phi) = g(\phi)$ for every bounded or nonnegative measurable function $\phi : M \rightarrow \mathbb{R}$. The following lemma shows that if the condition above holds for every characteristic function $\phi = \chi_B$, $B \in \mathcal{B}$, then it holds for every function under consideration.

Lemma 1.7. *Let \mathcal{F} be either the space of bounded or nonnegative measurable functions on (M, \mathcal{B}) , and let \mathcal{G} be an ordered vector space. Let $f, g : \mathcal{F} \rightarrow \mathcal{G}$ be positive linear operators such that*

1. *if $\exists \phi \in \mathcal{F}$ such that $\phi \neq \phi^+$, then $f(1) = g(1)$*
2. *$\forall B \in \mathcal{B}$, $f(\chi_B) \leq g(\chi_B)$*
3. *if $\phi_n \uparrow \phi \in \mathcal{F}$ then $f(\phi_n) \uparrow f(\phi)$.*

Then

$$\forall \phi \in \mathcal{F}, \quad f(\phi) \leq g(\phi).$$

Proof. If ϕ is bounded, there exists $c > 0$ such that $\phi + c$ is nonnegative, hence we can assume ϕ nonnegative; moreover by linearity and positivity $f(\phi) \leq g(\phi)$ holds for all positive simple functions. Let $n \in \mathbb{N}$ and let ϕ be a nonnegative function, define

$$\phi_n(x) = \sum_{i=0}^{\infty} \frac{1}{2^n} \chi_{(i, i+1/2^n]} \circ \phi(x) = \sum_{i=0}^{\infty} \frac{1}{2^n} \chi_{\phi^{-1}((i, i+1/2^n])}(x),$$

which is simple functions because the sum contains $n2^n$ non-zero terms. Observe that $\phi_n \uparrow \phi$ because if $\phi(x) \in [i/2^n, (i+1)/2^n)$, for some $i \in \{0, \dots, n2^n - 1\}$, then $\phi_n(x) = i/2^n$, otherwise $\phi_n(x) = 0$. Since for every $n \in \mathbb{N}$, $f(\phi_n) \leq g(\phi)$ and $f(\phi_n) \uparrow f(\phi)$ we conclude that $f(\phi) \leq g(\phi)$. \square

proof of lemma 1.6. Let us prove that for every $\phi = \chi_B$, $B \in \mathcal{B}$, $f(\phi) = g(\phi)$; by the last lemma this will extend to all bounded or nonnegative measurable functions $\phi : M \rightarrow \mathbb{R}$. By proposition 1.4, and 1.2, 1.3

$$\begin{aligned} f(\phi) &= \int U\chi_B(x) d\eta = \int \mu(F^{-1}(\Omega \times B)_x) d\eta(x) = (\mu \times \eta)(F^{-1}(\Omega \times B)) \\ &= \int \eta(F^{-1}(\Omega \times B)^\omega) d\mu(\omega) = U^*\eta(B) = g(\phi). \end{aligned}$$

\square

1.1.2/ STATIONARY MEASURES

Definition 1.8. A probability measure η for M is called **stationary** for the random transformation F , if

$$U^*\eta = \eta,$$

i.e. it is a fixed point of the operator U^* .

In the deterministic set there is the concept of invariant measures:

Definition 1.9. Let (D, \mathcal{D}) be a measure space, and $T : D \rightarrow D$ a measurable function, then μ is an **invariant measure** for T if for every $E \in \mathcal{D}$, $\mu(T^{-1}(E)) = \mu(E)$.

Remark 1.10. If η is invariant for every F_ω , then η is stationary for F . The converse is not true: consider example 1.3, and set $d = 2$, $M = P\mathbb{R}^2$, $F_1([x]) = [A_1x]$ and $F_2([x]) = [A_2x]$, where A_1 and A_2 are invertible matrices that admit a decomposition in invariant sets, i.e.

$$R^3 = E_i^s \oplus E_i^u, \quad i = 1, 2$$

such that for all $x^s \in E_i^s$ and $x^u \in E_i^u$,

$$\|A_i x^s\| \leq \lambda_i \|x^s\|, \quad \|A_i^{-1} x^u\| \leq \mu_i \|x^u\|, \quad \lambda_i, \mu_i \leq 1,$$

and such that E_i^s and E_i^u are all distinct.

Let η be a measure which is both F_1 and F_2 invariant and let V be a neighborhood of $[E_1^u]$. Since $V \subset F_1^{-1}(V) \subset \dots \subset \cup_{k=0}^{\infty} F_1^{-k}(V) = M \setminus [E_1^s]$, the invariance with respect to F_1 implies

$$\eta(V) = \eta(\cup_{k=0}^{\infty} F_1^{-k}(V)) = \eta(M \setminus [E_1^s]).$$

This holds for every neighborhood V , therefore $\eta(M \setminus ([E_1^s] \cup [E_1^u])) = 0$. On the other side this holds also for $\eta(M \setminus ([E_2^s] \cup [E_2^u])) = 0$ and since we are assuming the eigenspaces to be distinct, this implies $\eta(M) = 0$, which is a contradiction.

The following proposition shows the link between invariant and stationary measures.

Proposition 1.11 ([53], lemma 2.1). *For every probability measure η on M , $F_*(\mu \times \eta) = \mu \times U^*(\eta)$. In particular, η is stationary if and only if $\mu \times \eta$ is invariant.*

Proof. Let $A \times B \subset \Omega \times M$ be a measurable set,

$$\begin{aligned} (\mu \times \eta)(F^{-1}(A \times B)) &= \int \int \chi_A \circ \sigma(\omega) \chi_B \circ F_\omega(x) d\eta(x) d\mu(\omega) \\ &= \int \chi_A \circ \sigma(\omega) \int \int \chi_B \circ F_\omega(x) d\eta(x) dp(\omega_0) d\mu(\sigma(\omega)) \\ &= \mu(A) \int \int \chi_B F_\omega(x) d\eta(x) dp(\omega_0) \\ &= \mu(A) \int \int \chi_B F_\omega(x) d\mu(x) d\eta(\omega) \\ &= \mu(A) U^* \eta(B). \end{aligned}$$

□

However not every invariant measure for F which projects to μ can be written as a product measure of the kind above; an explicit example can be found in [3], p.52, ex. 2.1.2.

1.2/ EXISTENCE AND REGULARITY OF STATIONARY DENSITIES

In this section we investigate conditions which guarantee the existence of stationary densities for a random transformation $F : \Omega \times M \rightarrow \Omega \times M$ that we assume continuous, i.e. each $F_\omega = F(\omega, \cdot)$ is continuous. The existence of these measures is guaranteed by the standard Krylov-Bogolyubov procedure, in fact we take advantage of the topological properties of U and U^* , to deduce (in abstract) the existence of stationary densities, via a suitable fixed point theorem.

Remark 1.12 ([19], pag.262). The space of regular probability measures on M , denoted by $\mathcal{P}(M)$, can be identified with a subspace of $C(M)^*$, the dual of the bounded continuous

real function on M ; under this identification, for all $\phi \in C(M)$, for all $\eta \in \mathcal{P}(M)$

$$\eta(\phi) = \int \phi d\eta.$$

In the following, we will consider the weak*-topology on $\mathcal{P}(M)$ as the subspace topology obtained when $C^*(M)$ is endowed with the weak*-topology.

Before to restate lemma 1.6 in this new taste, let us prove a variant of theorem 16.8 in [10]

Theorem 1.13. *Let O be an open subspace of a metric space, let μ be a probability measure on Ω and let $f(\cdot, x) : \Omega \rightarrow \mathbb{R}$ be a measurable function for each $x \in O$. Then:*

1. *If $f(\omega, \cdot)$ is continuous at x_0 for a.e. $\omega \in \Omega$ and there exists a neighborhood $O_{x_0} \ni x_0$ where $|f(\omega, \cdot)| \leq g(\omega)$ for some $g \in L^1(\mu)$; then $\int f(\cdot, x)d\mu$ is continuous at x_0 .*
2. *Suppose that $O \subset \mathbb{R}$, that for every $x_0 \in O$ there exists a full measure set Ω_{x_0} such that $f(\omega, \cdot)$ and $\partial_x f(\omega, \cdot)$ are continuous at x_0 and that there exists some neighborhood $O_{x_0} \ni x_0$ where $|\partial_x f(\omega, \cdot)| \leq g(\omega, x_0)$ for some $g(\cdot, x_0) \in L^1(\mu)$. Then*

$$\partial_x \int_{\Omega} f(\omega, \cdot) d\mu(\omega) = \int_{\Omega} \partial_x f(\omega, \cdot) d\mu(\omega),$$

and $\int_{\Omega} f(\omega, \cdot) d\mu(\omega) \in C^1(O)$.

Proof. For the first item, observe that

$$\lim_{x \rightarrow x_0} f(\omega, x) = f(\omega, x_0),$$

for a.e. $\omega \in \Omega$ and $|f(\omega, x)| \leq g(\omega)$, for all $x \in O_{x_0}$; we obtain the thesis by Lebesgue's dominated convergence theorem.

Let Ω_{x_0} and O_{x_0} as in the hypothesis if item 2. We have that

$$\lim_{x \rightarrow x_0} \frac{f(\omega, x) - f(\omega, x_0)}{x - x_0} = \partial_x f(\omega, x_0)$$

and that for every $\omega \in \Omega_{x_0}$ and for every $x \in O_{x_0}$

$$\left| \frac{f(\omega, x) - f(\omega, x_0)}{x - x_0} \right| \leq g(\omega, x_0).$$

Since O_{x_0} is a full measure set, we may take $\partial_x f(\omega, x_0) = 0$ where the partial derivative does not exists, and by this argument we can extend $\partial_x(\cdot, x_0)$ to the whole Ω . By applying

Lebesgue's dominated convergence theorem

$$\lim_{x \rightarrow 0} \int_{\Omega_{x_0}} \left| \frac{f(\omega, x) - f(\omega, x_0)}{x - x_0} - \partial_x f(\omega, x_0) \right| d\mu(\omega) = 0,$$

thus

$$\lim_{x \rightarrow 0} \frac{\int_{\Omega} f(\omega, x) d\mu(\omega) - \int_{\Omega} f(\omega, x_0) d\mu(\omega)}{x - x_0} = \int_{\Omega} \partial_x f(\omega, x_0) d\mu(\omega).$$

Moreover, for every $\omega \in \Omega_{x_0}$, $\lim_{x \rightarrow x_0} \partial_x f(\omega, x) = \partial_x f(\omega, x_0)$, and for all $x \in O_{x_0}$, $|\partial_x f(\omega, x)| \leq g(\omega, x_0)$. Hence we can conclude by applying item 1 to $\partial_x f(\cdot, x)$.

□

Let us state the extension of this theorem to function defined on \mathbb{R}^n , recall that we use the multindex notation.

Corollary 1.14. *Let $f(\cdot, x) : \Omega \rightarrow \mathbb{R}$ be a measurable function for each $x \in O \subset \mathbb{R}^n$, where O is open, and let μ be a probability measure on Ω . Suppose that for every $x_0 \in O$ there exists a full measure set Ω_{x_0} such that $\partial^\alpha f(\omega, \cdot)$ is a continuous function at x_0 for every $\omega \in \Omega_{x_0}$ and $\alpha \in \mathbb{N}_0^n$ is such that $|\alpha| \leq k$. Suppose that there exists a neighborhood O_{x_0} of x_0 such that for every $\alpha \in \mathbb{N}_0^n$, with $|\alpha| \leq k$, $|\partial^\alpha f(\omega, \cdot)| \leq g_\alpha(\omega, x_0)$ for some $g_\alpha(\cdot, x_0) \in L^1(\mu)$. Then $\int_{\Omega} f(\omega, \cdot) d\mu(\omega) \in C^k(O)$ and $\partial^\alpha \int_{\Omega} f(\omega, \cdot) d\mu(\omega) = \int_{\Omega} \partial^\alpha f(\omega, \cdot) d\mu(\omega)$ for $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leq k$.*

Proposition 1.15. *The operator U maps $C(M)$ to itself and U^* is weak* continuous.*

Proof. For the first part observe that we can apply item 1 of theorem 1.13 to the continuous functions $f(\omega, \cdot) := \phi \circ F_\omega$. For the second part let $\eta \in \mathcal{P}(M)$ and consider the following basic neighborhood of $U^*\eta$,

$$V = \{\xi \in \mathcal{P}(M) : \forall \phi \in \Phi (|\xi(\phi) - U^*\eta(\phi)| \leq \epsilon)\},$$

where Φ is a finite subset of $C(M)$ and $\epsilon > 0$. Consider now

$$W = \{\xi \in \mathcal{P}(M) : \forall \psi \in U(\Phi) (|\xi(\psi) - \eta(\psi)| \leq \epsilon)\},$$

which is a basic neighborhood of η , because $U(\Phi)$ is finite subset of $C(M)$ and $U^*(W) \subset V$, because $U^*\xi(\phi) = \xi(U\phi)$.

□

Corollary 1.16. *If M is compact, then there exists a stationary measure on M .*

Proof. M is compact, hence the set of the signed Borel measures can be identify via Riesz representation theorem with $C^0(M)^*$, which it is, if endowed with the weak*-topology, locally convex ([3], pag 31) Under this identification

$$\mathcal{P}(M) = \cap_{\phi \in C^0(M)_+} \{\eta \in C^0(M)^* : \eta(\phi) \geq 0\} \cap \{\eta \in C^0(M)^* : \eta(1) = 1\} \subset B(0, 1),$$

is weak*-compact, since $B(0, 1)$ is compact by Banach-Alaoglu's theorem. Trivially, it is also convex, therefore U^* admits a fixed point by Schauder-Tychonoff fixed point theorem [19], pag. 456. \square

A somewhat standard procedure to obtain stationary measures, is the so-called Krylov-Bogolyubov procedure ([3], pag 29).

Theorem 1.17. *Let $n \in \mathbb{N}$, let ν be a probability measure on M , and define*

$$\nu_n = \frac{1}{n} \sum_{j=0}^{n-1} U^{*j} \nu.$$

Then every limit point of $\{\nu_n\}_{n \in \mathbb{N}}$ is weak-invariant, and ν invariant arise in this way.*

Proof. Assume without loss of generality that $\{\nu_n\}$ weak*-converge to $\bar{\nu}$; since U^* is weak*-continuous, also $\{U^* \nu_n\}$ weak*-converge to $U^* \bar{\nu}$. Moreover, if $\phi \in C(M)$ then

$$|U^* \nu_n(\phi) - \nu_n(\phi)| = \frac{1}{n} |U^{*n} \nu(\phi) - \nu(\phi)| \leq \frac{2}{n} \|\phi\|,$$

therefore $U^* \bar{\nu} = \bar{\nu}$.

For the last part, observe that if ν is invariant, the (Birkhoff) averages ν_n are equal to ν . \square

ERGODIC THEOREMS

2.1/ BIRKHOFF ERGODIC THEOREM

In the last chapter we described how the deterministic concept of invariant measure can be translated into the random context; another example which is of main importance, is the **ergodicity**, which naturally extend to random dynamical systems.

Let us recall that our setting is made by a measurable space (M, \mathcal{B}) and a random transformation $F : \Omega \times M \rightarrow \Omega \times M$.

Definition 2.1. Let η be a probability measure on M

- a set B is called η -stationary if $U_{\chi_B} = \chi_B$.
- a function $\phi \in L^1(\eta)$, $p \in [1, +\infty]$ is called η -stationary, if $U\phi = \phi$.

Definition 2.2. A probability measure η on M is said to be **ergodic** for a random transformation F if every η -stationary set has either full or null measure.

The following is a technical lemma, which shows that in our setting monotone sequences converge. It will be used in lemma 2.14 to show that the invariant sets are a σ -algebra on M .

Lemma 2.3. *If $\phi_n \uparrow \phi \in L^p(\eta)$, then $U\phi_n \uparrow U\phi$.*

Proof. Let $n \in \mathbb{N}$ and let

$$E_n = \{x \in M : |U\phi_n(x) - U\phi(x)| > \epsilon\}.$$

$\{U\phi_n\}_{n \in \mathbb{N}}$ is a monotone sequence, by the positivity of U , thus $E_1 \supset E_2 \dots$; we claim that $\inf_{n \in \mathbb{N}} \eta(E_n) = 0$. If $p = \infty$, $\eta(E_n) = 0$ whenever $\|\phi_n - \phi\| \leq \epsilon/\|U\|$. If $p < \infty$ we use Markov inequality. In fact E_n may be written as

$$E_n = \{x \in M | |U\phi_n(x) - U\phi(x)|^p > \epsilon^p\},$$

$$\eta(E_n) \leq \epsilon^{-p} \int_M |U\phi_n(x) - U\phi(x)|^p d\eta(x) \leq \frac{1}{\epsilon^p} \|U\phi_n - U\phi\|^p \leq \frac{\|U\|^p}{\epsilon^p} \|\phi_n - \phi\|^p$$

□

A central role is played by **positive, contractive** operators, i.e. operators P such that

- If $f \geq 0$ then $Pf \geq 0$
- $\|P\| \leq 1$

Theorem 2.4 ([20], pag. 8). *Let η a probability measure on M and let U be a linear, positive, contractive operator on $L^p(\eta)$, $p \in [1, +\infty]$. The set S of the stationary sets $B \subset M$, is the smallest σ -ring on M such that the stationary functions, which are bounded or non negative, are measurable.*

We will now deal with the main feature of this section, (a variant of) the famous **Birkhoff Ergodic Theorem**. To formalize the idea that if η is stationary and ergodic for some random transformation F , then the **space average** is equal to the **time average** almost everywhere, we may want to introduce the **Birkhoff sums**:

Definition 2.5. Let T be a linear, positive, contractive operator on $L^1(\eta)$, then the n -th Birkhoff sum with respect to T , is the positive, linear, contractive operator $S_n : L^1(\eta) \rightarrow L^1(\eta)$ such that

$$S_n(\phi) = \sum_{i=0}^{n-1} T^i \phi = \phi + T\phi + \dots + T^{n-1}\phi. \quad (2.1)$$

We define also the positive, linear contractive operator $A_n : L^1(\eta) \rightarrow L^1(\eta)$, called the n -th Birkhoff average, defined by

$$A_n \phi = \frac{S_n \phi}{n}.$$

Our main result in this section is the following theorem, which shows the pointwise convergence of the Birkhoff averages:

Theorem 2.6. *Let T be a linear, positive, contractive operator on $L^1(\eta)$. If there exists $\phi_0 \in L^1(\eta)$ such that $\phi_0 > 0$ and the sequence $\{A_n \phi_0\}$ admits a weakly convergent subsequence, then for every $\phi \in L^1(\eta)$ the limit*

$$\bar{\phi}(x) = \lim_{n \rightarrow +\infty} A_n \phi(x) = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} T^i \phi(x) \quad (2.2)$$

exists for η -a.e. $x \in M$, and it defines a function $\bar{\phi} \in L^1(\eta)$ such that $T\bar{\phi} = \bar{\phi}$.

Furthermore, if C is the σ -algebra of the invariant sets, then for every $B \in C$, $\bar{\phi}$ satisfies

$$\int_B \bar{\phi} d\eta = \int_B H_C \phi d\eta, \quad (2.3)$$

where

$$H_C = \phi(x) = \chi_C \sum_{k=0}^{\infty} (TI_D)^k \phi(x), \quad \text{with } I_D \phi(x) = \chi_D(x) \phi(x). \quad (2.4)$$

Remark 2.7. In some particular case the link between time and space averages (which may be not clear in the last statement), is glaring; for instance, if T is the operator P_η in proposition 3.11, then

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{i=0}^{n-1} T^i \phi(x) = \int_M \phi d\eta,$$

for η -a.e. $x \in M$.

The rest of this section is devoted to the proof of theorem 2.6, which requires two general statements, namely **the Hopf maximal ergodic lemma** and the **mean ergodic theorem**.

2.1.1/ HOPF ERGODIC MAXIMAL LEMMA AND THE σ -ALGEBRA OF THE INVARIANT SETS

The Hopf lemma gives us a first way to investigate the Birkhoff averages, in particular it is used to decompose M into **conservative** and **dissipative** parts; for a proof see [20], pag. 9.

Lemma 2.8 (Hopf maximal ergodic lemma). *Let $T : L^1(\eta) \rightarrow L^1(\eta)$, be a linear, positive, contractive operator and let $\psi \in L^1(\eta)$. Then*

$$\int_E \psi d\eta \geq 0, \quad \text{where } E = \{x \in M : \sup_{n \in \mathbb{N}} S_n \psi(x) > 0\}. \quad (2.5)$$

Definition 2.9 ([32], pag. 32). Let $\phi_0 \in L^1(\eta)$, $\phi_0 > 0$; the conservative and dissipative parts of M are defined respectively as

$$C = \{x \in M : \sum_{j=0}^{\infty} T^j \phi_0(x) = +\infty\}, \quad (2.6)$$

$$D = \{x \in M : \sum_{j=0}^{\infty} T^j \phi_0(x) < +\infty\}. \quad (2.7)$$

Lemma 2.10. *The definition above is well posed, in particular:*

1. C, D do not depend of the choice of $\phi_0 > 0$ and $T\chi_C \leq \chi_C, \chi_D \leq T\chi_D$
2. the Banach operator adjoint of T, T^* , is positive and $T^*1 \leq 1$.

Proof. Let $\phi \in L_+^1(\eta)$ and let

$$C' = \{x \in C : 0 < \sum_{j=0}^{\infty} T^j \phi(x) < +\infty\}$$

$$D' = \{x \in D : \sum_{j=0}^{\infty} T^j \phi(x) = +\infty\}.$$

Let $j \in \mathbb{N}$ such that

$$\eta(\{x \in C' : T^j \phi(x) > 0\}) > 2^{-j-1} \eta(C')$$

If we apply 2.8 to $\psi_C = \phi_0 - aT^j \phi_0$ and $\psi_D = \phi_0 - a\phi_0, a > 0$, then for $I \in \{C, D\}$ we have

$$\begin{aligned} I &\subset E_I := \{x \in M : \sup_{\mathbb{N}} S_n \psi_I(x) > 0\} \\ 0 &\leq \int_{E_C} \psi_C d\eta \leq \int_M \phi_0 d\eta - a \int_{C'} T^j \phi d\eta \\ 0 &\leq \int_{E_D} \psi_D d\eta \leq \int_M \phi d\eta - a \int_{D'} \phi_0 d\eta. \end{aligned}$$

By arbitrarily choosing $a > 0$, we conclude that $\eta(C') = \eta(D') = 0$. Hence

$$C \subset \{x \in M : \sum_{j=0}^{\infty} T^j \phi(x) = 0 \quad \text{or} \quad \sum_{j=0}^{\infty} T^j \phi(x) = +\infty\} \quad (2.8)$$

$$D \subset \{x \in M : \sum_{j=0}^{\infty} T^j \phi(x) < +\infty\}. \quad (2.9)$$

In particular C and D are independent from the choice of ϕ_0 ; moreover if $x \in C$ then $\sum_{j=0}^{\infty} T^j \chi_C(x) = +\infty$ and also $\sum_{j=0}^{\infty} T^j(T\chi_C(x)) = +\infty$, which implies the second part of item 1.

For item 2, fix $\epsilon > 0$ and let $B \in \mathcal{B}$ defined as $B := \{x \in M : T^*1 > 1 + \epsilon\}$; then

$$(1 + \epsilon)\eta(B) \leq \int_M (T^*1)\chi_B d\eta = \int_M T\chi_B d\eta \leq \|T\chi_B\| \leq \|chi_B\| = \eta(B),$$

thus $\eta(B) = 0$ and $T^*1 \leq 1$.

□

Definition 2.11. A set $B \in \mathcal{B}$ is called **invariant** if $T^*\chi_B = \chi_B$ on C .

The following from [32], is an easy criterion to verify if a set is invariant.

Lemma 2.12. Let $\phi \in L^\infty(\eta)$ (or $\phi \in L^1_+(\eta)$) such that $T^*\phi \leq \phi$ on C (or $T\phi \leq \phi$ on C); then $T^*\phi = \phi$ on C (or $T\phi = \phi$ on C).

Proof. Let $\phi \in L^\infty(\eta)$ such that $T^*\phi \leq \phi$ on C , let $\psi = \phi - T^*\phi$ and $\phi_0 = 1$. We observe that

$$\psi + \dots + T^{*(n-1)}\psi = \phi - T^{*n}\phi,$$

hence

$$\int_M \psi(\phi_0 + \dots + T^{n-1}\phi_0) d\eta = \int_M (\phi - T^{*n}\phi)\phi_0 d\eta \leq 2\|\phi\|_{L^\infty}\|\phi_0\|_{L^1}.$$

Trivially $\sum_{i=0}^{\infty} T^{i-1}\phi_0 = +\infty$ on C thus $\psi = 0$ on C , which proves the first part of the theorem. For the second part let $\phi \in L^1_+(\eta)$ such that $T\phi \leq \phi$ and suppose $T\phi \neq \phi$ on C . Then there exists $\epsilon > 0$ such that $F = \{x \in C : \phi(x) - T\phi(x) > \epsilon\}$ has a positive measure. Let us prove that $\sum_{j=0}^{\infty} T^{*j}\chi_F = +\infty$ on F : let $c > 0$ and let $F' = \{x \in F : \sum_{j=0}^{\infty} T^{*j}\chi_F < c\}$, then

$$\int_M \left(\sum_{j=0}^{\infty} T^j\chi_{F'}\right)\chi_F d\eta = \int_M \chi_{F'} \left(\sum_{j=0}^{\infty} T^{*j}\chi_F\right) d\eta < +\infty,$$

thus $\eta(F') = 0$ and since $F' \subset F \subset C$ equation 2.8 implies $(\sum_{j=0}^{\infty} T^j\chi_{F'})\chi_F = +\infty$ on F' . But this implies that on F

$$(\phi - T\phi)\left(\sum_{j=0}^{\infty} T^{*j}\chi_F\right) = +\infty,$$

which implies a contradiction

$$\int_M (\phi - T\phi)\left(\sum_{j=0}^{\infty} T^{*j}\chi_F\right) d\eta = \int_M \left(\sum_{j=0}^{\infty} T^j(\phi - T\phi)\right)\chi_F \leq \int_M \phi\chi_F d\eta < +\infty.$$

□

Remark 2.13. If M is an invariant set, using that $T^*1 \leq 1$, we may deduce

$$\chi_C \leq T^* \chi_C \quad \text{and} \quad T^* \chi_D \leq \chi_D;$$

in fact $T^* \chi_C \leq T^* 1 \leq 1$ on C , which implies that $T^* \chi_C = 1$ on C , i.e. $T^* \chi_C \leq \chi_C$. Moreover $T^* \chi_D = T^*(1 - \chi_C) \leq 1 - \chi_C = \chi_D$.

Lemma 2.14. *The set of the invariant sets is a σ -algebra \mathcal{C} on M and every $\phi \in L^\infty(\eta)$ such that $T^* \phi = \phi$ on C , is C -measurable.*

Proof. We seek a probability η_C on (M, \mathcal{B}) such that $T^* \phi = \phi$ on B if and only if $T^* \phi = \phi$ η_C -a.e. The case $\eta(C) = 0$ is trivial, since we can set $\eta = \eta_C$ and $C = \mathcal{B}$; so let $\eta(C) > 0$ and define for every $B \in \mathcal{B}$

$$\eta_C(B) = \frac{\eta(B \cap C)}{\eta(C)}.$$

Let $B_1, B_2 \in \mathcal{B}$ such that $\chi_{B_1} = \chi_{B_2}$ on C , observe that $\chi_C T^* \chi_D = 0$ by remark 2.13, then

$$\chi_C(T^* \chi_{B_1}) = \chi_C(T^* \chi_{B_1 \cap C} + T^* \chi_{B_1 \cap D}) = \chi_C T^* \chi_{B_1 \cap C} = \chi_C T^* \chi_{B_2}.$$

By lemma 2.3 it follows that if $\phi = \psi$ on C then $T^* \phi = T^* \psi$ on C . Hence we can regard at T^* as an operator defined on $L^1(\eta_C)$, on which it is still linear, positive and contractive. Moreover $T^* 1 = 1$ η_C -a.e., thus C is a σ -algebra for which every invariant $\phi \in L^\infty(\eta_C)$ is measurable. \square

2.1.2/ CONVERGENCE OF BIRKHOFF AVERAGES

The last ingredient of the Birkhoff theorem's proof it is the famous **Mean ergodic theorem** (for a proof refer to [41], th. 1.1).

Theorem 2.15 (Mean ergodic theorem). *Let T be a bounded operator on a Banach space V such that the Birkhoff averages are uniformly bounded (as operators). Then for any $v \in V$ such that*

$$\lim_{n \rightarrow +\infty} n^{-1} T^{n-1} v = 0,$$

and for any $w \in V$, the following are equivalent:

1. $T w = w$ and $w \in \text{cl}\{T^{n-1} v : n \in \mathbb{N}\}$
2. $\{A_n v\}_{n \in \mathbb{N}}$ strongly converges to w
3. $\{A_n v\}_{n \in \mathbb{N}}$ weakly converges to w

4. $\{A_n v\}_{n \in \mathbb{N}}$ admit a subsequences that converges to w .

