RIGOROUS COMPUTATIONS OF DYNAMICAL QUANTITIES

A Doctoral Thesis

XIAOLONG NIU

Submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy of Loughborough University

Acknowledgements

In the course of writing up my thesis, many people deserve thanks for their assistance.

Special thanks to my supervisor Dr. Wael Bahsoun for his continuous encouragement throughout my Ph.D. I greatly appreciate his guidance, enthusiasm and patience that made this thesis possible. I also thank my second supervisor Professor Huaizhong Zhao for his support.

Many thanks to Dr. Stefano Galatolo and Dr. Isaia Nisoli for their help and collaboration during the 'School on Computation and Computability in Dynamics', 21 July-2 August 2014, ICTP, Trieste, Italy. In this regards, I would also like to thank ICTP and Professor Stefano Luzzatto for their hospitality during this school. It had a big impact on my thesis.

I would like to thank my office mates for creating a good atmosphere. I would also like to thank Loughborough University and the Department of Mathematical Sciences for their financial support.

Last but not least, I would like to thank my parents for their continuous support. Without them, I would not have been able to study in the UK.

Abstract

This thesis is concerned with rigorous computation of dynamical quantities. In particular, we provide rigorous computation of diffusion coefficients for uniformly expanding maps of the interval. Moreover, we provide a rigorous computational scheme for linear response and we apply it in the case of uniformly expanding circle maps. Our results have been implemented successfully on a computer. Examples are included to illustrate the computer implementation. The new outcomes of this thesis are based on our work in [5, 6].

Contents

Introduction

Dynamical systems is a modern mathematical field which is concerned with studying phenomena that evolve over time. In dynamical systems, the time-evolution of the process is modelled by iterates of a map (or a flow). Most systems of interest are chaotic; i.e., their orbits are sensitive dependent to initial conditions. Thus, although chaotic dynamical systems are deterministic, their long-term behaviour is impossible to predict by following the orbit of the map.

Ergodic theory provides the probabilistic solution to this problem: the Birkhoff Ergodic Theorem states that if a map has an ergodic invariant measure then the time average of an integrable observable along the orbits of the map converges almost everywhere, with respect to the invariant measure, to the space average of the observable. Thus, it is important to get quantitative information on invariant measures to forecast statistically the long-term dynamics.

To make use of the Ergodic Theorem, one would need to select among many possible invariant measures, the most meaningful ones. This means invariant measures that provide information for a large set of initial conditions. Such invariant measures are called physical measures. For example, when the reference measure is Lebesgue, absolutely continuos invariant measures are physical measures since they provide information for a large set of initial conditions.

Recently, there have been remarkable advances in studying the existence of physical measures for different classes of uniformly hyperbolic dynamical systems [10, 13, 29, 50], and references therein. The main technique employed in [10, 13, 29, 50] is to prove that the transfer operator associated with the system is quasi-compact when acting on a suitable Banach space. Consequently, several researchers exploited the nice functional analytic results on transfer operators associated with uniformly hyperbolic systems and computed rigorously physical measures, in appropriate topologies, and found computer assisted proofs to approximate associated spectral data. The computational approach is usually based on finding a suitable finite rank approximation of the transfer operator associated

with the original system. Such techniques have proved to be robust computationally and to be successful when approximating physical measures of uniformly expanding systems [3, 26, 40, 45], uniformly hyperbolic systems [22, 25], and one-dimensional non-uniformly expanding maps [4, 26, 46]. It has also proved to be a successful approach in approximating spectral data [2, 18, 23, 24, 27, 40].

In this thesis, we will be concerned with the rigorous computation of two quantities for systems whose transfer operator admits a spectral gap on a suitable Banach space. The first quantity is the diffusion coefficient, which is the variance of the normal distribution that the corrected Birkhoff averages of an observable converges to. It is well known that in a setting like this the Central Limit Theorem holds [39]. This is done in Chapter 2 and is based on our work in [5]. In Chapter 2 we use Ulam's method [52] to provide rigorous approximations of diffusion coefficients for expanding maps. Such coefficients are focal in the study of limit theorems for dynamical systems (see [17, 33, 39, 44] and references therein).

In [47], following the approach of [34], Pollicott used a Fourier approximation scheme to estimate diffusion coefficients for expanding maps. The approach of [47] requires the map to have a Markov partition and to be piecewise analytic. Although the result of [47] provides an order of convergence, it does not compute the constant hiding in the rate of convergence. In our approach, we do not require the map to admit a Markov partition and we only assume it is piecewise C^2. More importantly, our approximation is rigorous; i.e., given a map, an observable, and a pre-specified tolerance on error, we approximate the diffusion coefficient rigorously up to the pre-specified error (see Theorem 2.1.2).

The second quantity that we compute in this thesis is the so called 'linear response', or the derivative with respect to noise, of a physical measure. This is done in Chapter 3 and is based on our work in [6]. A question of central interest from both theoretical and applied point of views in dynamical systems is the following: given a deterministic dynamical system that admits a physical measure, how does the physical measure change if the original system gets perturbed, perhaps randomly? It is known that in certain situations the physical measure changes smoothly and a formula of such a "derivative" can be obtained [8, 11, 17, 29, 36, 48]. This is called the Linear Response formula. We refer to [9] for a recent survey and progress in this area of research. However, from a rigorous computational point of view there are no results in the literature that approximate the derivative of a physical measure up to a pre-specified error in a suitable topology. In Chapter 3 we pioneer this direction of research. This work can be considered as a starting point of rigorous numerical approaches that aim to identify tipping points in the statisti-

cal behaviour of systems studied in applications, such as the the comprehensive climate dynamics models considered in [43].

The thesis is organised as follows. In Chapter 1 we review tools from measure theory, functional analysis, ergodic theory and dynamical systems. Moreover, this chapter includes some background about stochastic matrices and computer arithmetic. These two ingredients are essential for our rigorous computations in this thesis.

Chapter 2 is based on our work in [5]. It is concerned with the rigorous computation of diffusion coefficients for uniformly expanding maps of the interval. In particular, In Section 2.1, we first introduce our system and the assumptions on it. We then state the problem and introduce the method of approximation. The statement of the main result (Theorem 2.1.2) is given in Section 2.1. Section 2.2 contains the proofs and an algorithm. Section 2.3 contains an example that illustrates the implementation of the algorithm of Section 2.2.

Chapter 3 is based on our work in [6]. It is concerned with the rigorous computation of linear response. In particular, in Section 3.1 we present a general setting in which the formula corresponding to the linear response can be obtained. In this section we also show how the formula of such derivative can be rigorously computed using a computer. In Section 3.2 we apply our results to expanding circle maps. In particular, in this section we first find suitable Banach spaces and suitable discretization schemes that can be used to compute linear response. In Section 3.3 we present an example where we compute, up to a pre-specified error in the L^∞-norm, the derivative of the physical measure of an expanding circle map under stochastic perturbations.

Finally, Appendix A includes proofs and tools used in the computations in the example of Section 3.3 of Chapter 3.

Chapter 1

Preliminaries

In this chapter, we review some basic definitions and results from measure theory, functional analysis, ergodic theory, and matrix analysis. This chapter also includes background on computer arithmetic. For measure theory we mainly use [30], for functional analysis we mainly use [54], for ergodic theory we mainly use [15], [53], for matrix analysis we use [1], [12], and for computer arithmetic we mainly use [51].

1.1 Mathematical Background

1.1.1 Measure Theory

Definition 1.1.1 ($\sigma-$**algebra**). *A collection $\mathfrak{B}$ of subsets of X is called $\sigma-$algebra if and only if:*

(1) $X \in \mathfrak{B}$;

(2) *for any* $B \in \mathfrak{B}$, $\quad X \setminus B \in \mathfrak{B}$;

(3) *if* $B_n \in \mathfrak{B}$, *for* $n = 1, 2, ...,$ *then* $\bigcup_{n=1}^{\infty} B_n \in \mathfrak{B}$.

Elements of $\mathfrak{B}$ are usually referred to as *measurable sets.*

The simplest example of a $\sigma-$algebra of a set X is the collection of empty set and the set itself $\{\emptyset, X\}$. Let us see one more example of $\sigma-$algebra:

Example 1.1.1. *For a subset A of X, the collection $\{\emptyset, A, A^c, X\}$ is a $\sigma-$algebra, denoted by $\mathfrak{B}$. In this case, it is easy to see that $\mathfrak{B}$ satisfies Definition 1.1.1.*

Definition 1.1.2 (**Measure**). *A function $\mu : \mathfrak{B} \to \mathbb{R}^+$ is called a measure on $(X, \mathfrak{B})$ if and only if :*

(1) $\mu(\emptyset) = 0$;

(2) *for any sequence of $\{B_n\}$ with $B_n \in \mathfrak{B}$ and $B_n \cap B_m = \emptyset$, $m \neq n$, we have*

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

If $\mu(X) = 1$, we say μ is a probability measure and $(X, \mathfrak{B}, \mu)$ is a *probability space,* or a normalized measure space.

One of the most natural measures is called Lebesgue measure. It is defined using Borel $\sigma-$algebra. The formal definition of Borel $\sigma-$algebra and Lebesgue measure are presented below:

Definition 1.1.3 (**Topological Space**). *Let X be a set. A topology Y on X is a collection of subsets of X (called open sets) with the following properties:*

1. $\emptyset, X \in Y$.
2. *The union of any collection of open sets is an open set.*
3. *The intersection of any finite collection of open sets is an open set.*

Definition 1.1.4 (**Borel $\sigma-$algebra**). *Let X be a topological space. Then the smallest $\sigma-$algebra generated by open sets is called the Borel $\sigma-$algebra of X. Elements of $\mathfrak{B}$ are called Borel subsets of X.*

Definition 1.1.5 (**Lebesgue Outer Measure**). *Let $\mathfrak{B}$ be the Borel $\sigma-$algebra on $\mathbb{R}$. The Lebesgue outer measure of any set $A \subseteq \mathbb{R}$ is the non-negative real-number*

$$m^*(A) = \inf \left\{ \sum_{n=1}^{\infty} l(I_n) : I_n \text{ are intervals}, A \subseteq \cup_{n=1}^{\infty} I_n \right\}$$

where $l(I) = |b - a|$, for interval $I = [a, b]$.

Definition 1.1.6 (**Lebesgue Measurable**). *A set $E \subseteq \mathbb{R}$ is Lebesgue measurable if for every set $A \subseteq \mathbb{R}$ we have*

$$m^*(A) = m^*(A \cap E) + m^*(A \cap E^c)$$

where $E^c = \mathbb{R} \setminus E$. For Lebesgue measurable sets, Lebesgue outer measure is called just Lebesgue measure.

If a function $f : X \to Y$ is measurable, it preserves the structure between measurable spaces. This means the preimage of measurable set in Y is also a measurable set on X. The formal definition is the following:

Definition 1.1.7 (**Measurable Function**). *Let $(X, \mathfrak{B}, \mu)$ be a measure space. A function $f : X \to \mathbb{R}$ is said to be measurable if for all $c \in \mathbb{R}, f^{-1}(c, \infty) \in \mathfrak{B}$, equivalently, if $f^{-1}(A) \in \mathfrak{B}$ for any Borel set $A \subset \mathbb{R}$.*

Absolutely continuous measures play a crucial role in this thesis. The definition of absolutely continuous measure is the following.

Definition 1.1.8 (**Absolutely Continuous Measure**). *Let ν and μ be two measures on the same measure space $(X, \mathfrak{B})$. We say that ν is absolutely continuous with respect to μ if for any $A \in \mathfrak{B}$, such that $\mu(A) = 0$, we have $\nu(A) = 0$. We write $\nu \ll \mu$.*

A useful property of absolutely continuous measures is illustrated by the Radon-Nikodym Theorem which we state below.

Theorem 1.1.2 (**Radon-Nikodym Theorem**). *[15]*

Let $(X, \mathfrak{B})$ be a measure space and let μ and m be two normalized measures on $(X, \mathfrak{B})$. If $\mu \ll m$, then there exists a unique $f \in L^1(X, \mathfrak{B}, m)$ such that for every $A \in \mathfrak{B}$,

$$\mu(A) = \int_A f dm.$$

f is called the Radon-Nikodym derivative and is denoted by $\frac{d\mu}{dm}$.

1.1.2 Matrix Analysis

In my research matrix theory is a bridge that connects the mathematical theory with the computer implementations. The objects that we study, like L^1 functions, BV functions, the Frobenius-Perron operator, invariant density functions, are all objects that can not be dealt with directly on a computer. Computers can only store finite dimensional elements. A matrix can be stored and run very well on a computer and can be used to approximate operators. In my study, I use stochastic matrices to approximate Frobenius-Perron operators. In this section, I will give the definition of a stochastic matrix and explain its properties (See [12] for more details).

First of all, every element in a stochastic matrix represent a probability. Thus a stochastic matrix is supposed to be a nonnegative matrix.

Definition 1.1.9 (**Nonnegative Matrix**). *A matrix $A \in \mathbb{R}_{m\times n}$ is said to be nonnegative iff no element of A is negative.*

Example 1.1.3. *This concept is straight forward, a matrix A defined by:*

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \tag{1.1.1}$$

is a nonnegative matrix.

Nonsingular is an important concept when we study matrices. There are many ways to define it. Here is one way to define a nonsingular matrix [12]:

Definition 1.1.10 (**Nonsingular Matrix**). *A matrix A is said to be nonsingular if its determinant is nonzero.*

We can also say an $n\times n$ matrix A is a nonsingular matrix if all the columns contained in this matrix are linearly independent. In other words, every column in this matrix contains useful information. Thus, a nonsingular matrix can be understood as every column of it contains useful information.

Example 1.1.4. *Recall the matrix* (1.1.1) *in the previous example. Its determinant* $\det(A) = 1 \cdot 0 - 2 \cdot 0 = 0$*, so it is a singular matrix.*

Let

$$B = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix}. \tag{1.1.2}$$

Its determinant $\det(B) = 1 \cdot 0 - 2 \cdot 1 = -2 \neq 0$*, thus matrix* B *is nonsingular.*

In matrix analysis, there are two kinds of eigenvectors associated with eigenvalue λ: the left eigenvector, which is row vector, and the right eigenvector, which is a column vector.

Definition 1.1.11 (**Right Eigenvector**). *Let* A *be a matrix. If* $x \neq 0$ *is a vector, such that there exists a number* λ *with* $Ax = \lambda x$*, then* x *is called a right eigenvector of* A *and* λ *is the corresponding eigenvalue.*

Definition 1.1.12 (**Spectrum**). *Let* A *be a square matrix. The set of all eigenvalues ofA it is called the spectrum of* A*.*

Remark 1.1.1. *Unlike Definition 1.1.25, the operator here, represented by the matrix, is finite dimensional. Hence the spectrum consists only of discrete spectrum represented by the eigenvalues.*

Definition 1.1.13 (**Spectral Radius**). *Let* $\lambda_1, \lambda_2, ..., \lambda_n$ *be the eigenvalues of a square matrix* $A \in \mathbb{C}^{n \times n}$*. Then its spectral radius* $\rho(A)$ *is defined as*

$$\rho(A) = \max\{|\lambda_1|, ..., |\lambda_n|\}.$$

Definition 1.1.14 (**Permutation Matrix**). *Let unit vector* e_j *be the vector with one in the* j*-th position and zero elsewhere.*
An $n \times n$ *matrix* P *is said to be permutation matrix if it has the form*

$$P = [e_{i_1}, e_{i_2}, ..., e_{i_n}]$$

where $i_1, i_2, ...i_n$ *is a permutation of* $1, 2, ..., n$*.*

Example 1.1.5. *Let* $n = 3$*,*

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } P = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

are both permutation matrices.

Definition 1.1.15 (**Irreducible Matrix**). *Let* A *be a square matrix with* $n \geq 2$*. Suppose there exists is a permutation matrix* P *such that,*

$$P'AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$$

where A_{11}, A_{22} *are square matrices of dimension less than* n*, then* A *is called reducible. If no such* P *exists, then* A *is irreducible.*

Here is an example of a reducible matrix.

Example 1.1.6. *Let matrix A be a nonnegative matrix*

$$A = \begin{bmatrix} 0 & 0 & 0.5 & 0 \\ 0 & 0 & 0.5 & 1 \\ 0.5 & 0 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \end{bmatrix} \tag{1.1.3}$$

The permutation matrix P and its transpose

$$P = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad P' = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Compute the matrix $P'AP$

$$P'AP = \begin{bmatrix} 0 & 1 & 0.5 & 0 \\ 1 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0.5 & 0 \end{bmatrix}.$$

It can be written as

$$P'AP = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

where

$$A_{11} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, A_{22} = \begin{bmatrix} 0 & 0.5 \\ 0.5 & 0 \end{bmatrix} \text{ and } A_{12} = \begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}$$

Thus, the matrix A is reducible.

Remark 1.1.2. *The irreducibility of a matrix is a important condition to use the Perron-Frobenius Theorem (See [12]). As we can see in the above example, a nonnegative matrix may be reducible, but a positive matrix is irreducible (See [12]).*

Theorem 1.1.7 (**Perron-Frobenius Theorem**)**.** *If the matrix $A \in \mathbb{R}_{m\times n}$ is nonnegative and irreducible then:*

1. *A has a positive eigenvalue λ, equal to the spectral radius of A.*
2. *There is a positive right eigenvector associated with the eigenvalue λ.*
3. *The eigenvalue λ has (algebraic) multiplicity* 1.

Definition 1.1.16 (**Stochastic Matrices**)**.** *Matrix $P \in \mathbb{R}_{n\times n}$ is said to be a stochastic matrix iff P is nongegative and*

$$\sum_{j=1}^{n} P_{ij} = 1, \ i = 1, 2, ..., n.$$

Theorem 1.1.8. *A nonnegative matrix A is stochastic iff it has dominant eigenvalue* 1 *with right eigenvector given by $v' = [1, 1, ..., 1]$. In particular, the spectral radius of a stochastic matrix is* 1.

The above theorem follows from Theorem 1.1.7.

Theorem 1.1.9. *Let A be a nonnegative matrix with maximal real eigenvalue λ. If there is a positive right eigenvector z associated with λ, then*

$$A = \lambda ZPZ^{-1}$$

where P is a stochastic matrix and $Z = diag\{z_1, ..., z_n\}$.

1.1.3 Functional Analysis

We first introduce the notion of a linear space.

Definition 1.1.17 (**Linear Space**). *A linear space (also known as a vector space) is a set $\mathcal{V}$ over a field K together with two operations and multiply by elements of K, that satisfy the following:*

1. *For any $v_1, v_2 \in \mathcal{V}$ we have $v_1 + v_2 \in \mathcal{V}$.*
2. *(Scalar multiplication) For any $v \in \mathcal{V}$ and $k \in K$, $kv \in \mathcal{V}$.*
3. *(Commutative law of vector addition) $v_1 + v_2 = v_2 + v_1$ for each pair of vectors $v_1, v_2 \in \mathcal{V}$.*
4. *(Associative law of vector addition) $(v_1 + v_2) + v_3 = v_1 + (v_2 + v_3)$ for each triple of vectors $v_1, v_2, v_3 \in \mathcal{V}$.*
5. *(Identity element of addition) There is a unique vector $\mathbf{0}$, called the zero vector, such that $v_1 + \mathbf{0} = v_1$ for every vector $v_1 \in \mathcal{V}$.*
6. *(Inverse elements of addition) For each vector $v_1 \in \mathcal{V}$ there is a unique vector $-v_1$ such that $v_1 + (-v_1) = \mathbf{0}$.*
7. *(Distributivity of scalar multiplication with respect to vector addition) $k(v_1 + v_2) = kv_1 + kv_2$ for each $k \in K$ and each pair of vectors $v_1, v_2 \in \mathcal{V}$.*
8. *(Distributivity of scalar multiplication with respect to field addition) $(r + k)v_1 = rv_1 + kv_1$ for each pair $r, k \in K$ and each vector $v_1 \in \mathcal{V}$.*
9. *(Compatibility of scalar multiplication with field multiplication) $(rk)v_1 = r(kv_1)$ for each pair $r, k \in K$ and each vector $v_1 \in \mathcal{V}$.*
10. *For each vector $v_1 \in \mathcal{V}$, $\mathbf{1}v_1 = v_1$.*

We now define the notion of a norm:

Definition 1.1.18 (**Norm**). *Let $\mathcal{F}$ be a linear space. A function $\|\cdot\| : \mathcal{F} \to \mathbb{R}^+$, where $\mathbb{R}^+ = [0, \infty)$ is called a norm if it has the following properties:*

$$\|f\| = 0 \Leftrightarrow f \equiv 0,$$
$$\|\alpha f\| = |\alpha| \|f\|,$$
$$\|f + g\| \leq \|f\| + \|g\|.$$

For $f, g \in \mathcal{F}$ and $\alpha \in \mathbb{R}$, the space $\mathcal{F}$ endowed with a norm $\|\cdot\|$ is called a normed linear space.

Definition 1.1.19 (**Cauchy sequence**). *A sequence* $\{f_n\}$ *in a normed linear space* $\mathcal{F}$ *is called a Cauchy sequence if, for any* $\epsilon > 0$*, there exists an* $N > 1$ *such that for any* $n, m \geq N$*,*

$$\|f_n - f_m\| < \epsilon.$$

A normed linear space $\mathcal{F}$ is complete if every Cauchy sequence converges, i.e. for each Cauchy sequence $\{f_n\}$ there exists $f \in \mathcal{F}$ such that

$$\lim_{n\to\infty} \|f_n \to f\| = 0.$$

Now, we can define a Banach space.

Definition 1.1.20 (**Banach space**). *A complete normed space is called a Banach space.*

In our research, the spectrum of transfer operator and the dynamical quantities are all define on a suitable Banach spaces. We now present examples of Banach spaces that will be used later in this thesis.

Examples of Banach spaces:

Example: L^p space and L^∞ space Let $(X, \mathfrak{B}, \mu)$ be a measure space. Let $1 \leq p < \infty$. The family of real valued measurable functions $f : X \to \mathbb{R}$ satisfying

$$\int_X |f(x)|^p d\mu < \infty$$

is called the $L^p(X, \mathfrak{B}, \mu)$ space.

When equipped with the norm

$$\|f\|_p = \left(\int_X |f(x)|^p d\mu \right)^{\frac{1}{p}},$$

$(L^p, \|\cdot\|_p)$ is a Banach space.

The space of almost everywhere bounded measurable functions on $(X, \mathfrak{B}, \mu)$ is denoted by L^∞. The L^∞ norm $\|\cdot\|_\infty$ is given by:

$$\|f\|_\infty = \text{essup}|f(x)| = \inf\{M : \mu\{x : f(x) > M\} = 0\}.$$

The L^∞ space with the norm $\|\cdot\|_\infty$ is a Banach space.

Example: Smooth function spaces Let $r \geq 0$. $C^r(X)$ denotes the space of r-times continuously differentiable real functions $f : X \to \mathbb{R}$ with the norm

$$\|f\|_{C^r} = \max_{0\leq k\leq r} \sup_{x\in X} |f^{(k)}(x)|,$$

where $f^{(k)}(x)$ is the k-th derivative of $f(x)$ and $f^0(x) = f(x)$. $(C^r, \|\cdot\|_{C^r})$ forms a Banach space.

Example: Functions of Bounded Variation We now present an example of a Banach space that will turn out to be very useful in this thesis. We will give more details in this example.

Definition 1.1.21 (**Total Variation**). *Let $f : [a, b] \to \mathbb{R}$ be a function and let $\mathcal{P} = \{a = x_0 < x_1 < ... < x_k = b\}$ be a partition of $[a, b]$. The number*

$$\bigvee_{[a,b]} f = \sup_{\mathcal{P}} \left\{ \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \right\}$$

is called total variation of f on $[a, b]$.

Definition 1.1.22 (**Bounded Variation**). *Let $f : [a, b] \to \mathbb{R}$. If there exists a positive number M such that*

$$\bigvee_{[a,b]} f \leq M$$

then f is said to be of bounded variation on $[a, b]$.

We now present an example of a function of bounded variation.

Example 1.1.10. *Let $f(x) = \sin(x)$. We compute its variation on $[-\pi, \pi]$.*

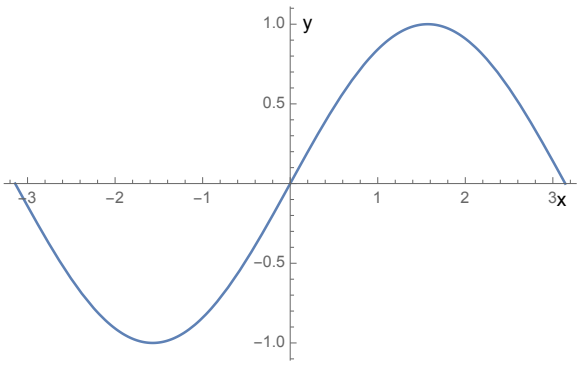

Figure 1.1: $f(x) = \sin(x)$ on $[-\pi, \pi]$

Note that, the function $f(x)$ is monotonic on $[-\pi, -\frac{\pi}{2}]$, $[-\frac{\pi}{2}, \frac{\pi}{2}]$ and $[\frac{\pi}{2}, \pi]$. We have

$$\begin{aligned}
\bigvee_{[a,b]} f(x) =& \sup_{\mathcal{P}} \{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} \\
\leq& \sup_{\mathcal{P}\cap[-\pi,-\frac{\pi}{2}]} \{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} + \sup_{\mathcal{P}\cap[-\frac{\pi}{2},\frac{\pi}{2}]} \{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} \\
&+ \sup_{\mathcal{P}\cap[\frac{\pi}{2},\pi]} \{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} \\
=& |\sin(-\pi) - \sin(-\frac{\pi}{2})| + |\sin(-\frac{\pi}{2}) - \sin(\frac{\pi}{2})| + |\sin(\frac{\pi}{2}) - \sin(\pi)| \\
=& 1 + 2 + 1 = 4
\end{aligned}$$

Thus, $f(x) = \sin(x)$ is of bounded variation on $[-\pi, \pi]$.

Example 1.1.11. *Consider $f(x) = \tan(x)$, let us compute its total variation on $[0, \pi]$. We will show that it is not a function of bounded variation.*

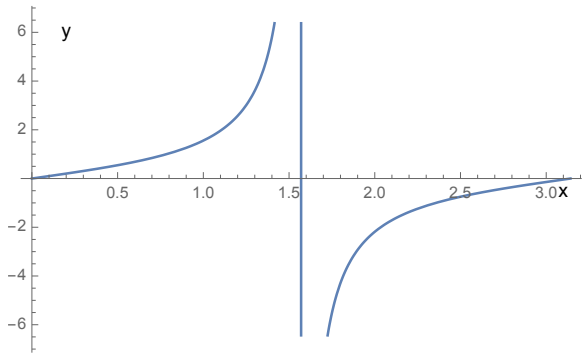

Figure 1.2: $f(x) = \tan(x)$ on $[0, \pi]$

Note that the function $f(x)$ is monotonic on $[0, \frac{\pi}{2}]$ and $[\frac{\pi}{2}, \pi]$. By the definition of total variation, we have

$$\begin{aligned}\bigvee_{[0,\pi]} f(x) =& \sup_{\mathcal{P}}\{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} \\ \leq& \sup_{\mathcal{P}\cap[0,\frac{\pi}{2}]}\{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} + \sup_{\mathcal{P}\cap[\frac{\pi}{2},\pi]}\{\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\} \\ =& |\tan(0) - \lim_{x_k \to \frac{\pi}{2}} \tan(x_k)| + |\lim_{x_k \to \frac{\pi}{2}} \tan(x_k) - \tan(\pi)| = \infty.\end{aligned}$$

Thus, $f(x) = \tan(x)$ on $[0, \pi]$ is not a function of bounded variation.

Definition 1.1.23 (**Bounded Variation Norm**). *Let $f \in L^1([0,1])$. Let*

$$\|f\|_{BV} = \bigvee_{[0,1]} f + \|f\|_{L^1},$$

where

$$\bigvee_{[0,1]} f = \inf_{\overline{f}} \left\{ \bigvee_{[0,1]} \overline{f} \ : f = \overline{f} \ a.e. \right\}.$$

Then $(BV, \|\cdot\|_{BV})$ is a Banach space [20].

The spectral properties of linear operator We recall some basic definitions from spectral theory of bounded linear operators. We refer the reader to [20] and [35] for more details.

Definition 1.1.24 (**Linear Operator**). *Let B_1, B_2 be two Banach spaces. A function P that sends $u \in B_1$ into $v = Pu \in B_2$ is called a linear operator if P preserves linear relations:*

$$P(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 P u_1 + \alpha_2 P u_2,$$

for all u_1, u_2 of B_1 and all scalars α_1, α_2.

Example 1.1.12. *Let B be a Banach spaces. Let $f(x), g(x) \in B$ be complex-valued functions of bounded variation over $[a, b]$. The Stieltjes integral [20] of f with respect to*

g, denoted by P

$$Pg(x) = \int_a^b g(x)df(x)$$

defines a linear operator on the space $C([a, b])$.

Let B_1, B_2 be two Banach spaces, and $P : B_1 \to B_2$ be a bounded linear operator, such that the operator norm is bounded. A complex number λ is called an eigenvalue of P if there is a non-zero $u \in B_1$, $B_1 = B_2$ such that

$$Pu = \lambda u.$$

u is called an eigenvector of P associated with the eigenvalue λ.

Definition 1.1.25 (**Resolvent and Spectrum of** P)**.** *For a bounded linear operator* $P : B \to B$*, we define the spectrum of it as:*

$$\sigma(P) := \{\lambda : (\lambda I - P) \text{ has no bounded inverse}\}.$$

The complementary set of $\sigma(P)$ *is the resolvent set of* P*, denoted by* $Res(P)$.

Definition 1.1.26. *The operator norm is given by*

$$\|P\| = \sup_{v \in \mathfrak{B}, \|v\|=1} \|Pv\|.$$

Definition 1.1.27 (**Spectral radius**)**.** *For a bounded linear operator* P *acting on a Banach space* $(B, \|\cdot\|)$*, the spectral radius is defined as:*

$$\rho(P) = \lim_{n\to\infty} (\|P^n\|)^{1/n}.$$

Definition 1.1.28 (**Essential spectral radius**)**.** *The essential spectral radius* $\rho_{ess}(P)$ *of* P *is the smallest number* $\rho_{ess} \geq 0$ *such that any* $\lambda \in \sigma(P)$ *with modulus* $|\lambda| > \rho_{ess}$ *is an isolated eigenvalue of finite multiplicity.*

Now, we can discuss the notion of a spectral gap of a bounded linear operator. It is a very important property that is related to exponential mixing in dynamical systems. I will first explain the notion of quasi-compactness.

Theorem 1.1.13 (**Theorem of Ionescu-Tulcea and Marinescu**)**.** *[15]*

Let $(B_1, ||\cdot||_{B_1})$ *and* $(B_2, \|\cdot\|_{B_2})$ *be two complex Banach space with* $B_1 \subset B_2$*, and a linear operator* P *from* B_1 *into* B_1 *be bounded with respect to both* $||\cdot||_{B_1}$ *and* $\|\cdot\|_{\hat{B_1}}$*, which is the restriction of* $\|\cdot\|_{B_2}$ *to* B_1*, where*

$$\|P\|_{\hat{B_1}} = \sup\left\{\frac{\|Pf\|_{B_2}}{\|f\|_{B_2}}, f \in B_1, f \neq 0\right\}.$$

Assume that

1. *If* $f_n \in B_1$, $f \in B_2$, $\lim_{n\to\infty} \|f_n - f\|_{B_2} = 0$, *and* $||f_n||_{B_1} \leq C$ *for all* n*, then* $f \in B_1$ *and* $||f||_{B_1} \leq C$,

2. $H = \sup_{n\geq 0} \|P_n\|_{\hat{B_1}} < \infty$,

3. *There exist* $k \geq 1$, $0 < r < 1$, *and* $R < \infty$ *such that for* $f \in B_1$,
$$||P^k f||_{B_1} \leq r||f||_{B_1} + R\|f\|_{B_2}, \tag{1.1.4}$$

4. *If* B_3 *is a bounded subset of* $(B_1, ||\cdot||_{B_1})$, *then the closure of* $P^k B_3$ *is compact in* $(B_2, \|\cdot\|_{B_2})$.

For any complex number λ, *we introduce the following notation:*
$$E(\lambda) = \{f \in B_1 . Pf = \lambda f, f \neq 0\}.$$
λ *is an eigenvalue of* P *if and only if* $E(\lambda) \neq \emptyset$.

Under the above conditions $(1) - (4)$, *the intersection of the spectrum of operator* P *with the unit circle is a set* G *of eigenvalues of* P *of modulus* 1 *which has only a finite number of elements. For each* $\lambda \in G$, $E(\lambda)$ *is finite-dimensional. Furthermore, there exist bounded linear operator* Π_λ,$\lambda \in G$, *and* Q *on* B_1 *such that*

$$P^n = \sum_{\lambda \in G} \lambda^n \Pi_\lambda + Q^n \tag{1.1.5}$$
$$\Pi_{\lambda_1}\Pi_{\lambda_2} = 0, \text{ if } \lambda_1 \neq \lambda_2, \Pi_{\lambda_1}^2 = \Pi_{\lambda_1} \tag{1.1.6}$$
$$\Pi_\lambda Q = Q\Pi_\lambda = 0 \tag{1.1.7}$$
$$\Pi_\lambda B_1 = E(\lambda) \tag{1.1.8}$$
$$\rho(Q) < 1, \tag{1.1.9}$$
where $\rho(Q) = \lim_{n\to\infty} \|Q^n\|^{1/n}$ *is the spectral radius of* Q.

