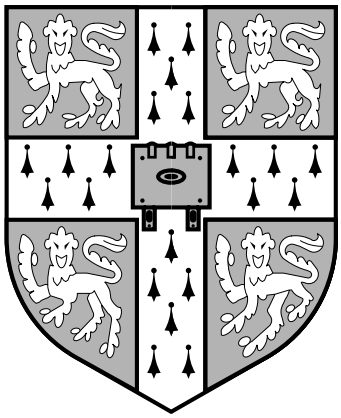
# Discrete approximation of invariant densities



**Rua Douglas Alan Murray**

**Darwin College**

A dissertation submitted for the degree of Doctor of Philosophy
at the University of Cambridge, August 1997

# Summary

A central and important problem in the application of Ergodic Theory to Dynamical Systems is the specification of an appropriate *invariant measure*. This dissertation is about methods of numerical approximation of invariant measures for *expanding transformations* which support an *invariant density*. The analysis centres around a certain Markov approximation scheme, introduced by Ulam[1]. The main results are *explicit quantitative bounds* on the approximation error (previous work[2] has focused on the convergence of the scheme in the limit); the key theorems are illustrated with numerical calculations to emphasise their applicability.

The principal techniques of proof involve casting classical existence results[3] for invariant densities into the framework of *invariant cones*. Then, using *constructive* arguments, these structures are shown to be robust to the *discretisation* induced by computation. A sample of the kind of theorem proved is:

> **Theorem** (Approximation error for expanding circle maps) *Let $T$ be a sufficiently expanding circle map, and let $h$ be the unique invariant density for $T$. There is a constant $C$ (with an explicit formula) such that whenever $n$ is sufficiently large and $h_n$ is the nth Ulam approximation*
>
> $$\|h - h_n\| \leq C \frac{\log n}{n}.$$

Similar results are proved for more general expanding maps of the interval, and a large class of multi–dimensional maps (including the classical *Jacobi–Perron* transformation, and other number theoretical transformations of recent interest[4]). As a by–product of the proofs, good rates of *decay of correlations* are obtained for the dynamical systems in question.

Finally, several new numerical approaches to the calculation of invariant densities are considered. These include: computer assisted proofs for error bounds, an automated process for adapting the Ulam approximation to the dynamics of the maps under consideration, and *Monte–Carlo* type simulations of Ulam's model.

[1] Ulam. *A collection of mathematical problems.* Interscience (1960)

[2] Li. *J. Approx. Theory* **17**:177–186 (1976). Ding and Zhou. *Physica D* **92**:61–68 (1996)

[3] Krzyżewski and Szlenk. *Stud. Math.* **33**:83–92 (1969). Lasota and Yorke. *Trans. Amer. Math. Soc.* **186**:481–488 (1973). Góra and Boyarsky. *Israel J. Math.* **67**:272–286 (1989)

[4] Baladi and Nogueira. *Nonlinearity* **9**:1529–1546 (1996)

# Preface

I hereby declare that this dissertation is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other University.

I further state that no part of this dissertation has already been or is being concurrently submitted for any such degree, diploma or other qualification.

This dissertation is the result of my own work and includes nothing that is the outcome of work done in collaboration.

Rua Murray
August 22, 1997

# Acknowledgements

For their assistance in the course of writing this thesis, many people deserve thanks.

Special mention must go to Colin Sparrow (my supervisor) and Anthony Quas of the Statistical Laboratory. Colin's tolerant reading of my work, careful criticism, continual encouragement and inexhaustible repertoire of humourous advice made writing this thesis possible. I thank Anthony for his ongoing interest in my progress, frequent discussions and down–to–earth attitude.

I also thank: Mike Keane for his enthusiasm and generous hospitality at both TU Delft and BRIMS, for many helpful discussions, and for involving me in joint work; Sebastian van Strien for his friendly discussions, helpful ideas and invitations to Warwick; and Chris Bose for rekindling my interest in multi–dimensional bounded variation during his brief visit to the Statistical Laboratory.

David Rose merits particular thanks for his uncomplaining camaraderie for the last three years, and Qun Yan Yu for her efficient management of the local computer systems. All other members of both the Statistical Laboratory and the Nonlinear Centre in DAMTP deserve credit for their maintenance of a friendly environment.

Many others have made a more personal contribution.

Finally, I thank the Cambridge Commonwealth Trust and the Committee of Vice Chancellors and Principals' ORS scheme whose financial support enabled me to come to Cambridge.

# Contents

# List of Figures

# Part I

# Introduction and background review

# Introduction

The Ergodic Theory of Dynamical Systems offers rigorous foundations for many interesting and important dynamical phenomena. These include Lyapounov Exponents, the "statistical stability" of dynamics and the long-term distribution of orbits in phase space. Yet—relative to many other branches of dynamics—the study of ergodic phenomena under *numerical simulation* remains surprisingly open.

## Motivation for the thesis

One important and accessible problem is the numerical approximation of invariant measures. Many existence results for invariant measures are known [59, 43, 47] (to name a few classical examples), and there is some history of numerical approaches to the problem [64, 48, 9, 24]. Mostly, these references adopt a fairly rigorous viewpoint by discussing the *asymptotic convergence* of approximation schemes. Many other papers have a more numerical outlook. In this thesis, we seek to exploit the virtue in both approaches: the results presented are completely rigorous (they are stated as theorems), but the development of the arguments is geared towards numerically relevant conclusions.

## Type of result

Mostly, we analyse an approximation scheme for invariant densities which was originally proposed by Ulam [64]. The scheme is known as *Ulam's method*, and its convergence to an invariant density has been established for various classes of transformations [48, 40, 11, 21]. However, the analysis of *approximation error* for Ulam's method has—thus far—remained almost untouched. The kind of result proved here is:

> **Prototype Theorem** *Let $T : X \to X$ be a suitable transformation, supporting an invariant density $h$. Then, there exists a constant $C$ (which has an explicit formula depending only on $T$) such that if $h_n$ denotes the* $n$th Ulam approximate density*,*
>
> $$\|h - h_n\| \leq C\frac{\log n}{n}.$$

The Prototype Theorem is proved in several settings; each is described in further detail below. These results conform very well to our motivations: Ulam's method is relatively easy to implement on a computer (many numerical calculations are included in the thesis), and the analytical results from precise versions of the Prototype Theorem give explicit and rigorous bounds on the approximation error.

## Specific results

The thesis contains a mixture of theorems and numerical examples, together with discussions of the interactions between them. Connections with the published literature are pointed out as they arise.

CHAPTER 1 This is a survey chapter, which begins with some comments about the effects of computer representation of phase space. The concept of a *stochastic discretisation* is introduced, and it is argued that this provides a natural framework for both dynamical simulation, and measure approximations. Several aspects of the relevant literature are surveyed, and Ulam's method is recalled. A method of proving contraction rates for Markov operators (related to Doeblin's condition) is introduced for later use.

CHAPTER 2 The main result is Theorem 2.2, the Prototype Theorem for uniformly expanding circle maps. The key technique of proof is to "discretise" a family of cones which are invariant under the Perron–Frobenius operator, thus obtaining a sequence of cones acted on by the Ulam approximation. The argument is related to one developed in collaboration with Mike Keane and Lai–Sang Young [39].

CHAPTER 3 The Prototype Theorem is proved for general expanding maps of the interval (Theorem 3.2). One important innovation is the use of a new family of $BV$ norms. As a by–product of the proofs, good rates of *decay of correlations* are obtained for the maps. The proofs can be adapted to get rates of mixing for the Markov chains corresponding to Ulam's method, and a Central Limit Theorem based analysis of *Monte–Carlo* type simulations is given.

CHAPTER 4 A method of getting numerically assisted versions of the Prototype Theorem is introduced. The argument is independent of the dimension of the phase space, but one–dimensional examples are presented (for simplicity). Buoyed by the success of this numerical approach, an *adaptive* version of Ulam's method is proposed; numerical results for the family of logistic maps are presented as an illustration.

CHAPTER 5 The Prototype Theorem is proved for certain multi–dimensional transformations (Theorem 5.5). Because the existence theory of invariant measures for multi–dimensional transformations is still developing [30, 66, 1], some effort is devoted to "cleaning up" several results from the literature. In particular, some new lemmas about multi–dimensional *bounded variation* are proved. The results are illustrated with the classical *Jacobi–Perron* transformation [62, 3].

All the arguments in the thesis are constructive; this is necessitated by the desire to obtain applicable results, but has the unfortunate affect of making some of the proofs longer, and less elegant than they might otherwise be. Furthermore, each of Chapters 2, 3 and 5 begins by reviewing the appropriate existence theory for invariant densities. This expository structure has been employed to highlight the fact that our analysis of discrete density approximations entails a certain amount of "discretisation" of the underlying mathematical concepts!

## Navigating the thesis

Each chapter is introduced with a summary of the main results and topics to be covered. The major results in the thesis—the theorems and propositions—are numbered by chapter. For ease of reference, the page numbers for these results are listed in an index.

All other intermediate results (lemmas, corollaries and cross-referenced remarks) are numbered in a continuous sequence, by section. For example, Lemma 3.2.7 is the 7th lemma, corollary or remark in Section 2 of Chapter 3. Because most sections are rather short, such results are easy to locate.

Some non–standard definitions and notations are compiled in the index, referenced by the page on which the definition appears. More standard notational conventions are summarised at the end of Chapter 1.

Finally, there are several appendices to the text. These contain either proofs or supplementary discussions whose inclusion in the main body of the thesis would be a distraction.

For the numerical results, all calculations were done using purpose written C code, and the output was displayed using SPLUS.

# Chapter 1

# Invariant measures and discretisation

In this chapter, the basic framework for the thesis is established. Because our interest in computing invariant measures is motivated by general questions about the "computability" of dynamical phenomena, we begin in Section 1.1 with a discussion of the *discretisation* entailed in computer representations of phase–space. Section 1.2 reviews those prerequisite facts about *absolutely continuous invariant measures* that are needed for the thesis, and contains a brief discussion of *stochastic stability* of dynamical systems. This is facilitated via general *Markov operators*, and some of the obstacles to translating these structures into a "discretised" framework are mentioned. Then, in Section 1.3, the problems of discretised dynamics on the one hand, and discretised "statistics" on the other, are linked together by *stochastic discretisation*, a natural framework for the computation of invariant measures. Indeed, one of the most obvious choices of stochastic discretisation is equivalent to a well known [64, 48, 32, 24] *Markov approximation* scheme: *Ulam's method*. The analysis of Ulam's method occupies most of the thesis.

Therefore, Sections 1.1—1.3 have several purposes: to motivate the problems considered in the thesis, to introduce the notation and language that will be used throughout, and to review relevant background from the literature. As such, the style is rather informal. The final section of the chapter—Section 1.4—lists other miscellaneous observations and notations, and establishes the *method of lower bounds for invariant cones*, an elementary but powerful technique for proving contraction rates for Markov operators. The method, a generalisation of Doeblin's condition, is stated as Theorem 1.1, and is used repeatedly in later chapters.

## 1.1 Discretisation of phase space

A key problem in computer simulations of dynamical systems is the extent to which the results of finite precision numerical calculations present a reliable picture of the "real" dynamics. Here, we are interested in the behaviour under discretisation of *invariant sets* and the *invariant measures* they support. Many of these problems reduce to the fact that a computer necessarily represents the phase space of a dynamical system as a *discrete* and *finite* set.

### Discretisation as a partition of phase space

Let $(X, \rho)$ be a compact metric space, equipped with the Borel $\sigma$–algebra $\mathcal{B}(X)$. (In all cases considered in this thesis, $X$ will be a Riemannian manifold, with Riemannian metric $\rho$.) A computer representation of $X$ will be a finite subset of $X$, together with a "projection" of $X$ onto that privileged finite set.

DEFINITION (**Spatial discretisation**) Let $\eta$ be a finite partition of $X$ into sets from $\mathcal{B}(X)$, and for each $B \in \eta$ pick a distinguished point $x_B \in B$. Let

$$X_\eta = \{x_B : B \in \eta\}$$

and define a *projection* $\pi_\eta : X \to X_\eta$ by

$$\pi_\eta(x) = x_B \qquad \Longleftrightarrow \qquad x \in B.$$

The finite set $X_\eta$, together with the partition $\eta$ and projection $\pi_\eta$ is called a *spatial discretisation* of $X$. Let

$$\delta(\eta) = \max\{\mathrm{diam}(B) : B \in \eta\}$$

be the *diameter* of the partition $\eta$. □

There have been many attempts to account for the effects of spatial discretisation on dynamics. A few pre-eminently relevant papers are [7, 10, 29, 17] (this list is very far from complete). Blank's papers [7, 10] concern the robustness of periodic orbits to spatial discretisation, while Diamond et. al. [17] show that certain global attractors persist under discretisation. Góra and Boyarsky [29] consider the effects of discretisation on approximations to invariant measures. Each of these papers takes the *finite set* $X_\eta$ as the fundamental object, and treats the partitioning of phase space as incidental. While some of these papers are further mentioned below, we argue that it is more appropriate to make the *partition* $\eta$ the main focus of discretisation when discussing *invariant measures*.

REMARK. The partition $\eta$ will **always** consist of finitely many sets $B$, each of which is connected, and has $m(\partial B) = 0$. □

## Invariant measures

Although we have not yet finished discussing the discretisation of dynamical phenomena, we now recall some basic definitions in measure preserving dynamics. Walters [65] is a good general reference.

Let $\mathcal{M}(X)$ denote the collection of Borel probability measures[1] on $X$. $\mathcal{M}(X)$ is a convex set, compact in the topology of weak–$*$ convergence.

DEFINITION (**Invariant measure [65]**) Let $T : X \to X$ be a transformation (map). A measure $\mu \in \mathcal{M}(X)$ is an *invariant (probability) measure* if

$$\mu \circ T^{-1}(A) \triangleq \mu(T^{-1}A) = \mu(A)$$

for every $A \in \mathcal{B}(X)$, and is *ergodic* if

$$\mu\left((A \setminus T^{-1}A) \cup (T^{-1}A \setminus A)\right) = 0 \quad \Longrightarrow \quad \mu(A) \in \{0, 1\}. \quad \square$$

In view of the *Birkhoff Ergodic Theorem* (e.g. [65]), there are many important dynamical applications for ergodic invariant measures. We are therefore concerned about how to calculate them.

REMARK 1.1.1. If $\mu$ is an ergodic invariant probability measure for $T$, then

$$\mu(A) = \int_X \chi_A \, d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \chi_A \circ T^k(x) = \lim_{n\to\infty} \frac{\#\{k < n : T^k x \in A\}}{n} \quad (1.1)$$

for $\mu$–a.e. $x \in X$. Consequently, a common strategy for approximating the measure of $A$ is to take an initial point $X$, "compute" a long segment of its orbit under $T$ and approximate $\mu(A)$ as suggested by the rhs of (1.1). Even ignoring the possibly slow convergence in (1.1) and the practical difficulties in choosing an appropriate initial point $x$, the Birkhoff Ergodic Theorem requires that the sequence $\{x, Tx, T^2x, \ldots\}$ be computed *exactly*. The truncation effects of finite precision machine arithmetic mean that this will rarely be possible. Therefore, (1.1) does not constitute a reliable method of invariant measure approximation. $\square$

## Discretisation induced periodicity

For the remainder of the section, we discuss the effects of discretisation on a dynamical system. This may be treated as something of an aside to the main body of the thesis, but provides some dynamical motivation for employing stochastic models in computer simulations.

Suppose that $\eta$ is a partition of phase space, and that $X_\eta$ is the corresponding collection of points from $X$. If $T : X \to X$ is a map, then let

$$T_\eta : x_B \mapsto \pi_\eta \circ T(x_B) \qquad \forall x_B \in X_\eta.$$

[1] A measure $\mu$ on $X$ is a probability measure if $\mu(X) = 1$.

The pair $(X_\eta, T_\eta)$ will be called an *$\eta$–discretisation* of the dynamical system $(X, T)$[2]. An orbit of the map $T_\eta$ will be called a *discretised orbit*, and is automatically a $\delta(\eta)$–pseudo orbit of $T$, although we mention this only in passing (*shadowing* will not even be discussed).

In computer models, $X_\eta$ will normally consist of the set of representable numbers (dependent on the programming of the computer), and the pair $(\eta, \pi_\eta)$ will be implicitly determined by the configuration of the machine. One conclusion of our arguments below is that (for invariant measure approximations) the set $X_\eta$ should be more "coarse" than the maximum precision deliverable by the machine.

REMARK 1.1.2. Any $\eta$–discretisation can also be viewed as a map between elements of the partition $\eta$. In this light, it is clear that the dependence of the map $T_\eta$ on the collection $\{x_B\}_{B\in\eta}$ (chosen without regard for the affect their privilege confers on the discretised dynamics) may be unreasonable. In Section 1.3 below, we seek to redress this by making an arbitrary choice of the point $x_B$ each time the set $B \in \eta$ is visited by $T_\eta$. □

For $\eta$–discretisations, one fundamental observation is immediate: because $X_\eta$ is finite, every discretised orbit is ultimately periodic. Indeed, there is very little that can be said about discretised dynamics that is more general than this.

EXAMPLE: Many authors (e.g. [9]) mention the circle doubling map $T : x \mapsto 2x \pmod 1$. Suppose that the computer represented phase space is the set $X_\eta = \{0, 2^{-N}, \ldots, i \times 2^{-N}, \ldots, 1\}$ for some $N > 0$, then[3] $T_\eta^n(x) = 0$ for every initial $x \in X_\eta$ and $n \geq N$. Of course, the map $T$ is chaotic, and leaves Lebesgue measure invariant, so the "collapsing" of the dynamics under discretisation is particularly extreme in this case. □

Now, once it is accepted that all discretised orbits are ultimately periodic, there are essentially two questions to ask about the discretised dynamics:

1. Are the periodic orbits of $T_\eta$ "close" to orbits of $T$?
2. How is the discretised orbit of a typical point distributed in phase space *before* it becomes periodic?

The first question has been studied extensively by Blank [7, 10]. In these papers, the phenomenon of *period multiplication* (where an orbit of $T_\eta$ has period a multiple of a nearby orbit of $T$) is found to be ubiquitous for certain transformations. However, the earlier paper [7] points out that a coherent theory of the persistence of other ergodic phenomena[4] is unlikely. For the second question, Góra and Boyarsky [29] give conditions under which the transient part

[2]This should not be confused with an $\epsilon$–discretisation from [7, 9, 10], where $\epsilon$ refers to a quantity related to $\delta(\eta)$.

[3]With $\pi_\eta$ being "round–down" to the nearest element of $X_\eta$, this is a realistic model for finite precision machine arithmetic.

[4]A periodic orbit can be thought of as an ergodic phenomenon insofar as it supports an invariant measure.

of a discretised orbit might be expected to mimic the distribution of an absolutely continuous invariant measure, but their hypotheses would be impossible to verify in practice.

In view of all this, the outlook for a coherent and general theory of robust dynamics under discretisation induced perturbations appears bleak. In Section 1.2, some concepts about the robustness of *ergodic* properties to *random* perturbations are recalled. Then, in Section 1.3, these concepts are "discretised", leading to Ulam's method.

We conclude the discussion of discretisation induced periodicity with a further example.

### Example: smooth unimodal maps

Consider the two parameter family of smooth unimodal maps

$$x \mapsto c(1 - |2x - 1|^l) \quad \forall x \in [0, 1] \quad \text{where } l \geq 1, 0 < c \leq 1.$$

The ergodic properties of these maps have been well studied (e.g. [57, 16]). Many maps from the family are known to possess unique ergodic invariant measures, yet the method discussed in Remark 1.1.1 is unlikely to exhibit them:

With $c = 1$, Beck and Roepstorff [5] performed numerical experiments for a variety of values of $l$. If $[0, 1]$ is discretised onto a grid of $N$ equally spaced points, they found that the maximum period of a discretised orbit scaled as $O(N^{1/l})$, and was therefore extremely un-typical (Lebesgue almost every orbit of the undiscretised transformation is dense in $[0, 1]$).

Further anecdotal evidence is easily obtained. Fixing $l = 4$, and letting $c$ vary between 0.965 and 1 in increments of $10^{-4}$, I found periodic orbits of periods between 3 and 93106. The low period orbits are close to attracting periodic orbits of the unperturbed system, but the longer period orbits are almost certainly discretisation induced. The calculations were performed in double precision arithmetic (on an HP work-station), which in this case corresponds to a partition of $[0, 1]$ into subintervals of length approximately $10^{-17}$. All these "computer orbits" are therefore very sparse indeed.

FINAL NOTE: "Discretised" definitions of "dense", "ergodic", and related concepts have not been attempted. This is because *every* orbit of $T_\eta$ is distributed according to an invariant measure (concentrated uniformly on the ultimately attained periodic orbit), and there appears to be no meaningful basis from which to claim that any such invariant measure is more distinguished then any other.

## 1.2 Invariant measures and Markov operators

Above, we discussed the behaviour of dynamics under discretisation. As a precursor to the discussion of *invariant measures* under discretisation, we recall

some standard machinery for analysing invariant measures and their robustness to perturbations.

Again, $X$ will be a Riemannian manifold. Corresponding to the Riemannian metric is a Lebesgue measure $m(\cdot)$. Assume $m(\cdot)$ is a probability measure. From a dynamical viewpoint, ergodic invariant measures which are *absolutely continuous* with respect to Lebesgue measure are of particular importance. For our purposes, it is sufficient to remark that a measure $\mu \in \mathcal{M}(X)$ is absolutely continuous with respect to Lebesgue if there exists an integrable function $\phi$ such that

$$\mu(A) = \int_A \phi \, dm$$

for every $A \in \mathcal{B}(X)$. The function $\phi$ is the *density* of $\mu$ with respect to $m$, and if $\mu$ is an invariant measure (for $T$), then $\phi$ is an *invariant density* (for $T$). Henceforth, all "absolutely continuous" measures will be absolutely continuous with respect to $m(\cdot)$, and "absolutely continuous invariant (probability) measure" will be abbreviated to *acim*. The set of absolutely continuous measures from $\mathcal{M}(X)$ is isomorphic to the non–negative elements of the unit ball in the Banach space $(L^1(X), \|\cdot\|)$.

A discussion of the physical importance of acims can be found in [12]. Roughly speaking, absolutely continuous invariant measures are important because they describe the asymptotic distribution in phase space of a large set of initial conditions, and because they can be robust to *random* perturbations in the dynamics; this is discussed more below.

REMARK. In this thesis, other physical measures are not mentioned. These include natural (or SRB) measures for hyperbolic systems, measures which are only continuous with respect to Lebesgue (such as are found on the Feigenbaum attractor) and $\sigma$–finite absolutely continuous invariant measures. □

### Perron–Frobenius operator

Many classes of transformations are known to have absolutely continuous invariant measures. For expanding transformations of the interval, the papers of Renyi [59] and Lasota and Yorke [47] are classical. Krzyżewski and Szlenk's paper [43] deals with $\mathcal{C}^2$ expanding endomorphisms of a compact manifold. Many more recent papers have dealt with increasingly complicated cases; results exist for transformations which are not $\mathcal{C}^2$, not expanding, or act in more than one dimension. Throughout the thesis, we consider the numerical approximation of acims in many of these scenarios.

Much of the success in the programme of establishing existence for invariant measures (and of our analysis of approximation schemes) rests on the fact that invariant densities have a relatively tractable analytic characterisation: as fixed points of a *Perron–Frobenius* operator. Rather than motivate the definition, we proceed directly to the expression for the operator; many excellent accounts are available, [65, 46] are particularly accessible.

DEFINITION (**Perron–Frobenius operator**) Let $T : X \to X$ be a nonsingular, piecewise differentiable transformation, with tangent map $DT$. Suppose also that there exists a partition $\xi = \{B_\alpha\}$ of $X$ into countably many sets—the monotonicity components of $T$—such that for each $\alpha$, $T|_{B_\alpha}$ is a 1-1 transformation, and $\det DT \neq 0$ on the interior of $B_\alpha$. For each $B_\alpha \in \xi$ let the corresponding *inverse branch* of $T$ be denoted by $T_\alpha^{-1}$. Then, if $f \in L^1(X)$, put

$$\mathcal{L}f = \sum_{\{\alpha : B_\alpha \in \xi\}} \frac{f \circ T_\alpha^{-1}}{\left|\det(DT) \circ T_\alpha^{-1}\right|} \chi_{TB_\alpha}.$$

The operator $\mathcal{L} : L^1(X) \to L^1(X)$ defined by this pointwise action is called the *Perron–Frobenius operator (for $T$)*. The rhs of the above expression will often be abbreviated as

$$\mathcal{L}f = \sum_\alpha \frac{f \circ T_\alpha^{-1}}{|\det(DT) \circ T_\alpha^{-1}|}. \tag{1.2}$$

By induction,

$$\mathcal{L}^n f = \sum_{\alpha^{(n)}} \frac{f \circ T_{\alpha^{(n)}}^{-n}}{|\det(DT^n) \circ T_{\alpha^{(n)}}^{-n}|}, \tag{1.3}$$

where $\alpha^{(n)}$ indexes the monotonicity components of $T^n$. □

If $\mu \in \mathcal{M}(X)$ is absolutely continuous with density $f$, then the measure $\mu \circ T^{-1}$ has density $\mathcal{L}f$. It follows immediately that $\phi \geq 0$ is an invariant density for $T$ if and only if

$$\mathcal{L}\phi = \phi \qquad \text{and} \qquad \int_X \phi \, dm = 1.$$

One standard technique for proving the existence of a fixed point for $\mathcal{L}$ rests on showing that some *cone* in $L^1(X)$ is invariant under the action of $\mathcal{L}$. In each of the situations considered in the thesis, invariant densities arise in this way, and in each case, our approximation arguments rely on the robustness of the invariant cones to discretisation.

## General Markov operators

In recent years, there has been much interest in the robustness of dynamics to *random* perturbations. This has lead many authors [60, 44, 45, 68, 51, 46] to define a new sort of structural stability for dynamical systems: *statistical structural stability*. The basic idea is that a system has statistically structurally stability if it admits an invariant measure describing the asymptotic behaviour of a large[5] set of orbits and which persists under the addition of a small amount of random "noise". A transformation $T$ under i.i.d. noise would be replaced by a *Markov chain*, described by a family of *transition probabilities* $\{\mathbb{P}_x(\cdot)\}_{x \in X}$. Kifer's books [41, 42] offer a thorough exposition, including the (weak–$*$) convergence of the invariant measures for the Markov chains to invariant measures

[5] Positive Lebesgue measure.

for $T$ as $\mathbb{P}_x(\cdot) \to \delta_{T(x)}$ (the point mass on $T(x)$). Whether the limiting invariant measures are of physical interest—in our case, acims—is an altogether more difficult question.

If the transition probabilities are absolutely continuous, then each $\mathbb{P}_x(\cdot)$ has a density $\phi(x,\cdot)$. Under the action of the Markov chain, $f \in L^1(X)$ is acted on according to

$$f \mapsto \hat{\mathcal{L}} f \qquad \text{where} \qquad \hat{\mathcal{L}} f(y) = \int_{x \in X} f(x)\, \phi(x,y)\, dm(x).$$

If the support of each $\phi(x,\cdot)$ is sufficiently localised, then the operator $\hat{\mathcal{L}}$ is a small perturbation of the Perron–Frobenius operator $\mathcal{L}$. Further explanation of this can be found in any of the papers referred to above, most of which give examples of classes of systems which have statistical structural stability. Some particularly good results—about the persistence of acims and spectral properties of $\mathcal{L}$ under this type of perturbation—may be found in [4].

Therefore, the study of the robustness of acims to small random perturbations has become fairly advanced. In fact, the operators $\hat{\mathcal{L}}$ share a *Markov* type structure with $\mathcal{L}$ which facilitates much of the analysis. For our results about invariant measure approximations, it is convenient to recall the class to which these operators belong; c.f. [46].

DEFINITION (**Markov operator**) Let $(X, \|\cdot\|_X)$ be a partially ordered Banach space, and let the partial order be denoted by $\precsim$. If $\mathcal{T} : X \to X$ satisfies

$$f \precsim g \Rightarrow \mathcal{T} f \precsim \mathcal{T} g \quad \forall f, g \in X$$

and

$$\|\mathcal{T} f\|_X \leq \|f\|_X \quad \forall f \in X$$

with equality if $0 \precsim f$, then $\mathcal{T}$ is a *Markov operator*. The composition of two Markov operators is a Markov operator. □

It is easy to check that the Perron–Frobenius operator $\mathcal{L}$ is a Markov operator on $L^1(X)$, as is the operator $\hat{\mathcal{L}}$ corresponding to a Markov chain with absolutely continuous transition probabilities $\{\mathbb{P}_x(\cdot)\}_{x \in X}$. Indeed, Markov operators are the natural analogue of the transition matrix of a countable state space Markov chain. Motivated by the possibility of applying the analytic tractability of Markov operators to discrete invariant measure approximations, we show below how to "discretise" a Markov operator to get a finite state space Markov chain.

## 1.3 Stochastic discretisation and Ulam's method

We first consider how to discretise a probability measure. Let $\eta$ be a partition of phase space $X$. When discretising $\mathcal{M}(X)$, we prefer to assign mass to the subsets of the partition $\eta$, rather than to a somewhat arbitrary collection of points $X_\eta$.

## Histogram measures

An absolutely continuous measure $\mu$ on $X$ will be called a *histogram measure* over $\eta$ if its density $f$ is constant on each $B \in \eta$. Then, $f$ has the form

$$f = \sum_{B \in \eta} f_B \, \chi_B$$

where each $f_B \in \mathbb{R}^+$, and $\sum_{B \in \eta} f_B \, m(B) = \|f\| = \mu(X)$. All the measure approximations discussed in the thesis will be approximations by histogram measures. As such, it is useful to have a notation for the collection of histogram measures over $\eta$, and it is most convenient to describe histogram measures in terms of their densities:

$$\mathcal{D}_\eta(X) = \left\{ f \in L^1(X) : f = \sum_{B \in \eta} f_B \chi_B, \;\; f_B \in \mathbb{R}^+ \;\; \forall B \in \eta \right\}.$$

Histogram measures are in a one to one correspondence with elements of $\mathcal{D}_\eta(X)$, and if $\eta = \{B_1, \ldots, B_N\}$ for some $N = N(\eta)$ then the vector $(\mu(B_1), \ldots, \mu(B_N))$ (the *vector representation* of $\mu$) is an element of $\mathbb{R}^N$. Therefore, the "discrete" space of histogram measures over $\eta$ is isomorphic to the positive cone of the Euclidean space $(\mathbb{R}^N, |\cdot|_1)$.

Throughout the thesis, the representations of a histogram measure as a measure, a density in $\mathcal{D}_\eta(X)$ or a vector in $\mathbb{R}^N$ will be used interchangeably.

Finally, let $\mu$ be an arbitrary absolutely continuous measure on $X$, and let the measure $\mu_\eta$ be defined by

$$\mu_\eta(A) = \sum_{B \in \eta} \frac{m(B \cap A)}{m(B)} \mu(B)$$

for every $A \in \mathcal{B}(X)$. Then $\mu_\eta$ is a histogram measure over $\eta$. We describe this "discretisation" operation in terms of its action on densities:

DEFINITION (**Projection onto $\mathcal{D}_\eta(X)$**) For each $f \geq 0$ in $L^1$, let the operator $\Pi_\eta$ be defined pointwise by

$$\Pi_\eta f = \sum_{B \in \eta} \frac{\int_B f \, dm}{m(B)} \chi_B. \tag{1.4}$$

Then $\Pi_\eta$ is the *projection* onto $\mathcal{D}_\eta(X)$, and if $\mu$ is an absolutely continuous measure with density $f$, then the histogram measure $\mu_\eta$ has density $\Pi_\eta f$. $\square$

We have chosen this *discretisation* operation for measures partly because it leaves the Banach space $(L^1(X), \|\cdot\|)$ invariant (allowing the possibility of analysis), but also because if $f$ is sufficiently regular, then $\|f - \Pi_\eta f\| = O(\delta(\eta)\|f\|)$ (c.f. Lemmas 2.4.1 and 3.1.1 and Proposition 5.1 below). Therefore, the "discretisation error" can be made arbitrarily small by choosing a sufficiently fine partition $\eta$. Also important is the following easy lemma:

**Lemma 1.3.1** *Embedding $\mathcal{D}_\eta(X)$ in $L^1(X)$, $\Pi_\eta : (L^1(X), \|\cdot\|) \to (L^1(X), \|\cdot\|)$ is a Markov operator.*

Now, if $\mathcal{T}$ is any Markov operator on $(L^1, \|\cdot\|)$, then

$$\Pi_\eta \circ \mathcal{T}$$

is a Markov operator on $\mathcal{D}_\eta(X)$. In particular, each operator $\Pi_\eta \circ \mathcal{T}$ can be represented as a *stochastic matrix*, acting on vectors from $\mathbb{R}^{N(\eta)}$ by left-multiplication. In this way, a "discretised Markov operator" is precisely a Markov chain on the elements of the partition $\eta$. This will be called the *Markov chain representation of* $\Pi_\eta \circ \mathcal{T}$. Consequently, $\Pi_\eta$ offers a concrete translation of the "statistical stability" formalism of general Markov operators into a discrete computational setting.

## Markov chain models

Now, having taken an analytical route to Markov chain models, we consider the reasonableness of this strategy from a dynamical point of view.

In Remark 1.1.2, it was observed that once $\eta$ is fixed, the essentially arbitrary choice of the discrete set $X_\eta$ may unreasonably affect the discretised dynamics. Our proposal is simple:

Suppose that $x \in B \in \eta$. Rather than defining $T_\eta(x)$ in terms of a fixed $x_B \in X_\eta$, we propose to let $T_\eta(x)$ be the image under $T$ of *any* point from $B$ and only use $X_\eta$ as a collection of "representatives" of sets in $\eta$. In practice, this means making a *random* choice of a point $x' \in B$, and applying $T$ with sufficient accuracy to determine which element of $\eta$ contains $T(x')$. If the distribution of $x' \in B$ depends only on $B$, then the sequence of sets in $\eta$ which are "visited" by a trajectory of this process is Markovian. When embedded in $X$, each such sequence is a $\delta(\eta)$–pseudo orbit of $T$.

An *$\eta$–stochastic discretisation* of $T$ is a Markov chain $\{Z_k\}$ on $\eta$ such that for all $B, B' \in \eta$,

$$\mathbb{P}(Z_{k+1} = B' | Z_k = B) > 0 \Rightarrow m(B \cap T^{-1}B') > 0.$$

Therefore, there is a positive probability of a transition from the set $B$ to $B'$ only if there is a positive measure subset of $B$ which is mapped into $B'$ under one application of $T$. From the point of view of capturing the "dynamics" of $T$, the intuitive appeal of such a model is clear.

REMARK. Several authors have suggested similar Markov models for computation; e.g. [58, 9]. In [58], the addition of noise was suggested to overcome the problem of discretisation induced periodicity. Blank [8, 9] has been more concerned with finding effective models for invariant measure computation, although his method and results are very different in character from ours. □

### Ulam's method

We now recall Ulam's method. In keeping with the "dual view" developed above, it is described first as a stochastic discretisation (Markov Chain) model for discrete dynamics, and then as a discretised Perron–Frobenius operator. Both views are important for the thesis. The precise treatment given here is not borrowed from any particular source, but is similar to many papers on the subject. References to the literature are given at the end.

Given a partition $\eta$ of $X$, let the elements be labelled $\{B_1, \ldots, B_N\}$ and let the transition probabilities $B_i \mapsto B_j$ be determined by the stochastic matrix $P = P_\eta$ with entries

$$P_{ij} = \frac{m(B_i \cap T^{-1}B_j)}{m(B_i)}.$$

Ulam [64] proposed this Markov chain as a model for the dynamics of $T$, and suggested that the invariant probability vector for the matrix $P$ could be embedded in $\mathcal{D}_\eta(X)$ (our notation) to obtain an approximation to an invariant measure for $T$. This procedure has become known as Ulam's method.

While the intuitive attraction of Ulam's method is clear—the probability of making a transition from the set $B_i$ to $B_j$ is precisely the (Lebesgue) proportion of mass from $B_i$ which maps into $B_j$ under one application of $T$—there is also a profound analytical attraction: the Markov chain arises as the *discretisation* of the Perron–Frobenius operator for $T$. In particular, if $\mathcal{L}$ is the Perron–Frobenius operator for $T$, let

$$\mathcal{P}_\eta = \Pi_\eta \circ \mathcal{L}.$$

Then $\mathcal{P}_\eta$ is a *Ulam approximation*, and a normalised fixed point of $\mathcal{P}_\eta$ is a *Ulam approximate density*. If

$$f = \sum_{B_i \in \eta} f_{B_i} \chi_{B_i} = \sum_{B_i \in \eta} \frac{f_{B_i} m(B_i)}{m(B_i)} \chi_{B_i},$$

a routine calculation combining the definitions of $\Pi_\eta$ and $\mathcal{L}$ (equations (1.4) and (1.2)) yields:

$$\mathcal{P}_\eta f = \sum_{B_j \in \eta} \frac{\sum_{B_i \in \eta} (f_{B_i} m(B_i))\, P_{ij}}{m(B_j)} \chi_{B_j},$$

so that the Markov chain representation of $\mathcal{P}_\eta$ has transition probabilities $(P_{ij}) = P_\eta$. Throughout the thesis, Ulam's method may be described either in the context of the operator $\mathcal{P}_\eta$ (when the outlook is analytical) or in terms of the matrix $P_\eta$ (when the emphasis is more numerical); frequently, the notational dependence on $\eta$ will be suppressed, or denoted another way.

Some comments about the practicalities of implementing Ulam's method are made in subsection 1.4.1 below.

HISTORICAL REMARKS: This scheme was originally proposed by Ulam in a collection of mathematical conjectures [64]. The problem lay dormant for many

years, and in 1976 Li [48] revived the method by proving that for Lasota–Yorke type transformations of the interval [47], the Ulam approximate densities converge strongly to an invariant density for $T$ as the partition $\eta$ is refined. In Chapter 3, we discuss Li's result, extending the class of transformations to which it applies and analysing the approximation error. In Chapter 5, the use of Ulam's method for multi–dimensional transformations is considered, and further references to the literature are given there.

Finally, we note that there are many other studies of Ulam's method, both analytical and numerical. Although they are tangentially relevant, and many have independent interest, they have not had substantial affect on the results presented in the thesis. The interested reader may care to consult some (or all) of [20, 15, 32, 23, 34].

## 1.4 Some preliminary facts and techniques

Before launching into the main body of the thesis, we collect some preliminary facts for future use.

First of all, we make some observations about the matrices corresponding to Ulam's method that will be recalled during the discussions of numerical implementations. A second (and independent) subsection introduces the *method of lower bounds for invariant cones*. This technique for proving rates of contraction for Markov operators is used in the major results of Chapters 2, 3 and 5. To avoid repetition, it is fully explained here (introducing it this early also helps motivate the construction in those chapters of lower bound functions for iterates of Perron–Frobenius operators and their Ulam approximations). Some notational conventions for the thesis are established in the last subsection.

### 1.4.1 Observations about Ulam's method

Suppose that the partition $\eta = \{B_i\}_{i=1}^{N(\eta)}$ has been fixed, and let $P = (P_{ij})$ be the matrix representation of the Ulam approximation over $\eta$ for $T$. Then

$$P_{ij} = \frac{m(B_i \cap T^{-1}B_j)}{m(B_i)}.$$

Because $m(\cdot)$ is a probability measure and $T^{-1}\eta$ is a partition of $X$, it is easy to see that the matrix $P$ is row stochastic, so possesses an invariant probability vector. Moreover, because $P_{ij} > 0$ only if a positive proportion of the mass in $B_i$ is mapped into $B_j$ under one iteration of $T$, the pattern of non–zero entries of the matrix reflects the structure of the graph of $T$ in $X \times X$. In particular, if $X$ is one–dimensional, and the elements of $\eta$ are arranged in the same order as they lie in $X$, then a plot of the zero/non–zero pattern of the entries of $P$ will resemble the graph of $T$. Some examples are depicted in Figures 4.4 and 4.5.

The same is true of iterates of $P$: the entry $(P^n)_{ij}$ is positive only if there is a sequence $i_1, \ldots, i_{n-1}$ such that each of $P_{ii_1}, P_{i_1i_2}, \ldots, P_{i_{n-1}j}$ is non-zero. It

follows that

$$m(B_i \cap T^{-n}B_j) > 0 \quad \Longrightarrow \quad (P^n)_{ij} > 0. \tag{1.5}$$

Unfortunately, the converse does not hold unless $T$ is a Markov map, and $\eta$ a Markov partition[6]. However, if the map $T$ is expansive, (1.5) easily implies the existence of an $n_0$ such that $(P^n)_{ij} > 0$ whenever $n \geq n_0$. Therefore, such a $P$ has a unique invariant probability vector.

One of the main attractions of Ulam's method is that it is relatively easy to implement. No particular care is required in the choice of partition $\eta$, and the transition probabilities can be calculated as the Lebesgue measure of the appropriate sets. This typically requires only one inverse iteration of $T$, and we assume that the entries of $P$ can be computed to arbitrary accuracy[7]. Also, if the rate of expansion of $T$ is bounded above, and the elements of $\eta$ are of roughly equivalent size, then the number of non–zero entries of each row of $P$ is bounded, irrespective of the number of elements in $\eta$. This sparse structure can be exploited to minimise memory usage. Moreover, because $P$ generally has a *unique* invariant probability vector (c.f. above), it can quickly be approximated by iterated multiplication of some initial vector by $P$. The computational requirements of Ulam's method are therefore rather modest.

Finally, if the transformation $T$ and partition $\eta$ have a sufficiently complicated structure that the above remarks do not apply, then one can resort to *Monte–Carlo* type simulations of the Markov Chain governed by $P$. Assuming that it is possible to generate a uniform (i.e. Lebesgue) distribution on each $B \in \eta$, sample paths of the Markov chain can easily be simulated; this approach is discussed in Chapter 3.

We therefore see Ulam's method as a good candidate for realistic approximation of invariant densities, and dedicate the thesis to controlling the approximation error.

### 1.4.2 Cones of functions and rates of mixing

Many of the arguments in subsequent chapters rely on estimating contraction rates for certain *Markov operators*. There is one method that will be used repeatedly, we call it the *method of lower bounds for invariant cones*. The technique can be described fairly generally, and (to avoid repetition later on) we present it below as Theorem 1.1. First of all:

[6]This is one of the principal difficulties in analysing Ulam's method, and is the main reason why so much effort in Chapter 2 is devoted to the construction of the *central component* of iterates of initial densities under $P$.

[7]The error analyses in later chapters can be adapted to account for possible errors in this part of the calculation.

DEFINITION (**Cone**) An *additive cone* (cone) is a closed set $\mathcal{C} \subset \{f \in X : 0 \precsim f\}$ such that

$$f \in \mathcal{C} \ \Rightarrow \mu f \in \mathcal{C} \ \forall \mu \geq 0$$

and

$$f + g \in \mathcal{C} \ \text{ whenever } \ f, g \in \mathcal{C}.$$

Every additive cone is convex. □

Schmitt [22] introduced cone based methods for proving exponential decay of correlations. Liverani [49] recently popularised the technique by proving exponential decay of correlations for several classes of dynamical systems. His methods involve estimating rates of contraction in a *projective Hilbert metric*. The relationship between Liverani's techniques and the method of lower bounds for invariant cones is described in Remark 1.4.2 below. For the moment, it suffices to point out that the cones considered in this thesis always lie in a Banach space $(X, \|\cdot\|_X)$, and we will require contraction rates in a norm related to $\|\cdot\|_X$, rather than the projective Hilbert metric in [49].

Now, suppose we are given a partially ordered Banach space $(X, \|\cdot\|_X)$ and a Markov operator $\mathcal{T} : X \to X$ (usually a Perron–Frobenius operator).

DEFINITION: Let the partial order on $(X, \|\cdot\|_X)$ be denoted by $\precsim$. We say that $\|\cdot\|_X$ *respects the order* if

$$0 \precsim f \precsim g \Rightarrow \|g\|_X = \|g - f\|_X + \|f\|_X. \ \square$$

Let $\mathcal{C} \subset X$ be a cone. Before defining the norm for which the method of lower bounds implies contraction for $\mathcal{T}$, we give couple of examples of cones in Banach spaces with order respecting norms.

EXAMPLE 1: Let $X = L^1([0,1])$ (the space of integrable functions on the interval $[0,1]$) and let $\|\cdot\|_X$ be the usual norm

$$\|f\|_X = \|f\| = \int_{[0,1]} |f| \, dm.$$

With the partial order $\leq$, $\|\cdot\|$ respects the order. Let

$$\mathcal{C} = \{f \in X : f \geq 0\}.$$

Then $\mathcal{C} \subset X$ is a cone. □

EXAMPLE 2: The set $\mathcal{D}_n([0,1])$ of positive functions which are constant on each subinterval $[i/n, (i+1)/n)$ is a cone. The set of positive vectors in $\mathbb{R}^n$ with the usual vector 1–norm is a cone. Equipped with the partial order $(x_1, \ldots, x_n) \precsim (y_1, \ldots, y_n)$ if and only if $x_i \leq y_i$ for each $i$, the norm respects the order. (Of course, these cones are isomorphic.) □

We now define the norms in which the method of lower bounds gives contraction.

DEFINITION (**Difference cone**) Suppose that $(X, \|\cdot\|_X)$ is a partially ordered Banach Space, and that $\|\cdot\|_X$ respects the order. Let $\mathcal{C} \subset X$ be a cone and let

$$\Gamma = \{f - g : f, g \in \mathcal{C} \text{ and } \|f\|_X = \|g\|_X\}.$$

The space $\Gamma$ is called a *difference cone.* □

DEFINITION (**Difference norm**) For $f \in \Gamma$ define

$$\|f\|_\Gamma = \inf\{\|f^{(1)}\|_X = \|f^{(2)}\|_X : f = f^{(1)} - f^{(2)}, f^{(1)} \in \mathcal{C}, f^{(2)} \in \mathcal{C}\}.$$

In view of Lemma 1.4.1 below, $\|\cdot\|_\Gamma$ will be called the *difference norm* for the (difference) cone $\Gamma$. □

**Lemma 1.4.1** $(\Gamma, \|\cdot\|_\Gamma)$ *is a normed linear space over* $\mathbb{R}$.

*Proof:* Because $\mathcal{C}$ is an additive cone, $\Gamma$ is invariant under addition and scalar multiplication. We verify the triangle inequality: Let $f, g \in \Gamma$ and let $\epsilon > 0$ be given. Let $f^{(1)}, f^{(2)}, g^{(1)}, g^{(2)} \in \mathcal{C}$ be such that

$$f = f^{(1)} - f^{(2)}, \quad \|f^{(1)}\|_X = \|f^{(2)}\|_X < \|f\|_\Gamma + \epsilon$$

and

$$g = g^{(1)} - g^{(2)}, \quad \|g^{(1)}\|_X = \|g^{(2)}\|_X < \|g\|_\Gamma + \epsilon.$$

Then

$$f + g = (f^{(1)} + g^{(1)}) - (f^{(2)} + g^{(2)})$$

and because $\|\cdot\|_X$ respects order,

$$\|f^{(1)} + g^{(1)}\|_X = \|f^{(2)} + g^{(2)}\|_X = \|f^{(1)}\|_X + \|g^{(1)}\|_X < \|f\|_\Gamma + \|g\|_\Gamma + 2\epsilon.$$

The triangle inequality follows from the definition of $\|\cdot\|_\Gamma$. The other properties are obvious. □

REMARK. If the unit ball in $(\mathcal{C}, \|\cdot\|_X)$ is compact, then the infimum in the definition of the difference norm on $\Gamma_\mathcal{C}$ is attained. □

EXAMPLE: Let $\mathcal{C}$ be the cone of integrable positive functions on $[0, 1]$. Then

$$\Gamma = \{f \in L^1 : \int_{[0,1]} f \, dm = 0\}$$

and by putting $f^{(1)} = f^+ = \max\{f, 0\}$ and $f^{(2)} = f^- = \max\{-f, 0\}$ it follows that

$$\|f\|_\Gamma = \|f^+\| = \|f^-\| = \|f\|/2. \quad \square$$

We can now state the theorem on which the method of lower bounds is based.

**Theorem 1.1 (Strict contraction for $\mathcal{T}|_\Gamma$)** *Let $\mathcal{C} \subset X$ be a cone, and let $\mathcal{T}$ be a Markov operator such that $\mathcal{T}(\mathcal{C}) \subset \mathcal{C}$. Suppose that $0 < \alpha < 1$ and that for every $f, g \in \mathcal{C}$ with $\|f\|_X = \|g\|_X$ there exists a* lower bound function *$\psi_{f,g} \in X$ such that*

$$0 \precsim \psi_{f,g}, \;\; \psi_{f,g} \precsim \mathcal{T}f, \;\; \psi_{f,g} \precsim \mathcal{T}g, \;\; \|\psi_{f,g}\|_X \geq \alpha \|f\|_X$$

*and both*

$$(\mathcal{T}f - \psi_{f,g}), (\mathcal{T}g - \psi_{f,g}) \in \mathcal{C}.$$

*Then for every $\phi \in \Gamma$ such that $\|\phi\|_\Gamma < \infty$,*

$$\mathcal{T}\phi \in \Gamma \qquad \textit{and} \qquad \|\mathcal{T}\phi\|_\Gamma \leq (1-\alpha)\|\phi\|_\Gamma.$$

*Proof:* We show that for every $\epsilon > 0$ there exist $\phi_\epsilon^{(1)}, \phi_\epsilon^{(2)} \in \mathcal{C}$ such that

$$\mathcal{T}\phi = \phi_\epsilon^{(1)} - \phi_\epsilon^{(2)}, \;\; \|\phi_\epsilon^{(1)}\|_X = \|\phi_\epsilon^{(2)}\|_X \leq (1-\alpha)\|\phi\|_\Gamma + \epsilon. \tag{1.6}$$

Let $f_\epsilon^{(1)}, f_\epsilon^{(2)} \in \mathcal{C}$ be such that

$$\phi = f_\epsilon^{(1)} - f_\epsilon^{(2)}, \;\; \|f_\epsilon^{(1)}\|_X = \|f_\epsilon^{(2)}\|_X \leq \|\phi\|_\Gamma + \epsilon.$$

Let $\psi_\epsilon = \psi_{f_\epsilon^{(1)}, f_\epsilon^{(2)}}$ be a *lower bound function* corresponding to $f_\epsilon^{(1)}, f_\epsilon^{(2)}$ and put

$$\phi_\epsilon^{(1)} = \mathcal{T}f_\epsilon^{(1)} - \psi_\epsilon, \phi_\epsilon^{(2)} = \mathcal{T}f_\epsilon^{(2)} - \psi_\epsilon.$$

By construction, both $\phi_\epsilon^{(1)}, \phi_\epsilon^{(2)} \in \mathcal{C}$ and $\mathcal{T}\phi = \phi_\epsilon^{(1)} - \phi_\epsilon^{(2)}$. Since $0 \precsim \phi_\epsilon^{(1)} \precsim \mathcal{T}f_\epsilon^{(1)}$,

$$\|f_\epsilon^{(1)}\|_X = \|\mathcal{T}f_\epsilon^{(1)}\|_X = \|\phi_\epsilon^{(1)}\|_X + \|\psi_\epsilon\|_X$$

because $\mathcal{T}$ is Markov and $\|\cdot\|_X$ respects $\precsim$. Since the same equality holds with $f_\epsilon^{(2)}, \phi_\epsilon^{(2)}$ replacing $f_\epsilon^{(1)}, \phi_\epsilon^{(1)}$ we have

$$\begin{aligned} \|\phi_\epsilon^{(2)}\|_X = \|\phi_\epsilon^{(1)}\|_X &= \|f_\epsilon^{(1)}\|_X - \|\psi_\epsilon\|_X \\ &\leq \|f_\epsilon^{(1)}\|_X - \alpha\|f_\epsilon^{(1)}\|_X \\ &< (1-\alpha)\|\phi\|_\Gamma + \epsilon. \end{aligned}$$

This establishes (1.6) and the theorem follows. □

EXAMPLE (**Doeblin's condition**) Let $\mathcal{C}$ be the cone of positive integrable functions on the interval $[0,1]$ and let $\mathcal{T}$ be a Markov operator. Suppose that there exists a function $\psi \in \mathcal{C}$ such that $0 < \|\psi\| = \alpha < 1$ and

$$\mathcal{T}f \geq \psi \quad \text{whenever } f \in \mathcal{C} \cap \{\|g\| = 1\}.$$

Then for every $\phi \in L^1([0,1])$ for which $\int_{[0,1]} \phi \, dm = 0$,

$$\|\mathcal{T}\phi\| = 2\|\mathcal{T}\phi\|_\Gamma \leq 2(1-\alpha)\|\phi\|_\Gamma = (1-\alpha)\|\phi\|.$$

This contraction result is equivalent to Doeblin's condition; e.g. see [52]. □

REMARK 1.4.2. In [49], Liverani uses a projective Hilbert metric on a cone $\mathcal{C}$ to prove strict contraction for iterates of certain Markov operators. Without going into details, if $\Theta$ denotes the projective Hilbert metric [49], and $\sup_{x,y\in\mathcal{C}} \Theta(\mathcal{T}x, \mathcal{T}y) < \infty$, then a classical argument (due to Birkhoff) implies that $\mathcal{T}$ is a strict contraction. Liverani's method is therefore to estimate $\Theta(\mathcal{T}x, \mathcal{T}y)$ for arbitrary $x, y \in \mathcal{C}$. The definition of the projective metric means that suitable estimates of lower (and upper) bounds for $\mathcal{T}x \in \mathcal{TC}$ must be obtained. It is in this respect that the method of lower bounds is related to Liverani's work. □

REMARK. The "lower bound" condition in the statement of Theorem 1.1 is similar to Lasota and Mackey's *lower functions* [44],[46]. There, a lower function is a non–negative function $h_{\mathcal{T}}$ such that

$$\liminf_{n\to\infty} \left\| (\mathcal{T}^n f - h_{\mathcal{T}})^- \right\| = 0$$

for every positive (normalised) $f \in L^1$. Markov operators for which a lower function exists are *asymptotically stable* and every initial density converges to a unique periodic orbit under iterated application of $\mathcal{T}$. □

Throughout the thesis, Theorem 1.1 will be applied to prove strict contraction on certain cones for appropriate Markov operators. The operator will usually be either a power of the Perron–Frobenius operator for a prescribed transformation, or a discrete approximation to such a Perron–Frobenius operator. In any case, most of the work will be in choosing a suitable collection of cones, and proving the existence of the lower bounds required in the theorem.

FINAL NOTE: If the usefulness of Theorem 1.1 seems doubtful at a first reading, it may be helpful to look back at this section when it is applied later on. Once the specific cones and lower bounds have been constructed for each class of transformations, the argument will seem more natural.

### 1.4.3 Some basic notation

For the sake of completeness, we now list some standard notations which are used throughout the thesis (some of them have already been used):

If $f : X \to \mathbb{R}$, then the *positive* and *negative* parts of $f$ are (respectively)

$$f^+ = \max\{f, 0\} \qquad \text{and} \qquad f^- = \max\{-f, 0\}.$$

The *usual norm* on $L^1(X)$ will be denoted by $\|\cdot\|$ (unless otherwise stated), and

$$\|f\| = \int_X |f|\, dm,$$

where $m(\cdot)$ denotes the Lebesgue measure on $X$.

The *support* of a function $f$ is the smallest closed set containing all points at which $f \neq 0$, and is denoted by

$$supp(f).$$

The *characteristic* or *indicator* function of a measurable set $A$ is

$$\chi_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

The *restriction* of a function $f$ to a measurable set $A$ is

$$f|_A = f\,\chi_A.$$

The *essential infimum* of a function $f$ over a set $A$ is

$$\operatorname*{ess\,inf}_A(f) = \sup\left\{\inf_A \tilde{f} : f = \tilde{f} \text{ a.e.}\right\},$$

and the *essential supremum* $\operatorname*{ess\,sup}_A f$ is defined similarly.

The *diameter* of a subset $A$ in the metric space $(X, \rho)$ is defined (as usual) by

$$\operatorname{diam}(A) = \sup\{\rho(x, y) : x, y \in A\}.$$

Finally, if $x \in \mathbb{R}$ then the *integer part* or *floor* of $x$ is denoted by $\lfloor x \rfloor$, and

$$\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \leq x\}.$$

For the sake of consistency, we adopt the convention

$$\lceil x \rceil = \lfloor x \rfloor + 1.$$

# Part II

# Invariant measures for one–dimensional maps

# Chapter 2

# Invariant measures for expanding circle maps

In this chapter, a method for constructing error bounds for invariant measure approximations is presented. By taking continuous structures (cones) related to the existence theory for invariant densities and "discretising" them, the argument may be regarded as a prototype for the rest of the thesis.

