Polynomial Decay of Correlations for Generalized Baker's Transformations via Anisotropic Banach Spaces Methods and Operator Renewal Theory

by

Seth William Chart
B.Sc. Mathematics, Montana State University, 2009
M.Sc. Mathematics, Montana State University, 2011

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Mathematics and Statistics

Polynomial Decay of Correlations for Generalized Baker's Transformations via Anisotropic Banach Spaces Methods and Operator Renewal Theory

by

Seth William Chart
B.Sc. Mathematics, Montana State University, 2009
M.Sc. Mathematics, Montana State University, 2011

Supervisory Committee

Dr. Christopher Bose, Supervisor
(Department of Mathematics and Statistics at the University of Victoria)

Dr. Anthony Quas, Departmental Member
(Department of Mathematics and Statistics at the University of Victoria)

Dr. Bruce Kapron, Outside Member
(Department of Computer Science at the University of Victoria)

Supervisory Committee

Dr. Christopher Bose, Supervisor
(Department of Mathematics and Statistics at the University of Victoria)

Dr. Anthony Quas, Departmental Member
(Department of Mathematics and Statistics at the University of Victoria)

Dr. Bruce Kapron, Outside Member
(Department of Computer Science at the University of Victoria)

ABSTRACT

We apply anisotropic Banach space methods together with operator renewal theory to obtain polynomial rates of decay of correlations for a class of generalized baker's transformations. The polynomial rates were proved for a smaller class of observables in [5] by fundamentally different methods. Our approach provides a direct analysis of the Frobenius-Perron operator associated to a generalized baker's transformation in contrast to [5] where decay rates are obtained for a factor map and lifted to the full map.

Contents

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

My Supervisor Christopher Bose for his patience, guidance, and support thoughout my Ph.D.

My Mother Leslie Chart for her unwavering support, for homeschooling me for twelve years, and for laying the educational foundation upon which this work is built.

My Father Thomas Chart for being my role model both as a scientist and as a man.

My Wife Salimah Ismail for staying positive through the frustrations and setbacks, for all of her support, and for chosing to be with me.

Chapter 1

Outline and Statement of Results

The main result of this thesis is Theorem 5.9.17. The theorem is a statement about the rate of decay of correlation for a class of maps of the unit square called generalized baker's transformations (GBTs) that were introduced in [6]. A particular class of GBTs were identified in [5] that are piecewise non-uniformly hyperbolic and possess lines of indifferent fixed points. Because orbits that pass near indifferent fixed points only escape a neighborhood of the fixed point intermittently we refer to these maps as *intermittent Baker's maps* (IBTs). In [5] the authors proved a decay of correlations result for IBTs with Hölder data. The method of proof is based on the Young tower method introduced in [24] and [25]. In this thesis we recover the rate of decay of correlations for IBTs obtained in [5] for more general spaces of functions by a completely different proof. We use anisotropic Banach space methods as introduced in [4] and further applied in [14, 3, 8] among others. We also apply operator renewal theory as introduced in [23] and refined in [13]. Others have recently applied anisotropic Banach space methods in concert with operator renewal theory, see for example [19, 18].

Each IBT has a parameter $\alpha > 0$ associated to it. Very roughly this parameter controls the intensity of intermittency caused by the indifferent fixed points. A larger value of α indicates that orbits will on average be trapped in the neighborhood of an indifferent fixed point for longer. We will also introduce spaces of functions $\mathfrak{L}_u$ and $\mathfrak{C}_a$. Both spaces contain the space of Lipschitz functions that are supported away from the indifferent fixed points. With this rough description of what is to come we state our main theorem in a preliminary form.

Theorem 1.0.1. *If* $B\colon [0,1]^2 \circlearrowleft$ *is the Intermittent Baker's Transformation with* *parameter* $\alpha > 0$ *and* λ *is Lebesgue measure restricted to* $[0,1]^2$*, then for all* $\eta \in \mathfrak{L}_u$ *and* $\psi \in \mathfrak{C}_a$ *we have,*

$$\left| \int \eta \, \psi \circ B^n \, d\lambda - \int \eta \, d\lambda \int \psi \, d\lambda \right| = O\left(\left(\frac{1}{n} \right)^{1/\alpha} \right).$$

The remainder of this thesis is is organized as follows. In Chapter 2 we provide an introduction to some of the methods that we will use. We apply the methods to some simple examples and collect what we believe to be a new variation on a familiar result in Theorem 3.2.3. In Chapter 4 we review the literature and previous results that inform this work. In Chapter 5 we prove the main theorem, this chapter is separated into nine sections each dealing with a particular aspect of the proof.

Chapter 2

Introduction and Background

In this thesis belongs to an area of mathematics often refered to as *smooth ergodic theory*. Roughly speaking smooth ergodic theory is the study differentiable maps using methods from ergodic theory. The central objects of ergodic theory are *measurable dynamical systems*. The remainder of this section is an overview of the rudimentary terminology and concepts pertaining to measurable dynamical systems that will be important through out the remainder of this thesis. We emphasise that after this section we will be operating within the realm of smooth ergodic theory where differentiability plays an important role.

To begin, a *dynamical system* is a function T from a set X back into itself. We refer to T as the *map*, X as the *state space*, and the pair as a dynamical system. We will use the notation $T\colon X \circlearrowleft$ to indicate that T is a map on X.

We think an element of the state space, which we call a *state*, as a description of some object at a particular time. A classic example is a particle in three dimensional space described by position and momentum. The position and momentum of a particle can be represented by a six dimensional real valued vector, therefore we can identify the state space of the particle as $\mathbb{R}^6$. A second example, that we will investigate more carefully in the next section, describes a real number by its fractional part. For example the fractional part of 2.0732 is 0.0732. The state space is $[0, 1)$.

A map provides a rule of evolution. An object in state $x \in X$ transitions to state $T(x)$. For example, consider a particle with mass m, position q, and momentum p described by $x = (q, p) \in \mathbb{R}^6$ moving in the absence of external forces for one unit of time. The particle transitions from x to $T(x) = (q + \frac{p}{m}, p)$. Since T is defined for any state we have a map $T\colon \mathbb{R}^6 \circlearrowleft$. We refer to $T(x)$ as *the state after one step of the dynamics*. After two steps of the dynamics the particle is in state $T(T(x))$. It is convenient to introduce a more compact notation for multiple steps of the dynamics, we define $T^0(x) = x$ and for $n \geq 1$, $T^n(x) = T\left(T^{n-1}(x)\right)$. If a particle is in state x then $T^n(x)$ is the state of the particle after n steps of the dynamics. The sequence

$$x, T(x), T^2(x), \cdots$$

is the *orbit* of a particle initially in state x, it is a chronological list of the states that the particle visits as it is carried from state to state. Dynamical systems are models of discrete time deterministic evolution.

Given a dynamical system we can add the notion of an *observable*, which is a function $\psi\colon X \to \mathbb{R}$. An observable represents a quantifiable property of states. For example suppose that a particle moves through an empty region of space and is bathed in the light of a distant star. This intensity can be represented by an observable $\psi\colon \mathbb{R}^6 \to \mathbb{R}$. It is natural to consider the sequence of light intensities that a particle witnesses as it passes through space. If the particle begins its journey in the state x then the sequence is just the value of ψ at each point along the orbit of x, that is

$$\psi(x), \psi\left(T(x)\right), \psi\left(T^2(x)\right), \cdots$$

This leads us to consider the *Koompan operator* acting on observables defined by

$$\psi \mapsto \psi \circ T$$

where $\circ$ denotes composition of functions. Given an observable ψ the Koopman operator produces the observable $\psi \circ T$ which associates to each state the value of ψ witnessed after one step of the dynamics. We will often consider $\psi \circ T^k$ for $k \geq 1$ which can be interpreted similarly.

It is often desirable to consider ensembles of states, in other words subsets E of the state space X. For example $E \subset \mathbb{R}^6$ could be the set of all states that are positioned less then one unit away from the origin. It is convenient to define $T^{-1}E = \{x \in X : T(x) \in E\}$. Notice that this defines a map on subsets of X by

$E \mapsto T^{-1}E$ which we call the *pre-image operator*. The set $T^{-1}E$ is the set of all states that land in E after one step of the dynamics. We will often consider $T^{-k}E$ for $k \geq 0$ which can be interpreted similarly. The familiar operations of union, intersection, and complement are the logical connectives *and*, *or*, and *not* as they apply to set membership. For any state space X the power-set $\wp(X) = \{E \subseteq X\}$ is a collection of sets that is closed under the operations of union, intersection, and complementation. Further if $T\colon X \circlearrowleft$ then $\wp(X)$ is closed under the pre-image operator. A *σ-algebra* on a state space X is a collection $\mathcal{X}$ of subsets of X that contains X as one of its elements and is closed under any countable sequence of applications of the operations of union, intersection, and complement. A σ-algebra on a state space may be strictly smaller then the power set, and for all of our applications we will need for this to be the case to avoid measure theoretic pathologies. Given a state space X with a σ-algebra $\mathcal{X}$ we say that a map $T\colon X \circlearrowleft$ is *measurable with respect to* $\mathcal{X}$ if for every E in $\mathcal{X}$ the pre-image $T^{-1}E$ is also in $\mathcal{X}$. A state space X together with a σ-algebra $\mathcal{X}$ and a map $T\colon X \circlearrowleft$ that is measurable with respect to $\mathcal{X}$ form a *measurable dynamical system.*

The final ingredient in our framework is a representation of the objects that are moved from state to state by the dynamics. A very flexible perspective is to consider a measure on the σ-algebra of a measurable dynamical system as such a representation. If one imagines a particle that is initially at state $x \in X$, then this particle can be represented by the measure δ_x defined for any $E \in \mathcal{X}$ by

$$\delta_x(E) = \left\{ \begin{array}{ll} 1 & \text{if } x \in E \\ 0 & \text{if } x \notin E \end{array} \right\}.$$

From this definition it is easy to check that $\delta_{T(x)}(E) = \delta_x(T^{-1}E)$. This suggests the following more general definition. Given a measurable dynamical system consisting of X, $\mathcal{X}$, and T, and a measure μ on $\mathcal{X}$ define $T_*\mu$ by $T_*\mu(E) = \mu(T^{-1}E)$. Notice that this is a well defined measure on $\mathcal{X}$ since T is measurable and therefore for any $E \in \mathcal{X}$ we have $T^{-1}E \in \mathcal{X}$. We refer to the map $\mu \mapsto T_*\mu$ on measures as the *transfer operator*. There are many useful ways to interpret a measure μ on the state space. For example we can think of μ as representing a collection of particles distributed among the states in X so that for any $E \in \mathcal{X}$, the quantity $\mu(E)$ is the proportion of the collection of particles that are in states contained in E. If $\mu(X) = 1$, then one can view $\mu(E)$ as representing the probability that a particle is in a state contained in E. This perspective will be investigated for a simple example in the next section.

2.1 Doubling Map

In this section we will consider a very simple piecewise smooth map of the unit interval $[0,1]$. The *doubling map* $T\colon [0,1] \circlearrowleft$ is defined by $T(x) = 2x \mod 1$. When $[0,1]$ is equipped with the Borel σ-algebra this map can be viewed as a measurable dynamical system. This simple map of the unit interval exhibits complicated behavior.

To see what we mean consider the question of predicting $T^{10}(x)$ given some initial point $x \in [0,1]$. Suppose that $x = 0.0732$, we easily compute

$$T^{10}(0.0732) = 2^{10}(0.0732) \mod 1 = 0.9568.$$

Therefore we would predict that $T^{10}(x) = 0.9568$. We think of T as a model of some physical process and the quantity 0.0732 as an imperfect measurement of x. Perhaps the true value of x is 0.07326, then the true value of $T^{10}(x)$ is

$$T^{10}(0.07326) = 2^{10}(0.07326) \mod 1 = 0.01824.$$

We see that a small error in the initial data has resulted in a large error in our prediction.

This observation illustrates a property called *sensitive dependence on initial conditions*, which is by no means unique to the map T. It can be observed in models of physical phenomena such as weather and in economic models. The doubling map indicates that even very simple models can suffer from sensitive dependence on initial conditions.

Any measurement of a physical phenomena has some associated amount of uncertainty. When we study systems with sensitive dependence on initial conditions it would be useful to incorporate uncertainty into our predictions. One way to manage the uncertainty in our measurement $x = 0.0732$ is to reinterpret the meaning of the measurement. Rather than treating 0.0732 as the true value of x, we could estimate that our measurement is only precise up to an error of ± 0.05. We could then take a probabilistic perspective and say that given our measurement, the true value of x could be any point in the interval $[0.0232, 0.1232]$ with equal probability. In other words, x is an $[0,1]$ valued random variable and the probability density function (density) of x is $\eta_0 \colon [0,1] \to [0,\infty)$ defined by

$$\eta_0(t) = \begin{cases} \frac{1}{0.1}, & t \in [0.0232, 0.1232] \\ 0, & \text{otherwise.} \end{cases} \tag{2.1.1}$$

The question of predicting $T^{10}(x)$ given x changes slightly. We would like to know the probability that $T^{10}(x) = y$, given that x is a random variable with some density η. The following elementary probability calculation gives us the cumulative distribution function for $T(x)$ given that x is distributed according to some density η

$$\begin{aligned} P\left(T(x) \le y\right) &= P\left(T(x) \le y, x \in [0,1/2)\right) + P\left(T(x) \le y, x \in [1/2,1]\right) \\ &= P\left(x \le T^{-1}(y), x \in [0,1/2)\right) + P\left(x \le T^{-1}(y), x \in [1/2,1]\right) \\ &= \int_0^{y/2} \eta(t)\, dt + \int_{1/2}^{(y+1)/2} \eta(t)\, dt \end{aligned}$$

Differentiation yields the density that represents the location of $T(x)$ given that the density representing the location of x is η

$$\tfrac{1}{2}\left(\eta\left(\tfrac{t}{2}\right) + \eta\left(\tfrac{t+1}{2}\right)\right).$$

This calculation allows us to translate the map T on points into a map $\mathcal{P}$ on densities defined by

$$\left(\mathcal{P}\eta\right)(t) := \tfrac{1}{2}\left(\eta\left(\tfrac{t}{2}\right) + \eta\left(\tfrac{t+1}{2}\right)\right). \tag{2.1.2}$$

The map $\mathcal{P}$ is called the *Frobenius-Perron operator* associated to T. In fig. 2.1 we have plotted $\mathcal{P}^n\eta_0$ for $n = 0, 1, 2, 4, 7, 10$ where η_0 is the density defined in eq. (2.1.1). From this diagram we see that as n increases $\mathcal{P}^n\eta_0$ seems to approach the uniform density. Although the initial measurement of the position of x was precise (the assumed error was ± 0.05) the position of $T^{10}(x)$ is essentially equally likely to be any point in the interval $[0,1]$. Hopefully the reader will not be surprised to hear that improving the precision of the measurement of the initial position of x does very little to improve the situation. Regardless of the precision of initial data after some number of steps n the position of $T^n(x)$ will be essentially random and distributed uniformly on $[0,1]$. The

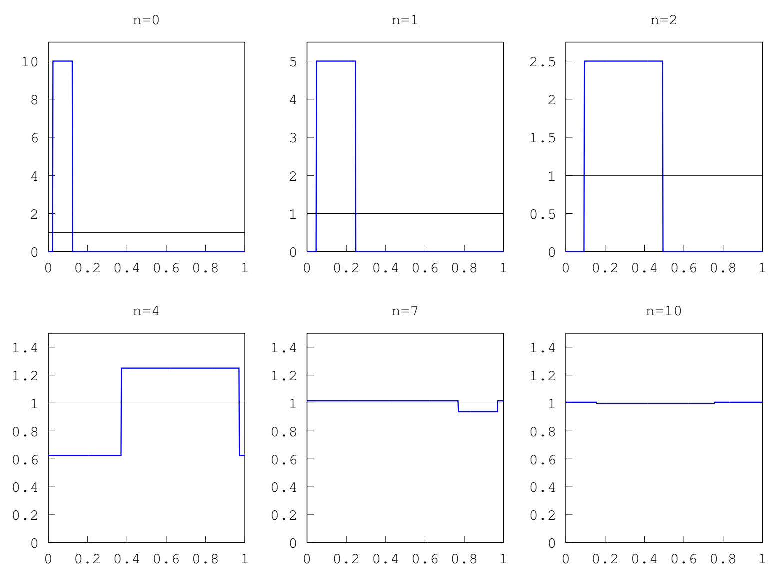

Figure 2.1: Plots of $\mathcal{P}^n\eta_0$ for $n = 0, 1, 2, 4, 7, 10$ with guide lines $y = 1$.

uniform distribution is the best description of the position of $T^n(x)$ for large n. This does not give us a prediction like our first calculation but it is in many ways more informative since it incorporates the unavoidable error in measurement of initial data.

The remainder of this thesis is in part aimed at describing methods for identifying limiting densities, like the uniform density that we observed numerically above, and determining how quickly these limiting densities begin to overwhelm the initial data.

2.2 Frobenius-Perron Operators

In the last section we described a map $\mathcal{P}$ on probability densities that encoded the dynamics of the doubling map T. In this section we will discuss the Frobenius-Perron operator $\mathcal{P}$ associated to a general measurable map T on a probability space (X, μ). We will also describe the Perron-Frobenius operator in more detail. Before we give a general definition let us consider an example that highlights one of the key obstructions to defining $\mathcal{P}$.

Example 2.2.1. Consider the map $S\colon [0,1] \circlearrowleft$ defined by

$$S(x) = \left\{ \begin{array}{ll} 2x, & x \in \left[0, \frac{1}{2}\right); \\ 1 & x \in \left[\frac{1}{2}, 1\right]. \end{array} \right\} \tag{2.2.1}$$

If we think of x being uniformly distributed in $[0,1]$ then $S(x) = 1$ with probability $1/2$. There is no density that represents this distribution for $S(x)$. Even though there is no density, we can define a measure that represents this distribution of mass. The map S is piecewise linear and hence measurable, meaning that for every measurable set E the set $S^{-1}E$ is measurable. Given any probability measure μ we can define $S_*\mu(E) := \mu\left(S^{-1}E\right)$). The map S_* acts on measures and corresponds to the map $\mathcal{P}$ on densities (see Lemma 2.2.4). Note that since $S_*\mu(E) = \mu(S^{-1}E)$ the map S_* carries the measure of the initial set $S^{-1}(E)$ on to the future set E. The uniform distribution of x corresponds to Lebesgue measure on $[0,1]$ i.e. $P(x < t) = \mu[0,t] = t$. Let δ_1 denote the measure defined for measurable E by

$$\delta_1(E) := \left\{ \begin{array}{ll} 0, & 1 \notin E; \\ 1, & 1 \in E. \end{array} \right.$$

It is not hard to see that $S_*\lambda = \frac{1}{2}\delta_1 + \frac{1}{2}\lambda$.

The Radon-Nikodym theorem states that a measure ν can be represented by a density with respect to another measure μ if and only if ν is absolutely continuous with respect to μ. Recall that a measure ν is *absolutely continuous* with respect to another measure μ if, for every measurable set E, if $\mu(E) = 0$, then $\nu(E) = 0$. The measure δ_1 is not absolutely continuous with respect to λ, it follows that $S_*\lambda$ is not absolutely continuous with respect to λ. Therefore, $S_*\lambda$ cannot be represented by a density with respect to λ. We conclude that the Frobenius-Perron operator associated to S cannot be defined.

With the previous example in mind we make the following definition.

Definition 2.2.2. A measurable transformation T of a measure space (X, μ) is *non-singular* if $T_*\mu$ is absolutely continuous with respect to μ.

If T is non-singular with respect to Lebesgue measure, then we can define $\mathcal{P}$ acting on densities. Any density η defines a measure $\eta\lambda$ by the formula $\eta\lambda(E) := \int_E \eta \, d\lambda$. It follows from the definition that the measure $\eta\lambda$ is absolutely continuous with respect to λ. If T is a nonsingular map with respect to Lebesgue measure, then $T_*(\eta\lambda)$ is absolutely continuous with respect to λ. By the Radon-Nikodym Theorem, there there exists a density $\frac{dT_*(\eta\lambda)}{d\lambda}$ such that

$$T_*(\eta\lambda)(E) = \int_E \frac{dT_*(\eta\lambda)}{d\lambda} \, d\lambda. \tag{2.2.2}$$

Notice that the equation above makes sense for any $\eta \in L^1([0,1], \lambda)$. In fact the equation above makes sense for any measure μ and map T that is non-singular with respect to μ. With the formula above in mind we make the following general definition of the Frobenius-Perron operator.

Definition 2.2.3. If T is a non-singular transformation on a probability space (X, μ), then $\mathcal{P}\colon L^1(X, \mu) \circlearrowleft$ is defined by

$$\mathcal{P}\eta = \frac{dT_*(\eta\mu)}{d\mu}.$$

If we wish to study the action of a non-singular map on absolutely continuous measures then eq. (2.2.2) indicates that it is equivalent to study either of the operators T_* or $\mathcal{P}$. Before we move on, we will record a few useful facts about general Frobenius-Perron operators.

Lemma 2.2.4. *If $T\colon X \circlearrowleft$ is a nonsingular transformation of a probability space (X,μ) and $\mathcal{P}\colon L^1(X,\mu) \circlearrowleft$ is the associated Frobenius-Perron operator, then*

1. *$\mathcal{P}$ is a linear operator.*

2. *$\mathcal{P}$ is the unique linear operator such that for all $\eta \in L^1(X,\mu)$ and $\psi \in L^\infty(X,\mu)$,*

$$\int_X \mathcal{P}\eta\, \psi \, d\mu = \int_X \eta\, \psi \circ T \, d\mu. \tag{2.2.3}$$

3. *$\mathcal{P}$ is a positive operator, meaning that if η is non-negative μ almost everywhere, then $\mathcal{P}\eta$ is also.*

4. *$\mathcal{P}$ preserves integrals: for all $\eta \in L^1(X,\mu)$, $\int_X \mathcal{P}\eta\, d\mu = \int_X \eta\, d\mu$.*

5. *$\mathcal{P}$ has norm 1: for all $\eta \in L^1(X,\mu)$, $\|\mathcal{P}\eta\|_1 \leq \|\eta\|_1$, and there exists $\eta \neq 0$ such that $\|\mathcal{P}\eta\|_1 = \|\eta\|_1$.*

6. *The following equivalent statements highlight the relationship between $\mathcal{P}$ and T-invariant measures.*

   - *A measure ν is T-invariant ($\nu \circ T^{-1} = \nu$) and absolutely continuous with respect to μ, if and only if $\frac{d\nu}{d\mu}$ is $\mathcal{P}$-invariant $\left(\mathcal{P}\frac{d\nu}{d\mu} = \frac{d\nu}{d\mu}\right)$.*
   - *A density $\eta \in L^1(X,\mu)$ is $\mathcal{P}$-invariant, if and only if the measure $\eta\mu$ is T-invariant.*

7. *If $\eta\colon [0,1] \to \mathbb{R}$ is an integrable function, and T is a non-singular transformation of $([0,1],\lambda)$ that is almost everywhere differentiable, then*

$$\mathcal{P}\eta(x) = \sum_{y \in T^{-1}(x)} \frac{\eta(y)}{DT(y)}$$

   is a pointwise formula for $\mathcal{P}$ viewed as an operator on functions rather than $L^1(I)$ classes.

2.3 Spectral Theory for the Doubling Map

In section 2.1 we considered the doubling map $T(x) = 2x \mod 1$. We noticed that, given a localized initial density η_0 the densities $\mathcal{P}^n\eta_0$ seemed to spread out and approach the constant density. In this section we will see that by studying the spectrum of $\mathcal{P}$ we can prove that for sufficiently nice densities we have $\left\|\mathcal{P}^n\eta - \mathbf{1}_{[0,1]}\right\|_1 \to 0$ as $n \to \infty$ and that the rate of convergence is exponential. This is a strong quantitative restatement of our qualitative observation about "spreading out".

An easy calculation shows that $\mathcal{P}\mathbf{1}_{[0,1]} = \mathbf{1}_{[0,1]}$. Therefore $\mathbf{1}_{[0,1]}$ is an eigenvector with eigenvalue 1 for $\mathcal{P}$. From Lemma 2.2.4 it follows that Lebesgue measure is preserved by T. What is the rest of the spectrum of $\mathcal{P}$ and what can it reveal about the map T? The following lemma gives a disappointing answer for general densities.

Lemma 2.3.1. *The spectrum of* $\mathcal{P}\colon L^1(\lambda) \circlearrowleft$ *is equal to* $\{z \in \mathbb{C} : |z| \leq 1\}$.

Proof. Define the functions $E_0 = \mathbf{1}_{[0,\frac{1}{2}]} - \mathbf{1}_{[\frac{1}{2},1]}$ and $E_n = E_0 \circ T^n$. The functions E_n have a copy of E_0 on each interval $[j2^{-n}, (j+1)2^{-n}]$. A direct calculation shows that $\mathcal{P}E_n = E_{n-1}$ and that $\mathcal{P}E_0 = 0$. Now define the functions $\eta_z = \sum_{k=0}^{\infty} z^k E_k$ for each $z \in \mathbb{C}$ with $|z| < 1$. For each n we have $\|E_n\|_1 = 1$, and thus $\|\eta_z\|_1 \leq (1-|z|)^{-1} < \infty$. We see that

$$\mathcal{P}\eta_z = \mathcal{P}\sum_{k=0}^{\infty} z^k E_k = \sum_{k=0}^{\infty} z^k \mathcal{P}E_k = \sum_{k=1}^{\infty} z^k E_{k-1} = z\eta_z$$

This shows that every z in the open unit disk $\{|z| < 1\}$ is an eigenvalue of $\mathcal{P}$. From Lemma 2.2.4 we know that $\|\mathcal{P}\|_{op} = 1$, it follows easily that $\|\mathcal{P}^n\|_{op} = 1$. By the Gelfand spectral radius formula, $\rho(\mathcal{P}) = 1$. We conclude that $\sigma(\mathcal{P})$ is contained in $\{|z| \leq 1\}$. A standard result in functional analysis is that the spectrum of a bounded linear operator on a Banach space is closed. The spectrum of $\mathcal{P}$ must contain the closure of $\{|z| < 1\}$ and is contained in $\{|z| \leq 1\}$. Therefore, the spectrum of $\mathcal{P}$ acting on $L^1([0,1])$ is $\{|z| \leq 1\}$. □

This result tells us that the spectrum of $\mathcal{P}$ does not provide any usable information about the asymptotic behavior of $\mathcal{P}$ since there are no distinguished dominant eigenvalues. The lemma above illustrates a phenomena that persists in great generality. In [9] the authors show that, when the Frobenius-Perron operator of a non-singular map satisfying a weak technical condition[1] acts on L^1, then $\sigma(\mathcal{P})$ is either the closed unit disk or contained in the unit circle.

From the proof of Lemma 2.3.1 we can see that $L^1([0,1], \lambda)$ contains functions that are "arbitrarily detailed". If these functions represented measured data, then they

[1] To be precise, $(X, \mathcal{A}, \mu, T)$ is a non-singular transformation of a σ-finite measure space such that $\forall E \in \mathcal{A}$, $\mu(E) > 0 \implies \mu(T^{-1}E) > 0$.

would require measurements with infinite precision. Our motivation for introducing the probabilistic paradigm was to address the fact that we can not make measurements with infinite precision. We could make this explicit by restricting $\mathcal{P}$ to a subspace of $L^1([0,1],\lambda)$ functions with some amount of regularity. We will chose to consider the following fairly weak form of regularity.

Definition 2.3.2. Given $\eta\colon \mathbb{R}\to\mathbb{R}$ define

$$\begin{aligned} V(\eta) &= \sup\left\{\sum_{j=1}^{N}|\eta(x_j)-\eta(x_{j-1})| : -\infty < x_0 < \cdots < x_N < \infty\right\}, \\ \operatorname{var}(\eta) &= \inf\left\{V(\hat{\eta}) : \hat{\eta} = \eta\ \lambda - a.e.\right\}. \end{aligned}$$

From the definition we see that var is constant on $L^1([0,1])$ classes. With this in mind, we define the following Banach space.

Definition 2.3.3. Define $\|\eta\|_{\mathcal{BV}} = \operatorname{var}(\eta) + \|\eta\|_1$ and let

$$\mathcal{BV} = \left\{\eta \in L^1([0,1],\lambda) : \|\eta\|_{\mathcal{BV}} < \infty\right\}.$$

Note that $\mathcal{BV}$ is a closed subspace of $L^1([0,1],\lambda)$. It is not to hard to show using eq. (2.1.2) that $\mathcal{PBV}\subseteq\mathcal{BV}$ so $\mathcal{P}|_{BV}\colon \mathcal{BV} \circlearrowleft$ is a well defined linear operator. Now let us consider the spectrum of $\mathcal{P}|_{\mathcal{BV}}$.

Lemma 2.3.4. *The operator $\mathcal{P}\colon \mathcal{BV}\circlearrowleft$ has the following properties:*

1. *The function $\mathbf{1}_{[0,1]}$ is the unique eigenvector of $\mathcal{P}$ with eigenvalue 1.*
2. *If $F = span\left(\mathbf{1}_{[0,1]}\right)$ and $H = \left\{\eta\in\mathcal{BV} : \int_{[0,1]}\eta\, d\lambda = 0\right\}$, then $\mathcal{BV} = F\oplus H$ and this splitting is $\mathcal{P}$-invariant meaning that $\mathcal{P}F\subseteq F$ and $\mathcal{P}H\subseteq H$.*
3. *If $\Pi\colon\mathcal{BV}\to F$ is defined by $\Pi\eta = \mathbf{1}_{[0,1]}\int_{[0,1]}\eta\, d\lambda$ and $Q = \mathcal{P}-\Pi$, then for all $k\geq 1$,*
$$\mathcal{P}^k = \Pi + Q^k.$$
4. *The spectrum of $\mathcal{P}$ is the union of $\sigma(\Pi) = \{0,1\}$ and $\sigma(Q)$, which is contained in the closed disk of radius $1/2$ centered at the origin in $\mathbb{C}$.*

Proof. To verify that $\mathbf{1}_{[0,1]}$ is an eigenvector of $\mathcal{P}$ with eigenvalue 1 we apply eq. (2.1.2). For any $x\in[0,1]$,

$$\mathcal{P}\mathbf{1}_{[0,1]}(x) = \tfrac{1}{2}\left[\mathbf{1}_{[0,1]}\left(\tfrac{x}{2}\right) + \mathbf{1}_{[0,1]}\left(\tfrac{x+1}{2}\right)\right] = 1 = \mathbf{1}_{[0,1]}(x).$$

It remains to show that this eigenvector is unique.

Next we check that $F \oplus H$ is an invariant splitting of $\mathcal{BV}$. By standard arguments both F and H are closed subspaces of $\mathcal{BV}$ and $F \cap H = \{0\}$, therefore $\mathcal{BV} = F \oplus H$. From eigenvector equation above and linearity of $\mathcal{P}$ we see that $\mathcal{P}F \subseteq F$. In Lemma 2.2.4 we observed that $\mathcal{P}$ preserves integrals. Therefore, $\mathcal{P}H \subseteq H$ and we have verified that the splitting is $\mathcal{P}$-invariant.

We verify the formula for $\mathcal{P}^k$ as follows. From the definition of Π it is easy to check that $\Pi^2 = \Pi$. We claim that $\mathcal{P}\Pi = \Pi = \Pi\mathcal{P}$. To prove this we apply the eigenvector equation for the first equality and preservation of integrals for the second. For all $\eta \in L^1(\lambda)$,

$$\mathcal{P}\left(\mathbf{1}_{[0,1]} \int \eta \, d\lambda\right) = \mathbf{1}_{[0,1]} \int \eta \, d\lambda = \mathbf{1}_{[0,1]} \int \mathcal{P}\eta \, d\lambda.$$

Next we claim that $Q\Pi = \Pi Q = 0$. We verify this algebraically as follows

$$\Pi Q = \Pi(\mathcal{P} - \Pi) = \Pi\mathcal{P} - \Pi^2 = \Pi - \Pi = 0,$$

with a similar calculation for $Q\Pi = 0$. Finally, by the definition of Q we have $\mathcal{P} = \Pi + Q$, and the calculation

$$\mathcal{P}\left(\Pi + Q^k\right) = (\Pi + Q)(\Pi + Q^k) = \Pi^2 + Q\Pi + (\Pi Q)Q^{k-1} + Q^{k+1} = \Pi + Q^{k+1}$$

proves by induction that $\mathcal{P}^k = \Pi + Q^k$.

We proceed to verify the decomposition of the spectrum of $\mathcal{P}$. We will first show that $\sigma(Q)$ is contained in the closed disk of radius $1/2$ centered at the origin in $\mathbb{C}$. To do so we will show that for all η in $\mathcal{BV}$ we have $\left\|Q^k\eta\right\|_{\mathcal{BV}} \leq 2^{1-k} \|\eta\|_{\mathcal{BV}}$. From this bound we conclude that $\|Q^k\|_{op} \leq 2^{1-k}$ and hence by the Gelfand spectral radius formula we have

$$\rho(Q) = \liminf_{k\to\infty} \|Q^k\|_{op}^{1/k} \leq \tfrac{1}{2}.$$

To prove the desired bound we claim that it suffices to show that for all bounded measurable ξ we have

$$V\left(\mathcal{P}^k\xi\right) \leq \tfrac{1}{2^k} V(\xi).$$

To see why note that $\mathcal{P}(I - \Pi) = \mathcal{P} - \Pi = Q$. It is a standard result that for all

$\eta \in \mathcal{BV}$,

$$\|\eta\|_\infty \leq \mathrm{var}(\eta) + \left| \int \eta \, d\lambda \right|.$$

Since $\Pi Q = 0$ we know that for all ξ we have $\int Q\eta \, d\lambda = 0$. Therefore, we have $\left\|Q^k\eta\right\|_1 \leq \left\|Q^k\eta\right\|_\infty \leq \mathrm{var}(Q^k\eta)$, hence $\left\|Q^k\eta\right\|_{\mathcal{BV}} = \left\|Q^k\eta\right\|_\infty + \mathrm{var}Q^k\eta \leq 2\mathrm{var}(Q^k\eta)$. Now by definition we have

$$\begin{aligned} \mathrm{var}(Q^k\eta) &= \inf\left\{V(\nu) : \nu = Q^k\eta \ \lambda\text{-a.e.}\right\} \leq \inf\left\{V(Q^k\xi) : \xi = \eta \ \lambda\text{-a.e.}\right\} \\ &\leq \tfrac{1}{2^k} \inf\left\{V(\xi) : \xi = \eta \ \lambda\text{-a.e.}\right\} \end{aligned}$$

where the first inequality comes from noting that the second set is contained in the first and the second inequality will follow from our bound on $V\left(\mathcal{P}^k\xi\right)$. This completes the proof of sufficiency.

Next we compute the desired bound on $V\left(\mathcal{P}^k\xi\right)$. Fix $x_0 < x_1 < \cdots < x_n$. Let $x_0^k < x_1^k < \cdots < x_{(n+1)2^k-1}^k$ be the left to right enumeration of $T^{-k}\left\{x_0, \cdots, x_n\right\}$. Applying eq. (2.1.2), induction, and the triangle inequality we obtain

$$\sum_{i=1}^{n} \left|\mathcal{P}^k\xi(x_i) - \mathcal{P}^k\xi(x_{i-1})\right| \leq \tfrac{1}{2^k} \sum_{i=1}^{(n+1)2^k-1} \left|\xi(x_i^k) - \xi(x_{i-1}^k)\right|.$$

Since the partition $x_0 < x_1 < \cdots < x_n$ was arbitrary we obtain

$$V(\mathcal{P}^k\xi) \leq \tfrac{1}{2^k} V(\xi)$$

as desired. This completes the proof that $\rho(Q) \leq \frac{1}{2}$.

A consequence of the bound $\left\|\mathcal{P}^k(I - \Pi)\eta\right\|_{\mathcal{BV}} = \left\|Q^k\eta\right\|_{\mathcal{BV}} \leq \frac{1}{2^{k-1}}\mathrm{var}(\eta)$ is that if $\mathcal{P}\eta = \eta$ then $\|\eta - \Pi\eta\|_{\mathcal{BV}} = 0$. We conclude that $\mathbf{1}_{[0,1]}$ is the unique eigenvector with eigenvalue 1.

Lastly, note that from the identity $\mathcal{P}^k = \Pi + Q^k$ we obtain

$$\sum_{k=0} z^k\mathcal{P}^k = \frac{\Pi}{1-z} + \sum_{k=0}^{\infty} z^k Q^k.$$

Therefore the resolvent set of $\mathcal{P}$ is contained in the intersection of the resolvent sets of Π and Q which proves the claimed decomposition of the spectrum of $\mathcal{P}$.

□

2.4 Quasi-Compactness

The spectral decomposition for the doubling map is prototypical for class operators called *quasi-compact operators*.

Definition 2.4.1. Let L be an operator on a Banach space $\mathfrak{B}$. L is *quasi-compact*, if there exists $p \in [0, \rho(L))$, such that there are closed subspaces F and H in $\mathfrak{B}$ that satisfy the following conditions:

1. $\mathfrak{B} = F \oplus H$,
2. $0 < \dim(F) < \infty$, $LF \subseteq F$, and $|\lambda| > p$ for all $\lambda \in \sigma(L|_F)$,
3. $LH \subseteq H$ and $\rho(L|_H) \leq p$.

We define the *essential spectral radius* of L, denoted $\rho_{ess}(L)$, to be the infimum over the set containing $\rho(L)$ and all $p \geq 0$ satisfying the hypotheses above.

A quasi-compact operator L splits into the direct sum of a dominant finite dimensional operator $L|_F$ and a transient operator $L|_H$, which decays exponentially fast. Quasi-compact operators generalize compact operators. From the spectral theory of compact operators, one can see that every compact operator is a quasi-compact operator with essential spectral radius equal to zero. An alternate characterization of compact operators is that the image of any bounded set is a totally bounded set. The next definition gives a way to quantify how far an operator on a Banach space is from being compact.

Definition 2.4.2. Let $\mathfrak{B}_1$ and $\mathfrak{B}_2$ be Banach spaces. Consider a bounded subset A of $\mathfrak{B}_i$, let $r(A)$ denote the infimum of the set of non-negative numbers d such that there exist a finite cover of A by sets of diameter at most d. We refer to $r(A)$ as the *measure of non-compactness* of A. Given a mapping $L\colon \mathfrak{B}_1 \to \mathfrak{B}_2$ we say that L is a C-set-contraction if for all bounded sets $A \subseteq \mathfrak{B}$ we have $r(L(A)) \leq Cr(A)$. We abuse notation and define $r(L)$ to be the infimum over non-negative numbers C such that L is a C-set-contraction. We refer to $r(L)$ as the *measure of non-compactness* of L.

If $L\colon \mathfrak{B}_1 \to \mathfrak{B}_2$ is a compact operator and $A \subset \mathfrak{B}_1$ is bounded, then $r(L(A)) = 0$. This is because, $L(A)$ is a totally bounded set. Since A was arbitrary it follows that $r(L) = 0$.

The following theorem of Nussbaum shows that quasi-compact operators can also be viewed as a generalization of compact operators from the perspective of the alternate characterization.

Theorem 2.4.3 ([20]). *Let $\mathfrak{B}$ be a complex Banach space and $L \in Hom(\mathfrak{B}, \mathfrak{B})$. Then $\rho_{ess}(L) = \lim_{n\to\infty} r(L^n)^{1/n}$.*

The following theorem of Hennion leverages the alternate characterization of quasi-compactness provided by the Nussbaum theorem to provide a flexible and very useful sufficient condition for an operator to be quasi-compact.