The properties $(1.1.5) - (1.1.9)$ of P is one of definitions of a quasi-compact operator. Thus, under the assumptions $(1) - (4)$, the operator P is quasi-compact. In Figure 1.3, we present an example of the spectral picture of a quasi-compact operator. Inside the small circle is essential spectrum, which we don't know the eigenvalues. There are isolated eigenvalues between the circle with$r < \rho(Q) < 1$.

Remark 1.1.3. *If* P *is a matrix,* Π_λ *is a projection matrix onto the eigenspace associated with eigenvalue* λ. *If* P *is an infinite dimensional operator,* Π_λ *is the spectral projection associated with* λ *([35], [20]).*

Remark 1.1.4. *It is well known that* r *in inequality* (1.1.4) *is an upper bound on the essential spectral radius of* P *[31].*

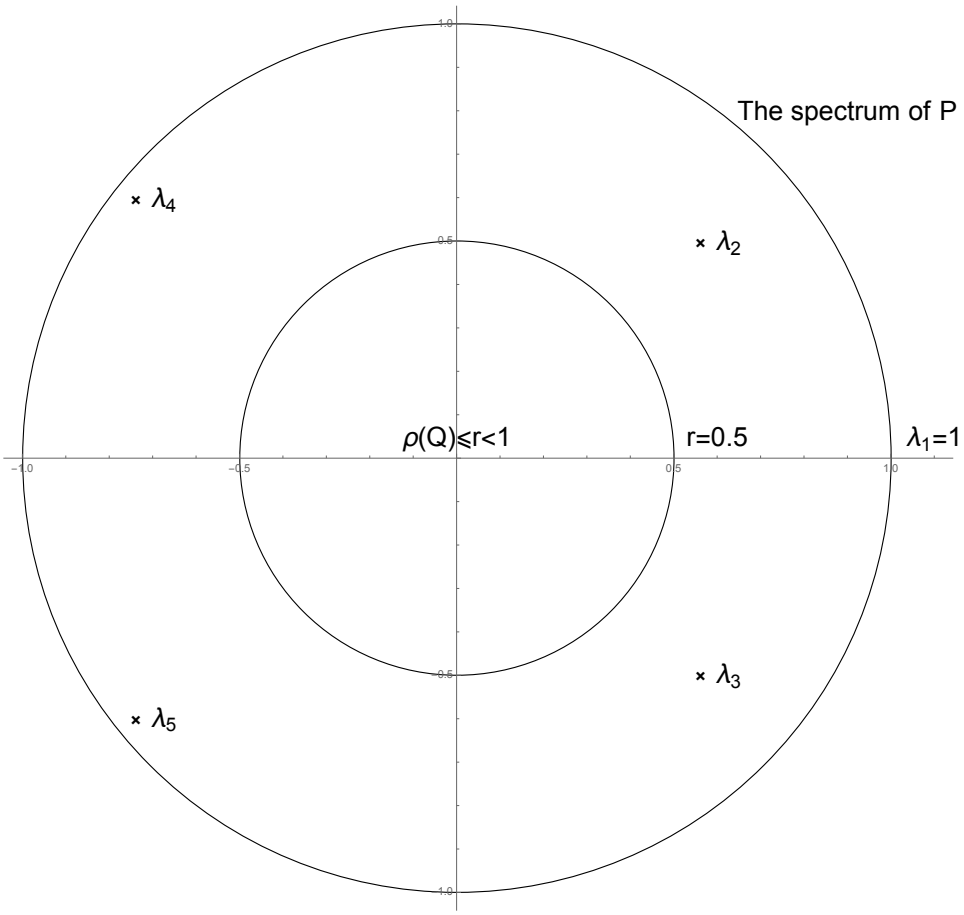

Figure 1.3: An example of spectral picture of a quasi-compact operator

Definition 1.1.29 (**Spectral Gap**). *Let B be a Banach space. A bounded linear operator $P : B \to B$ has spectral gap if it satisfied the following:*

1. *$\rho_{ess}(P) < \rho(P)$;*
2. *The eigenvalue λ with $|\lambda| = \rho(P)$ is a simple eigenvalue of P and any other P eigenvalues of modulus strictly smaller than $\rho(P)$.*

Remark 1.1.5. *In fact, Figure 1.3 shows an operator that has a spectral gap.*

Keller-Liverani's Theorem In [37], Keller and Liverani provided conditions that ensure stability of the spectrum of certain bounded linear operators[1] Their result plays an important role in the analytical background of this thesis. I will state their theorem here. It will be used later in the thesis.

[1]Similar to the Ionescu-Tulcea and Marinescu Theorem 1.1.13, the Keller-Liverani result [37] can be seen as an abstract functional analytic result. Thus, it can be introduced even before mentioning dynamical systems and ergodic theory.

Let $(B, \|\cdot\|)$ be a Banach space equipped with a second norm $|\cdot|$ such that $|\cdot| \leq \|\cdot\|$. For any bounded linear operator $P : B \to B$, consider the set

$$V_{\delta,r}(P) = \{z \in \mathbb{C} : |z| \leq r \text{ or } \operatorname{dist}(z, \sigma(P)) \leq \delta\},$$

where $\rho(P)$ is the spectrum of P with respect to $(B, \|\cdot\|)$, and define

$$H_{\delta,r}(P) := \sup\{\|(zI - P)^{-1}\| : z \in \mathbb{C} \setminus V_{\delta,r}\} < \infty.$$

Further, we define the operator norm by following:

$$|||P||| = \sup_{\|f\| \leq 1} |Pf|.$$

Theorem 1.1.14. *Let $P_i : B \to B$ be two bounded linear operators, $i = 1, 2$. Assume that: there are $C_1, M > 0$ such that for all $x \in \mathbb{N}$*

$$|P_i^n| \leq C_1 M^n;$$

and $\exists \alpha \in (0, 1)$, $\alpha < M$, and $C_2, C_3 > 0$ such that

$$\|P_i^n f\| \leq C_2 \alpha^n \|f\| + C_3 M^n |f| \quad \forall n \in \mathbb{N}\ \forall f \in B, i = 1, 2;$$

moreover, if $|z| > \alpha$, then z is not in the residual spectrum of P_i, $i = 1, 2$.
For $r \in (\alpha, M)$, let

$$n_1 = \left\lceil \frac{\ln 2C_2}{\ln r/\alpha} \right\rceil$$

$$n_2 = \left\lceil \frac{\ln 8C_3C_2(C_2 + C_3 + 2)MH_{\delta,r}(P_1)}{\ln r/\alpha} + n_1 \frac{\ln(M/r)}{\ln r/\alpha} \right\rceil.$$

If

$$|||P_1 - P_2||| \leq \frac{(r/M)^{n_1+n_2}}{8C_3(H_{\delta,r}(P_1)C_3 + \frac{C_1}{M-r})} := \epsilon_1(P, r, \delta)$$

then for each $z \in \mathbb{C} \setminus V_{\delta,r}(P_1)$, we have

$$\|(z_2 - P_2)^{-1} f\| \leq \frac{4(C_2 + C_3)}{M - r} (\frac{M}{r})^{n_1} \|f\| + \frac{1}{2\epsilon_1} |f|.$$

Set

$$\gamma = \frac{\ln(r/\alpha)}{\ln(M/\alpha)},$$

$$a = \frac{8M(C_2 + C_3)^2}{M - r} (\frac{M}{r})^{n_1} [2C_2(C_2 + C_3) + \frac{C_1}{M - r}] + \frac{2C_1}{M - r},$$

$$b = \frac{8M}{M - r} [MC_2(C_2 + C_3 + 2) + C_3](C_2 + C_3)^2 (\frac{M}{r})^{n_1} + 2C_3,$$

and

$$\epsilon_2(P_1, r, \delta) := [\frac{1}{4C_3} (\frac{M}{r})^{n_1} (\frac{1}{H_{\delta,r}(P_1)[C_2(C_2 + C_3 + 2)M + C_3] + 2C_2(C_2 + C_3) + \frac{C_1}{M-r}})]^{\frac{1}{\gamma}}.$$

If

$$|||P_1 - P_2||| \leq \min\{\epsilon_1(P_1, r, \delta), \epsilon_2(P_1, r, \delta)\} := \epsilon_0(P_1, r, \delta)$$

then for each $z \in \mathbb{C} \setminus V_{\delta,r}(P_1)$, we have

$$|||(z - P_2)^{-1} - (z - P_1)^{-1}||| \leq |||P_1 - P_2|||^{\gamma} (a\|(z - P_1)^{-1}\| + b\|(z - P_1)^{-1}\|^2).$$

Corollary 1.1.15. *If $|||P_1 - P_2||| \leq \epsilon_1(P_1, r, \delta)$, then $\sigma(P_2) \subset V_{\delta,r}(P_1)$. In addition, if $|||P_1 - P_2||| \leq \epsilon_0(P_1, r, \delta)$, then in each connected component of $V_{\delta,r}(P_1)$ that does not contain 0, both $\sigma(P_1)$ and $\sigma(P_2)$ have the same multiplicity; i.e. the associated spectral projections have the same rank.*

Figure 1.4 presents an example and an illustration of the set $V_{\delta,r}(P)$.

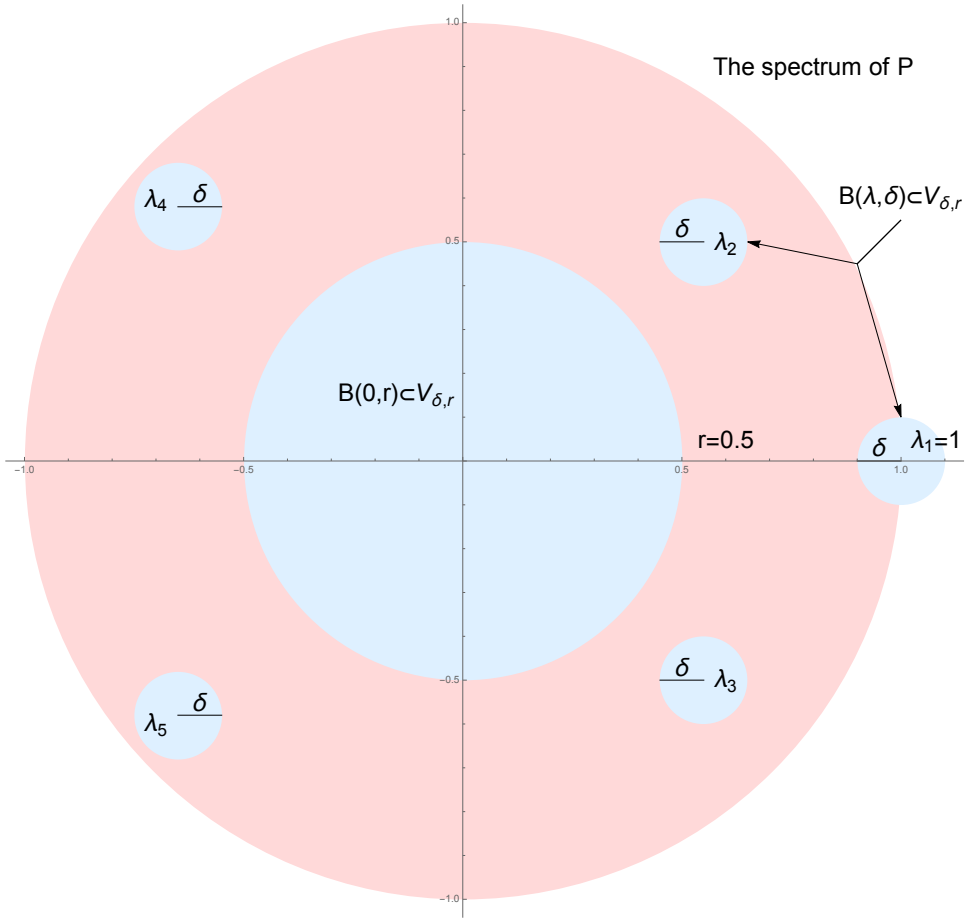

Figure 1.4: The $V_{\delta,r}$ set in the spectral picture of a quasi-compact operator, where $B(\lambda, \delta)$ is the small balls centre on eigenvalues with radius δ

Remark 1.1.6. *In the section of rigorous computation, the Keller-Liverani's Theorem ensure that the error of approximation of transfer operator can be as small as we want, as long as we choose a small enough η.*

1.1.4 Ergodic Theory

Ergodic theory studies statistical aspects of the long term behaviour of a dynamical system. In this section, I will state some definitions and theorems that will be used often in my thesis. Let us start from the formal definition of a dynamical system [15, 53] .

Definition 1.1.30 (Dynamical system). *Let $(X, \mathfrak{B}, \mu)$ be a normalised measure space and let $T : X \to X$ be measurable. The quadruple $(X, \mathfrak{B}, \mu, T)$ is called a dynamical system.*

Definition 1.1.31 (π-system). *A family $\mathcal{P}$ of subsets of X is called a π-system if and only if for any A, B in $\mathcal{P}$, their intersection is also in $\mathcal{P}$.*

Definition 1.1.32 (**Invariant Measure**). *The measurable transformation $T : X \to X$ preserves measure μ or that μ is T-invariant if $\mu(T^{-1}(B)) = \mu(B)$ for all $B \in \mathfrak{B}$.*

Theorem 1.1.16. *Let $(X, \mathfrak{B}, \mu)$ be a normalised measure space and let $T : X \to X$ be measurable. Let $\mathcal{P}$ be a π-system that generates $\mathfrak{B}$. If $\mu(T^{-1}(A)) = \mu(A)$ for any $A \in \mathcal{P}$, then the measure μ is T-invariant.*

Now, we can formally define invariant measure.

Here is an example of invariant measure.

Example 1.1.17. *Let $I = [0, 1]$, $\mathfrak{B}$ be Borel σ-algebra of I, m be the Lebesgue measure on I. Consider the measurable transformation $T : I \to I$, which is given by $T(x) = 3x$ (mod 1).*
Let $[a, b] \subset I$, its pre-image is three disjoint intervals I_1, I_2, I_3, with $m(I_1) = m(I_2) = m(I_3) = \frac{1}{3}(b-a)$. Thus, $m(T^{-1}([a, b])) = m(I_1 \cup I_2 \cup I_3) = 3 \cdot \frac{1}{3}(b-a) = b-a = m([a, b])$. Since $[a, b]$ is any subinterval in I, by Theorem 1.1.16, Lebesgue measure m is T-invariant.

Definition 1.1.33 (**Ergodic**). *Let $T : X \to X$ and μ be T-invariant. Then T is ergodic if for any $B \in \mathfrak{B}$, such that $T^{-1}B = B$, $\mu(B) = 0$ or $\mu(X \setminus B) = 0$.*

Theorem 1.1.18 (**Birkhoff Ergodic Theorem**). *[15] Let $T : X \to X$ and μ be T-invariant ergodic probability measure. Let $f \in L^1(\mu)$. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i(x)) = \int_X f d\mu$$

for μ-a.e. $x \in X$.

Definition 1.1.34 (**Mixing**). *Let $T : X \to X$ and μ be T-invariant probability measure. Then T is weakly mixing if for all $A, B \in \mathfrak{B}$,*

$$\frac{1}{n} \sum_{i=0}^{n-1} |\mu(T^{-i}A \cap B) - \mu(A)\mu(B)| \to 0 \text{ as } n \to +\infty.$$

T is strongly mixing if for all $A, B \in \mathfrak{B}$,

$$\mu(T^{-n}A \cap B) \to \mu(A)\mu(B) \text{ as } n \to +\infty.$$

Remark 1.1.7. *Both strongly mixing and weakly mixing are stronger properties than ergodicity, strong mixing means asymptotic independence. It can be shown that:*

$$\textit{Strongly Mixing} \Rightarrow \textit{Weakly Mixing} \Rightarrow \textit{Ergodic}.$$

See [53].

1.1.5 The Frobenius-Perron Operator

The Frobenius-Perron operator is a very useful tool to study existence and properties of absolutely continuous invariant measures. Below I will explain in details the definition and properties of the Frobenius-Perron Operator.
First of all, we recall the definition of a nonsingular transformation.

Definition 1.1.35 (**Nonsingular Transformation**). *Let $(X, \mathfrak{B}, \mu)$ be a measure space. A transformation $T : X \to X$ is measurable. T is nonsingular if $\mu(T^{-1}(A)) = 0$ whenever $\mu(A) = 0$, $A \in \mathfrak{B}$.*

Now, assume $T : I \to I$ is a nonsingular transformation with respect to Lebesgue measure m. Let Y be a random variable on I, and it has probability density function f. We apply T to every point in the space, then for any set $E \subset X$, the probability of points land at E is

$$\begin{aligned}\text{Prob}(T(Y) \in E) &= \int I_E(T(Y)) f dm = \int I_{T^{-1}E} f dm = \int I_{T^{-1}E} d\mu_f \\ &= \int I_E d\mu_f \circ T^{-1} = \int_E (\frac{d\mu_f \circ T^{-1}}{dm}) dm,\end{aligned}$$

where, I_E is a characteristic function.

$$d\mu_f := f dm.$$

Since T is non-singular, we have

$$\mu_f \circ T^{-1} \ll m.$$

Therefore, by the Radon-Nikodym Theorem, there exists a unique $P_T f := \frac{d\mu_f \circ T^{-1}}{dm}$ in $L^1(m)$. We called P_T the Frobenius-Perron operator associated with T.

The following properties of P_T are well known. Proofs can be found in [15].

Proposition 1.1.19. *P_T is the unique element of $L^1(m)$, that for all test functions $\varphi \in L^\infty$, $\int \varphi \cdot (P_T f) dm = \int (\varphi \circ T) \cdot f dm$.*

Proposition 1.1.20. ***(Properties of Frobenius-Perron operator)***

1. *(Linearity) $P_T : L^1 \to L^1$ is a linear operator.*

2. *(Positivity) Let $f \in L^1$ with $f \geq 0$. Then $P_T f \geq 0$.*

3. *(Preservation of Integrals) For $f \in L^1$, $\int_I P_T f dm = \int_I f dm$.*

4. *(Contraction Property) $P_T : L^1 \to L^1$ is a contraction, i.e., $\|P_T f\|_1 \leq \|f\|_1$ for any $f \in L^1$.*

5. *(Composition Property) Let $T : I \to I$ and $S : I \to I$ be nonsingular. Then $P_{T \circ S} f = P_T \circ P_S f$.*

6. *Let $T : I \to I$ be nonsingular. $P_T f^* = f^*$ if and only if it is the density of a T-invariant measure μ, $\mu(A) = \int_A f^* dm$. μ is called absolutely continuous invariant measure.*

Example 1.1.21. *Let $T : I \to I$ be given by $T(x) = 2x \pmod 1$. Let $A = [a, b]$ be any interval.*

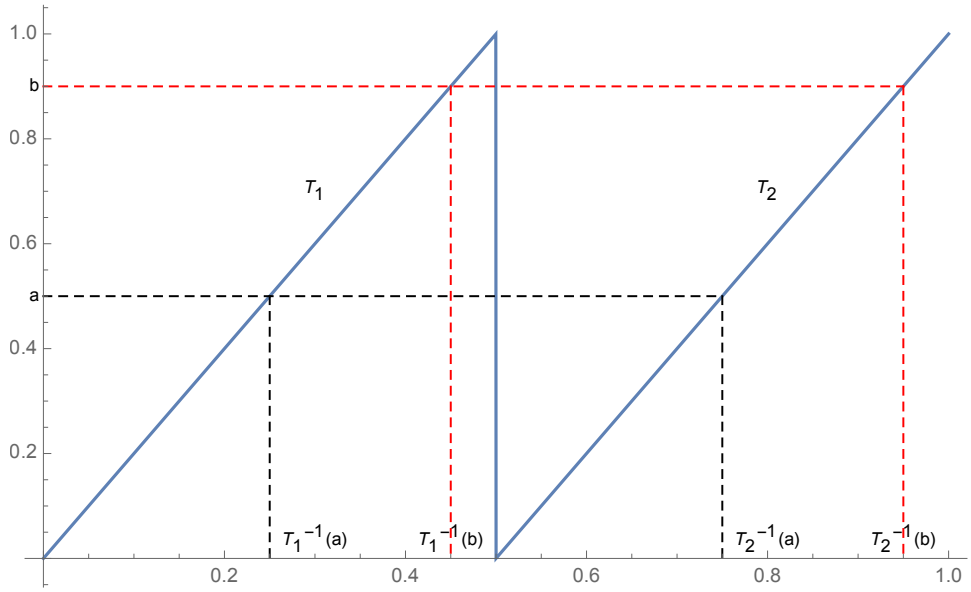

Figure 1.5: The Map T

We have

$$\int_{[a,b]} P_T f dm = \int_{T^{-1}([a,b])} f dm = \int_{(T^{-1}([a,b])\cap[0,\frac{1}{2}])\cup(T^{-1}([a,b])\cap[\frac{1}{2},1])} f dm.$$

We check that $f^* = 1$ is a fixed point of P_T. We have

$$\int_{[a,b]} f^* dm = \int_{[a,b]} 1 dm = m([a, b]).$$

On the other hand,

$$\begin{aligned}\int_{(T^{-1}([a,b])\cap[0,\frac{1}{2}])\cup(T^{-1}([a,b])\cap[\frac{1}{2},1])} f^* dm &= \int_{(T^{-1}([a,b])\cap[0,\frac{1}{2}])\cup(T^{-1}([a,b])\cap[\frac{1}{2},1])} 1 dm \\ &= (m(T^{-1}([a, b]) \cap [0, \frac{1}{2}]) + m(T^{-1}([a, b]) \cap [\frac{1}{2}, 1])) \\ &= (\frac{1}{2}m([a, b]) + \frac{1}{2}m([a, b])) = m([a, b]).\end{aligned}$$

Therefore, we have $\int_{[a,b]} f^* dm = \int_{[a,b]} P_T f^* dm$. Since this is true over any interval $[a, b]$, we get $P_T f^* = f^*$ a.e. Thus, we have used P_T to show that m is T-invariant.

Representation of the Frobenius-Perron operator If our system is piecewise monotonic, there is a powerful representation of the Frobenius-Perron operator.

Definition 1.1.36 (Piecewise Monotonic interval map). *Let $I = [0, 1]$. A transformation $T : I \to I$ is called piecewise monotonic if there exists a partition of $I, 0 = a_0 < a_1 < ... < a_q = 1$, and a number $r > 1$ such that*

1. *$T_{|(a_{i-1},a_i)}$ is a differentiable, $i = 1, ..., q$ which can be extended to a differentiable map on $[a_{i-1}, a_i]$, $i = 1, ..., q$.*

2. *$|T'(x)| > 0$ on (a_{i-1}, a_i), $i = 1, ..., q$.*

For piecewise monotonic map $T : I \to I$, the Frobenius-Perron operator P_T associated with T has a piecewise representation, i.e. for a.e x.

$$P_T f(x) = \sum_{i=1}^{q} \frac{f(T_i^{-1}(x))}{|T'(T_i^{-1}(x))|} \chi_{T(a_{i-1},a_i)}(x).$$

Also, it can be writen as

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

Here I will give an example of a piecewise monotonic map and its Frobenius-Perron operator.

Example 1.1.22. *Let $I = [0, 1]$ and $T : I \to I$ be defined as*

$$T(x) = \begin{cases} \frac{5x}{3-x} & \text{if } x \in [0, \frac{1}{2}), \\ 2x - 1 & \text{if } x \in [\frac{1}{2}, 1]. \end{cases}$$

As shown in Figure 3.3.1,

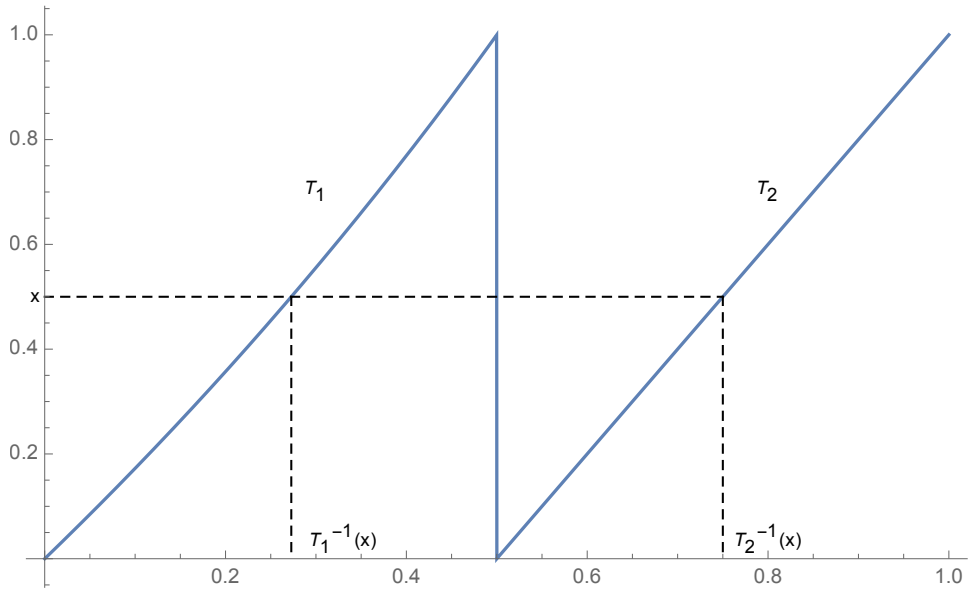

Figure 1.6: The Map T

In our example the map T is piecewise monotonic. Then

$$P_T f(x) = \sum_{y \in T^{-1}(x)} \frac{f(y)}{|T'(y)|} = \frac{f(T_1^{-1}(x))}{|T_1'(T_1^{-1}(x))|} + \frac{f(T_2^{-1}(x))}{|T_2'(T_2^{-1}(x))|} \tag{1.1.10}$$

where

$$T_1^{-1}(x) = \frac{3x}{5+x}, T_2^{-1}(x) = \frac{x+1}{2},$$

and

$$T_1'(x) = \frac{5}{3-x} + \frac{5x}{(3-x)^2}, T_2'(x) = 2.$$

From (1.1.10) *we get*

$$P_T f(x) = \frac{15 f(\frac{3x}{5+x})}{(5+x)^2} + \frac{f(\frac{x+1}{2})}{x+1}.$$

1.1.6 Ulam's Method and Approximating Dynamical Quantities

Suppose we know that a dynamical system has an invariant density; i.e, its Frobenius-Perron operator P_T has a fixed point f^*. Often, it is very difficult to find f^* explicitly. Moreover, it is also difficult to determine analytically other dynamical quantities associated with f^*. To approximate such quantities, a well known technique is to approximate P_T by a matrix. This idea goes back to Ulam [52]. The main theme of this thesis is to use Ulam (or Ulam-like) scheme and implemented rigorously on a computer.

Ulam's scheme Let $\eta := \{I_k\}_{k=1}^{d(\eta)}$ be a partition of $[0,1]$ into intervals of size $m(I_k) \leq \eta$. Let $\mathfrak{B}_\eta$ be the σ-algebra generated by η and for $f \in L^1$ define the projection

$$\Pi_\eta f = E(f|\mathfrak{B}_\eta),$$

and

$$P_\eta = \Pi_\eta \circ P \circ \Pi_\eta.$$

P_η, which is called Ulam's approximation of P, is finite rank operator which can be represented by a (row) stochastic matrix acting on vectors in $\mathbb{R}^{d(\eta)}$ by left multiplication. Its entries are given by

$$P_{kj} = \frac{m(I_k \cap T^{-1}(I_j))}{m(I_k)}.$$

Example 1.1.23. *In this example, I will show in details how the Ulam scheme works. Let $T : I \to I$ be given by $T(x) = 2x \pmod 1$.*

$$P_T f(x) = \sum_{y \in T^{-1}(x)} \frac{f(y)}{|T'(y)|} = \frac{f(T_1^{-1}(x))}{|T_1'(T_1^{-1}(x))|} + \frac{f(T_2^{-1}(x))}{|T_2'(T_2^{-1}(x))|} \tag{1.1.11}$$

The Ulam approximation P_η can be represented by the following matrix:

$$P_\eta = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix}$$

As we will see in Chapters 2 and 3, we will be able to compute rigorously important dynamical quantities, namely diffusion coefficients (Chapter 2) and for the first time linear response (Chapter 3).

1.2 Numerical Background

In this section, I will introduce definitions and operations that I used in the computer codes to produce rigorous approximations of dynamical quantities. For more information on computer arithmetic, we refer to [51].

In our work we aim to provide rigorous approximation of dynamical quantities. By rigorous we mean the following:

Suppose we want to compute f^*, the invariant density of a certain dynamical systems, up to a pre-specified error τ in a certain topology. How do we rigorously achieve such an error? There are two important ingredients to doing this:

1. We should be able to track all the constants that come from the theoretical part, for example from the Keller-Liverani Theorem 1.1.14, and compute them[2].

2. We should be able to track all the roundoff errors made by the computer throughout the process.

This section provides a background on roundoff errors.

1.2.1 Computer Arithmetic

Positional system

The differences between usual arithmetic and the computer one begin from the positional system. Instead of 10, the computer uses positional system in base 2. But we can present any real number in a positional system with an arbitrary integer base $\beta \geq 2$ as follows:

$$(-1)^{\gamma}(b_n b_{n-1}...b_0 b_{-1} b_{-2}...)_{\beta}, \tag{1.2.1}$$

where $b_n b_{n-1}, ...$ are integers in the range $[0, \beta - 1]$, and $\gamma \in \{0, 1\}$ provides the sign of the number. The real number corresponding to (1.2.1) is

$$x = (-1)^{\gamma} \sum_{i=-\infty}^{n} b_i \beta^i = (-1)^{\gamma}(b_n \beta^n + b_{n-1}\beta^{n-1} + \cdots b_0 \beta^0 + b_{-1}\beta^{-1} \cdots).$$

Floating point numbers

The floating point numbers system provides a more convenient way to present real number than (1.2.1). We define the set of floating point numbers in base β as:

$$\mathbb{F}_{\beta} = \{(-1)^{\sigma} m \times \beta^e : m = (b_0.b_1 b_2....)_{\beta}\},$$

where, as before, we request that β is an integer not less than 2, and that $0 \leq b_i \leq \beta - 1$ for all i, and $0 \leq b_i \leq \beta - 2$, $i = n, n+1....$ The exponent e may be any integer. But this

[2]This point will become clear in Chapters 2, 3 when we deal will specific systems.

set is uncountably infinite, computer can only store finite digits, we need to build a finite set to approximate real numbers. The first step, define the set

$$\mathbb{F}_{\beta,p} = \{x \in \mathbb{F}_\beta : m = ((b_0.b_1b_2....b_{p-1})_\beta\},$$

where p is called the precision of the floating point system. Then, in order to form a finite set, we need to specify four integers: the base β, the precision p and the minimal and maximal exponents $\check{e}$ and $\hat{e}$. Define the parameterised sets of computer representable floating point numbers,

$$\mathbb{F}_{\beta,p}^{\check{e},\hat{e}} = \{x \in \mathbb{F}_{\beta,p} : \check{e} < e < \hat{e}\}.$$

Rounding

As computer can only store the number set $\mathbb{F}$ instead of the real number set $\mathbb{R}$, we have to find a number in set $\mathbb{F}$ as the approximation of the number in $\mathbb{R}$. We call this approximation Rounding. We usually have four ways to round numbers: round to zero, rounded up, rounded down and round to nearest.

Round to zero Round to zero, also known as "truncation", this name reflects the idea that discard the significand digits beyond position $p-1$. The formal definition given by an operator $\square_z : \mathbb{R}^* \to \mathbb{F}^*$

$$\square_z(x) = sign(x) \max\{y \in \mathbb{F}^* : y \leq |x|\},$$

where $sign(x)$ is the sign of x. This is the simplest way but not the most accurate way to round and we do not use it often.

Directed rounding There are two rounding modes called directed, round toward minus infinity (also known as "round down") and round toward plus infinity (also known as"round up"). They are denoted as $\bigtriangledown(x)$ and $\bigtriangleup(x)$ respectively, and defined as

$$\bigtriangledown(x) = \max\{y \in \mathbb{F}^* : y \leq x\} \text{ and } \bigtriangleup(x) = \min\{y \in \mathbb{F}^* : y \geq x\}.$$

Round to nearest For the previous two rounding modes, round down and round up, the maximum of its error is the length of the interval $[\bigtriangledown(x), \bigtriangleup(x)]$. Round to nearest, denote as $\square$, is a more accurate rounding mode, it makes the error down to $\frac{1}{2}[\bigtriangledown(x), \bigtriangleup(x)]$. It is defined as:

$$x > 0 \Rightarrow \square_n(x) = \begin{cases} \bigtriangledown(x), & \text{if } x \in [\bigtriangledown(x), \frac{1}{2}(\bigtriangledown(x) + \bigtriangleup(x))), \\ \bigtriangleup(x), & \text{if } x \in [\frac{1}{2}(\bigtriangledown(x) + \bigtriangleup(x)), \bigtriangleup(x)], \end{cases}$$
$$x < 0 \Rightarrow \square_n(x) = -\square_n(-x)$$

1.2.2 Interval Arithmetic

When we implement a function $f(x), x \in \mathbb{R}$ in a computer, the rounding error appears since we can only store set $\mathbb{F}$ in computer to approximate set $\mathbb{R}$. To rigorously estimate

the rounding error, we can use interval arithmetic. We can get an interval $F([x])$ that contains the exact value of $f(x)$, and the length of interval $F([x])$ is the bound of rounding error. The general idea of interval arithmetic is apply F to a set of intervals $([x])$, instead of the classical way of mapping a number x to number $f(x)$. To accomplish that, we need to extend real functions to interval functions.