The class of transformations considered are expanding circle maps. The existence of a unique absolutely continuous invariant measure for such maps is the simplest case in the classical paper of Krzyżewski–Szlenk [43]. After reviewing their argument in the framework of *uniformly Lipschitz cones*, we proceed to show how these continuous structures can be *discretised*. This permits a quantitative analysis of Ulam's Markov approximation scheme, culminating in the proof of the following:

> **Theorem** (Approximation error for expanding circle maps) *Let $T$ be a sufficiently expanding circle map, and let $h$ be the unique invariant density for $T$. There exists a constant $C$ (constructively defined) such that whenever $n$ is sufficiently large and $h_n$ is the $n$th Ulam approximation*
>
> $$\|h - h_n\| \leq C\frac{\log n}{n}.$$

The precise statement of the Theorem allows the specification of explicit quantitative error bounds for Ulam's Markov approximation scheme, and is therefore of substantial computational relevance; this is illustrated with numerical examples.

Let $\mathcal{P}_n$ be the $n$th Ulam approximate operator (defined below) and let $\phi$ be an "initial density". Sections 2.1–2.3 are devoted to constructing a constant $\delta > 0$ such that for every sufficiently large $n$ there exists an $n_0 = O(\log n)$ for which

$\mathcal{P}_n{}^{n_0}\phi > \delta$. In Section 2.4, we use the *method of lower bounds for invariant cones* to get a rate of mixing for $\mathcal{P}_n$; this leads to the Theorem above.

Material related to this chapter was developed in collaboration with Mike Keane and Lai–Sang Young [39].

## 2.1 Lipschitz cones and classical existence theory

We begin with a precise definition of an expanding circle map $T$, and then briefly describe the proof of the existence of an absolutely continuous invariant measure (acim). The description of this classical [43] argument in terms of *uniformly Lipschitz cones* is reminiscent of Liverani [49], and provides a good basis for understanding the extension of cone based methods to the analysis of Ulam's Markov method.

### 2.1.1 Cones for uniformly expanding circle maps

DEFINITION: Let $T : S^1 \to S^1$ be a $\mathcal{C}^2$ circle map and let $DT$ be its tangent map (derivative). The map $T$ will be called *uniformly expanding* if there exist constants $\lambda > 1$ and $K < \infty$ such that for every inverse branch $T_\alpha^{-1}$ of $T$

$$\rho(T_\alpha^{-1}(x), T_\alpha^{-1}(y)) \le \lambda^{-1}\rho(x,y) \tag{2.1}$$

and

$$\Big|\log|DT(x)| - \log|DT(y)|\Big| \le K\rho(x,y) \tag{2.2}$$

for every $x, y \in S^1$ where $\rho$ is the usual metric[1] on $S^1$. □

Throughout this chapter, $T$ will always be a uniformly expanding circle map, and $\lambda$ and $K$ will be the corresponding constants.

Let $\mathcal{L}$ be the Perron–Frobenius operator for $T$. In this simple setting, the main idea behind the existence of a fixed point for $\mathcal{L}$ (invariant density) is that $\mathcal{L}$ preserves a certain cone of functions.

DEFINITION (**Cone** $\mathcal{C}_a$) For each $a > 0$ put

$$\mathcal{C}_a = \left\{\phi : S^1 \to \mathbb{R}^+ : \frac{\phi(x)}{\phi(y)} \le e^{a\rho(x,y)} \ \ \forall x, y \in S^1\right\}. \quad \square$$

[1] If $S^1 = \mathbb{R}/\mathbb{Z}$ is identified with the interval $[0,1)$, then $\rho(x,y) = \min\{|x-y|, 1-|x-y|\}$. For the *lift* to $\mathbb{R}$, it will be more convenient to work with $|x-y|$.

**Lemma 2.1.1** *For $a > 0$*

1. *if $\phi \in \mathcal{C}_a$ then $\phi$ is uniformly continuous;*
2. *$\mathcal{C}_a$ is a cone, and for each $c \in \mathbb{R}^+$ the set $\mathcal{C}_a \cap \{\phi : \|\phi\| = c\}$ is compact in the topology of uniform convergence;*
3. *If $\phi \in \mathcal{C}_a$, then*
$$e^{-a}\|\phi\| \leq \phi(x) \leq e^a\|\phi\| \quad \forall x \in S^1;$$
4. *$\mathcal{L}\mathcal{C}_a \subset \mathcal{C}_{\lambda^{-1}(a+K)}$ where $\mathcal{L}$ is the Perron–Frobenius operator for $T$. In particular,*
$$\mathcal{L}\mathcal{C}_{\frac{K}{\lambda-1}} \subset \mathcal{C}_{\frac{K}{\lambda-1}}.$$

*Proof:* Let $\phi \in \mathcal{C}_a$ and let $x_0 \in S^1$ be such that $\phi(x_0) \leq \|\phi\|$. Then $\phi(x) \leq \phi(x_0)e^{\rho(x,x_0)} \leq \|\phi\|e^a$ because $\rho(x,y) \leq 1$ for every $x, y \in S^1$. The lower bound is similar, and uniform continuity follows because

$$|\phi(x) - \phi(y)| \leq \phi(y)\left(e^{a\rho(x,y)} - 1\right) \leq \|\phi\|e^a\left(e^{\rho(x,y)} - 1\right).$$

It is easy to check that $\mathcal{C}_a$ is closed, and invariant under addition and scalar multiplication. Moreover, the above inequality implies that each $\mathcal{C}_a \cap \{\phi : \|\phi\| = c\}$ is equicontinuous, so that compactness follows from the Arzela–Ascoli theorem.

The proof of the last part is deferred until after Lemma 2.1.4 below. □

Each $\mathcal{C}_a$ is called a *uniformly Lipschitz cone* on $S^1$. In view of Lemma 2.1.1 (4), uniformly Lipschitz cones are ideally suited to Perron–Frobenius operators for uniformly expanding circle maps. In particular, notice how elements of $\mathcal{C}_a$ are uniformly bounded below. This property of Lipschitz cones turns out to be vital to the method of "lower bounding" that is described in Section 2.4. Finally, for the sake of completeness, it is important to know that $\mathcal{L}$ has a fixed point.

**Corollary 2.1.2 ([43])** *Let $T$ be a uniformly expanding circle map, and let $\mathcal{L}$ be its Perron–Frobenius operator. There exists an invariant density $h \in \mathcal{C}_{K/(\lambda-1)}$.*

*Proof:* By the lemma,

$$\mathcal{L}\left(\mathcal{C}_{K/(\lambda-1)} \cap \{\phi : \|\phi\| = 1\}\right) \subset \mathcal{C}_{K/(\lambda-1)} \cap \{\phi : \|\phi\| = 1\}$$

and $\mathcal{C}_{K/(\lambda-1)} \cap \{\phi : \|\phi\| = 1\}$ is convex and compact. □

REMARK. Liverani [49] uses uniformly Lipschitz cones to prove good bounds on the rate of decay of correlations for certain circle maps. Unfortunately, the loss of continuous structure involved in discrete invariant measure approximations means that his arguments are not applicable to the situation here. However, we also prove a rate of exponential mixing (for the discrete approximation to $\mathcal{L}$) but in a much more "bare–hands" way. □

### 2.1.2 Lifted cones for lifted circle maps

Because the inverse of the map $T$ has more than one branch, it can be difficult to keep track of all contributions to the Perron–Frobenius operator $\mathcal{L}$. This is especially true of discrete approximations to $\mathcal{L}$, and the complications may be dealt with by working entirely with a *lift* of the map $T$ to $\mathbb{R}$ (the universal cover of $S^1$). Therefore, the majority of the following analysis is done for the lift, and the results are projected back to the circle at the end.

Identify $S^1$ with the interval $[0,1)$, and define $\pi : \mathbb{R} \to S^1$ by

$$\pi(x) = x - \lfloor x \rfloor.$$

The inverse maps to $\pi$ are indexed by $l \in \mathbb{Z}$, and we write

$$\pi_l^{-1}(x) = x + l.$$

DEFINITION: A $\mathcal{C}^2$ diffeomorphism $\tilde{T} : \mathbb{R} \to \mathbb{R}$ satisfying

$$\pi \circ \tilde{T} = T \circ \pi$$

is called a *lift* of $T$. $\square$

REMARK. Because $T$ is a uniformly expanding circle map, it follows immediately from (2.1) and (2.2) that

$$|\tilde{T}'| \geq \lambda \quad \text{and} \quad \left|\tilde{T}''/\tilde{T}'\right| \leq K$$

for any lift $\tilde{T}$ of $T$. $\square$

Henceforth, we fix a lift $\tilde{T}$ of $T$ such that $\tilde{T}(0) \in [0,1)$, and establish some elementary properties necessary for the correspondence between the Perron–Frobenius operators for $\tilde{T}$ and $T$.

The definition of the lift implies that there exists a unique integer $M$ such that

$$\tilde{T}(x+1) = \tilde{T}(x) + M$$

for every $x \in \mathbb{R}$. It is easy to see that $|M|$ is the number of inverse branches of $T$. Let the inverse branches $T_\alpha^{-1}$ be indexed by the integers $0, \ldots, |M|-1$. Then

$$\pi_l^{-1} \circ T_\alpha^{-1} = \tilde{T}^{-1} \circ \pi_{l|M|+\alpha}^{-1}$$

for each $l \in \mathbb{Z}$ and inverse branch $T_\alpha^{-1}$.

Now, we need the Perron–Frobenius operator for $\tilde{T}$. This will be denoted by $\tilde{\mathcal{L}}$, and is defined in precisely the same way as the Perron–Frobenius operator

for $T$. Because $\tilde{T}$ is a diffeomorphism, $\tilde{T}^{-1}$ has a unique branch, and for every positive function $\tilde{\phi}$ on $\mathbb{R}$

$$\tilde{\mathcal{L}}\tilde{\phi}(x) = \frac{\tilde{\phi} \circ \tilde{T}^{-1}(x)}{|\tilde{T}' \circ \tilde{T}^{-1}(x)|}.$$

DEFINITION: Let $\tilde{\phi} : \mathbb{R} \to \mathbb{R}^+$, and let $\hat{\phi} : S^1 \to \mathbb{R}^+$ be defined according to the formula

$$\hat{\phi}(x) = \sum_{l \in \mathbb{Z}} \tilde{\phi} \circ \pi_l^{-1}(x).$$

Then $\hat{\phi}$ is called the *projection* of $\tilde{\phi}$ to $S^1$. In general, this projection operation will be denoted by a $\hat{\cdot}$. □

**Lemma 2.1.3 (Correspondence between $\mathcal{L}$ and $\tilde{\mathcal{L}}$)** *Let $\tilde{\phi} : \mathbb{R} \to \mathbb{R}^+$ and let $\hat{\phi}$ be its projection to $S^1$. Then $\int_{\mathbb{R}} \tilde{\phi} = \int_{S^1} \hat{\phi}$ and*

$$\mathcal{L}\hat{\phi} = \sum_{l \in \mathbb{Z}} \left( \tilde{\mathcal{L}}\tilde{\phi} \right) \circ \pi_l^{-1}.$$

*Proof:* The equality of integrals is obvious from the definition. The other part is almost as easy:

$$\begin{aligned} \mathcal{L}\hat{\phi} = \mathcal{L}\left( \sum_{l \in \mathbb{Z}} \tilde{\phi} \circ \pi_l^{-1} \right) &= \sum_{\alpha=0}^{|M|-1} \sum_{l \in \mathbb{Z}} \frac{\tilde{\phi} \circ \pi_l^{-1} \circ T_\alpha^{-1}}{|T' \circ T_\alpha^{-1}|} \\ &= \sum_{\alpha=0}^{|M|-1} \sum_{l \in \mathbb{Z}} \frac{\tilde{\phi} \circ \tilde{T}^{-1} \circ \pi_{l|M|+\alpha}^{-1}}{|\tilde{T}' \circ \tilde{T}^{-1} \circ \pi_{l|M|+\alpha}^{-1}|} \\ &= \sum_{\alpha=0}^{|M|-1} \sum_{l \in \mathbb{Z}} \left( \frac{\tilde{\phi} \circ \tilde{T}^{-1}}{|\tilde{T}' \circ \tilde{T}^{-1}|} \right) \circ \pi_{l|M|+\alpha}^{-1}, \end{aligned}$$

as required. □

By Lemma 2.1.3, all important properties of $\mathcal{L}$ can be deduced by working with the lift, and projecting. In particular, uniformly Lipschitz cones may be lifted.

DEFINITION (**Cone** $\mathcal{C}_{a,A}$) Let $a > 0$ and let $A \subset \mathbb{R}$ be an interval. Put

$$\mathcal{C}_{a,A} = \left\{ \tilde{\phi} : \mathbb{R} \to \mathbb{R}^+ \, : \tilde{\phi}|_{\mathbb{R} \setminus A} = 0, \frac{\tilde{\phi}(x)}{\tilde{\phi}(y)} \leq e^{a|x-y|} \;\; \forall x, y \in A \right\}.$$

Each $\mathcal{C}_{a,A}$ is a *uniformly Lipschitz cone* on $\mathbb{R}$. □

**Lemma 2.1.4 (Basic properties of $\mathcal{C}_{a,A}$)** *Let $a > 0$. Then*

1. *if $m(A) < \infty$ then elements of $\mathcal{C}_{a,A}$ are uniformly continuous, and $\mathcal{C}_{a,A}$ is compact;*

2. *if $\exists N \in \mathbb{Z}$ such that $A = [z, z+N)$ and $\tilde{\phi} \in \mathcal{C}_{a,A}$ satisfies $\tilde{\phi}(z) = \tilde{\phi}(z+N)$ then $\hat{\phi} \in \mathcal{C}_a$, where $\hat{\phi}$ is the projection of $\tilde{\phi}$ to $S^1$;*

3. *for every interval $A$,*

$$\tilde{\mathcal{L}}(\mathcal{C}_{a,A}) \subset \mathcal{C}_{\lambda^{-1}(a+K),\tilde{T}A}.$$

*Proof:* The first part follows by the same proof as in Lemma 2.1.

For the second part, suppose that $x < y \in S^1 \equiv [0,1)$ and put $l_0 = \inf\{l \in \mathbb{Z} : z \le x + l\}$. Then

$$\tilde{\phi}(x) = \sum_{l=l_0}^{l_0+N-1} \tilde{\phi}(x+l)$$

and either $l_0 = \inf\{l \in \mathbb{Z} : z \le y+l\}$ or $l_0 - 1 = \inf\{l \in \mathbb{Z} : z \le y+l\}$. Since the first case is easier, we prove this part of the lemma for the second case. Then:

$$\begin{aligned}
\frac{\hat{\phi}(x)}{\hat{\phi}(y)} &= \frac{\sum_{l=l_0}^{l_0+N-1} \tilde{\phi}(x+l)}{\sum_{l=l_0-1}^{l_0+N-2} \tilde{\phi}(y+l)} \\
&\le \max\left\{ \max_{l_0 \le l \le l_0+N-2} \frac{\tilde{\phi}(x+l)}{\tilde{\phi}(y+l)}, \frac{\tilde{\phi}(x+l_0+N-1)}{\tilde{\phi}(y+l_0-1)} \right\} \\
&\le \max\left\{ e^{a|x-y|}, \frac{\tilde{\phi}(x+l_0+N-1)}{\tilde{\phi}(z+N)} \frac{\tilde{\phi}(z+N)}{\tilde{\phi}(z)} \frac{\tilde{\phi}(z)}{\tilde{\phi}(y+l_0-1)} \right\} \\
&\le \max\left\{ e^{a|x-y|}, e^{a(z+N-(x+l_0+N-1))} e^{a(y+l_0-1-z)} \right\} \\
&= e^{a|x-y|}
\end{aligned}$$

because $\tilde{\phi} \in \mathcal{C}_{a,A}$. A similar argument shows that $\frac{\hat{\phi}(x)}{\hat{\phi}(y)} \le e^{a(1-|x-y|)}$.

The final part of the lemma follows immediately from the definitions of $\tilde{\mathcal{L}}$, $\mathcal{C}_{a,A}$, $K$ and $\lambda$. □

*Proof of Lemma 2.1.1 (4):* If $\phi \in \mathcal{C}_a$, then $\tilde{\phi} \triangleq \phi \circ \pi_0^{-1} \in \mathcal{C}_{a,[0,1)}$. By Lemma 2.1.4 (3), $\tilde{\mathcal{L}}\tilde{\phi} \in \mathcal{C}_{\lambda^{-1}(a+K),A}$ where $A = \tilde{T}[0,1)$ is an interval of length $|M|$. By Lemma 2.1.3, the projection of $\tilde{\mathcal{L}}\tilde{\phi}$ onto $S^1$ is exactly $\mathcal{L}\phi$, and part (2) of Lemma 2.1.4 proves the inclusion in Lemma 2.1.1. □

Henceforth, the $\tilde{\cdot}$ on functions in the lift will be omitted for notational convenience.

## 2.2 Discrete Lipschitz cones

The $n$th Ulam Markov approximate operator for $T$ arises as the composition of the Perron–Frobenius operator for $T$ with a projection onto the space of functions which are piecewise constant on intervals of length $\frac{1}{n}$. For the remainder of the chapter, we suppose that $n$ is fixed.

This projection onto piecewise constant functions is easily extendable from $[0,1)$ to $\mathbb{R}$. The extension to $\mathbb{R}$ will be denoted by $\Pi_n$:

DEFINITION (**Ulam's method**) Let $\phi : \mathbb{R} \to \mathbb{R}^+$ be integrable, and let $n$ be a fixed natural number. Define a projection $\Pi_n$ by

$$\Pi_n \phi(x) = \int_{\lfloor nx \rfloor}^{\lceil nx \rceil} \phi(n^{-1}y)\, dm(y).$$

It is obvious that $\Pi_n \phi$ is constant on each interval $[i/n, (i+1)/n)$ and the *nth Ulam Markov approximation* corresponding to $\tilde{T}$ is defined as

$$\tilde{\mathcal{P}}_n = \Pi_n \circ \tilde{\mathcal{L}}. \quad \square$$

By Lemma 2.1.3, $\Pi_n$ commutes with projection onto $S^1$, so it is entirely reasonable to regard $\tilde{\mathcal{P}}_n$ as the *lift* of $\mathcal{P}_n$ (the $n$th Ulam approximate operator on $S^1$). Indeed, if $\phi : \mathbb{R} \to \mathbb{R}^+$ and $\hat{\phi}$ is its projection to $S^1$,

$$\mathcal{P}_n \hat{\phi} = \sum_{l \in \mathbb{Z}} \left( \tilde{\mathcal{P}}_n \phi \right) \circ \pi_l^{-1}. \tag{2.3}$$

The analysis of the mixing properties of $\mathcal{P}_n$ is therefore carried out via $\tilde{\mathcal{P}}_n$. Hence, we next define cones for $\tilde{\mathcal{P}}_n$ to act on.

While the Perron–Frobenius operators for $T$ and $\tilde{T}$ preserve certain uniformly Lipschitz cones, the operator $\Pi_n$ introduces discontinuities which the uniformly Lipschitz cones are not robust enough to absorb. This is elaborated by the following definition, and subsequent lemma.

DEFINITION (**Cone** $\mathcal{C}_{a,A,\frac{k}{n}}$) Let $a > 0$, $A \subset \mathbb{R}$ an interval and $k \geq 0$. Put

$$\mathcal{C}_{a,A,\frac{k}{n}} = \left\{ \phi : \phi|_A > 0, \frac{\phi(x)}{\phi(y)} \leq e^{a(|x-y|+k/n)} \ \ \forall x, y \in \mathbb{R} \right\}.$$

Each $\mathcal{C}_{a,A,\frac{k}{n}}$ is called a *discrete Lipschitz cone*. $\square$

It is obvious that any element of a discrete Lipschitz cone is allowed to have small discontinuities, and the following lemma shows that the projection $\Pi_n$ allows only a small increase in these discontinuities.

**Lemma 2.2.1** *Fix $a = K/(\lambda - 1)$. Then*

1. *for each $k \geq 0$ and $A \subset \mathbb{R}$,*
$$\tilde{\mathcal{L}}\mathcal{C}_{a,A,\frac{k}{n}} \subset \mathcal{C}_{a,\tilde{T}A,\frac{k}{n}};$$
2. *if $A = [\frac{i}{n}, \frac{j}{n})$ for some $i, j \in \mathbb{Z}$ then*
$$\Pi_n \mathcal{C}_{a,A,\frac{k}{n}} \subset \mathcal{C}_{a,A,\frac{k+1}{n}}.$$

*Proof:* The first part follows from the definitions of $\tilde{\mathcal{L}}$ and discrete Lipschitz cones. To prove the other part, let $\phi \in \mathcal{C}_{a,A,\frac{k}{n}}$ and let $x, y \in A$. Let $i_1, i_2 \in \{i, \ldots, j\}$ be the unique integers such that $x \in [\frac{i_1}{n}, \frac{i_1+1}{n})$ and $y \in [\frac{i_2}{n}, \frac{i_2+1}{n})$. Then
$$\frac{|i_1 - i_2| - 1}{n} \leq |x - y| \leq \frac{|i_1 - i_2| + 1}{n}.$$
But
$$\begin{aligned}
\frac{\Pi_n\phi(x)}{\Pi_n\phi(y)} &= \frac{\int_{i_1}^{i_1+1} \phi(n^{-1}x')\, dm(x')}{\int_{i_2}^{i_2+1} \phi(n^{-1}y')\, dm(y')} \\
&= \frac{\int_{i_2}^{i_2+1} \phi(n^{-1}(y' - (i_1 - i_2)))\, dm(y')}{\int_{i_2}^{i_2+1} \phi(n^{-1}y')\, dm(y')} \\
&\leq \max_{y' \in [\frac{i_2}{n}, \frac{i_2+1}{n})} \frac{\phi(y' - \frac{(i_1 - i_2)}{n})}{\phi(y')} \\
&\leq e^{a\left(\frac{|i_1 - i_2|}{n} + \frac{k}{n}\right)} \leq e^{a\left(|x-y| + \frac{1}{n} + \frac{k}{n}\right)},
\end{aligned}$$
so the proof is complete. □

REMARK. Just as uniformly Lipschitz cones on $\mathbb{R}$ can be projected to uniformly Lipschitz cones on $S^1$ as specified in Lemma 2.1.4, discrete Lipschitz cones can be projected onto $S^1$. In Proposition 2.2 below, we will do this. □

In the next section, the two parts of Lemma 2.2.1 are combined to examine how quickly initial densities are spread out under iterative application of $\tilde{\mathcal{P}}_n$.

## 2.3 Iteration of discrete initial densities

The aim of this section is to lay the ground–work for proving an exponential rate of mixing for the operator $\mathcal{P}_n$. To simplify things later on, assume that the expansivity constant $\lambda > 2$. This restriction is not problematic, because any uniformly expanding circle map remains a uniformly expanding circle map under iteration, while the expansivity constant grows exponentially. Consequently, there is no loss of generality in replacing $T$ by a power of $T$ with $\lambda > 2$.

The Ulam Markov approximate operator $\mathcal{P}_n$ acts on density functions which are piecewise constant on subintervals of length $\frac{1}{n}$ of the unit circle. Each such density may be written as a convex combination of the basis functions

$$\psi_i = n\chi_{[i/n,(i+1)/n)} \quad i = 0, \ldots, n-1.$$

Therefore, to estimate how quickly an arbitrary initial density mixes to equilibrium under the action of $\mathcal{P}_n$, it suffices to estimate how long it takes for each $\psi_i$ to mix to equilibrium. This is accomplished via the lift.

REMARK. This is equivalent to analysing the speed at which each row of the stochastic matrix representation of $\mathcal{P}_n$ converges to the invariant probability vector under iterative multiplication. □

### 2.3.1 Control of central component

Fix an integer $i_0 \in \{0, \ldots, n-1\}$ and let $\phi_0 = \psi_{i_0} \circ \pi_0^{-1}$ (so $\phi_0$ is a function on $\mathbb{R}$). Let $A_0 = [\frac{i_0}{n}, \frac{i_0+1}{n})$ be the support of $\phi_0$, so that

$$\phi_0 \in \mathcal{C}_{\frac{K}{\lambda-1}, A_0, \frac{0}{n}}.$$

By Lemma 2.2.1 (1),

$$\tilde{\mathcal{L}}\phi_0 \in \mathcal{C}_{\frac{K}{\lambda-1}, \tilde{T}A_0, \frac{0}{n}}.$$

The next step we would like to take is to apply Lemma 2.2.1 (2) to find a cone containing $\tilde{\mathcal{P}}_n\phi_0 = \Pi_n \circ \tilde{\mathcal{L}}\phi_0$. Because $\tilde{T}A_0$ will not usually cover the $\frac{1}{n}$–subintervals containing its endpoints, Lemma 2.2.1 (1) may not be directly applicable. This difficulty can be avoided by restricting to a subset:

DEFINITION: Let $A \subset \mathbb{R}$ be a finite interval with $m(A) \geq \frac{2}{n}$. The *central component* of $A$ is defined as

$$A^{cen} = \bigcup \left\{ \left[ \frac{i}{n}, \frac{i+1}{n} \right) : i \in \mathbb{Z}, [i/n, (i+1)/n) \subset A \right\},$$

and the *covering component* for $A$ by

$$A^{cov} = \bigcup \left\{ \left[ \frac{i}{n}, \frac{i+1}{n} \right) : i \in \mathbb{Z}, [i/n, (i+1)/n) \cap A \neq \emptyset \right\}.$$

Obviously, $A^{cen} \subset A \subset A^{cov}$. □

For a function $\phi$, if $supp(\phi) = A$, then $supp(\Pi_n\phi) = A^{cov}$ and

$$\Pi_n\left(\phi|_{A^{cen}}\right) = (\Pi_n\phi)\Big|_{A^{cen}}.$$

We are now ready to deal with iterates of $\phi_0$. The key idea is that $\tilde{\mathcal{P}}_n^{\,k}\phi_0$ has a *central component* which is contained in a discrete Lipschitz cone. The remainder of this subsection is dedicated to the construction.

DEFINITION (**Cone $\mathcal{C}_k^*(i_0)$**) Let $i_0 \in \mathbb{Z}$, let $A_0 = [i_0/n, (i_0+1)/n) \subset \mathbb{R}$, and put

$$\mathcal{R}_0(i_0) = A_0^{cov} = A_0 \quad \text{and} \quad \mathcal{R}_0^*(i_0) = A_0^{cen} = A_0.$$

For each $k > 0$, put

$$\mathcal{R}_k(i_0) = \left(\tilde{T}\mathcal{R}_{k-1}(i_0)\right)^{cov} \quad \text{and} \quad \mathcal{R}_k^*(i_0) = \left(\tilde{T}\mathcal{R}_{k-1}^*(i_0)\right)^{cen}.$$

Finally, define a sequence of discrete Lipschitz cones:

$$\mathcal{C}_k^*(i_0) = \mathcal{C}_{\frac{K}{\lambda-1}, \mathcal{R}_k^*(i_0), \frac{k}{n}}. \quad \square$$

DEFINITION: For each $k \geq 0$, let

$$\phi_k = \left(\tilde{\mathcal{P}_n}^k \phi_0\right) \Big|_{\mathcal{R}_k^*(i_0)}.$$

Each $\phi_k$ is the *central component* of $\tilde{\mathcal{P}_n}^k \phi_0$. $\square$

**Proposition 2.1 (Control of $\phi_k$)** *For $n, i_0, \{\phi_k\}$ as above,*

$$supp\left(\tilde{\mathcal{P}_n}^k \phi_0\right) = \mathcal{R}_k(i_0) \quad \textit{and} \quad \phi_k \in \mathcal{C}_k^*(i_0).$$

*Proof:* The proof is by induction. For $k = 0$, the statement is true by definition. Suppose now that

$$supp\left(\tilde{\mathcal{P}_n}^l \phi_0\right) = \mathcal{R}_l(i_0) \quad \text{and} \quad \phi_l \in \mathcal{C}_l^*(i_0) = \mathcal{C}_{\frac{K}{\lambda-1}, \mathcal{R}_l^*(i_0), \frac{l}{n}}.$$

Obviously, $supp(\tilde{\mathcal{L}} \circ \tilde{\mathcal{P}_n}^l \phi_0) = \tilde{T}\mathcal{R}_l(i_0)$, and the remarks following the definition of the covering component imply that

$$supp\left(\tilde{\mathcal{P}_n}^{l+1} \phi_0\right) = supp\left(\Pi_n \circ \tilde{\mathcal{L}} \circ \tilde{\mathcal{P}_n}^l \phi_0\right) = \left(\tilde{T}\mathcal{R}_l(i_0)\right)^{cov} = \mathcal{R}_{l+1}(i_0).$$

Similarly

$$\mathcal{R}_{l+1}^*(i_0) = \left(\tilde{T}\mathcal{R}_l^*(i_0)\right)^{cen} \subset \left(\tilde{T}\mathcal{R}_l^*(i_0)\right)^{cov} = supp\left(\tilde{\mathcal{P}}_n \phi_l\right),$$

so that

$$\phi_{l+1} = \left(\tilde{\mathcal{P}_n}^{l+1} \phi_0\right) \Big|_{\mathcal{R}_{l+1}^*(i_0)} = \left(\tilde{\mathcal{P}}_n \phi_l\right) \Big|_{\mathcal{R}_{l+1}^*(i_0)} = \Pi_n \left(\left(\tilde{\mathcal{L}} \phi_l\right) \Big|_{\mathcal{R}_{l+1}^*(i_0)}\right).$$

Now, because $\phi_l \in \mathcal{C}_{\frac{K}{\lambda-1}, \mathcal{R}_l^*(i_0), \frac{l}{n}}$, Lemma 2.2.1 (1) implies that

$$\tilde{\mathcal{L}} \phi_l \in \mathcal{C}_{\frac{K}{\lambda-1}, \tilde{T}\mathcal{R}_l^*(i_0), \frac{l}{n}}$$

so that

$$\left(\tilde{\mathcal{L}} \phi_l\right) \Big|_{\mathcal{R}_{l+1}^*(i_0)} \in \mathcal{C}_{\frac{K}{\lambda-1}, \mathcal{R}_{l+1}^*(i_0), \frac{l}{n}}.$$

Because $\mathcal{R}^*_{l+1}(i_0)$ is an interval of the form $[i/n, j/n)$, Lemma 2.2.1 (2) implies that

$$\phi_{l+1} = \Pi_n \left( \left( \tilde{\mathcal{L}} \phi_l \right) \Big|_{\mathcal{R}^*_{l+1}(i_0)} \right) \in \mathcal{C}_{\frac{K}{\lambda - 1}, \mathcal{R}^*_{l+1}(i_0), \frac{l+1}{n}}.$$

The proposition follows by induction. □

**Corollary 2.3.1** *If $x, y \in \mathcal{R}^*_k(i_0)$ then*

$$\frac{\tilde{\mathcal{P}_n}^k \phi_0(x)}{\tilde{\mathcal{P}_n}^k \phi_0(y)} \leq e^{\frac{K}{\lambda - 1}\left(|x-y| + \frac{k}{n}\right)}.$$

### 2.3.2 Control of non–central components

By Proposition 2.1, the central component of $\tilde{\mathcal{P}_n}^k \phi_0$ has a regular structure. In the next section, it is shown how the regularity guaranteed by the discrete cone condition can be used to estimate a lower bound for the density $\mathcal{P}_n{}^k \phi_0 \circ \pi$ on the circle. First, we obtain control of $\tilde{\mathcal{P}_n}^k \phi_0$ on all of $\mathcal{R}_k(i_0)$, rather than just that part which is supported on $\mathcal{R}^*_k(i_0)$. This is done by decomposing $\mathcal{R}_k(i_0)$ into sets like $\mathcal{R}^*_k(i_0)$.

**Lemma 2.3.2 (Decomposition of $\mathcal{R}_k(i_0)$)** *Fix $k$. For each $j = 0, \ldots, k$, there exists a finite collection of intervals in $\mathbb{R}$, indexed by $\{i_r^{(j)}\}_{r=1}^{r_j} \subset \mathbb{Z}$ such that*

1. *$r_j \leq 2^j$;*
2. *each $[i_r^{(j)}/n, (i_r^{(j)}+1)/n) \subset \mathcal{R}_j(i_0)$ (we make the obvious labelling $i_r^{(0)} = i_0$ for $r = r_0 = 1$);*
3. *and*

$$\mathcal{R}_k(i_0) = \bigcup_{j=0}^{k} \bigcup_{r=1}^{2^j} \mathcal{R}^*_{k-j}(i_r^{(j)}).$$

*We call each $\mathcal{R}^*_{k-j}(i_r^{(j)})$ a* component *of the decomposition.*

*Proof:* The lemma is true more or less by definition. To help understand the decomposition, we elaborate the construction.

First of all, if $A_1$ and $A_2$ are two adjacent intervals in $\mathbb{R}$, then either $\exists\, p \in \mathbb{Z}$ such that

$$A_1^{cov} \cap A_2^{cov} = \left[ \frac{p}{n}, \frac{p+1}{n} \right)$$

or the intersection is empty. Hence, if $A_1, \dots A_l$ is a family of adjacent intervals, then

$$\begin{aligned}(A_1 \cup \ldots \cup A_l)^{cov} &= \left[\frac{p_0}{n}, \frac{p_0+1}{n}\right) \cup A_1^{cen} \cup \left[\frac{p_1}{n}, \frac{p_1+1}{n}\right) \cup A_2^{cen} \ldots \\ &\quad \ldots \cup A_{l-1}^{cen} \cup \left[\frac{p_{l-1}}{n}, \frac{p_{l-1}+1}{n}\right) \cup A_l^{cen} \cup \left[\frac{p_l}{n}, \frac{p_l+1}{n}\right)\end{aligned}$$

where each $[p_r/n, (p_r+1)/n) = A_r^{cov} \cap A_{r+1}^{cov}$ (with the obvious interpretation where the intersection is empty).

The decomposition can now be constructed inductively. Set $i_1^{(0)} = i_0$ and let

$$\mathcal{R}_1(i_0) \setminus \mathcal{R}_1^*(i_0) = [i_1^{(1)}/n, (i_1^{(1)}+1)/n) \cup [i_2^{(1)}/n, (i_2^{(1)}+1)/n),$$

so that

$$\mathcal{R}_1(i_0) = \mathcal{R}_0^*(i_1^{(1)}) \cup \mathcal{R}_1^*(i_1^{(0)}) \cup \mathcal{R}_0^*(i_2^{(1)}).$$

This provides a basis for the induction.

Suppose the conclusion of the lemma is satisfied for $j = l$. Then the decomposition contains at most $2^{l+1} - 1$ intervals:

$$\mathcal{R}_l(i_0) = \bigcup_{j=0}^{l} \bigcup_{r=1}^{2^j} \mathcal{R}_{l-j}^*(i_r^{(j)}).$$

By the remarks at the beginning of the proof, there exists a finite sequence $\{i_r^{(l+1)}\} \subset \mathbb{Z}$ containing at most $2^{l+1}$ elements such that

$$\left(\tilde{T}\mathcal{R}_l(i)\right)^{cov} = \left\{\bigcup_{j=0}^{l} \bigcup_{r=1}^{2^j} \left(\tilde{T}\mathcal{R}_{l-j}^*(i_r^{(j)})\right)^{cen}\right\} \cup \left\{\bigcup_{r=1}^{2^{l+1}} \mathcal{R}_0^*(i_r^{(l+1)})\right\}.$$

This equation can be rewritten to be exactly the conclusion of the lemma for $j = l+1$, so the result follows by induction. □

Having decomposed $\mathcal{R}_k(i_0)$ into components on which Proposition 2.1 holds, the final step before projecting back to $S^1$ is to identify those components containing "complete copies" of the circle $S^1$.

**Lemma 2.3.3** *Fix $k$. There exist disjoint subsets $L_1, L_2 \subset \mathbb{Z}$ (depending on $i_0$ and $k$) with the following properties:*

1. *if $x \in \mathcal{R}_k(i_0)$, then $\exists l \in L_1 \cup L_2$ such that $x \in [l, l+1)$;*
2. *if $l \in L_1$ then there exists a unique $i = i(l)$ such that*
$$i(l) \in \bigcup_{j=0}^{k} \bigcup_{r=0}^{2^j} \{i_r^{(j)}\} \text{ and } [l, l+1) \subset \mathcal{R}^*_{k-j}(i(l));$$
3. $\mathrm{card}\{L_2\} \leq 2^{k+1}$;
4. $\left\| \left( \tilde{\mathcal{P}_n}^k \phi_0 \right) \Big|_{\Lambda_k} \right\| \geq \left(1 - n2^{k+1}\lambda^{-k}\right)$, *where* $\Lambda_k = \bigcup_{l \in L_1} [l, l+1)$.

*Proof:* Put

$$L_1 = \left\{ \begin{array}{r} l \in \mathbb{Z}: \quad \exists j \leq k, i(l) \in \mathbb{Z} \text{ such that } [i(l)/n, (i(l)+1)/n) \subset \mathcal{R}_j(i_0) \\ \text{and } \pi_l^{-1} S^1 \subset \mathcal{R}^*_{k-j}(i(l)) \end{array} \right\}$$

and

$$L_2 = \left\{ l \in \mathbb{Z} : \mathcal{R}_k(i_0) \cap \pi_l^{-1} S^1 \neq \emptyset \right\} \setminus L_1.$$

The set $L_1$ indexes those lifted copies of $S^1$ which are wholly contained in one of the components $\mathcal{R}^*_{k-j}(i_r^{(j)})$ of the decomposition in Lemma 2.3.2, while $L_2$ indexes those copies which intersect more than one component. Properties (1) and (2) are obvious from the definition.

To see why (3) is true, notice that $L_2$ consists of those $l$ such that the "copy" $\pi_l^{-1} S^1$ of $S^1$ overlaps more than one interval in the decomposition in Lemma 2.3.2. Because $\mathcal{R}_k(i_0)$ can be written both as a sequence of adjacent intervals from the above decomposition, and as a sequence of adjacent copies of $S^1$, the number of copies of $S^1$ which overlap more than one interval must be bounded by the number of boundary points of intervals from the decomposition. Because there are no more than $2^{k+1} - 1$ subintervals in the decomposition, (3) follows.

Finally,

$$\left\| \left( \tilde{\mathcal{P}_n}^k \phi_0 \right) \Big|_{\Lambda_k} \right\| = 1 - \sum_{l \in L_2} \left\| \left( \tilde{\mathcal{P}_n}^k \phi_0 \right) \Big|_{[l, l+1)} \right\|.$$

Because $\phi_0 \leq n$ and $|\tilde{T}'| \geq \lambda$, it follows that

$$\tilde{\mathcal{P}_n}^k \phi(x) \leq n\lambda^{-k} \quad \forall\, x \in \mathbb{R}.$$

By Part (3) of the lemma, the sum is bounded above by $2^{k+1} \times n \times \lambda^{-k}$. □

DEFINITION: Put

$$\phi_{i_0,k} = \left( \tilde{\mathcal{P}_n}^k \phi_0 \right) \Big|_{\Lambda_k}$$

and let $\psi_{i_0,k}$ be the projection of $\phi_{i_0,k}$ onto $S^1$. □

**Proposition 2.2 (Properties of $\psi_{i_0,k}$)** *For every $k$, the function $\psi_{i_0,k}$ has the following properties:*

1. $\|\psi_{i_0,k}\| \geq \left(1 - n2^{k+1}\lambda^{-k}\right)$;
2. $\psi_{i_0,k} \leq {\mathcal{P}_n}^k \psi_{i_0}$.

   *If $k > \log(2n)/\log(\lambda/2)$, then also:*
3. *for every $x, y \in S^1$,*
$$\frac{\psi_{i_0,k}(x)}{\psi_{i_0,k}(y)} \leq e^{\frac{K}{\lambda-1}\left(1+\frac{k}{n}\right)}.$$

*Proof:* The norm estimate follows directly from Lemma 2.3.3 (4) because $\psi_{i_0,k}$ is the projection of $\phi_{i_0,k}$ onto $S^1$.

For the second part, note that by Lemma 2.3.3,

$$supp\left(\phi_{i_0,k}\right) = \left\{ \bigcup_{l \in L_1} [l, l+1) \right\} \subset \mathcal{R}_k(i_0) = supp\left(\tilde{\mathcal{P}_n}^k \phi_0\right).$$

Consequently, by equation (2.3),

$$\begin{aligned} \psi_{i_0,k} = \sum_{l \in L_1} \phi_{i_0,k} \circ \pi_l^{-1} &= \sum_{l \in L_1} \left(\tilde{\mathcal{P}_n}^k \phi_0\right) \circ \pi_l^{-1} \\ &\leq \sum_{l \in \mathbb{Z}} \left(\tilde{\mathcal{P}_n}^k \phi_0\right) \circ \pi_l^{-1} \\ &= {\mathcal{P}_n}^k \left( \sum_{l \in \mathbb{Z}} \phi_0 \circ \pi_l^{-1} \right) = {\mathcal{P}_n}^k \psi_{i_0}. \end{aligned}$$

Next, if $k$ is sufficiently large (as specified) then it follows from Part (1) of the proposition that $\psi_{i_0,k} > 0$. Since $\psi_{i_0,k}$ is the lift of $\tilde{\mathcal{P}_n}^k \phi_0 \Big|_{\Lambda_k}$, it follows that $\Lambda_k \neq \emptyset$. Therefore, the set $L_1$ of indices is non–empty.

Now, for each $l \in L_1$, there exists (by Lemma 2.3.3) a unique $j \leq k$ and $i(l) \in \mathbb{Z}$ such that

$$[l, l+1) \subset \mathcal{R}^*_{k-j}(i(l)).$$

Let $\phi_{i(l),j}$ denote the restriction of $\tilde{\mathcal{P}_n}^j \phi_0$ to the interval $[i(l)/n, (i(l)+1)/n)$. Then

$$\left(\tilde{\mathcal{P}_n}^k \phi_0\right) \Big|_{[l,l+1)} = \left(\tilde{\mathcal{P}_n}^{k-j} \phi_{i(l),j}\right) \Big|_{[l,l+1)}$$

and Proposition 2.1 implies that

$$\left(\tilde{\mathcal{P}_n}^{k-j} \phi_{i(l),j}\right) \Big|_{\mathcal{R}^*_{k-j}(i(l))} \in \mathcal{C}^*_{k-j}(i(l)).$$

Therefore, if $x_l, y_l \in [l, l+1)$ then by Corollary 2.3.1

$$\frac{\phi_{i_0,k}(x_l)}{\phi_{i_0,k}(y_l)} = \frac{\tilde{\mathcal{P}_n}^{k-j}\phi_{i(l),j}(x_l)}{\tilde{\mathcal{P}_n}^{k-j}\phi_{i(l),j}(y_l)} \leq e^{\frac{K}{\lambda-1}(|x_l-y_l|+\frac{k-j}{n})} \leq e^{\frac{K}{\lambda-1}(1+k/n)}.$$

Since $\pi_l^{-1}(x), \pi_l^{-1}(y) \in [l, l+1)$ whenever $x, y \in S^1$,

$$\frac{\psi_{i_0,k}(x)}{\psi_{i_0,k}(y)} = \frac{\sum_{l\in L_1}\phi_{i_0,k}\circ\pi_l^{-1}(x)}{\sum_{l\in L_1}\phi_{i_0,k}\circ\pi_l^{-1}(y)} \leq \max_{l\in L_1}\frac{\phi_{i_0,k}\circ\pi_l^{-1}(x)}{\phi_{i_0,k}\circ\pi_l^{-1}(y)} \leq e^{\frac{K}{\lambda-1}(1+k/n)}.$$

This completes the proof of the proposition. □

### 2.3.3 Uniform lower bounds

Fix $k_n = \lceil \frac{\log 4n}{\log \lambda/2} \rceil$, and let $\psi \in \mathcal{D}_n(S^1)$ be non–negative.

**Corollary 2.3.4 (Lower bound for iterates of $\psi$)** *Let $n > 0$. Then if $\psi \in \mathcal{D}_n$ and $k > \log(2n)/\log(\lambda/2)$,*

$$\mathcal{P}_n{}^k\psi \geq c_{n,k},$$

*where*

$$c_{n,k} = e^{-\frac{K}{\lambda-1}(1+k/n)}\left(1 - n2^{k+1}\lambda^{-k}\right).$$

*In particular, $c_{n,k} \geq \delta = e^{-3K/(\lambda-1)}/2$ whenever $k > k_n$, and $c_{n,k_n} \to e^{-K/(\lambda-1)}/2$ exponentially fast as $n$ increases.*

*Proof:* Because $\psi$ is constant on each interval $[i/n, (i+1)/n) \subset S^1$, it suffices to establish the corollary for each basis function $\psi_{i_0} \in \mathcal{D}_n(S^1)$.

By the Mean Value Theorem and Part (1) of the proposition, there exists $x_0 \in S^1$ such that

$$\psi_{i_0,k}(x_0) \geq \|\psi_{i_0,k}\| \geq 1 - n2^{k+1}\lambda^{-k}.$$

By Part (3) of the proposition,

$$\psi_{i_0,k}(x) \geq e^{-\frac{K}{\lambda-1}(1+k/n)}\psi_{i_0,k}(x_0)$$

for each $x \in S^1$. The first part of the corollary now follows from Proposition 2.2 (2). The other part is easy, because the choice of $k_n$ guarantees that $(1 - n2^{k+1}\lambda^{-k}) \geq \frac{1}{2}$ whenever $k \geq k_n$. □

Corollary 2.3.4 provides the required "lower bound", and we are now in a position to derive a rate of mixing for $\mathcal{P}_n$ and an error bound for the invariant density approximation.

## 2.4 Rate of mixing and an error bound Theorem

We use the method of *lower bounds for invariant cones* to get a rate of mixing estimate for ${\mathcal{P}_n}^k$.

### 2.4.1 Rate of mixing

Recall that $\mathcal{D}_n(S^1)$ is the collection of positive functions which are piecewise constant on subintervals of $S^1$ of length $\frac{1}{n}$. Denote by $\Delta_n$ the collection of functions which are piecewise constant on each subinterval, but whose mean is zero. Then

$$\Delta_n = \left\{ \phi : S^1 \to \mathbb{R} : \int_{S^1} \phi \, dm = 0, \quad \phi^+, \phi^- \in \mathcal{D}_n \right\}.$$

**Theorem 2.1 (Rate of mixing for $\mathcal{P}_n$)** *For each $n > 0$, let $\mathcal{P}_n$ be the nth Ulam approximate operator and let $c_{n,k}$ be as in Corollary 2.3.4. Then for every $\phi \in \Delta_n$,*

$$\|{\mathcal{P}_n}^k \phi\| \leq (1 - c_{n,k}) \|\phi\|,$$

*whenever $k > \log 2n / \log(\lambda/2)$.*

*Proof:* It is easy to check that $\mathcal{D}_n(S^1)$ equipped with the usual norm $\|\cdot\|$ is a cone. Let $\Gamma_{\mathcal{D}_n}$ be the corresponding *difference cone*. Then

$$\Gamma_{\mathcal{D}_n} = \{ f - g : f, g \in \mathcal{D}_n(S^1) \text{ and } \|f\| = \|g\| \} = \Delta_n$$

and obviously

$$\|\phi\|_{\Gamma_{\mathcal{D}_n}} = \|\phi\|/2 \quad \forall \phi \in \Delta_n. \tag{2.4}$$

Now, let $f \in \mathcal{D}_n(S^1)$ so that by Corollary 2.3.4,

$${\mathcal{P}_n}^k f \geq c_{n,k} \|f\|.$$

Then clearly,

$${\mathcal{P}_n}^k f - c_{n,k} \|f\| \in \mathcal{D}_n(S^1) \quad \forall f \in \mathcal{D}_n(S^1),$$

so that putting $\mathcal{T} = {\mathcal{P}_n}^k$ in Theorem 1.1,

$$\|{\mathcal{P}_n}^k \phi\|_{\Gamma_{\mathcal{D}_n}} \leq (1 - c_{n,k}) \|\phi\|_{\Gamma_{\mathcal{D}_n}}$$

for every $\phi \in \Delta_n$. The theorem follows from (2.4). □

### 2.4.2 Error bounds for Ulam's method

Before the error bound theorem, we need one more lemma.

**Lemma 2.4.1** *For each $n > 0$ let $\Pi_n$ be the projection onto $\mathcal{D}_n(S^1)$. Then for any function $f \in \mathcal{C}_a$,*

$$\|f - \Pi_n f\| \leq \left(e^{a/n} - 1\right) \|f\|.$$

*Proof:* Since $f \in \mathcal{C}_a$, $f$ is continuous by Lemma 2.1.1. Therefore, by the Mean Value Theorem, for each $i = 0, \ldots, n-1$ there exists $x_i \in [i/n, (i+1)/n)$ such that

$$f(x_i) = n \int_{i/n}^{(i+1)/n} f \, dm = (\Pi_n f)(x)$$

for each $x \in [i/n, (i+1)/n)$. For each such $x$,

$$\begin{aligned} |(f - \Pi_n f)(x)| = |f(x) - f(x_i)| &\leq |f(x_i)| \left(e^{a|x - x_i|} - 1\right) \\ &\leq |(\Pi_n f)(x)| \left(e^{a/n} - 1\right). \end{aligned}$$

The lemma follows because $\|\Pi_n f\| \leq \|f\|$. □

**Theorem 2.2 (Error bounds for Ulam's method)** *Let $T : S^1 \to S^1$ satisfy (2.1) and (2.2), where the expansivity constant $\lambda > 2$. Then if $h$ is the unique invariant density for the transformation $T$, and $h_n$ is the normalised fixed point of the nth Ulam approximation,*

$$\|h - h_n\| \leq \left( k \frac{e^{\frac{K}{\lambda - 1}(1 + k/n)}}{1 - n 2^{k+1} \lambda^{-k}} + 1 \right) \left(e^{\frac{1}{n} \frac{K}{\lambda - 1}} - 1\right)$$

*whenever $k \geq \lceil \frac{\log 2n}{\log \lambda/2} \rceil$. In particular, by putting $k = k_n = \lceil \frac{\log 4n}{\log \lambda/2} \rceil$, there exists a constant $C$, independent of $n$, such that*

$$\|h - h_n\| \leq C \frac{\log n}{n}.$$

*Proof:* Recall that $\mathcal{P}_n = \Pi_n \circ \mathcal{L}$, where $\mathcal{L}$ is the Perron–Frobenius operator for $T$, and $\Pi_n$ is the projection onto $\mathcal{D}_n$. Since $h \in \mathcal{C}_{K/(\lambda-1)}$, Lemma 2.4.1 implies that

$$\|h - \Pi_n h\| \leq \left(e^{\frac{1}{n} \frac{K}{\lambda - 1}} - 1\right)$$

and since $\mathcal{P}_n h = \Pi_n \circ \mathcal{L} h = \Pi_n h$,

$$\begin{aligned} \|\Pi_n h - h_n\| &= \|\mathcal{P}_n h - h_n\| \\ &\leq \sum_{i=1}^{k} \left\|\mathcal{P}_n{}^i (h - \mathcal{P}_n h)\right\| + \left\|\mathcal{P}_n{}^k (\mathcal{P}_n h - h_n)\right\| \\ &\qquad + \left\|\mathcal{P}_n{}^k h_n - h_n\right\| \\ &\leq \sum_{i=1}^{k} \|h - \mathcal{P}_n h\| + (1 - c_{n,k}) \|\mathcal{P}_n h - h_n\| + 0 \\ &\leq k \left(e^{\frac{1}{n} \frac{K}{\lambda - 1}} - 1\right) + (1 - c_{n,k}) \|\Pi_n h - h_n\|. \end{aligned}$$

Therefore

$$\|\Pi_n h - h_n\| \leq \frac{k}{c_{n,k}} \left( e^{\frac{1}{n}\frac{K}{\lambda-1}} - 1 \right),$$

and the first part of the theorem follows.

The second part follows by putting $k = k_n$. The first term on the right is then bounded by $2e^{3K/(\lambda-1)} \left( \frac{\log 4n}{\log \lambda/2} + 1 \right)$ (c.f. Corollary 2.3.4), while the second term is obviously bounded by $\left( e^{K/(\lambda-1)} - 1 \right) \frac{1}{n}$. □

Theorem 2.2 is the sought after *error bound theorem*. Before proceeding to some examples, notice that the theorem has been stated without explicitly writing down the constant $C$. For the purposes of description, it is convenient to be able to say "the approximation error is bounded by $C \log n/n$", but for actually getting error bounds, it is desirable to get the tightest bound possible:

For each $z > z_0 = \log 2n / \log(\lambda/2)$ let

$$e_n(z) = \left( z \frac{e^{\frac{K}{\lambda-1}(1+z/n)}}{1 - n2^{z+1}\lambda^{-z}} + 1 \right) \left( e^{\frac{1}{n}\frac{K}{\lambda-1}} - 1 \right).$$

Then Theorem 2.2 says that whenever $k > z_0$,

$$\|h - h_n\| \leq e_{n,k} = e_n(k).$$

To obtain the optimal bound, one must therefore minimise $e_n(z)$ on the interval $(z_0, \infty) \cap \mathbb{Z}$. Since the function $e_n(z) : (z_0, \infty) \to \mathbb{R}^+$ is convex, a simple procedure for minimising $e_{n,k}$ is to calculate $e_n(k)$ for each $k \in \{\lceil z_0 \rceil, \lceil z_0 \rceil + 1, \ldots\}$ until $e_n(k)$ stops decreasing. By convexity, the minimal $k_*$ for which $e_n(k_*) \leq e_n(k_* + 1)$ will minimise $e_{n,k}$.

REMARK 2.4.2. When it comes to calculating the fixed point of $\mathcal{P}_n$, the same argument as in the proof of Theorem 2.2 can be used to account for numerical errors: suppose that $h'_n$ is such that

$$\|h'_n - \mathcal{P}_n h'_n\| = \epsilon \tag{2.5}$$

for some small $\epsilon$. Then, as in the proof of the theorem,

$$\|h_n - h'_n\| \leq \frac{k}{c_{n,k}} \epsilon.$$

Because equation (2.5) is easily enforced, $h_n$ can be approximated to any desired level of accuracy. □

### 2.4.3 Examples

Here, we present several examples. Although a little contrived, they serve well as illustrations of the usefulness and weaknesses of the error bounds obtained above. More examples will be presented in the next two chapters.

EXAMPLE 1: (THE TRIVIAL CASE) The canonical expanding circle maps are $z \mapsto z^m$, for $m \geq 2$. Each such map preserves Lebesgue measure on the circle, so $h = 1$ is the invariant density. Moreover, the lift is a linear function for which the distortion constant $K = 0$. By Theorem 2.2, the $n$th Ulam approximation must always be identically equal to the Lebesgue measure. □

EXAMPLE 2: (A LITTLE CONTRIVED) We now consider making small perturbations to the canonical expanding circle maps. We pick $m = 5$ and let $S_5 : x \mapsto 5x (\mathrm{mod}\ 1)$ be the corresponding interval map. Next, for each $\epsilon > 0$ put

$$g_\epsilon(x) = (1+\epsilon)x - 4\epsilon(x - \frac{1}{2})^3 - \epsilon/2,$$

and

$$T_\epsilon(x) = g_\epsilon^{-1} \circ S_5 \circ g_\epsilon(x) \quad \forall x \in [0,1).$$

Because $S_5$ preserves Lebesgue measure (with density 1), it follows easily that each $T_\epsilon$ has

$$h_\epsilon(x) = g_\epsilon'(x) = 1 + \epsilon - 12\epsilon(x - \frac{1}{2})^2 \tag{2.6}$$

as an invariant density.

In order to apply Theorem 2.2, we need to know about the constants $\lambda_\epsilon > 2$ and $K_\epsilon \geq 0$ corresponding to $T_\epsilon$. Obviously

$$1 - 2\epsilon \leq g_\epsilon' \leq 1 + \epsilon$$

so that

$$T_\epsilon' = (g_\epsilon^{-1})' \circ S_5 \circ g_\epsilon \times 5 \times g_\epsilon' = 5\frac{h_\epsilon}{h_\epsilon \circ T_\epsilon} \geq 5\frac{1-2\epsilon}{1+\epsilon}.$$

Next, recall that $K_\epsilon$ is the minimal Lipschitz constant of

$$\log\left|T_\epsilon'\right| = \log\left|(g_\epsilon^{-1})' \circ S_5 \circ g_\epsilon\right| + \log 5 + \log\left|g_\epsilon'\right|$$

and by differentiation

$$\left|\frac{T_\epsilon''}{T_\epsilon'}\right| = \left|\frac{h_\epsilon'}{h_\epsilon} - 5\frac{h_\epsilon}{h_\epsilon \circ T_\epsilon}\frac{h_\epsilon' \circ T_\epsilon}{h_\epsilon \circ T_\epsilon}\right| \leq \frac{36\epsilon(2+\epsilon)}{(1-2\epsilon)^2}.$$

Hence,

$$\lambda_\epsilon \geq 5\frac{1-2\epsilon}{1+\epsilon} \quad \text{and} \quad K_\epsilon \leq \max\left|\frac{T_\epsilon''}{T_\epsilon'}\right| \leq \frac{36\epsilon(2+\epsilon)}{(1-2\epsilon)^2}.$$

Then, provided $\epsilon < 1/4$, we can be sure that $\lambda_\epsilon > 2$ and Theorem 2.2 can be applied.