Proof of theorem 2.6. We observe that the assumptions of the mean ergodic theorem are verified. In particular, by contraction property $\|T^n\| \leq 1$; furthermore for any $\phi \in L^1(\eta)$ the sequence $\{A_n \phi\}_{n \in \mathbb{N}} = \{S_n \phi / n\}_{n \in \mathbb{N}}$ has a weakly convergent subsequence, which is equivalent to show ([19], cor. 11) that uniformly on n the following holds

$$\lim_{\eta(E) \rightarrow 0} \int_E \frac{S_n(\phi)}{n} d\eta = 0.$$

Fix $\epsilon > 0$ and let ϵ_0 such that if $\eta(E) < \epsilon_0$ then $\int_E |\phi(\cdot)| d\eta < \epsilon$. Let $t > 0$ such that $\eta(B_t) < \min\{\epsilon, \epsilon_0\}$, where $B_t = \{x \in M : t\phi_0(x) < |\phi(x)|\}$.

Let ϕ_0 such that $\{S_n \phi_0 / n\}$ has a weakly convergent subsequence (such ϕ_0 exists by hypothesis); then there exists $\delta > 0$ such that if $\eta(E) < \delta$ then $\int_E S_n(\phi_0) / n d\eta < \epsilon / t$.

Thus if $\eta(E) < \delta$, then

$$\begin{aligned} \left| \int_E \frac{S_n \phi}{n} d\eta \right| &= \left| \int_{E \setminus B_t} \frac{S_n(\phi + t\phi_0)}{n} d\eta + \int_{B_t} \frac{S_n(\phi + t\phi_0)}{n} d\eta - \int_E \frac{S_n(t\phi_0)}{n} d\eta \right| \\ &= \left| \int_{E \setminus B_t} \frac{S_n(\phi + t\phi_0)^+}{n} d\eta + \int_{B_t} \frac{S_n(\phi)}{n} d\eta - \int_{E \setminus B_t} \frac{S_n(t\phi_0)}{n} d\eta \right| \\ &\leq \int_{E \setminus B_t} \frac{S_n(\phi + t\phi_0)^+}{n} d\eta + \int_{B_t} \frac{S_n|\phi(\cdot)|}{n} d\eta + t \int_{E \setminus B_t} \frac{S_n(\phi_0)}{n} d\eta \\ &\leq \int_E |\phi(\cdot)| d\eta + 2t \int_{E \setminus B_t} \phi_0 d\eta < 4\epsilon. \end{aligned} \tag{2.10}$$

Hence the assumptions of the mean ergodic theorem are verified and we may apply it to obtain a function $\bar{\phi} \in L^1(\eta)$, such that $T\bar{\phi} = \bar{\phi}$ and

$$\lim_{n \rightarrow +\infty} \left\| \frac{S_n \phi}{n} - \bar{\phi} \right\|_{L^1(\eta)} = 0.$$

We want to prove that

$$\lim_{n \rightarrow +\infty} \frac{S_n \phi(x)}{n} = \bar{\phi}(x), \quad \text{for } \eta - \text{a.e. } x \in M. \tag{2.11}$$

It is sufficient to show that for any given $\epsilon > 0$, for η -a.e. $x \in M$

$$\limsup_{n \rightarrow +\infty} \frac{S_n \phi(x)}{n} \leq \bar{\phi}(x) + \epsilon, \tag{2.12}$$

in fact the latter applied to $-\phi$ gives the opposite bound with the liminf. The set of the

points $x \in M$ such that 2.12 does not hold is contained in

$$G = \{x \in M : \sup_{n \in \mathbb{N}} S_n(\phi(x) - \bar{\phi}(x)) = +\infty\};$$

we claim that $\eta(G) = 0$. Let $\psi = \phi - \bar{\phi} - a$, with $a > 0$. Using the Hopf maximal ergodic lemma 2.5

$$\int_E (\phi - \bar{\phi}) d\eta \geq a\eta(E), \quad \text{where} \quad E = \{x \in M : \sup_{n \in \mathbb{N}} S_n\psi(x) > 0\}.$$

We observe that $E \subset G$ and thus

$$\|\phi\| + \|\bar{\phi}\| \geq \int_E (\phi - \bar{\phi}) d\eta \geq a\eta(E) \geq a\eta(B),$$

hence $\eta(G) = 0$ by the arbitrariness of a . We now turn to the second part of the theorem (2.3). Observe that $S_n\bar{\phi}$ is bounded on D by 2.9, which implies that $\bar{\phi} = 0$ on D . Since

$$\chi_C T^k \phi = \chi_C T^k (\chi_C \phi) = \chi_C T^{k-1} (T I_D) \phi = \dots = \chi_C \sum_{j=0}^k T^{k-j} (\chi_C (T I_D)^j \phi),$$

then for any $B \in \mathcal{C}$

$$\begin{aligned} \int_B \bar{\phi} d\eta &= \int_B \chi_C \bar{\phi} d\eta = \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_B \chi_C T^k \phi d\eta \\ &= \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_B \chi_C \sum_{j=0}^k T^{k-j} (\chi_C (T I_D)^j \phi) d\eta \\ &= \lim_{n \rightarrow +\infty} \frac{1}{n} \int_M T^{*(k-j)} (\chi_B \chi_C) \sum_{j=0}^k \chi_C (T I_D)^j \phi d\eta \quad (2.13) \\ &= \lim_{n \in \mathbb{N}} \frac{1}{n} \int_{<} \chi_B \chi_C \sum_{j=0}^k \chi_C (T I_D)^j \phi d\eta \\ &= \int_B H_C \phi d\eta. \end{aligned}$$

□

We will end this section showing that since $S_n 1 = n$, the η -a.e. convergence of $A_n \phi$ is a particular case of the following theorem

Theorem 2.16 (Chacon-Ornstein [20], pag. 26). *Let $\phi \in L^1(\eta)$, $\phi_0 \in L^1_+(\eta)$, and define*

$$Q_n(x) = \frac{S_n \phi(x)}{S_n \phi_0(x)}.$$

Then Q_n converges to a finite limit η -a.e. on the set

$$B = \{x \in M : \sum_{j=0}^{\infty} S_n \phi_0(x) > 0\}.$$

We first need the following lemma

Lemma 2.17. *Let $\phi, \phi_0 \in L_+^1(\eta)$, and Q_n as defined above, then*

$$\sup_{n \in \mathbb{N}} |Q_n(\cdot)| < +\infty$$

η -a.e. on $\{x \in M : \phi_0(x) > 0\}$ /

Proof. Let $a > 0$, and let

$$A := \{x \in M : \phi_0 > 0, \sup_{n \in \mathbb{N}} |Q_n(x)| = +\infty\}$$

$$E := \{x \in M : \sup_{n \in \mathbb{N}} S_n(\phi - a\phi_0)(x) > 0\}.$$

Observe that $A \subset E$; Hopf maximal ergodic lemma (2.5) applied to $\psi = \phi - a\phi_0$ implies

$$n\|\phi\| \geq \int_E S_n \phi(x) d\eta \geq a \int_E S_n \phi_0(x) d\eta \geq a \int_A S_n \phi_0(x) d\eta.$$

Since a is arbitrary, it must hold

$$\int_A S_n \phi_0(x) d\eta = 0.$$

But $S_n \phi_0(x) > 0$ on A , hence $\eta(A) = 0$. □

proof of theorem 2.16. Without loss of generality we may assume ϕ non-negative, because we can always split

$$Q_n(x) = \frac{S_n \phi^+(x)}{S_n \phi_0(x)} - \frac{S_n \phi^-(x)}{S_n \phi_0(x)}.$$

By triangle inequality,

$$\left| \frac{S_n \phi(x)}{S_n \psi(x)} \bar{\psi}(x) - \bar{\phi}(x) \right| \leq \left| \frac{S_n \phi(x)}{S_n \psi(x)} (\bar{\psi}(x) - \frac{S_n \psi(x)}{n}) \right| + \left| \frac{S_n \phi(x)}{n} - \bar{\phi}(x) \right|.$$

The right-hand side converges to 0 for η -a.e. $x \in B$ by the latter lemma and Birkhoff ergodic theorem, hence for η -a.e. $x \in B$

$$\lim_{n \rightarrow +\infty} \frac{S_n \phi(x)}{S_n \psi(x)} \bar{\psi}(x) = \bar{\phi}(x).$$

Because $\bar{\psi}(x)$ is positive on B , we have

$$\lim_{n \rightarrow +\infty} \frac{S_n \phi(x)}{S_n \psi(x)} = \frac{\bar{\phi}(x)}{\bar{\psi}(x)} < +\infty.$$

□

So far from now we deal with several properties of linear positive contractive operators, in particular the Birkhoff Ergodic Theorem; in fact a non singular random transformation induced an operator of this kind, hence we can study some of the properties of a random transformation, by taking a look to the induced operator. Since the latter is a linear positive contractive operator, we can apply the results described in this chapter and in the next one we will see how to do it.

PERRON-FROBENIUS OPERATOR

In the last chapter we deal with the pointwise convergence of the Birkhoff averages of an operator U ; turns out that a non singular transformation F induced an operator (the so-called **Perron-Frobenius operator**) which satisfies the hypothesis of the Birkhoff theorem. This chapter is devoted to the study of this class of operators and then to the study of criteria for the existence of stationary densities for this particular class.

3.1/ PERRON-FROBENIUS OPERATOR

Let λ be a fixed Borel regular probability on M .

Definition 3.1. A random transformation $F : \Omega \times M \rightarrow \Omega \times M$ is **non singular** with respect to λ if for every probability measure $q \ll \mu \times \lambda$ then $F^*q \ll \mu \times \lambda$.

Let us recall a sufficient condition for a transformation to be nonsingular.

Lemma 3.2. *If $(F_\omega)_*\lambda \ll \lambda$ for μ -a.e. $\omega \in S$, then F is non singular.*

Proof. The sets of the form $A \times B$ generate the σ -algebra of $\Omega \times M$, and the null sets form a σ -algebra. Let $A \times B \subset \Omega \times M$ be a $\mu \times \lambda$ -null set, then either $\mu(\sigma^{-1}(A)) = \mu(A) = 0$ or $\lambda(F_\omega^{-1}(B)) = \lambda(B) = 0$ for a.e. $\omega \in \Omega$; thus

$$F_*(\mu \times \lambda)(A \times B) = \int_{\sigma^{-1}(A)} \lambda(F_\omega^{-1}(B))d\mu(\omega) = 0.$$

□

If F is a non singular transformation (from now on, we will omit **with respect to λ**), and $\eta \ll \lambda$ then

$$\pi_M^*F^*(\mu \times \eta) \ll \pi_M^*(\mu \times \lambda) = \lambda,$$

and by proposition 1.11

$$U^*\eta = \pi_M^*F^*(\mu \times \eta).$$

Thus U^* may act on a space of densities instead of a space of measures; before to formalize the last statement, let us show with the following example that if F is non singular, U can be viewed as an operator acting on $L^\infty(M)$, furthermore if η is stationary for F , U can be defined on $L^1(\eta)$.

Example ^{3.3}. Let F a non singular transformation and let ϕ and ψ two bounded measurable functions such that $A = \{x \in M : \psi(x) \neq \phi(x)\}$ has null measure; then

$$\begin{aligned} \int |U\phi - U\psi|d\lambda &\leq \int_M \int_M |\phi(F_\omega x) - \psi(F_\omega x)|d\mu(\omega)d\lambda(x) \\ &\leq (\|\phi\|_\infty + \|\psi\|_\infty)F^*(\mu \times \lambda)(\Omega \times A) = 0. \end{aligned} \quad (3.1)$$

Thus also the set $\{x \in M : U\phi(x) \neq U\psi(x)\}$ has null measure and U is well defined as an operator $U : L^\infty(M) \rightarrow L^\infty(M)$. If η is stationary for F , then for all $\phi \in L^\infty(\eta)$, $\|U\phi\|_{L^1(\eta)} \leq \|\phi\|_{L^1(\eta)}$ and $\overline{L^\infty(\eta)} = L^1(\eta)$. Thus we may consider $U : L^1(\eta) \rightarrow L^1(\eta)$.

Remark ^{3.4}. If $\phi \in L^1(M)$, we denote by $\phi\lambda$ the Borel measure with density ϕ , i.e. the unique measure such that

$$\forall \psi \in C^0(M) : \int \psi d(\phi\lambda) = \int \psi\phi d\lambda.$$

Definition ^{3.5}. A non singular random transformation F induced an operator $L : L^1(M) \rightarrow L^1(M)$ defined by

$$L\phi = \frac{d(U^*(\phi\lambda))}{d\lambda}.$$

L is called the **Perron Frobenius operator** associated to F or induced by F . Furthermore, for every $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$,

$$\int_M (L\phi)\psi d\lambda = \int_M \phi(U\psi) d\lambda,$$

thus identifying $L^\infty(M)$ as the dual space $L^1(M)$, we may write $L^* = U$.

We remark that the definition above naturally extends the definition in the deterministic context, via identification of a deterministic map $T : M \rightarrow M$ with a random map $F : \Omega \times M \rightarrow \Omega \times M$, where Ω is the trivial probability space. The following proposition introduces some standard properties of the Perron-Frobenius operator

Proposition ^{3.6}. Let F be a random transformation and let $L : L^1(M) \rightarrow L^1(M)$ be the induced Perron Frobenius operator, the following hold:

1. L is linear
2. L is positive, i.e. $L\phi \geq 0$ if $\phi \geq 0$

3. for every $\phi \in L^1(M)$, $\int_M L\phi d\lambda = \int_M \phi d\lambda$
4. F_ω , the random transformation on $\Omega \times M$ defined by $(\omega, x) \mapsto F_M^n(x)$, induced a Perron-Frobenius operator L_M such that $L_M = L^n$
5. if M is compact, then for any $\phi \in L^1(M)$, $\|L\phi\|_{L^1} \leq \|\phi\|_{L^1}$.

Proof. 1. L can be written as the composition of the linear operators, $\phi \mapsto \phi\lambda$, U^* and $\eta \mapsto d\eta/dm$; the image of the first is the subspace of absolutely continuous measures, the second preserves this subspace by nonsingularity and the third is defined by absolutely continuous measures.

2. We claim that if $\phi \geq 0$ then $U^*(\phi\lambda)$ is a non negative measure, because the Radon-Nikodym derivative of an unsigned measure is unsigned; this is true because U is positive, and thus U^* preserves the subspace of nonnegative measures.
3. We observe that $U1 = 1$ and

$$\int_M L\phi d\lambda = \int_M (U1)\phi d\lambda = \int_M \phi d\lambda.$$

4.
$$U^n\phi(x) = \int_M \dots \int_M \phi \circ F_{\omega_N} \circ \dots \circ F_{\omega_1}(x) d\mu(\omega_1) \dots d\mu(\omega_n).$$

5. For any $\phi \in L^\infty(M)$ and for any $B \in \mathcal{B}$

$$\int_M \chi_B L\phi d\eta = \int_M U(\chi_B)\phi d\eta \leq \int_M |\phi| d\eta = \|\phi\|_{L^\infty}.$$

□

Definition 3.7. Let (M, \mathcal{B}, η) be a measure space; an operator $P : L^1(\eta) \rightarrow L^1(\eta)$ which satisfies items 1, 2, 3 of proposition 3.6, is called **Markov operator**.

Proposition 3.8. Let (M, \mathcal{B}, η) be a measure space, let $\phi \in L^1(\eta)$ and let P be a Markov operator, then

- $(P\phi)^+(x) \leq P\phi^+(x)$, for η -a.e. $x \in M$
- $(P\phi)^-(x) \leq P\phi^-(x)$, for η -a.e. $x \in M$
- $|(P\phi)(x)| \leq P|\phi(\cdot)|(x)$ for η -a.e. $x \in M$
- P is contractive, i.e. $\|P\phi\|_{L^1} \leq \|\phi\|_{L^1}$.

Proof. The positivity of P implies that $\phi^+ - \phi \geq 0$, then $P(\phi^+ - \phi) \geq 0$ and $P\phi^+ \geq P\phi$; observing that $(P\phi)^+ \geq 0$, we conclude that $(P\phi)^+ \leq P\phi^+$. Symmetrically, $(P\phi)^- \leq P\phi^-$. By linearity

$$|(P\phi)(\cdot)| = (P\phi)^+ + (P\phi)^- \leq P\phi^+ + P\phi^- = P(\phi^+ + \phi^-) = P|\phi(\cdot)|.$$

As last step, since for all $\phi \in L^1(\eta)$, $\int_M P\phi d\eta = \int_M \phi d\eta$,

$$\|P\phi\| = \int_M |P\phi(\cdot)| d\eta \leq \int_M P|\phi(\cdot)| d\eta = \int_M |\phi(\cdot)| d\eta = \|\phi\|.$$

□

If η is equal to the reference measure, i.e. $\eta = \lambda$, and we denote by \mathcal{D} the set of positive densities $h \in L^1(M)$ such that $\|h\|_1 = 1$, then for any Markov operator P we have $P(\mathcal{D}) \subset \mathcal{D}$.

Definition 3.9. Let $h \in \mathcal{D}$ such that $L(h\lambda) = h\lambda$, then h is called a **stationary density** for F .

Stationary densities play a central role, because every fixed point of L can be obtained from them.

Proposition 3.10. Let P a Markov operator and let $\phi \in L^1(\eta)$ such that $P\phi = \phi$, then $P\phi^+ = \phi^+$ and $P\phi^- = \phi^-$.

Proof. By hypothesis we have

$$\begin{aligned} \phi^+ &= (P\phi)^+ \leq P\phi^+ \quad \text{and} \quad \phi^- = (P\phi)^- \leq P\phi^-; \\ \int (P\phi^+ - \phi^+) d\eta + \int (P\phi^- - \phi^-) d\eta &= \int P(\phi^+ + \phi^-) - (\phi^+ + \phi^-) d\eta \\ &+ \int P|\phi(\cdot)| - |\phi(\cdot)| d\eta = \int P|\phi(\cdot)| d\eta - \int |\phi(\cdot)| d\eta = 0. \end{aligned}$$

Since $(P\phi^+ - \phi^+) \geq 0$ and $(P\phi^- - \phi^-) \geq 0$, we conclude that $P\phi^+ - \phi^+ = 0$ and $P\phi^- - \phi^- = 0$.

□

3.1.1/ ERGODICITY AND MIXING

If h is a stationary density for a random transformation F , we can naturally define a Markov operator P_h with fixed point 1; throughout this particular class of operators we will define a mixing property. From now on we will use Lebesgue's measure as reference measure.

Proposition 3.11. *Let F be a non singular random transformation, let U^* be its adjoint transfer operator; let $h \in \mathcal{D}$ and $\eta = h\lambda$. Then the operator $P_\eta : L^1(\eta) \rightarrow L^1(\eta)$ defined as*

$$P_\eta \phi = \frac{d(U^*(\phi\eta))}{d\eta}$$

is a Markov operator. If h is also stationary, then $P_\eta 1 = 1$ and for every $\phi \in L^1(\eta)$ and $n \in \mathbb{N}$

$$P_\eta^n(\phi)h = L^n(\phi h).$$

Proof. P_η is a Markov operator by the same arguments used in proposition 3.6 and it is well defined by proposition 1.11. If h is stationary then $U^*\eta = \eta$ which implies $P_\eta 1 = 1$; furthermore for every $\phi \in L^1(\eta)$, $\phi h \in L^1(M)$ and for every $B \in \mathcal{B}$,

$$\int_B P_\eta^n(\phi)h d\lambda = (U^*)^n(\phi\eta)(B) = \int_B L^n(\phi h) d\lambda.$$

□

Definition 3.12. A non singular random transformation F is said to be **mixing** for $\eta = h d\lambda$, if for every $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$.

$$\lim_{n \rightarrow +\infty} \int_M P_\eta^n(\phi)\psi d\eta = \int_M \phi d\eta \int_M \psi h d\lambda. \quad (3.2)$$

Proposition 3.13. *If for every $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$*

$$\lim_{n \rightarrow \infty} \int_M L^n(\phi)\psi d\lambda = \int_M \phi d\lambda \int_M \psi h d\lambda, \quad (3.3)$$

then F is mixing for η . The opposite is true if $\text{supp}(\eta) = M$.

Proof. Suppose 3.3 holds; by hypothesis $\phi h \in L^1(M)$ and $\psi \chi_{h>0} \in L^\infty(M)$. Evaluating 3.3 in these functions we obtain

$$\lim_{n \in \mathbb{N}} \int_M L^n(\phi h)\psi \chi_{h>0} d\lambda = \int_M \phi h d\lambda \int_M \psi \chi_{h>0} h d\lambda = \int_M \phi d\eta \int_M \psi d\eta,$$

and this implies 3.2 because the first integral is equal to $\int_M P_\eta^n(\phi)\psi d\eta$.

Now suppose 3.2 and that $\text{supp}(\eta) = M$; if $\psi \in L^\infty(M)$ then $\psi \in L^\infty(\eta)$ and if $\phi \in L^1(M)$, there exists $\phi' \in L^1(\eta)$ such that $\phi = \phi' h$. Using 3.2

$$\lim_{n \in \mathbb{N}} \int_M P_\eta^n(\phi')\psi d\eta = \int_M \phi' d\eta \int_M \psi d\eta = \int_M \phi d\lambda \int_M \psi h d\lambda.$$

□

We will end this section with a characterization of ergodicity and mixing which clarifies the link between these two notions ([42], th. 4.2.2 and th. 4.4.1).

Proposition 3.14. *Let F be a non singular random transformation. If η is ergodic for F then there is at most one stationary density of P_η . Moreover, if there is a unique stationary density h of η and $h > 0$ η -a.e., then η is ergodic.*

Proof. Suppose η ergodic, and let ϕ_1, ϕ_2 be two stationary densities for F . Consider $\psi = \phi_1 - \phi_2$, then $P_\eta \psi = \psi$ and $P_\eta \psi^+ = \psi^+$, $P_\eta \psi^- = \psi^-$ (proposition 3.10). cannot

Let $A^\pm := \{x \in M : \psi^\pm = 0\}$ and observe that $U\chi_{A^\pm} = \chi_{A^\pm}$. In fact $U^*(\phi^\pm \eta) \ll \phi \eta$ by the non singularity condition. But η is ergodic, hence either $\eta(\chi_{A^+}) = 0$ or 1, the same for χ_{A^-} . Since their union lies in M they cannot have both null measure, thus they both have full measure, and the same must hold for the intersection $\{x \in M : \psi(x) = 0\}$, i.e. $\phi_1 = \phi_2$.

For the second part of the proposition, suppose that P_η has a unique positive stationary density $h > 0$. Let $B \in \mathcal{B}$ be any η -stationary set, let $B' = M \setminus B$ and rewrite $h = \chi_B h + \chi_{B'} h$. Now we can apply L

$$\chi_B h + \chi_{B'} h = L(\chi_B h) + L(\chi_{B'} h).$$

We observe that $L(\chi_{B'} h) = 0$ on B' because $L(\chi_B h) \geq 0$ and

$$\int_{B'} L(\chi_B h) d\eta = \int_M (U\chi_{B'}) \chi_B h d\eta = \int_M \chi_{B'} \chi_B h d\eta = 0.$$

Hence $L(\chi_B h) = \chi_B h$ and either $\chi_B = 0$ or $\chi_B = 1$ (because the stationary density is unique). \square

Proposition 3.15. *Let F be a random transformation and η a stationary measure for F . Then F is ergodic (mixing) for η if and only if the sequence $\{P_\eta^j \phi\}_{j \in \mathbb{N}}$ converge in the sense of Cesàro (weakly converge) to 1, for all $\phi \in \mathcal{D}_\eta := \{\psi \in L^1(\eta) : \int_M \psi d\eta = 1\}$.*

Proof. The Césaro convergence of $\{P_\eta^j \phi\}$ is equivalent to the pointwise convergence of the Birkhoff averages $\{A_n \phi\}$; we also observe that since η is stationary the conservative part is just M . By Birkhoff ergodic theorem, we deduce that the pointwise limit of these averages, say ϕ' , it is η -stationary and $\int \phi d\eta = \int \phi' d\eta$; furthermore it is equal to 1 for all $\phi \in \mathcal{D}_\eta$, if and only if every η -stationary ϕ' it is equal to 1, if and only if η is ergodic. Combining the latter with the fact that $L^1(\eta)$ can be identified with $L^\infty(\eta)$, give us the proof of the claim in the mixing case. \square

3.1.2/ EXISTENCE OF STATIONARY MEASURE FOR POSITIVE CONTRACTIONS

We end this section by discussing the existence of stationary measures for a random transformation, which are absolutely continuous with respect to a reference measure λ (e.g. Lebesgue measure). A first criteria is given by the Krylov-Bogolyubov procedure, which shows that a necessary and sufficient condition for the existence of stationary densities is the existence of some $\phi \in \mathcal{D}$ such that $\{\frac{1}{n} \sum_{j=0}^{n-1} L^j \phi\}_{n \in \mathbb{N}}$ is a weakly sequentially compact set.

However positive contractions satisfy the following criteria for the existence of nonzero positive fixed points ([41], th 4.2), which can be utilized for the Perron-Frobenius operators since it is a particular case of positive contractions.

Theorem 3.16. *Let P be a positive contraction on $L^1(\eta)$; there exists $f \in L^1_+(\eta)$ such that $Pf = f \neq 0$ if and only if for every strictly positive $h \in L^\infty_+(\eta)$,*

$$\inf_{n \geq 0} \int P^{*n} h d\eta > 0. \quad (3.4)$$

Moreover, there exists a strictly positive $f \in L^1_+(\eta)$ such that $Pf = f$ if and only if 3.4 holds for all non zero $h \in L^\infty_+(\eta)$.

Proof. Let $f \in L^1_+(\eta)$ such that $Pf = f \neq 0$ and let $0 \neq h \in L^\infty_+(\eta)$. If either f or h is strictly positive, then there exists $\epsilon > 0$ such that for all $n \in \mathbb{N}$

$$\int f P^{*n}(h \wedge 1) d\eta = \int (P^n f)(h \wedge 1) d\eta = \int f(h \wedge 1) d\eta > 2\epsilon \|f\|_{L^1(\eta)}.$$

Let $A_n := \{P^{*n}(h \wedge 1) \geq \epsilon\}$ and observe that $P^{*n}(h \wedge 1) \leq P^{*n}1 \leq 1$ implies

$$2\epsilon \|f\|_{L^1(\eta)} < \int f P^{*n}(h \wedge 1) d\eta < \epsilon \int_{M \setminus A_n} f d\eta + \int_{A_n} f d\eta,$$

thus for every $n \in \mathbb{N}$, $\int_{A_n} f d\eta > \epsilon \|f\|_{L^1(\eta)}$. But $f \in L^1_+(\eta)$, hence there exists a $\delta > 0$ such that $\eta(A_n) > \delta$, which implies

$$\inf_{n \geq 0} \int P^{*n} h d\eta \geq \inf_{n \geq 0} \int P^{*n}(h \wedge 1) d\eta \geq \delta \epsilon > 0.$$

Let us split the other direction of the statement in two different cases

1. If there exists a strictly positive $h \in L^\infty_+(\eta)$ such that

$$\inf_{n \geq -} \int P^{*n} h d\eta = 0,$$

then for every $f \in L^1_+(\eta)$ it holds $\inf_{n \geq 0} \int f P^{*n} h d\eta = 0$. In fact for every $n \in \mathbb{N}$

and $\epsilon > 0$ there exists $a > 0$ such that $\|(f - a)^+\|_{L^1(\eta)} < \epsilon$ and by observing that $f \leq a + (f - a)^+$ we obtain

$$\int f P^{*n} h . d\eta \leq a \int P^{*n} h d\eta + \int P^{*n} (f - a)^+ h d\eta,$$

thus

$$\int f P^{*n} h . d\eta \leq a \int P^{*n} h d\eta + \epsilon \|h\|_{L^\infty(\eta)}.$$

Hence for every $\epsilon > 0$ we have $\inf_{n \geq 0} \int f P^{*n} h d\eta \leq \epsilon \|h\|_{L^\infty(\eta)}$. If now $f \in L^1_+(\eta)$ is such that $Pf = f$ we have

$$\int f h d\eta = \inf_{n \geq 0} \int (P^n f) h d\eta = \inf_{n \geq 0} \int f P^{*n} h d\eta = 0,$$

but h is strictly positive, which implies $f = 0$.

2. If 3.4 is satisfied by every non zero $h \in L^1_+(\eta)$, then $M = C$. Otherwise there exists $c > 0$ such that $B = \{\sum_{j=0}^\infty P^j 1 < c\}$ has positive measure, which implies

$$\sum_{n=0}^\infty \int P^{*j} \chi_B d\eta = \int_B \sum_{n=0}^\infty P^j 1 < c,$$

which contradicts 3.4.