Let us start with some notation that will be used often in this section.

Let $A \subseteq \mathbb{R}$ and $f : A \to \mathbb{R}$. Denote the range of f over A by

$$R(f, A) = \{f(x) : x \in A\}.$$

For $a \in \mathbb{R}$, the interval $[a]$ is:

$$[a] = [\underline{a}, \overline{a}] = \{a \in \mathbb{R} : \underline{a} \leq a \leq \overline{a}\}.$$

The set of all intervals $[a]$ of real line:

$$\mathcal{R} = \{[\underline{a}, \overline{a}] : \underline{a} \leq \overline{a}; \underline{a}, \overline{a} \in \mathbb{R}\}.$$

The set relations of elements in $\mathcal{R}$ define as:

$$[a] = [b] \Leftrightarrow \underline{a} = \underline{b} \text{ and } \overline{a} = \overline{b},$$
$$[a] \subseteq [b] \Leftrightarrow \underline{b} \leq \underline{a} \text{ and } \overline{a} \leq \overline{b},$$
$$[a] \subset [b] \Leftrightarrow [a] \subseteq [b] \text{ and } [a] \neq [b],$$
$$[a] \mathring{\subset} [b] \Leftrightarrow \underline{b} < \underline{a} \text{ and } \overline{a} < \overline{b},$$
$$[a] \leq [b] \Leftrightarrow \underline{a} \leq \underline{b} \text{ and } \overline{a} \leq \overline{b},$$
$$a \in [b] \Leftrightarrow \underline{b} < a \text{ and } a < \overline{b}.$$

To get one and only one new interval from unions, we define an operation called hull:

$$[a] \sqcup [b] = [\min\{\underline{a}, \underline{b}\}, \max\{\overline{a}, \overline{b}\}].$$

Define the intersection operation as:

$$[a] \cap [b] = \begin{cases} [\emptyset] & \text{if } \overline{a} < \underline{b} \text{ or } \overline{b} < \underline{a} \ , \\ [\max\{\underline{a}, \underline{b}\}, \min\{\overline{a}, \overline{b}\}] & \text{otherwise} \end{cases}$$

For $[a] \in \mathcal{R}$, we define the following real-valued functions, the middle point and the radius function of interval $[a]$:

$$Rad([a]) = \frac{1}{2}(\overline{a} - \underline{a}) \text{ (the Radius of } [a]),$$
$$Mid([a]) = \frac{1}{2}(\overline{a} + \underline{a}) \text{ (the Midpoint of } [a]).$$

The mignitude and magnitude function of interval $[a]$:

$$Mig([a]) = \min\{|a| : a \in [a]\} \text{ (the Mignitude of } [a]),$$
$$Mag([a]) = \max\{|a| : a \in [a]\} \text{ (the Magnitude of } [a]).$$

The absolute value of interval $[a]$:

$$Abs([a]) = [Mig([a]), Mag([a])].$$

Furthermore, we can define arithmetic on the elements of $\mathcal{R}$ by:

$$[a] + [b] = [\underline{a} + \underline{b}, \overline{a} + \overline{b}],$$
$$[a] - [b] = [\underline{a} - \overline{b}, \overline{a} - \underline{b}],$$
$$[a] \times [b] = [\min\{\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{ab}\}, \max\{\underline{ab}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{ab}\}],$$

$$[a] \div [b] = [a] \times \left[\frac{1}{\overline{b}}, \frac{1}{\underline{b}}\right], \text{if } 0 \notin [b].$$

Example 1.2.1. *The example will show how computer arithmetic works on some particular intervals.*

$$[-2,1] + [3,5] = [-2+3, 1+5] = [1,6]$$
$$[0,3] - [-3,4] = [0-4, 3-(-3)] = [-4,6]$$
$$[1,\pi] \times [\sqrt{5},5] = [\min\{\sqrt{5},5,\pi\sqrt{5},5\pi\}, \max\{\sqrt{5},5,\pi\sqrt{5},5\pi\}] = [\sqrt{5},5\pi]$$
$$[-\sqrt{2},2] \div [e,4] = [-\sqrt{2},2] \times [\frac{1}{4},\frac{1}{e}]$$
$$= [\min\{-\sqrt{2}\frac{1}{4}, -\sqrt{2}\frac{1}{e}, 2\frac{1}{4}, 2\frac{1}{e}\}, \max\{-\sqrt{2}\frac{1}{4}, -\sqrt{2}\frac{1}{e}, 2\frac{1}{4}, 2\frac{1}{e}\}] = [-\frac{\sqrt{2}}{4}, \frac{2}{e}]$$

Additionally, the addition and multiplication are both associative and commutative: for $[a], [b], [c] \in \mathcal{R}$,

$$[a] + ([b] + [c]) = ([a] + [b]) + [c];\ [a] + [b] = [b] + [a],$$
$$[a] \times ([b] \times [c]) = ([a] \times [b]) \times [c];\ [a] \times [b] = [b] \times [a].$$

There are two important properties of interval arithmetic: sub-distributivity and inclusion isotonicity.

The sub-distributivity is a weaker rule than distributive law, for $[a], [b], [c] \in \mathcal{R}$, it states:

$$[a]([b] + [c]) \subseteq [a][b] + [a][c].$$

The inclusion isotonicity is presented in the following theorem:

Theorem 1.2.2. *If* $[a] \subseteq [a']$, $[b] \subseteq [b']$, *and* $\star \in \{+, -, \times, \div\}$, *then*

$$[a] \star [b] \subseteq [a'] \star [b'],$$

where $0 \notin [b']$ *for division.*

Now, we are ready to extend real functions to interval functions. In interval analysis, we have a theorem to extend the real function to interval functions. The way we extend is to substitute all real numbers x with intervals $[x]$ to get an interval function $f([x])$, called the natural interval extension of f. The complete statement of the theorem is following:

Theorem 1.2.3. *Given a real-valued, rational function* f*, and its natural interval-extension* F *such that* $F([x])$ *is well-defined for some* $[x] \in \mathbb{R}$*. We have*

1. $[z] \subseteq [z'] \Rightarrow F([z]) \subseteq F([z'])$, *(Inclusion Isotonicity)*
2. $R(f, [x]) \subseteq F([x])$. *(Range Enclosure)*

This theorem can be extended to elementary functions. Elementary functions will be defined below. Let us recall the set of standard functions $\mathcal{S}$:

$$\mathcal{S} = \{a^x, \log_a x, x^{p/q}, |x|, \sin(x), \cos(x), \tan(x)...\}.$$

Definition 1.2.1. *Any real-valued function expressed as a finite number of standard functions combined with constants, arithmetic operations and compositions is called an elementary function. We denote the class of elementary function by $\mathcal{E}$.*

Then, we can extend the above theorem to elementary function, it is called the fundamental theorem of interval analysis.

Theorem 1.2.4. *(The Fundamental theorem of Interval Analysis) Given a real-valued, elementary function f, and its natural interval-extension F such that $F([x])$ is well-defined for some $[x] \in \mathbb{R}$. We have*

1. $[z] \subseteq [z'] \Rightarrow F([z]) \subseteq F([z'])$, *(Inclusion Isotonicity)*
2. $R(f,[x]) \subseteq F([x])$. *(Range Enclosure)*

1.2.3 Automatic Differentiation

When we use the Newton's method to find a root, we need to compute each $f'(x_i)$, where $i = 1, 2, 3, ..., n$ is the iteration time. We usually have two ways to compute $f'(x_i)$. One is the formula:

$$f'(x_i) = \lim_{h \to 0} \frac{f(x_i + h) - f(x_i)}{h}.$$

The other way is to deduce the exact formula of f', then substitute x_i to the formula f'. However, both ways are proved not effective in computation. The first one has poor precision and for the second one it is too hard to find a formula for f' for some function f. Also, they are memory-consuming and time-consuming.

In this section, a technique called the differentiation arithmetic will be presented. The implementation of differentiation arithmetic only involve the computation of elementary functions, like sin, cos, exp, log, and elementary arithmetic, like addition, subtraction, multiplication, division. The key of this technique is to repeatedly apply chain rule, then we will get a precise differentiation with efficient computation. I will state the automatic differentiation in detail and give an example that makes it more clear.

The first order case

In the first order case, we only consider the first derivative of $f(x)$. To compute $f'(x_0)$, we should introduce some notation.

For any function $u : \mathbb{R} \to \mathbb{R}$, let $u_0 = u(x_0)$ and $u'_0 = u'(x_0)$. Define the pair of real numbers as:

$$\vec{u} = (u_0, u'_0).$$

Here is some pairs for standard functions:

$$\begin{aligned}
\sin(\vec{u}) &= \sin(u_0, u_0') = (\sin(u_0), u_0' \cos(u_0)),\\
\cos(\vec{u}) &= \cos(u_0, u_0') = (\cos(u_0), -u_0' \sin(u_0)),\\
e^{(\vec{u})} &= e^{(u_0, u_0')} = (e^{(u_0)}, u_0' e^{(u_0)}),\\
\log(\vec{u}) &= \log(u_0, u_0') = (\log(u_0), \frac{u_0'}{u_0}),\ \text{for } (u_0 > 0),\\
\vec{u}^{\alpha} &= (u_0, u_0')^{\alpha} = (u_0^{\alpha}, u_0' \alpha u_0^{\alpha-1}),\ \text{for } (\alpha \neq 0),\\
|\vec{u}| &= |(u_0, u_0')| = (|u_0|, u_0' sign(u_0)),\ \text{for } (u_0 \neq 0).
\end{aligned} \tag{1.2.2}$$

Define the arithmetic rules by:

$$\begin{aligned}
\vec{u} + \vec{v} &= (u_0 + v_0, u_0' + v_0'),\\
\vec{u} - \vec{v} &= (u_0 - v_0, u_0' - v_0'),\\
\vec{u} \times \vec{v} &= (u_0 v_0, v_0 u_0' + u_0 v_0'),\\
\vec{u} \div \vec{v} &= (\frac{u_0}{v_0}, \frac{(u_0' - (\frac{u_0}{v_0}) v_0')}{v_0}).
\end{aligned} \tag{1.2.3}$$

where $v_0 \neq 0$ for the last rule.

Example 1.2.5. *For function $f(x) = \frac{(x+3)(x-5)}{x-8}$, compute the value of $f(x_0)$ and $f'(x_0)$ when $x_0 = 4$.*

For a variable x, the pair $\vec{x}$ is:

$$\vec{x} = (x, 1),$$

and for constants c, where $c = 3, 5, 8$ in this example. The pair $\vec{c}$ is

$$\vec{c} = (c, 0).$$

Applying the arithmetic rules we just defined above:

$$\vec{f}(\vec{x}) = \frac{(\vec{x} + \vec{3})(\vec{x} - \vec{5})}{\vec{x} - \vec{8}} = \frac{((x,1) + (3,0)) \times ((x,1) - (5,0))}{(x,1) - (8,0)}.$$

When $x_0 = 4$, $\vec{x} = (4, 1)$.

$$\begin{aligned}
\vec{f}(4,1) &= \frac{((4,1) + (3,0)) \times ((4,1) - (5,0))}{(4,1) - (8,0)} = \frac{(7,1) \times (-1,1)}{(-4,1)}\\
&= \frac{(-7,6)}{(-4,1)} = (\frac{7}{4}, -\frac{17}{16}).
\end{aligned}$$

Then, the first element of $\vec{f}(4,1)$ is the value of $f(4) = \frac{7}{4}$, the second element of $\vec{f}(4,1)$ is the value of $f'(4) = -\frac{17}{16}$.

Let us check the result by using usual differentiation:

$$f'(x) = 1 + \frac{6}{x-8} - \frac{6x-15}{(x-8)^2}.$$

Substitute $x_0 = 4$ to the formulae $f(x)$ and $f'(x)$.

$$f(4) = \frac{(4+3)(4-5)}{4-8} = \frac{7}{4},$$

and

$$f'(4) = 1 + \frac{6}{4-8} - \frac{6 \cdot 4 - 15}{(4-8)^2} = -\frac{17}{16}.$$

Result checked.

When we use interval arithmetic to control the rounding error, we can use automatic

differentiation to intervals as well. Define the pair with intervals by:

$$\vec{u} = ([u_0], [u_0']).$$

The result will be two intervals. They contain the exact value of $f(x_0)$ and $f'(x_0)$ respectively, and the length of these two intervals are the rounding error of the computation of value of $f(x_0)$ and $f'(x_0)$.

Let us see an example that shows how automatic differentiation works for intervals.

Example 1.2.6. *Let $f(x) = e^x + \sin(x) + x + 2$, find the enclosure of $f'([0, \frac{\pi}{2}])$.*
Define the extend function:

$$\vec{f}([\vec{x}]) = e^{[\vec{x}]} + \sin[\vec{x}] + [\vec{x}] + \vec{2},$$

the interval pair $[\vec{x}] = ([0, \frac{\pi}{2}], [1])$.

$$\begin{aligned}
\vec{f}([\vec{x}]) &= e^{[\vec{x}]} + \sin[\vec{x}] + [\vec{x}] + \vec{2} \\
&= e^{([0,\frac{\pi}{2}],1)} + \sin([0, \frac{\pi}{2}], [1]) + ([0, \frac{\pi}{2}], [1]) + ([2], [0]) \\
&= (e^{[0,\frac{\pi}{2}]}, e^{[0,\frac{\pi}{2}]}) + (\sin([0, \frac{\pi}{2}]), \cos([0, \frac{\pi}{2}])) + ([0, \frac{\pi}{2}], [1])) + ([2], [0]) \\
&= ([1, e^{\frac{\pi}{2}}], [1, e^{\frac{\pi}{2}}]) + ([0, 1], \cos([0, \frac{\pi}{2}])) + ([0, \frac{\pi}{2}], [1]) + ([2], [0]) \\
&= ([1, e^{\frac{\pi}{2}}], [1, e^{\frac{\pi}{2}}]) + ([0, 1 + \frac{\pi}{2}], \cos([0, \frac{\pi}{2}]) + [1]) + ([2], [0]) \\
&= ([3, e^{\frac{\pi}{2}} + 3 + \frac{\pi}{2}], [1, e^{\frac{\pi}{2}}] + \cos([0, \frac{\pi}{2}]) + [1]) \\
&= ([3, e^{\frac{\pi}{2}} + 3 + \frac{\pi}{2}], [3, 1 + e^{\frac{\pi}{2}}]).
\end{aligned}$$

Thus, $f([0, \frac{\pi}{2}]) = [3, e^{\frac{\pi}{2}} + 3 + \frac{\pi}{2}]$ and $f'([0, \frac{\pi}{2}]) = [3, 1 + e^{\frac{\pi}{2}}]$.

The higher order case

When we want to compute the 2nd, the 3rd or even 10th derivative of a function $f(x)$, the method we have implemented in the first order case still works. For example, to compute the second derivative of $f(x)$, we define triples instead of pairs as: $\vec{u} = (u, u', u'')$ and its arithmetic rules should define as following:

$$\begin{aligned}
\vec{u} + \vec{v} &= (u + v, u' + v', u'' + v'') \\
\vec{u} - \vec{v} &= (u - v, u' - v', u'' - v'') \\
\vec{u} \times \vec{v} &= (uv, vu' + uv', vu'' + 2u'v' + uv'') \\
\vec{u} \div \vec{v} &= (u/v, (u' - (u/v)v')/v, (u'' - 2(u/v)'v' - (u/v)v'')/v).
\end{aligned}$$

However, I will not go further with the triples method, it is possible to compute, but very tedious and needed patience. Furthermore, what if we need 5th derivative or even 10th? That seems impossible to define such complicated arithmetic rules, and then apply it. Fortunately, we have a more effective way to compute higher order derivative. The new method will use the Taylor series. For a real-valued function $f \in C^\infty$, we study the k-th

derivative of $f(x)$ at $x = x_0$. Firstly, we can write Taylor expansion of $f(x)$ at $x = x_0$ as:

$$f(x) = f_0 + f_1(x - x_0) + ... + f_k(x - x_0)^k + ...$$

where $f_k = f_k(x_0) = \frac{f^{(k)}(x_0)}{k!}$ and $f^{(k)} = \frac{d^k f}{dx^k}$.

From the Taylor expansion, we can see the derivative of $f(x)$ and the Taylor coefficient are related. The k-th derivative of f at x_0 is equal to the Taylor coefficient f_k times k factorial, e.g. $f^{(k)}(x_0) = k!f_k$. For a function $f(x)$, as long as we know the Taylor expansion of it at x_0, we will know the value of any order of its derivative at x_0.

Now, rather than studying the derivative of $f(x)$ at x_0 directly, we turn to study the Taylor expansion of $f(x)$ at x_0, especially the coefficient of its Taylor expansion. However, for some functions, computing their Taylor expansion and coefficient directly is time-consuming as well. Note that functions that consist of standard functions, and studying the Taylor expansion of standard functions is much easier.

For two functions f and g, let f_k and g_k be the Taylor coefficients of functions f and g respectively. But how f_k and g_k are related to the Taylor coefficients of functions $(f \star g)_k$? where $\star$ means the operation $+, -, \times, \div$.

The rules of the Taylor arithmetic are derived in details in [51], I only state the result here:

$$(f + g)_k = f_k + g_k \tag{1.2.4}$$

$$(f - g)_k = f_k - g_k \tag{1.2.5}$$

$$(f \times g)_k = \sum_{i=0}^{k} f_i g_{k-i} \tag{1.2.6}$$

$$(f \div g)_k = \frac{1}{g_0}(f_k - \sum_{i=0}^{k-1}(f \div g)_i g_{k-i}). \tag{1.2.7}$$

Remark 1.2.1. *The following is how we do the Taylor expansion of constants c and independent variable x.*

$$x = x_0 + 1 \cdot (x - x_0) + 0 \cdot (x - x_0)^2 + ... + 0 \cdot (x - x_0)^k + ...,$$
$$c = c + 0 \cdot (x - x_0) + 0 \cdot (x - x_0)^2 + ... + 0 \cdot (x - x_0)^k + ...$$

Now, we are prepared to consider the derivative of standard function $e^{g(x)}$, where the Taylor series of function g is known. The first derivative of $e^{g(x)}$ can be written as:

$$\frac{d}{dx}e^{g(x)} = g'(x)e^{g(x)}. \tag{1.2.8}$$

The Taylor series of $g(x)$, $e^{g(x)}$ and $\frac{d}{dx}e^{g(x)}$ expand as following:

$$g(x) = \sum_{k=0}^{\infty} g_k(x - x_0)^k, \tag{1.2.9}$$

$$g'(x) = \sum_{k=1}^{\infty} k g_k(x - x_0)^{k-1}, \tag{1.2.10}$$

$$e^{g(x)} = \sum_{k=0}^{\infty} (e^g)_k (x - x_0)^k, \tag{1.2.11}$$

$$\frac{d}{dx} e^{g(x)} = \sum_{k=1}^{\infty} k(e^g)_k (x - x_0)^{k-1}. \tag{1.2.12}$$

Substitue the Taylor expansion of $g'(x)$, $e^{g(x)}$ and $\frac{d}{dx}e^{g(x)}$ into (1.2.8), we have

$$\sum_{k=1}^{\infty} k(e^g)_k (x - x_0)^{k-1} = \sum_{k=1}^{\infty} k g_k (x - x_0)^{k-1} \sum_{k=0}^{\infty} (e^g)_k (x - x_0)^k$$

$$\sum_{k=1}^{\infty} k(e^g)_k (x - x_0)^{k} = \sum_{k=1}^{\infty} k g_k (x - x_0)^{k} \sum_{k=0}^{\infty} (e^g)_k (x - x_0)^k$$

Right side, change the Taylor series back to functions,

$$\sum_{k=1}^{\infty} k(e^g)_k (x - x_0)^k = kg \times e^g$$

Taylor expanding to the right side,

$$\sum_{k=1}^{\infty} k(e^g)_k (x - x_0)^k = \sum_{k=1}^{\infty} (kg \times e^g)_k (x - x_0)^k.$$

Using the Taylor arithmetic to $(kg \times e^g)_k$, we have

$$k(e^g)_k = (kg \times e^g)_k = \sum_{i=1}^{k} i g_i (e^g)_{k-i}.$$

Then, we have the Taylor coefficients of $e^{g(x)}$,

$$(e^g)_k = \begin{cases} e^{g_0} & \text{if } k = 0, \\ \frac{1}{k} \sum_{i=1}^{k} i g_i (e^g)_{k-i} & \text{if } k > 0. \end{cases} \tag{1.2.13}$$

Let us see how this method works for a specific function:

Example 1.2.7. *Let $f(x) = e^{x^2+x+1}$, compute the third derivative of $f(x)$ at $x_0 = 1$. For $g(x) = x^2 + x + 1$, its Taylor expansion at x_0 is:*

$$g(x) = 3 + 3(x - 1) + (x - 1)^2,$$

The Taylor expansion of $e^{g(x)}$ at x_0,

$$e^{g(x)} = e^3 + 3e^3(x - 1) + \frac{11}{2}e^3(x - 1)^2 + ...$$

By (1.2.13),

$$(e^{g(x)})_k = \frac{1}{3} \sum_{i=1}^{3} i g_i (e^g)_{3-i} = \frac{1}{3}(g_1(e^g)_2 + 2g_2(e^g)_1 + 3g_3(e^g)_0)$$
$$= \frac{1}{3}(3 \cdot \frac{11}{2}e^3 + 2 \cdot 1 \cdot 3e^3 + 0 \cdot e^3) = \frac{15}{2}e^3.$$

Then, by definition of $(e^{g(x)})_k$,

$$f'''(x_0) = (e^{g(x)})_k \cdot k! = 45e^3.$$

Check the result by substituting x_0 into the third derivative function of $f(x)$,

$$f'''(x) = 6(2x + 1)e^{x^2+x+1} + (2x + 1)^3 e^{x^2+x+1},$$

then, $f'''(1) = 18e^3 + 27e^3 = 45e^3$, which is the same result as the Taylor method.

Using the same steps to exponential functions, Tucker derived other standard functions in [51], I only state his result here.

$$(\ln(g))_k = \begin{cases} \ln(g) & \text{if } k = 0, \\ \frac{1}{g_0}(g_k - \frac{1}{k}\sum_{i=1}^{k-1} i \ln(g)_i g_{k-i} & \text{if } k > 0. \end{cases}$$

$$(g^a)_k = \begin{cases} g_0^a & \text{if } k = 0, \\ \frac{1}{g_0}\sum_{i=1}^{k}(\frac{(a+1)i}{k} - 1)g_i(g^a)_{k-i} & \text{if } k > 0. \end{cases}$$

$$(\sin g)_k = \begin{cases} \sin g_0 & \text{if } k = 0, \\ \frac{1}{k}\sum_{i=1}^{k} i g_i(\cos g)_{k-i} & \text{if } k > 0. \end{cases}$$

$$(\cos g)_k = \begin{cases} \cos g_0 & \text{if } k = 0, \\ -\frac{1}{k}\sum_{i=1}^{k} i g_i(\sin g)_{k-i} & \text{if } k > 0. \end{cases}$$

$$(\tan g)_k = \begin{cases} \tan g_0 & \text{if } k = 0, \\ \frac{1}{\cos^2 g_0}(g_k - \frac{1}{k}\sum_{i=1}^{k-1} i(\tan g)_i(\cos^2 g)_{k-i}) & \text{if } k > 0. \end{cases}$$

$$(\arcsin g)_k = \begin{cases} \arcsin g_0 & \text{if } k = 0, \\ \frac{1}{\sqrt{1-(g_0)^2}}(g_k - \frac{1}{k}\sum_{i=1}^{k-1} i(\arcsin g)_i(\sqrt{1-g^2})_{k-i}) & \text{if } k > 0. \end{cases}$$

$$(\arccos g)_k = \begin{cases} \arccos g_0 & \text{if } k = 0, \\ \frac{-1}{\sqrt{1-(g_0)^2}}(g_k + \frac{1}{k}\sum_{i=1}^{k-1} i(\arccos g)_i(\sqrt{1-g^2})_{k-i}) & \text{if } k > 0. \end{cases}$$

$$(\arctan g)_k = \begin{cases} \arctan g_0 & \text{if } k = 0, \\ \frac{1}{1+(g_0)^2}(g_k - \frac{1}{k}\sum_{i=1}^{k-1} i(\arctan g)_i(1+g^2)_{k-i}) & \text{if } k > 0. \end{cases}$$

In order to show the full strength of the Taylor method, we will apply it to a more complicated example following.

Example 1.2.8. *Let $f(x) = (1+x^2)^{(3+x)}$. Compute the fourth derivative of $f(x)$ at $x_0 = 1$.*

The form of function $f(x)$ is not a standard function, it can not use any of the formulas above directly. However, we can write $f(x)$ as:

$$f(x) = e^{(3+x)\ln(1+x^2)}.$$

It is the form we can apply the derivative formula of standard functions. Let $g(x) = (3+x)\ln(1+x^2)$. We need the Taylor expansion of $g(x)$ and e^g at $x_0 = 1$:

$$g(x) = 4\ln(2) + (4+\ln(2))(x-1) + (x-1)^2 - \frac{2}{3}(x-1)^3 + \frac{1}{3}(x-1)^4 + ...$$

$$e^{g(x)} = 16 + (64 + 16\ln(2))(x-1) + (8\ln(2)^2 + 64\ln(2) + 144)(x-1)^2$$
$$+(224 + 144\ln(2) + 32\ln(2)^2 + \frac{8}{3}\ln(2)^3)(x-1)^3$$
$$+(\frac{808}{3} + 224\ln(2) + 72\ln(2)^2 + \frac{32}{3}\ln(2)^3 + \frac{2}{3}\ln(2)^4)(x-1)^4 + ...$$

Substitute the coefficient into formulae (1.2.13)*, we have*

$$\begin{aligned}(e^{g(x)})_4 &= \frac{1}{4}\sum_{i=1}^{4} i g_i(e^g)_{4-i} \\ &= (4+\ln(2))(224 + 144\ln(2) + 32\ln(2)^2 + \frac{8}{3}\ln(2)^3) + 16\ln(2)^2 + 96\ln(2) + \frac{544}{3} \\ &= 1851.588286.\end{aligned}$$

By the definition of f_k, we have

$$f^{(4)}(1) = (e^{g(x)})_4 \cdot 4! = 1851.588286 \times 4! = 44438.11886.$$

Remark 1.2.2. *The example above shows that for a general real function, we can change its form, then we can apply the formulae of derivative of standard function and the rules of Taylor arithmetic to compute its Taylor coefficient. The most important advantage of this method is its efficiency, even if we see many additions and multiplications of numbers in our example, the computer can deal with it accurately.*

In the numerical implement, we use fadbad library to implement automatic differentiation.

1.2.4 Newton's Method

In numerical analysis, Newton's method is used to find a successively better approximation to the roots of a real-valued function.

Let $f : \mathbb{R} \to \mathbb{R}$ be differentiable and $x \in \mathbb{R}$ such that $f(x) = 0$. For any x_0, we define

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Repeating this process, we have

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad n = 0, 1, 2, ..., m, m \in \mathbb{N} \tag{1.2.14}$$

Here is a graph to explain how Newton's Method work.

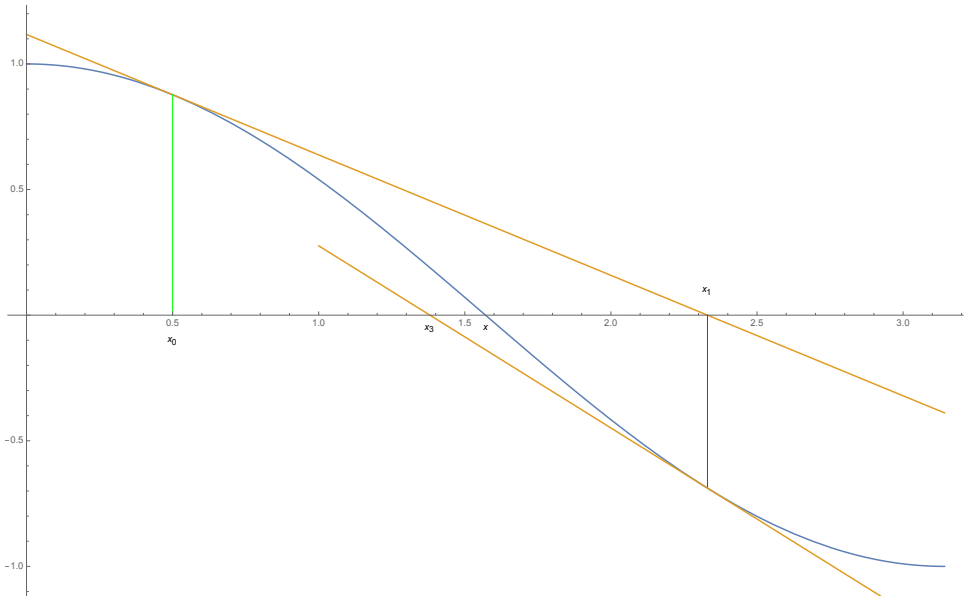

Figure 1.7: Find root of function $f(x) = \cos(x)$ by Newton's Method

Define x_{n+1} as the approximation of root of function f.

When we use the Newton's Method in our rigours computation, the most concern is the error. In [28], there is a theorem to estimate this error, here I state the theorem without proof:

Theorem 1.2.9. *Let $x \in \mathbb{R}$ such that $f(x) = 0$ and x is a simple root. Let $I_\epsilon = \{y \in \mathbb{R} : |x - y| \leq \epsilon\}$. Assume that $f \in C^2[I_\epsilon]$. Define*

$$M(\epsilon) = \max_{s,t \in I_\epsilon} \left| \frac{f''(s)}{2f'(t)} \right|.$$

If ϵ is so small that

$$2\epsilon M(\epsilon) < 1,$$

then for every $x_0 \in I_\epsilon$, Newton's Method is well defined and converges quadratically to the only root $x \in I_\epsilon$.

Can we extend Newton's method to intervals? In [51], Tucker proves a theorem of the Interval Newton Method that makes Newton's method work for intervals. I will explain his idea and theorem briefly here.

Let function $f : [x] \to \mathbb{R}$ be a differentiable function. Let x^* be root of f, $x^* \in [x]$, such that $f(x^*) = 0$. Let $F'([x])$ be the first derivative function of $f([x])$ and $0 \notin F'([x])$. By the Mean Value Theorem, for any $x \in [x]$,

$$f'(\xi) = \frac{f(x^*) - f(x)}{x - x^*}, \tag{1.2.15}$$

where ξ is some number in the interval with endpoints x and x^*.
We define a notation $N([x], x)$ as:

$$N([x], x) := x - \frac{f(x)}{F'([x])}.$$

As $f(x^*) = 0$ and $f'(\xi) \neq 0$, by (1.2.15)

$$x^* = x - \frac{f(x)}{f'(\xi)} \in N([x], x).$$

Recall that we assume $x \in [x]$, and $x \in N([x], x)$ as well. We have $x^* \in N([x], x) \cap [x]$, for any $x \in [x]$.
Define an operator $N([x])$ called the Interval Newton Operator as:

$$N([x]) := N([x], m) = m - \frac{f(m)}{F'([x])},$$

where $m = Mid([x])$. As $m \in [x]$, $x^* \in N([x]) \cap [x]$ as well.
Now, we define a sequence of intervals

$$[x_{k+1}] = N([x_k]) \cap [x_k], \quad k = 0, 1, 2, ...$$

We have defined all the notation that we need to present the theorem of the Interval Newton Method. Here is the formal statement:

Theorem 1.2.10. *Assume that $N[x_0]$ is well-defined. If $[x_0]$ contains a root x^* of f, then x^* is in all iterates of $[x_k]$, $k \in \mathbb{N}$. Furthermore, the intervals $[x_k]$ form a nested sequence converging to x^*.*

Furthermore, discussing the existence and the numbers of roots, Tucker proved another theorem about this question.

Theorem 1.2.11. *Let $f \in C^2([x], \mathbb{R})$, and assume that $N([x])$ is well-defined for some $[x] \in \mathbb{R}$. Then the following statements hold:*

```
/** This is a newton method we use to find roots for the unquotiented dynamic, identical to the newton method
    for the dynamic
 * of the class Newton_x
 */
Interval Newton_x_unquotiented::operator()(const dynamic_x & f,Interval domain, const Interval& value,const Real
    &precision)
{
    image=f.rest_branch(domain);
    if (!(proper_subset(value,image)))
    {
        return Interval();
    }
    else
    {
        // Declaration of variables

        // Start the loop
        while( width(domain)>precision )
        {
            midpoint=median(domain);

            //Automatic differentiation
            aut_diff_domain=domain;
            aut_diff_domain.diff(0,1);
            independent_variable=f.rest_branch(aut_diff_domain);
            derivative=independent_variable.d(0);

            //The new enclosure
            domain=intersect(domain,midpoint-(f.rest_branch(midpoint)-value)*multiplicative_inverse(derivative))
                ;
        }
    }
    return domain;
}
```

Figure 1.8: The implement of Newton's method

1. *If* $N([x]) \cap [x] = \emptyset$*, then* $[x]$ *contains no roots of* f*;*

2. *If* $N([x]) \subseteq [x]$*, then* $[x]$ *contains exactly one root of* f*.*

In Figure 1.8 is the related code that the implement of automatic differentiation, newton method and interval arithmetic in our rigorous computation.