For selected values of $\epsilon$, the Ulam approximations have been calculated on partitions of $n = 100, 1000, 10000$ equal subintervals of $[0,1)$. The error data are presented in Table 2.1. Because each $T_\epsilon$ has been constructed so that the invariant density is given by (2.6), it easy to calculate the exact $L^1$ error for each Ulam approximation. These are displayed in the columns headed "Error". Theorem 2.2 was used to get *a priori* error bounds. The procedure there was

| $\epsilon$ | $n = 100$ | | $n = 1000$ | | $n = 10000$ | |
|---|---|---|---|---|---|---|
| | Error | Bound | Error | Bound | Error | Bound |
| 0.00 | $4.42E{-}8$ | $9.21E{-}7$ | $4.00E{-}8$ | $1.20E{-}6$ | $4.56E{-}8$ | $1.48E{-}6$ |
| 0.02 | $3.92E{-}4$ | $7.15E{-}2$ | $3.41E{-}5$ | $8.88E{-}3$ | $3.39E{-}6$ | $1.08E{-}3$ |
| 0.04 | $6.69E{-}4$ | $3.39E{-}1$ | $6.52E{-}5$ | $4.05E{-}2$ | $6.72E{-}6$ | $4.88E{-}3$ |
| 0.06 | $9.99E{-}4$ | $1.60E0$ | $9.88E{-}5$ | $1.77E{-}1$ | $1.01E{-}5$ | $2.12E{-}2$ |
| 0.08 | $1.39E{-}3$ | $9.91E0$ | $1.37E{-}4$ | $9.65E{-}1$ | $1.35E{-}5$ | $1.14E{-}1$ |

Table 2.1: Error bounds for invariant density calculations for selected maps $T_\epsilon$. For each $T_\epsilon$, the Ulam approximation was calculated on partitions of $[0,1)$ into $n = 100, 1000, 10000$ equal subintervals. Because the invariant density is known for each $T_\epsilon$, the accuracy of the approximation can be computed exactly; these errors are displayed in the columns headed "Error". The best possible error bound from Theorem 2.2 is displayed in the adjacent column headed "Bound". The notation $4.42E{-}8$ should be read $4.42 \times 10^{-8}$.

to use the values of $K_\epsilon$ and $\lambda_\epsilon$ above, and find the optimal $k_\epsilon$ such that $e_{n,k}$ is minimal. This yields the error bounds displayed in the columns headed "Bounds".

When $\epsilon = 0.0$, the $n$th Ulam approximation should reproduce Lebesgue measure (c.f. Example 1 above). However, our numerical implementation finds a density $h'_n$ satisfying

$$\|h'_n - \mathcal{P}_n h'_n\| \leq 10^{-7}.$$

Hence, it is not necessarily the case that $h_n = h'_n$, but we can still obtain error bounds as in Remark 2.4.2.

Finally, notice how the error bounds from Theorem 2.2 get progressively worse as $\epsilon$ increases. This is largely because the bounds in the theorem are extremely sensitive to the distortion constant $K$. As $\epsilon$ increases, the nonlinearity in the map becomes greater, making the bounds from Theorem 2.2 less useful. □

EXAMPLE 3: (LESS CONTRIVED) The final example in this section is also a nonlinear perturbation of a canonical expanding circle map. For each $\Omega \in [0,1)$ let

$$T_\Omega(x) = 2x + \Omega + \frac{\sin 2\pi x}{4\pi} \pmod 1.$$

Then each $T_\Omega$ is an expanding circle map, and the invariant densities are unknown. By differentiation,

$$3/2 \leq \lambda_\Omega \leq 5/2$$

and $T''_\Omega / T'_\Omega$ is maximised when $\cos 2\pi x = \frac{-1}{4}$. Hence $K_\Omega \leq \frac{2\pi}{\sqrt{15}}$. However, Theorem 2.2 requires that $\lambda > 2$. Therefore, we must replace $T_\Omega$ by $T^2_\Omega$. Differentiating,

$$\frac{(T^2_\Omega)''}{(T^2_\Omega)'} = \frac{T''_\Omega \circ T_\Omega}{T'_\Omega \circ T_\Omega} T'_\Omega + \frac{T''_\Omega}{T'_\Omega} \leq (5/2+1)K_\Omega \leq \frac{7\pi}{\sqrt{15}}.$$

| | $n = 10^4$ | $n = 10^5$ | $n = 10^6$ |
|---|---|---|---|
| Error bound | $5.14 \times 10^0$ | $5.83 \times 10^{-1}$ | $6.82 \times 10^{-2}$ |

Table 2.2: Error bounds for $n$th Ulam approximation to the invariant densities for $T_\Omega^2$.

Because the constants $\lambda = 2.25$, $K = \frac{7\pi}{\sqrt{15}}$ are independent of $\Omega$, the optimal error bounds from Theorem 2.2 are the same for each $\Omega$. For $n = 10^4, 10^5, 10^6$ these are calculated as above and presented in Table 2.2.

Notice that the error bounds for this example are very bad in comparison with the previous example. This is because the conclusion of Theorem 2.2 is exponentially sensitive to the distortion constant $K$. In fact, the bound for $n = 10000$ contains no information whatsoever, because the maximum possible distance between any two probability measures is 2. As we shall see in the next chapter, much better estimates are possible using *bounded variation* techniques. There, we are able to show that the $n = 10000$ approximations are in very good agreement with the exact invariant densities for the maps $T_\Omega$. For a grey–scale plot of the Ulam approximate densities with $n = 10000$ for a selection of values of $\Omega$, see Chapter 3. □

### 2.4.4 Working at the matrix level

Recall that the $n$th Ulam approximate operator $\mathcal{P}_n$ can be represented by the $n \times n$ stochastic matrix with entries:

$$P_{ij} = \frac{m\left(\left[\frac{i-1}{n}, \frac{i}{n}\right) \cap T^{-1}\left[\frac{j-1}{n}, \frac{j}{n}\right)\right)}{m\left(\left[\frac{i-1}{n}, \frac{i}{n}\right)\right)}.$$

Any initial function $\psi \in \mathcal{D}_n(S^1)$ can be written as

$$\psi = n \sum_{i=1}^{n} \psi_i \chi_{[(i-1)/n, i/n)}$$

where $\sum_{i=1}^n \psi_i = 1$. Then

$$\mathcal{P}_n \psi = n \sum_{i,j=1}^{n} \psi_i P_{ij} \chi_{[(j-1)/n, j/n)}.$$

In particular, if $\psi = n\chi_{[(i_0-1)/n, i_0/n)}$, then the coefficients of ${\mathcal{P}_n}^k \psi$ are precisely the elements of the $i_0$th row of the matrix $P^k$. Therefore, all of the analysis in the preceeding three sections can be translated into statements about the *regularity* of the rows of the matrices $P^k$. This is the emphasis in [39]. The following proposition is a sample from the results one can expect for the matrix versions:

**Proposition 2.3 (Rows of $P^k$)** *Let $n$ be fixed and let $T$ be a uniformly expanding circle map with $\lambda > 2$. Let $P(n) = (P_{ij})$ be the matrix representation of the $n$th Ulam approximation, and suppose that $k > \log 2n / \log(\lambda/2)$. Then for each $i_0 = 1 \dots n$ and $j_1, j_2 = 1 \dots n$,*

$$\frac{(P^k)_{i_0 j_1}}{(P^k)_{i_0 j_2}} \leq 1/c_{n,k},$$

*where $c_{n,k}$ is defined in Corollary 2.3.4.*

*Proof:* Let $\psi_{i_0} = n\chi_{[(i_0-1)/n, i_0/n)}$, and let $\phi_{i_0}$ be it's lift. Fix $k$, let $L_1, L_2$ be the two subsets from Lemma 2.3.3 and let $\psi_{i_0,k}$ be as in Proposition 2.2. Let $\mathcal{P}_n$ be the $n$th Ulam approximate operator, and $\tilde{\mathcal{P}}_n$ its lift. Then

$$\begin{aligned}
\frac{(P^k)_{i_0 j_1}}{(P^k)_{i_0 j_2}} &= \frac{{\mathcal{P}_n}^k \psi_{i_0}(j_1/n)}{{\mathcal{P}_n}^k \psi_{i_0}(j_2/n)} \\
&= \frac{\sum_{l \in L_1 \cup L_2} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_1/n + l)}{\sum_{l \in L_1 \cup L_2} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_2/n + l)} \\
&\leq \frac{\sum_{l \in L_1} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_1/n + l)}{\sum_{l \in L_1} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_2/n + l)} + \frac{\sum_{l \in L_2} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_1/n + l)}{\sum_{l \in L_1} {\tilde{\mathcal{P}}_n}^k \phi_{i_0}(j_2/n + l)} \\
&\leq \frac{\psi_{i_0,k}(j_1/n)}{\psi_{i_0,k}(j_2/n)} + \frac{2^{k+1} \times n \times \lambda^{-k}}{\psi_{i_0,k}(j_2/n)}
\end{aligned}$$

where the last inequality follows from the definition of $\psi_{i_0,k}$ and Lemma 2.3.3. From Proposition 2.2 and Corollary 2.3.4 we have

$$\frac{\psi_{i_0,k}(j_1/n)}{\psi_{i_0,k}(j_2/n)} \leq e^{\frac{K}{\lambda - 1}\left(1 + \frac{k}{n}\right)}$$

and

$$\psi_{i_0,k} \geq e^{-\frac{K}{\lambda-1}\left(1+\frac{k}{n}\right)} \left(1 - 2n(2/\lambda)^k\right) = c_{n,k}.$$

Putting this together, the proposition follows. □

Just as Theorem 2.1 was deduced from Proposition 2.2, an analogous mixing result for the matrix $P(n)$ follows from Proposition 2.3. By working with a matrix representation of $\mathcal{L}$, Theorem 2.2 can be derived in this setting; see [39] for details.

In Chapter 4, we will return to the analysis of matrix representations to construct a *numerical* approach to obtaining error bounds for Ulam's method. The purpose of this section has been to show that the behaviour of the Ulam approximation can be written down in terms of the stochastic matrix representations.

## 2.5 Structure of the argument

Some of the arguments used in this chapter to obtain error bounds for Ulam's method will be repeated in subsequent chapters. We now conclude by emphasising this structure.

First of all, recall that the Ulam approximations $\mathcal{P}_n$ are small perturbations of the Perron–Frobenius operator $\mathcal{L}$ in the usual $\|\cdot\|$–norm on $L^1$. Indeed, for any Lipschitz function $f$,

$$\|\mathcal{P}_n f - \mathcal{L}f\| = \|\Pi_n \circ \mathcal{L}f - \mathcal{L}f\| = \|(\Pi_n - Id)\mathcal{L}f\| = O(1/n).$$

Later on, we will see that the same order of perturbative error holds for functions of *bounded variation.*

The other part in the proof of the error bounds is the estimation of a rate of mixing for the Ulam approximate operator. This is the part which uses the method of lower bounds for invariant cones. Here, by employing *discrete Lipschitz cones*, we were able to construct lower bound functions for iterates of $\mathcal{P}_n$. Theorem 1.1 implied a contraction rate in the *difference norm* for $\Gamma_{\mathcal{D}_n}$. In Chapters 3 and 5, cones of *uniformly bounded variation* are used to construct lower bound functions for iterates of $\mathcal{L}$; Theorem 1.1 is then applied to get contraction rates for $\mathcal{L}$.

In this way, we are always able to prove an $O(\log n/n)$ rate for the error in Ulam's scheme. However, the constants in the $O(\cdot)$ may sometimes be quite bad. Indeed, in the examples above, we have already seen that the quantitative error bounds may be many orders of magnitude worse than the actual approximation error.

In Chapter 4, *numerically assisted* versions of the analytic results are explored, and much better error bounds are obtained. Consequently, Chapter 2 should be regarded as a prototype for the rest of the thesis, rather than as a definitive analysis of approximation error for invariant densities of expanding circle maps.

# Chapter 3

# Bounded Variation and approximation error

The principal purpose in this chapter is to generalise the results of Chapter 2 to much larger classes of one–dimensional transformations. The main approximation result is the same: there exists a constant $C$ such that the approximation error in the $n$th Ulam approximation is bounded by $C \log n/n$. Only a few of the estimates that we make in the course of the proof are intrinsic to one–dimensional transformations. With suitable adaptations to manage the more complicated definition of *variation* in more dimensions, the essential arguments remain unchanged in the extension to the multi–dimensional situation. This is the subject of Chapter 5.

Initially, we prove the result for piecewise–onto expanding maps of the interval under a bounded distortion assumption. The coarse structure of the proof is similar to Chapter 2: we obtain explicit lower bounds for the iterates of initial densities under the Perron–Frobenius operator, use these to deduce quantitative bounds on the rate of mixing to equilibrium in a *bounded variation* norm, and get error bounds by showing that the Ulam approximation is a small perturbation of the Perron–Frobenius operator. Again we rely on an invariant cone structure, and the "rate of mixing" calculations are facilitated by explicit use of certain *difference norms*. One important difference from the situation in Chapter 2 is that both the Perron–Frobenius operator *and* its Ulam approximations preserve the cone structure (c.f. the necessity to define *discrete Lipschitz cones* to cope with the discontinuities introduced by the projections $\Pi_n$).

Now, the lower bound estimates for iterated initial densities rely explicitly on the geometry of iterates of the transformation $T$. Because we work directly with monotonicity branches of the map (rather than a "lift"), our estimates are very amenable to relaxing the "onto" assumption for each branch. Indeed, we illustrate this by proving the rate of mixing theorem for the well known class of $\beta$–transformations, and a class of Markov transformations with a *finite range structure* [37]. The arguments will be interspersed with examples.

To illustrate the flexibility of the method, we then show how the lower bound

estimates for the Perron–Frobenius operator also imply lower bound estimates for the Ulam approximations. Because the Ulam approximations are representable by stochastic matrices, we can deduce *uniform* bounds on the spectra of the corresponding matrix approximations. This facilitates the derivation of a *Central Limit Theorem* for "Monte–Carlo" type simulations of the Ulam approximation.

The structure of the chapter is as follows: in Section 3.1 we define the class of transformations under consideration, review the notion of *one–dimensional variation* and recall the classical *Lasota–Yorke inequality* [47]. In fact, we give a slight generalisation of the Lasota–Yorke inequality (needed for the "non–onto" applications). We also define the cones $\mathcal{C}_a$ of *uniformly bounded variation.* Initially, we work with "every branch onto" transformations: Section 3.2 contains the estimates on the non–peripheral spectral radius, and Section 3.3 applies these to get error bounds for Ulam's method. The latter section concludes with a few remarks about the history of attempts to get error bounds for Ulam's method. In Section 3.4, relaxations of the onto–condition are discussed. Then, as a second application of the spectral estimates from Section 3.2, we prove in Section 3.5 bounds on the rate of mixing for the Ulam approximation. For the sake of being self–contained, the proofs of some easy and standard lemmas are included as Appendix A. Finally, it is worth pointing out that throughout Sections 3.2, 3.3 and 3.5 the constants corresponding to the transformation $T$ are carried through the proofs. While slightly cumbersome, this allows the approximation theorems to be presented in a way which makes explicit the dependence on parameters. It is hoped that this makes the results easier to apply in practice.

A paper based on this chapter has been submitted for publication [55].

## 3.1 Inequalities for one–dimensional variation

This introductory section contains the pre-requisite details for the arguments in subsequent sections. Some of the results are slightly non–standard, but the overall purpose is to gather the necessary tools in one place.

### The transformations

Let $T : [0,1] \to [0,1]$ be a piecewise $\mathcal{C}^2$ transformation satisfying the following conditions:

**Piecewise Monotonicity:** There exists a partition $\xi = \{I_\alpha\}$ of $I = [0,1]$ into countably many subintervals such that

$$I = \bigcup_\alpha I_\alpha \quad \text{and} \quad I_\alpha \cap I_{\alpha'} = \emptyset \quad \text{if} \quad \alpha \neq \alpha' \tag{3.1}$$

and each map $T_\alpha = T|_{I_\alpha} : I_\alpha \to I$ is strictly monotone. Each $I_\alpha$ is an *interval of monotonicity* and the corresponding inverse map will be denoted by $T_\alpha^{-1}$ :

$T(I_\alpha) \to I_\alpha$. For iterates of $T$, denote $\xi^{(1)} = \xi$ and $\xi^{(n)} = T^{-1}\xi^{(n-1)} \vee \xi^{(1)}$. Letting $\xi^{(n)}$ be indexed by $\alpha^{(n)}$, each $I_{\alpha^{(n)}} \in \xi^{(n)}$ is an interval of monotonicity for $T^n$, and the inverse branches of $T^n$ are denoted by $T^{-n}_{\alpha^{(n)}} : T^n(I_{\alpha^{(n)}}) \to I_{\alpha^{(n)}}$.

**Uniform Expansion:** There exist *expansivity* and *distortion* constants $\lambda > 1$ and $s \geq 0$ (respectively) such that

$$|T'(x)| \geq \lambda \quad \forall x \in I, \tag{3.2}$$

and

$$|T''(x)|/|T'(x)|^2 \leq s \quad \forall x \in I. \tag{3.3}$$

**Uniformly Bounded Distortion:** Later on, it will be useful to express the bounded distortion property in terms of the *Renyi condition*:

$$\exists D \geq 0 \quad \text{such that} \quad |T^{-n}_{\alpha^{(n)}}{}'(x)|/|T^{-n}_{\alpha^{(n)}}{}'(y)| \leq D \quad \forall x, y \in I \tag{3.4}$$

for any $n \geq 1$ and inverse branch $T^{-n}_{\alpha^{(n)}}$ of $T^n$. By using (3.2) and (3.3) to estimate the Lipschitz constant of $\log |T^{-n}_{\alpha^{(n)}}{}'|$, it follows that $D \leq e^{s\lambda/(\lambda-1)}$.

For the first few sections of this chapter, we assume that every branch of $T$ is *onto*. That is,

$$\overline{T(I_\alpha)} = I \quad \forall I_\alpha \in \xi.$$

## Invariant measures, Ulam's method and Bounded Variation

We are concerned with approximating absolutely continuous invariant measures (acim) for expanding maps. The treatment of Ulam's method given here represents a blend of the work Li [48] with some easy observations about invariant cones.

Recall that every absolutely continuous probability measure has a positive $L^1$ density $h$, arising as a fixed point of $\mathcal{L}$, the *Perron–Frobenius operator* for $T$. Fix $n$, let $\eta = \{B_1, \ldots, B_n\}$ be the partition of $I$ into intervals of length $\frac{1}{n}$ and put

$$\mathcal{D}_n \triangleq \mathcal{D}_\eta(I) = \left\{ f \geq 0 : f = \sum_{i=1}^{n} f_i \chi_{B_i}, f_1, \ldots, f_n \in \mathbb{R}^+, \int_I f = 1 \right\}.$$

Let $\Pi_n \triangleq \Pi_\eta : \{f \in L^1(X) : f \geq 0\} \to \mathcal{D}_n$, so that

$$\mathcal{P}_n \triangleq \Pi_n \circ \mathcal{L} \tag{3.5}$$

is the *nth Ulam approximation*.

To analyse Ulam's method, we follow Li's original approach [48] of working with functions of (one–dimensional) *bounded variation*. There are many equivalent definitions of variation for one–dimensional functions. For ease of generalisation (if not transparency) we have chosen:

DEFINITION (**Bounded Variation [27]**) Let $A \subset \mathbb{R}$ be an interval, and let $C_0^1(A;\mathbb{R})$ denote the collection of compactly supported smooth functions on $A$. For each $f \in L^1(A)$ let $\cdot'$ denote the ordinary derivative and put

$$V_A(f) = \sup\left\{ \int_A f(x) w'(x)\, dm(x) : w \in C_0^1(A;\mathbb{R}), |w(x)| \leq 1 \ \forall x \in A \right\}.$$

Then $V_A(f)$ is the (one–dimensional)–*variation*[1] of $f$ on $A$, and it is usual to write $f \in BV(A)$. The set of $f \in BV(A)$ equipped with the norm

$$\|f\|_{BV} = \|f\| + V_A(f)$$

is a Banach space.

If $f \in BV(A)$, then $f$ is differentiable a.e. Letting the *generalised derivative* [69] be denoted by $df$,

$$V_A(f) = \int_A |df|. \quad \square$$

REMARK. Keller's paper [40] contains a similar definition of variation in one–dimension, although Giusti's book [27] appears to be the first exhaustive treatment of this approach. $\square$

When calculating the variation of a function by the formula above, it is important to remember that discontinuities in $f$ contribute signed point masses to the generalised derivative $df$:

EXAMPLE: Let

$$f(x) = \begin{cases} \sin x & x \in [-\pi/2, \pi/2], \\ 0 & \text{otherwise.} \end{cases}$$

Then

$$V_{(-\pi/2,\pi/2)}(f) = \int_{-\pi/2}^{\pi/2} |\cos x|\, dm(x) = 2$$

whereas

$$\begin{aligned} V_{\mathbb{R}}(f) &= \int_{-\infty}^{\infty} \Big( |\sin -\pi/2|\, d\delta_{-\pi/2} + |\cos x|\, \chi_{(-\pi/2,\pi/2)}\, dm(x) \\ &\qquad\qquad + |\sin \pi/2|\, d\delta_{\pi/2} \Big) \\ &= 4. \quad \square \end{aligned}$$

Functions in $BV$ are important for the analysis of Ulam's method because variation is robust to discretisation:

[1] An alternative definition is to let

$$\mathrm{var}_{[a,b]}(\tilde{f}) = \sup\left\{ \sum_{i=1}^{m-1} |\tilde{f}(x_{i+1}) - \tilde{f}(x_i)| : a = x_0 < x_1 < \ldots < x_m = b \right\}.$$

Then

$$V_A(f) = \inf\{\mathrm{var}_A(\tilde{f}) : f = \tilde{f} \text{ a.e.}\}.$$

**Lemma 3.1.1 (Basic properties of $\Pi_n$)** *Letting $\|\cdot\|$ be the usual $L^1$ norm:*

1. *$V_I(\Pi_n f) \leq V_I(f)$ and hence $V_I(\Pi_n f - f) \leq 2V_I(f)$;*
2. *$\|f - \Pi_n f\| \leq \frac{1}{n} V_I(f)$.*

*Proof:* See Appendix A. □

These two properties of variation under projection are fundamental to the analysis of Ulam's method. The first implies that

$$V_I(\mathcal{P}_n f) = V_I(\Pi_n \circ \mathcal{L} f) \leq V_I(\mathcal{L} f), \tag{3.6}$$

while the second says that for $BV$ functions, the Ulam approximation is a "small perturbation" of the Perron–Frobenius operator. Therefore, by establishing rates of contraction for iterates of the Perron–Frobenius operator, a simple Banach space argument will control the effect of the perturbative error introduced by the Ulam approximation.

Next, we recall a fundamental tool for the obtaining rates of contraction for iterates of $BV$ functions: the *Lasota–Yorke inequality* [47]. The version here is a slight modification of the original in [47] because we want to allow the transformation $T$ to have infinitely many branches. Consequently, we include a proof[2].

**Lemma 3.1.2 (Lasota–Yorke Inequality [47])** *Let $T : I \to I$ satisfy (3.1)—(3.3). Let $\mathcal{L}$ be the Perron–Frobenius operator for $T$, and let $f \in L^1$ satisfy $V_I(f) \leq a\|f\|$ for some finite $a > 0$. Then*

1. *if every branch of $T$ is onto,*

$$V_I(\mathcal{L} f) \leq V_I(f)/\lambda + s\|\mathcal{L} f\| \leq (a/\lambda + s)\|f\|;$$

2. *if there exists $c > 0$ such that $m(T(I_\alpha)) > c$ for each $I_\alpha \in \xi$ then*

$$V_I(\mathcal{L} f) \leq \frac{2}{\lambda} V_I(f) + \left(2s + \frac{2}{c}\right)\|f\|.$$

*Proof:* The proof is similar to Lasota and Yorke's original in [47], except that their version has a "long–intervals" condition $m(I_\alpha) > c'$ for some $c' > 0$. By taking some of the steps in a different order, our "long–images" condition $m(T(I_\alpha)) > c$ gives a more general result.

First of all, recall that

$$\mathcal{L} f(x) = \sum_\alpha \frac{f \circ T_\alpha^{-1}}{|T' \circ T_\alpha^{-1}|} \chi_{T(I_\alpha)}.$$

[2] As an additional justification for including the proof, when we describe our extensions to Góra and Boyarsky's [30] multi–dimensional version, it may be helpful to have the one–dimensional proof for reference.

Then, for each $\alpha$, recalling (3.2) and (3.3)

$$\begin{aligned}
\int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| &= \int_{T(I_\alpha)} \left| \frac{d(f\circ T_\alpha^{-1})}{|T'_\alpha\circ T_\alpha^{-1}|} + f\circ T_\alpha^{-1} d\left(\frac{1}{|T'_\alpha\circ T_\alpha^{-1}|}\right)\right| \\
&\le \int_{T(I_\alpha)} \left|\frac{(df)\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|^2}\right| + \int_{T(I_\alpha)} \left| f\circ T_\alpha^{-1}\frac{T''_\alpha\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|^3}\right| \\
&\le \frac{1}{\lambda}\int_{T(I_\alpha)} \left|\frac{(df)\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| + s\int_{T(I_\alpha)} \left|\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| \\
&= \frac{1}{\lambda}\int_{I_\alpha} |df| + s\int_{I_\alpha} |f|.
\end{aligned}$$

If every branch is onto, then

$$V_I(\mathcal{L}f) \le \sum_\alpha \int_I \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| = \sum_\alpha \int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right|,$$

and the result follows immediately from the above. Suppose now that not every branch is onto. For each $\alpha$ let $T(I_\alpha) = [l_\alpha, r_\alpha]$. Then,

$$\begin{aligned}
\int_I \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| &\le \int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| + \left|\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}(l_\alpha)\right| \\
&\qquad + \left|\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}(r_\alpha)\right| \\
&\le \int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| + \int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| \\
&\qquad + 2\operatorname*{ess\,inf}_{[l_\alpha,r_\alpha]} \left|\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| \\
&\le 2\int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| + \frac{2}{|r_\alpha - l_\alpha|}\int_{T(I_\alpha)} \left|\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| \\
&\le 2\int_{T(I_\alpha)} \left| d\frac{f\circ T_\alpha^{-1}}{|T'_\alpha\circ T_\alpha^{-1}|}\right| + \frac{2}{c}\int_{I_\alpha} |f|
\end{aligned}$$

because $|r_\alpha - l_\alpha| = m(T(I_\alpha)) > c$. The lemma follows by summing over $\alpha$. □

EXAMPLE (CONTINUED FRACTIONS) With the extension of the Lasota–Yorke inequality to transformations with infinitely many branches, the results of this chapter apply to the classical *continued fraction* algorithm. That is, $X = (0,1)$ and

$$T : x \mapsto \frac{1}{x} - \left\lfloor \frac{1}{x}\right\rfloor$$

is the classical Gauss transformation. One can easily check that

$$\left|\frac{T''}{(T')^2}\right| \le 3 \quad \text{and} \quad |(T^{-n})'| \le 4^{-\lfloor n/2\rfloor},$$

so that Lemma 3.1.2 and all our subsequent analysis applies to $T$. Of course, it is a standard fact that $T$ preserves the measure with density

$$\phi(x) = \frac{\log 2}{1+x},$$

but it is reassuring to know that this classical example fits into the framework developed below. □

Throughout, we will work on *cones* of functions of uniformly bounded variation (cf. [50]). We define these now, and briefly explain why such cones are the natural spaces in which to analyse convergence to invariant densities.

DEFINITION: Fix $a > 0$ and consider the following collection of $L^1$ functions:

$$\mathcal{C}_a = \{0 \le f \in L^1 : V_I(f) \le a\|f\|\}.$$

It is easy to check that $\mathcal{C}_a$ is an *additive cone*. We call $\mathcal{C}_a$ a cone of *uniformly bounded variation*. □

Remembering that we are dealing with transformations where every branch is onto, it follows immediately from Lemma 3.1.2 that

$$\mathcal{L}\mathcal{C}_a \subset \mathcal{C}_{(a/\lambda+s)}.$$

In particular, if $a \ge s\lambda/(\lambda-1)$, then $\mathcal{L}(\mathcal{C}_a \cap \{f : \|f\| = 1\}) \subset \mathcal{C}_a \cap \{f : \|f\| = 1\}$. Because each $\mathcal{C}_a \cap \{f : \|f\| = \text{constant}\}$ is compact and convex, this implies that $\mathcal{L}$ has a fixed point $h \in \mathcal{C}_{(s\lambda/(\lambda-1))}$. See [47] for details.

Moreover, it follows from (3.5), (3.6) and Lemma 1.3.1 that

$$\mathcal{L}\mathcal{C}_a \subset \mathcal{C}_a \quad \Rightarrow \quad \mathcal{P}_n\mathcal{C}_a \subset \mathcal{C}_a.$$

By the same argument as for $\mathcal{L}$, each $\mathcal{P}_n$ has a fixed point $h_n \in \mathcal{C}_{s\lambda/(\lambda-1)}$. The density $h_n$ is the Ulam approximation, and every limit point of the sequence $\{h_n\}$ is an invariant density for $T$ [48].

## 3.2 Lower bound functions and quantitative mixing

The purpose of this section is to obtain an explicit bound on the non–peripheral part of the spectrum of the Perron–Frobenius operator restricted to functions of uniformly bounded variation. This is Theorem 3.1. We use the method of *lower bounds for invariant cones*.

Fundamentally, we are interested in rates of mixing (contraction) for functions in

$$BV_0 = \left\{ g \in BV(I) : \int_I g\, dm = 0 \right\}.$$

We get contraction rates by using the *difference cones* for certain $\mathcal{C}_a$s. The exposition as is follows: First of all, lower bounds for iterates of $\mathcal{L}$ applied

to initial densities from $\mathcal{C}_a$ are proved. Then, Theorem 1.1 implies a rate of contraction in a particular *difference norm*. The application to Ulam's method requires estimates in the *usual norm* on $BV_0$. Consequently, some care is taken to exhibit the relationship between the usual norm on $BV_0$ and the difference norms corresponding to $\mathcal{C}_a$. The extra complexity from working with difference norms is rewarded by a simpler statement, proof and application of Theorem 3.1 than would otherwise be possible.

### 3.2.1 Lower bound functions

The first step is to get a lower bound for iterates of $\mathcal{L}$. Some aspects of the construction are similar to [50], and discussion of this may be found in subsection 3.4.4 below.

**Lemma 3.2.1** *Let $f \in \mathcal{C}_a$, and let $\epsilon > 0$. Let $\eta = \{I_\beta\}$ be a partition of $I$ into subintervals of length less than or equal to $\epsilon$, and let $\tilde{f}$ be a function which is constant on each subinterval $I_\beta$ with*

$$\operatorname*{ess\,inf}_{I_\beta} f \leq \tilde{f}_\beta \leq \operatorname*{ess\,sup}_{I_\beta} f.$$

*Then*

$$\|f - \tilde{f}\| \leq \epsilon a \|f\|.$$

*Proof:* See Appendix A. □

**Lemma 3.2.2** *Let $m(\cdot)$ denote the Lebesgue measure on $I$, and let $I_{\alpha^{(n)}}$ be any subinterval from the partition $\xi^{(n)}$. If $D$ is the constant in (3.4) then*

$$|T^{n\prime} \circ T^{-n}_{\alpha^{(n)}}(x)|^{-1} \geq m(I_{\alpha^{(n)}})/D \quad \forall x \in I_{\alpha^{(n)}}.$$

*Proof:* See Appendix A. □

**Proposition 3.1 (Lower bound for $\mathcal{L}^n$)** *Let $f \in \mathcal{C}_a$ and let $n \geq 1$. Then*

$$\mathcal{L}^n f \geq \|f\|(1 - a\lambda^{-n})/D.$$

*Proof:* Let $\eta = \xi^{(n)}$, the partition of $I$ into monotonicity intervals of $T^n$. Because $|T^{n\prime}(x)| \geq \lambda^n$ (c.f. equation (3.2)), each subinterval in $\eta$ has length less than or equal to $\lambda^{-n}$. Put

$$\tilde{f} = \sum_{\alpha^{(n)}} \left( \operatorname*{ess\,inf}_{I_{\alpha^{(n)}}} f \right) \chi_{I_{\alpha^{(n)}}}.$$

Then Lemma 3.2.1 implies that $\|f - \tilde{f}\| \leq \lambda^{-n} a \|f\|$. But $\|f\| = \|f - \tilde{f}\| + \|\tilde{f}\|$ (since $f \geq \tilde{f} \geq 0$ a.e.). Therefore $\|\tilde{f}\| \geq \|f\|(1 - \lambda^{-n} a)$.

Now, from (1.3)

$$\begin{aligned}\mathcal{L}^n f(x) = \sum_{\alpha^{(n)}} \frac{f \circ T^{-n}_{\alpha^{(n)}}(x)}{|T^{n\prime} \circ T^{-n}_{\alpha^{(n)}}(x)|} &\geq \sum_{\alpha^{(n)}} \left( \left( \operatorname*{ess\,inf}_{I_{\alpha^{(n)}}} f \right) \frac{1}{|T^{n\prime} \circ T^{-n}_{\alpha^{(n)}}(x)|} \right) \\ &\geq \sum_{\alpha^{(n)}} \tilde{f}_{\alpha^{(n)}} m(I_{\alpha^{(n)}})/D \\ &= \|\tilde{f}\|/D.\end{aligned}$$

Lemma 3.2.2 was used to get the last inequality. The result follows. □

REMARK. The construction in Proposition 3.1 substantially improves similar estimates given by Liverani [50, Lemma 3.5]. □

### 3.2.2 Uniform $BV$ cones

With Proposition 3.1 established, we need a suitable cone of functions on which $\mathcal{L}$ is a strict contraction.

For each $a \geq 0$, put

$$\Gamma_a \triangleq \Gamma_{\mathcal{C}_a} = \Big\{ f \in L^1 : \int_I f = 0 \quad \text{and} \quad \exists f^{(1)}, f^{(2)} \in \mathcal{C}_a \text{ s.t.} \\ f = f^{(1)} - f^{(2)} \text{ and } \|f^{(1)}\| = \|f^{(2)}\| \Big\}.$$

Each $\Gamma_a$ is the *difference cone* for $\mathcal{C}_a$, and for each $\Gamma_a$ let:

$$\|f\|_a = \inf \Big\{ \|f^{(1)}\| : f = f^{(1)} - f^{(2)}, f^{(1)}, f^{(2)} \in \mathcal{C}_a, \|f^{(1)}\| = \|f^{(2)}\| \Big\}.$$

Then $\|\cdot\|_a$ is the *difference norm* for $\mathcal{C}_a$, and $(\Gamma_a, \|\cdot\|_a)$ is a normed linear space (c.f. Lemma 1.4.1).

EXAMPLE: Let $f(x) = \sin 2\pi x$. Then $\|f\| = 2/\pi$ and $V_{[0,1]}(f) = 4$. Hence $f \in BV_0$ and $\|f\|_{BV} = 4 + 2/\pi$. On the other hand, both $f^+, f^- \in \mathcal{C}_a$ whenever $a \geq 2\pi$. One can check that

$$\|f\|_a = \begin{cases} 2/a & \text{if } a < 2\pi, \\ 1/\pi & \text{if } a \geq 2\pi. \end{cases} \quad \square$$

### 3.2.3 Strict contraction on $\Gamma_{a_*}$

Put

$$a_* = \frac{\frac{1}{2} + \frac{s\lambda}{\lambda - 1}}{1 - \frac{1}{2D}} \text{ and } n_* = \left\lceil \frac{\log 2a_*}{\log \lambda} \right\rceil.$$

**Theorem 3.1 (Mixing Theorem for $\mathcal{L}$)** *Let $a_*, n_*$ be as just defined, and let $f \in \Gamma_{a_*}$. Then*

$$\|\mathcal{L}^{n_*} f\|_{a_*} \leq (1 - 1/2D)\|f\|_{a_*}.$$

*Proof:* By Theorem 1.1, it suffices to establish that,

$$\mathcal{L}^{n_*} g \geq \|g\|/2D \quad \text{and} \quad \mathcal{L}^{n_*} g - \|g\|/2D \in \mathcal{C}_{a_*} \quad \text{whenever } g \in \mathcal{C}_{a_*}. \tag{3.7}$$

Let $g \in \mathcal{C}_{a_*}$. By Proposition 3.1,

$$\mathcal{L}^{n_*} g \geq \|g\| \left(1 - a_* \lambda^{-n_*}\right)/D \geq \|g\|/2D. \tag{3.8}$$

Therefore,

$$\|\mathcal{L}^{n_*} g - \|g\|/2D\| = \|\mathcal{L}^{n_*} g\| - \|g\|/2D = (1 - 1/2D)\|g\|. \tag{3.9}$$

Next, repeated application of Lemma 3.1.2 yields

$$\begin{aligned} V_I\left(\mathcal{L}^{n_*} g - \|g\|/2D\right) &= V_I(\mathcal{L}^{n_*} g) \\ &\leq \lambda^{-n_*} V_I(g) + \frac{s\lambda}{\lambda - 1}\|g\| \\ &\leq \left(\frac{1}{2} + \frac{s\lambda}{\lambda - 1}\right)\|g\| \\ &= (1 - 1/2D) a_* \|\mathcal{L}^{n_*} g\| \end{aligned} \tag{3.10}$$

because $g \in \mathcal{C}_{a_*}$ and $\lambda^{-n_*} a_* \leq 1/2$ (by the choice of $n_*$). Combining equations (3.9) and (3.10), we have that

$$\mathcal{L}^{n_*} g - \|g\|/2D \in \mathcal{C}_{a_*}.$$

This, together with (3.8), establishes (3.7) and the theorem follows by the method of lower bounds for invariant cones. □

Theorem 3.1 will be called a *Mixing Theorem*.

REMARK. The Mixing Theorem gives an explicit rate at which $BV$ densities converge to an invariant density under iterates of $\mathcal{L}$. Compare with Liverani's Theorem 3.6 [50]. Liverani comments that his estimates are probably not optimal, and by sacrificing some generality (requiring each branch to be onto), Theorem 3.1 represents a considerable improvement; c.f. subsection 3.4.4 below. □

EXAMPLE: For a fixed integer $k$, consider the transformation $x \mapsto kx \pmod 1$. In the notation above, $\lambda = k$, $s = 0$, $D = 1$. Then $a_* = 1$ and $n_* = 1$. By applying Theorem 3.1,

$$\|\mathcal{L}^n f\|_{a_*} \leq 2^{-n}\|f\|_{a_*}$$

for each $n \geq 0$, $f \in BV_0$. In this case, a rather better estimate can be obtained: by Proposition 3.1, $\mathcal{L}g \geq \|g\|\frac{k-1}{k}$ whenever $g \in \mathcal{C}_{a_*}$. Reproducing the argument in the proof of Theorem 3.1:

$$\|\mathcal{L}^n f\|_{a_*} \leq k^{-n}\|f\|_{a_*}$$

for each $f \in BV_0$. $\square$

EXAMPLE: For each $\Omega \in [0,1)$ consider the circle map

$$T_\Omega : x \mapsto 2x + \frac{\sin 2\pi x}{4\pi} + \Omega \pmod 1.$$

Each $T_\Omega$ is a nonlinear perturbation of $z \mapsto e^{2\pi i\Omega} z^2$, which leaves Lebesgue measure invariant. Obviously, each branch of $T_\Omega$ is onto, and $1.5 \le T'_\Omega \le 2.5$ for each $\Omega$. Moreover, $|T''_\Omega(x)|/|T'_\Omega(x)|^2$ is maximised when $\cos 2\pi x = 2 - \sqrt{6}$. Hence,

$$\lambda = 1.5, \quad s \le 0.891, \quad D \le e^{s\lambda/(\lambda-1)} \le 14.5.$$

With these constants it follows that

$$a_* = 3.29, \quad n_* = 5.$$

Therefore, by Theorem 3.1, for each $k \ge 0$ and $f \in \Gamma_{a_*}$,

$$\|\mathcal{L}_\Omega^{5k} f\|_{a_*} = \|\mathcal{L}_\Omega^{kn_*} f\|_{a_*} \le \left(\frac{28}{29}\right)^k \|f\|_{a_*},$$

where $\mathcal{L}_\Omega$ is the Perron–Frobenius operator corresponding to $T_\Omega$. $\square$

### 3.2.4 Equivalence of $BV$ norms

To complete the preparation for the analysis of Ulam's method, we need to know about the relationship between the spaces $(\Gamma_a, \|\cdot\|_a)$ and $(BV_0, \|\cdot\|_{BV})$.

**Lemma 3.2.3** *For each $a \ge 0$:*

1. *the infimum in the definition of $\|\cdot\|_a$ is attained;*
2. *if $f \in \Gamma_a$ and $a < b$ then $f \in \Gamma_b$ and $\|f\|_b \le \|f\|_a$;*
3. *if $\mathcal{T}$ is a Markov operator, and $n > 0$, then*

$$\mathcal{T}^n \mathcal{C}_b \subset \mathcal{C}_a \quad \Rightarrow \quad \|\mathcal{T}^n f\|_a \le \|f\|_b \quad \forall f \in \Gamma_b;$$

4. *if $f \in \Gamma_a$ then*

$$\|f\| \le 2\|f\|_a \quad \text{and} \quad V_I(f) \le 2a\|f\|_a;$$

5. *if $f \in \Gamma_a$ then*

$$\|f\|_a \le \max\left\{\frac{\|f\|}{2}, \frac{V(f)}{a}\right\};$$

6. *if $f \in \Gamma_a$ then*

$$\min\{2, a\}\|f\|_a \le \|f\|_{BV} \le 2(1+a)\|f\|_a,$$

*so that $f \in BV_0$ if and only if $f \in \Gamma_a$ and $\|f\|_a < \infty$ for every $a > 0$.*

*Proof:* See Appendix A. □

REMARK. The norms $\|\cdot\|_a$ can be shown to be equivalent to certain of the *balanced norms* $\|\cdot\|_\gamma$ used in [4, §5] (and subsequent papers) to estimate rates of decay of correlations for random perturbations of $BV$ transformations. □

Although Lemma 3.2.3 implies that $(BV_0, \|\cdot\|_{BV})$ and $(\Gamma_a, \|\cdot\|_a)$ consist of the same functions with the same topology, we have obtained contraction rates in $\|\cdot\|_a$ with relative ease. Attempting a construction directly in $\|\cdot\|_{BV}$ is both difficult, and not guaranteed to succeed. See [31]. On the other hand, the combination of Theorem 3.1 and Lemma 3.2.3 is potentially very powerful.

**Corollary 3.2.4** *Let $f \in BV_0$ and suppose that $a_*, n_*$ are as in Theorem 3.1. Then for each $k \geq 0$*

$$\|\mathcal{L}^k f\| \leq 2\|\mathcal{L}^k f\|_{a_*} \leq 2(1-1/2D)^{\lfloor \frac{k}{n_*} \rfloor} \|f\|_{a_*}.$$

*Proof:* Follows directly from Lemma 3.2.3 and Theorem 3.1. □

EXAMPLE: Recall from above the circle maps $\{T_\Omega\}_{\Omega \in [0,1)}$. By Lemma 3.2.3 and Theorem 3.1, for each $k \geq 0$ and $f \in BV_0$

$$\begin{aligned} \|\mathcal{L}_\Omega^{5k} f\|_{BV} = \|\mathcal{L}_\Omega^{kn_*} f\|_{BV} &\leq 8.58\, \|\mathcal{L}_\Omega^{kn_*} f\|_{a_*} \\ &\leq 8.58 \left(\frac{28}{29}\right)^k \|f\|_{a_*} \leq 4.29 \left(\frac{28}{29}\right)^k \|f\|_{BV} \end{aligned}$$

where $\mathcal{L}_\Omega$ is the Perron–Frobenius operator for $T_\Omega$. □

### 3.2.5 Aside: Decay of correlations of certain functions

One standard application of contraction rates for Perron–Frobenius operators is to the *decay of correlations* of test functions. While Theorem 3.1 gives good contraction rates for $\mathcal{L}$, the formalism of $\|\cdot\|_a$–norms allows a very clean application to the decay of correlations. The following result should be compared with [50, Theorem 3.6].

**Decay of Correlations** *Let $\phi \in BV(I)$, $\psi \in L^1(I) \cap L^\infty(I)$ and let $n_*$ be as in Theorem 3.1. Then, letting $\mu$ denote the acim for $T$, for any $k > 0$,*

$$\left| \int_I \phi \cdot \psi \circ T^{kn_*}\, d\mu - \int_I \phi\, d\mu \int_I \psi\, d\mu \right| \leq 4\|\psi\|_\infty V_I(\phi)\,(1-1/2D)^k.$$

*Proof:* Letting $\bar{\phi} = \int_I \phi\, d\mu/\mu(I)$, and $f$ be the density of the invariant measure, a standard calculation yields

$$\int_I \phi \cdot \psi \circ T^{kn_*}\, d\mu - \int_I \phi\, d\mu \int_I \psi\, d\mu = \int_I \psi \cdot \mathcal{L}^{kn_*} \left((\phi - \bar{\phi}) f\right)\, dm.$$

The result will follow if we can show that

$$\left\| \mathcal{L}^{kn_*} \left( (\phi - \bar{\phi}) f \right) \right\| \leq 4\, V_I(\phi)\, (1 - 1/2D)^k. \tag{3.11}$$

Now, by definition of $\bar{\phi}$, $\int_I (\phi - \bar{\phi})\, f\, dm = 0$, and

$$\begin{aligned} V\left((\phi - \bar{\phi}) f\right) &\leq \|f\|_\infty V(\phi - \bar{\phi}) + V(f) \|\phi - \bar{\phi}\|_\infty \\ &\leq (1 + V(f))\, V(\phi) + V(f)\, V(\phi) \end{aligned}$$

by standard properties of variation. Hence, $(\phi - \bar{\phi}) f \in BV_0$, and Lemma 3.2.3 (5) implies that for any $a > 0$,

$$\|(\phi - \bar{\phi}) f\|_a \leq \max \left\{ \frac{\|(\phi - \bar{\phi}) f\|}{2}, \frac{V((\phi - \bar{\phi}) f)}{a} \right\}.$$

Since $\|(\phi - \bar{\phi}) f\| \leq \|(\phi - \bar{\phi})\|_\infty \|f\| \leq V(\phi) \times 1$, this last estimate reduces to

$$\|(\phi - \bar{\phi}) f\|_a \leq V(\phi) \max \left\{ \frac{1}{2}, \frac{1 + 2V(f)}{a} \right\}.$$

However, because $f \in \mathcal{C}_{s\lambda/(\lambda-1)}$,

$$1 + 2V(f) \leq 2 \left( \frac{1}{2} + \frac{s\lambda}{\lambda - 1} \right) < 2a_*,$$

where $a_*$ is as in Theorem 3.1. Therefore, by Lemma 3.2.3 (4), Theorem 3.1 and the above estimates,

$$\begin{aligned} \left\| \mathcal{L}^{kn_*} \left( (\phi - \bar{\phi}) f \right) \right\| &\leq 2 \left\| \mathcal{L}^{kn_*} \left( (\phi - \bar{\phi}) f \right) \right\|_{a_*} \\ &\leq 2(1 - 1/2D)^k \left\| (\phi - \bar{\phi}) f \right\|_{a_*} \\ &\leq 2(1 - 1/2D)^k V(\phi) \max \left\{ \frac{1}{2}, \frac{2a_*}{a_*} \right\}. \end{aligned}$$

This establishes (3.11), and the decay of correlations follows. □

## 3.3 Error bounds for density approximation

We now deduce error bounds for Ulam's method. The strategy is relatively straight–forward: first of all, a standard argument shows that strict contraction rates for iterates of $\mathcal{L}$ imply that the operator $(Id - \mathcal{L})|_{\Gamma_{a_*}}$ has a bounded inverse. We then deduce a bound on the $L^1$ distance from the invariant density for $T$ to another $BV$ density in terms of a certain $\|\cdot\|_a$–norm (Corollary 3.3.2). The final part of the proof consists in choosing a value of $a$ such that the $\|\cdot\|_a$–norm is sufficiently small.

### 3.3.1 Theoretical error bounds

The operator $(Id - \mathcal{L})|_{\Gamma_{a_*}}$ is invertible:

**Lemma 3.3.1 (Bounded inverse for $(Id - \mathcal{L})$)** *Let $n_*, a_*$ be defined as preceding Theorem 3.1. Then the operator*

$$(Id - \mathcal{L}) : (\Gamma_{a_*}, \|\cdot\|_{a_*}) \to (\Gamma_{a_*}, \|\cdot\|_{a_*})$$

*is invertible, and*

$$\|(Id - \mathcal{L})^{-1} f\|_{a_*} \leq 2Dn_* \|f\|_{a_*}$$

*for every $f \in \Gamma_{a_*}$.*

*Proof:* Using Corollary 3.2.4, a standard Banach space argument shows that the sum

$$\sum_{k=0}^{\infty} \mathcal{L}^k$$

is well defined, and satisfies $(Id - \mathcal{L}) \sum_{k=0}^{\infty} \mathcal{L}^k = Id|_{BV_0}$. The norm estimate follows immediately. □

**Proposition 3.2** *Let $f \in BV_0$. Then whenever $a \geq a_*$,*

$$\|(Id - \mathcal{L})^{-1} f\| \leq \left\lceil \frac{\log 2a}{\log \lambda} \right\rceil \|f\| + 2(2D - 1)n_* \|f\|_a.$$

*Proof:* Assume that $a \geq a_*$ and put $n_a = \lceil \frac{\log 2a}{\log \lambda} \rceil$. Let $f \in BV_0$, so that $f \in \Gamma_{a_*}$ by Lemma 3.2.3. Next, observe that

$$(Id - \mathcal{L})^{-1} f = \sum_{k=0}^{\infty} \mathcal{L}^k f = \sum_{i=0}^{n_a - 1} \mathcal{L}^i f + (Id - \mathcal{L})^{-1} \mathcal{L}^{n_a} f.$$

Then, by Lemma 3.2.3 (4) and Lemma 3.3.1,

$$\begin{aligned} \left\|(Id - \mathcal{L})^{-1} f\right\| &\leq n_a \|f\| + 2 \left\|(Id - \mathcal{L})^{-1} \mathcal{L}^{n_a} f\right\|_{a_*} \\ &\leq n_a \|f\| + 2 \times 2Dn_* \|\mathcal{L}^{n_a} f\|_{a_*}. \end{aligned} \tag{3.12}$$

Now, let $f^{(1)}, f^{(2)} \in \mathcal{C}_a$ satisfy

$$f = f^{(1)} - f^{(2)} \quad \text{and} \quad \|f^{(1)}\| = \|f^{(2)}\| = \|f\|_a.$$

By the same argument as in the proof of Theorem 3.1,

$$g^{(1)} = \mathcal{L}^{n_a} f^{(1)} - \|f^{(1)}\|/2D \in \mathcal{C}_{a_*} \quad \text{and} \quad g^{(2)} = \mathcal{L}^{n_a} f^{(2)} - \|f^{(2)}\|/2D \in \mathcal{C}_{a_*}.$$

Since $\mathcal{L}^{n_a} f = \mathcal{L}^{n_a} f^{(1)} - \mathcal{L}^{n_a} f^{(2)} = g^{(1)} - g^{(2)}$,

$$\|\mathcal{L}^{n_a} f\|_{a_*} \leq \|g^{(1)}\| = \|g^{(2)}\| = (1 - 1/2D)\|f\|_a,$$

and the proposition follows from (3.12). □

**Corollary 3.3.2** *Let $h$ be the normalised fixed point of $\mathcal{L}$, and suppose that $g \in BV$ satisfies $\|g\| = \|h\|$. Let $a \geq a_*$, and let $n_*$ be as usual. Then*

$$\|g - h\| \leq \left\lceil \frac{\log 2a}{\log \lambda} \right\rceil \|g - \mathcal{L}g\| + 2(2D-1)n_* \|g - \mathcal{L}g\|_a.$$

*Proof:* Since

$$(Id - \mathcal{L})(g - h) = g - \mathcal{L}g,$$

Proposition 3.2 implies the result. □

To apply the corollary to get an error bound for Ulam's approximation, let $g = h_n = \mathcal{P}_n h_n$. The remaining part of this section consists in finding an $a$ such that $h_n - \mathcal{L}h_n \in \Gamma_a$ and $\|h_n\|_a$ is sufficiently small.

### 3.3.2 Choice of cone

Let $h_n$ be the normalised fixed point of the $n$th Ulam approximation.

**Lemma 3.3.3** *For each $n \geq 0$,*

$$\|h_n - \mathcal{L}h_n\| \leq \frac{1}{n} \frac{s\lambda}{\lambda - 1} \quad \textit{and} \quad \|h_n - \mathcal{L}h_n\|_{4n} \leq \frac{1}{2n} \frac{s\lambda}{\lambda - 1}.$$

*where $\|\cdot\|_{4n}$ denotes the difference norm for $\Gamma_a$ when $a = 4n$.*

*Proof:* Recall that $h_n, \mathcal{L}h_n \in \mathcal{C}_{s\lambda/(\lambda-1)}$. Therefore, by Lemma 3.1.1,

$$\|h_n - \mathcal{L}h_n\| = \|\Pi_n \circ \mathcal{L}h_n - \mathcal{L}h_n\| = \|(\Pi_n - Id)\mathcal{L}h_n\| \leq \frac{s\lambda}{\lambda - 1} \frac{1}{n},$$

since $\|\mathcal{L}h_n\| = \|h_n\| = 1$. Also,

$$V(h_n - \mathcal{L}h_n) \leq V(h_n) + V(\mathcal{L}h_n) \leq 2 \frac{s\lambda}{\lambda - 1}.$$

By Lemma 3.2.3 (5),

$$\|h_n - \mathcal{L}h_n\|_{4n} \leq \max \left\{ \frac{\|h_n - \mathcal{L}h_n\|}{2}, \frac{V(h_n - \mathcal{L}h_n)}{4n} \right\} \leq \frac{1}{2n} \frac{s\lambda}{\lambda - 1}.$$

□

Therefore, by putting $a = 4n$ in Corollary 3.3.2, we have

**Theorem 3.2** *Let $T$ satisfy (3.1)—(3.3), let $\mathcal{L}$ be the Perron–Frobenius operator for $T$ and let $h \in \mathcal{C}_{s\lambda/(\lambda-1)}$ be such that $\mathcal{L}h = h$, $\|h\| = 1$. For each $n \geq a_*/4$ (where $a_*, n_*$ as in Theorem 3.1) let $h_n \in \mathcal{C}_{s\lambda/(\lambda-1)}, \|h_n\| = 1$ be the Ulam approximation: $h_n = \mathcal{P}_n h_n$. Then*

$$\|h - h_n\| \leq \left( \left\lceil \frac{\log 8n}{\log \lambda} \right\rceil + (2D-1)n_* \right) \frac{1}{n} \frac{s\lambda}{\lambda - 1} = O\left( \frac{\log n}{n} \right).$$

EXAMPLE: Recall the family $\{T_\Omega\}$ of transformations from the end of Section 3.2. If $h_{n,\Omega}$ is the $n$th Ulam approximation to $h_\Omega$ (the unique invariant density for $T_\Omega$) then Theorem 3.2 implies

$$\|h_\Omega - h_{n,\Omega}\| \leq \frac{2.67}{n}\left(\left\lceil \frac{\log 8n}{\log \lambda}\right\rceil + 28 \times 5\right).$$

Putting $n = 10000$,

$$\|h_\Omega - h_{10000,\Omega}\| \leq 0.0449.$$

For 100 values of $\Omega$ between 0 and 0.99 I have performed the calculation. For each $\Omega$, most of the work consists in calculating the entries of the transition matrix $P(10000, \Omega)$; finding the invariant probability vector by repeated multiplication of some random initial vector is extremely fast because each transition matrix is exponentially mixing (cf. Remark 3.5.4 below) and has a sparse structure (since $P_{ij} \neq 0 \Rightarrow T_\Omega B_i \cap B_j \neq \emptyset$). In Figure 3.1, the densities are represented by a grey–scale plot; darker grey denotes higher density.

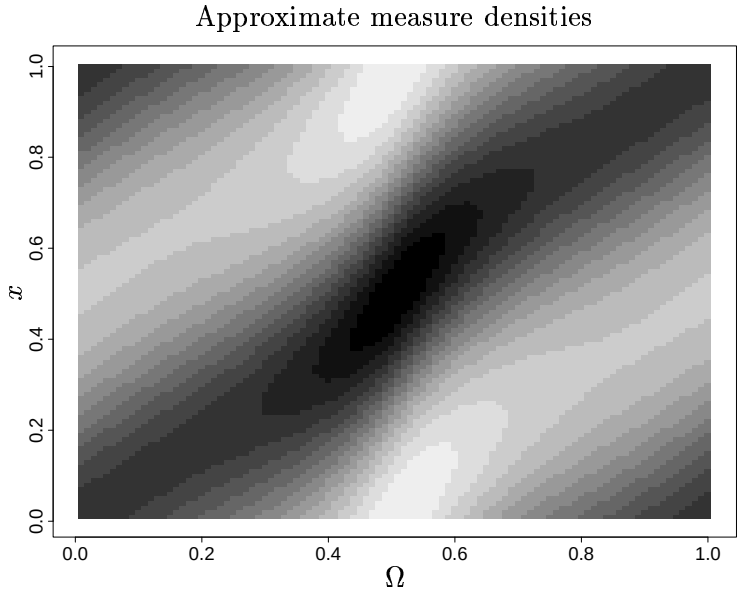



Figure 3.1: Grey scale image of approximate invariant densities for the maps $T_\Omega$.