Theorem 2.4.4 (Hennion [15] via Liverani [17]). *If $\mathfrak{B} \subseteq \mathfrak{B}_w$ are Banach spaces with norms $\|\cdot\|_{\mathfrak{B}}$ and $\|\cdot\|_{\mathfrak{B}_w}$ respectively, such that $\|\cdot\|_{\mathfrak{B}_w} \leq \|\cdot\|_{\mathfrak{B}}$, and $L : \mathfrak{B} \circlearrowleft$ is a bounded linear operator such that:*

1. *$L : \mathfrak{B} \to \mathfrak{B}_w$ is a compact operator;*
2. *There exists $\theta, A, B, C > 0$ such that forall $n \in \mathbb{N}$ there exists $M_n > 0$ such that for all $f \in \mathfrak{B}$, we have*
   (a) $\|L^n f\|_{\mathfrak{B}_w} \leq CM_n \|f\|_{\mathfrak{B}_w}$,
   (b) $\|L^n f\|_{\mathfrak{B}} \leq A\theta^n \|f\|_{\mathfrak{B}} + BM_n \|f\|_{\mathfrak{B}_w}$.

Then $L\colon \mathfrak{B} \circlearrowleft$ is quasi compact and $\rho_{ess}(L) \leq \theta$. We will refer to the second inequality above as the Lasota-Yorke inequality*.*

When we apply the theorem above we will refer to the norm that will play the role of $\|\cdot\|_{\mathfrak{B}}$ as the *strong norm*, and the norm that will play the role of $\|\cdot\|_{\mathfrak{B}_w}$ as the *weak norm*. Similarly we will refer to the Banach that play the roles of $\mathfrak{B}$ and $\mathfrak{B}_w$, as the *strong space* and *weak space* respectively. The proof of Theorem 2.4.4 is both short and provides intuition so we include a sketch.

Proof. Fix $n \geq 1$ and select $\epsilon_n \geq 0$ such that $\epsilon_n \leq \frac{\theta^n}{M_n}$. Let B_1 denote the unit ball $\mathfrak{B}$. By the first hypothesis of Theorem 2.4.4 the set $L(B_1)$ is compact in the topology of $\mathfrak{B}_w$. Select a finite set $f_1, \cdots, f_k \in A$ such that the sets $U_j =$

$\left\{f \in L(B_1) : \|f - f_j\|_{\mathfrak{B}_w} < \epsilon_n\right\}$ cover $L(B_1)$. By the Lasota-Yorke inequality and the fact that each U_j is contained in the L-image of the unit ball of $\mathfrak{B}$ we have, for all $f \in U_j$,

$$\|f - f_j\|_{\mathfrak{B}} \le \|f\|_{\mathfrak{B}} + \|f_j\|_{\mathfrak{B}} \le 2(A\theta + BM_1).$$

The sets $L^{n-1}(U_j)$ cover $L^n(B_1)$ and a second application of the Lasota-Yorke inequality provides the following bound on their $\|\cdot\|_{\mathfrak{B}}$-diameter. For all $f \in U_j$

$$\begin{aligned}\left\|L^{n-1}(f - f_j)\right\|_{\mathfrak{B}} &\le A\theta^{n-1} \|f - f_j\|_{\mathfrak{B}} + BM_{n-1} \|f - f_j\|_{\mathfrak{B}_w} \\ &\le A\theta^{n-1} \left(A\theta + BM_1\right) + BM_{n-1}\epsilon_n \\ &\le \frac{A\left(A\theta + BM_1\right) + B}{\theta}\, \theta^n =: D\theta^n\end{aligned}$$

We conclude that $L^n(B_1)$ is covered by finitely many subsets of $\mathfrak{B}$ of $\|\cdot\|_{\mathfrak{B}}$-diameter at most $D\theta^n$ and hence by Theorem 2.4.3 we have

$$\rho_{ess}\left(L\right) = \lim_{n\to\infty} r(L^n)^{1/n} \le \lim_{n\to\infty} \left(D\theta^n\right)^{1/n} = \theta.$$

□

Notice that in the proof above the numbers ϵ_n are chosen independently for each n and can therefore accommodate any rate of growth in the sequence $\|L^n(f - f_j)\|_{\mathfrak{B}_w}$. From the proof we see that the key ingredients are a pairing of strong and weak Banach spaces, and a Lasota-Yorke inequality.

Chapter 3

Expanding Interval Maps

3.1 Expanding Interval Maps

The doubling map discussed in section 2.1 fits into the class of *expanding interval maps*. In this section, we will define expanding interval maps and introduce some notation that will be useful when we work with them. In section 3.2 we will apply Theorem 2.4.4 to show that all expanding interval maps have Frobenius-Perron operators that are quasi-compact on the space of functions of bounded variation.

We will refer to *branches* of a map $T\colon I \to I$, by this we mean a continuous function obtained by restricting T to a subinterval of I. An expanding interval map is a map of the unit interval I that consists of finitely many smooth[1], injective, and uniformly expanding[2] branches.[3] Let $T_1, \cdots, T_N$ denote the functions obtained by extending each branch of T to the closure of its domain, and let I_j denote the interior of the domain of T_j. Each I_j is an open interval and $\lambda\left(I - \bigcup_{j=1}^{N} I_j\right) = 0$.

We will say that an expanding interval map is *full branched* if each of its branches maps onto I. Consider the action of the Frobenius-Perron operator of a full branched

[1] S is a *smooth* branch of T if, there exists an interval J such that $S = T|_J$, the restriction of S to the interior of J is C^2, and D^2S extends continuously to the closure of J.

[2] The branches $S_1, \cdots, S_k$ of T are uniformly expanding if, there exists a constant $C > 1$ such that for all $1 \le j \le k$ and $x \in I$, $DS_j(x) > C$.

[3] The phrase 'consists of' is ambiguous. By this we mean that there exists a partition of I into intervals such that each of the functions obtained by restricting T to a cell of the partition is a smooth branch.

expanding interval map on a smooth PDF η. From Lemma 2.2.4 we have

$$\mathcal{P}\eta(x) = \sum_{y\in T^{-1}(x)} \frac{\eta(x)}{DT(y)} = \sum_{j=1}^{N} \frac{\eta\left(T_j^{-1}(x)\right)}{DT_{j^{-1}}\left(T_j^{-1}(x)\right)}. \tag{3.1.1}$$

We have used the fact that each branch T_j is onto to ensure that each term in the second sum is well defined for all $x \in I$. It follows from the formula above that if η is smooth and T is full branched, then the function $\mathcal{P}\eta$ is smooth. We compute the derivative of $\mathcal{P}\eta$ below.

$$D\left[\mathcal{P}\eta\right](x) = \sum_{j=0}^{N} \frac{D\eta\left(T_j^{-1}(x)\right)}{\left[DT_j\left(T_j^{-1}(x)\right)\right]^2} - \frac{\eta\left(T_j^{-1}(x)\right)}{DT_j\left(T_j^{-1}(x)\right)} \frac{D^2T_j\left(T_j^{-1}(x)\right)}{\left[DT_j\left(T_j^{-1}(x)\right)\right]^2}$$

Inspecting the formula above indicates that there are two important quantities that control the effect of $\mathcal{P}$ on the derivative of η. The first is $\beta(T) < 1$ defined by

$$\beta(T) = \sup\left\{\frac{1}{DT_j(t)} : 1 \leq j \leq N, t \in Cl(I_j)\right\}. \tag{3.1.2}$$

This is just the multiplicative inverse of the expansion rate for T. The second important quantity associated to an expanding interval map is its *distortion*, which is defined by

$$\kappa(T) = \sup\left\{\frac{D^2T_j(x)}{\left(DT_j(t)\right)^2} : 1 \leq j \leq N, t \in Cl(I)\right\}. \tag{3.1.3}$$

Since T has finitely many branches and D^2T_j is continuous on a compact set for each $1 \leq j \leq N$ the distortion of an expanding interval map is always finite. With the two quantities defined we obtain the following bound on the derivative of $\mathcal{P}\eta$.

$$\left\|D\left[\mathcal{P}\eta\right]\right\|_\infty \leq \beta(T)\left\|\mathcal{P}\left[D\eta\right]\right\|_\infty + \kappa(T)\left\|\mathcal{P}\eta\right\|_\infty \tag{3.1.4}$$

Let us consider two examples.

Example 3.1.1. Consider the doubling map, which has two branches over the intervals $[0, 1/2)$ and $[1/2, 1]$. The branches of Td are clearly smooth, injective, and expanding so the doubling map is an expanding interval map. For the doubling map we have $\beta(Td) = 1/2$ and $\kappa(Td) = 0$. Since Td has two branches and derivative 2 everywhere, it follows easily from eq. (3.1.1) that $\left\|\mathcal{P}\eta\right\|_\infty \leq \left\|\eta\right\|_\infty$ for all $\eta \in L^\infty(\lambda)$. Note that any piecewise linear map T will have the property that $\kappa(T) = 0$. Applying eq. (3.1.4) we find

$$\|D[\mathcal{P}\eta]\|_\infty \le \tfrac{1}{2}\|D\eta\|_\infty$$

This shows that $\mathcal{P}$ has a rather strong smoothing effect on PDFs. By iterating this bound, we see that if η is smooth $\mathcal{P}^n f$ must converge to a constant function. Since η is a PDF and $\mathcal{P}$ preserves Lebesgue integrals we see that this constant is necessarily 1.

Example 3.1.2. Consider the map

$$f(x) = \begin{cases} \frac{3-\sqrt{9-16x}}{2} & x \in \left[0, \frac{1}{2}\right), \\ \frac{-1+\sqrt{16x-7}}{2} & x \in \left[\frac{1}{2}, 1\right]. \end{cases}$$

Let f_0 and f_1 denote the left and right branches of f. For reasons which will become

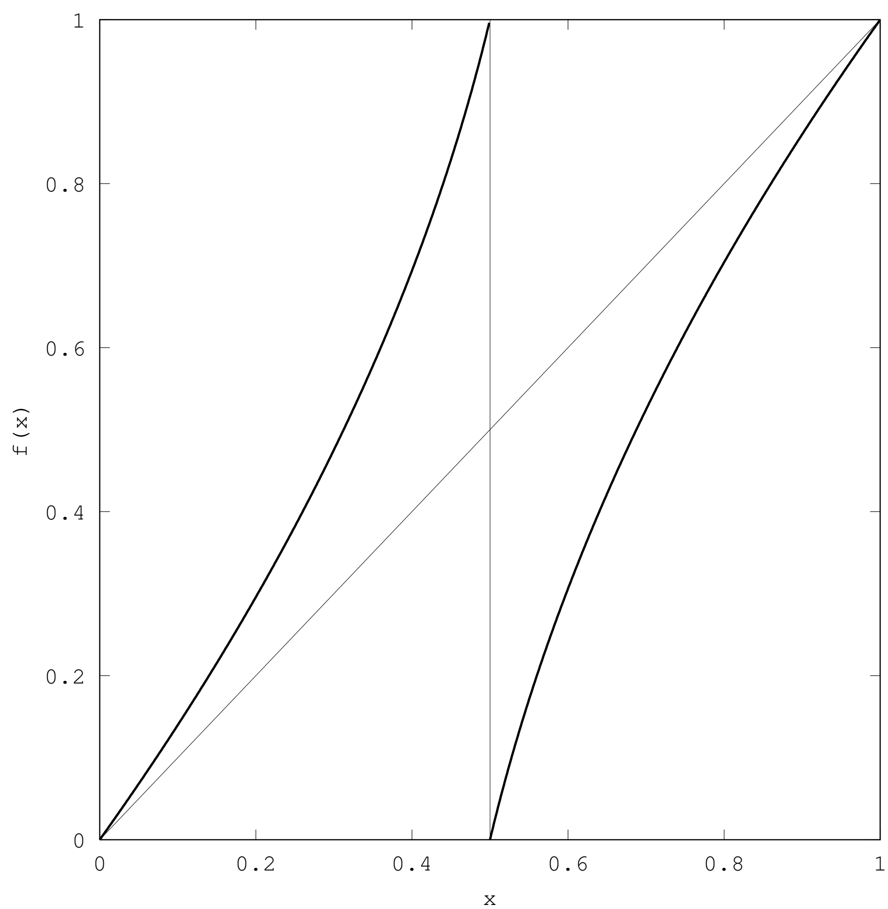

Figure 3.1: Plot of $y = f(x)$ with guide lines $x = 1/2$ and $y = x$.

apparent in section 5.1, we have the following functional relationship between the branches and the function $\phi = \frac{1}{4}(3-2x)$,

$$Df_0(x) = \frac{1}{\phi(f_0(x))}$$
$$Df_1(x) = \frac{1}{1-\phi(f_1(x))}$$

Applying the relations above, one can obtain $\beta(f) = \frac{3}{4}$ and $\kappa(f) = \frac{2}{3}$. It also follows from the relations that, if x_0 and x_1 are the preimages of a point x under f_0 and f_1 respectively, then we have

$$\frac{1}{Df(x_0)} + \frac{1}{Df(x_1)} = \frac{1}{Df_0(x_0)} + \frac{1}{Df_1(x_1)} = \phi(x) + 1 - \phi(x) = 1.$$

It follows from the equation above and eq. (3.1.1) that $\|\mathcal{P}\eta\|_\infty \leq \|\eta\|_\infty$ for all $g \in L^\infty(I,\lambda)$. Finally, applying the bound in eq. (3.1.4) we obtain.

$$\|D[\mathcal{P}\eta]\|_\infty \leq \tfrac{3}{4}\|D\eta\|_\infty + \tfrac{2}{3}\|\eta\|_\infty .$$

The non-linearity of the branches introduces distortion and it becomes less clear what effect $\mathcal{P}$ has on the derivative of a PDF.

The last example left us with a bound on $\|D[\mathcal{P}\eta]\|_\infty$ that was not so instructive. It would be nice to obtain a bound on $\|D[\mathcal{P}^n\eta]\|_\infty$. A nice way to obtain such a bound is to analyze the map f^n directly.

We claim that, if T is an expanding interval map, then for any n, T^n is also an expanding interval map. To see this, let $\mathcal{P}$ denote the partition[4] into the intervals I_j, and let $T^{-k}\mathcal{P} = \left\{T^{-k}I_j : I_j \in \mathcal{P}\right\}$. Finally, let $\mathcal{P}_n$ denote the common refinement of the partitions $T^{-k}\mathcal{P}$ for $0 \leq k < n$. It follows from the continuity and injectivity of the branches of T, that the partition[5] $\mathcal{P}_n$ consists of finitely many open intervals. The map T^n, restricted to an interval contained in $\mathcal{P}_n$, is the composition of n branches of T all of which are smooth, injective and expanding. Therefore the resulting branch of T^n is also smooth, injective and expanding. We conclude that T^n is in fact an

[4]This collection of disjoint sets does not contain the finitely many endpoints of the intervals I_j and thus is not a true partition. We could insist that $\mathcal{P}$ be the partition by right open left closed intervals containing the I_j, but it makes little difference in what follows and would only serve to confuse the notation.

[5]Again this this collection covers up to a finite set of points.

expanding interval map.

Since T^n is an expanding interval map we can compute $\beta(T^n)$ and $\kappa(T^n)$. The quantity $\beta(T^n)$ can be bounded in terms of $\beta(T)$ by applying the chain rule,

$$\left|\frac{1}{DT^n(x)}\right| = \prod_{k=0}^{n-1}\left|\frac{1}{DT\left(T^k(x)\right)}\right|,$$

so that we have

$$\beta(T^n) \leq \beta(T)^n. \tag{3.1.5}$$

We can also obtain the following bound on the distortion of T^n,

$$\kappa(T^n) \leq (1-\beta(T))^{-1}\kappa(T). \tag{3.1.6}$$

We will verify the bound by induction as follows. Apply the chain rule and the bound on β to obtain

$$\begin{aligned}
DT^{n+1}(x) &= DT^n\left(T(x)\right)DT(x)\\
D^2T^{n+1}(x) &= D^2T^n\left(T(x)\right)\left[DT(x)\right]^2 + DT^n\left(T(x)\right)D^2T(x)\\
\frac{D^2T^{n+1}(x)}{\left[DT^{n+1}(x)\right]^2} &= \frac{D^2T^n\left(T(x)\right)}{\left[DT^n\left(T(x)\right)\right]^2} + \frac{1}{DT^n\left(T(x)\right)}\frac{D^2T(x)}{\left[DT(x)\right]^2}\\
&\leq \kappa\left(T^n\right) + \beta\left(T^n\right)\kappa\left(T\right)\\
&\leq \kappa\left(T^n\right) + \beta\left(T\right)^n\kappa\left(T\right).
\end{aligned}$$

From the last inequality and a geometric series calculation we obtain the claimed bound.

Let us return to our example map f.

Example 3.1.3. By analyzing f^n directly we obtain

$$\begin{aligned}
\left\|D\left[\mathcal{P}^n\eta\right]\right\|_\infty &\leq \beta\left(f^n\right)\left\|D\eta\right\|_\infty + \kappa\left(f^n\right)\left\|\eta\right\|_\infty\\
&\leq \beta\left(f\right)^n\left\|D\eta\right\|_\infty + \frac{\kappa\left(f\right)}{1-\beta(f)}\left\|\eta\right\|_\infty\\
&= \left(\tfrac{2}{3}\right)^n\left\|D\eta\right\|_\infty + \tfrac{8}{3}\left\|\eta\right\|_\infty\\
&\leq \left(\tfrac{2}{3}\right)^n\left(\left\|D\eta\right\|_\infty + \left\|\eta\right\|_\infty\right) + \tfrac{8}{3}\left\|\eta\right\|_\infty
\end{aligned}$$

By the Arzelá-Ascoli Theorem, $C^1([0,1])$ is compactly embedded into $C^0([0,1])$. We have observed that $\mathcal{P}$ is a norm one operator on $L^\infty(I,\lambda)$, and hence on $C^0([0,1])$. The bound above readily shows that $\mathcal{P}$ is also a bounded operator on $C^1([0,1])$ and that $\mathcal{P}$ satisfies the following Lasota-Yorke inequality where we view the C^0-norm $\|\cdot\|_\infty$ as the weak norm and the C^1-norm defined by $|g|_{C^1} := \|D\eta\|_\infty + \|\eta\|_\infty$ as the strong norm.

$$|\mathcal{P}^n\eta|_{C^1} \leq \left(\tfrac{2}{3}\right)^n |\eta|_{C^1} + \tfrac{11}{3}\|\eta\|_\infty$$

Applying Theorem 2.4.4 with strong norm $|\cdot|_{C^1}$ and weak norm $\|\cdot\|_\infty$, we conclude that $\mathcal{P}$ is quasi-compact as an operator on $C^1([0,1])$ with essential spectral radius less than or equal to $\frac{2}{3}$.

While the example above is instructive and motivates the definitions of β and κ, not all expanding interval maps have Frobenius-Perron operators that restrict to C^1. The examples that we gave above did because they were full branched and preserved Lebesgue measure. To study expanding interval maps in general, we will need to study the action of the Frobenius-Perron operator on the less restrictive class of $\mathcal{BV}$ functions. In the next section we will analyze general expanding interval maps and obtain quasi-compactness of the associated Frobenius-Perron operators acting on $\mathcal{BV}$.

3.2 Spectral Theory for Expanding Interval Maps

In this section we will study the Frobenius-Perron operator associated to an arbitrary expanding interval map $T\colon I \circlearrowleft$ acting on $\mathcal{BV}$. We do not assume that the map is full branched or that Lebesgue measure is preserved.

The first step in this analysis is to recast the norm on $\mathcal{BV}$ so that it is more compatible with the functional analytic tools from section 2.4. The following lemma, which is proved in appendix A.1, provides us with two new formulas for computing var from Definition 2.3.2.

Lemma 3.2.1. *Let* $\mathbf{C}$ *denote the space of continuously differentiable functions from* I *to* $\mathbb{R}$ *and let* $\mathbf{L}$ *denote the space of Lipschitz functions from* I *to* $\mathbb{R}$*. For all absolutely* *integrable* $\eta\colon I \to \mathbb{R}$*,*

$$\operatorname{var}(\eta) = \sup\left\{\int_I \eta\, D\psi\, d\lambda : \psi \in \mathbf{C}, \|\psi\|_{\sup} \le 1\right\}$$
$$\operatorname{var}(\eta) = \sup\left\{\int_I \eta\, D\psi\, d\lambda : \psi \in \mathbf{L}, \|\psi\|_{\sup} \le 1\right\}.$$

In the second formula we recall that the derivative of a Lipschitz function ψ *exists at almost every point and is bounded by the Lipschitz constant of* ψ *where it exists. Therefore we interpret* $D\psi$ *as the* $L^\infty(I,\lambda)$ *class of functions almost everywhere equal to the derivative of* ψ *where it is defined.*

Our goal for the remainder of this section is to apply Theorem 2.4.4 and the formulas for var obtained in Lemma 3.2.1 to show that the Frobenius-Perron operator associated to T acting on $\mathcal{BV}$ is quasi-compact. To do so we identify $\mathfrak{B} = \mathcal{BV}([0,1],\lambda)$ and $\mathfrak{B}_w = L^1([0,1],\lambda)$ in the notation of Theorem 2.4.4. The map T is non-singular with respect to Lebesgue measure, which follows easily from the fact that $\beta(T) < 1$. Therefore the Frobenius-Perron operator $\mathcal{P}$ is defined, and by Lemma 2.2.4 we have $\|\mathcal{P}\eta\|_1 \le \|\eta\|_1$, which verifies the first inequality in Theorem 2.4.4. By Definition 2.3.2 we have $\|\cdot\|_1 \le \|\cdot\|_{\mathcal{BV}}$ so that the inequality between the weak and strong norms is satisfied. By Helly's Selection Theorem the space $\mathcal{BV}([0,1],\lambda)$ is compactly embedded in $L^1([0,1],\lambda)$. To apply Theorem 2.4.4 it remains to show that $\mathcal{P}$ is a bounded operator on $\mathcal{BV}$ and that $\mathcal{P}$ satisfies a Lasota-Yorke inequality. Notice that boundedness of $\mathcal{P}$ will follow from the Lasota-Yorke inequality since

$$\|\mathcal{P}\eta\|_{\mathcal{BV}} \le A\theta\,\|\eta\|_{\mathcal{BV}} + BM_1\,\|\eta\|_1 \le (A\theta + BM_1)\,\|\eta\|_{\mathcal{BV}}.$$

Therefore all that remains is to prove a Lasota-Yorke inequality for $\mathcal{P}$.

The first step in proving a Lasota-Yorke inequality in this setting is to bound $\operatorname{var}(\mathcal{P}\eta)$ for $\eta \in \mathcal{BV}$. To do this we will apply the first formula in Lemma 3.2.1. The next lemma is a well known identity that will be the foundation of our Lasota-Yorke inequality.

Lemma 3.2.2. *If $\eta \in \mathcal{BV}$ and $\psi \in \mathbf{C}$ with $\|\psi\|_{\sup} \leq 1$, then*

$$\int_I \mathcal{P}\eta \, D\psi \, d\lambda = \underbrace{\int_I \eta \, D\left(\sum_{j=1}^{N} \frac{\psi \circ T_j}{DT_j} \mathbf{1}_{I_j}\right) d\lambda}_{I} - \underbrace{\int_I \eta \, (\psi \circ T) \, \frac{D^2T}{(DT)^2} \, d\lambda}_{II}$$

Proof. Fix $\eta \in \mathcal{BV}$ and $\psi \in C^1$. We apply eq. (2.2.3) to obtain

$$\int_I \mathcal{P}\eta \, D\psi \, d\lambda = \int_I \eta \, (D\psi \circ T) \, d\lambda = \sum_{j}^{N} \int_{I_j} \eta \, (D\psi \circ T_j) \, d\lambda$$

Applying the chain rule interval by interval we obtain for each j

$$D\psi \circ T_j = \frac{D(\psi \circ T_j)}{DT_j}.$$

An application of the product rule yields

$$(D\psi) \circ T_j = D\left(\frac{\psi \circ T_j}{DT_j}\right) - (\psi \circ T_j) \, \frac{D^2T}{(DT)^2}.$$

$\square$

In order to control $\operatorname{var}(\mathcal{P}\eta)$ we must bound both terms in Lemma 3.2.2 uniformly with respect to ψ. Recall that $\left\|\frac{D^2T}{(DT)^2}\right\|_\infty < \kappa(T)$ and $\|\psi \circ T\|_{\sup} \leq \|\psi\|_{\sup} \leq 1$. Applying Hölder's inequality gives the following bound on term II.

$$\left| \int_I \eta \, (\psi \circ T) \, \frac{D^2T}{(DT)^2} \, d\lambda \right| \leq \kappa(T) \, \|\eta\|_1.$$

Term I requires slightly more care. Let ψ be fixed and define

$$\xi = \sum_{j=1}^{N} \frac{\psi \circ T_j}{DT_j} \mathbf{1}_{I_j},$$

so that Term I becomes $\int_I \eta \, D\xi \, d\lambda$. Note that ξ is doubly defined at each endpoint of an interval I_j. This integral appears to be of the form used to compute $\operatorname{var}(\psi)$ so we might hope to bound Term II by $\operatorname{var}(\eta)$. Unfortunately ξ may have jump discontinuities at the boundaries of the intervals I_j (see fig. 3.2).

In order to bound an integral of the form $\int_I \eta \, D\xi \, d\lambda$ by $\operatorname{var}(\eta)$ we must have ξ at

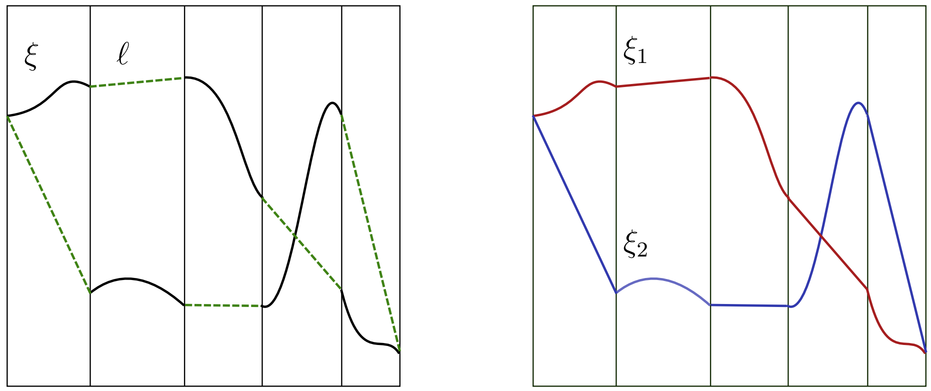

Figure 3.2: On the left we see the function ξ which is discontinuous and a discontinuous affine function ℓ that connects branches of ξ. On the right we see the functions ξ_1 and ξ_2 obtained by alternating between ξ and ℓ so that the resulting functions are continuous.

least Lipschitz continuous. The issue of discontinuities in ξ is well know and has been addressed in many different ways by various authors. Below we present a method for managing the discontinuities in ξ that is to to our knowledge previously unknown.

To resolve the issue of discontinuities in ξ we introduce a piecewise affine function $\ell\colon I \to \mathbb{R}$. We allow ℓ to be doubly defined at the endpoints of the intervals I_j. The function ℓ is uniquely determined by the following requirements,

1. $\ell|_{I_j}$ is affine,
2. $\ell(0) = \xi(0)$,
3. $\ell(1) = \xi(1)$,
4. The values of $\ell|I_j$ and $\xi|_{I_{j\pm 1}}$ agree at the shared endpoints of I_j and $I_{j\pm 1}$ whenever $I_{j\pm 1}$ exists.

Since an explicit formula for the values of ℓ is cumbersome and not needed we leave its computation to the reader as an instructive exercise in applying the definitions so far (see fig. 3.2). Let $r(T) = \min_j \{|I_j|\}$, which is the minimum length of the intervals I_j. Form the definition of ℓ we see that the values of ℓ are always convex combinations of values of ξ. From this observation we obtain two important bounds,

$$\|\ell\|_{\text{sup}} \leq \|\xi\|_{\text{sup}}$$

and

$$\left\| D\ell|_{I_j} \right\|_{\sup} \leq 2 \left\| \xi \right\|_{\sup} r(T),$$

where $r(T) = \min_j |I_j|$ is the minimum length of the intervals I_j. Next define $\zeta_1 \colon I \to \mathbb{R}$ and $\zeta_2 \colon I \to \mathbb{R}$ by

$$\zeta_1(x) = \begin{cases} \xi_j(x) & \text{if } x \in Cl(I_j) \text{ and } j \text{ is odd.} \\ \ell_j(x) & \text{if } x \in Cl(I_j) \text{ and } j \text{ is even.} \end{cases}$$

$$\zeta_2(x) = \begin{cases} \ell_j(x) & \text{if } x \in Cl(I_j) \text{ and } j \text{ is odd.} \\ \xi_j(x) & \text{if } x \in Cl(I_j) \text{ and } j \text{ is even.} \end{cases}$$

We have constructed ℓ so that both of the functions ζ_1 and ζ_2 are continuous. We note that the bound $\|\ell\|_{\sup} \leq \|\xi\|_{\sup}$ implies that for $i = 1, 2$ we have

$$\|\zeta_i\|_{\sup} \leq \|\xi\|_{\sup} \leq \max_j \|\psi \circ T_j\|_{\sup} \left\| \frac{1}{DT_j} \right\|_{\sup} \leq \beta(T)$$

We further note that

$$\|D\zeta_i\|_{\sup} \leq \max_j \left\{ \left\| D\xi|_{I_j} \right\|_{\sup}, \left\| D\ell|_{I_j} \right\|_{\sup} \right\} < \infty$$

and thus ζ_1 and ζ_2 are Lipschitz. From the definitions of ξ, ℓ, ζ_1, and ζ_2 we see that for all $x \in I$ we have

$$\xi(x) + \ell(x) = \zeta_1(x) + \zeta_2(x).$$

Solving for $\xi(x)$ and substituting we obtain the following bound on Term I.

$$\int_I \eta \, D\xi \, d\lambda \leq 2\beta(T)\mathrm{var}(\eta) + \frac{2\beta(T)}{r(T)} \|\eta\|_1 .$$

This bound is verified by the following computation

$$\begin{aligned}
\int_I \eta\, D\xi\, d\lambda =& \sum_{j=1}^{N} \int_{I_j} \eta\, D\xi_j\, d\lambda \\
=& \int_I \eta\, D\zeta_1\, d\lambda + \int_I \eta\, D\zeta_2\, d\lambda - \sum_{j=1}^{N} \int_{I_j} \eta\, D\ell\, d\lambda \\
=& \|\zeta_1\|_{\sup} \int_I \eta\, D\left(\frac{\zeta_1}{\|\zeta_1\|_{\sup}}\right) d\lambda \\
&+ \|\zeta_2\|_{\sup} \int_I \eta\, D\left(\frac{\zeta_2}{\|\zeta_2\|_{\sup}}\right) d\lambda \\
&+ \sum_{j=1}^{N} \left\|D\ell|_{I_j}\right\|_{\sup} \int_{I_j} |\eta|\ d\lambda \\
\le& 2\beta(T)\mathrm{var}(\eta) + \frac{2\beta(T)}{r(T)} \|\eta\|_1 .
\end{aligned}$$

We are now in a position to prove the following Lasota-Yorke inequality.

$$\|\mathcal{P}^n\eta\|_{\mathcal{BV}} \le 2\beta(T)^n \|\eta\|_{\mathcal{BV}} + \left[\left(\frac{2\beta(T)^n}{r(T^n)}\right) + \frac{\kappa(T)}{1-\beta(T)} + 1\right] \|\eta\|_1 . \qquad (3.2.1)$$

First note that by the bounds on terms I and II above we have,

$$\mathrm{var}(\mathcal{P}\eta) \le 2\beta(T)\mathrm{var}(\eta) + \left[\left(\frac{2\beta(T)}{r(T)}\right) + \kappa(T)\right] \|\eta\|_1 .$$

We apply the bound above to the case $n = 1$.

$$\begin{aligned}
\|\mathcal{P}\eta\|_{\mathcal{BV}} =& \mathrm{var}(\mathcal{P}\eta) + \|\mathcal{P}\|_1 \\
\le& 2\beta(T)\mathrm{var}(\eta) + \left[\left(\frac{2\beta(T)}{r(T)}\right) + \kappa(T)\right] \|\eta\|_1 + \|\eta\|_1 \\
\le& 2\beta(T) \|\eta\|_{\mathcal{BV}} + \left[\left(\frac{2\beta(T)}{r(T)}\right) + \kappa(T) + 1\right] \|\eta\|_1
\end{aligned}$$

Recall that $\beta(T^n) \le \beta(T)^n$ and $\kappa(T^n) \le \frac{\kappa(T)}{1-\beta}$. Applying the one step inequality above to T^n yields the result.

We are now in position to apply Theorem 2.4.4 to obtain the following theorem.

Theorem 3.2.3. *If $T\colon [0,1] \circlearrowleft$ is an expanding interval map, then the associated Frobenius-Perron operator $\mathcal{P}$ acting on $\mathcal{BV}$ satisfies the Lasota-Yorke inequality in eq.* (3.2.1) *and is quasi-compact with spectral radius* 1 *and essential spectral radius $\beta(T)$.*

Proof. The Lasota-Yorke inequality has been proved over the course of this section. $\mathcal{BV}$ is compactly embedded into $L^1([0,1],\lambda)$. Therefore, Theorem 2.4.4 applies and we obtain the claimed essential spectral radius. $\square$

Chapter 4

Historical Interlude

In the last two sections we have investigated the statistical properties of certain expanding interval maps. Historically expanding interval maps were the first examples to be analyzed within the framework that we have described. The seminal paper in the area is [16] where in Lasota and Yorke show that for an expanding interval map T with finitely many branches the associated Frobenius-Perron operator $\mathcal{P}$ viewed as an operator on functions of bounded variation has a unique fixed point η^*. Further for any function η in L^1

$$\frac{1}{n}\sum_{k=0}^{n-1}\mathcal{P}^k\eta \to \eta^*.$$

Many authors proceeded to refine the results for expanding interval maps, in [22] Marek Rychlik showed that the Frobenius-Perron operator associated to a expanding interval map with countably many branches is quasi-compact on BV and further characterized the peripheral spectrum.

Any smooth expanding map from the circle to the circle induces an expanding interval map by cutting the circle at a point and viewing the result as the unit interval. The next development that we would like to highlight was the consideration of smooth multidimensional maps. Here the maps were required to be diffeomorphisims as in the case of circle maps but manifolds of dimension greater then one were considered. In [21] David Ruelle showed that for a diffeomorphism f of a manifold M acting in the neighborhood of an Axiom-A attractor Λ, there exists an invariant measure μ supported on Λ such that f exhibits exponential decay of correlations on $C^1(M))$ with respect to μ. The argument proceeds by constructing a collection of test functions which are continuous on cells of a Markov partition and constant along stable manifolds. Measures are viewed as functionals on this space of test functions and the action on measures adjoint to the Koopman operator is shown to have a uniform limit on probability measures, which is a projection onto a simplex with finitely many ergodic extreme points. This approach begins to hint at the importance of functional analytic tools paired with anisotropic spaces of observables.

The study of diffeomorphisims acting on manifolds partially generalized the expanding interval map examples to higher dimensions however the requirement that maps be globally smooth is a notable restriction. In [24] Young introduced a method for piecewise hyperbolic maps that allowed for the computation of mixing rates. Here a product structure is identified. Projecting the dynamics along stable manifolds yields a Markov expanding factor which can be analyzed via a coupling argument. The decay rates for the expanding factor are then lifted to the full dynamics. This method proved to be quite flexible and allowed for the treatment of maps that failed to be uniformly hyperbolic. For these maps a set Λ is identified such that returns to this set enjoy uniform hyperbolicity. The product structure, expanding factor, and coupling argument then yield subexponential rates of decay of correlations. In [25] the methods pertaining to subexponential decay rates were generalized to an abstract framework which influenced much of the following work on maps that display non-uniform hyperbolicity or non-uniform expansion. One key feature that this paper brings to the forefront is that for a non-uniform map if one can identify a set such that returns to this set are uniform (that is uniformly hyperbolic or expanding) then the rate of decay of correlations for the map can be bounded above in terms of the tail probabilities for the time of return. The methods introduced by Young in both papers are commonly referred to as the *Young tower method*.

One feature of the Young tower method that is noteworthy is the identification and analysis of the expanding factor. Decay rates for a space of observables supported on the domain of a full hyperbolic map must be obtained through a lifting argument which connects decay rates for the full map to the decay rates of the expanding factor. The method does not allow for observables on the full space to be analyzed directly. In [4] Blank, Keller, and Liverani analyzed Anosov diffeomorphisims. They constructed Banach spaces of observables that possessed Hölder regularity along a stable direction but were only bounded and measurable on the whole space. Distributions in the dual space of the observables were constructed such that the Frobenius-Perron operator associated to the diffeomorphism extended to a quasi-compact operator on the space of distributions. Both the space of observables and the space of distributions were supported on the domain of the diffeomorphism. Each space had different regularity properties with respect to stable and unstable directions for the diffeomorphism and as such are referred to as *anisotropic Banach spaces*. These spaces generalize the work of Ruelle in that observables are no longer constant along the stable directing which can be viewed as passing to an expanding factor since values of observables depend only on the “unstable coordinate”. Note that once again the muti-dimensional case has been treated with a new method by first addressing the globally smooth case.

In [7] Demers and Liverani introduced anisotropic Banach spaces which were flexible enough to accommodate a class of piecewise uniformly hyperbolic maps. Here we see the anisotropic Banach space method progressing from globally smooth and uniformly hyperbolic examples to piecewise smooth uniformly hyperbolic examples.

A major technical issue related to applying the anisotropic Banach space method to maps that are non-uniformly hyperbolic is finding the appropriate tool to relate uniformly hyperbolic returns to some set to the global decay rates. The Young tower method seems to depend critically on passing to the expanding factor and then lifting results to the full map. In the Young tower method recurrence to the “good set” is treated via a coupling argument as we have mentioned. In [23] Sarig addressed the issue of recurrence to a “good set” from the prospective of renewal theory. Classical probability methods relating to renewal theory analyze return time probabilities for associated to a renewal process via generating functions. In [11] Gelfond proved a renewal theorem which provided estimates of return time probabilities with very tight error bounds. Sarig proved an analog of this theorem that pertained to operators on a Banach space rather than probabilities. The renewal theorem was shown to be applicable to Frobenius-Perron operator associated to Markov maps with a cell of the Markov partition possessing return times with polynomial tail probabilities. The renewal theorem then provides a polynomial estimate for the decay rate of the map and the error control is sufficient to prove that the rate is sharp. Note that the Young tower method in general only provides an upper bound on the rate of decay. In [13] Gouezël refined the results of Sarig.

The renewal theorem has recently been applied in conjunction with the anisotropic Banach space method to treat non-uniformly hyperbolic maps, see for example [19, 18].

The next chapter of this thesis is an application of the anisotropic Banach space method in conjunction with renewal theory to a class of Generalized Baker's Transformations that are piecewise smooth maps of the unit square originally introduced in [6] by Bose. In [5] Bose and Murray obtained rates of decay of correlations for Generalized Baker's via Young tower methods. In this thesis we recover the rates for a space of anisotropic observables supported on the unit square.

Chapter 5

Generalized Baker's Transformations

A generalized baker's transformation (GBT) is a piecewise continuous area preserving transformation of the unit square. Figure 5.1 depicts a GBT and a few important features.