1.2.5 The Power Method

In numerical computation, the power method is an effective way to approximate the dominant eigenvalue λ of a matrix $n \times n$ matrix A and the eigenvector corresponding to λ.

The algorithm of the power method is:

1. Take an arbitrary normalised vector x_0.

2. Apply the matrix A to vector x_0 to get the first approximation of dominant eigenvector x_1 and normalised it. The first approximation of dominant eigenvalue λ_1 can be computed by $Ax_1 \cdot x_1$.

3. Repeat step two n times, we can have the approximation of the dominant eigenvector x_n and the approximation of the dominant eigenvalue λ_n. Each iteration provides a better approximation of the dominant eigenvector and dominant eigenvalue of matrix A.

4. Stop the process when the given error condition is satisfied.

There are two ways to normalise the eigenvector at step two: the Euclidean scaling and the maximum entry scaling. In book [1], Anton and Rorres present two theorems for

this.

For the Euclidean scaling, the theorem states as:

Theorem 1.2.12. *Let A be a $n \times n$ matrix with a positive dominant eigenvalue λ. If x_0 is a unit vector in $\mathbb{R}^n$ that is not orthogonal to the eigenspace corresponding to λ, then the normalised power sequence*

$$x_0, x_1 = \frac{Ax_0}{\|Ax_0\|}, x_2 = \frac{Ax_1}{\|Ax_1\|}, ..., x_k = \frac{Ax_{k-1}}{\|Ax_{k-1}\|}, ...$$

converges to a unit dominant eigenvector, and the sequence

$$Ax_1 \cdot x_1, Ax_2 \cdot x_2, Ax_3 \cdot x_3, ..., Ax_k \cdot x_k, ...$$

converges to the dominant eigenvalue λ.

For the maximum entry scaling, the theorem states:

Theorem 1.2.13. *Let A be a $n \times n$ matrix with a positive dominant eigenvalue λ. If x_0 is a unit vector in $\mathbb{R}^n$ that is not orthogonal to the eigenspace corresponding to λ, then the normalised power sequence*

$$x_0, x_1 = \frac{Ax_0}{\max Ax_0}, x_2 = \frac{Ax_1}{\max Ax_1}, ..., x_k = \frac{Ax_{k-1}}{\max Ax_{k-1}}, ...$$

converges to a unit dominant eigenvector, and the sequence

$$\frac{Ax_1 \cdot x_1}{x_1 \cdot x_1}, \frac{Ax_2 \cdot x_2}{x_2 \cdot x_2}, \frac{Ax_3 \cdot x_3}{x_3 \cdot x_3}, ..., \frac{Ax_k \cdot x_k}{x_k \cdot x_k}, ...$$

converges to the dominant eigenvalue λ.

The rate of convergence for the power method depends on the ratio of the absolute value of the largest eigenvalue and the second largest eigenvalue. For a matrix A, we can arrange the eigenvalues of A as follows:

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq ..... \geq |\lambda_k|,$$

Then, the ratio $\frac{|\lambda_1|}{|\lambda_2|}$ will determine the rate convergence. If this ratio is close to 1, the convergence rate is slow; the greater the ratio is the faster the convergence is.

As the power method is a method to approximate the dominant eigenvalue and the corresponding eigenvector of matrix A, when we use this method in rigorous computation, the most important thing is the error approximation. The relative error in the k-th iteration is defined by:

$$\text{Relative error in } \lambda^{(k)} = |\frac{\lambda - \lambda^{(k)}}{\lambda}|$$

where $\lambda^{(k)}$ is the k-th approximation of eigenvalue λ.

In most cases, we can not know the exact dominant eigenvalue λ of matrix A. In non-rigorous computation the iteration process is stopped when the relative error:

$$\text{Estimate relative error in } \lambda^{(k)} = |\frac{\lambda^{(k)} - \lambda^{(k-1)}}{\lambda^{(k)}}| < E$$

where E is the given error tolerance.

However, the relative error is not enough for the rigorous computation, we have to know the difference between the exact dominant eigenvalue and the approximation dominant eigenvalue, or find an upper bound of this error. In [21], Friedman has found a bound on the error, but only for a positive symmetric matrix. Our matrix approximation of a Perron-Frobenius operator is only stochastic but not symmetric.

In [26], Galatolo and Nisoli rigorously estimated the error for the stochastic matrix case. Based on the fact that Ulam matrix P is a stochastic matrix and it contracts the simplex of vectors v, with $\|v\|_1 = 1$ and $\int v = 0$: for any $v \in V$, we can find $n > 0$ and $0 < \alpha < 1$, such that

$$\|P^n v\|_1 \leq 2 \max_i \|P^n(e_1 - e_i)\|_1 := \alpha.$$

Thus, the maximum error of power computation are bounded by α.

Chapter 2

Rigorous approximation of diffusion coefficients for uniformly expanding maps of the interval

In this Chapter we use Ulam's method to provide rigorous approximations of diffusion coefficients for expanding maps. Such coefficients are focal in the study of limit theorems for dynamical systems (see [17, 33, 39, 44] and references therein). This chapter is based on our work in [5].

2.1 The setting

The system and its transfer operator

Let $(I, \mathcal{B}, m)$ be the measure space, where $I := [0,1]$, $\mathcal{B}$ is Borel σ-algebra, and m is the Lebesgue measure on I. Let $T : I \to I$ be piecewise monotonic[1], piecewise C^2 and uniformly expanding ; i.e., $\inf_x |T'x| \geq \beta > 1$ (see [38] for original reference and first results on the existence on absolutely continuous invariant measure). We recall that the transfer operator (Perron-Frobenius) [7] associated with T, $P : L^1 \to L^1$ is defined by duality: for $f \in L^1$ and $g \in L^\infty$

$$\int_I f \cdot g \circ T dm = \int_I P(f) \cdot g dm.$$

Moreover, for $f \in L^1$ we have

$$Pf(x) = \sum_{y=T^{-1}x} \frac{f(y)}{|T'(y)|}.$$

For $f \in L^1$, we define

$$Vf = \inf_{\overline{f}} \{\mathrm{var}\overline{f} \ : f = \overline{f} \text{ a.e.}\},$$

where

$$\mathrm{var}\overline{f} = \sup\{\sum_{i=0}^{l-1} |\overline{f}(x_{i+1}) - \overline{f}(x_i)| \ : 0 = x_0 < x_1 < \cdots < x_l = 1\}.$$

[1] This was defined in Definition 1.1.36.

We denote by BV the space of functions of bounded variation on I equipped with the norm $||\cdot||_{BV} = V(\cdot) + ||\cdot||_1$. Further, we introduce the mixed operator norm which will play a key role in our approximation:

$$|||P|||_{BV\to L^1} = \sup_{||f||_{BV}\leq 1} ||Pf||_1.$$

Assumptions

We assume[2]:

(A1) $\exists\ \alpha \in (0,1)$, and $B_0 \geq 0$ such that $\forall f \in BV$

$$VPf \leq \alpha Vf + B_0||f||_1;$$

(A2) P, as operator on BV, has 1 as a simple eigenvalue. Moreover P has no other eigenvalues whose modulus is unity; i.e., P has a spectral gap on BV.

Remark 2.1.1. *It is important to remark that the constants α and B_0 in* **(A1)** *depend only on the map T and have explicit analytic expressions (see [38]).*

The above assumptions imply that T admits a unique absolutely continuous invariant measure ν, such that $\frac{d\nu}{dm} := h \in BV$. Moreover, the system $(I, \mathcal{B}, \nu, T)$ is mixing and it enjoys exponential decay of correlations for observables in BV; i.e., for all $n > 1$, $\phi_1 \in L^\infty$ and $\phi_2 \in BV$, there are $0 < \lambda < 1$ and $C_{\phi_1,\phi_2} > 0$ so that ,

$$|\int (\phi_1 \circ T^n)\phi_2 d\nu - \int \phi_1 d\nu \int \phi_2 d\nu| \leq C_{\phi_1,\phi_2}\lambda^n. \tag{2.1.1}$$

See [7] for a profound background on this topic.

The problem

Let $\psi \in BV$ and define

$$\sigma^2 := \lim_{n\to\infty} \frac{1}{n} \int_I \left(\sum_{i=0}^{n-1} \psi(T^i x) - n \int_I \psi d\nu \right)^2 d\nu. \tag{2.1.2}$$

Under our assumptions the limit in (2.1.2) exists (see [33]), and by using (2.1.1) and the duality property of P, one can rewrite σ^2 as

$$\sigma^2 := \int_I \hat{\psi}^2 h dm + 2\sum_{i=1}^{\infty} \int_I P^i(\hat{\psi}h)\hat{\psi} dm, \tag{2.1.3}$$

where

$$\hat{\psi} := \psi - \mu \text{ and } \mu := \int_I \psi d\nu.$$

The number σ^2 is called the variance, or the diffusion coefficient, of $\sum_{i=0}^{n-1} \psi(T^i x)$. In particular, for the systems under consideration, it is well known (see [33]) that the Central

[2]It is well known that the systems under consideration satisfy a Lasota-Yorke inequality. What we are assuming in **(A1)** is that there is no constant in front of α. Such an assumption is satisfied for instance when $\inf_x |T'(x)| > 2$ or when T is piecewise onto. When the original map T does not satisfy the assumption **(A1)**, one can find an iterate of T where **(A1)** is satisfied, and then apply the results of this paper.

Limit Theorem holds:

$$\frac{1}{\sqrt{n}}\left(\sum_{i=0}^{n-1}\psi(T^i x) - n\int_I \psi d\nu\right) \xrightarrow{\text{law}} \mathcal{N}(0,\sigma^2)$$

and $\sigma^2 > 0$ if and only if $\psi \neq c + \phi \circ T - \phi$, $\phi \in BV$, $c \in \mathbb{R}$.

The goal of this paper is to provide an algorithm whose output approximates σ^2 with rigorous error bounds. The first step in our approach will be to discretize P as follows:

2.1.1 Ulam's scheme

Let $\eta := \{I_k\}_{k=1}^{d(\eta)}$ be a partition of $[0,1]$ into intervals of size $m(I_k) \leq \varepsilon$ where $d(\eta) = 1/\varepsilon$, we define the ε as mesh(η). Let $\mathfrak{B}_\eta$ be the σ-algebra generated by η and for $f \in L^1$ define the projection

$$\Pi_\eta f = E(f|\mathfrak{B}_\eta),$$

and

$$P_\eta = \Pi_\eta \circ P \circ \Pi_\eta.$$

P_η, which is called Ulam's approximation of P, is finite rank operator which can be represented by a (row) stochastic matrix acting on vectors in $\mathbb{R}^{d(\eta)}$ by left multiplication. Its entries are given by

$$P_{kj} = \frac{m(I_k \cap T^{-1}(I_j))}{m(I_k)}.$$

The following lemma collects well known results on P_ε. See for instance [40] for proofs of (1)-(4) of the lemma, and [40, 26] and references therein for statement (5) of the lemma.

Lemma 2.1.1. *For $f \in BV$ we have*

1. $V(\Pi_\varepsilon f) \leq V(f)$;
2. $||f - \Pi_\varepsilon f||_1 \leq \varepsilon V(f)$;
3. $$VP_\varepsilon f \leq \alpha V f + B_0||f||_1,$$
   where α and B_0 are the same constants that appear in **(A1)**;
4. $|||P_\varepsilon - P||| \leq \Gamma\varepsilon$, *where* $\Gamma = \max\{\alpha + 1, B_0\}$;
5. *P_ε has a unique fixed point $h_\varepsilon \in BV$. Moreover, $\exists$ a computable constant K_* such that*
   $$||h_\varepsilon - h||_1 \leq K_* \varepsilon \ln \varepsilon^{-1}.$$
   In particular, for any $\tau > 0$, there exists ε_* such that $||h_{\varepsilon_*} - h||_1 \leq \tau$.

Statement of the main result

Define

$$\hat{\psi}_\varepsilon := \psi - \mu_\varepsilon \text{ and } \mu_\varepsilon := \int_I \psi h_\varepsilon dm.$$

Set

$$\sigma^2_{\varepsilon,l} := \int_I \hat{\psi}^2_\varepsilon h_\varepsilon dm + 2\sum_{i=1}^{l-1} \int_I P^i_\varepsilon(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon dm.$$

Theorem 2.1.2. *For any $\tau > 0$, $\exists\ l_* > 0$ and $\varepsilon_* > 0$ such that [5]*

$$|\sigma^2_{\varepsilon_*,l_*} - \sigma^2| \leq \tau.$$

Remark 2.1.2. *Theorem 2.1.2 says that given a pre-specified tolerance on error $\tau > 0$, one finds $l_* > 0$ and $\varepsilon_* > 0$ so that $\sigma^2_{\varepsilon_*,l_*}$ approximates σ up to the pre-specified error τ. In subsection 2.2.1 we provide an algorithm that can be implemented on a computer to find l_* and ε_*, and consequently $\sigma^2_{\varepsilon_*,l_*}$.*

To illustrate the issue of the rate of convergence and to elaborate on why we define the approximate diffusion by $\sigma^2_{\varepsilon,l}$ as a truncated sum, let us define

$$\sigma^2_\varepsilon := \int_I \hat{\psi}^2_\varepsilon h_\varepsilon dm + 2\sum_{i=1}^{\infty} \int_I P^i_\varepsilon(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon dm.$$

Theorem 2.1.3. *$\exists$ a computable constant $\tilde{K}_*$ such that [5]*

$$|\sigma^2_\varepsilon - \sigma^2| \leq \tilde{K}_* \varepsilon (\ln \varepsilon^{-1})^2.$$

Remark 2.1.3. *To compute the constant $\tilde{K}_*$, we need the constants α and B_0 from Lasota-Yorke inequality and the ϵ from the partition η.*

Remark 2.1.4. *Note that σ^2_ε can be written as*

$$\begin{aligned} \sigma^2_\varepsilon &= \int_I \hat{\psi}^2_\varepsilon h_\varepsilon dm + 2\sum_{i=1}^{\infty} \int_I P^i_\varepsilon(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon dm \\ &= -\int_I \hat{\psi}^2_\varepsilon h_\varepsilon + 2\int_I \hat{\psi}_\varepsilon(\mathbf{1} - P_\varepsilon)^{-1}(\hat{\psi}_\varepsilon h_\varepsilon) dm. \end{aligned} \tag{2.1.4}$$

Since P_ε has a matrix representation, and consequently $(I - P_\varepsilon)^{-1}$ is a matrix, one may think that σ^2_ε provides a more sensible formula to approximate σ^2 than $\sigma^2_{\varepsilon,l}$. However, from the rigorous computational point of view one has to take into account the errors that arise at the computer level when estimating $(I - P_\varepsilon)^{-1}$. Indeed $(I - P_\varepsilon)^{-1}$ can be computed rigorously on the computer by estimating it by a finite sum plus an error term coming from estimating the tail of the sum[3]. This is what we do in Theorem 2.1.2.

Remark 2.1.5. *In [14] an example of a* highly regular *expanding map (piecewise affine) was presented where the* exact *rate of Ulam's method for approximating the invariant density h is $\varepsilon \ln \varepsilon^{-1}$. In Theorem 2.1.3 the rate for approximating σ^2 is $\varepsilon(\ln \varepsilon^{-1})^2$. This is due to the fact that $||h - h_\varepsilon||_1$ is an essential part in estimating σ^2 and the extra $\ln \varepsilon^{-1}$ appears because of the infinite sum in the formula of σ^2.*

[3]Of course, usual computer software would give an estimated matrix of $(I - P_\varepsilon)^{-1}$, but it does not give the errors it made in its approximation.

Remark 2.1.6. *By using the representation* (2.1.4) *of* σ_ε^2*, it is obvious that the main task in the proof of Theorem 2.1.3 is to estimate*

$$|||(\mathbf{1}-P)^{-1}-(\mathbf{1}-P_\varepsilon)^{-1}|||_{BV_0\to L^1},$$

where $BV_0=\{f\in BV \ s.t. \int f dm=0\}$*. Thus, it would be tempting to use estimate (9) in Theorem 1 of [37] (also stated in Theorem 1.1.14 in this thesis), which reads:*

$$\begin{aligned}&|||(\mathbf{1}-P)^{-1}-(\mathbf{1}-P_\varepsilon)^{-1}|||_{BV_0\to L^1}\\ &\qquad\leq |||P-P_\varepsilon|||^{\theta}_{BV_0\to L^1}(c_1||(\mathbf{1}-P_\varepsilon)^{-1}||_{BV_0}+c_2||(\mathbf{1}-P_\varepsilon)^{-1}||^2_{BV_0}),\end{aligned}\tag{2.1.5}$$

where $\theta=\frac{\ln(r/\alpha)}{\ln(1/\alpha)}$*,* $r\in(\alpha,1)$*, and* c_1,c_2 *are constants that dependent only on* α*,* B_0 *and* r*. On the one hand, this would lead to a shorter proof than the one we present in section 2.2; however, estimate* (2.1.5) *would lead to a convergence rate of order* ε^θ*, where* $0<\theta<1$ *which is slower than the rate obtained in Theorem 2.1.3. Naturally, this have led us to opt for using the proofs of section 2.2.*

2.2 Proofs and an Algorithm

Lemma 2.2.1. *For* $\psi\in BV$*, we have [5]*

1. $||\hat{\psi}||_\infty\leq 2||\psi||_\infty$ *and* $||\hat{\psi}_\varepsilon||_\infty\leq 2||\psi||_\infty$*;*
2. $|\int_I(\hat{\psi}^2h-\hat{\psi}_\varepsilon^2h_\varepsilon)dm|\leq 8||\psi||^2_\infty||h_\varepsilon-h||_1$.

Proof. Using the definition of $\hat{\psi}$, $\hat{\psi}_\varepsilon$ we get (1). We now prove (2). We have

$$\begin{aligned}|\int_I(\hat{\psi}_\varepsilon^2-\hat{\psi}^2)hdm|=&|\int_I(\hat{\psi}_\varepsilon-\hat{\psi})(\hat{\psi}_\varepsilon+\hat{\psi})hdm|=|\int_I(\mu-\mu_\varepsilon)(2\psi-\mu-\mu_\varepsilon)hdm|\\ &\leq 4||\psi||_\infty|\mu_\varepsilon-\mu|\int_I hdm\leq 4||\psi||^2_\infty||h_\varepsilon-h||_1.\end{aligned}\tag{2.2.1}$$

We now use (1) and (2.4.7) to get

$$\begin{aligned}|\int_I(\hat{\psi}^2h-\hat{\psi}_\varepsilon^2h_\varepsilon)dm|&\leq|\int_I(\hat{\psi}^2h-\hat{\psi}_\varepsilon^2h)dm|+|\int_I(\hat{\psi}_\varepsilon^2h-\hat{\psi}_\varepsilon^2h_\varepsilon)dm|\\ &\leq 8||\psi||^2_\infty||h_\varepsilon-h||_1.\end{aligned}$$

□

Lemma 2.2.2. *For any* $l\geq 1$ *we have [5]*

$$\begin{aligned}&|\sum_{i=1}^{l-1}\int_I\left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon-P^i(\hat{\psi}h)\hat{\psi}\right)dm|\leq 8(l-1)\cdot||\psi||^2_\infty\cdot||h_\varepsilon-h||_1\\ &+2||\psi||_\infty|||P_\varepsilon-P|||\sum_{i=1}^{l-1}\sum_{j=0}^{i-1}\left(2||\psi||_\infty(B_j+1+\frac{\alpha^jB_0}{1-\alpha})+\frac{\alpha^j(B_0+1-\alpha)}{1-\alpha}V\psi\right),\end{aligned}$$

where $B_j=\sum_{k=0}^{j-1}\alpha^kB_0$.

Proof.

$$\begin{aligned}
&\left|\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon - P^i(\hat{\psi}h)\hat{\psi}\right)dm\right| \\
&\leq |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon - P_\varepsilon^i(\hat{\psi}h)\hat{\psi}\right)dm| + |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}h)\hat{\psi} - P^i(\hat{\psi}h)\hat{\psi}\right)dm| \\
&\leq |\sum_{i=1}^{l-1}\int_I P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon - \hat{\psi}h)\psi dm| + |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\mu_\varepsilon - P_\varepsilon^i(\hat{\psi}h)\mu\right)dm| \\
&\qquad\qquad + |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}h)\hat{\psi} - P^i(\hat{\psi}h)\hat{\psi}\right)dm| \\
&:= (I) + (II) + (III).
\end{aligned}$$

We have

$$\begin{aligned}
(I) &\leq ||\psi||_\infty \sum_{i=1}^{l-1}\int_I |\hat{\psi}_\varepsilon h_\varepsilon - \hat{\psi}h|dm \\
&= ||\psi||_\infty \cdot (l-1)\int_I |\hat{\psi}_\varepsilon h_\varepsilon - \hat{\psi}_\varepsilon h + \hat{\psi}_\varepsilon h - \hat{\psi}h|dm \\
&\leq ||\psi||_\infty \cdot (l-1)\left(||\hat{\psi}_\varepsilon||_\infty ||h_\varepsilon - h||_1 + |\mu - \mu_\varepsilon|\right) \\
&\leq 3||\psi||_\infty^2 \cdot (l-1)\cdot ||h_\varepsilon - h||_1.
\end{aligned} \tag{2.2.2}$$

We estimate (II):

$$\begin{aligned}
(II) &\leq |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\mu_\varepsilon - P_\varepsilon^i(\hat{\psi}h)\mu_\varepsilon\right)dm| + |\sum_{i=1}^{l-1}\int_I \left(P_\varepsilon^i(\hat{\psi}h)\mu_\varepsilon - P_\varepsilon^i(\hat{\psi}h)\mu\right)dm| \\
&\leq (l-1)|\mu_\varepsilon|\int_I \left|\hat{\psi}_\varepsilon h_\varepsilon - \hat{\psi}h\right|dm + 2(l-1)\cdot||\psi||_\infty|\mu_\varepsilon - \mu| \\
&\leq 3||\psi||_\infty^2\cdot(l-1)\cdot||h_\varepsilon - h||_1 + 2(l-1)\cdot||\psi||_\infty^2||h_\varepsilon - h||_1 \\
&= 5||\psi||_\infty^2\cdot(l-1)\cdot||h_\varepsilon - h||_1.
\end{aligned} \tag{2.2.3}$$

Finally we estimate (III):

$$\begin{aligned}
(III) &\leq 2||\psi||_\infty \sum_{i=1}^{l-1}\sum_{j=0}^{i-1}||P_\varepsilon^{i-1-j}(P_\varepsilon - P)P^j(\hat{\psi}h)||_1 \\
&\leq 2||\psi||_\infty \cdot |||P_\varepsilon - P||| \cdot \sum_{i=1}^{l-1}\sum_{j=0}^{i-1}||P^j(\hat{\psi}h)||_{BV} \\
&\leq 2||\psi||_\infty \cdot |||P_\varepsilon - P||| \cdot \sum_{i=1}^{l-1}\sum_{j=0}^{i-1}\left(\alpha^j V(\hat{\psi}h) + (B_j+1)||\hat{\psi}h||_1\right) \\
&\leq 2||\psi||_\infty |||P_\varepsilon - P||| \sum_{i=1}^{l-1}\sum_{j=0}^{i-1}\left(2||\psi||_\infty(B_j + 1 + \frac{\alpha^j B_0}{1-\alpha}) + \frac{\alpha^j(B_0+1-\alpha)}{1-\alpha}V\psi\right),
\end{aligned} \tag{2.2.4}$$

where in the above estimate we have used **(A1)** and its consequence that $Vh \leq \frac{B_0}{1-\alpha}$. Combining estimates (2.2.2),(2.2.3) and (2.2.4) completes the proof of the lemma. □

Proof. (Proof of Theorem 2.1.2)

$$|\sigma^2_{\varepsilon,l} - \sigma^2| \leq |\int_I (\hat{\psi}^2 h - \hat{\psi}^2_\varepsilon h_\varepsilon) dm| + 2|\sum_{i=1}^{l-1} \int_I \left(P^i_\varepsilon(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon - P^i(\hat{\psi}h)\hat{\psi} \right) dm|$$
$$+ 4||\psi||_\infty \sum_{i=l}^{\infty} ||P^i(\hat{\psi}h)||_{BV}$$
$$:= (I) + (II) + (III).$$

We start with (III). Since $\int_I \hat{\psi} h dm = 0$, there exists a computable constant C_* and a computable number[4] ρ_*, where $\alpha < \rho_* < 1$, such that

$$||P^i(\hat{\psi}h)||_{BV} \leq ||\hat{\psi}h||_{BV} C_* \rho^i_* \leq (2||\psi||_\infty + V(\psi)) \frac{B_0 + 1 - \alpha}{1 - \alpha} C_* \rho^i_*.$$

Consequently,

$$(III) \leq 4||\psi||_\infty (2||\psi||_\infty + V(\psi)) \frac{B_0 + 1 - \alpha}{(1-\alpha)(1-\rho_*)} C_* \rho^l_*.$$

Thus, choosing l_* such that

$$l_* := \left\lceil \frac{\log(\tau/2) - \log\left(4||\psi||_\infty (2||\psi||_\infty + V(\psi)) \frac{B_0+1-\alpha}{(1-\alpha)(1-\rho_*)} C_*\right)}{\log \rho_*} \right\rceil \tag{2.2.5}$$

implies

$$4||\psi||_\infty \sum_{i=l_*}^{\infty} ||P^i(\hat{\psi}h)||_{BV} \leq \frac{\tau}{2}.$$

Fix l_* as in (2.3.5). Now using Lemmas 2.1.1, 2.2.1 and 2.2.2, we can find ε_* such that

$$|\int_I (\hat{\psi}^2 h - \hat{\psi}^2_{\varepsilon_*} h_{\varepsilon_*}) dm| + 2|\sum_{i=1}^{l_*-1} \int_I \left(P^i_{\varepsilon_*}(\hat{\psi}_\varepsilon h_{\varepsilon_*})\hat{\psi}_\varepsilon - P^i(\hat{\psi}h)\hat{\psi} \right) dm| \leq \frac{\tau}{2}.$$

This completes the proof of the theorem. □

2.2.1 Algorithm [5]

Theorem 2.1.3 suggests an algorithm as follows. Given T that satisfies **(A1)** and **(A2)** and $\tau > 0$ a tolerance on error:

1. Find l_* such that
$$4||\psi||_\infty \sum_{i=l_*}^{\infty} ||P^i(\hat{\psi}h)||_{BV} \leq \frac{\tau}{2}.$$

2. Fix l_* from (1).

3. Find $\varepsilon_* = \text{mesh}(\eta)$ such that
$(16(l_* - 1) + 8) \cdot ||\psi||^2_\infty \cdot ||h_{\varepsilon_*} - h||_1$
$+ 4||\psi||_\infty \sum_{i=1}^{l_*-1} \sum_{j=0}^{i-1} \left(2||\psi||_\infty (B_j + 1 + \frac{\alpha^j B_0}{1-\alpha}) + \frac{\alpha^j (B_0 + 1 - \alpha)}{1-\alpha} V\psi \right) |||P_{\varepsilon_*} - P||| \leq \frac{\tau}{2}.$

4. Output $\sigma^2_{\varepsilon_*,l_*} := \int_I \hat{\psi}^2_{\varepsilon_*} h_{\varepsilon_*} dm + 2\sum_{i=1}^{l_*-1} \int_I P^i_{\varepsilon_*}(\hat{\psi}_{\varepsilon_*} h_{\varepsilon_*})\hat{\psi}_{\varepsilon_*} dm$.

[4]There are many ways to approximate (III). In the proof we follow [2, 3] which is based on the spectral perturbation result of [37]. Other techniques are possible, see for instance [42] and the recent work of [27]. We will leave it to reader to decide which one is more convenient to use when implementing the algorithm.

Remark 2.2.1. *Note that the split of $\frac{\tau}{2}$ between items (1) and (2) in Algorithm 2.2.1 to lead to an error of at most τ can be relaxed in following way. One can compute the error in item (1) to be at most $\frac{\tau}{k}$ and in item (2) to be $\frac{k-1}{k}\tau$ for any integer $k \geq 2$. We exploit this fact in the implementation in section 2.3.*

Proof. (Proof of Theorem 2.1.3)

$$\begin{aligned}|\sigma_\varepsilon^2 - \sigma^2| \leq |\int_I (\hat{\psi}^2 h - \hat{\psi}_\varepsilon^2 h_\varepsilon) dm| + 2|\sum_{i=1}^{l-1} \int_I \left(P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)\hat{\psi}_\varepsilon - P^i(\hat{\psi}h)\hat{\psi}\right) dm| \\ + 4||\psi||_\infty \sum_{i=l}^{\infty} ||P^i(\hat{\psi}h)||_{BV} + 4||\psi||_\infty \sum_{i=l}^{\infty} ||P_\varepsilon^i(\hat{\psi}_\varepsilon h_\varepsilon)||_{BV} \\ := (I) + (II) + (III) + (IV).\end{aligned}$$

We first get an estimate on (III) and (IV). There exists a computable constant C_* and a computable number ρ_*, where $\alpha < \rho_* < 1$, such that

$$(III) + (IV) \leq 8||\psi||_\infty \left(2||\psi||_\infty + V(\psi)\right) \frac{B_0 + 1 - \alpha}{(1-\alpha)(1-\rho_*)} C_* \rho_*^l.$$

For (II), as in Lemma 2.2.2, in particular (2.2.4), and by using Lemma 2.1.1, we have

$$\begin{aligned}(II) &\leq 4||\psi||_\infty \sum_{i=1}^{l-1} \sum_{j=0}^{i-1} ||P_\varepsilon^{i-1-j}(P_\varepsilon - P)P^j(\hat{\psi}h)||_1 + 16(l-1) \cdot ||\psi||_\infty^2 \cdot ||h_\varepsilon - h||_1 \\ &\leq 4||\psi||_\infty \Gamma \cdot \left(\alpha V(\psi) \frac{B_0 + 1 - \alpha}{1-\alpha} + ||\psi||_\infty \frac{2B_0 + \alpha B_0}{1-\alpha}\right)(l-1)\varepsilon \\ &\quad + K_* 16(l-1)\varepsilon \ln \varepsilon^{-1}.\end{aligned}$$

For (I) we use Lemmas 2.1.1 and 2.2.1 to obtain

$$(I) \leq 8||\psi||_\infty^2 ||h_\varepsilon - h||_1 \leq 8||\psi||_\infty^2 K_* \varepsilon \ln \varepsilon^{-1}.$$

Finally, choosing $l = \lceil \frac{\ln \varepsilon}{\ln \rho_*} \rceil$ leads to the rate $\tilde{K}_* \varepsilon (\ln \varepsilon^{-1})^2$. □

2.3 Example of rigorous computation of the diffusion coefficient

Let $T : I \to I$ be given by

$$T(x) = \begin{cases} \frac{9x}{1-x} \text{ if } x \in [0, \frac{1}{10}), \\ 10x - i \text{ if } x \in [\frac{i}{10}, \frac{i+1}{10}). \end{cases} \tag{2.3.1}$$

where $i = 1, 2..., 9$. We first show that T satisfies assumptions **(A1)** and **(A2)**. Notice that $\forall f \in BV$ we have,

$$VPf \leq \alpha Vf + B_0 \|f\|_1, \tag{2.3.2}$$

where $\alpha = \frac{1}{9}$ and $B_0 = \frac{2}{9}$. Thus, assumption **(A1)** is satisfied. Moreover, since T is piecewise onto, the system is mixing. Consequently assumption **(A2)** is satisfied. We now use Algorithm 2.2.1 to approximate the variance σ^2 of this system when the observable $\psi(x) = x^2$. In particular we want to perform the approximation up to a pre-specified error $\tau = 0.001$.

2.3.1 The implementation of Algorithm 2.2.1

Computation of Item (1) in Algorithm 2.2.1

In this step, we find l^* so that item (1) of Algorithm 2.2.1 is satisfied. In particular, we want to find l^* such that

$$4||\psi||_\infty \sum_{i=l_*}^{\infty} ||P^i(\hat{\psi}h)||_{BV} \leq \frac{\tau}{1000}. \tag{2.3.3}$$

As explained in Remark 2.2.1, instead of verifying item (1) to be smaller than $\frac{\tau}{2}$, we verify that it is smaller than $\frac{\tau}{1000}$. This will give us more room in verifying item (2) so that the sum of the errors from both items is smaller than τ. Since the system satisfies **(A2)**, for any $g \in BV_0$ and any $k \in \mathbb{N}$, there exist $0 < \rho_* < 1$, and $C_* > 0$ such that

$$\|P^k g\|_{BV} \leq C_* \rho_*^k \|g\|_{BV}. \tag{2.3.4}$$

We will compute a $0 < \rho_* < 1$ and a $C_* > 0$ so that (2.3.4) is satisfied. Once these two numbers are computed, we can easily find l_* is satisfied.

