Computation of stationary densities of systems with additive noise

Ricardo Bioni Liberalquino

Contents

Introduction

The purpose of this work is to introduce some quantitative methods that can be used in the study of random dynamical systems, that is, systems whose laws of evolution are stochastically determined. This is distinct from the qualitative treatment of dynamical systems, which can be applied to systems whose precise evolution is poorly understood. Instead, a quantitative approach is useful to study specific systems for which, for example, a numerical prediction is desired.

The concept of random dynamical systems arises from the problem of generalizing the ergodic theorem, which states that, if F is a measure preserving transformation, then the Birkhoff averages of any $\phi \in L^1$, given by the arithmetic mean of the first k iterates, converge almost everywhere and in L^1 to some invariant $\overline{\phi} \in L^1$. Instead of applying the same transformation in each step to obtain the Birkhoff averages, a generalization is obtained by considering random iterates, where the transformation is chosen in a family according to some probability distribution.

In [30], it was verified that the generalization reduces to the former theorem if the random system is considered as a skew-product that preserves some measure whose marginal is stationary. Roughly speaking, a measure is stationary if it is equal to the expected value of its pushforward, averaged over the deterministic components of the system (definition 1.2).

The skew-product representation of the random transformation provides a way to treat a random dynamical system as a deterministic dynamical systems over a larger space. Many concepts, like ergodicity, admit a natural extension to random dynamical systems considering the skew-product representation.

A classical condition on the existence of invariant measures of a deterministic transformation requires only compactness of the space and continuity of the transformation. Since the skew-product is a deterministic transformation (over a larger space), this leads to a condition on the existence of stationary measures. However, this criterion is based on a fixed point theorem that does not guarantee absolute continuity with respect to some reference measure, because the set of absolutely continuous measures is not necessarily closed in the weak-* topology; consider, for example, an approximate identity, that converges to a Dirac measure, which is not absolutely continuous with respect to Lebesgue.

Absolute continuity with respect to Lebesgue measure is important for numerical investigation, thus another approach to the existence of stationary measures with this property is necessary. Instead of taking pushforward measures, the Perron-Frobenius operator can be used to study the effect of the random transformation on densities, in the sense of Radon-Nikodym derivatives of absolutely continuous measures with respect to a reference measure, provided the transformation satisfies a nonsingularity condition.

The Perron-Frobenius operator is a Markov operator, that is, positive, contractive in L^1 and preserves the set of L^1 functions of norm 1. An existence criterion of absolutely continuous stationary measures is then a particular case of a known theorem on the existence of fixed points of Markov operators. In corollary 3.2, we exhibit an explicit condition that guarantees the existence of a stationary density for systems on the torus with additive noise. That is, random dynamical systems whose deterministic components are given by a translation of some fixed transformation on the state space.

Computation of the stationary densities is in general not simple. For example, trying to directly use the ergodic theorem can lead to a very slow convergence if additional assumptions aren't made [29]. Some relatively general criteria that guarantee fast convergence exist, for example, [35] in the setting of Markov chains. However, verifying that a system satisfies these conditions can be a difficult task.

We will discuss a method used in [14] to study an one-dimensional model related to Belousov-Zhabotinsky reaction, a system with a marked oscillatory behavior (see also [16, 17, 15] for applications to deterministic dynamics or iterated functions systems). This method relies on a mixing condition on the system that is verified numerically, which guarantees that the iterates of densities converge exponentially fast to a unique stationary density, and provides a way to calculate stationary densities, with L^1 bounds on the approximation. A "coarse-fine" strategy, where a discretization using a coarser partition is used in computationally more expensive tasks and is then related to a finer partition to obtain tighter bounds, makes the algorithm sufficiently fast to be run on a personal computer.

This method was used to study the effect of varying levels of additive noise on the Lyapunov exponent of the stochastically perturbed system. That is, since a transformation with an additive noise induces a family of such transformations, considering the approximate identity generated by the noise distribution, one may consider the Lyapunov exponent for each of these random transformations. The L^1 approximation error in the stationary density permitted to calculate these observables within a narrow margin of error and provided a numerical approach to determine "noise-induced order for this system.

We note that in the original setting in [14], the one-dimensionality of the system played an important role on the numerical estimates. Here, we prove that the estimates can be generalized to higher dimensions (theorem B) by using an appropriate higher-dimensional definition of functions of BV type defined on the torus and a range of results available in the literature, for example the Poincaré inequality for convex domains. This new result indicates the possibility to extend the method to higher dimensional systems.

Beyond existence, a natural question that arises is the regularity of the stationary densities. In [49], smoothness of the stationary densities was obtained assuming smoothness of the density of the transition probability of a random dynamical system. In the context of noises of BV type, in particular in the numerically simulated noise, this result can't be used.

We show in theorem 3.2 that, if the BV noise is Lipschitz in its support, assumed to be a ball, we can assure local Lipschitz continuity in every point that satisfies a condition of local boundedness of the noiseless Perron-Frobenius operator in a sphere centered in the point.

We illustrate the (one-dimensional) method with a model that arises in the context of neural networks, proposed in [31] as an example of a system exhibiting "chaotic itinerancy", a property that it suggests to hold in the mesoscopic dynamics of the hippocampus.

Although this map is related to mesoscopic brain dynamics, it shares properties with microscopic brain dynamics. For example, a similar map is obtained in [47], in the context of the BvP neuron driven by a sinusodial external stimulus, which belongs to a family known as Arnold circle maps (named after [3]), an useful family in physiology (see [18, equation (3)]). Like the more complicated models from which it is derived, we have the presence of quasi attractors (it can be seen as a stochastic perturbation of a system with Milnor attractors).

This model is obtained in a similar way to the one related to the Belousov-Zhabotinsky reaction that is, derived from the Lorentz plot of an oscillatory process, which in this case is the collective firing activity of neurons. Therefore it is natural to apply the method from [14] to study this system, which is represented as a random dynamical system with additive noise, providing a new application to the algorithm.

We determine both the existence of regions where the iterates of the system concentrates and the mixing property of the system, a contribution towards the mathematical formalization of the chaotic itinerancy property. A local Lipschitz continuity at every point outside a finite set is obtained from theorem A. These results are summarized in theorem C.

The structure of this work is as follows. In the first chapter, we define random dynamical systems, stationary measures and other basic notions, and present a version of the ergodic theorem, along with some results that are generalizations of classical theorems from ergodic theory, for example the ergodic decomposition theorem. In our context, in which random dynamical systems are essentially Markov processes, these results are essentially rephrasing of results from the ergodic theory of Markov processes (see, for example, [13]).

In the second chapter, we present the Perron-Frobenius operator which allows the study of stationary measures that are absolutely continuous with respect to a fixed reference measure, by viewing the action of the random transformation on densities of measures. We also present the context of systems perturbed by an additive noise, in which an explicit expression of the operator is available in the literature. We also study an appropriate notion of BV space which will play an important role in chapter 4.

In the third chapter, we discuss general results on the existence of the stationary measures and densities with respect to a reference measure and the effect of the Perron-Frobenius operator on the regularity of densities. We also establish theorem A on the regularity of stationary densities for these systems.

Finally, in the fourth chapter, we present the framework used in [14] to rigorously compute stationary densities (with respect to Lebesgue measure) of systems perturbed by an additive BV noise. Except where specified, we don't assume that the system is one-dimensional, and provide a generalization of estimates that were previously shown to hold only in the one-dimensional case.

Up to the author's knowledge, theorems A, B and C are new, except parts of C that were obtained in the joint work [6].

This thesis is submitted in partial satisfaction of the requirements for the doctoral degree in Mathematics from Universidade Federal do Rio de Janeiro — UFRJ, with the supervision of Maria José Pacifico, from UFRJ, and Stefano Galatolo, from UNIPI.

Acknowledgements

I express my gratitude to my tutors Zezé, who made possible this work and has always motivated me, and Stefano, who presented the subject of this thesis and has guided my studies, especially during my stay in Pisa; to Maurizio Monge, who helped me in many parts of this work, and to Isaia Nisoli.

I received financial support from CNPq during my studies in Universidade Federal do Rio de Janeiro and from CAPES, under the PDSE scholarship, during my studies in Università di Pisa.

Chapter 1

Random dynamical systems

Throughout the text, we denote by $(S, \mathcal{S}, p)$ a probability space and by $(\Omega, \mathcal{A}, \mu)$ the corresponding space of (one-sided) sequences with the product σ-algebra $\mathcal{A} = \mathcal{S}^{\otimes \mathbb{N}_0}$ and probability measure $\mu = p^{\mathbb{N}_0}$. Also we denote by σ the shift map on Ω, $\sigma(\{\omega_i\}_{i\in\mathbb{N}_0}) = \{\omega_{i+1}\}_{i\in\mathbb{N}_0}$.

Definition 1.1 ([44, p. 60])**.** Let $(M, \mathcal{B})$ be measurable space. Endow $\Omega \times M$ with the product σ-algebra $\mathcal{A} \otimes \mathcal{B}$. A *random transformation* over σ is a measurable transformation of the form

$$F : \Omega \times M \to \Omega \times M, \quad F(\omega, x) = (\sigma(\omega), F_\omega(x)),$$

where $\omega \mapsto F_\omega$ depends only on the zeroth coordinate of ω. By abuse of notation, we denote this common mapping as F_{ω_0}.

A *random orbit* for F and a starting point x is a sequence given by

$$x_0 = x, \quad x_n = F^n_\omega(x) := F_{\omega_{n-1}} \circ \cdots \circ F_{\omega_0}(x), \qquad n \geq 1.$$

If q is a measure on $\Omega \times M$, and $\pi_M : \Omega \times M \to \Omega$, $\pi_M : \Omega \times M \to M$ are the canonical projections, we say that q *projects* to $\pi_{\Omega,*}q$ and that $\pi_{M,*}q$ is the *marginal* of q.

Example 1.1. Let $S = \{1, \dots, d\}$, $\Omega = S^{\mathbb{N}_0}$, and suppose $F_1, \dots, F_d$ are homeomorphisms over a locally compact metric space M. Define $F(\omega, x) = (\sigma(\omega), F_{\omega_0}(x))$. Then F is a random transformation over σ.

A random orbit with starting point x_0 can be defined inductively by independently choosing (according to the distribution given by p) a function $F_{\omega_i} \in \{F_1, \dots, F_d\}$ and taking $v_{i+1} = F_{\omega_i}(x_i)$. Equivalently, this can be viewed as the projection in M of the orbit $\{F^i(\omega, x_0)\}_{i\in\mathbb{N}_0}$, where $\omega = \{\omega_i\}_{i\in\mathbb{N}_0}$.

It is sometimes useful to consider the *slices* of a measurable $E \subset \Omega \times M$, defined by $E_x := \{\omega \in \Omega : (\omega, x) \in E\}$, for $x \in M$, and $E^\omega := \{x \in M : (\omega, x) \in E\}$, for $\omega \in \Omega$. The following proposition shows that the process of slicing is well-behaved.

Proposition 1.1 ([44, exercise 5.1])**.** *The operators $P^\omega : M \to \Omega \times M$, where $\omega \in \Omega$, and $P_x : \Omega \to \Omega \times M$, where $x \in M$, defined by $P^\omega(x) = (\omega, x)$ and $P_x(\omega) = (\omega, x)$ are measurable. Moreover, for measurable $E \subset \Omega \times M$,*

1. $E_x = \{\omega \in \Omega : (\omega, x) \in E\} \in \mathcal{A}$ *and* $E^\omega = \{x \in M : (\omega, x) \in E\} \in \mathcal{B}$*;*
2. $x \mapsto \mu(E_x)$ *and* $\omega \mapsto \eta(E^\omega)$ *are measurable for any probability measures* μ *on* $(\Omega, \mathcal{A})$ *and* η *on* $(M, \mathcal{B})$*;*
3. $\int_M \mu(E_x)\, d\eta(x) = (\mu \times \eta)(E) = \int_M \eta(E^\omega)\, d\mu(\omega)$*.*

Proof. P_x is measurable because $\pi_M \circ P_x$ and $\pi_M \circ P_x$ are both measurable, for the first is the identity on Ω and the second a constant function. Similarly, P^ω is measurable. Thus $E^\omega = (P^\omega)^{-1}(E)$ and $E_x = (P_x)^{-1}(E)$ are measurable.

To prove the second and third items, we will show that the family

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

is a λ-system, that is, it contains $\Omega \times M$ and is closed under complements and finite unions of pairwise disjoint sets; and contains the sets $A \times B \in \mathcal{A} \otimes \mathcal{B}$, which form a π-system, that is, a family closed under finite intersections.

The π-λ theorem [5, theorem 3.2] states that, if a λ-system $\mathcal{L}$ contains a π-system $\mathcal{P}$, then $\mathcal{L}$ contains the σ-algebra generated by $\mathcal{P}$. Therefore all measurable $E \subset \Omega \times M$ belong to $\mathcal{E}$, because the sets $A \times B$ generate this σ-algebra. The proof of the statement concerning $\omega \mapsto \eta(E^\omega)$ is analogous.

We first prove that $A \times B \in \mathcal{A} \otimes \mathcal{B}$ constitute a π-system and belong to $\mathcal{E}$. The π-system condition follows from $(A \times B) \cap (A' \times B') = (A \cap A') \times (B \cap B')$. The inclusion in $\mathcal{E}$ follows from

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

and

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

To verify that $\mathcal{E}$ is a λ-system, we need to prove that $\Omega \times M \in \mathcal{E}$; if $E \in \mathcal{E}$, then $E^c := (\Omega \times M) \setminus E \in \mathcal{E}$; and if $\{E_i\}_{i\in\mathbb{N}}$ is a sequence of pairwise disjoint sets in $\mathcal{E}$, then $\cup_{i\in\mathbb{N}} E_i \in \mathcal{E}$. The first condition holds, because $x \mapsto \mu((\Omega \times M)_x) = \mu(\Omega)$ is constant, in particular measurable, and

$$\int_M \mu((\Omega \times M)_x)\, d\eta(x) = \mu(\Omega)\eta(M) = (\mu \times \eta)(\Omega \times M).$$

The second condition holds, since $x \mapsto \mu((E^c)_x) = \mu(\Omega \setminus E_x) = \mu(\Omega) - \mu(E_x)$ is measurable if $x \mapsto \mu(E_x)$ is measurable, and

$$\int_M \mu((E^c)_x)\, d\eta(x) = \mu(\Omega)\eta(M) - \int_M \mu(E_x)\, d\eta(x) = (\mu \times \eta)(E^c).$$

Now, let $\{E(i)\}_{i\in\mathbb{N}}$ be a sequence of pairwise disjoint sets in $\mathcal{E}$ and note that $(\cup_{i\in\mathbb{N}} E(i))_x = \cup_{i\in\mathbb{N}} E(i)_x$. Then $x \mapsto \mu((\cup_{i\in\mathbb{N}} E(i))_x) = \sum_{i\in\mathbb{N}} \mu(E(i)_x)$ is given by a series of positive measurable functions, hence measurable. Further

$$\int_M \mu((\cup_{i\in\mathbb{N}} E(i))_x)\, d\eta(x) = \sum_{i\in\mathbb{N}} \int_M \mu(E(i)_x)\, d\eta(x) = \sum_{i\in\mathbb{N}} (\mu \times \eta)(E(i))$$
$$= (\mu \times \eta)((\cup_{i\in\mathbb{N}} E(i))_x). \square$$

1.1 The transition operator

Suppose we have a random transformation $F : \Omega \times M \to \Omega \times M$ and an *observable* that quantifies some property of M, in general $\phi : M \to \mathbb{R}$ or $\phi \in L^p(\eta)$, $p \in [1, \infty]$. Given a random orbit $\{x_i\}_{i\in\mathbb{N}}$ for which we know the value of $x_0 = x \in M$, a natural question is to predict the value of $\phi(x_1)$, that is, the expected value $\mathbb{E}(\phi(x_1) \,|\, x_0 = x)$. Since the iterate depends on an outcome $\omega \in \Omega$, which is distributed according to μ, this is formally calculated as

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

If $\phi = \chi_B$ for some $B \in \mathcal{B}$, the integral is well-defined and

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

thus $U\phi(\cdot)$ defines a bounded, measurable function on M according to proposition 1.1. This extends to every measurable ϕ that is bounded or measurable by a standard measure theory argument (see proof of lemma 1.1).

Thus (1.1) defines the *transition operator* associated to F, which maps bounded (or nonnegative) measurable functions to bounded (or nonnegative) measurable functions.

Dually, we can consider its *adjoint transition operator* that acts in the space of signed or probability measures η on M, defined by

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

for $B \in \mathcal{B}$.

Proposition 1.1 shows that the integral is well-defined and

$$U^*\eta(B) = (\mu \times \eta)(F^{-1}(\Omega \times B)) = F_*(\mu \times \eta)(\Omega \times B), \tag{1.3}$$

a special case of proposition 1.2. Therefore $U^*\eta$ is a well-defined probability.

The operators U and U^* are related by the following property.

Lemma 1.1. *For any bounded or nonnegative measurable $\phi : M \to \mathbb{R}$,*

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

Proof. Set f, g the positive linear functionals defined by

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

If $\{a_n\}_{n\in\mathbb{N}}$ is an increasing sequence that converges pointwise to a, denote this by $a_n \uparrow a$. Lebesgue's monotone convergence theorem tells that if $\phi_n \uparrow \phi$, then $f(\phi_n) \uparrow f(\phi)$ and $g(\phi_n) \uparrow g(\phi)$.

The following lemma shows that we need to check $f(\phi) = g(\phi)$ only for characteristic functions $\phi = \chi_B$, where $B \in \mathcal{B}$.

Lemma 1.1.1. *Let $f, g : \mathcal{F} \to \mathcal{G}$ be positive linear operators, where $\mathcal{G}$ is an ordered vector space and $\mathcal{F}$ is either the space of nonnegative or bounded measurable functions on $(M, \mathcal{B})$. Suppose that*

$$f(1) = g(1) \quad \textit{if} \quad (\exists \phi \in \mathcal{F} : \phi \neq \phi^+)$$
$$\forall B \in \mathcal{B} : f(\chi_B) \leq g(\chi_B);$$
$$\forall \{\phi_n\}_{n\in\mathbb{N}} \subset \mathcal{F} : (\phi_n \uparrow \phi \in \mathcal{F} \implies f(\phi_n) \uparrow f(\phi)).$$

Then

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

Proof. By linearity and positivity, $f(\phi) \leq g(\phi)$ for all positive simple functions $\phi = \sum_{i=1}^{n} a_i \chi_{B_i}$.

If ϕ is bounded, then there is $c > 0$ such that $\phi + c$ is nonnegative and $f(\phi) \leq g(\phi)$ if and only if $f(\phi + c) \leq g(\phi + c)$. Hence we can assume that ϕ is nonnegative.

Define, for each $n \in \mathbb{N}$,

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

which are simple functions because the sums have finite, $n2^n$, nonzero terms. If $\phi(x) \in [i/2^n, (i+1)/2^n)$ for some $i \in \{0, \ldots, n2^n - 1\}$, then $\phi_n(x) = i/2^n$; otherwise, $\phi_n(x) = 0$. Hence $\phi_n \uparrow \phi$.

Since $f(\phi_n) \leq g(\phi)$ for every $n \in \mathbb{N}$ and $f(\phi_n) \uparrow f(\phi)$, we conclude that $f(\phi) \leq g(\phi)$. □

Let $\phi = \chi_B$, $B \in \mathcal{B}$. By (1.2), (1.3) and proposition 1.1,

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

Then lemma 1.1.1 allows to extend the equality to all bounded or measurable $\phi : M \to \mathbb{R}$. □

Of particular importance are the fixed points of the operator U^*.

Definition 1.2. A probability measure η for M is called *stationary* for the random transformation F if $U^*\eta = \eta$.

We recall the deterministic concept of invariant measures.

Definition 1.3. If $h : C \to C$ is a measurable mapping in the measurable space $(C, \mathcal{C})$, we say μ is an *invariant measure* for h if $\mu(h^{-1}(E)) = \mu(E)$ for any measurable $E \subset C$.

Notice that, if η is invariant for every F_ω, then η is stationary for F. The converse does not hold, as shown by the following example.

Example 1.2. Let $d = 2$, $M = \mathrm{P}\mathbb{R}^2$, $F_1([x]) = [A_1 x]$, $F_2([x]) = [A_2 x]$ in the example 1.1, where A_1 and A_2 are invertible matrices that admit a decomposition into invariant sets

$$\mathbb{R}^3 = \mathrm{E}_i^s \oplus E_i^u, \qquad i = 1, 2$$

satisfying, for all $x^s \in \mathrm{E}_i^s$ and $x^u \in \mathrm{E}_i^u$,

$$\|A_i x_i^s\| < \lambda_i \|x^s\|, \|A_i^{-1} x^u\| < \mu_i \|x^u\|, \qquad \lambda_i, \mu_i < 1,$$

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

Suppose a measure η is both F_1 and F_2 invariant and take V a neighborhood of $[E_1^u \setminus \{0\}]$ such that $[(\lambda x^s, x^u)] \in V$ for every $\lambda < 1$ and $[(x^s, x^u)] \in V$. We have

$$V \subset F_1^{-1}(V) \subset \{[(x^s, x^u)] : [(\mu_1 \lambda_1 x^s, x^u)] \in V\}.$$

Inductively, $V \subset \cup_{k=0}^{\infty} F_1^{-k}(V)$. Clearly, $\cup_{k=0}^{\infty} F_1^{-k}(V) \subset M \setminus [E_1^s]$, and conversely, $M \setminus [E_1^s] \subset \cup_{k=0}^{\infty} F_1^{-k}(V)$ because for any $[(x^s, x^u)] \in M \setminus [E_1^s]$, we have $[(\lambda_1^k \mu_1^k x^s, x^u)] \in V$ for some $k \in \mathbb{N}$. By the invariance of F_1,

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

This holds for arbitrarily small neighborhoods V of $[E_1^u]$, therefore $\eta(M \setminus ([\mathrm{E}_1^u] \cup [\mathrm{E}_1^s])) = 0$. The same reasoning implies $\eta(M \setminus ([\mathrm{E}_2^u] \cup [\mathrm{E}_2^s])) = 0$ and since we assume the eigenspaces to be distinct, we have $\eta(M) = 0$, a contradiction.

Invariant measures and stationary measures for a one-sided random transformation F are related by the following proposition.

Proposition 1.2 ([30, lemma 2.1]). *$F_*(\mu \times \eta) = \mu \times (U^* \eta)$ for every probability measure η on M. In particular, η is stationary if and only if $\mu \times \eta$ is invariant.*

Proof. For any measurable $A \times B \subset \Omega \times M$,

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

Therefore $F_*(\mu \times \eta) = \mu \times (U^* \eta)$ and $F_*(\mu \times \eta) = \mu \times \eta$ if and only if $U^* \eta = \eta$. $\square$

It should be noted, however, that not every invariant measure for F that projects to μ can be written as this kind of product measure. To show this, we first give a general characterization of such measures, based on their disintegration.

Definition 1.4. A *disintegration* of a probability measure q on $\Omega \times M$ with respect to μ is a function $q : \Omega \times \mathcal{B} \to [0, 1]$ such that

1. for all $B \in \mathcal{B}$, $\omega \mapsto q_\omega(B)$ is $\mathcal{A}$-measurable;
2. for μ-a.e. $\omega \in \Omega$, $B \mapsto q_\omega(B)$ is a probability measure on $(M, \mathcal{B})$;
3. for all $S \in \mathcal{A} \otimes \mathcal{B}$,

$$q(S) = \int_\omega \int_M \chi_S(\omega, x)\, dq_\omega(x)\, d\mu(\omega)$$

When M is a complete separable metric space, with its Borel σ-algebra, then such a disintegration exists and is unique up to μ-null sets. For any sub-σ-algebra $\mathcal{S}$ of $\mathcal{A}$, the restriction $q|_{\mathcal{S}\otimes\mathcal{B}}$ also admits a disintegration with respect to $\mu|_{\mathcal{S}}$, the *conditional expectation* of q with respect to $\mathcal{S}$. It is denoted by $\omega \mapsto \mathbb{E}(q_\cdot \,|\, \mathcal{S})_\omega$ and satisfies

$$\mathbb{E}(q_\cdot \,|\, \mathcal{S})_\omega(B) = \mathbb{E}(q_\cdot(B) \,|\, \mathcal{S})(\omega) \qquad \mu\text{-a.e.},$$

where for a measurable and integrable function $\Phi : \Omega \to \mathbb{R}$, the conditional expectation with respect to $\mathcal{S}$ is the unique $\mathcal{S}$-measurable function $\mathbb{E}(\Phi \,|\, \mathcal{S})$ such that

$$\forall A \in \mathcal{S} : \int_A \mathbb{E}(\Phi \,|\, \mathcal{S})\, d\mu = \int_A \Phi\, d\mu. \tag{1.4}$$

Proposition 1.3. *[2, p. 23, lemma 1.4.4] Let $F : \Omega \times M \to \Omega \times M$ be a random transformation, where Ω is a standard space, and let q be a probability measure on $\Omega \times M$ that projects to μ. Then q is F-invariant if and only if, for all $n \in \mathbb{N}_0$:*

$$\mathbb{E}(F^n_{\cdot,*} q_\cdot | \sigma^{-n}\mathcal{A})_\omega = q_{\sigma^n(\omega)} \qquad \mu\text{-}a.e. \tag{1.5}$$

Proof. For any $A \in \mathcal{A}$ and $B \in \mathcal{B}$, the σ-algebras associated to Ω and M respectively,

$$\begin{aligned} F^n_* q(A \times B) = q(F^{-n}(A \times B)) &= \int_{\sigma^{-n}(A)} q_\omega(F^{-n}_\omega B)\, d\mu(\omega) \\ &= \int_{\sigma^{-n}(A)} F^n_{\omega,*} q_\omega(B)\, d\mu(\omega) \end{aligned}$$

and

$$q(A \times B) = \int_A q_\omega(B)\, d\mu(\omega) = \int_A q_\omega(B)\, d(\sigma_*^n \mu)(\omega)$$
$$= \int_{\sigma^{-n}(A)} q_{\sigma^n(\omega)}(B)\, d\mu(\omega)$$

Thus q is F-invariant if and only if, for all $n \in \mathbb{N}_0$,

$$\forall A \in \mathcal{A}, B \in \mathcal{B} : \int_{\sigma^{-n}(A)} F^n_{\omega,*} q_\omega(B)\, d\mu(\omega) = \int_{\sigma^{-n}(A)} q_{\sigma^n(\omega)}(B)\, d\mu(\omega),$$

which is equivalent to (1.5). □

Definition 1.5. [2, p. 25] An F-invariant probability measure q is called a *random Dirac measure* if $q_\omega = \delta_{x(\omega)}$ μ-a.e. for some random variable $x : \Omega \to M$. The function x is called the base point of the random Dirac measure.