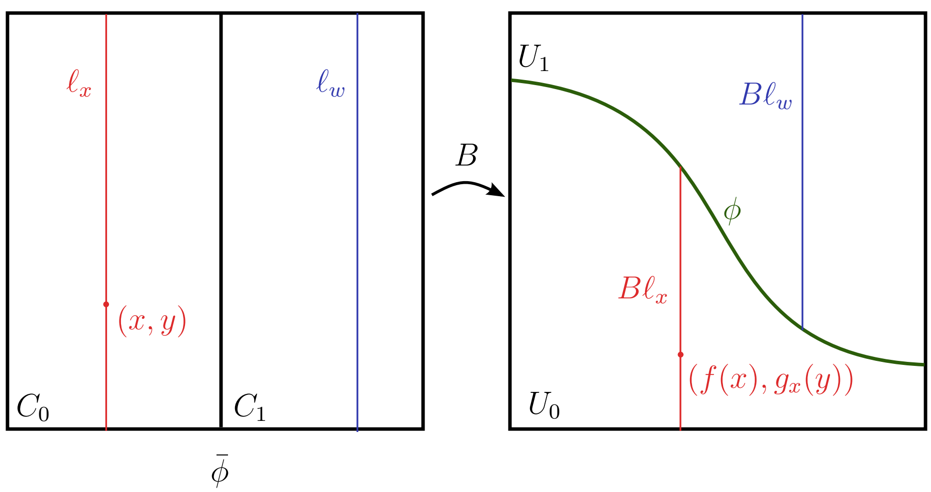

Figure 5.1: The key structures required to define a GBT

The fundamental object in fig. 5.1 is the *cut function* $\phi\colon [0,1] \to [0,1]$. We define several objects in terms of ϕ. Let $\bar{\phi}$ denote the area below ϕ. Let C_0 and C_1 denote the closed rectangles respectively to the left and right of the line $x = \bar{\phi}$. Let U_0 and

U_1 denote the closed regions above and below the graph of the cut function. The GBT B associated to ϕ, is a piecewise continuous map with bijective branches B_0 mapping C_0 onto U_0 and B_1 mapping C_1 onto U_1.

The map B is also a skew product meaning that there exist functions f and g such that $B(x, y) = (f(x), g(x, y))$. We will refer to f as the *factor map* and g as the *fibre map*. To emphasize the skew product structure we define for each $x \in [0, 1]$ the map $g_x(y) = g(x, y)$. The skew product equation then becomes $B(x, y) = (f(x), g_x(y))$. We will prefer this notation from here forward. The branches of B will also split as $B_i(x, y) = (f_i(x), g_{x,i}(y))$. An important consequence of the skew product structure is that B maps the vertical line over x into the vertical line over $f(x)$. It will be convenient to let ℓ_x denote the vertical line over x, with this notation our last comment becomes $B\ell_x \subset \ell_{f(x)}$. Since $BC_i = U_i$, $B\ell_x \subset \ell_{f(x)}$ and B_i is invertible we have

$$B\ell_x = B\left(\ell_x \cap C_i\right) = B\ell_x \cap BC_i = \ell_{f(x)} \cap U_i.$$

We will impose the additional restrictions that f is non-decreasing and for each $x \in [0, 1]$ the map g_x is affine.

The rectangle bounded on the left by ℓ_x in fig. 5.1 maps on to the region bounded by $B\ell_x$ and the graph of ϕ. The area of the first region is x. Since $B\ell_x \cap U_0 = \ell_{f(x)}$ the area of the second region is $\int_0^{f(x)} \phi(t)\, dt$. Since B is area preserving we have for any $x \in \left[0, \bar{\phi}\right]$

$$x = \int_0^{f_0(x)} \phi(t)\, dt; \tag{5.0.1}$$

Similarly, for any $x \in \left[\bar{\phi}, 1\right]$ we have

$$x - \bar{\phi} = \int_0^{f_1(x)} 1 - \phi(t)\, dt. \tag{5.0.2}$$

Note that $f_0\left(\bar{\phi}\right) \neq f_1\left(\bar{\phi}\right)$ and that the top of C_0 and the bottom of C_1 both map onto the graph of $\bar{\phi}$. It follows from these observations that we cannot obtain a well-defined bijective map B with branches B_0 and B_1 on all of $[0, 1]^2$. We will see in the next section that we can define such a map B on $[0, 1)^2$.

To motivate the name 'baker's map' one thinks of the unit square as two dimensional bread dough that is kneaded by a baker who first slices the dough into two pieces, flattens the left half, and then flatteners the right half on top of the already flattened left to make a new square of dough.

5.1 GBTs Defined

In this section we make the definition of a GBT precise and make a few easy observations.

Definition 5.1.1. Given a Borel measurable function [1] $\phi\colon [0,1] \to [0,1]$, which we call a *cut function*, let $\bar{\phi} = \int_0^1 \phi(t)\,dt$ and define sets

$$C_0 = \left[0, \bar{\phi}\right] \times [0,1], \quad U_0 = \left\{(x,y) \in [0,1]^2 : 0 \le y \le \phi(x)\right\},$$
$$C_1 = \left[\bar{\phi}, 1\right] \times [0,1], \quad U_1 = \left\{(x,y) \in [0,1]^2 : \phi(x) \le y \le 1\right\}.$$

For $i = 0, 1$ define $B_i\colon C_i \to U_i$ by $B_i(x,y) = (f_i(x), g_{x,i}(y))$, where f_0 satisfies [2] eq. (5.0.1), f_1 satisfies [3] eq. (5.0.2), and

$$g_{x,0}(y) = \phi\left(f(x)\right) y, \tag{5.1.1}$$
$$g_{x,1}(y) = \left[1 - \phi\left(f(x)\right)\right] y + \phi\left(f(x)\right). \tag{5.1.2}$$

Finally let

$$\tilde{C}_0 = \left[0, \bar{\phi}\right) \times [0,1), \quad \tilde{U}_0 = \left\{(x,y) \in [0,1)^2 : 0 \le y \le \phi(x)\right\},$$
$$\tilde{C}_1 = \left[\bar{\phi}, 1\right) \times [0,1), \quad \tilde{U}_1 = \left\{(x,y) \in [0,1)^2 : \phi(x) \le y \le 1\right\},$$

and define [4] $B\colon [0,1)^2 \circlearrowleft$ by

$$B|_{\tilde{C}_i}(x,y) = \left(f(x), g_x(y)\right)|_{\tilde{C}_i} = \left(f_i(x), g_{x,i}(y)\right) = B_i(x,y).$$

The map $B\colon [0,1)^2 \circlearrowleft$ is a *generalized baker's transformation* associated to ϕ, The map f is the *expanding factor* and g_x are the *fibre maps.*

[1] It is best to exclude *trivial* cut functions such that $\lambda\left\{\phi(t) = 0\right\} = 1$ or $\lambda\left\{\phi(t) = 1\right\} = 1$.

[2] Applying the intermediate value theorem to $s \mapsto \int_0^s \phi(t)\,dt$ shows that this function maps $[0,1]$ onto $\left[0, \bar{\phi}\right]$ and therefore possesses at least one right inverse.

[3] A similar argument for existence applies.

[4] The defining equation for B implicitly defines $f\colon [0,1) \circlearrowleft$ and $g_x\colon [0,1) \to [0,1)$ by $f(x) = f_0(x)$ and $g_x(y) = g_{x,0}(y)$ for all $x \in \left[0, \bar{\phi}\right)$, and $f(x) = f_1(x)$ and $g_x(y) = g_{x,1}(y)$ for all $x \in \left[\bar{\phi}, 1\right)$.

Remark 5.1.2. The iterates of B are also skew products. For each $k \geq 1$ we have the following skew product formula $B^k(x,y) = \left(f^k(x), g_x^{(k)}(y)\right)$, where the functions $g_x^{(k)} \colon [0,1) \to [0,1)$ satisfy the recurrence relation $g_x^{(k)}(y) = g_{f^{k-1}(x)}\left(g_x^{(k-1)}(y)\right)$, with $g_x^{(0)}(y) = y$. Applying eqs. (5.1.1) and (5.1.2) the recurrence splits into the following cases,

- if $f^{k-1}(x)$ is in $[0,\bar{\phi})$, then

$$g_x^{(k)}(y) = \phi\left(f^k(x)\right) g_x^{(k-1)}(y); \tag{5.1.3}$$

- if $f^{k-1}(x)$ is in $[\bar{\phi},1]$, then

$$\begin{aligned} g_x^{(k)}(y) = &\left[1 - \phi\left(f^k(x)\right)\right] g_x^{(k-1)}(y) \\ &+ \phi\left(f^k(x)\right). \end{aligned} \tag{5.1.4}$$

In the following chapters it will be convenient to have some notation for partitions. Before making definitions that are specific to GBTs we will briefly review the standard terminology and notation associated with partitions.

A partition $\mathcal{Q}$ of a set X is a collection of non-empty pairwise disjoint subsets of X such that $X = \bigcup \mathcal{Q}$. If $T \colon X \circlearrowleft$ is a map on X and $\mathcal{Q}$ is a partition of X then we define

$$T^{-1}\mathcal{Q} = \left\{T^{-1}E : E \in \mathcal{Q}\right\},$$

this collection of subsets of X is also a partition [5] which we call the *pull-back* of $\mathcal{Q}$ under T. If $T \colon X \circlearrowleft$ is invertible [6] and $\mathcal{Q}$ is a partition of X, then

$$T\mathcal{Q} = \{TE : E \in \mathcal{Q}\}$$

is also a partition, [7] called the *push-forward* of $\mathcal{Q}$ under T. If $\mathcal{Q}$ and $\mathcal{R}$ are both

[5] If E and F are in $\mathcal{Q}$ then $E \cap F = \emptyset$ and thus $T^{-1}E \cap T^{-1}F = T^{-1}(E \cap F) = T^{-1}\emptyset = \emptyset$ so the collection is pairwise disjoint. The covering property is a simmilar application of the fact that T^{-1} is an automorphism of the boolean algebra $(\wp(X), \cup, \cap, \cdot^c)$.

[6] If T is not invertible then the collection $T\mathcal{Q}$ need not be a partition. The doubling map and the partition $\{[0,1/2],[1/2,1)\}$ provide and example where the pairwise disjointness property fails. For any injective map $T\mathcal{Q}$ is a partition of TX.

[7] If T is invertible, then $T^{-1} \colon X \circlearrowleft$ is a map on X and $\left(T^{-1}\right)^{-1}\mathcal{Q}$ is a partition by previous arguments.

partitions of a set X, then we define

$$\mathcal{Q} \vee \mathcal{R} = \{E \cap F : E \in \mathcal{Q}, F \in \mathcal{R}, E \cap F \neq \emptyset\},$$

this collection is once again a partition which is called the *join* of $\mathcal{Q}$ and $\mathcal{R}$. [8]

Definition 5.1.3. If ϕ is a cut function with associated GBT B, then define [9]

$$\begin{aligned} \tilde{\mathcal{Z}}_B^1 &= \left\{\tilde{C}_0, \tilde{C}_1\right\}, & \tilde{\mathcal{Z}}_B^{-1} &= B\tilde{\mathcal{Z}}_B^1 = \left\{\tilde{U}_0, \tilde{U}_1\right\}, \\ \mathcal{Z}_B^1 &= \left\{Cl(\tilde{C}) : \tilde{C} \in \tilde{\mathcal{Z}}_B^1\right\} = \{C_0, C_1\}, & \mathcal{Z}_B^{-1} &= B\mathcal{Z}_B^1 = \{U_0, U_1\}, \end{aligned}$$

and for each $k \geq 2$, define [10]

$$\begin{aligned} \tilde{\mathcal{Z}}_B^k &= \tilde{\mathcal{Z}}_B^1 \vee B^{-1}\tilde{\mathcal{Z}}_B^{k-1}, & \tilde{\mathcal{Z}}_B^{-k} &= \tilde{\mathcal{Z}}_B^{-1} \vee B\tilde{\mathcal{Z}}_B^{-(k-1)}, \\ \mathcal{Z}_B^k &= \left\{Cl(\tilde{C}) : \tilde{C} \in \tilde{\mathcal{Z}}_B^k\right\} & \mathcal{Z}_B^{-k} &= B^k\mathcal{Z}_B^k, \end{aligned}$$

For convenience we define

$$\tilde{\mathcal{Z}}_B^0 = \left\{[0,1)^2\right\} \quad \mathcal{Z}_B^0 = \left\{Cl(\tilde{E}) : \tilde{E} \in \tilde{\mathcal{Z}}_B^0\right\} = \left\{[0,1]^2\right\}.$$

For any $k \in \mathbb{Z}$ and $(x,y) \in [0,1)^2$ we abuse notation and define $\tilde{\mathcal{Z}}_B^k(x,y)$ to be the unique [11] cell of $\tilde{\mathcal{Z}}_B^k$ that contains (x,y). For each $k \in \mathbb{Z}$ and $(x,y) \in [0,1)^2$ define $\mathcal{Z}_B^k(x,y)$ to be the unique [12] set in $\mathcal{Z}_B^k$ such that $\mathcal{Z}_B^k(x,y) \supset \tilde{\mathcal{Z}}_B^k(x,y)$.

Remark 5.1.4. Notice that all of the objects in Definition 5.1.3 are determined by $\tilde{\mathcal{Z}}_B^1$ and $\tilde{\mathcal{Z}}_B^0$. All of the objects are well defined as long as for all $k \geq 1$ and $\tilde{C} \in \tilde{\mathcal{Z}}_B^k$, if $\tilde{E} \in \tilde{\mathcal{Z}}_B^k$ and $\tilde{E} \subseteq Cl(\tilde{C})$, then $\tilde{E} = \tilde{C}$. We could define all of the objects with

[8] Pairwise disjointness: $(E_1 \cap F_1) \cap (E_2 \cap F_2) = (E_1 \cap E_2) \cap (F_1 \cap F_2) = \emptyset \cap \emptyset = \emptyset$. Covering property:

$$\bigcup \mathcal{Q} \vee \mathcal{R} = \bigcup_{E \in \mathcal{Q}} \bigcup_{F \in \mathcal{R}} E \cap F = \bigcup_{E \in \mathcal{Q}} \left[E \cap \bigcup_{F \in \mathcal{R}} F\right] = \bigcup_{E \in \mathcal{Q}} [E \cap X] = X \cap X = X.$$

[9] The sets $\tilde{C}_i$, C_i, $\tilde{U}_i$, and U_i are the sets in Definition 5.1.1

[10] Here $Cl(\tilde{C})$ denotes the closure of $\tilde{C}$ viewed as a subset of $\mathbb{R}^2$ with the standard topology.

[11] The existence and uniqueness of $\tilde{\mathcal{Z}}_B^k(x,y)$ is equivalent to the fact that $\tilde{\mathcal{Z}}_B^k$ is a partition of $[0,1)^2$.

[12] The for $k \leq -1$ existence of $\mathcal{Z}_B^k(x,y)$ is clear from the definition. Uniqueness of $\mathcal{Z}_B^k(x,y)$ could fail if there were a cell in $\tilde{\mathcal{Z}}_B^k$ with empty interior, however this only occurs for trivial cut functions with $\lambda\{\phi(t) = 0\} = 1$ or $\lambda\{\phi(t) = 1\} = 1$.

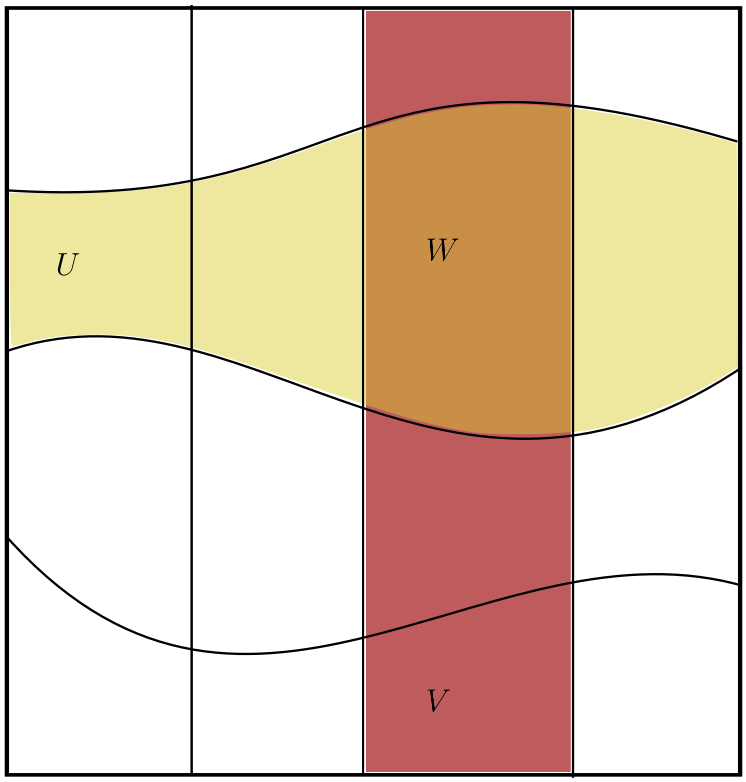

Figure 5.2: In the figure above the closed rectangle V is an element of $\mathcal{Z}_B^2$. Removing the top and right edges of V yields the set $\tilde{V}$ which is an element of $\tilde{\mathcal{Z}}_B^2$. Similarly, the closed strip U is an element of $\mathcal{Z}_B^{-2}$. By removing the top curve and right edge of U we obtain $\tilde{U}$, which is an element of $\tilde{\mathcal{Z}}_B^{-2}$. The set $\tilde{W} = \tilde{U} \cap \tilde{V}$ is an element of $\tilde{\mathcal{Z}}_B^2 \vee \tilde{\mathcal{Z}}_B^{-2}$. Lastly $W = U \cap V$.

non-negative exponent for any map T and associated partitions $\tilde{\mathcal{Z}}_T^1$ and $\tilde{\mathcal{Z}}_T^0$ provided the closure property is satisfied. All of the objects make sense if T is invertible.

Lemma 5.1.5. *Below we record a few useful properties of the objects from Definition 5.1.3. Let k be grater than or equal to* 0*.*

1. *Every set in $\tilde{\mathcal{Z}}_B^k$ is a column of the form $\tilde{C} = [a, b) \times [0, 1)$ for some $a \neq b \in [0, 1)$.*

2. *If $\tilde{C} \in \tilde{\mathcal{Z}}_B^k$, then for each $0 \leq j \leq k$, the set $B^j\tilde{C}$ is a cell of the partition $\tilde{\mathcal{Z}}_B^{k-j} \vee \tilde{\mathcal{Z}}_B^{-j}$.*

3. *If $\tilde{U} \in \tilde{\mathcal{Z}}_B^{-k}$, then for each $0 \leq j \leq k$, the set $B^{-j}\tilde{U}$ is a cell of the partition $\tilde{\mathcal{Z}}_B^{j} \vee \tilde{\mathcal{Z}}_B^{-(k-j)}$.*

4. *If ℓ_x is the vertical line over x and $\tilde{U} \in \tilde{\mathcal{Z}}_B^{-k}$ then $\ell_x \cap \tilde{U} = \{x\} \times [c,d)$ for some $c \neq d \in [0,1)$.*

5.2 Intermittent Baker's Transformations

In [5] a class of GBTs were identified that display polynomial rates of decay of correlations for holder observables. The class is defined by restricting the allowed cut functions.

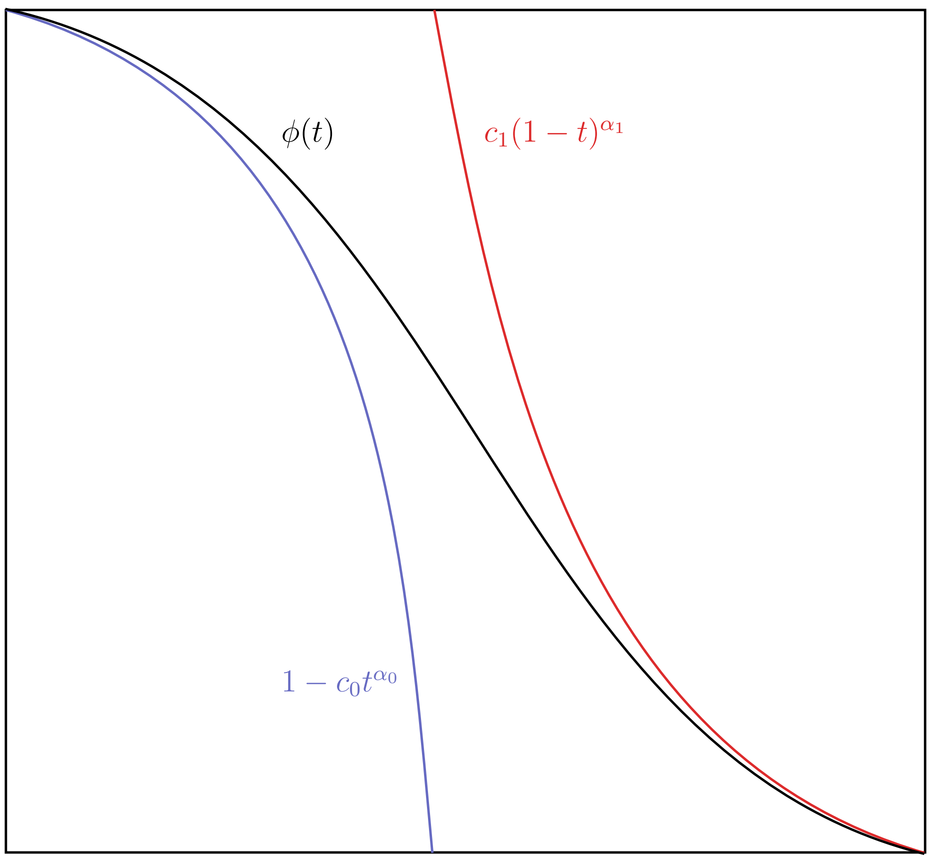

Figure 5.3: An intermittent cut function ϕ is a smooth decreasing map of the interval that first order contacts with power functions at zero and one.

Definition 5.2.1. We call a function $\phi\colon [0,1] \to [0,1]$ an *intermittent cut function* if,

1. ϕ is smooth and strictly decreasing.

2. There exist $\alpha_0 > 0$, $c_0 > 0$, and $h_0\colon [0,1] \to [0,1]$ with $h_0(0) = 0$ and $Dh_0(t) \in o\left(t^{\alpha_0 - 1}\right)$ such that
$$1 - \phi(t) = c_0 t^{\alpha_0} + h_0(t). \tag{5.2.1}$$

3. There exist $\alpha_1 > 0$, $c_1 > 0$ and $h_1\colon [0,1] \to [0,1]$ with $h_1(0) = 0$ and $Dh_1(t) \in o\left(t^{\alpha_1 - 1}\right)$ such that
$$\phi(1 - t) = c_1 t^{\alpha_1} + h_1(t). \tag{5.2.2}$$

Henceforth we will only consider GBTs with intermittent cut functions.

Definition 5.2.2. If ϕ is a intermittent cut function, then the associated GBT $B(x, y) = (f(x), g_x(y))$ is a *intermittent baker's transformation* (IBT).

Lemma 5.2.3. *If $B(x, y) = (f(x), g_x(y))$ is an IBT, then f is uniquely determined by ϕ and has two smooth onto branches $f_0 = f|_{[0,\bar{\phi})}$ and $f_1 = f|_{[\bar{\phi},1]}$ such that $f_0(0) = 0$ and*
$$Df_0(x) = \left[\phi\left(f(x)\right)\right]^{-1}; \tag{5.2.3}$$
$f_1(1) = 1$ and
$$Df_1(x) = \left[1 - \phi\left(f(x)\right)\right]^{-1}. \tag{5.2.4}$$

Proof. By Definition 5.2.1 we know that ϕ is strictly decreasing and thus, for all $t \in (0, 1)$, we have $\phi(t) > 0$ and $1 - \phi(t) > 0$. If f and $\hat{f}$ both satisfy eq. (5.0.1), then for all $x \in [0, \bar{\phi})$,
$$0 = \int_{f(x)}^{\hat{f}(x)} \phi(t)\, dt$$
Since $\phi(t) > 0$ this can only occur if $f(x) = \hat{f}(x)$. An analogous argument shows that for all $x \in [\bar{\phi}, 1]$, $f(x) = \hat{f}(x)$, therefore the uniqueness claim is proved.

The function $A\colon [0, 1] \to \mathbb{R}$ defined by
$$A(x) = \int_0^x \phi(t), dt$$
is strictly increasing, smooth, and has image $[0, \bar{\phi}]$. A is invertible and $A^{-1}\colon [0, \bar{\phi}] \to [0, 1]$ satisfies eq. (5.0.1), therefore $f_0 = A^{-1}|_{[0,\bar{\phi})}$. The derivative condition in eq. (5.2.3) follows from the inverse function theorem. An analogous argument proves bijectivity of f_1 and eq. (5.2.4). $\square$

Remark 5.2.4. Differentiating eqs. (5.1.3) and (5.1.4) with respect to y and applying eqs. (5.2.3) and (5.2.4) and the chain rule yields

$$\partial_y g_x^{(k)}(y) = \left[Df^k\right]^{-1}(x). \tag{5.2.5}$$

If $f^{k-1}(x) \in [0, \bar{\phi})$, then differentiating eq. (5.1.3) with respect to x yields,

$$\begin{aligned} \partial_x g_x^{(k)}(y) = D\phi\left(f^k(x)\right) Df^k(x) g_x^{(k-1)}(y) \\ + \phi\left(f^k(x)\right) \partial_x g_x^{(k-1)}(y). \end{aligned} \tag{5.2.6}$$

If $f^{k-1}(x) \in [\bar{\phi}, 1]$, then differentiating eq. (5.1.4) with respect to x yields,

$$\begin{aligned} \partial_x g_x^{(k)}(y) = D\phi\left(f^k(x)\right) Df^k(x) \left[1 - g_x^{(k-1)}(y)\right] \\ + \left[1 - \phi\left(f^k(x)\right)\right] \partial_x g_x^{(k-1)}(y). \end{aligned} \tag{5.2.7}$$

Lemma 5.2.5. *If $B(x, y) = (f(x), g_x(y))$ is an IBT, then f has a unique period 2 orbit. Specifically there exist $0 < p < \bar{\phi} < q < 1$ such that $f(p) = q$ and $f(q) = p$.*

Proof. Apply eq. (5.0.1) with $x = 0$ and recall that $\phi(t) > 0$ for all $t \in (0, 1)$ to conclude that $f(0) = 0$. Argue analogously with $x = 1$ and eq. (5.0.2) to conclude that $f(1) = 1$.

It follows from eqs. (5.2.3) and (5.2.4) that for $i = 0, 1$, $Df_i > 1$. Therefore 0 and 1 are the only fixed points of f. Since f has two increasing onto branches, f^2 has four increasing onto branches. By the intermediate value theorem applied to $f^2(x) - x$ each branch has a fixed point. On each branch of f^2, $Df^2 > 1$ by the chain rule. This bound implies that the fixed point of each branch is unique.

The left most branch of f^2 fixes 0 and the right most branch fixes 1. Let $p < q$ denote the fixed points of the central branches of f^2. Let L and R denote the closures of the domains of the left and right central branches of f^2 respectively. L and R are closed intervals contained in $(0, 1)$ with $\sup L = \inf R = \bar{\phi}$. Since each branch of f^2 is increasing and onto we have $\left(f^2|_L\right)^{-1} L \subset int\,(L)$ and $\left(f^2|_R\right)^{-1} R \subset int\,(R)$. Since f^2 fixes p and q we have $p \in \left(f^2|_L\right)^{-1} L$ and $q \in \left(f^2|_R\right)^{-1} R$, therefore $0 < p < \bar{\phi} < q < 1$ as desired. □

5.3 Associated Induced Map

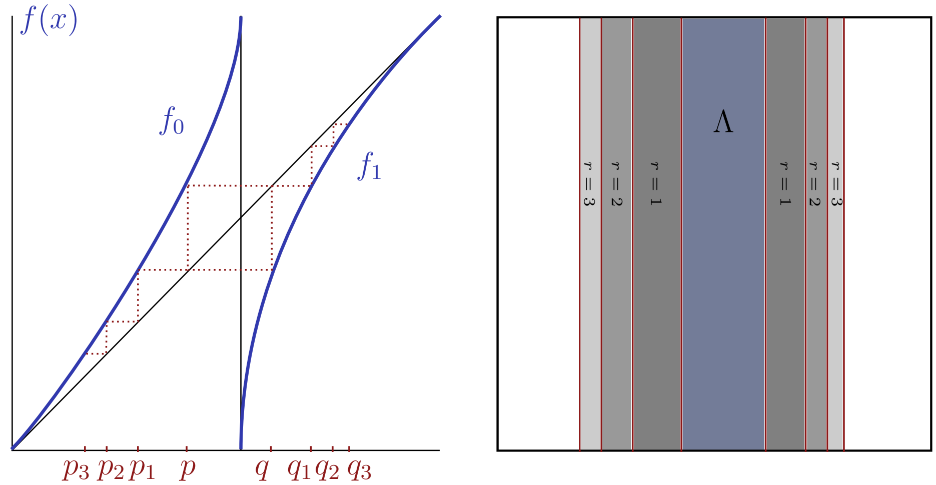

Figure 5.4: On the left we see a period-2 orbit $\{p, q\}$ for the map f and sequences p_k and q_k that are mapped by B^k onto p and q respectively. On the right we see the inducing set Λ flanked on either side by vertical columns that return to Λ under B^r.

Definition 5.3.1. If $B(x, y) = (f(x), g_x(y))$ is an IBT, and $\{p, q\}$ is the unique period-2 orbit of f, then we define,

- A return set $\Lambda = [p, q) \times [0, 1)$, which we refer to as the *base*.
- The *return time function* $r \colon \Lambda \to \mathbb{N}$ defined by

$$r(x, y) = \inf \{n \in \mathbb{N} : B^n(x, y) \in \Lambda\}. \tag{5.3.1}$$

- The *induced map* $T \colon \Lambda \circlearrowleft$ defined by

$$T(x, y) = B^{r(x,y)}(x, y). \tag{5.3.2}$$

- The *base measure* μ defined by

$$\mu(E) = \frac{\lambda(E \cap \Lambda)}{\lambda(\Lambda)}. \tag{5.3.3}$$

Lemma 5.3.2. *The value of the return time function $r(x, y)$ is independent of y.*

Proof. Since the membership of a point (x, y) in Λ is independent of y and each vertical line ℓ_t is mapped into the vertical line $\ell_{f(t)}$ we have that $r|_{\ell_t}$ is constant for each $t \in [p, q]$, which completes the proof. □

Remark 5.3.3. In light of Lemma 5.3.2 we will abuse notation and define $r\colon [p, q] \to \mathbb{N}$ by $r(t) = r|_{\ell_t}$ and feel free to use $r(x)$ to denote the return time of either a point $(x, y) \in \Lambda$ to Λ under B or the return time of a point $x \in [p, q]$ to $[p, q]$ under f.

Lemma 5.3.4. *The map T is a skew product with factor map $u\colon [p, q] \circlearrowleft$ and fibre map $v_x\colon \Lambda \to [0, 1]$ defined by*

$$u(x) = f^{r(x)}(x) \tag{5.3.4}$$

$$v_x(y) = g_x^{(r(x))}(y) \tag{5.3.5}$$

Proof. Fix a point $(x, y) \in \Lambda$. By Lemma 5.3.2 and Remark 5.1.2 we see that

$$T(x, y) = B^{r(x,y)}(x, y) = B^{r(x)}(x, y) = \left(f^{r(x)}(x), g_x^{(r(x))}(y)\right),$$

as desired. □

Remark 5.3.5. Combining eqs. (5.2.5), (5.3.4) and (5.3.5) we obtain

$$\partial_y T_x^k = \left[Du^k\right]^{-1}. \tag{5.3.6}$$

Definition 5.3.6. Let $p_0 = p$ and $p_{j+1} = f_0^{-1}(p_j)$ for $j \geq 0$ and define

$$J_j = [p_j, p_{j-1}), \quad j \geq 1; \tag{5.3.7}$$

$$I_j = f_1^{-1} J_k, \quad j \geq 1; \tag{5.3.8}$$

$$\phi_j = \phi|_{J_j}, \quad j \geq 1. \tag{5.3.9}$$

Similar definitions could be made for the with q and the roles of f_0 and f_1 reversed, however we will only prove lemmas about the objects we have just defined and argue that the corresponding results follow by analogous arguments.

Lemma 5.3.7. *If $x_0 \in [0, \bar{\phi})$ and $x_{j+1} = f_0^{-1}(x_j)$ for all $j \geq 0$, then (x_j) is strictly decreasing and convergent to 0. In particular p_k decreases to 0.*

Proof. For all $t \in (0, 1)$, $1 - \phi(t) > 0$. From eq. (5.2.3) it follows that for all $t \in (0, \bar{\phi})$, $Df_0(t) > 1$. We have also observed that $f_0(0) = 0$. It follows from a standard mean value theorem argument that $f_0(t) > t$ for all $t \in (0, \bar{\phi})$. Since the domain of f_0 is $[0, \bar{\phi})$ we have, for all $t \in (0, \bar{\phi})$, $0 \leq f_0^{-1}(t) < t$. Therefore (x_j) is strictly decreasing and by the monotone convergence theorem for sequences $x_j \to 0$ as $j \to \infty$. □

Corollary 5.3.7.1. *For $j \geq 1$, $J_j \subset [0, p)$, J_{j+1} is to the left of and adjacent to J_j, and J_1 is to the left of and adjacent to $[p, q]$. For $j \geq 1$, $I_k \subset (\bar{\phi}, q)$ is a left closed right open interval, I_{k+1} is to the left of and adjacent to I_k, and I_1 is to the left of and adjacent to $[q, 1]$.*

Proof. The adjacency conditions for J_j follow directly from the fact that the sequence (p_j) is decreasing and the inclusions into $[0, p)$ follow from $p_0 = p$. The conditions on I_k follow by noting that f_1 is continuous, strictly increasing, $f_1(\bar{\phi}) = 0$, and $f_1(q) = p$, therefore order, adjacency, and topology are preserved by the homeomorphism $f_1^{-1}|_{[0,p)} \colon [0, p) \to [\bar{\phi}, q)$. □

Corollary 5.3.7.2. *For each $j \geq 1$ $f_1|_{I_j} \colon I_j \to J_j$ and $f_0|_{J_j} \colon J_j \to J_{j-1}$ are smooth bijections, as is $f_0|_{J_1} \colon J_1 \to [p, q]$*

Proof. The smoothness and bijectivity follow from the inclusions $I_j \subset (\bar{\phi}, q)$ and $J_j \subset (0, p)$. The matching of domains and images follows from the definitions of p, q, and (p_j). □

Lemma 5.3.8. *For all $(x, y) \in \left[\bar{\phi}, q\right] \times [0, 1]$,*

$$Du(x) = [1 - \phi(f(x))]^{-1} \prod_{i=2}^{r(x)} \left[\phi\left(f^i(x)\right)\right]^{-1} \tag{5.3.10}$$

and

$$\begin{aligned} \partial_x v_x(y) = &-\frac{D^2u(x)}{[Du(x)]^2}\left[y + \frac{\phi\left(f(x)\right)}{1 - \phi\left(f(x)\right)}\right] \\ &+ \frac{1}{Du(x)}\left[\frac{D\phi\left(f(x)\right)}{\left[1 - \phi\left(f(x)\right)\right]^3}\right]. \end{aligned} \tag{5.3.11}$$

For all $(x, y) \in \left[p, \bar{\phi}\right) \times [0, 1]$,

$$Du(x) = [\phi(f(x))]^{-1} \prod_{i=2}^{r(x)} \left[1 - \phi\left(f^i(x)\right)\right]^{-1} \tag{5.3.12}$$

and

$$\partial_x v_x(y) = \frac{D^2u(x)}{[Du(x)]^2}\left[\frac{1-\phi\left(f(x)\right)y}{\phi\left(f(x)\right)}\right] - \frac{1}{Du(x)}\left[\frac{D\phi\left(f\left(x\right)\right)}{\left[\phi\left(f(x)\right)\right]^3}\right]. \tag{5.3.13}$$

For all $(x,y) \in [p,q]$,

$$\partial_y v_x(y) = \left[Du(x)\right]^{-1}. \tag{5.3.14}$$

Proof. Fix $j \geq 1$ and $x \in I_j$. By Corollary 5.3.7.1 $x \in (\bar{\phi}, q)$. By Corollary 5.3.7.2 we have that, $f(x) \in J_j \subset [0,p)$, for all $i = 2, \cdots, j+1$, $f^{i-1}(x) \in J_{j+2-i} \in [0,\bar{\phi})$, and $f^{j+1}(x) \in [p,q]$. We conclude that $r(x) = j+1$.

Applying eq. (5.3.4) we obtain $u(x) = f_0^j \circ f_1(x)$. Differentiation using the chain rule and applying eq. (5.2.3) we obtain eq. (5.3.10).

Applying eq. (5.1.3) we obtain for all $y \in [0,1]$ and $i = 2, \cdots, j+1$

$$g_x^{(j)}(y) = \phi\left(f^i\left(x\right)\right) g_x^{(j-1)}(y).$$

By applying eq. (5.3.5) it follows that

$$v_x(y) = g_x(y) \prod_{i=2}^{j+1} \phi\left(f^i(x)\right),$$

and by applying eqs. (5.1.1) and (5.3.10) we see that

$$v_x(y) = \frac{\left[1-\phi\left(f(x)\right)\right]y + \phi\left(f(x)\right)}{\left[1-\phi\left(f(x)\right)\right]Du(x)}.$$

Differentiating the equation above with respect to x yields eq. (5.3.11).