Let δ be a linear functional on $L^\infty(\eta)$ defined by

$$\delta(h) := l\left(\left\{\int P^{*n} h d\eta\right\}_{n \in \mathbb{N}}\right),$$

where l is a Banach limit. We observe that for any $h \in L^1_+(\eta)$, $\delta(h) \geq 0$ and $\delta(P^*h) = \delta(h)$. For $h \in L^1_+$ we define

$$\tilde{\nu}(h) := \inf\left\{\sum_{n=1}^\infty \delta(h_n) : h = \sum_{n=1}^\infty h_n, h_n \in L^1_+(\eta)\right\}.$$

Let us prove now that $\nu(B) := \tilde{\nu}(\chi_B)$ defines a σ -additive finite measure on \mathcal{B} . For every $b \in \mathcal{B}$, $\chi_B \in L^1_+(\eta)$ and it holds $\nu(B) \leq \delta(\chi_B) < \infty$, thus ν is a finite function; for sure it is σ -additive, because given a sequence of disjoint sets $\{B_i\}_{i \in \mathbb{N}}$ and their union $B = \cup B_i$, for every $\epsilon > 0$ there exist sequences $\{h_{i,j}\}_{j \in \mathbb{N}}$ such that

$$\chi_{B_i} = \sum_{j=1}^\infty h_{i,j} \quad \text{and} \quad \nu(B_i) > \sum_{j=1}^\infty \delta(h_{i,j}) - \frac{\epsilon}{2^i}$$

hence

$$\sum_{i=1}^\infty \nu(B_i) > \sum_{i,j=1}^\infty \delta(h_{i,j}) \geq \delta(\chi_B) \geq \nu(B).$$

the reverse inequality holds as well, because if $\{h_j\}_{j \in \mathbb{N}}$ is a sequence in $L_+^\infty(\eta)$ such that $\chi_B = \sum_{j=1}^\infty h_j$, we have

$$\sum_{i=1}^n \nu(B_i) \leq \sum_{i=1}^n \sum_{j=1}^\infty \delta(\chi_{B_i} h_j) = \sum_{j=1}^\infty \delta(\chi_{B_1 \cup \dots \cup B_n} h_j) \leq \sum_{j=1}^\infty \delta(h_j),$$

which implies using that δ is linear and positive

$$\sum_{i=1}^\infty \nu(B_i) \geq \sum_{j=1}^\infty \delta(h_j).$$

So ν is a σ -additive measure on \mathcal{B} ; let $f = \frac{dy}{d\eta}$, we want to show that $Pf = f$. In fact $\delta(P^*h) = \delta(h)$ which implies $\tilde{\nu}(P^*h) \leq \tilde{\nu}(h)$, hence $Pf \leq f$. On the other side, since $M = C$ we conclude $Pf = f$ by lemma 2.12. Now suppose that the set $B = \{f = 0\}$ has a positive measure; we observe that $\tilde{\nu}(\chi_B) = 0$ which implies that for every $m \in \mathbb{N}$ there exists a sequence $\{h_{m,n}\}_{n \in \mathbb{N}}$ such that $\chi_B = \sum_{n=1}^\infty h_{m,n}$ and $\sum_{n=1}^\infty \delta(h_{m,n}) < 1/m$. Take $k(m)$ large enough to have

$$\sum_{n=k(m)+1}^\infty \int h_{m,n} d\eta < 2^{-m} \eta(B),$$

and define $h^* = \inf_{m \in \mathbb{N}} \sum_{n=1}^{k(m)} h_{m,n}$.

Since $h_{m,n} \in L_+^\infty(\eta)$ also $h^* \in L_+^\infty(\eta)$ and moreover

$$\int h^* d\eta + \sum_{m=1}^\infty \sum_{n=k(m)+1}^\infty \int h_{m,n} d\eta \geq \eta(B),$$

hence $\int h^* d\eta > 0$ which implies in particular that $h^* \neq 0$. However for all $m \in \mathbb{N}$

$$\delta(h^*) \leq \delta\left(\sum_{n=1}^{k(m)} h_{m,n}\right) = \sum_{n=1}^{k(m)} \rho(h_{m,n}) < \frac{1}{m},$$

thus $\delta(h^*) = 0$. In particular $\liminf_{n \rightarrow \infty} \int P^n h^* d\eta = 0$, which contradicts 3.4, implying that B has null measure, i.e f is strictly positive. □

Example 3.17. Let $T : [0, 1] \rightarrow [0, 1]$ the square map $T(x) = x^2$; then T does not admit any invariant measure absolutely continuous with respect to Lebesgue. Let U be the transfer operator of T and $h(x) = x$; observe that $h \in L_+^\infty$, it is strictly positive and for every $n \in \mathbb{N}$, for every $\epsilon > 0$

$$\int U^n h(x) dx = \int_0^{\sqrt{2n\epsilon}} h(T^n(x)) dx + \int_{\sqrt{2n\epsilon}}^1 U^n h(T^n(x)) dx \leq$$

$$\leq \epsilon \sqrt{2n\epsilon} + (1 - \sqrt{2n\epsilon}) \rightarrow \epsilon,$$

as n tends to infinity, and this implies

$$\inf_{n \in \mathbb{N}} \int U^n h(x) dx = 0.$$

The next proposition shows that there is a large class of random transformations which admit stationary densities absolutely continuous with respect to Lebesgue ([20]).

Proposition 3.18. *Let $F : \Omega \times M \rightarrow \Omega \times M$ be a non singular random transformation with respect to λ . Suppose there exists $\delta > 0$ such that*

$$\lambda(B) > 1 - \delta \Rightarrow \inf_{x \in M} \mu(\{\omega \in \Omega : F_\omega(x) \in B\}) > 0,$$

then F admits a stationary density $\eta \ll \lambda$.

Proof. Let us show that there exists $f \in L_+^1(\lambda)$ such that $Lf = f \neq 0$, where L is the Perron-Frobenius operator of F . If $h \in L_+^\infty(M)$ is strictly positive, fix $\epsilon > 0$ such that $B = \{h > \epsilon\}$ satisfies $\lambda(B) > 1 - \delta$. For every $x \in M$ we have

$$\int_\Omega U h(x) d\mu(\omega) = \int_\Omega h(F_\omega(x)) d\mu(\omega) > \epsilon \inf_{x \in M} \mu(\{\omega \in \Omega : F_\omega(x) \in B\}) = \beta > 0.$$

We conclude by applying theorem 3.16, observing that for all $n \in \mathbb{N}$

$$\int U^n h d\lambda \geq \beta > 0.$$

□

3.1.3/ REGULARITY OF PERRON-FROBENIUS ITERATES

We first introduce an alternative approach to the Perron-Frobenius operator based on transition probabilities.

Definition 3.19. Let $F : \Omega \times M \rightarrow \Omega \times M$ be a random transformation and let $B \subset M$ and $x \in M$. The transition probability of x to B is

$$p(x, B) = U_{x_B}(x) = \mu(\{\omega \in \Omega : f_\omega(x) \in B\}).$$

Proposition 3.20. *The map $(x, B) \mapsto p(x, B)$, for every fixed $x \in M$ defines a probability on M and for every fixed $B \in \mathcal{B}$ defines a measurable map (sometimes such a map is called a **transition kernel**). Furthermore, if F is non singular and \mathcal{B} countably generated, then there exists a transition density $p \in L^1(M \times M)$ such that*

$$p(x, B) = \int_B p(x, \omega) d\lambda(\omega),$$

for λ -a.e. $x \in M$.

Proof. For every fixed $x \in M$, proposition 1.4 implies that the map $P_x : \omega \mapsto (\omega, x)$ is measurable; $\mu_x : B \mapsto p(x, B)$ is a probability measure on M , in fact $\mu_x(M) = \mu(\Omega) = 1$ and $\mu_x = (\pi_2 \circ F \circ P_x)_* \mu$, and for every fixed $B \in \mathcal{B}$, $x \mapsto \mu_x(B) = U_{\chi_B}(x)$ is measurable. Thus

$$\mu(A \times B) = \int_A \mu_x(B) d\lambda(x),$$

defines a measure μ over $M \times M$. μ is absolutely continuous with respect to $\lambda \times \lambda$: if $A \times B$ has zero $\lambda \times \lambda$ measure, then either $\lambda(A) = 0$ or $\lambda(B) = 0$, i.e. $\mu(A \times B) = 0$ or $p(x, B) = 0$ for a.e. $x \in M$. Then μ admits a density $p \in L^1(M \times M)$. Moreover, for every fixed $B \in \mathcal{B}$

$$\mu_x(B) = \int_B p(x, \omega) d\lambda(\omega), \quad \text{for } \lambda\text{-a.e. } x \in M.$$

By hypothesis \mathcal{B} admits a countable generator, hence there exists a set $S \subset M$ of full measure such that the expression above is true for every $B \in \mathcal{B}$.

□

Proposition 3.21. *Let F be a non singular random transformation and suppose \mathcal{B} to be countably generated. If $\phi \in L^\infty(M)$, then for λ -a.e. $x \in M$,*

$$U\phi(x) = \int_M p(x, \cdot) \phi d\lambda \quad \text{and} \quad L\phi(x) = \int_M p(\cdot, x) \phi d\lambda.$$

Proof. U is a bounded linear operator, and because the first identity is true for $\phi = \chi_B$, it follows for any $\phi \in L^\infty(M)$. On the other hand, let $B \subset M$ be a measurable set, then by Fubini's theorem

$$\int_B L\phi d\lambda = \int_M (U\chi_B) \phi d\lambda = \int_M p(\cdot, B) \phi d\lambda = \int_B \int_M p(\cdot, x) \phi d\lambda d\lambda.$$

□

By the proposition above, if F is non singular, \mathcal{B} is countably generated and $\phi \in L^1(M)$, the Perron-Frobenius operator $L : L^1(M) \rightarrow L^1(M)$ admits the following expression

$$L\phi(x) = \int p(t, x) \phi(t) d\lambda(t).$$

Therefore if we set $\Omega = M$ and $\mu = \lambda$ in corollary 1.14 we have the following

Proposition 3.22. *If $p \in L^1(M \times M)$ is of class C^k , then for every $\phi \in L^\infty(M)$, $L\phi$ is also of class C^k .*

Proof. The function $x \mapsto f(t, x) = p(t, x)\phi(t)$ is C^k for any $t \in M$. Moreover, for every $x_0 \in M$, for all $\alpha \in \mathbb{N}_0^n$, $|\alpha| \leq k$,

$$|\partial^\alpha f(t, x)| \leq \|\partial^\alpha p\|_{L^\infty(B(x_0))} |\phi(t)|,$$

where $x \in B(x_0)$, a sufficiently small neighborhood of x_0 such that $\partial^\alpha p$ are bounded. Hence corollary 1.14 applies to f and the claim follows. \square



RIGOROUS SCHEMES OF APPROXIMATION

SYSTEMS WITH ADDITIVE NOISE

From now on the rest of the thesis will focus on systems with additive noise. In this chapter, in addition to introducing the basic concepts, we will see how adding noise to a deterministic dynamical system regulates its behavior (for example, we will prove that in such random systems, every stationary measure is Lipschitz-continuous). This regularization property will allow us to rigorously approximate the transfer operator associated with a dynamical system with additive noise, as we will explain in the next chapter. A random dynamical system with additive noise is formally defined in the following way.

Definition 4.1. Let $T : M \rightarrow M$ be a non-singular transformation and let $\rho_\xi \in BV(M)$ be a probability density ($\xi \in [0, 1]$ is a fixed parameter, which represents the "size of the noise").

A random dynamical system with additive noise is the random system $T_\xi : \Omega \times M \rightarrow \Omega \times M$ defined by

$$T_\xi(\omega, x) = (\sigma(\omega), \pi(T(x) + \omega_0)),$$

where Ω is the symbolic space of the one-sided sequences on M endowed with the probability measure p with density ρ_ξ . As required in the definition of random system, the map $(\omega, x) \mapsto \pi(T(x) + \omega_0)$ depends only from the 0-th coordinate of ω ; the noise is modeled as an i.i.d. stochastic process distributed according to the probability density ρ_ξ .

In concrete applications, it might be difficult to model a natural phenomenon with a single system; it makes more sense to take into consideration a family of dynamical systems, dependent on a parameter, and to study how the corresponding statistical properties (e.g. the stationary measures) vary with this parameter: the linear response theory answers this question, and will be introduced in the second section of this chapter. In particular the linear response of a system can be described by a suitable derivative which represent the rate of change of the invariant measure with respect to the perturbation.

In the same spirit, in the last section, we will see some property of stability under suitable perturbation for random system: in a nutshell, if a system satisfies a particular notion of

mixing, there is a whole neighborhood of nearby systems satisfying the same notion.

Let $M = \mathbb{T}^n = (\mathbb{S}^1)^n$ the n -dimensional torus identified with the n -dimensional interval $[0, 1]$, as measure spaces with the Lebesgue measure λ . Let us define the convolution between two functions $f, g \in L^1(\lambda)$, which will play a central role.

Definition 4.2. Let $t = (t_1, \dots, t_n) \in \mathbb{R}^n$, let $\pi : \mathbb{R} \rightarrow [0, 1]^n$ the retraction

$$\pi(x_1, \dots, x_n) := (x_1 - \lfloor x_1 \rfloor, \dots, x_n - \lfloor x_n \rfloor),$$

and consider the family of rotations $\tau_t : L^1(\lambda) \rightarrow L^1(\lambda)$ defined as

$$\tau_t f(x_1, \dots, x_n) = f(\pi(x_1 - t_1, \dots, x_n - t_n)).$$

Let $f, g \in L^1(\lambda)$, then the convolution $f * g$ of f and g is defined in the following way

$$(f * g)(x) := \int_M (\tau_t f)(x) g(t) d\lambda(t) = \int_M f(t) (\tau_t g)(x) d\lambda(t),$$

for every $x \in M$, where in the equality above we used translational invariance of the Lebesgue measure.

4.1/ THE BOUNDED VARIATION SPACE

We will discuss random dynamical system with additive noise, where the noise belongs to the space of bounded variation functions, $BV(M)$, which will now define. Our definition takes in account that we are working on a torus.

In the following recall that $C_c^1(U, \mathbb{R}^n)$ is the space of all continuously differentiable functions $\psi : U \rightarrow \mathbb{R}^n$ such that $\overline{\{\psi > 0\}}$ is compact and $\|\psi\|_{L^\infty} = \sup_U |\psi(x)| \leq 1$.

In perspective, the work of this thesis will be applied to multidimensional contexts. As we will see in the next chapter, a fundamental role will be played by Fourier's analysis, which will allow to have estimates of errors of exponential order. With this motivation, and because it is not technically complicated, we now introduce the BV space in several dimensions, but we remark that we will work in dimension 1.

Definition 4.3. Let $U \subset \mathbb{R}^n$ be an open set, let $f \in L^1(U)$, the **variation** of f on U is defined by

$$\|Df\|(U) := \sup_{\psi \in C_c^1(U, \mathbb{R}^n)} \int_U f(x) \operatorname{div} \psi(x) d(x).$$

Let $f \in L^1(\lambda)$, let $\hat{f} = f \circ \pi$, then the **variation** of f is

$$\text{Var}(f) := \lim_{U \downarrow M} \|D\hat{f}\|(U) = \inf_{U \supset M} \|D\hat{f}\|(U), \quad (4.1)$$

where the limit and the infimum are taken over the open sets $U \supset M$. If $I \subset M$, then $\text{var}_I(f) := \text{Var}(\chi_I)$. Finally

$$BV(M) := \{f \in L^1(\lambda) : \text{Var}(f) < +\infty\}.$$

Observe that if $I \subset M$, then $\text{var}_I(f) := \text{Var}(\chi_I)$. In the following when we write $f \in C^1(M)$, we mean that $f \circ \pi$ belongs to $C^1(\mathbb{R}^n)$.

Proposition 4.4. *Let $f \in BV(M)$, the following holds:*

1. *Var is a seminorm on $BV(M)$.*
2. *let $f \in C^1(M) \cap BV(M)$, then $\text{Var}(f) = \int |\nabla f| d\lambda$:*
3. *there exists a sequence of $f_i \in C^\infty(M)$ such that $f_i \rightarrow f$ in $L^1(\lambda)$ and $\text{Var}(f_i) \rightarrow \text{Var}(f)$.*

Proof. 1. Let $f, g \in BV(M)$ and $U \supset M$ an open set, then

$$\begin{aligned} \|D(\hat{f} + \hat{g})\|(U) &= \sup_{\psi \in C_c^1(U, \mathbb{R}^n)} \int_U (\hat{f} + \hat{g})(x) \text{div } \psi(x) dx \\ &\leq \sup_{\psi \in C_c^1(U, \mathbb{R}^n)} \int_U \hat{f}(x) \text{div } \psi(x) dx + \sup_{\phi \in C_c^1(U, \mathbb{R}^n)} \int_U \hat{g} \text{div } \phi(x) dx \\ &= \|D\hat{f}\|(U) + \|D\hat{g}\|(U). \end{aligned}$$

Taking the limit $U \downarrow M$ we obtain $\text{Var}(f + g) \leq \text{Var}(f) + \text{Var}(g)$. Analogously we see that for $a \in \mathbb{R}$, $\text{Var}(af) = |a| \text{Var}(f)$; thus $\text{Var}(f)$ is a seminorm on $BV(M)$.

2. Let $U \supset M$ be an open set and let $g \in L^1(\mathbb{R}^n, \mathbb{R}^n)$ defined by

$$g(x) = \begin{cases} \frac{\nabla f(x)}{|\nabla f(x)|}, & \text{if } x \in M, |\nabla f(x)| > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Observe that $\sup_{x \in \mathbb{R}^n} |g(x)| \leq 1$; fix $\rho \in C^\infty(\mathbb{R}^n)$ such that $\int \rho d\lambda = 1$ and $\overline{\{\rho > 1\}} = B(0, 1/2)$. For every $\xi > 0$, let $\rho_\xi(t) = \xi^n \rho(t/\xi)$. For every $i \in \mathbb{N}$ let

$$\phi_i := \int \rho_{1/i}(t) g(x-t) dt = \int \rho_{1/i}(x-t) g(t) dt.$$

We remark that $\sup_{x \in U} |\phi_i(x)| \leq \int \rho_{1/i}(t) \sup |g(y)| dt \leq 1$; since M is compact, there exists $i_0 \in \mathbb{N}$ such that $B_{1/i_0}(x) \subset U$ for every $x \in M$. Let $i > i_0$, then corollary 1.14

implies that $\phi \in C_c^1(U, \mathbb{R}^N)$ and Lebesgue's density theorem ensures that $\phi_i \rightarrow g$ in $L^1(U, \mathbb{R}^n)$. Hence

$$\int_U \hat{f} \operatorname{div} \phi_i(x) dx = \int_U \nabla \hat{f}(x) \cdot \phi_i(x) dx \rightarrow \int_M |\nabla f(x)| dx.$$

Thus $\|D\hat{f}\|(U) \geq \int_M |\nabla f(x)| dx$; at the same time if $\phi \in C^1(U, \mathbb{R}^n)$, with $\|\phi\|_\infty \leq 1$, then

$$\int_U \hat{f} \operatorname{div} \phi(x) dx = \int_U \nabla \hat{f}(x) \cdot \phi(x) dx \leq \int_U |\nabla \hat{f}(x)| dx,$$

and by taking the supremum over all such ϕ we obtain $\|D\hat{f}\|(U) \leq \int_U |\nabla \hat{f}(x)| dx$. Putting together the two inequalities and taking the limit $U \downarrow M$, we get $\operatorname{Var}(f) = \int_M |\nabla f(x)| dx$.

3. For $i \in \mathbb{N}$, define the open sets $U_i := \cup_{x \in M} B_{1/i}(x)$; consider

$$f_i(x) := \int \rho_{1/i}(t) \hat{f}(x-t) dt = \int \rho_{1/i}(x-t) \hat{f} dt.$$

For each fixed i_0 , we have that for $i > i_0$, $f_i \in C^\infty(U_{i_0})$, $f_i \rightarrow \hat{f}$ in $L^1(U_{i_0})$ and in particular $f_i|_M \rightarrow f$ in $L^1(\lambda)$. Furthermore, for any $t \in \mathbb{R}^n$, for every $\alpha \in \mathbb{N}_0^n$ and for every $x, y \in \mathbb{R}^n$, since $f(x-t) = f(y-t)$, we have $\partial^\alpha f_i(x) = \partial^\alpha f_i(y)$; thus $f_i|_M \in C^\infty(M)$. By item 2,

$$\begin{aligned} \operatorname{Var}(f_i|_M) &= \int |\nabla f_i(x)| d\lambda(x) = \int \int \rho_{1/i}(t) \nabla \hat{f}(x-t) dt d\lambda(x) \\ &\leq \int_{U_i} |\nabla \hat{f}(x)| dx \int \rho_{1/i}(t) dt = \|D\hat{f}\|(U), \end{aligned}$$

in fact $x-t \in U_i$ and $t \in \overline{\{\rho_{1/i} > 0\}}$. We just proved that $\operatorname{Var}(f_i|_M) \leq \|D\hat{f}\|(U_i) \rightarrow \operatorname{Var}(f)$. For the converse, let $\epsilon > 0$, let $i_0 \in \mathbb{N}$, then there exists $\phi \in C_c^1(U_{i_0}, \mathbb{R}^N)$ such that

$$\int_{U_{i_0}} f_i|_M(\pi(x)) \operatorname{div} \phi(x) dx \rightarrow \int_{U_{i_0}} \hat{f}(x) \operatorname{div} \phi(x) dx \geq \|D\hat{f}\|(U_{i_0}) - \epsilon$$

thus

$$\liminf_{i \rightarrow +\infty} \|D(f_i|_M)\|(U_{i_0}) \geq \|D\hat{f}\|(U_{i_0}) - \epsilon \geq \operatorname{Var}(f) - \epsilon,$$

which implies that $\liminf_{i \rightarrow +\infty} \operatorname{Var}(f_i|_M) \geq \operatorname{Var}(f) - \epsilon$.

□

We now describe some properties which hold in the one-dimensional case; we will make use of them in the next section, when we will approach the problem of the regularity of the stationary densities in systems with additive noise.

Proposition 4.5. Let $f : [0, 1] \rightarrow \mathbb{R}$; x is said to be a *point of approximate continuity* of f if

$$f(x) = \frac{f(x^-) + f(x^+)}{2} \quad \text{where} \quad f(x^\pm) = \lim_{h \rightarrow 0^+} f(\pi(x \pm h)).$$

Then

1. If $g \in C^\infty(\mathbb{S}^1)$, then $V(g) = \text{Var}(g)$, where

$$V(g) := \sup_{\{x_j\}_{j=0}^{k-1}} \sum_{j=0}^{k-1} |g(x_{j+1}) - g(x_j)|, \quad x_k := x_0,$$

where all the $\{x_j\}_{j=0}^{k-1}$ are increasing sequences in $[0, 1]$ of points of approximate continuity of g .

2. if $f \in L^1(\lambda)$, then $V(f) = \text{Var}(f)$.

3. if $f \in BV(\mathbb{S}^1)$ then there exist increasing functions $f_1, f_2 : [0, 1] \rightarrow \mathbb{R}$ such that $f(x) = f_1(x) - f_2(x)$, for every x point of approximate continuity of f

Proof. Consider the periodic extension $\hat{g} \in C^\infty(\mathbb{R})$ of g and note that $g(x_0) = \hat{g}(1 + x_0)$, for any sequence $\{x_j\}_{j=0}^{k-1}$ satisfying the hypothesis. Let $y_j = x_j$ for $j \leq k-1$ and $y_k = 1 + x_0$. Then

$$\begin{aligned} \sum_{j=0}^{k-1} |g(x_{j+1}) - g(x_j)| &= \sum_{j=0}^{k-1} \left| \int_{y_j}^{y_{j+1}} \nabla \hat{g}(y) dy \right| \leq \sum_{j=0}^{k-1} \int_{y_j}^{y_{j+1}} |\nabla \hat{g}(y)| dy \\ &\leq \int_{y_0}^{1+y_0} |\nabla \hat{g}(y)| dy + \int_0^1 |\nabla \hat{h}(y)| dy = \int |\nabla g| d\lambda = \text{Var}(g). \end{aligned}$$

Thus $V(g) \leq \text{Var}(g)$ and moreover the Lebesgue integral is equivalent to a Riemann one; hence for every fixed $\epsilon > 0$ there exists $\delta > 0$ such that for any increasing sequence $\{y_j\}_{j=0}^{k+1}$ with $y_0 = 0$, $y_{k+1} = 1$, $|y_{j+1} - y_j| < \delta$ for $0 \leq j \leq k$, we have

$$\sum_{j=0}^k |g(y_{j+1}) - g(y_j)| > \int |\nabla g| d\lambda - \epsilon,$$

and since $[0, 1]$ is compact and g is continuous, we can choose δ such that every term of the above sum is less than ϵ . Almost everywhere any point x of $[0, 1]$ is a point of approximate continuity of g , hence there exists an increasing sequences $\{x_j\}_{j=0}^{k-1}$ of such points such that $x_0 < \delta$, $x_k > 1 - \delta$, $x_{j+1} - x_j < \delta$ for $0 \leq j \leq k-2$. We deduce that

$$\sum_{j=0}^{k-1} |g(x_{j+1}) - g(x_j)| > \sum_{j=0}^k |g(y_{j+1}) - g(y_j)| - 2\epsilon > \int |\nabla g| d\lambda - 3\epsilon,$$

where $x_k = x_0, y_0 = 0, y_{k+1} = 1$ and $y_j = x_{j+1}$ for $1 \leq j \leq k$; $V(g) \geq \text{Var}(g)$ and the claim follows.

By item 3 of the last proposition, there exists a sequence $\{f_i\}_{i \in \mathbb{N}}$ such that $f_{i|M} \rightarrow f$, in $L^1(\lambda)$. Take any sequence of points $\{x_j\}_{j=0}^k$ of approximate continuity of f , then since λ -a.e. $x \in [0, 1]$ is a point of approximate continuity of f , also for λ -a.e. $s \in [0, 1]$, $\pi(x_j - s)$ are points of approximate continuity of f . Furthermore, up to cyclical permutation, $\{\pi(x_j - s)\}_{j=0}^{k-1}$ is increasing, thus

$$\begin{aligned} \sum_{j=0}^{k-1} |f_i(x_{j+1}) - f_i(x_j)| &= \sum_{j=0}^{k-1} \left| \int_{-1/2}^{1/2} \rho_{1/i}(s) (f(\pi(x_{j+1} - s)) - f(\pi(x_j - s))) ds \right| \\ &\leq \int_{-1/2}^{1/2} \rho_{1/i}(s) \sum_{j=0}^{k-1} |f(\pi(x_{j+1} - s)) - f(\pi(x_j - s))| ds \\ &\leq V(f) \int_{-1/2}^{1/2} \rho_{1/i}(s) = V(f). \end{aligned}$$

From the arbitrariness of the increasing sequence follows that $V(f_{i|M}) \leq V(f)$ and we can apply item 1 to obtain $\text{Var}(f_{i|M}) \leq V(f)$; in fact $f_{i|M} \rightarrow f$ in $L^1(\lambda)$ and thus

$$\text{Var}(f) = \lim_{i \rightarrow +\infty} \text{Var}(f_{i|M}) \leq V(f).$$

Now since for any x point of approximate continuity of f holds $f_i(x) \rightarrow f(x)$, we deduce that

$$\begin{aligned} \sum_{j=0}^{k-1} |f(x_{j+1}) - f(x_j)| &= \lim_{i \rightarrow +\infty} \sum_{j=0}^{k-1} |f_i(x_{j+1}) - f_i(x_j)| \\ &\leq \lim_{i \rightarrow +\infty} \text{Var}(f_{i|M}) = \text{Var}(f), \end{aligned}$$

thus $V(f) \leq \text{Var}(f)$, which implies $V(f) = \text{Var}(f)$.