Note that, by the invariance condition, the random variable x in the definition must satisfy, for all $n \in \mathbb{N}_0$:

$$F^n_\omega(x(\omega)) = x(\sigma^n(\omega)) \qquad \mu\text{-a.e.}$$

If x is not a.e. constant, then q can not be written in the form $\mu \times \eta$, because a measure with this form has constant disintegration. The following example, adapted from [2, p. 52, exercise 2.1.2], shows a situation in which this happens.

Example 1.3. Take $\Omega = S^{\mathbb{N}_0}$, $b \in L^1(S)$, where S is a probability space, $M = \mathbb{R}^2$ and define the random dynamical system

$$F : \Omega \times M \to \Omega \times M, \quad F(\omega, x) = (\sigma(\omega), Ax + b(\omega_0)), \qquad A = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}.$$

Now, consider the disintegration $\omega \to \delta_{x(\omega)}$, where

$$x(\omega) = \Big(0, -\sum_{i=0}^{\infty} 2^{-i-1} b(\omega_i)\Big).$$

This defines a (F-invariant) random Dirac measure because

$$F_\omega(x(\omega)) = \Big(0, b(\omega_0) - \sum_{i=-1}^{\infty} 2^{-i-1} b(\omega_{i+1})\Big) = x(\sigma(\omega)).$$

If $S = \{0, 1\}$ with uniform distribution and $b(\omega) = \omega$, the range of x is the interval $[0, 1]$ and this measure can not be written in the form $\mu \times \eta$.

Another concept from deterministic dynamical systems that can be naturally extended to random dynamical systems is that of ergodicity.

1.2 Ergodicity

Definition 1.6. We say that a (not necessarily stationary) probability η on M is *ergodic* for a random transformation $F : \Omega \times M \to \Omega \times M$ if every set B that is η-stationary, i.e. $U\chi_B = \chi_B$ η-a.e., has full or null η-measure.

In case η is stationary, ergodicity is equivalent to the stronger condition that every $\phi \in L^\infty(\eta)$ that is η-stationary, i.e. $U\phi = \phi$, is constant in some full η-measure set [25, lemma 2.4]. In fact, if ϕ is η-stationary, then $B(c) = \{x \in M : \phi(x) > c\} = \phi^{-1}(c, +\infty)$ is η-stationary for every $c \in \mathbb{R}$ (theorem 1.1). Therefore $\{x \in M : \phi(x) = \overline{c}\}$ has full measure, where $\overline{c}$ is the supremum over the set of $c \in \mathbb{R}$ such that $\eta(B(c)) = 1$.

Theorem 1.1. *Let η be a probability on M and T be a linear, positive, contractive (that is, $\|T\|_{L^p} \leq 1$) operator on $L^p(\eta)$, $p \in [1, \infty]$, such that $T1 \leq 1$. Let $\mathcal{S}$ be the family of sets $B \subset M$ such that $T\chi_B = \chi_B$.*

Then $\mathcal{S}$ is a σ-ring on M. Moreover, $\mathcal{S}$ is the smallest σ-ring on M for which the nonnegative or bounded fixed points of T are measurable.

Proof. Suppose η is a probability on M. We verify that the family of stationary sets $\mathcal{S}$ is a σ-ring on M as in [13, p. 8]. If $A, B \in \mathcal{S}$, then $A \cup B \in \mathcal{S}$ because

$$T\chi_{A\cup B} \leq \min\{1, T\chi_A + T\chi_B)\} = \min\{1, \chi_A + \chi_B\} = \chi_{A\cup B},$$
$$T\chi_{A\cup B} \geq \max\{T\chi_A, T\chi_B\} = \max\{\chi_A, \chi_B\} = \chi_{A\cup B}.$$

Also, $A \setminus B \in \mathcal{S}$ because

$$T\chi_{A\setminus B} = T(\chi_{A\cup B} - \chi_B) = \chi_{A\cup B} - \chi_B = \chi_{A\setminus B}.$$

Finally, for any sequence $\{B_i\}_{k\in\mathbb{N}}$ of sets in $\mathcal{S}$ and $B = \cup_{i\in\mathbb{N}} B_i$,

$$T\chi_B = \lim_{k\to+\infty} T\chi_{\cup_{j=1}^k B_j} = \lim_{i\to+\infty} \chi_{\cup_{j=1}^k B_j} = \chi_B,$$

where the limits hold η-a.e. according to the following result, where $\psi_n \uparrow \psi$ indicates pointwise convergence of the increasing sequence $\{\psi_n\}_{n\in\mathbb{N}}$.

Lemma 1.1.2. *If $\phi_n \uparrow \phi \in L^p(\eta)$, then $T\phi_n \uparrow T\phi$.*

Proof. For each $n \in \mathbb{N}$, let

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

$\{T\phi_n\}_{n\in\mathbb{N}}$ is a monotone sequence because $\{\phi_n\}_{n\in\mathbb{N}}$ is monotone and T is positive. Thus $E_1 \supset E_2 \supset \cdots$ and we have to prove that $\inf_{n\in\mathbb{N}} \eta(E_n) = 0$.

In case $p = \infty$, $\eta(E_n) = 0$ if $\|T\phi_n - T\phi\| < \epsilon$, so $\eta(E_n) = 0$ whenever $\|\phi_n - \phi\| < \epsilon/\|T\|$.

In case $p \in [1, \infty)$, we write $E_n = \{x \in M : |T\phi_n(x) - T\phi(x)|^p > \epsilon^p\}$ and use the Markov inequality

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

so $\eta(E_n)$ whenever $\|\phi_n - \phi\| < \epsilon^p/\|T\|^p$.

In either case, such an n exists because $\phi_n \to \phi$ in $L^p(\eta)$. $\square$

We treat now the second part of the theorem. Any σ-ring for which the nonnegative or bounded fixed points of T are measurable contains $\mathcal{S}$, for $A = \{x \in M : \chi_A(x) > 0\}$ belongs to $\mathcal{S}$ if and only if $T\chi_A = \chi_A$.

Furthermore, every nonnegative or bounded fixed point ϕ is $\mathcal{S}$-measurable, because any set of the form $B(c) := \{x \in M : \phi(x) > c\}$ belongs to $\mathcal{S}$. To show this, the idea from [34, p. 92–93] is to construct a monotone sequence $\{\phi_n\}_{n\in\mathbb{N}}$ of nonnegative stationary functions such that ϕ_n converges pointwise to $\chi_{B(c)}$. It follows that $T\chi_{B(c)} = \chi_{B(c)}$, because

$$T\chi_{B(c)} = \lim_{n\to+\infty} T\phi_n = \lim_{n\to+\infty} \phi_n = \chi_{B(c)}.$$

For the sequence $\{\phi_n\}_{n\in\mathbb{N}}$, we can take

$$\phi_n(x) = \min\{1, n \max\{0, \phi(x) - c\}\}.$$

Clearly, $\{\phi_n\}_{n\in\mathbb{N}}$ is a monotone sequence that converges pointwise to $\chi_{B(c)}$. Each ϕ_n is a fixed point because, for every fixed points ψ_1 and ψ_2,

$$\min\{\psi_1, \psi_2\} = \frac{\psi_1 + \psi_2}{2} - \frac{|\psi_1 - \psi_2|}{2}, \quad \max\{\psi_1, \psi_2\} = \frac{\psi_1 + \psi_2}{2} + \frac{|\psi_1 - \psi_2|}{2}$$

are fixed points.

In fact, the fixed points of T form a linear subspace and $|\psi(\cdot)|$ is fixed if ψ is fixed, by proposition 2.6. $\square$

When $F : \Omega \times M \to \Omega \times M$ is a random transformation and ϕ is some appropriate function, we have defined $U\phi$ as the expected value of $\phi(x_1)$ when we know the initial point x_0 of a random orbit $\{x_i\}_{i\in\mathbb{N}}$. In the same spirit, consider averages

$$\frac{\phi(x) + U\phi(x) + \cdots + U^{n-1}\phi(x)}{n} = \frac{\mathbb{E}(\phi(x_0) + \ldots, \phi(v_{n-1}) \,|\, x_0 = x)}{n}.$$

Intuitively, when n is large, and R is some small region, the contributions of $x_i \in R$ to the average are the product of the frequency in which they appear and the expected value of ϕ on R. In case η is ergodic, this frequency is proportional to $\eta(R)$, thus the contribution is $\int_R \phi \, d\eta$. Summing over all small regions, we find that

$$\frac{\phi(x) + U\phi(x) + \cdots + U^{n-1}\phi(x)}{n} \approx \int \phi \, d\eta.$$

This idea is formalized by the ergodic theorems, which we present in the context of positive contractions (cp. with [28]).

1.3 Ergodic theorems

The main result in this section is theorem 1.2, which deals with the pointwise convergence of the Birkhoff averages. The results in this section apply to both the transition operator U and the Perron-Frobenius operator L (definition 2.1) of a nonsingular random transformation $F : \Omega \times M \to \Omega \times M$, by remark 2.1.

Definition 1.7. If T is a linear, positive, contractive operator on $L^1(\eta)$, then the n-th *Birkhoff sum* and the n-th *Birkhoff average*, with respect to T, are the linear, positive, contractive operators $S_n : L^1(\eta) \to L^1(\eta)$ and $A_n : L^1(\eta) \to L^1(\eta)$ given by

$$S_n\phi = \sum_{j=0}^{n-1} T^j\phi = \phi + T\phi + \cdots + T^{n-1}\phi \quad \text{and} \quad A_n\phi = \frac{S_n\phi}{n} \tag{1.6}$$

We start our investigation on the convergence of the Birkhoff averages with the following lemma. For a proof, see [13, p. 9].

Lemma 1.2 (Hopf maximal ergodic lemma)**.** *Let $T : L^1(\eta) \to L^1(\eta)$ be a linear, positive, contractive operator. If $\psi \in L^1(\eta)$, then*

$$\int_E \psi \, d\eta \geq 0, \quad \textit{where } E = \Big\{x \in M : \sup_{n\in\mathbb{N}} S_n\psi(x) > 0\Big\}. \tag{1.7}$$

In case ψ is of the form $\phi - \epsilon$, where $\epsilon > 0$, we obtain

$$\int_E \phi \, d\eta \geq \epsilon \eta(E), \quad \text{where } E = \Big\{ x \in M : \sup_{n \in \mathbb{N}} A_n \phi(x) > \epsilon \Big\}. \tag{1.8}$$

This lemma is traditionally used to decompose M into *conservative* and *dissipative* parts [20, p. 32], respectively

$$C = \Big\{ x \in M : \sum_{j=0}^{\infty} T^j \phi_0(x) = +\infty \Big\}, \tag{1.9}$$

$$D = \Big\{ x \in M : \sum_{j=0}^{\infty} T^j \phi_0(x) < +\infty \Big\}. \tag{1.10}$$

where $\phi_0 \in L^1(\eta)$, $\phi_0 > 0$.

To show the decomposition is well-defined, let $\phi \in L^1_+(\eta)$ and consider

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

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

Take $j \in \mathbb{N}_0$ such that $\eta(\{x \in C' : T^j \phi(x) > 0\}) \geq 2^{-j-1} \eta(C')$, which exists because $\sum_{j=0}^{\infty} T^j \phi(x) > 0$ implies

$$\eta(C') \leq \sum_{j=0}^{\infty} \eta(\{x \in C' : T^j \phi(x) > 0\}).$$

Apply (1.7) to $\psi_C = \phi_0 - aT^j \phi$ and $\psi_D = \phi - a\phi_0$, where $a > 0$. Then $I' \subset E_I := \{x \in M : \sup_{n \in \mathbb{N}} S_n \psi_I(x) > 0\}$ for $I \in \{C, D\}$ and

$$0 \leq \int_{E_C} (\phi_0 - aT^j \phi) \, d\eta \leq \int_M \phi_0 \, d\eta - a \int_{C'} T^j \phi \, d\eta,$$
$$0 \leq \int_{E_D} (\phi - a\phi_0) \, d\eta \leq \int_M \phi \, d\eta - a \int_{D'} \phi_0 \, d\eta.$$

Since $a > 0$ was arbitrary, the conclusion is that $\eta(C') = 0 = \eta(D')$. Therefore,

$$C \subset \Big\{ x \in M : \sum_{j=0}^{\infty} T^j \phi(x) = 0 \quad \text{or} \quad \sum_{j=0}^{\infty} T^j \phi(x) = +\infty \Big\}, \tag{1.11}$$

$$D \subset \Big\{ x \in M : \sum_{j=0}^{\infty} T^j \phi(x) < +\infty \Big\}, \tag{1.12}$$

In particular, C and D do not depend on the choice of $\phi_0 > 0$.

If we decompose the averages as

$$A_n\phi = A_n\phi^+ - A_n\phi^-,$$

then we see that $\lim_{n\to+\infty} A_n\phi = 0$ η-a.e. on D.

Consider now the Banach operator adjoint $T^* \in L^\infty(\eta)$ of T. Note that T^* is positive and $\|T^*\|_{L^\infty} = \|T\|_{L^1} \leq 1$. Moreover,

$$T^*1 \leq 1. \tag{1.13}$$

In fact, for any $\epsilon > 0$ and $B \in \mathcal{B}$, $B = \{x \in M : T^*1 > 1+\epsilon\}$ satisfies

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

thus $\eta(B) = 0$, which proves (1.13).

Definition 1.8. $B \in \mathcal{B}$ is an *invariant set* if $T^*\chi_B = \chi_B$ on the conservative part C of T.

The next lemma, adapted from [20, lemma 9.4], provides a way to verify if a set is invariant.

Lemma 1.3. *If $\phi \in L^\infty(\eta)$ satisfies $T^*\phi \leq \phi$ on C, then $T^*\phi = \phi$ on C. If $\phi \in L^1_+(\eta)$ satisfies $T\phi \leq \phi$ on C, then $T\phi = \phi$ on C.*

Proof. Suppose $\phi \in L^\infty(\eta)$, $T^*\phi \leq \phi$ on C and fix $\psi = \phi - T^*\phi$, $\phi_0 = 1$. Since $\psi + \ldots + T^{*(n-1)}\psi = \phi - T^{*n}\phi$, we have

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

It results from $\sum_{i=0}^\infty T^{n-1}\phi_0 = +\infty$ on C that $\psi = 0$ on C. Thus the first part of the lemma is proven.

Now suppose $\phi \in L^1_+(\eta)$, $T\phi \leq \phi$ but $T\phi \neq \phi$ on C. In particular, there exists $\epsilon > 0$ such that $B = \{x \in C : \phi(x) - T\phi(x) > \epsilon\}$ has positive measure.

We claim that $\sum_{j=0}^\infty T^{*j}\chi_B = +\infty$ on B. In fact, for any $c > 0$, $B' = \{x \in B : \sum_{j=0}^\infty T^{*j}\chi_B < c\}$ satisfies

$$\int_M \Big(\sum_{j=0}^\infty T^j\chi_{B'}\Big)\chi_B\, d\eta = \int_M \chi_{B'}\Big(\sum_{j=0}^\infty T^{*j}\chi_B\Big)\, d\eta < +\infty,$$

thus $\eta(B')=0$, for $B'\subset B\subset C$ and (1.11) imply $(\sum_{j=0}^{\infty} T^j\chi_{B'})\chi_B=+\infty$ on B'. Note that the duality extends to infinite sums by monotone convergence. Then $(\phi-T\phi)\sum_{j=0}^{\infty} T^{*j}\chi_B=+\infty$ on B, which contradicts

$$\int_M (\phi - T\phi)\Big(\sum_{j=0}^{\infty} T^{*j}\chi_B\Big)\,d\eta = \int_M \Big(\sum_{j=0}^{\infty} T^j(\phi - T\phi)\Big)\chi_B\,d\eta \le \int_B \phi\,d\eta < +\infty.$$

Thus $T\phi=\phi$ on C. □

An immediate consequence of the lemma is that M is invariant, by (1.13). Moreover,

$$\chi_C \le T^*\chi_C \quad \text{and} \quad T^*\chi_D \le \chi_D. \tag{1.14}$$

This follows from the fact that $T^*\chi_C\le T^*1\le 1$ on C, thus $T^*\chi_C=1$ on C, that is, $T^*\chi_C\ge\chi_C$; and $T^*\chi_D=T^*(1-\chi_C)\le 1-\chi_C=\chi_D$. We are ready to prove the following.

Lemma 1.4. *The invariant sets (definition 1.8) form a σ-algebra $\mathcal{C}$ on M. Moreover, every $\phi\in L^\infty(\eta)$ such that $T^*\phi=\phi$ on C is $\mathcal{C}$-measurable.*

Proof. In view of theorem 1.1, we search a probability η_C on $(M,\mathcal{B})$ with the property that $T^*\phi=\phi$ on C if and only if $T^*\phi=\phi$ η_C-a.e. If $\eta(C)=0$, we can simply take $\eta=\eta_C$, then $\mathcal{C}=\mathcal{B}$; otherwise, we define η_C on C by $\eta_C(B)=\frac{\eta(B\cap C)}{\eta(C)}$ for $B\in\mathcal{B}$.

Let $B^1,B^2\in\mathcal{B}$ such that $\chi_{B^1}=\chi_{B^2}$ η-a.e. on C. Then

$$\chi_C(T^*\chi_{B^1})=\chi_C(T^*\chi_{B^1\cap C}+T^*\chi_{B^1\cap D})=\chi_C T^*\chi_{B^1\cap C}=\chi_C T^*\chi_{B_2},$$

η-a.e. on M, where $\chi_C T^*\chi_D=0$ by (1.14).

Taking monotone sequences, it follows that $T^*\phi=T^*\psi$ on C whenever $\phi=\psi$ on C. Thus T^* can be regarded as an operator on $L^1(\eta_C)$, and is also linear, positive and contractive. Since $T^*1=1$ η_C-a.e., the lemma shows that $\mathcal{C}$ is a σ-algebra and every $\phi\in L^\infty(\eta_C)$ such that $T^*\phi=\phi$ is $\mathcal{C}$-measurable. □

We recall that a sequence $\{x_n\}_{n\in\mathbb{N}}$ in a Banach space X is *weakly convergent* if there exists $x\in X$ such that $\lim_{n\to+\infty} f(x_n)=f(x)$ for every continuous linear functional f on X. We are ready to state the following.

Theorem 1.2 (Pointwise ergodic theorem)**.** *Let T be a linear, positive, contractive operator on $L^1(\eta)$.*

If there exists $\phi_0 \in L^1(\eta)$ with $\phi_0 > 0$ such that $\{A_n\phi_0\}_{n\in\mathbb{N}}$ admits a weakly convergent subsequence, then for every $\phi \in L^1(\eta)$, the limit

$$\overline{\phi}(x) = \lim_{n\to+\infty} A_n\phi(x) = \lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} T^j\phi(x) \tag{1.15}$$

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

Furthermore, $\overline{\phi}$ satisfies

$$\forall B \in \mathcal{C} : \int_B \overline{\phi}\, d\eta = \int_B H_C\phi\, d\eta, \tag{1.16}$$

where $\mathcal{C}$ is the σ-algebra of invariant sets (in the sense of definition 1.8) and

$$H_E\phi = I_E \sum_{k=0}^{\infty}(TI_{M\setminus E})^k\phi, \quad I_E\phi(x) = \chi_E(x)\phi(x), \qquad E \subset M. \tag{1.17}$$

In particular, if F is a random transformation with stationary ergodic probability η, then

$$\lim_{n\to+\infty} \frac{1}{n}\sum_{j=0}^{n-1} \phi(F^j_\omega x) = \int_M \phi\, d\eta$$

for $\mu \times \eta$-a.e. $(\omega, x) \in \Omega \times M$.

Remark 1.1. If L^1 limits of Birkhoff averages exist for every $\phi \in L^1(\eta)$, then we recover the hypothesis: there exists $\phi_0 \in L^1(\eta)$ with $\phi_0 > 0$ such that $\{A_n\phi_0\}_{n\in\mathbb{N}}$ admits a weakly convergent subsequence. In fact, we can take any $\phi_0 \in L^1(\eta)$ such that $\phi_0 > 0$, because L^1 convergence implies weak convergence. This hypothesis appears in [26] and is related to the Krylov-Bogolyubov procedure (proposition 3.2).

An alternative hypothesis is given in [**?**]: $\lim_{n\to+\infty} T^{*n}\chi_D = 0$ a.e. and there exists $\phi_0 \in L^1_+(\eta)$ such that $\{x \in M : \phi_0(x) > 0\} = C$ and $\{A_n\phi_0\}_{n\in\mathbb{N}}$ is weakly sequentially compact. Moreover, ϕ_0 can be chosen in order that $T\phi_0 = \phi_0$ [28, p. 175, theorem 3.3].

In the general case, we have convergence of Birkhoff averages in the stochastic sense [28, p. 143, theorem 4.9]. This follows from the existence of a decomposition $M = \tilde{C} \cup \tilde{D}$ such that $\tilde{C} = \{x \in M : \phi_0(x) > 0\} \in \mathcal{C}$, $\tilde{D} = \{x \in M : h_0(x) > 0\}$, $T^*\chi_{\tilde{D}} \le \chi_{\tilde{D}}$, for some $\phi_0 \in L^1_+(\eta)$ and

$h_0 \in L^\infty_+(\eta)$ satisfying $T\phi_0 = \phi_0$ and $\lim_{n\to+\infty} \|A_n^* h_0\|_{L^\infty} = 0$ [28, p. 141–142, theorem 4.6]. In fact, this implies $T(L^1(\tilde{C})) \subset L^1(\tilde{C})$, thus, on $\tilde{C}$, $A_n\phi = A_n(I_{\tilde{C}}\phi)$ converges a.e. to $H_{\tilde{C}}\phi$ on $\tilde{C}$ by theorem 1.2; on the other hand, given any $\alpha > 0$ and $\phi \in L^1_+(\eta)$, there exist $\delta > 0$ and $N \in \mathbb{N}$ such that $B = \{x \in M : h_0(x) > \delta\}$ satisfies $\eta(\tilde{D} \setminus B) < \alpha/2$ and $\|A_n^*\chi_B\| \leq \delta^{-1}\|A_n^* h_0\|_{L^\infty} < \alpha\epsilon\|\phi\|^{-1}_{L^1}/2$ for $n \geq N$. Therefore, for all $n \geq N$,

$$\begin{aligned}\eta(\{x \in \tilde{D} : A_n(I - H_{\tilde{C}})\phi(x) > \epsilon\}) &\leq \alpha/2 + \epsilon^{-1} \int_B A_n(I - H_{\tilde{C}})\phi(x)\, d\eta(x) \\ &\leq \alpha/2 + \epsilon^{-1}\|A_n^*\chi_B\|_{L^\infty}\|\phi\|_{L^1} < \alpha.\end{aligned}$$

It follows that $A_n\phi$ converges stochastically to 0 on $\tilde{D}$, for all $\phi \in L^1(\eta)$.

If one is interested only in a.e. convergence, a sufficient condition is given in [28, p. 132, theorem 3.12]: there exists $\phi \in L^1_+(\eta)$ such that $T\phi \leq \phi$ and $\{x \in M : \phi(x) > 0\} = C$. This also provides an equivalence. In fact, lemma 1.3 implies that $T\phi = \phi$ on C, and evidently $T\phi = 0 = \phi$ on D; hence $\{A_n\phi\}_{n\in\mathbb{N}}$ is constant, thus weakly convergent. Moreover, Conversely, given $\phi_0 \in L^1(\eta)$ with $\phi_0 > 0$ such that $\{A_n\phi_0\}_{n\in\mathbb{N}}$ admits a weakly convergent subsequence, let $\phi \in L^1_+(\eta)$ be the weak limit of this subsequence. We have $T\phi \leq \phi$ as the pointwise limit of

The following proof requires two general results, lemma 1.2 and theorem 1.3 [28, theorem 1.1]. Alternate proofs of similar statements are found in [13], [28]. Here, η denotes any probability measure on $(M, \mathcal{B})$.

Theorem 1.3 (Mean ergodic theorem)**.** *Let T be a bounded linear operator on a Banach space X such that the Birkhoff averages defined by* (2.3) *are uniformly bounded operators. Then, for any $x \in X$ such that*

$$\lim_{n\to+\infty} n^{-1}T^{n-1}x = 0$$

and for any $y \in X$, the following assertions are equivalent.

1. *$Ty = y$ and $y \in \overline{\mathrm{ch}}\{T^{n-1}x : n \in \mathbb{N}\}$;*
2. *$\{A_n x\}_{n\in\mathbb{N}}$ converges strongly to y;*
3. *$\{A_n x\}_{n\in\mathbb{N}}$ converges weakly to y;*
4. *$\{A_n x\}_{n\in\mathbb{N}}$ admits a subsequence that converges weakly to y.*

Proof of theorem 1.2. We first verify the assumptions of the mean ergodic theorem. Clearly, $\|T^n\| \leq 1$ for every $n \in \mathbb{N}$, by the contractive property.

Also, given any $\phi \in L^1(\eta)$, the sequence $\{\frac{S_n\phi}{n}\}_{n\in\mathbb{N}}$ has a weakly convergent subsequence if and only if [10, p. 294, corollary 11]

$$\lim_{\eta(E)\to 0} \int_E \frac{S_n(\phi)}{n}\, d\eta = 0 \tag{1.18}$$

uniformly on n. Given any $\epsilon > 0$, take $\epsilon_0 > 0$ such that, for measurable $E \subset M$,

$$\eta(E) < \epsilon_0 \implies \int_E |\phi(\cdot)|\, d\eta < \epsilon.$$

Now, take $t > 0$ such that

$$\eta(B_t) < \min\{\epsilon, \epsilon_0\}, \quad \text{where } B_t = \{x \in M : t\phi_0(x) < |\phi(x)|\}.$$

By hypothesis, there is $\phi_0 > 0$ such that $\{\frac{S_n\phi_0}{n}\}_{n\in\mathbb{N}}$ has a weakly convergent subsequence. Thus, there exists $\delta > 0$ such that

$$\eta(E) < \delta \implies \int_E \frac{S_n(\phi_0)}{n} < \frac{\epsilon}{t}.$$

We conclude that if $\eta(E) < \delta$, then

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

Thus, $\{\frac{S_n\phi}{n}\}_{n\in\mathbb{N}}$ has a weakly convergent subsequence, and the assumptions of theorem 1.3 are verified.