We note that for $x \in I'_j$ we can apply eq. (5.1.4) to obtain for all $i = 2, \cdots, j+1$,

$$1 - g_x^{(i)}(y) = \left[1-\phi\left(f\left(x\right)\right)\right]\left[1-g_x^{(i-1)}(y)\right].$$

The identities in eqs. (5.3.12) and (5.3.13) follow by similar arguments utilizing the recurrence relation above. ∎

Definition 5.3.9. Given two $\mathbb{R}$-valued functions on f and g and a point a, we say that f approximately g near a if there exists a neighborhood $N(a)$ of a and $C > 0$ such that for all $x \in N(a)$ we have $C^{-1}g(x) \leq f(x) \leq Cg(x)$, we denote this by

$$f \approx g.$$

Lemma 5.3.10 ([5] Lemma 1)**.** .
If $x_0 \in [0, \bar{\phi})$ and for all $j \geq 0$, $x_{j+1} = f_0^{-1}(x_j)$, then for all $k \geq 0$,

$$x_j \approx \left[\frac{1}{j}\right]^{1/\alpha_0}. \tag{5.3.15}$$

Proof. By Definition 5.2.1 we know that $Dh_0(x) \in o(x^{\alpha_0 - 1})$ near 0 and that $h_0(0) = 0$, from these facts it follows easily that $h_0(x) \in o(x^{\alpha_0})$. It follows from eqs. (5.0.1) and (5.2.1) that for $x \in \left[0, \bar{\phi}\right)$,

$$x - f_0^{-1}(x) = \int_0^x 1 - \phi(t)\, dt = \int_0^x c_0 t^{\alpha_0} \left[1 + \frac{h_0(t)}{c_0 t^{\alpha_0}}\right] dt \approx x^{\alpha_0 + 1}.$$

Therefore,

$$\frac{1}{\left[\frac{1}{y}\right]^{1/\alpha_0} - f_0^{-1}\left(\left[\frac{1}{y}\right]^{1/\alpha_0}\right)} \approx y^{1 + 1/\alpha_0}, \tag{5.3.16}$$

for y sufficiently large. By the mean value theorem applied to $t \mapsto t^{1/\alpha_0}$, for all $y, z > 0$, there exists $\theta \in [0, 1]$ such that

$$\left[\frac{1}{y}\right]^{1/\alpha_0} - \left[\frac{1}{y + z}\right]^{1/\alpha_0} = \frac{1}{\alpha_0}\left[\frac{z}{y(y+z)}\right]\left[\frac{1}{y + \theta z}\right]^{1/\alpha_0 - 1}.$$

It follows from the expression above that

$$\frac{z}{2^{1/\alpha_0}\alpha_0 y^{1 + 1/\alpha_0}} \leq \left[\left[\frac{1}{y}\right]^{1/\alpha_0} - \left[\frac{1}{y + z}\right]^{1/\alpha_0}\right] \leq \frac{z}{\alpha_0 y^{1 + 1/\alpha_0}}, \tag{5.3.17}$$

where the lower bound holds for $y \geq z$. Combining the two relations above we obtain that for some $C \geq 1$

$$\left(\frac{1}{2^{1/\alpha_0}C}\right) z \leq \frac{\left[\frac{1}{y}\right]^{1/\alpha_0} - \left[\frac{1}{y+z}\right]^{1/\alpha_0}}{\left[\frac{1}{y}\right]^{1/\alpha_0} - f_0^{-1}\left(\left[\frac{1}{y}\right]^{1/\alpha_0}\right)} \leq \left(\frac{C}{\alpha_0}\right) z \tag{5.3.18}$$

for y sufficiently large. Let $z = \frac{\alpha_0}{C}$ so that the upper bound in eq. (5.3.18) is 1 and rearrange to obtain

$$f_0^{-1}\left(\left[\frac{1}{y}\right]^{1/\alpha_0}\right) \leq \left[\frac{1}{y+z}\right]^{1/\alpha_0}.$$

Induction yields,

$$f_0^{-k}\left(\left[\frac{1}{y}\right]^{1/\alpha_0}\right) \leq \left[\frac{1}{y+kz}\right]^{1/\alpha_0} \leq \left[\frac{1}{kz}\right]^{1/\alpha_0} = \left[\frac{C}{\alpha_0}\right]^{1/\alpha_0} \left[\frac{1}{k}\right]^{1/\alpha_0}$$

Similarly let $z = 2^{\alpha_0}C$ so that the lower bound in eq. (5.3.18) is 1. Rearrange and apply induction to obtain

$$f_0^{-k}\left(\left[\frac{1}{y}\right]^{1/\alpha_0}\right) \geq \left[\frac{1}{2^{1/\alpha_0}C}\right]^{1/\alpha_0} \left[\frac{1}{k}\right]^{1/\alpha_0}$$

for large y. Finally by Lemma 5.3.7 we may select $N > 0$ such that for any $x_0 \in [0, \bar{\phi})$ we have x_N sufficiently small that setting $y = x_N^{-\alpha_0}$ we may apply the bounds above to obtain the asymptotics in eq. (5.3.15). □

Lemma 5.3.11 (Lemma 2 from [5]). *The constant*

$$\beta = \sup_{t \in [p,q]} \max\{\phi(t), 1 - \phi(t)\} \tag{5.3.19}$$

is less then 1 *and*

$$\left\| [Du]^{-1} \right\|_{\sup} \leq \beta. \tag{5.3.20}$$

Proof. The cut function ϕ is strictly decreasing by Definition 5.2.1 and thus $\phi|_{[p,q]} < 1$ and $(1-\phi)|_{[p,q]} < 1$, therefore $\beta < 1$. By Definition 5.1.1 we have $\phi(t) \leq 1$ and $1 - \phi(t) \leq 1$ for all $t \in [0, 1]$. For all $x \in [p, q]$ we have $f^{r(x)}(x) \in [p, q]$ by Lemma 5.3.2 and Remark 5.3.3. By inspecting eq. (5.3.10) and applying the bounds that we have just observed we see that for all $x \in [p, \bar{\phi})$ we have

$$\left|[Du]^{-1}\right| \leq \left|1 - \phi\left(f^{r(x)}(x)\right)\right| \leq \beta.$$

Similarly by inspecting eq. (5.3.12) we see that for all $x \in [\bar{\phi}, q]$ we have

$$\left|[Du]^{-1}\right| \leq \left|\phi\left(f^{r(x)}(x)\right)\right| \leq \beta.$$

Combining the two bounds above completes the proof. □

Lemma 5.3.12 (Lemma 3 from [5])**.** *There exists $\kappa < \infty$ such that*

$$\left\| \frac{D^2 u}{[Du]^2} \right\|_{\sup} \leq \kappa; \tag{5.3.21}$$

and for any $I \in \mathcal{Z}_u^k$ and $s, t \in I$,

$$\left| \frac{Du^k(t)}{Du^k(s)} - 1 \right| \leq \kappa \left| u^k(t) - u^k(s) \right|. \tag{5.3.22}$$

Proof. By the last displayed expression in Appendix A of [5] there exists a constant $C > 0$ such that for all $s, t \in [p, q]$

$$\left| \log\left(\frac{Du(s)}{Du(t)} \right) \right| \leq C \left| u(x) - u(y) \right|. \tag{5.3.23}$$

From eq. (5.3.23) we see that for each branch w of u we have that $\log \circ Du \circ w^{-1}$ is Lipschitz with constant C. Therefore for each branch w of u and all $t \in [p, q]$

$$\left| D \log\left(Du\left(w^{-1}(t)\right)\right)\right| = \left| \frac{D^2 u}{[Du]^2} \left(w^{-1}(t)\right) \right| < C$$

Since the domains of the branches of u cover $[p, q]$ we have eq. (5.3.21) with C in the place of κ. It follows by applying the mean value theorem that,

$$\begin{aligned}\left|\log\left(\frac{Du^k(s)}{Du^k(t)}\right)\right| &\leq C\left|u^k(x)-u^k(y)\right|\sum_{i=1}^{k}\frac{\left|u^{k-i}(x)-u^{k-i}(y)\right|}{|u^k(x)-u^k(y)|}\\ &\leq C\left|u^k(x)-u^k(y)\right|\sum_{j=0}^{k-1}\left[Du^j(\theta_j)\right]^{-1}\\ &\leq \frac{C}{1-\beta}\left|u^k(x)-u^k(y)\right|\end{aligned}$$

If $|\log(x)| < A$, then $|\log(x)| > \frac{A}{e^A-1}|x-1|$. Therefore,

$$\left|\frac{Du^k(x)}{Du^k(y)}-1\right| \leq \frac{C}{[1-\beta]\left[e^{\frac{C}{1-\beta}}-1\right]}\left|u^k(x)-u^k(y)\right|.$$

Letting $\kappa = \max\left\{C, \frac{C}{[1-\beta]\left[e^{\frac{C}{1-\beta}}-1\right]}\right\}$ completes the proof. □

Lemma 5.3.13. *There exists a constant $\Theta < \infty$ such that*

$$\left|\frac{\partial_x v_x}{Du}\right| \leq \Theta. \tag{5.3.24}$$

Proof. Fix $(x,y) \in \left[\bar{\phi}, q\right] \times [0,1]$. Recall eq. (5.3.11),

$$\begin{aligned}\partial_x v_x(y) = &-\frac{D^2u(x)}{[Du(x)]^2}\left[y+\frac{\phi(f(x))}{1-\phi(f(x))}\right]\\ &+\frac{1}{Du(x)}\left[\frac{D\phi(f(x))}{[1-\phi(f(x))]^3}\right].\end{aligned}$$

Dividing both sides of the equation above by $Du(x)$ yields

$$\begin{aligned}\frac{\partial_x v_x(y)}{Du(x)} = &-\frac{D^2u(x)}{[Du(x)]^2}\frac{1}{Du(x)}y-\frac{D^2u(x)}{[Du(x)]^2}\frac{\phi(f(x))}{Du(x)[1-\phi(f(x))]}\\ &+\left[\frac{D\phi(f(x))}{[1-\phi(f(x))]^3\,[Du(x)]^2}\right].\end{aligned}$$

Recall from eqs. (5.3.20) and (5.3.21) that we have

$$\left\| [Du]^{-1} \right\|_{\sup} \le \beta,$$
$$\left\| \frac{D^2 u}{[Du]^2} \right\|_{\sup} \le \kappa.$$

Applying the triangle inequality and the bounds above we obtain

$$\left| \frac{\partial_x v_x(y)}{Du(x)} \right| \le \kappa\beta + \kappa \left| \frac{\phi\left(f(x)\right)}{Du(x)[1-\phi\left(f(x)\right)]} \right| + \left| \frac{D\phi\left(f(x)\right)}{\left[1-\phi\left(f(x)\right)\right]^3 \left[Du(x)\right]^2} \right|.$$

Recall from eq. (5.3.10) we have

$$Du(x) = [1-\phi(f(x))]^{-1} \prod_{i=2}^{r(x)} \left[\phi\left(f^i(x)\right)\right]^{-1}.$$

Substituting the right hand side of the identity above yields

$$\left| \frac{\partial_x v_x(y)}{Du(x)} \right| \le \kappa\beta + \kappa \left| \prod_{i=1}^{r(x)} \phi\left(f^i(x)\right) \right| + \|D\phi\|_{\sup} \left| \frac{\left[\prod_{i=2}^{r(x)} \phi\left(f^i(x)\right)\right]^2}{1-\phi\left(f(x)\right)} \right| \tag{5.3.25}$$

From Definition 5.2.1 it follows that

$$\prod_{i=1}^{r(x)} \phi\left(f^i(x)\right) \le 1.$$

It remains to bound the ratio in the last term of eq. (5.3.25).

Note that f^n maps J_n onto $[p, q]$ and hence by the mean value theorem there exists $\Theta_n \in J_n$ such that $|J_n|\, Df^n(\Theta) = |[p, q]|$. By the chain rule and eq. (5.2.3) we have that for any $x \in J_n$

$$Df^n(x) = \left[\prod_{i=1}^{n} \phi\left(f^i(x)\right)\right]^{-1}$$

and hence

$$\prod_{i=1}^{n} \phi\left(f^i(\Theta_n)\right) = \frac{|J_n|}{|[p,q]|}.$$

Since ϕ is strictly decreasing and positive f is increasing and hence the product on the left in the equation above is strictly decreasing, from this it follows directly that for all $x \in J_n$,

$$\frac{|J_{n+1}|}{|[p,q]|} \leq \prod_{i=1}^{n} \phi\left(f^i(\Theta_n)\right) \leq \frac{|J_{n-1}|}{|[p,q]|}.$$

Note that,

$$x - f_0^{-1}(x) = \int_0^x 1 - \phi(t)\,dt = \int_0^x c_0 t^{\alpha_0}\left[1 + \frac{h_0(t)}{c_0 t^{\alpha_0}}\right] dt \sim \frac{c_0}{\alpha_0 + 1} x^{\alpha_0+1}.$$

Applying asymptotic above and eq. (5.3.15) it follows that

$$|J_n| \approx \left[\frac{1}{n}\right]^{1+1/\alpha_0}$$

Combining this with the bound on the product above we obtain for all $x \in J_n$

$$\prod_{i=1}^{n} \phi\left(f^i(\Theta_n)\right) \approx \left[\frac{1}{n}\right]^{1+1/\alpha_0}.$$

Similar arguments show that

$$1 - \phi\left(f(x)\right) \approx \frac{1}{n}.$$

We conclude that

$$\frac{\left[\prod_{i=2}^{r(x)} \phi\left(f^i(x)\right)\right]^2}{1 - \phi\left(f(x)\right)} \approx \frac{\left[\left[\frac{1}{n}\right]^{1+1/\alpha_0}\right]^2}{\frac{1}{n}} = \left[\frac{1}{n}\right]^{1+2/\alpha_0},$$

and hence the ratio is bounded. This completes the proof.

$\square$

5.4 Unstable Partitions

Lemma 5.4.1. *Let I'_k denote the intervals in $[p, \bar{\phi})$ that are analogous to $I_k \subset [\bar{\phi}, q)$. Letting*

$$\tilde{\mathcal{Z}}^1_T = \{I_k \times [0,1) : k \geq 1\} \cup \{I'_k \times [0,1) : k \geq 1\}\,, \quad \tilde{\mathcal{Z}}^0_T = \{[p,q) \times [0,1)\} \tag{5.4.1}$$

induces partitions for the induced map $T = B^r$ that are analogous to the ones in Definition 5.1.3 for B.

Proof. The sets in $\tilde{\mathcal{Z}}^1_T$ have the closure property described in Remark 5.1.4, to check that the property holds for $\tilde{\mathcal{Z}}^k_B$ with $k > 1$ is a straightforward application of the definitions. Since T is invertible all of the objects are well defined for negative exponents. □

Lemma 5.4.2. *If $k \geq 1$ and $U \in \mathcal{Z}^{-k}_T$, then for all $t \in [p,q]$, the set*[13] *$\ell_t \cap U$ is non-empty and homeomorphic to $[0,1]$.*

Proof. By Definition 5.1.3 the map $\left(T^{-k}\right)|_U$ is a smooth homeomorphism onto some $C \in \mathcal{Z}^k_T$. It follows from Lemmas 5.1.5, 5.3.4 and 5.4.1 that there exists $s \in [p,q]$ such that

$$T^{-k}\left(\ell_t \cap U\right) = \ell_s. \tag{5.4.2}$$

It follows from eq. (5.3.14) that $T^k|_{\ell_s}$ is affine. Since T^k is invertible we conclude that $\ell_t \cap U = T^k\left(\ell_s\right)$ is the affine image of ℓ_s and hence homeomorphic to $[0,1]$ as desired. □

Definition 5.4.3. For each $k \geq 1$ and $U \in \mathcal{Z}^{-k}_T$ define $U^x = \ell_x \cap U$ and let $|U^x|$ denote the length of this line segment.

Lemma 5.4.4. *For each $k \geq 1$ and $U \in \mathcal{Z}^{-k}_T$, there exists $I \in \mathcal{Z}^k_u$ such that for all $x \in [p,q]$,*

$$|U^x| = \frac{1}{\left[Du^k \circ \left(u^k|_I\right)^{-1}\right](x)} \tag{5.4.3}$$

Proof. We apply eq. (5.4.2) and a change of variables

$$\begin{aligned}|U^x| &= \int_{\ell_x \cap U} dy = \int_{\ell_s} \partial_y v^{(k)}_s(y)\, dy = \int_{\ell_s} \left[Du^k(s)\right]^{-1} dy \\ &= \left[Du^k(s)\right]^{-1}.\end{aligned}$$

[13] recall that ℓ_t is the vertical line over the horizontal coordinate t.

Let $C = T^{-k}U$ which by Lemmas 5.1.5 and 5.4.1 can be written as a product $I \times [0,1]$ where $I \in \mathcal{Z}_u^k$ and thus $u^k|_I$ is invertible. We conclude that $s = \left(u^k|_I\right)^{-1}(t)$, which completes the proof. □

Corollary 5.4.4.1. *If $k \geq 1$ and $U \in \mathcal{Z}_T^{-k}$, then*

$$|U^x| \leq \beta^k. \tag{5.4.4}$$

and

$$\left|\frac{|U^t|}{|U^s|} - 1\right| \leq \kappa \, |t - s| \, . \tag{5.4.5}$$

Proof. The bounds in eqs. (5.4.4) and (5.4.5) follow from eqs. (5.3.20), (5.3.22) and (5.4.3). □

Definition 5.4.5. We say that two points (x, y) and (w, z) in Λ are *backward equivalent* if and only if for all $n \geq 1$, $\tilde{\mathcal{Z}}_T^{-n}(x, y) = \tilde{\mathcal{Z}}_T^{-n}(w, z)$.

Remark 5.4.6. Backward equivalence is an equivalence relation on Λ.

Definition 5.4.7. Let Γ^u denote the partition of Λ into backward equivalence classes. Let $\pi_u \colon \Lambda \to \Gamma^u$ by

$$(x, y) \in \pi_u(x, y).$$

Remark 5.4.8. For all $(x, y) \in \Lambda$,

$$T^{-1}\pi_u(x, y) \subset \pi_u\left(T^{-1}(x, y)\right). \tag{5.4.6}$$

Definition 5.4.9. For each branch ω of u^k we define $G_\omega \colon C^1\left([p, q], [0, 1]\right) \circlearrowleft$ by

$$G_\omega\left(\varphi\right)(x) = v^{(k)}_{\omega^{-1}(x)}\left(\varphi\left(\omega^{-1}(x)\right)\right). \tag{5.4.7}$$

Lemma 5.4.10. *For all branches ω of u^k and $\varphi, \varphi' \in C^1\left([p, q], [0, 1]\right)$,*

$$\left|G_\omega\left(\varphi\right) - G_\omega\left(\varphi'\right)\right| \leq \beta^k \left|\varphi - \varphi'\right|. \tag{5.4.8}$$

Proof. Fix $x \in [p, q]$. By the mean value theorem applied to $v_k x$ together with eqs. (5.3.14) and (5.3.20) we have for some θ between $\varphi\left(\omega^{-1}(x)\right)$ and $\varphi'\left(\omega^{-1}(x)\right)$,

$$\begin{aligned}
\left|G_\omega(\varphi)(x) - G_\omega(\varphi')(x)\right| &= \left|\partial_y v^{(k)}_{\omega^{-1}(x)}(\theta)\right| \left|[\varphi - \varphi']\left(\omega^{-1}(x)\right)\right| \\
&= \frac{\left|[\varphi - \varphi']\left(\omega^{-1}(x)\right)\right|}{\left|Du^k\left(\omega^{-1}(x)\right)\right|} \\
&\leq \beta^k \left|\varphi - \varphi'\right|.
\end{aligned}$$

Since x was arbitrary this completes the proof. □

Lemma 5.4.11. *If ω is a branch of u and φ is in $C^1([p,q],[0,1])$, then*

$$D\left[G_\omega(\varphi)\right](x) = \frac{\partial_x v_x\left(\omega^{-1}(x), \varphi\left(\omega^{-1}(x)\right)\right)}{Du\left(\omega^{-1}(x)\right)} + \frac{D\varphi\left(\omega^{-1}(x)\right)}{\left[Du\left(\omega^{-1}(x)\right)\right]^2}. \tag{5.4.9}$$

Proof. Fix $x \in [p,q]$. By the multivariate chain rule, the inverse function theorem, and eq. (5.3.14), we have[14]

$$\begin{aligned}
D\left[G_\omega(\varphi)\right](x) &= Dv_x\left(\omega^{-1}(x), \varphi\left(\omega^{-1}(x)\right)\right) \\
&= \begin{bmatrix} \partial_x v_x & \partial_y v_x \end{bmatrix} \left(\omega^{-1}(x), \varphi\left(\omega^{-1}(x)\right)\right) \begin{bmatrix} 1 \\ D\varphi \end{bmatrix} \left(\omega^{-1}(x)\right) D\omega^{-1} \\
&= \frac{\partial_x v_x\left(\omega^{-1}(x), \varphi\left(\omega^{-1}(x)\right)\right)}{Du\left(\omega^{-1}(x)\right)} \\
&\quad + \frac{D\varphi\left(\omega^{-1}(x)\right)}{\left[Du\left(\omega^{-1}(x)\right)\right]^2},
\end{aligned}$$

which completes the proof. □

Lemma 5.4.12. *If $(\omega_i)_{i=1}^{\infty}$ is a sequence of branches of u, φ_0 is in $C^1([p,q],[0,1])$, and for $i \geq 1$ we define $\varphi_i = G_{\omega_i}(\varphi_{i-1})$, then*[15] *for all $i \geq 0$,*

$$\left|D\varphi_i\right| \leq \beta^{2i}\left|D\varphi_0\right| + \frac{\Theta}{1-\beta^2} \tag{5.4.10}$$

[14] Recall that $\partial_y v_x = [Du]^{-1}$.
[15] recall β from Lemma 5.3.11 and Θ from Lemma 5.3.13.

Proof. By eqs. (5.3.20), (5.3.24) and (5.4.9) we have for each $i \geq 1$

$$\begin{aligned} |D\varphi_i| &\leq \Theta + \beta^2 |D\varphi_{i-1}| \\ &= \beta^{2i} |D\varphi_0| + \Theta \sum_{k=0}^{i-1} \beta^{2k} \\ &\leq \beta^i |D\varphi_0| + \frac{\Theta}{1-\beta^2}, \end{aligned}$$

as desired. □

Lemma 5.4.13. *If ω is a branch of u and φ, φ' are in $C^1([p,q],[0,1])$, then*

$$|D[G_\omega(\varphi)] - D[G_\omega(\varphi')]| \leq \beta\kappa |\varphi - \varphi'| + \beta^2 |D\varphi - D\varphi'| \tag{5.4.11}$$

Proof. Note that by eqs. (5.3.11) and (5.3.13) we have $\partial_y \partial_x v_x(y) = \frac{D^2 u(x)}{[Du(x)]^2}$. Using eq. (5.4.9) and applying the mean value theorem to $\partial_x v_{\omega^{-1}(x)}$ we obtain for some θ between $\varphi(\omega^{-1}(x))$ and $\varphi'(\omega^{-1}(x))$,

$$\begin{aligned} [[D[G_\omega(\varphi)] - D[G_\omega(\varphi')]](x) &= \frac{\partial_y \partial_x v_x(\omega^{-1}(x), \theta)}{Du(\omega^{-1}(x))} [\varphi - \varphi'] \left(\omega^{-1}(x)\right) \\ &+ \left[\frac{D\varphi - D\varphi'}{[Du]^2}\right] \left(\omega^{-1}(x)\right) \\ &= \left[\frac{D^2 u}{[Du]^3} [\varphi - \varphi']\right] \left(\omega^{-1}(x)\right) \\ &+ \left[\frac{D\varphi - D\varphi'}{[Du]^2}\right] \left(\omega^{-1}(x)\right) \\ &\leq \beta\kappa |\varphi - \varphi'| + \beta^2 |D\varphi - D\varphi'|, \end{aligned}$$

as desired. □

Lemma 5.4.14. *If $(\omega_i)_{i=1}^{\infty}$ is a sequence of branches of u, φ_0 and φ_0' are in $C^1([p,q],[0,1])$, and for $i \geq 1$ we define $\varphi_i = G_{\omega_i}(\varphi_{i-1})$, then for all $i \geq 0$,*

$$|\varphi_i - \varphi_i'|_{C^1} \leq \beta^i \left[1 + \frac{\kappa}{1-\beta}\right] |\varphi_0 - \varphi_0'|_{C^1} \tag{5.4.12}$$

Proof. First we apply eqs. (5.4.8) and (5.4.11) to obtain and solve a recursive bound,

$$\begin{aligned}\left|D\varphi_i - D\varphi_i'\right| &\leq \beta\kappa\left|\varphi_{i-1} - \varphi_{i-1}'\right| + \beta^2\left|D\varphi_{i-1} - D\varphi_{i-1}'\right| \\ &\leq \beta^i\kappa\left|\varphi_0 - \varphi_0'\right| + \beta^2\left|D\varphi_{i-1} - D\varphi_{i-1}'\right| \\ &\leq \beta^{2i}\left|D\varphi_0 - D\varphi_0'\right| + \beta^i\kappa\left|\varphi_0 - \varphi_0'\right|\sum_{k=0}^{i}\beta^{k-1} \\ &\leq \beta^{2i}\left|D\varphi_0 - D\varphi_0'\right| + \beta^i\frac{\kappa}{1-\beta}\left|\varphi_0 - \varphi_0'\right|.\end{aligned}$$

Second, we combine the bound above with eq. (5.4.8) to verify eq. (5.4.12) □

Lemma 5.4.15. *Each cell of Γ^u is the graph of a $C^1([p,q],[0,1])$ function.*

Proof. Fix $\gamma \in \Gamma^u$. By Definition 5.4.7 we may select a sequence of sets $(U_k)_{k=1}^{\infty}$ such that for each $k \geq 1$, $U_k \in \mathcal{Z}_T^{-k}$ and $\gamma \subseteq U_k$. It follows from the definition of the sequence that the U_k are nested and $\gamma \subseteq \bigcap_k U_k$. If (x,y) and (w,z) are both in $\bigcap_k U_k$, then for all $k \geq 1$, $\mathcal{Z}_T^{-k}(x,y) = \mathcal{Z}_T^{-k}(w,z) = U_k$ so the points are backward equivalent, we conclude that $\gamma = \bigcap_k U_k$. It follows from eq. (5.4.4) and Lemma 5.4.2 that for each $t \in [p,q]$ the set $\ell_t \cap \bigcap_k U_k$ contains exactly one point. In other words $\bigcap_k U_k$ is the graph of some function from $[p,q]$ into $[0,1]$.

Let C_k be the unique cell in $\mathcal{Z}_T^1$ such that $T(C_k \cap U_{k-1}) = U_k$. By Lemma 5.4.1 there exists $I_k \in \mathcal{Z}_u^1$ such that $C_k = I_k \times [0,1]$. Let ω_k denote the branch of u with domain I_k. Given any function φ_0 in $C^1([p,q],[0,1])$ let $\varphi_k = G_{\omega_k}(\varphi_{k-1})$ for each $k \geq 1$. By eq. (5.4.12) and the contraction mapping principal there exists a unique limit point φ in $C^1([p,q],[0,1]))$ such that $|\varphi_k - \varphi|_{C^1} \to 0$, and φ is independent of φ_0. It follows from Definition 5.4.9 and the definition of the U_k that the graph of φ_k is contained in U_k. Since the U_k are nested we see that the graph of φ is contained in $\bigcap_k U_k$. We conclude that γ is the graph of φ. □

5.5 Observables

In this section we will define a space of observables on Λ that are compatible with the dynamics of T. There are several requirements that we will want our observables to satisfy. The first is for the space to be invariant under the action of the Koopman operator, in other words if ψ is an observable in our space, then $\psi \circ T$ is also.

Consider the function $\psi(x, y) = x$, if we apply the Koopman operator the resulting function $\psi \circ T$ will take the value p at the left edge of each stable column in $\mathcal{Z}_T^1$, and approach the value q at the right edge of each stable column. The Koopman operator has introduced countably many jump discontinuities. This issue gets worse as the Koopman operator is iterated, $\psi \circ T^k$ has jump discontinuities of magnitude $q - p$ at the boundary of each cell in $\mathcal{Z}_T^k$ and these boundaries become dense. We conclude from this example that we would need to make precise restrictions on function values at the left and right boundaries of Λ to isolate a space of continuous observables that are invariant under the action of the Koopman operator.

Notice that in our example iteration under the Koopman operator did not introduce any discontinuities along vertical lines. To be precise we introduce the following notation. Given a bounded measurable function $\psi\colon \Lambda \to \mathbb{C}$ and a point $x \in [p, q]$ let $\psi_x\colon \ell_x \to \mathbb{C}$ be defined by $\psi_x(y) = \psi(x, y)$. We will refer to the function ψ_x as the *vertical section* of ψ over x. For any observable ψ on Λ and $x \in [p, q]$, we have

$$(\psi \circ T)_x (y) = \left[\psi_{u(x)} \circ v_x\right](y) \tag{5.5.1}$$

For each $x \in [p, q]$, the function v_x is an affine contraction and hence C^∞, it follows that if the functions ψ_x are C^1 uniformly in x, then the functions $(\psi \circ T)_x$ will also be C^1 uniformly in x, since

$$\left|D (\psi \circ T)_x (y)\right| = \left|D\psi_{u(x)}(v_x(y))\partial_x v_x(y)\right| \leq \beta \left|D\psi_{u(x)}(v_x(y))\right|.$$

In fact this calculation shows that the Koopman operator has a regularizing effect on the vertical sections of ψ. From our discussion so far we see that if we wish to identify a space of observables that are invariant under the action of the Koopman operator, then we can not easily require continuity on Λ, however we can require as much regularity as we would like from the vertical sections of observables.

The second requirement is that the space of observables should be invariant under the following sequence of averaging operators.

Definition 5.5.1. For each $k \geq 1$ and bounded measurable $\psi\colon \Lambda \to \mathbb{C}$, define the *k-th unstable average of* ψ by

$$E_k\psi(x,y) = \frac{\int_{\mathcal{Z}_T^{-k}(x,y)} \psi \, d\mu}{\mu\left(\mathcal{Z}_T^{-k}(x,y)\right)} \tag{5.5.2}$$

These projections are useful in a compact embedding argument particularly the proof of Proposition 5.8.1. Consider the observable $\psi(x,y) = y$, the image $E_1\psi$ of this observable is constant on each cell of $\mathcal{Z}_T^{-1}$. The cells of $\mathcal{Z}_T^{-1}$ are connected strips, bounded above and below by C^1 curves, that extend horizontally across Λ. We will use a notation for sets that is analogous to the notation for functions, if A is a subset of Λ then let A_x denote the set $\ell_x \cap A$. If U and V are cells of $\mathcal{Z}_T^{-1}$ that touch along a common boundary and V is above U, then by the choice of ψ we have $\psi_x|_{V_x} > \psi_x|_{U_x}$, an application of Fubini's theorem shows that $E_1\psi|_V > E_1\psi|_U$. We conclude that E_1 may introduce discontinuities along the boundaries of the cells of $\mathcal{Z}_T^{-1}$. Similarly the operators E_k for $k > 1$ introduce discontinuities at the boundaries of the cells of $\mathcal{Z}_T^{-k}$. As we will see Lemma 5.5.10 the situation is slightly different than for the Koopman operator. The E_k operators introduce smaller and smaller discontinuities as k increases. This suggests that a Hölder space with respect to a symbolic metric, which we will introduce shortly, is preserved.

The third and final requirement on the space of observables is technical and is required for the proof of Proposition 5.9.10. For any $C \in \mathcal{Z}_T^k$ we need for the norm of $\mathbf{1}_C\left[\psi \circ T^k\right]$ to be on the order of the norm of ψ times the measure of C.

Having discussed at length the desired conditions on observables we would like to use let us get to the task of defining them. We begin by defining a symbolic metric on Λ with respect to the partitions $\mathcal{Z}_T^{-k}$ for $k \geq 1$.

Definition 5.5.2. Define the *stable separation time* $s\colon \Lambda \times \Lambda \to \mathbb{N} \cup \{\infty\}$ by

$$s\left((x,y),(w,z)\right) = \sup\left\{n \in \mathbb{N} : \tilde{\mathcal{Z}}_T^{-n}(x,y) = \tilde{\mathcal{Z}}_T^{-n}(w,z)\right\}. \tag{5.5.3}$$

Define the stable pseudometric $d\colon \Lambda \times \Lambda \to [0,\infty)$ by

$$d\left((x,y),(w,z)\right) = \beta^{s((x,y),(w,z))}. \tag{5.5.4}$$

where we follow the convention that $\beta^\infty = 0$. For each vertical line $\ell_x \subset \Lambda$, let d_x

denote the restriction of d to ℓ_x defined for $y, z \in \ell_x$ by,

$$d_x(y,z) = d\left((x,y),(x,z)\right).$$

Before moving on we observe a few easy facts about s and d.

Remark 5.5.3. The point (x,y) is backward equivalent to the point (w,z), meaning that for all $k \geq 0$ $\mathcal{Z}_T^{-k}(x,y) = \mathcal{Z}_T^{-k}(w,z)$, if and only if $s\left((x,y),(w,z)\right) = \infty$. From the definition of d it follows that $d\left((x,y),(w,z)\right) = 0$ if and only if (x,y) and (w,z) are backward equivalent. In section 5.4 we showed that each backward equivalence class is the graph of a function mapping $[p,q]$ into $[0,1]$. If (x,y) and (x,z) are distinct points that lie in a common vertical line $\ell_x \subset \Lambda$, then they cannot lie on a common graph and hence cannot be backward equivalent. Therefore, $d\left((x,y),(x,z)\right) = 0$ if and only if $y = z$. We have just seen that for each vertical line ℓ_x the restriction d_x is a genuine metric. Finally note that if $(a,c),(a,d),(b,e),(b,f) \in \Lambda$, (a,c) is backward equivalent to (b,e), and (a,d) is backward equivalent to (b,f), then

$$d_a(c,d) = d_b(e,f).$$

This can be verified with the triangle inequality as follows,

$$\begin{aligned} d_a(c,d) &\leq d\left((a,c),(b,e)\right) + d_b(e,f) + d\left((a,d),(b,f)\right) = d_b(e,f) \\ d_b(e,f) &\leq d\left((a,c),(b,e)\right) + d_a(c,d) + d\left((a,d),(b,f)\right) = d_a(c,d). \end{aligned}$$

If (x,y) and (w,z) are backward equivalent[16], then $d\left((x,y),(w,z)\right) = 0$. Note $\mathcal{Z}_T^{-k}(x,y) = \mathcal{Z}_T^{-k}(w,z)$ if and only if $\mathcal{Z}_T^{-k-1}\left(T\left(x,y\right)\right) = \mathcal{Z}_T^{-k-1}\left(T\left(w,z\right)\right)$, therefore $s\left(T\left(x,y\right), T\left(w,z\right)\right) = s\left(\left(x,y\right), \left(w,z\right)\right) + 1$, and thus

$$d\left(T\left(x,y\right), T\left(w,z\right)\right) = \beta d\left(\left(x,y\right), \left(w,z\right)\right). \tag{5.5.5}$$

In Lemma 5.5.12 we will see that if ψ is Hölder continuous, in the sense of Definition 5.5.11, with respect to d, then for all $k \geq 1$, $E_k\psi$ is also.

Let d_x denote the restriction of d to ℓ_x. For each $x \in [p,q]$, the restriction d_x is a

[16] see Definition 5.4.5

metric. With this notation eq. (5.5.5) becomes

$$d_{u(x)}\left(v_x(y), v_x(z)\right) = \beta d_x(y, z) \tag{5.5.6}$$

Having equipped each ℓ_x with a symbolic metric that is adapted to the dynamics of T we proceed to measure Hölder continuity of the vertical sections of an observable.

Definition 5.5.4. If $x \in [p, q]$, $h \colon \ell_x \to \mathbb{C}$ is a bounded measurable function, and $a \in (0, 1]$, then define

$$H_x^a(h) = \sup_{y \neq z} \frac{|h(y) - h(z)|}{d_x(y, z)^a}, \tag{5.5.7}$$

and

$$|h|_x^a = \|h\|_{\sup} + H_x^a(h)\,. \tag{5.5.8}$$

Let $\mathcal{H}_x^a = \{h : |h|_x^a < \infty\}$. The set $\mathcal{H}_x^a$ is the space of a-Hölder functions on ℓ_x with respect to the metric d_x

Remark 5.5.5. The set ℓ_x together with d_x is a compact metric space and thus $\mathcal{H}_x^a$ is a Banach space with respect to $|\cdot|_x^a$.

Remark 5.5.6. Given a bounded measurable function $\psi \colon \Lambda \to \mathbb{C}$ and $x \in [p, q]$ we often compute $|\psi_x|_x^a$. Henceforth we will adhere to the convention that $H_x^a(\psi) := H_x^a(\psi_x)$ and similarly $|\psi|_x^a := |\psi_x|_x^a$.

Remark 5.5.7. Note that for all $x \in [p, q]$ and $y, z \in [0, 1]$ we have $d_x(y, z) \leq 1$. For any $0 < a < b < 1$, $x \in [p, q]$, $y, z \in [0, 1]$ and $h \in \mathcal{H}_x^b$, we have

$$\frac{|h(y) - h(z)|}{d_x(y, z)^a} \leq \frac{|h(y) - h(z)|}{d_x(y, z)^b} d_x(y, z)^{b-a} \leq \frac{|h(y) - h(z)|}{d_x(y, z)^b}.$$

We conclude that

$$|h|_x^a \leq |h|_x^b\,,$$

therefore there is a norm decreasing continuous inclusion of $\mathcal{H}_x^b$ into $\mathcal{H}_x^a$.

Before moving on we examine the relationship between this notion of regularity for vertical sections and the action of the Koopman operator. Applying eqs. (5.5.1) and (5.5.6) we compute

$$\begin{aligned}\left|(\psi \circ T)_x(y) - (\psi \circ T)_x(z)\right| &= \left|\psi_{u(x)}\left(v_x(y)\right) - \psi_{u(x)}\left(v_x(z)\right)\right| \\ &\leq H^a_{u(x)}(\psi)\ d_{u(x)}\left(v_x(y), v_x(z)\right)^a \\ &\leq \beta^a\, H^a_{u(x)}(\psi)\ d_x\left(y,z\right)^a .\end{aligned}$$

The calculation above yields

$$H^a_x(\psi \circ T) \leq \beta^a H^a_{u(x)}(\psi)\,, \tag{5.5.9}$$

which shows that the Koopman operator has a regularizing effect on functions with Hölder vertical sections. A similar calculation shows that

$$\left\|(\psi \circ T)_x\right\|_{\sup} \leq \left\|\psi_{u(x)}\right\|_{\sup} . \tag{5.5.10}$$

Combining these two bounds yields

$$|\psi \circ T|^a_x \leq |\psi|^a_{u(x)} . \tag{5.5.11}$$

Before proceeding we will need the following elementary lemma

Lemma 5.5.8. *If $k \geq 1$ and $I \in \mathcal{Z}^k_u$, then*[17]

$$\left\|\frac{\mathcal{P}^k_u \mathbf{1}_I}{\bar{\mu}(I)} - 1\right\|_{\sup} \leq \kappa.$$

In the following proof and throughout the remainder of the section we will need to view $[p,q]$ as a factor measure space of Λ. For this reason we make the following definitions.

Definition 5.5.9. Let $\pi\colon \Lambda \to [p,q]$ by $\pi(x,y) = x$ and let $\bar{\mu} = \mu \circ \pi^{-1}$.

Proof of Lemma 5.5.8. Note that $u^k|_I$ is smooth, invertible, and onto $[p,q]$.

$$\mathcal{P}^k_u \mathbf{1}_I(x) = \frac{1}{Du^k\left(\left[u^k|_I\right]^{-1}(x)\right)},$$

[17]Recall Lemma 5.3.12

and

$$\bar{\mu}(I) = \int_{[p,q]} \mathcal{P}_u^k \mathbf{1}_I \, d\bar{\mu}.$$

It follows by the mean value theorem that there exists $t \in [p,q]$ such that

$$\bar{\mu}(I) = \mathcal{P}_u^k \mathbf{1}_I(t).$$

We conclude that

$$\frac{\mathcal{P}_u^k \mathbf{1}_I(x)}{\bar{\mu}(I)} = \frac{Du^k\left(\left[u^k|_I\right]^{-1}(t)\right)}{Du^k\left(\left[u^k|_I\right]^{-1}(x)\right)},$$

and therefore the claim follows from Lemma 5.3.12 □

The next lemma captures the relationship between the unstable projections and Hölder regularity of the vertical sections of an observable.