Let $c = \liminf_{y \rightarrow 0} f(y)$, where y belongs to the set of the points of approximate continuity of f . Let $f_1, f_2 : [0, 1] \rightarrow \mathbb{R}$ be

$$\begin{aligned} f_1(x) &= c + \sup_{\{x_j\}} \sum_{j=0}^{k-1} \max\{f(x_{j+1}) - f(x_j), 0\} \\ f_2(x) &= \sup_{\{x_j\}} \sum_{j=0}^{k-1} \max\{f(x_j) - f(x_{j+1}), 0\}, \end{aligned}$$

where as before, we take the supremum over all increasing sequence $\{x_j\}_{j=0}^{k-1}$ in $[0, 1]$ of points of continuity of f such that $x_j \leq x$. f_1, f_2 are increasing functions; fix $\epsilon > 0$, let x be any point of approximate continuity and let $\{x_j^1\}_{j=0}^{k-1}, \{x_j^2\}_{j=0}^{l-1}$ be increasing sequences in $[0, 1]$ upper bounded by x , such that

$$f_1(x) \geq c + \sum_{j=0}^{k-1} \max\{f(x_{j+1}^1) - f(x_j^1), 0\} > f_1(x) - \epsilon$$

$$f_2(x) \geq \sum_{j=0}^{l-1} \max\{f(x_j^2) - f(x_{j+1}^2), 0\} > f_2(x) - \epsilon.$$

Let $x_0 < \min\{x_0^1, x_0^2\}$ such that $|f(x_0) - c| < \epsilon$. Since $\max\{a, +b, 0\} \leq \max\{a, 0\} + \max\{b, 0\}$ and $\max\{a, 0\} - \max\{-a, 0\} = a$, we can consider an increasing sequence $\{x_j\}_{j=0}^{m-1}$, made by joining together the two sequences $\{x_j^1\}_{j=0}^{k-1}, \{x_j^2\}_{j=0}^{l-1}$, such that

$$(f_1(x) - \epsilon) - f_2(x) < c + \sum_{j=0}^{m-1} f(x_{j+1}) - f(x_j) < f_1(x) - (f_2(x) - \epsilon).$$

Finally

$$|f(x) - (f_1(x) - f_2(x))| \leq |c + \sum_{j=0}^{m-1} f(x_{j+1}) - f(x_j) - (f_1(x) - f_2(x))| + |f(x_0) - c| < 2\epsilon.$$

□

Let $N_\xi : L^1(M) \rightarrow L^1(M)$ be the operator defined by the convolution

$$N_\xi f = \rho_\xi * f.$$

Proposition 4.6. *The Perron-Frobenius operator L_ξ of T_ξ is $N_\xi L$, where L is the (deterministic) Perron-Frobenius operator of T .*

Proof. Let $\phi \in L^1(M)$ and let $B \subset N$ be a measurable set; then

$$\begin{aligned} \int_B L_\xi \phi d\lambda &= \int_\Omega \int_{T^{-1}(B-\omega_0)} \phi d\mu(x) = \int_S \int_{T^{-1}(B-t)} \phi(v) \rho_\xi(t) d\lambda(x) dt \\ &= \int_S \int_B L\phi(x+t) \rho_\xi(-t) d\lambda(x) dt = \int_B N_\xi L\phi d\lambda. \end{aligned} \tag{4.2}$$

□

4.2/ EXISTENCE AND REGULARITY OF STATIONARY DENSITIES

We first investigate the existence of a stationary measure for system with additive noise, which is a consequence of proposition 3.18.

Theorem 4.7. *Let $F : \Omega \times M \rightarrow \Omega \times M$, $M = [0, 1]^n$, be a random dynamical system with additive noise, where the noise has density $\rho \in L^1(\lambda)$ with respect to Lebesgue and $\lambda(\{\rho > 0\}) > 0$. Then there is a stationary measure $\eta \ll \lambda$.*

Proof. Let $\delta = \frac{1}{2}\lambda(\{\rho > 0\})$, we want to apply proposition 3.18. Let B be a set such that $\lambda(B) > 1 - \delta$, let $x \in M$

$$\begin{aligned} \lambda(\{f_\omega(x) \in B : \rho(\omega) > 0\}) &\geq \lambda(B) + \lambda(\{F_\omega x : \rho(\omega) > 0\}) - 1 > \delta \\ &\Rightarrow \lambda(\{\omega \in \{\rho > 0\} : F_\omega(x) \in B\}) > \delta. \end{aligned}$$

Let $\epsilon > 0$ such that

$$\lambda(\{\omega \in \{\rho > \epsilon\} : F_\omega(x) \in B\}) > \lambda(\{\omega \in \{\rho > 0\} : F_\omega(x) \in B\}) - \frac{\delta}{2},$$

thus for every $x \in M$

$$\mu(\{\omega \in \Omega : F_\omega(x) \in B\}) \geq \epsilon \lambda(\{\omega \in \{\rho > \epsilon\} : F_\omega(x) \in B\}) > \frac{\epsilon\delta}{2}$$

and the hypothesis of proposition 3.18 are satisfied. □

Lemma 4.8. *Let $f \in BV(M)$ and let $h \in \mathbb{R}^n$, then $\|\tau_h f - f\|_{L^1} \leq |h| \text{Var}(f)$.*

Proof. Let $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ and $e = h/|h| \in \mathbb{R}^n$ and let $\hat{\phi} = \phi e \in C^1(\mathbb{R}^n, \mathbb{R}^n)$; for $t = |h|$ we have

$$\begin{aligned} \int (\tau_h f(x) - f(x))\phi(x)d\lambda(x) &= \int f(x)(\phi(x+h) - \phi(x))d\lambda(x) \\ &= \int f(x) \int_0^t \nabla \phi(x+se) \cdot e ds d\lambda(x) \\ &= \int_0^t \int f(x) \text{div} \hat{\phi}(x+se)d\lambda(x) ds \leq t \text{Var}(f). \end{aligned}$$

Since $C^1(\mathbb{R}^n, \mathbb{R})$ is dense in $L^\infty(\mathbb{R}^n)$ we deduce $\|\tau_h f - f\|_{L^1} \leq |h| \text{Var}(f)$. □

Lemma 4.9. *If ρ in $BV(M)$ and $f \in L^1(M)$ then $\rho * f \in BV(M)$.*

Proof. $BV(M) \subset L^1(M)$, hence $\rho * f \in L^1(M)$ by Young's inequality. Let $\phi \in C_c^1(M)$ with

$\|\phi\|_\infty \leq 1$, then

$$\begin{aligned} \int (\rho * f) \operatorname{div} \phi d\lambda &= \int \int \tau_t \rho(x) f(t) \operatorname{div} \phi(x) d\lambda(t) d\lambda(x) \\ &= \int \left(\int \tau_t \rho(x) \operatorname{div} \phi(x) d\lambda(x) \right) f(t) d\lambda(t) \\ &\leq \int \operatorname{Var}(\tau_t \rho) f(t) d\lambda(t) \leq \operatorname{Var}(\tau_t \rho) \|f\|_{L^1}. \end{aligned}$$

ρ has bounded variation and $\operatorname{Var}(\tau_t \rho) = \operatorname{Var}(\rho) < +\infty$, hence we deduce that $\rho * f \in BV(M)$. \square

Lemma 4.10. *Let $BV'(m)$ be the set of bounded functions $g : [0, 1] \rightarrow \mathbb{R}$ such that for every interval $[a, b] \subset [0, 1]$, $g^{-1}([a, b])$ can be expressed as an union of m intervals, where m is a fixed integer. Then $BV'(m) \subset BV(M)$.*

Proof. Let $h \in BV'(M)$ and let $c = \sup_{x \in [0, 1]} |h(x)|$. Our claim is that $\operatorname{Var}(h) \leq 4c(m+1)$. Suppose for an absurdum to have $\operatorname{Var}(h) > 4c(m+1)$, then there exists an increasing sequence $\{x_j\}_{j=0}^{k-1}$ such that

$$\sum_{j=0}^{k-1} |h(x_{j+1}) - h(x_j)| > 4c(m+1), \quad x_k = x_0.$$

Every interval of the form $[h(x_j), h(x_{j+1})]$ or $[h(x_{j+1}), h(x_j)]$ is contained in $[-m, m]$, thus there exists an interval $[a', b'] \subset [-m, m]$ that is contained in at least $2(m+1)$ of this intervals, and at least $m+1$ are of the same type $[h(x_j), h(x_{j+1})]$ or $[h(x_{j+1}), h(x_j)]$; assuming the first type without loss of generality we have

$$[a', b'] \subset \bigcap_{i=0}^m [h(x_{j(i)}), h(x_{j(i)+1})],$$

where $j(0), \dots, j(m)$ is now an increasing sequence in $\{0, \dots, k-1\}$. Let $a \in (a', b')$ and let $b > \max\{h(x_{j(1)+1}), \dots, h(x_{j(m)+1})\}$, then $h(x_{j(i)}) \notin [a, b]$ and $h(x_{j(i)+1}) \in [a, b]$ for all $0 \leq i \leq m-1$. We deduce that $h^{-1}([a, b])$ contains at least $m+1$ intervals, each one of them contains one of the $x_{j(i)}$, and thus $g \notin BV'(M)$, which is a contradiction. \square

We are now ready to show that every stationary density is Lipschitz continuous.

Theorem 4.11. *Let L be the Perron Frobenius operator associated to a non-singular transformation T , let $\rho_\xi \in L^\infty(M)$ and let $L_\xi = N_\xi L$ the convolution operator as defined above. Then the following holds:*

1. *If $\phi \in L^1(M)$ is such that $L\phi \in BV(M)$, then $L_\xi\phi$ is Lipschitz-continuous*

2. if $\rho_\xi \in C^k(M)$ then for any $\phi \in L^1(M)$, $L_\xi \phi \in C^k(M)$
3. if $\rho_\xi \in BV(M)$ and $L(BV(M)) \subset BV(M)$, then every stationary density is Lipschitz-continuous
4. let $M = \mathbb{S}^1$ and $\rho_\xi \in BV(M)$. If for every interval $[a, b] \subset [0, 1]$, T^{-1} can be expressed as a union of m intervals, for some fixed integer m , then every stationary density is Lipschitz-continuous.

Proof. 1. If $L\phi \in BV(M)$, then for every $x, h \in M$,

$$\begin{aligned} |\tau_{-h} N_\xi L\phi(x) - N_\xi L\phi(x)| &= \left| \int \rho_\xi(t) (\tau_t \tau_{-h} L\phi(x) - \tau_t L\phi(x)) d\lambda(t) \right| \\ &\leq \int |\rho_\xi(t) (\tau_{-t} \tau_t L\phi(x) - \tau_t L\phi(x))| d\lambda(t) \\ &\leq \|\rho_\xi\|_{L^\infty} |h| \text{Var}(\tau_t L\phi) = |h| \|\rho_\xi\|_{L^\infty} \text{var}(L\phi), \end{aligned}$$

where we used 4.8, thus $N_\xi L\phi$ is Lipschitz continuous.

2. We apply corollary 1.14 to $f(\omega, x) = \rho(x - \omega)\phi(\omega)$.
3. If f is stationary then $f = L_\xi f = N_\xi L\phi$ is in $BV(M)$ by lemma 4.9, thus it is Lipschitz continuous by item 1.
4. We claim that if $f \in BV(M)$ then $Lf = f \circ T$ is in $BV'(m) \subset BV(M)$, as in lemma 4.10. By proposition 4.5, up to a null measure set, there are two increasing functions f_1 and f_2 such that $f = f_1 - f_2$. If $[a, b] \subset [0, 1]$, then $f^{-1}([a, b]) = [c, d]$ and

$$(Lf)^{-1}([a, b]) = T^{-1}(f^{-1}([a, b])) = T^{-1}([c, d])$$

is union of m intervals, thus $Lf \in BV'(m)$.

□

4.3/ LINEAR RESPONSE

The goal of this section is to provide an introduction to the linear response theory and to enunciate theorem 4.21 ([22]) which we will apply in the study of the one-dimensional Arnold map in the next chapter. To do so we need to introduce the space of Borel signed measure and the associated regularization properties of the convolution, in the same taste of the last section. Most of the proof are standard or contained in [22], and they will be omitted.

Definition 4.12. Let $BS(\mathbb{S}^1)$ be the vector space of the Borel signed measures on \mathbb{S}^1 endowed with the Wasserstein-Kantorovich norm defined on BS as

$$\|\mu\|_W = \sup_{\|g\|_\infty \leq 1, Lip(g) \leq 1} \int_{\mathbb{S}^1} g(x) d\mu(x),$$

where $Lip(g)$ is the best Lipschitz constant of g .

By some abuse of notation, in the following we will identify a measure which is absolutely continuous with respect to the Lebesgue measure with its density.

Remark 4.13. BS is not complete with respect to $\|\cdot\|$, the completion leads to a distributions space, that is the dual of the space of the Lipschitz function.

Definition 4.14. Let $f, g \in BS(\mathbb{S}^1)$, we define the convolution $f * g \in BS(\mathbb{S}^1)$ as

$$f * g(A) := \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} \chi_A(x+y) df(x) dg(y).$$

Remark 4.15. If $f \in BS(\mathbb{S}^1)$ is absolutely continuous with respect to λ , then

$$f * g(t) = \int_{\mathbb{S}^1} f(t - \tau) dg(\tau),$$

and in this case the convolution $f * g$ is a L^1 function.

Definition 4.16. Let μ be a finite Borel measure with sign on \mathbb{S}^1 . The total variation of μ is defined as

$$\|\mu\| = \mu^+(\mathbb{S}^1) + \mu^-(\mathbb{S}^1),$$

where $\mu = \mu^+ - \mu^-$ is the decomposition of μ as the difference of two positive measures.

The next lemmas provide regularization properties of the convolution, as in the last section.

Lemma 4.17. *Under the above conditions, the following inequality holds:*

$$\|f * g\|_1 \leq \|g\| * \|f\|_1$$

Proof. Suppose first that g is positive

$$\begin{aligned}
\|f * g\|_1 &\leq \sup_{\|h\|_\infty \leq 1} \left| \int_{\mathbb{S}^1} h(t) (f * g)(t) dt \right| \\
&= \sup_{\|h\|_\infty \leq 1} \left| \int_{\mathbb{S}^1} h(t) \int_{\mathbb{S}^1} f(t - \tau) dg(\tau) dt \right| \\
&\leq \sup_{\|h\|_\infty \leq 1} \left| \int_{\mathbb{S}^1} \int_{\mathbb{S}^1} h(t) f(t - \tau) dt dg \right| \\
&\leq \int_{\mathbb{S}^1} \|f\|_1 dg(\tau) \leq \|f\|_1 \cdot g(X)
\end{aligned}$$

The general case follows by the linearity of the convolution considering $g = g^+ - g^-$ and the fact that g^-, g^+ are positive measures. \square

Lemma 4.18. *Let $f \in BS(\mathbb{S}^1)$ such that $f(\mathbb{S}^1) = 0$ and $g \in L^1$. We have*

$$\|f * g\|_W \leq \|f\|_W \cdot \|g\|_1. \quad (4.3)$$

Lemma 4.19. *Let $f \in BS(\mathbb{S}^1)$, such that $f(\mathbb{S}^1) = 0$, $g \in BV$*

$$\|f * g\|_1 \leq 2\|f\|_W \cdot \|g\|_{BV}. \quad (4.4)$$

Lemma 4.20. *Let $f \in L^1$, $g \in BV$*

$$\|f * g\|_{BV} \leq \|f\|_1 \cdot \|g\|_{BV}. \quad (4.5)$$

Proof. Let us suppose first that $f, g \in C^1(\mathbb{S}^1)$. In this case $f * g \in C^1$ and

$$\text{Var}(f * g) = \int_{\mathbb{S}^1} |(f * g)'(t)| dt = \int_{\mathbb{S}^1} |f * g'(t)| dt$$

and by Lemma 4.17

$$\text{Var}(f * g) \leq \|f\|_1 \|g'\|_1$$

from which we get directly the statement.

Now suppose $f \in C^1$ and $g \in BV$, let us consider as before $g_\varepsilon \in C^1$ such that $\|g_\varepsilon - g\|_1 \leq \varepsilon$ and $\|g_\varepsilon\|_{BV} = \|g\|_{BV}$.

Now

$$\begin{aligned}
\text{Var}(f * g) &\leq \text{Var}(f * g - g_\varepsilon) + \text{Var}(f * g_\varepsilon) \\
&\leq \text{Var}(f * g - g_\varepsilon) + \|f\|_1 \|g'_\varepsilon\|_1
\end{aligned}$$

but $\text{Var}(f * g - g_\varepsilon) \leq \text{Var}(f) \|g - g_\varepsilon\| \leq \varepsilon \text{Var}(f)$, which is arbitrarily small; hence we proved the partial statement when $f \in C^1$ and $g \in BV$.

Now by approximating $f \in L^1$ by a $f_\varepsilon \in C^1$ such that $\|f_\varepsilon - f\|_1 \leq \varepsilon$ and using again Lemma 4.17 we get the full statement. \square

We are finally ready to state the following linear response result. In the statement, $(B_w, \|\cdot\|_w)$ play the role of a **weak** Banach space; considering a weak and a strong space (in this case L^1) is a common strategy in many areas of ergodic theory, for instance in the study of Lasota-York inequalities. Please remark that we use the following notation: if A, B are two normed vector spaces $T : A \rightarrow B$, we write

$$\|T\|_{A \rightarrow B} := \sup_{f \in A, \|f\|_A \leq 1} \|Tf\|_B$$

Theorem 4.21. *Suppose that the family of operators L_δ satisfies the following four conditions:*

(LR0) $f_\delta \in BV(S^1)$ is a probability measure such that $L_\delta f_\delta = f_\delta$ for each $\delta \in [0, \bar{\delta})$. Moreover there is $M \geq 0$ such that $\|f_\delta\|_{BV} \leq M$ for each $\delta \in [0, \bar{\delta})$.

(LR1) (mixing for the unperturbed operator) For each $g \in BV(S^1)$ with $\int g \, dm = 0$

$$\lim_{n \rightarrow \infty} \|L_0^n g\|_{L^1} = 0.$$

(LR2) (regularization of the unperturbed operator) L_0 is regularizing from B_w to L^1 and from L^1 to Bounded Variation i.e. $L_0 : (B_w, \|\cdot\|_w) \rightarrow L^1$, $L_0 : L^1 \rightarrow BV$ are continuous.

(LR3) (small perturbation and derivative operator) There is $K \geq 0$ such that $\|L_0 - L_\delta\|_{L^1 \rightarrow (B_w, \|\cdot\|_w)} \leq K\delta$, and $\|L_0 - L_\delta\|_{BV \rightarrow V} \leq K\delta$. There is $\dot{L}f_0 \in V_w$ such that

$$\lim_{\delta \rightarrow 0} \left\| \frac{(L_0 - L_\delta)}{\delta} f_0 - \dot{L}f_0 \right\|_w = 0. \quad (4.6)$$

Then $(Id - L_0)^{-1} : V_w \rightarrow V_w$ is a continuous operator and we have the following linear response formula

$$\lim_{\delta \rightarrow 0} \left\| \frac{f_\delta - f_0}{\delta} - (Id - L_0)^{-1} \dot{L}f_0 \right\|_w = 0. \quad (4.7)$$

Thus $(Id - L_0)^{-1} \dot{L}f_0$ represents the first-order term in the change of equilibrium measure for the family of systems L_δ .

Remark 4.22. Observe that condition LR0 is always satisfied by systems with additive noise and furthermore, the stationary measure f_0 has a density of bounded variation (see [22], Lemma 23), which is coherent with the theory developed so far from now.

Moreover by conditions LR1 and LR2 it follows that f_0 is the unique fixed probability measure of L_0 in $BS(S^1)$.

We stress on the fact, that condition LR2 (regularization property) is required only for the **unperturbed** operator L_0 as well as condition LR1 (mixing assumption).

We finally remark, that condition LR1 can be verified via computer-aided proof for a certain class of systems, as we shall see in the next chapter.

4.4/ RATE OF MIXING AND PERTURBATIONS

In concrete applications, see for instance Chapter 6, we need to verify that a families of random systems are mixing in the sense of Assumption LR1 of Theorem 4.21. Hence this assumption must be satisfied in some large set of examples, not in just one. A solution to this problem is to show that if a system is mixing, then there exists an open set of nearby systems which are mixing. In the next subsection we provide the theoretical estimates, and in the next chapter we will make advantage of them to study the rotation number of a family of Arnold maps (see for instance proposition 6.8). These estimates are from [48] and more details may be found there.

4.4.1/ PERTURBING THE MAP

In this subsection we show that the aforementioned notion of mixing is stable under suitably perturbation. In particular if a given system with additive noise is proved to be mixing, then we can extend the mixing to nearby system, and we provide quantitative estimates in the $\|\cdot\|_\infty$ norm.

Definition 4.23. A piecewise continuous map T on $[0, 1]$ is a function $T : [0, 1] \rightarrow [0, 1]$ such that there is partition $\{I_i\}_{1 \leq i \leq k}$ of $[0, 1]$ made of intervals I_i such that T has a continuous extension to the closure \bar{I}_i of each interval.

Proposition 4.24. Let T_1 and $T_2 : [0, 1] \rightarrow [0, 1]$ be piecewise continuous nonsingular maps and $\rho \in BV$. With the notations introduced above, for any $f \in L^1$ it holds

$$\|(NL_{T_1})^n f - (NL_{T_2})^n f\|_{L^1} \leq 2n \|T_1 - T_2\|_{L^\infty} \cdot \|\rho\|_{BV} \cdot \|f\|_{L^1}.$$

Let us state some preliminary lemmas.

Lemma 4.25. Let T_1 and $T_2 : [0, 1] \rightarrow [0, 1]$ be piecewise continuous nonsingular maps and let L_1, L_2 the associated deterministic transfer operators, let $f \in L^1$. Then

$$\|L_{T_1}(f) - L_{T_2}(f)\|_W \leq \|T_1 - T_2\|_{L^\infty} \cdot \|f\|_{L^1}.$$

Lemma 4.26. *Let T_1 and $T_2 : [0, 1] \rightarrow [0, 1]$ be piecewise continuous nonsingular maps and $\rho \in BV$. Let the associated transfer operators with additive noise given by the kernel ρ be NL_{T_1}, NL_{T_2} , then for any $f \in L^1$ it holds*

$$\|NL_{T_1}(f) - NL_{T_2}(f)\|_{L^1} \leq 2\|T_1 - T_2\|_{L^\infty} \cdot \|\rho\|_{BV} \cdot \|f\|_{L^1}.$$

Proof of Proposition 4.24. We have that

$$\begin{aligned} (NL_{T_1})^n - (NL_{T_2})^n &= \sum_{k=1}^n (NL_{T_1})^{n-k} (NL_{T_1} - NL_{T_2}) (NL_{T_2})^{k-1} \\ \|(NL_{T_1})^n - (NL_{T_2})^n\|_{L^1 \rightarrow L^1} &\leq n \|NL_{T_1} - NL_{T_2}\|_{L^1 \rightarrow L^1}. \end{aligned}$$

We can estimate $\|NL_{T_1} - NL_{T_2}\|_{L^1 \rightarrow L^1}$ by Lemma 4.26 we get the statement. \square

The following corollary shows how to estimate the rate of mixing of a nearby system.

Corollary 4.27. *If*

$$\|(NL_{T_1})^n f\|_{V \rightarrow L^1} \leq \alpha < 1 \quad (4.8)$$

then

$$\|(NL_{T_2})^n f\|_{V \rightarrow L^1} \leq \alpha + 2n\|T_1 - T_2\|_{L^\infty} \cdot \|\rho\|_{BV}.$$

Indeed Corollary 4.27 implies that, if we have computed n, α for which 4.8 is satisfied for the operator NL_{T_1} , then all the operators T_2 such that

$$\|T_1 - T_2\|_{L^\infty} < \frac{1 - \alpha}{2n\|\rho\|_{BV}} \quad (4.9)$$

are still mixing. We will use the latter equation in Proposition 6.8.

4.4.1.1/ PERTURBING THE NOISE

Another thing we can do, is to provide mixing and mixing rate of a system when the noise distribution is changed: the Fourier-based algorithm that we will present in this thesis can be applied to random dynamical system with additive gaussian noise, i.e. noise with a gaussian kernel. The result of this subsection shows how to recover the mixing rate for the same system with a different noise kernel, with a small perturbation in L^1 .

If $\rho_1, \rho_2 \in BV$ are two kernels we denote by N_1, N_2 the associated convolution operators.

Proposition 4.28. *For each $n \in \mathbb{N}$*

$$\|(N_1 L)^n f - (N_2 L)^n f\|_{L^1} \leq n\|\rho_1 - \rho_2\|_{L^1} \|f\|_{L^1}.$$

Proof. The proof is straightforward

$$(N_1L)^n - (N_2L)^n = \sum_{k=1}^n (N_1L)^{n-k} (N_1L - N_2L) (N_2L)^{k-1}$$

$$\|(N_1L)^n - (N_2L)^n\|_{L^1} \leq n \|N_1 - N_2\|_{L^1 \rightarrow L^1} \leq n \|\rho_1 - \rho_2\|_{L^1}.$$

□

Using Proposition 4.28 we can show how to estimate the rate of mixing of a perturbation with respect to the noise of the operator.

Corollary 4.29. *If*

$$\|(N_1L)^n f\|_{V \rightarrow L^1} \leq \alpha < 1$$

then

$$\|(N_2L)^n f\|_{V \rightarrow L^1} \leq \alpha + n \|\rho_1 - \rho_2\|_{L^1}.$$

5

RIGOROUS FINITE SCHEMES OF APPROXIMATION

In this chapter we introduce two rigorous finite schemes of approximations and their applications; let us emphasize with a list what is, in a nutshell, the core-idea of a rigorous approximation

- we choose a finite base of functions of a suitable subspace (for instance in the Ulam method they will be indicator functions as basis of V , the zero-average space of L^1), and we compose an infinite dimensional operator L with the projection operator on this base, yielding to a finite dimensional operator, i.e. a matrix L_δ
- now we search for a bound on the norm of the iterates of the discretized operator; indeed this kind of bound, which quantifies in a sense the **rate of mixing** of the operator L_δ , implies a bound on the norm of the difference between the stationary density of L and the fixed point of L_δ (see lemma 5.6). Furthermore, this bound can be established with a rigorous computation, in particular we will search for an $\alpha < 1$ and $n \in \mathbb{N}$ such that $\|L_\delta^n\| \leq \alpha$.
- all the computation are made in the framework of **Interval Arithmetic**, and not in floating point arithmetic, which means that the results of our computations are not floating point numbers, but **intervals** which contain the exact solution to the problem. If these intervals are really small (and they will be), we can then speak properly of rigorous approximation.

We stress on the fact that since we are working with interval arithmetic, the results of our computations are **mathematical statements** proved with a computer aided proof, and not **simulations**. In the next chapter we will briefly introduce interval arithmetic, but let us remark this concept with an example.

Example 5.1. When we say "we compute n, α , with $\alpha < 1$, such that $\|L_\delta\| \leq \alpha$, we actually mean that, based on the properties of the interval arithmetic, we proved the following

theorem

Theorem 5.2. *There exists $a, b \in \mathbb{Q}$, with $b < 1$, such that $\|L_\delta\| \in [a, b]$.*

The latter implies $\|L_\delta\| \leq \alpha$, with $\alpha = b$.

In the following, we first introduce a rigorous Ulam approximation ([23]), which yield to a finite stationary density (a vector) and a rigorous (small) error bound on it. Then we introduce a Fourier finite scheme of approximation which gives a rigorous approximation of the Perron-Frobenius operator (i.e. a matrix, with a rigorous error bound on it) with exponential bounds; thus, at least empirically, we should be able to manage system of dimension more than 1, which is not possible with the quadratic bounds of the Ulam approximation. The study of the multidimensional setting will be a future work.

In the next chapters we then describe three original applications, in which we use finite rigorous finite scheme of approximation to investigate (family of) random dynamical systems, using both the Ulam and Fourier approximation (the Ulam-based applications appeared in [44; 48]). Working at the crossroad of mathematics (linear response theory) and computer science (interval arithmetic) yields not only to **concrete** results (e.g. we compute up to a certain approximation the Lyapunov exponents of a random system) but also to **theoretical results** (e.g. the rotation number of the Arnold map is a differentiable map).

5.1/ ULAM APPROXIMATION

In this section we describe part of the work contained in [23] [24], which is at the core of the thesis, and we will use this framework to present the work of the next section. In particular we will try to explain, without going through all the details (which again, are contained in [23], [24]), the ingredients of a rigorous computation, so that the reader can be prepared to approach the next section, in which we present a new scheme of approximation (in its entirety).

Setting 5.3. For the rest of this section, $L_\xi = N_\xi L$ will be the Perron Frobenius operator of the random dynamical system with additive noise T_ξ (see proposition 4.6), and the noise is distributed according to a kernel $\rho_\xi \in BV(M)$, with support in $[-\frac{\xi}{2}, \frac{\xi}{2}]$. We also know that there exists a stationary density f_ξ , such that $L_\xi f_\xi = f_\xi$.

In 1960 (see [59]) Ulam introduces a scheme of approximation, which perhaps is still the most popular one, for the discretization of a Perron-Frobenius operator.

Definition 5.4. Let $\{B_1, \dots, B_k\}$ be a partition of M by a finite number of disjoint sets, let T a non singular transformation over M , then the Ulam's matrix of order k is the matrix

$(p_{ij}) = P \in M(k, \mathbb{R})$ whose entries are defined by the following relation

$$p_{ij} = \frac{\lambda(T^{-1}(B_j) \cap B_i)}{\lambda(B_i)}. \quad (5.1)$$

The definition above is justified from the fact that one can look at the entries p_{ij} as the probabilities of being mapped from box B_i to box B_j by the dynamical system T . P is a row-stochastic matrix since the denominator $\lambda(B_i)$ normalizes the entries and thus P defines a finite Markov chain and has a left eigenvector with eigenvalues 1; whenever or not this eigenvector approximates the invariant measure of the Perron-Frobenius operator associated to T , is an open question, called the Ulam's conjecture, which has been solved, for instance, in the one-dimensional case ([16], [21], [43], [52]).

Actually the Ulam's method is a Galerkin projection of the Perron-Frobenius operator to the subspace spanned by the functions $\psi_i = \frac{1}{\lambda(B_i)}\chi_{B_i}$. Someone could ask if using higher-order functions in the Galerkin projection could improve the approximation of the operator: the answer is that in some case it is convenient, however high order approximations do have the disadvantage that the discretized operator is not, in general, a Markov operator (i.e. a stochastic matrix, [39], Section 3). In [34] specific Petrov-Galerkin methods which use piecewise first or second order polynomial as bases functions, are able to recover the Markov property for the approximated operator.