Finally, we reiterate that the densities depicted in Figure 3.1 are *guaranteed*[3] by Theorem 3.2 to be within $4\frac{1}{2}\%$ of the exact invariant density for each $T_\Omega$. □.

REMARK. The choice of $a_* = \frac{1/2+s\lambda/(\lambda-1)}{1-1/2D}$ is essentially arbitrary, and has been made for the sake of getting a simple formula for the contraction rate. Below, we discuss a strategy for choosing "more optimal" values of $a_*$. Further discussion may be found in Appendix B. □

[3]Assuming that the numerics are accurate.

### 3.3.3 Historical remarks

In 1976, Li [48] was able to answer an original conjecture of Ulam [64] by proving that as $n \to \infty$, the sequence $\{h_n\}$ converges to an invariant density for any expanding $BV$ transformation of the interval. The issue of how quickly the convergence occurs has taken much longer to resolve.

In Hunt [31], a careful choice of $BV$–norm was used to get bounds on the approximation error for an invariant density calculation. The technique is less general than ours, but nevertheless returns *explicit quantitative information* in the limited cases where it applies. Because the interest in [31] was in getting a single error bound, Hunt did not consider the rate of convergence of his scheme.

Now, the main result in this chapter has been the construction of an explicit constant $C$ such that $L^1$ distance from the the $n$th Ulam approximation to the correct invariant density is bounded by $C\frac{\log n}{n}$. Keller [40] has proved an $O(\log n/n)$ rate in a more general setting, but relies upon spectral information from the Ionescu–Tulcea and Marinescu Ergodic Theorem. The novelty in our approach has been to obtain explicit quantitative information about the constants in the $O(\cdot)$ notation.

One natural question (separate to getting quantitative control) is whether the $O(\log n/n)$ rate is optimal. Chiu et. al. [15] estimated the rate of convergence of the scheme at $O(1/n)$, using an argument which can be translated into our cone–based setting. Unfortunately, their proof is flawed [33], and our cones notation allows a relatively straight–forward explanation of where their error lies:

While the details of their argument involve careful analysis of *resolvent operators* to deal with some analytical details, the essence of their Theorem 2 (on which the error estimates depend) is to write

$$\|h - h_n\| \leq \|(Id - \mathcal{L})^{-1}\| \|\mathcal{L}h_n - h_n\| \tag{3.13}$$

where $(Id - \mathcal{L})$ is regarded as an operator

$$(Id - \mathcal{L}) : (BV_0, \|\cdot\|_{BV}) \to (BV_0, \|\cdot\|).$$

In Lemma 10 of [15] the authors prove that for each $C > 0$ there exists a constant $M(C)$ such that if $g \in BV_0$ is such that $\|g\| = 1$ and $V_I(g) \leq C$ then

$$\|(Id - \mathcal{L})^{-1} g\| \leq M(C). \tag{3.14}$$

(In our notation, it is easy to see that both $g^+, g^- \in \mathcal{C}_{2C}$ so that Proposition 3.2 implies that $M(C) = O(\log C)$.) Now, we have seen that there exists a constant $c$ such that $V_I(\mathcal{L}h_n - h_n) \leq c$ for all $n$, while $\|\mathcal{L}h_n - h_n\| = O(1/n)$. The authors put

$$C = c$$

and (erroneously) use (3.13) to deduce

$$\|h - h_n\| \leq M(c)\|\mathcal{L}h_n - h_n\| = M(c)O(1/n).$$

This expression hides the error in [15]: the "$C$" for which it is valid to apply (3.14) to $g_n = \mathcal{L}h_n - h_n$ may depend on $n$. Since $V(g_n) \leq c = O(1)$ and $\|g_n\| = O(1/n)$, the best that can be deduced about $C_n \triangleq V(g_n)/\|g_n\|$ is that it scales like $O(n)$. By (3.14),

$$\|h - h_n\| \leq M(C_n)\|\mathcal{L}h_n - h_n\| = M(C_n)O(1/n).$$

From our results, $M(C_n)$ scales like $O(\log n)$, thereby recovering the $O(\log n/n)$ rate. Unfortunately, a proof for the conjectured $O(1/n)$ rate remains elusive.

## 3.4 Generalisations to non–onto transformations

We now consider generalisations of the arguments in the previous two sections. For Ulam's approximation, the error bounds and rates of mixing that are proved above rely on spectral estimates for the Perron–Frobenius operator $\mathcal{L}$. For more general transformations, the arguments in Section 3.3 can be repeated to give similar approximation rates for Ulam's method once a Mixing Theorem is available. Therefore, most of the discussion in this section is concentrated on a generalised Mixing Theorem: Theorem 3.3.

The onto Mixing Theorem (Theorem 3.1) had two key ingredients: the Lasota–Yorke inequality (Lemma 3.1.2) and the construction of explicit *lower bound functions* for iterates of $\mathcal{L}$ (Proposition 3.1). For non–onto transformations with $\lambda > 2$, Lemma 3.1.2 provides a sufficient Lasota–Yorke inequality. Obtaining lower bound functions is much harder in the non–onto case, and occupies most of subsection 3.4.1. In subsection 3.4.2, the Mixing Theorem is applied to the well known class of $\beta$–transformations, and in subsection 3.4.3 to a class of transformations with *Finite Range Structure* [37]. In each of these situations, error bounds for Ulam's method can be deduced as in Section 3.3.

### 3.4.1 Non-onto transformations

We no longer assume that every branch of the transformation $T$ maps onto the whole interval, but we do suppose that there exists a constant $c > 0$ such that

$$m\left(T(I_\alpha)\right) \geq c \quad \forall I_\alpha \in \xi.$$

Again, let $s \geq |T''|/|T'|^2$, but strengthen the expansivity assumption to $|T'| \geq \lambda > 2$ (see [6] for a discussion of the necessity of this stronger mixing condition for statistical stability of non–onto maps). For ease of notation, put

$$\sigma = 2/\lambda \quad \text{and} \quad A = 2(s + 1/c).$$

Then

$$V_I(\mathcal{L}f) \leq \sigma V_I(f) + A\|f\|$$

whenever $f \in BV$.

**Theorem 3.3 (Mixing Theorem for certain non–onto maps)** *Suppose that there exists* $n_* > \frac{\log A/(1-\sigma)}{\log \lambda}$, *and an interval* $U_* = U(n_*) \subset I$ *such that*

$$U_* \subset T^{n_*}(I_{\alpha^{(n_*)}})$$

*for every monotonicity interval* $I_{\alpha^{(n_*)}}$ *of* $T^{n_*}$. *Then there exists an* $a_* \in (A/(1-\sigma), \lambda^{n_*})$ *such that whenever* $a \in [a_*, \lambda^{n_*})$,

$$\|\mathcal{L}^{n_*} f\|_a \leq \left(1 - \frac{m(U_*)}{D}\left(1 - a\lambda^{-n_*}\right)\right)\|f\|_a$$

*for every* $f \in BV_0$. *In fact* $a_*$ *is the positive root of the equation*

$$a_*\left(1 - \frac{m(U_*)}{D}\left(1 - \frac{a_*}{\lambda^{n_*}}\right)\right) = \sigma^{n_*} a_* + \frac{1-\sigma^{n_*}}{1-\sigma}A + \frac{2}{D}\left(1 - \frac{a_*}{\lambda^{n_*}}\right).$$

*Proof:* The proof is in two parts. We first of all establish the existence of a lower bound function $\psi_g$ for $\mathcal{L}^{n_*} g$ whenever $g \in \mathcal{C}_a$. The second part consists in showing how the choice of $a_*$ guarantees that $\mathcal{L}^{n_*} g - \psi_g \in \mathcal{C}_a$ whenever $a \geq a_*$.

Let $\eta = \xi^{(n_*)}$ be the partition of $I$ into monotonicity branches of $T^{n_*}$. Then, analogously to Lemma 3.2.2, whenever $x \in I_{\alpha^{(n_*)}} \in \eta$,

$$1/|T^{n_*\prime}(x)| \geq m(I_{\alpha^{(n_*)}})/D.$$

Since every branch of $T^{n_*}$ covers $U_*$,

$$\mathcal{L}^{n_*} g\Big|_{U_*} \geq \sum_{I_{\alpha^{(n_*)}} \in \eta} \left(\operatorname*{ess\,inf}_{I_{\alpha^{(n_*)}}} g\right) m(I_{\alpha^{(n_*)}})/D = \|\tilde{g}\|/D \tag{3.15}$$

where

$$\tilde{g} = \sum_{I_{\alpha^{(n_*)}} \in \eta} \left(\operatorname*{ess\,inf}_{I_{\alpha^{(n_*)}}} g\right) \chi_{I_{\alpha^{(n_*)}}}.$$

By Lemma 3.2.1, if $g \in \mathcal{C}_a$ then

$$\|\tilde{g}\| \geq (1 - \lambda^{-n_*} a)\|g\|. \tag{3.16}$$

Therefore, putting

$$\psi_g = \left((1 - \lambda^{-n_*} a)\|g\|/D\right)\chi_{U_*},$$

(3.15) and (3.16) imply

$$\mathcal{L}^{n_*} g \geq \psi_g \quad \text{and} \quad \|\psi_g\| = \frac{m(U_*)}{D}\left(1 - \lambda^{-n_*} a\right)\|g\|.$$

It now follows that

$$\|\mathcal{L}^{n_*} g - \psi_g\| = \left(1 - \frac{m(U_*)}{D}\left(1 - \lambda^{-n_*} a\right)\right)\|g\|.$$

To apply Theorem 1.1 we must check that $\mathcal{L}^{n_*}g - \psi_g \in \mathcal{C}_a$. Because $\psi_g$ is a multiple of the characteristic function of an interval,

$$\begin{aligned} V_I\left(\mathcal{L}^{n_*}g - \psi_g\right) \le V_I\left(\mathcal{L}^{n_*}g\right) + V_I(\psi_g) &= V_I\left(\mathcal{L}^{n_*}g\right) + |\psi_g| V_I(\chi_{U_*}) \\ &= V_I\left(\mathcal{L}^{n_*}g\right) + 2|\psi_g|. \end{aligned}$$

By iterated application of the Lasota–Yorke inequality,

$$V_I\left(\mathcal{L}^{n_*}g - \psi_g\right) \le \sigma^{n_*} V_I(g) + \frac{1-\sigma^{n_*}}{1-\sigma} A\|g\| + \frac{2}{D}\left(1 - \lambda^{-n_*}a\right)\|g\|.$$

Recalling that $g \in \mathcal{C}_a$, and combining the norm and variation estimates just derived, we have that $\mathcal{L}^{n_*}g - \psi_g \in \mathcal{C}_a$ provided that

$$\frac{\sigma^{n_*}a + \frac{1-\sigma^{n_*}}{1-\sigma}A + \frac{2}{D}\left(1 - \lambda^{-n_*}a\right)}{1 - \frac{m(U_*)}{D}\left(1 - \lambda^{-n_*}a\right)} \le a.$$

Since $a = A/(1-\sigma)$ does not satisfy this inequality, whereas $a = \lambda^{n_*}$ does, it follows that the positive root of the quadratic equation in the statement of the theorem is the minimal $a$ for which the condition is satisfied. Denoting this value by $a_*$, the theorem follows by the method of lower bounds for invariant cones whenever $a_* \le a < \lambda^{n_*}$. □

To use Theorem 3.3 to get rates of mixing for non–onto transformations, the following procedure is natural: choose the minimal $n_*$ such that $\lambda^{n_*} > A/(1-\sigma)$. This guarantees that the partition of $I$ into monotonicity branches of $T^{n_*}$ is sufficiently fine that

$$\left\| \sum_{I_{\alpha^{(n_*)}} \in \xi^{(n_*)}} \left( \operatorname*{ess\,inf}_{I_{\alpha^{(n_*)}}} g \right) \chi_{I_{\alpha^{(n_*)}}} \right\| > 0 \quad \text{whenever} \quad g \in \mathcal{C}_a$$

for $a < \lambda^{n_*}$. The next step is to find an interval $U$ which is contained in each $T^{n_*}(I_{\alpha^{(n_*)}})$. If none exists, then one can try and find such an interval for $n_*+1$, and so on. Of course, there is no guarantee that a $U_*$ will exist for an arbitrary transformation, but if one does, then Theorem 3.3 can be applied. Supposing that the hypotheses of the theorem are satisfied for $n_*, U_*$, the fastest rate of mixing guaranteed by the method will be for functions in $\mathcal{C}_{a_*}$, where $a_*$ is the minimal value of $a$ in the conclusion of the theorem.

REMARK. When using Theorem 3.3 to estimate rates of mixing, obtaining an optimal rate is a balancing act between choosing $n_*$ and $a_*$ as small as possible, and the interval $U_*$ as large as possible. □

### 3.4.2 $\beta$–transformations

Theorem 3.3 can be applied to the well–known class of $\beta$–transformations. The choice of a sequence of values of $\beta$ between 2 and 3 illustrates both situations

in which Theorem 3.3 gives good mixing estimates, and situations where the estimates are not so good. An alternative approach to mixing rates for BV transformations is described in [55].

Fix $\beta > 2$ and let $T_\beta : [0,1] \to [0,1]$ be defined by the formula

$$T_\beta(x) = \beta x \pmod 1.$$

The graph of $T = T_\beta$ consists of $\lfloor \beta \rfloor$ linear branches mapping over $[0,1)$ and one linear branch mapping over $[0, \beta - \lfloor \beta \rfloor)$. For each $n > 1$, the graph of $T^n$ consists of both onto branches, and branches which fail to cover the whole interval; all the branches have slope $\beta^n$ and have 0 as the image of the left hand endpoint. It follows easily that the expansivity and distortion constants for the maps are

$$\lambda = \beta, \quad s = 0, \quad D = 1.$$

Finally, note the non–onto branch has length $c_1 = \beta - \lfloor \beta \rfloor$. Then, Lemma 3.1.2 implies that for all $g \in BV$,

$$V_I(\mathcal{L}g) \leq \frac{2}{\beta} V_I(g) + \frac{2}{c_1} \|g\|.$$

Now, since each branch of $T_\beta^n$ has zero as the image of its left–hand endpoint, $T^n(I_{\alpha^{(n)}}) = [0, m(T^n(I_{\alpha^{(n)}})))$. Therefore, to find an interval which is covered by every branch of $T^n$, we need to calculate the minimum length of any branch of $T^n$:

Recall that $c_1 = \beta - \lfloor \beta \rfloor = \lim_{x \nearrow 1} T(x) \triangleq T(1^-)$. For each $n > 1$ put

$$c_n = \lim_{x \nearrow c_{n-1}} T(x) \triangleq T^n(1^-).$$

Then:

**Lemma 3.4.1** *For each $n > 0$ and monotonicity interval $I_{\alpha^{(n)}}$ of $T^n$,*

$$[0, b_n) \subset T^n(I_{\alpha^{(n)}})$$

*where $b_n = \min\{1, c_1, \ldots, c_n\}$.*

*Proof:* For each $n > 0$ let $B_n \subset (0,1]$ be the set of lengths of images of monotonicity branches of $T^n$. That is

$$x \in B_n \Leftrightarrow [0,x) = T^n(I_{\alpha^{(n)}}) \quad \exists I_{\alpha^{(n)}} \in \xi^{(n)}.$$

The lemma will follow if we can show that

$$B_n \subset \{1, c_1, \ldots, c_n\}.$$

The hypothesis is obvious for $n = 1$, so we suppose that it is true for $n = k$. Let $J$ be a monotonicity interval of $T^{k+1}$, and let $J'$ be the unique monotonicity

interval of $T^k$ such that $J \subset J'$. Then $T^k(J') = [0, c_j)$ for some $j \leq k$ and $T^{k+1}(J) = T(T^k(J))$ is the image of a monotonicity interval of $T|_{[0,c_j)}$. Hence

$$T^{k+1}(J) = [0,1) \quad \text{or} \quad T^{k+1}(J) = [0, T(c_j^-)) = [0, c_{j+1}).$$

Therefore, $B_{k+1} \subset B_k \cup \{c_{k+1}\}$, and the lemma follows by induction. □

Suppose that $\beta > 2$ is given. Put $n = \lceil \frac{\log A/(1-\sigma)}{\log \beta} \rceil$ and $U_*(n) = [0, b_n)$. By Lemma 3.4.1, we can apply Theorem 3.3 to obtain a rate of mixing for functions in $\mathcal{C}_{a_*}$, where $a_* = a_*(n)$ is defined in the statement of the theorem. We then obtain

$$\|\mathcal{L}^n g\|_{a_*} \leq r_{n,a_*} \|g\|_{a_*}$$

for each $g \in BV_0$, where $r_{n,a_*} < 1$ is from Theorem 3.3. Let $R_n = (r_{n,a_*})^{1/n}$. Then $R_n = R_n(\beta)$ is an upper bound for the spectral radius of $\mathcal{L}|_{BV_0}$. It may be that a larger value of $n$ will lead to a smaller value of $R_n$, and with this in mind it may be worth incrementally increasing $n$ until the sequence $R_{n+k}$ stops decreasing. For a selection of values of $\beta \in (2,3)$ we have performed this procedure. That is:

1. Use the Lasota–Yorke inequality to determine the minimal value of $n_*$ for which Theorem 3.3 will guarantee strict contraction for $\mathcal{L}^{n_*}$ in a $\|\cdot\|_a$–norm.

2. Calculate $U_*$ as described in Lemma 3.4.1, and let $a_*$ and $R_{n_*}(\beta)$ be derived from Theorem 3.3.

3. Increase $n_*$ by 1, and return to Step 2 until the sequence of values $\{R_{n_*+k}(\beta)\}$ stops decreasing.

The upper bounds on mixing rates for $\mathcal{L}|_{BV_0}$ given by this procedure are depicted in Figure 3.2.

Notice how some of the mixing rates depicted are extremely close to 1, and that the graph jumps sharply near those points. This happens when $\beta$ is very close to a value for which the transformation has a finite Markov partition. It is obvious that this corresponds to $T_\beta^m(1^-) = 1$ for some $m$, and one can check that such a $\beta$ satisfies a Diophantine equation

$$\beta^m = n_m + n_{m-1}\beta + \ldots + n_1\beta^{m-1} \tag{3.17}$$

where $0 \leq n_i \leq \lfloor \beta \rfloor$. (In fact, $\{n_1, n_2, \ldots\}$ encodes the sequence of monotonicity intervals of $T$ visited by the orbit of $1^-$: $T^i(1^-) \in (n_i/\beta, (n_i+1)/\beta) \cap (0,1)$.)

We give a phenomenological explanation for the apparent discontinuities in Figure 3.2: Suppose that $T_\beta^m(1^-) = 1$ for some $m > 0$. Then, $T_{\beta+\epsilon}^k(x) = T_\beta^k(x) + O(\epsilon) \pmod 1$ for most points $x$, sufficiently small $\epsilon$, and $k \leq m$. If $k < m$, then this implies that the monotonicity branches of $T_{\beta+\epsilon}^k$ are very similar in length to those of $T_\beta^k$, but if $k = m$ then $T_{\beta+\epsilon}^m(1^-) = O(\epsilon)$. Therefore,

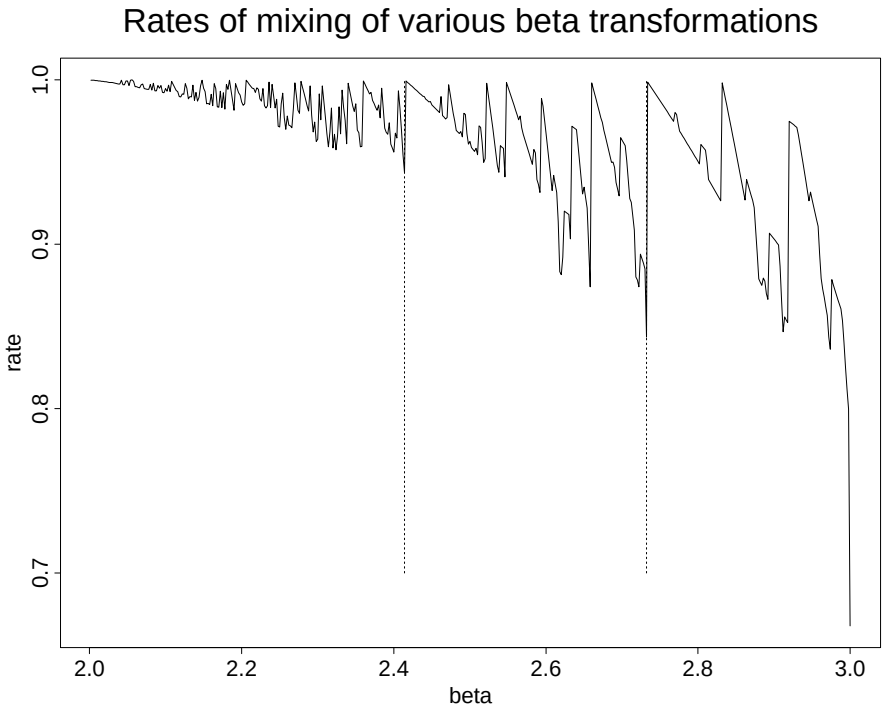



Figure 3.2: Rates of mixing for various $\beta$ transformations, as derived from Theorem 3.3. The vertical lines at $\beta = 1 + \sqrt{2}, 1 + \sqrt{3}$ are to accentuate the difficulties in using the Theorem to approximate mixing rates near Markov transformations.

the interval $U_*(n)$ in Theorem 3.3, is very small for $k \geq m$. Thus, the lower bound function used to construct the rate of mixing in the Theorem captures very little of the mass of an arbitrary initial density, giving a very slow rate of mixing[4]. On the other hand, when $\epsilon > 0$ is sufficiently small, all of the branches of $T^k_{\beta-\epsilon}$ are very similar in length to those of the transformation $T^k_\beta$, and the theorem should estimate the rate of mixing for $T_{\beta-\epsilon}$ at much the same value as it would for $T_\beta$.

Therefore, one would expect to find a solution of one of the Diophantine equations (3.17) near the sharp jumps in Figure 3.2. For example, the fact that $2 < \beta < 3$ implies that any equation (3.17) will have $n_1 = 2$. For $n_2 = 0, 1, 2$ and $m = 2$ we expect jumps at $\beta = 2, 1 + \sqrt{2}, 1 + \sqrt{3}$; these are clearly visible in the figure!

REMARK. Since the set of $\beta$ for which $T_\beta$ is Markov is dense, one is always "close" to a Markov transformation. The reason that this does not seem problematic for all of our attempts to use Theorem 3.3 is that the value of $n_*$ which is used for the estimates depicted in Figure 3.2 is relatively small. Therefore, the collection of $\beta$ for which $T^m_\beta(1) = (1)$ for $m \leq n_*$ is relatively sparse. $\square$

[4]Rather than implying that these transformations mix slowly, this only means that our method does not always do so well.

By following the arguments in Section 3.3, one can obtain an exponential decay of correlations result, and $O(\log n/n)$ error bounds for Ulam's method applied to $\beta$–transformations.

### 3.4.3 Transformations with Finite Range Structure

We briefly describe another "non–onto" application of Theorem 3.3. Following [37], consider the class of transformations satisfying a *finite range structure* condition. Here, we just describe the kind of mixing result that Theorem 3.3 implies; it will not be until the application to the Jacobi–Perron transformation in chapter 5 that the idea is used.

DEFINITION (**Finite Range Structure [37]**) Suppose there exists a finite collection $\{U_1, \ldots, U_k\}$ of intervals in $[0,1]$ such that for every $n > 0$ and monotonicity interval $I_{\alpha^{(n)}}$ of $T^n$

$$T^n(I_{\alpha^{(n)}}) = U_i \quad \exists i = i(\alpha^{(n)}) \in 1 \ldots k.$$

Then $T$ has a *Finite Range Structure.* □

REMARK. Ito and Yuri's [37] definition is not restricted to expanding transformations. They prove that under some bounded distortion and transitivity assumptions there exists a unique ergodic absolutely continuous invariant measure whose density is bounded above and below. Indeed, in subsequent papers, Yuri [66, 67] generalises the definitions and constructions to multi–dimensional transformations, proving both the existence of acim and decay of correlations (exponential mixing). □

The full generality of Ito and Yuri's [37] constructions is much greater than that which we now use. Instead, we assume that $T$ satisfies the expansivity and bounded distortion conditions above, and that $T$ has a *finite range structure.* If the collection of range sets is $\{U_1, \ldots, U_k\}$, then put

$$A = \cap_{i=1}^{k} U_i. \tag{3.18}$$

If $m(A) \neq 0$, then Theorem 3.3 implies an explicit rate of mixing for $T$. Because the application of the theorem is so direct, and any discussion of weaker conditions than (3.18) would be impossible without a substantial complication of our method, we do not attempt further generalisation.

### 3.4.4 Liverani's technique

Finally, we comment on Liverani's [50] constructions of contraction rates for non–onto maps, and compare it to ours:

In [50], the cones $\mathcal{C}_a$ are also used. Liverani's interest is in obtaining rates of contraction for $\mathcal{L}$ in a *projective Hilbert metric*, and this necessitates the construction of lower bounds for functions in $\mathcal{L}^{N_0}\mathcal{C}_a$ (for some $N_0$); c.f. [50, Lemma 3.5].

The lower bounds constructed in Proposition 3.1 and Theorem 3.3 are much better than Liverani's because we sum up contributions to $\mathcal{L}^n f$ from *all* inverse branches of $T^n$, whereas Liverani uses the contribution from just one onto branch. On the other hand, Liverani proves (non–constructively) that a rather weak mixing assumption is sufficient to guarantee the existence of at least one onto branch for some iterate of a non–onto transformation [50, Theorem 4.4]. His arguments can be applied to obtain *decay of correlations* for these transformations. It is therefore reasonable to suppose that our Mixing Theorems may also be extended to more general transformations.

## 3.5 Reworking the method in terms of the matrix approximation

For the final part of the chapter, we depart from the theme of obtaining error bounds for the Ulam approximation, and show instead how the spectral bounds for $\mathcal{L}$ in Theorem 3.1 can be used to get explicit spectral bounds for the approximate operator $\mathcal{P}_n$. For the sake of simplicity, we confine the discussion once more to onto transformations. As well as being an encouraging sign for the robustness of our method, the persistence of spectral bounds under Ulam approximation leads to bounds on the sample path variance in a *Central Limit Theorem* for the associated Markov chain.

### 3.5.1 Spectral bounds for $\mathcal{P}_n$

Recall the approximate operator $\mathcal{P}_n$ and in analogy with Theorem 3.1, consider the action of $\mathcal{P}_n$ on

$$\Delta_n = \Big\{ f \in L^1 : \int_I f = 0 \quad \text{and} \quad \exists f^{(1)}, f^{(2)} \in \mathcal{D}_n \ \text{ s.t.} \\ f = f^{(1)} - f^{(2)} \ \text{ and } \ \|f^{(1)}\| = \|f^{(2)}\| \Big\}.$$

The space $\Delta_n$ has particularly nice properties:

**Lemma 3.5.1** *For each $n$, recall the definition of the norm $\|\cdot\|_{2n}$ on $\Gamma_{2n}$.*

1. *$\mathcal{D}_n \subset \mathcal{C}_{2n}$ so that $\Delta_n \subset \Gamma_{2n}$ and*
2. *if $f \in \Delta_n$ then $\|f\| = 2\|f\|_{2n}$.*

*Proof:* See Appendix A. □

REMARK. Because $\mathcal{D}_n$ is preserved by $\mathcal{P}_n$, it follows immediately that $\Delta_n$ is preserved by $\mathcal{P}_n$. Of course, this is equivalent to the fact that the matrix representation on $\mathcal{P}_n$ preserves $\{(x_1, \ldots, x_n) \in \mathbb{R}^n : \sum_{i=1}^n x_i = 0\}$. □

To get contraction for functions in $\Delta_n$ we are going to use the method of lower bounds for invariant cones.

Put

$$a_* = 5\frac{\frac{1}{2} + \frac{s\lambda}{\lambda-1}}{1 - \frac{1}{3D}}, \quad \text{and} \quad n_* = \left\lceil \frac{\log 2a_*}{\log \lambda} \right\rceil .$$

**Proposition 3.3 (Lower bound for powers of $\mathcal{P}_n$)** *Let $n \geq 6Dn_*a_*$, and suppose that $f, g \in \mathcal{C}_{a_*} \cap \mathcal{D}_n$ satisfy $\|f\| = \|g\|$. Then*

$$\mathcal{P}_n^{n_*} f, \mathcal{P}_n^{n_*} g \in \mathcal{C}_{(1/2+s\lambda/(\lambda-1))} \cap \mathcal{D}_n$$

*and there exists a function $\psi = \psi_{f,g} \in \mathcal{D}_n$ such that*

$$\mathcal{P}_n{}^{n_*} f \geq \psi, \mathcal{P}_n{}^{n_*} g \geq \psi, \quad \|\psi\| \geq \|f\|/3D$$

*and*

$$V_I(\psi) \leq 4(1/2 + s\lambda/(\lambda-1))\|f\|.$$

*Proof:* First of all, repeated alternate application of Lemmas 3.1.2 and 3.1.1 gives

$$\begin{aligned} V_I(\mathcal{P}_n{}^{n_*} f) &\leq \lambda^{-n_*} V_I(f) + s\lambda/(\lambda-1)\|f\| \\ &\leq (a_*\lambda^{-n_*} + s\lambda/(\lambda-1))\|f\| \\ &\leq (1/2 + s\lambda/(\lambda-1))\|f\|. \end{aligned} \tag{3.19}$$

Since the same is true for $g$, this establishes the first part of the proposition (it is obvious that $\mathcal{P}_n{}^{n_*}\mathcal{D}_n \subset \mathcal{D}_n$).

Now, because $f \in \mathcal{C}_{a_*}$ and $\lambda^{-n_*}a_* \leq 1/2$, Proposition 3.1 implies that

$$\mathcal{L}^{n_*} f \geq \|f\|/2D. \tag{3.20}$$

Next,

$$\begin{aligned} \|\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f\| &\leq \sum_{i=1}^{n_*} \|\mathcal{L}^i \mathcal{P}_n{}^{(n_*-i)} f - \mathcal{L}^{i-1}\mathcal{P}_n\mathcal{P}_n{}^{(n_*-i)} f\| \\ &\leq \sum_{i=1}^{n_*} \|\mathcal{L}\mathcal{P}_n{}^{(n_*-i)} f - \Pi_n\mathcal{L}\mathcal{P}_n{}^{(n_*-i)} f\| \\ &\leq \frac{1}{n}\sum_{i=1}^{n_*} V_I(\mathcal{L}\mathcal{P}_n{}^{(n_*-i)} f) \\ &\leq \frac{n_*}{n} a_* \|f\| \leq \|f\|/6D \end{aligned}$$

by Lemma 3.1.1 and the restriction on $n$. Because $\|\mathcal{L}^{n_*} f\| = \|\mathcal{P}_n{}^{n_*} f\|$, the above implies that

$$\|(\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f)^+\| = \|\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f\|/2 \leq \|f\|/12D. \tag{3.21}$$

Moreover, similar to the proof of (3.19):

$$\mathcal{L}^{n_*} f \in \mathcal{C}_{(1/2+s\lambda/(\lambda-1))}.$$

Thus

$$V_I((\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f)^+) \le V_I(\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f) \le 2(1/2 + s\lambda/(\lambda-1))\|f\|. \quad (3.22)$$

Equations (3.20)—(3.22) hold with $g$ replacing $f$. Then, put

$$\begin{aligned} \psi' = \psi'_{f,g} = 1/2D\|f\| \quad &- \quad (\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f)^+ \\ &- \quad (\mathcal{L}^{n_*} g - \mathcal{P}_n{}^{n_*} g)^+. \end{aligned}$$

Together, (3.20) and (3.21) imply that $\|\psi'\| \ge 1/3D\|f\|$, and (3.22) implies that

$$V_I(\psi') \le 4(1/2 + s\lambda/(\lambda-1))\|f\|.$$

Also,

$$\begin{aligned} \mathcal{P}_n{}^{n_*} f &= \mathcal{L}^{n_*} f - (\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f) \\ &\ge 1/2D\|f\| - (\mathcal{L}^{n_*} f - \mathcal{P}_n{}^{n_*} f)^+ \\ &\ge \psi' \end{aligned}$$

and the same relation holds for $g$. Putting $\psi = \Pi_n \psi'$, the result follows. □

We now prove strict contraction for powers of $\mathcal{P}_n$ applied to any function in $\Delta_n$.

For each $a > 0$ let

$$n_a = \begin{cases} 0 & \text{if } a \le a_*, \\ \left\lceil \frac{\log 2a}{\log \lambda} \right\rceil & \text{if } a > a_*. \end{cases}$$

**Theorem 3.4 (Mixing for $\mathcal{P}_n$)** *Suppose that $f \in \Delta_n$, and let $n, a_*, n_*, n_a$ be as given in Proposition 3.3. Then, for any $a \ge 0$:*

$$\|\mathcal{P}_n{}^{(n_*+n_a)} f\|_{a_*} \le (1 - 1/3D)\|\mathcal{P}_n{}^{n_a} f\|_{a_*} \le (1 - 1/3D)\|f\|_a.$$

*Proof:* First of all, assume that $n_a = 0$. Then $a \le a_*$ and there exist functions $f^{(1)}, f^{(2)} \in (\mathcal{D}_n \cap \mathcal{C}_a) \subset (\mathcal{D}_n \cap \mathcal{C}_{a_*})$ such that $f = f^{(1)} - f^{(2)}$ and $\|f^{(1)}\| = \|f^{(2)}\| = \|f\|_a \ge \|f\|_{a_*}$. By Proposition 3.3,

$$\mathcal{P}_n{}^{n_*} f^{(1)}, \mathcal{P}_n{}^{n_*} f^{(2)} \in \mathcal{C}_{(1/2+s\lambda/(\lambda-1))},$$

and there exists $\psi = \psi_{f^{(1)}, f^{(2)}} \in \mathcal{D}_n$ such that

$$\mathcal{P}_n{}^{n_*} f^{(1)} \ge \psi, \mathcal{P}_n{}^{n_*} f^{(2)} \ge \psi, \quad \|\psi\| \ge \|f^{(1)}\|/3D$$

and

$$V_I(\psi) \le 4(1/2 + s\lambda/(\lambda-1))\|f^{(1)}\|.$$

It follows immediately that

$$\|\mathcal{P}_n{}^{n_*} f^{(1)} - \psi\| = \|\mathcal{P}_n{}^{n_*} f^{(1)}\| - \|\psi\| \geq (1 - 1/3D)\|f^{(1)}\|$$

(and the same is true with $f^{(2)}$ replacing $f^{(1)}$). Therefore,

$$\begin{aligned} V_I(\mathcal{P}_n{}^{n_*} f^{(1)} - \psi) &\leq V_I(\mathcal{P}_n{}^{n_*} f^{(1)}) + V_I(\psi) \\ &\leq (1/2 + s\lambda/(\lambda - 1))\|f^{(1)}\| + 4(1/2 + s\lambda/(\lambda - 1))\|f^{(1)}\| \\ &= a_*(1 - 1/3D)\|f^{(1)}\| \\ &\leq a_*\|\mathcal{P}_n{}^{n_*} f^{(1)} - \psi\| \end{aligned}$$

(and the same is true with $f^{(2)}$ replacing $f^{(1)}$). By Theorem 1.1,

$$\|\mathcal{P}_n{}^{n_*} f\|_{a_*} \leq (1 - 1/3D)\|f\|_{a_*}.$$

Suppose now that $n_a > 0$. Then $a > 5a_*$, and repeated alternate application of Lemmas 3.1.2 and 3.1.1 implies

$$\mathcal{P}_n{}^{n_a}\mathcal{C}_a \subset \mathcal{C}_{\lambda^{-n_a}a + s\lambda/(\lambda-1)} \subset \mathcal{C}_{1/2 + s\lambda/(\lambda-1)} \subset \mathcal{C}_{a_*},$$

where the last inclusions follow from Lemma 3.2.3 (2). By Lemma 3.2.3 (3),

$$\|\mathcal{P}_n{}^{n_a} f\|_{a_*} \leq \|f\|_a,$$

and the proposition follows from the $n_a = 0$ case. □

REMARK. The constants $a_*, n_*$ and the restriction on $n$ have been chosen (somewhat arbitrarily) to get the simply expressed contraction rate $(1 - 1/3D)$. □

**Corollary 3.5.2** *Let $n_*$ be as in Proposition 3.3 and for sufficiently large $n$ put $n_0 = \lceil \frac{\log 4n}{\log \lambda} \rceil$. Then for $f \in \Delta_n$, $k > 0$*

$$\|\mathcal{P}_n{}^{(kn_* + n_0)} f\| \leq (1 - 1/3D)^k \|f\|.$$

*Proof:* By Lemma 3.5.1, Lemma 3.2.3 and repeated application of Theorem 3.4,

$$\begin{aligned} \|\mathcal{P}_n{}^{(kn_* + n_0)} f\| = 2\|\mathcal{P}_n{}^{(kn_* + n_0)} f\|_{2n} &\leq 2\|\mathcal{P}_n{}^{kn_*}\mathcal{P}_n{}^{n_0} f\|_{a_*} \\ &\leq 2(1 - 1/3D)^k \|\mathcal{P}_n{}^{n_0} f\|_{a_*}. \end{aligned}$$

By the usual argument, $(\mathcal{P}_n{}^{n_0} f)^{\pm} \in \mathcal{C}_{a_*}$, so that

$$\|\mathcal{P}_n{}^{n_0} f\|_{a_*} = \|(\mathcal{P}_n{}^{n_0} f)^+\| = \|f^+\| = \frac{1}{2}\|f\|.$$

The corollary follows. □

Finally, recall that the Ulam approximate operator $\mathcal{P}_n$ can be expressed as a stochastic matrix $P = P(n)$ with entries

$$P_{ij} = m(T^{-1}B_j \cap B_i)/m(B_i)$$

where $\{B_i\}_{i=1}^n$ is the partition of $I$ into $n$ equal subintervals. Since there is a bijective correspondence between elements of $\mathcal{D}_n$ and probability vectors in $\mathbb{R}^n$, Corollary 3.5.2 implies immediate bounds on the eigenvalues of the matrices $P(n)$. Let

$$\rho_* = (1 - 1/3D)^{1/n_*}.$$

**Corollary 3.5.3** *For large enough $n$ (as specified by Theorem 3.4), let $\rho$ be any eigenvalue of the matrix $P(n)$. Then either*

$$\rho = 1 \quad \text{or} \quad |\rho| \leq \rho_*.$$

REMARK 3.5.4. In view of Corollary 3.5.3, the numerical computation of the invariant probability vector for $P(n)$ is relatively straightforward. Because the matrix $P(n)$ is ergodic, and the rate of exponential mixing is known to be bounded above by $\rho_*$, a uniform initial vector will converge to within $\epsilon$ of the invariant probability vector after $O(\log(\epsilon))$ iterative multiplications by the matrix $P(n)$. Consequently, this part of the calculation is extremely fast once the matrix $P(n)$ is stored in the computer. □

### 3.5.2 Central Limit Theorem

The bounds on approximation error for Ulam's method for computing the invariant measure lead automatically to rigorous estimates of ergodic averages of observables. The main idea in this section is that an average over a finite length sample path of the Markov Chain corresponding to the Ulam approximation provides a good estimate of the average with respect to the exact invariant measure for the transformation. We state this as a Central Limit Theorem (CLT) for averages over sample paths of the Ulam approximating Markov chain. Theorem 3.4 is used to obtain a rigorous quantitative bound on the variance appearing in the CLT.

Let $\phi : I \to \mathbb{R}$ be bounded, and for each $x \in I$ let

$$S_n(\phi)(x) = \sum_{i=0}^{n-1} \phi(T^i x),$$

and suppose that it is of interest to know

$$\bar{\phi} = \lim_{n\to\infty} \frac{1}{n} S_n(\phi)$$

for a.e. $x \in I$. By the Birkhoff Ergodic Theorem, if $\mu$ is the unique absolutely continuous invariant measure for $T$,

$$\bar{\phi} = \int_I \phi \, d\mu.$$

Since $\mu$ has density $h$, by Theorem 3.2

$$\begin{aligned}\left|\bar{\phi} - \int_I \phi(x)h_n(x)dm(x)\right| &\leq \int_I |\phi(x)||h(x) - h_n(x)|dm(x) \\ &\leq \left(\sup_I |\phi|\right)\|h - h_n\| \leq C\|\phi\|_\infty \log(n)/n\end{aligned}$$

where the constant $C$ can be written down in terms of the expansivity and distortion constants of the map $T$.

Therefore, one strategy for approximating $\bar{\phi}$ is to compute the approximate invariant density $h_n$ (for some sufficiently large $n$) and then calculate the mean of $\phi$ with respect to $h_n$. Since the function $h_n$ arises as the invariant probability vector of a stochastic matrix $P(n)$, an alternative to doing this calculation is to perform a "Monte Carlo" simulation of the corresponding Markov chain. We describe this now:

To avoid ambiguity with the standard Central Limit Theorem notation, we will work with the $N$th Ulam approximation. Let $\{B_i\}_{i=1}^N$ be the partition of $[0,1]$ into $N$ equal subintervals. The idea is to simulate a sample path of the Markov chain with states $B_i$ by generating a pseudo–orbit of $T$ on $[0,1]$ and recording the sequence of $B_i$s which are visited.

Consider the Markov chain on $[0,1]$ with transition probabilities

$$\mathbb{P}_x(A) = \sum_{i,j=1}^N \chi_{B_i}(x)P_{ij}\frac{m(A \cap B_j)}{m(B_j)}$$

where $P_{ij}$ are the entries of $P(N)$, the stochastic matrix representation of $\mathcal{P}_N$. Let $\{x_k\}$ be a sample path of the Markov chain. The sequence of states $\{B_{i_k}\}$ such that $x_k \in B_{i_k}$ is a sample path of the Markov chain representation of $\mathcal{P}_n$:

$$\mathbb{P}(x_{k+1} \in B_j | x_k \in B_i) = \sum_{j'=1}^N P_{ij'}\frac{m(B_{j'} \cap B_j)}{m(B_{j'})} = P_{ij}.$$

Moreover, a typical sample path is very easy to generate: If $x_k \in B_i$, let $x'_k \in B_i$ be distributed according to Lebesgue measure on $B_i$. Put $x_{k+1} = T(x'_k)$. Denoting the transition probabilities of this process by $\mathbb{P}'$:

$$\mathbb{P}'(x_{k+1} \in B_j | x_k \in B_i) = \mathbb{P}'(Tx'_k \in B_j | x'_k \in B_i) = m(T^{-1}B_j \cap B_i)/m(B_i) = P_{ij},$$

as required.

Now, for each $n$, put

$$S'_n(\phi) = \sum_{i=0}^{n-1} \phi(x_i).$$

By the Ergodic Theorem for finite state Markov chains ,

$$\lim_{n\to\infty} \frac{1}{n}S'_n(\phi) = \int_I \phi(x)h_N(x)dm(x) = \phi'.$$

Now, let $g = \Pi_N \phi - \phi'$ and consider the quantity

$$\gamma_g^2 = \int g^2 h_N dm + 2 \sum_{k=1}^{\infty} \int g \cdot \mathcal{P}_N^k g \cdot h_N dm.$$

**Corollary 3.5.5** *Let $n_*$ be as in Theorem 3.4 and let $N$ be sufficiently large that the other hypotheses of Theorem 3.4 hold. Put $N_0 = \lceil \frac{\log 4N}{\log \lambda} \rceil$. Then*

$$\gamma_g^2 \leq 2\Big(N_0 + 3Dn_*\Big)\|g \cdot h_N\|_\infty \|g\|.$$

*Proof:* By Corollary 3.5.2

$$\int g \cdot \mathcal{P}_N^k g \cdot h_N dm \leq \|g \cdot h_N\|_\infty \|\mathcal{P}_N^k g\| \leq \|g \cdot h_N\|_\infty \min\{1, (1 - 1/3D)^{\lfloor \frac{k-N_0}{n_*} \rfloor}\} \|g\|.$$

The result follows by summation. □

Finally, we can write down a Central Limit Theorem.

**Central Limit Theorem** [52, 17.0.1] *In the established notation, for every $t \in \mathbb{R}$*

$$\lim_{n \to \infty} \mathbb{P}\left( \frac{S_n'(\phi) - n\phi'}{\sqrt{n\gamma_g^2}} < t \right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-x^2/2} dx. \ \square$$

By Corollary 3.5.5, $\gamma_g^2 = O(\log N)$ is very small in comparison to $N$. Therefore, with high probability

$$\left| \frac{1}{n} S_n'(\phi) - \phi' \right| = O(\sqrt{\log N}/N) \ \text{ if } n = O(N^2).$$

This statement could be made precise with more detailed knowledge of the rate of convergence in the CLT. In combination with the discussion at the beginning of this subsection, we have with high probability

$$\left| \frac{S_{N^2}'(\phi)}{N^2} - \bar{\phi} \right| \leq C' \|\phi\|_\infty \log(N)/N$$

where $\bar{\phi}$ is the desired ergodic average for the exact invariant measure, and the sample path average is for the Markov approximation. The constant $C'$ may be slightly larger than is given by Theorem 3.2 to allow for the approximation error in the CLT.

EXAMPLE: Consider the estimation of Lyapounov Exponents for the maps $T_\Omega$ introduced in Section 3.2. In this case, the function $g = \log |T_\Omega'|$. Let $\sigma_\Omega = \int \log |T_\Omega'| \, h_\Omega \, dm$ and if $h_{N,\Omega}$ denotes the Ulam approximation to $h_\Omega$ by

using a partition of $N = 10000$ subintervals, let $\sigma_{10000,\Omega} = \int \log |T'_\Omega| h_{N,\Omega}\, dm$. Then

$$|\sigma_\Omega - \sigma_{10000,\Omega}| \leq \| \log |T'_\omega| \|_\infty \| h_\Omega - h_{N,\Omega} \| \leq \log 2.5 \times 0.0449 \approx 0.0411. \quad (3.23)$$

Using the density approximations $h_{10000,\Omega}$, the approximate Lyapounov Exponents $\sigma_{10000,\Omega}$ can be calculated exactly; the data are displayed in the left–hand plot of Figure 3.3. Alternatively, $\sigma_{10000,\Omega}$ could be approximated by averaging $\log |T'_\Omega|$ over a *Monte–Carlo* sample path of the Markov chain, as described above. For paths of lengths $n = 10^5$ and $n = 10^6$ the result of the procedure is depicted in the right–hand plot of Figure 3.3. The Monte–Carlo method produces estimates which agree well with the exact integrated version.

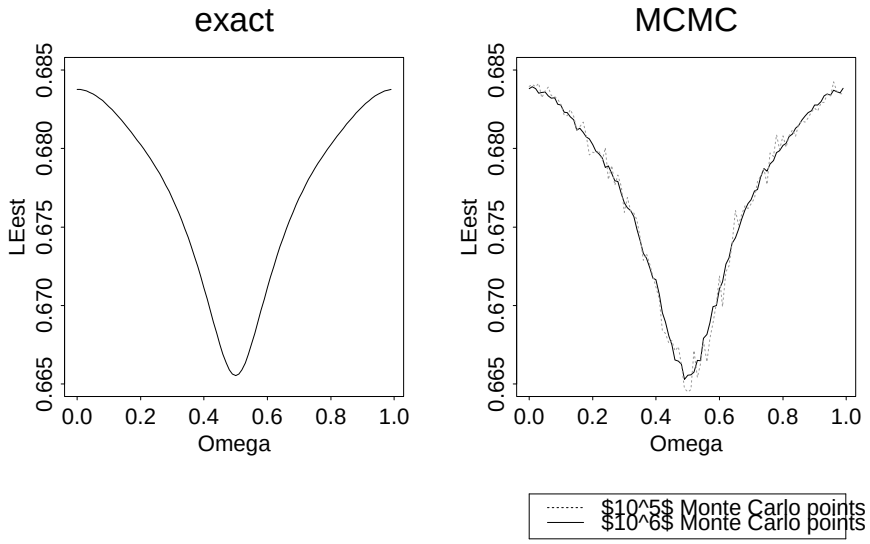



Figure 3.3: Comparison of integration based and Monte Carlo approximation of Lyapounov Exponents for the maps $T_\Omega$.

In our implementation, simulating a sample path of length $10^6$ takes approximately the same amount of time as calculating the approximate invariant measure (over 10000 subintervals) directly (about 10 seconds). By bounding $\gamma_g^2$ as in Corollary 3.5.5, and assuming that the convergence in the CLT happens quickly, $n = 10^6$ should produce (with high probability) a Lyapounov Exponent estimate of the same order of accuracy as in (3.23). In fact, the picture for $n = 10^4$ (which is omitted to avoid a cluttered appearance in Figure 3.3) is also in good agreement with the left–hand plot, suggesting that the convergence is very fast indeed. If this occurs generally, the computational advantages are obvious: the Monte–Carlo method is almost as accurate, much faster, and

requires almost no memory. These considerations become especially important for higher dimensional maps.

## Summary

This chapter has collected a reasonable breadth of results about expanding interval maps. First of all, a new method for constructing rates of mixing for $\mathcal{L}$ (the Perron–Frobenius operator of such a map) was presented. These spectral estimates are better than previous approaches [61, 50], apply automatically to maps with infinitely many branches, and are readily generalisable to multi–dimensions. In the course of the proof, the classical Lasota–Yorke [47] was extended to allow for infinitely many branches, and a new family $\|\cdot\|_a$ of $BV$ norms were introduced. Most importantly, the method of proof is adaptable to the situation where not all the branches of the map $T$ are onto, and this was illustrated by bounding mixing rates for the classical $\beta$–transformations. However, the main result of the chapter was the derivation of an $O(\log n/n)$ error bound for Ulam's method in all the situations where the mixing rates apply. This generalises the main result of Chapter 2, and because most of the argument does not rely on $T$ being one–dimensional, the approach can be generalised to more dimensions. The final application was to prove uniform mixing rates for the Ulam approximations of $\mathcal{L}$. From an aesthetic point of view, the persistence of our argument under "discretisation" is very pleasing. Yet the real benefit is practical: the existence of uniform spectral estimates for the corresponding Markov matrices implies rigorous bounds on the sample path variance of "Monte–Carlo" simulations of the underlying Markov chain. One can therefore perform numerical experiments with some confidence in the results.

Having achieved this variety of theoretical results for expanding maps in one dimension, there are two natural tasks ahead: to investigate how the existing constructions behave under numerical implementation, and to generalise the techniques to more dimensions. These are the subjects of the remaining two chapters of the thesis.

# Chapter 4

# Numerically assisted error bounds

One of the main themes of this thesis is the robustness to discretisation of analytical constructions for invariant densities. As such, it is appropriate to devote some attention to the translation of our constructions onto a computational level. The main theorem of this chapter (Theorem 4.1) implies a procedure for computer assisted proofs of error bounds for Ulam's method. The examples from Chapters 2 and 3 are revisited in this new context, and some further examples are introduced.

Now, with the emphasis on *numerical* determination of invariant density approximations and error bounds, the technical requirements are less restrictive than in the previous two chapters. In particular, more general transformations can be considered. To begin with, we still require fairly uniform expansivity, but the transformations may be multi–dimensional. Even the uniform expansivity is relaxed later on.

The first three sections of the chapter are devoted to constructing a numerical procedure for bounding the approximation error in Ulam's method. The technique draws only on linear algebra and elementary Markov Chains theory.

In the first section, we derive for the invariant measure an *exact asymptotic formula* in terms of a certain collection of matrices. We then show that each matrix in the collection is uniformly close to the matrix representation of an appropriate Ulam approximate operator $\mathcal{P}$. In the second section, we explain how to numerically compute a bound on $\|(Id-\mathcal{P})^{-1}\|$ (c.f. the analytical result in Lemma 3.3.1). Then, in Section 4.3, this "mixing" result is put together with the perturbation results from Section 4.1 to get numerically computable error bounds for the Ulam approximation.

After revisiting the examples from Chapters 2 and 3 in this new context, we apply the method to the class of smooth one–dimensional maps mentioned in the Chapter 1. The trick employed in that section suggests an easily implementable procedure for improving the accuracy of Ulam's method for non–

expansive maps. This is discussed in the final section of the chapter.

For the sake of simplicity, the examples presented in this chapter are all one–dimensional, even though the results apply to multi–dimensional transformations.

## 4.1 Matrix formula for invariant measure

In this section we describe the class of (multi–dimensional) maps under consideration, and derive a matrix formula for the invariant measure restricted to a partition $\eta$ (Corollary 4.1.2). We then show how the matrices appearing in the formula are small perturbations of the matrix representation of the Ulam approximation over $\eta$.

### Setup

Let $X$ be a compact Riemannian manifold, with Riemannian metric $\rho$. Let $m(\cdot)$ be the Lebesgue measure corresponding to $\rho$; we assume that $m$ is normalised to be a probability measure ($m(X) = 1$).

Next, let $T : X \to X$ be a piecewise $\mathcal{C}^2$ transformation satisfying the following conditions:

**Piecewise Monotonicity:** Let $\xi = \{B_\alpha\}$ be a partition of $X$ into *monotonicity components*. That is, each $T|_{B_\alpha}$ is 1-1, and each $B_\alpha$ is a connected set such that the tangent map[1]

$$DT : \mathcal{T}_x X \to \mathcal{T}_{T(x)} X$$

is invertible at each point $x \in B_\alpha \setminus \partial B_\alpha$.

**Uniform Expansivity:** Let $\lambda > 1$ and $K \geq 0$ be such that

$$|DT(x)v| \geq \lambda |v| \qquad \forall v \in \mathcal{T}_x X \tag{4.1}$$

and for each monotonicity set $B_\alpha$

$$\left| \log \left| \det \left( D(T_\alpha^{-1})(x) \right) \right| - \log \left| \det \left( D(T_\alpha^{-1})(y) \right) \right| \right| \leq K \rho(x, y) \tag{4.2}$$

where $x, y \in T(B_\alpha)$ and $T_\alpha^{-1}$ denotes the appropriate inverse branch of $T$.

**Finite Range Structure [37]:** There exists a finite collection $\{U_1, \ldots, U_k\}$ of subsets of $X$ such that for any $n > 0$ and monotonicity set $B_{\alpha^{(n)}}$ of $T^n$

$$T^n(B_{\alpha^{(n)}}) = U_i \quad \exists i = i(\alpha^{(n)}) \in 1 \ldots k. \tag{4.3}$$

DEFINITION: A partition $\eta$ of $X$ will be said to *respect the (finite) range structure (mod 0)* if for every $A \in \eta$ and $U_i \in \{U_1, \ldots, U_k\}$ either

$$m(A \cap U_i) = m(A) \text{ or } m(A \cap U_i) = 0.$$

[1] The notation $\mathcal{T}_x X$ denotes the tangent space to $X$ at $x$, and for every $v \in \mathcal{T}_x X$, $|v|$ is its length, measured in the norm corresponding to the Riemannian metric $\rho$.

It is easy to see that any refinement of a partition which respects the range structure must also respect the range structure. □

EXAMPLE: Any expanding, $\mathcal{C}^2$ endomorphism of a compact manifold has a finite range structure, with $\{U_1, \ldots, U_k\} = \{U_1\} = \{X\}$, so that any partition respects the range structure. Every transformation with a finite Markov partition has a finite range structure, and it is automatic that any Markov partition respects the range structure. □

**Theorem (Folklore)** *For $(X, T)$ as above, there exists an absolutely continuous invariant measure $\mu$. If the transformation $T$ is ergodic with respect to Lebesgue measure, then the acim $\mu$ is unique, and for every Borel set $A$*

$$\mu(A) = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} m\left(T^{-k}(A)\right).$$

REMARK. The Folklore Theorem dates back to Renyi [59], and is true far more generally than for the transformations $T$ just described (see [16, 66] and many others). I am grateful to Sebastian van Strien for pointing out this characterisation of the acim. □

## Probability vectors and a formula for the invariant measure

Let $\eta$ be a fixed partition which respects the range structure of $T$ (mod 0). Write $\eta = \{X_1, \ldots, X_N\}$ for some $N = N(\eta)$. The aim in this subsection is to use the Folklore Theorem to write down an "exact" asymptotic expression for the invariant measure $\mu$ restricted to the partition $\eta$.