Here is the formula given in the proof of the Theorem 2.1.3:

$$l_* := \left\lceil \frac{\log(\tau/2) - \log\left(4||\psi||_\infty (2||\psi||_\infty + V(\psi)) \frac{B_0+1-\alpha}{(1-\alpha)(1-\rho_*)} C_*\right)}{\log \rho_*} \right\rceil \tag{2.3.5}$$

To compute ρ^* and C_*, we use Ulam's method and follow the work of [27]. By equation (2.3.2), for $f \in BV$ we get

$$\|P^n f\|_{BV} \leq \alpha^n \|f\|_{BV} + (\frac{B_0}{1-\alpha})\|f\|_1 \tag{2.3.6}$$
$$\|P^n\|_1 \leq \|P^n\|_{BV} \leq 1 + \frac{B_0}{1-\alpha} := M.$$

Now, we have the constants for the first inequality, for the second one, we need a lemma to estimate its constants.

Lemma 2.3.1. *Suppose there are two norms $||\ ||_s \geq ||\ ||_w$, such that $\forall f \in \mathcal{B}, \forall n \geq 1$*

$$||P^n f||_s \leq A\lambda_1^n ||f||_s + B||f||_w. \tag{2.3.7}$$

Let π_η be a finite rank operator satisfying:

- *$P_\eta = \pi_\eta P \pi_\eta$ with $||\pi_\eta v - v||_w \leq K\eta ||v||_s$;*
- *π_η, P^i and P_η^i are bounded for the norm $||\ ||_w : ||\pi_\eta||_w \leq P$ and $\forall i > 0$, $||P^i||_w \leq M$.*

Then

$$\begin{aligned} ||(P - P_\eta)f||_w &\leq K\eta(A\lambda_1 + P)||f||_s + K\eta B||f||_w \\ ||P^n f - P_\eta^n f||_w &\leq \eta K M \left(\frac{(A\lambda_1 + P)A}{1-\lambda_1} ||f|| + nB(A\lambda_1 + P + M)||f||_w \right). \end{aligned}$$

Proof. We have

$$||(P - P_\eta)f||_w \leq ||\pi_\eta P \pi_\eta f - \pi_\eta P f||_w + ||\pi_\eta P f - P f||_w,$$

but

$$\pi_\eta P \pi_\eta f - \pi_\eta P f = \pi_\eta P(\pi_\eta f - f).$$

Since $||\pi_\eta v - v||_w \leq K\eta||v||_s$

$$||\pi_\eta P(\pi_\eta f - f)||_w \leq P||\pi_\eta f - f||_w \leq PK\eta||f||_s.$$

On the other hand

$$||\pi_\eta Pf - Pf||_w \leq K\eta||Pf||_s \leq K\eta(A\lambda_1||f||_s + B||f||_w)$$

which gives

$$||(P - P_\eta)f||_w \leq K\eta(A\lambda_1 + P)||f||_s + K\eta B||f||_w \tag{2.3.8}$$

Now let us consider $(P_\eta^n - P^n)f$. We have

$$\begin{aligned}
||(P_\eta^n - P^n)f||_w &\leq \sum_{k=1}^{n} ||P_\eta^{n-k}(P_\eta - P)P^{k-1}f||_w \leq M\sum_{k=1}^{n} ||(P_\eta - P)P^{k-1}f||_w \\
&\leq K\eta M\sum_{k=1}^{n}(A\lambda_1 + P)||P^{k-1}f||_s + B||P^{k-1}f||_w \\
&\leq K\eta M\sum_{k=1}^{n}(A\lambda_1 + P)(A\lambda_1^{k-1}||f||_s + B||f||_w) + BM||f||_w \\
&\leq K\eta M(\frac{(A\lambda_1 + P)A}{1-\lambda_1}||f|| + Bn(A\lambda_1 + P + M)||f||_w).
\end{aligned}$$

□

We have

$$\begin{cases} \|P^{n_1}f\|_{BV} \leq \alpha^{n_1}\|f\|_{BV} + (\frac{B_0}{1-\alpha})\|f\|_1 \\ \|P^{n_1}f\|_1 \leq \alpha_2\|f\|_1 + \eta M((\frac{1+\alpha}{1-\alpha})\|f\|_{BV} + B_0 n_1(1+\alpha+M)\|f\|_1. \end{cases} \tag{2.3.9}$$

Next, we identify and compute η and α_2 in $\mathcal{M}$.

Let n_0 be the smallest integer such that

$$\alpha^{n_0} < 1.$$

For this example, since $\alpha = \frac{1}{9}$, n_0=1. Next we identify and compute α_2. Let P_η be the Ulam operator. We assume $\eta = \frac{1}{m}$, where $m \in \mathbb{N}$. We want to find η and $n_1 \geq n_0$ such that $\forall v \in BV_0$

$$\|P_\eta^{n_1}v\|_1 \leq \alpha_2\|v\|_1 \tag{2.3.10}$$

with $\alpha_2 < 1$.

We now explain how we find η and how α_2 is computed. We follow the idea of [26]. Since the Ulam matrix P_η is a stochastic matrix and it contracts the simplex of vectors v, with $\|v\|_1 = 1$ and $\int v = 0$. Thus, for any $v \in BV_0$, we can find $n_1 > 0$ and $0 < \alpha_2 < 1$, such that

$$\|P_\eta^{n_1}v\|_1 \leq 2\max_i \|P_\eta^{n_1}(e_1 - e_i)\|_1 := \alpha_2$$

where vectors $e_1, ...e_i$ are basis of $\mathbb{R}^m$. Then we use the following algorithm:

1. For a set of vectors u_j with size m, we assign 0.5 as its first component and -0.5 as its j^{th}.

2. For $k \in \mathbb{N}$, apply P_η^k to each vector we have, then compute its L^1-norm.

3. Check if the biggest norm we computed from the previous step is smaller than 1, if so, take k as n_1 and the biggest norm as α_2, if not, add 1 to m and repeat previous step.

In our example, the out put of this algorithm outputs $n_1 = 3$ and $\alpha_2 = 0.1$ for $\eta = 1/200000$.

Remark 2.3.1. *When we compute the norm of vectors in $\|P_\eta^{n_1} v\|_1$, we consider the numerical roundoff error of matrix operation. According to [32], this error can be bounded as follows:*

$$\|float(Pv) - Pv\|_1 \leq \gamma_{N_z} \cdot \|P\| \cdot \|v\|_1$$

where

$$\gamma_{N_z} = \frac{N_z u}{1 - N_z u}$$

and $float(Pv)$ is the computed Pv value, N_z is the number of nonzero element of vector v and u is the machine precision. In fact, after taking this roundoff error into account, we computed $\alpha_2 = 0.099995$, but we will still use an upper bound on this , namely $\alpha_2 = 0.1$.

We have computed all the constants we need to compute the following matrix now.

$$\mathcal{M} = \begin{pmatrix} \alpha^{n_1} & \frac{B_0}{1-\alpha} \\ \eta M(\frac{1+\alpha}{1-\alpha}) & \eta M B_0 n_1(1+\alpha+M) + \alpha_2 \end{pmatrix},$$

where the (positive) entries of $\mathcal{M}$ will be defined below. It is given by:

$$\mathcal{M} = \begin{pmatrix} 0.0014 & 0.25 \\ 0.00015625 & 0.1002 \end{pmatrix}.$$

with (2.4.4) (below) is satisfied with $n_1 = 3$, $\alpha = \frac{1}{9}$, $B_0 = \frac{2}{9}$, $\eta = 1/200000$, $M = 1.25$, $\alpha_2 = 0.1$.

In particular, it is shown that

$$\begin{pmatrix} \|P^{in_1} g\|_{BV} \\ \|P^{in_1} g\|_{L^1} \end{pmatrix} \preceq \mathcal{M}^i \begin{pmatrix} \|g\|_{BV} \\ \|g\|_{L^1} \end{pmatrix}, \tag{2.3.11}$$

where $\preceq$ means component-wise inequalities, e.g. for vectors $\overrightarrow{x} = (x_1, x_2)$ and $\overrightarrow{y} = (y_1, y_2)$, $\overrightarrow{x} \preceq \overrightarrow{y}$, iff , $x_1 \leq y_1$ and $x_2 \leq y_2$; and as a consequence it is also proved in [27] that for any $k \in \mathbb{N}$

$$\|P^k g\|_{BV} \leq (A/a + B/b)\rho_*^{\llcorner \frac{k}{n_1} \lrcorner} \|g\|_{BV}, \tag{2.3.12}$$

where ρ_* is the dominant eigenvalue of $\mathcal{M}$ and (a, b) is the corresponding eigenvector.

Thus, $\rho_* = 0.1006$ and the eigenvector (a, b) associated to the eigenvalue ρ_* is given by:

$a = 0.3969$, $b = 0.6031$.Thus, by (2.4.6), we obtain

$$\|P_*^l g\|_{BV} \leq (2.9342) \times 0.1006^{\llcorner \frac{l_*}{n_1} \lrcorner} \|g\|_{BV}.$$

Consequently, $C_* = 2.9342$ and we compute l_* using the formula in (2.3.5) to obtain

$$\begin{aligned} l_* &= n_1 \cdot \left\lceil \frac{\log(\tau/1000) - \log(4\|\psi\|_\infty(2\|\psi\|_\infty + V(\psi))\frac{B_0+1-\alpha}{(1-\alpha)(1-\rho_*)}C_*)}{\log \rho_*} \right\rceil \\ &= 24. \end{aligned}$$

Computation of Item (2) of Algorithm 2.2.1

From now on l_* is fixed and it is equal to 24. So far, we executed the first loop of the Algorithm 2.2.1; i.e.,

$$4\|\psi\|_\infty \sum_{i=24}^{\infty} \|P^i(\hat{\psi})\|_{BV} \leq \frac{\tau}{1000}.$$

Computation of Item (3) of Algorithm 2.2.1

In this step, we have to find η_*, a mesh size of the Ulam discretization, such that

$$\begin{aligned} &(16(l_*-1)+8)\cdot\|\psi\|_\infty^2 \cdot \|h_{\eta_*} - h\|_1 \\ &+ 4\|\psi\|_\infty \sum_{i=1}^{l_*-1}\sum_{j=0}^{i-1}(2\|\psi\|_\infty(B_0+1+\frac{\alpha^j B_0}{1-\alpha}) + \frac{\alpha^j(B_0+1-\alpha)}{1-\alpha}V\psi)|||P_{\eta_*} - P||| \\ &\leq \frac{999}{1000}\tau. \end{aligned} \tag{2.3.13}$$

The main issue in this estimate is the rigorous approximation of the T-invariant density, h, in the L^1-norm. In the next subsection, we explain how we perform such approximation. We follow the ideas of [26].

Approximation of the invariant density Let h_η be the invariant density of the Ulam's approximation P_η, $\tilde{h}_\eta$ be the invariant density of the Ulam matrix $\tilde{P}_\varepsilon$, which is the output of our computer program, and h_{η_*} be the vector that we finally use to approximate h. We have

$$\|h - h_{\eta_*}\|_1 \leq \|h - h_\eta\|_1 + \|h_\eta - \tilde{h}_\eta\|_1 + \|\tilde{h}_\eta - h_{\eta_*}\|_1, \tag{2.3.14}$$

where $\|h-h_\eta\|_1$ is the discretization error, $\|h_\eta-\tilde{h}_\eta\|_1$ is the approximation error, $\|\tilde{h}_\eta-h_{\eta_*}\|_1$ the numerical error. Now, we will estimate these three terms respectively.

$\|h - h_\eta\|_1$ **estimates**

$$\begin{aligned} \|h - h_\eta\|_1 \leq \|P_\eta^N h_\eta - P^N h\|_1 &\leq \|P_\eta^N h_\eta - P_\eta^N h\|_1 + \|P_\eta^N h - P^N h\|_1 \\ &\leq \|P_\eta^N(h_\eta - h)\|_1 + \|P_\eta^N h - P^N h\|_1 \end{aligned}$$

$\exists N \in \mathbb{N}$ such that $\|P_\eta^N(h_\eta - h)\|_1 \leq \frac{1}{2}\|h - h_\eta\|_1$. That iteration N can be compute by a program. Thus,

$$\|h - h_\eta\|_1 \leq \frac{1}{2}\|h - h_\eta\|_1 + \|P_\eta^N h - P^N h\|_1 \leq 2\|P_\eta^N h - P^N h\|_1$$

Consider the term $\|P_\eta^N h - P^N h\|_1$,

$$\|P_\eta^N h - P^N h\|_1 \le \|\sum_{k=1}^{l} P_\eta^{N-k}(P_\eta - P)h\|_1 \le \sum_{k=1}^{N} \|P_\eta^{N-k}\|_1 \|(P_\eta - P)h\|_1$$

where $\|P_\eta^{N-k}\|_1 \le 1$. Next, estimate the term $\|(P_\eta - P)h\|_1$,

$$\|(P_\eta - P)h\|_1 = \|(\Pi_\eta P \Pi_\eta - P)h\|_1 \le \|h - \Pi_\eta h\|_1 + \|\Pi_\eta (P - P\Pi_\eta)h\|_1$$

For $\|h - \Pi_\eta h\|_1$, we have:

$$\|h - \Pi_\eta h\|_1 \le \eta V(f)$$

where η is the length of partition size. For $\|\Pi_\eta(P - P\Pi_\eta)h\|_1$, we estimate as:

$$\|\Pi_\eta(P - P\Pi_\eta)h\|_1 \le \|(P - P\Pi_\eta)h\|_1 \le \|P(h - \Pi_\eta h)\|_1 \le \|P\|_1 \|(h - \Pi_\eta h)\|_1$$

In our example, we have

$$V(f) = V(Pf) \le \alpha V(f) + B_0 \|f\|_1$$

where $B_0 = 2/9$, $\lambda = 1/9$. Then,

$$V(f) \le \frac{B_0 \|f\|_1}{1 - \alpha}$$

Also, we have

$$\|P^n\|_1 \le \|P^n\|_{BV} \le 1 + \frac{B_0}{1 - \alpha} := M = 1.25$$

then, we have

$$\|h - \Pi_\eta h\|_1 \le \eta \frac{B_0 \|f\|_1}{1 - \alpha}$$

$$\|\Pi_\eta(P - P\Pi_\varepsilon)h\|_1 \le M\eta \frac{B_0 \|f\|_1}{1 - \alpha}$$

$$\|(P_\eta - P)h\|_1 \le \eta \frac{B_0 \|f\|_1}{1 - \alpha} + M\eta \frac{B_0 \|f\|_1}{1 - \alpha} = \eta(1 + M) \frac{B_0 \|f\|_1}{1 - \alpha}$$

Thus,

$$\|h - h_\eta\|_1 \le 2 \cdot N \cdot \eta \cdot (1 + M) \frac{B_0 \|f\|_1}{1 - \alpha}$$

$\|h_\eta - \tilde{h_\eta}\|_1$**estimates** Using Interval arithmetic, follow the algorithm in [26]. We can construct Ulam's matrix in computer and its error being estimated as following:

$$\|f_\eta - \tilde{f}_\eta\|_1 \le 2N_\eta \|P_\eta - \tilde{P}_\eta\|_1 \|v\|_1 \le 4N_\varepsilon \cdot N_Z \cdot \eta,$$

where N_Z=number of non-zero element in each vector, N_η is the integer such that $\|\tilde{P}_\varepsilon^N v\|_1 \le \frac{1}{2}\|v\|_1$, ε is the maximum of the error $|\tilde{P}_\eta^N - \tilde{P}^N|$, which is estimate by interval arithmetic.

$\|\tilde{h_\eta} - h_{\eta_*}\|_1$**estimates** Ulam's matrix $\tilde{P}_\eta$ is a Markov matrix and it contracts the simplex of vectors v, with $\|v\|_1 = 1$ and $\int v = 0$. This simplex is the combination of the base. Let $Diam_1$ denote the distance induced by norm 1.

$$\begin{aligned} &Diam_1(\tilde{P_\eta}^k u) \\ &\le \max_{i,j} \|\tilde{P_\eta}^k (e_i - e_j)\|_1 \le \max_{i,j} \|\tilde{P_\eta}^k (e_1 - e_j)\|_1 + \max_{i,j} \|\tilde{P_\eta}^k (e_i - e_1)\|_1 \\ &\le 2 \max_j \|\tilde{P_\eta}^k (e_1 - e_j)\|_1. \end{aligned}$$

Thus, in the computation, we can give our program a threshold ε_{num}, it will outcome a integer k, which is the maximum iteration time makeing $\|\tilde{h_\eta} - h_{\eta_*}\|_1$ enclosed by ε_{num}.

$$\|h - h_{\eta_*}\|_1 \leq 2 \cdot N \cdot \eta \cdot (1 + M) \cdot \frac{B_0 \|f\|_1}{1 - \alpha} + 4N_\eta \cdot N_z \cdot \varepsilon + \varepsilon_{num}. \tag{2.3.15}$$

The N, N_η, ε, N_z, ε_{num} will be explained in the following list:

1. N is the integer such that $\|P_\eta^N v\|_1 \leq \frac{1}{2}\|v\|_1$, v is zero average vector. This is similar to the computation we did in Item (2).

2. N_η is the integer such that $\|\tilde{P}_\eta^{N_\eta} v\|_1 \leq \frac{1}{2}\|v\|_1$.

3. ε is the maximum of the error $|\tilde{P}_\eta^N - \tilde{P}^N|$, which is estimated using interval arithmetic.

4. N_z is the maximum number of nonzero elements in each row of the Ulam matrix $\tilde{P}_\eta^N$. Since we define Ulam matrix as $P_{ij} = m(T^{-1}(I_j) \cap I_i)/m(I_i)$, we can bound $N_z \leq \sup |T'| + 4$.

5. ε_{num} is the error we give to the program as a threshold to compute the iteration time. Then, we know for sure, after $k \in \mathbb{N}$ iterate, our numerical error will be smaller than ε_{num}.

Trial of several mesh sizes to achieve (2.3.13) **and consequently Item 3**.
We try $\eta = 1/8000000$. We obtain $N = 8, N_\eta = 8, k = 14, \varepsilon_{num} = 0.0000002, \varepsilon = 1.999958 \times 10^{-9}, N_z = 20$,

$$\|h - h_{\eta_*}\|_1 \leq 0.000002425;$$

$$(16(l_* - 1) + 8) \cdot \|\psi\|_\infty^2 \cdot \|h_{\eta_*} - h\|_1 \leq 0.00091179;$$

$$|||P_{\eta_*} - P||| \leq \eta_*(1 + \alpha);$$

$$4\|\psi\|_\infty \sum_{i=1}^{l_*-1} \sum_{j=0}^{i-1} (2\|\psi\|_\infty (B_0 + 1 + \frac{\alpha^j B_0}{1 - \alpha}) + \frac{\alpha^j (B_0 + 1 - \alpha)}{1 - \alpha} V\psi) |||P_{\eta_*} - P||| < 0.000032726.$$

Thus, desired error 999×10^{-6} has been finally achieved, and set $\eta_* = 1/8000000$. All the above iterations are summarized in Table 1.

Computation of Item (4) in Algorithm 2.2.1

$$|\sigma^2_{\eta_*, l_*} - \sigma^2| \leq 0.0001/1000 + 0.00091179 + 0.000032726 \leq 0.00094552,$$

We fix $\eta_* = \frac{10^{-6}}{8}$ and $l_* = 24$ in the computation of $\sigma^2_{\eta_*, l_*}$, and let function $\psi = x^2$ be the observable. We compute:

$$\sigma^2_{\eta_*,l_*} := \int_I \hat{\psi}^2_{\eta_*} h_{\eta_*} dm + 2\sum_{i=1}^{l_*-1} \int_I P^i_{\eta_*}(\hat{\psi}_{\eta_*} h_{\eta_*})\hat{\psi}_{\eta_*} dm \in [0.108176, 0.109107].$$

I will explain key steps in the computation of $\sigma^2_{\eta_*,l_*}$ in the following remark.

Remark 2.3.2.

1. *By the definition of $\hat{\psi}$, we have to compute $\int_I \psi h_\eta dm$ rigorously. This means we need to compute rigorously the projection of ψ on the Ulam basis. Let Π be the Ulam projection, for any observable ψ,*
$$\Pi\psi = \sum_{i=0}^{n-1} v_i \cdot \frac{\chi_{I_i}}{m(I_i)},$$
where $\{v_i\} = \{v_0, ..., v_{n-1}\}$ are the coefficients.
To compute these coefficients, we use the interval arithmetics of rigorous integration [51]
$$v_i = \int_{I_i} \psi dm.$$
Then, the integral with respect to Lebesgue measure of an observable ψ projected on the Ulam basis,
$$\int_0^1 \Pi\psi dm = \int_0^1 \sum_{i=0}^{n} v_i \frac{\chi_{I_i}}{m(I_i)} dm = \sum_i v_i.$$

2. *The computed approximation h_{η_*} is a vector that is consist of the coefficient $\{w_0, ..., w_{n-1}\}$ on Ulam projection. We compute*
$$\Pi(\psi h_{\eta_*})_i = \frac{v_i \cdot w_i}{m(I_i)}.$$
Recall the property of projection, we have $\chi^2_{I_i} = \chi_{I_i}$ and $\chi_{I_i} \cdot \chi_{I_j} = 0$ for $j \neq i$, then we have
$$\Pi(\psi \cdot h_{\eta_*})(x) = \sum_i \frac{v_i \cdot w_i}{m(I_i)} \cdot \frac{\chi_{I_i}(x)}{m(I_i)} = \sum_i v_i \cdot \frac{\chi_{I_i}(x)}{m(I_i)} \sum_i w_i \cdot \frac{\chi_{I_i}(x)}{m(I_i)} = (\Pi\psi)(x) \cdot h_{\eta_*}(x).$$
Note that h_{η_*} is constant on each I_i and equal to w_i, we have
$$(\Pi(\psi h_{\eta_*}))_i = \int_{I_i} h_{\eta_*} \psi dm = \int_{I_i} w_i \cdot \frac{\chi_{I_i}(x)}{m(I_i)} \psi dm = \frac{w_i}{m(I_i)} \cdot \int_{I_i} \psi dm = \frac{v_i \cdot w_i}{m(I_i)}.$$
Then, we have
$$\int_0^1 \psi \cdot h_{\eta_*} dm = \sum_i \frac{v_i \cdot w_i}{m(I_i)}.$$

3. *By the fact of $\Pi^2 = \Pi$, when we compute $P_{\eta_*}(\hat{\psi}_{\eta_*} h_{\eta_*})$, we can compute as following:*
$$P_{\eta_*}(\hat{\psi}_{\eta_*} h_{\eta_*}) = \Pi P \Pi(\hat{\psi}_{\eta_*} h_{\eta_*}) = \Pi P \Pi\Pi(\hat{\psi}_{\eta_*} h_{\eta_*}) = P_{\eta_*}(\Pi\hat{\psi}_{\eta_*} h_{\eta_*})$$

2.4 Another Example

Let

$$T_0(x) = S(x) + P(x) + 0.005\sin(64\pi x),$$

τ	0.001	0.001	0.001	0.001	0.001
ρ_*	0.1006	0.1006	0.1006	0.1006	0.1006
C_*	2.9342	2.9342	2.9342	2.9342	2.9342
l_*	24	24	24	24	24
η	10^{-4}	$\frac{10^{-6}}{2}$	$\frac{10^{-6}}{3}$	$\frac{10^{-6}}{4}$	$\frac{10^{-6}}{8}$
$\lvert\sigma_\eta - \sigma\rvert$	0.1531	0.0018	0.0013	0.0013	0.00091179
Check error $\lvert\sigma_\eta - \sigma\rvert < \tau$	η failed	η failed	η failed	η failed	η works
η_*	try $\eta = \frac{10^{-6}}{2}$	try $\eta = \frac{10^{-6}}{3}$	try $\eta = \frac{10^{-6}}{4}$	try $\eta = \frac{10^{-6}}{8}$	$\eta_* = \frac{10^{-6}}{8}$
$\sigma^2_{\eta_*, l_*}$	none	none	none	none	[0.108176,0.109107]

Table 2.1: Summary and output of our computations for Example 2.3.1

where

$$S(x) = \frac{31x}{1-x}$$

and $P(x)$ is the polynomial that satisfies

$$P(0) = P(1/32) = 0,$$
$$P'(0) = 32 - S'(0), \quad P'(1/32) = 32 - S'(1/32),$$
$$P''(0) = -S''(0), \quad P''(1/32) = -S''(1/32),$$
$$P'''(0) = -S'''(0), \quad P'''(1/32) = -S'''(1/32).$$

The coefficients of this polynomial are computed by inverting symbolically (therefore without numerical errors) a Vandermonde matrix. We use interval dynamics to enclose the coefficients. The computed polynomial, with rational coefficients, is given by:

$$P(x) = -\frac{1048576}{29791}x^7 - \frac{917504}{29791}x^6 - \frac{923648}{29791}x^5 - \frac{923520}{29791}x^4 - 31x^3 - 31x^2 + x.$$

Let

$$T_1(x) = 32x - 1 + 0.005 \cdot \sin(64\pi \cdot x),$$

and define

$$T(x) := \begin{cases} T_0(x - 2k/32), & x \in [2k/32, (2k+1)/32] \\ T_1(x - (2k+1)/32), & x \in [(2k+1)/32, (2k+2)/32], \end{cases} \tag{2.4.1}$$

where $k = 0, 1, \ldots, 15$. We will revisit this map in Chapter 3 to compute linear response. This is to illustrate that for smooth maps (basically C^3 circle maps) we can compute rigorously the dynamical quantities we need. We compute the diffusion coefficient for the same observable $\phi(x) = x^2$ up to a pre-specified error $\tau = 0.01$. We follow the same steps as in the previous example.

2.4.1 Item (1) in Algorithm 2.2.1

In this step, we find l_* such that item (1) of Algorithm 2.2.1 is satisfied. In particular we want to find l_* such that

$$4||\phi||_\infty \sum_{i=l_*}^{+\infty} ||P^i((\hat{\phi}\cdot h))||_{BV} \leq \frac{\tau}{256}. \tag{2.4.2}$$

As explained in Remark 2.2.1, instead of verifying item (1) to be smaller than $\frac{\tau}{2}$, we verify that it is smaller than $\frac{\tau}{256}$. This will give us more room in verifying item (2) so that the sum of the errors from both items is smaller than τ. Since the system satisfies **(A2)**, for any $g \in BV_0$ and any $k \in \mathbb{N}$, there exist $0 < \rho_* < 1$ and $C_* > 0$ such that

$$\|P^k g\|_{BV} \leq C_* \rho_*^k \|g\|_{BV}. \tag{2.4.3}$$

We want to find $0 < \rho_* < 1$ and $C_* > 0$ so that (2.4.3) is satisfied.

Once these two numbers are computed, we can easily find l_* (see (2.3.5)) so that (2.4.2) is satisfied. To compute ρ^* and C_* we follow the work of [27] whose main idea is to build a system of iterated inequalities governed by a positive matrix $\mathcal{M}$ such that:

$$\begin{pmatrix} \|P^{in_1} g\|_{BV} \\ \|P^{in_1} g\|_{L^1} \end{pmatrix} \preceq \mathcal{M}^i \begin{pmatrix} \|g\|_{BV} \\ \|g\|_{L^1} \end{pmatrix}, \tag{2.4.4}$$

where $\preceq$ means component-wise inequalities, e.g. for vectors $\overrightarrow{x} = (x_1, x_2)$ and $\overrightarrow{y} = (y_1, y_2)$, if $\overrightarrow{x} \preceq \overrightarrow{y}$, then, $x_1 \leq y_1$ and $x_2 \leq y_2$.

By using equation (2.3.6), we get that, if $||P_\varepsilon^n|_{BV_0}||_1 \leq \alpha_2$, the following inequalities are satisfied:

$$\begin{cases} \|P^{n_1} f\|_{BV} \leq \alpha^{n_1} \|f\|_{BV} + (\frac{B_0}{1-\alpha})\|f\|_1 \\ \|P^{n_1} f\|_1 \leq \alpha_2 \|f\|_1 + \varepsilon((\frac{1+\alpha}{1-\alpha})\|f\|_{BV} + B_0 n_1 (2+\alpha)\|f\|_1. \end{cases} \tag{2.4.5}$$

Using the inequalities above we have that:

$$\mathcal{M} = \begin{pmatrix} \alpha^{n_1} & \frac{B_0}{1-\alpha} \\ \varepsilon(\frac{1+\alpha}{1-\alpha}) & \varepsilon B_0 n_1 (2+\alpha) + \alpha_2 \end{pmatrix}.$$

Following the ideas of [27] we have that

$$\|P^k n_1 g\|_{BV} \leq \frac{1}{a} \rho_*^k \|g\|_{BV}, \tag{2.4.6}$$

where ρ_* is the dominant eigenvalue of $\mathcal{M}$ and (a, b) is the corresponding left eigenvector.

Thus, our main task now is to identify all the entries of the above matrix. The two constants α_2 and n_1 in $\mathcal{M}$ are two constants that give us an estimate of the speed at which P_ε contracts the space BV_0, with respect to the L^1-norm. Let P_ε be the discretized Ulam operator and fix $\alpha_2 < 1$; we want to find and $n_1 \geq 0$ such that $\forall v \in BV_0$

$$\|P_\varepsilon^{n_1} v\|_1 \leq \alpha_2 \|v\|_1 \tag{2.4.7}$$

with $\alpha_2 < 1$. We follow the idea of [26] and use the computer to estimate n_1 numerically; we refer to their paper for the algorithm used to certify n_1 and the corresponding numerical estimates and methods.

Consequently, (2.4.4) is satisfied with $n_1 = 3$, $\alpha \leq 0.033$, $B_0 \leq 0.212$, $\varepsilon = 1/1024$, $\alpha_2 = 1/25$; i.e.,

$$\mathcal{M} = \begin{pmatrix} 0.0000338 & 0.2115 \\ 0.001042 & 0.04126 \end{pmatrix}.$$

Thus, $\rho_* = 0.047$ and the eigenvector (a, b) associated to the eigenvalue ρ_* is given by $a \in [0.0221, 0.0222]$, $b \in [0.978, 0.979]$.

Thus, by (2.4.6), we obtain

$$\|P_\varepsilon^{3k} g\|_{BV} \leq (45.17) \cdot 0.047^k \|g\|_{BV}$$

Consequently we can compute $l_* \geq 15$.

Remark 2.4.1. *Using equation* (2.4.6)*, we see that, for any ϕ in BV_0 and $l_* = k \cdot n_1$ we have that:*

$$\sum_{i=l_*}^{+\infty} ||P^i(\phi)|||_{BV} \leq ||\phi||_{BV} \frac{1}{a} \cdot n_1 \sum_{i=k}^{+\infty} \rho_*^i \leq ||\phi||_{BV} \frac{1}{a} n_1 \frac{\rho_*^k}{1-\rho_*}.$$

2.4.2 Item (2) of Algorithm 2.2.1

From now on l_* is fixed and it is equal to 15. So far, we executed the first loop of the Algorithm 2.2.1; i.e.,

$$4\|\psi\|_\infty \sum_{i=15}^{\infty} \|P^i(\hat{\psi})\|_{BV} \leq \frac{\tau}{256}.$$

2.4.3 Item (3) of Algorithm 2.2.1

In this step, we have to find ε_*, a mesh size of the Ulam discretization, such that

$$\begin{aligned} &(8(l_*-1)+8) \cdot \|\psi\|_\infty^2 \cdot \|h_{\varepsilon_*} - h\|_1 \\ &+ 4\|\psi\|_\infty \sum_{i=1}^{l_*-1} \sum_{j=0}^{i-1} \left(2\|\psi\|_\infty (B_j + 1 + \frac{\alpha^j B_0}{1-\alpha}) + \frac{\alpha^j (B_0 + 1 - \alpha)}{1-\alpha} V\psi \right) |||P_{\varepsilon_*} - P||| \\ &\leq \frac{255}{256}\tau. \end{aligned} \tag{2.4.8}$$

To bound this term we need a rigorous approximation of the T-invariant density, h, in the L^1-norm; we follow the ideas (and refer for the algorithm) to the paper [26]. Set

$$\kappa := 4\|\psi\|_\infty |||P_{\varepsilon_*} - P||| \sum_{i=1}^{l_*-1} \sum_{j=0}^{i-1} \left(2\|\psi\|_\infty (B_j + 1 + \frac{\alpha^j B_0}{1-\alpha}) + \frac{\alpha^j (B_0 + 1 - \alpha)}{1-\alpha} V\psi \right).$$

Remark 2.4.2. *Note that κ gives the biggest contribution to the error. Fixed l_*, we are summing $(l_*)^2/2$ terms and the contribution of each term is bigger than $4\|\psi\|_\infty| \cdot |||P_{\varepsilon_*} - P|||$. Since ε_* can be chosen independently from l_* this is not a theoretical issue for our algorithm.*

In the following table we collected data from different mesh sizes:

ε	2^{-10}	2^{-16}	2^{-18}
$\|h_{\varepsilon_*} - h\|_1$	0.0014	$4.07 \cdot 10^{-5}$	$1.064 \cdot 10^{-5}$
$(8(l_*-1)+8) \cdot \|\psi\|_\infty^2 \cdot \|h_{\varepsilon_*} - h\|_1$	0.17	0.0049	0.00128
κ	1.57	0.0244	0.00425

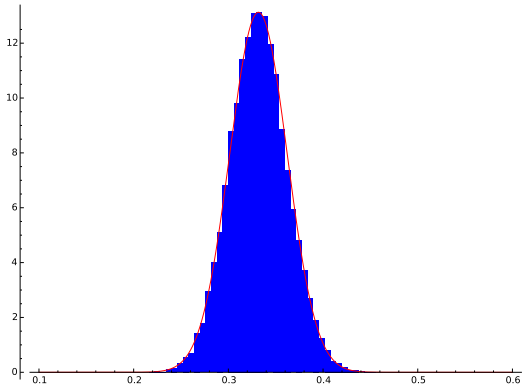

Figure 2.1: Distribution of the averages $A_{100}(x_i)$, $i = 0, \ldots, 19999$. Horizontal axis indicate the A_k value we compute, Vertical axis indicate how many times we get the same A_k

2.4.4 Item (4) in Algorithm 2.2.1

$$|\sigma^2_{\varepsilon_*,l_*} - \sigma^2| \leq 0.01/256 + 0.00128 + 0.00425 \leq 0.0056,$$

and we compute $\sigma^2_{\varepsilon_*,l_*}$

$$\sigma^2_{\varepsilon_*,l_*} := \int_I \hat{\psi}^2_{\varepsilon_*} h_{\varepsilon_*} dm + 2 \sum_{i=1}^{l_*-1} \int_I P^i_{\varepsilon_*}(\hat{\psi}_{\varepsilon_*} h_{\varepsilon_*}) \hat{\psi}_{\varepsilon_*}\, dm \in [0.092, 0.093]. \tag{2.4.9}$$

2.4.5 A non-rigorous experiment

We also perform a non-rigorous experiment to compute σ^2 in the above example. Let $\mathcal{F}_\zeta$ be the set of floating point numbers in $[0, 1]$ with ζ binary digits.