Thus let us redefine the Ulam's method as a Galerkin projection, to adapt it to our context.

Definition 5.5. Let $M = \mathbb{T}^n$, let $\mathcal{I}_\delta = \{I_j\}_{j=1}^l$ be a δ -partition of M (i.e. for every j , the I_j -s are convex and $\text{diam}(I_j) \leq \delta$); given an operator $L : L^1(M) \rightarrow L^1(M)$ the associated discretized (Ulam) operator is $L_\delta : L^1(\lambda) \rightarrow L^1(\lambda)$ defined by

$$L_\delta = \pi_\delta L \pi_\delta,$$

where $\pi_\delta : L^1(\lambda) \rightarrow L^1(\lambda)$ is the projection

$$\pi_\delta h(x) = \sum_{i=1}^l \mathbb{E}(h|I_i)\chi_{I_i}.$$

Thus, our purpose is to study the properties of the discretized operator $\pi_\delta L_\xi \pi_\delta$; given the particular nature of L_ξ , which is the product of the two operators N_ξ and L , for computational purpose we consider the following discretization

$$L_{\delta,\xi} := \pi_\delta N_\xi \pi_\delta L \pi_\delta = (\pi_\delta N_\xi \pi_\delta)(\pi_\delta L \pi_\delta).$$

Indeed with this definition, we compute once and for all the matrix $\pi_\delta L \pi_\delta$, which is expensive from a computational point of view, but leads to a sparse matrix. We then apply the

operator $\pi_\delta N_\xi \pi_\delta$ which is independent of the dynamics. Another reasonable discretization could be

$$\pi_\delta N_\xi \pi_\delta$$

, but this definition implies that we must recompute the discretized operator for each size of the noise.

Observe that the invariant measure of $L_{\delta,\xi}$ (the eigenvector of eigenvalue 1), $f_{\delta,\xi}$, can be computed up to any precision, because $L_{\delta,\xi}$ is a finite operator: hence the problem becomes how to estimate effectively

$$\|f_{\delta,\xi} - f_\xi\|_{L^1}.$$

5.1.1/ BOUNDING THE RATE OF MIXING

Let V be the zero average subspace of L^1 . We start with a very simple but powerful lemma, which explains why it is useful, for our purpose, to have a bound on the norms of the iterates of a Markov operator (see definition 3.7).

Lemma 5.6. *Suppose that there exists $\bar{n} \in \mathbb{N}$ and $\alpha > 0$ such that*

$$\|L_{\delta,\xi}^{\bar{n}}|_V\|_{L^1 \rightarrow L^1} \leq \alpha < 1,$$

then

$$\|f_\xi - f_{\delta,\xi}\|_{L^1} \leq \frac{1}{1-\alpha} \|(L_{\delta,\xi}^{\bar{n}} - L_\xi^{\bar{n}})f_\xi\|_{L^1}. \quad (5.2)$$

Proof. Since both f_ξ and $f_{\delta,\xi}$ are fixed points

$$\begin{aligned} \|f_{\delta,\xi} - f_\xi\|_{L^1} &= \|L_{\delta,\xi}^{\bar{n}} f_{\delta,\xi} - L_\xi^{\bar{n}} f_\xi\|_{L^1} \\ &+ \|L_{\delta,\xi}^{\bar{n}} f_{\delta,\xi} - L_{\delta,\xi}^{\bar{n}} f_\xi + L_{\delta,\xi}^{\bar{n}} f_\xi - L_\xi^{\bar{n}} f_\xi\|_{L^1} \\ &\leq \|L_{\delta,\xi}^{\bar{n}}(f_{\delta,\xi} - f_\xi)\|_{L^1} + \|(L_{\delta,\xi}^{\bar{n}} - L_\xi^{\bar{n}})f_\xi\|_{L^1}. \end{aligned} \quad (5.3)$$

Then the claim follows because

$$\|f_\xi - f_{\delta,\xi}\|_{L^1} \leq \alpha \|f_\xi - f_{\delta,\xi}\|_{L^1} + \|(L_{\delta,\xi}^{\bar{n}} - L_\xi^{\bar{n}})f_\xi\|_{L^1}.$$

□

Indeed with the useful lemma we reduce the problem to finding a good upper bound on $\|(L_{\delta,\xi}^{\bar{n}} - L_\xi^{\bar{n}})f_\xi\|_{L^1}$. The following lemma, via telescopic decomposition, will show that this is equivalent to the study of three quantities, which are way easier to be bounded.

Lemma 5.7. Let $\|L_{\delta,\xi}^i|_V\|_{L^1} \leq C_i$ and let f_ξ be an invariant probability measure for L_ξ , it holds

$$\begin{aligned} \|L_{\delta,\xi}^n - L_\xi^n\|_{L^1} f_\xi &\leq \|(\pi_\delta - 1)f_\xi\|_{L^1} \\ &+ \left(\sum_{i=0}^{n-1} C_i\right) \cdot (\|N_\xi(\pi_\delta - 1)Lf_\xi\|_{L^1} \\ &+ \|N_\xi\pi_\delta L(\pi_\delta - 1)f_\xi\|_{L^1}). \end{aligned} \quad (5.4)$$

Proof. We proceed by telescopic decomposition

$$\begin{aligned} L_{\delta,\xi} &= \pi_\delta N_\xi \pi_\delta - N_\xi L \\ &= N_\xi L - \pi_\delta N_\xi L \\ &+ \pi_\delta N_\xi L - \pi_\delta N_\xi \pi_\delta L \\ &+ \pi_\delta N_\xi \pi_\delta L - \pi_\delta N_\xi \pi_\delta L \pi_\delta. \end{aligned}$$

With a similar decomposition on $L_{\delta,\xi}^n = (\pi_\delta N_\xi \pi_\delta L)^n \pi_\delta$, we obtain:

$$\begin{aligned} \|(L_{\delta,\xi}^n - L_\xi^n)f_\xi\|_{L^1} &= \|((\pi_\delta N_\xi \pi_\delta L)^n \pi_\delta - (N_\xi L)^n)f_\xi\|_{L^1} \\ &\leq \sum_{i=0}^n \|(\pi_\delta N_\xi \pi_\delta L)^i (\pi_\delta - 1) (N_\xi L)^{n-i} f_\xi\|_{L^1} \\ &+ \sum_{i=0}^{n-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta N_\xi (\pi_\delta - 1) L (N_\xi L)^{n-i-1} f_\xi\|_{L^1} \\ &= \sum_{i=0}^n \|(\pi_\delta N_\xi \pi_\delta L)^i (\pi_\delta - 1) f_\xi\|_{L^1} \\ &+ \sum_{i=0}^{n-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1}. \end{aligned}$$

We can shift the indexes by 1 in the first sum, thus

$$\begin{aligned} \|(L_{\delta,\xi}^n - L_\xi^n)f_\xi\|_{L^1} &\leq \|(\pi_\delta - 1)f_\xi\|_{L^1} + \sum_{i=0}^{n-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta|_V\|_{L^1 \rightarrow L^1} \cdot \|N_\xi \pi_\delta L(\pi_\delta - 1)f_\xi\|_{L^1} \\ &+ \sum_{i=0}^{n-1} \| * \pi_\delta N_\xi \pi_\delta L)^i \pi_\delta|_V\|_{L^1 \rightarrow L^1} \cdot \|N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1} \\ &= \|(\pi_\delta - 1)f_\xi\|_{L^1} + \left(\sum_{i=0}^{n-1} C_i\right) \cdot (\|N_\xi (\pi_\delta - 1) L f_\xi\|_{L^1} + \|N_\xi \pi_\delta L(\pi_\delta - 1)f_\xi\|_{L^1}). \end{aligned}$$

□

Looking at the three terms, it is easy to realize that what we really need to bound are the

quantities

$$\|(1 - \pi_\delta)N_\xi\|_{L^1},$$

$$\|N_\xi(1 - \pi_\delta)\|_{L^1}.$$

The following theorem (5.8), provides these bounds; we report it in its completeness since it provides other bounds that may be useful to the reader in other situations. Within lemma 5.7 and lemma 5.6, we can use part of theorem 5.8, to deduce an initial bound on $\|f_{\delta,\xi} - f_\xi\|_{L^1}$.

Theorem 5.8. *Let \mathcal{I}_δ be a partition of M into convex sets, let π_δ be the associated Ulam projection, let I be a finite union of sets from \mathcal{I}_δ and let N_ξ be the convolution operator; then*

1. $\|1 - \pi_\delta\|_{\text{var} \rightarrow L^1} \leq \delta/2$
2. $\|1 - \pi_\delta\|_{L^1 \rightarrow W} \leq \delta/2;$
3. $\|1 - \pi_\delta\|_{\text{var}_I \rightarrow W(I)} \leq \delta^2/8$
4. $\|N_\xi\|_{L^1 \rightarrow \text{var}} \leq \xi^{-1} \text{Var}(\rho_\xi)$
5. $\|N_\xi\|_{W \rightarrow L^1} \leq \xi^{-1} \text{Var}(\rho_\xi)$
6. $\|(1 - \pi_\delta)N_\xi\|_{L^1} \leq \frac{1}{2}\delta\xi^{-1} \text{Var}(\rho_\xi)$
7. $\|N_\xi(1 - \pi_\delta)\|_{L^1} \leq \frac{1}{2}\delta\xi^{-1} \text{Var}(\rho_\xi).$

Lemma 5.9. *If there exist $\bar{n}, \alpha < 1$ such that*

$$\|L_{\delta,\xi}^{\bar{n}}|V\|_{L^1} \leq \alpha < 1$$

then if $0 \leq C_i \leq 1$ are such that $\|L_{\delta,\xi}^i\|_{L^1} \leq C_i$, we have

$$\|f_{\delta,\xi} - f_\xi\|_{L^1} \leq \frac{1 + 2 \sum_{i=0}^{\bar{n}-1} C_i}{2(1 - \alpha)} \delta \xi^{-1} \text{Var}(\rho). \quad (5.5)$$

Proof. Theorem 5.8 implies that

$$\|N_\xi(1 - \pi_\delta)\|_{L^1 \rightarrow L^1} \leq \frac{1}{2} \delta \xi^{-1} \text{Var}(\rho)$$

$$\|(1 - \pi_\delta)N_\xi\|_{L^1 \rightarrow L^1} \leq \frac{1}{2} \delta \xi^{-1} \text{Var}(\rho).$$

We observe that on the right hand side of 5.4 all the items have either $N_\xi(1 - \pi_\delta)$ or $(1 - \pi_\delta)N_\xi$. Indeed since $\|f_\xi\|_{L^1} \leq 1$,

$$\|(\pi_\delta - 1)f_\xi\|_{L^1} = \|(\pi_\delta - 1)N_\xi L f_\xi\|_{L^1} \leq \frac{1}{2}\delta\xi^{-1} \text{Var}(\rho),$$

and similarly for the other summands. Finally applying 5.6 we get the statement. \square

5.1.2/ STRONGER BOUNDS AND COARSE-FINE STRATEGY

Usually we are going to work with large matrices, hence it would be useful to compute some of the properties of these matrices (for instance its rate of mixing) on smaller matrices. Indeed, if the noise has already been applied to the deterministic system, this is possible; the following lemma, which bounds the distance between the powers of L_ξ and $L_{\delta,\xi}$ (provided that the noise has been applied at least once), yields to a bound on the contraction rate of a certain operator using the computed contraction rate of a coarser one.

Furthermore, assuming some regularity on the deterministic map T , it is possible to strengthen the previous bounds via a **coarse-fine strategy**.

This strategy depends from several estimates, and we will resume them without reporting their proof, which can be found in [23]; please remark, that we will develop a similar strategy (in a different taste) in the next section, in the Fourier context, and there we will provide the full details.

Lemma 5.10. *Let $\|L_{\delta,\xi}^i|V\|_{L^1} \leq C_i$; let σ be a linear operator such that $\|\sigma\|_{L^1} \leq 1$, $\sigma^2 = \sigma$ and $\sigma\pi_\delta = \pi_\delta\sigma = \pi_\delta$; let $\Lambda = \sigma N_\xi \sigma L$. Then for all $n \geq 0$*

$$\|(L_{\delta,\xi}^n - \Lambda^n)N_\xi\|_{L^1} \leq \frac{\delta}{\xi} \cdot (2 \sum_{i=0}^{n-1} C_i + 1).$$

In particular the hypothesis are satisfied if

1. $\sigma = Id$ and $\Lambda = L_\xi$
2. $\sigma = \pi_{\delta'}$ and $\Lambda = L_{\delta',\xi}$.

Proof. Since $\sigma g - \pi_\delta g = \sigma(1 - \pi_\delta)g$, applying theorem 5.8, we have

$$\|\sigma g - \pi_\delta g\|_{L^1} \leq \text{Var}(g)\delta/2, \quad \|\sigma g - \pi_\delta g\|_W \leq \|g\|_{L^1}\delta/2.$$

Then we have

$$\begin{aligned}
\|(L_{\delta,\xi}^n - \Lambda^n)N_\xi\|_{L^1} &\leq \|(\pi_\delta N_\xi \pi_\delta L)^n\|_{L^1} \cdot \|(\pi_\delta - \sigma)N_\xi\|_{L^1} \\
&+ \sum_{i=0}^{n-1} \|(\pi_\delta N_\xi \pi_\delta L)^i\|_{L^1} \|(\pi_\delta - \sigma)N_\xi\|_{L^1} \cdot \|(\sigma L \sigma N_\xi)^{n-i}\|_{L^1} \\
&+ \sum_{i=0}^{n-1} \|(\pi_\delta N_\xi \pi_\delta L)^i \pi_\delta\|_{L^1} \cdot \|N_\xi(\pi_\delta - \sigma)\|_{L^1} \cdot \|L \sigma N_\xi (\sigma L \sigma N_\xi)^{n-i-1}\|_{L^1} \\
&\leq \|(\pi_\delta - \sigma)N_\xi\|_{L^1} + \sum_{i=0}^{n-1} C_i \|(\pi_\delta - \sigma)N_\xi\|_{L^1} \cdot \|(\sigma L \sigma N_\xi)^{n-i}\|_{L^1} \\
&+ \sum_{i=0}^{n-1} C_i \|N_\xi(\pi_\delta - \sigma)\|_{L^1} \cdot \|L \sigma N_\xi (\sigma L \sigma N_\xi)^{n-i-1}\|_{L^1},
\end{aligned} \tag{5.6}$$

and the thesis follows because $\|L\|_{L^1} \leq 1$, $\|N_\xi\|_{L^1} \leq 1$ and $\|\sigma\|_{L^1} \leq 1$. \square

The estimate given at 5.9 depend on the ratio $\delta\xi^{-1}$ and is obtained only with the information about the approximated transfer operator $L_{\delta,\xi}$ (in particular its contraction rate). In order to obtain stronger bounds we need to assume some regularity of the deterministic part T .

Setting 5.11. In this subsection, the map T is assumed to be piecewise $C^{1+\gamma}$, $\gamma > 0$, and monotonic on a partition $\{C_j\}$ of $[0,1]$. If we denote $T_i = T|_{C_i}$, then for each density $g \in L^1(M)$ we denote by $L_i g = L(g\chi_i)$ the component of Lg coming from the i -th branch, in order to have

$$Lg = \sum_i L_i g.$$

We can now use theorem 5.8 and the regularity of T to give better bounds on the terms appearing in 5.4.

Theorem 5.12. *Each term in 5.4 admit a bound of the form*

$$A_i \|f_\xi - \hat{f}\|_{L^1} + B_i,$$

for $i = 1, 2, 3$. This implies the following bound

$$\|f_\xi - \hat{f}\|_{L^1} \leq \frac{1}{1-D} \cdot (\|f_{\delta,\xi} - \hat{f}\|_{L^1}) + C$$

Proof. The three bounds for each term of 5.4 will be proven in the next three lemmas. Plugging these estimates in 5.4 lead to

$$\|(L_{\delta,\xi}^{\bar{n}} - L_{\xi}^{\bar{n}})f_{\xi}\|_{L^1} \leq A \cdot \|f_{\xi} - \hat{f}\|_{L^1} + B,$$

where

$$A = A_1 + (A_2 + A_3) \sum_{i=0}^{\bar{n}-1} C_i \quad B = B_1 + (B_2 + B_3) \sum_{i=0}^{\bar{n}-1} C_i.$$

If $C = A/(1 - \alpha)$ and $D = B/(1 - \alpha)$ (where α comes from 5.6), using 5.2 we obtain

$$\|f_{\xi} - f_{\delta,\xi}\|_{L^1} \leq C + D \cdot \|f_{\xi} - \hat{f}\|_{L^1}$$

and then

$$\|f_{\xi} - \hat{f}\|_{L^1} \leq \|f_{\delta,\xi} - \hat{f}\|_{L^1} + C_D \cdot \|f_{\xi} - \hat{f}\|_{L^1}$$

which implies

$$\|f_{\xi} - \hat{f}\|_{L^1} \leq \frac{1}{1 - D} \cdot (\|f_{\delta,\xi} - \hat{f}\|_{L^1} + C).$$

□

For the sake of completeness we report the constants A_i, B_i , and we remark again that they are taken from [23].

- $A_1 = \frac{\delta}{2} \xi^{-1} \text{Var}(\rho)$ and $B_1 = \frac{\delta}{2} \text{Var}(N_{\xi} L f \hat{f})$
- $A_2 = \frac{\delta}{2} \text{Var}(\rho_{\xi})$ and $B_2 = \frac{\delta}{2} \text{Var}(\rho_{\xi}) \sum_{I \in \Pi} \sum_i \min \{ \frac{\delta}{4}, \text{Var}_I(L_i \hat{f}), \|L_i \hat{f}\|_{L^1(I)} \}$
- $A_3 = \frac{\delta}{2} \xi^{-1} \text{Var}(\rho)$
- $B_3 = \sum_{I \in \Pi} \min \{ \frac{\delta^2}{8} \xi^{-1} \text{Var}(\rho) \cdot \|T'\|_{L^{\infty}(I)}, \frac{\delta}{2} \} \text{Var}_I(N_{\xi} L f \hat{f}) + \frac{\delta^2}{4} \xi^{-1} \text{Var}(\rho) \cdot \text{Var}(N_{\xi} L f \hat{f}).$

Please remark that in the second and third item Π is a uniform partition of $[0, 1]$ whose parts have size that is an integer multiple of δ and T is piecewise monotonic with L_i defined above.

5.2/ FOURIER APPROXIMATION

The Ulam method proposed in the last section is not suitable to work with multidimensional systems, because the size of the matrices involved becomes too large to be handle. Moreover, one of the advantage of the Ulam method, which is that it does not require regularity of the map, it is also its disadvantage, since this does it means also that the Ulam method cannot exploit the regularity of the map if there is any. The idea is then to switch bases of function and instead of considering characteristic functions, we will consider the classic Fourier basis; the regularity of the system will come from the convolution

with the noise, that we suppose, in this section, to have a Gaussian kernel. As we will see, this approach, which is slightly different from a computational point of view from the Ulam one, leads to exponential bounds, which are exactly what we want in the perspective of a multidimensional applications. This is a work in progress: here we present the theoretical bounds and in the last chapter we will see an application of this method; further examples will be presented in [47].

Let us introduce this new setting.

Let ℓ_∞ be the space of bilateral bounded sequences and let \mathcal{F} be the operator

$$\mathcal{F}(f)[k] = \int_{S^1} f(x)e^{-2k\pi ix} dx.$$

This operator, is a version of the well known **Fourier transform**.

Lemma 5.13. *The operator \mathcal{F} is continuous from L^1 to ℓ_∞ and*

$$\|\mathcal{F}\|_{L^1 \rightarrow \ell_\infty} = 1.$$

Proof. Since $|e^{2k\pi ix}| = 1$ we have that for all k :

$$|\mathcal{F}(f)[k]| = \left| \int_{S^1} f(x)e^{-2k\pi ix} dx \right| \leq \int_{S^1} |f(x)| |e^{2k\pi ix}| dx \leq \|f\|_{L^1}.$$

□

To simplify the notations, let $\gamma_\sigma = \frac{1}{\sqrt{2\pi\sigma^2}}$ and let $\rho_\sigma(x) = \gamma_\sigma e^{-\frac{x^2}{2\sigma^2}}$. These constants will appear through the rest of this section.

Let $f \in L^1$ and let $N_\sigma f$ be the convolution of ρ_σ with the periodic extension of f , i.e.

$$N_\sigma f = \rho_\sigma \hat{*} f = \sum_{j=-\infty}^{+\infty} \int_0^1 \rho_\sigma(x + j - y) f(y) dy,$$

By definition $N_\sigma f$ is a periodic function and $\forall x \in S^1$ we have that

$$N_\sigma f(x) \leq 2 \sum_{j=0}^{+\infty} \rho_\sigma(x + j) \|f\|_{L^1},$$

which we can bound uniformly by

$$2\|f\|_{L^1} \sum_{j=0}^{+\infty} \rho_\sigma(j),$$

since the sum is upper bounded, for instance, by $\int_0^\infty \rho_\sigma < +\infty$.

The following lemma explains the relationship between the Fourier transform and the operator N_σ .

Lemma 5.14. *There exists an operator $D_\sigma : \ell_\infty \rightarrow \ell^\infty$ such that*

$$D_\sigma \mathcal{F} = \mathcal{F} N_\sigma$$

Proof. By direct computation we have

$$\mathcal{F} N_\sigma f(x) = \int_0^1 \sum_{j=-\infty}^{+\infty} \int_0^1 f(x-y) \rho_\sigma(y+2kj) dy e^{-2k\pi i x} dx.$$

Since the series in the integral is uniformly bounded, the integral above is equal to

$$\sum_{j=-\infty}^{+\infty} \int_0^1 \int_0^1 f(x-y) \rho_\sigma(y+2kj) e^{-2k\pi i x} dy dx.$$

Let $z = x - y$ then

$$\begin{aligned} & \sum_{j=-\infty}^{+\infty} \int_0^1 \int_0^1 f(z) \rho_\sigma(y+2kj) e^{-2k\pi i(z-y)} dy dx \\ &= \sum_{j=-\infty}^{+\infty} \int_0^1 \int_0^1 f(z) e^{-2k\pi i z} \rho_\sigma(y+2kj) e^{+2k\pi i y} dy dz \\ &= \int_0^1 f(z) e^{-2k\pi i z} dz \sum_{j=-\infty}^{+\infty} \int_0^1 \rho_\sigma(y+2kj) e^{+2k\pi i y} dy. \end{aligned}$$

Which implies that

$$D_\sigma e_j = \int_0^1 \rho_\sigma(y+2kj) e^{2k\pi i y} dy e_j = e^{-\frac{k^2 \sigma^2}{2}} e_j.$$

□

Indeed, we have something stronger:

Corollary 5.15. *Let*

$$\ell^{\text{exp}} = \{(a_k)_{k \in \mathbb{Z}} \mid a_k \leq C e^{-|k|}\}.$$

Then D_σ maps ℓ_∞ into ℓ^{exp} .

Proof. Since each sequence in ℓ_∞ is bounded, this follows from the fact that

$$e^{-\frac{k^2 \sigma^2}{2}}$$

decays exponentially.

□

Now recall (it is a standard result in Fourier analysis) that if a sequence belongs to ℓ^{exp} , then the associated Fourier series converges uniformly to an analytic function.

Theorem 5.16. *Let $\{a_k\}_{k \in \mathbb{Z}}$ be a sequence in ℓ^{exp} . Then*

$$\mathcal{F}^{-1}(\{a_k\}) = \sum_{-\infty}^{+\infty} a_k e^{2\pi i k x}$$

converges uniformly to an analytic function. Moreover, if f is in L^2 , we have that $\mathcal{F}^{-1}\mathcal{F}(f) = f$.

We are now ready to define the projection π of a function over a finite Fourier basis, and to exploit the properties of π in relation to the operators \mathcal{F} and D_σ .

Definition 5.17. Let $\pi_k : L^1 \rightarrow L^\infty$ be defined as

$$\pi_k f = \sum_{j=-k}^k \left(\int_0^1 f(x) e^{-2i\pi j x} dx \right) e^{-2i\pi j x} = \sum_{j=-k}^k \mathcal{F}(f)[j] e^{-2i\pi j x}.$$

Definition 5.18. We denote by $\tilde{\pi}_k : \ell_\infty \rightarrow \ell_\infty$ the operator that sends $\{a_j\}_{j \in \mathbb{Z}}$ to the sequence $\{b_j\}_{j \in \mathbb{Z}}$ such that $b_j = a_j$ for $|j| \leq k$, $b_j = 0$ for $|j| > k$.

Lemma 5.19. *The following are true*

1. D_σ commutes with $\tilde{\pi}_k$ and $(1 - \tilde{\pi}_k)$,
2. $\mathcal{F}^{-1}\tilde{\pi}_k\mathcal{F} = \pi_k f$.

Remark 5.20. Since it has only a finite number of Fourier coefficients, the convergence of

$$\mathcal{F}^{-1}\tilde{\pi}_k\mathcal{F}f$$

is trivial.

As we already saw in the last section, a central role is played by the quantities, $\|N_\sigma(1 - \pi_k)\|$ and $\|(1 - \pi_k)N_\sigma\|$, and our goal is to bound them. We have now all the tools to prove the following theorem, which provides exponential bounds.

Theorem 5.21. *Let $\Gamma_{\sigma,k} := \gamma_\sigma \frac{e^{-\frac{k^2\sigma^2}{2}}}{k}$; the following are true*

1. $\|N_\sigma(1 - \pi_k)\|_{L^1 \rightarrow L^\infty} \leq 2 \sum_{j=k}^{+\infty} e^{-\frac{j^2\sigma^2}{2}} \leq \Gamma_{\sigma,k}$,
2. $\|(1 - \pi_k)N_\sigma\|_{L^1 \rightarrow L^\infty} \leq 2 \sum_{j=k}^{+\infty} e^{-\frac{j^2\sigma^2}{2}} \leq \Gamma_{\sigma,k}$.

Proof. We prove the first item; we proceed with a formal computation, which will be proved to be well defined then, by showing that \mathcal{F}^{-1} is defined.

$$\begin{aligned} 3N_\sigma(1 - \pi_k) &= \mathcal{F}^{-1}\mathcal{F}N_\sigma(1 - \pi_k) \\ &= \mathcal{F}^{-1}D_\sigma\mathcal{F}(1 - \pi_k) = \\ &= \mathcal{F}^{-1}D_\sigma(1 - \tilde{\pi}_k)\mathcal{F}. \end{aligned}$$

Since D_σ maps ℓ_∞ into ℓ^{exp} we have that $\mathcal{F}^{-1}D_\sigma(1 - \tilde{\pi}_k)\mathcal{F}f$ is an analytic function whose Fourier coefficients are such that $a_j = 0$ for $|j| \leq k$ and

$$|a_j| \leq \|f\|_{L^1} e^{-\frac{j^2\sigma^2}{2}}$$

for $|j| > k$. This proves also that

$$\|N_\sigma(1 - \pi_k)\|_{L^1 \rightarrow L^\infty} = \|\mathcal{F}^{-1}D_\sigma(1 - \tilde{\pi}_k)\mathcal{F}\|_{L^1 \rightarrow L^\infty} \leq 2 \sum_{j=k}^{+\infty} e^{-\frac{j^2\sigma^2}{2}} \leq \frac{e^{-\frac{k^2\sigma^2}{2}}}{k\sigma\sqrt{2\pi}} = \Gamma_{\sigma,k}.$$

Observe now that

$$\begin{aligned} (1 - \pi_k)N_\sigma &= \mathcal{F}^{-1}\mathcal{F}(1 - \pi_k)N_\sigma \\ &= \mathcal{F}^{-1}(1 - \tilde{\pi}_k)D_\sigma\mathcal{F} \\ &= \mathcal{F}^{-1}D_\sigma(1 - \tilde{\pi}_k)\mathcal{F} \end{aligned}$$

hence the proof can be adapted also to the second item. □

5.2.1/ BOUNDING THE RATE OF MIXING

In this section we do the analogous of what we did for the Ulam method, in a new taste, using the new bounds we just obtained.

We start by proving the following lemma.

Lemma 5.22. $\pi_k N_\sigma \pi_k L \pi_k = \pi_k N_\sigma L \pi_k$

Proof. Let use the Fourier transform,

$$\begin{aligned}
\pi_k N_\sigma \pi_k L \pi_k &= \mathcal{F}^{-1} \mathcal{F} \pi_k N_\sigma \pi_k L \pi_k \\
&= \mathcal{F}^{-1} \tilde{\pi}_k D_\sigma \tilde{\pi}_k \mathcal{F} L \pi_k \\
&= \mathcal{F}^{-1} \tilde{\pi}_k \tilde{\pi}_k D_\sigma \mathcal{F} L \pi_k \\
&= \mathcal{F}^{-1} \tilde{\pi}_k D_\sigma \mathcal{F} L \pi_k \\
&= \pi_k N_\sigma L \pi_k.
\end{aligned}$$

We used that D_σ commutes with $\tilde{\pi}_k$ and $\tilde{\pi}_k^2 = \tilde{\pi}_k$. \square

To proceed in our treatment, we need a bound on the quantity $N\pi_k$, provided by the next lemma.