If $\phi \in L^1(\eta)$, apply the mean ergodic theorem to obtain $\overline{\phi} \in L^1(\eta)$ such that $T\overline{\phi} = \overline{\phi}$ and

$$\lim_{n\to+\infty} \Big\|\frac{S_n\phi}{n} - \overline{\phi}\Big\|_{L^1(\eta)} = 0.$$

We claim that

$$\lim_{n\to+\infty} \frac{S_n\phi(x)}{n} = \overline{\phi}(x) \tag{1.19}$$

for η-a.e. $x \in M$.

To prove (1.19), we show that for any given $\epsilon > 0$,

$$\limsup_{n\to+\infty} \frac{S_n\phi(x)}{n} \leq \overline{\phi}(x) + \epsilon \tag{1.20}$$

for η-a.e. $x \in M$, because this bound applied to $-\phi$ is equivalent to

$$\liminf_{n\to+\infty} \frac{S_n\phi(x)}{n} \geq \overline{\phi}(x) - \epsilon.$$

For any given $\delta > 0$, fix $m \in \mathbb{N}$ such that

$$\|A_m\phi - \overline{\phi}\|_{L^1(\eta)} < \epsilon\delta. \tag{1.21}$$

We write $\psi = A_m\phi - \overline{\phi} - \epsilon$,

$$E = \left\{x \in M : \sup_{n\in\mathbb{N}} S_n\psi(x) > 0\right\} = \left\{x \in M : \sup_{n\in\mathbb{N}} \frac{S_n(A_m\phi - \overline{\phi})(x)}{n} > \epsilon\right\}$$

and apply the maximal inequality as in (1.8) to obtain

$$\int_E (A_m\phi - \overline{\phi})\, d\eta > \epsilon\eta(E).$$

Combining with (1.21), we find that $\eta(E) < \delta$. To finish the proof in the case $\phi \in L^\infty(\eta)$, we proceed as in [11, p. 46–47] and use the following lemma.

Lemma 1.5. *If $\phi \in L^\infty(\eta)$ and $n \in \mathbb{N}$, then*

$$\lim_{n\to+\infty} \frac{1}{n}|S_n(A_m\phi)(x) - S_n\phi(x)| = 0. \tag{1.22}$$

for η-a.e. $x \in M$.

Proof. By a telescopic sum argument,

$$T^j\phi - \phi = S_j(T\phi - \phi).$$

We use this identity, the linearity and commutativity of the operators S_j, and the fact that the norm of T is 1, to obtain

$$S_n\Big(\frac{S_m\phi}{m}\Big)(x) - S_n\phi(v) = \frac{S_n}{m}(S_m\phi - m\phi)(x) = \frac{S_n}{m}\Big(\sum_{j=1}^{m-1} S_j(T\phi - \phi)(x)\Big)$$

$$= \frac{1}{m}\sum_{j=1}^{m-1} S_j(T^n\phi - \phi)(x) \leq \frac{1}{m}\frac{m(m-1)}{2} 2\|\phi\|_{L^\infty} = (m-1)\|\phi\|_{L^\infty}$$

for η-a.e. $x \in M$. Dividing by n and letting $n \to +\infty$, we obtain the desired inequality. $\square$

Using the fact that $\overline{\phi}$ is invariant by T, we have that, for η-a.e. $x \in M \backslash E$,

$$\limsup_{n\to+\infty} \frac{S_n\phi(x)}{n} = \limsup_{n\to+\infty} \frac{S_n(A_m\phi(x))}{n} = \limsup_{n\to+\infty} \frac{S_n(A_m(\phi - \overline{\phi}))(x)}{n} + \overline{\phi}(x)$$
$$\leq \overline{\phi}(x) + \epsilon.$$

Taking $\delta \to 0$, we obtain (1.20) in the case $\phi \in L^\infty(\eta)$. To extend to $\phi \in L^1(\eta)$, take $\phi_0 \in L^\infty(\eta)$ such that $\|\phi - \phi_0\|_{L^1} < \frac{\delta\epsilon}{2}$. We have $\|\overline{\phi} - \overline{\phi}_0\|_{L^1} < \delta\epsilon$ because $\|A_n\phi - A_n\phi_0\|_{L^1} \leq \|\phi - \phi_0\|_{L^1}$ for every $n \in \mathbb{N}$. Moreover, $A_n\phi(x) - \overline{\phi}(x) = A_n(\phi - \phi_0)(x) + (A_n\phi_0(x) - \overline{\phi}_0(x)) + (\overline{\phi}_0(x) - \overline{\phi}(x))$ and $\limsup_{n\to+\infty} |A_n\phi_0(x) - \phi_0(x)| = 0$ imply

$$\eta(\{x \in M : \limsup_{n\to+\infty} A_n\phi(x) > \overline{\phi}(x) + \epsilon\}) \leq \eta(\{x \in M : \overline{\phi}_0(x) - \overline{\phi}(x) > \tfrac{\epsilon}{2}\})$$
$$+\eta(\{x \in M : \sup_{n\in\mathbb{N}} A_n(\phi - \phi_0)(x) > \tfrac{\epsilon}{2}\}) < 2\delta,$$

where we applied Markov inequality to $\overline{\phi}_0 - \overline{\phi}$ and (1.8) to $A_n(\phi - \phi_0)$. Taking $\delta \to 0$ as before, this implies (1.20) and consequently (1.19), because $\epsilon > 0$ was arbitrary.

We turn now to (1.16). First, $\overline{\phi} = 0$ on D because $S_n\overline{\phi}$ is bounded on D by (1.12). Since

$$\chi_C T^k \phi = \chi_C T^k(\chi_C\phi) + \chi_C T^{k-1}(TI_D)\phi = \cdots = \chi_C \sum_{j=0}^{k} T^{k-j}(\chi_C(TI_D)^j\phi),$$

and $T^{*k}(\chi_B)\chi_C = \chi_B\chi_C$ for all $k \in \mathbb{N}$ and $B \in \mathcal{C}$, we conclude that

$$
\begin{aligned}
\int_B \overline{\phi}\, d\eta = \int_B \chi_C \overline{\phi}\, d\eta &= \lim_{n\to+\infty} \frac{1}{n} \int_B \chi_C \sum_{k=0}^{n-1} T^k \phi\, d\eta \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_B \chi_C \sum_{j=0}^{k} T^{k-j}(\chi_C (TI_D)^j \phi)\, d\eta \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_M \sum_{j=0}^{k} T^{*(k-j)}(\chi_B\chi_C)(\chi_C (TI_D)^j \phi)\, d\eta \\
&= \lim_{n\to+\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_M \sum_{j=0}^{k} \chi_B\chi_C (TI_D)^j \phi\, d\eta \\
&= \lim_{n\to+\infty} \int_M \sum_{j=0}^{n} \chi_B\chi_C (TI_D)^j \phi\, d\eta = \int_B H_C \phi\, d\eta,
\end{aligned}
$$

where we used the fact that H_C is a well-defined operator on L^1. This follows from the identity [28, p. 125, (3.3)]

$$
(I_C + TI_D)^k = \sum_{j=0}^{k-1} I_C (TI_D)^j + (TI_D)^k. \tag{1.23}
$$

In fact, there exist $\phi_j \in L^1(\eta)$ such that $\|\phi_j\|_{L^1} = 1$ and $I_C(TI_D)^j \phi_j \geq \|I_C(TI_D)^j\|_{L^1} - 2^{-j}$, hence $\psi_k = \max\{\phi_0, \ldots, \phi_k\}$ satisfies

$$
\begin{aligned}
\sum_{j=0}^{k-1} \|I_C(TI_D)^j\|_{L^1} &\leq (I_C + TI_D)^k \psi_k - (TI_D)^k \psi_k + \sum_{j=0}^{k-1} 2^{-j} \\
&\leq \|I_C + TI_D\|_{L^1}^k + \|TID\|_{L^1}^k + 2 \leq 4 < +\infty.
\end{aligned}
$$

Therefore, the series that defines H_C converges in the space of bounded operators on L^1.

Since η is an ergodic stationary probability if and only if it is ergodic and invariant for F (theorems 1.2 and 1.6), we can apply the result to T the transition operator $\tilde{\phi} \mapsto \tilde{\phi} \circ T$ and $\phi = \tilde{\phi} \circ \pi_M$, so that $\phi(F_\omega^j(x)) = T^j \tilde{\phi}(x)$. □

The pointwise ergodic theorem has many consequences, for example

Corollary 1.1. *If η_1 and η_2 are distinct ergodic stationary probabilities, then they're mutually singular (notation: $\eta_1 \perp \eta_2$), that is, there exist measurable sets B_1 and B_2 such that*

$$\eta_i(B_j) = \delta_{ij} := \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

Proof. η_1 and η_2 being distinct probabilities, there exists $\phi \in L^1(\eta_1) \cap L^1(\eta_2)$ such that

$$\int_M \phi \, d\eta_1 \neq \int_M \phi \, d\eta_2.$$

The pointwise ergodic theorem assures, for $i \in \{1, 2\}$, the existence of measurable sets B_i such that

$$\forall x \in B_i : \lim_{n \to +\infty} \frac{S_n \phi(x)}{n} = \int_M \phi \, d\eta_i$$

and $\eta_1(B_1) = 1 = \eta_2(B_2)$. The limits are distinct according to whether $x \in B_1$ or $x \in B_2$, therefore $B_1 \cap B_2 = \emptyset$ and $\eta_1(B_2) = 0 = \eta_2(B_1)$. □

Corollary 1.2. *If $\eta \ll \eta_0$, where η is a stationary probability and η_0 is an ergodic stationary probability of F, then $\eta = \eta_0$.*

Proof. Given an arbitrary $\phi \in L^1(\eta) \cap L^1(\eta_0)$, we apply the pointwise ergodic theorem to ϕ and η_0, and obtain a measurable set B with $\eta_0(B) = 1$ such that

$$\forall \omega \in B : \lim_{n \to +\infty} \frac{S_n \phi(\omega)}{n} = \int_M \phi \, d\eta_0.$$

If $\eta \ll \eta_0$, then $\eta(B) = 1$. Thus the theorem applied to ϕ and η gives

$$\int_M \phi \, d\eta = \int_M \int_M \phi \, d\eta_0 \, d\eta = \int_M \phi \, d\eta_0.$$

We conclude that $\eta = \eta_0$. □

These corollaries suggest that we may be able to decompose a stationary measure in terms of ergodic stationary measures. We will describe such a decomposition as a consequence of the following general theorem from convex analysis [8, p. 316, theorem 4.2].

Theorem 1.4 (Choquet)**.** *Let K be a compact convex metrizable subset of a locally convex topological linear space E. Let K_e be the set of extreme points* *of K. Then K_e is a G_δ in K and every $x \in K$ has a representation of the form*

$$x = \int y \, d\xi(y) \tag{1.24}$$

for some non-negative Baire measure ξ satisfying

$$\xi(K_e) = \xi(K) = 1.$$

Remark 1.2. The integral $\int y \, d\xi(y)$ denotes the unique element $w \in K$ that satisfies

$$F(w) = \int F(y) \, d\xi(y)$$

for every $F \in E^*$.

Theorem 1.5 (Ergodic decomposition)**.** *Suppose $F : \Omega \times M \to \Omega \times M$ is a random transformation, where M is a compact Hausdorff space. Then there is a probability measure ξ on $\mathcal{M}_s(M)$, the space of stationary probabilities on M with the weak-* topology, such that $\xi(K_e) = 1$ for K_e the set of ergodic stationary measures and*

$$\eta(B) = \int \kappa(B) \, d\xi(\kappa),$$

for every measurable $B \in \mathcal{B}$.

Proof. $\mathcal{M}_s(M)$ can be identified as a closed subset of the unit ball $B(0)$ of $C^0(M)^*$ with the weak-* topology. In fact, Riesz representation theorem [10, p. 265, theorem 3] identifies the space of signed Borel regular measures with $C^0(M)^*$ and

$$\mathcal{M}_s(M) = (I - U^*)^{-1}\{0\} \cap \{\eta \in C^0(M)^* : \eta(1) = 1, \eta(\phi) \geq 0 \text{ if } \phi \geq 0\} \cap B(0),$$

where U^* is the weak-* continuous operator given by $U^*\eta(\phi) = \eta(U\phi)$.

We verify that the assumptions of the Choquet's theorem hold.

1. $\mathcal{M}_s(M) \subset C^0(M)^*$, where $C^0(M)^*$ is a locally convex topological linear space [10, p. 419, lemma 3];

2. $M_s(M)$ is compact, because $B(0)$ is compact by Alaoglu's theorem [10, p. 424, theorem 2];

3. $M_s(M)$ is metrizable, because $B(0)$ is metrizable [10, p. 426, theorem 1], for $C^0(M)^*$ is a separable normed space by Stone-Weierstrass theorem [10, p. 437, exercise 17];

4. $\mathcal{M}_s(M)$ is convex, because for every $\eta_1, \eta_2 \in \mathcal{M}_s(M)$, $t \in (0,1)$ and $B \in \mathcal{B}$, $((1-t)\eta_1 + t\eta_2)(B) = (1-t)\eta_1(B) + t\eta_2(B) \in [0,1]$;

5. The set of ergodic stationary measures K_e is the set of extreme points of $\mathcal{M}_s(M)$.

Proof. If $\eta \in \mathcal{M}_s(M)$ is ergodic, suppose $\eta = (1-t)\eta_1 + t\eta_2$ for some $t \in (0,1)$ and η_1, η_2 stationary probabilities. We need to verify that $\eta_1 = \eta_0 = \eta_2$. This is a direct consequence of corollary 1.2, because the non-negativity of the measures and $t \in (0,1)$ imply $\eta_1 \ll \eta_0$ and $\eta_2 \ll \eta_0$.

If $\eta \in \mathcal{M}_s(M)$ is not ergodic, let S be an η-stationary set with $t = \eta(S) \in (0,1)$. Define a probability η_S by $\eta_S(B) = t^{-1}\eta(B \cap S)$ for every $B \in \mathcal{B}$. η_S is stationary because, for any bounded measurable ϕ,

$$\int U\phi \, d\eta_S = \int \chi_S U\phi \, d\eta = \int (U\chi_S)(U\phi) \, d\eta \overset{*}{=} \int \chi_S \phi \, d\eta = \int \phi \, d\eta_S,$$

where the equality marked with $*$ follows from the facts that $\|U\| = 1$ and $\chi_S, \phi \in L^2(\eta)$. Then $\eta = t\eta_S + (1-t)(\eta - \eta_S)$. □

Choquet's theorem guarantees the existence of a Baire measure ξ with the desired properties. $\mathcal{M}_s(M)$ is separable, for it is compact and metrizable [10, p. 22, theorem 15]. Thus Baire and Borel sets are the same [19, p. 218-219] and the measure is Borel regular [19, p. 239]. □

Corollary 1.3. *Theorem 1.5 applies also in the case where M is a complete and separable metric space (or a Polish space).*

Proof. Let $h : M \to h(M) \subset Q$ be a homeomorphism, where $Q = [0,1]^{\mathbb{N}}$ is the Hilbert cube and $h(M)$ is a G_δ subset of Q [38, p. 55, remark 2.2.8]. In particular, $h : M \to Q$ is measurable and $h(B) \subset Q$ is measurable for every $B \in \mathcal{B}$. Given any probability η on M, we can define a probability $h_*\eta$ on Q by $h_*\eta(C) = \eta(h^{-1}(C))$ for measurable $C \subset Q$.

We use the following lemma.

Lemma 1.6. *Let $F_Q : \Omega \times Q \to \Omega \times Q$ be a random transformation such that*

$$\forall x \in M : F_Q(\omega, h(x)) = (\sigma(\omega), h(F_\omega(x))).$$

Then η is stationary (or ergodic stationary) for F if and only if $h_*\eta$ is stationary (or ergodic stationary) for F_Q.

Proof. If $\psi : Q \to \mathbb{R}$ is bounded and measurable, then $\psi \circ h$ is bounded, with the same bound as ψ, and measurable because $h : M \to Q$ is measurable. Thus, for every $x \in M$,

$$U_Q(\psi)(h(x)) = \int_M \psi(h(F_\omega(x)))\, d\mu(\omega) = U(\psi \circ h)(x).$$

If $C \subset Q$ is measurable and η is a probability on M, then

$$\begin{aligned} {U_Q}^*(h_*\eta)(C) = \int_Q U_Q(\chi_C)\, d(h_*\eta) = \int_M U_Q(\chi_C) \circ h\, d\eta = \int_M U(\chi_C \circ h)\, d\eta \\ = \int_M U(\chi_{h^{-1}(C)})\, d\eta = U^*\eta(h^{-1}(C)) = h_*(U^*\eta)(C). \end{aligned}$$

It follows immediatly that η is stationary for F if and only if $h_*\eta$ is stationary for F.

For the statement concerning ergodicity, we note that if $\psi : Q \to \mathbb{R}$ is bounded and measurable, then

$$h_*\eta(\{w \in Q : U_Q\psi(w) = \psi(w)\}) = \eta(\{x \in M : U_Q(\psi)(h(x)) = \psi(h(x))\})$$

Along with $U_Q(\psi) \circ h = U(\psi \circ h)$, this implies that ψ is $(h_*\eta)$-stationary if and only if $\psi \circ h$ is η-stationary. In particular, a set C is $(h_*\eta)$-stationary if and only if $h^{-1}(C)$ is η-stationary. Since $(h_*\eta)(C) = \eta(h^{-1}(C))$, we have that $h_*\eta$ is ergodic if and only if η is ergodic. $\square$

From the random transformation F, we can obtain a random transformation as in the lemma. Let, for example,

$$F_Q(\omega, w) = \begin{cases} (\sigma(\omega), (h \circ F_\omega \circ h^{-1})(w))) & \text{if } w \in h(M); \\ (\sigma(\omega), w) & \text{otherwise.} \end{cases}$$

If η is a stationary probability of F, then $h_*\eta$ is a stationary probability of F_Q. Theorem 1.5 gives a probability measure ξ on $\mathcal{M}_s(Q)$ such that $\xi(K_e) = 1$, where K_e is the set of ergodic stationary measures of F_Q, and

$$h_*\eta(C) = \int \kappa(C)\, d\xi(\kappa) \tag{1.25}$$

for every measurable $C \subset Q$.

Denote by $\mathcal{M}_s^*(Q)$ the space of stationary probabilities κ of F_Q such that $\kappa(h(M)) = 1$. Since $h_*\eta(h(M)) = 1$, equation (1.25) implies $\xi(\mathcal{M}_s^*(Q)) = 1$. By lemma 1.6, we can view $\eta \mapsto h_*\eta$ as a mapping $h_* : \mathcal{M}_s(M) \to \mathcal{M}_s^*(Q)$

such that $(h_*)^{-1}(K_e)$ is the set of ergodic stationary probabilities of F. h_* is bijective, with inverse given by $\iota(\kappa)(B) = \kappa(h(B))$, because h is a homeomorphism.

Let $\iota_*\xi$ be the measure on $\mathcal{M}_s(M)$ defined by $\iota_*\xi(Z) = \xi(\iota^{-1}(Z))$. If we put $C = h(B)$ in (1.25), taking into account $\xi(\mathcal{M}_s^*(Q)) = 1$,

$$\eta(B) = \int \kappa(h(B))\, d\xi(\kappa) = \int_{\mathcal{M}_s^*(Q)} \iota(\kappa)(B)\, d\xi(\kappa) = \int \mu(B)\, d(\iota_*\xi)(\mu). \quad \square$$

A relationship analogous to proposition 1.2 between ergodicity in the context of deterministic and random dynamical systems holds.

Theorem 1.6 ([30, theorem 3.1])**.** *η is an ergodic probability of F if and only if $m := \mu \times \eta$ is an ergodic probability.*

Proof. Suppose a probability η is not ergodic for F, denote $m = \mu \times \eta$. Take B an η-stationary set with $\eta(B) \in (0, 1)$ and let $m_B = \mu \times (\eta|_B)$, where $\eta|_B$ is defined by $\eta|_B(B') = \eta(B' \cap B)$ for $B' \in \mathcal{B}$. m_B is invariant for F and $m_B \ll m$, with $m_B \neq m$. Then m is not ergodic for F, as corollary 1.2 applies to the deterministic case.

Now suppose instead that η is ergodic for F. Take E a $\mu\times\eta$-invariant set with positive measure. For $x \in M$, denote by E_x the slice $\{\omega \in \Omega : (\omega, x) \in E\}$. We claim that $G = \{x \in M : \mu(E_x) = 1\}$ has measure $\eta(G) = m(E)$. In fact, $\eta(G) = \int_M \chi_G(x)\, d\eta(x) \leq \int_M \mu(E_x)\, d\eta(x) = m(E)$ by proposition 1.1. The opposite inequality is due to the following lemma.

Lemma 1.7. *Let η be a probability measure on M and $m = \mu \times \eta$. For any measurable $E \subset \Omega \times M$ with $m(E) > 0$ and $\epsilon > 0$, there exists $n_0 \in \mathbb{N}$ such that*

$$\eta(\{x \in M : \mu((F^n E)_x) > 1 - \epsilon\}) \geq (1 - \epsilon) m(E)$$

for all $n \geq n_0$.

Proof. Let $\mathcal{A}_n = \mathcal{S}^{\otimes n} \otimes \{\varnothing, \Omega\}^{\otimes \mathbb{N}}$, the σ-algebra generated by the first n coordinates $\omega_0, \ldots, \omega_{n-1}$ of $\omega \in \Omega$. Let $\Phi_A^n(\omega) = \mu(\sigma^n(A \cap C_n(\omega)))$, where $C_n(\omega) := \{\omega' \in \Omega : \omega'_0 = \omega_0, \ldots, \omega'_{n-1} = \omega_{n-1}\} \in \mathcal{A}_n$.

$\Phi_A^n(\omega) = \mu(A \,|\, \mathcal{A}_n)(\omega)$ for μ-a.e. $\omega \in \Omega$, that is,

$$\forall C \in \mathcal{A}_n : \int_C \Phi_A^n(\omega)\, d\mu(\omega) = \mu(A \cap C),$$

because $\mu(C_n(\omega))\mu(\sigma^n(A \cap C_n(\omega))) = \mu(A \cap C_n(\omega))$, thus this equality holds for a σ-algebra that contains $\{C_n(\omega)\}_{\omega\in\Omega} \subset \mathcal{A}_n$. Therefore $\lim_{n\to+\infty} \Phi_A^n(\omega) =$

$\chi_A(\omega)$ for μ-a.e. $\omega \in \Omega$, by the forward martingale convergence theorem [9, p. 195].

Let $E_x^n = \{\omega \in \Omega : \Phi_{E_x}^n(\omega) > 1 - \epsilon\}$ and $G^n = \{(\omega, x) \in E : \omega \in E_x^n\}$. Note that $E_x^n \in \mathcal{A}_n$ and $G_x^n = E_x \cap E_x^n$. Then

$$m(G^n) = \int_M \mu(G_x^n)\, d\eta(x) = \int_M \int_{E_x^n} \mu(E_x \,|\, \mathcal{A}_n)(\omega)\, d\mu(\omega)\, d\eta(x)$$

Letting $n \to +\infty$, we obtain

$$\lim_{n \to +\infty} m(G^n) = \int_M \mu(E_x)\, d\eta(x) = m(E).$$

For any $w \in M$ such that $(F^n(G^n))_w \neq \varnothing$, let $(z, w) \in F^n(G^n)$. Then there exist $x \in M$ and $\omega \in G_x^n$ such that $z = \sigma^n(\omega)$ and $w = F_M^n(x)$. $\sigma^n(C_n(\omega)) = \{\sigma(\omega)\}$, so $(F^n(G^n))_w \supset \sigma^n(G_x^n \cap C_n(\omega))$ and we may assume

$$\Phi_{G_x^n}^n(\omega) = \mu(G_x^n \,|\, \mathcal{A}_n)(\omega) = \chi_{E_x^n}(\omega)\mu(E_x \,|\, \mathcal{A}_n)(\omega) = \chi_{E_x^n}(\omega)\Phi_{E_x}^n(\omega).$$

because $\mu(C_n(\omega)) > 0$ and these relations hold for μ-a.e. $\omega \in \Omega$. Hence

$$\mu((F^n(G^n))_w) \geq \mu(\sigma^n(G_x^n \cap C_n(\omega))) = \Phi_{G_x^n}^n(\omega) = \chi_{E_x^n}(\omega)\Phi_{E_x}^n(\omega) > 1 - \epsilon.$$

Let $\pi_M(G^n) = \{x \in M : G_x^n \neq \varnothing\}$, the projection of G^n onto M. Then $\pi_M(G^n) \subset \{x \in M : \mu((F^n E)_x) > 1 - \epsilon\}$. Since

$$\eta(\pi_M(G^n)) \geq m(G^n) \geq (1 - \epsilon)m(E)$$

for sufficiently large n, we obtain the lemma. □

The relation

$$m(E) = \int_M \mu(E_x)\, d\eta(x) = \mu(G) + \int_{M \setminus G} \mu(E_x)\, d\eta(x)$$

shows that $\mu(E_x) = 0$ for every x in some subset $G' \subset M \setminus G$ with measure $\eta(G') = 1 - \eta(G)$. Since

$$\mu(E_x) = \mu((F^{-1}E)_x) = \int_M \mu(E_{F_\omega(x)})\, d\mu(\omega),$$

$x \in G$ and $w \in G'$ if and only if $F_\omega(x) \in G$ and $F_\omega(w) \in G'$ for μ-a.e. $\omega \in \Omega$, which means that, for every x in the full η-measure set $G \cup G'$,

$$U(\chi_G)(x) = \int_M \chi_G(F_\omega(x))\, d\mu(\omega) = \chi_G(x).$$

We conclude that $\eta(G) = 1$, by the ergodicity of η. □

Chapter 2

The Perron-Frobenius operator

Fix some (Borel regular) probability of reference λ on M. If a random transformation $F : \Omega \times M \to \Omega \times M$ is *nonsingular* with respect to λ, that is, if for every probability measure q absolutely continuous to $\mu\times\lambda$ (notation: $q \ll \mu \times \lambda$), we have $F_* q \ll \mu \times \lambda$, then we can consider the effect of U^* on densities instead of measures because $\eta \ll \lambda$ implies $U^*\eta = \pi_{M,*}F_*(\mu \times \eta)$ by proposition 1.2 and $\pi_{M,*}F_*(\mu \times \eta) \ll \pi_{M,*}(\mu \times \lambda) = \lambda$.

Remark 2.1. If F is nonsingular, then for any ϕ and ψ bounded, measurable functions such that $B = \{x \in M : \phi(x) \neq \psi(x)\}$ has null measure,

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

This implies that $\{x \in M : U\phi(x) \neq U\psi(x)\}$ has null measure and U can be seen as an operator $U : L^\infty(M) \to L^\infty(M)$.