Note that the system has high entropy, so we have to be careful in our computation and choose ζ big. Due to high expansion of the system, in few iterations the ergodic average along the simulated orbit may have little in common with the orbit of the real system. So, we have to do computations with a really high number of digits ($\zeta = 2048$ binary digits).

Let $\{x_0, \ldots, x_{n-1}\}$ be n random floating points in $\mathcal{F}_\zeta$; fix k and let

$$A_k(x) = \frac{1}{k} \sum_{i=0}^{k-1} \phi(T^i(x)).$$

Let μ be an approximation of the average of ϕ with respect to the invariant measure, obtained by integrating the observable using the approximation of the invariant density:

$$\mu \in [0.33175, 0.33176].$$

Now, for each point $\{x_0, \ldots, x_n\}$ we compute the value $A_k(x_0), \ldots, A_k(x_m)$ and from these the following two estimators:

$$\tilde{\mu} = \frac{1}{n} \sum_{i=0}^{n-1} A_k(x_i)$$

$$\tilde{\sigma}^2 = \frac{1}{n} \sum_{i=0}^{n-1} \frac{(k \cdot A_k(x_i) - k\mu)^2}{k}.$$

In the table the result of an experiment with $n = 20000$ is given. In Fgure 2.1, a histogram plot of the distribution of $A_k(x_i)$ for $k = 100$, $n = 20000$; in red we have the normal distribution with average μ and variance $\sigma^2_{\varepsilon_*,l_*}/\sqrt{k}$.

k	$\tilde{\mu}$	$\tilde{\sigma}^2$
90	$[0.33162, 0.33163]$	$[0.09245, 0.09245]$
95	$[0.33155, 0.33156]$	$[0.09259, 0.0926]$
100	$[0.33157, 0.33158]$	$[0.09236, 0.09237]$

The out-put of this non-rigourous experiment is in line with the output from our rigorous computation in (2.4.9).

Chapter 3

A rigorous computational approach for linear response

In this chapter we provide suitable discretization schemes that can be used to approximate linear response (the derivative of an invariant density with respect to noise). The problem that we face in our rigorous approximation is is two-fold. The first is functional analytic. In particular, we need to find suitable discretization schemes that preserve the regularity of the function space(s) where the transfer operator acts, and which can approximate the original transfer operator. The second is computational. In particular, the computational approach should be amenable to tracking all the round-off errors made by the computer. This chapter is based on our work in [6]. We first start by a theorem that proves linear response in a general setting.

3.1 Differentiation of invariant densities

We present a general setting in which the formula corresponding to the derivative of a fixed point[1] of a family of positive operators P_ϵ can be obtained[2]. We consider the action of the operators on different spaces. Let B_w, B_s, B_{ss} denote abstract Banach spaces of Borel measures on X equipped with norms $||\ ||_w, ||\ ||_s, ||\ ||_{ss}$, respectively, such that $||\ ||_w \leq ||\ ||_s \leq ||\ ||_{ss}$. We suppose that P_ϵ, $\epsilon \geq 0$, has a unique fixed point $h_\epsilon \in B_{ss}$. Let $P := P_0$ be the unperturbed operator and $h \in B_{ss}$ be its fixed point. Let $V_s^0 = \{v \in B_s, v(X) = 0\}$, $V_w^0 = \{v \in B_w, v(X) = 0\}$.

The following proposition is essentially proved in [42]. Since we adapted the assumptions to a general setting we include a proof.

Proposition 3.1.1. *Suppose that the following assumptions hold: [6]*

[1] In applications to dynamical systems, such a fixed point corresponds to the density of an absolutely continuous invariant measure.

[2] The differentiation is done with respect to the variable ε in a suitable norm. This will be clear in the statement of Proposition 3.1.1 below.

1. *The norms* $||P^k||_{B_w \to B_w}$ *and* $||P_\epsilon^k||_{B_w \to B_w}$ *are uniformly bounded with respect to* k *and* $\epsilon > 0$.

2. P_ϵ *is a perturbation of* P *in the following sense*
$$||P_\epsilon - P||_{B_s \to B_w} \leq C\epsilon. \tag{3.1.1}$$

3. *The operators* P_ϵ, $\varepsilon \geq 0$, *have a uniform rate of contraction on* V_s^0*: there are* $C_1 > 0$, $0 < \rho < 1$, *such that*
$$||P_\epsilon^n||_{V_s^0 \to B_s} \leq C_1 \rho^n. \tag{3.1.2}$$

4. *There is an operator* $\hat{P} : B_{ss} \to B_s$ *such that*
$$\lim_{\epsilon \to 0} ||\epsilon^{-1}(P - P_\epsilon)f - \hat{P}f||_s = 0 \ \forall f \in B_{ss}. \tag{3.1.3}$$

Let
$$\hat{h} = (Id - P)^{-1}\hat{P}h.$$
Then
$$\lim_{\epsilon \to 0} ||\epsilon^{-1}(h - h_\epsilon) - \hat{h}||_w = 0; \tag{3.1.4}$$
i.e. $\hat{h}$ *represents the derivative of* h_ϵ *with respect to* ϵ.

Before the proof, we state what we mean by 'linear response'.

Remark 3.1.1. *We call* $\hat{h}$ *in* (3.1.4) *the derivative of* h_ϵ *with respect to* ϵ. *In particular, in the context of dynamical systems; i.e., if* P *is a Perron-Frobenious operator associated with a map, whenever such a limit exists, we say the system has linear response.*

Proof. Notice that by its definition, $\hat{P}h \in V_s^0$. Let h_ϵ be the invariant density of P_ϵ, such that $h_\epsilon = P_\epsilon h_\epsilon$. We have
$$(Id - P_\epsilon)(h_\epsilon - h) = (P_\epsilon - P)h$$
and since $\hat{h} = (Id - P)^{-1}\hat{P}h$, we obtain
$$\begin{aligned}
\lim_{\epsilon \to 0} ||\epsilon^{-1}(h_\epsilon - h) - \hat{h}||_w &= \lim_{\epsilon \to 0} ||\epsilon^{-1}(Id - P_\epsilon)^{-1}(P_\epsilon - P)h - (Id - P)^{-1}\hat{P}h||_w \\
&\leq \lim_{\epsilon \to 0} ||(Id - P_\epsilon)^{-1}[\epsilon^{-1}(P_\epsilon - P)h - \hat{P}h]||_w \\
&\qquad + \lim_{\varepsilon \to 0} ||(Id - P_\epsilon)^{-1}\hat{P}h - (Id - P)^{-1}\hat{P}h||_w \\
&:= (I) + (II).
\end{aligned} \tag{3.1.5}$$
Notice that by assumption (3), $||(Id - P_\epsilon)^{-1}||_{V_s^0 \to B_w}$ are uniformly bounded. Moreover, since $\lim_{\epsilon \to 0} ||\epsilon^{-1}(P_\epsilon - P)h - \hat{P}h||_s = 0$, we obtain
$$(I) = \lim_{\epsilon \to 0} ||(Id - P_\epsilon)^{-1}[\epsilon^{-1}(P_\epsilon - P)h - \hat{P}h||_w = 0.$$
Now we consider (II). By assumption (3), on the space V_1^0, $(Id - P_\epsilon)^{-1} = \sum_0^\infty P_\epsilon^k$. Notice that by assumptions (2) and (3) we have:
$$||P^k - P_\epsilon^k||_{V_s^0 \to V_w^0} \leq \sum_{j=0}^{k-1} ||P_\epsilon^j (P_\epsilon - P) P^{k-1-j}||_{V_s^0 \to V_w^0}$$

$$\begin{aligned} &\leq \sup_j ||P_\epsilon^j||_w \sum_{j=0}^{k-1} ||(P_\epsilon - P)P^{k-1-j}||_{V_s^0 \to V_w^0} \\ &\leq \epsilon \sup_j ||P_\epsilon^j||_w \sum_{j=0}^{k-1} ||P^{k-1-j}||_{V_s^0} \\ &\leq \epsilon \sup_j ||P_\epsilon^j||_w C_1 \frac{1-\rho^k}{1-\rho}. \end{aligned}$$

Consequently,

$$\begin{aligned} &||(Id - P_\epsilon)^{-1}\hat{P}h - (Id-P)^{-1}\hat{P}h||_w \\ &\leq ||\hat{P}h||_s [\sum_{k=0}^{l-1} ||P^k - P_\epsilon^k||_{V_s^0 \to V_w^0} + \sum_{l}^{\infty} (||P^k||_{V_s^0 \to V_w^0} + ||P_\epsilon^k||_{V_s^0 \to V_w^0})] \\ &\leq ||\hat{P}h||_s [\epsilon \sup ||P_\epsilon^j||_w C_1 \frac{1-\rho^l}{1-\rho} + 2C_1 \rho^l \frac{1}{1-\rho}]. \end{aligned}$$

Choosing $l = \lceil |\log \epsilon| \rceil$ implies

$$(II) = \lim_{\epsilon \to 0} ||(Id - P_\epsilon)^{-1}\hat{P}h - (Id-P)^{-1}\hat{P}h||_w = 0.$$

Hence, $\lim_{\epsilon \to 0} ||\epsilon^{-1}(h_\epsilon - h) - \hat{h}||_w = 0$. □

The operator $\hat{P}$ depends on the kind of perturbation we consider (deterministic, stochastic, etc.). In the following, we suppose that $\hat{P}h$ is computable with a small error in the B_s norm. Then we show that this leads to the rigorous computation of $\hat{h}$ in the B_w norm. The computation will be performed by approximating P with a finite rank operator P_η which can be implemented on a computer. Let us consider a finite rank discretization

$$\Pi_\eta : B_s \to W_\eta,$$

where $W_\eta \subseteq B_s$ is a finite dimensional space of measures, such that for $f \in B_s$,

$$\lim_{\eta \to 0} ||(\Pi_\eta - Id)f||_w = 0.$$

The operator P_η is defined as

$$P_\eta = \Pi_\eta P \Pi_\eta.$$

Let us denote by $f_\eta \in V_s^0$ a family of approximations of $\hat{P}h$ in the weak norm $||\cdot||_w$.

Theorem 3.1.2. *Suppose that: [6]*

1. *$||f_\eta||_s$ are uniformly bounded and $||f_\eta - \hat{P}h||_w \underset{\eta \to 0}{\to} 0$.*
2. *P_η is an approximation of P in the following sense*
$$||P_\eta - P||_{B_s \to B_w} \leq C\eta.$$
3. *The operators P_η, $\eta > 0$, have uniform rate of contraction on V_s^0: there are $C_2 > 0$, $0 < \rho_2 < 1$, such that*
$$||P_\eta^n||_{V_s^0 \to B_s} \leq C_2 \rho_2^n. \tag{3.1.6}$$

Then, for any $\tau > 0$, there are $\eta > 0$ and $l^* \in \mathbb{N}$ such that

$$||(Id - P)^{-1}\hat{P}h \ - \sum_{k=0}^{l^*-1} P_\eta^k f_\eta||_w < \tau.$$

Proof. Notice that $(Id - P)^{-1}\hat{P}h$ is well defined since $\hat{P}h$ is of zero average. We have

$$\begin{aligned}
\sum_{k=0}^{l^*-1} ||(P^k - P_\eta^k)f_\eta||_w &\leq \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1} ||P^j(P - P_\eta)P_\eta^{k-1-j} f_\eta||_w \\
&\leq M \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1} ||(P - P_\eta)P_\eta^{k-1-j} f_\eta||_w \\
&\leq M||(P - P_\eta)||_{B_s \to B_w} \cdot \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1} ||P_\eta^{k-1-j} f_\eta||_{B_s}, \qquad (3.1.7)
\end{aligned}$$

where $M = \sup_k ||P^k||_{B_w \to B_w}$. Consequently, we obtain

$$\begin{aligned}
||(Id - P)^{-1}\hat{P}h - \sum_{k=0}^{l^*-1} \mathcal{L}_\eta^k f_\eta||_w &= ||\sum_{k=0}^{\infty} P^k\hat{P}h - \sum_{k=0}^{l^*-1} P_\eta^k f_\eta||_w \\
\leq ||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w &+ ||\sum_{k=0}^{l^*-1} P^k\hat{P}h - \sum_{k=0}^{l^*-1} P_\eta^k f_\eta||_w \\
\leq ||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w &+ \sum_{k=0}^{l^*-1} ||(P^k - P_\eta^k)f_\eta||_w + \sum_{k=0}^{l^*-1} ||P^k(\hat{P}h - f_\eta)||_w \qquad (3.1.8) \\
\leq ||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w &+ M||(P - P_\eta)||_{B_s \to B_w} \cdot \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1} ||P_\eta^{k-1-j} f_\eta||_s \\
&+ M\ l^*||\hat{P}h - f_\eta||_w.
\end{aligned}$$

Now, choose l^* big enough so that $||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w \leq \frac{\tau}{2}$. Since for each η $||P_\eta^{k-1-j} f_\eta||_s$ are uniformly bounded, by assumptions (2) and (3) we can choose η small enough such that

$$M||(P - P_\eta)||_{B_s \to B_w} \cdot \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1} ||P_\eta^{k-1-j} f_\eta||_s + M\ l^*||\hat{P}h - f_\eta||_w < \frac{\tau}{2}. \qquad (3.1.9)$$

□

Remark 3.1.2. *For computational purposes it is important to have an algorithm to find first l^* and then η. Let us comment on each summand in Equation (3.1.8):*

1. *The first summand of (3.1.8), $||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w$ can be estimated by (3.1.2). However, it is enough to have an estimation on the weak norm. In Section 3.2 (see also Section A.0.10) we will see how to find in systems satisfying a Lasota-Yorke inequality, upper bounds C_3, ρ_3 of C_2 and ρ_2, such that: $||P^k\hat{P}h||_w \leq C_3\rho_3^k||\hat{P}h||_s$. Once the constants are found, we can bound $||\sum_{k=l^*}^{\infty} P^k\hat{P}h||_w \leq \frac{C_2\rho_2^{l^*}||\hat{P}f||_s}{1-\rho_2}$ and find a suitable l^* to make this summand as small as wanted.*

2. *For the second summand of (3.1.8)*
$$M||(P-P_\eta)||_{B_s\to B_w}\cdot\sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1}||P_\eta^{k-1-j}f_\eta||_s$$
we need an estimate on M which can be recovered by a Lasota-Yorke inequality (see Proposition 3.2.4). $||(P-P_\eta)||_{B_s\to B_w}$ will be estimated by condition (2) of Theorem 3.1.2. The summands $||P_\eta^{k-1-j}f_\eta||_s$ can be approximated by the fact that P_η is of finite rank; i.e., by computing the matrix representing it. $||P_\eta^{k-1-j}f_\eta||_s$ will be estimated by the computer.

3. *For M $l^*||\hat{P}h-f_\eta||_w$ of (3.1.8), we have to find a suitable approximation of $\hat{P}h$ such that $||\hat{P}h-f_\eta||_w$ is as small as wanted. Note that this depends on the properties of $\hat{P}$ and consequently on the kind of perturbations P_ϵ that we consider.*

In the following we will discuss in details how the above results can be applied for a certain class of maps. We also present an example of the computer implementation based on the algorithm of Remark 3.1.2.

3.2 Expanding circle map and random perturbations

Let $(\mathbb{T},\mathfrak{B},m)$ be the measure space where $\mathbb{T}$ is the unit circle, $\mathfrak{B}$ is Borel σ-algebra and m is the Lebesgue measure on $\mathbb{T}$. Let $C^k(\mathbb{T})$ be the space of continuous functions with k continuous derivatives, equipped with the norm $||f||_{C^k}=\sum_{i=0}^k||f^{(i)}||_\infty$, where $||g||_\infty=\max_{x\in\mathbb{T}}|g(x)|$. Let $T:\mathbb{T}\to\mathbb{T}$ be a C^3 uniformly expanding circle map; i.e. $\inf_{x\in\mathbb{T}}|D_xT|>1$. Let
$$\lambda=1/\inf_{x\in\mathbb{T}}|D_xT|.$$
Without loss of generality we assume that T is orientation preserving. Recall that the transfer operator associated with T (the Perron-Frobenius operator, which we still denote as $P:L^1\to L^1$,)
$$Pf(x)=\sum_{y=T^{-1}x}\frac{f(y)}{|T'(y)|}.$$
It is well known that such a transfer operator satisfies several Lasota-Yorke inequalities (see for instance [42]). In particular, the following proposition holds:

Proposition 3.2.1. *$\exists\ A_1,A_2>0$ and $0<\beta<1$, such that for any $f\in C^2(\mathbb{T})$ and $n\in\mathbb{N}$ we have [6]*

1.
$$||P^nf||_{C^1}\le A_1\beta^n||f||_{C^1}+A_2||f||_\infty,$$

2.
$$||P^nf||_{C^2}\le A_1\beta^{2n}||f||_{C^2}+A_2||f||_{C^1}.$$

The above inequalities imply (see e.g. [7]) along with the properties of the system, that P has a spectral gap on $C^1(\mathbb{T})$ and on $C^2(\mathbb{T})$. Moreover, 1 is a simple dominant eigenvalue. In particular, this implies that T admits a unique invariant density $h \in C^2(\mathbb{T})$ and the system $(T, \mathbb{T}, \mu)$, where $\mu := h \cdot m$, is mixing.

3.2.1 A stochastic perturbation

Let $\varepsilon \in (0,1)$. For $f \in L^\infty(\mathbb{T})$ let K_ε denote the operator defined as:

$$K_\varepsilon f(x) = \int_{\mathbb{T}} \varepsilon^{-1} j(\varepsilon^{-1}(x-y)) f(y) dy,$$

where $j \in C^\infty(\mathbb{R}, \mathbb{R}^+)$, $\text{supp}(j) \subset [-1/2, 1/2]$ and $\int_{\mathbb{R}} j(y)dy = 1$.

Lemma 3.2.2 (Properties of K_ϵ [6])**.**

1. *For $f \in C^k$, $k \in \{1,2\}$*
$$||K_\epsilon f||_{C^k} \leq ||f||_{C^k};$$
2. *for $f \in C^1(\mathbb{T})$*
$$||K_\epsilon f - f||_\infty \leq \epsilon ||f||_{C^1};$$
3. *for $f \in C^2(\mathbb{T})$*
$$||K_\epsilon f - f||_{C^1} \leq \epsilon ||f||_{C^2}.$$

Proof. The first assertion is a standard property of a convolution. For (2), we have

$$|K_\epsilon f(x) - f(x)| = |\int_{\mathbb{T}} \epsilon^{-1} j(\epsilon^{-1}(x-y)) f(y) dy - f(x)|$$

By $\int_{\mathbb{R}} j(y)dy = 1$ and mean value theorem,

$$= |\int_{\mathbb{T}} \epsilon^{-1} j(\epsilon^{-1}(x-y))(f(y) - f(x)) dy| \leq \epsilon ||f'||_\infty.$$

Since the support of j is contained in $[-\epsilon/2, \epsilon/2]$. To prove (3), observe that

$$\frac{\partial}{\partial x} j(\epsilon^{-1}(x-y)) = -\frac{\partial}{\partial y} j(\epsilon^{-1}(x-y)).$$

Therefore,

$$|(K_\epsilon f(x))' - f'(x)| = |\int_{\mathbb{T}} \epsilon^{-1} j(\epsilon^{-1}(x-y)) \frac{\partial}{\partial y} f(y) dy - f'(x)|.$$

Using integration by parts and the compactness of the support of j, we obtain:

$$|(K_\epsilon f(x))' - f'(x)| \leq \epsilon ||f''||_\infty.$$

□

We now define a family of operators by setting

$$P_\varepsilon := K_\varepsilon P.$$

and define

$$\hat{P} f := \gamma (Pf)' \quad \gamma := \int j(\xi)|\xi| d\xi. \tag{3.2.1}$$

Proposition 3.1.1 can be applied to get a linear response formula for the perturbations P_ε using the function spaces $B_w = L^\infty(\mathbb{T})$, $B_s = C^1(\mathbb{T}), B_{ss} = C^2(\mathbb{T})$. Let us check that the required assumptions hold.

From (3.2.1) one obtains (see [42])

$$\lim_{\epsilon\to 0} ||\epsilon^{-1}(P - P_\epsilon)f - \hat{P}f||_{C^1} = 0 \ \forall f \in C^2(\mathbb{T});$$

i.e. assumption (4) of Proposition 3.1.1 is satisfied. For the other assumptions we refer to the following remark.

Remark 3.2.1. *By Item (1) of Lemma 3.2.2 it follows that P_ϵ and P, satisfy a uniform Lasota-Yorke inequality. This implies that assumption (1) of Proposition 3.1.1 holds. Moreover, by the stability result of [37] (see Theorem 1.1.14) for sufficiently small $\epsilon > 0$, assumption (3) of Proposition 3.1.1 holds. Finally, by Item (2) of Lemma 3.2.2 we obtain the approximation assumption (Item (2)) of Proposition 3.1.1.*

Thus, by Proposition 3.1.1, the linear response holds:

$$\lim_{\epsilon\to 0} ||\frac{h_\epsilon - h}{\epsilon} - \hat{h}||_\infty = 0, \tag{3.2.2}$$

where $\hat{h} := (\mathrm{Id} - P)^{-1}\hat{P}h$.

Remark 3.2.2. *Our aim is to compute $\hat{h}$ up to a pre-specified error. This will be done by Theorem 3.1.2. We note that in Theorem 3.1.2 the approximation of P and the approximation of $\hat{P}h$ require different assumptions to be satisfied. Thus, for this purpose we will use two different discretization schemes (see Subsection 3.2.3) and small modifications of the spaces $L^\infty(\mathbb{T}), C^1(\mathbb{T}), C^2(\mathbb{T})$ to achieve our goal.*

3.2.2 Modified function spaces

To compute $\hat{h}$, we apply Theorem 3.1.2 by using the function spaces $B_w = L^\infty(\mathbb{T})$, $B_s = \tilde{C}^1(\mathbb{T}), B_{ss} = \tilde{C}^2(\mathbb{T})$ where the two latter spaces are essentially the spaces $C^k(\mathbb{T})$, $k = 1, 2$, allowing a discontinuity at zero. For $k \in \{1, 2\}$,

$$\tilde{C}^k(\mathbb{T}) = \{f \mid f(x) = a + \int_0^x g(\xi)d\xi, g \in C^{k-1}(\mathbb{T}), a \in \mathbb{R}\}.$$

Let

$$J(f) = |\lim_{x\to 0^-} f(x) - \lim_{x\to 0^+} f(x)| = |\int_0^1 f'(\xi)d\xi|.$$

We define the following norms

$$||f||_{\tilde{C}^1} = J(f) + ||f||_\infty + ||f'(x)||_\infty,$$
$$||f||_{\tilde{C}^2} = ||f||_{\tilde{C}^1} + ||f''(x)||_\infty.$$

Working with the above spaces simplifies the computational part of our work. In particular it allows us to use discretization schemes that can be easily implemented on the computer and will be described in the following sections.

Lemma 3.2.3. *We have [6]*

1. *P preserves $\tilde{C}^k(\mathbb{T})$, $k \in \{1, 2\}$;*

2. *$J(Pf) \leq \lambda J(f)$.*

Proof. Since T is an expanding circle map, it can be conjugated to a full branched expanding map on the interval which is orientation preserving. We can suppose $T(0) = 0$. Therefore Pf has a discontinuity at 0 and it is C^k otherwise. To prove the second statement, let us denote by d_i the preimages of 0 that are contained inside the interval $(0, 1)$. By continuity of f on $(0, 1)$ we have

$$\lim_{x\to 0^+} Pf(x) - \lim_{x\to 0^-} Pf(x) = \frac{1}{T'(0)}(\lim_{x\to 0^+} f(x) - \lim_{x\to 0^-} f(x)) + \sum_i \frac{f(d_i)}{T'(d_i)} - \sum_i \frac{f(d_i)}{T'(d_i)}.$$

□

Before introducing our discretization schemes, we state Lasote-Yorke inequalities for P when acting on $\tilde{C}^1(\mathbb{T}), \tilde{C}^2(\mathbb{T})$. These inequalities are useful to show that P and P_η, which will be defined in (3.2.5) below, satisfy the assumptions of Theorem 3.1.2 and approximates P as an operator from $\tilde{C}^1$ to L^∞, and to show that $\tilde{P}_\eta$, which is defined in (3.2.11) below, approximates P as an operator from $\tilde{C}^2$ to $\tilde{C}^1$. Since these inequalities will be used in the computer implementation, we also give estimates for the constants involved. For the proof of Proposition 3.2.4, see Section A.0.7 in the appendix.

Proposition 3.2.4. *[6]*

1. Let $M := \sup_n ||P^n||_{L^\infty \to L^\infty}$. Then
$$M \leq 1 + \frac{B}{1-\lambda},$$
where $\lambda := (\inf_{x\in\mathbb{T}} |D_x T|)^{-1} < 1$ and $B = ||T''/(T'^2)||_\infty$.

2. For $f \in \tilde{C}^1(\mathbb{T})$ we have
$$||Pf||_{\tilde{C}^1} \leq \lambda M ||f||_{\tilde{C}^1} + C||f||_\infty,$$
where
$$C = \lambda ||\frac{T''}{(T')^2}||_\infty + (1-\lambda)M.$$

3. For $f \in \tilde{C}^2(\mathbb{T})$ we have
$$||Pf||_{\tilde{C}^2} \leq \lambda^2 M ||f||_{\tilde{C}^2} + D||f||_{\tilde{C}^1}.$$
where
$$D = \lambda M + C + 3\max\{1, \left|\left|\frac{T''}{(T')^2}\right|\right|_\infty^2\} M + M\left|\left|\frac{T'''}{(T')^3}\right|\right|_\infty.$$

The above inequalities, along with the properties of the system, imply that P has a spectral gap on $\tilde{C}^1(\mathbb{T})$ and on $\tilde{C}^2(\mathbb{T})$. Moreover, 1 is a simple dominant eigenvalue. In particular, this implies that T admits a unique invariant density h in $\tilde{C}^2(\mathbb{T})$ and the system $(T, \mathbb{T}, \mu)$, where $\mu := h \cdot m$, is mixing.

3.2.3 Finite rank approximations of P

We introduce two finite rank approximations of P which will be called P_η and $\tilde{P}_\eta$. As noted before (Remark 3.2.2)[3] both operators are needed, and in some sense complementary, to achieve the rigorous approximation of the linear response.

An approximation of P as an operator from $\tilde{C}^1 \to L^\infty$.

We start by defining a suitable partition of unity. Let us consider the partition of unity $\{\phi_i\}_{i=0}^m$ defined in the following way: for $i = 1, \dots, m-1$, let $a_i = i/m$, and let $a_0 = 0$, $a_m = 1$. For $i = 0, \dots, m$ set

$$\phi_i(x) = \phi(m \cdot x - i), \tag{3.2.3}$$

where

$$\phi(x) = \begin{cases} 1 - 3x^2 - 2x^3 & x \in [-1, 0] \\ 1 - 3x^2 + 2x^3 & x \in [0, 1] \end{cases}. \tag{3.2.4}$$

Note that for $i = 0$ and $i = m$, the bump function is restricted to half of its support. Also note that $\phi_i(a_j) = \delta_{ij}$ (where $\delta_{ij} = 1$ if $i = j$, 0 in all the other cases) and that $||\phi_i||_\infty = 1$, $||\phi_i'(x)||_\infty = 3m/2$. To ensure that our discretization preserves integrals, we define an auxiliary function $\kappa(x)$ in the following way. Let $\{\tilde{\phi}_j\}_{j=0}^{2m}$ be the partition of unity associated to the partition of size $1/2m$ and let

$$\kappa(x) = 2 \cdot \sum_{j=0}^{m-1} \tilde{\phi}_{2j+1}(x).$$

Note that $\kappa(a_i) = 0$ for all i and that $\int_0^1 \kappa dm = 1$, $||\kappa||_\infty = 2$, $||\kappa'||_\infty = 3m$. Set $\eta := 1/m$ and define

$$\Pi_\eta(f)(x) := \sum_i f(a_i) \cdot \phi_i(x) + \left(\int_0^1 f dm - \sum_{i=0}^m f(a_i) \int_0^1 \phi_i dm \right) \kappa(x).$$

We set

$$P_\eta := \Pi_\eta P \Pi_\eta. \tag{3.2.5}$$

We now prove properties of Π_η that will be used to verify the assumptions of Theorem 3.1.2.

Lemma 3.2.5. *For $f \in \tilde{C}^1(\mathbb{T})$, supposing $m > 2$, we have [6]*

1. $||\Pi_\eta f||_\infty \leq 5||f||_\infty$;
2. $||\Pi_\eta f||_\infty \leq ||f||_\infty + 2\frac{||f'||_\infty}{m}$;
3. $||\Pi_\eta f||_{\tilde{C}^1} \leq \frac{11}{2}||f||_{\tilde{C}^1}$;
4. $||\Pi_\eta f - f||_\infty \leq 3\frac{||f'||_\infty}{m}$;

[3]The reason behind this 'two step' approximation is also detailed in Section 3.2.5.

Proof. The following approximation inequality holds:

$$\begin{aligned} |f(x) - \sum f(a_i)\phi_i(x)| &= |\sum_i (f(x) - f(a_i))\phi_i(x)| \\ &= |\sum_i f'(\xi_i)(x - a_i)\phi_i(x)| \leq \frac{||f'||_\infty}{m}. \end{aligned} \tag{3.2.6}$$

This implies that

$$\left| \int_0^1 f dm - \sum_{i=0}^m f(a_i) \int_0^1 \phi_i dm \right| \leq \frac{||f'||_\infty}{m}. \tag{3.2.7}$$

By (3.2.7), we have

$$\begin{aligned} ||\Pi_\eta f||_\infty &= \max_{x\in[0,1]} |\sum_{i=0}^n f(a_i)\phi_i(x)| + 2|\int f(x) - \sum_{i=0}^n f(a_i)\phi_i(x)dx| \\ &\leq ||f||_\infty |\sum_i \phi_i(x)| + 2|\int f(x) - \sum_{i=0}^n f(a_i)\phi_i(x)dx|, \end{aligned} \tag{3.2.8}$$

which implies (1) and (2) of the lemma. We now prove (3). First, since the $\{\phi_i\}_{i=0}^m$ is a partition of unity, we have

$$\sum_{i=0}^n \phi_i'(x) = 0.$$

Therefore

$$\begin{aligned} ||(\Pi_\eta f)'||_\infty &\leq |\sum_{j=0}^{m/2} f(a_{2j}) - f(a_{2j+1})\phi_j'(x)| + 3m|\int f(x) - \sum_{i=0}^n f(a_i)\phi_i(x)dx| \\ &\leq \frac{3}{2}\max_j |\frac{f(a_{2j}) - f(a_{2j+1})}{\eta}| + 3m\frac{||f'||}{m} \leq \frac{9}{2}||f'||_\infty. \end{aligned} \tag{3.2.9}$$

Further,

$$J(\Pi_\eta f) = |f(1) - f(0)| \leq ||f'||_\infty. \tag{3.2.10}$$

Therefore (3) follows from (3.2.8), (3.2.9) and (3.2.10). Also note that (4) of the lemma follows from (3.2.6) and (3.2.7). □

Remark 3.2.3. *By (4) Lemma 3.2.5 assumption (2) of Theorem 3.1.2 is satisfied. By the Lasota-Yorke inequalities*[4] *given in Proposition 3.2.4, and a similar reasoning as in Remark 3.2.1, P and P_η satisfy assumption (3) of Theorem 3.1.2.*

An approximation of P as an operator from $\tilde{C}^2 \to \tilde{C}^1$.