Lemma 5.5.10. *If* $k \geq 1$, $a \in (0,1]$, $x \in [p,q]$, *and* $\psi\colon \Lambda \to \mathbb{C}$ *is bounded measurable, then*

$$H_x^a(E_k\psi) \leq [\kappa+1] \int_{[p,q]} \left(2\kappa \left\|\psi_x\right\|_{\sup} + [\kappa+1]\, H_x^a(\psi)\right) \, d\bar{\mu} \tag{5.5.12}$$

Proof. Select points y and z in $[0,1]$ and let $j = s\left(\left(x,y\right),\left(x,z\right)\right)$. Applying eq. (5.5.2) and the fact that T preserves μ we calculate

$$\begin{aligned}\left(E_k\psi\right)_x(y) - \left(E_k\psi\right)_x(z) =& \frac{\int_{\mathcal{Z}_T^{-k}(x,y)} \psi \, d\mu}{\mu\left(\mathcal{Z}_T^{-k}(x,y)\right)} - \frac{\int_{\mathcal{Z}_T^{-k}(x,z)} \psi \, d\mu}{\mu\left(\mathcal{Z}_T^{-k}(x,z)\right)} \\ =& \frac{\int_{T^{-j}\mathcal{Z}_T^{-k}(x,y)} \psi \circ T^j \, d\mu}{\mu\left(T^{-j}\mathcal{Z}_T^{-k}(x,y)\right)} \\ & - \frac{\int_{T^{-j}\mathcal{Z}_T^{-k}(x,z)} \psi \circ T^j \, d\mu}{\mu\left(T^{-j}\mathcal{Z}_T^{-k}(x,z)\right)}\end{aligned} \tag{5.5.13}$$

To reduce the abundance of unsightly notation let us make a few observations and notations. Let $U = T^{-j}\mathcal{Z}_T^{-k}(x,y)$ and $V = T^{-j}\mathcal{Z}_T^{-k}(x,z)$ so that eq. (5.5.13) becomes

$$\frac{\int_U \psi \circ T^j \, d\mu}{\mu\left(U\right)} - \frac{\int_V \psi \circ T^j \, d\mu}{\mu\left(V\right)} \tag{5.5.14}$$

If $j \geq k$, then $U = V$ and the expression above is identically zero. If $j < k$, then U and V are rectangles that can be written as the intersection of distinct unstable strips in $\mathcal{Z}_T^{-(k-j)}$ and a common stable column in $\mathcal{Z}_T^j$. Therefore $\pi U = \pi V$ is a interval in

$\mathcal{Z}_u^j$ which we will denote by I. By applying Fubini's theorem we can rewrite the last displayed expression as in iterated integral as follows [18],

$$\int_I \left[\frac{\int_{U_x} \left(\psi \circ T^j\right)_x (y)\, dy}{\mu\,(U)} - \frac{\int_{V_x} \left(\psi \circ T^j\right)_x (y)\, dy}{\mu\,(V)} \right] d\bar{\mu}(x). \tag{5.5.15}$$

Now consider the bracketed integrand. Each of the integrals inside can be bounded by replacing $\left(\psi \circ T^j\right)_x$ with its global extrema producing an upper bound of

$$\frac{|U_x|}{\mu(U)} \sup_y \left(\psi \circ T^j\right)_x - \frac{|V_x|}{\mu(V)} \inf_y \left(\psi \circ T^j\right)_x . \tag{5.5.16}$$

Reversing the roles of sup and inf in eq. (5.5.16) provides a lower bound on the bracketed integrand of eq. (5.5.15). Note that,

$$\mu(V) = \int_I |V_x| \; d\bar{\mu}(x).$$

By eq. (5.4.3) The function $|V_x|$ is continuous and hence by the mean value theorem there exists $t \in I$ such that

$$\mu(V) = \bar{\mu}(I)\, |V_t| \, ,$$

and similarly there is an $s \in I$ associated to U. Next we rearrange eq. (5.5.16) to the equivalent expression

$$\begin{aligned} &\frac{1}{\bar{\mu}(I)} \left[\left[\frac{|U_x|}{|U_s|} - 1 \right] - \left[\frac{|V_x|}{|V_t|} - 1 \right] \right] \sup \left(\psi \circ T^j\right)_x \\ +&\frac{1}{\bar{\mu}(I)} \left[\left[\frac{|V_x|}{|V_t|} - 1 \right] + 1 \right] \left[\sup \left(\psi \circ T^j\right)_x - \inf \left(\psi \circ T^j\right)_x \right] . \end{aligned} \tag{5.5.17}$$

Next we bound each of the terms above separately. We apply eq. (5.5.9) to obtain

$$\left[\sup \left(\psi \circ T^j\right)_x - \inf \left(\psi \circ T^j\right)_x \right] \le H_x^a\left(\psi \circ T^j\right) \le \left(\beta^a\right)^j H_{u^j(x)}^a(\psi) \, . \tag{5.5.18}$$

Similarly we apply eq. (5.5.10) to obtain

$$\sup \left(\psi \circ T^j\right)_x \le \left\| \left(\psi \circ T^j\right)_x \right\|_{\sup} \le \left\| \psi_{u^j(x)} \right\|_{\sup} \tag{5.5.19}$$

[18]Note that we use x and y as variables of integration so that the geometry is clear even though these symbols have already been used in eq. (5.5.13).

By eq. (5.4.5) we see that

$$\left[\frac{|V_x|}{|V_t|} - 1\right] \leq \kappa\,|x-t| \leq \kappa\beta^j \tag{5.5.20}$$

where the last inequality follows from the fact that x and t are in I and eq. (5.3.20). The same bound holds for U. Combining eqs. (5.5.18) to (5.5.20) we see that eq. (5.5.17) is bounded above by

$$\frac{1}{\bar{\mu}(I)}\left[2\kappa\beta^j \left\|\psi_{u^j(x)}\right\|_{\sup} + \left[\kappa\beta^j+1\right](\beta^a)^j\, H^a_{u^j(x)}(\psi)\right],$$

which provides a uniform bound on the integrand in eq. (5.5.15), applying Lemma 5.5.8 we obtain the following upper bound on eq. (5.5.15),

$$\begin{aligned}
&\int_{[p,q]} \frac{\mathbf{1}_I}{\bar{\mu}(I)}\left[2\kappa\beta^j \left\|\psi_{u^j(x)}\right\|_{\sup} + \left[\kappa\beta^j+1\right](\beta^a)^j\, H^{u^j(x)}_a(\psi)\right] d\bar{\mu}\\
=&\int_{[p,q]} \frac{\mathcal{P}^j_u\mathbf{1}_I}{\bar{\mu}(I)}\left[2\kappa\beta^j \left\|\psi_x\right\|_{\sup} + \left[\kappa\beta^j+1\right](\beta^a)^j\, H^x_a(\psi)\right] d\bar{\mu}\\
\leq&\frac{\|\mathcal{P}^j_u\mathbf{1}_I\|_{\sup}}{\bar{\mu}(I)}\int_{[p,q]} \left[2\kappa\beta^j \left\|\psi_x\right\|_{\sup} + \left[\kappa\beta^j+1\right](\beta^a)^j\, H^x_a(\psi)\right] d\bar{\mu}\\
\leq&\left[\kappa+1\right]\int_{[p,q]} \left[2\kappa\beta^j \left\|\psi_x\right\|_{\sup} + \left[\kappa\beta^j+1\right](\beta^a)^j\, H^x_a(\psi)\right] d\bar{\mu}.
\end{aligned}$$

Dividing by $(\beta^a)^j = d_x(y,z)^a$ and replacing $(\beta^{1-a})^j$ and β^j with 1 completes the proof. □

The lemma above indicates that the E_k operators are compatible with the following semi-norm[19] on bounded measurable observables on Λ.

Definition 5.5.11. Given a bounded measurable function $\psi\colon \Lambda \to \mathbb{C}$ and $a \in (0,1]$ define

$$|\psi|_a = \int_{[p,q]} |\psi|^a_x \; d\bar{\mu}(x). \tag{5.5.21}$$

To be explicit the operators E_k are uniformly bounded with respect to $|\cdot|_a$ for all $a \in (0,1]$ as we prove in the following Lemma.

---

[19]Suppose that $\psi \neq \hat{\psi}$ and for $\bar{\mu}$ almost every x, we have $\psi_x = \hat{\psi}_x$. Then $\left|\psi - \hat{\psi}\right|_a = 0$ even though $\psi - \hat{\psi} \neq 0$

Lemma 5.5.12. *For all $k \geq 1$, $a \in (0,1]$, and bounded measurable $\psi \colon \Lambda \to \mathbb{C}$,*

$$\left|E_k\psi\right|_a \leq [2\kappa+1]^2 \left|\psi\right|_a \tag{5.5.22}$$

Proof. Note that for all $x \in [p,q]$

$$\begin{aligned}
\left(E_k\psi\right)^x(y) &= \frac{\int_{\mathcal{Z}_T^{-k}(x,y)} \psi \, d\mu}{\mu\left(\mathcal{Z}_T^{-k}(x,y)\right)} \\
&= \frac{\int_{T^{-k}\mathcal{Z}_T^{-k}(x,y)} \psi \circ T^k \, d\mu}{\mu\left(T^{-k}\mathcal{Z}_T^{-k}(x,y)\right)}
\end{aligned}$$

Let $C = T^{-k}\mathcal{Z}_T^{-k}(x,y)$ and note that C is in $\mathcal{Z}_T^k$ and thus is a stable column. Let $J = \pi C$ and note that $J \in \mathcal{Z}_u^k$, and since C is a stable column we have $\mu(C) = \bar{\mu}(J)$. We can rewrite the last integral above using the notation we have just introduced and apply Fubini's theorem and Lemma 5.5.8 as follows

$$\begin{aligned}
\frac{\int_C \psi \circ T^k \, d\mu}{\mu(C)} &= \frac{1}{\bar{\mu}(J)} \int_J \int_{C_x} \left(\psi \circ T^k\right)_x(y) \, dy \, d\bar{\mu}(x) \\
&\leq \frac{1}{\bar{\mu}(J)} \int_J \left|C_x\right| \left\| \left(\psi \circ T^k\right)_x \right\|_{\sup} d\bar{\mu}(x) \\
&\leq \int_{[p,q]} \frac{\mathbf{1}_J}{\bar{\mu}(J)} \left\|\psi_{u^k(x)}\right\|_{\sup} d\bar{\mu}(x) \\
&= \int_{[p,q]} \frac{\mathcal{P}_u^k \mathbf{1}_J}{\bar{\mu}(J)} \left\|\psi_x\right\|_{\sup} d\bar{\mu}(x) \\
&\leq \frac{\left\|\mathcal{P}_u^k \mathbf{1}_J\right\|_{\sup}}{\bar{\mu}(J)} \int_{[p,q]} \left\|\psi_x\right\|_{\sup} d\bar{\mu}(x) \\
&\leq [\kappa+1] \int_{[p,q]} \left\|\psi_x\right\|_{\sup} d\bar{\mu}(x)
\end{aligned}$$

Combining this bound with eq. (5.5.12) we obtain

$$\left|E_k\psi\right|_x^a \leq [\kappa+1][2\kappa+1] \left|\psi\right|_a \leq [2\kappa+1]^2 \left|\psi\right|_a$$

We obtain the claimed bound by integrating. □

Next we verify that the Koopman operator is compatible with $\left|\cdot\right|_a$ for all $a \in (0,1]$.

Lemma 5.5.13. *For all $a \in (0,1]$ and bounded measurable $\psi\colon \Lambda \to \mathbb{C}$,*

$$|\psi \circ T|_a \leq |\psi|_a \,. \tag{5.5.23}$$

Proof. Apply the fact that $\bar{\mu}$ is preserved by u and eq. (5.5.11) as follows,

$$\begin{aligned} |\psi \circ T|_a &= \int_{[p,q]} |\psi \circ T|_x^a \, d\bar{\mu}(x) \\ &\leq \int_{[p,q]} |\psi|_{u(x)}^a \, d\bar{\mu}(x) \\ &= \int_{[p,q]} |\psi|_x^a \, d\bar{\mu}(x) \\ &= |\psi|_a \,. \end{aligned}$$

□

Next we verify that $|\cdot|_a$ interacts with the indicator functions of stable columns as desired.

Lemma 5.5.14. *For each $k \geq 1$, stable column $C \in \mathcal{Z}_T^k$, and observable ψ with $|\psi|_x^a < \infty$,*

$$\left|\mathbf{1}_C \left[\psi \circ T^k\right]\right|_a \leq [\kappa + 1]\, \mu(C)\, |\psi|_a \,. \tag{5.5.24}$$

Proof. Let $I = \pi C$ and note that

$$\left|\mathbf{1}_C \left[\psi \circ T^k\right]\right|_x^a = \mathbf{1}_I \left|\left[\psi \circ T^k\right]\right|_x^a \leq \mathbf{1}_I \, |\psi|_{u^k(x)}^a$$

We compute as follows,

$$\begin{aligned} \left|\mathbf{1}_C \left[\psi \circ T^k\right]\right|_a &\leq \int_{[p,q]} \mathbf{1}_I \, |\psi|_{u^k(x)}^a \, d\bar{\mu}(x) \\ &= \bar{\mu}(I) \int_{[p,q]} \frac{\mathcal{P}_u \mathbf{1}_I}{\bar{\mu}(I)} \, |\psi|_x^a \, d\bar{\mu}(x) \\ &\leq \bar{\mu}(I) \frac{\|\mathcal{P}_u \mathbf{1}_I\|_{\sup}}{\bar{\mu}(I)} \int_{[p,q]} |\psi|_x^a \, d\bar{\mu}(x) \\ &\leq [\kappa + 1]\, \bar{\mu}(I)\, |\psi|_a \,, \end{aligned}$$

since $\bar{\mu}(I) = \mu(C)$ this completes the proof. □

Having verified that for each $a \in (0,1]$ the semi-norm $|\cdot|_a$ satisfies all of our requirements we proceed to select Banach spaces of observables. Define the usual equivalence with respect to $|\cdot|_a$ by $\psi \sim \hat{\psi}$ if and only if $\left|\psi - \hat{\psi}\right|_a = 0$. We note that $\psi \sim \hat{\psi}$ if an only if $\psi_x = \hat{\psi}_x$ for $\bar{\mu}$ almost every x in $[p,q]$. If we let $C^a(\ell_x)$ denote the space of a-Hölder functions on ℓ_x with respect to the metric d_x we see that every class of bounded measurable ψ with $|\psi|_a < \infty$ can be identified with an element of the $L^1([p,q],\bar{\mu})$ direct sum

$$\bigoplus_{x\in[p,q]} C^a(\ell_x)$$

by sending ψ to $\oplus_{x\in[p,q]}\psi_x$. The bounded measurable observables ψ with $|\psi|_a < \infty$ are mapped to a dense subset of the direct sum. We conclude that the usual completion of the quotient construction with respect to the $|\cdot|_a$ semi-norm is isomorphic to the aforementioned direct sum of Hölder spaces.

Definition 5.5.15. Fix $a \in (0,1)$. Let Q^a denote the quotient of the set of bounded measurable functions on Λ with respect to the equivalence relation $\psi \sim \psi'$ if $|\psi - \psi'|_a = 0$. Let $\mathfrak{C}_a$ denote the completion of $\{[\psi] \in Q^a : |[\psi]|_a < \infty\}$ with respect to $|\cdot|_a$. We refer to $\mathfrak{C}_a$ as the *strong space of observables* and let $\|\cdot\|_{\mathfrak{C}_a}$ denote the norm on $\mathfrak{C}_a$. Let $\mathfrak{C}_1$ be constructed in the same way with $a = 1$, and let $\|\cdot\|_{\mathfrak{C}_1}$ denote the norm on $\mathfrak{C}_1$. We refer to $\mathfrak{C}_1$ as the *weak space of observables*.

The naming convention for the spaces above may seem counter intuitive at first. The weak space $\mathfrak{C}_1$ is a subset of the strong space $\mathfrak{C}_a$ and the norms are related by $\|\cdot\|_{\mathfrak{C}_a} \leq \|\cdot\|_{\mathfrak{C}_1}$. The reason for the teminology is that the spaces $\mathfrak{C}_a$ and $\mathfrak{C}_1$ will be used to define norms through integrationand we will consider spaces $\mathfrak{B}$ and $\mathfrak{B}_w$ that resemble the dual spaces $(\mathfrak{C}_a)'$ and $(\mathfrak{C}_1)'$ respectivly. Note that $(\mathfrak{C}_a)' \subseteq (\mathfrak{C}_1)'$.

5.6 The Unstable Expectation Operator

In this section we will investigate some properties of the E_k operators acting on $\mathfrak{C}_a$. We will prove that their limit coincides with the conditional expectation with respect to the σ-algebra of Γ^u saturated Borel sets, which we refer to as the *unstable expectation operator*.

We will then be able to show that this conditional expectation operator is bounded on $\mathfrak{C}_a$. The results of this section are applied in the proof of Proposition 5.8.1.

Lemma 5.6.1. *For all $\psi \in \mathfrak{C}_a$, the sequence $(E_k\psi)_{k=1}^{\infty}$ is Cauchy with respect to the uniform norm.*

Proof. The argument is very similar to the proof of Lemma 5.5.10 so we will only outline the argument. Fix $(x,y) \in \Lambda$ and let $k > j \geq 1$. Note that $\mathcal{Z}_T^{-k}(x,y) \subset \mathcal{Z}_T^{-j}(x,y)$. Define $C = T^{-j}\mathcal{Z}_T^{-k}(x,y)$ and $V = T^{-j}\mathcal{Z}_T^{-j}(x,y)$ and let I denote the interval $\pi C = \pi V$. We then calculate

$$\begin{aligned}
|E_j\psi(x,y) - E_k\psi(x,y)| &= \left| \frac{\int_V \psi \circ T^j \, d\mu}{\mu(V)} - \frac{\int_C \psi \circ T^j \, d\mu}{\mu(C)} \right| \\
&\leq \frac{1}{\bar{\mu}(I)} \int_I \left| \frac{|V^x|}{|V^t|} \sup \left(\psi \circ T^j\right)^x - \inf \left(\psi \circ T^j\right)^x \right| d\bar{\mu}(x) \\
&\leq \int_{[p,q]} \frac{\mathbf{1}_I}{\bar{\mu}(I)} \left[\kappa \beta^j \left| \psi^{u^j(x)} \right| + (\beta^a)^j H^a_{u^j(x)}(\psi) \right] d\bar{\mu}(x) \\
&\leq \kappa \left[\kappa + 1\right] (\beta^a)^j \|\psi\|_{\mathfrak{C}_a}
\end{aligned}$$

Since $(x,y) \in \Lambda$ was arbitrary we see that the sequence $E_j\psi$ is uniformly Cauchy. □

The previous lemma shows that the E_k operators have a well defined limit.

Definition 5.6.2. Given $\psi \in \mathfrak{C}_a$ define

$$E^u\psi(x,y) = \lim_{k\to\infty} E_k\psi(x,y) \tag{5.6.1}$$

Next we investigate regularity properties of the limit.

Corollary 5.6.2.1. *For all ψ in $\mathfrak{C}_a$, the function $E^u\psi$ is also in $\mathfrak{C}_a$. The operator norm of $E^u\colon \mathfrak{C}_a \circlearrowleft$ is bounded above by $[2\kappa + 1]^2$.*

Proof. By Lemma 5.6.1 the sequence $E_k\psi$ is uniformly convergent. Fix a point $x \in [p,q]$. For each $k \geq 0$ define[20] $h_k = (E_k\psi)^x$. By Definition 5.5.4 there exists $C > 0$ such that for all $k \geq 0$, $H^a(h_k) = H^a_x(E_k\psi) \leq C$. Let h denote the uniform limit of (h_k). Select $\epsilon > 0$ and chose $k \geq 0$ such that $\|h_k - h\|_{\sup} < \epsilon/2$. For all $u, v \in [0,1]$ we have,

$$\begin{aligned}
|h(u) - h(v)| &\leq |h(u) - h_k(u)| + |h_k(u) - h_k(v)| + |h_k(v) - h(v)| \\
&\leq H^a(h_k)\,|u - v| + \epsilon \\
&\leq C\,|u - v| + \epsilon
\end{aligned}$$

[20]Each h_k is a function on $[0,1]$.

Since ϵ was arbitrary we see that for all $u, v \in [0,1]$,

$$|h(u) - h(v)| \leq C\,|u - v|\,.$$

Therefore, $H^a(h) \leq C$. Since x was arbitrary the same argument holds for every $x \in [p, q]$. We conclude that $E^u\psi$ has holder sections with uniformly bounded Hólder constant. The bound in eq. (5.5.22) is easily seen to hold in the limit.

□

We wish to compare E^u to the unstable expectation. We will show that E^u is a version of the unstable expectation.

Definition 5.6.3. Let $\mathcal{B}^u$ denote the σ-algebra of Borel sets that are saturated[21] with respect to Γ^u.

Lemma 5.6.4. *For all bounded measurable $\psi\colon \Lambda \to \mathbb{C}$ and $A \in \mathcal{B}^u$,*

$$\int_A E^u\psi\, d\mu = \int_A \psi\, d\mu. \tag{5.6.2}$$

Proof. The proof will be an application of the π-λ theorem [].

Recall that given a set X a π system P on X is a non-empty collection of subsets of X that is closed under finite intersections. A λ-system D on X is a collection of subsets of X that contains X, is closed under complements, and is closed under countable disjoint unions. The π-λ theorem states that if P is contained in D, then the sigma algebra generated by P is contained in D.

In this proof we will show that:

1. The collection $\mathrm{P} = \left\{A \subseteq \Lambda : \exists k \geq 1 \text{ s.t. } A \in \mathcal{Z}_T^{-k}\right\} \cup \{\emptyset\}$ is a π-system.
2. The collection $\mathrm{D} = \left\{A \subseteq \Lambda : \int_A E^u\psi\, d\mu = \int_A \psi\, d\mu\right\}$ is a λ-system.
3. The collection P is contained in D.
4. The collection P generates $\mathcal{B}^u$.

The π-λ theorem then implies that $\mathcal{B}^u \subset \mathrm{D}$ and hence the claim.

[21] For more details see appendix A.2 in particular Definition A.2.13 and Lemma A.2.14. In the notation of the appendix we would let $\mathcal{B}$ denote the Borel σ-algebra on $[0,1]^2$ and define $\mathcal{B}^u = \mathcal{B}/\Gamma^u$.

Proof of item 1: The collection P is clearly non-empty. Given A and B in P there exists j and k greater then zero such that A is in $\mathcal{Z}_T^{-j}$ and B is in $\mathcal{Z}_T^{-k}$. With out loss of generality we assume that $j \geq k$. Since $\mathcal{Z}_T^{-j}$ and $\mathcal{Z}_T^{-k}$ are partitions and $\mathcal{Z}_T^{-j}$ refines $\mathcal{Z}_T^{-k}$ it is either the case that A and B are disjoint or $A \cap B = A$. In either case $A \cap B$ is in P. Therefore, P is a π-system.

It is convenient to prove item 3 next.

Proof of item 3: Note that if A is in $\mathcal{Z}_T^{-k}$, then for all (x, y) in A we have

$$E_k\psi(x,y) = \frac{\int_A \psi \, d\mu}{\mu A},$$

Integrating over A we obtain

$$\int_A E_k\psi \, d\mu = \int_A \psi \, d\mu.$$

For all $j > k$ we have observed that $\mathcal{Z}_T^{-j}$ refines $\mathcal{Z}_T^{-k}$. The partitions $\mathcal{Z}_T^{-j}$ are countable. Therefore, for all A in $\mathcal{Z}_T^{-k}$ and $j > k$ there exists a countable disjoint sequence A_i^j in $\mathcal{Z}_T^{-j}$, where $i = 1, \cdots$, such that $A = \cup_i A_i^j$. Applying this notation we see that for all A in $\mathcal{Z}_T^{-k}$ and $j > k$,

$$\begin{aligned} \int_A E_j\psi \, d\mu &= \sum_i \int_{A_i^j} E_j\psi \, d\mu \\ &= \sum_i \int_{A_i^j} \psi \, d\mu \\ &= \int_A \psi \, d\mu, \end{aligned}$$

Where we have applied monotone convergence in the first and last equalities and our previous observation in the middle equality. By construction, for all $A \in$ P, there exists k such that A is in $\mathcal{Z}_T^{-k}$. Therefore, for all $j \geq k$,

$$\int_A E_j\psi \, d\mu = \int_A \psi \, d\mu.$$

By Lemma 5.6.1 we may apply the dominated convergence theorem, which yields for all A in P,

$$\int_A E^u \psi \, d\mu = \lim_{j\to\infty} \int_A E_j \psi \, d\mu = \int_A \psi \, d\mu.$$

We conclude that P is contained in D.

Proof of item 2: Note that for all $k \geq 1$ we have

$$\begin{aligned}\int_\Lambda E_k \psi \, d\mu &= \sum_{A \in \mathcal{Z}_T^{-k}} \int_A E_k \psi \, d\mu \\ &= \sum_{A \in \mathcal{Z}_T^{-k}} \int_A \psi \, d\mu \\ &= \int_\Lambda \psi \, d\mu.\end{aligned}$$

Therefore Λ is in D. If A is in D, then

$$\begin{aligned}\int_{A^c} E_k \psi \, d\mu &= \int_\Lambda E_k \psi \, d\mu - \int_A E_k \psi \, d\mu \\ &= \int_\Lambda \psi \, d\mu - \int_A \psi \, d\mu \\ &= \int_{A^c} \psi \, d\mu\end{aligned}$$

Therefore, D is closed under complements. If A_i is a countable disjoint sequence in D, then by applying the monotone convergence theorem we obtain

$$\begin{aligned}\int_{\cup_i A_i} E^u \psi \, d\mu &= \sum_i \int_{A_i} E^u \psi \, d\mu \\ &= \sum_i \int_{A_i} \psi \, d\mu \\ &= \int_{\cup_i A_i} \psi \, d\mu.\end{aligned}$$

Therefore, D is closed under countable disjoint unions. We conclude that D is a λ-system.

Proof of item 4: Let B be a closed set in $\mathcal{B}^u$. For each $k \geq 1$ let

$$B_k = \bigcup \left\{ A \in \mathcal{Z}_T^{-k} : A \cap B \neq \emptyset \right\}.$$

It follows from the refinement properties of the $\mathcal{Z}_T^{-k}$ partitions that the sets B_k form a decreasing nested sequence. It follows from the definitions that each B_k is in the σ-algebra generated by P. By definition $B \subseteq \cap_k B_k$. We claim that in fact $B = \cap_k B_k$. If (x, y) is in $\cap_k B_k$, then for all $k \geq 1$ there exists (w_k, z_k) in $\mathcal{Z}_T^{-k}(x, y) \cap B$. Let (w, z) be a limit point of this sequence and note that since B is closed we have that (w, z) is in B. Since the sets B_k are nested we see that (w, z) is backward equivalent[22] to (x, y) and hence that $\pi_u(x, y) \cap B$ is a non-empty. Since B is in $\mathcal{B}^u$ we know that B is saturated with respect to Γ^u meaning that for all $\gamma \in \Gamma^u$, if $\gamma \cap B$ is non-empty, then γ is contained in B. It follows that $\pi_u(x, y)$ is contained in B and hence that (x, y) is contained in B. Since (x, y) in $\cap_k B_k$ was arbitrary we see that that $B = \cap_k B_k$. Since all of the sets B_k were in the σ-algebra generated by P we see that B is in the σ-algebra generated by P. We conclude that every closed set that is saturated with respect to Γ^u is in the σ-algebra generated by P. The closed subsets of $[0, 1]^2$ generate the Borel σ-algebra. It follows from the elementary Lemma A.2.15 that the closed Γ^u-saturated sets generate $\mathcal{B}^u$. We conclude that **P** generates $\mathcal{B}^u$.

□

Let μ^u denote the restriction of μ to $\mathcal{B}^u$ and let $\psi\mu$ denote the measure defined for all Borel sets A by

$$\psi\mu(A) = \int_A \psi \, d\mu.$$

The Radon-Nikodym Theorem together with Lemma 5.6.4 show that $E^u\psi$ is a version of the expectation of ψ with respect to $\mathcal{B}^u$, which is defined by

$$\frac{d\,(\psi\mu)}{d\mu^u}.$$

This observation along with the operator norm bound observed in Corollary 5.6.2.1 provide all of the technical tools regarding E^u that we will need.

[22] see Definition 5.4.5.

5.7 Densities

In this section we define a space of densities that are bounded measurable functions $\eta\colon \Lambda \to \mathbb{C}$. We will require that the space of densities be invariant under the Frobenius-Perron operator. We will also require these densities to induce linear functionals on our space of observables via integration.

Note that since T is invertible and measure preserving we have the following simple pointwise formula for the Frobenius-Perron operator,

$$\mathcal{P}\eta(x,y) = \left[\eta \circ T^{-1}\right](x,y). \tag{5.7.1}$$

As was the case for the Koopman operator acting on observables, the Frobenius-Perron operator introduces discontinuities. A simple example illustrates the problem. If $\eta(x,y) = y$, then $\mathcal{P}\eta(x,y) = \eta(T^{-1}(x,y))$. If U is in $\mathcal{Z}_T^{-1}$ and (x,y) is on the lower boundary of U, then $T^{-1}(x,y)$ is on the lower boundary of the stable column $T^{-1}U$ and thus $\mathcal{P}\eta(x,y) = 0$. If (x,y) approaches the upper boundary of U then $T^{-1}(x,y)$ approaches the upper boundary of $T^{-1}U$ and thus $\mathcal{P}\eta(x,y)$ approaches 1. We conclude that $\mathcal{P}\eta$ has a jump discontinuity of magnitude 1 at the boundary of each $U \in \mathcal{Z}_T^{-1}$.

The Frobenius-Perron operator also mirrors the regularizing behavior of the Koopman operator acting on vertical sections of observables. In section 5.4 we identified the partition Γ^u that consisted of graphs of smooth functions. The Frobenius-Perron operator has a familiar regularizing effect on Γ^u-sections of densities. Motivating examples of the regularizing effect of $\mathcal{P}$ on Γ^u-sections of densities are more technically involved then our examples with the Koopman operator acting on vertical sections, but have essentially the same content. Restricted to a unstable disk γ of Γ^u the map T^{-1} is a strict contraction. Precomposition of a regular function by a contraction induces a multiplicative decrease in any reasonable measure of regularity; for example Lipschitz constants enjoy this property.

We also wish to select a space of densities such that for all densities η the function from $\mathfrak{C}_a$ to $\mathbb{C}$ defined by

$$\psi \mapsto \int_\Lambda \eta\psi \, d\mu$$

is a bounded linear functional on $\mathfrak{C}_a$.

To achieve both of these goals simultaneously we introduce the following operations on bounded measurable functions that will allow us to define a norm on bounded measurable $\eta\colon \Lambda \to \mathbb{C}$ that measures the Lipschitz constant of Γ^u-sections of η through integration against observables ψ in $\mathfrak{C}_a$. This will provide us with a norm that enjoys a regularizing property with respect to $\mathcal{P}$ and is compatible with our goal of having η induce a linear functional on $\mathfrak{C}_a$.

Definition 5.7.1. Given a bounded measurable function $\eta\colon \Lambda \to \mathbb{R}$ and points s and t in $[p,q]$ we define

$$\delta_t(\eta)\,(x,y) = \eta\,(\ell_t \cap \pi_u(x,y))\,, \tag{5.7.2}$$

$$\Delta_s^t(\eta) = \delta_t(\eta) - \delta_s(\eta)\,. \tag{5.7.3}$$

Note that both $\delta_t(\eta)$ and $\Delta_s^t(\eta)$ are bounded measurable functions on Λ that are constant along unstable disks. The value of $\Delta_s^t(\eta)$ along a particular unstable disk γ is the increment of η between the points $\gamma\cap\ell_s$ and $\gamma\cap\ell_t$. The mappings $\eta \mapsto \delta_t(\eta)$ and $\eta \mapsto \Delta_s^t(\eta)$ both define bounded linear operators on the space of bounded measurable functions on Λ with respect to the supremum norm.

It will be important to determine how the operations that we have just defined interact with the Frobenius-Perron operator. It will be convenient to have the following notation.

Definition 5.7.2. Given $t \in [p,q]$, $(x,y) \in \Lambda$ and $k \geq 1$, let $t_k(x,y) \in [p,q]$ denote the unique point such that

$$T^{-k}\,(\ell_t \cap \pi_u(x,y)) \in \ell_{t_k(x,y)}. \tag{5.7.4}$$

When the point (x,y) is clear we will often abbreviate $t_k(x,y)$ as t_k. The next lemma shows the relationship between the notation that we have just defined and the fact that u is a uniformly expanding map.

Lemma 5.7.3. *If s and t are in $[p,q]$ and $(x,y) \in \Lambda$, then for all $k \geq 1$,*

$$|s_k(x,y) - t_k(x,y)| \leq \beta^k\,|s-t| \tag{5.7.5}$$

Proof. From Definition 5.7.2 it follows[23] that for all $k \geq 1$,

$$\mathcal{Z}_u^k(s_k(x,y)) = \mathcal{Z}_u^k(t_k(x,y)) = \mathcal{Z}_u^k\left(\pi\left(T^{-k}(x,y)\right)\right) =: I_k.$$

An application of the mean value theorem to $u^k|_{I_k}$ together with eq. (5.3.20) yields the claimed bound. □

We are now ready to analyze the interaction between $\delta_t(\cdot)$ and $\mathcal{P}$ the next lemma provides a commutation relation that will be important in the proof of the Lasota-Yorke inequality.

Lemma 5.7.4. *For all bounded measurable $\eta\colon \Lambda \to \mathbb{R}$ and $t \in [p,q]$,*

$$\delta_t(\mathcal{P}\eta)\,(x,y) = \mathcal{P}\delta_{t_1(x,y)}(\eta)\,(x,y). \tag{5.7.6}$$

Proof. We compute,

$$\begin{aligned}
\delta_t(\mathcal{P}\eta)\,(x,y) &= \delta_t\big(\eta \circ T^{-1}\big)\,(x,y) \\
&= \left[\eta \circ T^{-1}\right](\ell_t \cap \pi_u\,(x,y)) \\
&= \eta\left(\ell_{t_1(x,y)} \cap \pi_u\left(T^{-1}(x,y)\right)\right) \\
&= \left[\delta_{t_1(x,y)}(\eta) \circ T^{-1}\right](x,y) \\
&= \mathcal{P}\delta_{t_1(x,y)}(\eta)\,(x,y)
\end{aligned}$$

The key step above is in the third line where we have used the equivariance of Γ^u observed in Remark 5.4.8. □

The analogous commutation relation for $\Delta_s^t(\cdot)$ and $\mathcal{P}$ is the content of the next corollary.

Corollary 5.7.4.1. *For all bounded measurable $\eta\colon \Lambda \to \mathbb{R}$ and $t, s \in [p,q]$,*

$$\Delta_s^t(\mathcal{P}\eta)\,(x,y) = \mathcal{P}\Delta_{s_1(x,y)}^{t_1(x,y)}(\eta)\,(x,y). \tag{5.7.7}$$

Proof. This follows directly from eqs. (5.7.3) and (5.7.6). □

[23]For each $k \geq 0$ the points $\ell_s \cap \pi_u(x,y)$ and $\ell_t \cap \pi_u(x,y)$ are contained in the unstable strip $\mathcal{Z}_T^{-k}(x,y)$, thus for every $k \geq 0$ both points have their k-th pre-image in $\mathcal{Z}_T^k(x,y)$, which is the stable column over $\mathcal{Z}_u^k(x,y)$

Next we define a strong notion of unstable Lipschitz constant that makes no reference to our space of observables. Functions with this form of unstable regularity will from a dense subset of our final space of functionals.

Definition 5.7.5. Given a bounded measurable function $\eta\colon \Lambda \to \mathbb{R}$ define

$$L_u(\eta) = \sup\left\{\frac{|\Delta_s^t(\eta)|}{|t-s|} : t, s \in [p,q],\, (x,y) \in \Lambda\right\} \tag{5.7.8}$$

$$\|\eta\|_{\mathfrak{L}_u} = |\eta| + L_u(\eta) \tag{5.7.9}$$

Having defined a norm on bounded measurable functions we introduce notation for the finite norm elements.

Definition 5.7.6. Define the following subset of all bounded measurable functions on Λ

$$\mathfrak{L}_u = \left\{\eta : \|\eta\|_{\mathfrak{L}_u} < \infty\right\}, \tag{5.7.10}$$

which we refer to as the space of *unstable Lipschitz densities.*

The space $\mathfrak{L}_u$ is in fact a Banach space that is very similar to our space of observables $\mathfrak{C}_a$. It can be viewed as a sup-norm direct sum of Lipschitz spaces $C^1(\gamma)$ for $\gamma \in \Gamma^u$ with respect to the horizontal distance metric on γ. We will not use this perspective so we do not develop it any further.

Next we verify that $\mathcal{P}$ has a regularizing effect on $\mathfrak{L}_u$.

Lemma 5.7.7. *For all $\eta \in \mathfrak{L}_u$, $\mathcal{P}\eta \in \mathfrak{L}_u$ and $L_u(\mathcal{P}f) \leq \beta L_u(f)$.*

Proof. Fix $\eta \in \mathfrak{L}_u$, $t, s \in [p,q]$, and $(x,y) \in \Lambda$. We compute,

$$\begin{aligned}\frac{|\Delta_s^t(\mathcal{P}\eta)(x,y)|}{|t-s|} &= \frac{\mathcal{P}\left|\Delta_{s_1}^{t_1}(\eta)\right|(x,y)}{|t-s|}\\ &= \frac{\left|\Delta_{s_1}^{t_1}(\eta)\left(T^{-1}(x,y)\right)\right|}{|t-s|}\\ &= \frac{L_u(\eta)\,|t_1-s_1|}{|t-s|}\\ &\leq \beta L_u(\eta),\end{aligned}$$

The key steps are the first and last lines where we apply the commutation relation from eq. (5.7.7) and the bound on $|t_1 - s_1|$ from eq. (5.7.5) respectively. □

Now we proceed to define norms that are adapted to our goal of having densities induce functionals on $\mathfrak{C}_a$.

Definition 5.7.8. For all bounded measurable functions $\eta\colon \Lambda \to \mathbb{R}$ define

$$\|\eta\|_{\mathfrak{B}_w} = \sup\left\{\int_\Lambda \delta_t(\eta)\ \psi\, d\mu : t \in [p,q], \psi \in \mathfrak{C}_1, \|\psi\|_{\mathfrak{C}_1} \le 1\right\} \tag{5.7.11}$$

$$\|\eta\|_s = \sup\left\{\int_\Lambda \delta_t(\eta)\ \psi\, d\mu : t \in [p,q], \psi \in \mathfrak{C}_a, \|\psi\|_{\mathfrak{C}_a} \le 1\right\} \tag{5.7.12}$$

$$\mathbb{L}(\eta) = \sup\left\{\frac{\int_\Lambda \Delta_s^t(\eta)\ \psi\, d\mu}{|t-s|} : t, s \in [p,q], \psi \in \mathfrak{C}_a, \|\psi\|_{\mathfrak{C}_a} \le 1\right\} \tag{5.7.13}$$

$$\|\eta\|_{\mathfrak{B}} = \|\eta\|_s + \mathbb{L}(\eta) \tag{5.7.14}$$

We see that the operations $\delta_t(\cdot)$ and $\Delta_s^t(\cdot)$ allow for the unstable Lipschitz constant to be reformulated to incorporate integration against an observable which yields a "weak" formulation of the unstable Lipschitz constant which will extend to functionals on $\mathfrak{C}_a$ that may not have a density. The next lemma shows that the "weak" formulation of the unstable Lipschitz constant is dominated by our first formulation.

Remark 5.7.9. For all bounded measurable functions $\eta\colon \Lambda \to \mathbb{R}$,

$$\mathbb{L}(\eta) \le L_u(\eta). \tag{5.7.15}$$

Remark 5.7.10. By Remark 5.5.7 the unit ball of $\mathfrak{C}_1$ can be viewed as a subset of the unit ball of $\mathfrak{C}_a$, therefore for all $\eta \in \mathfrak{B}$

$$\|\eta\|_{\mathfrak{B}_w} \le \|\eta\|_s \le \|\eta\|_{\mathfrak{B}}.$$

Having verified that the norms that we have defined give suitable bounds on functional norms we proceed to construct Banach spaces with respect to these norms.

Definition 5.7.11. Define $\mathfrak{B}_w$ to be to be the $\|\cdot\|_{\mathfrak{B}_w}$-completion of $\mathfrak{L}_u$ and $\mathfrak{B}$ to be the $\|\cdot\|_{\mathfrak{B}}$-completion of $\mathfrak{L}_u$.