Lemma 5.23.

$$\|N_\sigma \pi_k\|_{L^1} = \|\pi_k N_\sigma\|_{L^1} \leq (1 + \Gamma_{\sigma,k})$$

Proof. We observe that

$$\pi_k N_\sigma = N_\sigma + (\pi_k - 1)N_\sigma.$$

Then

$$\|\pi_k N_\sigma\|_{L^1} = \|N_\sigma\|_{L^1} + \|(\pi_k - 1)N_\sigma\|_{L^1} \leq 1 + \Gamma_{\sigma,k},$$

where we used

$$\|(1 - \pi_k)N_\sigma\|_{L^1} \leq \|(1 - \pi_k)N_\sigma\|_{L^1 \rightarrow L^\infty}.$$

\square

Thus if we consider f_σ , the fixed point of L_σ this implies the following corollary.

Corollary 5.24.

$$\|\pi_k f_\sigma\|_{L^1} = \|\pi_k N_\sigma L f_\sigma\|_{L^1} \leq (1 + \Gamma_{\sigma,k}) \|f_\sigma\|_{L^1}$$

Lemma 5.25.

$$\|(L_{\sigma,k}^n - L_\sigma^n) f_\sigma\|_{L^\infty} \leq \sum_{i=0}^{n-1} C_i (1 + \Gamma_{\sigma,k} + \|\rho_\sigma\|_\infty) \Gamma_{\sigma,k}$$

Proof. We start by observing:

$$(\pi_k N_\sigma L \pi_k - N_\sigma L) f_\sigma = (\pi_k N_\sigma L \pi_k - N_\sigma L \pi_k + N_\sigma L \pi_k - N_\sigma L) f_\sigma,$$

which gives us

$$(1 - \pi_k) N_\sigma L \pi_k f_\sigma - N_\sigma L (1 - \pi_k) f_\sigma$$

that can be estimated using

$$\|(1 - \pi_k) N_\sigma\|_{L^1 \rightarrow L^\infty} \leq \Gamma_{\sigma,k}$$

which implies

$$\|(1 - \pi_k)N_\sigma L \pi_k f_\sigma\|_{L^\infty} \leq (1 + \Gamma_{\sigma,k}) \Gamma_{\sigma,k},$$

and the fact that

$$\|(1 - \pi_k)f_\sigma\|_{L^1} \leq \|(1 - \pi_k)N_\sigma L f_\sigma\|_{L^\infty} \leq \Gamma_{\sigma,k}$$

which implies that

$$\|N_\sigma L(1 - \pi_k)f_\sigma\|_\infty \leq \|\rho_\sigma\|_\infty \Gamma_{\sigma,k}.$$

With a telescopic argument we obtain

$$L_{\sigma,k}^n - L_\sigma = \sum_{i=0}^{N-1} L_{\sigma,k}^i (L_{\sigma,k} - L_\sigma) L_\sigma^{N-i-1}$$

So, thanks to $L_\sigma f_\sigma = f_\sigma$ we have

$$\left\| \sum_{i=0}^{N-1} L_{\sigma,k}^i (L_{\sigma,k} - L_\sigma) L_\sigma^{N-i-1} f_\sigma \right\|_\infty \leq \sum_{i=0}^{N-1} C_i (1 + \Gamma_{\sigma,k} + \|\rho_\sigma\|_\infty) \Gamma_{\sigma,k}$$

□

Lemma 5.26. *Let f_σ be the unique fixed point of L_σ and $f_{\sigma,k}$ be the unique fixed point of $L_{\sigma,k}$.*

Then

$$\|f_\sigma - f_{\sigma,k}\|_\infty \leq \frac{1}{1 - \alpha} \sum_{i=0}^{N-1} C_i (1 + \Gamma_{\sigma,k} + \|\rho_\sigma\|_\infty) \Gamma_{\sigma,k}$$

Proof. This follows from

$$\|f_\sigma - f_{\sigma,k}\|_\infty \leq \|L_{\sigma,k}^n (f_\sigma - f_{\sigma,k})\|_\infty + \|L_{\sigma,k}^n - L_\sigma^n\|_{L^\infty}.$$

□

So far from now, we obtained an effective way to bound the quantity $\|f_\sigma - f_{\sigma,k}\|_\infty$, which was our purpose. Let us prove, again, a coarse-fine inequality; as we already explained this kind of inequalities are extremely useful during the computations.

Lemma 5.27.

$$\|(L_{\sigma,k}^n - L_{\sigma,k'}^n)N_\sigma\|_{L^1 \rightarrow L^\infty} \leq \sum_{i=0}^{n-1} C_i (1 + \|\rho_\sigma\|_\infty + \Gamma_{\sigma,k'} + \Gamma_{\sigma,k}) \Gamma_{\sigma,k} (1 + \Gamma_{\sigma,k'})^n$$

Proof. We may use again a telescopic decomposition

$$(L_{\sigma,k}^n - L_{\sigma,k'}^n)N_\sigma = \sum_{i=0}^{n-1} L_{\sigma,k}^i (L_{\sigma,k} - L_{\sigma,k'}) L_{\sigma,k'}^{n-i-1} N_\sigma.$$

We already proved that

$$\|(1 - \pi_k)N_\sigma\|_{L^1 \rightarrow L^\infty} \leq \Gamma_{\sigma,k},$$

which in turn implies that

$$\begin{aligned} \|\pi_k N_\sigma f\|_{L^1} &\leq \|N_\sigma f\|_{L^1} + \|(1 - \pi_k)N_\sigma f\|_{L^1} \\ &\leq \|f\|_{L^1} + \|(1 - \pi_k)N_\sigma f\|_{L^\infty} \\ &\leq (1 + \Gamma_{\sigma,k}) \|f\|_{L^1}. \end{aligned}$$

We rewrite now

$$(\pi_{k'} N_\sigma L \pi_{k'})^j N_\sigma = \pi_{k'} N_\sigma (L \pi_{k'} N_\sigma)^j.$$

The computation above implies that

$$\|(L \pi_{k'} N_\sigma)^j\|_{L^1} \leq (1 + \Gamma_{\sigma,k'})^{j+1}.$$

We observe now that $\pi_k \pi_{k'} = \pi_k$. We bound now

$$\begin{aligned} (L_{k,\sigma} - L_{k',\sigma}) \pi_{k'} N_\sigma &= \pi_k N_\sigma L \pi_k N_\sigma - \pi_{k'} N_\sigma L \pi_{k'} N_\sigma \\ &= \pi_k N_\sigma L \pi_k N_\sigma - \pi_{k'} N_\sigma L \pi_k N_\sigma \\ &\quad + \pi_{k'} N_\sigma L \pi_k N_\sigma - \pi_{k'} N_\sigma L \pi_{k'} N_\sigma. \end{aligned}$$

We observe that

$$\|(\pi_k - \pi_{k'}) N_\sigma\|_{L^1} \leq \|(\pi_k - \pi_{k'}) N_\sigma\|_{L^1 \rightarrow L^\infty} \leq \Gamma_{\sigma,k}$$

Then

$$\begin{aligned} \|\pi_k N_\sigma L \pi_k N_\sigma - \pi_{k'} N_\sigma L \pi_k N_\sigma\|_{L^1 \rightarrow L^\infty} &\leq \|(\pi_k - \pi_{k'}) N_\sigma\|_{L^1 \rightarrow L^\infty} \|L\|_{L^1} \|\pi_k N_\sigma\|_{L^1} \\ &\leq \Gamma_{\sigma,k} (1 + \Gamma_{\sigma,k}) \end{aligned}$$

and

$$\begin{aligned}
& \|\pi_{k'} N_\sigma L \pi_k N_\sigma - \pi_{k'} N_\sigma L \pi_{k'} N_\sigma\|_{L^1 \rightarrow L^\infty} \\
& \leq \|\pi_{k'} N_\sigma\|_{L^1 \rightarrow L^\infty} \|L\|_{L^1} \|(\pi_k - \pi_{k'}) N_\sigma\|_{L^1} \\
& \leq (\|N_\sigma\|_{L^1 \rightarrow L^\infty} + \|(1 - \pi_{k'}) N_\sigma\|_{L^1 \rightarrow L^\infty}) \Gamma_{\sigma, k} \\
& \leq (\|\rho_\sigma\|_\infty + \Gamma_{\sigma, k'}) \Gamma_{\sigma, k}
\end{aligned}$$

□

5.2.2/ ERROR BY USING THE DFT

The Fourier transform of course cannot be exactly implemented in a laptop, anyway we may use the Discrete Fourier Transform instead, which provides a fast way (thanks to the FFT algorithm, **Fast Fourier Transform**) to compute an approximation of the Fourier coefficients we want. Thus, we need to take care of the error we are making by using a FFT of order say N , to compute a coefficient of order $k \ll N$. We end this section providing an exponential bound on the so-called **aliasing error**, i.e. the difference between a Fourier coefficient and the same coefficient computed via DFT.

Lemma 5.28. *Let $f \in BV$, let $\{x_i\}_{i=0}^{N-1}$ be equispaced points on S^1 ; consider the two coefficients*

$$X_k = \sum_{n=0}^{N-1} f\left(\frac{n}{N}\right) e^{-2\pi i k n / N}$$

and

$$\hat{f}(k) = \mathcal{F} f(x) e^{-2\pi i k x}$$

Then we have that

$$|X_k - \hat{f}(k)| \leq 2 \sum_{p=1}^{+\infty} e^{-\frac{(k+pN)^2 \sigma^2}{2}} \leq \Gamma_{\sigma, N} \quad (5.7)$$

Proof.

$$\begin{aligned}
X_k &= \frac{1}{N} \sum_{n=0}^{N-1} \sum_{j=-\infty}^{\infty} \hat{f}(j) e^{2j\pi i n / N} e^{-2k\pi i n / N} \\
&= \frac{1}{N} \sum_{j=-\infty}^{\infty} \left(\sum_{i=0}^{N-1} e^{2j\pi i n (j-k) / N} \right)
\end{aligned}$$

Now, the sum into parenthesis is different from 0 only when $j - k$ is a multiple of N .

Therefore

$$X_k = \sum_{p=-\infty}^{+\infty} \hat{f}(k + pN) = \hat{f}(k) + \sum_{p=1}^{+\infty} \hat{f}(k + pN) + \hat{f}(k - pN).$$

We can choose N such that $\hat{f}(k + N)$ is really small, since our Fourier coefficients decay fast. We have indeed

$$|X_k - \hat{f}(k)| \leq 2 \sum_{p=1}^{+\infty} e^{\frac{-(k+pN)^2 \sigma^2}{2}} \leq \Gamma_{\sigma, N}$$

□



EXPERIMENTS AND RESULTS

APPLICATIONS

In this chapter and in the last section of chapter 7 we will apply the rigorous schemes of approximation described so far, to three significant examples. The first two dynamical systems under consideration belong to the family of the so-called **Arnold Maps** ([44],[48]), and they come from two different natural phenomena; indeed Arnold's standard circle maps are widely used to study the quasi-periodic route to chaos. We will use the first one as an introductory example to the Ulam method, explaining how the theory made so far, is being applied in a concrete example. We will turn then to a more complicated example, in which rigorous computations, co-aided by linear response theory, will lead to non-trivial results, such as the differentiability of the rotation number. Arnold's standard circle maps are widely used to study the quasi-periodic route to chaos and other phenomena associated with nonlinear dynamics in the presence of two rationally unrelated periodicities.

Finally in the last example we apply the Fourier scheme of approximation (section 5.2) to a perturbed doubling map; as we said before this is a work in progress and many, more meaningful, other examples will be produced in the future. Still, even in this easy but non-trivial case, we can appreciate the power of the Fourier method, obtaining bounds of the order of 10^{-80} . We choose to report this last example at the end of the numerical discussion of the Fourier algorithm, 7.2.3.

6.1/ ASSOCIATIVE MEMORY IN THE HIPPOCAMPUS AS RANDOM DYNAMICAL SYSTEM

The first example that we want to study is a random dynamical system T_ξ , with a deterministic component given by an Arnold circle map (see Figure 6.2) and a stochastic part given by a random additive noise, defined as

$$x_{n+1} = T(x_n) + \xi_n \pmod{1}, \text{ where } T(v) = v + A \sin(4\pi v) + C, \quad (6.1)$$

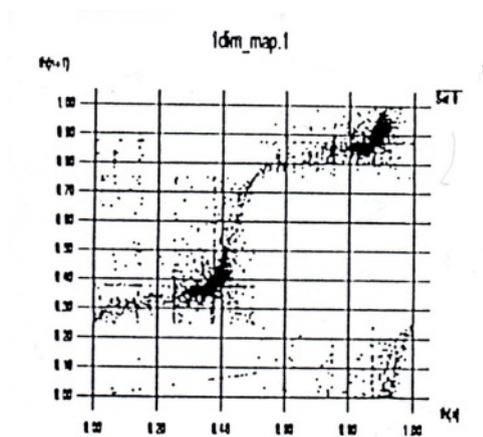


Figure 6.1: Data collected in [33].

for $A = 0.08$, $C = 0.1$ and ξ_n an i.i.d. sequence of random variables with a distribution assumed uniform over $[-\xi, \xi]$.

This map was proposed in [33; 57] as a model of the evolution of a macrovariable, related to the activity of an asynchronous neural network in absence of external stimulus, related to a rule of successive association of memory. Indeed a macrovariable, related to the “activity” of the network (see Figure 6.1), was observed to evolve as a noisy one dimensional map in the case that the network receives no external stimulus (its definition of can be found in [33] (p. 6).). This was regarded in [57] as related to a rule of successive association of memory, exhibiting chaotic dynamics.

We recall that the Perron-Frobenius operator associated to this system is given by $L_\xi = N_\xi L$, where N_ξ is a convolution operator and L is the Perron-Frobenius operator of T .

$$N_\xi f(t) = \xi^{-1} \int_{-\xi/2}^{\xi/2} f(t - \tau). \quad (6.2)$$

As explained in the lasts chapters, this operator admits an invariant probability density in L^1 , which can be rigorously approximated. We want to compute, up to some error, this density, using lemma 5.7, lemma 5.9 and theorem 5.12; we recall that in these statements there are bounds which depend from properties of the discretized operator, e.g. its rate of mixing. Since the discretized operator is a matrix, we can rigorously compute via interval arithmetics these quantities, leading to rigorous bounds on the error on the computed stationary density (lemma 5.7). All the algorithm and the numeric details are reported in the next chapter.

If we apply this **black box** described so far, we obtained data summarized in Table 6.1, 6.2; in figure 6.3d, 6.4d, you may see the plot of the computed stationary densities. In the tables,

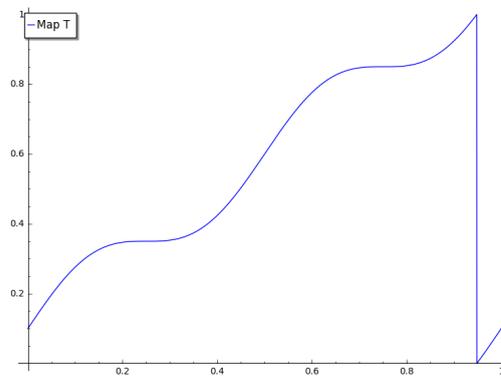


Figure 6.2: Deterministic component of Equation (6.1).

- ξ is the size of the noise
- n_{contr} and α_{contr} are such that $\|L_{\delta', \xi}^{n_{\text{contr}}}\| \leq \alpha_{\text{contr}}$, these are the bounds obtained on a coarser partition δ'
- via coarse-fine strategy we obtain the bounds for the finer partition: α , and the related $\sum C_i$
- finally $\|l_{\text{apriori}}$ and $\|l_{\text{err}}$, are the bounds on the L^1 error.

Remark 6.1. The images below show a rigorous estimate of the density of probability of visits of typical trajectories near the attractors, showing that this is relatively high with respect to the density of probability of visits in other parts of the space. In particular we can observe that the trajectories concentrate in certain “weakly attracting” and “low chaotic” regions of the space.

The presence of these “weakly attracting” and “low chaotic” regions is in concordance with the concept of **chaotic itinerancy**, a concept which still did not have a complete mathematical formalization. The results presented in this section aim to bring some light on this concept, which is so widely used in the applied literature: chaotic itinerancy is used to refer to a dynamical behavior in which typical orbits visit a sequence of regions of the phase space called “quasi attractors” or “attractor ruins” in some irregular way. This behavior was observed in several models and experiments related to the dynamics of neural networks and related to neurosciences. We invite the reader to consult [58] for a wider introduction to the subject and to its literature.

ξ	n_{contr}	α_{contr}	α	$\sum C_i$	llapriori	llerr
0.732×10^{-1}	126	0.027	0.05	56.64	0.313×10^{-2}	0.715×10^{-4}
0.610×10^{-1}	167	0.034	0.067	78.66	0.530×10^{-2}	0.105×10^{-3}
0.488×10^{-1}	231	0.051	0.1	120.56	0.106×10^{-1}	0.184×10^{-3}
0.427×10^{-1}	278	0.068	0.14	156.45	0.163×10^{-1}	0.268×10^{-3}
0.366×10^{-1}	350	0.087	0.19	213.17	0.273×10^{-1}	0.432×10^{-3}
0.305×10^{-1}	453	0.12	0.26	307.03	0.523×10^{-1}	0.813×10^{-3}
0.275×10^{-1}	532	0.14	0.32	380.64	0.776×10^{-1}	0.122×10^{-2}
0.244×10^{-1}	596	0.19	0.41	467.70	0.124	0.202×10^{-2}

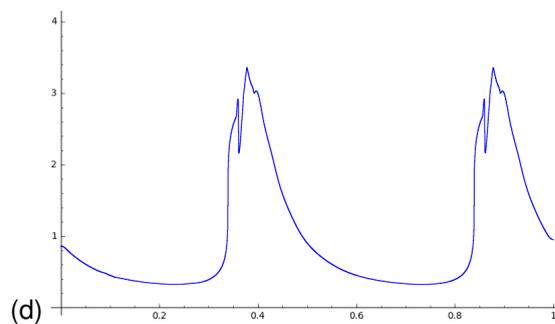
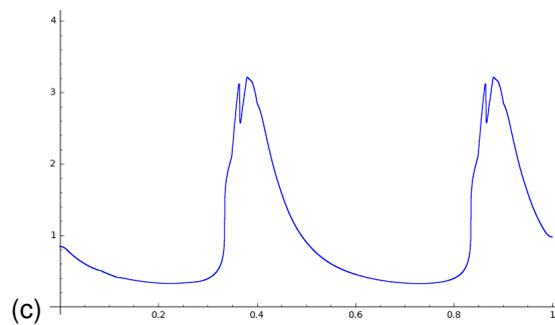
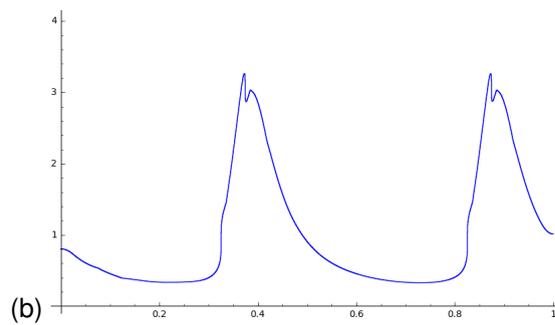
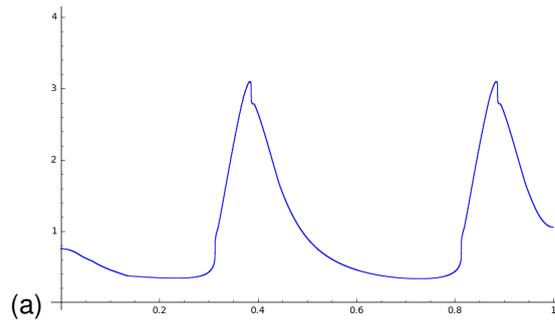
$\delta = 2^{-19}$ is used to compute the invariant density via power method; $\delta_{\text{contr}} = 2^{-14}$, is used to find the estimates on the contraction rate and $\delta_{\text{est}} = 2^{-12}$ is used to estimate the L^1 error of the invariant density

Table 6.1: Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ , with parameters $A = 0.8$, $C = 0.01$.

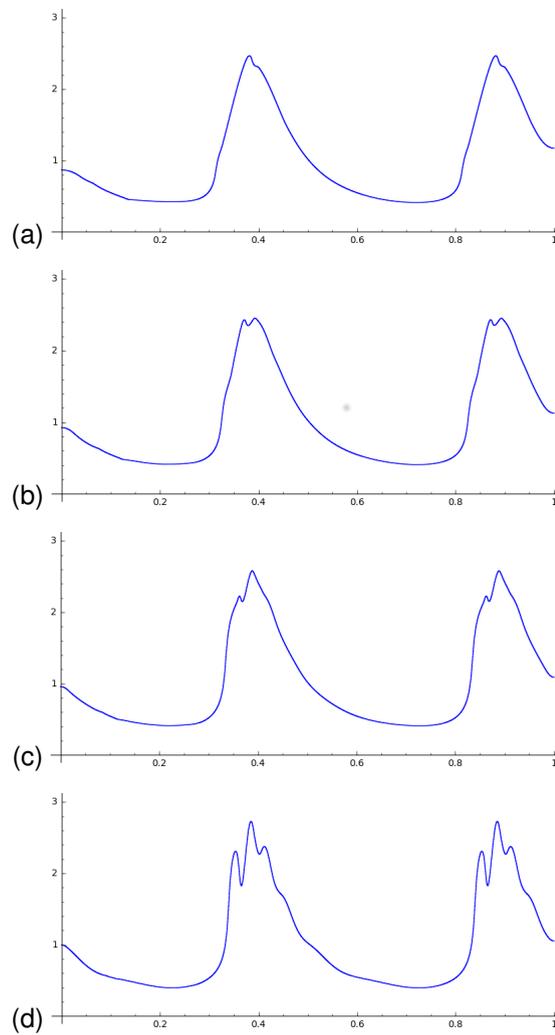
ξ	n_{contr}	α_{contr}	α	$\sum C_i$	llapriori	llerr
0.732×10^{-1}	183	0.03	0.059	83.57	0.466×10^{-2}	0.255×10^{-4}
0.610×10^{-1}	237	0.046	0.089	119.31	0.822×10^{-2}	0.282×10^{-4}
0.488×10^{-1}	332	0.069	0.14	186.80	0.170×10^{-1}	0.323×10^{-4}
0.427×10^{-1}	406	0.087	0.18	244.95	0.267×10^{-1}	0.358×10^{-4}
0.366×10^{-1}	494	0.12	0.25	330.89	0.459×10^{-1}	0.419×10^{-4}
0.305×10^{-1}	500	0.3	0.46	419.92	0.974×10^{-1}	0.646×10^{-4}
0.275×10^{-1}	596	0.32	0.52	517.97	0.151	0.807×10^{-4}
0.244×10^{-1}	600	0.49	0.73	573.04	0.326	0.189×10^{-3}

$\delta = 2^{-19}$ is used to compute the invariant density via power method; $\delta_{\text{contr}} = 2^{-14}$, is used to find the estimates on the contraction rate and $\delta_{\text{est}} = 2^{-12}$ is used to estimate the L^1 error of the invariant density

Table 6.2: Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ , with parameters $A = 0.8$, $C = 0.01$.



Approximated stationary densities $f_{\xi, \delta}$ for T_{ξ} , with $\delta = 2^{-19}$ and $A = 0.08$. **(a)** $\xi = 0.732 \times 10^{-1}$; **(b)** $\xi = 0.488 \times 10^{-1}$; **(c)** $\xi = 0.305 \times 10^{-1}$; **(d)** $\xi = 0.214 \times 10^{-1}$. See remark 6.1.



Approximated stationary densities $f_{\xi, \delta}$ for T_{ξ} , with $\delta = 2^{-19}$ and $A = 0.07$. (a) $\xi = 0.732 \times 10^{-1}$; (b) $\xi = 0.488 \times 10^{-1}$; (c) $\xi = 0.305 \times 10^{-1}$; (d) $\xi = 0.214 \times 10^{-1}$. See remark 6.1.

6.2/ THE ROTATION NUMBER OF THE ARNOLD MAP

The El Niño–Southern Oscillation (ENSO) phenomenon is a crucial component of climate variability on interannual time scales and it is dominated by the seasonal cycle, on the one hand, and an intrinsic oscillatory instability with a period of a few years, on the other. The role of meteorological phenomena on much shorter time scales, such as westerly wind bursts, has also been recognized and modeled as additive noise. Arnold’s standard circle maps are widely used to study this kind of phenomena, in which two rationally unrelated periodicities are in competition; they are defined in the following way: let $T_{\tau,\epsilon} : S^1 \rightarrow S^1$ be the deterministic circle map defined by

$$T_{\tau,\epsilon}(x) := x + \tau - \frac{\epsilon}{2\pi} \sin(2\pi x) \pmod{1}. \quad (6.3)$$

Here $\tau := 2\pi/\omega$ and $\omega := \omega_i/\omega_e$ is the driving frequency, while $\epsilon \geq 0$ parameterizes the magnitude of nonlinear effects. By Arnold map with additive noise we mean the stochastic process $\{X_n\}_{n \in \mathbb{N}}$ on S^1 defined by

$$X_{n+1} = T_{\tau,\epsilon}(X_n) + \Omega_n \pmod{1}. \quad (6.4)$$

At each iterate of $T_{\tau,\epsilon}(x)$, an independent identically distributed (i.i.d.) noise Ω_n , that is uniformly distributed on $[-\xi/2, \xi/2]$, is added to the deterministic term on the right-hand side of equation 6.4; in particular, the noise is independent of the point $X_n \in S^1$.

In the case $\epsilon = 0, \xi = 0$, the system is simply a rotation of the circle. In the deterministic case, where $\xi = 0$, and $\epsilon \in (0, 1)$ we get the classical Arnold circle map, which is one of the most used models in several scientific areas: models of coupled oscillators [60; 38], model for cardiac arrhythmias [61; 28] or sleep-wake regulation cycle [7].

In the deterministic case, one of the most remarkable feature of this model is the **mode-locking** phenomenon, but let us first introduce the concept of **rotation number**.

Definition 6.2. The rotation number $\rho = \rho_\tau$ measures the average rotation per iterate of 6.3 on S^1 and is defined as

$$\rho_\tau := \lim_{n \rightarrow \infty} \frac{\tilde{T}_{\tau,\epsilon}^n(x)}{n}$$

where $\tilde{T}_{\tau,\epsilon}^n : \mathbb{R} \rightarrow \mathbb{R}$ is the lift of $T_{\tau,\epsilon}$ to \mathbb{R} , defined by

$$\tilde{T}_{\tau,\epsilon}(x) := x + \tau - \frac{\epsilon}{2\pi} \sin(2\pi x). \quad (6.5)$$

We need to adapt this definition to our random context, by considering a stochastic pro-

cess \tilde{X} on \mathbb{R} defined by $\tilde{X}_{n+1} = \tilde{T}_{\tau,\epsilon}(\tilde{X}_n) + \Omega_n$ and

$$\rho_\tau := \lim_{n \rightarrow \infty} \frac{\tilde{X}_n}{n}.$$

Let us recall two general classical results about the properties of the rotation number, in the particular case of deterministic and orientation-preserving diffeomorphism of the circle; for further details we refer to [30; 37].

- the rotation number is independent of $x \in S^1$.
- the Denjoy theorem : it states that an orientation-preserving diffeomorphism (for our maps $T_{\tau,\epsilon}$ this corresponds to the weakly nonlinear case $\epsilon < 1$) having an irrational rotation number is topologically conjugate to an irrational rotation in S^1 (see [37; 62]).

Something surprising happens around the rational values of the driving frequency: the rotation number is locally constant. In other words: the map "rotation number vs driving frequency" is a devil's staircase (see Fig. 6.5 for one example). Furthermore the graph of the rotation number seems to go through a "smoothing" process, i.e the map $\tau \mapsto \rho_\tau$ becomes smooth.

In the case $\epsilon \in (0, 1)$ (what we called weak nonlinearity), this was rigorously proved in [64]. Hence we may ask whether or not the map ρ_τ admits a derivative with respect to the parameter τ . A classical result asserts that this derivative is defined in a point τ_0 when the corresponding rotation number is irrational ($\rho_{\tau_0} \notin \mathbb{Q}$) [30]); other results can be found in [49],[46].

If $\epsilon > 1$, i.e. in the case of strong nonlinearity, the methods of [64] cannot work. As in the diffeomorphism case we can numerically observe that the action of the noise makes the rotation number smoother. In [48], with the methods illustrated so far from now, it was proved that

- the rotation number is differentiable at every value τ for which the system is mixing (and the existence of intervals for which this assumption is verified is proved as well); this was done by using linear response theory, in particular theorem 4.21 and using the Ulam algorithm for the rigorous computation of the rate of mixing
- the rotation number is not always monotonic. Here a certified approximation of the stationary measure, with a small error on L^1 , leads to a certified estimate on the rotation number.