Let $\{\phi_i\}_{i=0}^m$ be the partition of unity defined in (3.2.3); in the following we will denote by $\eta := 1/m$. Define:

$$\tilde{g}(x) := \sum_i f'(a_i) \cdot \phi_i(x).$$

Let

$$I_i = \int_0^1 \left(\int_0^x \phi_i(\xi)d\xi \right) dx$$

and

$$I(f) = \sum_{i=0}^m f'(a_i) I_i.$$

[4]See also the Appendix for a proof of a uniform Lasota-Yorke inequality of P and P_η on $\tilde{C}^1$.

Let $f(0^+) := \lim_{x\to 0^+} f(x)$. We have

$$f(0^+) + I(f) = \int_0^1 \left(f(0^+) + \int_0^x \tilde{g}(\xi)d\xi \right) dx.$$

We define the operator

$$\tilde{\Pi}_\eta(f)(x) := \left(\int f dx - I(f) \right) + \int_0^x \tilde{g}(\xi)d\xi,$$

Note that

$$(\tilde{\Pi}_\eta f)'(x) = \tilde{g}(x), \quad (\tilde{\Pi}_\eta f)''(x) = \tilde{g}'(x).$$

We set

$$\tilde{P}_\eta := \tilde{\Pi}_\eta P \tilde{\Pi}_\eta. \tag{3.2.11}$$

Remark 3.2.4. *An interesting property of $\tilde{\Pi}_\eta$ is the following. Suppose we have a function f such that:*

$$f(x) = a + \int_0^x \sum_{i=0}^n b_i \cdot \phi_i(\xi)d\xi;$$

then

$$f'(x) = \sum_{i=0}^n b_i \cdot \phi_i(x).$$

This means that, if we find a fixed point of $\tilde{P}_\eta$, we can easily find its derivative with respect to the basis $\{\phi_i(x)\}$.

We now prove properties of $\tilde{\Pi}_\eta$. We start with a preparatory lemma.

Lemma 3.2.6. *Let f be in $\tilde{C}^2(\mathbb{T})$, then [6]*

1. $|f'(x) - \tilde{g}(x)| \leq \frac{||f''||_\infty}{m}$;
2. $|\int f dx - I(f)| \leq ||f||_\infty + 2||f'||_\infty$;
3. $|\int f dx - I(f)| \leq ||f||_\infty + \frac{||f''||}{m}$.

Proof. For (1), we have

$$|f'(x) - \tilde{g}(x)| = |\sum_i (f'(x) - f'(a_i))\phi_i(x)| \leq \frac{||f''||_\infty}{m}.$$

Moreover,

$$\begin{aligned}
\int f dx - I(f) &= f(0^+) + \int f dx - f(0^+) - I(f) \\
&= f(0^+) + \int_0^1 f dx - \int_0^1 \left(f(0^+) + \int_0^x \tilde{g}(\xi)d\xi \right) dx \\
&= f(0^+) + \int_0^1 f(x) - \left(f(0^+) + \int_0^x \tilde{g}(\xi)d\xi \right) dx \\
&= f(0^+) + \int_0^1 \int_0^x f'(\xi) - \tilde{g}(\xi)d\xi dx.
\end{aligned}$$

From the last equality and (1) we get (2) and (3). □

Lemma 3.2.7. *For $f \in \tilde{C}^1(\mathbb{T})$ we have [6]*

1. $||\tilde{\Pi}_\eta f||_{\tilde{C}^1} \leq 3||f||_{\tilde{C}^1} + 2||f'||_\infty \leq 5||f||_{\tilde{C}^1}$.

   Moreover, for $f \in \tilde{C}^2(\mathbb{T})$, we have

2. $||\tilde{\Pi}_\eta f||_\infty \leq ||f||_\infty + ||f'||_\infty + \eta||f''||_\infty$;
3. $||\tilde{\Pi}_\eta f||_{\tilde{C}^2} \leq 3||f||_{\tilde{C}^1} + 2||f'||_\infty + \frac{3}{2}||f''||_\infty \leq 5||f||_{\tilde{C}^2}$;
4. $||\tilde{\Pi}_\eta f - f||_\infty \leq \frac{||f''||_\infty}{m}$;
5. $||\tilde{\Pi}_\eta f - f||_{\tilde{C}^1} \leq \frac{3||f''||_\infty}{m}$;
6. $\int_0^1 \tilde{\Pi}_\eta f dx = \int_0^1 f dx$.

Proof. We have, from Lemma 3.2.6, item 2

$$\begin{aligned} ||\tilde{\Pi}_\eta f||_\infty &= \sup_{x\in(0,1)} |\int f dx - I(f) + \int_0^x \tilde{g}\ d\xi| \\ &\leq ||f||_\infty + 2||f'||_\infty + \int_0^x ||f'||_\infty d\xi \leq 3||f||_{\tilde{C}^1}, \end{aligned} \tag{3.2.12}$$

and

$$||(\tilde{\Pi}_\eta f)'||_\infty = \sup_{x\in(0,1)} |\sum_i f'(a_i)\phi_i(x)| \leq ||f'||_\infty |\sum_i \phi_i(x)| = ||f'||_\infty. \tag{3.2.13}$$

Similarly, from Lemma 3.2.6, item 3 we get:

$$\begin{aligned} ||\tilde{\Pi}_\eta f||_\infty &= \sup_{x\in(0,1)} |\int f dx - I(f) + \int_0^x \tilde{g}\ d\xi| \\ &\leq ||f||_\infty + \frac{||f''||}{m} + ||f'||_\infty. \end{aligned}$$

Moreover,

$$J(\tilde{\Pi}_\eta f) = |\int_0^1 \sum_i f'(a_i)\phi_i(x)dx| \leq ||f'||_\infty \int_0^1 |\sum \phi_i(x)dx| \leq ||f'||_\infty. \tag{3.2.14}$$

Therefore (1) follows from (3.2.12),(3.2.13) and (3.2.14). To prove (2), observe that

$$||(\tilde{\Pi}_\eta f)''||_\infty = \max_i \frac{3}{2} \cdot |f'(a_{i+1}) - f'(a_i)| \cdot m \leq \frac{3m}{2} \int_{a_i}^{a_{i+1}} ||f''||_\infty dx \leq \frac{3}{2}||f''||_\infty, \tag{3.2.15}$$

Thus, by (1) and (3.2.15) we obtain (2) and (3). Now, to prove (4), observe that

$$|f'(x) - \sum f'(a_i)\phi_i(x)| = |\sum_i (f'(x) - f'(a_i))\phi_i(x)| \tag{3.2.16}$$

$$= |\sum_i f''(\xi)(x - a_i)\phi_i(x)| \leq \frac{||f''||_\infty}{m}. \tag{3.2.17}$$

Therefore, for (4), using (3.2.16), we have

$$|\tilde{\Pi}_\eta f(x) - f(x)| = |\int_0^x f'(\xi) - \tilde{g}(\xi)d\xi| \leq \sup_x |f'(x) - \sum f'(a_i)\phi_i(x)| \leq \frac{||f''||_\infty}{m}.$$

For (5), by using (4) and (3.2.16), we obtain

$$||\tilde{\Pi}_\eta f - f||_{\tilde{C}^1} \leq \frac{3||f''||_\infty}{m}.$$

(6) is true because

$$\int_0^1 \int_0^x \tilde{g}(\xi)\ d\xi = I(f).$$

□

Remark 3.2.5. *By (5) of Lemma 3.2.7 and the uniform Lasota-Yorke inequality (see Appendix), the Keller-Liverani Theorem [37] implies, for sufficiently small η, P and $\tilde{P}_\eta$ have a uniform rate of contraction when acting on zero average function of $\tilde{C}^2(\mathbb{T})$.*

3.2.4 A matrix representation of $\tilde{P}_\eta$

In equation (3.2.11) we defined the finite rank operator $\tilde{P}_\eta$. To treat it numerically it is natural to represent it by a matrix. To do so, we have to choose a basis for the domain and a basis for the range.

A natural choice would be to choose the same basis for its domain and its range: the basis

$$\mathcal{B} := \{e_i := \int_0^x \phi_i(\xi)d\xi\}_{i=0}^n \cup \{\mathbf{1}\},$$

where by $\mathbf{1}$ we denote the constant function 1. This choice, which may seem natural, however, due to the fact that the functions e_i have big support, this implies that the number of coefficients we have to compute is very large. Thus, storing and operating with these matrices may be complicated. To overcome this difficulty we take a basis with compact support in the domain of the operator, the basis

$$\mathcal{B}' := \{f_i = a_i \cdot e_i - b_i \cdot e_{i+1}\}_{i=0}^m \cup \{\mathbf{1}\},$$

where $a_i = 1/2$ for $i = 1, \dots, m-1$, $a_0 = 1$, $a_m = 1$ and $b_i = 1/2$ for $i = 0, \dots, m-2$, $b_{m-1} = 1$, $b_m = 0$. Using basis $\mathcal{B}'$ will reduce the the number of nonzero elements in the matrix that we have to iterate, that makes we numerical implement much easier.

3.2.5 The rigorous computation of the response $\hat{h}$

Now we show how to compute the linear response of the class of systems described in this section. A consequence of Theorem 3.1.2 is the following:

Corollary 3.2.8. *Given $\tau > 0$, $\exists\ l^* \in \mathbb{N}$ and $\eta > 0$ such that [6]*

$$||\hat{h} - \gamma \sum_{k=0}^{l^*-1} P_\eta^k \tilde{h}'_\eta||_\infty < \tau,$$

where $\tilde{h}_\eta$ is the eigenfunction corresponding to the dominant eigenvalue of $\tilde{P}_\eta$, $\tilde{h}'_\eta$ is its first derivative, γ is as in (3.2.1) *and P_η is the operator[5] defined in* (3.2.5).

Proof. In Remark 3.2.3 we established that assumptions (2), (3) of Theorem 3.1.2 hold. Since we $\hat{P}f := \gamma(Pf)'$ (see (3.2.1)) we will apply it with $f_\eta = \gamma\tilde{h}'_\eta$. Note also that assumption (1) of Theorem 3.1.2 is a consequence of the stability result of [37]. We note that by Lemma 3.2.7 and the stability result of [37], $\tilde{h}_\eta$, the eigenfuction of $\tilde{P}_\eta$, converges to the invariant density h in the $\tilde{C}^1$ norm. □

In the implementation we have to first find a suitable $l^*$and η. We follow Remark 3.1.2. For the summand $||\sum_{k=l^*}^\infty P^k\hat{P}h||_w$ we use the uniform contraction, whose coefficients can

[5]It is important to note that $\tilde{h}_\eta$ is the invariant density of $\tilde{P}_\eta$ and not of P_η.

be estimated as it will be explained in Section A.0.10. By (3.1.2) and the Lasota-Yorke inequality of item (4) of Proposition 3.2.4, we have

$$||P^k h'||_\infty \leq C_1\rho^k||h'||_{\tilde{C}^1} \leq C_1\rho^k||h||_{\tilde{C}^2}, \tag{3.2.18}$$

$$||P^k h'||_\infty \leq C_1\rho^k \cdot \frac{D}{1-\theta}. \tag{3.2.19}$$

After this estimate, following the proof of Theorem 3.1.2 (see (3.1.8)), we then obtain

$$||\hat{h} - \gamma \sum_{k=0}^{l^*-1} P_\eta^k \tilde{h}'_\eta||_\infty \leq \tag{3.2.20}$$

$$\begin{array}{c} |\gamma|(\frac{\lambda^{l^*}}{1-\lambda} \cdot \frac{D}{1-\theta} + M||(P-P_\eta)||_{\tilde{C}^1\to C^0} \cdot \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1}||P_\eta^{k-1-j}\tilde{h}'_\eta||_{\tilde{C}^1}) \\ +|\gamma|M\ l^*||\hat{P}h - \hat{P}\tilde{h}_\eta||_\infty \end{array}. \tag{3.2.21}$$

As explained in Remark 3.1.2 $M||(P-P_\eta)||_{\tilde{C}^1\to\tilde{C}^0} \cdot \sum_{k=0}^{l^*-1}\sum_{j=0}^{k-1}||P_\eta^{k-1-j}h'_\eta||_{\tilde{C}^1}$ will be estimated by the matrix associated with P_η. $||(P-P_\eta)||_{\tilde{C}^1\to\tilde{C}^0}$ will be estimated by the approximation inequality (3.1.1). The summand $M\ l^*||\hat{P}h - \hat{P}\tilde{h}_\eta||_\infty = \gamma M\ l^*||h' - \tilde{h}'_\eta||_\infty$ will be bounded by estimating $||h' - \tilde{h}'_\eta||_\infty \leq ||h - \tilde{h}_\eta||_{\tilde{C}^1}$, which can be done by Lemma 3.2.7 and the general technique of [26]. More details on this will be given in the next section.

Remark 3.2.6. *It is now clear why we use two discretization schemes to achieve our goal. One the one hand P_η satisfies the assumptions of Theorem 3.1.2. On the other hand, however, P_η cannot be used to approximate h in the $\tilde{C}^1$-norm. For $\tilde{P}_\eta$, on the one hand it does not satisfy the assumptions, in particular (2), of Theorem 3.1.2. On the other hand $\tilde{P}_\eta$ can be used to approximate h in the $\tilde{C}^1$-norm.*

3.3 Implementation and an example

3.3.1 A C^3 expanding circle map

In this section we implement the algorithm of Remark 3.1.2 and compute the linear response of a given circle map up to a pre-specified error τ. Let

$$T_0(x) = S(x) + P(x) + 0.005\sin(64\pi x),$$

where

$$S(x) = \frac{31x}{1-x}$$

and $P(x)$ is the polynomial that satisfies

$$\begin{gathered} P(0) = P(1/32) = 0, \\ P'(0) = 32 - S'(0), \quad P'(1/32) = 32 - S'(1/32), \\ P''(0) = -S''(0), \quad P''(1/32) = -S''(1/32), \\ P'''(0) = -S'''(0), \quad P'''(1/32) = -S'''(1/32). \end{gathered}$$

The coefficient of this polynomial are computed by inverting symbolically (therefore without numerical errors) a Vandermonde matrix, and therefore using interval dynamics to

enclose the coefficients. The computed polynomial, with rational coefficients, is as follow:

$$P(x) = -\frac{1048576}{29791}x^7 - \frac{917504}{29791}x^6 - \frac{923648}{29791}x^5 - \frac{923520}{29791}x^4 - 31x^3 - 31x^2 + x.$$

Let

$$T_1(x) = 32x - 1 + 0.005 \cdot \sin(64\pi \cdot x),$$

and define

$$T(x) := \begin{cases} T_0(x - 2k/32), & x \in [2k/32, (2k+1)/32] \\ T_1(x - (2k+1)/32), & x \in [(2k+1)/32, (2k+2)/32], \end{cases} \tag{3.3.1}$$

where $k = 0, 1, \ldots, 15$. Notice that T is a C^3 circle map. Its plots on $[0, 1]$ and a restricted plot on $[0, 1/16]$ are shown in Figure 3.1.

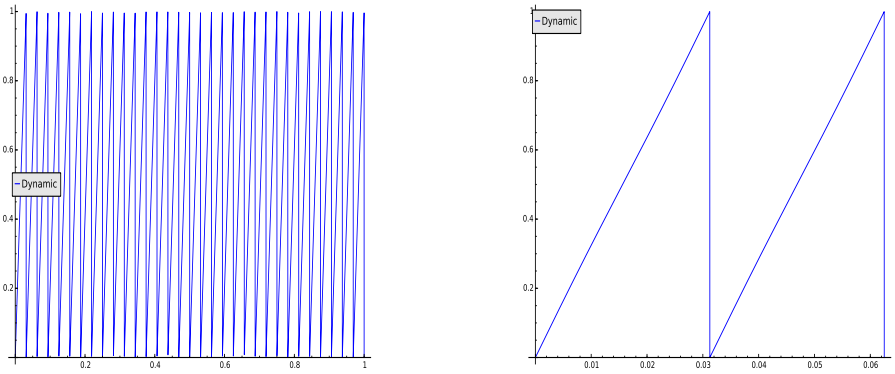

Figure 3.1: The graph of the map T as defined in (3.3.1).Horizontal axis indicate the value of x, Vertical axis indicate value of $T(x)$.

We want to compute the linear response of T which is given by

$$\hat{h} := \gamma(I - P)^{-1}h',$$

where h is the T-invariant density, h' is its derivative with respect to x, P is the transfer operator associated with T, $\gamma = \int j(\xi)|\xi|d\xi$, and ξ is a smooth density as defined in Subsection 3.2.1. Our computation will be done up to the pre-specified error $\tau = 0.044\gamma$.

3.3.2 Computing the Lasota-Yorke inequalities for P

We use the computer to estimate the following quantities, by enclosing the ranges of $|T'|$, of $|T''/(T')^2|$ and of $|T'''/(T')^3|$, using a bisection method as in , Example 3.1.3 of [51].

$$|T'(x)| > 30.93, \quad \left|\left|\frac{T''(x)}{(T')^2(x)}\right|\right|_\infty \leq 0.22, \quad \left|\left|\frac{T'''(x)}{(T')^3(x)}\right|\right|_\infty \leq 1.374.$$

Thus P satisfies the following Lasota-Yorke inequalities:

$$||Pf||_{BV} \leq 0.0324||f||_{BV} + 1.18||f||_{L^1}$$

$$||Pf||_{\tilde{C}^1} \leq 0.0394||f||_{\tilde{C}^1} + 1.45||f||_\infty$$

$$||Pf||_{\tilde{C}^2} \leq 0.00128||f||_{\tilde{C}^2} + 3.29||f||_{\tilde{C}^1}.$$

Therefore, see the Appendix, we can estimate that $||P^n||_\infty \leq 1.22$, for any n. From the first Lasota-Yorke inequality we have $||h||_{BV} \leq 1.22$, $||h||_\infty \leq 1.22$, and consequently $||h||_{\tilde{C}^1} \leq 1.84$ and $||h||_{\tilde{C}^2} \leq 6.03$.

Remark 3.3.1. *In our computations the constants are enclosed numerically using interval arithmetics with high precision. For the sake of readability, we write upper bounds (or lower bounds, depending on the sign of the inequality) with less significant digits. As an example, in the Lasota-Yorke inequalities above, we use upper bounds. If $\lambda < \tilde{\lambda} < 1$ and $B < \tilde{B}$, it is easy to see that*

$$||Pf||_s \leq \lambda||f||_s + B||f||_w \leq \tilde{\lambda}||f||_s + \tilde{B}||f||_w.$$

Note that our computation is carried on with higher precision, therefore some bounds may appear to be tighter than what expected by the written constants.

3.3.3 Approximating h'

Our first task is to approximate h', the spatial derivative of h. Obviously, to approximate h' in the L^∞ norm it is enough to approximate h in the $\tilde{C}^1$-norm. To do so, we use the approximation $\tilde{P}_\eta$ of P as explained in Remark 3.2.6. This will be done in few steps.

Convergence to equilibrium in the $\tilde{C}^1$-norm

First, using the Lasota-Yorke on $\tilde{C}^1$ we have $||P^n||_{\tilde{C}^1} \leq 1.49$ for all n. For $\eta = 1/16384$ (and for all finer mesh sizes, see Appendix) the following uniform Lasota-Yorke inequality is satisfied:

$$||\tilde{P}_\eta f||_{\tilde{C}^2} \leq 0.0041||f||_{\tilde{C}^2} + 5.092||f||_{\tilde{C}^1}. \tag{3.3.2}$$

We use the result in Lemma A.0.8 to get an approximation inequality for operator $\tilde{P}_\eta$, for any function g in $\tilde{C}^1$:

$$||P^n g - \tilde{P}_\eta^n g||_{\tilde{C}^1} \leq \eta\Big(707.34||g||_{\tilde{C}^2} + 3006.6 \cdot n||g||_{\tilde{C}^1}\Big). \tag{3.3.3}$$

By using the $\tilde{C}^1$ Lasota-Yorke inequality and the approximation inequalities we have that, for all n, $||\tilde{P}_\eta^n||_{\tilde{C}^1} \leq 47.084$. We computed the discretized matrix on the given partition, and let the computer estimate that $||\tilde{P}_\eta^3|_V||_{\tilde{C}^1} \leq 1/2048$, where V is the set of zero average functions. From this, the Lasota-Yorke inequality (3.3.2), and the approximation inequality (3.3.3), we can build a matrix as in Appendix A.0.10, such that for any g in V we have that, denoting by $g_i = P^i g$:

$$\begin{pmatrix} ||g_{i+1}||_{\tilde{C}^2} \\ ||g_{i+1}||_{\tilde{C}^1} \end{pmatrix} \preceq \begin{pmatrix} 2.067 \cdot 10^{-9} & 3.29 \\ 0.044 & 0.56 \end{pmatrix} \begin{pmatrix} ||g_i||_{\tilde{C}^2} \\ ||g_i||_{\tilde{C}^1} \end{pmatrix}$$

Therefore, for $g \in V$ we have:

$$||P^{3k} g||_{a,b} \leq (0.742)^k ||g||_{a,b},$$

with $a \in [0.055, 0.056], b \in [0.94, 0.95]$, which implies:

$$||P^{3k} g||_{\tilde{C}^1} \leq 1.059 \cdot (0.742)^k ||g||_{\tilde{C}^2} \quad ||P^{3k} g||_{\tilde{C}^2} \leq 18.19 \cdot (0.742)^k ||g||_{\tilde{C}^2}.$$

Approximating h in the $\tilde{C}^1$-norm

An important step in our approach to compute the linear response is to approximate the invariant density in the $\tilde{C}^1$ norm. To do so, we use the algorithm developed in [26] and

approximate the invariant density h by a fixed point of the discretized operator $\tilde{P}_\eta$, with $\eta = 1/4194304$.

One of the fundamental steps in this approximation is the estimation of the contraction time, i.e., the N such that $||\tilde{P}_\eta^N|_V||_{\tilde{C}^1} \leq \alpha$. We get $||\tilde{P}_\eta^5|_V||_{\tilde{C}^1} \leq 0.155$, and obtain therefore an approximation $\tilde{h}$ such that

$$||h_\eta - h||_{\tilde{C}^1} \leq 0.00199.$$

3.3.4 Convergence to equilibrium in the L^∞-norm

We first construct the matrix representation of the operator P_η. For $\eta = 1/65536$ (and for all finer mesh sizes) the following uniform Lasota-Yorke inequality is satisfied:

$$||P_\eta f||_{\tilde{C}^1} \leq 0.9753||f||_{\tilde{C}^1} + 7.96||f||_\infty. \tag{3.3.4}$$

We use the result in Lemma A.0.8 to get an approximation inequality for operator P_η, for any function g in $\tilde{C}^1$:

$$||P^n g - P_\eta^n g||_\infty \leq \eta\Big(109.02||g||_{\tilde{C}^1} + 188.97 \cdot n||g||_\infty\Big). \tag{3.3.5}$$

By using (3.3.4) and (3.3.5) we get, for all $n \geq 1$,

$$||P_\eta^n||_\infty \leq 6.93. \tag{3.3.6}$$

We computed the discretized matrix on the given partition, and let the computer estimate that $||P_\eta^5|_V||_\infty \leq 1/256$. From (3.3.4), (3.3.5) and (3.3.6), we can construct a matrix as in Appendix A.0.10, such that for any g in V we have that, denoting by $g_i = P^i g$:

$$\begin{pmatrix} ||g_{i+1}||_{\tilde{C}^1} \\ ||g_{i+1}||_\infty \end{pmatrix} \preceq \begin{pmatrix} 9.49 \cdot 10^{-8} & 1.45 \\ 0.00167 & 0.019 \end{pmatrix} \begin{pmatrix} ||g_i||_{\tilde{C}^1} \\ ||g_i||_\infty \end{pmatrix}$$

Therefore, for $g \in V$ we have:

$$||P^{5k} g||_{a,b} \leq (0.059)^k ||g||_{a,b},$$

with $a \in [0.0274, 0.0275], b \in [0.9725, 0.9726]$, which implies:

$$||P^{5k} g||_\infty \leq 1.029 \cdot (0.059)^k ||g||_{\tilde{C}^1} \quad ||P^{5k} g||_{\tilde{C}^1} \leq 36.47 \cdot (0.059)^k ||g||_{\tilde{C}^1}.$$

3.3.5 Computing the linear response

In this subsection we use the estimates obtained in subsections 3.3.3 and 3.3.4 to compute the linear response with a pre-specified error $\tau = 0.044\gamma$. Recall that h is the T-invariant density. We have $f := h' \in V$. Therefore:

$$\begin{aligned} ||P^{5k}\hat{P}h||_\infty &\leq |\gamma|||P^{5k} f||_\infty \\ &\leq \gamma \cdot 1.029 \cdot (0.059)^k ||f||_{\tilde{C}^1} \leq \gamma \cdot 1.029 \cdot (0.059)^k ||h||_{\tilde{C}^2} \\ &\leq \gamma \cdot 1.029 \cdot (0.059)^k \cdot 6.03. \end{aligned}$$

In the following table we list the error for different choices of l^*:

l^*	k	$\sum_{k=l*}^{+\infty} ||P^k \hat{P}h||_\infty$
5	1	$\frac{5}{1-0.059} \cdot 1.029 \cdot (0.059) \cdot 6.03\gamma \leq 1.95\gamma$
10	2	$\frac{5}{1-0.059} \cdot 1.029 \cdot (0.059)^2 \cdot 6.03\gamma \leq 0.12\gamma$
15	3	$\frac{5}{1-0.059} \cdot 1.029 \cdot (0.059)^3 \cdot 6.03\gamma \leq 0.0068\gamma$
20	4	$\frac{5}{1-0.059} \cdot 1.029 \cdot (0.059)^4 \cdot 6.03\gamma \leq 0.0004\gamma.$

In subsection 3.3.3 we found, $\tilde{h}$ such that,

$$||h_\eta - h||_{\tilde{C}^1} \leq 0.00199.$$

We let

$$f_\eta := \gamma \cdot (h_\eta)',$$

and we also use

$$\tilde{f}_\eta := \frac{f_\eta}{\gamma} = (h_\eta)'.$$

We have that

$$||\tilde{f}_\eta||_\infty \leq 0.6, \qquad ||\tilde{f}_\eta||_{\tilde{C}^1} \leq 4.83.$$

Now, let[6] $\eta = 1/4194304$, and compute the discretized operator P_η.

n	$\|\|P_\eta^n \tilde{f}_\eta\|\|_{\tilde{C}^1}$	n	$\|\|P_\eta^n \tilde{f}_\eta\|\|_{\tilde{C}^1}$	n	$\|\|P_\eta^n \tilde{f}_\eta\|\|_{\tilde{C}^1}$	n	$\|\|P_\eta^n \tilde{f}_\eta\|\|_{\tilde{C}^1}$
1	$1.275 \cdot 10^{-7}$	6	$2.69 \cdot 10^{-14}$	11	$2.69 \cdot 10^{-14}$	16	$2.69 \cdot 10^{-14}$
2	$8.045 \cdot 10^{-9}$	7	$2.69 \cdot 10^{-14}$	12	$2.69 \cdot 10^{-14}$	17	$2.69 \cdot 10^{-14}$
3	$1.43 \cdot 10^{-10}$	8	$2.69 \cdot 10^{-14}$	13	$2.69 \cdot 10^{-14}$	18	$2.69 \cdot 10^{-14}$
4	$7.95 \cdot 10^{-13}$	9	$2.69 \cdot 10^{-14}$	14	$2.69 \cdot 10^{-14}$	19	$2.69 \cdot 10^{-14}$
5	$3.09 \cdot 10^{-14}$	10	$2.69 \cdot 10^{-14}$	15	$2.69 \cdot 10^{-14}$	20	$2.69 \cdot 10^{-14}$.

Therefore $\sum_{i=0}^{l^*} ||P_\eta^i \tilde{f}_\eta||_{\tilde{C}^1} \leq 4.83$, for $l^* = \{5, 10, 15, 20\}$. Note that the rounding applied in the first estimate of $||f_\eta||_{\tilde{C}^1}$ was already bigger than the sum of the norms of the other terms.

Remark 3.3.2. *Due to the fact that the matrix-vector product is approximated on the computer, a small component which does not lie in V may appear; therefore, at some point this component will converge to a really small multiple of the fixed point of P_η. This is the reason why the computed bound for the $\tilde{C}^1$-norm stabilizes.*

Remark 3.3.3. *In fact, one can use the results of Section A.0.10 to estimate $||P_\eta^n f_\eta||_{\tilde{C}_1}$. If η is small enough, we can estimate the contration rate of the zero average space using the uniform Lasota-Yorke of P_η instead of the Lasota-Yorke for the operator P. With the data of Subsection 3.3.4 we have:*

$$||P_\eta^n \tilde{f}_\eta||_{\tilde{C}_1} \leq 10.05 \cdot (0.898)^{\lfloor n/5 \rfloor} ||\tilde{f}_\eta||_{\tilde{C}_1},$$

where $||\tilde{f}_\eta||_{\tilde{C}_1}$ can be estimated from the computation (since f_η lives in a finite dimensional space, it is possible to get this bound) or from the uniform Lasota-Yorke.

Using Lemma A.0.8, we now estimate:

$$||P - P_\eta||_{\tilde{C}^1 \to C^0} \leq \frac{3}{4194304} \cdot (0.0394 + 5 + 1.45) \leq 4.75 \cdot 10^{-6}.$$

Therefore:

$$\gamma \sum_0^{l^*-1} ||(P^k - P_\eta^k)\tilde{f}_\eta||_\infty \leq 1.22 \cdot 4.75 \cdot 10^{-6} \cdot 4.83\gamma \leq 0.000028\gamma.$$

Recalling that

$$||P^n||_\infty \leq 1.22$$

[6] In this example, we chose the same η to compute $\tilde{P}_\eta$ and P_η. Note that such a choice is not necessary to carryout the algorithm and the computations. In fact we could have choose different η's in the two steps.

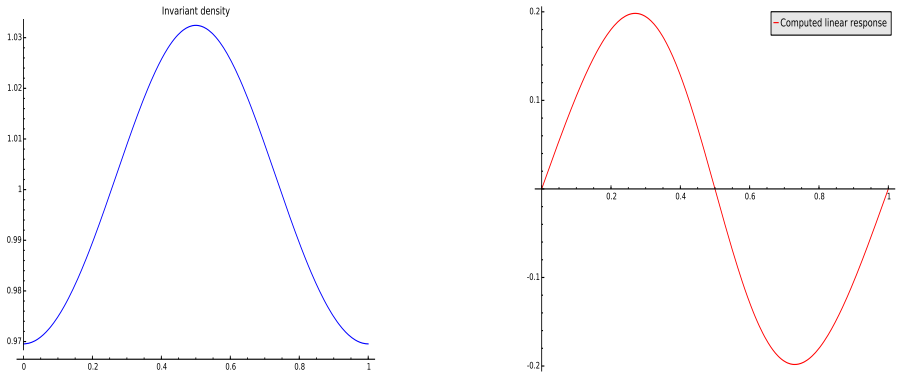

Figure 3.2: The density and its linear response. Horizontal axis indicate variable $x \in [0,1]$, Vertical axis on the left graph indicate value of invariant density function, Vertical axis on the right graph indicate value of linear response.

and

$$||h' - \tilde{f}_\eta||_\infty \leq ||h - h_\eta||_{\tilde{C}^1} \leq 0.00199,$$

we fix $l^* = 15$, and obtain

$$\sum_{i=0}^{l^*-1} ||P^k(\hat{P}h - \gamma\tilde{f}_\eta)||_\infty \leq \gamma \cdot 1.22 \cdot 15 \cdot 0.00199 \leq \gamma \cdot 0.0365.$$

Recapping the items of Remark 3.1.2 and the output of our computation

We summarize our estimates:

$$\sum_{k=15}^{+\infty} ||P^k\hat{P}h||_\infty \leq 0.0068\gamma,$$

$$\gamma\sum_{0}^{14} ||(P^k - P_\eta^k)\tilde{f}_\eta||_\infty \leq 0.000028\gamma,$$

$$\sum_{i=0}^{14} ||P^k(\hat{P}h - \gamma\tilde{f}_\eta)||_\infty \leq 0.0365\gamma.$$

Summing the errors we finally have our estimate:

$$||\hat{h} \ - \gamma\sum_{k=0}^{14} P_\eta^k\tilde{f}_\eta||_\infty \leq 0.044\gamma.$$

We plot the graph of the approximation of the invariant density in the $\tilde{C}^1$ norm on the left hand side of Figure 3.2. On the right hand side of Figure 3.2 we plot the graph of the approximation of the linear response in the L^∞-norm.