Note that by construction $\mathfrak{L}_u$ is dense in both spaces. In general we will prove lemmas about $\mathfrak{L}_u$ with the understanding that they may be extended to $\mathfrak{B}$ or $\mathfrak{B}_w$ by passing through a Cauchy sequence. We begin by investigating the interaction between the Frobenius-Perron operator and the norms.

Lemma 5.7.12. *If $\eta \in \mathfrak{L}_u$, then*

$$\mathbb{L}\left(\mathcal{P}\eta\right) \leq \beta\mathbb{L}\left(\eta\right) \tag{5.7.16}$$

$$\left\|\mathcal{P}\eta\right\|_{\mathfrak{B}_w} \leq \left\|\eta\right\|_{\mathfrak{B}_w} \tag{5.7.17}$$

$$\left\|\mathcal{P}\eta\right\|_s \leq \left\|\eta\right\|_s \tag{5.7.18}$$

Proof. We will prove eq. (5.7.16), the other two bounds are similar. Fix $\eta \in \mathfrak{L}_u$, $s, t \in [p, q]$ and $\psi \in \mathfrak{C}_a$ with $\|\psi\|_{\mathfrak{C}_a} \leq 1$. We compute,

$$\begin{aligned}
\frac{\left|\int_\Lambda \Delta_s^t(\mathcal{P}\eta)\ \psi\, d\mu\right|}{|t-s|} &= \frac{\left|\int_\Lambda \mathcal{P}\Delta_{s_1}^{t_1}(\eta)\ \psi\, d\mu\right|}{|t-s|} \\
&= \frac{\left|\int_\Lambda \Delta_{s_1}^{t_1}(\eta)\ \psi\circ T\, d\mu\right|}{|t-s|} \\
&\leq \frac{\mathbb{L}\left(\eta\right)|t_1 - s_1|\, \|\psi\circ T\|_{\mathfrak{C}_a}}{|t-s|} \\
&\leq \beta\mathbb{L}\left(\eta\right).
\end{aligned}$$

□

The remainder of this section is devoted to the proof of the following Lasota-Yorke inequality.

Proposition 5.7.13. *For all $\eta \in \mathfrak{L}_u$ and $k \geq 1$*

$$\left\|\mathcal{P}^k\eta\right\|_{\mathfrak{B}} \leq 3\left(\beta^a\right)^k \|\eta\|_{\mathfrak{B}} + \|\eta\|_{\mathfrak{B}_w}. \tag{5.7.19}$$

The next corollary of Lemma 5.7.12 provides a first step toward the Lasota-Yorke inequality.

Corollary 5.7.13.1. *If $\eta \in \mathfrak{L}_u$ and $k \geq 1$ then*

$$\left\|\mathcal{P}^k\eta\right\|_{\mathfrak{B}} \leq \beta^k \|\eta\|_{\mathfrak{B}} + \left\|\mathcal{P}^k\eta\right\|_s. \tag{5.7.20}$$

Proof. First note that if we can prove the following inequality,

$$\left\|\mathcal{P}^k\eta\right\|_{\mathfrak{B}} \leq \beta^k \|\eta\|_{\mathfrak{B}} + [1-\beta]\sum_{j=0}^{k-1} \beta^j \left\|\mathcal{P}^{k-j}\eta\right\|_s,$$

then by eq. (5.7.18) we have

$$\left\|\mathcal{P}^k\eta\right\|_{\mathfrak{B}} \leq \beta^k \left\|\eta\right\|_{\mathfrak{B}} + [1-\beta] \left\|\mathcal{P}^k\eta\right\|_s \sum_{j=0}^{k-1} \beta^j \leq \beta^k \left\|\eta\right\|_{\mathfrak{B}} + \left\|\mathcal{P}^k\eta\right\|_s ,$$

which proves the claimed inequality.

We will prove the preliminary inequality by induction. Fix $\eta \in \mathfrak{L}_u$. We compute,

$$\begin{aligned}\left\|\mathcal{P}\eta\right\|_{\mathfrak{B}} &\leq \beta \mathbb{L}\left(\eta\right) + \beta \left\|\eta\right\|_s + [1-\beta] \left\|\mathcal{P}\eta\right\|_s \\ &\leq \beta \left\|\eta\right\|_{\mathfrak{B}} + [1-\beta] \left\|\mathcal{P}\eta\right\|_s ,\end{aligned}$$

which varifies the preliminary inequality for $k = 1$. Next we verify the inductive step from k to $k+1$.

$$\begin{aligned}\left\|\mathcal{P}^{k+1}\eta\right\|_{\mathfrak{B}} &\leq \beta^k \left\|\mathcal{P}\eta\right\|_{\mathfrak{B}} + [1-\beta] \sum_{j=0}^{k-1} \beta^j \left\|\mathcal{P}^{k-j}\mathcal{P}\eta\right\|_s \\ &\leq \beta^{k+1} \left\|\eta\right\|_{\mathfrak{B}} + [1-\beta]\beta^k \left\|\mathcal{P}\eta\right\|_s + [1-\beta] \sum_{j=0}^{k-1} \beta^j \left\|\mathcal{P}^{(k+1)-j}\eta\right\|_s \\ &\leq \beta^{(k+1)} \left\|\eta\right\|_{\mathfrak{B}} + [1-\beta] \sum_{j=0}^{(k+1)-1} \beta^j \left\|\mathcal{P}^{(k+1)-j}\eta\right\|_s .\end{aligned}$$

Therefore the preliminary inequality holds by induction and the proof is complete.

□

Next we treat the last term in eq. (5.7.20).

Lemma 5.7.14. *If* $\eta \in \mathfrak{L}_u$ *and* $k \geq 1$, *then*

$$\left\|\mathcal{P}^k\eta\right\|_s \leq 2\left(\beta^a\right)^k \left\|\eta\right\|_s + \left\|\eta\right\|_{\mathfrak{B}_w} . \tag{5.7.21}$$

Proof. Fix $\eta \in \mathfrak{L}_u$, $t \in [p, q]$, and $\psi \in \mathfrak{C}_a$ with $\|\psi\|_{\mathfrak{C}_a} \leq 1$. Let $\psi_0(x, y) = [\psi \circ T](x, 0)$ and note that ψ_0 is constant on each vertical line so that

$$H^a_x(\psi_0) = H^1_x(\psi_0) = 0.$$

We compute using the definition of ψ_0 as follows

$$\begin{aligned} \|(\psi \circ T - \psi_0)^x\|_{\sup} &= \sup_y |(\psi \circ T)(x, y) - (\psi \circ T)(x, 0)| \\ &\leq \sup_y H_x^a(\psi \circ T)\, |y|^a \\ &\leq \beta^a H_{u(x)}^a(\psi)\,. \end{aligned}$$

Applying the bound above we obtain

$$\begin{aligned} \|\psi \circ T - \psi_0\|_{\mathfrak{C}_a} &= \int_{[p,q]} \|(\psi \circ T - \psi_0)_x\|_{\sup} + H_x^a(\psi \circ T - \psi_0)\; d\bar{\mu}(x) \\ &= \int_{[p,q]} \|(\psi \circ T - \psi_0)_x\|_{\sup} + H_x^a(\psi \circ T)\; d\bar{\mu}(x) \\ &\leq \int_{[p,q]} \beta^a H_{u(x)}^a(\psi)\,. + \beta^a H_{u(x)}^a(\psi)\; d\bar{\mu}(x) \\ &\leq 2\beta^a \|\psi\|_{\mathfrak{C}_a}\,. \end{aligned}$$

Simmilarly we compute

$$\begin{aligned} \|\psi_0\|_{\mathfrak{C}_1} &= \int_{[p,q]} \|(\psi_0)_x\|_{\sup} + H_x^1(\psi_0)\; d\bar{\mu}(x) \\ &= \int_{[p,q]} \|(\psi_0)_x\|_{\sup}\; d\bar{\mu}(x) \\ &\leq \int_{[p,q]} \|\psi_x\|_{\sup}\; d\bar{\mu}(x) \\ &\leq \int_{[p,q]} \|\psi_x\|_{\sup} + H_x^a(\psi)\; d\bar{\mu}(x) \\ &\leq \|\psi\|_{\mathfrak{C}_a}\,. \end{aligned}$$

Applying the bounds above we obtain our result for $k = 1$,

$$\begin{aligned} \left| \int_\Lambda \delta_t(\mathcal{P}\eta)\; \psi\, d\mu \right| &= \left| \int_\Lambda \delta_{t_1}(\eta)\; [\psi \circ T - \psi_0]\; d\mu \right| \\ &\quad + \left| \int_\Lambda \delta_{t_1}(\eta)\; \psi_0\, d\mu \right| \\ &\leq \|\psi \circ T - \psi_0\|_{\mathfrak{C}_a} \|\eta\|_s + \|\psi_0\|_{\mathfrak{C}_1} \|\eta\|_{\mathfrak{B}_w} \\ &\leq 2\beta^a \|\eta\|_s + \|\eta\|_{\mathfrak{B}_w}\,. \end{aligned}$$

Applying an analogous argument to $\mathcal{P}^k$ yields the claimed bound. $\square$

Finally by combining Corollary 5.7.13.1 and Lemma 5.7.14 we are able to obtain the Lasota-Yorke inequality.

Proof of Proposition 5.7.13. Fix $\eta \in \mathfrak{L}_u$. We compute,

$$\begin{aligned}\left\|\mathcal{P}^k\eta\right\|_{\mathfrak{B}} &\le \beta^k \|\eta\|_{\mathfrak{B}} + 2\left(\beta^a\right)^k \|\eta\|_s + \|\eta\|_{\mathfrak{B}_w} \\ &\le 3\left(\beta^a\right)^k \|\eta\|_{\mathfrak{B}} + \|\eta\|_{\mathfrak{B}_w}.\end{aligned}$$

$\square$

5.8 Compactness

In the last section we obtained a Lasota-Yorke inequality for $\mathcal{P}$ acting on $\mathfrak{B}$. In this section we check the other main hypothesis of Theorem 2.4.4, namely the compact embedding of our strong space into our weak space.

Proposition 5.8.1. *The space $\mathfrak{B}$ is compactly embedded into $\mathfrak{B}_w$.*

In order to prove this proposition we will show that the unit ball of $\mathfrak{B}$ is totally bounded with respect to the norm on $\mathfrak{B}_w$. We begin by recalling relevant notation and results from previous sections. We will then prove a sequence of lemmas and will finish with the proof of the proposition.

5.8.1 Previous Notation and Results

Recall from Definition 5.5.2 and Remark 5.5.3 that that for each $x \in [p, q]$ we have defined a metric d_x on the vertical line $\ell_x \subset \Lambda$. Recall from Definitions 5.4.5 and 5.4.7 that Γ^u is a partition of Λ into graphs of functions mapping $[p, q]$ into $[0, 1]$ and that two points that lie in the same partition element of Γ^u are said to be *backward equivalent*. In Remark 5.5.3 it was observed that if (a, c) and (b, e) are backward equivalent and (a, d) and (b, f) are backward equivalent, then

$$d_a(c, d) = d_b(e, f).$$

In Definition 5.5.4 we introduced for each point $x \in [p,q]$ and Hölder exponent $a \in (0,1]$ a space $\mathcal{H}_x^a$ of a-Hölder functions on ℓ_x. This space is equipped with a Hölder norm $|\cdot|_x^a$, which is defined for $h\colon \ell_x \to \mathbb{C}$ in terms of the metric d_x on ℓ_x by

$$H_x^a(h) = \sup_{y \neq z} \frac{|h(y) - h(z)|}{d_x(y,z)^a},$$
$$|h|_x^a = \|h\|_{\sup} + H_x^a(h)\,.$$

In section 5.5 we introduced some notation related to functions on Λ. Given a function $\psi\colon \Lambda \to \mathbb{C}$ and a point $x \in [p,q]$ we introduced the *vertical section* of ψ, which is the function $\psi_x\colon \ell_x \to \mathbb{C}$ defined by $\psi_x(y) = \psi(x,y)$. We also introduced the convention that $H_x^a(\psi) := H_x^a(\psi_x)$ and $|\psi|_x^a := |\psi_x|_x^a$. The following semi-norm on bounded measurable functions $\psi\colon \Lambda \to \mathbb{C}$ was introduced in Definition 5.5.11,

$$|\psi|_a = \int_{[p,q]} |\psi|_x^a \, d\bar{\mu}(x).$$

In Definition 5.5.15 a Hölder exponent $a \in (0,1)$ was fixed and a standard "quotient and complete" construction yielded the Banach spaces $(\mathfrak{C}_a, \|\cdot\|_{\mathfrak{C}_a})$ and $(\mathfrak{C}_1, \|\cdot\|_{\mathfrak{C}_1})$ for which $\|\cdot\|_{\mathfrak{C}_a}$ and $\|\cdot\|_{\mathfrak{C}_1}$ are genuine norms.

Recall that in Definition 5.7.1 we defined for each $t \in [p,q]$ an operator δ_t on functions with domain Λ. If $\eta\colon \Lambda \to \mathbb{C}$, then $\delta_t(\eta)$ is constant on each unstable disk $\gamma \in \Gamma^u$ and the value of $\delta_t(\eta)$ on γ is[24] $\eta(\gamma \cap \ell_t)$. We also defined for each t and s in $[p,q]$ the operator defined by $\Delta_s^t(\eta) = \delta_t(\eta) - \delta_s(\eta)$. We then applied these operators in Definition 5.7.8 to introduce the norms,

$$\|\eta\|_{\mathfrak{B}_w} = \sup\left\{ \int_\Lambda \delta_t(\eta)\ \psi \, d\mu : t \in [p,q], \psi \in \mathfrak{C}_1, \|\psi\|_{\mathfrak{C}_1} \leq 1 \right\},$$
$$\|\eta\|_s = \sup\left\{ \int_\Lambda \delta_t(\eta)\ \psi \, d\mu : t \in [p,q], \psi \in \mathfrak{C}_a, \|\psi\|_{\mathfrak{C}_a} \leq 1 \right\},$$
$$\mathbb{L}(\eta) = \sup\left\{ \frac{\int_\Lambda \Delta_s^t(\eta)\ \psi \, d\mu}{|t-s|} : t,s \in [p,q], \psi \in \mathfrak{C}_a, \|\psi\|_{\mathfrak{C}_a} \leq 1 \right\},$$
$$\|\eta\|_{\mathfrak{B}} = \|\eta\|_s + \mathbb{L}(\eta)\,.$$

In Remark 5.7.10 we noted that $\|\cdot\|_{\mathfrak{B}_w} \leq \|\cdot\|_s \leq \|\cdot\|_{\mathfrak{B}}$.

[24]Since γ is the graph of a function and ℓ_t is a vertical line, the intersection $\gamma \cap \ell_x$ is a single point

Finally we recall a few ideas from section 5.6 that are described in more detail in appendix A.2. A set is E *saturated* with respect to Γ^u if for every $\gamma \in \Gamma^u$ such that $E \cap \gamma \neq \emptyset$ we have $\gamma \subseteq E$. The collection of all Borel subsets of Λ that are saturated with respect to Γ^u form a σ-algebra, which we denote by $\mathcal{B}^u$. In Definition 5.3.1 we defined μ which is normalized Lebesgue measure restricted to Λ. Let μ^u denote the restriction of μ to the σ-algebra $\mathcal{B}^u$. Given a bounded measurable function ψ on Λ let $\psi\mu$ denote the measure defined by $\psi\mu(E) = \int_E \psi\, d\mu$. The Radon-Nikodym derivative $\frac{d(\psi\mu)}{d\mu^u} \in L^1(\Lambda, \mathcal{B}^u, \mu^u)$ is referred to as the *unstable expectation* of ψ with respect to $\mathcal{B}^u$. In Definition 5.6.2 we defined an operator E^u. In Lemma 5.6.4 we showed that E^u is a version of the unstable expectation operator. The operator E^u viewed as an operator on either $\mathfrak{C}_a$ or $\mathfrak{C}_1$ has norm at most $[2\kappa+1]^2$ by Corollary 5.6.2.1, where κ is the distortion of the expanding factor map u as defined in Lemmas 5.3.4 and 5.3.12.

5.8.2 Preliminary Lemmas

Lemma 5.8.2. *If η and ψ are Borel measurable and $t \in [p, q]$, then*

$$E^u\left(\delta_t(\eta)\ \psi\right) = \delta_t(\eta)\ E^u\psi.$$

Proof. The function $\delta_x(\eta)$ is $\mathcal{B}^u$-measurable and E^u is a version of the conditional expectation with respect to $\mathcal{B}^u$. The identity is a standard property of conditional expectation. □

Definition 5.8.3. Let $\mathfrak{C}_a^u$ denote the set of all functions in $\mathfrak{C}_a$ that are $\mathcal{B}^u$-measurable.

Lemma 5.8.4. *If $\psi \in \mathfrak{C}_a^u$, then for all $\gamma \in \Gamma^u$, $\psi|_\gamma$ is a constant function.*

Proof. Fix $\gamma \in \Gamma^u$ and $(x, y) \in \gamma$. Let $a = \psi(x, y)$. The set $\psi^{-1}\{a\}$ is in $\mathcal{B}^u$ since ψ is $\mathcal{B}^u$-measurable. Every set in $\mathcal{B}^u$ is saturated with respect to Γ^u, therefore $\gamma \subseteq \psi^{-1}\{a\}$. We conclude that $\psi(w, z) = a$ for all $(w, z) \in \gamma$. □

Lemma 5.8.5. *If $\psi\colon \Lambda \to \mathbb{C}$ is constant along unstable disks, then for all $x, w \in [p, q]$,*

$$H_x^a(\psi) = H_w^a(\psi)\,,$$
$$|\psi|_x^a = |\psi|_w^a\,.$$

Proof. If (x, y) is backward equivalent to (w, y') and (x, z) is backward equivalent to (w, z'), then $\psi(x, y) = \psi(w, y')$ and $\psi(x, z) = \psi(w, z')$. This follows from the fact that unstable disks are backward equivalence classes. It has already been observed in the previous subsection that $d_x(y,z) = d_w(y',z')$. This yields for any $a \in (0,1]$,

$$\frac{\psi_x(y) - \psi_x(z)}{d_x(y,z)^a} = \frac{\psi_w(y') - \psi_x(z')}{d_w(y',z')^a}.$$

Since every point in ℓ_x is backward equivalent to a unique point in ℓ_w the identity above showed that the following sets are equal,

$$\left\{\frac{\psi_x(y) - \psi_x(z)}{d_x(y,z)^a} : (x,y) \neq (x,z) \in \ell_x\right\} = \left\{\frac{\psi_w(y) - \psi_w(z)}{d_w(y',z')^a} : (w,y') \neq (w,z') \in \ell_w\right\}.$$

Taking the supremum of each of the sets above yields

$$H_x^a(\psi) = H_w^a(\psi)\,.$$

This completes the proof of the first equality. The proof of the second equality is similar to the foregoing proof and is omitted.

□

Lemma 5.8.6. *Given $x \in [p,q]$ the mapping $\psi \mapsto \psi_x$ restricts to an isometric isomorphism of Banach spaces from $\mathfrak{C}_a^u$ onto $\mathcal{H}_x^a$. The same map provides an isometric isomorphism of $\mathfrak{C}_1^u$ onto $\mathcal{H}_x^1$.*

Proof. That the mapping $\psi \mapsto \psi_x$ is linear follows easily from the definitions. The mapping $\psi \mapsto \psi_x$ is onto. To verify this we will show that every function in $\mathcal{H}_x^a$ has an inverse $h \in \mathcal{H}_x^1$, define $\tau\colon \Lambda \to \mathbb{C}$ by setting the value of τ on each unstable disk $\gamma \in \Gamma^u$ to be $h(\gamma \cap \ell_x)$. Since τ is constant along unstable disks Lemma 5.8.5 applies and we see that for all $t \in [p,q]$

$$|\tau|_t^a = |\tau|_x^a = |h|_x^a\,.$$

Therefore, we have

$$\|\tau\|_{\mathfrak{C}_a} = \int_{[p,q]} |\tau|_t^a \, d\bar{\mu}(t) = |\tau|_x^a = |h|_x^a\,,$$

so $\tau \in \mathfrak{C}_a$. By definition $\tau_x(y) = h(y)$, so τ is the desired preimages of h and we have verified surjectivity.

To verify the isometric property fix $\psi \in \mathfrak{C}_a^u$ and note that by Lemma 5.8.4 ψ is constant along unstable disks. Calculating as before we obtain, $\|\psi\|_{\mathfrak{C}_a} = |\psi|_x^a = |\psi_x|_x^a$, which shows that $\psi \mapsto \psi_x$ is isometric.

Since $\psi \mapsto \psi_x$ is isometric it is also an injective map. This completes the proof. $\square$

Lemma 5.8.7. *Fix $x \in [p, q]$. Any ball of finite radius in $\mathcal{H}_x^1$ is totally bounded with respect the norm of $\mathcal{H}_x^a$.*

Proof. The result is standard, see for example Lemma 6.33 in [12]. We sketch a proof here for the convenience of the reader. Select a sequence $(h_k)_{k=1}^{\infty}$ contained in a ball of finite radius in $\mathcal{H}_x^1$. This sequence is bounded and equicontinuous, therefore by Arzela-Ascoli there exists a uniformly convergent subsequence $(b_k)_{k=1}^{\infty}$.

It remains to show that the sequence converges in $\mathcal{H}_x^a$. Begin by calculating

$$\begin{aligned}\left[\frac{|(b_j - b_k)(y) - (b_j - b_k)(z)|}{d_x(y,z)^a}\right]^{1/a} &= \frac{|(b_j - b_k)(y) - (b_j - b_k)(z)|}{d_x(y,z)} \\ &\quad \times |(b_j - b_k)(y) - (b_j - b_k)(z)|^{1/a-1} \\ &\leq 2H_x^1(b_j - b_k)\, \|b_j - b_k\|_{\sup}^{1/a-1}\end{aligned}$$

From this we deduce that

$$H_x^a(b_j - b_k)^{1/a} \leq 2H_x^1(b_j - b_k)\, \|b_j - b_k\|_{\sup}^{1/a-1}.$$

Since the sequence b_k is bounded in the $|\cdot|_x^1$-norm we have that $H_x^1(b_j - b_k)$ is bounded uniformly in j and k. Since b_k is a uniformly convergent sequence $\|b_j - b_k\|_{\sup}$ can be made arbitrarily small by selecting j and k sufficiently large. We conclude that (b_k) is Cauchy in the $|\cdot|_x^a$-norm. $\square$

Corollary 5.8.7.1. *Any ball of finite radius in $\mathfrak{C}_1^u$ is totally bounded with respect to the norm of $\mathfrak{C}_a$.*

Lemma 5.8.8. *For each $\epsilon > 0$, there exists a finite set $A_\epsilon \subseteq \mathfrak{C}_1^u$ such that for all $\psi \in \mathfrak{C}_1$, there exists $\xi \in A_\epsilon$ such that for all $\eta \in \mathfrak{L}_u$ with $\|\eta\|_{\mathfrak{B}} \leq 1$ and all $t \in [p, q]$,*

$$\left| \int_\Lambda \delta_t(\eta)\ \psi\, d\mu - \int_\Lambda \delta_t(\eta)\ \xi\, d\mu \right| < \epsilon\, \|\eta\|_s .$$

Proof. By Lemma 5.8.2 we have for all $t \in [p,q]$ and $\psi \in \mathfrak{C}_1$.

$$\int_\Lambda \delta_t(\eta)\ \psi\, d\mu = \int_\Lambda \delta_t(\eta)\ E^u\psi\, d\mu$$

By Corollary 5.6.2.1 we have that $E^u\psi \in \mathfrak{C}_1$ and $\|E^u\psi\|_{\mathfrak{C}_1} \leq [2\kappa+1]^2\,\|\psi\|_{\mathfrak{C}_1}$. By Lemma 5.6.4 $E^u\psi$ is a version of the unstable expectation of ψ and thus $E^u\psi \in \mathfrak{C}_1^u$. Let $E = \left\{E^u\psi : \psi \in \mathfrak{C}_1, \|\psi\|_{\mathfrak{C}_1} \leq 1\right\}$ and $F = \left\{\psi \in \mathfrak{C}_1^u : \|\psi\|_{\mathfrak{C}_1} \leq [2\kappa+1]^2\right\}$ and note that $E \subseteq F$. The set F is a bounded ball in $\mathfrak{C}_1^u$, therefore by Corollary 5.8.7.1 we may select a finite set $A_\epsilon \subseteq F$ that is ϵ-dense with respect to the norm on $\mathfrak{C}_a$.

Fix $\psi \in \mathfrak{C}_1$ and select $\xi \in A_\epsilon$ such that $\|\xi - E^u\psi\|_{\mathfrak{C}_a} < \epsilon$. Now we compute as follows,

$$\begin{aligned}\left|\int_\Lambda \delta_t(\eta)\ \psi\, d\mu - \int_\Lambda \delta_t(\eta)\ \xi\, d\mu\right| &= \left|\int_\Lambda \delta_t(\eta)\ E^u\psi\, d\mu - \int_\Lambda \delta_t(\eta)\ \xi\, d\mu\right| \\ &\leq \|\eta\|_s\,\|E^u\psi - \xi\|_{\mathfrak{C}_a} \\ &\leq \epsilon\,\|\eta\|_s\,.\end{aligned}$$

□

Definition 5.8.9. For each $\epsilon > 0$ let $B_\epsilon \subseteq [p,q]$ be a finite ϵ-dense subset.

Lemma 5.8.10. *For all $\epsilon > 0$, $t \in [p,q]$, and $\psi \in \mathfrak{C}_1$ with $\|\psi\|_{\mathfrak{C}_1} \leq 1$ there exist $s \in B_\epsilon$ and $\xi \in A_\epsilon$ such that for all $\eta \in \mathfrak{L}_u$ with $\|\eta\|_{\mathfrak{B}} \leq 1$,*

$$\left|\int_\Lambda \delta_t(\eta)\ \psi\, d\mu - \int_\Lambda \delta_s(\eta)\ \xi\, d\mu\right| < \|\eta\|_{\mathfrak{B}}\,\epsilon.$$

Proof. Fix $\epsilon > 0$, $t \in [p,q]$ and $\psi \in \mathfrak{C}_1$ with $\|\psi\|_{\mathfrak{C}_1} \leq 1$. Select $\xi \in A_\epsilon$ such that $\|E^u\psi - \xi\|_{\mathfrak{C}_a} \leq \epsilon$ and $s \in B_\epsilon$ such that $|t-s| < \epsilon$. We compute

$$\begin{aligned}\left|\int_\Lambda \delta_t(\eta)\ \psi\, d\mu - \int_\Lambda \delta_s(\eta)\ \xi\, d\mu\right| &\leq \left|\int_\Lambda \Delta_s^t(\eta)\ \psi\, d\mu\right| + \left|\int_\Lambda \delta_s(\eta)\ [\psi - \xi]\ d\mu\right| \\ &= \left|\int_\Lambda \Delta_s^t(\eta)\ \psi\, d\mu\right| + \left|\int_\Lambda \delta_s(\eta)\ [E^u\psi - \xi]\ d\mu\right| \\ &\leq \mathbb{L}\,(\eta)\,|t-s| + \|\eta\|_s\,\|E^u\psi - \xi\|_{\mathfrak{C}_a} \\ &\leq \epsilon\,\|\eta\|_{\mathfrak{B}}\end{aligned}$$

□

Definition 5.8.11. Let $\{(\xi_j, s_j) : j = 1, \cdots, |A_\epsilon|\,|B_\epsilon|\}$ be an enumeration of $A_\epsilon \times B_\epsilon$. Define $v_j \colon \{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\} \to \mathbb{R}$ by

$$v_j(\eta) = \int_\Lambda \delta_{s_j}(\eta)\ \xi_j\, d\mu$$

and $v \colon \{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\} \to \mathbb{R}^{|A_\epsilon||B_\epsilon|}$ by letting $v(\eta)$ be the vector with $v_j(\eta)$ in the j-th coordinate. Let V denote the image of v.

Lemma 5.8.12. *The set V is a totally bounded subset of $\mathbb{R}^{|A_\epsilon||B_\epsilon|}$ with respect to the max norm $|\cdot|_{\max}$.*

Proof. Note that for all $j = 1, \cdots, |A_\epsilon|\,|B_\epsilon|$,

$$|v_j(\eta)| = \left|\int_\Lambda \delta_{s_j}(\eta)\ \ \xi_j\, d\mu\right| \leq \|\eta\|_s\, \|\xi_j\|_{\mathfrak{C}_a} \leq [2\kappa + 1]^2$$

and thus $|v(\eta)|_{\max} < [2\kappa + 1]^2$. By the Heine-Borel theorem the closure of V is compact and thus V is totally bounded. □

Definition 5.8.13. For $\epsilon > 0$ let V_ϵ be a finite ϵ-dense subset of V and let $\left\{w^k : k = 1, \cdots, |V_\epsilon|\right\}$ be an enumeration of V_ϵ.

Definition 5.8.14. For $\epsilon > 0$ and each $k = 1, \cdots, |V_\epsilon|$ define

$$U_k = \left\{\eta \in \mathfrak{L}_u : \left|v(\eta) - w^k\right|_{\max} < \epsilon\right\}$$

Lemma 5.8.15. *The collection $\{U_k : k = 1, \cdots, |V_\epsilon|\}$ is a cover of $\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$. Furthermore, for all $k = 1, \cdots, |V_\epsilon|$ and $\eta, \nu \in U_k$, we have $\|\eta - \nu\|_{\mathfrak{B}_w} < 4\epsilon$.*

Proof. First we verify that the collection covers. If $\eta \in \mathfrak{L}_u$ with $\|\eta\|_{\mathfrak{B}} \leq 1$, then by Definitions 5.8.11 and 5.8.13 there exists $w^k \in V_\epsilon$ such that

$$\left|v(\eta) - w^k\right|_{\max} < \epsilon.$$

Therefore, $\eta \in U_k$. Since η was arbitrary this completes the proof that the collection of U_k's cover.

Next we verify the bound on the diameter of the U_k sets. Fix $t \in [p,q]$ and $\psi \in \mathfrak{C}_1$ with $\|\psi\|_{\mathfrak{C}_1} \leq 1$. By Lemma 5.8.10 there exists $1 \leq j \leq |V_\epsilon|$ such that

$$\left| \int_\Lambda \delta_t(\eta - \nu)\ \psi\, d\mu - \int_\Lambda \delta_{s_j}(\eta - \nu)\ \xi_j\, d\mu \right| < \|\eta - \nu\|_{\mathfrak{B}}\, \epsilon \leq 2\epsilon.$$

Since $\eta, \nu \in U_k$, we have

$$\left| \int_\Lambda \delta_{s_j}(\eta - \nu)\ \xi_j\, d\mu \right| \leq \left| \int_\Lambda \delta_{s_j}(\eta)\ \xi_j\, d\mu - w_j^k \right| + \left| \int_\Lambda \delta_{s_j}(\nu)\ \xi_j\, d\mu - w_j^k \right| < 2\epsilon.$$

Therefore,

$$\left| \int_\Lambda \delta_t(\eta - \nu)\ \psi\, d\mu \right| \leq \left| \int_\Lambda \delta_t(\eta - \nu)\ \psi\, d\mu - \int_\Lambda \delta_{s_j}(\eta - \nu)\ \xi_j\, d\mu \right| + \left| \int_\Lambda \delta_{s_j}(\eta - \nu)\ \xi_j\, d\mu \right| < 4\epsilon,$$

Since t and ψ were arbitrary we conclude that $\|\eta - \nu\|_{\mathfrak{B}_w} < 4\epsilon$ as desired. □

5.8.3 Conclusion

Proof of Proposition 5.8.1. Given $\epsilon > 0$ Lemma 5.8.15 shows that the collection

$$\{U_k : k = 1, \cdots, |V_\epsilon|\}$$

is a finite cover of

$$\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

by sets of diameter at most 4ϵ with respect to the norm on $\mathfrak{B}_w$. Select a point $\eta_k \in U_k$ for each $k = 1, \cdots, |V_\epsilon|$. Let

$$B_k = \left\{\eta \in \mathfrak{B}_w : \|\eta - \eta_k\|_{\mathfrak{B}_w} < 5\epsilon\right\}$$

and note that $U_k \subseteq B_k$. The collection

$$\{B_k : k = 1, \cdots |V_\epsilon|\}$$

is a cover of

$$\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

by balls of radius 5ϵ. Therefore, the set

$$\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

is totally bounded with respect to the norm on $\mathfrak{B}_w$ and thus is precompact. From Definition 5.7.11 we see that

$$\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

is dense in

$$\{\eta \in \mathfrak{B} : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

with respect to the norm on $\mathfrak{B}_w$. Therefore the $\mathfrak{B}_w$-closure of

$$\{\eta \in \mathfrak{B} : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

is contained in the $\mathfrak{B}_w$-closure of

$$\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

which is compact. Therefore,

$$\{\eta \in \mathfrak{B} : \|\eta\|_{\mathfrak{B}} \leq 1\}$$

is precompact in $\mathfrak{B}_w$. We conclude that $\mathfrak{B}$ is compactly embedded into $\mathfrak{B}$. □

5.8.4 A Point of Interest

Remark 5.8.16. The collection $\{U_k : k = 1, \cdots, |V_\epsilon|\}$ is an open cover of $\{\eta \in \mathfrak{L}_u : \|\eta\|_{\mathfrak{B}} \leq 1\}$ in the topology of $\mathfrak{B}_w$.

Proof. We have already proved that the collection covers in Lemma 5.8.15. We verify that each U_k is open in the topology of $\mathfrak{B}_w$. Fix $\eta \in U_k$ and select $\delta > 0$ such that

$$\left|v(\eta) - w^k\right| + \delta < \epsilon.$$

If $\nu \in \mathfrak{L}_u$ with $\|\nu\|_{\mathfrak{B}} \leq 1$ and $[2\kappa+1]^2 \|\eta - \nu\|_{\mathfrak{B}_w} < \delta$, then

$$\begin{aligned} |v(\eta) - v(\nu)|_{\max} &= \max_j \left| \int_\Lambda \delta_{t_j}(\eta - \nu)\ \psi_j \, d\mu \right| \\ &\leq \max_j \|\eta - \nu\|_{\mathfrak{B}_w} \|\psi_j\|_{\mathfrak{C}_1} \\ &\leq [2\kappa+1]^2 \|\eta - \nu\|_{\mathfrak{B}_w} \\ &< \delta \end{aligned}$$

Finally,

$$\left|v(\nu) - w^k\right|_{\max} \leq |v(\nu) - v(\eta)|_{\max} + \left|v(\eta) - w^k\right|_{\max} < \epsilon,$$

as desired. □

5.9 Renewal Theory and Decay Rates for B

In the previous two sections we have collected a Lasota-Yorke inequality and compact embedding result for the Frobenius-Perron operator of the induced map T acting on Banach spaces $\mathfrak{B}$ and $\mathfrak{B}_w$. By Theorem 2.4.4 this is sufficient to deduce an exponential rate of decay of correlation for the induced map T. In this section we will use operator renewal theory to connect properties of the Frobenius-Perron operator of the original intermittent baker's transformation B with properties of the Frobenius-Perron operator of the induced map T. In particular we will deduce a polynomial rate of decay of correlations for IBT B.

5.9.1 Previous Notation and Results

Throughout this section $\mathcal{P}_B$ will denote the Frobenius-Perron operator of the intermittent baker's map B and $\mathcal{P}$ will denote the Frobenius-Perron operator of the induced map T.

5.9.2 Outline of the Argument

We will study operators that capture the return time structure of the map T. Before we define these operators we describe intuitively what they are doing.

There are two events of interest. First, if a point (x,y) is in Λ and at time $n \geq 0$ the orbit of (x,y) returns for the first time[25] to Λ, then we say that (x,y) has a first return at time n. We will define a *first return time operator* R_n for each $n \geq 1$ that is related to this event. If we view $\eta\colon \Lambda \to \mathbb{R}$ as a probability density function, then $R_n\eta$ will be a probability density function representing the probability that a point $(w,z) \in \Lambda$ is the first return at time n of a point (x,y) selected from Λ according to the distribution of η. Second, if (x,y) is a point in Λ and at time $n \geq 1$ the orbit of (x,y) returns to Λ, not necessarily for the first time then we say that (x,y) has *returned at time n*. Given a probability density η on Λ we will define an operator Q_n such that $Q_n\eta$ is the density representing the probability that a point (w,z) is a return at time n for a point $(x,y) \in \Lambda$ selected according to the distribution of η. Finally, If $1 \leq j \leq k$, then we interpret the product $Q_{k-j}R_j\eta$ as the probability that a point (w,z) is a return to Λ at time k of some point (x,y) selected according to the distribution of η given that (x,y) made its first return to Λ at time j. With this interpretation it is plausible that by the probability that (w,z) is a return at time n of some point (x,y) chosen according to the distribution of η would be equal to the sum of the conditional probabilities of the disjoint first return times of (x,y), that is

$$Q_n = R_n + \sum_{j=1}^{n-1} Q_{n-j}R_j. \tag{5.9.1}$$

The equation above is often referred to as the *renewal equation*, it can be reformulated in terms of generating functions as we will see below. The generating function formulation of the renewal equation is verified in Lemma 5.9.8 by applying Proposition 5.9.4.

Having given an intuitive description of the operators R_n and Q_n we now provide a rigorous definition.

Definition 5.9.1. Let B be an IBT with Frobenius-Perron operator $\mathcal{P}_B$ and $T\colon \Lambda \circlearrowleft$ be be the associated induced map with return time $r\colon \Lambda \to \mathbb{N} \cup \{\infty\}$ as in Definition 5.3.1 and Frobenius-Perron operator $\mathcal{P}$. For each $n \geq 1$ define operators Q_n and R_n on $L^1(\Lambda,\mu)$ by

$$Q_n\eta = \mathbf{1}_\Lambda \left[\mathcal{P}_B^n \left(\mathbf{1}_\Lambda \eta\right)\right],$$
$$R_n\eta = \mathcal{P}\left(\mathbf{1}_{\{r=n\}}\eta\right).$$

[25] $T^n(x,y) \in \Lambda$ and for $1 \leq k < n$, $T^k(x,y) \notin \Lambda$.

We will investigate the asymptotic properties of the operators R_n and Q_n by analyzing their generating functions

$$R(z) = \sum_{n=1}^{\infty} z^n R_n$$
$$Q(z) = I + \sum_{n=1}^{\infty} z^n Q_n$$

If we formally multiply the power series above keeping eq. (5.9.1) in mind we see that

$$\begin{aligned} Q(z)R(z) &= R(z) + \sum_{n=1}^{\infty} z^n \sum_{j=1}^{n-1} Q_{n-j} R_j \\ &= \sum_{n=1}^{\infty} z^n \left[R_n + \sum_{j=1}^{n-1} Q_{n-j} R_j \right] \\ &= Q(z) - I. \end{aligned}$$

Rearranging the identity above we obtain the identity $Q(z)(I - R(z)) = I$. The next lemma shows that this formal manipulation is justified provided some basic dynamical conditions are satisfied. Before we state the lemma we recall two standard definitions.

Definition 5.9.2. Given a measure space $(X, \mathcal{A}, \nu)$ and a measurable map $S\colon X \circlearrowleft$, we say that S is *non-singular* if for all $A \in \mathcal{A}$ such that $\nu(A) = 0$, we have $\nu\left(S^{-1}(A)\right) = 0$.

Definition 5.9.3. Given a measure space $(X, \mathcal{A}, \nu)$ and a measurable map $S\colon X \circlearrowleft$ that is non-singular, we say that a set $A \in \mathcal{A}$ is *wandering* if the sets $S^{-k}(A)$ for $k \geq 0$ are disjoint. If every wandering set is a null set, then we say that S is *conservative*.