If η is stationary for F, then U can be either seen as an operator $U : L^\infty(\eta) \to L^\infty(\eta)$ by the same arguments above, taking into consideration proposition 1.2, or as an operator $U : L^1(\eta) \to L^1(\eta)$ because $\|U\phi\|_{L^1(\eta)} \leq \|\phi\|_{L^1(\eta)}$ for any $\phi \in L^\infty(\eta)$ and $\overline{L^\infty(\eta)} = L^1(\eta)$. Then interpolation of L^p measures shows that we can consider $U : L^p(\eta) \to L^p(\eta)$ for any $p \in [1, \infty]$.

Definition 2.1. The *Perron-Frobenius operator* associated to the nonsingular random transformation F is the operator $L : L^1(M) \to L^1(M)$ given by

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

where $\phi\lambda$ denotes the Borel measure with density ϕ, defined by

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

Remark 2.2. This definition extends the Perron-Frobenius operator associated to a deterministic $T : M \to M$, which can be viewed as a random transformation $F : \Omega \times M \to \Omega \times M$ where Ω is a trivial probability space.

An immediate consequence of the definition is that, for every $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$,

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

Identifying $L^\infty(M)$ as the dual space of $L^1(M)$, we can write $L^* = U$.

A sufficient condition for nonsingularity is given as follows.

Proposition 2.1. *If $(F_\omega)_*\lambda \ll \lambda$ for μ-a.e. $\omega \in S$, then F is nonsingular.*

Proof. If $A \times B \subset \Omega \times M$ is a $\mu \times \lambda$-null set, then $\mu(\sigma^{-1}(A)) = \mu(A) = 0$ or $\lambda(F_\omega^{-1}(B)) = \lambda(B) = 0$ for a.e. $\omega \in \Omega$. Thus

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

Remark 2.3. Since the null sets form a σ-algebra and the sets of the form $A \times B$ generate the σ-algebra of $\Omega \times M$, it suffices to verify the condition of absolute continuity for sets of this form.

An alternative approach for the Perron-Frobenius operator is given by transition probabilities.

Definition 2.2. Let $F : \Omega \times M \to \Omega \times M$ be a random transformation. If $x \in M \supset B$, the *transition probability* of x to B is given by

$$p(x, B) = U\chi_B(x) = \mu(\{\omega \in \Omega : F_\omega(x) \in B\}).$$

Proposition 2.2. *$(x, B) \in M \times \mathcal{B} \to p(x, B)$ defines a* transition kernel, *that is, for every $x \in M$, $B \mapsto p(x, B)$ defines a probability on M and for every $B \in \mathcal{B}$, $x \mapsto p(x, B)$ is measurable. If F is nonsingular and $\mathcal{B}$ is countably generated, then there exists a* transition density $p \in L^1(M \times M)$ *such that $p(x, B) = \int_B p(x, w) \, d\lambda(w)$ for λ-a.e. $x \in M$.*

Proof. We recall that, for every $x \in M$, the mapping $P_x : \omega \mapsto (\omega, x)$ is measurable (proposition 1.1). Since $\mu_x : B \mapsto p(x, B)$ satisfies $\mu_x(M) = \mu(\Omega) = 1$ and $\mu_x = (\pi_2 \circ F \circ P_x)_* \mu$, it is a probability on M. Moreover, for fixed $B \in \mathcal{B}$, $x \mapsto \mu_x(B) = U\chi_B(x)$ is measurable. We can thus define a measure μ over $M \times M$ by

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

We claim that $\mu \ll \lambda \times \lambda$. In fact, if $A \times B$ is a $\lambda \times \lambda$-null set, then $\lambda(A) = 0$ or $\lambda(B) = 0$. In the former case, $\mu(A \times B) = 0$ because A is a null set, in the latter because $p(x, B) = 0$ for a.e. $x \in M$ (remark 2.1). We conclude that μ admits a density $p \in L^1(M \times M)$.

The definition of μ implies that, for fixed $B \in \mathcal{B}$, $\mu_x(B) = \int_B p(x, w)\, d\lambda(w)$ for λ-a.e. $x \in M$. Since $\mathcal{B}$ admits a countable generator, we conclude that there exists a full measure set S of $x \in M$ such that $\mu_x(B) = \int_B p(x, w)\, d\lambda(w)$ for every $B \in \mathcal{B}$. □

The next proposition shows that transition densities define integral kernels for the transition and the Perron-Frobenius operator.

Proposition 2.3. *Suppose F is nonsingular and $\mathcal{B}$ is countably generated. If $\phi \in L^\infty(M)$, then for λ-a.e. $x \in M$,*

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

Proof. The first identity follows immediately for $\phi = \chi_B$ and thus for any $\phi \in L^\infty(M)$ because U is a bounded linear operator.

For the second identity, we note that for every measurable $B \subset M$,

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

where we've applied Fubini's theorem in the last step. The claim follows. □

Remark 2.4. Note that we require $\phi \in L^\infty(M)$ also in the second identity. In case $p \in L^\infty(M \times M)$, we may extend it to $\phi \in L^1(M)$ by continuity of L.

The following proposition from [32, remark 3.2.2] summarizes basic properties of the Perron-Frobenius operator. For any (η-a.e. defined) functions $\phi : M \to \mathbb{R}$, $\psi : M \to \mathbb{R}$, we say that $\phi \geq \psi$ or $\psi \leq \phi$ if $\phi(x) \geq \psi(x)$ for (η-almost) every $x \in M$ and denote the identically $c \in \mathbb{R}$ function by c.

Proposition 2.4. *The Perron-Frobenius L associated to a nonsingular transformation F has the following properties.*

1. *L is a linear operator on $L^1(M)$;*

2. *L is a positive operator, that is, $L\phi \geq 0$ if $\phi \geq 0$;*

3. *$\forall \phi \in L^1(M) : \int_M L\phi \, d\lambda = \int_M \phi \, d\lambda$;*

Proof. For the first item, we note that L is the composition of the linear operators $\phi \mapsto \phi\lambda$, which has as image the subspace of absolutely continuous measures; U^*, which preserves this subspace by the nonsingularity condition; and $\eta \mapsto \frac{d\eta}{dm}$, which is defined for absolutely continuous measures.

For the second item, it suffices to show that $U^*(\phi\lambda)$ is a nonnegative measure if $\phi \geq 0$, because the Radon-Nikodym derivative of an unsigned measure is unsigned. This follows readily from the fact that U is a positive operator, and thus U^* preserves the subspace of nonnegative measures.

For the third item, we note that $U1 = 1$ and calculate

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

□

2.1 Markov operators

Definition 2.3. A *Markov operator* on the measure space $(M, \mathcal{B}, \eta)$ is a mapping $P : L^1(\eta) \to L^1(\eta)$ that satisfies items 1–3 of proposition 2.4, that is, P is a positive, linear operator and $\int_M P\phi \, d\eta = \int_M \phi \, d\eta$ for all $\phi \in L^1(\eta)$.

For any function $\phi : M \to \mathbb{R}$, the positive and negative parts are defined by $\phi^+(x) = \max\{\phi(x), 0\}$ and $\phi^-(x) = \max\{-\phi(x), 0\}$. This definition extends naturally to η-a.e. defined functions.

Proposition 2.5 ([32, proposition 3.1.1])**.** *If $(M, \mathcal{B}, \eta)$ is a measure space and P is a Markov operator, then, for every $\phi \in L^1(\eta)$,*

1. *$(P\phi)^+(x) \leq P\phi^+(x)$ for η-a.e. $x \in M$;*

2. *$(P\phi)^-(x) \leq P\phi^-(x)$ for η-a.e. $x \in M$;*

3. *$|(P\phi)(x)| \leq P|\phi(\cdot)|(x)$ for η-a.e. $x \in M$;*

4. *$\|P\phi\| \leq \|\phi\|$, that is, P is* contractive.

Proof. Since $\phi^+ - \phi+ \geq 0$, positivity of P implies $P(\phi^+ - \phi) \geq 0$, thus $P\phi^+ \geq P\phi$. Along with $(P\phi)^+ \geq 0$, this proves that $(P\phi)^+ \leq P\phi^+(x)$. For the same reason, $(P\phi)^- \leq P\phi^-$. Linearity of P implies

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

Finally, as a consequence of $\int_M P\phi \, d\eta = \int_M \phi \, d\eta$ for all $\phi \in L^1(\eta)$,

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

Remark 2.5. In the case of the Perron-Frobenius operator L associated to $F : \Omega \times M \to \Omega \times M$, where M is compact, we also have $\|L\phi\|_{L^\infty} \leq \|\phi\|_{L^\infty}$ for $\phi \in L^\infty(M)$. In fact, for any $\phi \in L^\infty(M)$,

$$\forall B \in \mathcal{B} : \int_M \chi_B L\phi \, d\eta = \int_M U(\chi_B)\phi \, d\eta \leq \int_M |\phi| \, d\eta = \|\phi\|_{L^\infty}.$$

In the case $\eta = \lambda$, we denote by $\mathcal{D}$ the set of positive densities $h \in L^1(M)$ with L^1 norm $\|h\|_1 = 1$. Clearly, $P(\mathcal{D}) \subset \mathcal{D}$ for any Markov operator P.

Definition 2.4. A density $h \in \mathcal{D}$ such that $L(h\lambda) = h\lambda$ is called a *stationary density* for F. More generally, if P is a Markov operator, a density $h \in \mathcal{D}$ such that $Ph = h$ is called a stationary density of P.

The following proposition shows that every fixed point of L can be obtained from stationary densities.

Proposition 2.6 ([32, proposition 3.1.3])**.** *If P is a Markov operator and $P\phi = \phi$, then $P\phi^+ = \phi^+$ and $P\phi^- = \phi^-$.*

Proof. From $P\phi = \phi$, we have

$$\phi^+ = (P\phi)^+ \leq P\phi^+ \quad \text{and} \quad \phi^- = (P\phi)^- \leq P\phi^-.$$

Hence,

$$\begin{aligned}\int (P\phi^+ - \phi^+) \, d\eta + \int (P\phi^- - \phi^-) \, d\eta &= \int P(\phi^+ + \phi^-) - (\phi^+ + \phi^-) \, d\eta \\ &= \int P|\phi(\cdot)| - |\phi(\cdot)| \, d\eta = \int P|\phi(\cdot)| \, d\eta - \int |\phi(\cdot)| \, d\eta = 0.\end{aligned}$$

As a consequence of $(P\phi^+ - \phi^+) \geq 0$ and $(P\phi^- - \phi^-) \geq 0$, we have that $P\phi^+ - \phi^+ = 0$ and $P\phi^- - \phi^- = 0$. $\square$

Given a stationary density h for F, we can naturally obtain a Markov operator $P : L^1(h\lambda) \to L^1(h\lambda)$ with fixed point $1 := \chi_M$ in a similar fashion to the definition of the Perron-Frobenius operator.

Proposition 2.7. *Let F be a nonsingular random transformation with adjoint transition operator U^*. Take $h \in \mathcal{D}$ and $\eta = h\lambda$. Then $P_\eta : L^1(\eta) \to L^1(\eta)$ defined by*

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

is a Markov operator. If h is stationary, then $P_\eta 1 = 1$ and

$$\forall \phi \in L^1(\eta), n \in \mathbb{N} : P_\eta^n(\phi)h = L^n(\phi h).$$

Proof. The same arguments from the proof of proposition 2.4 show that P_η is a Markov operator (recall that by proposition 1.2, they are well-defined). If h is stationary, $P_\eta 1 = 1$ because $U^*\eta = \eta$; and, for every $\phi \in L^1(\eta)$, we have $\phi h \in L^1(M)$ and

$$\forall B \in \mathcal{B} : \int_B P_\eta^n(\phi)h \, d\lambda = (U^*)^n(\phi\eta)(B) = \int_B L^n(\phi h) \, d\lambda.$$

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

These operators can be used to define a mixing property.

Definition 2.5. We say that F is *mixing* for $\eta = h\, d\lambda$ if

$$\lim_{n\to\infty} \int_M P_\eta^n(\phi)\psi \, d\eta = \int_M \phi \, d\eta \int_M \psi \, d\eta \tag{2.1}$$

for every $\phi \in L^1(\eta)$ and $\psi \in L^\infty(\eta)$

A more natural and often more convenient formulation is given by

$$\lim_{n\to\infty} \int_M L^n(\phi)\psi \, d\lambda = \int_M \phi \, d\lambda \int_M \psi h \, d\lambda \tag{2.2}$$

for every $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$.[1].

These notions are related as follows.

Proposition 2.8. (2.2) *implies* (2.1) *and they're equivalent if* $\operatorname{supp} \eta = M$.

[1] This definition is adapted from [4] (Equation (1.5)).

Proof. Suppose 2.2 holds. If $\phi \in L^1(\eta)$ and $\psi \in L^\infty(\eta)$, then $\phi h \in L^1(M)$ and $\psi\chi_{h>0} \in L^\infty(M)$. Substituting these functions into 2.2, we obtain

$$\lim_{n\to\infty} \int_M L^n(\phi h)\psi\chi_{h>0}\, d\lambda = \int_M \phi h\, d\lambda \int_M \psi\chi_{h>0} h\, d\lambda = \int_M \phi\, d\eta \int_M \psi\, d\eta.$$

The first integral is equal to $\int_M P^n(\phi)\psi\, d\eta$ and this implies 2.1.

Reciprocally, suppose 2.1 holds and $\operatorname{supp}\eta = M$. If $\phi \in L^1(M)$ and $\psi \in L^\infty(M)$, then $\phi = \hat{\phi}h$ for some $\hat{\phi} \in L^1(\eta)$, because $\operatorname{supp}\eta = M$, and $\psi \in L^\infty(\eta)$. Substituting these functions into 2.1, we obtain

$$\lim_{n\to\infty} \int_M P^n(\hat{\phi})\psi\, d\eta = \int_M \hat{\phi}\, d\eta \int_M \psi\, d\eta = \int_M \phi\, d\lambda \int_M \psi h\, d\lambda.$$

The first integral is equal to $\int_M L^n(\phi)\psi\, d\lambda$ and this concludes the proof. □

In view of Radon-Nikodym theorem, the next theorem, presented in [32, theorem 4.2.2] in the deterministic case, but valid also in the random case, generalizes corollary 1.2. It shows that if η_0 is an ergodic probability, then there is at most one stationary probability such that $\eta \ll \eta_0$. Additionally, if a probability η_0 has the property that $\eta_0 \ll \eta$ for every stationary $\eta \ll \eta_0$, then η_0 is ergodic.

Proposition 2.9. *Let F be a nonsingular random transformation. If η is ergodic for F (definition 1.6), then there is at most one stationary density of P_η. Further, if there is a unique stationary density h of η and $h > 0$ η-a.e., then η is ergodic.*

Proof. Suppose η is ergodic and ϕ_1, ϕ_2 are stationary densities of F. Then $\psi = \phi_1 - \phi_2$ satisfies $P_\eta\psi = \psi$ and proposition 2.6 implies that $P_\eta\psi^+ = \psi^+$ and $P_\eta\psi^- = \psi^-$. Set $B^\pm = \{x \in M : \psi^\pm(x) = 0\}$ and note that $U\chi_B\pm = \chi_B^\pm$. This follows from the fact that $U^*(\phi^\pm\eta) \ll \phi\eta$ for all $\phi \in L^1(m)$ by the nonsingularity condition. Ergodicity implies that $\eta(\chi_{B^+}) = 0$ or 1 and similarly for χ_{B^-}. Since their union is M, they can't both have null measure. Together with the fact that $\int_M \phi_1\, d\eta = \int_M \phi_2\, d\eta$, we conclude that both have full measure, as well as the intersection $\{x \in M : \psi(x) = 0$. Thus $\phi_1 = \phi_2$.

Now, suppose P_η admits a unique stationary density $h > 0$. Given any η-stationary $B \in \mathcal{B}$, set $B' = M \setminus B$ and write $h = \chi_B h + \chi_{B'} h$. Then

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

$L(\chi_B h) = 0$ on B' because $L(\chi_B h) \geq 0$ and

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

Thus $L(\chi_B h) = \chi_B h$ and the uniqueness of the stationary density implies that $\chi_B = 0$ or $\chi_B = 1$. □

The following characterization of ergodicity and mixing, analogous to [32, theorem 4.4.1], shows that mixing is stronger than ergodicity.

Proposition 2.10. *Suppose F is a random transformation and η is a stationary measure for F. Then F is ergodic (or mixing) for η if and only if the sequence $\{P_\eta^j \phi\}_{j\in\mathbb{N}}$ is Cesàro (or weakly) convergent to 1 for all $\phi \in \mathcal{D}_\eta := \{\psi \in L^1(\eta) : \int_M \psi\, d\eta = 1\}$.*

Proof. Cesàro convergence of $\{P_\eta^j \phi\}_{j\in\mathbb{N}}$ means pointwise convergence of the Birkhoff averages

$$A_n\phi := \frac{1}{n}\sum_{j=0}^{n-1} P_\eta^j \phi. \tag{2.3}$$

Theorem 1.2 implies that the pointwise limit $\overline{\phi}$ of these averages is a stationary density of P_η and $\int \phi\, d\eta = \int \overline{\phi}\, d\eta$. Thus it is 1 for all $\phi \in \mathcal{D}_\eta$ if and only if every η-stationary $\overline{\phi}$ is equal to $\int \overline{\phi}\, d\eta$; equivalently, if and only if η is ergodic.

As for the mixing property, since the dual of $L^1(\eta)$ is identified with $L^\infty(\eta)$, the claim is simply a rephrasing of (2.1). □

In our application, we shall verify the following stronger condition.

$$\|L^k|_V\|_{L^1(M)\to L^1(M)} \to 0, \qquad V = \Big\{f \in L^1(M) : \int f\, dm = 0\Big\}. \tag{2.4}$$

Remark 2.6. It suffices to verify that $\|L^k|_V\|_{L^1(M)\to L^1(M)} < 1$ for some $k \in \mathbb{N}$, because $\|L_\xi\|_{L^1} = 1$.

2.2 Systems perturbed by an additive noise

Let $M = \mathbb{T}^n = (\mathbb{S}^1)^n$, the n-dimensional torus, identified, as a measure space, with $[0,1]^n$ with the Lebesgue measure λ. Consider the family of rotations $\tau_t : L^1(\lambda) \to L^1(\lambda)$, $\tau_t f(x_1, \ldots, x_n) = f(\pi(x_1 - t_1, \ldots, x_n - t_n))$, for $t = (t_1, \ldots, t_n) \in \mathbb{R}^n$ and $\pi(x_1, \ldots, x_n) := (x_1 - \lfloor x_1 \rfloor, \ldots, x_n - \lfloor x_n \rfloor)$, where $\lfloor a \rfloor = \min\{k \in \mathbb{Z} : k \le a\}$. Note in particular that τ_t is the identity for any $t \in \mathbb{Z}^n$.

If $f, g \in L^1(\lambda)$, we define the convolution $f * g$ of f and g by

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

Note that the rightmost equality follows from the translational invariance of Lebesgue measure. Moreover, $f * g \in L^1(\lambda)$ as a particular case of the inequality [37, theorem 7.14]

$$\forall p, q \in L^1(\mathbb{R}^n) : \int_{\mathbb{R}^n} \Big| \int_{\mathbb{R}^n} p(x-y) q(y)\, dy \Big|\, dx \le \|p\|_{L^1(\mathbb{R}^n)} \|q\|_{L^1(\mathbb{R}^n)}. \quad (2.5)$$

We will discuss the special case of a random dynamical system that can be viewed as a perturbation by an additive noise of a deterministic dynamical system [32, section 10.5]. That is, given $r \in L^1(\mathbb{R}^n)$ such that $\operatorname{supp} r \subset B_{1/2}(0)$, the open ball of radius 1/2 centered at 0, with mean value 1 and $r(x) = r(-x)$ for $x \in \mathbb{R}^n$, define a one-parameter family of functions in $L^1(\mathbb{R}^n)$ with mean value 1 and support contained in $B_{\xi/2}(0)$ by

$$r_\xi(x) = \xi^{-1} r(x/\xi), \qquad \xi \in (0, 1].$$

Let ρ_ξ be the unique function in $L^1(\lambda)$ such that $\rho_\xi \circ \pi = r$ on $B_{1/2}(0)$. Define a one-parameter family of operators $N_\xi : L^1(\lambda) \to L^1(\lambda)$ by the convolution

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

For a fixed nonsingular $T : M \to M$, we obtain a family of random dynamical systems by the skew-product

$$T_\xi : \Omega \times M \to \Omega \times M, \quad T_\xi(\omega, x) = (\sigma(\omega), \pi(T(x) + \omega_0)),$$

where we take S the ball with radius $\xi/2$ and p the probability measure with density given by ρ_ξ, with the same notations as in chapter 1. In this case, we can obtain the following expression of the Perron-Frobenius operator.

Proposition 2.11. *The Perron-Frobenius operator L_ξ of T_ξ is $N_\xi L_0$, where L_0 is the Perron-Frobenius operator of T (see remark 2.2).*

Proof. For any $\phi \in L^1(M)$ and measurable $B \subset M$, keeping the usual notations for $F = T_\xi$,

$$\int_B L_\xi \phi\, d\lambda = \int_\Omega \int_{T^{-1}(B-\omega_0)} \phi\, d\lambda d\mu(\omega) = \int_S \int_{T^{-1}(B-t)} \phi(x) r_\xi(t)\, d\lambda(x)\, dt$$
$$= \int_S \int_B L_0 \phi(x+t) r_\xi(-t)\, d\lambda(x)\, dt = \int_B N_\xi L_0 \phi\, d\lambda. \quad \square$$

In order to calculate $L_0\phi$ for $\phi \in L^1(\lambda)$, we consider the following smooth manifold structure on M, identified as a set with $\pi(\mathbb{R}^n) = [0,1)^n$. We take as open sets the images $\pi(O)$ of open $O \subset \mathbb{R}^n$ and, given any $x \in M$, for the neighborhood $\pi(B_{1/2}(x))$, we have a chart $\zeta_x : \pi(B_{1/2}(x)) \to B_{1/2}(x)$ because $\pi|_{B_{1/2}(x)}$ is injective.

We can thus naturally consider the set $C^1(M,M)$ of C^1 functions from M to M, that is, of functions $T : M \to M$ such that

$$\zeta_{T(x)} \circ T \circ \pi|_{B_s(x)} : B_s(x) \to \mathbb{R}^n$$

is of class C^1, where $s > 0$ is taken so small that $T(\pi(B_s(x))) \subset \pi(B_{1/2}(T(x)))$.

Proposition 2.12. *Suppose $T \in C^1(M,M)$. If $\phi \in L^1(\lambda)$, then*

$$L_0\phi(x) = \sum_{T(y)=x} \frac{\phi(y)}{|\det T'(y)|} \tag{2.6}$$

for a.e. $x \in M$ such that $\#T^{-1}(x) < +\infty$ and $|\det T'(y)| > 0$ for $y \in T^{-1}(x)$.

Proof. Lift T to some $S \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ such that $\pi(S(x)) = \pi(S(\pi(x)))$ for all $x \in \mathbb{R}^n$ and extend ϕ to $L^1(\mathbb{R}^n)$ setting $\phi(x) = 0$ if $x \notin M$.

Take $B \in \mathcal{B}$ a ball centered at x sufficiently small that $|\det T'(y)| > \epsilon$ for all $y \in T^{-1}(B)$. Changing variables as in [12, theorem 3.9], we get, for any $B \in \mathcal{B}$,

$$\int_B L_0\phi(x)\, d\lambda(x) = \int_{T^{-1}(B)} \phi(x)\, d\lambda(x) = \int_B \sum_{T(y)=x} \frac{\phi(y)}{|\det T'(y)|}\, d\lambda(x),$$

because $\phi(y) = 0$ if $y \notin M$. □

2.2.1 Noises of BV type

In chapter 4, the noise will be considered in the space $BV(M)$ defined as follows.

Definition 2.6. If $f \in L^1(U)$, where $U \subset \mathbb{R}^n$ is an open set, the *variation* of f on U is defined by [12, definition 5.1]

$$\|Df\|(U) = \sup_{\phi} \int_U f(x)\,\mathrm{div}\,\phi(x)\, dx,$$

where the supremum is taken over all $\phi \in C_c^1(U, \mathbb{R}^n)$, that is, continuosly differentiable functions $\phi : U \to \mathbb{R}^n$ such that $\overline{\{\phi > 0\}}$ is compact, that satisfy $\sup_{x\in U} |\phi(x)| \leq 1$. Given $f \in L^1(\lambda)$, consider its extension $\hat{f} = f \circ \pi$ to $\mathbb{R}^n$, where $\pi : \mathbb{R}^n \to [0,1)^n$ is the retraction used in the definition of the rotations $\tau_\cdot$. We define

$$\operatorname{var}(f) = \lim_{U \downarrow M} \|D\hat{f}\|(U) = \inf_{U \supset M} \|D\hat{f}\|(U), \tag{2.7}$$

where the limit and infimum are taken over the open sets $U \supset M$. The variation of f on $I \subset M$, $\operatorname{var}_I(f)$, is defined as the variation of $f\chi_I$. We denote by $BV(M)$ the set of $f \in L^1(\lambda)$ such that $\operatorname{var}(f) < +\infty$.

Remark 2.7. Our definition is made in order that, when "gluing" oposite sides of the n-dimensional cube, the variation of f in the passage is taken into account. It is well defined because $\|Df\|(U) \subset \|Df\|(V)$ if $U \subset V$.

In the following, we list some properties of the variation.