In the next sections we are going to review these results, showing the effectiveness and the power of our rigorous computations.

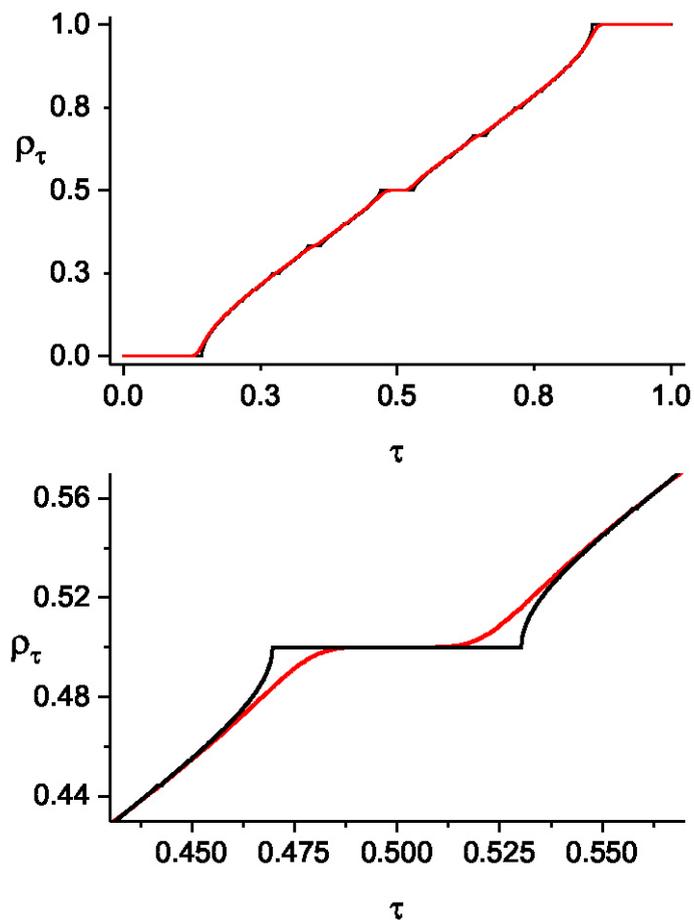
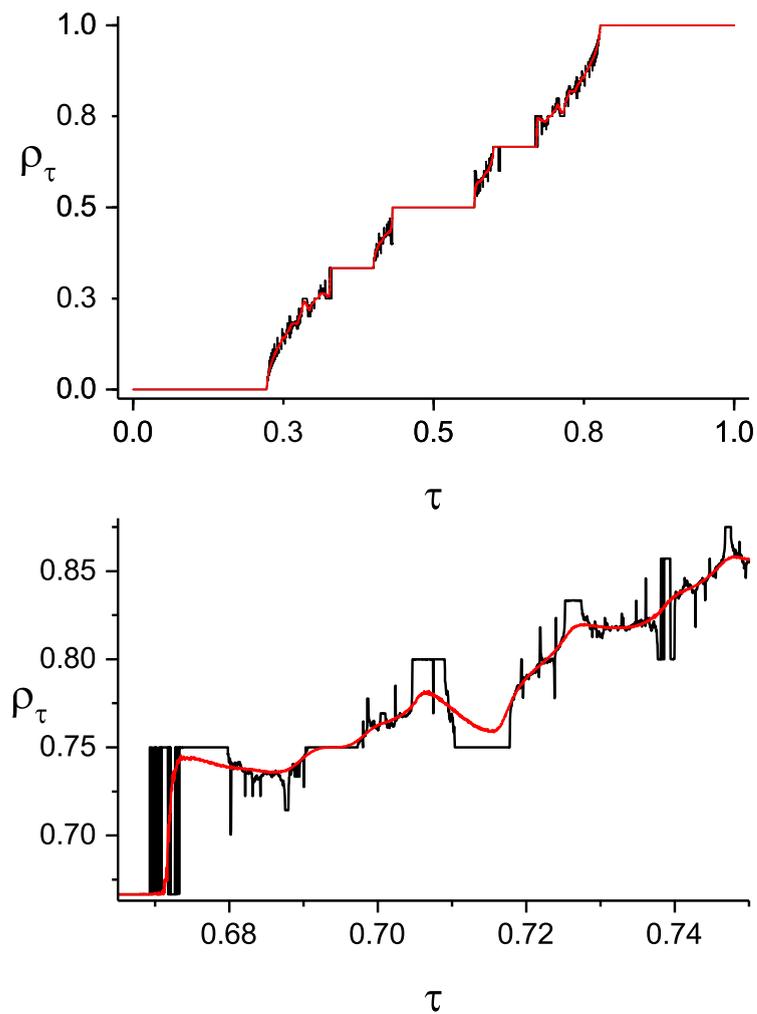


Figure 6.5: Plot of the rotation number $\rho = \rho_\tau$ for $\epsilon = 0.9$. The black line corresponds to the absence of uniformly distributed noise, while the red line shows this dependence in the presence of such noise, with amplitude $\xi = 0.05$. Recall that $\tau = 2\pi/\omega$, where ω is the driving frequency. The lower panel is a blow-up of the dependence near the value $\tau = 0.5$.

Figure 6.6: Same as Fig. 6.5 but for $\epsilon = 1.4$ and for noise amplitude $\xi = 0.01$. Here the lower panel is a blow-up of the dependence near $\tau = 0.7$.



6.3/ DIFFERENTIABILITY OF THE ROTATION NUMBER FOR STRONG NONLINEARITY

Adapting a general linear response result, theorem 4.21, we will show that the rotation number is differentiable with respect to the parameter τ . We will use a computer assisted proof to compute an interval in which the assumption of theorem 4.21 are satisfied. One of the keys is that the linear response with respect to $\|\cdot\|_W$ is sufficient to deduce the smoothness of the rotation number, because it has an associated observable which is Lipschitz. Let us start by recalling that we are considering a family of operators L_δ , $\delta \in [0, \bar{\delta}]$, i.e. the transfer operators associated with the Arnold maps with additive noise; as already pointed out the transfer operator associated to the system with additive noise $L_\delta : L^1 \rightarrow L^1$ is given by

$$[L_\delta f](t) = [\rho_\xi * L_{T_\delta}(f)](t). \quad (6.6)$$

We start now to verify the assumption of theorem 4.21, this is done in the following two lemmas. Assumption LR1 will be computationally verified, and we are going to show it at the end of the section.

Lemma 6.3. *The limit defined at (4.6) exists in BS and the limit converges in the W-norm.*

$$\lim_{\delta \rightarrow 0} \left\| \frac{(L_0 - L_\delta)}{\delta} f_0 - \frac{[\delta_{-\xi} - \delta_\xi]}{2\xi} * L_{T_0}(f_0) \right\|_W = 0.$$

Proof. We remark that $L_{T_0}(f_0) \in L^1(S^1)$ furthermore

$$L_0 f_0(t) = [\rho_\xi * L_{T_0}(f_0)](t) = \int_{\mathbb{S}^1} \rho_\xi(t - \tau) [L_{T_0}(f_0)](\tau) d\tau$$

and

$$\begin{aligned} L_\delta f_0(t) &= [\rho_\xi * L_{T_\delta}(f_0)](t) = \int_{\mathbb{S}^1} \rho_\xi(t - \tau) [L_{T_0}(f_0)](\tau - \delta) d\tau \\ &= \int_{\mathbb{S}^1} \rho_\xi(t - c - \delta) [L_{T_0}(f_0)](c) dc \end{aligned}$$

hence

$$\left[\frac{L_\delta - L_0}{\delta} f_0 \right](x) = \int_{\mathbb{S}^1} \frac{\rho_\xi(t - \delta - \tau) - \rho_\xi(t - \tau)}{\delta} [L_{T_0}(f_0)](\tau) d\tau. \quad (6.7)$$

Denoting by $R_\delta : L^1 \rightarrow L^1$ the translation operator given by $[R_\delta f](x) = f(x - \delta)$. It holds that

$$\frac{L_\delta - L_0}{\delta} f_0 = \frac{[R_{-\delta} \rho_\xi - \rho_\xi]}{\delta} * [L_{T_0}(f_0)] \quad (6.8)$$

but

$$\lim_{\delta \rightarrow 0} \left\| \frac{[R_{-\delta} \rho_\xi - \rho_\xi]}{\delta} - \frac{[\delta_{-\xi} - \delta_\xi]}{2\xi} \right\|_W = 0$$

where $\delta_{-\xi}$ and δ_ξ are the delta measures placed on $\pm\xi$.

Then by (4.3)

$$\lim_{\delta \rightarrow 0} \frac{L_0 - L_\delta}{\delta} f_0 = \frac{[\delta_{-\xi} - \delta_\xi]}{2\xi} * L_{T_0}(f_0)$$

(recall that T_0 is nonsingular and then $L_{T_0}(f_0) \in L^1$) with convergence in the $\|\cdot\|_W$ norm. \square

Lemma 6.4. *The remaining assumptions of Item LR3 of Theorem 4.21 are satisfied:*

$$\|L_0 - L_\delta\|_{L^1 \rightarrow (BS, \|\cdot\|_W)} \leq K\delta, \text{ and } \|L_0 - L_\delta\|_{BV \rightarrow V} \leq K\delta.$$

Proof. Let $f \in L^1$. By (6.8) it holds

$$[L_\delta - L_0]f = [R_{-\delta} \rho_\xi - \rho_\xi] * [L_{T_0}(f)].$$

Since there is a K such that $\|[R_{-\delta} \rho_\xi - \rho_\xi]\|_{L^1} \leq K\delta$ by Lemmas 4.17 and 4.18 we directly get the statement. \square

These estimates are leading to the following linear response statement adapted to our situation.

Proposition 6.5. *Let $T_0 : S^1 \rightarrow S^1$ be a nonsingular map. Let T_δ defined as $T_\delta(x) = T_0(x) + \delta$, let $L_\delta : L^1 \rightarrow L^1$ be the transfer operator defined as in (6.6). Let $f_\delta \in L^1$ be such that $L_\delta f_\delta = f_\delta$ (a stationary measure for the system L_δ).*

Suppose L_0 is mixing: for every $g \in BV[0, 1]$ with $\int_1 g \, dm = 0$, then

$$\lim_{n \rightarrow \infty} \|L_0^n g\|_{L^1} = 0.$$

(see Assumption LR1 of Theorem 4.21) Then $(Id - L_0)^{-1}$ is a continuous operator on the space of zero average Borel measures equipped with the $\|\cdot\|_W$ norm and

$$\lim_{\delta \rightarrow 0} \left\| \frac{f_\delta - f_0}{\delta} - (Id - L_0)^{-1} \left[\frac{[\delta_{-\xi} - \delta_\xi]}{2\xi} * L_{T_0}(f_0) \right] \right\|_W = 0. \quad (6.9)$$

To conclude our proof, we still need to exploit the fact that the rotation number can be computed as the integral of a Lipschitz observable with respect to the stationary measure;

moreover we need to verify that there exists an interval on which assumption LR1 of theorem 4.21 is true. For the first fact, we report the following proposition from [48], but this result was already in [64].

Proposition 6.6. *Let $T_{\tau,\epsilon}$ be the Arnold map with parameters (τ, ϵ) and uniformly distributed noise of size ξ , suppose the system satisfies the assumption LR1, let μ_τ be the corresponding stationary measure and ρ_τ be the associated rotation number. Then*

$$\rho_\tau(\omega, x) = \int_{\mathbb{S}^1} \varphi_\tau d\mu_\tau. \quad (6.10)$$

In particular, ρ_τ is $\mathbb{P} \otimes \mu_\tau$ almost surely constant.

Hence, so far from now, we have shown that

Corollary 6.7. *The rotation number of the Arnold maps with uniformly distributed additive noise is differentiable at each value of the parameter τ for which the associated system is mixing (in the sense stated in assumption LR1 and Proposition 6.5). Furthermore if τ_0 is such a parameter we get the following formula for the derivative of the rotation number computed at τ_0*

$$\left[\frac{d}{d\tau} \rho_\tau \right](\tau_0) = 1 + \int_{\mathbb{S}^1} \varphi_{\tau_0} d[(Id - L_{T_{\tau_0, \epsilon}})^{-1} \frac{[\delta_{-\xi} - \delta_\xi]}{2\xi} * L_{T_{\tau_0, \epsilon}}(\mu_{\tau_0})] \quad (6.11)$$

where $L_{T_{\tau_0, \epsilon}}$ is the pushforward operator of the map $T_{\tau_0, \epsilon}$.

Proof. By Proposition 6.6 the rotation number is the integral of a Lipschitz observable. Considering the increment of ρ_τ we get

$$\begin{aligned} \frac{\rho_\tau(\tau_0 - h) - \rho_\tau(\tau_0)}{h} &= \frac{1}{h} \left[\int_{\mathbb{S}^1} \varphi_{\tau_0+h} d\mu_{\tau_0+h} - \int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0} \right] \\ &= \frac{1}{h} \left[\int_{\mathbb{S}^1} \varphi_{\tau_0+h} d\mu_{\tau_0+h} - \int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0+h} \right] \\ &\quad + \frac{1}{h} \left[\int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0+h} - \int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0} \right]. \end{aligned}$$

Here

$$\frac{1}{h} \left[\int_{\mathbb{S}^1} \varphi_{\tau_0+h} d\mu_{\tau_0+h} - \int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0+h} \right] = \int_{\mathbb{S}^1} \frac{\varphi_{\tau_0+h} - \varphi_{\tau_0}}{h} d\mu_{\tau_0+h} = 1$$

and

$$\frac{1}{h} \left[\int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0+h} - \int_{\mathbb{S}^1} \varphi_{\tau_0} d\mu_{\tau_0} \right] = \int_{\mathbb{S}^1} \varphi_{\tau_0} d\left[\frac{\mu_{\tau_0+h} - \mu_{\tau_0}}{h} \right]$$

and the statement directly follows from Proposition 6.5. \square

We will now compute an explicit example of case in which the system is mixing and (6.11) holds. As already pointed out in chapter 4 when considering a family of systems (as in

this case), it is not sufficient to provide the rate of mixing of just one of them. This is why we provided results about the rate of mixing of a perturbed system (section 4.4), and we are going now to use them: we use the Ulam-based strategy of chapter 5 to verify assumption *LR1*, i.e to find n and α such that

$$\|L_\xi^n\|_{V \rightarrow L^1} \leq \alpha < 1, \quad (6.12)$$

and we take advantage of section 4.4 to deduce that this assumption is satisfied by nearby systems.

Proposition 6.8. *Let $\epsilon = 1.4$ and $\xi = 0.1$. Then, for each $\tau \in [0.75, 0.8]$, the corresponding Arnold map with noise and parameters (ϵ, τ, ξ) satisfies assumption *LR1* of Theorem 4.21.*

Proof. Suppose we have proved the mixing for a system with parameters $(\epsilon_0, \tau_0, \xi_0)$ and we have α and n such that (6.12) is satisfied, then (4.9) implies that there exists $\theta_0 > 0$ and a whole interval $I_0 = [\tau_0 - \theta_0, \tau_0 + \theta_0]$, such that all the systems with parameters $(\epsilon_0, \tau, \xi_0)$, $\tau \in I_0$, are still mixing. Moreover (4.9) gives an explicit formula for θ_0 which depends from α and n . These constants are explicitly computable, hence we can explicitly compute θ_0 .

To show that the mixing property holds for every system of parameters $(1.4, \tau, 0.1)$, with $\tau \in [0.75, 0.8]$, a strategy is to consider a finite sequence of points $\{\tau_i\} \subset [0.75, 0.8]$ such that the systems with parameters $(1.4, \tau_i, 0.1)$ are mixing, for every i , and the associated intervals $I_i := [\tau_i - \theta_i, \tau_i + \theta_i]$, defined as above, cover the interval $[0.75, 0.8]$.

In Table 6.3 we show the computer aided estimates about the values of θ_0 by the method described above for each example. As it can be seen, since the union of all this computed intervals is equal to (a, b) , with $a = 0.749399418088000$ and $b = 0.800715949198087$, we have then proved the desired property in the whole interval $[0.75, 0.8]$. \square

Once we have Assumption *LR1* of Theorem 4.21 satisfied for this family of systems, applying Corollary 6.7 we directly get

Corollary 6.9. *Let $\epsilon = 1.4$, $\xi = 0.1$ then, for each $\tau \in [0.75, 0.8]$ the rotation number corresponding to the Arnold map with noise and parameters (ϵ, τ, ξ) is differentiable as τ varies and (6.11) holds.*

Table 6.3: Given the Arnold map with noise of magnitude ξ and parameters (τ_0, ϵ_0) , for which we have already proved mixing, the table shows the computed intervals $I_0 = [\tau_0 - \theta_0, \tau_0 + \theta_0]$, such that if $\tau \in I_0$ then the Arnold map with parameters (τ, ϵ_0) is mixing

(τ_0, ϵ_0)	ξ	$[\tau_0 - \theta_0, \tau_0 + \theta_0]$
(0.7502, 1.4)	0.1	[0.749399418088000, 0.751000581912001]
(0.7516, 1.4)	0.1	[0.750823157412727, 0.752376842587275]
(0.7532, 1.4)	0.1	[0.752340409081139, 0.754059590918862]
(0.7548, 1.4)	0.1	[0.753967757739070, 0.755632242260931]
(0.7564, 1.4)	0.1	[0.755472331622558, 0.757327668377443]
(0.7582, 1.4)	0.1	[0.757304135181951, 0.759095864818050]
(0.7598, 1.4)	0.1	[0.758934791149505, 0.760665208850496]
(0.7616, 1.4)	0.1	[0.760590252726379, 0.762609747273622]
(0.7632, 1.4)	0.1	[0.762227848809647, 0.764172151190354]
(0.7648, 1.4)	0.1	[0.763863841059881, 0.765736158940120]
(0.7668, 1.4)	0.1	[0.765692173699252, 0.767907826300749]
(0.7688, 1.4)	0.1	[0.767737148314721, 0.769862851685280]
(0.7708, 1.4)	0.1	[0.769780384398816, 0.771819615601185]
(0.7728, 1.4)	0.1	[0.771572710558602, 0.774027289441399]
(0.7748, 1.4)	0.1	[0.773628600754358, 0.775971399245643]
(0.7768, 1.4)	0.1	[0.775681672076595, 0.777918327923406]
(0.7788, 1.4)	0.1	[0.777626636524504, 0.779973363475497]
(0.7812, 1.4)	0.1	[0.779896782767258, 0.782503217232743]
(0.7836, 1.4)	0.1	[0.782295331693974, 0.784904668306028]
(0.7860, 1.4)	0.1	[0.784762771410257, 0.787237228589744]
(0.7884, 1.4)	0.1	[0.787166089347180, 0.789633910652821]
(0.7908, 1.4)	0.1	[0.789629611581057, 0.791970388418944]
(0.7928, 1.4)	0.1	[0.791684956386327, 0.793915043613674]
(0.7948, 1.4)	0.1	[0.793685568440505, 0.795914431559496]
(0.7968, 1.4)	0.1	[0.795685318203188, 0.797914681796813]
(0.7988, 1.4)	0.1	[0.797684476459441, 0.799915523540560]
(0.7996, 1.4)	0.1	[0.798484050801914, 0.800715949198087]

Table 6.4: This table shows the computed intervals in which the rotation number lies for each value of the parameters in consideration.

(τ, ϵ)	ξ	ρ_τ
(0.707, 1.4)	0.01	[0.780594, 0.780604]
(0.708, 1.4)	0.01	[0.778348, 0.778361]
(0.709, 1.4)	0.01	[0.775291, 0.775302]
(0.710, 1.4)	0.01	[0.771833, 0.771844]
(0.711, 1.4)	0.01	[0.768335, 0.768348]
(0.712, 1.4)	0.01	[0.765170, 0.765183]
(0.713, 1.4)	0.01	[0.762568, 0.762590]
(0.714, 1.4)	0.01	[0.760585, 0.760612]
(0.715, 1.4)	0.01	[0.759288, 0.759344]
(0.716, 1.4)	0.01	[0.759915, 0.759970]

6.3.1/ NON-MONOTONIC ROTATION NUMBER FOR STRONG NONLINEARITY

For $\epsilon < 1$, the monotonicity of the rotation number of the Arnold map is a classical result: it is well known that if $\tau_2 > \tau_1$ then $\rho_{\tau_2} > \rho_{\tau_1}$ (see [62]). This is false if $\epsilon > 1$ and we are going to show it now for a particular example with $\epsilon = 1.4$ and $\xi = 0.01$.

Again we can rigorously approximating the value of the rotation number for several values of τ , since it can be computed as the integral of a Lipschitz observable with respect to the (rigorously) computed stationary measure. Indeed, there exist $\tau_1 < \tau_2$ with corresponding rotation numbers $\rho_1 \in I_1, \rho_2 \in I_2$, where I_1 and I_2 are the rigorous computed intervals in which the rotation numbers lie, and furthermore, these intervals are such that $\max(I_2) < \min(I_1)$. We conclude that ρ_2 must be smaller than ρ_1 .

Proposition 6.10. *Let $\epsilon = 1.4$ and $\xi = 0.01$, then the rotation number ρ_τ , as function of the parameter τ , is not monotonic in the interval $[0.707, 0.716]$.*

Proof. As explained above, we use the Ulam based algorithm to estimate the stationary measure for $\epsilon = 1.4$ and for each $\tau \in \{0.707, 0.708, \dots, 0.716\}$. We estimate the expected value of the observable $\tilde{T}_{\tau, \epsilon}(x) - x$ with respect to the stationary measure for each example. This gives a certified interval in which the rotation number ρ_τ of each example lies (see Proposition 6.6). The results are reported in Table 6.4. The inspection of these, disjoint, decreasing, intervals shows that the rotation number decreases for $\tau \in \{0.707, 0.708, \dots, 0.715\}$. The last estimate at $(0.716, 1.4)$ shows an increasing behavior, showing non monotonicity. \square

NUMERICAL ASPECTS

In this last chapter we discuss the computational details associated to the theory developed so far; we will try to make this part of the thesis accessible, as far as possible, even to those who do not have a mathematical background and are more interested in computer science aspects, or simply want to learn how to use the code we will present without being able or need to go into the details of the theory.

There are two sections, dedicated respectively to the approximation scheme based on Ulam, and the one based on Fourier. We have already seen some theoretical differences, but the algorithms differ, quite substantially, even in numerical aspects, for instance the matrices of Ulam are sparse, those of Fourier are not. We will focus more on the exposure of the second method, for several reasons:

- the Ulam-based strategy, which is at the basis of this work, has already been described in full details in [23],[24] and then in [44], [48].
- the Ulam-based algorithm, used for example to obtain the results in Chapter 7, is implemented in SageMath¹, meanwhile we have chosen to work with Julia² with the new Fourier-based algorithm.

Before moving on to the description of the two algorithms, let us introduce Julia and the main packages that were used during this work, and the concept of **interval arithmetic**, which is the basis of our rigorous calculations.

Working at the crossroad of mathematics, computer science and applied sciences, it can sometimes be difficult to find the right programming language to accomplish your tasks. In the author's opinion, an excellent compromise between facility of use, versatility, speed and community presence has been achieved by Julia in the last years, the programming

¹SAGE, "System for Algebra and Geometry Experimentation", is a computer algebra system with features covering many aspects of mathematics; the first version of SageMath was released on 24 February 2005

²Julia is a high-level, high-performance, dynamic programming language. While it is a general-purpose language and can be used to write any application, many of its features are well suited for numerical analysis and computational science.

language in which the Fourier approximation method we present has been implemented. The packages used in this work are

- LinearAlgebra.jl, Plots.jl
- IntervalArithmetic.jl, IntervalRootFinding.jl
- ArbNumerics.jl

LinearAlgebra.jl and Plots.jl, implement the obvious basic functions we expect. We will discuss ArbNumerics.jl directly in section 2, while as you can guess from the name, IntervalArithmetic.jl and IntervalRootFinding.jl implement the aforementioned Interval Arithmetic. To explain what it is, let's start with an example from ([51]), first given by Rump.

Example 7.1. Let

$$f = 333.75b^6 + a^2(11a^2b^2) - b^6 - 121b^4 - 2) + 5.5b^8 + \frac{a}{2b}$$

with $a = 77617.0$ and $b = 33096.0$.

If we compute the powers by successive multiplications using, single, doubling or extended precision (this was done by Rump on an IBM 370 system), we obtain

- **single precision:** $f = 1.17260361 \dots$
- **double precision:** $f = 1.17260394005317847 \dots$
- **extended precision:** $f = 1.17260394005317863185 \dots$

The correct result is $f = -0.827396 \dots$, not even a positive number.

This example shows us how, a priori, we cannot trust numerical simulations based on floating point arithmetic. This is normal, for who is used to the scientific simulations: roundoff error can destroy a numerical simulation. A possible solution to this problem is to represent intervals instead of numbers, a bit like the polygonal approximations of Archimede for π ; in this way, the result of a numerical calculation, it will be an interval, instead of a number, and not an ordinary interval, but an interval that contains the exact solution of our problem.

The idea behind this method, just stated, is quite clear; however, it is clear that the goodness of this strategy depends strictly on the ability to obtain a very small interval at the end. The Fourier based method presented in this thesis, is able to obtain errors of the order of 10^{-86} (in the next section we will specify what we mean by error, since the final error in our situation does not arise just from the computation itself, you must take in account the error coming from the discretization of the transfer operator, for instance, for which we developed bounds in the last chapters).

We have already mentioned the packages used in our algorithms, and we will talk about some aspects of them later; as for the theory underlying the interval arithmetic (algebraic properties, interval functions and so on) we refer to [51].

Before starting the study of the algorithms, let us remember what these algorithms do and what are the errors that come into action, to which we must be careful. In a nutshell

- Input : map T , size of the noise ξ , size of the discretization δ , size of a coarser discretization δ'
- Output: a function ν and a number ϵ ; ν is represented by the vector of its basis coefficients and if f is the stationary density of the system, then $\|f - \nu\| \leq \epsilon$, for some suitable norm.

The intermediate steps, will then serve to calculate a discretization of the transfer operator, calculate its fixed point ν , and its bound on the error, ϵ . As we pointed out before, we have three sources of error, we must take in account:

- **Discretization error:** we project the transfer operator on a finite base, thus there is an error on the output coming from this discretization process
- **Approximation error:** the discretization of the transfer operator is a finite linear dimensional operator which we represent with a matrix that cannot be exactly computed; hence we have an error on the output coming from this approximation.
- **Eigen error:** the error coming from the numerical computation of the eigenvector.

The third part is somewhat a solved problem: given a matrix, we may compute rigorously its steady vector, and the bound on the error depends from the **machine epsilon**, which can be arbitrarily increase up to any reasonable level. We refer to the section 8.3 of [24].

Hence in the next sections, providing the full details and the pseudocode for the Fourier based method, we will describe the algorithms for the computation of the aforementioned approximation of the stationary density and the relative bounds on the error, with the respect to the "true" stationary density.

7.1/ ULAM-BASED SCHEME OF APPROXIMATION

In the next sections we provide the details of the rigorous approximation algorithm. This algorithm can be found in different versions, one of them for instance in https://bitbucket.org/luigimarangio/arnold_map/. As we have already observed, for several reasons, we will focus on the full exposition of the Fourier based method; all the details that are omitted here, can be found in [24], [23].

Remember that the steps to perform are :

- construct a matrix, which is an approximation of the transfer operator associated with our random system
- calculate the discretization error and approximation error; this is done by rigorously estimating the rate of mixing of the approximate transfer operator.

In the next two subsections, we will describe how to deal with the above mentioned steps. Finally, we will briefly describe how to estimate the average of an observable, algorithm used for example of the rigorous computation of the rotation number of the Arnold maps.

7.1.1/ BUILDING THE ULAM MATRIX

Let P_k be the matrix of the Ulam approximation, where k is the size of the approximation; if we partition the interval $[0, 1]$ in k intervals I_1, \dots, I_k then, the i, j -th entry of P_k is given by equation 5.1. As already pointed out we cannot compute exactly P_k , but we can compute a matrix \tilde{P}_k which is equal to P_k up to an approximation error. Indeed, we first fix a small $\beta > 0$ and we compute a row matrix \tilde{P}'_k : for each I_i , consider a partition in m smaller intervals J_1, \dots, J_m ; we look at the images $T(J_l)$, and we test whether or not these images are contained in I_j . The only problems comes if $T(J_l)$ overlaps I_j but is not contained. In this case we partition again, until the "problematic" interval has a measure smaller than β ; we denote this measure by $\epsilon_{i,j}$, and at the end we will obtain a row matrix $\tilde{P}'_{i,j}$ with an associated error $\epsilon = \max_{i,j} \epsilon_{i,j}$.

Furthermore observe that we can always split the interval $[0, 1]$ in monotonic branches for the map T ; hence we can suppose that either T is monotonic on I_j or has a discontinuity

point.

Algorithm 1: Computing the row Ulam Matrix \tilde{P}'_{ij} : T is monotone on I_i .