Remark 3.3.4. *There are three zeros occur on the graph of linear response function. Let us see the graph of invariant density on the left, its spacial derivative h' has zero value on $0, 0.5, 1$. Recall the definition of linear response $\hat{h} = (Id - P)^{-1}\hat{P}h$, the value of linear response is given by h'. Thus the zeros occur on the graph of linear response function at $0, 0.5, 1$.*

Appendix A

Useful inequalities used in the computer implementation of Chapter 3

A.0.6 Useful estimates

Lemma A.0.1. *Let T be a C^3 circle map. For some $k \geq 1$ let $G_k := T^k$. We have*

$$\left\| \frac{G_k''}{(G_k')^2} \right\|_\infty \leq \frac{1-\lambda^{k+1}}{1-\lambda} \left\| \frac{T''}{(T')^2} \right\|_\infty.$$

In particular, it is true for every k:

$$\left\| \frac{G_k''}{(G_k')^2} \right\|_\infty \leq \frac{1}{1-\lambda} \left\| \frac{T''}{(T')^2} \right\|_\infty = \frac{B}{1-\lambda}.$$

Moreover,

$$\left\| \frac{G_k'''}{(G_k')^3} \right\|_\infty \leq \frac{1}{1-\lambda^2} \left(\left\| \frac{T'''}{(T')^3} \right\|_\infty + \frac{3\lambda}{1-\lambda} \left\| \frac{T''}{(T')^2} \right\|_\infty^2 \right) =: Z$$

Proof. Write $G_k(x) = T(G_{k-1}(x))$. Therefore:

$$G_k(x)' = T'(G_{k-1}(x))G_{k-1}'(x),$$

$$G_k(x)'' = T''(G_{k-1}(x))(G_{k-1}'(x))^2 + T'(G_{k-1}(x))G_{k-1}''(x).$$

Using these two expressions we have

$$\frac{G_k(x)''}{(G_k(x)')^2} = \frac{T''(G_{k-1}(x))}{(T'(G_{k-1}(x)))^2} + \frac{1}{T'(G_{k-1}(x))} \frac{G_{k-1}''(x)}{(G_{k-1}'(x))^2},$$

which implies the first inequality. We now compute

$$G_k(x)''' = T'''(G_{k-1}(x))(G_{k-1}'(x))^3 + 3T''(G_{k-1}(x))G_{k-1}'(x)G_{k-1}''(x) + T'(G_{k-1}(x))G_{k-1}'''(x).$$

Using this last expression and the computations above we have:

$$\frac{G_k(x)'''}{(G_k(x)')^3} = \frac{T'''(G_{k-1}(x))}{(T'(G_{k-1}(x)))^3} + 3\frac{1}{T'(G_{k-1}(x))} \frac{T''(G_{k-1}(x))}{(T'(G_{k-1}(x)))^2} \frac{G_{k-1}''(x)}{(G_{k-1}'(x))^2} + \frac{1}{(T'(G_{k-1}(x)))^2} \frac{G_{k-1}'''(x)}{(G_{k-1}'(x))^3}.$$

□

A.0.7 Lasota-Yorke inequalities

Proposition A.0.2. *Let $T : \mathbb{T} \to \mathbb{T}$ be a C^3 uniformly expanding circle map. Set $Var(\mu) := \sup_{\substack{\phi \in C^1 \\ ||\phi||_\infty \leq 1}} |\mu(\phi')|$, where $\mu := f dx$. Then*

$$Var(P\mu) \leq \lambda Var(\mu) + B||f||_{L^1},$$

where $B = ||T''/(T')^2||_\infty$ and $\lambda = (\inf_{x \in \mathbb{T}} |D_x T|)^{-1} < 1$.

Proof. Let ϕ be a $C^1(\mathbb{T})$ observable, we have

$$\phi' \circ T(x) = \left(\frac{\phi \circ T(x)}{T'(x)} \right)' + (\phi \circ T(x)) \cdot \frac{T''(x)}{(T')^2}.$$

Consequently if $||\phi||_\infty \leq 1$,

$$\begin{aligned} |P\mu(\phi')| = |\mu(\phi' \circ T)| &\leq \left| \mu\left(\left(\frac{\phi \circ T(x)}{T'(x)} \right)' \right) \right| + \left| \mu\left(\phi \circ T(x)) \cdot \frac{T''(x)}{(T')^2} \right) \right| \\ &\leq \frac{1}{\inf_x |T'(x)|} Var(\mu) + \left|\left| \frac{T''}{(T')^2} \right|\right|_\infty ||f||_{L^1}. \end{aligned}$$

□

Lemma A.0.3 (Uniform bound for $||P^n||_\infty$)**.** *For $n \in \mathbb{N}$ we have*

$$||P^n||_\infty \leq M := 1 + B/(1 - \lambda).$$

Proof. The operator P^n is positive. Therefore $||P^n||_\infty = \sup_{x \in \mathbb{T}} P^n 1$. By Proposition A.0.2, we have

$$\mathrm{Var}(P^n 1) \leq \frac{B}{1 - \lambda},$$

and therefore $||P^n 1||_\infty \leq M$. □

Proposition A.0.4. *Let T be a C^3 circle map. We have*

$$||P^n f||_{\tilde{C}^1} \leq M \cdot \lambda^n ||f||_{\tilde{C}^1} + C||f||_\infty,$$

where

$$C = \frac{BM}{1 - \lambda} + M.$$

In particular, there exists an iterate $G := T^k$ of T such that

$$||P_G f||_{\tilde{C}^1} \leq \theta ||f||_{\tilde{C}^1} + C||f||_\infty,$$

where $\theta \leq \lambda^k M < 1$.

Proof. Denote by $G := T^n$. For $x \in (0, 1)$ we have

$$\frac{\partial}{\partial x}(P^n f)(x) = \frac{\partial}{\partial x}\left(\sum_{y = G^{-1}(x)} \frac{f(y)}{G'(y)} \right) = \sum_{y = G^{-1}(x)} \frac{f'(y)}{(G')^2} - f(y) \frac{G''(y)}{(G')^2} \frac{1}{T'(y)}. \tag{A.0.1}$$

By Lemma A.0.3, lemma A.0.1 and (A.0.1)

$$||(Pf)'||_\infty \leq \lambda^n ||Pf'||_\infty + \left|\left| \frac{G''}{(G')^2} \right|\right|_\infty ||Pf||_\infty \leq M\lambda^n ||f'||_\infty + \frac{BM}{1 - \lambda} ||f||_\infty. \tag{A.0.2}$$

Therefore, by (A.0.2) and Lemmas A.0.3, 3.2.3, we have

$$\begin{aligned} ||P^n f||_{\tilde{C}^1} &= ||P^n f||_\infty + J(P^n f) + ||(P^n f)'||_\infty \\ &\leq M||f||_\infty + \lambda^n J(f) + M\lambda^n ||f'||_\infty + \frac{BM}{1 - \lambda} ||f||_\infty \\ &\leq M\lambda^n ||f||_{\tilde{C}^1} + C||f||_\infty. \end{aligned}$$

□

Proposition A.0.5. *We have*

$$||P^n f||_{\tilde{C}^2} \leq M(\lambda^2)^n ||f||_{\tilde{C}^2} + D||f||_{\tilde{C}^1}.$$

where

$$D = \max\{3\frac{\lambda BM}{1-\lambda}, 3M\left(\frac{B}{1-\lambda}\right)^2 + MZ\} + M\lambda + C,$$

In particular, there exists an iterate $G := T^k$ of T such that

$$||P_G f||_{\tilde{C}^2} \leq \mu||f||_{\tilde{C}^2} + D||f||_{\tilde{C}^2},$$

where $\mu \leq \lambda^{2k} M < 1$.

Proof. We denote $G := T^n$. For $x \in (0,1)$ we have

$$\begin{aligned}(P^n f)''(x) = &\sum_{y=G^{-1}(x)} \frac{f''(y)}{(G'(y))^3} - 3f'(y)\frac{G''(y)}{(G'(y))^4} \\ &+ \sum_{y=G^{-1}(x)} -f(y)\frac{G'''(y)}{(G'(y))^4} + 3f(y)\frac{(G''(y))^2}{(G'(y))^5}.\end{aligned}$$

Therefore,

$$\begin{aligned}||(P^n f)''||_\infty \leq \lambda^{2n}||P^n(f'')||_\infty + 3\lambda^n \frac{B}{1-\lambda}||P^n(f')||_\infty \\ + 3\left(\frac{B}{1-\lambda}\right)^2 ||P^n f||_\infty + Z||P^n f||_\infty.\end{aligned} \tag{A.0.3}$$

In particular

$$||(P^n f)''||_\infty \leq M\lambda^{2n}||f''||_\infty + \max\{3\frac{\lambda^n BM}{1-\lambda}, 3M\left(\frac{B}{1-\lambda}\right)^2 + MZ\}||f||_{\tilde{C}^1}.$$

Thus, by Proposition A.0.4, we get

$$\begin{aligned}||Pf||_{\tilde{C}^2} &= ||Pf||_{\tilde{C}^1} + ||(Pf)''||_\infty \\ &\leq M\lambda^{2n}||f||_{\tilde{C}^2} \\ &+ \left(\max\{3\frac{\lambda^n BM}{1-\lambda}, 3M\left(\frac{B}{1-\lambda}\right)^2 + MZ\} + M\lambda^n(1-\lambda^n) + C\right)||f||_{\tilde{C}^1}.\end{aligned}$$

□

A.0.8 Uniform Lasota-Yorke inequalities for the discretized operators

Lemma A.0.6. *Let $G := T^k$. The discretized operator P_η associated with G satisfy a Lasota-Yorke inequality on the space $\tilde{C}^1(\mathbb{T})$:*

$$||P_{\tilde{\eta}} f||_{\tilde{C}^1} \leq \lambda_\eta ||f||_{\tilde{C}^1} + C_\eta ||f||_\infty$$

where:

$$\lambda_\eta = \left(\frac{11}{2} + \frac{2}{m}\right)\frac{9}{2}M\lambda^k + \frac{10M}{m},$$

and

$$C_\eta = \left(\frac{11}{2} + \frac{2}{m}\right)5\frac{BM}{1-\lambda} + 5M - \lambda_\eta.$$

Proof. Following the proof of Proposition A.0.4, we have

$$\begin{aligned}||(P\Pi_\eta f)'||_\infty &\leq M\lambda^k||(\Pi_\eta f)'||_\infty + \frac{BM}{1-\lambda}||\Pi_\eta f||_\infty \\ &\leq \frac{9}{2}M\lambda^k||f'||_\infty + 5\frac{BM}{1-\lambda}||f||_\infty.\end{aligned}$$

By Lemma 3.2.5 we have:

$$||(P_\eta f)'||_\infty \leq \frac{9}{2}||(P\Pi_\eta f)'||_\infty \quad , J(P_\eta f) \leq ||(P\Pi_\eta f)'||_\infty,$$

and

$$\begin{aligned}||P_\eta f||_\infty &\leq 5||P\Pi_\eta f|| + \frac{2}{m}||(P\Pi_\eta f)'|| \\ &\leq 5M\bigg(||f||_\infty + \frac{2}{m}||f'||_\infty\bigg) + \frac{2}{m}||(P\Pi_\eta f)'||.\end{aligned}$$

Then

$$||P_\eta f||_{\tilde{C}^1} \leq \bigg(\frac{9}{2} + 1 + \frac{2}{m}\bigg)||(P\Pi_\eta f)'|| + 5M\bigg(||f||_\infty + \frac{2}{m}||f'||_\infty\bigg).$$

□

Remark A.0.5. *Using the above estimates, if f is a fixed point of P*

$$\begin{aligned}||Pf - P_\eta f||_\infty &\leq ||P(f - \Pi_\eta f)||_\infty + ||P\Pi_\eta f - \Pi_\eta P\Pi_\eta f||_\infty \\ &\leq M||f - \Pi_\eta f||_\infty + \frac{3}{m}||(P\Pi_\eta f)'||_\infty \\ &\leq \frac{3M}{m}||f'||_\infty + \frac{3}{m}\bigg(\frac{9}{2}\theta||f'||_\infty + 5\frac{BM}{1-\lambda}||f||_\infty\bigg) \\ &\leq \frac{3}{m}\bigg(3M + \frac{9}{2}\theta\bigg)\frac{C}{1-\theta} + \frac{15}{m}\frac{BM}{1-\lambda}\frac{B}{1-\lambda}.\end{aligned}$$

Lemma A.0.7. *Let $G := T^k$. The discretized operator $\tilde{P}_\eta$ associated with G satisfy a Lasota-Yorke inequality on the space $\tilde{C}^2(\mathbb{T})$:*

$$||\tilde{P}_\eta f||_{\tilde{C}^2} \leq \mu_\eta||f||_{\tilde{C}^2} + D_\eta||f||_{\tilde{C}^1},$$

where

$$\mu_\eta = 3\lambda^{2k}M + (M + 3\lambda^k B + 3B^2M + MZ)\eta,$$

and

$$\begin{aligned}D_\eta &= \max(a, b) - \mu_\eta. \\ a &= 3\lambda^k B + 2(3B^2M + MZ) + M, \\ b &= 3\lambda^k M + 3\lambda^k B + 6\lambda^k BM + 2(3B^2M + MZ) + M.\end{aligned}$$

Proof. To prove this result we will use some inequalities proved in the proof of lemma 3.2.7, Proposition A.0.5 and Lemma A.0.4. From the proof of A.0.4 we have:

$$\begin{aligned}||(P\tilde{\Pi}_\eta f)'|| &\leq \lambda^k M||(\tilde{\Pi}_\eta f)'||_\infty + \frac{BM}{1-\lambda}||\tilde{\Pi}_\eta f||_\infty \\ &\leq \lambda^k M||f'||_\infty + \frac{BM}{1-\lambda}(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty).\end{aligned}$$

This implies that:

$$\begin{aligned}||(\tilde{P}_\eta f)'||_\infty &\leq \lambda^k M||f'||_\infty + \frac{BM}{1-\lambda}(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty) \\ J(\tilde{P}_\eta f) &\leq \lambda^k M||f'||_\infty + \frac{BM}{1-\lambda}(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty).\end{aligned}$$

From the proof of Proposition A.0.5, equation (A.0.3) we have:

$$\begin{aligned}||(P\tilde{\Pi}_\eta f)''||_\infty \leq{}& \frac{3\lambda^{2k}M}{2}||f''||_\infty + 3\lambda^k BM||f'||_\infty \\ &+ (3B^2M + MZ)(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty).\end{aligned}$$

We now observe that

$$\begin{aligned}||\tilde{P}_\eta f||_\infty &\leq ||P\tilde{\Pi}_\eta f||_\infty + ||(P\tilde{\Pi}_\eta f)'||_\infty + ||(P\tilde{\Pi}_\eta f)''||_\infty \\ &\leq M(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty) \\ &+ \lambda^k M||f'||_\infty + \lambda^k B(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty) \\ &+ \frac{3\lambda^{2k}}{2}M||f''||_\infty + 3\lambda^k BM||f'||_\infty \\ &+ (3B^2M + MZ)(||f||_\infty + ||f'||_\infty + \eta||f''||_\infty).\end{aligned}$$

Summing the inequalities we obtain that the coefficient in front of $||f''||_\infty$ is

$$\mu_\eta = 3\lambda^{2k}M + (M + 3\lambda^k B + 3B^2M + MZ)\eta;$$

please note that if η is small enough, this coefficient is smaller than 1. The coefficient of $||f||_\infty$ is

$$3\lambda^k B + 2(3B^2M + MZ) + M$$

and the coefficient of $||f'||$ is

$$3\lambda^k M + 3\lambda^k B + 6\lambda^k BM + 2(3B^2M + MZ) + M.$$

By a standard argument we get the result. □

A.0.9 Some approximation inequalities

We show how a discretized operator satisfying a Lasota-Yorke inequality satisfies useful inequalities that are used in the paper.

Lemma A.0.8. *Suppose there are two norms $||\ ||_s \geq ||\ ||_w$, such that $\forall f \in \mathcal{B}, \forall n \geq 1$*

$$||P^n f||_s \leq A\lambda_1^n||f||_s + B||f||_w. \tag{A.0.4}$$

Let π_δ be a finite rank operator satisfying:

- *$P_\delta = \pi_\delta P \pi_\delta$ with $||\pi_\delta v - v||_w \leq K\delta||v||_s$;*
- *π_δ, P^i and P_δ^i are bounded for the norm $||\ ||_w : ||\pi_\delta||_w \leq P$ and $\forall i > 0$, $||P^i||_w \leq M$, $||P_\delta^i||_w \leq M_\delta$.*

Then

$$\begin{aligned}||(P - P_\delta)f||_w &\leq K\delta(A\lambda_1 + P)||f||_s + K\delta B||f||_w \\ ||P^n f - P_\delta^n f||_w &\leq \delta K M_\delta\left(\frac{(A\lambda_1 + P)A}{1 - \lambda_1}||f|| + nB(A\lambda_1 + P + M)||f||_w\right).\end{aligned}$$

Proof. We have

$$||(P - P_\delta)f||_w \leq ||\pi_\delta P\pi_\delta f - \pi_\delta Pf||_w + ||\pi_\delta Pf - Pf||_w,$$

but

$$\pi_\delta P\pi_\delta f - \pi_\delta Pf = \pi_\delta P(\pi_\delta f - f).$$

Since $||\pi_\delta v - v||_w \leq K\delta||v||_s$

$$||\pi_\delta P(\pi_\delta f - f)||_w \leq P||\pi_\delta f - f||_w \leq PK\delta||f||_s.$$

On the other hand

$$||\pi_\delta Pf - Pf||_w \leq K\delta||Pf||_s \leq K\delta(A\lambda_1||f||_s + B||f||_w)$$

which gives

$$||(P - P_\delta)f||_w \leq K\delta(A\lambda_1 + P)||f||_s + K\delta B||f||_w \tag{A.0.5}$$

Now let us consider $(P_\delta^n - P^n)f$. We have

$$\begin{aligned}
||(P_\delta^n - P^n)f||_w &\leq \sum_{k=1}^{n} ||P_\delta^{n-k}(P_\delta - P)P^{k-1}f||_w \leq M_\delta \sum_{k=1}^{n} ||(P_\delta - P)P^{k-1}f||_w \\
&\leq K\delta M_\delta \sum_{k=1}^{n} (A\lambda_1 + P)||P^{k-1}f||_s + B||P^{k-1}f||_w \\
&\leq K\delta M_\delta \sum_{k=1}^{n} (A\lambda_1 + P)(A\lambda_1^{k-1}||f||_s + B||f||_w) + BM||f||_w \\
&\leq K\delta M_\delta (\frac{(A\lambda_1 + P)A}{1 - \lambda_1}||f|| + Bn(A\lambda_1 + P + M)||f||_w).
\end{aligned}$$

□

Lemma A.0.9. *Suppose there are two norms $||\ ||_s \geq ||\ ||_w$, such that $\forall f \in \mathcal{B}, \forall n \geq 1$*

$$||\mathcal{P}^n f||_s \leq A\lambda_1^n||f||_s + B||f||_w \tag{A.0.6}$$

Let π_δ be a finite rank operator satisfying:

- *$\mathcal{P}_\delta = \pi_\delta \mathcal{P} \pi_\delta$ with $||\pi_\delta v - v||_w \leq K\delta||v||_s$*
- *π_δ, $\mathcal{P}^i$ and $\mathcal{P}_\delta^i$ are bounded for the norm $||\ ||_w : ||\pi_\delta||_w \leq P$ and $\forall i > 0$, $||\mathcal{P}^i||_w \leq M$.*

Then if f is a fixed point of $\mathcal{P}$, we have

$$||Pf - P_\delta f|| \leq K\delta(1 + PM)||f||_s.$$

Proof. The proof is almost identical to the one above:

$$||Pf - P_\delta f||_w \leq ||Pf - \pi_\delta Pf||_w + ||\pi_\delta Pf - \pi_\delta P\pi_\delta f||_w,$$

since f is fixed point:

$$\begin{aligned}
||Pf - P_\delta f||_w &\leq ||f - \pi_\delta f||_w + ||\pi_\delta Pf - \pi_\delta P\pi_\delta f||_w \\
&\leq K\delta||f||_s + P||Pf - P\pi_\delta f||_w \\
&\leq K\delta||f||_s + PM||f - \pi_\delta f||_w \\
&\leq K\delta||f||_s + PMK\delta||f||_s.
\end{aligned}$$

□

A.0.10 Recursive convergence to equilibrium estimation for maps satisfying a Lasota-Yorke inequality

Here we recall an algorithm introduced in [27] to compute the convergence to equilibrium of a measure preserving system satisfying a Lasota-Yorke inequality. We will see how, the Lasota-Yorke inequality together with a suitable approximation of the system by a finite

dimensional one can be used to deduce finite time and asymptotic upper bounds on the contraction of the zero average space.

Consider two vector subspaces of the space of signed measures $B_s \subseteq B_w$ with norms $||\ ||_s \geq ||\ ||_w$. Let us suppose that there are operators P_δ approximating P satisfying an approximation inequality of the following type: there are constants C, D such that $\forall g \in B_s, \forall n \geq 0$:

$$||(P_\delta^n - P^n)g||_w \leq \delta(C||g||_s + nD||g||_w). \tag{A.0.7}$$

We note that in the systems and the discretizations which are considered in the paper this inequality follows from Lemma A.0.8. Now let us consider as before the space V of zero total mass measures

$$V = \{\mu \in B_s | \mu(X) = 0\}$$

and let us suppose that there exists δ and n_1 such that

$$\forall v \in V,\ ||P_\delta^{n_1}(v)||_w \leq \lambda_2 ||v||_w \tag{A.0.8}$$

with $\lambda_2 < 1$. Let us consider a starting measure: $g_0 \in V$, let us denote $g_{i+1} = P^{n_1} g_i$. If the system is as above, putting together the Lasota-Yorke inequality, (A.0.7) and (A.0.8)

$$\begin{cases} ||P^{n_1} g_i||_s \leq \lambda_1^{n_1} ||g_i||_s + B||g_i||_w \\ ||P^{n_1} g_i||_w \leq ||P_\delta^{n_1} g_i||_w + \delta(C||g_i||_s + n_1 D||g_i||_w) \end{cases}, \tag{A.0.9}$$

$$\begin{cases} ||P^{n_1} g_i||_s \leq A\lambda_1^{n_1} ||g_i||_s + B||g_i||_w \\ ||P^{n_1} g_i||_w \leq \lambda_2 ||g_i||_w + \delta(C||g_i||_s + n_1 D||g_i||_w) \end{cases}.$$

Compacting it in a vector notation,

$$\begin{pmatrix} ||g_{i+1}||_s \\ ||g_{i+1}||_w \end{pmatrix} \preceq \begin{pmatrix} A\lambda_1^{n_1} & B \\ \delta C & \delta n_1 D + \lambda_2 \end{pmatrix} \begin{pmatrix} ||g_i||_s \\ ||g_i||_w \end{pmatrix} \tag{A.0.10}$$

where $\preceq$ indicates the component-wise $\leq$ relation (both coordinates are less or equal). The relation $\preceq$ can be used because the matrix is positive. The relation (A.0.10) and the assumptions allow to estimate explicitly the contraction rate, by approximating the matrix and its iterations. Let $\mathcal{M} = \begin{pmatrix} A\lambda_1^{n_1} & B \\ \delta C & \delta n_1 D + \lambda_2 \end{pmatrix}$. Consequently, we can bound $||g_i||_s$ and $||g_i||_w$ by a sequence

$$\begin{pmatrix} ||g_i||_s \\ ||g_i||_w \end{pmatrix} \preceq \mathcal{M}^i \begin{pmatrix} ||g_0||_s \\ ||g_0||_w \end{pmatrix}$$

which can be computed explicitly. This gives an explicit estimate on the speed of convergence for the norms $||\ ||$ and $||\ ||_w$ at a given time[1].

We need an asymptotic estimation as the one given in (3.1.6) and in particular an estimation for C_1 and ρ. This can be done by the eigenvalues and eigenvectors of $\mathcal{M}$.

Indeed, let the leading eigenvalue be denoted by $\rho_{\mathcal{M}}$ and a left positive eigenvector (a, b), such that $a + b = 1$. For each pair of values (a, b) such that $a + b = 1$ we can define a norm

$$||g||_{(a,b)} = a||g||_s + b||g||_w.$$

[1]Moreover, $\lambda_1^{n_1}$, $\lambda_2 < 1$ and the quantities δC, $\delta n_1 D$ have a chance to be very small when δ is very small. This is not automatic because n_1 depend on δ. However, in the case of piecewise expanding maps, with P_δ being an Ulam-type approximation of P, as we consider in this paper, $\delta n_1 D$ can be made sufficiently small (see [26], Theorem 12).

We have that

$$||Pg||_{(a,b)} = a||Pg||_s + b||Pg||_w \leq (a,b) \cdot \mathcal{M} \cdot \begin{pmatrix} ||g||_s \\ ||g||_w \end{pmatrix}$$

then

$$||P^{kn_1}g||_{(a,b)} \leq \rho_{\mathcal{M}}^k ||g||_{(a,b)}.$$

By estimating of $\rho_{\mathcal{M}}$ and the coefficients (a,b) we can have upper estimates for C_1 and ρ.

Bibliography

[1] Anton, H and Rorres, C. : *Elementary Linear Algebra*. John Wiley Sons (Asia) Pte Ltd, 2011.

[2] Bahsoun, W., Rigorous numerical approximation of escape rates, *Nonlinearity*, 19 (2006), no. 11, 2529-2542.

[3] Bahsoun, W., Bose, C., *Invariant Densities and Escape Rates: Rigorous and Computable Approximations in The L^∞-norm*. Nonlinear Analysis, 2011, vol. 74, 4481–4495.

[4] Bahsoun, W. and Bose, C. and Duan, Y., Rigorous Pointwise approximations for invariant densities of nonuniformly expanding maps, Ergodic Theory and Dynamical Systems, 2015, 35, 1028–1044.

[5] Bahsoun, W., Galatolo, S., Nisoli, I. and Niu, X. Rigorous Approximation of Diffusion Coefficients for Expanding Maps. http://arxiv.org/pdf/1409.6909v3.pdf

[6] Bahsoun, W., Galatolo, S., Nisoli, I. and Niu, X. A Rigorous Computational Approach to Linear Response. http://arxiv.org/pdf/1506.08661.pdf

[7] Baladi, V. *Positive transfer operators and decay of correlations*. Advanced Series in Nonlinear Dynamics, 16. World Sci. Publ., NJ, 2000.

[8] Baladi, V., On the susceptibility function of piecewise expanding interval maps, *Comm. Math. Phy.*, (2007) 839-859.

[9] Baladi, V., Linear response, or else.
Available at http://arxiv.org/pdf/1408.2937v1.pdf

[10] Baladi, V., Gouëzel, S., Good Banach spaces for piecewise hyperbolic maps via interpolation. *Ann. Inst. H. Poincare Anal. Non Lineaire* (2009), 1453–1481.

[11] Baladi, V., Smania, D., Linear response formula for piecewise expanding unimodal maps. *Nonlinearity*, (2008) 677–711.

[12] Bhatia, R., *Matrix Analysis*.Springer-Verlag, New York, 1997.

[13] Blank, M., Keller, G., Liverani, C., Ruelle-Perron-Frobenius spectrum for Anosov maps. *Nonlinearity* (2002), 1905–1973.

[14] Bose, C., Murray, R., *The exact rate of approximation in Ulam's method*. Discrete Contin. Dynam. Systems 7 (2001), no. 1, 219–235.

[15] Boyarsky, A. and Góra, P. : *Laws of Chaos: invariant measures and dynamical systems in one dimension*,Birkhäuser Boston, 1997.

[16] Capinski, M and Kopp, E., *Measure, Integral and Probability*, Springer-Verlag, London, 1999.

[17] Dolgopyat, D., *Limit theorems for partially hyperbolic systems.* Trans. Amer. Math. Soc. 356 (2004), no. 4, 1637–1689.

[18] Dellnitz, M., Junge, O., On the approximation of complicated dynamical behavior. *SIAM J. Num. Anal.* 36 (1999) 491–515.

[19] Dolgopyat, D., On differentiability of SRB states for partially hyperbolic systems. *Invent. Math.* (2004), 389–449.

[20] Dunford,N and Schwartz,J., *Linear operators*. Wiley, USA, 1988.

[21] Friedman, J., *Error bounds on the power method for determining the largest eigenvalue of a symmetric, positive define matrix*. Linear Algebra and its Applications 280(1998) 199-216.

[22] Froyland, G., *Finite approximation of Sinai-Bowen-Ruelle measures for Anosov systems in two dimensions*. Random Comput. Dynam. 3 (1995), no. 4, 251–263.

[23] Froyland, G. *Using Ulam's method to calculate entropy and other dynamical invariants*. Nonlinearity 12 (1999), no. 1, 79–101.

[24] Froyland, G., Computer-assisted bounds for the rate of decay of correlations. *Comm. Math. Phys.* 189 (1997), no. 1, 237-257.

[25] Galatolo, S., Nisoli, I., *Rigorous computation of invariant measures and fractal dimension for piecewise hyperbolic maps: 2D Lorenz like maps* http://arxiv.org/pdf/1402.5918v1.pdf

[26] Galatolo, S. and Nisoli, I. : *An elementary approach to rigorous approximation of invariant measures.. SIAM J. Appl. Dyn. Syst.* 13 (2014), no. 2, 958–985.

[27] Galatolo, S., Nisoli, I., Saussol, S., *An elementary way to rigorously estimate convergence to equilibrium and escape rates*. Journal of Computational Dynamics, V. 2, Issue 1, 2015 Pages 51-64

[28] Gautschi,W., *Numerical Analysis*. Birkhäuser Boston, 1997.

[29] Gouëzel, S., Liverani, C., Banach spaces adapted to Anosov systems, *Ergodic Theory Dynam. Systems* 26 (2006), 189–217.

[30] Halmos, P. R., *Measure Theory*.Springer-Verlag, New York, 1974.

[31] Hennion, H., *Sur un théorème spectral et son application aux noyaux lipchitziens*. Proc.Amer. Math. Soc. 118 (1993), no. 2, 627-634.

[32] Higham., *Accuracy and Stability of Numerical Algorithms*, 2nd edition (2002) SIAM publishing, Philadelphia (PA), US, ISBN 0-89871-521-0.

[33] Hofbauer, F., Keller, G. *Ergodic properties of invariant measures for piecewise monotonic transformations*. Math. Z. 180 (1982), no. 1, 119–40.

[34] Jenkinson, O., Pollicott, M. *Orthonormal expansions of invariant densities for expanding maps*, Advances in Mathematics 192 (2005), 1–34.

[35] Kato,T. : *Perturbation Theory for Linear Operators*.Springer, USA, 1980.

[36] Katok, A., Knieper, G., Pollicott, M., Weiss, H., Differentiability and analyticity of topological entropy for Anosov and geodesic flows. *Invent. Math.* 98 (1989), no. 3, 581–597.

[37] Keller, G. and Liverani, C. : *Stability of the spectrum for transfer operators*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28**, (1999), no. 1, 141–152.

[38] Lasota, A., Yorke, James A. *On the existence of invariant measures for piecewise monotonic transformations*. Trans. Amer. Math. Soc. 186 (1973), 481–488.

[39] Liverani, C., *Central Limit Theorem for Deterministic Systems*, International Conference on Dynamical Systems, Montevideo 1995, a tribute to Ricardo Mane, Pitman Research Notes in Mathematics Series, 362, editor F. Ledrappier, J. Levovicz, S.Newhouse, (1996).

[40] Liverani, C., *Rigorous numerical investigation of the statistical properties of piecewise expanding maps. A feasibility study*, Nonlinearity, **14**, (2001), no. 3, 463–490.

[41] Liverani C., *Decay of correlations for piecewise expanding maps*. J. Statist. Phys. 78 (1995), no. 3-4, 1111–1129.

[42] Liverani, C., *Invariant measures and their properties. A functional analytic point of view*. Dynamical systems. Part II, 185–237, Pubbl. Cent. Ric. Mat. Ennio Giorgi, Scuola Norm. Sup., Pisa, 2003.

[43] Lucarini, V., Blender, R., Herbert, C., Pascale, S., Wouters, J., Mathematical and Physical Ideas for Climate Science. *Rev. Geophys.*, 52 (2014), 809–859.

[44] Melbourne, I. and Nicol, M. *Large deviations for nonuniformly hyperbolic systems.* Trans. Amer. Math. Soc. 360 (2008), no. 12, 6661–6676.

[45] Murray, R., *Existence, mixing and approximation of invariant densities for expanding maps on R^r*. Nonlinear Anal. 45 (2001), no. 1, 37–72.

[46] Murray, R., *Ulam's method for some non-unoformly expanding maps*, Discrete. Contin. Dyn. Syst. **26**, (2010), no. 3, 1007-1018.

[47] Pollicott, M., *Estimating variance for expanding maps.* Preprint available at http://homepages.warwick.ac.uk/ masdbl/preprints.

[48] Ruelle, D., Differentiation of SRB states, *Comm. Math. Phys.* 187 (1997) 227–241.

[49] Rynne. B. and Youngson,M *linear Functional Analysis*, Springer-Verlag, London, 2000.

[50] Saussol, B., Absolutely continuous invariant measures for multidimensional expanding maps. *Israel J. Math.* 116 (2000), 223–248.

[51] Tucker. W. *Auto-validating numerical methods*, Uppsala University, Sweden,

[52] Ulam S. M., *A Collection of Mathematical Problems* (Interscience Tracts in Pure and Applied Math. vol 8) (New York: Interscience), 1960.

[53] Walters, P., *An introduction to Ergodic Theory*, Springer-Verlag, New York, 1982.

[54] Yosida, K., *Functional Analysis*.Springer-Verlag, Berlin Heidelberg, 1995.