Let

$$\mathbb{R}^N_{\text{prob}} = \left\{ x \in \mathbb{R}^N : \sum_{i=1}^{N} x_i = 1 \text{ and } x_i \geq 0 \quad \forall i = 1 \ldots N \right\}$$

denote the collection of probability vectors in $\mathbb{R}^N$. Then $\mathbb{R}^N_{\text{prob}}$ is a level set of the cone of positive vectors in $\mathbb{R}^N$, and there is a one–to–one correspondence between probability measures restricted to $\eta$, and elements of $\mathbb{R}^N_{\text{prob}}$: if $\nu \in \mathcal{M}(X)$, then

$$\nu_{(\eta)} \triangleq (\nu(X_1), \ldots, \nu(X_N))$$

is the *vector representation of $\nu$ (restricted to $\eta$).*

For each $n \geq 0$, let $x^{(n)}$ be the vector representation of $m \circ T^{-n}$ restricted to $\eta$. Then

$$(x^{(n)})_i = m\left(T^{-n}(X_i)\right), \tag{4.4}$$

and $x^{(n)} \in \mathbb{R}^N_{\text{prob}}$. If $\mu$ denotes the restriction of the acim (from the Folklore Theorem) to $\eta$, then $\mu \in \mathbb{R}^N_{\text{prob}}$ and

$$\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} x^{(k)}. \tag{4.5}$$

Next, we write down a recurrence relation satisfied by the vectors $x^{(k)}$:

DEFINITION (**The matrices $P^{(n)}$**) Let $P^{(0)}$ be the $N \times N$ identity matrix, and for each $n > 0$ and $i, j = 1 \dots N$, put

$$P_{ij}^{(n)} = \frac{m\left(T^{-n}X_j \cap T^{-(n-1)}X_i\right)}{m\left(T^{-(n-1)}X_i\right)}.$$

It is easy to check that each $P^{(n)} = \left(P_{ij}^{(n)}\right)$ is a stochastic matrix. □

**Lemma 4.1.1** *For each $n > 0$,*

$$x^{(n)} = x^{(n-1)}P^{(n)}$$

*where the vectors $x^{(n)}$ are defined by (4.4).*

*Proof:* For each $n > 0$, $1 \le i \le N$,

$$\begin{aligned}(x^{(n)})_j = m\left(T^{-n}X_j\right) &= \sum_{i=1}^{N} m\left(T^{-n}X_j \cap T^{-(n-1)}X_i\right) \\ &= \sum_{i=1}^{N} m(T^{-(n-1)}X_i)\frac{m\left(T^{-n}X_j \cap T^{-(n-1)}X_i\right)}{m(T^{-(n-1)}X_i)} \\ &= \left(x^{(n-1)}P^{(n)}\right)_j\end{aligned}$$

because $T^{-(n-1)}\eta = \{T^{-(n-1)}X_1, \dots, T^{-(n-1)}X_N\}$ is a partition of $X$. □

REMARK. Notice that if $\mathcal{L}$ is the Perron–Frobenius operator for $T$, and $P$ is the matrix representation of the Ulam approximation $\mathcal{P}_\eta$, then $P = P^{(1)}$. Moreover, $P^{(n)}$ is the matrix representation of the Ulam approximation over the partition $T^{-(n-1)}\eta$. □

Finally, it follows from the Folklore Theorem and Lemma 4.1.1 that:

**Corollary 4.1.2 (Formula for $\mu$)** *Let $\mu$ denote the acim from the Folklore Theorem, and let the vector $x^{(0)}$ and the matrices $P^{(n)}$ be defined with respect to a partition $\eta$. Then*

$$\mu(X_i) = \left(\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} x^{(0)}P^{(0)}P^{(1)}\dots P^{(k)}\right)_i$$

*for each $X_i \in \eta$.*

At first glance, it may not seem that writing down the Folklore Theorem as a matrix equation (Corollary 4.1.2) is necessarily useful. However, we show next

that each of the matrices $P^{(n)}$ is *uniformly close* to the matrix $P^{(1)}$. This is helpful because an explicit rate of mixing for powers of the matrix $P^{(1)}$ will allow control of the perturbative error incurred by replacing each $P^{(k)}$ by $P^{(1)}$. In this way, a bound on the distance between the vector $\mu$ and the invariant probability vector for $P^{(1)}$ (the Ulam approximation) will be deduced.

## Growth of the matrices $P^{(n)}$

Let $T$ satisfy (4.1)—(4.3) and let $\eta = \{X_i\}_{i=1}^N$ be a partition of $X$ which respects the range structure (mod 0). For the partition $\eta$, recall from Chapter 1 that

$$\delta \triangleq \delta(\eta) = \operatorname{diam}(\eta) = \max_{B \in \eta} \operatorname{diam}(B)$$

is called the *diameter* of the partition.

**Proposition 4.1** *Let $\delta = \delta(\eta)$ and let the matrices $P^{(n)}$ be as above. Then for each $n \geq 0$ and $i, j = 1 \ldots N$,*

$$e^{-\delta \frac{K\lambda}{\lambda - 1}} \leq \frac{P_{ij}^{(n)}}{P_{ij}^{(1)}} \leq e^{\delta \frac{K\lambda}{\lambda - 1}}.$$

*Proof:* Fix $i, j \in 1 \ldots N$ and $n > 0$. Let $A = X_i$ and $B = T^{-1}X_j \cap X_i$. Then

$$P_{ij}^{(1)} = \frac{m(B)}{m(A)} \quad \text{and} \quad P_{ij}^{(n)} = \frac{m\left(T^{-(n-1)}B\right)}{m\left(T^{-(n-1)}A\right)}.$$

Without loss of generality, assume that both $A$ and $B$ are connected (the following argument extends easily to the general case). Let $\{T_\alpha^{-(n-1)}\}$ be the collection of inverse branches of $T^{n-1}$ whose images intersect $A$, and let $\{C_\alpha\}$ be the corresponding subsets of monotonicity. Therefore,

$$P_{ij}^{(n)} = \frac{\sum_\alpha m\left(T_\alpha^{-(n-1)}B\right)}{\sum_\alpha m\left(T_\alpha^{-(n-1)}A\right)} = \frac{\sum_\alpha m\left(C_\alpha \cap T^{-(n-1)}B\right)}{\sum_\alpha m\left(C_\alpha \cap T^{-(n-1)}A\right)},$$

and in particular

$$\min_\alpha \frac{m\left(T_\alpha^{-(n-1)}B\right)}{m\left(T_\alpha^{-(n-1)}A\right)} \leq P_{ij}^{(n)} \leq \max_\alpha \frac{m\left(T_\alpha^{-(n-1)}B\right)}{m\left(T_\alpha^{-(n-1)}A\right)}. \tag{4.6}$$

Now, because $T$ has a finite range structure which is respected by $\eta$ (mod 0), it follows that for each $\alpha$,

$$A \cap T^{(n-1)}C_\alpha = A \pmod 0.$$

Furthermore, because each monotonicity branch is differentiable, there exist $x_\alpha \in T_\alpha^{-(n-1)}A$ and $y_\alpha \in T_\alpha^{-(n-1)}B$ such that

$$\left(\det D(T^{n-1})(x_\alpha)\right) m\left(T_\alpha^{-(n-1)}A\right) = m(A)$$

and

$$\left(\det D(T^{n-1})(y_\alpha)\right) m\left(T_\alpha^{-(n-1)}B\right) = m(B).$$

Therefore,

$$\frac{m\left(T_\alpha^{-(n-1)}B\right)}{m\left(T_\alpha^{-(n-1)}A\right)} = \frac{\det\left(D(T_\alpha^{-(n-1)})\right)\left(T^{n-1}(x_\alpha)\right)}{\det\left(D(T_\alpha^{-(n-1)})\right)\left(T^{n-1}(y_\alpha)\right)} \times \frac{m(B)}{m(A)}.$$

The proof is completed by estimating the distortion ratio in this inequality: Notice that for each $\alpha$, $T_\alpha^{n-1}(x_\alpha) \in A$ and $T_\alpha^{n-1}(y_\alpha) \in B \subset A$. Hence

$$\rho(T_\alpha^{n-1}(x_\alpha), T_\alpha^{n-1}(y_\alpha)) \leq \delta.$$

An easy calculation shows that $Lip(\log|\det D\,(T^{-n})|) \leq K\lambda/(\lambda-1)$ for each $n > 0$, so that

$$e^{-\delta\frac{K\lambda}{\lambda-1}} \leq \frac{\det\left(D(T_\alpha^{-(n-1)})\right)\left(T^{n-1}(x_\alpha)\right)}{\det\left(D(T_\alpha^{-(n-1)})\right)\left(T^{n-1}(y_\alpha)\right)} \leq e^{\delta\frac{K\lambda}{\lambda-1}}.$$

Therefore,

$$e^{-\delta\frac{K\lambda}{\lambda-1}} \times \frac{m(B)}{m(A)} \leq \frac{m\left(T_\alpha^{-(n-1)}B\right)}{m\left(T_\alpha^{-(n-1)}A\right)} \leq e^{\delta\frac{K\lambda}{\lambda-1}} \times \frac{m(B)}{m(A)}$$

and the proposition follows from (4.6). □

**Corollary 4.1.3** *Letting the partition $\eta$ and vectors $\{x^{(k)}\}$ be as above,*

$$\left\|x^{(k)} - x^{(k-1)}P^{(1)}\right\|_1 \leq \left(e^{\delta\frac{K\lambda}{\lambda-1}} - 1\right) = O(\delta).$$

*Proof:* By Lemma 4.1.1,

$$\left(x^{(k)} - x^{(k-1)}P^{(1)}\right)_j = \left(x^{(k-1)}\left(P^{(k)} - P^{(1)}\right)\right)_j = \sum_{i=1}^{n}(x^{(k-1)})_i P_{ij}^{(1)}\left(\frac{P_{ij}^{(k)}}{P_{ij}^{(1)}} - 1\right).$$

The result follows from the Proposition because $\|x^{(k-1)}P^{(1)}\|_1 = 1$. □

REMARK. In the case where the map $T$ is a piecewise linear Markov map, and the partition $\eta$ is a Markov partition such that $T$ is linear on each $A \in \eta$, Proposition 4.1 shows that each $P^{(n)} = P^{(1)}$: the Lipschitz continuity of $\log\det DT^{-n}_{\alpha^{(n)}}$

is used only to compare the value of $\det DT^{-n}_{\alpha^{(n)}}$ at points which lie within monotonicity components of $T^n$. Since $T^n$ is linear on all such components, one can take the constant $K = 0$. Therefore, $P^{(n)} = P^{(n-1)} = \ldots = P^{(1)}$. If the map $T$ is ergodic, then the matrix $P^{(1)}$ has a unique invariant probability vector $p$, and $x^{(0)}\left(P^{(1)}\right)^k \to p$ as $k \to \infty$. Therefore, Corollary 4.1.2 implies that $\mu = p$, and Ulam's method will reproduce the exact invariant measure (which is well–known). □

Now, by choosing the partition $\eta$ to be sufficiently fine ($\delta(\eta)$ sufficiently small), the exponential in Proposition 4.1 can be made arbitrarily close to 1, implying that each $P^{(n)}$ is a small perturbation of $P^{(1)}$. Since $\mu$ arises as the limit of a multiplicative sequence of of $P^{(n)}$s, the analysis of the error in approximating $\mu$ by Ulam's method[2] reduces to estimating the sensitivity to small perturbations of the asymptotic product.

REMARK. The problem of the asymptotic matrix product we describe is similar to a situation considered by Diamond et. al. [19]. They define an *interval stochastic matrix* to be the collection of all stochastic matrices $Q$ which satisfy $A_{ij} \leq Q_{ij} \leq B_{ij}$ for some fixed positive matrices $A$ and $B$, and characterise the set of invariant vectors for matrices $Q$ by a sequence of inequalities. Unfortunately, their results do not imply concrete bounds on the diameter of the set of probability vectors which arise as limits of products of matrices from the collection. □

## 4.2 Numerical calculation of rates of mixing

In Chapters 2 and 3, the key ingredients in obtaining analytic error bounds for Ulam's method were explicit bounds on the rate of mixing for the Ulam approximation and Perron–Frobenius operator (respectively) when restricted to a suitable *difference cone*. Since the current scenario involves dealing with stochastic matrices and their action on the positive cone in $\mathbb{R}^N$, we consider

$$\Gamma = \Gamma(\eta) = \left\{ x \in \mathbb{R}^N : \sum_{i=1}^{N} x_i = 0 \right\}. \tag{4.7}$$

$\Gamma$ is the *difference cone* corresponding to the cone of positive vectors in $\mathbb{R}^N$.

Our objective is to estimate

$$\left\| \left( Id - P^{(1)} \right)^{-1} \Big|_{\Gamma} \right\|_1 .$$

Assume that $P^{(1)}$ has a unique invariant probability vector. Then, the invariant subspace corresponding to the eigenvalue 1 is one–dimensional, and it is easy

[2] That is, by the invariant vector for $P^{(1)}$.

to see that $\Gamma$ is a complementary space, invariant under left multiplication by $P^{(1)}$. Moreover

$$z \in \Gamma \quad \Rightarrow z\left(P^{(1)}\right)^k \to 0$$

exponentially fast as $k \to \infty$.

Let $\{e_i\}_{i=1}^N$ denote the usual basis for $\mathbb{R}^N$ and for each $i = 1 \ldots N-1$ put

$$\gamma_i = e_i - e_N.$$

Then, $\{\gamma_i\}_{i=1}^{N-1}$ is a basis for $\Gamma$ because if $z \in \Gamma$,

$$z = \sum_{i=1}^{N} z_i e_i = \sum_{i=1}^{N-1} z_i \gamma_i + \left(z_N + \sum_{i=1}^{N-1} z_i\right) e_N = \sum_{i=1}^{N-1} z_i \gamma_i$$

since $\sum_{i=1}^N z_i = 0$.

**Proposition 4.2 (Bound on $(Id - P^{(1)})^{-1}$)** *With $P^{(1)}$ and $\{\gamma_i\}$ as above, put*

$$C = \max_{i=1\ldots N-1} \left\| \sum_{j=0}^{\infty} \gamma_i \left(P^{(1)}\right)^j \right\|_1 .$$

*Then $C < \infty$ and*

$$\frac{C}{2} \leq \left\| \left(Id - P^{(1)}\right)^{-1} \Big|_\Gamma \right\|_1 \leq C.$$

*Proof:* The fact that $C$ is finite follows from our comments before the statement of the proposition. It is a standard fact that

$$\left(Id + P^{(1)} + \left(P^{(1)}\right)^2 + \cdots\right)\left(Id - P^{(1)}\right) = Id\Big|_\Gamma,$$

so it only remains to demonstrate the bounds. Let $z \in \Gamma$. Then

$$\begin{aligned}
\left\| \sum_{j=0}^{\infty} z \left(P^{(1)}\right)^j \right\|_1 &= \left\| \sum_{j=0}^{\infty} \sum_{i=1}^{N-1} z_i \gamma_i \left(P^{(1)}\right)^j \right\|_1 \\
&\leq \sum_{i=1}^{N-1} |z_i| \left\| \sum_{j=0}^{\infty} \gamma_i \left(P^{(1)}\right)^j \right\|_1 \\
&\leq C \sum_{i=1}^{N-1} |z_i| \leq C\|z\|_1,
\end{aligned}$$

and the upper bound follows. For the other inequality, let $\gamma_{i^*}$ be such that

$$C = \left\| \sum_{j=0}^{\infty} \gamma_{i^*} \left(P^{(1)}\right)^j \right\|_1 .$$

Then $\gamma_{i^*} \in \Gamma$ and the lower bound follows because $\|\gamma_{i^*}\|_1 = 2$. □

The number $C$ in Proposition 4.2 is very amenable to computation. Because each vector $\gamma_i$ converges to zero exponentially fast under iterated multiplication by $P^{(1)}$, the calculation of each sum $\sum_j \gamma_i (P^{(1)})^j$ is numerically stable. However, when $N$ is large, there are many such sums to compute. These could easily be done in parallel.

REMARK. (RELATION WITH PREVIOUS CHAPTERS) In Chapters 2 and 3, much of the analytical work was devoted to obtaining *a priori* upper bounds for the rate of mixing of the Ulam approximations $\mathcal{P}_n$. Since each $\mathcal{P}_n$ may be completely represented by a stochastic matrix, it is clear that our previous estimates give bounds on the constants $C$: If $\eta_N$ is a partition of the interval $[0,1]$ into subintervals of length $1/N$ and $T$ is of the class considered in Chapter 2 or 3, then

$$C(\eta_N) \leq O(\log N).$$

Moreover, in Section 3.5 it was explained how mixing rates for the matrices $P$ can be used to bound the sample path variance for the *Monte–Carlo* estimation of functionals. If $\gamma_g^2$ denotes the asymptotic variance of a functional $g$, then

$$\gamma_g^2 \leq 2\, C(\eta)\, \|g\|_\infty\, p(g)$$

where $p(g)$ denotes the expected value of $|g|$ with respect to $p$, the invariant vector for $P^{(1)}$. □

REMARK. In investigating the convergence of Ulam's method, Froyland [24] has considered a similar object to the operator $(Id - P^{(1)})^{-1}|_\Gamma$: the *fundamental matrix* of a Markov chain which is "close" to the one governed by $P^{(1)}$. Under restricted circumstances, he found that the norm of the fundamental matrix scales like $O(\log N)$. □

## 4.3 Computer assisted error bounds

We now put together the ingredients from Sections 4.1 and 4.2 to obtain numerically assisted error bounds for Ulam's method.

### 4.3.1 Numerics based Theorem

Suppose that $T$ is a transformation satisfying (4.1)—(4.3), that $\eta$ is a partition which respects the range structure, and that $\delta(\eta)$ is the diameter of the partition. If $T$ is ergodic, and $P^{(1)}$ is the matrix representation of the Ulam approximate operator over $\eta$, then $P^{(1)}$ has a unique invariant probability vector $p$. Let $\mu$ denote the vector representation (over $\eta$) of the (unique) acim, and let $C = C(\eta)$ be the constant from Section 4.2.

**Theorem 4.1 (Numerical error bounds)** *For the setup described above,*

$$\|\mu - p\|_1 \leq C(\eta)\left(e^{\delta(\eta)\frac{K\lambda}{\lambda-1}} - 1\right),$$

*where* $\|\cdot\|_1$ *is the usual vector* 1*–norm on* $\mathbb{R}^N$.

*Proof:* Let the vectors $\{x^{(i)}\}_{i\geq 0}$ and the matrices $\{P^{(k)}\}_{k>0}$ be as defined in (4.4) and the subsequent definition. For simplicity of notation, we set $P \triangleq P^{(1)}$. Then

$$\begin{aligned} x^{(k)} - p &= \sum_{j=0}^{k-1}(x^{(k-j)}P^j - x^{(k-j-1)}P^{j+1}) + x^{(0)}P^k - p \\ &= \sum_{j=0}^{k-1}\left(x^{(k-j)} - x^{(k-j-1)}P\right)P^j + (x^{(0)} - p)P^k \end{aligned}$$

since $p$ is an invariant vector for $P^{(1)}$.

Next,

$$\begin{aligned} \left\|\frac{1}{n}\sum_{k=0}^{n-1} x^{(k)} - p\right\|_1 &= \left\|\frac{1}{n}\sum_{k=1}^{n-1}\sum_{j=0}^{k-1}\left(x^{(k-j)} - x^{(k-j-1)}P\right)P^j + \frac{1}{n}\sum_{k=0}^{n-1}(x^{(0)} - p)P^k\right\|_1 \\ &= \left\|\frac{1}{n}\sum_{k=1}^{n-1}\sum_{j=0}^{n-1-k}\left(x^{(k)} - x^{(k-1)}P\right)P^j + \frac{1}{n}\sum_{k=0}^{n-1}(x^{(0)} - p)P^k\right\|_1 \\ &\leq \frac{1}{n}\sum_{k=1}^{n-1}\left\|\sum_{j=0}^{n-1-k}(x^{(k)} - x^{(k-1)}P)P^j\right\|_1 + \frac{1}{n}\left\|\sum_{k=0}^{\infty}(x^{(0)} - p)P^k\right\|_1. \end{aligned}$$

Now, recall from (4.5) that

$$\mu = \lim_{n\to\infty}\frac{1}{n}\sum_{k=0}^{n-1} x^{(k)}.$$

Therefore

$$\begin{aligned} \|\mu - p\|_1 &= \lim_{n\to\infty}\left\|\frac{1}{n}\sum_{k=0}^{n-1} x^{(k)} - p\right\|_1 \\ &\leq \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}\left\|\sum_{j=0}^{n-1-k}(x^{(k)} - x^{(k-1)}P)P^j\right\|_1 \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}\left\|\sum_{j=0}^{\infty}(x^{(k)} - x^{(k-1)}P)P^j\right\|_1 \\ &= \lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n-1}\left\|(x^{(k)} - x^{(k-1)}P)(Id - P)^{-1}\right\|_1 \\ &\leq \left\|(Id - P)^{-1}\right\|_1 \sup_{n\geq 1}\|x^{(n)} - x^{(n-1)}P\|_1. \end{aligned}$$

The theorem now follows from Corollary 4.1.3 and Proposition 4.1. □

A short calculation shows that if $h$ is the exact invariant density for $T$, and $\|\cdot\|$ denotes the usual $L^1$ norm, then

$$\|h - \Pi_\eta h\| \leq \left(e^{\delta(\eta)\frac{K\lambda}{\lambda-1}} - 1\right).$$

But $p$ is the vector representation of the Ulam approximation $h_\eta$, and $\mu$ is the vector representation of $\Pi_\eta h$. Thus $\|\Pi_\eta h - h_\eta\| = \|\mu - p\|_1$ and Theorem 4.1 implies that

$$\|h - h_\eta\| \leq \|h - \Pi_\eta h\| + \|\Pi_\eta h - h_\eta\| \leq (1 + C(\eta)) \left(e^{\delta(\eta)\frac{K\lambda}{\lambda-1}} - 1\right).$$

REMARK. If $T$ is as in Chapters 2 or 3 and if $\eta_N$ is a partition of $[0,1]$ into $N$ equal subintervals then

$$C(\eta_N) \leq O(\log N).$$

Since $\delta(\eta_N) = 1/N$, Theorem 4.1 implies that

$$\|h - h_n\| \leq O(\log N/N),$$

recovering our earlier results. □

### 4.3.2 Two examples

We now apply the *numerical approach* to error bounds to the simple examples introduced in Chapter 2.

EXAMPLE 1: Recall the family $\{T_\epsilon\}_{\epsilon \geq 0}$ introduced in Chapter 2. Each is smoothly conjugate to the map $x \mapsto 5x \pmod 1$ and has a known invariant density. Somewhat arbitrarily, we pick $\epsilon = 0.08$ (this is the "most nonlinear" of the maps considered in Chapter 2) and compare our analytically obtained error bounds with the numerically assisted approach described in this chapter.

There are four sets of data. Each set consists of error bounds for Ulam's method based on a partition of $[0,1]$ into $N = 10, 20, 40, \ldots 2^{10} \times 10$ equal subintervals (the corresponding partitions are denoted by $\eta_N$).

1. *Method from Chapter 2.* Using the expansivity and distortion constants from Chapter 2, an optimal choice of $k = k(N)$ is made to minimise $e_N(k)$, the first factor of the product in the conclusion of Theorem 2.2. Denoting the optimal such constant by $C'_N$,

   $$C(\eta_N) + 1 \leq C'_N,$$

   and an error bound follows.

| $N$ | Chapter 2 | | Chapter 3 | | Chapter 4 | | Error |
|---|---|---|---|---|---|---|---|
| | $C'_N$ | Bound | $C''_N$ | Bound | $C(\eta_N)$ | Bound | |
| 10 | $O(10^3)$ | $O(10^2)$ | 19.2 | $O(1)$ | 4.17 | $O(1)$ | $1.32 \times 10^{-2}$ |
| 20 | 598 | $O(10)$ | 20.1 | $O(1)$ | 4.51 | $8.71 \times 10^{-1}$ | $7.33 \times 10^{-3}$ |
| 40 | 413 | $O(10)$ | 20.2 | $9.79 \times 10^{-1}$ | 5.54 | $4.23 \times 10^{-1}$ | $3.11 \times 10^{-3}$ |
| 80 | 345 | $O(10)$ | 21.1 | $5.09 \times 10^{-1}$ | 6.54 | $2.83 \times 10^{-1}$ | $1.59 \times 10^{-3}$ |
| 160 | 318 | $O(1)$ | 21.2 | $2.57 \times 10^{-1}$ | 7.09 | $1.50 \times 10^{-1}$ | $8.01 \times 10^{-4}$ |
| 320 | 312 | $O(1)$ | 22.1 | $1.34 \times 10^{-1}$ | 8.38 | $8.66 \times 10^{-2}$ | $4.12 \times 10^{-4}$ |
| 640 | 320 | $O(1)$ | 22.9 | $6.95 \times 10^{-2}$ | 8.52 | $4.38 \times 10^{-2}$ | $2.10 \times 10^{-4}$ |
| 1280 | 334 | 0.768 | 23.1 | $3.49 \times 10^{-2}$ | 10.2 | $2.57 \times 10^{-2}$ | $1.07 \times 10^{-4}$ |
| 2560 | 350 | 0.403 | 23.9 | $1.82 \times 10^{-2}$ | 10.5 | $1.32 \times 10^{-2}$ | $5.25 \times 10^{-5}$ |
| 5120 | 368 | 0.212 | 24.1 | $9.09 \times 10^{-3}$ | 12.1 | $7.50 \times 10^{-3}$ | $2.65 \times 10^{-5}$ |
| 10240 | 387 | 0.112 | 25.0 | $4.72 \times 10^{-3}$ | 12.7 | $3.92 \times 10^{-3}$ | $1.32 \times 10^{-5}$ |

Table 4.1: Comparison of various methods of obtaining error bounds for Ulam's approximation applied to the map $T_\epsilon$ when $\epsilon = 0.08$. The calculations and the constants $C'_N, C''_N, C(\eta_N)$ are explained in the text.

2. *Method from Chapter 3.* By differentiation, bounds on the constants $s, \lambda, D$ corresponding to $T$ can be obtained. Then, exploiting the constructions in the proof of Theorem 3.2, a bound on the norm of

$$(Id - \mathcal{L})^{-1} : (BV_0, \|\cdot\|_{4N}) \to (BV_0, \|\cdot\|)$$

follows by an optimal choice of constants (c.f. Appendix B). Denoting this number by $C''_N$, Theorem 3.2 implies that

$$\|h - h_N\| \leq C''_N \frac{s\lambda}{\lambda - 1} \frac{1}{N}.$$

3. *Numerical method.* The constants $K, \lambda$ required by Theorem 4.1 are precisely the constants $s, \lambda$ from Chapter 3. Therefore, to get a quantitative error bound out of Theorem 4.1, it suffices to compute the numbers $C(\eta_N)$. The calculations have been performed as described in Proposition 4.2, and Theorem 4.1 implies the corresponding error bound for Ulam's method:

$$\|h - h_N\| \leq (C(\eta_N) + 1) \left( e^{\delta(\eta) \frac{K\lambda}{\lambda - 1}} - 1 \right).$$

4. *Exact error.* Because the exact invariant density is known (by construction), the three methods of bounding error can be compared with the true approximation error.

The results of the various calculations are presented in Table 4.1. The error bounds are calculated as described above. For the methods from Chapters 2 and 3 this consists in choosing appropriate constants with which to apply the relevant Theorems, whereas the numerically assisted approach can require lengthy calculations:

Each $C(\eta_N)$ is obtained by calculating $\|\gamma_i(Id+P+P^2+\ldots)\|$ for each $\gamma_i$ in a basis for the collection of zero–sum vectors in $\mathbb{R}^N$. Because of exponential mixing, these sums converge very quickly (in practice, we choose some tolerance—$10^{-7}$ in the data depicted above—and sum until $\|\gamma_i P^n\|$ is smaller than the chosen tolerance; the errors thus incurred are not significant, and are easy to account for). Typically, this convergence requires $O(\log N)$ iterated multiplications by $P$. Since the transformation $T$ has bounded slope, the number of non–zero entries in each row of $P$ is bounded (independently of $N$). Therefore, each iterated multiplication of a vector by $P$ takes $O(N)$ arithmetic operations. Combining these observations, it follows that calculating each constant $C(\eta_N)$ requires $O(N^2 \log N)$ arithmetic operations.

Turning to the results depicted in Table 4.1, notice that use of the numerically obtained constants $C(\eta)$ consistently out-performs the analytic methods. However, comparison with the final column shows that even the numerical method does relatively badly at *estimating* the approximation error.

Finally, it is clear that the BV based method produces tighter error bounds than the method from Chapter 2. This phenomenon is visible in a much more drastic fashion in the next example, and an explanation is given there. □

EXAMPLE 2. Recall the family of expanding maps

$$T_\Omega : x \mapsto 2x + \frac{\sin 2\pi x}{4\pi} + \Omega \pmod 1$$

introduced in Chapter 2. As above, the various techniques for obtaining error bounds for a sequence of Ulam approximations have been applied. The error bounds from Chapters 2 and 3 are independent of the value of $\Omega$; for the explicit numerical calculations, a fixed value of $\Omega$ must be chosen—I made the arbitrary choice $\Omega = 0.5$. For the BV and numerical methods, the values of $s, \lambda$ from Chapter 3 are used. The necessity of having $\lambda > 2$ for the technique from Chapter 2 means that $T^2$ must be used. The results are presented in Table 4.2.

The most obvious feature of the data for this example is that the method from Chapter 2 gives meaningless error bounds (the maximal distance between any two densities is 2). This phenomenon is relatively easy to explain: the constant $C'_N$ (a bound on $\|(Id - P(N))^{-1}\|$) has an exponential dependence on the distortion $K$ (see Corollary 2.3.4 and Theorem 2.1). Moreover, the $K$ corresponding to the $k$th iterate of an expanding map $T$ can grow *exponentially* in $k$. Because the method from Chapter 2 requires $\lambda > 2$, we are forced to use $T^2_\Omega$, thereby increasing $K$ and leading to a uselessly large value of $C'_N$ for moderate sized $N$ (of course, $C'_N$ is *guaranteed* by Theorem 2.2 to be no worse than $O(\log N)$, so that the asymptotic rate still holds).

By contrast, the distortion constant $s$ used in the method from Chapter 3 is uniformly bounded by $s\lambda/(\lambda - 1)$ if $T$ is replaced by any iterate. This makes the BV error bounds relatively stable.

For this class of transformations, no expression for the invariant density is known, so no comparison with the actual approximation error is possible. □

| $N$ | Chapter 2 | | Chapter 3 | | Chapter 4 | |
|---|---|---|---|---|---|---|
| | $C'_N$ | Bound | $C''_N$ | Bound | $C(\eta_N)$ | Bound |
| 10 | $O(10^9)$ | $O(10^9)$ | 128 | $O(10)$ | 5.45 | $O(1)$ |
| 20 | $O(10^7)$ | $O(10^7)$ | 130 | $O(10)$ | 7.34 | $9.80 \times 10^{-1}$ |
| 40 | $O(10^6)$ | $O(10^5)$ | 132 | $O(1)$ | 8.79 | $5.87 \times 10^{-1}$ |
| 80 | $O(10^5)$ | $O(10^3)$ | 133 | $O(1)$ | 10.5 | $3.50 \times 10^{-1}$ |
| 160 | $O(10^4)$ | $O(10^3)$ | 135 | $O(1)$ | 12.3 | $2.05 \times 10^{-1}$ |
| 320 | $O(10^4)$ | $O(10^2)$ | 137 | $O(1)$ | 14.0 | $1.17 \times 10^{-1}$ |
| 640 | $O(10^4)$ | $O(10^2)$ | 138 | $5.76 \times 10^{-1}$ | 15.8 | $6.58 \times 10^{-2}$ |
| 1280 | $O(10^4)$ | $O(10)$ | 140 | $2.92 \times 10^{-1}$ | 17.5 | $3.65 \times 10^{-2}$ |
| 2560 | $O(10^4)$ | $O(10)$ | 142 | $1.48 \times 10^{-1}$ | 19.2 | $2.00 \times 10^{-2}$ |
| 5120 | $O(10^4)$ | $O(1)$ | 143 | $7.46 \times 10^{-2}$ | 20.8 | $1.09 \times 10^{-2}$ |
| 10240 | $O(10^4)$ | $O(1)$ | 145 | $3.78 \times 10^{-2}$ | 22.4 | $5.84 \times 10^{-3}$ |

Table 4.2: Comparison of various methods of obtaining error bounds for Ulam's approximation applied to the map $T_\Omega$ when $\Omega = 0.5$. The calculations and the constants $C'_N, C''_N, C(\eta_N)$ are as in the previous example. The notation $O(10^k)$ implies that the actual value is between $10^k$ and $10^{k+1}$.

From these two simple one–dimensional examples, we see that the analytical techniques we have collected for bounding the approximation error in Ulam's method are rather sensitive to the distortion of the transformation. By contrast, the *numerically assisted* approach appears to do consistently better than the analytical techniques. While the first example suggests that all techniques produce rather poor *estimates* of the approximation error, the rigorous nature of our arguments *guarantees* that arbitrarily good approximations can be obtained in principle. In practice, the best bounds will come from the numerical method. With this in mind, we do not concentrate on optimising the constants in the analytical techniques for the remainder of the thesis.

### 4.3.3 If the density is sufficiently regular....

The possibilities for numerical error bounding are much greater than those described above. In this brief subsection, we show how the numerical procedure suggested by Proposition 4.2 can yield error bounds for Ulam's method in much more generality.

**Assumption:** *Let $X$ be any manifold of finite diameter, and suppose that $T : X \to X$ is any piecewise $\mathcal{C}^1$ transformation which has an absolutely continuous invariant probability measure whose density $h$ has* bounded variation[3]. *Let*

$$V_X(h) \leq A. \quad \square$$

[3]For an appropriate interpretation of this statement in multi–dimensions, see the next chapter.

Suppose that a finite partition $\eta$ of $X$ is given, and that $\delta(\eta)$ is known. Let $\Pi_\eta$ denote the projection onto densities which are piecewise constant on elements of $\eta$, and let

$$\mathcal{P}_\eta = \Pi_\eta \circ \mathcal{L}$$

be the corresponding Ulam approximation. Suppose that $\mathcal{P}_\eta$ has a unique eigenvalue of modulus 1, and denote the corresponding normalised eigenvector—the Ulam approximate density—by $h_\eta$. Let $P_\eta$ denote the stochastic matrix representation of $\mathcal{P}_\eta$. Then $P_\eta$ has a unique eigenvalue of modulus 1, and the constant $C(\eta)$ defined in Proposition 4.2 is finite.

**Proposition 4.3 (Numerical error bounds revisited)** *For $T, \eta$ as above,*

$$\|h - h_\eta\| \leq (C(\eta) + 1)\, A\delta(\eta).$$

*Proof:* First of all, recall that the finite dimensional operator $\mathcal{P}_\eta$ is exactly represented by the matrix $P_\eta$. Moreover, if $\eta$ consists of $N$ sets, then the natural isomorphism between $L^1 \cap \mathcal{D}_\eta(X)$ and $\mathbb{R}^N$ equipped with the vector 1–norm is an isometry. In particular, it follows from Proposition 4.2 that

$$\frac{C(\eta)}{2} \leq \left\| (Id - \mathcal{P}_\eta)^{-1} \Big|_{\mathcal{D}_\eta(X) \cap \{f \in L^1 : \int_X f = 0\}} \right\| \leq C(\eta).$$

Since $h_\eta$ is the fixed point of $\mathcal{P}_\eta$ and

$$(Id - \mathcal{P}_\eta)(h_\eta - \Pi_\eta h) = \mathcal{P}_\eta \circ \Pi_\eta h - \Pi_\eta h,$$

we have

$$\|h_\eta - \Pi_\eta h\| \leq C(\eta) \|\mathcal{P}_\eta \circ \Pi_\eta h - \Pi_\eta h\|.$$

But $h = \mathcal{L}h$ and $\mathcal{P}_\eta = \Pi_\eta \circ \mathcal{L}$ is a Markov operator, so

$$\|\mathcal{P}_\eta \circ \Pi_\eta h - \Pi_\eta h\| = \|\mathcal{P}_\eta(\Pi_\eta h - h)\| \leq \|(\Pi_\eta - Id)h\|.$$

By Lemma 3.2.1,

$$\|(\Pi_\eta - Id)h\| \leq \delta(\eta) \times A$$

and the proposition follows. □

There is no guarantee that Proposition 4.3 will provide useful error bounds for a given transformation. However, the above examples suggest that this numerically oriented approach will do at least as well as the analytic ones described previously, and may still yield some information in the cases where the theoretical constructions do not apply.

Yet Proposition 4.3 illustrates more: The uniform expansion of a map—or whatever other conditions are used to get quantitative control of mixing rates—is not essential to Ulam's method. The central feature is more that the map admits a relatively regular invariant density (moderate value of $A$): whether

this is accomplished by uniform expansion, or by a mixture of very strong expansion and localised contraction is irrelevant so long as the map mixes up fairly uniformly in the long term (moderate value of $C(\eta)$). While formalised results in this direction are not well developed, the numerical results presented in the remainder of the chapter provide good cause for optimism.

## 4.4 Example: smooth unimodal maps

We now apply our techniques to some more maps of theoretical interest where the invariant densities are unknown.

Recall the families of smooth unimodal maps discussed in the introduction. Let $X = [0, 1]$, and for $l \in [2, \infty)$ we have

$$S_l(x) = 1 - |2x - 1|^l.$$

These maps are the simplest representatives of a class of negative Schwarzian [16] unimodal maps studied extensively by van Strien and others [57, 56, 16]. In these references, it is shown that subject to a summability condition on the critical orbit (the orbit of the unique critical point of the map), there exists a unique ergodic absolutely continuous invariant probability measure. Moreover, the corresponding distributions have singularities of order $\frac{1}{l}$ along the critical orbit. For the family $\{S_l\}_{l \geq 2}$, the critical point (at $\frac{1}{2}$) is pre-periodic, landing on the fixed point at 0 in two iterations. Therefore, the singularities of the density are at 0 and 1. (For $l = 2$, the map $S_2$ is the fully developed logistic map $x \mapsto 4x(1 - x)$, which has $\pi^{-1}(x(1 - x))^{-1/2}$ as an invariant density. For all other values of $l$, the invariant density is unknown.)

Now, the results we have obtained so far for Ulam's method are applicable only to maps with fairly uniform expansion, and the $l$th order critical point at $\frac{1}{2}$ implies that the maps $S_l$ are *not* uniformly expanding. However, each $S_l$ has some expansive structure [16], and a little experimentation shows that this may be exhibited by working with a *conjugate* family:

By way of motivation, recall that the maps $T_\epsilon$ considered above were contrived to have a given invariant density by using the distribution of the desired density to conjugate with the uniform expansion $x \mapsto 5x \pmod 1$. Here, we try to reverse this argument: That is, to guess a function which has qualitative similarity to the distribution of the invariant measure, and use this to conjugate the maps $S_l$. We hope that the resulting collection of maps are uniformly expanding, or have relatively regular invariant densities (at the very least). Because the invariant distribution for $S_l$ has singularities of order $1/l$ along the critical orbit [56], any conjugacy which is to "straighten out" the contraction caused at critical point at $\frac{1}{2}$ must also have singularities of order $1/l$ along the critical orbit. With this in mind, put

$$h_l(x) = \frac{1 + x^{1/l} - (1 - x)^{1/l}}{2},$$

and

$$T_l = h_l \circ S_l \circ h_l^{-1}.$$

Then, if $\psi_l$ is an invariant density for $T_l$,

$$\phi_l = h_l' \times \psi_l \circ h_l$$

is an invariant density for $S_l$. Moreover, if $f$ is *any* density then

$$\|h_l' \times f \circ h_l - \phi_l\| = \|f - \psi_l\|.$$

Hence, for any error bounded approximate invariant density for $T_l$, the *pull–back* by $h_l$ is an approximate invariant density for $S_l$, *with the same error bound*. We exploit this fact below.

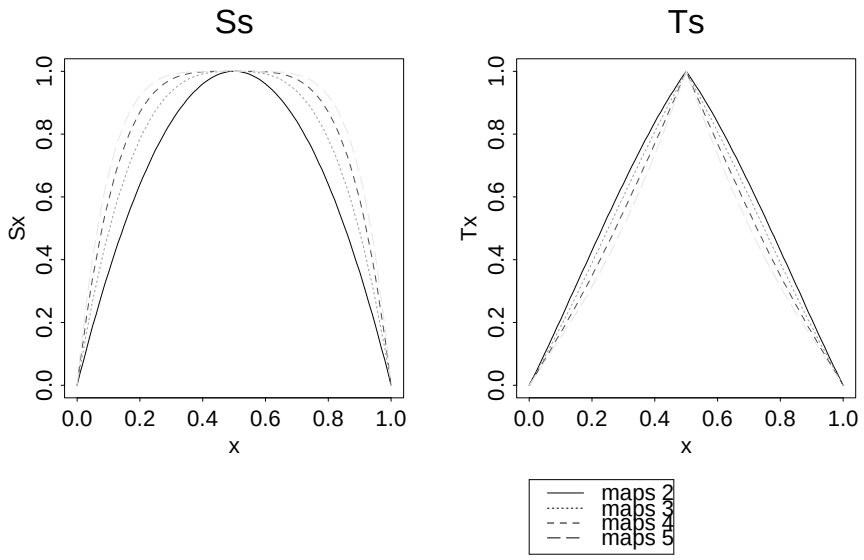



Figure 4.1: The maps $S_l$ and their conjugates $T_l$, for $l = 2, 3, 4, 5$.

Indeed, the maps $T_l$ turn out to be uniformly expanding, and for selected values of $l$, the straightening affected by the conjugacy is illustrated in Figure 4.1. The constants $\lambda_l$, $K_l$ (the minimal modulus derivative and Lipschitz constant of $\log|(T_l^{-1})'|$ respectively) can be calculated by a fine numerical search.

For each value of $l$, $[0, 1]$ was partitioned into $N = 10, 20, \ldots, 2^8 \times 10$ equal subintervals, and the Ulam approximate density for $T_l$ was computed. For each choice of $l$ and partition $\eta_N$, the constant $C(l, N) = C(\eta_N)$ from Proposition 4.2 was calculated. As in the previous examples, $C(l, N) = O(\log N)$, (these data are not displayed). For each $l$, let $f_l$ be the density arising from Ulam's method

with $N = 2^8 \times 10$, and $\psi_l$ be the invariant density for $T_l$. By Theorem 4.1,

$$\|f_l - \psi_l\| \leq (C(l,N)+1)\left(e^{\frac{K_l\lambda_l}{\lambda_l - 1}\frac{1}{N}} - 1\right).$$

These error data are presented in Table 4.3.

| $l$ | $\lambda$ | $K$ | $C(l,2560)$ | $(C(l,2560)+1)\frac{K\lambda}{\lambda-1}\frac{1}{2560}$ |
|---|---|---|---|---|
| 2.0 | 1.414 | 2.001 | 22.19 | 0.0591 |
| 3.0 | 1.817 | 0.512 | 23.97 | 0.0107 |
| 4.0 | 1.682 | 0.640 | 26.25 | 0.0162 |
| 5.0 | 1.583 | 0.911 | 26.61 | 0.0257 |

Table 4.3: Data from Ulam approximations for the maps $T_l$ with $N = 2560$. The last column provides upper bounds on $\|f_l - \psi_l\|$.

The $N = 2560$ Ulam approximations for the maps $T_l$ are depicted in Figure 4.2 below. They are drawn as histograms over 100 equal bins. The right–hand column of Table 4.3 gives bounds on the $L^1$ distances from the depicted densities to the exact invariant densities for the maps $T_l$. As remarked above, the Ulam approximations can be pulled back by the conjugacies $h_l$ to give approximate invariant densities for the corresponding maps $S_l$. These are depicted in Figure 4.3.

We conclude with several observations.

First of all, our theoretical techniques for obtaining error bounds require at least some regularity of the invariant density. In this case, qualitative knowledge of the degeneracies of the invariant densities allowed us to make a reasonable guess of conjugacies $h_l$. Of course, this method is rather ad hoc (we may not have been able to find a suitable family of conjugacies), yet it worked in this situation, providing rigorous quantitative approximations to a previously unknown collection of invariant densities. In the next section, a procedure for automating this "conjugacy guessing" approach is presented.

Given that conjugating the family $\{S_l\}_{l\geq 2}$ was motivated by the possibility of transforming onto a uniformly expanding linear map, one natural question is whether *in principle* it is possible to guess a conjugacy which would transform $S_l$ onto $S_1$, the fully developed tent map. Our numerical results can be used to prove that unless $l = 2$, no such absolutely continuous conjugacy can exist. We illustrate the method with $l = 4$:

*Non-existence of a.c. conjugacy:* Suppose that $h$ is an absolutely continuous function such that

$$h \circ S_4 \circ h^{-1} = S_1$$

(the fully developed tent map). Because $\phi_1 \triangleq 1$ is the *unique* invariant density for $S_1$, $S_4$ has invariant density $h'$. Evaluating the almost everywhere Lyapounov exponent:

$$\lambda(S_4) = \int_0^1 \log|S_4'|\, h' dm(x) = \int_0^1 \log|S_1'|\, dm(x) = \log 2$$

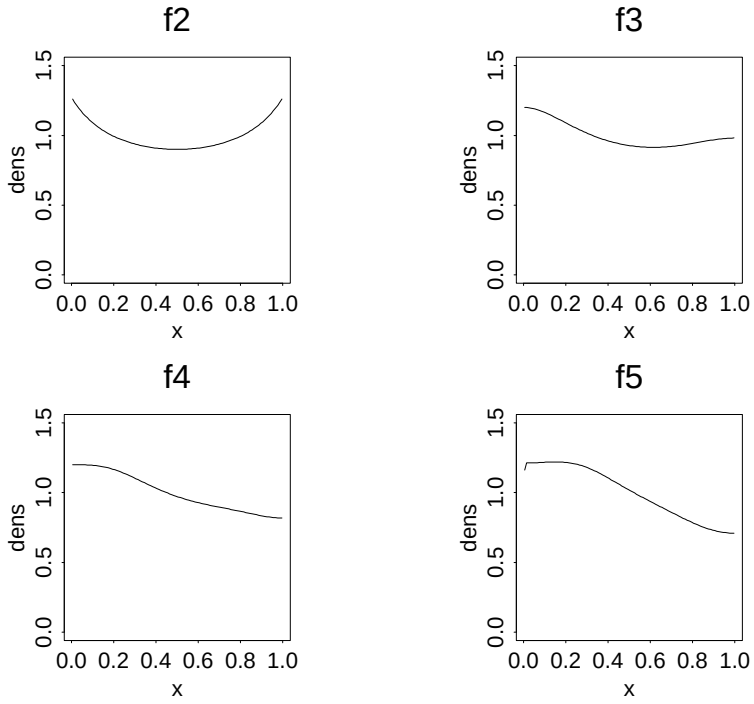



Figure 4.2: Approximate invariant densities for the maps $T_l$. The depicted histograms are Ulam approximations over $10 \times 2^8$ equal subintervals.

by conjugacy invariance. Also by conjugacy invariance, we must have that $\lambda(T_4) = \lambda(S_4)$. However, numerical evaluation of the Lyapounov exponent for $T_4$ via the Ulam approximation over 20480 equal subintervals produces an estimate

$$\lambda_{estimate} = \sum_{i=1}^{20480} \frac{f_4(x_i)}{20480} \log |T_4'(x_i)| = 0.684494$$

where $f_4$ is the Ulam approximate density and $x_i = (i - 0.5)/20480$. Applying Proposition F.3 from Appendix F

$$|\lambda(T_4) - \lambda_{estimate}| \leq Lip(\log |T_4'|)(\|f_4 - \psi_4\| + 1/20480),$$

where $\psi_4$ is the invariant density for $T_4$. Applying the $BV$ error bounds (see Appendix B), one obtains

$$\|f_4 - \psi_4\| \leq 0.003671.$$

Then[4]

$$\lambda_(T_4) \leq \lambda_{estimate} + 0.004475 = 0.688969 < \log 2 = \lambda(S_4),$$

a contradiction. Therefore, there is no absolutely continuous conjugacy between $S_4$ and the tent map, in spite of the fact that they are topologically conjugate. □

[4]The value of $Lip(T_4')$ is also obtained numerically [53].

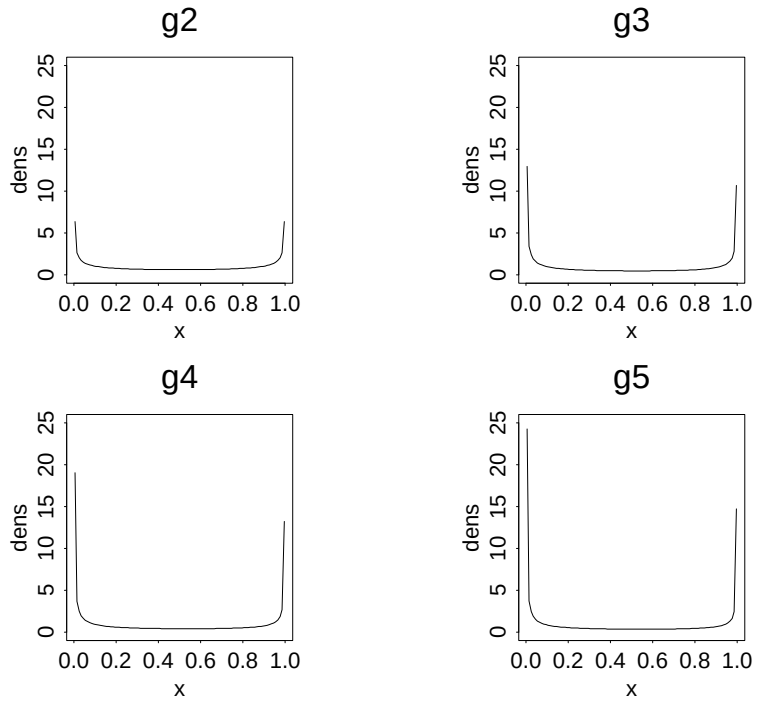



Figure 4.3: Approximate invariant densities for the maps $T_l$. The depicted histograms are the pull–backs by $h_l$ of the densities depicted in Figure 4.2.

Next, notice that the density approximations for the maps $S_l$ which are displayed in Figure 4.3 have a very noticeable asymmetry. This phenomenon is easily explained: Because the invariant measure has singularities of order $1/l$ near 0 and 1, we may write

$$\phi_l(x) \approx \begin{cases} a_l\, x^{1/l-1} & x \text{ near } 0, \\ b_l\, (1-x)^{1/l-1} & x \text{ near } 1, \end{cases}$$

for positive constants $a_l, b_l$. Substituting this approximation into the Perron–Frobenius operator equation near $x = 0$, one obtains the estimate

$$b_l \approx a_l \left( (2l)^{1/l} - 1 \right),$$

in good agreement with Figure 4.3.

Finally, consider what might happen if Ulam's method is applied directly to the maps $S_l$. Suppose that $N$ and $l$ are fixed, and let the matrix representation of the Ulam approximate operator over $N$ equal subintervals be denoted by $P = (P_{ij})$. If $N$ is sufficiently large,

$$P_{i1} \approx \begin{cases} 1/2l & \text{if } i = 1 \text{ or } N, \\ 0 & \text{otherwise.} \end{cases}$$

Let $p$ be an invariant vector for $P$. Since $pP = p$,

$$p_1 = p_1 P_{11} + p_N P_{N1} \approx (p_1 + p_N)/2l,$$

implying that

$$p_N \approx p_1(2l - 1). \tag{4.8}$$

However, comparison with the above estimates of the invariant density suggests that

$$\mu_N \approx \mu_1 \left( (2l)^{1/l} - 1 \right)$$

(where $(\mu_i)$ is the restriction of the invariant measure $\mu$ for $S_l$ to the partition into $N$ equal subintervals). Similar comparisons can be made for other components of the Ulam approximation. Therefore, the uniformity of the invariant density seems fundamental to the reliability of Ulam's method over *uniform partitions*.

## 4.5 Non–uniform partitions choice and an Adapted Ulam method

In Chapters 2 and 3 it was possible to exploit the uniform expansivity of the transformations considered to prove relatively uniform bounds on the *rate of mixing* of the corresponding Perron–Frobenius operators and Ulam approximations. In effect, these estimates measure the sensitivity of the invariant measures to perturbations of the underlying operator. For relatively uniform maps, we were able to show that the Ulam approximations over *uniform* partitions can be regarded as small perturbations of the Perron–Frobenius operator, so the mixing results implied error bounds for Ulam's method. In this section, we present numerical results which suggest that some of this structure can be imparted to non–uniform examples; the argument is rather informal.

### 4.5.1 Non–equal intervals

So far, we have always used *uniform*[5] *partitions* for analysis of Ulam's method. The examples above consistently show that the use of uniform partitions gives accurate invariant measure approximations for uniformly expanding maps. In the section above, we were able to apply this to some non–expanding maps by conjugating to an expanding family.

Now, a conjugacy corresponds to a change of variables—or "rescaling"—of the map, and it is reasonable to wonder whether similar results can be achieved by working with the original map, but a partition which is "scaled" to be more sensitive to the global dynamics.

Suppose that $S$ is a map of interest, and it is proposed to use some conjugacy $h$ to investigate

$$h \circ S \circ h^{-1}$$

[5]A partition $\eta = \{X_1, \ldots, X_N\}$ will be called *uniform* if $m(X_i) = 1/N$ for each $X_i \in \eta$.

via Ulam's method over a uniform partition. Denote the uniform partition by $\eta = \{X_1, \ldots, X_N\}$. An alternative approach is to apply Ulam's method directly to $S$ with the partition

$$\eta' = h^{-1}\eta = \{h^{-1}(X_1), \ldots, h^{-1}(X_N)\} \triangleq \{X'_1, \ldots, X'_N\}.$$

Suppose that $\mu$ is an invariant measure for $h \circ S \circ h^{-1}$ which is "close" to Lebesgue in the sense that $\mu(X_i) \approx m(X_i) = 1/N$. Then $\mu \circ h$ is an invariant measure for $S$, and $\mu \circ h(X'_i) = \mu \circ h \circ h^{-1}(X_i) = \mu(X_i) \approx 1/N$. Therefore, our *heuristic* aim is to find a conjugacy $h$ so that Ulam's method applied to $S$ over the partition $\eta'$ gives an *invariant probability vector* which is approximately uniform.

Assuming that $X = [0, 1]$, we propose:

**Algorithm: *Adapted Ulam method***

1. Fix a uniform partition $\eta = \{X_1, \ldots, X_N\}$.
2. Apply Ulam's method to $S$ using the given partition $\eta$; Let $p \in \mathbb{R}^N$ denote the invariant probability vector.
3. If $p$ is sufficiently close to uniform then **stop**; otherwise **continue**.
4. Put
$$h(x) = \int_0^x \sum_{X_i \in \eta} \frac{p_i}{m(X_i)} \chi_{X_i}\, dm.$$
5. Let $\eta' = h^{-1}\eta$, and return to Step 2 with $\eta := \eta'$.

Before proceeding to an example, note that the algorithm is not necessarily well defined: I have no proof that the sequence of vectors $p$ will converge to a uniform vector. Even if convergence occurs, there is no guarantee that the resulting measure approximation for $S$ will be close to invariant. However, the scheme is relatively easy to implement, and performs well in practice. Most importantly, the initial application of Ulam's method provides an estimate for the initial choice of conjugacy, removing the need to guess!

### 4.5.2 Example: the logistic family

Recall the family of logistic maps

$$x \mapsto cx(1-x), \qquad 0 < c \leq 4.$$

For $c = 4.0$, the invariant density is known to be $\pi^{-1}(x(1-x))^{-1/2}$. Numerical results are shown in Figure 4.4.

For each iteration of the algorithm, the results are depicted in three plots. The left-hand plot displays the pattern of the entries in the Ulam stochastic matrix.

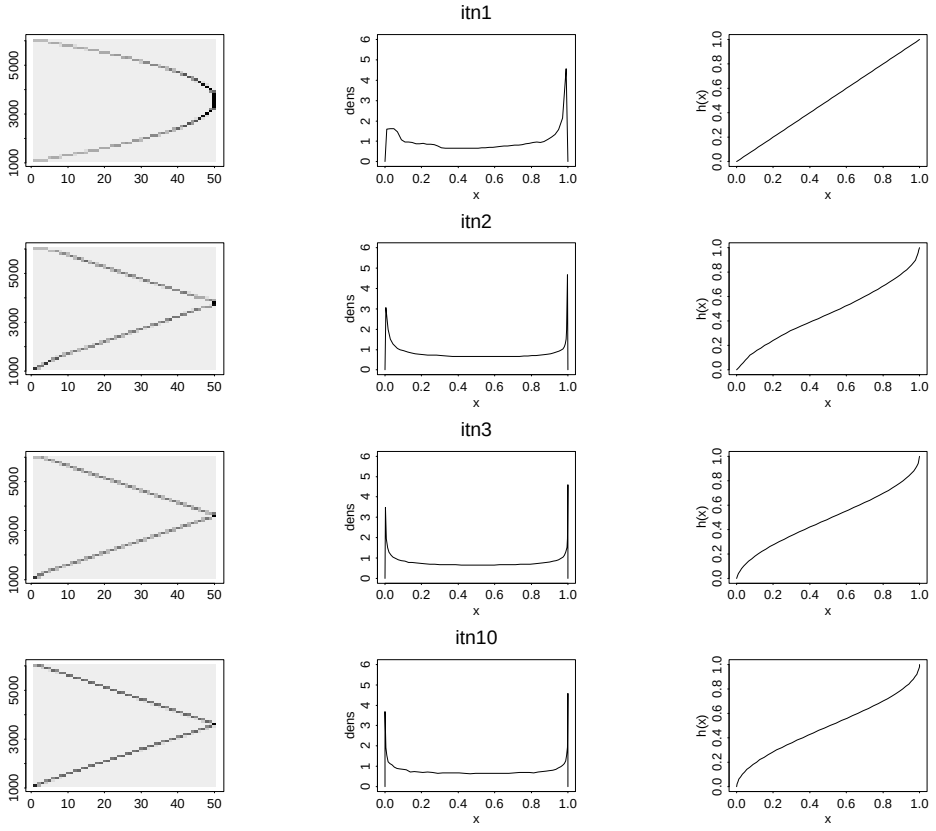



Figure 4.4: Sequence of $50 \times 50$ Ulam matrices, density approximations and dynamically chosen conjugacies for the logistic map with $c = 4.0$.