Result: $\tilde{P}'_{ij}, \epsilon$

$\tilde{P}'_{ij} = 0$;

$\epsilon_{ij} = 0$;

Partition I_i in m intervals $I_{i,k}$ for $k = 0, \dots, m-1$;

for $k = 0, \dots, m$ **do**

if $T(I_{i,k}) \subset I_j$ **then**

 | add $m(I_{i,k})$ to $\tilde{P}'_{ij} = 0$;

end

if $T(I_{i,k}) \subset (I_j)^c$ **then**

 | discard $I_{i,k}$;

end

if $T(I_{i,k}) \cap I_j \neq \emptyset$ and $T(I_{i,k}) \cap (I_j)^c \neq \emptyset$ and $m(I_{i,k}) \geq \beta$ **then**

 | divide $I_{i,k}$ in m intervals and repeat the procedure;

end

if $T(I_{i,k}) \cap I_j \neq \emptyset$ and $T(I_{i,k}) \cap (I_j)^c \neq \emptyset$ and $m(I_{i,k}) < \beta$ **then**

 | add $m(I_{i,k})$ to $\epsilon_{i,j}$, discard $I_{i,k}$;

end

end

$\epsilon = \max_{i,j} \epsilon_{i,j}$;

Algorithm 2: Computing the row Ulam Matrix \tilde{P}'_{ij} : T has a discontinuity in I_i .

Result: $\tilde{P}'_{ij}, \epsilon$

$\tilde{P}'_{ij} = 0$;

$\epsilon_{ij} = 0$;

Partition I_i in m intervals $I_{i,k}$ for $k = 0, \dots, m-1$;

for $k = 0, \dots, m$ **do**

if $I_{i,k}$ does not contain a discontinuity **then**

 | apply Algorithm 1;

end

if $I_{i,k}$ contains the discontinuity and $m(I_{i,k}) \geq \beta$ **then**

 | divide $I_{i,k}$ in m intervals and repeat the procedure;

end

if $I_{i,k}$ contains the discontinuity and $m(I_{i,k}) < \beta$ **then**

 | add $m(I_{i,k})$ to $\epsilon_{i,j}$, discard $I_{i,k}$;

end

end

$\epsilon = \max_{i,j} \epsilon_{i,j}$;

To obtain the matrix \tilde{P}_k , we are going to use a **markovization** process: compute the sum of the elements on each row, subtract 1 and spread the result uniformly on each nonzero elements of the row. In this way we have that for every i, j

$$|\tilde{P}_{ij} - P_{ij}| < 2 \times \epsilon,$$

where ϵ it was already defined (and computed) as the maximum of the errors $|\tilde{P}'_{ij} - P_{ij}|$. If nz_i is the number of nonzero elements of the i -th row and $nz = \max_i nz_i$, we have

$$\|P_k - \tilde{P}_k\|_1 < 2 \times nz \times \epsilon.$$

This process will ensure that the biggest eigenvalue of \tilde{P}_k is 1. We will now move to the estimation of the errors.

7.1.2/ BOUNDING THE ERRORS

Let us recall the notations

- f_ξ is the fixed point of L_ξ , the transfer operator associated to the random system
- v_k is the fixed point of P_k , the Ulam approximation of L_ξ (in chapter 5 we called it $L_{\delta, \xi}$)
- \tilde{v}_k is the fixed point of \tilde{P}_k , the matrix we computed
- w the numerical approximation of \tilde{v}_k .

With this notation, we have that

$$\|f - \tilde{v}\|_{L^1} \leq \|f_\xi - v_k\|_{L^1} + \|v_k - \tilde{v}_k\|_{L^1} + \|\tilde{v}_k - w\|_{L^1}.$$

Remark 7.2. In [24], methods to obtain similar estimates in the L^∞ norm are considered.

The third term of the latter equation, comes from the numerical computation of the eigenvector of the matrix \tilde{P}_k and we already treated this issue. The discretization error, $\|f - v_k\|_{L^1}$, was addressed in section 5.1.1, theorem 5.9, and it depends from the contraction rate of the discretized operator. We also see how to improve this estimates with a coarse fine strategy 5.1.2. For the approximation error, $\|v_k - \tilde{v}_k\|_{L^1}$, once we have computed the rate of mixing of \tilde{P}_k , i.e. \bar{n}, α such that $\|\tilde{P}_k^{\bar{n}}\|_{L^1} \leq \alpha$, using lemma 5.6 we have that

$$\|v_k - \tilde{v}_k\|_{L^1} \leq \frac{\bar{n}}{1 - \alpha} \|P_k - \tilde{P}_k\|_{L^1} \|v_k\|_{L^1} \leq \frac{2\bar{n}}{1 - \alpha} \times nz \times \epsilon.$$

As we can see, the only real computation which is needed in these estimates is the rate of mixing of a matrix (remember that if you are applying a coarse fine strategy, you need to compute the contraction rate of the two matrices). Remember that these computation are done in interval arithmetics, so the matrix \tilde{P}_k is an interval matrix. To compute \bar{n}, α such that $\|\tilde{P}^{\bar{n}}|_V\|_{L^1} \leq \alpha$, we have two strategies. The trivial one, we proceed via matrix multiplications: this way is only practicable if you are dealing with small size matrices (up to 2^{14}). The size of the discretization to use depends on the system you are considering; for example in [23], the matrices used are in the order of 2^{27} , very large, so not only a smarter strategy was used, but this strategy was implemented through OpenCL to be executed on gpu Nvidia. Still some computation has required a time of the order of a week, on an Asus GeForce GTX 1050Ti, 4GB of Ram GPU installed in a desktop computer with an AMD A4-6300 3.4 Ghz processor and 8 Gb of Ram.

In the case of the Arnold maps (section 6.2), the experiments were performed on the supercomputer facilities of the Mésocentre de calcul de Franche-Comté. The size of the matrices in this case is between 2^{16} and 2^{19} ; we chose to perform the computations on a supercomputer, because we were able to perform approximately one hundred rigorous calculations in parallel (see proposition 6.8). This parallel computation could not make advantage of the gpu, so each simulation was slower, but despite this there was a significant gain of time. We emphasize the fact that with matrices of size 2^{27} , this strategy would not have been feasible, the calculation without gpu is prohibitive.

The smarter strategy comes from [31], in which the following inequality about roundoff error in matrix vector multiplication is shown

$$\|float(Av) - Av\| \leq \gamma_k \cdot \|A\| \|v\|,$$

where

$$\gamma_k = \frac{ku}{1 - ku},$$

and U is the machine precision.

Hence, instead of a matrix multiplication, we consider the vector matrix products $\tilde{P}_k^n \cdot v_i$, where v_i are zero averages basis vectors, which can be computed just iterating a matrix vector product, without computing \tilde{P}_k^n . This is done in floating point arithmetics, thus we have to convert with a simple script the matrix \tilde{P}_k from interval arithmetic to floating point.

The rate of mixing can be then bounded as $\|\tilde{P}^{\bar{n}}|_V\|_{L^1} \leq \max_i \tilde{P}_k^{\bar{n}} \cdot v_i$. We implement a stopping criterion by choosing the desired rate of mixing α ; once we fixed α , when we find an integer s such that $\max_i \tilde{P}_k^s \cdot v_i + err \leq \alpha$ the algorithm stops (err is the roundoff error computed as explained before).

7.2/ FOURIER-BASED SCHEME OF APPROXIMATION

The algorithm presented in this section is a part of an original work still in progress; the code can be found at https://bitbucket.org/luigimarangio/phd_thesis/. We will now project the transfer operator associated to a random system on the Fourier basis, defined in Chapter 5. Let $b_j = e^{2j\pi ix}$, for $-k \leq j \leq k$, a Fourier basis of $2k + 1$ elements. We will use the fact that the j -th column of Lb_j is made by the "first" $2k + 1$ Fourier coefficients of Lb_j (from $-k$ to k); we will represent the matrix relabeling the frequencies in the following way $0, 1, 2, \dots, k - 1, k, -k, -k + 1, \dots, -1$. The computation of these columns is done by using the explicit action of the transfer operator, given by (see for instance [8], [12])

$$[Lf](x) = \sum_{y \in T^{-1}(x)} \frac{f(y)}{|T'(y)|}, \quad (7.1)$$

to evaluate the function $L_\sigma b_j$ at $N \gg k$ interpolation nodes. We stress on the fact that this is the only point in which we need some regularity assumption on the map T and in the future work we hope to remove this assumption by using the duality relation of the transfer operator to compute $L_\sigma b_j$. However, plenty of interesting random dynamical system have a smooth deterministic part. We then perform a Fast Fourier Transform to recover the Fourier coefficients of $L_\sigma b_j$; the coefficients obtained in this way are really close to the "true" coefficients, and the error is known as **aliasing error** and it was analyzed in section 5.2.2.

Once we have computed the matrix, we have to compute its rate of mixing, the approximation error and the discretization error; we now present the algorithms and the relative explanations.

7.2.1/ BUILDING THE FOURIER MATRIX

To compute the column of the Fourier matrix, we need an interval version of the Fast Fourier Transform; this problem was solved by **emulating** the *FFT* with the package `ArbComplex.jl`; this latter is a package that uses the Arb C Library, and adapts some C library interface work from Nemo; furthermore it exports the types needed to work in the interval analysis framework (e.g. `ArbReal` and `ArbComplex`). In the author's knowledge, until now there was not in the literature an interval version of the FFT, which has now been added to the package `ArbComplex.jl`.

Another important remark is about the enumeration of the frequencies. We choose to put at index 1 in the matrix the *0th*-frequency, then the positive frequencies from 1 to k and

finally the negative frequencies in the reverse order, from $-k$ to -1 .

Algorithm 3: Compute the Fourier matrix M_k via FFT

Result: An interval matrix M_k

Define N interpolation nodes $x_i = i/N$, for $i = 0, \dots, N - 1$;

Initialize an empty (interval) matrix M ;

for $j = 1, \dots, 2k + 1$ **do**

 Initialize a vector of zeros v_j ;

 Compute the frequency j' which corresponds to the index j ;

for $i = 0, \dots, N - 1$ **do**

 Compute $Lb_{j'}(x_i)$ with equation 7.1 and add it on $v_j[i]$;

end

 Apply the Fast Fourier Transform to v_j , obtaining a new vector w_j . Consider a new vector v'_j which contains the first $2k + 1$ frequencies, thus the first k elements and the last $N - k$ of w_j ;

v'_j is the j th-column of M

end

The computation of the noise matrix is particularly easy. Let us stress that in this section we are considering noises with gaussian kernel, meanwhile before we considered uniform distributed kernels; since the gaussian function is an eigenfunction of the Fourier transform, the noise matrix D_σ is diagonal with entries

$$D_{\sigma,k}(j, j) = e^{-(\sigma\pi j')^2},$$

where as usual we shift the indexes to j' to select the proper frequency. The finite dimensional operator $L_{\sigma,k}$ is then approximated by the matrix $H_k = D_{\sigma,k}M_k$. In the next subsection we end this work by describing how to estimate the errors arising from this scheme of approximation.

7.2.2/ BOUNDING THE ERRORS

In this first version of the algorithm, we are computing the contraction rate of H_k by iterative multiplications. This is possible because we have exponential bounds: thanks to coarse-fine strategy, it is possible to obtain bounds of order 10^{-80} with relative small matrices. In the example shown in Chapter 6 we consider a fine partition of size 2000 and a coarser partition of size 500: multiplying interval matrices of size 500 is not a big deal.

Remark 7.3. Of course we must be careful to the accumulation of the numeric error, which the interval analysis takes automatically in account. Indeed, we bound the L^∞ -norm on the zero average space with the sum of all the elements of H_k , except for the first row and the first column (these corresponds to the 0th-frequency); founding a more effective way

to bound this norm is a problem on which we are working on.

Let H_k and H_{kfine} be the two approximation matrices computed with algorithm 7.2.1 for two partitions of different size. The following algorithm bounds the mixing rate of H_{kfine} using information on H_k ; as said before, the norm of a matrix means the sum of all the elements of the matrix except for the first row and the first column.

Algorithm 4: Compute $\alpha, \sum C_i, \bar{n}$ as in section 5.2.1

Result: $\alpha, \text{sum}C_i, \bar{n}$

Set a maximum number of iteration l and $\alpha = 0$;

Set a tolerance value for α , say β ;

Set $\text{sum}C_i$ equal to the norm of H_{kfine} ;

Set $Q = H_k$ **for** $i = 1, \dots, l$ **do**

$Q = Q * H_k$;

we add to α the norm of H_k and the coarse fine estimate from equation 5.27;

if $\alpha < \beta$ **then**

return α, i and $\text{sum}C_i$;

end

add α to $\text{sum}C_i$;

end

Once we have computed the rate of mixing, we can take care of the discretization error and the approximation error. Nevertheless to say, with the theory developed so far, they are really use to compute:

the approximation error is given by lemma 5.26 and we just need to compute the following quantity

$$\frac{1}{1-\alpha} \sum_{i=0}^{N-1} C_i (1 + \Gamma_{\sigma,k} + \|\rho_{\sigma}\|_{\infty}) \Gamma_{\sigma,k};$$

the discretization error in the Fourier case is given by the aliasing error, equation 5.7, which is equal to

$$\Gamma_{\sigma,N}.$$

7.2.3/ AN INTRODUCTORY EXAMPLE

We are finally ready to apply this Fourier scheme to an easy but non trivial example; more examples will come in future works. As already mentioned at the beginning of Chapter 6, this is a first technical and explanatory example, which has no practical relevance (it is not a toy-model of a natural phenomenon); obviously then its complexity is not comparable to that of the models presented above, and this example is intended as a first application of

the theory presented so far.

We consider a perturbed doubling map

$$T(x) = 2x + 0.001 \sin(2\pi x), \quad (7.2)$$

and we choose the following parameters for the computation

- $k = 250$ is the size of the coarse discretization (thus we work with matrix of size 501), used to compute the C_i (see also remark 7.3)
- $k' = 2000$ is the size of a finer discretization, combined with the latter via lemma 5.27 to obtain better estimates
- $N = 10000$ is the number of interpolation points
- $\sigma = 0.01$, where σ^2 is the variance of the gaussian kernel.

Applying the results described so far we obtain the following estimate on the rate of mixing

$$\alpha = 0.0100808 \quad N = 6 \quad \text{and} \quad \sum_{i=0}^{N-1} C_i = 57.355,$$

which leads to an approximation error of

$$\|f_\sigma - f_{\sigma,k}\|_\infty \leq 6.5381 \times 10^{-86}$$

an error on the dft of

$$k \cdot \Gamma_{\sigma,N} \leq 10^{-300}.$$

This simple example allows us to observe the potential of the proposed method. As can be reasonably expected, figure 7.1 shows the density of the invariant measure of the system. We conclude by pointing out that the only thing that currently prevents us from applying this method to more complex systems is the computation of C_i , for which we need better theoretical estimates.

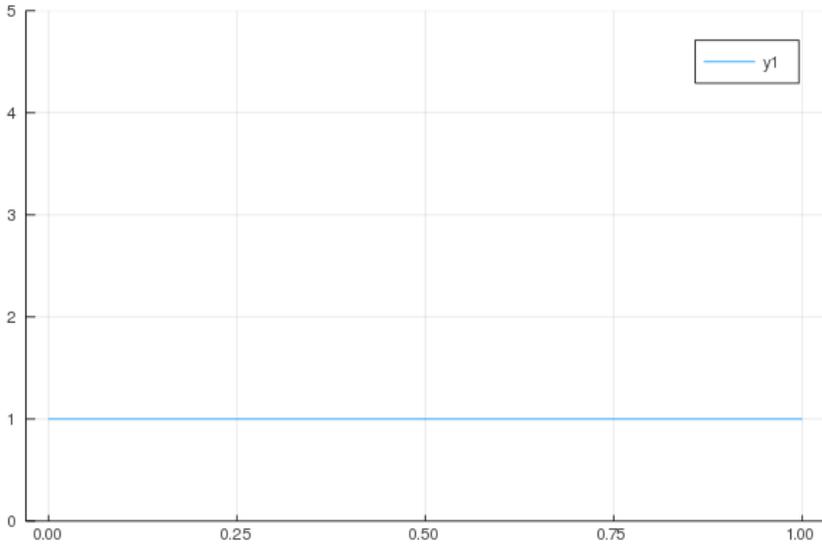


Figure 7.1: Density of the computed invariant measure of 7.2

BIBLIOGRAPHY

- [1] AIMINO, R. ; NICOL, M. ; VAIENTI, S.: **“Annealed and quenched limit theorems for random expanding dynamical systems.”**. In *Probab. Theory Relat. Fields* **162**, 233–274 (2015)
- [2] ARNOLD, L.: **“Random Dynamical Systems”**. In *Springer Monographs in Mathematics* (2013)
- [3] ARNOLD, V.I.: **“Small denominators I. Mapping the circumference onto itself.”**. (1965)
- [4] BAHOUN, W. ; GALATOLO, S. ; NISOLI, I. ; NIU, X.: **“A Rigorous Computational Approach to Linear Response”**. In *Nonlinearity* **31** (no.3), 1073–1109 (2018)
- [5] BAHOUN, W. ; RUZIBOEV, M. ; SAUSSOL, B.: **“Linear response for random dynamical systems”**. In *arXiv:1710.03706*
- [6] BAHOUN, W. ; SAUSSOL, B.: **“Linear response in the intermittent family: frentiation in a weighted C^0 -norm”**. In *Discrete Contin. Dyn. Syst.* **36**(12), 6657–6668 (2016)
- [7] BAILEY, M.P. ; DRKS, G. ; SHELDONM, A.C.: **“Circle maps with gaps: Understanding the dynamics of the two-process model for sleep–wake regulation”**. In *European Journal of Applied Mathematics* (2018)
- [8] BALADI, V.: **“Positive transfer operator and decay of correlations”**. In *vol.16, World Scientific, Shangai* (2000)
- [9] BALADI, V.: **“Linear Response, or else.”**. In *arXiv:1408.2937* (2014)
- [10] BILINGSLEY, P.: **“Probability and Measure”**. (1986)
- [11] BOSE, C. ; MURRAY, R.: **“The exact rate of approximation in Ulam’s method”**. In *Discrete Contin. Dynam. Systems* **7** no. 1, pp. 219–235 (2001)
- [12] BOYARSKY, A. ; GORA, P.: **“Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension”**. In *Probability and Its Applications Birkhauser Boston* (1997)
- [13] BOYARSKY, A. ; GÓRA, P.: *Laws of Chaos, Invariant measures and Dynamical Systems in one dimension.* 1997

- [14] CAPINSKI, M. ; SIMO, C.: **“Computer assisted proof for normally hyperbolic invariant manifolds.”**. In *Nonlinearity*, Vol. 25, No. 7, pp. 1997-2026 (2012)
- [15] CRIMMINS, H. ; FROYLAND, G.: **“Fourier approximation of the statistical properties on Anosov maps on tori.”**. In *Nonlinearity*, Volume 33, Number 11 (2020)
- [16] DELLNITZ, M. ; JUNGE, O.: **“On the approximation of complicated dynamical behavior.”**. In *SIAM Journal on Numerical Analysis*, 36(2):491–515 (1999)
- [17] DELLNITZ, M. ; JUNGE, O.: *Set Oriented Numerical Methods for Dynamical Systems*. 2002
- [18] DRAGICEVIC, D. ; FROYLAND, G. ; GONZÁLEZ-TOKMAN, C. ; VAIENTI, S.: **“A spectral approach for quenched limit theorems for random expanding dynamical systems”**.
- [19] DUNFORD, N. ; SCHWARTZ, J.T.: **“Linear operators. Part I: General Theory.”**. (1958)
- [20] FOGUEL, S. R.: **“The ergodic theory of Markov processes”**. (1969)
- [21] FROYLAND, G.: **“Approximating physical invariant measures of mixing dynamical systems.”**. In *Nonlinear Analysis, Theory, Methods & Applications*, 32:831–860 (1998)
- [22] GALATOLO, S. ; GIULIETTI, P.: **“Linear response for dynamical systems with additive noise.”**. In *Non linearity* 32, n. 6 pp. 2269-2301 (2019)
- [23] GALATOLO, S. ; MONGE, M. ; NISOLI, I.: **“Existence of Noise Induced Order, a Computer Aided Proof”**. In *Nonlinearity* 33 4237 (2020)
- [24] GALATOLO, S. ; NISOLI, I.: **“An elementary approach to rigorous approximation of invariant measures.”**. In *SIAM J. Appl. Dyn. Sys.* 13 pp. 958-985 (2014)
- [25] GALATOLO, S. ; NISOLI, I. ; ROJAS, C.: **“Probability, statistics and computation in dynamical systems”**. In *Math. Struct. in Comp. Science* , vol. 24, iss. 3 doi:10.1017/S0960129512000758 (2014)
- [26] GALIAS, Z. ; TUCKER, W.: **“Is the Hénon attractor chaotic?”**. In *Chaos* 25, 033102 (2015)
- [27] GARRIDO, P.L. ; GALLAVOTTI, G.: **“Billiards Correlation Functions”**. In *Journal of Statistical Physics*, Vol. 76, Nos. 1/2, pp. 549-585 (1994)
- [28] GLASS, L.: **“Cardiac arrhythmias and circle maps - A classical problem”**. In *Chaos: An Interdisciplinary Journal of Nonlinear Science* 1, 13–19 (1991)

- [29] GUIHENEUF, P.A.: **“Physical measures of discretizations of generic diffeomorphisms”**. In *to appear on Erg. Th. Dyn. Sys.* doi:10.1017/etds.2016.70
- [30] HERMAN, M.R.: **“Mesure de Lebesgue et nombre de rotation, in: Geometry and Topology (Proc. III Latin Amer. School of Math., Inst. Mat. Pura Aplicada CNPq, Rio de Janeiro, 1976).”**. In *Lecture Notes in Math., vol.597, Springer, Berlin, pp. 271-293* (1977)
- [31] HIGHAM, N.J.: **“Accuracy and Stability of Numerical Algorithms”**. In *2nd edition SIAM publishing, Philadelphia (PA), US, ISBN 0-89871-521-0* (2002)
- [32] HOPF, E.: **“The general temporally discrete Markoff process.”** (1954)
- [33] I. TSUDA, E. K. ; SHIMIZU, H.: **“Memory dynamics in asynchronous neural networks.”**. In *Prog. Theor. Phys., 51–71.* (1987)
- [34] J.DING ; LI, T.-Y.: **“Markov finite approximation of the Frobenius–Perron operator.”**. In *Nonlinear Analysis: Theory, Methods & Applications, 17:759–772* (1991)
- [35] JENKINSON, O. ; POLLICOTT, M.: **“Orthonormal expansions of invariant densities for expanding maps”**. In *Advances in Mathematics, 192, pp. 1–34.* (2005)
- [36] JENKINSON, O. ; POLLICOTT, M. ; VYTCNOVA, P.: **“Rigorous computation of diffusion coefficients for expanding maps”**. In *Journal of Statistical Physics, 170, pp. 221–253.* (2018)
- [37] KATOK, A. ; HSELBLATT, B.: **“Introduction to the modern theory of dynamical systems.”**. In *Cambridge University Press* (1995)
- [38] KEENER, J.P. ; GLASS, L.: **“Global bifurcations of a periodically forced nonlinear oscillator”**. In *J. Math. Biology 21, 175–190* (1984)
- [39] KOLTAI, P.: **“Efficient approximation methods for the global long-term behavior of dynamical systems – Theory, algorithms and examples.”**. In *PhD thesis, Technische Universitat Munchen* (2010)
- [40] KRENEGEL, U.: **“On the speed of convergence in the ergodic theorem”**. (1978)
- [41] KRENEGEL, U.: **“Ergodic Theorems”**. (1985)
- [42] LASOTA, A. ; MACKEY, M.: **“Chaos, Fractals and Noise. Stochastic Aspects of Dynamics, 2. ed.”**. (1994)
- [43] LI, T.-Y.: **“Finite approximation for the Frobenius–Perron operator. A solution to Ulam’s conjecture.”**. In *Journal of Approximation Theory, 17:177–186* (1976)

- [44] LIBERALQUINO, R. B. ; MONGE, M. ; GALATOLO, S. ; MARANGIO, L.: **“Chaotic Itinerary in Random Dynamical System Related to Associative Memory Models.”**. In *Mathematics 2018*, 6, 39 (2018)
- [45] LIVERANI, C.: **“Rigorous numerical investigations of the statistical properties of piecewise expanding maps-A feasibility study.”**. In *Nonlinearity* 14 pp. 463-490 (2001)
- [46] LUQUE, A. ; VILLANUEVA, J.: **“Computation of the derivatives of the rotation number for parametric families of circle diffeomorphisms.”**. In *Physica D* 237:2599–2615 (2008)
- [47] MARANGIO, L. ; NISOLI, I. ; GALATOLO, S.: **“A posteriori validated numerical method for the computation of stationary measures based on Fourier approximation”**. In *Work in progress* (2021)
- [48] MARANGIO, L. ; SEDRO, J. ; GALATOLO, S. ; GARBO, A. D. ; GHIL, M.: **“Arnold Maps with Noise: Differentiability and Non-monotonicity of the Rotation Number.”**. In *J. Stat. Phys.* 179, 1594–1624 (2020)
- [49] MATSUMOTO, S.: **“Derivatives of the rotation number of one parameter families of circle diffeomorphisms.”**. In *Kodai Mat. J.* 35:115–125 (2012)
- [50] M.CHEKROUN ; E.SIMONNET ; M.GHIL: **“Stochastic climate dynamics: random attractors and time-dependent invariant measures.”**. In *Physica D. Nonlinear Phenomena*, 240, 1685–1700 (2011)
- [51] MOORE, R. E. ; KEARFOTT, R.B. ; CLOUD, M.J.: **“Introduction to Interval Analysis”**. (2009)
- [52] MURRAY, R.: *Discrete approximation of invariant densities*. 1997
- [53] OHNO, T.: **“Asymptotic behaviors of dynamical systems with random parameters”**. (1981)
- [54] RUELLE, D.: **“Differentiation of SRB states.”**. In *Communications in Mathematical Physics*, 187:227–241 (1997)
- [55] RUELLE, D.: **“Nonequilibrium statistical mechanics near equilibrium: computing higher-order terms.”**. In *Nonlinearity*, 11(1):5 (1998)
- [56] SLIPANTSCHUK, J. ; RICHTER, M. ; CHAPPELL, D.J. ; G.TANNER ; JUST, W. ; BANDTLOW, F.: **“Transfer operator approach to ray-tracing in circular domains”**. In *Nonlinearity*, 33, 5773 (2020)

- [57] TSUDA, I.: **“Dynamic link of memory. Chaotic memory map in nonequilibrium neural networks.”**. In *Neural Netw.*, 5, 313–326. (1992)
- [58] TSUDA, I.: **“Chaotic Itinerancy.”**. In *Scholarpedia*, 8(1):4459 (2013)
- [59] ULAM, S. M.: **“A collection of mathematical problems.”**. In *Interscience Publisher NY* (1960)
- [60] V.I.ARNOLD: **“Small denominators. I: Mappings of the circumference onto itself”**. In *AMS Translations, Ser. 2*, 46, 213–284 (1965)
- [61] V.I.ARNOLD: **“Cardiac arrhythmias and circle mappings”**. In *Chaos: An Interdisciplinary J. Nonlin. Sci.* 1, 20–24 (1991)
- [62] WIGGINS, S.: **“Introduction to Applied Dynamical Systems and Chaos.”**. In *Springer NY* (2003)
- [63] WORMELL, C.L.: **“Spectral Galerkin methods for transfer operators in uniformly expanding dynamics.”**. In *Numerische Mathematik* 142 421–463 (2019)
- [64] ZMARROU, H. ; HOMBURG, A.J.: **“Bifurcations of stationary measures of random diffeomorphisms.”**. In *Ergodic Theory Dynam. Systems* 27(5):1651–1692 (2007)

LIST OF FIGURES

6.1	Data collected in [33].	80
6.2	Deterministic component of Equation (6.1).	81
6.5	Plot of the rotation number $\rho = \rho_\tau$ for $\epsilon = 0.9$. The black line corresponds to the absence of uniformly distributed noise, while the red line shows this dependence in the presence of such noise, with amplitude $\xi = 0.05$. Recall that $\tau = 2\pi/\omega$, where ω is the driving frequency. The lower panel is a blow-up of the dependence near the value $\tau = 0.5$	87
6.6	Same as Fig. 6.5 but for $\epsilon = 1.4$ and for noise amplitude $\xi = 0.01$. Here the lower panel is a blow-up of the dependence near $\tau = 0.7$	88
7.1	Density of the computed invariant measure of 7.2	106

LIST OF TABLES

6.1	Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ , with parameters $A = 0.8$, $C = 0.01$	82
6.2	Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ , with parameters $A = 0.8$, $C = 0.01$	82
6.3	Given the Arnold map with noise of magnitude ξ and parameters (τ_0, ϵ_0) , for which we have already proved mixing, the table shows the computed intervals $I_0 = [\tau_0 - \theta_0, \tau_0 + \theta_0]$, such that if $\tau \in I_0$ then the Arnold map with parameters (τ, ϵ_0) is mixing	93
6.4	This table shows the computed intervals in which the rotation number lies for each value of the parameters in consideration.	94

SPIM

école doctorale sciences pour l'ingénieur et microtechniques