Proposition 2.13. *In the context of definition 2.6, the following holds.*

1. var *is a seminorm on* $BV(M)$*; that is, for any* $f, g \in BV(M)$ *and* $\alpha \in \mathbb{R}$*, we have* $\operatorname{var}(f+g) \leq \operatorname{var}(f) + \operatorname{var}(g)$ *and* $\operatorname{var}(\alpha f) = |\alpha| \operatorname{var}(f)$*;*
2. *if* $f \in C^1(M)$ *(in the sense that* $f \circ \pi \in C^1(\mathbb{R}^n)$*), then* $\operatorname{var}(f) = \int |\nabla f|\, d\lambda$ *(cp. with [12, p. 197–198]);*
3. *if* $f \in BV(M)$*, then there exists a sequence of* $f_i \in C^\infty(M)$ *such that* $f_i \to f$ *in* $L^1(\lambda)$ *and* $\operatorname{var}(f_i) \to \operatorname{var}(f)$ *(cp. with [12, theorem 5.3]);*
4. *for all* $t \in \mathbb{R}^n$ *and* $f \in BV(M)$*,* $\tau_t f \in BV(M)$ *with* $\operatorname{var}(\tau_t f) = \operatorname{var}(f)$*;*
5. *in the one-dimensional case, for every* $f : [0,1] \to \mathbb{R}$*, we have*

$$\operatorname{var}(f) = \sup_{\{x_i\}_{j=0}^k} \sum_{j=0}^{k} |f(x_{j+1}) - f(x_j)|, \tag{2.8}$$

where the supremum is taken over all increasing sequences $\{x_j\}_{j=0}^k$ *in* $[0,1]$ *such that each* x_j *is a point of approximate continuity of* f*, that is,*

$$f(x_j) = \frac{f(x_j^-) + f(x_j^+)}{2}, \qquad f(x_j^\pm) = \lim_{\substack{h \to 0 \\ h > 0}} f(\pi(x_j \pm h)),$$

and $x_k := x_0$ *(cp. with [12, theorem 5.21]);*

6. *if $f \in BV(\mathbb{S}^1)$, then there exist increasing functions $f_1, f_2 : [0,1] \to \mathbb{R}$ such that $f(x) = f_1(x) - f_2(x)$ for every point x of approximate continuity of f (cp. with [36, p. 103]).*

Proof. 1. If $f, g \in BV(M)$ and $U \supset M$ is open, then

$$\begin{aligned}\|D(\hat{f}+\hat{g})\|(U) &= \sup_\phi \int_U (\hat{f}+\hat{g})(x) \operatorname{div} \phi(x)\, dx \\ &\leq \sup_\phi \int_U \hat{f}(x) \operatorname{div} \phi(x)\, dx + \sup_\psi \int_U \hat{g}(x) \operatorname{div} \psi(x)\, dx \\ &= \|D\hat{f}\|(U) + \|D\hat{g}\|(U),\end{aligned}$$

where the suprema, here and in the following, are taken over all $\phi, \psi \in C_c^1(U, \mathbb{R}^n)$ such that $\|\phi\|_{L^\infty} \leq 1$, and $\tilde{h} := h \circ \pi$ for $h \in C^1(M)$. Taking $U \downarrow M$, we obtain $\operatorname{var}(f+g) \leq \operatorname{var}(f) + \operatorname{var}(g)$.

If $f \in BV(M)$, $\alpha \geq 0$ and $U \supset M$ is open, then

$$\begin{aligned}\|D(\alpha\hat{f})\|(U) &= \sup_\phi \int_U (\alpha\hat{f})(x) \operatorname{div} \phi(x)\, dx \\ &= \sup_\phi \Big(\alpha \int_U \hat{f}(x) \operatorname{div} \phi(x)\, dx\Big) = \alpha \|D\hat{f}\|(x).\end{aligned}$$

Since $\phi \in C_c^1(U)$ if and only if $-\phi \in C_c^1(U)$, we have

$$\begin{aligned}\|D(-\hat{f})\|(U) &= \sup_\phi \int_U -\hat{f}(x) \operatorname{div} \phi(x)\, dx = \sup_\phi \int_U \hat{f}(x) \operatorname{div}(-\phi(x))\, dx \\ &= \sup_\psi \int_U \hat{f}(x) \operatorname{div} \psi(x)\, dx = \|D\hat{f}\|(U).\end{aligned}$$

Taking $U \downarrow M$ and combining these two identities, we conclude that $\operatorname{var}(\alpha f) = |\alpha| \operatorname{var}(f)$ for any $\alpha \in \mathbb{R}$.

2. Let $U \supset M$ and consider $g \in L^1(\mathbb{R}^n, \mathbb{R}^n)$ (in the sense that each coordinate function is in $L^1(\mathbb{R}^n)$) defined by

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

We have that $\sup_{x \in \mathbb{R}^n} |g(x)| \leq 1$. Fix $\rho \in C^\infty(\mathbb{R}^n)$ such that $\overline{\{\rho > 0\}} = B_{1/2}(0)$, the open ball of radius $1/2$ centered at 0, and $\int \rho\, d\lambda = 1$. Define, for every $\xi > 0$, $\rho_\xi \in C^\infty(\mathbb{R}^n)$ by $\rho_\xi(t) = \xi^n \rho(t/\xi)$.

Consider the sequence given by

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

Since M is compact, there exists $i_0 \in \mathbb{N}$ such that $B_{1/i_0}(x) \subset U$ for every $x \in M$. For $i \geq i_0$, $\phi_i \in C_c^1(U, \mathbb{R}^n)$ and $\phi_i \to g$ in $L^1(U, \mathbb{R}^n)$, a consequence of corollary 3.3 and Lebesgue's density theorem, respectively, with $\sup_{x \in U} |\phi_i(x)| \leq \int \rho_{1/i}(t) \sup_{y \in \mathbb{R}^n} |g(y)|\, dt = 1$. Thus, by Stoke's theorem,

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

Hence $\|D\hat{f}\|(U) \geq \int_M |\nabla f(x)|\, dx$. On the other hand, if $\phi \in C^1(U, \mathbb{R}^n)$ satisfies $\sup_{x \in U} |\phi(x)| \leq 1$, then

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

thus taking the supremum over all such ϕ, we get $\|D\hat{f}\|(U) \leq \int_U |\nabla \hat{f}(x)|\, dx$. Taking $U \downarrow M$, we conclude that $\operatorname{var}(f) = \int_M |\nabla f(x)|\, dx$.

3. Consider the open sets $U_i = \cup_{x \in M} B_{1/i}(x)$, for $k \in \mathbb{N}$ and the functions $\rho_{1/i} \in C^\infty(\mathbb{R}^n)$ defined in the previous item. Define a sequence of functions by

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

For each fixed $i_0 \in \mathbb{N}$, we have for $i \geq i_0$, $f_i \in C^\infty(U_{i_0})$ and $f_i \to \hat{f}$ in $L^1(U_{i_0})$, in particular $f_i|_M \to f$ in $L^1(\lambda)$. Moreover, if $x, y \in \mathbb{R}^n$, $\pi(x) = \pi(y)$, then $\partial^\alpha f_i(x) = \partial^\alpha f_i(y)$ for any $\alpha \in \mathbb{N}_0^n$, because $f(x-t) = f(y-t)$ for any $t \in \mathbb{R}^n$. Hence $f_i|_M \in C^\infty(M)$.

We have, for all $\phi \in C_c^1(U_{i_0}, \mathbb{R})$,

$$\begin{aligned}
\int_{U_{i_0}} f_i \operatorname{div} \phi\, dx &= \int_{U_{i_0}} \Big(\int \rho_{1/i}(t) \tau_t f(\cdot)\, dt \Big)(x) \operatorname{div} \phi(x)\, dx \\
&= \int_{U_{i_0}} \hat{f}(x) \operatorname{div} \Big(\int \rho_{1/i}(t) \tau_t \phi(\cdot)\, dt \Big)(x)\, dx \\
&\leq \|D\hat{f}\|(U_{i_0}),
\end{aligned}$$

hence $\operatorname{var}(f_i|_M) \leq \|D\hat{f}\|(U_{i_0}) \to \operatorname{var}(f)$.

Conversely, $\mathrm{var}(f) \leq \liminf_{i\to+\infty} \mathrm{var}(f_i|_M)$ because, for any $\epsilon > 0$, we can find, for each $i_0 \in \mathbb{N}$, some $\phi \in C_c^1(U_{i_0}, \mathbb{R}^n)$ such that

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

$$\therefore \liminf_{i\to+\infty} \|D(f_i|_M)\|(U_{i_0}) \geq \|D\hat{f}\|(U_{i_0}) - \epsilon \geq \mathrm{var}(f) - \epsilon,$$

from which we conclude that $\liminf_{i\to+\infty} \mathrm{var}(f_i|_M) \geq \mathrm{var}(f) - \epsilon$.

4. Given $f \in BV(M)$, take the sequence of approximations $\{f_i\}_{i\in\mathbb{N}}$ obtained in the proof of the previous item. Clearly, $\{\tau_t f_i\}_{i\in\mathbb{N}}$ is the corresponding sequence to $\tau_t f$, so that $\mathrm{var}(\tau_t f_i) \to \mathrm{var}(\tau_t f)$. Moreover,

$$\mathrm{var}(\tau_t f_i) = \int_M |\nabla(\tau_t f_i)|\, d\lambda = \int_M \tau_t |\nabla f_i|\, d\lambda = \int_M |\nabla f_i|\, d\lambda = \mathrm{var}(f_i)$$

by item 2. Therefore $\mathrm{var}(\tau_t f) = \mathrm{var}(f)$.

5. Denote by $V(f)$ the right-hand side of (2.8). Consider the sequence $\{f_i\}_{i\in\mathbb{N}}$ defined in item 3. We have that $f_i \to f$ in $L^1(\lambda)$. Take any sequence of points $\{x_j\}_{j=0}^k$ where f is approximately continuous and set $x_k = x_0$. Since λ-a.e. $x \in [0,1]$ is a point of approximate continuity of f, we have that, for λ-a.e. $s \in [0,1]$, all $\pi(x_j - s)$ are points of approximate continuity of f. Also, up to a cyclical permutation, $\{\pi(x_j - s)\}_{j=0}^{k-1}$ is increasing for $j \in \{0, \ldots, k-1\}$. Thus

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

Since this inequality holds for any increasing sequence $\{x_j\}_{j=0}^{k-1}$, it follows that $V(f_i|_M) \leq V(f)$. Thus $\mathrm{var}(f_i|_M) \leq V(f)$ by the following lemma.

Lemma 2.1. *If $g \in C^\infty(\mathbb{S}^1)$, then $V(g) = \mathrm{var}(g)$, where*

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

the supremum being taken over all increasing sequences $\{x_i\}_{j=0}^{k-1}$ in $[0,1]$ of points of approximate continuity of f.

Proof. Take any increasing sequence $\{x_j\}_{j=0}^{k-1}$ in $[0,1]$ of points of approximate continuity of f and set $x_k = x_0$. Note that $g(x_0) = \hat{g}(1+x_0)$ for the periodic extension $\hat{g} \in C^\infty(\mathbb{R})$ of g. Set $y_j = x_j$ for $j \leq k-1$ and $y_k = 1+x_0$. Then

$$\sum_{j=0}^{k-1} |g(x_{j+1}) - g(x_j)| = \sum_{j=0}^{k-1} \Big| \int_{y_j}^{y_{j+1}} \nabla \hat{g}(y)\, dy \Big| \leq \sum_{j=0}^{k-1} \int_{y_j}^{y_{j+1}} |\nabla \hat{g}(y)|\, dy$$
$$\leq \int_{y_0}^{1+y_0} |\nabla \hat{g}(y)|\, dy = \int_0^1 |\nabla \hat{g}(y)|\, dy = \int |\nabla g|\, d\lambda = \mathrm{var}(g).$$

Therefore

$$V(g) \leq \mathrm{var}(g).$$

Now fix $\epsilon > 0$. The equivalence of Riemann and Lebesgue integrals implies that there exists $\delta > 0$ such that, for any increasing sequence $\{y_j\}_{j=0}^{k+1}$ in $[0,1]$ satisfying $y_0 = 0$, $y_{k+1} = 1$ and $|y_{j+1} - y_j| < \delta$ for $j \in \{0, \ldots, k\}$, we have

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

Additionally, by the compactness of $[0,1]$ and continuity of g, δ can be chosen in order that each term in the sum is less than ϵ.

Since a.e. $x \in [0,1]$ is a point of approximate continuity of f, there is an increasing sequence $\{x_j\}_{j=0}^{k-1}$ of such points satisfying $x_0 < \delta$, $x_k > 1 - \delta$ and $|x_{j+1} - x_j| < \delta$ for $j \in \{0, \ldots, k-2\}$. Therefore

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

where $x_k = x_0$, $y_0 = 0$, $y_{k+1} = 1$ and $y_j = x_{j-1}$ for $j \in \{1, \ldots, k\}$. We conclude that $V(g) \geq \mathrm{var}(g)$, thus $V(g) = \mathrm{var}(g)$. □

Since $f_i|_M \to f$ in $L^1(\lambda)$ and $\mathrm{var}(f_i|_M) \leq V(f)$, we have that

$$\mathrm{var}(f) = \lim_{i \to +\infty} \mathrm{var}(f_i|_M) \leq V(f).$$

If x is a point of approximate continuity of f, then $f_i(x) \to f(x)$. Hence, for any increasing sequence $\{x_j\}_{j=0}^{k-1}$ in $[0,1]$ of points of approximate continuity of f and $x_k = x_0$,

$$\begin{aligned}\sum_{j=0}^{k-1} |f(x_{j+1}) - f(x_j)| &= \lim_{i\to+\infty} \sum_{j=0}^{k-1} |f_i(x_{j+1}) - f_i(x_j)| \\ &\le \lim_{i\to+\infty} \mathrm{var}(f_i|_M) = \mathrm{var}(f),\end{aligned}$$

from which we conclude that $V(f) \le \mathrm{var}(f)$, thus $V(f) = \mathrm{var}(f)$.

6. Set $c = \liminf_{y\to 0} f(y)$, where y runs over the points of approximate continuity of f. Define $f_1, f_2 : [0,1] \to \mathbb{R}$ by

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

where the suprema are taken over all increasing sequences $\{x_j\}_{j=0}^{k}$ in $[0,1]$ of points of continuity of f such that $x_j \le x$. Clearly, f_1 and f_2 are increasing functions. If x is any point of approximate continuity and $\epsilon > 0$ is arbitrary, take increasing sequences $\{x_j^1\}_{j=0}^{k-1}$ and $\{x_j^2\}_{j=0}^{l-1}$ in $[0,1]$, upper bounded by x, such that

$$\begin{aligned}f_1(x) &\ge c + \sum_{j=0}^{k-1} \max\{f(x_{j+1}^1) - f(x_j^1), 0\} > f_1(x) - \epsilon \\ f_2(x) &\ge \sum_{j=0}^{l-1} \max\{f(x_j^2) - f(x_{j+1}^2), 0\} > f_2(x) - \epsilon.\end{aligned}$$

Take $x_0 < \min\{x_0^1, x_0^2\}$ such that $|f(x_0) - c| < \epsilon$. Since $\max\{a+b, 0\} \le \max\{a, 0\} + \max\{b, 0\}$ and $\max\{a, 0\} - \max\{-a, 0\} = a$ for $a, b \in \mathbb{R}$, we can join these sequences and $\{x_0, x\}$ into one increasing sequence $\{x_j\}_{j=0}^{m-1}$ such that, setting $a_j = f(x_{j+1}) - f(x_j)$, we have

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

Therefore

$$|f(x) - (f_1(x) - f_2(x))| \le \left|c + \sum_{j=0}^{m-1} a_j - (f_1(x) - f_2(x))\right| + |f(x_0) - c| < 2\epsilon.$$

We conclude that $f(x) = f_1(x) - f_2(x)$. $\square$

The following proposition, a particular case of [24, proposition 1.4.7], shows that $L_\xi(BV(M)) \subset BV(M)$.

Proposition 2.14. *If $\rho \in BV(M)$ and $f \in L^1(M)$, then $\rho * f \in BV(M)$.*

Proof. Since $BV(M) \subset L^1(M)$, we have that $\rho * f \in L^1(M)$, by (2.5). Given any $\phi \in C^1_c(M)$, $|\phi| \le 1$, we have

$$\begin{aligned}
\int (\rho * f) \operatorname{div} \phi \, d\lambda &= \int \int \tau_t \rho(x) f(t) \operatorname{div} \phi(x) \, d\lambda(t) \, d\lambda(x) \\
&= \int \Big(\int \tau_t \rho(x) \operatorname{div} \phi(x) \, d\lambda(x) \Big) f(t) d\lambda(t) \\
&\le \int \operatorname{var}(\tau_t \rho) f(t) \, d\lambda(t) \le \operatorname{var}(\tau_t \rho) \|f\|_{L^1}
\end{aligned}$$

Since $\operatorname{var}(\tau_t \rho) = \operatorname{var}(\rho) < +\infty$ by item 4 of proposition 2.13, we conclude that $\rho * f \in BV(M)$. $\square$

Chapter 3

Existence and regularity of stationary densities

Assume initially that $F : \Omega \times M \to \Omega \times M$ is a continuous random transformation, in the sense that each $F_\omega = F(x, \cdot)$ is continuous, and M is a metric space. We first investigate conditions that guarantee the existence of stationary densities.

We remark that the space $\mathcal{P}(M)$ of regular probability measures on M can be identified with a subspace of $C(M)^*$, the dual of the bounded continuous real functions on M [10, p. 262]; under this identification, we have $\eta(\phi) = \int \phi \, d\eta$ for every $\eta \in \mathcal{P}(M)$ and $\phi \in C(M)$.

In the following proposition, we denote by *weak-* topology* on $\mathcal{P}(M)$ the subspace topology obtained when $C(M)^*$ is endowed with the weak-* topology. We recall that this topology is given by arbitrary unions of *basic neighborhoods*

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

where $\Phi \subset C(M)$ is a finite set.

Proposition 3.1. *The operator U maps $C(M)$ to itself and U^* is continuous in the weak-* topology.*

Proof. The first part of the proposition is a direct application of item 1 of theorem 3.2 to the functions $f(\omega, \cdot) = \phi \circ F_\omega$. Given any $\eta \in \mathcal{P}(M)$ and a basic neighborhood of $U^*\eta$,

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

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

$$V = \{\xi \in \mathcal{P}(M) : (\forall \psi \in U(\Phi) : |\xi(\psi) - \eta(\psi)| < \epsilon)\}.$$

Then $V \ni \eta$ is a basic neighborhood, because $U(\Phi)$ is a finite subset of $C(M)$, and $U^*(V) \subset W$ because $U^*\xi(\phi) = \xi(U\phi)$ for $\phi \in \Phi$. □

Remark 3.1. The first part of the proposition allows us to restate lemma 1.1 as $U^*\eta(\phi) = \eta(U\phi)$ for $\eta \in \mathcal{P}(M)$ and $\phi \in C(M)$.

3.1 Existence criteria

Corollary 3.1 (Existence of stationary measures [2, p. 31])**.** *If M is compact, then there exists a stationary measure on M.*

Proof. If M is compact, then the set of signed Borel measures can be identified with $C^0(M)^*$ by Riesz representation theorem, which is a locally convex topological vector space under weak-* topology (see proof of theorem 1.5). With this identification,

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

is compact in the weak-* topology, because the unit ball $B(0)$ is compact by Alaoglu's theorem. Clearly, it is a convex set. Therefore U^* admits a fixed point by Schauder-Tychonoff fixed point theorem [10, p. 456]. □

Stationary measures can be obtained in the following way, a version of Krylov-Bogolyubov procedure [2, p. 29]. Note that the choice of the initial probability is important in general, since convergence is only guaranteed in some weakly closed set [7, theorem 14, p. 392], which may be empty.

Proposition 3.2. *Define for an arbitrary probability ν on M and $n \in \mathbb{N}$,*

$$\nu_n = \frac{1}{n} \sum_{j=0}^{n-1} U^{*j} \nu. \tag{3.1}$$

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

Proof. Assume without loss of generality that $\{\nu_n\}_{n\in\mathbb{N}}$ converges weakly-* to $\overline{\nu}$. Since U^* is weak-* continuous, $\{U^*\nu_n\}_{n\in\mathbb{N}}$ converges weakly-* to $U^*\overline{\nu}$. On the other hand,

$$\forall\phi \in C(M) : |U^*\nu_n(\phi) - \nu_n(\phi)| = \frac{1}{n}|U^{*n}\nu(\phi) - \nu(\phi)| \leq \frac{2}{n}\|\phi\|.$$

Thus $U^*\overline{\nu} = \overline{\nu}$.

For the final assertion, note that for any invariant ν, the Birkhoff averages in (3.1) are equal to ν. □

We now consider stationary measures that are absolutely continuous with respect to a reference measure λ, or equivalently, stationary densities of the Perron-Frobenius operator, since the latter are Radon-Nikodym derivatives of the former.

A sequence of absolutely continuous probabilities converges weakly-* to some (absolutely continuous) probability if and only if the associated densities converge weakly to some density. Thus Krylov-Bogolyubov procedure shows that a necessary and sufficient condition for the existence of stationary densities is the existence of some $\phi \in \mathcal{D}$ such that $\{\frac{1}{n}\sum_{j=0}^{n-1} L^j\phi\}_{n\in\mathbb{N}}$ contains a weakly convergent subsequence, that is (see (1.18)),

$$\lim_{\lambda(E)\to 0} \frac{1}{n}\sum_{j=0}^{n-1}\int_E L^j\phi\, d\lambda = 0$$

uniformly on n.

Alternatively, we may consider the Perron-Frobenius operator as a particular case of positive contractions. Recall that the Banach dual of the Perron-Frobenius operator is the transition operator (remark 2.2). In this case, we have the following criterion for the existence of nonzero positive fixed points [33] (see also [28, p. 137, theorem 4.2]).

Theorem 3.1. *Let P be a positive contraction on $L^1(\eta)$. There exists $f \in L^1_+(\eta)$ with $Pf = f \neq 0$ if and only if, for every strictly positive $h \in L^\infty_+(\eta)$,*

$$\inf_{n\geq 0}\int P^{*n}h\, d\eta > 0. \tag{3.2}$$

Further, there exists a strictly positive $f \in L^1_+(\eta)$ such that $Pf = f$ if and only if (3.2) holds for all $0 \neq h \in L^\infty_+(\eta)$.

Proof. Suppose there is $f \in L^1_+(\eta)$ 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

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

for all $n \in \mathbb{N}$. Set $A_n = \{P^{*n}(h \wedge 1) \geq \epsilon\}$. Then $P^{*n}(h \wedge 1) \leq P^{*n}1 \leq 1$ implies

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

hence $\int_{A_n} f\, d\eta > \epsilon \|f\|_{L^1(\eta)}$ for every $n \in \mathbb{N}$. Since $f \in L^1_+(\eta)$, we conclude that there exists some $\delta > 0$ such that $\eta(A_n) > \delta$ for every $n \in \mathbb{N}$. Consequently,

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

For the converse, we divide into two cases.

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

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

Then $\inf_{n \geq 0} \int f P^{*n} h\, d\eta = 0$ for any $f \in L^1_+(\eta)$. Indeed, if we take $a > 0$ such that $\|(f - a)^+\|_{L^1(\eta)} < \epsilon$, then $f \leq a + (f - a)^+$ implies

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

for any $n \in \mathbb{N}$. We conclude that $\inf_{n \geq 0} \int f P^{*n} h\, d\eta \leq \epsilon \|h\|_{L^\infty(\eta)}$ for every $\epsilon > 0$ and the claim follows.

Suppose $f \in L^1_+(\eta)$ verifies $Pf = f$. We have that

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

therefore $f = 0$, because h is strictly positive.

2. Now suppose (3.2) holds for every $0 \neq h \in L^\infty_+(\eta)$. Note that $M = C$. Otherwise, there would be some $c > 0$ such that $B = \{\sum_{n=0}^\infty P^j 1 < c\}$ has positive measure. Therefore

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

which contradicts (3.2).

Define a linear functional ρ on $L^\infty(\eta)$ by $\rho(h) := l(\{\int P^{*n} h \, d\eta\}_{n\in\mathbb{N}})$, where l is an arbitrary Banach limit. Then $\rho(h) \geq 0$ for $h \in L^\infty_+(\eta)$ and $\rho(P^*h) = \rho(h)$. Define, for $h \in L^\infty_+$,

$$\tilde{\nu}(h) := \inf\Big\{\sum_{n=1}^\infty \rho(h_n) : h = \sum_{n=1}^\infty h_n, h_n \in L^\infty_+(\eta)\Big\}.$$

Lemma 3.1. *$\nu(B) := \tilde{\nu}(\chi_B)$ defines a σ-additive finite measure on $\mathcal{B}$.*

Proof. Since $\chi_B \in L^\infty_+(\eta)$ for every $B \in \mathcal{B}$, we have $\nu(B) \leq \rho(\chi_B) < +\infty$. Thus ν is a finite set function on $\mathcal{B}$.

ν is σ-additive because, given a sequence $\{B_i\}_{n\in\mathbb{N}}$ of disjoint sets in $\mathcal{B}$ and $B := \cup_{i=1}^\infty B_i$, we can find, for every $\epsilon > 0$, sequences $\{h_{i,j}\}_{j\in\mathbb{N}}$ such that

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

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

The reverse inequality holds because, given any sequence $\{h_j\}_{j=1}^\infty$ on $L^\infty_+(\eta)$ such that $\chi_B = \sum_{j=1}^\infty h_j$, we have, for each $n \in \mathbb{N}$,

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

$$\therefore \sum_{i=1}^\infty \nu(B_i) \leq \sum_{j=1}^\infty \rho(h_j),$$

by linearity and positivity of ρ. ∎

Set $f = \frac{d\nu}{d\eta}$. We claim that $Pf = f$. In fact, $\rho(P^*h) = \rho(h)$ for any $h \in L^\infty_+(\eta)$ implies that $\tilde{\nu}(P^*h) \le \tilde{\nu}(h)$ for any $h \in L^\infty_+(\eta)$, hence $Pf \le f$. Since $M = C$, we conclude that $Pf = f$ by lemma 1.3. Set $B = \{f = 0\}$ and suppose it has positive measure.

$\tilde{\nu}(\chi_B) = 0$ implies that, for each $m \in \mathbb{N}$, there is a sequence $\{h_{m,n}\}_{n\in\mathbb{N}}$ such that $\chi_B = \sum_{n=1}^\infty h_{m,n}$ and $\sum_{n=1}^\infty \rho(h_{m,n}) < m^{-1}$. Take $k(m)$ so large that

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

and set

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

$h^* \in L^\infty_+(\eta)$, because each $h_{m,n} \in L^\infty_+(\eta)$. Further,

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

from which follows that $\int h^*\, d\eta > 0$, hence $h^* \ne 0$. On the other hand, for all $m \in \mathbb{N}$,

$$\rho(h^*) \le \rho\Big(\sum_{n=1}^{k(m)} h_{m,n}\Big) = \sum_{n=1}^{k(m)} \rho(h_{m,n}) < m^{-1},$$

thus $\rho(h^*) = 0$. This implies that $\liminf_{n\to+\infty} \int P^{*n}h^*\, d\eta = 0$, which contradicts (3.2). We conclude that B has null measure, that is, f is strictly positive. $\square$

Example 3.1. $T : [0,1] \to [0,1]$, $T(x) = x^2$ does not admit an invariant measure that is absolutely continuous with respect to Lebesgue. In fact, consider U the transition operator of T and $\phi : [0,1] \to [0,1]$, $\phi(x) = x$. Then $\phi \in L^\infty_+$ is strictly positive and, for every $n \in \mathbb{N}$ and $\epsilon > 0$,

$$\begin{aligned}\int U^n\phi(x)\, dx &= \int_0^{\sqrt[2n]{\epsilon}} \phi(T^n(x))\, dx + \int_{\sqrt[2n]{\epsilon}}^1 U^n\phi(T^n(x))\, dx \\ &\le \epsilon \sqrt[2n]{\epsilon} + (1 - \sqrt[2n]{\epsilon}) \overset{n\to+\infty}{\longrightarrow} \epsilon.\end{aligned}$$

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

In contrast with the example, a large class of random transformations admits stationary densities.