The lightest grey denotes 0, the darkest represents a transition probability very close to 1. For ease of illustration, all computations used a partition of 50 subintervals. Notice in particular how the shape of the graph of the original map is reproduced in the top left plot. The right–hand column shows the map $h$ which is used to transform the 50 subinterval uniform partition into the partition over which the adjacent Ulam matrix is constructed[6]. The second column depicts the approximate density found by the adapted Ulam method: the invariant vector for the Ulam matrix is drawn as a histogram over the appropriate partition.

By examining the sequence in the left–hand column, it is clear that automated partition choice affects a gradual straightening of the contraction near the critical point of $S$ (remember that the indices for the matrix elements correspond to the index of the subinterval, rather than Euclidean distance along $[0, 1]$). Also, the density approximations appear to be converging to the correct *shape* for the

[6]Explicitly: the pre–images of the points $(i/50)_{i=0}^{50}$ under the map on the right are the boundary points of the partition used for the corresponding step of the Adapted Ulam method.

invariant density. Note how the erroneous asymmetry in the Ulam approximate densities predicted in (4.8) is particularly evident at the first iteration of the method. The final data in the sequence are from the 10th iteration. Unfortunately, the asymmetry mentioned above persists in this method (it is easy to check that it must), but the overall shape of the approximate densities is much more satisfactory than at the initial stage.

Another interesting question concerns the performance of the Adapted method when no invariant density exists. For $c = 3.82$, the logistic map has a stable period three orbit, with a very large basin of attraction. The results of applying the Adapted Ulam method are presented in Figure 4.5.

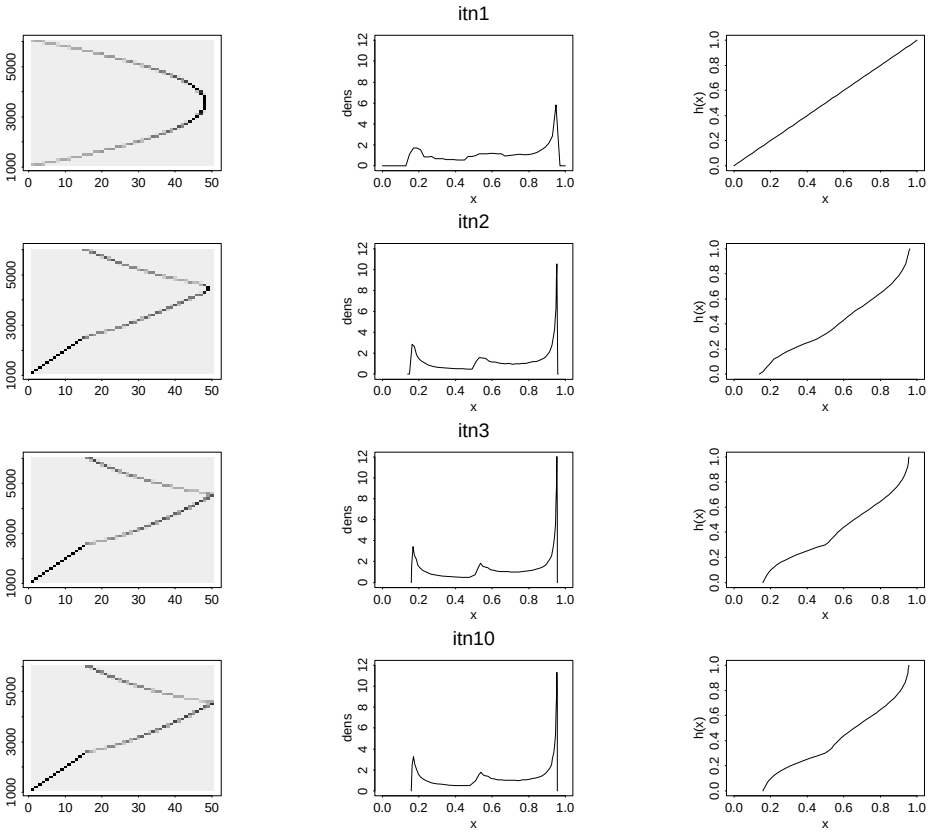



Figure 4.5: Sequence of $50 \times 50$ Ulam matrices, density approximations and dynamically chosen conjugacies for the logistic map with $c = 3.82$.

We can make the same qualitative observations as above about the second picture: the measure approximations appear to be "converging" under iteration of the scheme, and the dynamical partition choice leads to gradually "straighter" versions of the map.

In this case, the exact invariant measure is a sum of point masses located on the period three orbit, and the density approximations are strongly peaked

around the periodic orbit. However, there is one spurious effect: the relatively uniform concentration of mass between the points of the periodic orbit. This is an unavoidable consequence of the geometry of the scheme: this logistic map is not conjugate to a uniformly expanding map, but our scheme implicitly attempts to construct such a conjugacy, thereby "dragging" some of the mass that should be concentrated on each periodic point across the intervals between periodic points. This phenomenon becomes less significant as finer partitions are used, but persists nonetheless.

The two examples thus discussed have not been investigated in a rigorous manner. The principle purpose for including them has been to illustrate the relative robustness of Ulam's method to non–expansivity, and to display the sensitivity to partition choice.

A theory of how to construct a partition to minimise the approximation error in Ulam's method is lacking. These examples show that it is worthy of further exploration.

## Summary

In this chapter, the orientation has been towards computation. In the first few sections, we described a computer assisted procedure for constructing error bounds for Ulam's method which mirrored the structure of the proofs in Chapters 2 and 3. The essential point was to bound the norm $\|(Id - P)^{-1}\|$ *numerically* rather than *analytically*. The consequent procedure is independent of the dimension of the transformation, is easily implementable (if computationally expensive) and yields tighter error bounds than the methods of previous chapters. After illustrating the various approaches to error bounds on simple examples, the numerical procedure was applied to a certain conjugated family of smooth unimodal maps. This lead to rigorous bounds on the Ulam approximations to the (unknown) invariant densities, and motivated the definition of an *Adapted Ulam method*. The final section gave a brief discussion of this adapted algorithm, the basic idea of which is to automate the choice of a non–uniform partition for Ulam's method depending on the dynamics of the map. The numerical results with which the chapter is illustrated are extremely encouraging, suggesting that the algorithm is worthy of further investigation.

# Part III

# Invariant measures for multi–dimensional maps

# Chapter 5

# Multi–dimensional Bounded Variation and approximation error

In this chapter, the constructions from Chapter 3 are generalised to a multi–dimensional setting. The structure of the argument—using cones of uniformly bounded variation—remains unchanged, but the technical details require much more work. In due course, we are able to prove rates of mixing, and an $O(\log n/n)$ error bound for Ulam's method applied to expanding transformations in multi–dimensions. The approach is illustrated with an important example of ongoing theoretical interest [62, 36, 3, 14]: the Jacobi–Perron transformation in $\mathbb{R}^2$. The chapter concludes with a discussion of the relative merits and difficulties of precise implementations versus *Monte–Carlo* approximations of Ulam's method in multi–dimensions.

In the one–dimensional proofs in Chapter 3, the *variation* of an integrable function was a fundamental tool. The Lasota–Yorke inequality described the behaviour of variation under iteration by the Perron–Frobenius operator $\mathcal{L}$, and this facilitated the derivation of mixing rates in certain cones. Therefore, to generalise our methods to a multi–dimensional setting, it is first necessary to have a generalised definition of variation. (Historically, efforts to construct Lasota–Yorke type inequalities in multi–dimensions were frustrated by the intractability of variation in multi–dimensions; see the survey in [30].) In Section 5.1 we introduce *$p$–variation*—a slight modification of Giusti's multi–dimensional variation [27].

Now, for $p$–variation to be a useful tool for bounding the approximation error in Ulam's method, it is essential to have quantitative control of its behaviour under the projections $\Pi_\eta$. Section 5.2 is devoted to this problem, and a sharp result is obtained (Proposition 5.1). Although the proposition is unsurprising, the result does not seem to have appeared before[1]; since parts of the proof are

[1] A non–quantitative estimate appeared in [21].

straight–forward, but notationally intensive, they are included as Appendix D.

With Giusti's notion of variation, Góra and Boyarsky were able to construct a Lasota–Yorke type inequality in a multi–dimensional setting [30]. We call their result the *Góra–Boyarsky inequality* and Section 5.3 consists in presenting a slightly improved version (Theorems 5.1 and 5.2).

Recently, by exploiting the Góra–Boyarsky inequality and some non–constructive estimates about the behaviour of Giusti's variation under $\Pi_\eta$, Ding and Zhou [21] were able to generalise Li's convergence proof for Ulam's method [48] to a multi–dimensional setting. Unfortunately, their result is rather limited, and under a *finite range structure* (c.f. Section 3.4) assumption on $T$ we obtain a more complete generalisation of Li's work.

Consequently, the first three sections consist in gathering the ingredients for the analysis of Ulam's method. Indeed, much of our discussion is substantially similar to existing results in the literature. Yet our emphasis on *quantitative knowledge* necessitates enough adaptations of known results to justify a self–contained account. This effort is rewarded with many small improvements to multi–dimensional BV techniques.

In Section 5.4, the construction of lower bound functions from Chapter 3 is generalised, and concepts from Sections 5.1—5.3 are put to use. Again, fairly lengthy calculations are required, because the potentially complicated geometry of monotonicity components of $T^n$ makes the multi–dimensional construction more difficult. The argument is in two parts: First of all, lower bound functions are derived for iterates of certain *piecewise constant* initial densities. Then, an approximation argument (drawing on the results of Sections 5.2 and 5.3) is employed to build lower bound functions for iterates under $\mathcal{L}$ of arbitrary $BV$ initial densities.

Therefore, Sections 5.1—5.4 (which together comprise a majority of the chapter) culminate in the proof of Theorem 5.3, a lower bound theorem for iterates under $\mathcal{L}$ of pairs of $BV$ initial densities. Thereafter, the application to obtain a Mixing Theorem for $\mathcal{L}$ (Theorem 5.4) and error bounds for Ulam's method (Theorem 5.5) proceeds almost exactly as in Section 3.3. These are the main results of the chapter.

In Section 5.6, these results are applied to the classical Jacobi–Perron transformation [62, 3] on $\mathbb{R}^2$. This map is conjugate to a subshift of finite type, and is known to have a unique ergodic absolutely continuous measure, but no formula for its density exists [2]. We prove that Ulam's method will converge to the invariant density, and briefly comment on some of the issues surrounding the calculations required to implement this theoretical result in practice.

Finally, the chapter is concluded with a review of the various results that are proved, and the implications for calculation.

## 5.1 Bounded variation in $r$–dimensions

There have been several attempts to construct multi–dimensional generalisations of Lasota and Yorke's existence proof [47] for invariant densities, and Li's work on Ulam's method [48]. Boyarsky and Lou [13] use the *Tonelli variation* and a result of Jablonski [38] to prove weak convergence of Ulam's method for a limited class of multi–dimensional transformations. A concise proof of strong convergence appeared in [11] (also employing Tonelli variation). Independently, Blank [8] used a different definition of multi–dimensional variation to obtain some existence results for acims in multi–dimensions, and their robustness to perturbations. Further historical comment may be found in [30]. Most of these approaches have been limited by the choice of definition of multi–dimensional bounded variation.

The modern definition of *multi–dimensional variation* (found in Giusti's [27] book) is a natural extension of the one–dimensional version used in Chapter 3. Our applications to mixing rates and Ulam's method require a small modification.

Throughout the chapter, $m_k$ will denote $k$–dimensional Hausdorff measure. If $k = r$, then $m_k$ agrees with Lebesgue measure on $\mathbb{R}^r$. If $\Gamma$ is piecewise smooth of dimension $k$, then $m_{k,\Gamma}$ will denote the $k$–dimensional Hausdorff measure on $\Gamma$.

Next, we define the $p$–variation of a multi–dimensional function. The definition is a slight modification of the standard one from [27]. For $p \neq 2$, the definition is slightly unnatural (illustrated by examples below), but our application to Ulam's method requires $p = 1$.

DEFINITION (**Bounded $p$–Variation**) Let $A \subset \mathbb{R}^r$ be an open set, and let $C_0^1(A;\mathbb{R}^r)$ denote the collection of compactly supported smooth vector fields on $A$. Let div denote the divergence and for each $1 \leq p \leq \infty$, $f \in L^1(A)$ put

$$V_A^p(f) = \sup\left\{\int_A f(x)\text{div } w(x)\, dm(x) : w \in C_0^1(A;\mathbb{R}^r), |w(x)|_{p'} \leq 1 \ \forall x \in A\right\},$$

where $1/p + 1/p' = 1$ and $|(w_1,\ldots,w_r)|_{p'} = \left(\sum_{i=1}^r |w_i|^{p'}\right)^{1/p'}$ denotes the ordinary vector $p'$–norm. Then $V_A^p(f)$ is the *$p$–variation* of $f$ over $A$. One can easily check that if $1 \leq p \leq \infty$ then

$$V_A^\infty(f) \leq V_A^p(f) \leq V_A^1(f) \leq rV_A^\infty(f),$$

so that $V_A^p(f) < \infty$ for either all or no values of $p$; if $V_A^p(f) < \infty$, then $f$ is said to have *bounded variation over $A$* and it is usual to write $f \in BV(A)$. The set of $f \in BV(A)$ equipped with the norm

$$\|f\|_{BV,p} = \|f\| + V_A^p(f)$$

is a Banach space. Just as in the one–dimensional case, let

$$BV_0(A) = BV(A) \cap \left\{g \in L^1(A) : \int_A g\, dm = 0\right\}. \quad \square$$

REMARK. In our terminology, Giusti's definition would be 2–variation (so that the vector norm $|\cdot|_{p'}$ is the usual Euclidean norm on $\mathbb{R}^r$). We will work with $p = 1, 2$. Mostly, the usual 2–variation will be used, but the application to Ulam's method of rectangular partitions proceeds most easily using 1–variation, which is equivalent up to a constant:

$$V_A^2(f) \leq V_A^1(f) \leq \sqrt{r}\, V_A^2(f).$$

□

To calculate the $p$–variation of a given function, it will generally be useful to employ integration by parts. The remainder of this section is devoted to giving an intuitive explanation of how to compute with $V_A^p$.

If $f \in BV(A)$ then the *distributional derivative* [69] $df$ of $f$ is a finite vector–valued Radon measure [27, 1.5]. Then, we write

$$V_A^p(f) = \int_A |df|_p$$

The notation $|df|_p$ requires further explanation: If $f$ is differentiable at $x \in int(A)$ then

$$|df|_p(x) = \left( \sum_{i=1}^{r} \left| \frac{\partial f}{\partial x_i}(x) \right|^p \right)^{1/p} dm_r(x).$$

The only kind of non–differentiability which concerns us is a discontinuity along a surface of non–zero $(r-1)$–dimensional measure (any other singular points contribute nothing to the integral). We first describe how to restrict an $L^1$ function to an $(r-1)$–dimensional surface.

Let $\Gamma$ be a segment of a $\mathcal{C}^1$, $(r-1)$–dimensional manifold. If $x \in \Gamma$, let $v(x)$ be an oriented unit[2] vector orthogonal to the surface $\Gamma$ at $x$. If $B(x, \epsilon)$ is the ball of radius $\epsilon$ centred at $x$, and $\epsilon$ is sufficiently small, then $\Gamma$ partitions $B(x, \epsilon)$ into two connected components, one on either side of $\Gamma$; we denote these by $B_\epsilon^+$ and $B_\epsilon^-$ (for the sake of making a choice, let $B_\epsilon^+$ be on the positively oriented side of $\Gamma$, as described by $v$).

DEFINITION (**Trace of a function on an oriented surface [27, 2.5]**) For each point $x \in \Gamma$, put

$$\left(tr_\Gamma^+ f\right)(x) = \lim_{\epsilon \to 0} \frac{\int_{B_\epsilon^+} f\, dm_r}{m_r(B_\epsilon^+)} \quad \text{and} \quad \left(tr_\Gamma^- f\right)(x) = \lim_{\epsilon \to 0} \frac{\int_{B_\epsilon^-} f\, dm_r}{m_r(B_\epsilon^-)}. \qquad \square$$

If $f$ is discontinuous along a surface $\Gamma$, then $tr^+ f \neq tr^- f$ at points on $\Gamma$. See Figure 5.1 below.

[2] That is, $|v(x)|_2 = 1$.

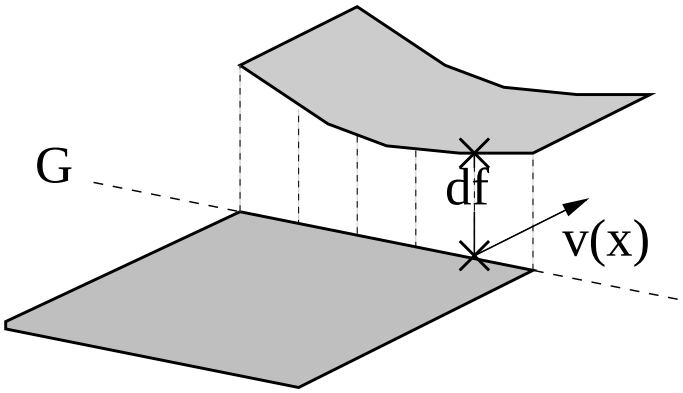



Figure 5.1: Discontinuity of a function $f$ along a surface $\Gamma$; $\Delta f = |(tr_\Gamma^+ f)(x) - (tr_\Gamma^- f)(x)|$.

Let $x \in \Gamma$. Then $df(x)$ is a vector–valued point mass, oriented in the direction specified by $v(x)$. For consistency of notation, we must have

$$|df|_p(x) = \left|\left(tr^+ f\right)(x) - \left(tr^- f\right)(x)\right| \, |v(x)|_p \, d\delta_{v(x)} \, dm_{r-1,\Gamma}(x)$$

where $d\delta_{v(x)}$ is interpreted as a point mass oriented in the direction specified by $v(x)$, and $dm_{r-1,\Gamma}(x)$ is the $(r-1)$–dimensional Hausdorff measure concentrated "orthogonal" to $v$ on $\Gamma$.

EXAMPLE [27, 1.4] Let $E$ be any set in $\mathbb{R}^r$ with piecewise $\mathcal{C}^2$ boundary, let

$$\chi_E(x) = \begin{cases} 1 & \text{if } x \in E, \\ 0 & \text{if } x \in \mathbb{R}^r \setminus E. \end{cases}$$

Then

$$V^2_{\mathbb{R}^r}(\chi_E) = m_{r-1}(\partial E),$$

where $m_{r-1}$ is $(r-1)$–dimensional Hausdorff measure, and $\partial E$ denotes the boundary of the set $E$. □

By adapting the argument in Giusti's proof of the above example, one obtains

$$V^p_{\mathbb{R}^r}(\chi_E) = \int_{\partial E} |v(x)|_p \, dm_{r-1}(x).$$

Therefore, the $p$–variation is sensitive to the orientation of a set, as the following simple example shows:

EXAMPLE. Let $E_1 = \{(x,y) \in \mathbb{R}^2 : 0 \le x, y \le 1\}$ and let $E_2 = \{(x,y) \in \mathbb{R}^2 : |x \pm y| \le 1/\sqrt{2}\}$. Then

$$V^1_{\mathbb{R}^r}(\chi_{E_1}) = 4 = V^2_{\mathbb{R}^r}(\chi_{E_1}) = V^2_{\mathbb{R}^r}(\chi_{E_2})$$

while

$$V^1_{\mathbb{R}^r}(\chi_{E_2}) = 4\sqrt{2}. \qquad \square$$

If $f$ is differentiable on the interior of a set $E$, but zero on the remainder of $A \setminus E$, then

$$V_A^p(f) = \int_E |\nabla f|_p \, dm_r + \int_{A \cap \partial E} |v|_p \left| tr^+ f - tr^- f \right| \, dm_{r-1, \partial E}.$$

More general functions are handled in a similar way. From time to time it will be convenient to ignore the contribution of the boundary to the variation. The convention

$$V_{int(E)}^p(f) = \int_{int(E)} |\nabla f|_p \, dm_r$$

will be adopted.

In Appendix C, more explanation is given of the relationship between $V_A^1(f)$ and $V_A^2(f)$. 1–variation is important for Ulam's method because it is well behaved with respect to projection onto partitions which comprise rectangles with faces parallel to coordinate axes; this is the main result of the next section.

REMARK. The *Tonelli variation* (see e.g. [13, 11]) is related to the 1–variation. In our notation, the Tonelli variation is given by

$$\mathrm{var}_A^{\mathrm{Tonelli}}(f) = \max_{1 \le i \le r} \int_A \left| \frac{\partial f}{\partial x_i} \right| \, dm_r,$$

where all partial derivatives are interpreted in a distributional sense. Obviously

$$\mathrm{var}_A^{\mathrm{Tonelli}}(f) \le V_A^1(f) \le r \, \mathrm{var}_A^{\mathrm{Tonelli}}(f),$$

for every $f \in BV$. $\square$

Finally, it is easy to check from the definition of $p$–variation that for any constant $c \in \mathbb{R}$, and $L^1$ functions $f, g$,

$$V_A^p(f+g) \le V_A^p(f) + V_A^p(g), \qquad V_A^p(cf) = |c| V_A^p(f)$$

and

$$V_A^p(f^{\pm}) \le V_A^p(f)$$

where $f^{\pm}$ are the positive and negative parts of $f$.

## 5.2 Behaviour of multi–dimensional variation under projection

This section generalises Lemma 3.1.1 to projections onto multi–dimensional rectangular partitions. The main result is Proposition 5.1. Blank [9] uses different definitions of multi–dimensional variation, and obtains results similar to Proposition 5.1. His definitions of variation are more complicated and difficult to work with, although one version agrees with $V_A^1$ in certain circumstances.

DEFINITION (**Rectangular partition**) If $X \in \mathbb{R}^r$ is a rectangular domain $X = (a_1, b_1) \times \ldots \times (a_r, b_r)$ then a partition $\eta$ of $X$ is a *rectangular partition* if every element of $\eta$ is a rectangle. Of particular interest is the case when there exist positive integers $N_1, \ldots, N_r$ such that if $B \in \eta$ then

$$B = B_{k_1 \ldots k_r} = [a_1 + k_1 h_1, a_1 + (k_1 + 1)h_1) \times \ldots \times [a_r + k_r h_r, a_r + (k_r + 1)h_r)$$

where[3] $0 \leq k_i < N_i$ and $h_i = (b_i - a_i)/N_i$. Such a partition will be called a *regular rectangular partition* and for each $x = (x_1, \ldots, x_r) \in X$, and $i = 1 \ldots r$ let $k_i(x)$ be the unique integer such that

$$a_i + k_i(x)h_i \leq x_i < a_i + (k_i(x) + 1)h_i. \qquad \square$$

REMARK 5.2.1. Put $h(\eta) = \min\{h_1, \ldots, h_r\}$ for $\eta$ a regular rectangular partition. If $f$ is constant on each subset from $\eta$, then

$$V^1(f) \leq \frac{2r\|f\|}{h(\eta)},$$

as is easily checked. $\square$

Let $\eta$ be a regular rectangular partition of a rectangular domain, and recall the projection operator $\Pi_\eta$ which acts on an $L^1$ function $f$ by the formula

$$\Pi_\eta f(x) = \sum_{B \in \eta} \frac{\int_B f \, dm_r}{m_r(B)} = \sum_{k_1, \ldots, k_r} \frac{\int_{B_{k_1 \ldots k_r}} f \, dm_r}{h_1 \ldots h_r}.$$

For each $i = 1 \ldots r$, define an operator

$$Q_i f(x) = \frac{\int_{x_i = a_i + k_i(x)h_i}^{a_i + (k_i(x)+1)h_i} f(x) \, dx_i}{h_i}$$

where $x = (x_1, \ldots, x_r)$. Each $Q_i$ averages over the $i$th coordinate direction, is a projection, and by Fubini's Theorem, $Q_i Q_j = Q_j Q_i$ for each $i, j$. Also by Fubini's Theorem,

$$\Pi_\eta = Q_1 Q_2 \ldots Q_r,$$

and the product may be taken in any order. The strategy for constructing a multi–dimensional version of Lemma 3.1.1 is to prove the result for each $Q_i$; the full result then follows.

First of all, some notation is needed. Recall that each $B_{k_1 \ldots k_r}$ is a rectangle

$$B_{k_1 \ldots k_r} = [a_1 + k_1 h_1, a_i + (k_1 + 1)h_1) \times \ldots \times [a_r + k_r h_r, a_r + (k_r + 1)h_r),$$

and for each $r$–tuple $k_1, \ldots, k_r$, $0 \leq k_i \leq N_i$ and $i = 1 \ldots r$ let

$$\Gamma^i_{k_1 \ldots k_r} = \left\{ \begin{array}{ll} (x_1, \ldots, x_r) \in \mathbb{R}^r : & a_j + k_j h_j \leq x_j < a_j + (k_j + 1)h_j \text{ if } j \neq i \\ & \text{and} \qquad x_i = a_i + k_i h_i \end{array} \right\}.$$

[3] If any $k_i = 0$, then the corresponding interval in the product should be $(a_i, a_i + h_i)$.

Then

$$\partial B_{k_1\ldots k_r} = \cup_{i=1}^{r} \left\{ \Gamma^i_{k_1\ldots k_r} \cup \Gamma^i_{k_1\ldots(k_i+1)\ldots k_r} \right\}.$$

Finally, let $\Gamma^i_{k_1\ldots k_r}$ be oriented in the *positive* direction of the unit normal vector $e_i$. Then, $tr^+_{\Gamma^i_{k_1\ldots k_r}} f$ is the restriction of a function $f$ to the surface $\Gamma^i_{k_1\ldots k_r}$ from the $B_{k_1\ldots k_r}$ side of $\Gamma^i_{k_1\ldots k_r}$, and $tr^-_{\Gamma^i_{k_1\ldots k_r}} f$ is the restriction of $f$ from the $B_{k_1\ldots(k_i-1)\ldots k_r}$ side of $\Gamma^i_{k_1\ldots k_r}$. A graph of the situation is depicted in Figure 5.2.

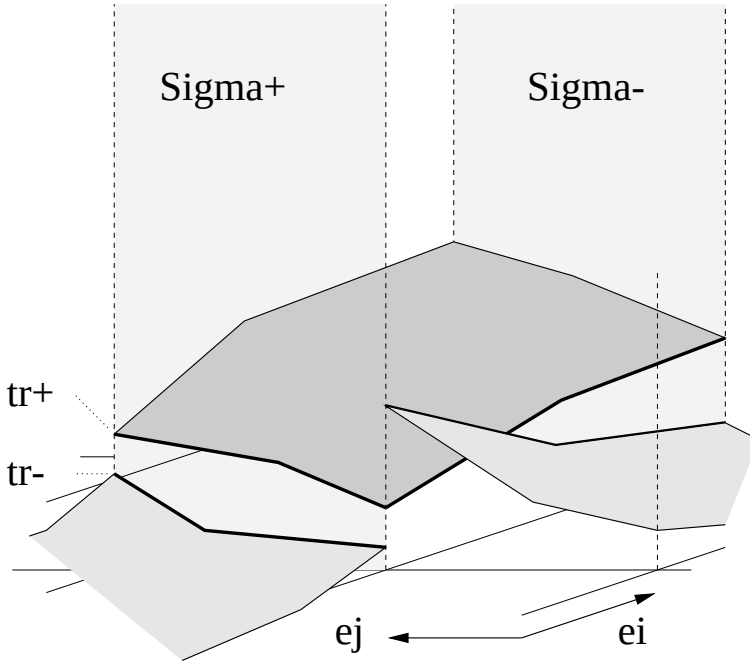



Figure 5.2: A function $f$ which is smooth on each rectangular partition element *before* the application of the projection $Q_i$. In the diagram, $\Gamma^- = \Gamma^i_{k_1\ldots k_r}$ and $\Gamma^+ = \Gamma^i_{k_1\ldots(k_i+1)\ldots k_r}$, where the rectangle in question is $B_{k_1\ldots k_r}$.

**Lemma 5.2.2** *For each collection of indices* $k_1, \ldots, k_r$ *and* $i = 1 \ldots r$*:*

1. *if* $j \neq i$ *then*

$$\int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial Q_i f}{\partial x_j} \right| \leq \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_j} \right| ;$$

2. *if* $j \neq i$*, let* $\Gamma^j = \Gamma^j_{k_1 \ldots k_r}$*. Then*

$$\int_{\Gamma^j} \left| tr^+_{\Gamma^j}(Q_i f) - tr^-_{\Gamma^j}(Q_i f) \right| \leq \int_{\Gamma^j} \left| tr^+_{\Gamma^j} f - tr^-_{\Gamma^j} f \right| ;$$

3. *put* $\Gamma^- = \Gamma^i_{k_1 \ldots k_r}$ *and* $\Gamma^+ = \Gamma^i_{k_1 \ldots (k_i+1) \ldots k_r}$*. Then*

$$\begin{aligned} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_i} \right| \quad &\geq \quad \int_{\Gamma^-} \left| tr^+_{\Gamma^-}(Q_i f) - tr^+_{\Gamma^-} f \right| \\ &\quad + \int_{\Gamma^+} \left| tr^-_{\Gamma^+}(Q_i f) - tr^-_{\Gamma^+} f \right| ; \end{aligned}$$

4. *Finally,*

$$\int_{B_{k_1 \ldots k_r}} |f - Q_i f| \leq h_i \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_i} \right| .$$

*Proof:* See Appendix D. □

Lemma 5.2.2 contains all the technical details required for controlling $V_A^1$ under projection onto a regular rectangular partition.

**Proposition 5.1** *Let* $A \subset \mathbb{R}^r$ *be a rectangle, let* $\eta = \{B_{k_1 \ldots k_r}\}_{k_i=0}^{N_i - 1}$ *be a regular rectangular partition of* $A$ *and let* $f \in BV(A)$*. Then*

1.

$$V_A^1(\Pi_\eta f) \leq V_A^1(f),$$

2. *and*

$$\|f - \Pi_\eta f\| \leq h \, V_A^1(f),$$

*where* $h = \max\{h_1, \ldots, h_r\}$.

*Proof:* Interpreting the partial derivatives $\partial / \partial x_i$ in the distributional sense and applying Fubini's theorem (cf. Appendix C), for any $g \in BV(A)$

$$\begin{aligned} V_A^1(g) \quad = \quad \sum_{i=1}^{r} \Big[ \sum_{k_1 \ldots k_r = 0}^{N_1, \ldots, N_r} \int_{\Gamma^i_{k_1 \ldots k_r}} \left| tr^+_{\Gamma^i_{k_1 \ldots k_r}} g - tr^-_{\Gamma^i_{k_1 \ldots k_r}} g \right| \\ + \sum_{k_1, \ldots, k_r = 0}^{N_1 - 1, \ldots, N_r - 1} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial g}{\partial x_i} \right| \Big] \end{aligned}$$

where $g$ is extended by setting $g = 0$ outside $A$. Now, consider the function $g = Q_j f$. By Lemma 5.2.2 (1) and (2), the terms in the above sum with $i \neq j$ are bounded by the equivalent integrals of $f$ and $tr^{\pm} f$. For the $j$ terms, $|\partial Q_j f / \partial x_j| = 0$ on the interior of each $B_{k_1 \ldots k_r}$, and for any $\Gamma$,

$$\left| tr^+_\Gamma Q_j f - tr^-_\Gamma Q_j f \right| \leq \left| tr^+_\Gamma Q_j f - tr^+_\Gamma f \right| + \left| tr^+_\Gamma f - tr^-_\Gamma f \right| + \left| tr^-_\Gamma f - tr^-_\Gamma Q_j f \right|.$$

Therefore, for fixed $k_1, \ldots, k_{j-1}, k_{j+1}, k_r$,

$$\begin{aligned}
&\sum_{k_j=0}^{N_j} \int_{\Gamma^j_{k_1 \ldots k_r}} \left| tr^+_{\Gamma^j_{k_1 \ldots k_r}} Q_j f - tr^-_{\Gamma^j_{k_1 \ldots k_r}} Q_j f \right| \\
&\quad \leq \sum_{k_j=0}^{N_j} \int_{\Gamma^j_{k_1 \ldots k_r}} \left| tr^+_{\Gamma^j_{k_1 \ldots k_r}} f - tr^-_{\Gamma^j_{k_1 \ldots k_r}} f \right| \\
&\qquad + \sum_{k_j=0}^{N_j-1} \int_{\Gamma^j_{k_1 \ldots k_r}} \left| tr^+_{\Gamma^j_{k_1 \ldots k_r}} f - tr^+_{\Gamma^j_{k_1 \ldots k_r}} Q_j f \right| \\
&\qquad\qquad + \int_{\Gamma^j_{k_1 \ldots (k_j+1) \ldots k_r}} \left| tr^-_{\Gamma^j_{k_1 \ldots (k_j+1) \ldots k_r}} f - tr^-_{\Gamma^j_{k_1 \ldots (k_j+1) \ldots k_r}} Q_j f \right| \\
&\quad \leq \sum_{k_j=0}^{N_j} \int_{\Gamma^j_{k_1 \ldots k_r}} \left| tr^+_{\Gamma^j_{k_1 \ldots k_r}} f - tr^-_{\Gamma^j_{k_1 \ldots k_r}} f \right| + \sum_{k_j=0}^{N_j-1} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_j} \right|
\end{aligned}$$

where Lemma 5.2.2 (3) was used for the second inequality. Summing up,

$$V^1_A(Q_j f) \leq V^1_A(f).$$

Because $\Pi_\eta = Q_1 \ldots Q_r$, successive application of this inequality for $j = 1 \ldots r$ proves the first part of the proposition.

For the other part, Lemma 5.2.2 (4) implies that

$$\int_A |f - Q_i f| \leq h_i \sum_{k_1 \ldots k_r} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_i} \right|,$$

and repeated application of Lemma 5.2.2 (1) implies that

$$\sum_{k_1 \ldots k_r} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial}{\partial x_i} Q_j \ldots Q_1 f \right| \leq \sum_{k_1 \ldots k_r} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial f}{\partial x_i} \right|$$

for each $i, j = 1 \ldots r$. Then

$$\begin{aligned}
\int_A |f - \Pi_\eta f| &\leq \sum_{j=1}^r \int_A |Q_{j-1} \ldots Q_1 f - Q_j \ldots Q_1 f| \\
&\leq \sum_{j=1}^r h_j \sum_{k_1 \ldots k_r} \int_{int(B_{k_1 \ldots k_r})} \left| \frac{\partial}{\partial x_j} Q_{j-1} \ldots Q_1 f \right| \\
&\leq h \sum_{j=1}^r \int_A \left| \frac{\partial f}{\partial x_j} \right|.
\end{aligned}$$

□

Proposition 5.1 is essential to the analysis of approximation error in Ulam's method. However, the result itself is unsurprising; the proof essentially consists in translating the one–dimensional equivalent (Lemma 3.1.1) into a multi–dimensional setting. The key feature is that the 1–variation can be decomposed into a collection of integrals in coordinate directions—a situation to which rectangular partitions are ideally suited. It is less obvious how to generalise the proposition to partitions without rectangular geometry, and there does not seem to be an easy direct proof of the following:

**Corollary 5.2.3** *For $\eta$ a regular rectangular partition of a rectangle $A \subset \mathbb{R}^r$ and function $f \in BV(A)$,*

$$V_A^2(\Pi_\eta f) \leq V_A^1(\Pi_\eta f) \leq V_A^1(f) \leq \sqrt{r}\, V_A^2(f),$$

*and*

$$\|\Pi_\eta f - f\| \leq h\sqrt{r} V_A^2(f).$$

Below, in the Góra–Boyarsky inequality, it is necessary to work with the 2–variation; 1–variation has merely been a useful tool to prove Corollary 5.2.3. Indeed, Ding and Zhou's paper [21] devoted substantial effort to proving the existence of a constant $c = c(r)$ such that

$$V_A^2(\Pi_\eta f) \leq c V_A^2(f).$$

For quantitative analysis, it is important to know what the constant $c$ actually is, and the estimate in Corollary 5.2.3 is sharp, as the following example shows:

EXAMPLE: Let $A = [0,1]^2 \subset \mathbb{R}^2$, and $f(x_1, x_2) = x_1 + x_2$. Then, $\nabla f = (1,1)$ and integration implies

$$V_{int(A)}^1(f) = |(1,1)|_1 = 2 \qquad \text{and} \qquad V_{int(A)}^2(f) = |(1,1)|_2 = \sqrt{2}.$$

Let $\eta$ be the regular rectangular partition obtained by dividing $A$ into squares with side–length $\frac{1}{n}$. Then, in the notation established in the previous section,

$$\Pi_\eta f(x) = \frac{k_1(x) + k_2(x) + 1}{n}.$$

Adding up the contributions to $V_{int(A)}^1(\Pi_\eta f)$ from all the boundaries of the sub–rectangles $B_{k_1 k_2}$ it follows that

$$V_{int(A)}^1(\Pi_\eta f) = V_{int(A)}^2(\Pi_\eta f) \geq 2\frac{n-1}{n},$$

proving the optimality of Corollary 5.2.3. □

REMARK. I am grateful to Dr. Chris Bose for discussions about multi-dimensional variation, during which he pointed out the above example. □

## 5.3 Extended Lasota–Yorke inequality and FRS improvements

We now discuss Góra and Boyarsky's [30] multi–dimensional Lasota–Yorke inequality, and present in subsection 5.3.1 a slight modification which allows the treatment of infinitely many monotonicity pieces. Although many of the calculations are routine, the objective of being able to treat transformations with infinitely many monotonicity pieces means that the estimates must be done in a slightly different order to [30]. Consequently, there is some overlap in *content* between this section and Góra and Boyarsky's paper. For the sake of a clear adaptation to our scenario, this repetition seems justified.

After deriving an improved Góra–Boyarsky inequality (Theorem 5.2), a brief exposition is given of Ding and Zhou's recent proof [21] of the convergence of Ulam's method in multi–dimensions. In fact, Ding and Zhou assume rather more of the Góra–Boyarsky inequality than was proved in [30], and we explain in subsection 5.3.2 how the consequent weakness in their assumptions can be filled by the existence of a *finite range structure* for $T$. Finally, in subsection 5.3.3 an explicit construction of the constants in the Góra–Boyarsky inequality is given; this keeps the application to real examples transparent.

### Notational convention

Henceforth, only the multi–dimensional 2–variation will be used, so that

$$V_A(f) \triangleq V_A^2(f).$$

Unless otherwise stated, the vector norm $|\cdot|$ will be the vector 2–norm

$$|(v_1, \ldots, v_r)| = \sqrt{\sum_{i=1}^{r} v_i^2}$$

and the matrix norm $|\cdot|$ is the corresponding operator norm:

$$|A| = \sup_{x \in \mathbb{R}^r} \frac{|Ax|}{|x|} = \sup_{x \in \mathbb{R}^r} \frac{|Ax|_2}{|x|_2},$$

where $A$ is an $r \times r$ matrix.

### The transformations $T$

The transformations $T$ of an $r$–dimensional set $X$ considered in this section must satisfy the following properties:

**Piecewise Monotonicity:** There exists a partition $\xi = \{B_\alpha\}$ of $X$ such that $T$ is $\mathcal{C}^2$ and non–singular on each $int(X_\alpha)$, and the boundary $\partial B_\alpha$ of each $B_\alpha$ is piecewise $\mathcal{C}^2$. Additional restrictions on the geometry of the pieces $B_\alpha$ will

be imposed below. The inverse branches will be denoted by $T_\alpha^{-1}$. The inverse branches of $T^n$ will be denoted

$$T_{\alpha^{(n)}}^{-n} = T_{\alpha_n \dots \alpha_1}^{-n} \triangleq T_{\alpha_n}^{-1} \circ \dots T_{\alpha_1}^{-1}.$$

**Expansivity:** There exist *global* constants $C \geq 0$, $\lambda > 1$ such that

$$\left| DT_{\alpha^{(n)}}^{-n} \right| \leq C\lambda^{-n} \tag{5.1}$$

where $DT_{\alpha^{(n)}}^{-n}$ is the Jacobian matrix of $T_{\alpha^{(n)}}^{-n}$.

**Bounded Distortion:** There exists a constant $K \geq 0$ such that

$$\left| \frac{\nabla \left( \det(DT_\alpha^{-1}) \right)}{\det(DT_\alpha^{-1})} \right| \leq K. \tag{5.2}$$

Writing

$$\det \left( DT_{\alpha^{(n)}}^{-n} \right) = \Pi_{k=1}^n \det \left( DT_{\alpha_k}^{-1} \right) \circ T_{\alpha_{k-1} \dots \alpha_1}^{-(k-1)},$$

differentiation yields

$$\nabla \left( \det(DT_{\alpha^{(n)}}^{-n}) \right) = \det(DT_{\alpha^{(n)}}^{-n}) \sum_{k=1}^n \frac{\nabla \left( \det(DT_{\alpha_k}^{-1}) \circ T_{\alpha_{k-1} \dots \alpha_1}^{-(k-1)} \right)}{\det(DT_{\alpha_k}^{-1}) \circ T_{\alpha_{k-1} \dots \alpha_1}^{-(k-1)}},$$

so that

$$\begin{aligned} \left| \frac{\nabla \left( \det(DT_{\alpha^{(n)}}^{-n}) \right)}{\det(DT_{\alpha^{(n)}}^{-n})} \right| &\leq \sum_{k=1}^n \left| DT_{\alpha_{k-1} \dots \alpha_1}^{-(k-1)} \left( \frac{\nabla \det(DT_{\alpha_k}^{-1})}{\det(DT_{\alpha_k}^{-1})} \circ T_{\alpha_{k-1} \dots \alpha_1}^{-(k-1)} \right) \right| \\ &\leq \sum_{k=1}^n \sup \left| DT_{\alpha^{(k-1)}}^{-(k-1)} \right| \left| \frac{\nabla \det(DT_{\alpha_k}^{-1})}{\det(DT_{\alpha_k}^{-1})} \right| \\ &\leq C \frac{1-\lambda^{-n}}{1-\lambda^{-1}} K. \end{aligned} \tag{5.3}$$

For constructing lower bound functions of iterates under the Perron–Frobenius operator for $T$, we will need a Renyi constant $D \geq 1$ such that

$$\left| \frac{\det DT_{\alpha^{(n)}}^{-n}(x)}{\det DT_{\alpha^{(n)}}^{-n}(y)} \right| \leq D \tag{5.4}$$

whenever $x, y$ are in the image of the same monotonicity component of $T^n$. It follows from (5.3) that

$$D \leq e^{\frac{CK}{1-\lambda^{-1}}}.$$

The coincidence of the notation for the distortion constant $D$ and the symbol for the Jacobian matrix of $T$ is unfortunate, but causes no ambiguity.

Let $\mathcal{L}$ be the Perron–Frobenius operator corresponding to $T$. Then

$$\mathcal{L}f(x) = \sum_\alpha \frac{f \circ T_\alpha^{-1}}{\left|\det(DT) \circ T_\alpha^{-1}\right|} \chi_{T(B_\alpha)},$$

and

**Lemma 5.3.1** *For each $f \in BV(X)$, and $T$ satisfying (5.1) and (5.2),*

*1.*

$$\begin{aligned} V_X(\mathcal{L}f) \;\leq\; \sum_\alpha \Big( & \int_{int(TB_\alpha)} \left| d\left( \frac{f \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right) \right| \\ & + \int_{\partial(TB_\alpha)} \left| tr^+_{\partial(TB_\alpha)} \left( \frac{f \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right) \right| \Big), \end{aligned}$$

*2. for each $\alpha$,*

$$\int_{int(TB_\alpha)} \left| d \frac{f \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right| \leq C\lambda^{-1} \int_{int(B_\alpha)} |df| + K \int_{B_\alpha} |f|.$$

*Proof:* The first part follows immediately from the definitions of variation and $\mathcal{L}$. The second part is similar to Lemma 3.1.2:

$$\begin{aligned} & \int_{int(TB_\alpha)} \left| d \frac{f \circ T_\alpha^{-1}}{|\det DT \circ T_\alpha^{-1}|} \right| \\ & \quad = \int_{int(TB_\alpha)} \left| \frac{d(f \circ T_\alpha^{-1})}{|\det DT \circ T_\alpha^{-1}|} + f \circ T_\alpha^{-1} d\left( \frac{1}{|\det DT \circ T_\alpha^{-1}|} \right) \right| \\ & \quad \leq \int_{int(TB_\alpha)} \left| DT_\alpha^{-1} \left( \frac{(df) \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right) \right| \\ & \qquad\qquad + \int_{int(TB_\alpha)} \left| \frac{f \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \frac{d\left(\det DT \circ T_\alpha^{-1}\right)}{\det DT \circ T_\alpha^{-1}} \right| \\ & \quad \leq C\lambda^{-1} \int_{int(TB_\alpha)} \left| \frac{(df) \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right| \\ & \qquad\qquad + K \int_{int(TB_\alpha)} \left| \frac{f \circ T_\alpha^{-1}}{\det DT \circ T_\alpha^{-1}} \right| \\ & \quad = C\lambda^{-1} \int_{int(B_\alpha)} |df| + K \int_{B_\alpha} |f|. \end{aligned}$$

□

Generalising the rest of the proof of Lemma 3.1.2 requires more work.

### Geometry of monotonicity pieces

The second half of the proof of Lemma 3.1.2 (the Lasota–Yorke inequality) consists in estimating the contribution to $V(\mathcal{L}f)$ from the boundary points of each image of a monotonicity interval. The major step in [30] is to generalise this procedure to multi–dimensions. In our notation, this involves bounding the second term in the sum in Lemma 5.3.1 (1). The key estimate from the one–dimensional version [47] is:

$$|\mathcal{L}f(a)| + |\mathcal{L}f(b)| \leq c_1 \int_{T_\alpha^{-1}[a,b]} |f'| + c_2 \int_{T_\alpha^{-1}[a,b]} |f|,$$

where $[a,b]$ is the image of a monotonicity interval of $T$.

REMARK. In both the original in [47] and our version in Lemma 3.1.2, $c_1 = 1$, but different estimates are obtained for $c_2$ (Lemma 3.1.2 allows infinitely many monotonicity intervals and in all cases the $c_2$ from [47] is an upper bound for the analogous quantity in Lemma 3.1.2). □

Now, let $B_\alpha$ be a monotonicity piece of $T$, and let the boundary of $TB_\alpha$ be oriented with a vector pointing inwards. The problem is to find constants $c_1, c_2$ such that

$$\int_{\partial(TB_\alpha)} \left| tr^+_{\partial(TB_\alpha)}(\mathcal{L}f) \right| \leq c_1 \int_{int(B_\alpha)} |df| + c_2 \int_{B_\alpha} |f|.$$

Góra and Boyarsky establish the following geometrical result:

**Lemma 5.3.2 ([30])** *Let $S$ be a closed domain with piecewise $\mathcal{C}^2$ boundary of finite $(r-1)$–dimensional measure, whose smooth faces meet at angles bounded away from zero. Then there exists a constant $a(S) > 0$, such that for any $\epsilon > 0$ there exists a positive $\delta = \delta_{\epsilon,S}$ with*

$$\int_{\partial S} |tr^+_{\partial S} g| \leq \frac{(1+\epsilon)^2}{a(S) - \epsilon} \left( \int_{int(S)} |dg| + \frac{1}{\delta} \int_S |g| \right)$$

*for any $g \in BV(S)$. The constants $a(S)$ and $\delta$ depend on the geometry of $S$.*

REMARK. Lemma 5.3.2 is essentially Lemma 3 from [30], with a slight modification to correct an error [28]. □

Góra and Boyarsky's proof [30] of Lemma 5.3.2 does not give an explicit construction for the constant $\delta = \delta(S, \epsilon)$, and no examples are presented. However, for explicit error bounds for Ulam's method, bounds on $a(S)$ and $\delta$ are essential. Fortunately, the geometric construction given in [30] may be explicitly realised in certain simple settings. An example of this (used in the application to the Jacobi–Perron transformation) is given is subsection 5.3.3 below.

DEFINITION: A closed set $S \subset \mathbb{R}^r$ with piecewise $\mathcal{C}^2$ boundary of finite $(r-1)$–dimensional measure whose faces meet at angles bounded away from 0 has $a(S) > 0$. Such an $S$ will be called *non–degenerate*. □

### 5.3.1 An improved Góra–Boyarsky inequality

Let $T$ satisfy (5.1) and (5.2), and suppose that each set $TB_\alpha$ is non–degenerate. Then Lemma 5.3.2 holds for each $\alpha$ with constants $a(TB_\alpha)$ and $\delta_{\epsilon,\alpha}$. Suppose that for any $\epsilon > 0$ there exist $a, \delta_\epsilon$ such that

$$\delta_\epsilon \leq \delta_{\epsilon,TB_\alpha} \qquad a \leq a(TB_\alpha)$$

for each $B_\alpha \in \xi$. Then

**Theorem 5.1 (Góra–Boyarsky Inequality [30, 28])** *If $T$ has sufficiently strong expansion that $C\lambda^{-1}(1+1/a) < 1$ then for every $\sigma \in (C\lambda^{-1}(1+1/a), 1)$ there exists a constant $A_\sigma > 0$ such that*

$$V_X(\mathcal{L}f) \leq \sigma V_X(f) + A_\sigma \|f\|$$

*whenever $f \in BV(X)$.*

*Proof:* Let $\epsilon > 0$ be such that

$$C\lambda^{-1}\left(1 + \frac{(1+\epsilon)^2}{a-\epsilon}\right) \leq \sigma$$

and let $\delta = \delta_\epsilon$ be such that Lemma 5.3.2 holds uniformly for all $S = TB_\alpha$ with $\delta_{\epsilon,S} \geq \delta_\epsilon$ and $a(S) \geq a$. Then, by Lemmas 5.3.1 and 5.3.2,

$$\begin{aligned} V_X(\mathcal{L}f) \quad \leq \quad & C\lambda^{-1}\sum_\alpha \left(1 + \frac{(1+\epsilon)^2}{a-\epsilon}\right)\int_{int(B_\alpha)} |df| \\ & + \sum_\alpha \left(K + \frac{(1+\epsilon)^2}{a-\epsilon}\frac{1}{\delta_\epsilon}\right)\int_{B_\alpha} |f|. \end{aligned}$$

By putting $A_\sigma = K + (1+\epsilon)^2/\delta_\epsilon(a-\epsilon)$ the proof is complete. □

REMARK. Rather than assuming that each $TB_\alpha$ is non–degenerate, Góra and Boyarsky assumed that each $B_\alpha$ is non–degenerate and set $a = \min_\alpha\{a(B_\alpha)\}$ and $\delta = \min_\alpha\{\delta_{\epsilon,B_\alpha}\}$. To obtain their result, they used $T_\alpha^{-1}$ to pull back the boundary integrals $\int_{\partial(TB_\alpha)} \left|tr^+_{\partial(TB_\alpha)}(\mathcal{L}f)\right|$ before using Lemma 5.3.2 to obtain upper bounds. Consequently, Theorem 5.1 generalises the Góra–Boyarsky inequality in [30] because the geometric conditions (existence of $a$ and $\delta$) are imposed on the *image sets* $TB_\alpha$ rather than the monotonicity sets $B_\alpha$ themselves. Certainly Góra and Boyarsky's assumptions are implied by ours, since a uniformly $\mathcal{C}^2$ expanding map cannot transform a non–degenerate set into a degenerate one[4]. Moreover, the Jacobi–Perron transformation—described below—has infinitely many monotonicity pieces, whose diameters and angle of boundary intersection go to zero. But since each monotonicity piece has

[4]The angle of intersection of boundary faces must remain bounded away from zero, and the boundaries must retain finite $(r-1)$–dimensional measure.

only two possible images under $T$, and both are non–degenerate, Theorem 5.1 applies. □

Finally:

**Theorem 5.2 (Improved Góra—Boyarsky inequality)** *Suppose that $T$ satisfies (5.1) and (5.2), and that $T$ has a* finite range structure $\{U_1, \ldots, U_k\}$ . *If each set $U_i$ is non–degenerate, then for every $\sigma \in (0,1)$ there exists $n_\sigma > 0$ and a constant $A_\sigma$ such that*

$$V_X(\mathcal{L}^{n_\sigma} f) \leq \sigma V_X(f) + A_\sigma \|f\|$$

*for every $f \in BV(X)$. Moreover, the constants $A_\sigma$ are uniformly bounded, and $n_*$ has an explicit formula.*

*Proof:* Let $a = \min_{i=1}^k a(U_i)$ and let $\epsilon > 0$ be such that $(1+\epsilon)^2/(a-\epsilon) = 2/a$. Let $\delta > 0$ be such that Lemma 5.3.2 holds with $\delta_{\epsilon,U_i} \geq \delta$ for each $i$. Then, let $n_\sigma$ be minimal such that $C\lambda^{-n_\sigma}(1+2/a) \leq \sigma$. Then, as in the proof of Theorem 5.1,

$$V_X(\mathcal{L}^{n_\sigma} f) \leq \sigma V_X(f) + \left( CK \frac{1-\lambda^{-n_\sigma}}{1-\lambda^{-1}} + \frac{2}{a\delta} \right) \|f\|,$$

(we have used (5.3) and the fact that for every monotonicity set $B_{\alpha^{(n_\sigma)}}$ of $T^{n_\sigma}$ there exists $i \in 1 \ldots k$ such that $T^{n_\sigma} B_{\alpha^{(n_\sigma)}} = U_i$). The theorem follows. □

Theorem 5.2 generalises the Góra–Boyarsky inequality from [30] in one further way: the finite range structure of $T$ allows sufficient iterates of $T$ to be taken to get arbitrarily strong contraction of the variation. This is not necessarily case with the original Góra–Boyarsky inequality, because as $n$ increases, the geometry of the monotonicity pieces of $T^n$ will likely get more and more degenerate. Consequently, the constant $a$ may decrease faster than the counteracting effect of the greater expansion induced by taking more iterates of $T$. Since Góra and Boyarsky's argument takes constants $a$ and $\delta$ as being from the monotonicity pieces, rather than their images, the following example illustrates that this approach cannot yield a result as strong as Theorem 5.2:

Example (Importance of Finite Range Structure) Let $X = T^2 = \mathbb{R}^2/\mathbb{Z}^2$ be the usual 2–torus, and let

$$T : \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \mapsto \begin{bmatrix} 2 & -5 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \pmod 1.$$

Each monotonicity piece of $T^n$ is a parallelogram, with one pair of sides parallel to the $x_1$–axis. Moreover,

$$|DT^{-n}| = 2^{-n},$$

for each positive $n$. The other important estimate for the Góra–Boyarsky inequality is the smallest interior angle of the boundary of a monotonicity piece of

$T^n$. Because each monotonicity piece of $T^n$ maps exactly onto $[0,1)^2$, it follows that the minimal interior angle of such a piece is the angle between

$$\left(\begin{bmatrix} 2 & -5 \\ 0 & 5 \end{bmatrix}^{-1}\right)^n \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \text{and} \quad \left(\begin{bmatrix} 2 & -5 \\ 0 & 5 \end{bmatrix}^{-1}\right)^n \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

It is easy to check that this angle scales as $(\frac{2}{5})^n$, so that the value of $a = a(n)$ for monotonicity pieces of $T^n$ scales as $(\frac{5}{2})^n$. Therefore the product of $|DT^{-n}|$ and $a(n)$ grows exponentially in $n$, preventing the construction of Theorem 5.2 by this method. In the analysis of Ulam's method below, it is important to be able to take many iterates of $T$ before applying the Góra–Boyarsky inequality. Fortunately, this example has a finite range structure (every monotonicity piece maps onto all of $X$), so that Theorem 5.2 applies. □

Returning to the main argument:

DEFINITION (**Cone $\mathcal{C}_M$**) In complete analogy with the situation in Chapter 3, let

$$\mathcal{C}_M = \{f \geq 0 : V_X(f) \triangleq V_X^2(f) \leq M\|f\|\}.$$

Then each $\mathcal{C}_M$ is a cone of *uniformly bounded (multi–dimensional) variation.* The (multi–dimensional) *difference cone* $\Gamma_M$ and *difference norm* $\|\cdot\|_M$ are defined in the same way as the one–dimensional equivalents, and Lemma 3.2.3 holds. □

For the sake of completeness, the conclusions of Theorem 5.2 and Corollary 5.2.3 can be summarised as:

**Corollary 5.3.3** *Let $T$ be a uniformly expanding map on a rectangle in $\mathbb{R}^r$ with a non–degenerate finite range structure, and let $a \geq 0$.*

1. *For any $0 < \sigma < 1$ there exists $n_\sigma > 0$ and $A_\sigma < \infty$ such that*

$$\mathcal{L}^{n_\sigma}\mathcal{C}_M \subset \mathcal{C}_{\sigma M + A_\sigma}.$$

2. *For any regular rectangular partition $\eta$,*

$$\Pi_\eta \mathcal{C}_M \subset \mathcal{C}_{\sqrt{r}M}.$$

REMARK 5.3.4. The existence of an invariant density for $T$ follows from Corollary 5.3.3. See [30] for more details. If $T$ has a *unique* invariant density (as is the case in the situation described below) then any fixed point of $\mathcal{L}^n$ is a multiple of the invariant density $f$. Therefore, we can apply the Góra–Boyarsky inequality to arbitrarily high iterates of $\mathcal{L}$ to get an estimate of $V(f)$: let $A_\infty$ be such that $A_\sigma \leq A_\infty$ for every $\sigma \in (0,1)$ (c.f. Theorem 5.2). Then for every $\sigma > 0$ there exists $n_\sigma$ such that

$$\mathcal{L}^{n_\sigma}\mathcal{C}_M \subset \mathcal{C}_{\sigma M + A_\infty}.$$

Therefore,

$$f \in \mathcal{C}_{A_\infty/(1-\sigma)} \qquad \forall \sigma \in (0,1),$$

so that

$$V(f) \leq A_\infty \|f\|$$

for the fixed point of $\mathcal{L}$. □

REMARK. A generalisation of the Góra–Boyarsky inequality to a situation where $T$ has finitely many monotonicity pieces, but some have cusp–like degeneracies is given in [1]. □

### 5.3.2 Convergence of Ulam's method

The convergence of a suitably modified version of Ulam's method follows immediately from Corollary 5.3.3. The argument is the same as in [21]:

Choose $\sigma$ such that $\sigma \times \sqrt{r} < 1$, and let $n_\sigma$ be such that Corollary 5.3.3 holds. Then, for any regular rectangular partition $\eta$,

$$\Pi_\eta \circ \mathcal{L}^{n_\sigma} \mathcal{C}_M \subset \mathcal{C}_{(\sqrt{r}\sigma M + \sqrt{r} A_\sigma)}.$$

Hence, for sufficiently large $M$, the cone $\mathcal{C}_M$ is preserved by the operator

$$\mathcal{P}_{\eta,n_\sigma} = \Pi_\eta \circ \mathcal{L}^{n_\sigma}.$$

A standard compactness argument, together with Corollary 5.2.3 implies that the fixed points of $\mathcal{P}_{\eta,n_\sigma}$ converge to fixed points of $\mathcal{L}^{n_\sigma}$ as the partition $\eta$ is refined. If $\|f_\eta\| = 1$ and $\mathcal{P}_{\eta,n_\sigma} f_\eta = f_\eta$ then put

$$g_\eta = \frac{1}{n_\sigma} \sum_{k=0}^{n_\sigma - 1} \mathcal{L}^k f_\eta.$$

Then $g_\eta$ converges to a fixed point of $\mathcal{L}$ as $\eta$ is refined.