We can now state the lemma.

Proposition 5.9.4 (Sarig [23] Proposition 1)**.** *Let (X, μ, τ) be a conservative non-singular transformation, and assume that $A \subseteq X$ has finite positive measure. If Q_n and R_n are given by $Q_n f = \mathbf{1}_A \mathcal{P}_\tau^n (\mathbf{1}_A f)$ and $R_n f = \mathbf{1}_A \mathcal{P}_\tau^n \left(\mathbf{1}_{\{r_A = n\}} f\right)$, then for all $|z| < 1$,*

$$Q(z) = (I - R(z))^{-1} \text{ where } R(z) = \sum_{n \geq 1} z^n R_n \text{ and } Q(z) = I + \sum_{n \geq 1} z^n Q_n.$$

Furthermore, $R(1)$ is the transfer operator of the return map τ^{r_A}.

We will see in Lemma 5.9.8 that B satisfies the hypothesis of the proposition above. In what follows we will need to analyze powers of the operator $R(z)$. The next definition and lemma provide some useful notation.

Definition 5.9.5. Given $k \geq 1$ and $(x,y) \in \Lambda$, let $r^{(k)}(x,y)$ denote the k-th time that the orbit of (x,y) returns to Λ. For each $n \geq 1$ and $k \geq 1$ define

$$R_n^{(k)}\eta = \mathcal{P}^k\left(\mathbf{1}_{\{r^{(k)}=n\}}\eta\right)$$

Remark 5.9.6. For each $k \geq 1$ the following identity holds for all $\eta \in L^1(\Lambda, \mu)$.

$$R(z)^k\eta = \sum_{n=\min r^{(k)}}^{\infty} z^n R_n^{(k)}\eta.$$

Having defined the generating functions $R(z)$ and $Q(z)$ we wish analyze them to obtain control over the asymptotic behavior of the Q_n operators. As we will see in Theorem 5.9.17 asymptotic control over the Q_n operators provides a sharp polynomial rate of decay of correlations for the map B. The following Theorem of Gouëzel relates asymptotic behavior of the Q_n operators to the asymptotic behavior of the R_n operators which are defined in terms of the induced map T which we have already analyzed at length in the preceding sections.

Theorem 5.9.7 (Gouëzel [13])**.** *Let Q_n be bounded operators on a Banach space $\mathcal{L}$ such that $Q(z) = I + \sum_{n\geq 1} z^n Q_n$ converges in $Hom(\mathcal{L},\mathcal{L})$[26] for every $z \in \mathbb{C}$ with $|z| < 1$. Assume that:*

1. **Renewal equation:** *for every $z \in \mathbb{C}$ with $|z| < 1$, $Q(z) = (I - R(z))^{-1}$ where $R(z) = \sum_{n\geq 1} z^n R_n$, $R_n \in Hom(\mathcal{L},\mathcal{L})$ and $\sum \|R_n\| < +\infty$.*
2. **Spectral Gap:** 1 *is a simple isolated eigenvalue of $R(1)$.*
3. **Aperiodicity:** *for every $z \neq 1$ with $|z| \leq 1$, $I - R(z)$ is invertible.*

Let Π_1 be the eigenprojection of $R(1)$ at 1. If $\sum_{k>n} \|R_k\| = O\left(1/n^\beta\right)$ for some $\beta > 1$ and $\Pi_1 R'(1)\Pi_1 \neq 0$, then for all n

$$Q_n = \frac{1}{m}\Pi_1 + \frac{1}{m^2}\sum_{k=n+1}^{+\infty} P_k + E_n$$

[26]With the strong operator topology

where m *is given*[27] *by* $\Pi_1 R'(1)\Pi_1 = m\Pi_1$, $P_n = \sum_{l>n} \Pi_1 R_l \Pi_1$ *and* $E_n \in Hom(\mathcal{L},\mathcal{L})$ *satisfy*

$$\|E_n\| = \begin{cases} O\left(1/n^{\beta}\right), & \text{if } \beta > 2; \\ O\left(\log(n)/n^2\right), & \text{if } \beta = 2; \\ O\left(1/n^{2\beta-2}\right), & \text{if } 2 > \beta > 1. \end{cases}$$

The remainder of this section is divided into subsections aimed at verifying that the return time operators from Definition 5.9.1 satisfy the hypotheses of Theorem 5.9.7 followed by a subsection in which decay rates are derived.

5.9.3 Renewal equation

In this section we will apply Proposition 5.9.4 to verify the renewal equation and prove the required bound on the sum of the norms of the R_n operators. We begin with a lemma verifying that Proposition 5.9.4 applies and guaranties that the renewal equation holds.

Lemma 5.9.8. *If* B *is a IBT, then* $([0,1]^2, \mu, B)$ *is conservative and non-singular, and* Λ *has finite positive measure.*

Proof. The map B is Lebesgue measure preserving, thus B is non-singular.

Suppose that W is a wandering set, then $\bigcup_{k=0}^{\infty} B^{-k}(W) \subset [0,1]^2$ and thus $\lambda\left(\bigcup_{k=0}^{\infty} B^{-k}(W)\right) \leq 1$. Since W is wandering the sets $B^{-k}(W)$ are disjoint and since B preserves Lebesgue measure $\lambda(B^{-k}W) = \lambda(W)$ for all k, therefore,

$$\lambda\left(\bigcup_{k=0}^{\infty} B^{-1}W\right) = \sum_{k=0}^{\infty} \lambda\left(B^{-k}(W)\right) = \sum_{k=0}^{\infty} \lambda(W)$$

The equation above and the inequality $\lambda(W) \leq 1$ can only be simultaneously satisfied if $\lambda(W) = 0$, thus W is a null-set. Since W was an arbitrary wandering set we conclude that B is conservative.

Since Λ is a non-degenerate rectangle it has positive Lebesgue measure. □

Next we collect a few technical lemmas relating the operators R_n, $\delta_t(\cdot)$, and $\Delta_s^t(\cdot)$.

[27] Here $R'(1)$ denotes the operator $\frac{d}{dz}R|_{z=1}$.

Lemma 5.9.9. *For all* $k \geq 1$, $n \geq \min r^{(k)}$, $t \in [p,q]$, $(x,y) \in \Lambda$, $z \in \mathbb{C}$ *with* $|z| < 1$, *and* $\eta \in \mathfrak{L}_u$,

$$\delta_t\big(R_n^{(k)}\eta\big)\,(x,y) = \mathbf{1}_{T^k\left\{r^{(k)}=n\right\}}(x,y)\,\delta_t\big(\mathcal{P}^k\eta\big)\,(x,y), \tag{5.9.2}$$

and

$$\delta_t\big(R(z)^k\eta\big)\,(x,y) = z^{r^{(k)}(T^{-k}(x,y))}\delta_t\big(\mathcal{P}^k\eta\big)\,(x,y). \tag{5.9.3}$$

Proof. We begin by computing,

$$\begin{aligned}
\delta_t\big(R_n^{(k)}\eta\big)\,(x,y) &= R_n^{(k)}\eta\,(\ell_t \cap \pi_u(x,y)) \\
&= \mathbf{1}_{\left\{r^{(k)}=n\right\}}\left(T^{-k}\left(\ell_t \cap \pi_u(x,y)\right)\right)\mathcal{P}^k\delta_{t_k}(\eta)\,(x,y) \\
&= \mathbf{1}_{T^k\left\{r^{(k)}=n\right\}}(x,y)\,\delta_t\big(\mathcal{P}^k\eta\big)\,(x,y).
\end{aligned}$$

The first step is an application of Definition 5.7.1. The second step is an application of Definition 5.9.1 and eq. (5.7.6). In the last step we have used the fact that $T^k\left\{r^{(k)}=n\right\} \in \mathscr{Z}_T^{-k}$ and hence $\ell_t \cap \pi_u(x,y) \in T^k\left\{r^{(k)}=n\right\}$ if and only if $(x,y) \in T^k\left\{r^{(k)}=n\right\}$. We have also applied eq. (5.7.6) in the last step. The value of $\mathbf{1}_{T^k\{r=n\}}(x,y)$ is 1 if and only if $r^{(k)}(T^{-k}(x,y)) = n$. The second claim now follows from Remark 5.9.6. □

Corollary 5.9.9.1. *For all* $k \geq 1$,$n \geq \min r^{(k)}$, $s,t \in [p,q]$, $(x,y) \in \Lambda$, $z \in \mathbb{C}$ *with* $|z| < 1$, *and* $\eta \in \mathfrak{L}_u$,

$$\Delta_s^t\big(R_n^{(k)}\eta\big)\,(x,y) = \mathbf{1}_{T^k\left\{r^{(k)}=n\right\}}(x,y)\,\Delta_s^t\big(\mathcal{P}^k\eta\big)\,(x,y), \tag{5.9.4}$$

and

$$\Delta_s^t\big(R(z)^k\eta\big)\,(x,y) = z^{r^{(k)}(T^{-k}(x,y))}\Delta_s^t\big(\mathcal{P}^k\eta\big)\,(x,y). \tag{5.9.5}$$

The next proposition verifies that the operator norms of the R_n operators are summable, which completes the verification of the renewal equation hypothesis.

Proposition 5.9.10. *For all* $k \geq 1$, $n \geq \min r^{(k)}$, *and* $\eta \in \mathfrak{L}_u$,

$$\left\|R_n^{(k)}\eta\right\|_{\mathfrak{B}} \leq [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\|\eta\|_{\mathfrak{B}} \tag{5.9.6}$$

$$\left\|R_n^{(k)}\eta\right\|_{\mathfrak{B}} \leq [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\left[3\left(\beta^a\right)^k\|\eta\|_{\mathfrak{B}} + \|\eta\|_{\mathfrak{B}_w}\right] \tag{5.9.7}$$

Proof. We will verify the inequalities below using arguments similar to those of section 5.7 that were used to prove Proposition 5.7.13.

$$\begin{aligned}
\left\|R_n^{(k)}\eta\right\|_{\mathfrak{B}_w} &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\|\eta\|_{\mathfrak{B}_w}\\
\left\|R_n^{(k)}\eta\right\|_s &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\|\eta\|_s\\
\mathbb{L}\left(R_n^{(k)}\eta\right) &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\beta^k\mathbb{L}(\eta)\\
\left\|R_n^{(k)}\eta\right\|_s &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\left[2(\beta^a)^k\|\eta\|_s+\|\eta\|_{\mathfrak{B}_w}\right]
\end{aligned}$$

Combing the second and third inequalities above yields eq. (5.9.6) and combining all four inequalities as in the proof of Proposition 5.7.13 yields eq. (5.9.7).

The key observations required to verify the inequalities above are the following integral equalities, which follow from eqs. (5.9.2) and (5.9.4). For $\psi\in\mathfrak{C}_a$ or $\mathfrak{C}_1$,

$$\begin{aligned}
\int_\Lambda \delta_t\big(R_n^{(k)}\eta\big)\;\psi\,d\mu &= \int_\Lambda \delta_{t_k}(\eta)\,\mathbf{1}_{\left\{r^{(k)}=n\right\}}\psi\circ T^k\,d\mu\\
\int_\Lambda \Delta_s^t\big(R_n^{(k)}\eta\big)\;\psi\,d\mu &= \int_\Lambda \Delta_{s_k}^{t_k}(\eta)\,\mathbf{1}_{\left\{r^{(k)}=n\right\}}\psi\circ T^k\,d\mu
\end{aligned}$$

Applying Lemma 5.5.14 we see that

$$\begin{aligned}
\left\|\mathbf{1}_{\left\{r^{(k)}=n\right\}}\psi\circ T^k\right\|_{\mathfrak{C}_1} &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\|\psi\|_{\mathfrak{C}_1},\\
\left\|\mathbf{1}_{\left\{r^{(k)}=n\right\}}\psi\circ T^k\right\|_{\mathfrak{C}_a} &\le [\kappa+1]\,\mu\left\{r^{(k)}=n\right\}\|\psi\|_{\mathfrak{C}_a}.
\end{aligned}$$

The first three claimed inequalities then follow by an argument similar to the proof of Lemma 5.7.12 and the fourth of the claimed inequalities follows from an argument similar to the proof of Lemma 5.7.14, where in each case $\mathbf{1}_{\left\{r^{(k)}=n\right\}}\psi\circ T^k$ replaces $\psi\circ T^k$ at the end of the argument. □

Lemma 5.9.11. *If $|z|\le 1$ and $\eta\in\mathfrak{L}_u$, then*

$$\|\mathbb{R}(z)\eta\|_{\mathfrak{L}_u}\le\|\eta\|_{\mathfrak{L}_u}.$$

Proof. Recall that

$$\|R(z)\eta\|_{\mathfrak{L}_u}=\|R(z)\eta\|_{\sup}+\sup\left\{\frac{\Delta_s^t(R(z)\eta)\,(x,y)}{|s-t|}:(x,y)\in\Lambda,s,t\in[p,q]\right\}.$$

We begin by computing the first term,

$$\begin{aligned}
\|R(z)\eta\|_{\sup} &= \sup_{(x,y)\in\Lambda}\left|\sum_{n=1}^{\infty} z^n\, R_n\eta(x,y)\right| \\
&= \sup_{(x,y)\in\Lambda}\left|\sum_{n=1}^{\infty} z^n\, \mathcal{P}_T\left(\mathbf{1}_{\{r=n\}}\,\eta\right)(x,y)\right| \\
&= \sup_{(x,y)\in\Lambda}\left|\sum_{n=1}^{\infty} z^n\, [\mathbf{1}_{\{r=n\}}\circ T^{-1}](x,y)\,[\eta\circ T^{-1}](x,y)\right| \\
&= \sup_{(x,y)\in\Lambda}\left|z^{r\left(T^{-1}(x,y)\right)}\,[\eta\circ T^{-1}](x,y)\right| \\
&\le \sup_{(x,y)\in\Lambda}\left|[\eta\circ T^{-1}](x,y)\right| \\
&\le \|\eta\|_{\sup}
\end{aligned}$$

For the second term fix $(x,y)\in\Lambda$ and $s,t\in[p,q]$, we apply Corollary 5.9.9.1 and Lemma 5.7.7,

$$\begin{aligned}
\left|\Delta_s^t(R(z)\eta)\,(x,y)\right| &= \left|z^{r\left(T^{-1}(x,y)\right)}\Delta_s^t(\mathcal{P}_T\eta)\right| \\
&\le \left|\Delta_s^t(\mathcal{P}_T\eta)\right| \\
&\le \beta L_u\,(\eta)\,|s-t|\,.
\end{aligned}$$

Since (x,y), s, and t were arbitrary $L_u\,(R(z)\eta)\le\beta L_u\,(\eta)$. We conclude that

$$\|R(z)\eta\|_{\mathfrak{L}_u}\le\|\eta\|_{\sup}+\beta L_u\,(\eta)\le\|\eta\|_{\mathfrak{L}_u}\,.$$

□

5.9.4 Preliminary Spectral Results

Before we can verify the spectral gap or aperiodicity hypothesis of Theorem 5.9.7 we will need a few lemmas establishing more elementary spectral properties of $R(z)$. We begin by verifying that $R(z)$ is quasi-compact for z in the complex unit disk. The first step is the following uniform Lasota-Yorke inequality.

Proposition 5.9.12. *For all $\eta\in\mathfrak{L}_u$ and $k\ge 1$*

$$\left\|R(z)^k\eta\right\|_{\mathfrak{B}}\le[\kappa+1]\,|z|^k\left[3\left(\beta^a\right)^k\|\eta\|_{\mathfrak{B}}+\|\eta\|_{\mathfrak{B}_w}\right]. \tag{5.9.8}$$

Proof. We note that $\min r^{(k)} \geq 2k$ and apply eq. (5.9.7) so that we have

$$\begin{aligned}
\left\|R(z)^k\eta\right\|_{\mathfrak{B}} &\leq \sum_{n=2k}^{\infty} |z^n| \left\|R_n^{(k)}\eta\right\|_{\mathfrak{B}} \\
&\leq |z|^k \sum_{n=2k}^{\infty} [\kappa+1]\mu\left\{r^k = n\right\}\left[3\left(\beta^a\right)^k \|\eta\|_{\mathfrak{B}} + \|\eta\|_{\mathfrak{B}_w}\right] \\
&= [\kappa+1]\,|z|^k \left[3\left(\beta^a\right)^k \|\eta\|_{\mathfrak{B}} + \|\eta\|_{\mathfrak{B}_w}\right].
\end{aligned}$$

Obviously we could have obtained $|z|^{2k}$ as a multiplier in the inequality above. We opt for the weaker bound as it makes no difference in what follows and is slightly less cumbersome. □

In the next lemma we record that the operators $R(z)$ are quasi-compact.

Lemma 5.9.13. *For each $|z| \leq 1$ the operator $R(z)\colon \mathfrak{B} \circlearrowleft$ is quasi-compact with spectral radius $\rho\left(R(z)\right) \leq |z|$ and essential spectral radius $\rho_{ess}\left(R(z)\right) \leq \beta^a |z|$.*

Proof. This follows from Theorem 2.4.4 and Propositions 5.8.1 and 5.9.12. □

We can refine the result above by obtaining more control over the peripheral spectrum[28] of $R(z)$. The following lemma is the first of three refinements that we will make.

Lemma 5.9.14. *For each z with $|z| = 1$,*

1. *the peripheral spectrum of $R(z)$ consists of semi-simple[29] eigenvalues.*

2. *given a peripheral eigenvalue ν of $R(z)$, the projection Π_ν onto the associated eigenspace is the uniform limit of the operators*

$$\frac{1}{n}\sum_{k=0}^{n-1} \nu^{-k} R(z)^k,$$

meaning that for all $\eta \in \mathfrak{B}$,

$$\lim_{n\to\infty} \left\| \frac{1}{n}\sum_{k=0}^{n-1} \nu^{-k} R(z)^k \eta - \Pi_\nu \eta \right\|_{\mathfrak{B}} = 0$$

[28] The peripheral spectrum of an operator A is the set of points $\nu \in \sigma(A)$ such that $|\nu| = \rho(A)$.

[29] An eigenvalue is semi-simple if its algebraic and geometric multiplicities match.

Proof. The argument is standard and can be found in a slightly different setting in [2] Proposition 3.5. We outline the argument here for the convenience of the reader.

First, by Lemma 5.9.13, for $|z| = 1$, $R(z)$ is quasi compact with spectral radius 1 and essential spectral radius β^a. By Definition 2.4.1 there is an $R(z)$-invariant splitting $\mathfrak{B} = F \oplus H$ such with $dim(F) < \infty$ such that $R(z)$ can be written as the sum of the operators $A = R(z)|_F$ and $B = R(z)|_H$, $\rho(B) < \beta^a$, $rank(A) < \infty$, and $AB = BA = 0$. Since $(A+B)(A^{k-1}+B^{k-1}) = A^k+B^k+ABB^{k-2}+BAA^{k-2} = A^k+B^k$, we see that for all $k \geq 0$,

$$R(z)^k = A^k + B^k.$$

Since $dim(F) < \infty$ and $A = R(z)|_F$ we see that A can be viewed as an operator on a finite vector space and as such has a Jordan decomposition

$$A = \sum_{\nu \in \sigma(A)} \nu \Pi_\nu + N_\nu,$$

where for each $\nu \in \sigma(A)$ the operator Π_ν is a projection on to the eigenspace for ν and N_ν is a nilpotent[30] matrix that commutes with Π_ν.

Second, if $\nu \in \sigma(A)$ and $|\nu| = 1$, then $N_\nu = 0$. To verify this we suppose that u is a vector in the generalized eigenspace for ν such that $N_\nu u = v \neq 0$ and $N_\nu v = 0$, and derive a contradiction. Since u is in the generalized eigenspace for ν we have $(A - \nu I)u = N_\nu u = v$ and $(A - \nu I)v = N_\nu v = 0$. A simple induction shows that

$$A^k u = k\nu^{k-1} v + \nu^k u$$

It follows from Lemma 5.7.12 that $\|Au\|_{\mathfrak{B}} \leq \|u\|_{\mathfrak{B}}$. Combining the bound with the displayed equation above we obtain

$$j \|v\|_{\mathfrak{B}} = \left\|A^k u - \nu^k u\right\|_{\mathfrak{B}} \leq 2 \|u\|_{\mathfrak{B}},$$

for all $k \geq 0$. The inequalities above can only hold for all $k \geq 0$ if $\|v\|_{\mathfrak{B}} = 0$, but this contradicts the assumption that $v \neq 0$. We have proved that $N_\nu = 0$, therefore the eigenspace and generalized eigenspace for ν agree and the eigenvalue ν is semi-simple. We conclude that for $\nu \in \sigma(A)$ with $|\nu| = 1$, the associated Jordan block is $\nu \Pi_\nu$. We

[30] A matrix M is nilpotent if there exists $j > 0$ such that $M^j = 0$

will show that for any $\eta \in \mathfrak{B}$ and $\nu \in \sigma(A)$ with $|\nu| = 1$,

$$\lim_{n\to\infty} \left\| \Pi_1 \eta - \frac{1}{n+1} \sum_{k=0}^{n} \nu^{-k} R^k(z)\eta \right\|_{\mathfrak{B}} = 0. \tag{5.9.9}$$

To verify the equation fix $\eta \in \mathfrak{B}$ and $\nu \in \sigma(A)$ with $|\nu| = 1$, and compute as follows,

$$\begin{aligned}
\frac{1}{n+1} \sum_{k=0}^{n} \nu^{-k} R(z)^k \eta &= \frac{1}{n+1} \sum_{k=0}^{n} \left[\nu^{-k} B^k \eta + \sum_{\omega\in\sigma(A)} \left(\frac{\omega}{\nu}\right)^k \Pi_\omega \eta + \nu^{-k} N_\omega^k \eta \right] \\
&= \frac{1}{n+1} \sum_{k=0}^{n} \left[\nu^{-k} B^k \eta + \sum_{\substack{\omega\in\sigma(A) \\ |\omega|<1}} \left(\frac{\omega}{\nu}\right)^k \Pi_\omega \eta + \nu^{-k} N_\omega^k \eta \right] \\
&\quad + \sum_{\substack{\omega\in\sigma(A)\setminus\{\nu\} \\ |\omega|=1}} \left[\frac{1}{n+1} \sum_{k=0}^{n} \left(\frac{\omega}{\nu}\right)^k \right] \Pi_\omega \eta \\
&\quad + \Pi_\nu \eta
\end{aligned} \tag{5.9.10}$$

Since $E = \{\nu \in \sigma(A) : |\nu| < 1\}$ is a finite set we may select $r \in (\beta^a, 1)$ such that E is contained in the complex disk of radius r, as is $\sigma(B)$. It follows by the spectral radius formula and the root test for convergence that

$$\sum_{k=0}^{n} \left[\left\|\nu^{-k} B^k \eta\right\|_{\mathfrak{B}} + \sum_{\substack{\omega\in\sigma(A) \\ |\omega|<1}} \left\| \left(\frac{\omega}{\nu}\right)^k \Pi_\omega + \nu^{-k} N_\omega^k \eta \right\|_{\mathfrak{B}} \right]$$

converges to a finite number so that

$$\lim_{n\to\infty} \frac{1}{n+1} \sum_{k=0}^{n} \left[\left\|\nu^{-k} B^k \eta\right\|_{\mathfrak{B}} + \sum_{\substack{\omega\in\sigma(A) \\ |\omega|<1}} \left\| \left(\frac{\omega}{\nu}\right)^k \Pi_\nu + \nu^{-k} N_\nu^k \eta \right\|_{\mathfrak{B}} \right] = 0.$$

For a complex number $\omega \neq \nu$ with $|\omega| = 1$ consider the quantity

$$\frac{1}{n+1} \sum_{k=0}^{n} \left(\frac{\omega}{\nu}\right)^k .$$

Select $\theta \in [0,1)$ such that $\frac{\omega}{\nu} = e^{i2\pi\theta}$ and let $S\colon : [0,1) \circlearrowleft$ by $S(x) = x + \theta \mod 1$ and $\psi(x) = e^{i2\pi x}$. We reinterpret the displayed sum above as

$$\frac{1}{n+1}\sum_{k=0}^{n}\psi(S^k(0)).$$

If θ is rational then every point in $[0,1)$ is periodic for S with the same period p, and we see that that

$$\sum_{k=0}^{p-1}\psi(S^k(0)) = \sum_{k=0}^{p-1} e^{i2\pi\theta k} = \frac{1-e^{i2\pi\theta p}}{1-e^{i2\pi\theta}} = \frac{\psi(0)-\psi(S^p(0))}{1-e^{i2\pi\theta}} = 0.$$

Therefore, the sequence

$$\sum_{k=0}^{n}\left(\frac{\omega}{\nu}\right)^k.$$

takes exactly p different values and thus

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\left(\frac{\omega}{\nu}\right)^k = 0.$$

If θ is irrational, then S is uniquely ergodic and measure preserving with respect to Lebesgue measure on $[0,1)$, therefore Birkhoff sums converge point-wise and we obtain

$$\lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\psi(S^k(x)) = \int_0^1 \psi(x)\,dx = \int_0^1 e^{i2\pi x}\,dx = 0,$$

for every x. In particular

$$\frac{1}{n+1}\sum_{k=0}^{n}\left(\frac{\omega}{\nu}\right)^k = \lim_{n\to\infty}\frac{1}{n+1}\sum_{k=0}^{n}\psi(S^k(0)) = 0.$$

We conclude that for $\omega \in \sigma(A)$ with $|\omega| = 1$

$$\lim_{n\to\infty}\left\|\left[\frac{1}{n+1}\sum_{k=0}^{n}\left(\frac{\omega}{\nu}\right)^k\right]\Pi_\omega\eta\right\|_{\mathfrak{B}} = \|\Pi_\omega\eta\|_{\mathfrak{B}}\lim_{n\to\infty}\left|\frac{1}{n+1}\sum_{k=0}^{n}\left(\frac{\omega}{\nu}\right)^k\right| = 0.$$

Finally, we subtract $\Pi_\nu\eta$ from both sides of eq. (5.9.10), take $\|\cdot\|_{\mathfrak{B}}$-norm of both sides of the result, and take the limit as $n\to\infty$ to verify eq. (5.9.9), which completes the proof. $\square$

We have just characterized points of the peripheral spectrum and their associated spectral projections. The next result characterises peripheral eigenvectors.

Lemma 5.9.15. *If* $|z| = 1$*,* $|\nu| = 1$*,* $\eta \in \mathfrak{B}$*, and*

$$R(z)\eta = \nu\eta,$$

then $\eta \in \mathfrak{L}_u$*.*

Proof. Let $V_\nu(z)$ denote the eigenspace of $R(z)$ associated to the eigenvalue ν and let Π_1 denote the eigenprojection of $\mathfrak{B}$ onto $V_\nu(z)$. Since $|\nu| = 1 = \rho(R(z))$, we have that ν is a peripheral eigenvalue and as we saw in the proof of Lemma 5.9.14 the quasi-compactness of $R(z)$ implies that $dim(V_\nu(z)) < \infty$. By definition the space $\mathfrak{L}_u$ is a dense linear subspace of $\mathfrak{B}$, thus the image of $\mathfrak{L}_u$ under $\Pi_\nu(z)$ is a dense linear subspace of $V_\nu(z)$. The only dense linear subspace of a finite dimensional vector space is the whole space. We conclude that $\Pi_\nu(z)$ maps $\mathfrak{L}_u$ onto $V_\nu(z)$, and therefore there exists $\eta_0 \in \mathfrak{L}_u$ such that

$$\Pi_\nu(z)\eta_1 = \eta.$$

By Lemma 5.9.11 we have $\left\|R(z)^k\right\|_{\mathfrak{L}_u} \leq \|\eta\|_{\mathfrak{L}_u}$ and thus letting

$$\eta_n = \frac{1}{n}\sum_{k=0}^{n-1} R(z)^k \eta_0$$

for $k \geq 1$ defines a Cauchy sequence in $\left(\mathfrak{L}_u, \|\cdot\|_{\mathfrak{L}_u}\right)$. Let $\eta^* \in \mathfrak{L}_u$ be defined by $\eta^* = \lim_{n\to\infty} \eta_n$.

Next we wish to show that η and η^* are equal as elements of $\mathfrak{B}$. Consider the following bound, which holds for all $n \geq 1$,

$$\|\eta - \eta^*\|_{\mathfrak{B}} \leq \|\eta - \eta_n\|_{\mathfrak{B}} + \|\eta_n - \eta^*\|_{\mathfrak{B}} \leq \|\eta - \eta_n\|_{\mathfrak{B}} + \|\eta_n - \eta^*\|_{\mathfrak{L}_u}.$$

By taking n sufficiently large, the quantity $\|\eta - \eta_n\|_{\mathfrak{B}}$ can be made arbitrarily small by Lemma 5.9.14 and the quantity $\|\eta_n - \eta^*\|_{\mathfrak{L}_u}$ can be made arbitrarily small since the η_n's converge to η^* in $\mathfrak{L}_u$. We conclude that $\|\eta - \eta^*\|_{\mathfrak{B}} = 0$ and hence that η and η^* are equal as elements in $\mathfrak{B}$, therefore $\eta \in \mathfrak{L}_u$. □

5.9.5 Spectral Gap and Aperiodicity

In the next lemma we further refine Lemmas 5.9.14 and 5.9.15 and verify both the spectral gap and aperiodicity conditions for Theorem 5.9.7.

Lemma 5.9.16. *For $z \neq 1$ with $|z| \leq 1$, $I - R(z)$ is invertible. 1 is a simple eigenvalue of $R(1)$ and the associated eigenspace is span $\{\mathbf{1}_\Lambda\}$.*

Proof. The proof of this lemma will be divided into several distinct parts.

Claim 1: For all $|z| \leq 1$ the operator $R(z) - I$ is invertible if and only if 1 is not and eigenvalue of $R(z)$.

Proof of Claim 1. If 1 is an eigenvalue of $R(z)$, then $R(z) - I$ is not invertible by the definition of an eigenvalue. Suppose that $R(z) - I$ is not invertible, then 1 is a point in the spectrum of $R(z)$. By Lemma 5.9.13 the operator $R(z)$ is quasi-compact with essential spectral radius less than $\beta^a |z|$, which is strictly less then 1, therefore 1 is a point in the spectrum of $R(z)$ that is outside the essential spectrum. It follows from Definition 2.4.1 that 1 is an eigenvalue of $R(z)$ and that the eigenvector associated to the eigenvalue 1 lies in a finite dimensional $R(z)$ invariant subspace of $\mathfrak{B}$.

■

Claim 2: If $|z| < 1$, then $R(z) - I$ is invertible.

Proof of Claim 2. Fix z such that $|z| < 1$. It follows from Lemma 5.9.13 that the spectral radius of $R(z)$ is at most $|z|$. By assumption $|z| < 1$, so 1 is not an eigenvalue of $R(z)$. By the previous claim $R(z) - I$ is invertible.

■

Claim 3: If $|z| = 1$ and $z \neq 1$, then $I - R(z)$ is invertible. The operator $R(1)$ has a simple eigenvalue at 1 and the associated eigenspace is *span* $\{\mathbf{1}_\Lambda\}$.

Proof of Claim 3. We will verify both parts of the claim simultaneously. Let z be a complex number such that $|z| = 1$ and let $\eta \in \mathfrak{B}$ be an eigenvector of $R(z)$ with eigenvalue 1, that is

$$R(z)\eta = \eta.$$

The proof relies on to observations about η:

Observation 1 η satisfies the following identity

$$[\eta \circ T](x, y)\, z^r = \eta(x, y). \tag{5.9.11}$$

Observation 2 η is a constant multiple of $\mathbf{1}_\Lambda$.

We will verify both observations after completing the proof of Claim 3.

We will show that, if $\eta \neq 0$, then $z = 1$. By Observation 2 η is constant, since T preserves Lebesgue measure $\eta \circ T = \eta$. It follows that eq. (5.9.11) reduces to

$$(z^{r(x)} - 1)\eta = 0.$$

The equation above is satisfied if $\eta = 0$ or if $z^{r(x)} = 1$.

The equation $z^{r(x)} = 1$ is satisfied if and only if for all $a \in \text{range}(r) \subseteq \mathbb{Z}$,

$$a \frac{\arg(z)}{2\pi} \in \mathbb{Z}.$$

The inclusion above can only hold if and only if there exists a rational number b/c such that $\frac{\arg(z)}{2\pi} = b/c$. Assuming that b/c is reduced we see that $ab/c \in \mathbb{Z}$ and if and only if c divides a. Therefore, $\frac{\arg(z)}{2\pi} = b/c$ and c divides a for all $a \in \text{range}(r)$. From Definition 5.3.6 it follows that $\text{range}(r) = \{n \in \mathbb{N} : n \geq 2\}$ and hence the greatest common divisor of range(r) is 1 so that $c = 1$ and hence $\frac{\arg(z)}{2\pi} \in \mathbb{Z}$. Therefore the principal value of the argument of z is 0 and hence $z = 1$.

Since T preserves Lebesgue measure on Λ and $R(1)$ is the Frobenius-Perron operator of T we have $R(1)\mathbf{1}_\Lambda = \mathbf{1}_\Lambda$. By Observation 2 any η that satisfies the eigenvector equation $R(1)\eta = \eta$ is a multiple of $\mathbf{1}_\Lambda$. We have verified that $\mathbf{1}_\Lambda$ is a basis for the eigenspace associated to the eigenvalue 1. By Lemma 5.9.14 the eigenvalue 1 is semi-simple. We conclude that 1 is a simple eigenvalue of $R(1)$.

We have observed that if $R(z)\eta = \eta$, then $\eta \neq 0$ implies that $z = 1$. By contraposition, If $R(z)\eta = \eta$ and $z \neq 1$, then $\eta = 0$. We conclude that for $z \neq 1$, the operator $R(z)$ does not have 1 as an eigenvalue. By our previous claim we conclude that $I - R(z)$ is invertible. ■

To complete the proof of the lemma it remains to verify Observation 1 and Observation 2 from the proof of the last claim.

Observation 1: If $|z| = 1$ and $\eta \in \mathfrak{B}$ such that $R(z)\eta = \eta$, then for almost every $(x, y) \in \Lambda$,

$$[\eta \circ T](x, y)\, z^r = \eta(x, y).$$

Proof of Observation 1. By Lemma 5.9.15 we have $\eta \in \mathfrak{L}_u$. Since $\|\eta\|_\infty \leq \|\eta\|_{\sup} \leq \|\eta\|_{\mathfrak{L}_u}$ we have $\eta \in L^\infty(\Lambda, \mu)$. For all ψ and η in $\mathfrak{L}_u$ we have

$$\begin{aligned}\int R(z)\eta\, \psi\, d\mu &= \int \sum_{n=1}^{\infty} z^n R_n \eta\, \psi\, d\mu = \sum_{n=1}^{\infty} \int z^n \mathcal{P}(\eta\, \mathbf{1}_{r=n})\, \psi\, d\mu \\ &= \sum_{n=1}^{\infty} \int \eta\, z^n \mathbf{1}_{r=n}\, \psi \circ T\, d\mu = \int \sum_{n=1}^{\infty} \eta\, z^n \mathbf{1}_{r=n}\, \psi \circ T\, d\mu \\ &= \int \eta\, z^r \psi \circ T\, d\mu,\end{aligned}$$

where we have applied the dominated convergence theorem to partial sums and eq. (2.2.3). Since $\eta \in L^\infty(\mu)$ we have $\eta \in L^2(\mu)$. Define $W(z)$ on $L^\infty(\mu)$ by $W(z)\psi = z^r\, \psi \circ T$. Now we compute as in [13]

$$\begin{aligned}\|W(z)\eta - \eta\|_2^2 &= \|W(z)\eta\|_2^2 - 2Re\langle W(z)\eta, \eta\rangle + \|\eta\|_2^2 \\ &= \|W(z)\eta\|_2^2 - 2Re\langle \eta, R(z)\eta\rangle + \|\eta\|_2^2 \\ &= \|W(z)\eta\|_2^2 - 2Re\langle \eta, \eta\rangle + \|\eta\|_2^2 \\ &= \|W(z)\eta\|_2^2 - \|\eta\|_2^2\end{aligned}$$

and note that

$$\|W(z)\eta\|_2^2 = \int |\eta|^2 \circ T\, d\mu = \int |\eta|^2\, d\mu = \|\eta\|_2^2,$$

from which we conclude that $W(z)\eta = [\eta \circ T]\, z^r = \eta$ except possibly on a μ null set.

We have verified eq. (5.9.11).

■

Observation 2: If $|z| = 1$ and $\eta \in \mathfrak{B}$ so that $\mathbb{R}(z)\eta = \eta$, then η is a constant multiple of $\mathbf{1}_\Lambda$.

Proof of Observation 2. We begin by showing that η is essentially constant along stable fibres. For each $j \geq 1$ select $\tau_j \in C^\infty$ such that $\|\tau_j - \eta\|_1 < 2^{-j}$. Note that $\|W(\tau_j - \eta)\|_1 = \|z^r(\tau_j - \eta) \circ T\|_1 = \|\tau_j - \eta\|_1 < 2^{-j}$. Let $\bar{\tau}_j(x, y) = \int \tau_j(x, y)\, dy$ and note that by the mean value theorem there exists $s \in (0, 1)$ and $t \in (y, s)$ such that

$$|\tau_j(x, y) - \bar{\tau}_j(x, y)| = |\tau_j(x, y) - \tau_j(x, s)| = |\partial_y \tau_j(x, t)||y - s| \leq \|\partial_y \tau_j\|_\infty |y - s|.$$

Further application of the mean value theorem yields

$$|W^n \tau_j(x, y) - W^n \bar{\tau}_j(x, y)| \leq \|\partial_y \tau_j\|_\infty \left\|\partial_y v_x^{(n)}\right\|_\infty \leq \|\partial_y \tau_j\|_\infty \beta^n.$$

For each $j \geq 1$ select $n = n(j)$ such that $\|\partial_y \tau_j\|_\infty \beta^n + 2^{-j} < 10 \cdot 2^{-j}$ and note that

$$\|\eta - \bar{\tau}_j\|_1 \leq \|W^n \eta - W^n \tau_j\|_1 + \|W^n \tau_j - W^n \bar{\tau}_j\|_1 \leq 10 \cdot 2^{-j}.$$

We see that η is the L^1-limit of functions that are constant along stable fibres. It follows that for $\bar{\mu}$-a.e. $x \in [p, q]$,

$$\text{for } \lambda\text{-a.e. } y,\ \eta(x, y) = \int_{\ell_x} \eta(x, z)\, d\lambda(z), \tag{5.9.12}$$

Next we will use the unstable regularity of η to show that Property 5.9.12 holds for every $x \in [p, q]$. To verify this suppose that x failed to satisfy Property 5.9.12. This can happen if and only if there exist sets $A_x, B_x \subset \ell_x$ and $\epsilon > 0$, such that $\lambda(A_x) > 0$, $\lambda(B_x) > 0$, and for all y in A_x and z in B_x

$$\eta(x, y) - \eta(x, z) \geq \epsilon. \tag{5.9.13}$$

For $w \neq x$ let $A_w \subset \ell_w$ be the set obtained by sliding[31] A_x along unstable disks into ℓ_w and let B_w be defined similarly. Note that $\lambda(A_w) > 0$ if and only if $\lambda(A_x) > 0$. Since η is in $\mathfrak{L}_u$ we have that

$$|\eta(x, y) - \eta(\ell_w \cap \pi_u(x, y))| \leq L_u(\eta)\, |x - w|\,.$$

[31] By sliding along unstable disks we mean $(x, y) \mapsto \pi_u(x, y) \cap \ell_w$

Choose $\delta > 0$ so that $L_u(\eta)\,\delta < \epsilon/3$. Fix $w \in [p,q]$ such that $|w-x| < \delta$. Select $(w,y) \in A_w$ and $(w,z) \in B_w$ and let $(x,y') \in A_x$ and $(x,z') \in B_x$ denote the points obtained by sliding along unstable disks back to ℓ_x. We compute,

$$\eta(w,y) - \eta(w,z) \geq \eta(x,y') - \eta(x,z') - 2L_u(\eta)\,|x-w| \geq \epsilon - 2L_u(\eta)\,\delta \geq \frac{\epsilon}{3}.$$

We have just shown that for every $w \in [p,q]$ with $|w-x| < \delta$ Property 5.9.13 holds at w, thus Property 5.9.12 fails at w. This contradicts our observation that eq. (5.9.12) holds for $\bar{\mu}$-a.e. $x \in [p,q]$. We conclude that eq. (5.9.12) holds for every $x \in [p,q]$.