Proposition 3.3. *Let $F : \Omega \times M \to \Omega \times M$ be a nonsingular random transformation with respect to λ. Suppose there exists $\delta > 0$ such that*

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

Then F admits a stationary $\eta \ll \lambda$.

Proof. It suffices to show that there is $f \in L^1_+(\lambda)$ such that $Lf = f \neq 0$, where L is the Perron-Frobenius operator of F. The transition operator of F is given by

$$Uh(x) = \int_\Omega h(F_\omega(x))\, d\mu(\omega).$$

If $h \in L^\infty_+(\lambda)$ is strictly positive, fix $\epsilon > 0$ such that $B = \{h > \epsilon\}$ satisfies

$$\lambda(B) > 1 - \delta.$$

For every $x \in M$, we have

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

Therefore

$$\forall n \in \mathbb{N} : \int U^n h\, d\lambda \geq \alpha > 0,$$

and the existence of f follows from theorem 3.1. □

Corollary 3.2. *Suppose $F : \Omega \times M \to \Omega \times M$, $M = [0,1]^n$, is obtained from $T : M \to M$ by an additive noise, as in section 2.2. Suppose the noise is distributed according to $\rho\lambda$, where $\rho \in L^1(\lambda)$ and $\lambda(\{\rho > 0\}) > 0$, for λ the Lebesgue measure. Then there is a stationary $\eta \ll \lambda$.*

Proof. Let $\delta = \frac{1}{2}\lambda(\{\rho > 0\})$. If $\lambda(B) > 1 - \delta$, then, for every $x \in M$,

$$\lambda(\{F_\omega(x) \in B : \rho(\omega_0) > 0\}) \geq \lambda(B) + \lambda(\{F_\omega(x) : \rho(\omega_0) > 0\}) - 1 > \delta$$
$$\therefore \lambda(\{\omega_0 \in \{\rho > 0\} : F_\omega(x) \in B\}) > \delta.$$

Fix $\epsilon > 0$ such that

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

We conclude that

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

for every $x \in M$. Thus the hypothesis of the proposition are satisfied. □

Although theorem 3.1 is not constructive, the existence of some strictly positive fixed point is enough to apply the ergodic theorem for any initial point and thus obtain fixed points.

An important issue is the speed of such convergence. In general, this can be arbitrarily slow [29]. In the following proposition, we show that condition (2.4) is strong enough that the convergence to a stationary density is exponentially fast in the L^1 rather than in Cesàro sense.

Proposition 3.4. *If the Perron-Frobenius operator L of a random dynamical system F is mixing in the sense of* (2.4)*, then L admits a unique stationary density h and any density f must converge exponentially fast to h. Precisely, there are $C > 0$ and $\lambda < 1$ such that $\|L^n(f) - h\|_{L^1(\lambda)} \leq C\lambda^n$.*

Proof. $\|L\|_{L^1(\lambda)} = 1$ by proposition 2.4 and $\|L^k|_V\|_{L^1(\lambda)} \leq \alpha$ for some $k \in \mathbb{N}$ and $\alpha < 1$, by (2.4). Take a density $f \in L^1(\lambda)$. We claim that $L^n(f) \to h$ for some density h. On the contrary, there would be $\epsilon > 0$ and a subsequence $\{L^{n_k}(f)\}_{k\in\mathbb{N}}$ such that

$$\forall k \in \mathbb{N} : \|L^{n_{k+1}}(f) - L^{n_k}(f)\|_{L^1(\lambda)} \geq \epsilon.$$

$L^{n_{k+1}-n_k}(f) - f \in V$ because $L^{n_{k+1}-n_k}(f)$ is a density, and $2\|L|_V\|^{n_k}_{L^1(\lambda)} \geq \epsilon$ for every $k \in \mathbb{N}$, a contradiction.

h is stationary because L is bounded, and unique because $g - h \in V$ implies $\|L^n(g) - h\|_{L^1(\lambda)} \to 0$ for any density g. Similarly, the mixing property follows from $f - (\int f\, dm)h \in V$.

Given any density f, $f - h \in V$ implies that

$$\|L^n(f) - h\|_{L^1(\lambda)} \leq \|L^k|_V\|^{\lfloor \frac{n}{k} \rfloor}_{L^1(\lambda)} \|L^{n - \lfloor \frac{n}{k} \rfloor k}(f - h)\|_{L^1(\lambda)} \leq C\lambda^n,$$

for $\lambda = \alpha^{1/k} < 1$ and $C = \max_{0\leq i\leq k-1} \frac{\|L^i(f-h)\|_{L^1(\lambda)}}{\lambda^i}$. □

In the following, we will use equation (2.4) as the definition of mixing property.

3.2 Regularity of Perron-Frobenius iterates

We wish to show that the Perron-Frobenius operator regularizes functions under some conditions on the transition density or on the noise, in the case of transformations obtained by additive noise. We state the following variant of [5, theorem 16.8].

Theorem 3.2. *Let $f(\cdot, x) : \Omega \to \mathbb{R}$ be a measurable function for each $x \in O$, where O is a metric space, and μ a probability measure on Ω.*

1. *Suppose $f(\omega, \cdot)$ 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 O is a subspace of $\mathbb{R}$ and, for every $x_0 \in O$, there exists a full measure set Ω_{x_0} such that $f(\omega, \cdot)$ and $\frac{\partial f}{\partial x}(\omega, \cdot)$ are continuous at x_0. Suppose also that there exists some neighborhood $O_{x_0} \ni x_0$ where $|\frac{\partial f}{\partial x}(\omega, \cdot)| \leq g(\omega, x_0)$ for some $g(\cdot, x_0) \in L^1(\mu)$. Then*

$$\frac{\partial}{\partial x} \int_\Omega f(\omega, \cdot)\, d\mu(\omega) = \int_\Omega \frac{\partial f}{\partial x}(\omega, \cdot)\, d\mu(\omega) \tag{3.3}$$

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

Proof. 1. We have $\lim_{x \to x_0} f(\omega, x) = f(\omega, x_0)$ for a.e. $\omega \in \Omega$ and $|f(\omega, x)| \leq g(\omega)$ for all $x \in O_{x_0}$. Thus, by Lebesgue's dominated convergence theorem, $\lim_{x \to x_0} \int f(\omega, x)\, d\mu(\omega) = \int f(\omega, x_0)\, d\mu(\omega)$.

2. Take Ω_{x_0} and O_{x_0} as in the hypothesis. We have $\lim_{x \to x_0} \frac{f(\omega,x) - f(\omega,x_0)}{x - x_0} = \frac{\partial f}{\partial x}(\omega, x_0)$ and $|\frac{f(\omega,x) - f(\omega,x_0)}{x - x_0}| \leq g(\omega, x_0)$ for every $\omega \in \Omega_{x_0}$ and $x \in O_{x_0}$. We may apply Lebesgue's dominated convergence theorem to obtain

$$\lim_{x \to x_0} \int_{\Omega_{x_0}} \Big| \frac{f(\omega, x) - f(\omega, x_0)}{x - x_0} - \frac{\partial f}{\partial x}(\omega, x_0) \Big|\, d\mu(\omega) = 0,$$

$$\therefore \lim_{x \to x_0} \frac{\int_\Omega f(\omega, x)\, d\mu(\omega) - \int_\Omega f(\omega, x_0)\, d\mu(\omega)}{x - x_0} = \int_\Omega \frac{\partial f}{\partial x}(\omega, x_0)\, d\mu(\omega),$$

where we tacitly extended $\frac{\partial f}{\partial x}(\cdot, x_0)$ to all Ω as a measurable function, because Ω_{x_0} is a full measure set, setting $\frac{\partial f}{\partial x}(\omega, x_0) = 0$ where the partial derivative doesn't exist.

We note that, for every $\omega \in \Omega_{x_0}$, $\lim_{x \to x_0} \frac{\partial f}{\partial x}(\omega, x) = \frac{\partial f}{\partial x}(\omega, x_0)$ and $|\frac{\partial f}{\partial x}(\omega, x)| \leq g(\omega, x_0)$ for all $x \in O_{x_0}$. Hence the previous item applies to $\frac{\partial f}{\partial x}(\cdot, x)$ and $\int_\Omega f(\omega, \cdot)\, d\mu(\omega) \in C^1(O)$. □

To extend the theorem to functions defined on $\mathbb{R}^n$, we use the multi-index notation: we define, for $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$, $|\alpha| = \sum_{i=1}^n \alpha_i$ and $\partial^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}$.

Corollary 3.3. *Let $f(\cdot,x):\Omega\to\mathbb{R}$ be a measurable function for each $x\in O$, where $O\subset\mathbb{R}^n$ is open, and μ 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 continuous at x_0 for every $\omega\in\Omega_{x_0}$ and $\alpha\in\mathbb{N}_0^n, |\alpha|\le k$. Suppose also that there exists some neighborhood $O_{x_0}\ni x_0$ where, for every $\alpha\in\mathbb{N}_0^n$ with $|\alpha|\le k$, $|\partial^\alpha f(\omega,\cdot)|\le g_\alpha(\omega,x_0)$ 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 $\alpha\in\mathbb{N}_0^n$, $|\alpha|\le k$.*

Proof. We proceed by induction on k. For $k=0$, the claim follows directly from Lebesgue's dominated convergence, because $\lim_{x\to x_0} f(\omega,x)=f(\omega,x_0)$ and $|f(\omega,\cdot)|\le g(\omega,x_0)$.

Suppose now that the claim is valid for some $k\in\mathbb{N}_0$ and denote the canonical basis of $\mathbb{R}^n$ by $\{e^1,\dots,e^n\}$. If the assumptions hold for $k+1$, then the theorem can be applied to each function $f_i^\alpha(\omega,\cdot):t\mapsto\partial^\alpha f(\omega,x_0+te^i)$, where $|\alpha|=k$. In fact, for every $\omega\in\Omega_{x_0+te^i}$, $f_i^\alpha(\omega,\cdot)$ and $\frac{\partial f_i^\alpha}{\partial t}(\omega,\cdot)$: $t\mapsto\frac{\partial(\partial^\alpha f)}{\partial x_i}(\omega,x_0+te^i)$ are continuous at each $t\in(-\epsilon,\epsilon)$, where $B_\epsilon(x_0)\subset O$, and $|f_i^\alpha(\omega,\cdot)|\le g_\alpha(\omega,x_0)$ in O_{x_0}. In particular, $\frac{\partial}{\partial t}\int_\Omega f_i^\alpha(\omega,t)\,d\mu(\omega)=\int_\Omega\frac{\partial f_i^\alpha}{\partial t}(\omega,t)\,d\mu(\omega)$ for $t\in(-\epsilon,\epsilon)$. It follows from the inductive hypothesis that, if $\alpha\in\mathbb{N}_0$ and $|\alpha|\le k$,

$$\begin{aligned}\frac{\partial(\partial^\alpha\int_\Omega f(\omega,x_0)\,d\mu(\omega))}{\partial x_i}&=\frac{\partial}{\partial x_i}\int_\Omega\partial^\alpha f(\omega,x_0)\,d\mu(\omega)=\frac{\partial}{\partial t}\int_\Omega f_i^\alpha(\omega,0)\,d\mu(\omega)\\&=\int_\Omega\frac{\partial f_i^\alpha}{\partial t}(\omega,0)\,d\mu(\omega)=\int_\Omega\frac{\partial(\partial^\alpha f)}{\partial x_i}(\omega,x_0)\,d\mu(\omega).\end{aligned}$$

Since $\partial^{\alpha+e^i}f(\omega,\cdot)$, where $(\alpha_1,\dots,\alpha_n)+e^i:=(\alpha_1,\dots,\alpha_i+1,\dots,\alpha_n)$, is continuous at x_0 for every $\omega\in\Omega_{x_0}$ and dominated by $g(\omega,x_0)$,

$$\lim_{x\to x_0}\int_\Omega\partial^{\alpha+e^i}f(\omega,x)\,d\mu(\omega)=\int_\Omega\partial^{\alpha+e^i}f(\omega,x_0)\,d\mu(\omega).$$

This concludes the inductive step, because any $\alpha\in\mathbb{N}_0^n$ with $|\alpha|\le k+1$ can be written as $\alpha'+e^i$ for some $\alpha'\in\mathbb{N}_0^n$ with $|\alpha'|\le k$. □

By proposition 2.3, the Perron-Frobenius operator $L:L^1(\lambda)\to L^1(\lambda)$ admits an expression of the form

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

if F is nonsingular, $\mathcal{B}$ is countably generated and $\phi\in L^\infty(\lambda)$. Therefore, we consider the case $\Omega=M$, $\mu=\lambda$ in corollary 3.3 and expect $L\phi$ to be at least as regular as p under mild conditions on ϕ. The following proposition develops this idea.

Proposition 3.5. *If $p \in L^1(M \times M)$ is of class C^k, then $L\phi$ is of class C^k for every $\phi \in L^\infty(\lambda)$.*

Proof. For each $t \in M$, $x \mapsto f(t,x) = p(t,x)\phi(t)$ is of class C^k. Further, for every $x_0 \in M$, we have

$$\forall \alpha \in \mathbb{N}_0^n, |\alpha| \leq k : |\partial^\alpha f(t,x)| \leq \|\partial^\alpha p\|_{L^\infty(B(v_0))}|\phi(t)|$$

for all $x \in B(x_0)$, a sufficiently small neighborhood of x_0 such that the $\partial^\alpha p$ are bounded.

Thus corollary 3.3 applies to f and we conclude that $L\phi$ is of class C^k. □

3.3 Regularity of stationary densities of systems with additive noise

We are mostly interested in the case that F is given by some map on $M = [0,1]^n \cong \mathbb{T}^n$ perturbed by a BV additive noise, as in subsection 2.2.1. If F is not a local diffeomorphism, p can admit discontinuities. Thus we take a different approach and consider properties of the convolution.

In the following, for an open set $O \subset M$, we say that $\phi \in L^\infty_{\text{loc}}(O)$ if $\|\phi\|_{L^\infty(K)} < +\infty$ for every compact $K \subset O$. And we say that ϕ is locally Lipschitz continuous at $x \in M$ if $\phi \circ \pi$ is Lipschitz continuous in some neighborhood of x.

Lemma 3.2. *Let $L_\xi = N_\xi L_0$, where N_ξ is the convolution operator given by $\phi \in L^1(\lambda) \mapsto \rho_\xi * \phi$, $\rho_\xi \in L^1(\lambda)$ is symmetric and supported on $B_{\xi/2}(0)$ as in section 2.2 and L_0 the Perron-Frobenius operator associated to a nonsingular T. If ρ_ξ is Lipschitz continuous in $\pi(B_{\xi/2}(0))$ and there is an open set $O \subset M$ such that $L_0\phi|_O \in L^\infty_{loc}(O)$, then $L_\xi\phi$ is locally Lipschitz continuous at every $x \in M$ such that $\pi(\partial B_{\xi/2}(x)) \subset O$.*

Proof. We use the following lemma.

Lemma 3.2.1. *If $f \in BV(M)$, then $\|\tau_h f - f\|_{L^1} \leq |h| \operatorname{var}(f)$ for every $h \in \mathbb{R}^n$.*

Proof. Let $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$, $\sup_{x\in\mathbb{R}^n} |\phi(x)| \leq 1$ and $e = \frac{h}{|h|} \in \mathbb{R}^n$. Define $\tilde{\phi} \in C^1(\mathbb{R}^n, \mathbb{R}^n)$ by $\tilde{\phi} = \phi e$. Then, for $t = |h|$,

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

$C^1(\mathbb{R}^n, \mathbb{R})$ is dense in $L^\infty(\mathbb{R}^n)$, therefore $\|\tau_h f - f\|_{L^1} \leq |h|\, \mathrm{var}(f)$. □

If $L_0\phi|_O \in L^\infty_{\mathrm{loc}}(O)$ and $\pi(\partial B_{\xi/2}(x)) \subset O$, fix $s > 0$ such that $K = \pi(\overline{B}_{\xi/2+s}(x) \setminus B_{\xi/2-s}(x)) \subset O$. Then, for every $y, y+h \in B_s(x)$, we have

$$\begin{aligned}|N_\xi L_0\phi(\pi(y+h)) - N_\xi L_0\phi(\pi(y))| &= \left| \int (\tau_{-h}\tau_y\rho_\xi(t) - \tau_y\rho_\xi(t))L_0\phi(t)\, d\lambda(t) \right| \\ &\leq |h|\, \mathrm{var}(\tau_y\rho_\xi)\|\|L_0\phi\|_{L^\infty(K)} + |h|\, \mathrm{Lip}(\rho_\xi|_{\pi(B_{\xi/2}(0))})\|L_0\phi\|_{L^1} \\ &\leq |h|(\mathrm{var}(\rho_\xi)\|L_0\phi\|_{L^\infty(K)} + \mathrm{Lip}(\rho_\xi|_{\pi(B_{\xi/2}(0))}),\end{aligned}$$

because $\pi(B_{\xi/2}(y))\Delta\pi(B_{\xi/2}(y+h)) \subset K$, thus both or none of $\pi(t-y)$ and $\pi(t-y+h)$ belong to $\pi(B_{\xi/2}(0))$ if $t \notin K$. We conclude that $N_\xi L\phi|_{B_r(x)}$ is Lipschitz continuous.

The second part of the theorem is a direct application of 3.3, taking $f(\omega, x) = \rho_\xi(x-\omega)\phi(\omega)$. □

Theorem A. *In the context of section 2.2, suppose that $T \in C^1(M, M)$, with $\#T^{-1}(x) \leq m$ for every $x \in M$, and ρ_ξ is Lipschitz continuous in $B_{\xi/2}(0)$. Set*

$$O = \{x \in M : \#T^{-1} \textit{ is locally constant at } x, |\det T'(y)| > 0 \textit{ for all } y \in T^{-1}(x)\}$$

Then every bounded stationary density ϕ, in particular every stationary density in the one-dimensional case, is locally Lipschitz continuous at every $x \in M$ such that $\pi(\partial B_{\xi/2}(x)) \subset O$.

Proof. For every compact $K \subset O$ and $\phi \in L^\infty(\lambda)$, we have, by proposition 2.12,

$$\|L_0\phi\|_{L^\infty(K)} \leq \frac{m\|\phi\|_{L^\infty}}{\inf_{y\in T^{-1}(K)} |T'(y)|} < +\infty.$$

We conclude that $L_0\phi|_O \in L^\infty_{\text{loc}}(O)$ and the hypothesis of lemma 3.2 is satisfied for $\phi = L_\xi \phi$.

In the one-dimensional case, since $L_\xi \phi = \phi$ implies that $\phi \in BV(M)$ (proposition 2.14) and $BV(M) \subset L^\infty(\lambda)$ (proposition 2.13), we have that every stationary density is bounded and the statement follows. $\square$

The next theorem from [14] shows that the stationary densities are stable under small changes of noise or the deterministic component, provided we have contraction on the zero average subspace

$$V = \{f \in L^1 : \int f \, d\lambda = 0\}.$$

Theorem 3.3. *In the context of theorem 3.2, denote by L_1, L_2 the Perron-Frobenius operators associated to continuous $T_1, T_2 : M \to M$, and N_1, N_2 the convolution operators associated to $\rho_1, \rho_2 \in BV(M)$, respectively. Suppose that $N_1 L_1$ satisfies the spectral gap condition*

$$\exists k \in \mathbb{N} : \|(N_1 L_1)^k|_V\|_{L^1} < 1.$$

If $L_i \phi_i = \phi_i$ for $i \in \{1, 2\}$, then there is a constant $C > 0$ such that

$$\|f_1 - f_2\|_{L^1} \leq C(\|\rho_1 - \rho_2\|_{L^1} + \min\{\text{var}(\rho_1), \text{var}(\rho_2)\} \||T_1(\cdot) - T_2(\cdot)|\|_{L^\infty}).$$

Proof. We use the following lemma, from which the constant is obtained.

Lemma 3.3. *Let P_1, P_2 be two Markov operators. Assume*

$$\forall j \in \{0, \dots, k\} : \|P_1^j|_V\| \leq C_j$$

and $C_k < 1$. If $P_i f_i = f_i$ for $i \in \{1, 2\}$, then

$$\|f_2 - f_1\| \leq \frac{\sum_{i=0}^{k-1} C_i}{1 - C_k} \|P_1 - P_2\|. \tag{3.4}$$

Proof. By triangle inequality,

$$\|f_1 - f_2\| = \|P_1^k f_1 - P_2^k f_2\| \leq \|P_1^k (f_1 - f_2)\| + \|(P_1^k - P_2^k) f_2\|.$$

The first term in the rightmost side is bounded by $C_k \|f_1 - f_2\|$. To estimate the second term, we write

$$P_1^k - P_2^k = \sum_{i=0}^{k-1} P_1^{k-1-j} (P_1 - P_2) P_2^j,$$

$$\therefore \|P_1^k - P_2^k\| \leq \sum_{i=0}^{k-1} C_i \|P_1 - P_2\|.$$

Therefore,

$$(1-C_k)\|f_1-f_2\| \le \sum_{i=0}^{k-1} C_i\|P_1-P_2\|\|f_2\| = \sum_{i=0}^{k-1} C_i\|P_1-P_2\|. \qquad \square$$

Thus it remains to bound $\|N_1L_1 - N_2L_2\|_{L^1}$ in terms of T_1, T_2, ρ_1, ρ_2, as follows. Take any $\phi \in L^1(\lambda)$ and estimate

$$\begin{aligned}\|N_1L_1 - N_2L_2\|_{L^1} &\le \|N_1L_1 - N_2L_1\|_{L^1} + \|N_2L_1 - N_2L_2\|_{L^1}\\ &\le \|\rho_1-\rho_2\|_{L^1}\|L_1\|_{L^1} + \|N_2\|_{W\to L^1}\|L_1-L_2\|_{L^1\to W}\\ &\le \|\rho_1-\rho_2\|_{L^1} + \mathrm{var}(\rho_2)\|L_1-L_2\|_{L^1\to W},\end{aligned}$$

where we used item 5 of theorem B. Together with the fact that, for all $f \in L^1(\lambda)$,

$$\begin{aligned}\|(L_1-L_2)f\|_W &= \sup_{\mathrm{Lip}(\phi)=1} \int f(x)[\phi(T_1(x)) - \phi(T_2(x))], d\lambda(x)\\ &\le \int |f(x)||T_1(x)-T_2(x)|\, d\lambda(x) \le \||T_1(\cdot)-T_2(\cdot)|\|_{L^\infty}\|f\|_{L^1},\end{aligned}$$

we obtain the desired estimate. $\square$

Chapter 4

Ulam approximation

Here we show how we can study the behavior of the Perron-Frobenius operator associated to a random dynamical system by approximating it by a finite rank operator. The finite rank approximation we use is known in literature as Ulam's method, named after [43]. For more details, see [14].

We assume the context of section 2.2, with $\rho \in BV(M)$.

Suppose we're given a δ-partition of M into convex sets $\mathcal{I}_\delta = \{I_i\}_{i=1}^{l}$, where $\operatorname{diam} I_i \leq \delta$, and denote the characteristic function of I_i by χ_i. An operator $P : L^1(M) \to L^1(M)$ can be discretized as

$$P_\delta : L^1(\lambda) \to L^1(\lambda), \quad P_\delta = \pi_\delta P \pi_\delta, \tag{4.1}$$

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

$$\pi_\delta h(x) = \sum_{i=1}^{l} \mathbb{E}(h|I_i)\chi_i. \tag{4.2}$$

This operator is completely determined by its restriction to the subspace generated by $\{\chi_1, \dots, \chi_l\}$, and thus may be represented by a matrix in this base, which we call the Ulam matrix. In the following, we assume that $\mathcal{I}_\delta$ is a δ-partition of M.

For computational purposes, the Perron-frobenius operator L_ξ of T_ξ is discretized as

$$L_{\delta,\xi} = \pi_\delta N_\xi \pi_\delta L_0 \pi_\delta. \tag{4.3}$$

This is simple to work with because it is the product of the discretized operators $\pi_\delta N_\xi \pi_\delta$ and $\pi_\delta L_0 \pi_\delta$.

4.1 Computation of the stationary density

Consider the zero average subspace

$$V = \{f \in L^1(\lambda) : \int f \, d\lambda = 0\} \tag{4.4}$$

and suppose $\|L^k_{\delta,\xi}|_V\| \leq \alpha < 1$ for some $k \in \mathbb{N}$. We have

$$\|L^i_\xi|_V\|_{L^1} \leq \|L^i_{\delta,\xi}|_V\|_{L^1} + \|L^i_{\delta,\xi}|_V - L^i_\xi|_V\|_{L^1}. \tag{4.5}$$

Denote by f_ξ and $f_{\delta,\xi}$ the stationary probability densities for L_ξ and $L_{\delta,\xi}$, respectively. Since (see the proof of lemma 3.3)

$$\|f_\xi - f_{\delta,\xi}\|_{L^1} \leq \frac{1}{1-\alpha}\|(L^k_{\delta,\xi} - L^k_\xi)f_\xi\|_{L^1}, \tag{4.6}$$

we search a good estimate of $\|(L^k_{\delta,\xi} - L^k_\xi)f_\xi\|_{L^1}$ to prove mixing of T_ξ and give a rigorous estimate of $\|f_\xi - f_{\delta,\xi}\|_{L^1}$.

The calculus of $\|L^i_{\delta,\xi}|_V\|$ is computationally complex, thus an alternative approach is used in [14]. First, a coarser version of the operator is considered, $L_{\delta_{\text{contr}},\xi}$, where δ_{contr} is a multiple of δ, for which $k_{\text{contr}} \in \mathbb{N}$ and constants $C_{i,\text{contr}}$, for $i < k_{\text{contr}}$, and $\alpha_{\text{contr}} < 1$ are calculated in order that

$$\|L^i_{\delta_{\text{contr}},\xi}\| \leq C_{i,\text{contr}}, \quad \|L^{k_{\text{contr}}}_{\delta_{\text{contr}},\xi}|_V\| \leq \alpha_{\text{contr}}. \tag{4.7}$$

Finally, the following lemma is used to relate the estimates of the coarser and finer partition.