REMARK. Ding and Zhou's version [21, Theorem 3.1] of Corollary 5.3.3 (2) is not as strong as ours. They prove only that there *exists* a constant $c < \infty$ such that $\Pi_\eta \mathcal{C}_M \subset \mathcal{C}_{cM}$. In [21, Theorem 3.1], they assume that the Góra–Boyarsky from [30] implies that it is always possible to have $\sigma c < 1$ by taking sufficiently many iterates of $\mathcal{L}$. The example following Theorem 5.2 shows that this is not necessarily the case. □

### 5.3.3 Proof of Lemma 5.3.2 on a triangle in $\mathbb{R}^2$

We now prove Lemma 5.3.2 for a triangle in $\mathbb{R}^2$. This case is needed for the Jacobi–Perron transformation, and is easily extended to arbitrary star-like regions in $\mathbb{R}^r$. The proof is essentially a realisation of the geometric construction in [30].

*Proof of lemma:* Let $S \subset \mathbb{R}^2$ be a non–degenerate triangle. We prove the lemma for $0 \leq g \in \mathcal{C}^1$; the general case follows by an approximation argument. First, we establish some notation: let $\{L_y\}_{y \in \partial S}$ be the field of line segments from points $y \in \partial S$ to the central point of the triangle; see Figure 5.3.

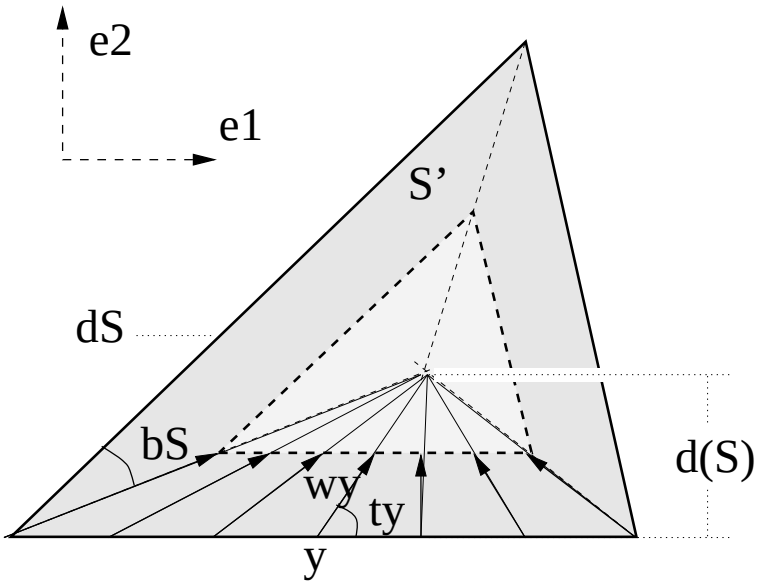



Figure 5.3: Construction of Lemma 5.3.2 on a triangle $S$ in $\mathbb{R}^2$.

For each $y \in \partial S$ there exists a vector $w_y$ such that

$$L_y = \{y + t\, w_y : t \in [0,1]\}.$$

Let

$$\delta(S) \triangleq \inf_{y \in \partial S} |w_y| \tag{5.5}$$

be the minimal length of a segment $L_y$. If the face of $\partial S$ containing $y$ has unit normal vector $v$, then $w_y$ meets $\partial S$ at an angle $\theta_y$ such that

$$\sin \theta_y = \frac{w_y \cdot v}{|w_y|}.$$

If $\beta(S)$ is half the minimal angle at a vertex of $S$, then

$$a(S) \triangleq \sin \beta(S) \leq |\sin \theta_y| \tag{5.6}$$

for each $y \in \partial S$. The constant $a(S)$ is the same as in [30].

Now, for each $y \in \partial S$,

$$tr^+_{\partial S} g(y) = \lim_{t \to 0} \frac{1}{t} \int_{s=0}^{t} g(z + s\, w_y)\, ds.$$

Letting $g_y(s) = g(y+s\,w_y)$, it follows from standard properties of one–dimensional variation var that for $c \in (0,1)$,

$$\begin{aligned} tr^+_{\partial S} g(y) &\leq \inf_{s\in[0,c]} g_y(s) + \mathrm{var}_{[0,c]}(g_y) \\ &\leq \frac{1}{c}\int_{s=0}^{c} g_y(s)\,ds + \int_{s=0}^{c} |g_y'|\,ds \\ &= \frac{1}{c}\int_{s=0}^{c} g_y(s)\,ds + \int_{s=0}^{c} |\nabla g \cdot w_y|\,ds \\ &\leq |w_y|\left(\frac{1}{c\delta(S)}\int_{s=0}^{c} g_y(s)\,ds + \int_{s=0}^{c} |\nabla g|\,ds\right), \end{aligned}$$

by (5.5). Therefore, by integrating over $\partial S$ and applying Fubini's Theorem, one obtains

$$\begin{aligned} \int_{\partial S} tr^+_{\partial S} g &\leq \frac{1}{c\delta(S)} \int_{\{(y,s)\in\partial S\times[0,c]\}} g(y+s\,w_y)\,|w_y|\,d\nu(y,s) \\ &\quad + \int_{\{(y,s)\in\partial S\times[0,c]\}} |\nabla g(y+s\,w_y)|\,|w_y|\,d\nu(y,s), \end{aligned} \tag{5.7}$$

where $d\nu(y,s) = dm_{2,\partial S\times[0,c]}(y,s) = dm_{1,[0,c]}(s) \times dm_{1,\partial S}(y)$.

Next, recall that $0 < c < 1$ and put

$$S' = S'(c) \triangleq \cup_{y\in\partial S}\{y + s\,w_y : s \in [0,c]\} \subset S,$$

and let $\Phi : \partial S \times [0,c] \to S'$ be the diffeomorphism defined by

$$\Phi(y,s) = y + s\,w_y.$$

To bound the rhs of (5.7) by integrals over $S' \subset S$, we must estimate the distortion of $\Phi$.

Suppose that the situation is as depicted in Figure 5.3, with $y \in \partial S$ contained a face aligned in the $e_1$ coordinate direction. Then

$$w_y = \delta_1\,(\cot\theta_y\,e_1 + e_2)$$

and

$$\Phi(y,s) = y + s\,\delta_1\,(\cot\theta_y\,e_1 + e_2)\,.$$

One can easily check that $\cot\theta_{y+\Delta y e_1} - \cot\theta_y = -\Delta y/\delta_1 + o(\Delta y)$ so that

$$\frac{\partial}{\partial y}\Phi(y,s) = \lim_{\Delta y\to 0}\frac{\Phi(y+\Delta y,s) - \Phi(y,s)}{\Delta y} = (1-s)e_1.$$

On the other hand,

$$\frac{\partial}{\partial s}\Phi(y,s) = \lim_{\Delta s\to 0}\frac{\Phi(y,s+\Delta s) - \Phi(y,s)}{\Delta s} = \delta_1\,(\cot\theta_y\,e_1 + e_2)\,.$$

Therefore,

$$\det D\Phi(y,s) = \begin{vmatrix} (1-s) & 0 \\ \delta_1\cot\theta_y & \delta_1 \end{vmatrix} = \delta_1(1-s) = \sin\theta_y\,|w_y|\,(1-s).$$

Then, for all $(y,s) \in \partial S \times [0,c]$,

$$\begin{aligned} |w_y|\, d\nu(y,s) &= \frac{|w_y|}{\det D\Phi(y,s)} dm_{2,S'}(\Phi(y,s)) \\ &= \frac{1}{1-s}\frac{1}{\sin\theta_y} dm_{2,S'}(\Phi(y,s)) \\ &\leq \frac{1}{1-c}\frac{1}{a(S)} dm_{2,S'}(\Phi(y,s)), \end{aligned} \tag{5.8}$$

by equation (5.6). Combining (5.7) and (5.8) we have (by the change of variables formula for integration):

$$\int_{\partial S} tr^+_{\partial S} g \leq \frac{1}{1-c}\frac{1}{a(S)} \int_{S'} |dg| + \frac{1}{c(1-c)}\frac{1}{a(S)\delta(S)} \int_{S'} g.$$

Since $S' \subset S$ and $c$ is arbitrary, the lemma follows. $\square$

## 5.4 Lower bound function for iterates under $\mathcal{L}$

Having established a Góra–Boyarsky inequality (generalised Lasota–Yorke inequality) for any uniformly expanding transformation with a finite range structure of non–degenerate sets, the next important step in the argument is to construct lower bound functions for iterates under $\mathcal{L}$ of initial densities.

Henceforth, we assume that $T$ has a *finite range structure* $\{U_1, \ldots, U_k\}$ , and that one of those sets $U_{i_0}$ is such that

$$U_{i_0} \subset T^n B_{\alpha^{(n)}}$$

for each monotonicity piece $B_{\alpha^{(n)}}$ of $T^n$. This is precisely the same assumption as was made in the one–dimensional case in Section 3.4. For the sake of simplicity, all calculations will be done $\mathbb{R}^2$; the methods are easily generalised.

Recall that in Chapter 3, only Lemmas 3.2.1 and 3.2.2 were required to construct a lower bound function for iterates under $\mathcal{L}$ of initial densities. Lemma 3.2.1 holds for *any* partition of a one–dimensional space into intervals, so is easily applied to the partition of $X$ into monotonicity intervals of $T^n$. Unfortunately, a precise analogue does not hold in multi–dimensions:

EXAMPLE (FAILURE OF LEMMA 3.2.1) Lemma 3.2.1 rests on the following fact:

$$\int_A |f - \operatorname*{ess\,inf}_A f| \leq \operatorname{diam}(A) \times V_{int(A)}(f), \tag{5.9}$$

for any interval $A \subset X$. To show that this result fails in more dimensions, let $A \subset \mathbb{R}^2$ be any connected open set, and let $B_\epsilon$ be a closed disk of radius $\epsilon$ which is contained in $int(A)$. Put

$$f = \chi_A - \chi_{B_\epsilon}.$$

Then, $\operatorname{ess\,inf}_A f = 0$ while

$$\int_A |f| = m(A) - m(B_\epsilon) = m(A) - \pi\epsilon^2 \quad \text{and} \quad V_{int(A)}(f) = m(\partial B_\epsilon) = 2\pi\epsilon.$$

Clearly, taking $\epsilon$ to be sufficiently small, (5.9) fails. □

We now consider how to adapt Lemma 3.2.1 to a multi–dimensional situation.

### 5.4.1 Discrete initial densities

In the proof of Proposition 3.1, an initial function $f$ was replaced by its infimal value on each monotonicity component of $T^n$. By using (5.9) and the bounded distortion properties of $T$, a lower bound for $\mathcal{L}^n f$ was easy to estimate. Because (5.9) may fail in a multi–dimensional setting, extra restrictions on $f$ and the geometry of the sets in the monotonicity partition are required. Therefore, we prove a generalised version of Proposition 3.1 in a *restricted* multi–dimensional setting. In the next subsection, the machinery from Sections 5.2 and 5.3 is exploited to generalise the proposition to an unrestricted case.

*Step I: A suitable class of initial densities.* For simplicity, suppose that $X = [0,1]^2$ and let $\eta$ be a regular rectangular partition, into squares with side–length $h$. We assume that $f \geq 0$ is constant on each rectangle in the partition $\eta$.

*Step II: A useful partition for constructing a lower bound function.* In Proposition 3.1, the contribution to $\mathcal{L}^n f$ from all the monotonicity pieces of $T^n$ is considered. However, in view of the above example, there is not a clear relationship between the amount of variation that $f$ has over a monotonicity piece $B$, and the amount of "mass" that as lost by replacing $f$ by its essential infimum over $B$ (c.f. the proof of Proposition 3.1). As such, we seek a partition such that the geometry of each partition element is sufficiently adapted to the function $f$ that a version of equation (5.9) can be derived.

With $h$ fixed, choose $n > 0$ such that

$$\operatorname{diam}(X)\, C\lambda^{-n} \leq h/5.$$

Let $\xi^{(n)}$ be the partition of $X$ into monotonicity components of $T^n$.

> *There exists a partition $\zeta$ of $X$, such that every $B \in \zeta$ can be written as a disjoint union of sets from $\xi^{(n)}$, and*
>
> $$\sum_{Y \in \zeta} \int_B |f - \operatorname{ess\,inf}_Y f| \leq \frac{8h}{5} V_X(f). \tag{5.10}$$

The proof of (5.10) and the details of the construction of $\zeta$ may be found in Appendix E. The constant $8h/5$ is surely far from optimal, but is sufficient for our purposes. In $r > 2$ dimensions, a similar construction can be done, but the constant $8/5$ would be larger.

**Proposition 5.2 (Lower bound functions in multi–dimensions)** *Let $T$ be as above, and let $f \geq 0$ be piecewise constant on a regular rectangular partition of $X \subset \mathbb{R}^2$ into cells of side–length $h$. Suppose that $n > 0$ is such that $diam(X)\, C\lambda^{-n} < h/5$. If*

$$\frac{8h}{5} V_X(f) < \|f\|$$

*then*

$$\mathcal{L}^n f \geq \left( \frac{1 - 8hV_X(f)/5\|f\|}{D} \right) \|f\|\, \chi_{U_{i_0}}.$$

*Proof:* Let $\zeta$ be the partition constructed above, and put

$$\tilde{f} = \sum_{Y \in \zeta} \operatorname*{ess\,inf}_Y f\, \chi_Y.$$

Then, by (5.10) and the fact that $f \geq \tilde{f}$,

$$\|\tilde{f}\| \geq \|f\| \left( 1 - \frac{8hV_X(f)}{5\|f\|} \right).$$

Next, since $\zeta$ is a refinement of $\xi^{(n)}$, we can write

$$\bigcup_{\alpha^{(n)}} \{B_{\alpha^{(n)}} \in \xi^{(n)}\} = \bigcup_{Y \in \zeta} \bigcup \left\{ B_{\alpha^{(n)}} \in \xi^{(n)} : B_{\alpha^{(n)}} \subset Y \right\},$$

and since every $x \in U_{i_0}$ has exactly one pre–image under $T^n$ in each monotonicity set $B_{\alpha^{(n)}}$,

$$\mathcal{L}^n f(x) = \sum_{Y \in \zeta} \sum_{\{B_{\alpha^{(n)}} \subset Y\}} \frac{f \circ T^{-n}_{\alpha^{(n)}}(x)}{|\det DT^n \circ T^{-n}_{\alpha^{(n)}}(x)|}.$$

Now, if $B_{\alpha^{(n)}} \subset Y \in \zeta$ then

$$f \circ T^{-n}_{\alpha^{(n)}}(x) \geq \operatorname*{ess\,inf}_Y f \triangleq \tilde{f}(Y).$$

Also, by the Mean Value Theorem and the fact that $m(X) = 1$,

$$m(B_{\alpha^{(n)}})\, \det DT^n \circ T^{-n}_{\alpha^{(n)}}(x) \leq D,$$

where $D$ is the distortion constant from (5.4).

Therefore, for $x \in U_{i_0}$,

$$\mathcal{L}^n f(x) \geq \sum_{Y \in \zeta} \operatorname*{ess\,inf}_Y f \sum_{\{B_{\alpha^{(n)}} \subset Y\}} \frac{m(B_{\alpha^{(n)}})}{D} = \frac{1}{D} \sum_{Y \in \zeta} \tilde{f}(Y) \times m(Y) = \frac{\|\tilde{f}\|}{D}.$$

Thus

$$\mathcal{L}^n f \geq \|f\| \left( \frac{1 - 8hV_X(f)/5\|f\|}{D} \right) \chi_{U_{i_0}}.$$

□

Proposition 5.2 is the crucial step in constructing lower bound functions for iterates under $\mathcal{L}$ of *arbitrary* initial functions from the cones $\mathcal{C}_M$.

REMARK 5.4.1. Notice that if $\epsilon$ is such that $\text{diam}(X)C\lambda^{-n_\epsilon} < \epsilon$, then a similar argument to the proof of the proposition implies that

$$\mathcal{L}^{n_\epsilon}(f) \geq \|f\| \left(1 - 8\epsilon V(f)\right) \chi_{U_{i_0}}.$$

In particular, since $V(f) \leq \frac{4\|f\|}{h}$ (c.f. Remark 5.2.1),

$$\mathcal{L}^{n_\epsilon}(f) \geq \|f\| \left(1 - \frac{32\epsilon}{h}\right) \chi_{U_{i_0}},$$

so that taking sufficiently many iterates of $\mathcal{L}$ will produce a lower bound function for *any* piecewise constant density on $\eta$. □

### 5.4.2 Initial densities from $\mathcal{C}_M$

We now prove a lower bound theorem for an arbitrary pair of $BV$ initial functions. The construction "boot–straps" off the case of piecewise constant initial functions, and comprises the final step in preparation for a multi–dimensional Mixing Theorem.

It is first convenient to establish some notation:

Let $T$ satisfy (5.1)—(5.4), and suppose that $T$ also has a non–degenerate finite range structure $\{U_1, \ldots, U_k\}$ with a set $U_{i_0} \subset T^n(B_{\alpha^{(n)}})$ for all $B_{\alpha^{(n)}} \in \xi^{(n)}$. Then, let $a, \delta$ be the constants in Lemma 5.3.2 such that for each $i$

$$\int_{\partial U_i} \left| tr^+_{\partial U_i} f \right| \leq \frac{2}{a} \left( \int_{int(U_i)} |df| + \frac{1}{\delta} \int_{U_i} |f| \right),$$

and for each $n \geq 0$ put

$$\sigma_n = C\lambda^{-n} \left(1 + \frac{2}{a}\right) \quad \text{and} \quad A_n = \left( CK \frac{1 - \lambda^{-n}}{1 - \lambda^{-1}} + \frac{2}{a\delta} \right).$$

Then, by Theorem 5.2, for any $f \in BV(X)$,

$$V_X(\mathcal{L}^n f) \leq \sigma_n V_X(f) + A_n \|f\|.$$

**Theorem 5.3 (Lower bound functions on $\mathbb{R}^2$)** *Let $M > 0$, and suppose that $f, g \in \mathcal{C}_M$ are such that $\|f\| = \|g\|$. If $n \geq 0$ satisfies*

$$diam(X)\, C\lambda^{-n} \leq \frac{m(U_{i_0})}{20M\sqrt{2}D},$$

*then there exists $\psi_{f,g} \geq 0$ such that*

$$\mathcal{L}^n f \geq \psi_{f,g} \quad and \quad \mathcal{L}^n g \geq \psi_{f,g}.$$

*Furthermore,*

$$\|f\|\frac{m(U_{i_0})}{2D} \geq \|\psi_{f,g}\| \geq \|f\|\frac{m(U_{i_0})}{4D}$$

*and*

$$V(\psi_{f,g}) \leq \left(\frac{m_1(\partial U_{i_0})}{2D} + 2\left(\sigma_n(1+\sqrt{2})M + A_n\frac{m(U_{i_0})}{4D}\right)\right)\|f\|.$$

*Proof:* The proof proceeds by approximating $f$ and $g$ by piecewise constant densities. Let $h > 0$ be such that

$$5\,\mathrm{diam}(X)\, C\lambda^{-n} \leq h \leq \frac{m(U_{i_0})}{4M\sqrt{2}D}, \tag{5.11}$$

and let $\eta$ be a regular rectangular partition of $X$ into squares of side–length $h$. Letting $\Pi_\eta$ be the corresponding projection onto piece–wise constant densities over $\eta$, we have that

$$\|\Pi_\eta f\| = \|f\| = \|g\| = \|\Pi_\eta g\|.$$

Moreover, by Corollary 5.2.3 and (5.11),

$$\|\Pi_\eta f - f\| \leq h\sqrt{2}M\|f\| \leq \frac{m(U_{i_0})}{4D}\|f\|, \quad \|\Pi_\eta g - g\| \leq \frac{m(U_{i_0})}{4D}\|g\|, \tag{5.12}$$

and

$$V_X(\Pi_\eta f) \leq \sqrt{2}M\|f\|, \quad V_X(\Pi_\eta g) \leq \sqrt{2}M\|g\|. \tag{5.13}$$

Therefore, by (5.11),

$$\frac{8hV_X(\Pi_\eta f)}{5\|f\|} \leq \frac{2}{5}\frac{m(U_{i_0})}{D} < \frac{1}{2}.$$

Then, by Proposition 5.2,

$$\mathcal{L}^n(\Pi_\eta f) \geq \frac{\|f\|}{2D}\chi_{U_{i_0}}, \quad \mathcal{L}^n(\Pi_\eta g) \geq \frac{\|g\|}{2D}\chi_{U_{i_0}}.$$

Now, because $\mathcal{L}$ is a linear operator,

$$\mathcal{L}^n f = \mathcal{L}^n(\Pi_\eta f) - \mathcal{L}^n(\Pi_\eta f - f) \geq \frac{\|f\|}{2D}\chi_{U_{i_0}} - (\mathcal{L}^n(\Pi_\eta f - f))^+.$$

Since the same relation holds with $f$ replaced by $g$, put

$$\psi_{f,g} = \max\left\{0, \frac{\|f\|}{2D}\chi_{U_{i_0}} - (\mathcal{L}^n(\Pi_\eta f - f))^+ - (\mathcal{L}^n(\Pi_\eta g - g))^+\right\}.$$

It is obvious that $\psi_{f,g}$ is a lower bound for both $\mathcal{L}^n f$ and $\mathcal{L}^n g$, the only remaining parts of the theorem to check are the norm and variation estimates:

For the norm, because $\int_X \mathcal{L}^n(\Pi_\eta f - f) = 0$,

$$\left\|(\mathcal{L}^n(\Pi_\eta f - f))^+\right\| = \frac{\|\mathcal{L}^n(\Pi_\eta f - f)\|}{2} \leq \frac{\|\Pi_\eta f - f\|}{2}.$$

Since the same is true with $g$ replacing $f$, it follows from (5.12) that

$$\|\psi_{f,g}\| \geq \|f\|\left(\frac{m(U_{i_0})}{2D} - 2\frac{m(U_{i_0})/4D}{2}\right).$$

Finally,

$$V_X(\psi_{f,g}) \leq m_1(\partial U_{i_0})\frac{\|f\|}{2D} + V_X(\mathcal{L}^n(\Pi_\eta f - f)) + V_X(\mathcal{L}^n(\Pi_\eta g - g)).$$

But the Góra–Boyarsky inequality implies that

$$V_X(\mathcal{L}^n(\Pi_\eta f - f)) \leq \sigma_n V_X(\Pi_\eta f - f) + A_n\|\Pi_\eta f - f\|.$$

However, since $V_X(\Pi_\eta f - f) \leq V_X(\Pi_\eta f) + V_X(f)$, the last estimate in the theorem follows from (5.12) and (5.13). □

REMARK. To generalise Theorem 5.3 to $\mathbb{R}^r$, $\sqrt{r}$ should replace $\sqrt{2}$, and a slightly larger value for $n$ would be required (c.f. the comments following (5.10)). □

Theorem 5.3 is a multi–dimensional version of Proposition 3.1, although the construction is considerably more complicated, drawing on arguments from the proofs of Theorem 3.3 and Proposition 3.3. The application to Ulam's method now follows exactly as before.

## 5.5 Rate of mixing and an error bound theorem

With important inequalities established in Theorems 5.2 and 5.3, the proofs of the two theorems—rate of mixing for $\mathcal{L}$ and approximation error for Ulam's method—involve almost the same manipulations as in Section 3.3. The error bound theorem for Ulam's method is Theorem 5.5 below, and a slight variation is stated as a corollary. The error bound for Ulam's method is far from optimal. However, for the sake of remaining parallel to the argument in Chapter 3, we present this less than optimal version.

Let the transformations $T$ be as in Theorem 5.3, and let $\sigma_n$ and $A_n$ be as defined above.

**Theorem 5.4 (Mixing Theorem for $\mathcal{L}$)** *Suppose that $M_* > 0$ and $n_* > 0$ are such that*

$$\frac{\sigma_{n_*} M_* + A_{n_*} + \frac{m_1(\partial U_{i_0})}{2D} + 2\left(\sigma_{n_*}(1+\sqrt{2})M_* + A_{n_*}\frac{m(U_{i_0})}{4D}\right)}{1 - \frac{m(U_{i_0})}{2D}} \leq M_*$$

*and*

$$diam(X)\, C\lambda^{-n_*} \leq \frac{m(U_{i_0})}{20M\sqrt{2}D}.$$

*Then for any function $f \in BV_0(X)$,*

$$\|\mathcal{L}^{n_*} f\|_{M_*} \leq (1 - m(U_{i_0})/4D)\|f\|_{M_*},$$

*where $\|\cdot\|_{M_*}$ is the difference norm of the difference cone $\Gamma_{M_*}$.*

*Proof:* Follows immediately from Theorems 1.1 and 5.3. □

REMARK. It is always possible to choose a pair $(n_*, M_*)$ to satisfy the hypothesis of the Mixing Theorem. Certainly $\sigma_n$ converges to zero exponentially fast as $n$ increases, and provided that

$$\sigma_n \left(\frac{3 + 2\sqrt{2}}{1 - m(U_{i_0})/2D}\right) < 1,$$

a sufficiently large value of $M_*$ will satisfy the first hypothesis of the theorem whenever $n_* \geq n$. Hence, after fixing $M_*$, one can simply increase $n$ to obtain an $n_*$ which satisfies the second condition. We will not discuss how to make an optimal choice! □

To obtain an error bound theorem for Ulam's method, let us start by establishing some notation. Let $n_0 > 0$ be such that

$$\sigma_{n_0}\sqrt{2} < 1,$$

where $\sigma_{n_0}$ corresponds to $n_0$ in Theorem 5.2. Let $\sigma_0 \triangleq \sigma_{n_0}$ and let $A_0$ be such that

$$V_X(\mathcal{L}^{n_0} f) \leq \sigma_0 V_X(f) + A_0\|f\|$$

for each $f \in BV(X)$. Let

$$M_0 \triangleq \frac{\sqrt{2}A_0}{1 - \sqrt{2}\sigma_0}$$

so that

$$\mathcal{L}^{n_0}\mathcal{C}_{M_0} \subset \mathcal{C}_{M_0/\sqrt{2}} \quad \text{and} \quad \Pi_\eta \circ \mathcal{L}^{n_0}\mathcal{C}_{M_0} \subset \mathcal{C}_{M_0} \tag{5.14}$$

whenever $\eta$ is a regular rectangular partition. Therefore, for every regular rectangular partition $\eta$,

$$\exists f_\eta \in \mathcal{C}_{M_0} \quad \text{such that} \quad \Pi_\eta \circ \mathcal{L}^{n_0} f_\eta = f_\eta, \quad \|f_\eta\| = 1.$$

Next, suppose that $(n_*, M_*)$ is a pair such that Theorem 5.4 holds, and assume without loss of generality that

$$n_* = k_* n_0$$

for some fixed integer $k_*$.

Finally, let $h(\eta)$ be the longest side–length of a regular rectangular partition $\eta$, put

$$M_\eta \triangleq \max\left\{\frac{4}{h(\eta)}, M_*\right\},$$

and let $k_\eta > 0$ be such that

$$\mathcal{L}^{k_\eta n_0}\mathcal{C}_{M_\eta} \subset \mathcal{C}_{M_*}$$

(c.f. Corollary 5.3.3).

**Theorem 5.5 (Error bound in multi–dimensions)** *Let the transformation $T$ be as above, let $n_0, M_0, k_*$ be as just described, let $\eta$ be a regular rectangular partition of $X$, with $h(\eta), k_\eta$ as above. If $f \geq 0$ is such that*

$$\mathcal{L}^{n_0} f = f \quad \textit{and} \quad \|f\| = 1,$$

*then*

$$\|f - f_\eta\| \leq \left(\frac{4D}{m(U_{i_0})}\, k_* + k_\eta\right) M_0 h(\eta)$$

*where $f_\eta \in \mathcal{C}_{M_0}$ is the Ulam approximate density (c.f. subsection 5.3.2).*

*Proof:* The proof of the theorem mirrors the proof of Theorem 3.2; some of the details are omitted here. First of all,

$$\begin{aligned}\|f - f_\eta\| &= \|(Id - \mathcal{L}^{n_0})^{-1}(f_\eta - \mathcal{L}^{n_0} f_\eta)\| \\ &\leq 2\|(Id + \mathcal{L}^{n_*} + \ldots)(Id + \mathcal{L}^{n_0} + \ldots + \mathcal{L}^{(k_*-1)n_0})\mathcal{L}^{k_\eta n_0}(f_\eta - \mathcal{L}^{n_0} f_\eta)\|_{M_*} \\ &\qquad + k_\eta \|f_\eta - \mathcal{L}^{n_0} f_\eta\|.\end{aligned}$$

By Corollary 5.2.3 and (5.14),

$$\|f_\eta - \mathcal{L}^{n_0} f_\eta\| = \|(\Pi_\eta - Id)\mathcal{L}^{n_0} f_\eta\| \leq h(\eta) M_0.$$

By the choice of $k_\eta$ and the definition of the norms $\|\cdot\|_M$ (c.f. Lemma 3.2.3 (5)),

$$\|\mathcal{L}^{k_\eta n_0}(f_\eta - \mathcal{L}^{n_0} f_\eta)\|_{M_*} \leq \|f_\eta - \mathcal{L}^{n_0} f_\eta\|_{M_\eta} \leq \frac{h(\eta) M_0}{2}.$$

The theorem now follows from Theorem 5.4. □

REMARK. If $\eta$ partitions $X$ into rectangles of side-length $1/N$, then $h(\eta) = 1/N$ and an $O(\log N)$ bound on $k_\eta$ follows easily from Theorem 5.2. Therefore, we have established an $O(\log N/N)$ error bound for Ulam's method. □

DEFINITION (**Ulam's method in multi–dimensions**) Suppose that $\eta(N)$ is a regular rectangular partition of $X$ of side–length $1/N$. Let $f_N$ be an invariant density of the operator

$$\Pi_{\eta(N)} \circ \mathcal{L}^{n_0},$$

and put

$$g_N = \frac{1}{n_0} \sum_{k=0}^{n_0-1} \Pi_{\eta(N)} \circ \mathcal{L}^k f_N.$$

Then $g_N$ will be called an *$N$th Ulam approximation for $T$.* □

**Corollary 5.5.1** *Suppose that $f$ is an invariant density for $T^{n_0}$, and that*

$$g = \frac{1}{n_0} \sum_{k=0}^{n_0-1} \mathcal{L}^k f.$$

*Then $g$ is an invariant density for $T$, and*

$$\|g - g_N\| \leq (c_T + k_{\eta(N)} + 1) M_0 \frac{1}{N} \leq O\left(\frac{\log N}{N}\right)$$

*whenever $g_N$ is an $N$th Ulam approximation for $T$ ($c_T$ is a constant depending only on $T$).*

Theorem 5.5 and Corollary 5.5.1 are the main results of the chapter. By comparing the development here with that of Chapter 3, it is clear that the only extra complications in the multi–dimensional situation are technical matters to do with the behaviour of variation. The fact that most of the effort in this chapter has been devoted to a constructive derivation of some key properties of multi–dimensional variation should not detract from the significance of the error bounds for Ulam's method in multi–dimensions. Although Keller [40] states that his results (which include a non–constructive $O(\log N/N)$ approximation rate for Ulam's method) extend to multi–dimensional cases, the results of this section appear to be unique: I am not aware of any other constructive attempt to obtain rigourous error bounds for multi–dimensional expanding transformations.

On the other hand, the error bounds from Theorem 5.5 may be quite bad (one may have to take an extremely fine grid to guarantee an accurate approximation). In practice, better bounds will be obtained by numerical computation like that described in Chapter 4: Theorem 5.5 *guarantees* an $O(\log N/N)$ approximation rate, but better constants in the $O(\cdot)$ notation can be got from Theorem 4.1. Furthermore, Proposition 4.3 can be easily generalised to a multi–dimensional situation, so that *a priori* control of the variation of an invariant density, together with *numerical* computation can also yield error bounds for Ulam's method. In Chapter 4, several (one–dimensional) examples were considered, and the numerical methods consistently produced the best error bounds.

Indeed, some of these were many orders of magnitude better than the corresponding error bounds from Chapter 2; it is reasonable to assume that this contrast would be reproduced—perhaps even amplified—in the multi–dimensional setting.

REMARK. Recall that the proof of mixing rates for $\mathcal{L}$ was proceeded by constructing lower bound functions which "boot–strapped" off the case of piecewise constant initial densities. For Ulam's method, this entails a slight inefficiency: a Ulam approximation is of the form

$$\Pi_\eta \circ \mathcal{L}^{n_0},$$

and each such operator acts only on piecewise constant densities. Now, if $h(\eta)$ is the shortest side–length of a regular rectangular partition $\eta$, Remark 5.4.1 implies that

$$\mathcal{L}^{k_\eta n_0} f \geq \|f\| \left(1 - 32h(\eta)\right) \chi_{U_{i_0}}$$

if $f$ is piecewise constant on $\eta$ and $\operatorname{diam}(X) C \lambda^{-k_\eta n_0} \leq (h(\eta))^2$. Then, as in the proof of Proposition 3.3, a lower bound can be obtained for $(\Pi_\eta \circ \mathcal{L}^{n_0})^{k_\eta} f$. An alternative error bound for Ulam's method can be constructed in this way (c.f. Theorem 2.2). This may give a smaller analytical bound on the approximation error, although it cannot out–perform the numerical method in Proposition 4.3. □

## 5.6 Example: the Jacobi–Perron transformation

In this section, we illustrate the main results of the chapter with the classical *Jacobi–Perron* transformation on $\mathbb{R}^2$. The quantitative results obtained are not particularly impressive, but the Jacobi–Perron transformation has long resisted this kind of analysis [36, 3].

The Jacobi–Perron transformation $T$ is one attempt to generalise Diophantine approximation by the one–dimensional continued fraction algorithm to a multi–dimensional setting. Hence, it is an example of a nontrivial multi–dimensional transformation, whose ergodic properties are of interest [36, 14, 62]. There are many interesting and deep mathematical questions surrounding the Jacobi–Perron transformation (and other algorithms for multi–dimensional Diophantine approximation), and the ideas in this chapter can be applied to make a small amount of progress towards one of these questions: the rate of Diophantine approximation of the Jacobi–Perron algorithm.

Now, there is a unique ergodic acim for the Jacobi–Perron transformation, but no exact formula exists [2]. Many other properties, both elementary and advanced, may be found in Schweiger's monograph [62] (which may be taken as a general reference for this section). A more up to date account is [14]. Exploiting our results, we show here that the invariant density of $T$ has bounded variation, and that it can be approximated by Ulam's method. This goes part way towards constructing a rigorous method for calculating the *rate of convergence*

of the associated *Jacobi–Perron algorithm* for simultaneous approximation of real numbers by rationals with a common denominator [36, 3]. The algorithm may be defined in arbitrary dimensions; our account is confined to $\mathbb{R}^2$.

## The transformation

Let $X = [0,1]^2$, and let

$$T(x_1, x_2) = \left( \left\{ \frac{x_2}{x_1} \right\}, \left\{ \frac{1}{x_1} \right\} \right),$$

where $\{\cdot\}$ denotes the fractional part of a number. Thus, there exist unique integers $k_2 \geq 1$ and $0 \leq k_1 \leq k_2$ such that

$$T(x_1, x_2) = \left( \frac{x_2}{x_1} - k_1, \frac{1}{x_1} - k_2 \right).$$

The pairs of integers $(k_1, k_2)$ index the monotonicity sets of the transformation $T$, we write this partition as $\{B_{(k_1,k_2)}\}$. The partition is illustrated in Figure 5.4.

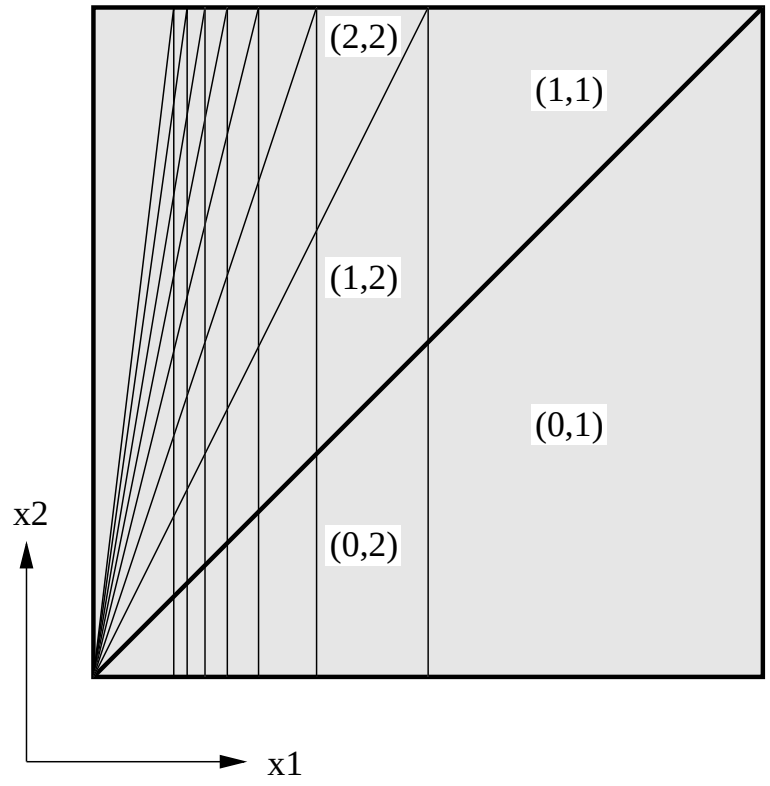



Figure 5.4: Partition of $[0,1]^2$ into monotonicity components of the Jacobi–Perron transformation.

The monotonicity components of $T$ consist of two kinds of pieces: trapezia (indexed by pairs $(k_1, k_2)$ with $k_1 < k_2$) and triangles (indexed by pairs $(k_1, k_2)$

with $k_1 = k_2$). Under one application of $T$, the trapezia map over the entire square, whilst the triangles map over the triangular region

$$S = \cup_{k_2 \geq k_1 \geq 1} B_{(k_1,k_2)} = \{(x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq x_2 \leq 1\}.$$

Therefore, $T$ has a *finite range structure* consisting of two subsets $\{S, [0,1]^2\}$. In fact, the transformation is Markov, and is conjugate to a subshift of finite type on the set of symbols

$$\{(k_1, k_2) : 0 \leq k_1 \leq k_2, 1 \leq k_2\}.$$

The transition

$$(k_1, k_2) \mapsto (k_1', k_2')$$

is *admissible* if either $k_1 < k_2$ or if $k_1 = k_2$ and $k_1' > 0$ (this is easily checked from the geometry of the map $T$). A string $\alpha^{(n)} = (k_1^{(0)}, k_2^{(0)}), \ldots, (k_1^{(n-1)}, k_2^{(n-1)})$ is called *admissible* if every transition $(k_1^{(j)}, k_2^{(j)}) \mapsto (k_1^{(j+1)}, k_2^{(j+1)})$ is admissible. It is obvious that $\alpha^{(n)}$ is an admissible string if and only if the cylinder set $T^{-n}_{\alpha^{(n)}}([0,1]^2)$ has non–zero measure. (To make the conjugacy with the subshift well–defined, all preimages of the vertical and integer slope lines which partition the monotonicity components of $T$ must be removed from the square (c.f. [62]). Since these together have Lebesgue measure zero, generic ergodic properties are unaffected.)

## Multi–dimensional Diophantine approximation

Before deriving the expansivity and distortion constants for $T$, we establish some notation, and describe the Jacobi–Perron algorithm for simultaneous approximation of two real numbers by rationals with a common denominator.

Fix $x = (x_1, x_2)$, and let $k(x) = (k_1(x), k_2(x))$ be the integers obtained by one application of the Jacobi–Perron transformation. Put

$$p_1 = 1, \quad q_1 = k_1(x), \quad r_1 = k_2(x)$$

and for each $n > 1$

$$\begin{bmatrix} p_{n+1} \\ q_{n+1} \\ r_{n+1} \end{bmatrix} = \begin{bmatrix} p_{n-2} & p_{n-1} & p_n \\ q_{n-2} & q_{n-1} & q_n \\ r_{n-2} & r_{n-1} & r_n \end{bmatrix} \begin{bmatrix} 1 \\ k_1(T^n(x)) \\ k_2(T^n(x)) \end{bmatrix},$$

where the initial conditions $p_{-1} = p_0 = q_0 = r_{-1} = 0$ and $q_{-1} = r_0 = 1$ have been adopted. Obviously, each of the sequences $\{p_n\}, \{q_n\}, \{r_n\}$ is increasing. Then the $n$th Jacobi–Perron approximation to $x$ is the pair

$$\left(\frac{p_n}{r_n}, \frac{q_n}{r_n}\right).$$

Each triple $(p_n, q_n, r_n)$ is in one to one correspondence with an admissible string $\alpha^{(n)} = (k_1(x), k_2(x)), \ldots, (k_1(T^{n-1}(x)), k_2(T^{n-1}(x)))$, so that each $x \in B_{\alpha^{(n)}} =$

$T^{-n}_{\alpha^{(n)}}[0,1]^2$ has the same $n$th Jacobi–Perron approximation. An open question [36, 3] is the *rate of approximation* in the algorithm. The quantity of interest is the value of the maximal constant $\gamma > 0$ such that

$$\max\left\{\left|\frac{p_n}{r_n} - x_1\right|, \left|\frac{q_n}{r_n} - x_2\right|\right\} \leq \frac{1}{{r_n}^{1+\gamma}}$$

for $m$–a.e. pair $(x_1, x_2)$. If $\gamma > 0$, then the Jacobi–Perron algorithm is said to *converge exponentially*. (In the case of one–dimensional Diophantine approximation by ordinary continued fractions, the analogous value of $\gamma$ is 1, so that the continued fraction algorithm converges exponentially). For a modified version of the Jacobi–Perron algorithm, Ito et. al. [36] prove exponential convergence, and numerical estimates of $\gamma$ (which suggest exponential convergence) are given in [3]. Indeed, a formula for $\gamma$ is known: Let $\lambda^{(2)} < \lambda^{(1)}$ be the almost everywhere Lyapounov exponents of $T$, and let $\lambda^{(0)}$ be defined as

$$\lambda^{(0)} = -\int \log x_1 \, d\mu(x_1, x_2)$$

where $\mu$ is the acim for $T$. Then

$$\lambda^{(1)} + \lambda^{(2)} = 3\lambda^{(0)} \quad \text{and} \quad \gamma + 1 = \frac{\lambda^{(2)}}{\lambda^{(0)}}.$$

Both these facts follow easily from the consideration of a certain matrix cocycle over the dynamical system induced by $T$; a particularly coherent account is in [3].

Therefore, calculating $\gamma$ requires the calculation of a Lyapounov exponent of $T$, and the number $\lambda^{(0)}$. Below, we show that Theorem 5.5 can be made to apply to the Jacobi–Perron transformation, implying a rigorous procedure for estimating the quantity $\lambda^{(0)}$.

## Expansivity constants for $T$

For quantitative estimates of the expansivity and distortion constants of $T$, the following relations are useful:

**Lemma 5.6.1 (Basic properties of Jacobi–Perron [62])** *Let $T$ be the Jacobi–Perron transformation, and for each $n > 0$ let $\alpha^{(n)}$ be an admissible string. If $(p_n, q_n, r_n)$ is the corresponding triple of integers, then*

1. *for each $x = (x_1, x_2)$,*

$$T^{-n}_{\alpha^{(n)}}(x) = \left(\frac{p_{n-2}x_1 + p_{n-1}x_2 + p_n}{r_{n-2}x_1 + r_{n-1}x_2 + r_n}, \frac{q_{n-2}x_1 + q_{n-1}x_2 + q_n}{r_{n-2}x_1 + r_{n-1}x_2 + r_n}\right),$$

2. *also*

$$|\det DT^{-n}_{\alpha^{(n)}}(x)| = \frac{1}{(r_{n-2}x_1 + r_{n-1}x_2 + r_n)^3},$$

3. *and*

$$\max\left\{\left|\frac{p_n}{r_n}-\frac{p_{n-1}}{r_{n-1}}\right|,\left|\frac{p_n}{r_n}-\frac{p_{n-2}}{r_{n-2}}\right|,\left|\frac{q_n}{r_n}-\frac{q_{n-1}}{r_{n-1}}\right|,\left|\frac{q_n}{r_n}-\frac{q_{n-2}}{r_{n-2}}\right|\right\}\leq\left(\frac{8}{9}\right)^{\lfloor n/2\rfloor}.$$

*Proof:* Part (1) is Lemma 1.2, Part (2) is Lemma 2.4 and Part (3) follows from the proof of Theorem 9.6 in [62]. □

We can now estimate the expansivity and distortion constants of the Jacobi–Perron transformation:

**Corollary 5.6.2** *For each inverse branch $T^{-n}_{\alpha^{(n)}}$ of $T^n$:*

1. *for any $x, y \in T^n(B_{\alpha^{(n)}})$,*

$$\left|\frac{\det DT^{-n}_{\alpha^{(n)}}(x)}{\det DT^{-n}_{\alpha^{(n)}}(y)}\right|\leq 3^3;$$

2. *while*

$$\left|DT^{-n}_{\alpha^{(n)}}\right|\leq 2\left(\frac{8}{9}\right)^{\lfloor n/2\rfloor},$$

3. *and for each $x$,*

$$\left|\nabla \det DT^{-n}_{\alpha^{(n)}}(x)\right|\leq 3\sqrt{2}\,|\det DT^{-n}_{\alpha^{(n)}}(x)|.$$

*Proof:* The first part follows immediately from Lemma 5.6.1 (2) since

$$0\leq x_1,x_2,y_1,y_2\leq 1 \qquad \text{and} \qquad r_{n-2}\leq r_{n-1}\leq r_n$$

(c.f. [62, Lemma 2.6]).

For Part (2), notice that

$$\frac{\partial}{\partial x_1}(T^{-n}_{\alpha^{(n)}}(x))_1=\frac{r_{n-2}}{r_{n-2}x_1+r_{n-1}x_2+r_n}\left(\frac{p_{n-2}}{r_{n-2}}-(T^{-n}_{\alpha^{(n)}}(x))_1\right).$$

By Lemma 5.6.1 (1), $(T^{-n}_{\alpha^{(n)}}(x))_1$ can be written as a convex combination of

$$\frac{p_{n-2}}{r_{n-2}},\frac{p_{n-1}}{r_{n-1}},\frac{p_n}{r_n}.$$

Therefore, since the sequence $\{r_n\}$ is increasing and $0\leq x_1,x_2\leq 1$,

$$\left|\frac{\partial}{\partial x_1}(T^{-n}_{\alpha^{(n)}}(x))_1\right|\leq\max\left\{\left|\frac{p_n}{r_n}-\frac{p_{n-1}}{r_{n-1}}\right|,\left|\frac{p_n}{r_n}-\frac{p_{n-2}}{r_{n-2}}\right|\right\}\leq\left(\frac{8}{9}\right)^{\lfloor n/2\rfloor}$$

by Lemma 5.6.1 (3). The other entries of $DT^{-n}_{\alpha^{(n)}}$ can be estimated in the same way, and this part of the corollary follows from standard properties of the matrix 2–norm.

The last part of the corollary also follows from Lemma 5.6.1 (1) by differentiation. □

The expansivity constants above are surely far from optimal. Rather than trying to improve them, the bounds have been lifted directly from Schweiger [62] to illustrate that the Jacobi–Perron transformation does indeed lie within the class we have considered.

## Application of results

With the above estimates, the main results of the chapter may be applied to the Jacobi–Perron transformation.

**Proposition 5.3** *Let $T$ be the Jacobi–Perron transformation on $\mathbb{R}^2$, and let $\mathcal{L}$ be its Perron–Frobenius operator. Then*

1. *For every $n > 0$ and $g \in BV$,*

$$\begin{aligned} V(\mathcal{L}^n g) \quad &\leq \quad 2\left(\frac{8}{9}\right)^{\lfloor n/2 \rfloor}\left(1 + 2\sqrt{4+2\sqrt{2}}\right) V(g) \\ &\qquad + \left(3\sqrt{2} + 4(2+\sqrt{2})\sqrt{4+2\sqrt{2}}\right) \|g\|. \end{aligned}$$

2. *$T$ has an invariant density $f$ with*

$$V(f) \leq \left(3\sqrt{2} + 4(2+\sqrt{2})\sqrt{4+2\sqrt{2}}\right) < 40.$$

3. *if $n_0 \geq 50$, then the Ulam's method with the operator*

$$\Pi_\eta \circ \mathcal{L}^{n_0}$$

*converges to an invariant density for $T$ as the regular rectangular partition $\eta$ is refined. The approximation error is at most $O(-h(\eta)\log h(\eta))$.*

*Proof:* The first part follows from Theorem 5.2 using the expansivity and distortion estimates from Corollary 5.6.2, together with the following estimates on the geometry of the finite range structure of $T$: Recall from above that each monotonicity piece of $T^n$ maps either over the whole square, or over the triangle

$$S = \{0 \leq x_1 \leq x_2 \leq 1\}.$$

From subsection 5.3.3, we can compute

$$\delta(S) = \frac{\sqrt{2}-1}{\sqrt{2}} \quad \text{and} \quad a(S) = \sin\frac{\pi}{8} = \frac{1}{\sqrt{4+2\sqrt{2}}}.$$

in the notation of subsection 5.3.3, we put $c = 1/2$, so that

$$\int_{\partial S} \left| tr^{+}_{\partial S} g \right| \leq \frac{2}{a(S)} \left( \int_{int(S)} |dg| + \frac{2}{\delta(S)} \int_{S} |g| \right)$$

for any $g \in BV(S)$. Since the geometry of the square $[0,1]^2$ is less restrictive, the same inequality holds for the square, and the Góra–Boyarsky inequality implies the result.

The second part follows from the first part and the uniqueness of the invariant density (c.f. Remark 5.3.4), and the final part follows from Theorem 5.5 because every monotonicity piece for $T^n$ maps over $S$. □

So far as I am aware, Proposition 5.3 is the first attempt to control the regularity of the invariant density for the Jacobi–Perron transformation. Moreover, the proof that Ulam's method will converge strongly to the invariant density is the first time that a numerical procedure has been shown to exhibit a good approximation to the invariant measure. The implications for the computation of the quantity $\lambda^{(0)}$ are obvious.

One obstacle in the implementation of Ulam's method in multi–dimensions is the volume of computation involved. In the one–dimensional examples presented in Chapter 3, it was seen that even for relatively regular expanding maps, good error bounds on the density approximations require quite fine partitions. For a grid of side–length $10^{-4}$, this corresponds to $10^8$ cells for a partition of the unit square in $\mathbb{R}^2$. If every row of the associated Ulam's matrix had only a handful of non–zero entries, then this level of precision would be implementable on a powerful work–station. However, the geometry of the Jacobi–Perron transformation is such that many of the cells in the grid map over the entire phase space in one iteration, requiring all entries of the corresponding row of the Ulam matrix to be non–zero. Therefore, the amount of storage required to represent the Ulam matrix may be very large.

On the other hand, the Monte–Carlo strategy discussed in Chapter 3 is no more difficult to implement in multi–dimensions than it is in the one–dimensional case: Recall that sample paths of the Markov Chain associated to Ulam's method can be calculated as a combination of exact iteration of $T$ and a small amount of noise to "spread" uniformly around the phase–space cell corresponding to the appropriate state of the Markov chain. Moreover, because the analysis of the mixing properties of $\mathcal{L}$ in multi–dimensions is in complete analogy with the one–dimensional situation, estimates of the sample path variance of the Monte–Carlo simulations can be done in the same way. Figure 5.5 depicts the results of such an experiment. No analysis is done of the sample path variance and approximation error; the figure is included as an illustration only:

The density depicted in Figure 5.5 is a histogram over $50 \times 50$ uniform cells of the the visits made by a sample path of length $10^7$ of the Markov chain corresponding to Ulam's method on the partition $\eta$ consisting of $10^4 \times 10^4$ equal sub-rectangles. The discontinuity along the diagonal $x_1 = x_2$ is because of the range structure of $T$. The undulations in the density shown in the

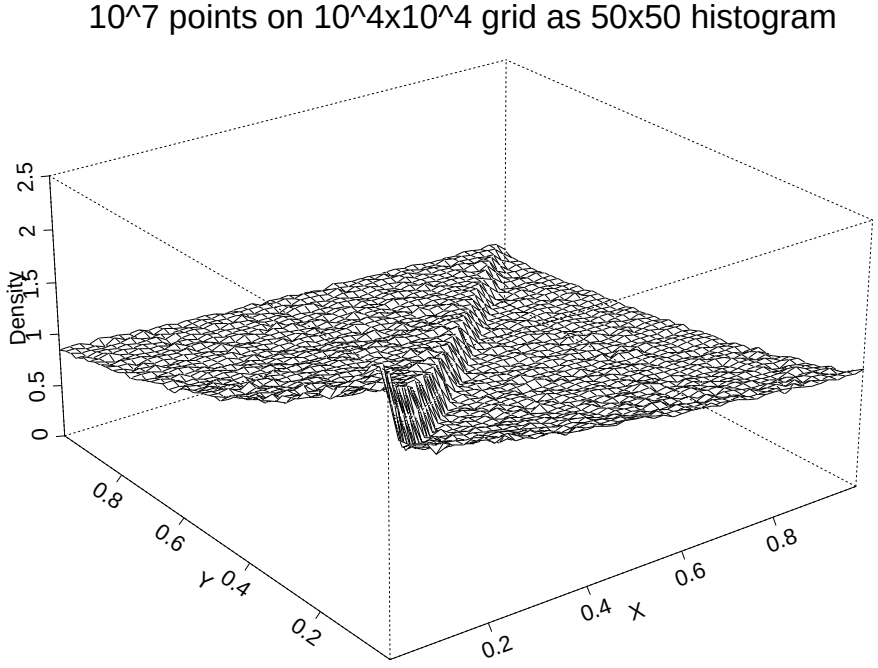



Figure 5.5: Histogram of Monte–Carlo simulation of invariant density for the $10000 \times 10000$ Ulam approximation for the Jacobi–Perron transformation.

picture are a natural feature of histograms of sample paths of Markov chains, and the smooth macroscopic structure is because the exact invariant measure has bounded variation (c.f. Proposition 5.3).

For the record, the value of $\lambda^{(0)}$ obtained by averaging $-\log x_1$ along the Monte–Carlo sample path is

$$\lambda^{(0)} \approx 1.201$$

(Baladi and Nogueira [3] estimate $\lambda^{(0)} \approx 1.206$ by an "exact" average of a long computed trajectory of $T$—rather than a Ulam approximation).

## 5.7 Final comments about Ulam's method in more dimensions

From a theoretical point of view, the bounded variation techniques developed in this chapter have provided very satisfactory convergence results for Ulam's method in multi–dimensions.

Substantial effort has been directed towards the explicit construction of $O(\log n/n)$ error bounds for the invariant density approximations, yet the results we have proved (Theorems 5.4 and 5.5) are still far from optimal. There is much that

could be done to refine the estimates in the proofs of error bounds, but such a project would be rather complicated and may not lead to significant improvements for general transformations.

At the end of Section 5.5, we noted that the numerical approach to error bounds (described in Chapter 4) is applicable to multi–dimensional transformations. Even though the computational requirements of this method may be very large for multi–dimensional transformations, it is likely that significantly better error bounds would be obtained than with the analytical models constructed above. In one sense, this suggests that the results of this chapter are mostly of theoretical interest. However, as well as providing a guarantee that the numerical scheme will converge, BV techniques can be employed to account for the effects of numerical errors[5]. Therefore, we consider our results to constitute a sound mathematical foundation for the rigorous analysis of Ulam's method.

Of course, the usefulness of these theoretical developments is limited by the extent to which they can be applied in practice. Throughout the thesis, the error bounds presented have been controlled by the *diameter* of the cells over which the Ulam approximation is constructed. In multi–dimensions, enforcing a small diameter requires a large number of cells, and the consequent memory requirements may become very demanding. While the necessary calculations (construction of transition matrices and iterated multiplication of vectors) are ideally suited to parallel implementation, Ulam's method remains computationally intensive.

As suggested several times in the thesis, *Monte–Carlo* simulations of the Markov chains corresponding to Ulam's method provide a computationally modest alternative to precise implementations. Our analysis of Ulam's method can be applied to guarantee Central Limit Theorem type convergence of such simulations (c.f. Section 3.5). For large Markov chain models, this convergence is rather slow, but (as illustrated in Figure 5.5 above) Monte–Carlo methods quickly provide a heuristic picture of the densities under approximation. However, rigorous statements about the accuracy of Monte–Carlo simulations are rather difficult to obtain [26]. One recent area of enquiry in Monte–Carlo methods is the *statistical estimation* of the rate of mixing for the underlying transition matrices [25]. Since the mixing properties of the Ulam approximations determine our error bounds (c.f. Proposition 4.3), this approach offers a promising direction for further enquiry.

Overall, we perceive the outlook for the approximation of ergodic objects in higher dimensions to be encouraging. In this thesis we have established several mathematical results about the approximation of invariant densities by Ulam's

[5]Our analysis above has consistently assumed that the transition probabilities $P_{ij}$ are computed exactly. The essential idea in accounting for numerical errors is that small errors in the computed transition probabilities result in a small diffusion of mass between adjacent cells of the partition $\eta$. In analogy with Proposition 5.1, this "averaging" phenomenon will not adversely affect the total variation of a BV function, so that arguments similar to subsection 5.3.2 will demonstrate the convergence of the "Ulam approximation" which is actually computed.

method, and illustrated these with numerical calculations. Some of these calculations[6] have exceeded the framework of our theory, and suggest possibilities for future work.

## 5.8 Summary of the chapter

This chapter has considered the approximation of invariant densities for expanding transformations in multi–dimensions.

Rather than the classical case of expanding endomorphisms of a boundaryless manifold [43], we have chosen to concentrate on the more general problem of piecewise $\mathcal{C}^2$ transformations on $\mathbb{R}^r$, with non–degenerate monotonicity pieces. The first proof of the existence of an acim for these transformations came relatively recently [30]. The fundamental tool is the *Góra–Boyarsky* inequality, which generalises the now classical *Lasota–Yorke* inequality to higher dimensions. Because of its importance to the constructions of this chapter, Section 5.3 was devoted to the Góra–Boyarsky inequality, which we generalised (Theorems 5.1 and 5.2). With this result, and a careful examination of the behaviour of *multi–dimensional variation* under the discretisation induced by the projections $\Pi_\eta$, we were able to prove Theorem 5.5, an error bound for Ulam's method in multi–dimensions. Just as in Chapter 3, bounds on the rate of mixing for the associated Perron–Frobenius operator were proved en route.

The techniques developed in the chapter were applied to the *Jacobi–Perron* transformation on $\mathbb{R}^2$. We proved that the invariant density (whose existence is well known [62], but explicit form is not [2]) has bounded variation, and that it can be approximated by Ulam's method.

However, the Jacobi–Perron transformation illustrated more: the relative infeasibility of exact implementation of Ulam's method in multi–dimensions. This motivated a discussion of the implications of our results for practical approximation of invariant densities.

[6] Especially the Monte–Carlo experiments of Sections 3.5 and 5.6, and Adapted Ulam method of Section 4.5.

# Appendix A

# Proofs of some Lemmas

*Proof of Lemma 3.2.1:* For each $I_\beta$, $V_{I_\beta}(f) \geq (\operatorname{ess\,sup}_{I_\beta} f) - (\operatorname{ess\,inf}_{I_\beta} f)$. Hence,

$$\int_{I_\beta} |f - \tilde{f}_\beta| \leq V_{I_\beta}(f)\, m(I_\beta) \leq \epsilon V_{I_\beta}(f),$$

so that

$$\int_I |f - \tilde{f}| = \sum_\beta \int_{I_\beta} |f - \tilde{f}_\beta| \leq \epsilon \sum_\beta V_{I_\beta}(f) \leq \epsilon V(f). \quad \square$$

*Proof of Lemma 3.2.2:* By the Mean Value Theorem, there exists $y \in I$ such that

$$|T^{n\prime} \circ T^{-n}_{\alpha^{(n)}}(y)| m(I_{\alpha^{(n)}}) = m(I) = 1.$$

The lemma follows from (3.4). $\square$

*Proof of Lemma 3.2.3:* The infimum in the definition of the norm is attained because the collection of functions with uniformly bounded variation is compact.

In the remainder of the proof, if we write $f = f^{(1)} - f^{(2)}$, it will be understood to mean that $f^{(1)}, f^{(2)} \in \mathcal{C}_a$ and $\|f^{(1)}\| = \|f^{(2)}\| = \|f\|_a$.

(2): If $a < b$ and $f \in \Gamma_a$, write $f = f^{(1)} - f^{(2)}$. Then $V_I(f^{(1)}) \leq a\|f^{(1)}\| < b\|f^{(1)}\|$ so $f^{(1)} \in \mathcal{C}_b$ and $\|f\|_b \leq \|f^{(1)}\| = \|f\|_a$.

(3): Let $f \in \Gamma_b$ and write $f = f^{(1)} - f^{(2)}$. Then, both $\mathcal{T}^n f^{(1)}, \mathcal{T}^n f^{(2)} \in \mathcal{T}^n \mathcal{C}_b \subset \mathcal{C}_a$ and $\mathcal{T}^n f = \mathcal{T}^n f^{(1)} - \mathcal{T}^n f^{(2)}$. Hence,

$$\|\mathcal{T}^n f\|_a \leq \|\mathcal{T}^n f^{(1)}\| \leq \|f^{(1)}\| = \|f\|_b.$$

(4): Write $f = f^{(1)} - f^{(2)}$. Then, for the norm estimate, $f^{(1)} \geq f^+$ and $f^{(2)} \geq f^-$, so that

$$\|f\| = 2\|f^+\| \leq 2\|f^{(1)}\| = 2\|f\|_a.$$

For the variation estimate, recall that both $f^{(1)}, f^{(2)} \in \mathcal{C}_a$, so that

$$V_I(f) \leq V_I(f^{(1)}) + V_I(f^{(2)}) \leq a\|f^{(1)}\| + a\|f^{(2)}\| = 2a\|f\|_a.$$

(5): Suppose first that

$$V(f) \le a\frac{\|f\|}{2} = a\|f^{\pm}\|.$$

Then, both $f^+, f^- \in \mathcal{C}_a$, so that $\|f\|_a = \|f^{\pm}\| = \|f\|/2$. Otherwise, put

$$g^{(1)} = f^+ + (V(f)/a - \|f\|/2) \quad \text{and} \quad g^{(2)} = f^- + (V(f)/a - \|f\|/2),$$

so that $g^{(1)} > f^+$ and $g^{(2)} > f^-$. Then

$$V(g^{(1)}) = V(f^+) \le V(f) = a\|g^{(1)}\|.$$

An identical estimate holds for $g^{(2)}$, so that $g^{(1)}, g^{(2)} \in \mathcal{C}_a$. Since $f = f^+ - f^- = g^{(1)} - g^{(2)}$, the definition of $\|\cdot\|_a$ implies that

$$\|f\| \le \|g^{(1)}\| = \frac{V(f)}{a}.$$

The bound on $\|f\|_a$ follows in either case.

(6): The inequalities follow immediately from Parts (4) and (5) above. The other part is obvious. □

*Proof of Lemma 3.1.1:* The first part is straight–forward, and the other part follows from Lemma 3.2.1 because

$$x \in B_i \Rightarrow \Pi_n f(x) = \int_{B_i} f/m(B_i) \in [\operatorname*{ess\,inf}_{B_i} f, \operatorname*{ess\,sup}_{B_i} f]. \ \square$$

*Proof of Lemma 3.5.1:* Let $f \in \mathcal{D}_n$. Because $f$ is constant on each subinterval $B_i = [\frac{i-1}{n}, \frac{i}{n})$,

$$f = \sum_{i=1}^{n} f_i \chi_{B_i}$$

where each $f_i \ge 0$ and $\|f\| = \sum_{i=1}^n f_i/n$. Also,

$$V_I(f) = f_1 + \sum_{i=1}^{n-1} |f_{i+1} - f_i| + f_n \le 2\sum_{i=1}^{n} f_i = 2n\|f\|,$$

so $f \in \mathcal{C}_{2n}$. The other part follows because if $g \in \Delta_n$, obviously $g^+, g^- \in \mathcal{D}_n$. □

# Appendix B

# Optimal choice of constants for BV error bounds

Recall from Chapter 3 that an essentially arbitrary choice of the constants $a_*, n_*, n_a$ was made to get the simply expressed error bounds in Theorem 3.2. We now describe how a more optimal choice can be made.

Suppose that the constants $\lambda, s, D$ are given for the map and that the order $n$ of the Ulam approximation is fixed. As in the proof of Theorem 3.2, take the function $h_n - \mathcal{L}h_n$ to be in the cone $\Gamma_{4n}$, and consider the estimates of the norm of

$$(Id - \mathcal{L})^{-1} : (BV_0, \|\cdot\|_{4n}) \to (BV_0, \|\cdot\|)$$

provided by Proposition 3.2.

Suppose that there exists $a_* \geq s\lambda/(\lambda - 1)$, $n_*, n_a > 0$ and constants $c_* \in (0,1), c' \in [0,1)$ such that for any $f \in BV_0$,

$$\begin{aligned} \|\mathcal{L}^{n_*} f\|_{a_*} &\leq (1 - c_*)\|f\|_{a_*}, \\ \|\mathcal{L}^{n_a} f\|_{a_*} &\leq (1 - c')\|f\|_{4n}. \end{aligned}$$

Then the argument in the proof of Proposition 3.2 implies that

$$\|(Id - \mathcal{L})^{-1} f\| \leq n_a \|f\| + 2\frac{n_*}{c_*}(1 - c')\|f\|_{4n}.$$

The error bound in Theorem 3.2 then follows. It is our objective to *minimise* the expression

$$n_a + \frac{n_*}{c_*}(1 - c') \tag{B.1}$$

subject to the above constraints, and available techniques.

From the proof of Theorem 3.1, any triple $(a_*, n_*, c_*)$ satisfying

$$\frac{\lambda^{-n_*} a_* + \frac{1-\lambda^{-n_*}}{1-\lambda^{-1}} s}{1 - c_*} \leq a_*, \tag{B.2}$$

where

$$c_* \leq (1 - \lambda^{-n_*} a_*)/D$$

will satisfy the first constraint above. Since it is desirable to have $c_*$ as large as possible, put

$$c_* = c_*(a_*) = (1 - \lambda^{-n_*} a_*)/D$$

for given choice of $a_*, n_*$.

Now, put

$$n_* = \left\lceil \frac{\log s\lambda/(\lambda - 1)}{\log \lambda} \right\rceil,$$

and choose $a_*(n_*)$ to be *minimal* such that

$$\frac{\lambda^{-n_*} a_* + \frac{1-\lambda^{-n_*}}{1-\lambda^{-1}} s}{1 - c_*(a_*)} \leq a_*$$

is satisfied (for fixed $n_*$, a larger value of $a_*$ yields a smaller value of $c_*$). Compare with the proof of Theorem 3.3. It is easy to see that this choice of $a_*$ and $c_*$ minimises $n_*/c_*(a_*(n_*))$, the first part of the product in (B.1) for a fixed choice of $n_*$. Because $n_*/c_*(n_*)$ is a *convex* function of $n_*$, it can be minimised by successively increasing $n_*$ by 1 until $n_*/c_*(n_*)$ stops decreasing. In this way, we minimise the first part of the product in (B.1), fixing $a_*$ in the process.

We can get a good estimate of the minimal value of (B.1) by fixing $n_*, a_*$ and choosing $n_a$ such that

$$\|\mathcal{L}^{n_a} f\|_{a_*} \leq \|f\|_{4n}.$$

Suppose that $f^{(1)}, f^{(2)} \in \Gamma_{4n}$ are such that

$$\|f^{(1)}\| = \|f^{(2)}\| = \|f\|_{4n}.$$

We need to choose a value of $n_a$ such that

$$V(\mathcal{L}^{n_a} f^{(1)}) \leq a_* \|f^{(1)}\|.$$

By iterated application of the Lasota–Yorke inequality,

$$n_a \geq \frac{\log\left(4n - s/(1 - \lambda^{-1})\right) / \left(a_* - s/(1 - \lambda^{-1})\right)}{\log \lambda}$$

will suffice. As above, once such an $n_a$ is chosen (say the minimal such $n_a$), then if $c' = c'(n_a)$ satisfies

$$\frac{\lambda^{-n_a} 4n + \frac{1-\lambda^{-n_a}}{1-\lambda^{-1}} s}{1 - c'} \leq a_*,$$

and

$$c' \leq (1 - \lambda^{-n_a} 4n)/D$$

then it follows that

$$\mathcal{L}^{n_a} f^{(1)} - c' \|f^{(1)}\| \in \mathcal{C}_{a_*}$$

and we deduce

$$\|\mathcal{L}^{n_a} f\|_{a*} \leq (1 - c')\|f\|_{4n}.$$

The minimal choice of $n_a$ and corresponding value of $c'$ satisfies the constraints listed at the beginning of the discussion, and provides a good estimate of the minimal value of

$$n_a + \frac{n_*}{c_*}(1 - c').$$

A slightly improved approximation may be obtained by successively incrementing $n_a$, until

$$n_a + \frac{n_*}{c_*}(1 - c'(n_a))$$

stops decreasing (of course, this may not result in any improvement).

Throughout the thesis, whenever we have claimed "an optimal choice of constants in Theorem 3.2", it is the above procedure that has been performed to optimise the error bounds.

# Appendix C

# More on 1–variation

If $f$ is a $BV$ function on a rectangle $A = (a_1, b_1) \times \ldots \times (a_r, b_r) \in \mathbb{R}^r$, we want to know that the 1–variation of $f$ can be computed by Fubini's theorem. That is:

$$\int_A |df|_1 = \sum_{i=1}^{r} \int_{A_i \cap \partial A} \left( \int_{x_i = a_i}^{b_i} \left| \frac{\partial f}{\partial x_i} \right| dx_i \right) dm_{r-1, A_i} \tag{C.1}$$

where $A_i = \{x \in \partial A : x_1 = a_i\}$ and the partial derivatives $\partial / \partial x_i$ are interpreted in a *distributional* sense. If $f$ is differentiable everywhere on $A$, then $|df|_1 = \sum_{i=1}^{r} |\partial f / \partial x_i| dm_r$ and the formula follows immediately from Fubini's Theorem. The only other case to consider is where $f$ has a discontinuity along a $\mathcal{C}^2$ surface $\Gamma$ of finite $(r-1)$–dimensional measure. Since $\Gamma$ and $f$ are at least $\mathcal{C}^1$, it suffices to establish the formula in the case where $f$ is piecewise constant and $\Gamma$ is linear:

*Sketch proof of (C.1) when $f$ is piecewise constant on a cube $A \subset \mathbb{R}^r$, with a jump along an $(r-1)$–dimensional plane $\Gamma$:* Let $\Gamma$ have unit normal $v$, and assume without loss of generality that

$$f(x_1, \ldots, x_r) = \begin{cases} 0 & \text{if } x \in A \text{ and } x \cdot v < 0, \\ 1 & \text{if } x \in A \text{ and } x \cdot v > 0 \end{cases}$$

where $\cdot$ denotes the usual scalar product.

Let $S = \Gamma \cap A$ be the portion of the surface $\Gamma$ which crosses $A$, and let $S_i$ be the orthogonal projection of $S$ onto the coordinate plane $A_i = \{x \in \mathbb{R}^r : x_i = a_i\}$. See Figure C.1. Since the vector $v$ is a unit normal to $\Gamma$,

$$m_{r-1, A_i}(S_i) = |v_i| \, m_{r-1, \Gamma}(S).$$

Now, for each $i = 1 \ldots r$, $z \in A_i \cap \partial A$ put

$$f_i(z, t) = f(z + te_i), \qquad t \in [0, b_i - a_i],$$

where $e_i$ is the unit vector in the $i$th coordinate direction. Then the one–dimensional distributional derivative $df_i$ is given by the formula

$$|df_i(z,t)| = \left| \frac{\partial f}{\partial x_i}(z + te_i) \right| dx_i = \begin{cases} 0 & \text{if } (z + te_i) \cdot v \neq 0, \\ d\delta & \text{if } (z + te_i) \cdot v = 0, \end{cases}$$

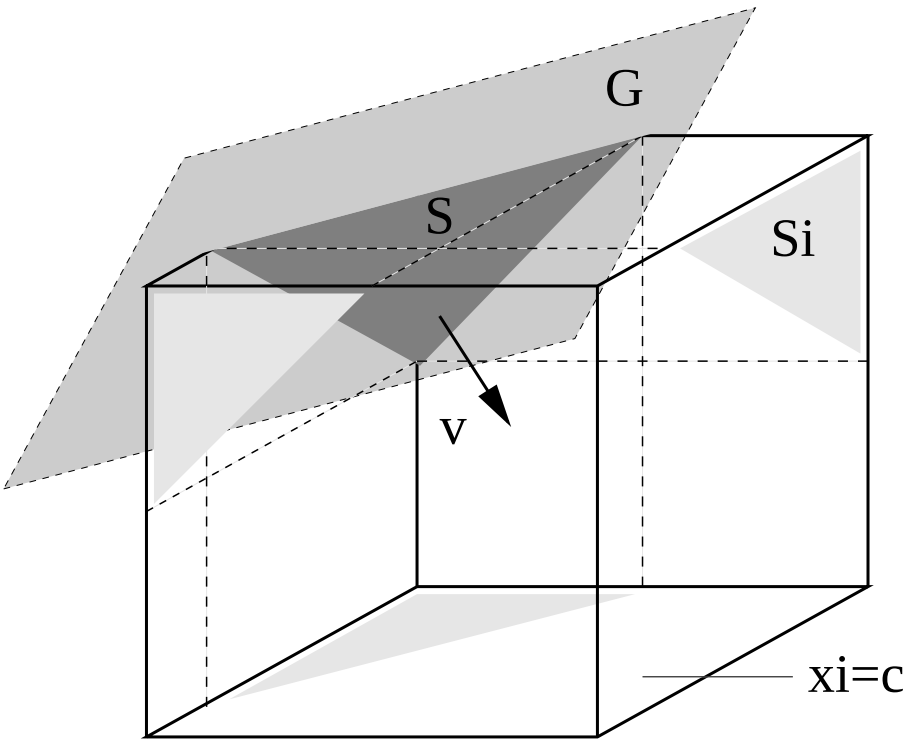



Figure C.1: Projection of surface area $S$ onto the faces of a coordinate box. A ray oriented in the $i$th coordinate direction only crosses $\Gamma$ inside the box if it also intersects $S_i$.

where $d\delta$ is the distribution of a point mass oriented in the $i$th coordinate direction. Hence,

$$\int_{t=0}^{b_i-a_i} |df_i(z,t)| = \begin{cases} 1 & \text{if } z \in S_i, \\ 0 & \text{otherwise.} \end{cases}$$

Integrating over $A_i \cap \partial A$, one obtains

$$\int_{z \in A_i \cap \partial A} \int_{t=0}^{b_i-a_i} |df_i(z,t)| dm_{r-1,A_i} = m_{r-1,A_i}(S_i).$$

Therefore,

$$\sum_{i=1}^{r} \int_{A_i \cap \partial A} \int_{a_i}^{b_i} \left| \frac{\partial f}{\partial x_i} \right| dx_i \, dm_{r-1,A_i} = \sum_{i=1}^{r} |v_i| \, m_{r-1,\Gamma}(S).$$

Formula (C.1) follows.

# Appendix D

# Proof of Lemma 5.2.2

All partial derivatives should be interpreted in a distributional sense, and the purpose of integrating over $int(B_{k_1 \ldots k_r})$ is to treat separately the contributions to the variation of any discontinuities of $f$ on the boundary. Figure D.1 illustrates the effect of $Q_i$ on the function $f$, and may be a helpful reference for the proof of the lemma. Without loss of generality, assume that $i = 1$.

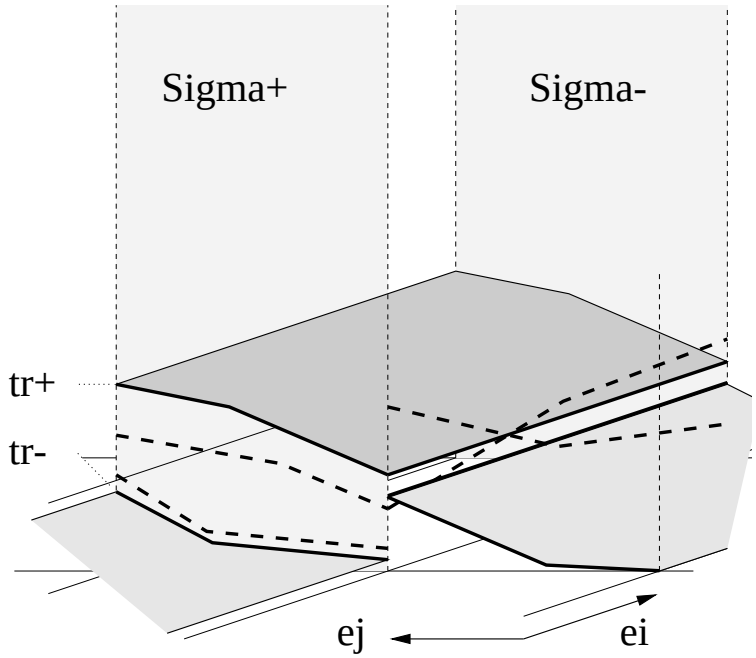



Figure D.1: The projected function $Q_i f$. The planes $\Gamma^{\pm}$ are as in Figure 5.2. The heavily dotted lines denote the trace of $f$ before application of the projection $Q_i$.

*Part (1):* First of all, let $u = a_1 + k_1(x)h_1$ so that

$$\left| \frac{\partial Q_1 f}{\partial x_j}(x) \right| \quad = \quad \left| \frac{\partial}{\partial x_j} \frac{1}{h_1} \int_{t=u}^{u+h_1} f(t, x_2, \ldots, x_r)\, dt \right|$$

$$\begin{aligned} &= \frac{1}{h_1}\left|\int_{t=u}^{u+h_1}\frac{\partial f}{\partial x_j}(t,x_2,\ldots,x_r)\,dt\right| \\ &\leq \frac{1}{h_1}\int_{t=u}^{u+h_1}\left|\frac{\partial f}{\partial x_j}(t,x_2,\ldots,x_r)\right|dt. \end{aligned}$$

Therefore, letting $\Gamma = \Gamma^i_{k_1\ldots k_r}$ and applying Fubini's Theorem (twice),

$$\begin{aligned} \int_{B_{k_1\ldots k_r}}\left|\frac{\partial Q_i f}{\partial x_j}\right| &= \int_{x_1=0}^{h_1}\left(\int_\Gamma\left|\frac{\partial Q_i f}{\partial x_j}(z+x_1e_1)\right|dm_{r-1,\Gamma}(z)\right)dx_1 \\ &\leq \frac{1}{h_1}\int_{x_1=0}^{h_1}\left(\int_\Gamma\int_{t=0}^{h_1}\left|\frac{\partial f}{\partial x_j}(z+te_1)\right|dt\,dm_{r-1,\Gamma}(z)\right)dx_1 \\ &= \frac{1}{h_1}\int_{x_1=0}^{h_1}\left(\int_{B_{k_1\ldots k_r}}\left|\frac{\partial f}{\partial x_j}\right|\right)dx_1 \\ &= \int_{B_{k_1\ldots k_r}}\left|\frac{\partial f}{\partial x_j}\right|. \end{aligned}$$

*Part (2):* For each $t\in[a_1+k_1h_1, a_1+(k_1+1)h_1)$ let

$$\Gamma_t = \Gamma^j\cap\{x : x_1 = t\}.$$

Then $\Gamma_t$ is $(r-2)$–dimensional and if $z\in\Gamma_t$ there exist unique $z'\in\Gamma_{(a_1+k_1h_1)}$ and $s = t-(a_1+k_1h_1)\in[0,h_1)$ such that

$$z = z'+se_1.$$

It is easy to see that $m_{r-1,\Gamma^j}$–a.e.

$$\left(tr^+_{\Gamma^j}Q_1f\right)(z'+se_1) = \frac{1}{h_1}\int_{s'=0}^{h_1}\left(tr^+_{\Gamma^j}f\right)(z'+s'e_1)\,ds',$$

so that the lhs of this expression depends only on $z'$, and not on $s$. Since the same holds for $tr^-_{\Gamma^j}Q_1f$,

$$\left(tr^+_{\Gamma^j}Q_1f - tr^-_{\Gamma^j}Q_1f\right)(z'+se_1) = \frac{1}{h_1}\int_{s'=0}^{h_1}\left(tr^+_{\Gamma^j}f - tr^-_{\Gamma^j}f\right)(z'+s'e_1)\,ds'.$$

Then, letting $u = a_1+k_1h_1$,

$$\begin{aligned} &\int_{\Gamma^j}\left|tr^+_{\Gamma^j}(Q_if) - tr^-_{\Gamma^j}(Q_if)\right|\,dm_{r-1,\Gamma^j} \\ &\quad= \int_{x_1=u}^{u+h_1}\int_{\Gamma_{x_1}}\left|tr^+_{\Gamma^j}(Q_if) - tr^-_{\Gamma^j}(Q_if)\right|(z)\,dm_{r-2,\Gamma_{x_1}}(z)\,dx_1 \\ &\quad= \frac{1}{h_1}\int_{x_1=u}^{u+h_1}\int_{\Gamma_{x_1}}\left|\int_{s'=0}^{h_1}\left(tr^+_{\Gamma^j}f - tr^-_{\Gamma^j}f\right)(z'+s'e_1)\,ds'\right|\,dm_{r-2,\Gamma_{x_1}}(z)\,dx_1 \\ &\quad\leq \frac{1}{h_1}\int_{x_1=u}^{u+h_1}\left(\int_{\Gamma_u}\int_{s'=0}^{h_1}\left|tr^+_{\Gamma^j}f - tr^-_{\Gamma^j}f\right|(z'+s'e_1)\,ds'\,dm_{r-2,\Gamma_u}(z')\right)dx_1 \\ &\quad= \frac{1}{h_1}\int_{x_1=u}^{u+h_1}\left(\int_{\Gamma^j}\left|tr^+_{\Gamma^j}f - tr^-_{\Gamma^j}f\right|\,dm_{r-1,\Gamma_j}\right)dx_1 \\ &\quad= \int_{\Gamma^j}\left|tr^+_{\Gamma^j}f - tr^-_{\Gamma^j}f\right|\,dm_{r-1,\Gamma^j}. \end{aligned}$$

*Part (3):* Fix $z \in \Gamma^-$ and consider the one–dimensional function

$$f_z(s) = f(z + se_1), \qquad s \in (0, h_1).$$

Extending $f_z$ such that

$$f_z(0) = \left(tr^+_{\Gamma^-} f\right)(z) \quad \text{and} \quad f_z(h_1) = \left(tr^-_{\Gamma^+} f\right)(z + h_1 e_1),$$

it is easy to see that for $m_{r-1,\gamma}$–a.e. $z$

$$\frac{1}{h_1} \int_{s=0}^{h_1} f_z(t)\, ds = Q_1 f(z) = \left(tr^+_{\Gamma^-} Q_1 f\right)(z) = \left(tr^-_{\Gamma^+} Q_1 f\right)(z + h_1 e_1)$$

and

$$|f_z(0) - Q_1 f(z)| + |f_z(h_1) - Q_1 f(z)| \leq V_{[0,h_1]}(f_z) = \int_{s=0}^{h_1} \left| \frac{\partial f}{\partial x_1}(z + se_1) \right| ds.$$

The result follows by integrating this point–wise inequality over the set $\Gamma^-$.

*Part (4):* In the same notation as above, for each $z \in \Gamma^1_{k_1 \ldots k_r}$ and $s \in [0, h_1]$,

$$\operatorname*{ess\,inf}_{[0,h_1]} f_z \leq f_z(s) \leq \operatorname*{ess\,sup}_{[0,h_1]} f_z.$$

Therefore,

$$|f(z + se_1) - Q_1 f(z + se_1)| \leq \operatorname*{ess\,sup}_{[0,h_1]} f_z - \operatorname*{ess\,inf}_{[0,h_1]} f_z \leq \int_{s=0}^{h_1} \left| \frac{\partial f}{\partial x_1} \right| ds.$$

The conclusion follows from Fubini's Theorem by integrating first over $s \in [0, h_1]$, and then over $\Gamma^1_{k_1 \ldots k_r}$.

# Appendix E

# Construction of the partition $\zeta$

Since $f$ is constant on each element of $\eta$, all of the variation of $f$ arises from the discontinuities on the boundary of elements of $\eta$. The idea is to group together elements of $\xi^{(n)}$ to obtain new sets which are approximately rectangular, and cover the boundaries between sets in $\eta$. The partition $\zeta$ is partially depicted in Figure E.1; the construction is described below.

Notice that the choice of $n$ guarantees that for any $B_{\alpha^{(n)}} \in \xi^{(n)}$, $\mathrm{diam}(B_{\alpha^{(n)}}) \leq h/5$.

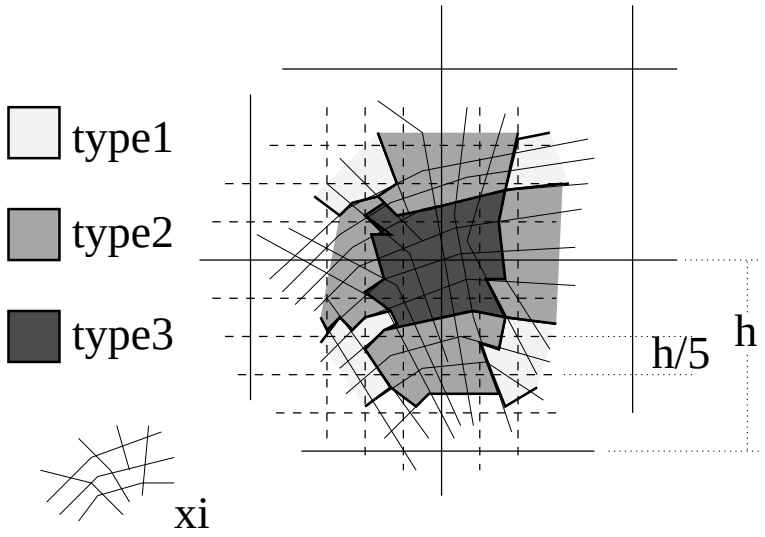



Figure E.1: The set $X$ is partition into three kinds of pieces: Types 1, 2, and 3. The partition $\eta$—a grid of side–length $h$—is depicted, as is the finer grid of side–length $h/5$, and the monotonicity partition $\xi^{(n)}$.

We work first on the interiors of cells from $\eta$. For $B \in \eta$, let $B' \subset B$ be the central rectangle of a further partition of $B$ into a $5 \times 5$ sub–grid. Let

$$Y_B = \{\cup B_{\alpha^{(n)}} : B_{\alpha^{(n)}} \in \xi^{(n)} \quad \text{and} \quad B' \cap B_{\alpha^{(n)}} \neq \emptyset\}.$$

Because each $B_{\alpha^{(n)}}$ has diameter less than or equal to $h/5$ while the distance from $\partial B$ to $B'$ is greater than equal to $2h/5$, the distance from $Y_B$ to $\partial B$ is *strictly greater* than $h/5$, so that

$$B' \subset Y_B \subset int(B).$$

Therefore, the function $f$ is constant on each $Y_B$. Having worked through all the cells $B \in \eta$, the collection $\{Y_B\}$ is the collection of *Type 1* pieces. The construction is illustrated in Figure E.2.

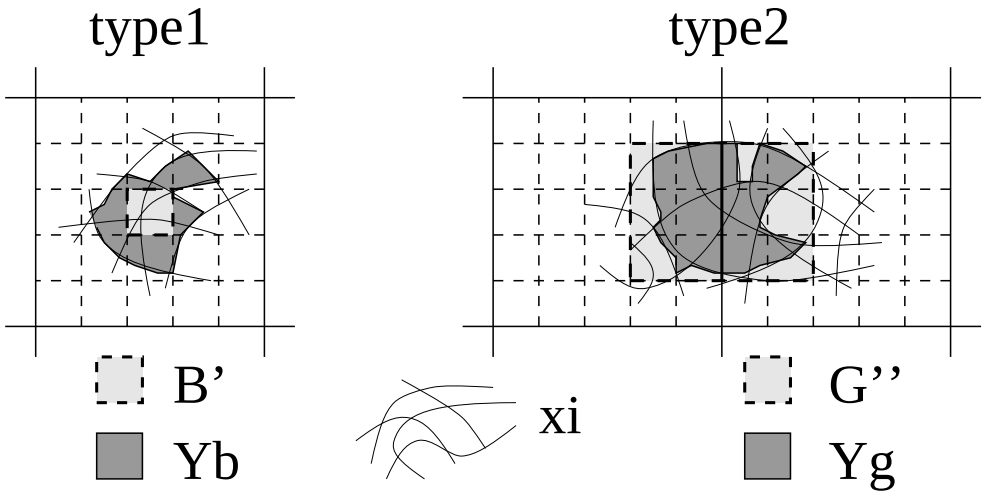



Figure E.2: Construction of Type 1 and 2 pieces of $\zeta$.

Next, we work on the boundaries of cells from $\eta$. Suppose that $\Gamma_0$ is the boundary of two adjacent cells from $\eta$. Let $\Gamma_0'$ be the rectangle obtained by taking the union of the two $h/5 \times h/5$ squares which border the middle 1/5th of $\Gamma_0$, and let $\Gamma_0''$ be the union of 12 $h/5 \times h/5$ squares in a $4 \times 3$ configuration which is bisected by the middle 3/5th of $\Gamma_0$; see Figure E.2. Then, put

$$Y_{\Gamma_0} = \{\cup B : B \in \xi^{(n)} \quad \text{and} \quad B \subset \Gamma_0''\}.$$

Then $Y_{\Gamma_0}$ is a *Type 2* piece. Now, let $\{\Gamma_i\}_{i=0}^N$ be the collection of all the boundaries $\Gamma_i$ of adjacent cells from $\eta$, and for each $i = 1 \ldots N$ put

$$Y_{\Gamma_i} = \{\cup B : B \in \xi^{(n)} \quad \text{and} \quad B \subset \Gamma_i''\} \setminus \{\cup_{j<i} Y_{\Gamma_j}\}.$$

Each Type 2 piece $Y_\Gamma$ satisfies

$$\frac{h}{5} \leq m_1(Y_\Gamma \cap \Gamma) \leq \frac{3h}{5} \quad \text{and} \quad \Gamma' \subset Y_\Gamma \subset \Gamma'', \tag{E.1}$$

because the diameter of each $B \in \xi^{(n)}$ is bounded by $h/5$.

Finally, each connected component of

$$X \setminus \{\{\cup_B Y_B\} \cup \{\cup_\Gamma Y_\Gamma\}\}$$

is a *Type 3* piece. By the construction of Type 1 and 2 pieces, each Type 3 piece contains a unique vertex $w$ at which 4 cells from $\eta$ meet. As such, a Type 3 piece will be denoted by $Y_w$. By construction, if $W'$ denotes the $2h/5 \times 2h/5$ square centred on $W$, and $W''$ denotes the $4h/5 \times 4h/5$ square centred on $w$, then

$$W' \subset Y_w \subset W''.$$

The situation is depicted in the left–hand diagram in Figure E.3.

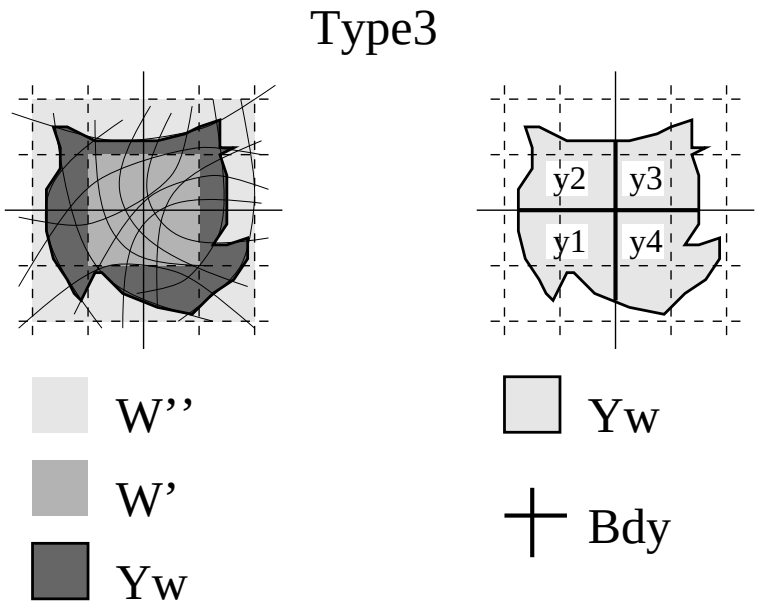



Figure E.3: The left–hand diagram depicts the construction of Type 3 pieces of $\zeta$. The right–hand diagram displays the partition of a Type 3 piece which is mentioned in the proof of (5.10).

The partition $\zeta$ consists of the collection of all Type 1, Type 2 and Type 3 pieces.

*Proof of (5.10):* To prove the inequality, it suffices to prove that for every $Y \in \zeta$,

$$\int_Y |f - \operatorname*{ess\,inf}_Y f| \leq \frac{8h}{5} V_{int(Y)}(f). \tag{E.2}$$

First of all, if $Y$ is a Type 1 piece, then $f$ is constant on $Y$ (by construction) so that both sides of (E.2) are zero.

Next, suppose that $Y = Y_\Gamma$ is a Type 2 piece. Then $f|_Y$ is piecewise constant, with a discontinuity along $\Gamma$. Denoting

$$J_\Gamma = \left| tr_\Gamma^+ f - tr_\Gamma^- f \right|,$$

it follows from (E.1) that

$$J_\Gamma \frac{h}{5} \leq V_{int(Y)}(f) \leq J_\Gamma \frac{3h}{5}$$

and

$$\int_Y |f - \operatorname*{ess\,inf}_Y f| \leq J_\Gamma \frac{6h^2}{25}.$$

Thus,

$$\int_Y |f - \operatorname*{ess\,inf}_Y f| \leq \frac{6h}{5} V_{int(Y)}(f).$$

Finally, suppose that $Y = Y_w$ is a Type 3 piece. Then partition $Y$ into quadrants $Y_1 \ldots Y_4$ as depicted in Figure E.3, and let

$$\Gamma_1 = \partial Y_1 \cap \partial Y_2, \ldots, \Gamma_4 = \partial Y_4 \cap \partial Y_1.$$

Denoting by $J_i = |tr^+_{\Gamma_i} f - tr^-_{\Gamma_i} f|$, it is easy to see that

$$\frac{h}{5} \sum J_i \leq V_{int(Y)}(f) \quad \text{and} \quad \frac{h^2}{25} \leq m(Y_i) \leq \frac{4h^2}{25}.$$

Similar to the Type 2 case, it follows that

$$\int_Y |f - \operatorname*{ess\,inf}_Y f| \leq (J_1 + J_3)\frac{4h^2}{25} + \max\{J_2, J_4\}\frac{8h^2}{25}.$$

Inequality (E.2) now follows easily. $\square$

# Appendix F

# A weak–$*$ metric

In this appendix, we recall a metric for the weak–$*$ topology on $\mathcal{M}(X)$, the set of all Borel probability measures on a compact set $X$. Some useful and interesting properties are stated; proofs are in [54, *In preparation*].

## The Lipschitz metric

The weak–$*$ metric is defined by integration against Lipschitz functions: Recall that a function $g : X \to \mathbb{R}$ is *Lipschitz continuous with Lipschitz constant* $L$ if

$$|g(x) - g(y)| \leq L\rho(x, y) \quad \forall x, y \in X;$$

we write $Lip(g) \leq L$. Let $\mathcal{C}(X)$ be the collection of all continuous real–valued functions on $X$, and denote

$$\mathcal{G}(X) = \{g \in \mathcal{C}(X) : Lip(g) \leq 1 \quad \text{and} \quad \|g\|_\infty \leq 1\}.$$

Then

DEFINITION (**Lipschitz metric**) Define a function $d^* : \mathcal{M}(X) \times \mathcal{M}(X) \to \mathbb{R}$ by

$$d^*(\mu, \nu) = \sup_{g \in \mathcal{G}(X)} \left| \int_X g \, d\mu - \int_X g \, d\nu \right|$$

for $\mu, \nu \in \mathcal{M}(X)$. $\square$

**Proposition F.1** *The function $d^*(\cdot, \cdot)$ is a metric on $\mathcal{M}(X)$, and the topology induced by $d^*(\cdot, \cdot)$ is the topology of weak–$*$ convergence.*

If $\mu, \nu \in \mathcal{M}(X)$ are absolutely continuous with respect to the Lebesgue measure $m$, then

$$d^*(\mu, \nu) \quad = \quad \sup_{g \in \mathcal{G}} \left| \int_X g \, d\mu - \int_X g \, d\nu \right|$$

$$\leq \sup_{\{h\in L^1:\|h\|=1\}} \left| \int_X h\,d\mu - \int_X h\,d\nu \right|$$
$$= \left| \int_X \phi(x)\,dx - \int_X \psi(x)\,dx \right| = \|\phi - \psi\|,$$

wher $\mu, \nu$ have densities $\phi, \psi$ respectively. Hence, if a function $g$ has $Lip(g) \leq L$,

$$\left| \int_X g\,d\mu - \int_X g\,d\nu \right| \leq L d^*(\mu,\nu) \leq L\|\phi - \psi\|. \tag{F.1}$$

## Computer representations of measures and estimation of functionals

Let $\eta$ be a measurable partition of $X$, and for any $\mu \in \mathcal{M}(X)$ recall the *histogram measure* $\mu_\eta$ defined according to the formula

$$\mu_\eta(A) = \sum_{B\in\eta} \frac{m(A\cap B)}{m(B)}\mu(B)$$

where $A \subset X$ is any Borel set, and $m(\cdot)$ is the Lebesgue measure on $X$. Then

**Proposition F.2** *Let $\mu, \nu$ be two probability measures on $X$. Let $\eta$ be a measurable partition of $X$, and let $\delta(\eta) = \max_{B\in\eta} diam(B) \triangleq diam(\eta)$. If $\mu_\xi = \nu_\xi$, then $d^*(\mu,\nu) \leq \delta(\eta)$.*

Therefore, the representation of arbitrary measures from $\mathcal{M}(X)$ by histogram measures is "continuous" in $d^*(\cdot,\cdot)$.

One useful and important corollary is:

**Proposition F.3** *Let $\mu_\eta$ be a histogram measure over $\eta$, and let $X_\eta \subset X$ be a collection of points, indexed by the elements of $\eta$ such that*

$$X_\eta = \{x_B : x_B \in B \quad \forall B \in \eta\}.$$

1. *If $g$ is a continuous function such that $Lip(g) \leq L$ then*

$$\left| \int_X g\,d\mu_\eta - \sum_{B\in\eta} \mu_\eta(B)\,g(x_B) \right| \leq L\,\delta(\eta).$$

2. *If $\mu_\eta$ has density $\phi_\eta$, and $\nu$ is an absolutely continuous measure with density $\psi$, then (using (F.1))*

$$\left| \int_X g\,d\nu - \sum_{B\in\eta} \mu_\eta(B)\,g(x_B) \right| \leq L(\|\psi - \phi_\eta\| + \delta(\eta)).$$

An application of Proposition F.3 to the estimation of Lyapounov exponents is given in Chapter 4.

## Convergence of stochastic discretisations

Recall from Chapter 1 the definition of a *stochastic discretisation* of a dynamical system $(X, T)$. The Lipschitz metric provides a straight–forward proof of the following:

**Proposition F.4** *Let $T : X \to X$ be continuous, and let $(\eta_n, \{\mathbb{P}_n(\cdot;\cdot)\})$ be a sequence of stochastic discretisations such that $\delta(\eta_n) \to 0$. If $\{\mu_n\}$ denotes the sequence of invariant measures for the stochastic discretisations, then any weak–$*$ limit point of the sequence $\{(\mu_n)_{\eta_n}\}$ is an invariant measure for $T$.*

In particular, Proposition F.4 implies that $\delta(\eta) \to 0$ is a sufficient condition for the weak–$*$ convergence of Ulam's method to *some* invariant measure. This result is neither surprising, nor new [23], but is very easy to prove with the Lipschitz metric.

## Historical Remarks

The Lipschitz metric has appeared several times in the literature [35, 63, 18]. In [35], it is used to proof the existence of an invariant measure for an *Iterated Function System* consisting of finitely many affine contractions applied in an i.i.d. fashion. Stark [63] proposed a Neural network model for computing the metric, and in [18], the authors were concerned with characterising the collection of invariant measures for a dynamical system which are amenable to numerical approximation. There, $d^*(\cdot,\cdot)$ was mentioned as an example of a weak–$*$ metric, although a stronger metric was used for their calculations. Finally, we note that many results about small random perturbations of dynamical systems [42] have easy proofs using the Lipschitz metric.

# Bibliography


[1] K Adl-Zarabi. Absolutely continuous invariant measures for piecewise expanding $\mathcal{C}^2$ transformations on $\mathbb{R}^n$ on domains with cusps on the boundaries. *Ergod. Th. & Dynam. Sys.*, 16:1–18, 1996.

[2] P Arnoux and A Noguiera. Mesures de Gauss pour des algorithmes do fractions continues multidimensionnelles. *Ann. scient. Éc. Norm. Sup.*, 26:645–664, 1993.

[3] V Baladi and A Nogueira. Lyapunov exponents for non–classical multi–dimensional continued fraction algorithms. *Nonlinearity*, 9:1529–1546, 1996.

[4] V Baladi and L-S Young. On the spectra of randomly perturbed expanding maps. *Comm. Math. Phys.*, 156:355–385, 1993. Erratum: *Comm. Math. Phys.* 166:219–220 (1994).

[5] C Beck and G Roepstorff. Effects of phase space discretisation on the long–time behaviour of dynamical systems. *Phys. D*, 25:173–180, 1987.

[6] M Blank and G Keller. Stochastic stability versus localisation in one–dimensional chaotic dynamical systems. *Nonlinearity*, 10:81–107, 1997.

[7] M L Blank. Ergodic properties of discretisations of dynamical systems. *Soviet Math. Dokl.*, 30:449–452, 1984.

[8] M L Blank. Stochastic properties of deterministic dynamical systems. *Sov. Sci. Rev. C: Math. Phys*, 6:243–271, 1987.

[9] M L Blank. Ergodic properties of a method of numerical simulation of a chaotic dynamical systems. *Math. Notes*, 45(3–4):267–273, 1989.

[10] M L Blank. Pathologies generated by round–off in dynamical systems. *Phys. D*, 78:93–114, 1994.

[11] C J Bose. On the existence and approximation of invariant densities for nonsingular transformations on $\mathbb{R}^d$. *J. Approx. Theory*, 79:260–270, 1994.

[12] A Boyarsky. On the significance of absolutely continuous invariant measures. *Phys. D*, 11:130–146, 1984.

[13] A Boyarsky and S Lou. Approximating measures invariant umder higher–dimensional chaotic tranformations. *J. Approx. Theory*, 65:231–244, 1991.

[14] A Broise. Fractions continues multidimensionnelles et lois stables. *Bull. Soc. Math. France*, 124:97–139, 1996.

[15] C Chiu, Q Du, and T-Y Li. Error estimates of the Markov finite approximation of the Perron–Frobenius operator. *Nonlinear Analysis*, 19(4):291–308, 1992.

[16] W de Melo and S van Strien. *One–dimensional dynamics*. Springer, 1993.

[17] P Diamond and P Kloeden. Spatial discretisation of mappings. *Comput. Math. Appl.*, 25(6):85–94, 1993.

[18] P Diamond, P Kloeden, and A Pokrovskii. Analysis of an algorithm for computing invariant measures. *Nonlinear Analysis*, 24(3):323–336, 1995.

[19] P Diamond, P Kloeden, and A Pokrovskii. Interval stochastic matrices: a combinatorial lemma and the computation of invariant measures of dynamical systems. *J. Dynamics and Differential Equations*, 7(2), 1995.

[20] Jiu Ding and T-Y Li. Markov finite approximation of Frobenius–Perron operator. *Nonlinear Analysis*, 17(8):759–772, 1991.

[21] Jiu Ding and Aihui Zhou. Finite approximations of Frobenius–Perron operators. A solution of Ulam's conjecture to multi–dimensional transformations. *Phys. D*, 92:61–68, 1996.

[22] P Ferrero and B Schmitt. Produits aléatoires d'opérateurs matrices de transfert. *Prob. Th. & Rel. Fields*, 79(2):227–248, 1988.

[23] G Froyland. Finite approximation of Sinai–Bowen–Ruelle measures for Anosov systems in two dimensions. *Random Comput. Dynam.*, 3(4):251–264, 1995.

[24] G Froyland. *Estimating physical invariant measures and space averages of dynamical systems indicators*. PhD thesis, University of Western Australia, 1996.

[25] S Garren and R Smith. Estimation of the second largest eigenvalue of a Markov transition matrix. Statistical Laboratory Research Report 95–18, University of Cambridge, 1995.

[26] W R Gilks, S Richardson, and D J Spiegelhalter (eds). *Markov Chain Monte Carlo in practice*. Chapman & Hall, 1996.

[27] E Giusti. *Minimal surfaces and functions of bounded variation*, volume 80 of *Monographs in Mathematics*. Birkhäuser, 1984.

[28] P Góra, 1997. Private communication.

[29] P Góra and A Boyarsky. Why computers like Lebesgue measure. *Comput. Math. Appl.*, 16(4):321–329, 1988.

[30] P Góra and A Boyarsky. Absolutely continuous invariant measures for piecewise expanding $\mathcal{C}^2$ transformations in $\mathbb{R}^n$. *Israel J. Math.*, 67(3):272–286, 1989.

[31] B Hunt. Estimating invariant measures and Lyapunov exponents. *Ergod. Th. & Dynam. Sys.*, 16:735–749, 1996.

[32] F Y Hunt. A Monte Carlo approach to the approximation of invariant measures. *Random Comput. Dynam.*, 2(1):111–133, 1994.

[33] F Y Hunt, 1997. Private communication.

[34] F Y Hunt. Finite precision representation of the Conley decomposition. Draft Manuscript, June 1997.

[35] J E Hutchinson. Fractals and self–similarity. *Indiana Univ. Math, J.*, 30:713–747, 1981.

[36] S Ito, M Keane, and M Ohtsuki. Almost everywhere exponential convergence of the modified Jacobi–Perron algorithm. *Ergod. Th. & Dynam. Sys.*, 13:319–334, 1993.

[37] S Ito and M Yuri. Number theoretical transformations with finite range structure and their ergodic properties. *Tokyo J. Math.*, 10(1):1–32, 1987.

[38] M Jablonski. On invariant measures for piecewise $\mathcal{C}^2$–transformations of the $n$–dimensional cube. *Ann. Polon. Math.*, 43:185–195, 1983.

[39] M S Keane, R D A Murray, and L-S Young. Computing invariant measures for expanding circle maps. *Nonlinearity*, 11:27–46, 1998.

[40] G Keller. Stochastic stability in some chaotic dynamical systems. *Monatsh. Mathematik*, 94:313–333, 1982.

[41] Y Kifer. *Ergodic theory of random transformations*, volume 10 of *Progress in probability and statistics*. Burkhäuser, 1986.

[42] Y Kifer. *Random perturbations of dynamical systems*, volume 16 of *Progress in probability and statistics*. Burkhäuser, 1988.

[43] K Krzyżewski and W Szlenk. On invariant measures for expanding differentiable mappings. *Studia Mathematica*, 33:83–92, 1969.

[44] A Lasota. Statistical stability of deterministic systemes. volume 1017 of *Lecture Notes in Mathematics*, pages 386–419. Springer, 1983.

[45] A Lasota and M C Mackey. Stochastic perturbation of dynamical systems: the weak convergence of measures. *J. Math. Anal. and Appl.*, 138:232–248, 1989.

[46] A Lasota and M C Mackey. *Chaos, Fractals and Noise: stochastic aspects of deterministic dynamics.* Springer, 2 edition, 1994.

[47] A Lasota and J A Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[48] T-Y Li. Finite approximation for the Perron–Frobenius operator. a solution to Ulam's conjecture. *Journal of approximation theory*, 17:177–186, 1976.

[49] C Liverani. Decay of correlations. *Annals of Mathematics*, 142:239–301, 1995.

[50] C Liverani. Decay of correlations for piecewise expanding maps. *J. Statistical Physics*, 78(3/4):1111–1129, 1995.

[51] R S Mackay. An extension of Zeeman's notion of structural stability to non–invertible maps. *Phys. D*, 52:246–253, 1991.

[52] S P Meyn and R L Tweedie. *Markov chains and stochastic stability.* Springer, 1993.

[53] R Murray. Approximation and concrete error bounds for acim density calculation for $C^{1+\alpha}$ expanding mappings. Statistical Laborartory Research Report. 96–22.

[54] R Murray. Weak convergence and the approximation of invariant measures. *In preparation.*

[55] R Murray. Approximation error for invariant density calculations. *J. Discrete & Continuous Dyn. Sys.*, 4:535–558, 1998.

[56] T Nowicki. Some dynamical properties of S–unimodal maps. *Fundamenta Mathematicae*, 142:45–57, 1993.

[57] T Nowicki and S van Strien. Invariant measures exist under a summability condition for unimodal maps. *Invent. Math.*, 105:123–136, 1991.

[58] J J Palmore and J L McCauley. Shadowing of computable chaotic orbits. *Phys. Lett. A*, 122:399–402, 1987.

[59] A Renyi. Representations of real numbers and their ergodic properties. *Acta Math. Akad. Sc. Hungar.*, 8:477–493, 1957.

[60] D Ruelle. Small random perturbations and the definition of attractors. *Comm. Math. Phys*, 82:137–151, 1981.

[61] M Rychlik. Regularity of the metric entropy for expanding maps. *Trans. Amer. Math. Soc.*, 315(2):833–847, 1989.

[62] F Schweiger. *The metrical theory of the Jacobi–Perron algorithm*, volume 334 of *LNM.* Springer, 1973.

[63] J Stark. A neural network to compute the hutchinson metric in fractal image processing. *IEEE Trans. on Neural Networks*, 2(1):156–158, 1991.

[64] S Ulam. *A collection of mathematical problems.* Interscience Publishers, 1960.

[65] P Walters. *An introduction to Ergodic Theory*, volume 79 of *Graduate Texts in Mathematics.* Springer, 1982.

[66] M Yuri. Invariant mesaures for certain multi–dimensional maps. *Nonlinearity*, 7:1093–1124, 1994.

[67] M Yuri. Decay of correlations for certain multi–dimensional maps. *Nonlinearity*, 9:1439–1461, 1996.

[68] E C Zeeman. Stability of dynamical systems. *Nonlinearity*, 1:115–155, 1988.

[69] W P Ziemer. *Weakly differentiable functions*, volume 120 of *GTM.* Springer, 1989.

# Index