Lemma 4.1. *Let $\|L^i_{\gamma,\xi}|_V\|_{L^1} \leq C_i(\gamma)$; let σ be a linear operator such that $\sigma^2 = \sigma$, $\|\sigma\|_{L^1} \leq 1$, and $\sigma\pi_\gamma = \pi_\gamma\sigma = \pi_\gamma$; let $\Lambda = \sigma N_\xi \sigma L_0 \sigma$. Then we have*

$$\|(L^j_{\gamma,\xi} - \Lambda^j)N_\xi\|_{L^1} \leq \frac{3\gamma}{\xi}\operatorname{var}(\rho)\sum_{i=0}^{j-1} C_i(\gamma). \tag{4.8}$$

The lemma is applied to two cases.

1. $\gamma = \delta_{\text{contr}}$, $\sigma = \pi_\delta$ and $\Lambda = L_{\delta,\xi}$ implies

$$\|(L^j_{\delta_{\text{contr}},\xi} - L^j_{\delta,\xi})N_\xi\|_{L^1} \leq \frac{3\delta_{\text{contr}}}{\xi}\operatorname{var}(\rho)\sum_{i=0}^{j-1} C_{i,\text{contr}}.$$

This is used to obtain $k \in \mathbb{N}$, $\alpha < 1$ and $C_i, i < k$, such that

$$\|L^i_{\delta,\xi}|_V\| \leq C_i, \quad \|L^k_{\delta,\xi}|_V\| \leq \alpha. \tag{4.9}$$

In fact, we have

$$\begin{aligned}\|L^i_{\delta,\xi}|_V\| &\leq \|L^{i-1}_{\delta,\xi}N_\xi \pi_\delta|_V\| \\ &\leq \|(L^{i-1}_{\delta,\xi} - L^{i-1}_{\delta_{\text{contr}},\xi})N_\xi \pi_\delta|_V\| + \|L^{i-1}_{\delta_{\text{contr}}}N_\xi \pi_\delta|_V\| \\ &\leq \|(L^{i-1}_{\delta,\xi} - L^{i-1}_{\delta_{\text{contr}},\xi})N_\xi\| + \|L^{i-1}_{\delta_{\text{contr}}}|_V\|.\end{aligned}$$

2. $\gamma = \delta$, $\sigma = \text{Id}$ and $\Lambda = L_\xi$ implies

$$\begin{aligned}\|L^{k+1}_\xi|_V\|_{L^1} &\leq \|L^k_{\delta,\xi}L_\xi|_V\|_{L^1} + \|(L^k_\xi - L^k_{\delta,\xi})L_\xi\|_{L^1} \\ &\leq \alpha + \frac{3\delta}{\xi}\operatorname{var}(\rho)\sum_{i=0}^{k-1} C_i(\delta).\end{aligned}$$

By remark 2.6, we conclude that the mixing condition is satisfied whenever

$$\alpha + \frac{3\delta}{\xi}\operatorname{var}(\rho)\sum_{i=0}^{k-1} C_i(\delta) < 1. \tag{4.10}$$

Proof of lemma 4.1. As in the proof of lemma 3.3, we have

$$(L^k_{\gamma,\xi} - \Lambda^k)N_\xi = \sum_{i=0}^{k-1} L^i_{\gamma,\xi}(L_{\gamma,\xi} - \Lambda)\Lambda^{k-1-i}N_\xi,$$

where the term $(L_{\gamma,\xi} - \Lambda)$ can be decomposed as

$$L_{\gamma,\xi} - \Lambda = \pi_\gamma N_\xi(\pi_\gamma - \sigma)L + (\pi_\gamma - \sigma)N_\xi \sigma L.$$

Therefore

$$\begin{aligned}\|(L^k_{\gamma,\xi} - \Lambda^k)N_\xi\|_{L^1} &\leq \sum_{i=0}^{k-1} C_i(\gamma)(\|\pi_\gamma N_\xi(\pi_\gamma - \sigma)\|_{L^1}\|L\Lambda^{k-i-1}\|_{L^1} \\ &\qquad + \|(\pi_\gamma - \sigma)N_\xi\|_{L^1}\|\sigma L\Lambda^{k-i-1}\|_{L^1}) \\ &\leq \sum_{i=0}^{k-1} C_i(\gamma)(\|N_\xi(\pi_\gamma - 1)\|_{L^1} + \|(\pi_\gamma - 1)N_\xi\|_{L^1}) \\ &\leq 2\sum_{i=0}^{k-1} C_i(\gamma)\Big(\frac{\gamma}{2} + \gamma\Big)\operatorname{var}(\rho_\xi) = \frac{3\gamma}{\xi}\operatorname{var}(\rho)\sum_{i=0}^{k-1} C_i(\gamma),\end{aligned}$$

where we use items 6 and 7 of theorem B. □

Remark 4.1. In the case that the partition is given by cubes, a better bound can be obtained by item 6 of theorem B. In particular, in the one-dimensional case, we get the constant $2\gamma/\xi$ instead of $3\gamma/\xi$.

We remark that a simple estimate to (4.6) is given by ([14, equation 4]). This follows from (4.8), because

$$\|(L^n_{\delta,\xi} - L^n_\xi)f_\xi\|_{L^1} \leq \|(L^n_{\delta,\xi} - L^n_\xi)N_\xi\|_{L^1}\|Lf_\xi\|_{L^1} = \|(L^n_{\delta,\xi} - L^n_\xi)N_\xi\|_{L^1}$$

$$\therefore \|f_\xi - f_{\delta,\xi}\|_{L^1} \leq \frac{3\delta \operatorname{var}(\rho)}{\xi(1-\alpha)} \sum_{i=0}^{k-1} C_i. \tag{4.11}$$

The analysis of data obtained from the numerical approximation $\tilde{f}$ of $f_{\delta,\xi}$, in particular its variance, permits to improve greatly this bound. This is done as follows: first, we verify that

$$\|(L^k_{\delta,\xi} - L^k_\xi)f_\xi\|_{L^1} \leq \|(1-\pi_\delta)f_\xi\|_{L^1} + \sum_{i=0}^{k-1} C_i(\delta)(\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1} + \|N_\xi\pi_\delta L(1-\pi_\delta)f_\xi\|_{L^1}),$$

where all the terms $\|(1-\pi_\delta)f_\xi\|_{L^1}$, $\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1}$ and $\|N_\xi\pi_\delta L(1-\pi_\delta)f_\xi\|_{L^1}$ admit bounds (for $i = 1, 2, 3$ respectively)

$$A_i\|f_\xi - \tilde{f}\|_{L^1} + B_i, \tag{4.12}$$

thus for $A = A_1 + (A_2 + A_3)\sum_{i=0}^{k-1} C_i$ and $B = B_1 + (B_2 + B_3)\sum_{i=0}^{k-1} C_i$,

$$\|(L^k_{\delta,\xi} - L^k_\xi)f_\xi\|_{L^1} \leq A\|f_\xi - f_{\delta,\xi}\|_{L^1} + B.$$

Using this estimate, (4.6) implies

$$\|f_\xi - f_{\delta,\xi}\|_{L^1} \leq C + D\|f_\xi - \tilde{f}\|_{L^1}, \qquad C = \frac{A}{1-\alpha},\ D = \frac{B}{1-\alpha}$$

Finally, adding the numeric error,

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

which can be rewritten as

$$\|f_\xi - \tilde{f}\|_{L^1} \leq \frac{1}{1-D}(\|f_{\delta,\xi} - \tilde{f}\|_{L^1} + C). \tag{4.13}$$

We proceed to prove the bounds on $\|(1-\pi_\delta)f_\xi\|_{L^1}$, $\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1}$ and $\|N_\xi\pi_\delta L(1-\pi_\delta)f_\xi\|_{L^1}$) in lemmas 4.19, 4.21 and 4.22. To do it, we will need the bounds summarized below.

4.1.1 Auxiliary bounds

Definition 4.1. The *Wasserstein-like norm* is defined on V by

$$\|f\|_W = \sup_{\mathrm{Lip}(\phi)=1} \int f(x)\phi(x)\,dx, \tag{4.14}$$

where Lip is defined on $C^0(M)$ by

$$\mathrm{Lip}(\phi) = \sup_{x,y\in M} \frac{|\phi(x)-\phi(y)|}{|x-y|}.$$

We generalize bounds obtained in [14] in the one-dimensional context.

Theorem B. *Let π_δ be the Ulam projection on a δ-partition $\mathcal{I}_\delta$ of M into convex sets and I a finite union of sets from $\mathcal{I}_\delta$. Let N_ξ be the convolution operator $N_\xi(f) = \rho_\xi * f$, where $\mathrm{var}(\rho_\xi) < +\infty$. Then*

1. $\|1-\pi_\delta\|_{\mathrm{var}\to L^1} \le \frac{\delta}{2}$;
2. $\|1-\pi_\delta\|_{L^1\to W} \le \delta$; *if $\mathcal{I}$ is a partition into cubes with side $\le d$, then* $\|1-\pi_\delta\|_{L^1\to W} \le d\int_{[0,1]^n} |x|\,d\lambda(x)$;
3. $\|1-\pi_\delta\|_{\mathrm{var}_I\to W(I)} \le n^{-1}\inf_{x_0\in I}\left(\lambda(B)^{-1}\int_I |x-x_0|^n\,d\lambda(x)\right)^{1/n}$, *where $\lambda(B)$ is the volume of the n-dimensional unit ball;*
4. $\|N_\xi\|_{L^1\to\mathrm{var}} \le \mathrm{var}(\rho_\xi)$;
5. $\|N_\xi\|_{W\to L^1} \le \mathrm{var}(\rho_\xi)$.

Consequently,

6. $\|(1-\pi_\delta)N_\xi\|_{L^1} \le \frac{\delta}{2}\,\mathrm{var}(\rho_\xi)$;
7. $\|N_\xi(1-\pi_\delta)\|_{L^1} \le \delta\,\mathrm{var}(\rho_\xi)$; *if $\mathcal{I}$ is a partition into cubes with side $\le d$, then* $\|N_\xi(1-\pi_\delta)\|_{L^1\to L^1} \le d\,\mathrm{var}(\rho_\xi)\int_{[0,1]^n}|x|\,d\lambda(x)$.

Proof. 1. Given any $f \in BV(M)$, we may assume that $f \in C^\infty(M)$ because there exists a sequence $\{f_k\}_k$ in $C^\infty(M)$ such that $f_k \to f$ in $L^1(M)$ and $\mathrm{var}(f_k) \to \mathrm{var}(f)$, by proposition 2.13. By the L^1 optimal Poincaré inequality for convex domains [1, theorem 3.2],

$$\forall I \in \mathcal{I}_\delta : \|f-\pi_\delta f\|_{L^1(I)} \le \frac{\delta}{2}\,\mathrm{var}_I(f). \tag{4.15}$$

Hence

$$\|f-\pi_\delta f\|_{L^1(M)} \le \sum_{I\in\mathcal{I}_\delta} \frac{\delta}{2}\,\mathrm{var}_I(f) = \frac{\delta}{2}\,\mathrm{var}(f).$$

2. Since $f \mapsto f\, d\lambda$ defines an isometric embedding of $L^1(M)$ into the space $\mathcal{M}(M)$ of signed Borel measures equipped with the TV norm, and $1 - \pi_\delta$ extends to $\mathcal{M}(M)$ taking

$$(1 - \pi_\delta)\eta = \eta - \mathbb{E}(\eta \,|\, \mathcal{I}_\delta),$$

it's sufficient to prove that $\|1 - \pi_\delta\|_{\mathcal{M}(M) \to W} \leq \delta$. To see this, we use Krein-Milman theorem and the following lemma.

Lemma 4.2. *The extreme set of* $(\mathcal{M}_{M+})_1 := \{\eta \in \mathcal{M}_{M+} : \|\eta\|_{TV} \leq 1\}$ *is the set of Dirac measures* $\{\delta_x\}_{x \in M}$.

Proof. Every Dirac measure δ_x is in the extreme set of $(\mathcal{M}_{M+})_1$, because if $\delta_x = (1-t)\eta_1 + t\eta_2$ with $0 < t < 1$, then $(1-t)\eta_1(\{x\}) + t\eta_2(\{x\}) = 1$; since $\eta_1(B), \eta_2(B) \leq 1$, we conclude that $\eta_1 = \delta_x = \eta_2$.

Conversely, every extreme point $\eta \in (\mathcal{M}_{M+})_1$ is a Dirac measure, otherwise there would be $B \in \mathcal{B}$ such that $0 < \eta(B) < 1$; thus $\eta = (1 - \eta(B))\eta_{M \setminus B} + t\eta_B$, where $\eta_B(C) := \frac{\eta(B \cap C)}{\eta(B)}$ and $\eta_{M \setminus B}(C) := \frac{\eta((M \setminus B) \cap C)}{\eta(M \setminus B)}$ are distinct measures in $(\mathcal{M}_{M+})_1$. □

If $\eta_1, \ldots, \eta_k \in (\mathcal{M}_N)_1$ have disjoint supports and the inequality $\|(1 - \pi_\delta)\eta_i\|_W \leq \delta \|\eta_i\|_{TV}$ is valid for $i \in \{1, \ldots, k\}$, then it is valid for any $\eta = t_1\eta_1 + \cdots + t_k\eta_k$, $t_i \in \mathbb{R}_+$. The set $\{\delta_x\}_{x \in M}$ constitutes of measures with pairwise disjoint supports. By continuity of the norms, we conclude that, if

$$\|(1 - \pi_\delta)\delta_x\|_W \leq \delta \|\delta_x\|_{TV} = \delta \tag{4.16}$$

for every $x \in M$, then $\|(1 - \pi_\delta)\eta\|_W \leq \delta\|\eta\|_{TV}$ for every $\eta \in \overline{\mathrm{ch}}(\{\delta_x\}_{x \in M}) = (\mathcal{M}_{M+})_1$. Finally, any $\eta \in \mathcal{M}(M)$ can be written as a difference of nonnegative measures with disjoint supports $\eta = \eta^+ - \eta^-$. We proceed to show (4.16). Take any δ_x, $x \in M$. We have

$$(1 - \pi_\delta)\delta_x = \delta_x - \mathbb{E}(\lambda \,|\, \mathcal{I}(x)),$$

where $\mathcal{I}(x)$ is the element of the partition $\mathcal{I}$ that contains x. Hence

$$\begin{aligned} \|(1 - \pi_\delta)\delta_x\|_W &= \sup_\phi \int \phi \, d((1 - \pi_\delta)\delta_x)\, d\lambda \\ &= \sup_\phi \Big(\phi(x) - \lambda(\mathcal{I}(x))^{-1} \int_{\mathcal{I}(x)} \phi \, d\lambda \Big), \end{aligned}$$

where the supremum is taken over all 1-Lipschitz $\phi \in C^0(M)$. Therefore,

$$\|(1-\pi_\delta)\delta_x\|_W \le \lambda(\mathcal{I}(x))^{-1} \int_{\mathcal{I}(x)} |x-y| \, d\lambda \le \delta.$$

If $\mathcal{I}(x)$ is a cube of side a, then

$$\lambda(\mathcal{I}(x))^{-1} \int_{\mathcal{I}(x)} |x-y| \, d\lambda(y) \le a \int_{[0,1]^n} |x| \, d\lambda(x).$$

which suffices to prove the second claim.

3. We have to bound $\|(1-\pi_\delta)f\|_{W(I)}$ for f such that $\mathrm{var}_I(f) < +\infty$. It suffices to consider the case $\int_I f \, d\lambda = 0$, because both $\mathrm{var}_I(\cdot)$ and $(1-\pi_\delta)$ are translation-invariant. As in the proof of item 1, we can assume $f \in C^\infty(I)$. We have, for any $x_0 \in I$,

$$\|f\|_{W(I)} = \sup_{\mathrm{Lip}(\phi)=1} \int_I f(x)\phi(x) \, d\lambda(x) \le \int_I |f(x)||x-x_0| \, d\lambda,$$

We apply the isoperimetric inequality [45, equation (6.41)] and obtain

$$\|f\|_{L^{n/(n-1)}(I)} \le \frac{\mathrm{var}_I(f)}{n\lambda(B)^{1/n}},$$

where B is the unit ball. Thus

$$\begin{aligned} \|f\|_{W(I)} &\le \|f\|_{L^{n/(n-1)}(I)} \inf_{x_0 \in I} \Big(\int_I |x-x_0|^n \, d\lambda(x) \Big)^{1/n} \\ &\le \frac{\mathrm{var}_I(f)}{n\lambda(B)^{1/n}} \inf_{x_0 \in I} \Big(\int_I |x-x_0|^n \, d\lambda(x) \Big)^{1/n}. \end{aligned}$$

4. As in the proof of item 1, we may assume that $\rho_\xi \in C^\infty(M)$ because there exists a sequence $\{r_k\}_k$ in $C^\infty(M)$ such that $r_k \to \rho_\xi$ in $L^1(M)$ and $\mathrm{var}(r_k) \to \mathrm{var}(\rho_\xi)$. Then

$$\nabla(N_\xi f) = (\nabla\rho_\xi) * f \in C^\infty(M). \tag{4.17}$$

Therefore, by (2.5),

$$\mathrm{var}(N_\xi f) = \|(\nabla\rho_\xi) * f\|_{L^1} \le \|\nabla\rho_\xi\|_{L^1}\|f\|_{L^1} = \mathrm{var}(\rho_\xi)\|f\|_{L^1}. \tag{4.18}$$

5. Let $f \in V$ and consider the sets $M^{\pm} = \{v \in M : f^{\pm}(v) > 0\}$. For each probability q on $M^+ \times M^-$ with marginals $q_{M^{\pm}} = f^{\pm}\, d\lambda$, we have

$$(f^+ - f^-)\lambda = \int_{M^+ \times M^-} \delta_x - \delta_y \, dq(x,y)$$

Therefore, extending N_ξ to $\mathcal{M}_N$ by $N_\xi \nu = (\rho_\xi \lambda) * \nu$, we have

$$\begin{aligned} N_\xi(f^+ - f^-)\lambda &= \int_{M^+ \times M^-} (N_\xi \delta_x - N_\xi \delta_y)\, dq(x,y) \\ &= \int_{M^+ \times M^-} [(\tau_x \rho_\xi)\lambda - (\tau_y \rho_\xi)\lambda]\, dq(x,y) \\ \therefore \|N_\xi f\|_{L^1} \le \int_{M^+ \times M^-} &\operatorname{var}(\rho_\xi)|x-y|\, dq(x,y) \le \operatorname{var}(\rho_\xi)\|f\|_W, \end{aligned}$$

where the estimate $\|\tau_x \rho_\xi - \tau_y \rho_\xi\|_{L^1} \le \operatorname{var}(\rho_\xi)$ follows from lemma 3.2.1.

Finally, items 6 and 7 are a consequence of items 1, 4 and 2, 5, respectively. □

4.1.2 Main estimates for stationary density

Based on the estimates presented in proposition B, we calculate the terms that appear in (4.13), thus establishing a bound for the approximation of the stationary density.

Lemma 4.3. *Let π_δ be the Ulam projection on a δ-partition. Then*

$$\begin{aligned} \|(1-\pi_\delta) f_\xi\|_{L^1} &\le A_1 \|f_\xi - f_{\delta,\xi}\|_{L^1} + B_1, \\ A_1 &= \frac{\delta}{2}\xi^{-1} \operatorname{var}(\rho),\ B_1 = \frac{\delta}{2}\operatorname{var}(N_\xi L \tilde{f}). \end{aligned} \tag{4.19}$$

Proof. We have, using item 6 of proposition B,

$$\begin{aligned} \|(1-\pi_\delta) f_\xi\|_{L^1} &= \|(1-\pi_\delta) N_\xi L f_\xi\|_{L^1} \\ &\le \|(1-\pi_\delta) N_\xi L (f_\xi - \tilde{f})\|_{L^1} + \|(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} \\ &\le \|(1-\pi_\delta) N_\xi\|_{L^1} \|f_\xi - \tilde{f}\|_{L^1} + \|1-\pi_\delta\|_{\operatorname{var}\to L^1} \operatorname{var}(N_\xi L \tilde{f}) \\ &\le \frac{\delta}{2}\xi^{-1} \operatorname{var}(\rho) \|f_\xi - \tilde{f}\|_{L^1} + \frac{\delta}{2} \operatorname{var}(N_\xi L \tilde{f}). \end{aligned}$$

□

We give now the estimates (4.12) for $i = 2, 3$.

Assumption 1. We assume T to be piecewise $C^{1+\beta}$ and monotonic on a partition $\{C_i\}$ of $[0,1]$, and let $T_i = T|_{C_i}$. For each L^1 density g, we let $L_i g$ be the component of Lg coming from the i-th monotone branch, that is,

$$L_i g = L(g\chi_i). \tag{4.20}$$

In this way, we have $Lg = \sum_i L_i g$.

Lemma 4.4. *Let Π be the partition whose elements are given by unions of k adjacent elements of the partition $\mathcal{I}_\delta = \{I_i\}_{i=1}^{l}$ into squares of side $\leq \delta/\sqrt{n} =: d$, where k divides l. Then*

$$\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1} \leq A_2\|f_\xi - \tilde{f}\|_{L^1} + B_2, \tag{4.21}$$

$$A_2 = \frac{d}{2}\operatorname{var}(\rho_\xi)\int_{[0,1]^n} |x|\, d\lambda(x),$$

$$B_2 = \frac{d}{2}\operatorname{var}(\rho_\xi)\sum_{I\in\Pi}\sum_i \min\left\{\frac{\operatorname{var}_I(L_i\tilde{f})}{2n\lambda(B)^{1/n}}\Big(\int_I |x|^n\, d\lambda(x)\Big)^{1/n}, \|L_i\tilde{f}\|_{L^1(I)}\right\},$$

where $\lambda(B)$ is the volume of the n-dimensional ball.

Proof. We can estimate as

$$\begin{aligned}
\|N_\xi(1-\pi_\delta)Lf_\xi\|_{L^1} &\leq \|N_\xi(1-\pi_\delta)L(f_\xi - \tilde{f})\|_{L^1} + \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1}\\
&\leq \|N_\xi(1-\pi_\delta)\|_{L^1}\cdot\|f_\xi - \tilde{f}\|_{L^1} + \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1}\\
&\leq \frac{d}{2}\operatorname{var}(\rho_\xi)\|f_\xi - \tilde{f}\|_{L^1} + \|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1}.
\end{aligned}$$

The first term corresponds to $A_2\|f_\xi - \tilde{f}\|_{L^1}$. The second term is estimated splitting the elements of Π as follows.

$$\begin{aligned}
&\|N_\xi(1-\pi_\delta)L\tilde{f}\|_{L^1} \leq \|N_\xi\|_{W\to L^1}\|(1-\pi_\delta)L\tilde{f}\|_W\\
&\quad\leq \|N_\xi\|_{W\to L^1}\sum_{I\in\Pi}\|(1-\pi_\delta)L\tilde{f}\chi_I\|_{W(I)}\\
&\quad\leq \|N_\xi\|_{W\to L^1}\sum_{I\in\Pi}\sum_i \min\{\|1-\pi_\delta\|_{\operatorname{var}_I\to W(I)}\operatorname{var}_I(L_i\tilde{f}),\\
&\qquad\qquad\qquad\qquad \|1-\pi_\delta\|_{L^1(I)\to W(I)}\|L_i\tilde{f}\|_{L^1(I)}\}\\
&\quad\leq \operatorname{var}(\rho_\xi)\sum_{I\in\Pi}\sum_i \min\left\{\frac{\operatorname{var}_I(L_i\tilde{f})}{2n\lambda(B)^{1/n}}\Big(\int_I |x|^n\, d\lambda(x)\Big)^{1/n}, \|L_i\tilde{f}\|_{L^1(I)}\right\},
\end{aligned}$$

where in the last step we used items 2, 3 and 5 of proposition B. □

Lemma 4.5. *Let* Π *be a partition as in lemma 4.4. It holds*

$$\|N_\xi \pi_\delta L(1-\pi_\delta) f_\xi\|_{L^1} \le A_3 \|f_\xi - \tilde{f}\|_{L^1} + B_3, \tag{4.22}$$

$$A_3 = \frac{d}{2\sqrt{n}} \operatorname{var}(\rho_\xi),$$

$$B_3 = \frac{d^2}{4\sqrt{n}} \operatorname{var}(\rho_\xi) \operatorname{var}(N_\xi L \tilde{f}) \int_{[0,1]^n} |x| \, d\lambda(x) + \sum_{I \in \Pi} \min \Big\{ \frac{d}{2\sqrt{n}},$$

$$\frac{\operatorname{var}(\rho_\xi)/2}{n \lambda(B)^{1/n}} \|T'\|_{L^\infty(I)} \Big(\int_I |x|^n \, d\lambda(x) \Big)^{1/n} \Big\} \operatorname{var}_I (N_\xi L \tilde{f}),$$

where $\lambda(B)$ *is the volume of the* n*-dimensional unit ball.*

Proof. We have

$$\begin{aligned} \|N_\xi \pi_\delta L(1-\pi_\delta) f_\xi\|_{L^1} &= \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L f_\xi\|_{L^1} \\ &\le \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L(f_\xi - \tilde{f})\|_{L^1} + \|N_\xi \pi_\delta L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1}. \end{aligned}$$

Since $\|L\|_{L^1} \le 1$, the first term is bounded by

$$\|N_\xi \pi_\delta\|_{L^1} \|(1-\pi_\delta) N_\xi\|_{L^1} \|f_\xi - \tilde{f}\|_{L^1} \le \frac{d}{2\sqrt{n}} \operatorname{var}(\rho_\xi) \|f_\xi - \tilde{f}\|_{L^1};$$

the second term, by

$$\begin{aligned} &\|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} + \|N_\xi (1-\pi_\delta)\|_{L^1} \|L\|_{L^1} \|1-\pi_\delta\|_{\mathrm{var} \to L^1} \operatorname{var}(N_\xi L \tilde{f}) \\ &\quad \le \|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1} + \frac{d^2}{4\sqrt{n}} \operatorname{var}(\rho_\xi) \operatorname{var}(N_\xi L \tilde{f}) \int_{[0,1]^n} |x| \, d\lambda(x). \end{aligned}$$

To estimate $\|N_\xi L(1-\pi_\delta) N_\xi L \tilde{f}\|_{L^1}$, we use the following lemma.

Lemma 4.6. *For each* $I \in \Pi$,

$$\|L\|_{W(I) \to W} = \|T'\|_{L^\infty(I)} \tag{4.23}$$

Proof. If $f = \chi_I f \in V$, then taking any $v_0 \in I$,

$$\begin{aligned} \sup_{\mathrm{Lip}(\phi)=1} \int_N L(f) \phi \, d\lambda &= \sup_{\mathrm{Lip}(\phi)=1} \int_N (\chi_I f)(\phi \circ T) \, d\lambda \\ &\le \int_I f(v) (\phi(T(v_0)) + \|T'\|_{L^\infty(I)} |v - v_0|) \, d\lambda(v) \\ &\le \|T'\|_{L^\infty(I)} \int_I f(v) |v - v_0| \, d\lambda(v) \le \|T'\|_{L^\infty(I)} \|f\|_{W(I)}, \end{aligned}$$

thus $\|Lf\|_W \le \|T'\|_{L^\infty(I)} \|f\|_{W(I)}$. $\square$

Now we split over the elements $I \in \Pi$.

$$
\begin{aligned}
\sum_{I\in\Pi} &\|N_\xi L(1-\pi_\delta)(N_\xi L\tilde{f}\cdot\chi_I)\|_{L^1} \\
&\le \sum_{I\in\Pi} \min\Big\{\|N_\xi\|_{W\to L^1}\|L\|_{W(I)\to W}\|1-\pi_\delta\|_{\mathrm{var}_I\to W(I)}, \\
&\qquad\qquad\qquad\qquad \|1-\pi_\delta\|_{\mathrm{var}_I\to L^1(I)}\Big\}\,\mathrm{var}_I(N_\xi L\tilde{f}) \\
&\le \sum_{I\in\Pi} \min\Big\{\frac{\mathrm{var}(\rho_\xi)/2}{n\lambda(B)^{1/n}}\|T'\|_{L^\infty(I)}\Big(\int_I |x|^n\,d\lambda(x)\Big)^{1/n},\ \frac{\delta}{2}\Big\}\,\mathrm{var}_I(N_\xi L\tilde{f}).
\end{aligned}
$$

proving the statement thanks to the lemma. □

4.2 Estimating the average of observables

Here we discuss how to average an observable h that has a finite number of *singularities* in the sense that, outside any neighborhood E of this finite set, the observable is in L^∞. This is done by approximating $\int_N hf_\xi\,d\lambda$ with $\int_{M\setminus E} hf_\xi\,d\lambda$ and bounding the error in terms of $\|f_\xi-\tilde{f}\|_{L^1}$, $\|f_\xi\|_{L^\infty(E)}$, $\|h\|_{L^1(E)}$ and $\|h\|_{L^\infty(M\setminus E)}$.

Lemma 4.7. *If $h\in L^\infty(S)$ and $v\in L^1(S)$ has zero average, then*

$$
\Big|\int_S hv\,d\lambda\Big| \le \frac{\sup h-\inf h}{2}\|v\|_{L^1}. \tag{4.24}
$$

Proof. For every $c\in\mathbb{R}$,

$$
\Big|\int_S hv\,d\lambda\Big| \le \Big|\int_S (h-c)v\,d\lambda\Big| + \Big|\int_S cv\,d\lambda\Big| \le \|h-c\|_{L^\infty}\|v\|_{L^1},
$$

because v has zero average. Then $\|h-c\|_{L^\infty}\ge\max\{|\sup h-c|,|\inf h-c|\}$ implies that the optimal bound is obtained with $c=(\sup h+\inf h)/2$, for which $\|h-c\|_{L^\infty}=(\sup h-\inf h)/2$ and we get the desired estimate. □

Corollary 4.1. *Let f and $\tilde{f}$ be bounded probability densities on M and $E\subset M$ a Borel subset for which the $\int_E f\,d\lambda=\int_E\tilde{f}\,d\lambda$. If h is an observable that is bounded on $M\setminus E$, then, for $a=(\sup_{M\setminus E}h-\inf_{M\setminus E}h)/2$,*

$$
\Big|\int_N hf\,d\lambda-\int_{M\setminus E}h\tilde{f}\,d\lambda\Big| \le \|h\|_{L^1(E)}\|f\|_{L^\infty(E)}+a\|f-\tilde{f}\|_{L^1}. \tag{4.25}
$$

Proof. We write

$$\Big|\int_N hf\,d\lambda - \int_{M\setminus E} h\tilde{f}\,d\lambda\Big| = \Big|\int_E hf\,d\lambda - \int_{M\setminus E} h(\tilde{f}-f)\,d\lambda\Big|$$
$$\leq \|h\|_{L^1(E)}\|f\|_{L^\infty(E)} + \Big|\int_{M\setminus E} h(\tilde{f}-f)\,d\lambda\Big|$$

and apply (4.24) to $v = \tilde{f} - f$ and $S = M \setminus E$. □

We proceed to obtain bounds for the term $\|f\|_{L^\infty(E)}$ that appears in (4.25) applied to f_ξ.

Lemma 4.8. *Let $\mathcal{I}$ be a partition as in proposition B. For every $I \in \mathcal{I}$,*

$$\|f_\xi\|_{L^\infty(I)} \leq \|N_\xi L\tilde{f}\|_{L^\infty(I)} + \|\rho_\xi\|_{L^\infty}\|\tilde{f} - f_\xi\|_{L^1}. \tag{4.26}$$

Proof. Since $f_\xi = N_\xi L f_\xi$, triangle inequality gives

$$\begin{aligned}\|f_\xi\|_{L^\infty(I)} &\leq \|N_\xi L\tilde{f}\|_{L^\infty(I)} + \|N_\xi L(\tilde{f}-f_\xi)\|_{L^\infty(I)} \\ &\leq \|N_\xi L\tilde{f}\|_{L^\infty(I)} + \|N_\xi\|_{L^1\to L^\infty}\|L(\tilde{f}-f_\xi)\|_{L^1},\end{aligned}$$

then (4.26) follows because $\|N_\xi\|_{L^1\to L^\infty} \leq \|\rho_\xi\|_{L^\infty}$. □

4.3 Application

We consider a random transformation on the circle, suggested by the behavior of a certain macrovariable in the neural networks studied in [31], modeling the neocortex with a variant of Hopfield's asynchronous recurrent neural network presented in [21].

In Hopfield's network, memories are represented by stable attractors and an unlearning mechanism is suggested in [22] to account for unpinning of these states (see also, e.g., [23]). In the network presented in [31], however, these are replaced by Milnor attractors, which appear due to a combination of symmetrical and asymmetrical couplings and some resetting mechanism.

The model we consider is made by a deterministic map T on the circle perturbed by a small additive noise. For a large enough noise, its associated random dynamical system exhibits an everywhere positive stationary density concentrated on a small region (see theorem C), which can be attributed to the "chaotic itinerancy" of the neural network. This concept still doesn't have a complete mathematical formalization, and deeper understanding of the systems where it was found is important to extract its characterizing mathematical aspects.

In the paper, with the help of a computer aided proof, we establish several results about the statistical and geometrical properties of the above system, with the goal to show that "the behavior of this system exhibits a kind of chaotic itineracy". We show that the system is (exponentially) mixing, hence globally chaotic. We also show a rigorous estimate of the density of probability (and then the frequency) 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. This is done by a computer aided rigorous estimate of the stationary probability density of the system. The computer aided proof is based on the approximation of the transfer operator of the real system by a finite rank operator which is rigorously computed and whose properties are estimated by the computer. The approximation error from the real system to the finite rank one is then managed using an appropriated functional analytic approach developed in [14] for random systems.

The transformation is given by

$$x_{i+1} = T(x_i) + \xi_n \pmod 1, \text{ where } T(x) = x + A\sin(4\pi x) + C, \quad (4.27)$$

for $A = 0.08$, $C = 0.1$ and ξ_n an i.i.d. sequence of random variables with a distribution assumed uniform over $[-\xi/2, \xi/2]$.

The Perron-Frobenius operator (definition 2.1) associated to this system is given by $L_\xi = N_\xi L_0$ (proposition 2.11),where N_ξ is a convolution operator (4.28) and L_0 is the Perron-Frobenius operator of T.

$$N_\xi \phi(x) = \xi^{-1} \int_{-\xi/2}^{\xi/2} \hat{\phi}(x-t)\, dt, \quad (4.28)$$

where $\hat{\phi}(x) = \phi(\pi(x))$ with $\pi : \mathbb{R} \to [0,1)$ as in subsection 2.2.1.

Since

$$T'(x) = 1 + 4\pi A\cos(4\pi x), \qquad 4\pi A = 1.005,$$

T satisfies assumption 1, with $\beta = \infty$ and 6 branches. Viewed as an operator on $\mathbb{S}^1$, it satisfies the hypothesis of theorem A.

In [6], we verified mixing and calculated the stationary density of the one dimensional system (4.27) using the numerical tools from the `compinv-meas` project [14], which implements the ideas presented in section 4.1. The data obtained is summarized in table 4.1.

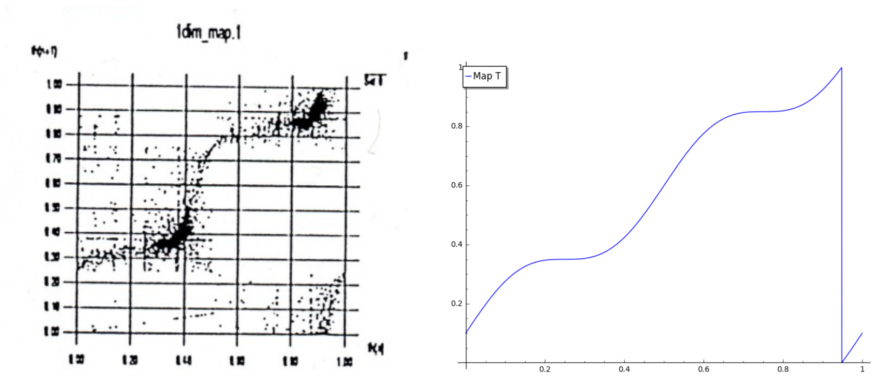

Figure 4.1: Data plot from Tsuda (2016, apud [6]) suggesting the model we studied and the deterministic component of (4.27).

We have shown how the numerical approach developed in [14] can be used to study dynamical properties for a one dimensional random dynamical system of interest in the areas of physiology and neural networks.

We have the following characterization of the stationary densities of the system.

Theorem C. *For every $\xi > 0$ such that the constants obtained by the algorithm, for some $\delta > 0$, satisfy*

$$\alpha + \frac{2\delta}{\xi} \sum_{i=0}^{n-1} C_i(\delta) < 1,$$

the system (4.27) *is mixing, with a unique everywhere positive, locally Lipschitz continuous outside a finite set, stationary density f such that $f(x) = f(x + 1/2)$ for $x \in [0, 1/2]$. In particular, this holds for the ξ in the table 4.1. The stationary densities obtained by the algorithm are presented in figure 4.2.*

Proof. The condition on the constants implies that (4.10) is satisfied, because in our case, $\mathrm{var}(\rho) = 2$; thus the system is mixing. In particular, it posssess a unique stationary density, by proposition 3.4. This density is a.e. locally Lipschitz continuous by theorem A, because $T'(x) = 0$ only for finitely many points $x \in [0, 1]$, and satisfies $f(x) = f(x+1/2)$ for $x \in [0, 1/2]$ because $T(x) = T(x + 1/2)$ for $x \in [0, 1/2]$, so that, for any $\phi \in L^\infty([0, 1])$,

ξ	k_{contr}	α_{contr}	α	$\sum C_i$	l1apriori	l1err
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}

$\delta = 2^{-19}$: used to calculate the invariant density. $\delta_{\text{contr}} = 2^{-14}$: used to find the estimates in (4.7). $\delta_{\text{est}} = 2^{-12}$: used to estimate the L^1 error of the invariant density.

Table 4.1: Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ, where k_{contr} and α_{contr} are chosen in order to satisfy (4.7) so that the values α and $\sum C_i$ obtained through lemma 4.1 attempt to minimize the error `l1err` obtained through the algorithm in [14].

denoting by $\hat{\phi}$ the periodic extension of ϕ to $\mathbb{R}$,

$$\begin{aligned}\int L_\xi f(x)\phi(x)\,d\lambda &= \int f(x)\int \hat{\phi}(T(x)+\omega)\,d\rho(\omega)\,d\lambda(x)\\ &= \int f(x)\int \hat{\phi}(T(x+1/2)+\omega)\,d\rho(\omega)\,d\lambda(x)\\ &= \int L_\xi f(x)\phi(x+1/2)\,d\lambda = \int L_\xi f(x+1/2)\phi(x)\,d\lambda.\end{aligned}$$

For the value $\alpha + \frac{2\delta}{\xi}\sum C_i$, we obtain, in the same order that the values appear in the table, 0.05, 0.07, 0.11, 0.15, 0.21, 0.30, 0.37, 0.48. Therefore the condition is satisfied in all these cases. □

We also studied the system Equation (4.27) in the case that $A = 0.07$ for the same range of noises (Table 4.2). In Figure 4.3, stationary densities obtained in this case are shown. We note that the same kind of "chaotic itinerancy" obtained in the main case is observed.

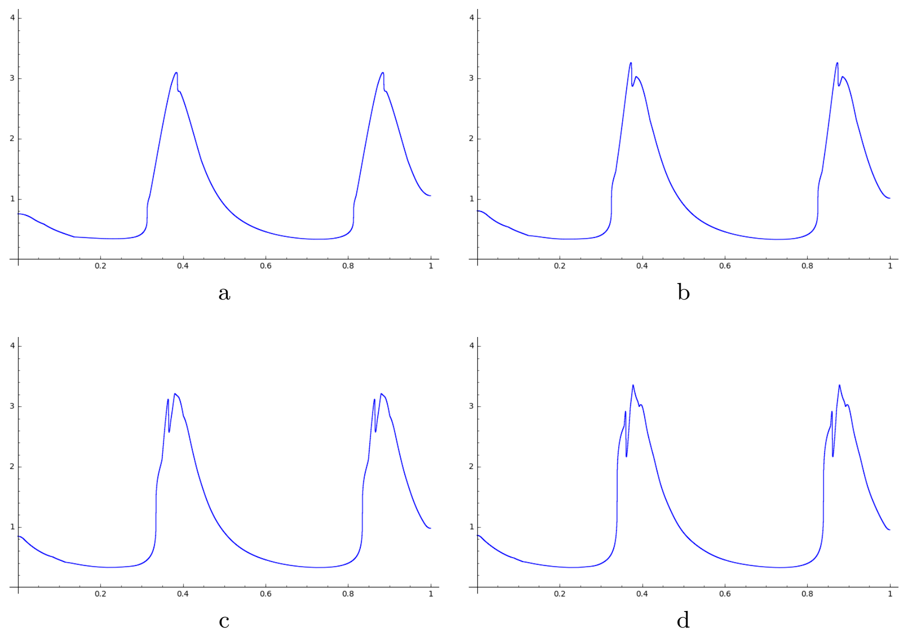

Figure 4.2: Approximated stationary densities $f_{\xi,\delta}$ for T_ξ, with $\delta = 2^{-19}$ and $A = 0.08$. (**a**) $\xi = 0.732 \times 10^{-1}$; (**b**) $\xi = 0.488 \times 10^{-1}$; (**c**) $\xi = 0.305 \times 10^{-1}$; (**d**) $\xi = 0.214 \times 10^{-1}$.

$\boldsymbol{\xi}$	$\boldsymbol{k}_{\mathbf{contr}}$	$\boldsymbol{\alpha}_{\mathbf{contr}}$	$\boldsymbol{\alpha}$	$\sum \boldsymbol{C_i}$	`l1apriori`	`l1err`
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}

Table 4.2: Summary of the L^1 bounds on the approximation error obtained for the range of noises ξ, for the system Equation (4.27) with the alternative value $A = 0.07$.

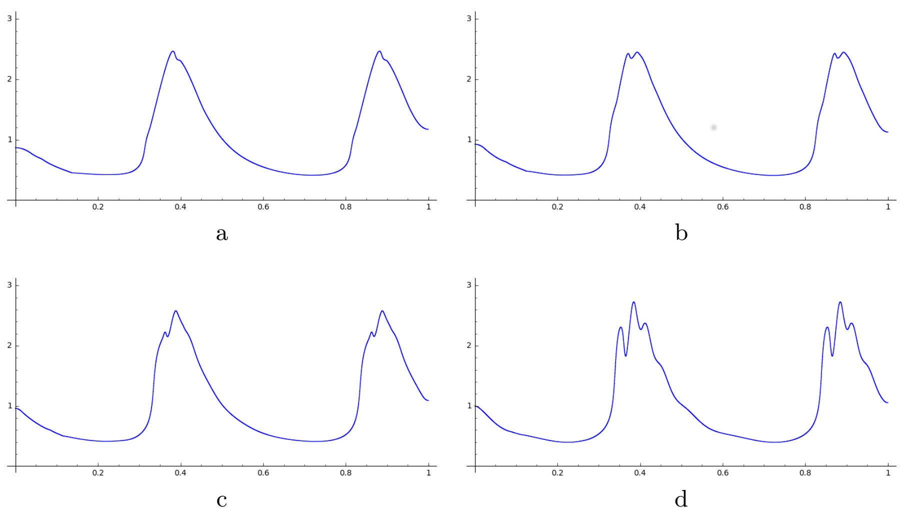

Figure 4.3: Approximated stationary densities $f_{\xi,\delta}$ for T_ξ, with $\delta = 2^{-19}$ and $A = 0.07$. (**a**) $\xi = 0.732 \times 10^{-1}$; (**b**) $\xi = 0.488 \times 10^{-1}$; (**c**) $\xi = 0.305 \times 10^{-1}$; (**d**) $\xi = 0.214 \times 10^{-1}$.

Conclusion

We have obtained theorem A on the regularity of the stationary density of a random dynamical system on the torus with additive noise, using elementary properties of an appropriate notion of BV space. We also proved theorem B, which provides higher-dimensional generalizations of estimates used to calculate the approximation error of the stationary density using the method developed in [14]. Finally, we summarized in theorem C results obtained in [6] and through application of theorem A.

These results are preliminary and can be further explored in future works. In the case of theorem A, it is likely that this can be used to determine an L^∞ bound in approximations of the stationary density, which should make more precise the "chaotic itinerancy" property, since the error in the calculated stationary density in theorem C holds only in L^1.

In theorem B, we have not investigated optimality of the bounds, which is important to the speed of the algorithm. Also an actual implementation would possibly require more sophisticated arguments, because higher dimensions would require much larger matrices in order to calculate the contraction rate of the discretized operator on the zero-average space.

In the application, to investigate further "chaotic itinerancy", it would be important to rigorously compute Lyapunov exponents and other chaos indicators and to investigate the robustness of the behavior of the system under various kinds of perturbations, including the zero noise limit. Another important direction is to refine the model to adapt it better to the experimental data shown in figure 4.1, with a noise intensity which depends on the point.

Bibliography

[1] G. Acosta, R. G. Durán: An optimal Poincaré inequality in L^1 for convex domains. *Proceedings of the american mathematical society*, v. 132, n. 1, p. 195–202, 2003.

[2] L. Arnold: *Random dynamical systems*. Berlin, Heidelberg: Springer Verlag, 1998. (Springer monographs in mathematics)

[3] V. I. Arnold: Small denominators. I. Mappings of the circumference onto itself. *Am. Math. Soc. Transl. Ser. 2*, v. 46, p. 213–284, 1965.

[4] V. Baladi: Decay of correlations. *Proceedings of symposia in pure mathematics*, v. 69, 2001.

[5] P. Billingsley: *Probability and Measure*. John Wiley & Sons, 1986. (Wiley series in probability and mathematical statistics)

[6] R. Bioni Liberalquino; M. Monge; S. Galatolo; L. Marangio: Chaotic itinerancy in random dynamical system related to associative memory models. *Mathematics*, v. 6, n. 3, p. 39, 2018. Available at `https://www.mdpi.com/2227-7390/6/3/39`.

[7] G. Birkhoff: *Lattice theory*, 3. ed. Providence: American Mathematical Society, 1967. (Colloquium publications, v. 25)

[8] E. Bishop, K. de Leeuw: The representations of linear functionals by measures on sets of extreme points. *Annales de l'institut Fourier*, v. 9, p. 305–331, 1959.

[9] J. L. Doob: *Measure Theory*. New York: Springer-Verlag, 1994. (Graduate Texts in Mathematics, v. 143)

[10] N. Dunford, J. T. Schwartz: *Linear Operators. Part I: General Theory*. New York: Interscience, 1958.

[11] M. Einsiedler, T. Ward: *Ergodic Theory with a view towards Number Theory.* London: Springer-Verlag. 2011. (Graduate Texts in Mathematics, v. 259)

[12] L. C. Evans, R. F. Gariepy: *Measure Theory and Fine Properties of Functions.* Boca Raton: CRC, 1992.

[13] S. R. Foguel: *The ergodic theory of Markov processes.* New York: Van Nostrand Reinhold Company. 1969.

[14] S. Galatolo, M. Monge, I. Nisoli: Existence of noise induced order, a computer aided proof. Preprint available at `http://arxiv.org/pdf/1702.07024v2.pdf`.

[15] S. Galatolo, M. Monge, I. Nisoli: Rigorous approximation of stationary measures and convergence to equilibrium for iterated function systems. *J. Phys. A*, v. 49, n. 27, p. 274001, 2016.

[16] S. Galatolo, I. Nisoli: An elementary approach to rigorous approximation of invariant measures. *SIAM J. Appl. Dyn. Syst.*, v. 13, p. 958–985, 2014.

[17] S. Galatolo, I. Nisoli: Rigorous computation of invariant measures and fractal dimension for maps with contracting fibers: 2D Lorenz-like maps. *Ergod. Theory Dyn. Syst.*, v. 36, p. 1865–1891, 2016.

[18] L. Glass: Synchronization and rhythmic processes in physiology. *Nature*, v. 410, n. 6825, p. 277–284, 2001.

[19] P. R. Halmos: *Measure Theory.* New York: Springer-Verlag, 1950.

[20] E. Hopf: The general temporally discrete Markoff process. *Journal of Rational Mechanics and Analysis*, v. 3, p. 13–45, 1954.

[21] J. J. Hopfield: Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci.*, v. 79, p. 2554–2558, 1982.

[22] J. J. Hopfield, D. I. Feinstein, R. G. Palmer: ‘Unlearning’ has a stabilizing effect in collective memories. *Nature*, v. 304, n. 5922, p. 158–159, 1983.

[23] O. Hoshino et al. Role of itinerancy among attractors as dynamical map in distributed coding scheme. *Neural Networks*, v. 10, n. 8, p. 1375–1390, 1997.

[24] E. Kaniuth: *A Course in Commutative Banach Algebras*. New York: Springer Science+Business Media, 2009. (Graduate texts in mathematics, v. 246)

[25] Y. Kifer: *Ergodic Theory of Random Transformations*. Boston: Birkhäuser, 1948. (Progress in probability and statistics, v. 10)

[26] C. Kim: A generalization of Ito's theorem concerning the pointwise ergodic theorem. *Ann. Math. Statist.*, v. 16, p. 2145–2148, 1968.

[27] J. Komorník: Asymptotic periodicity of Markov and related operators. In: C.K.R.T. Jones et al. *Dynamics reported: expositions in dynamical systems*. Berlin, Heidelberg: Springer-Verlag, 1993. (New series, v. 2)

[28] U. Krengel: *Ergodic theorems*. Berlin, New York: de Gruyter, 1985. (De Gruyter studies in mathematics, v. 6)

[29] U. Krengel: On the speed of convergence in the ergodic theorem. *Mh. Math.*, v. 86, p. 3–6, 1978.

[30] T. Ohno: Asymptotic behaviors of dynamical systems with random parameters. *Publ. RIMS, Kyoto Univ.*, v. 19, p. 83–98, 1983.

[31] E. Koerner, H. Shimizu, I. Tsuda: Memory dynamics in asynchronous neural networks. *Progr. Theoret. Phys.*, v. 78, n. 1, p. 51–71, 1987.

[32] A. Lasota, M. Mackey: *Chaos, Fractals, and Noise. Stochastic Aspects of Dynamics*, 2. ed. New York: Springer-Verlag, 1994. (Applied mathematical sciences, v. 97)

[33] J. Neveu: Existence of bounded invariant measures in ergodic theory. In: BERKELEY SIMPOSIUM ON MATHEMATICAL STATISTICS AND PROBABILITY, 5., 1965, California. *Contributions to probability theory...* Berkeley, Los Angeles: UNIVERSITY OF CALIFORNIA PRESS, 1967. v. 2, p. 461–472.

[34] M. Rosenblatt: *Markov processes. Structure and asymptotic behavior*. Berlin, Heidelberg: Springer-Verlag, 1971. (Die Grundlehren der mathematischen Wissenschaften, v. 184)

[35] J. S. Rosenthal: Quantitative convergence rates of Markov chains: a simple account. *Elect. Comm. in Probab.*, v. 7, p. 123–128, 2002.

[36] H. L. Royden: *Real Analysis*. New York: Macmillan Publishing Company, 1988.

[37] W. Rudin: *Real and Complex Analysis*. McGraw-Hill Inc. 1970.

[38] S. M. Srivastava: *A Course on Borel Sets*. Berlin, Heidelberg: Springer-Verlag, 1980. (Graduate Texts in Mathematics, v. 180)

[39] I. Tsuda: Chaotic itinerancy. *Scholarpedia*, v. 8, n. 1, p. 4459, 2013. Available at `http://www.scholarpedia.org/article/Chaotic_itinerancy`.

[40] I. Tsuda: Chaotic itineracy and its roles in cognitive neurodynamics. *Current Opinion in Neurobiology*, v. 31, p. 67–71, 2015.

[41] I. Tsuda: Hypotheses on the functional roles of chaotic transitory dynamics. *Chaos*, v. 19, p. 015113, 2009.

[42] J. W. Milnor *Attractor* Scholarpedia: http://www.scholarpedia.org/article/Attractor

[43] S. M. Ulam: *Problems in Modern Mathematics*. New York: John Wiley & Sons, Inc. 1964.

[44] M. Viana: *Lectures on Lyapunov Exponents*. Cambridge: Cambridge University Press, 2014. (Cambridge Studies in Advanced Mathematics)

[45] C. Villani: *Topics in Optimal Transportation*. Providence: American Mathematical Society, 2000. (Graduate Studies in Mathematics)

[46] K. Yosida, S. Kakutani: Operator-theoretical treatment of Markoff's process and mean ergodic theorem. *Annals of Mathematics*, v. 42, n. 1, p. 188–228, 1941.

[47] M. Zeller, M. Bauer, W. Martienssen: Neural dynamics modeled by one-dimensional circle maps. *Chaos Solitons & Fractals*, v. 5, p. 885–893, 1995.

[48] H. Zmarrou, A. J. Homburg: Dynamics and bifurcations of random circle diffeomorphism. *Discrete & Continuous Dynamical Systems - B*, v. 10, p. 719–731, 2008.

[49] H. Zmarrou: *Bifurcation of random maps*. Academisch proefschrift. Korteweg-de Vries Institute for Mathematics, Universiteit van Amsterdam. 2008.