Define $h(x) = \int_0^1 \eta(x,y)\,dy$. This function is Lipschitz. To verify this fix $x, w \in [p,q]$. Let $A_x \subset \ell_x$ denote the set of points in ℓ_x where eq. (5.9.12) fails and let A_w be defined similarly. By the previous paragraph both A_x and A_w are null sets. Let $B \subset \ell_x$ be the set obtained by sliding A_w along unstable disks into ℓ_x. The set B is a null set, therefore the set $G = \ell_x - (A_x \cup B)$ consisting of points in ℓ_x where $\eta(x,y) = h(x)$ and $\eta(\pi_u(x,y) \cap \ell_w) = h(w)$ has full measure. Choose $(x,y) \in G$ and note that

$$|h(x) - h(w)| = |\eta(x,y) - \eta\left(\pi_u(x,y) \cap \ell_x\right)| \leq L_u(\eta)\,|x-w|\,,$$

so h is Lipschitz with Lipschitz constant at most $L_u(\eta)$.

Next we would like to verify $[W(z)\eta](x,y) = z^r\,[h \circ u](x)$. Note that T maps ℓ_x into $\ell_{u(x)}$ affinely. We will apply the change of variable $y' = g_x(y)$ noting that $dy' = \partial_y g_x(y) dy$ and that $\partial_y g_x(y)$ is constant and exactly equal to the length of the interval $T\ell_x \subset \ell_{u(x)}$

$$\int_0^1 z^{r(x)}(\eta \circ T)(x,y)\,dy = z^{r(x)}\frac{1}{|T\ell_x|}\int_{T\ell_x} \eta(u(x),y')\,dy' = z^{r(x)}h\left(u(x)\right)$$

Applying eq. (5.9.11) we obtain

$$z^r\,[h \circ u](x) = h(x) \tag{5.9.14}$$

Next we deduce that h is an essentially constant function. We will apply Corollary 3.2 from [1]. We reformulate the Corollary in our notation for the convenience of the reader:

Suppose that:

- $u\colon [p,q] \circlearrowleft$ *is a probability preserving, almost onto Gibbs-Markov map with respect to the partition* $\alpha = \left\{I_j, I'_j : j = 2, \cdots, \infty\right\}$[32].
- $\varphi\colon [p,q] \to \{z \in \mathbb{C} : |z| = 1\}$ *is* α*-measurable.*
- $h\colon [p,q] \to \{z \in \mathbb{C} : |z| = 1\}$ *is Boreal measurable and* $\varphi(x) = h \cdot \bar{h} \circ u$

Then h *is essentially constant.*

Let us verify that u satisfies the first hypothesis of the Corollary. For each $a \in \alpha$ the map $u|_a$ is a homeomorphism onto $[p,q)$ with C^2 inverse $v_a\colon [p,q] \to a$. The map u is uniformly expanding by Lemma 5.3.11 and satisfies Adler's bounded distortion property by Lemma 5.3.12. By Example 2 of [1] it follows that u is a mixing Gibbs-Markov map. Since every branch of u is onto, u is almost onto as defined immediately after Theorem 3.1 of [1].

Since u is a Gibbs-Markov map, u is ergodic. Taking the complex modulus of eq. (5.9.14) yields $|h| = |h \circ u| = |h| \circ u$, thus $|h|$ is an essentially constant function. Since h is Lipschitz, we have that $|h|$ is Lipschitz and therefore point-wise constant. Without loss of generality assume that $|h| = 1$.

Since h is a circle valued function we have $\bar{h} = 1/h$. Let $\varphi(x) = h \cdot \bar{h} \circ T$. By eq. (5.9.14) we have

$$\varphi(x) = h \cdot \bar{h} \circ T = \frac{h}{h \circ T} = z^{r(x)}.$$

Since $r(x)$ is measurable with respect to the partition α we have that φ is circle valued and α-measurable. We have just verified that φ satisfies the second hypothesis above and that h and φ are related as required in the third hypothesis by definition.

Applying the Corollary we see that h is essentially constant. Since h is Lipschitz we conclude that h is point-wise constant. Let h_0 denote the constant value of h.

---

[32] see Definition 5.3.6

Define $H(x, y) = h_0$, this function is clearly in $\mathfrak{L}_u$. On each vertical line the function H agrees with η except possibly on a set of one dimensional Lebesgue measure zero. It follows that for all $t \in [p, q]$ there exists a μ-null set N_t such that for all $(x, y) \in \Lambda - N_t$ we have $\delta_t(\eta - H)(x, y) = 0$. With this fact it follows directly from Definition 5.7.8 that $\|\eta - H\|_s = 0$ and $\mathbb{L}(\eta - H) = 0$, thus $\|\eta - H\|_{\mathfrak{B}} = 0$. We conclude that η and H are in the same $\mathfrak{B}$-equivalence class. ■

Having verified Observation 1 and Observation 2 from the proof of Claim 3 we see that the lemma follows by combining Claim 2 and Claim 3. □

5.9.6 Rate of Decay of Correlations for B

We have verified the hypotheses of Theorem 5.9.7 for the operators of Definition 5.9.1 acting on the space $\mathfrak{B}$ from Definition 5.7.11. We have also identified the eigenspace associated to the eigenvalue 1 for the operator $R(1)$ to be the one dimensional space spanned by the Borel measure ν with density $\frac{d\nu}{d\mu} = \mathbf{1}_\Lambda$ that is $\nu = \mu$. Since $R(1) = \mathcal{P}$ we see that μ is the unique T-invariant functional in $\mathfrak{B}$.

Suppose that $\eta \in \mathfrak{L}_u$, then $\Pi_1 \eta = \int_\Lambda \eta \, d\mu$. Applying Definition 5.9.1 we see that

$$\begin{aligned}\Pi_1 R_l \Pi_1 \eta &= \Pi_1 R_l \int_\Lambda \eta \, d\mu = \Pi_1 \mathcal{P}\left(\mathbf{1}_{\{r=l\}}\right) \int_\Lambda \eta \, d\mu \\ &= \int_\Lambda \eta \, d\mu \int_\Lambda \mathcal{P}\left(\mathbf{1}_{r=l}\right) d\mu = \mu\{r = l\} \int_\Lambda \eta \, d\mu.\end{aligned}$$

Recall that $P_n = \sum_{l>n} \Pi_1 R_l \Pi_1$, we compute

$$P_n \eta = \sum_{l>n} \mu\{r = l\} \int_\Lambda \eta \, d\mu = \mu\{r > n\} \int_\Lambda \eta \, d\mu.$$

It follows that

$$\sum_{k=n+1}^{\infty} P_k \eta = \sum_{k=n+1}^{\infty} \mu\{r > k\} \int_\Lambda \eta \, d\mu.$$

Similarly,

$$\begin{aligned}\Pi_1 R'(1)\Pi_1\eta = \Pi_1 \sum_{k=1}^{\infty} kR_k\Pi_1\eta &= \sum_{k=1}^{\infty} k\mu\{r=k\}\int_\Lambda \eta\, d\mu \\ &= \frac{1}{\lambda(\Lambda)}\sum_{k=1}^{\infty} k\lambda\{r=k\}\int_\Lambda \eta\, d\mu \\ &= \frac{1}{\lambda(\Lambda)}\int_\Lambda \eta\, d\mu,\end{aligned}$$

where in the second to last line we have applied the definition $\mu(E) = \frac{\lambda(E)}{\lambda(\Lambda)}$ and in the last line we have applied Kac's lemma to conclude that $\sum_{k=1}^{\infty} k\lambda\{r=k\} = 1$. From the displayed equation above we obtain $m = \frac{1}{\lambda(\Lambda)}$.

As was observed in [5] Lemma 1 the asymptotics in Lemma 5.3.10 can be used to obtain

$$\lambda\{r=k\} = O\left(k^{-2-1/\alpha}\right)$$

where $\alpha = \max\{\alpha_0, \alpha_1\}$. It follows from *Proposition* 5.9.10 and the displayed equation above that for all $k \geq 1$,

$$\|R_k\|_{op} = O\left(1/k^{2+1/\alpha}\right)$$

and therefore

$$\sum_{k>n} \|R_k\|_{op} = O\left(1/n^{1+1/\alpha}\right).$$

From Theorem 5.9.7 we see that

$$\|E_n\|_{op} = \begin{cases} O\left(1/n^{1+1/\alpha}\right), & \text{if } \alpha > 1; \\ O\left(\frac{\log(n)}{n^2}\right), & \text{if } \alpha = 1; \\ O\left(1/n^{2/\alpha}\right), & \text{if } \alpha < 1. \end{cases}$$

We also note that

$$\sum_{k>n} \lambda\{r>k\} = O\left(1/n^{1/\alpha}\right).$$

If $\eta \in \mathfrak{L}_u$ and $\psi \in \mathfrak{C}_a$, then we can view both η and ψ as functions on $[0,1]^2$ that are supported on Λ. From this prospective we see that $\eta = \eta \mathbf{1}_\Lambda$ and $\psi = \psi \mathbf{1}_\Lambda$ so that

$$\int_{[0,1]^2} \mathcal{P}_B^n \eta \, \psi \, d\lambda = \int_{[0,1]^2} \mathcal{P}_B^n (\eta \mathbf{1}_\Lambda) \, \psi \mathbf{1}_\Lambda \, d\lambda = \int_\Lambda Q_n \eta \, \psi \, d\lambda$$

Applying Theorem 5.9.7

$$\begin{aligned}
\int_{[0,1]^2} \mathcal{P}_B^n \eta \, \psi \, d\lambda &= \int_{[0,1]^2} Q_n \eta \, \psi \, d\lambda \\
&= \int_\Lambda \left[\frac{1}{m} \Pi_1 + \frac{1}{m^2} \sum_{k=n+1}^{+\infty} P_k + E_n \right] \eta \, \psi \, d\mu \\
&= \int \eta \, d\lambda \int \psi \, d\lambda + \int \eta \, d\lambda \int \psi \, d\lambda \sum_{k=n+1} \lambda \{ r > k \} + \int_\Lambda E_n \eta \, \psi \, d\lambda \\
&= \int \eta \, d\lambda \int \psi \, d\lambda + O \left(1/n^{1/\alpha} \right),
\end{aligned}$$

and therefore,

$$\left| \int_{[0,1]^2} \eta \, \psi \circ B^n \, d\lambda - \int \eta \, d\lambda \int \psi \, d\lambda \right| = O \left(\left(\frac{1}{n} \right)^{1/\alpha} \right).$$

We conclude with the following theorem which captures the rates of decay of correlations for the IBT B.

Theorem 5.9.17. *If $\alpha_0, \alpha_1 > 0$ are exponents associated to an intermittent cut function ϕ as in Definition 5.2.1 and $B \colon [0,1]^2 \circlearrowleft$ is the Intermittent Baker's Transformation associated to ϕ, then for all $\eta \in \mathfrak{L}_u$ and $\psi \in \mathfrak{C}_a$ we have,*

$$\left| \int_{[0,1]^2} \eta \, \psi \circ B^n \, d\lambda - \int \eta \, d\lambda \int \psi \, d\lambda \right| = O \left(\left(\frac{1}{n} \right)^{1/\alpha} \right),$$

where $\alpha = \max \{ \alpha_0, \alpha_1 \}$ and the functions η and ψ are viewed as functions on $[0,1]^2$ that are supported on Λ.

Chapter 6

Conclusion

This dissertation contains two main results.

The first is Theorem 3.2.3 where we obtain a familiar bound on the essential spectral radius of an finite branched expanding interval map. This result is novel because of the proof of the Lasota-Yorke inequality, which to our knowledge is new. The proof is interesting because it very cleanly relates the expansion rate of the map and the complexity of the branch partition to the essential spectral radius of the associated Frobenius-Perron operator acting on functions of bounded variation. This approach may be applicable to multidimensional expanding maps where it would continue to provide a clear connection between dynamical quantities (rate of expansion and complexity of the branch partition) and the essential spectral radius of the associated Frobenius-Perron operator acting on functions of bounded variation.

The second result is Theorem 5.9.17 where we compute a sharp polynomial rate of decay of correlations for the Frobenius-Perron operator associated to an intermittent baker's map acting on an anisotropic Banach space containing functions on $[0,1]^2$. Intermittent baker's maps were introduced in [5] and are summarised in section 5.1 This result extends results in [5] and obtains decay rates via a fundamentally different analysis. Spectral properties of the Frobenius-Perron operator associated to the Baker's map are studied directly via operator renewal theory in section 5.9. The renewal arguments in section 5.9 rely on a careful analysis of the Frobenius-Perron operator associated to an induced map. The inducing scheme was introduced in [5] and is summarised in section 5.3. To analyze the Frobenius-Perron operator associated to the induced map we introduce anisotropic Banach spaces of observables in section 5.5 and densities in section 5.7. Properties of these spaces rely critically on a careful analysis of the unstable manifolds of the induced map which is carried out in section 5.4 and a conditional expectation operator related to the unstable manifolds which is analyzed in section 5.6. Once the spaces are defined a Lasota-Yorke inequality for the Frobenius-Perron operator associated to the induced map is proved in Proposition 5.7.13. A compact embedding for the spaces of densities is proved in section 5.8. The Lasota-Yorke inequality, compact embedding result, and a result of Hennion (see [15] and Theorem 2.4.4) provide the basic tools required to analyze the spectrum of the Frobenius-Perron operator of the induced map. In section section 5.9 a result of Gouëzel (see [13] and Theorem 5.9.7) is applied to analyze the Frobenius-Perron operator of the original Bakers map using properties of the induced map.

The second result is interesting because it provides decay rates for the full baker's map without passing to the expanding factor as was done in [5]. This makes statistical properties such as dynamical central limit theorem, Berry-Essen theorem, and convergence to stable laws more directly accessible. The analysis also provides a very explicit example of how one can apply both anisotropic Banach space methods and renewal methods to analyze a non-uniformly hyperbolic map with indifferent fixed points.

Appendix A

Additional Information

A.1 Functions of Bounded Variation Revisited

In this section we will introduce two closely related definitions of variation, which are alternatives to Definition 2.3.2.

Definition A.1.1. Given a function f in $L^1(\mathbb{R}, \lambda)$, define $\mathrm{var}_s \colon L^1(\mathbb{R}, \lambda) \to [0, \infty]$ by

$$\mathrm{var}_s(f) = \sup\left\{ \int_{\mathbb{R}} f\, D\phi\, d\lambda : \phi \in C_c^1, |\phi| \le 1 \right\}.$$

Definition A.1.2. Let Lip_c denote the set of all compactly supported absolutely continuous functions $\psi \colon \mathbb{R} \to \mathbb{R}$ such that $D\psi \in L^\infty(\mathbb{R}, \lambda)$.

Definition A.1.3. Define $\mathrm{var}_{ac} \colon L^1(\mathbb{R}, \lambda) \to [0, \infty]$ by

$$\mathrm{var}_{ac}(f) = \sup\left\{ \int_{\mathbb{R}} f\, D\psi\, d\lambda : \psi \in Lip_c, |g| \le 1 \right\}.$$

Our goal for the remainder of the section is to prove the following proposition.

Proposition A.1.4. *For all integrable* $f \colon \mathbb{R} \to \mathbb{R}$, $\mathrm{var}(f) = \mathrm{var}_{ac}(f) = \mathrm{var}_s(f)$.

Proof. This follows from Corollary A.1.22.1 and Lemmas A.1.23 and A.1.24. □

Before addressing this goal, we will review some standard facts relating increasing functions, functions of bounded variation, and Borel measures.

A.1.1 Background

Given a set X let $\wp^X$ denote the power set of X. We recall the definition of an algebra of sets.

Definition A.1.5. Given a set X, a collection of subsets $\mathcal{A} \subset \wp^X$ is an *algebra* if it is closed under complements and finite unions.

Note that every non-empty algebra on X contains both X (since $X = A \cup A^c$ for any $A \in \mathcal{A}$) and $\emptyset$ (since $\emptyset = X^c$). The following example will be important in proving that $\mathrm{var} = \mathrm{var}_{ac}$.

Example A.1.6. Given $S \subseteq \mathbb{R}$, let $I(S)$ denote the collection of all left open right closed intervals and rays in $\mathbb{R}$ with endpoints in S, that is

$$I(S) := \{(a, b] : a, b \in S\} \cup \{(a, \infty) : a \in S\} \cup \{(-\infty, b] : b \in S\} \cup \{\mathbb{R}, \emptyset\}.$$

Let $\mathcal{A}(S)$ denote the set of all finite unions of intervals in $I(S)$. Note that any finite union can be written as a finite disjoint union. The collection $\mathcal{A}(S)$ is an algebra of subsets of $\mathbb{R}$.

The following lemma connects these algebra to the Borel σ-algebra on $\mathbb{R}$.

Lemma A.1.7. *If S is a dense subset of $\mathbb{R}$ then the σ-algebra generated by $\mathcal{A}(S)$ contains the Borel σ-algebra on $\mathbb{R}$.*

Proof. Fix $(a, b) \subseteq \mathbb{R}$. By selecting a countable increasing sequence of intervals $(s_n, t_n] \subset (a, b)$ with end points in S, such that $\cup_{n=1}^{\infty}(s_n, t_n] = (a, b)$ we see that $(a, b) \in \sigma(\mathcal{A}(S))$. Since (a, b) was arbitrary we see that $\sigma(\mathcal{A}(S))$ contains all open intervals. By a standard result (see Proposition 1.2 of [10]) any σ-algebra containing the open intervals contains the Borel σ-algebra on $\mathbb{R}$. □

Next we recall the definition of a premeasure.

Definition A.1.8. Given an algebra $\mathcal{A} \subseteq \wp^X$, a function $\rho \colon \mathcal{A} \to [0, \infty]$ is a *premeasure* if $\rho(\emptyset) = 0$ and for any countable disjoint collection of sets $\{A_i : i \in \mathbb{N}\}$ in $\mathcal{A}$ such that $\cup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$,

$$\rho\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \rho(A_i).$$

Premeasures have a few notable properties which follow from the definitions.

Lemma A.1.9. *If $\rho\colon \mathcal{A} \to [0, \infty]$ is a premeasure, then ρ has the following properties:*

1. *Monotonicity: if $A \subseteq B$ are in $\mathcal{A}$, then $\rho(A) \leq \rho(B)$.*

2. *Subadditivity: if $\{A_i : i \in \mathbb{N}\}$ are in $\mathcal{A}$ and $\cup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$, then*
$$\rho\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} \rho(A_i).$$

3. *Continuity from below: If $\{A_i : i \in \mathbb{N}\}$ are in $\mathcal{A}$, $\cup_{i=1}^{\infty} A_i$ is in $\mathcal{A}$, and $A_i \subseteq A_{i+1}$ for all i, then*
$$\rho\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \lim_{i\to\infty} \rho(A_i).$$

4. *Continuity from above: If $\{A_i : i \in \mathbb{N}\}$ are in $\mathcal{A}$, $\cap_{i=1}^{\infty} A_i$ is in $\mathcal{A}$, and $A_i \supseteq A_{i+1}$ for all i, then*
$$\rho\left(\bigcap_{i=1}^{\infty} A_i\right) \leq \lim_{i\to\infty} \rho(A_i).$$

Proof. The proof is the same as the standard result for measures (see theorem 1.8 [10]). □

The next example shows how to produce a plethora of premeasures on algebras like those in Example A.1.6.

Example A.1.10. Given an increasing function $f\colon \mathbb{R} \to \mathbb{R}$, let S_f denote the set of continuity points for f and let $\mathcal{A}_f := \mathcal{A}(S_f)$. Let $\rho_f\colon \mathcal{A}_f \to [0, \infty]$ be the function defined on $I(S_f)$ by

$$\begin{aligned}
\rho_f\,(a, b] &= f(b) - f(a) \\
\rho_f\,(-\infty, b] &= f(b) - \lim_{t\to-\infty} f(t) \\
\rho_f\,(a, \infty) &= \lim_{t\to\infty} f(t) - f(a) \\
\rho_f \mathbb{R} &= \lim_{t\to\infty} f(t) - f(-t)
\end{aligned}$$

We then extend ρ_f to all of $\mathcal{A}_f$ by defining for any finite disjoint collection of intervals

$\{J_i : 1 \le i \le N\}$,

$$\rho\left(\bigcup_{i=1}^{N} J_i\right) = \sum_{i=1}^{N} \rho_f(J_i).$$

The function ρ_f is a premeasure on $\mathcal{A}_f$. This can be verified as in the proof of Proposition 1.15 of [10].

Notice that if we were to define ρ_f on all intervals $(a, b] \subset \mathbb{R}$ in the example above, then ρ_f could fail to be a premeasure. For example, assume that b is a continuity point for f but that $f(a) \ne \lim_{t\to a^+} f(t)$. Then, ρ_f will fail to be continuous from below since

$$\rho_f(a, b] = f(b) - f(a) \ne f(b) - \lim_{n\to\infty} f\left(a + \frac{1}{n}\right) = \lim_{n\to\infty} \rho_f\left(a + \frac{1}{n}, b\right].$$

The following theorem, which follows from the Carathéodory extension theorem, is our motivation for recalling the definitions of algebra and premeasure.

Theorem A.1.11. *Given an algebra $\mathcal{A} \subseteq \wp^X$ and a premeasure $\rho\colon \mathcal{A} \to [0, \infty]$, let $\sigma(\mathcal{A})$ denote the σ-algebra generated by $\mathcal{A}$. There exists a measure μ on $\sigma(\mathcal{A})$ such that $\mu|_{\mathcal{A}} = \rho$. If ρ is σ-finite, then μ is unique.*

It follows from this theorem that every increasing function $f\colon \mathbb{R} \to \mathbb{R}$ induces a measure on the σ-algebra generated by $\mathcal{A}_f$. In order to characterize this σ-algebra, we need to recall a few properties of increasing functions.

Lemma A.1.12. *If $f\colon \mathbb{R} \to \mathbb{R}$ is an increasing function, then f has the following properties:*

1. *Both of the limits $f(-\infty) := \lim_{t\to-\infty} f(t)$ and $f(\infty) := \lim_{t\to\infty} f(t)$ exist in $\mathbb{R} \cup \{\infty\}$.*
2. *f has, at most, countably many discontinuities.*
3. *At every point x in $\mathbb{R}$, the function f possess left and right limits.*
4. *f is differentiable at Lebesgue almost every point in $\mathbb{R}$.*

The Baire category theorem together with part 2 of the lemma above show that S_f is dense for any increasing f, and Lemma A.1.7 implies that $\sigma(\mathcal{A}_f)$ is the Borel σ-algebra on $\mathbb{R}$. We apply Theorem A.1.11 to obtain the following corollary.

Corollary A.1.12.1. *For each increasing function $f\colon \mathbb{R} \to \mathbb{R}$ there exists a unique Borel measure μ_f such that $\mu_f|_{\mathcal{A}_f} = \rho_f$.*

Notice that the assignment $f \mapsto \mu_f$ is not one-to-one.

Lemma A.1.13. *If $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ differ by a constant almost everywhere, then $\mu_f = \mu_g$.*

Proof. Since both f and g have at most countably many points of discontinuity, the set $S_f \cap S_g$ is dense by the Baire category theorem. In addition, for any $a, b \in S_f \cap S_g$, we have

$$\rho_g(a, b] = g(b) - g(a) = f(b) + c - f(a) - c = f(b) - f(a) = \rho_f(a, b].$$

Thus $(\rho_f - \rho_g)|_{\mathcal{A}(S_f \cap S_g)} = 0$ and thus the zero measure is a Borel extension of this premeasure. On the other hand, $(\mu_f - \mu_g)|_{\mathcal{A}(S_f \cap S_g)} = (\rho_f - \rho_g)|_{\mathcal{A}(S_f \cap S_g)}$ and hence $\mu_f - \mu_g$ is a Borel extensions of $(\rho_f - \rho_g)|_{\mathcal{A}(S_f \cap S_g)}$. By the uniqueness condition in Corollary A.1.12.1 we see that $\mu_f - \mu_g = 0$ and hence $\mu_f = \mu_g$. ☐

While there are many nondecreasing functions that can induce a particular Borel measure, the cumulative distribution function is distinguished as we will see in what follows.

Definition A.1.14. Given a finite Borel measure μ on $\mathbb{R}$ let the *cumulative distribution function* of μ be the function $P_\mu\colon \mathbb{R} \to [0, \infty)$ defined by $P_\mu(x) = \mu(-\infty, x]$.

It follows from the properties in Lemma A.1.9 that P_μ is increasing, right continuous, and $\lim_{t\to-\infty} P_\mu(t) = 0$. If $f\colon \mathbb{R} \to \mathbb{R}$ is a bounded increasing function, then the associated measure is bounded since $\mu_f(-\infty, \infty) = \rho_f(-\infty, \infty) = \lim_{t\to\infty} f(t) - f(-t) < \infty$. Therefore, the cumulative distribution function P_{μ_f} is well defined. For convenience we introduce the following notation.

Definition A.1.15. Given a bounded increasing function $f\colon \mathbb{R} \to \mathbb{R}$, let $\overline{f}\colon \mathbb{R} \to \mathbb{R}$ be defined by $\overline{f}(x) = P_{\mu_f}(x) + \lim_{t\to-\infty} f(t)$.

Lemma A.1.16. *If $f\colon \mathbb{R} \to \mathbb{R}$ is a bounded increasing function, then*

1. *$f(x) = \overline{f}(x)$ for all $x \in S_f$.*

2. *$\overline{f}(x) = \lim_{t\to x^+} f(x)$ for all $x \in \mathbb{R}$.*

3. *$\mu_f(a,b] = \overline{f}(b) - \overline{f}(a)$ for all $a, b \in \mathbb{R}$.*

We are now in a position to examine our definitions of variation. The following lemma allows us to apply all of our results about increasing functions to functions of bounded variation.

Lemma A.1.17. *Given $f\colon \mathbb{R} \to \mathbb{R}$, $V(f) < \infty$ if and only if f is the difference of two bounded increasing functions.*

Proof. This follows from Theorem 3.27 of [10]. □

Corollary A.1.17.1. *If $f\colon \mathbb{R} \to \mathbb{R}$ and $V(f) < \infty$, then f has the following properties:*

1. *Both of the limits $f(-\infty)$ and $f(\infty)$ exist and are finite.*
2. *At every point x in $\mathbb{R}$, the function f possess left and right limits.*
3. *f is bounded.*
4. *f has at most countably many discontinuities.*
5. *f is differentiable at Lebesgue almost every point in $\mathbb{R}$.*
6. *There exists a unique finite signed Borel measure μ_f such that for all continuity points a and b of f, $\mu_f((a,b]) = f(b) - f(a)$.*

The functions P_μ and $\overline{f}$ can be defined similarly to before for signed measures and functions of bounded variation. The following lemma connects $\mathrm{var}(f)$ to the measure μ_f.

Lemma A.1.18. *If $V(f) < \infty$, then $|\mu_f|(\mathbb{R}) = V(\overline{f})$.*

The following version of the integration by parts formula will be useful in what follows.

Lemma A.1.19. *If $f, g\colon \mathbb{R} \to \mathbb{R}$, $V(f) < \infty$, $V(g) < \infty$, and $\underline{g}(x) := \lim_{t\to x^-} g(t)$, then for any $a, b \in \mathbb{R}$,*

$$\int_{(a,b]} \overline{f}\, d\mu_g + \int_{(a,b]} \underline{g}\, d\mu_f = \overline{f}(b)\overline{g}(b) - \overline{f}(a)\overline{g}(a)$$

Proof. This follows from a slight modification of the proof of Theorem 3.30 in [10]. □

Note that if g is continuous then $\underline{g}$ and $\overline{g}$ can be replaced by g in the equation above, since all three functions coincide at points of continuity for g. Finally, we will make use of the following density theorem.

Lemma A.1.20. *The set C_c^1 of compactly supported continuously differentiable functions, is dense in $L^1(\mu, \mathbb{R})$ for every Borel measure μ on $\mathbb{R}$ that is finite on compact sets.*

Proof. By Theorem 7.8 of [10] every Borel measure on $\mathbb{R}$ that is finite on compact sets is a Radon measure. By Proposition 7.9 of [10] C_c is dense in $L^1(\mu, \mathbb{R})$ for every Radon measure μ. By the Stone-Weierstrass theorem (Theorem 4.52 of [10]) C_c^1 is dense in C_0 and hence in $L^1(\lambda, \mathbb{R})$. □

A.1.2 Equivalence of var and var_{ac}

Lemma A.1.21. *If $f : \mathbb{R} \to \mathbb{R}$ and $V(f) < \infty$ then $\mathrm{var}_{ac}(f) = V\left(\overline{f}\right)$.*

Proof. Fix $\psi \in Lip_c$ with $|\psi| \leq 1$. Let $P = supp(\mu_f^+)$ and $N = supp(\mu_f^-)$. we see that

$$\begin{aligned} -\int_{\mathbb{R}} \psi \, d\mu_f &= -\int_{\mathbb{R}} \psi|_P \, d\mu_f^+ + \int_{\mathbb{R}} \psi|_N \, d\mu_f^- \\ &\leq \left|\psi_{|P}\right| \int_{\mathbb{R}} d\mu_f^+ + \left|\psi_{|N}\right| \int_{\mathbb{R}} d\mu_f^- \\ &\leq |\mu_f|(\mathbb{R}) \\ &= V\left(\overline{f}\right). \end{aligned}$$

We apply Lemma A.1.19, the fact that ψ is compactly supported, and that $f = \overline{f}$ Lebesgue almost everywhere to obtain

$$-\int_{\mathbb{R}} \psi \, d\mu_f = \int_{\mathbb{R}} \overline{f} \, D\psi \, d\lambda = \int_{\mathbb{R}} f \, D\psi \, d\lambda.$$

Since ψ was arbitrary we have $\mathrm{var}_{ac}(f) \leq V(\overline{f})$.

Next we will show that the bound is attained. Let $\chi = \mathbf{1}_P - \mathbf{1}_N$. Note that $\int_{\mathbb{R}} \chi \, d\mu_f = |\mu_f|(\mathbb{R}) = V(\overline{f})$. Fix $\epsilon > 0$. By Lemma A.1.20 select $\phi \in C_c^1 \subset Lip_c$ such that $\|\chi - \phi\|_{L^1(\mathbb{R},\mu_f)} < \epsilon$ and $|\phi| \le 1$. We compute

$$\left| V(\overline{f}) - \int_{\mathbb{R}} f\ (-D\phi)\ d\lambda \right| = \left| \int_{\mathbb{R}} (\chi - \phi)\, d\mu_f \right| \le \|\chi - \phi\|_{L^1(\mathbb{R},\mu_f)} < \epsilon.$$

From the bound above it follows that

$$\mathrm{var}_{ac}(f) \ge \int_{\mathbb{R}} f D\phi\, d\lambda \ge V(\overline{f}) - \epsilon.$$

Since ϵ and f with $V(f) < \infty$ were arbitrary the claim is proved. ☐

Lemma A.1.22. *If $g, h \colon \mathbb{R} \to \mathbb{R}$, $g = h$ Lebesgue almost everywhere, and g is right continuous, then $V(g) \le V(h)$.*

Proof. Since $g = h$ almost everywhere, the set $\{x \in \mathbb{R} : g(x) = h(x)\}$ is dense in $\mathbb{R}$. Fix points $-\infty < x_0 < \cdots < x_N < \infty$ and $\epsilon > 0$. By right continuity of g, we may select points y_j such that $x_0 < y_0 < x_1 < \cdots < x_N < y_N$ such that

$$\sum_{j=0}^{N} |g(y_j) - g(x_j)| < \frac{\epsilon}{2}$$

and $g(y_j) = h(y_j)$ for all $1 \le j \le N$. Now note that by the triangle inequality

$$\begin{aligned} |g(x_j) - g(x_{j-1})| &\le |g(x_j) - g(y_j)| + |g(y_j) - g(y_{j-1})| + |g(y_{j-1}) - g(x_{j-1})| \\ &= |g(x_j) - g(y_j)| + |h(y_j) - h(y_{j-1})| + |g(y_{j-1}) - g(x_{j-1})|. \end{aligned}$$

and hence

$$|g(x_j) - g(x_j)| - |g(x_j) - g(y_j)| - |g(x_{j-1}) - g(y_{j-1})| \le |h(y_j) - h(y_{j-1})|.$$

Summing the inequality above over j yields

$$\left(\sum_{j=1}^{N} |g(x_j) - g(x_{j+1})| \right) - \epsilon \le \sum_{j=1}^{N} |h(y_j) - h(y_{j+1})|.$$

Since the x_j's and ϵ was arbitrary we conclude that $V(g) \le V(h)$. ☐

Corollary A.1.22.1. *If* $\mathrm{var}(f) < \infty$, *then* $\mathrm{var}_{ac}(f) = \mathrm{var}(f)$.

Proof. Assume that $V(f) < \infty$. Since $\overline{f}$ is right continuous and $f = \overline{f}$ Lebesgue almost everywhere, Lemma A.1.22 implies that $V(\overline{f}) \leq V(g)$ for every g that is equal to f Lebesgue almost everywhere. The claim follows from Lemma A.1.21 and that $\operatorname{var}(f) = V(\overline{f})$. If $\operatorname{var}(f) < \infty$ but $V(f) = \infty$, then select g such that $f = g$ Lebesgue almost everywhere and $V(g) < \infty$. Apply the previous argument to g to obtain middle inequality below,

$$\operatorname{var}(f) = \operatorname{var}(g) = \operatorname{var}_{ac}(g) = \operatorname{var}_{ac}(f).$$

□

We have shown that $\operatorname{var} = \operatorname{var}_{ac}$ on the set where var is finite. Next we will show that equality holds on the set where $\operatorname{var} = \infty$.

Lemma A.1.23. *If* $\operatorname{var}(f) = \infty$, *then* $\operatorname{var}_{ac}(f) = \infty$.

Proof. If $\operatorname{var}(f) = \infty$, then for all $M > 0$ and all g such that $g = f$ Lebesgue almost everywhere, there exist points $-\infty < x_0 < \cdots < x_N < \infty$ such that

$$\sum_{j=1}^{N} |g(x_j) - g(x_{j-1})| \geq M. \tag{$\star$}$$

By Lemma A.1.20 select a sequence $(\phi_n)_{n=1}^{\infty}$ such that $\phi_n \to f$ in $L^1(\mathbb{R}, \lambda)$. Select a subsequence $(\phi_{n_k})_{k=1}^{\infty}$ such that ϕ_{n_k} converges point-wise almost everywhere to f. Let g denote a point-wise limit of (ϕ_{n_k}). Note that $g = f$ Lebesgue almost everywhere. Select $x_0 < \cdots < x_N$ such that ($\star$) holds for g. Select $\delta_1 > 0$ such that $2N\delta < 1$. Fix K and $\phi := \phi_{n_K}$ such that for all $0 \leq j \leq N$

$$|\phi(x_j) - g(x_j)| < \delta_1$$

and $\|\phi - f\|_1 < \delta_2$, where δ_2 will be selected shortly. Apply the triangle inequality to obtain

$$\begin{aligned} \sum_{j=1}^{N} |\phi(x_j) - \phi(x_{j-1})| &\geq \left(\sum_{j=1}^{N} |g(x_j) - g(x_{j-1})| \right) - 2N\delta_1 \\ &\geq M - 1. \end{aligned}$$

The bound above shows that $\operatorname{var}(\phi) \geq M - 1$. Since $\phi \in C_c^1$, we have $\operatorname{var}(\phi) \leq$

$\int |\phi'| \, d\lambda < \infty$. By Corollary A.1.22.1, we have $\mathrm{var}_{ac}(\phi) = \mathrm{var}(\phi) \geq M - 1$. Select $\psi \in Lip_c$ with $|\psi| \leq 1$ such that $\left|\mathrm{var}_{ac}(\phi) - \int \phi D\psi \, d\lambda\right| \leq 1$. Now select $\delta_2 > 0$ such that $\|D\psi\|_\infty \delta_2 < 1$. Combining our bounds we obtain

$$\begin{aligned}
\mathrm{var}_{ac}(f) &\geq \int f D\psi \, d\lambda \\
&= \int \phi D\psi \, d\lambda - \int (\phi - f) D\psi \, d\lambda \\
&\geq (\mathrm{var}(\phi) - 1) - \|\phi - f\|_1 \|D\psi\|_\infty \\
&\geq M - 3.
\end{aligned}$$

Since M was arbitrary we see that $\mathrm{var}_{ac}(f) = \infty$, as desired. □

A.1.3 Equivalence of var_{ac} and var_s

Next we will show that $\mathrm{var}_{ac} = \mathrm{var}_s$.

Lemma A.1.24. *If* $f \in L^1(\mathbb{R}, \lambda)$, *then* $\mathrm{var}_s(f) = \mathrm{var}_{ac}(f)$.

Proof. Every ϕ in C_c^1 is in Lip_c, thus for every $f \in L^1(\mathbb{R}, \lambda)$ we have $\mathrm{var}_s(f) \leq \mathrm{var}_{ac}(f)$. Fix $\epsilon > 0$ and $\psi \in Lip_c$. Chose $s > 0$ such that

$$\int_{|f|>s} |f| \, d\lambda < \frac{\epsilon}{2\|D\psi\|_\infty}.$$

Select $\phi \in C_c^1$, such that $|D\phi| < \|D\psi\|_\infty$, $|\phi| < |\psi| + \epsilon$, and

$$\|D\phi - D\psi\|_1 < \frac{\epsilon}{s}.$$

Then

$$\begin{aligned}
\int_{\mathbb{R}} f\, D\psi\, d\lambda = &\int_{|f|>s} f\, D(\psi-\phi)\, d\lambda + \int_{|f|\le s} f\, D(\psi-\phi)\, d\lambda + \int_{\mathbb{R}} f\, D\phi\, d\lambda \\
\le &\left(\|D\psi\|_\infty + |D\phi|\right) \int_{|f|>s} |f|\, d\lambda + s\int_{\{|f|\le s\}} |D\psi - D\phi|\, d\lambda \\
&+ \int_{\mathbb{R}} f\, D\phi\, d\lambda \\
\le &2\epsilon + \int_{\mathbb{R}} f\, D\phi\, d\lambda \\
\le &2\epsilon + (1+\epsilon)\int_{\mathbb{R}} f\, D\left(\frac{\phi}{|\phi|}\right)\, d\lambda
\end{aligned}$$

Thus for $\psi \in Lip_c$ with $|\psi| \le 1$ we have

$$\int_{\mathbb{R}} f\, D\psi\, d\lambda \le 2\epsilon + (1+\epsilon)\mathrm{var}_s(f)$$

Since ψ and ϵ were arbitrary we have proved the claim. □

A.1.4 Restriction to I.

Given a function $f\colon I \to \mathbb{R}$ we define the variation of f by first extending the domain to $\mathbb{R}$ as follows.

$$\tilde{f} = \begin{cases} f(x), & x \in I \\ 0, & x \in I^c \end{cases}$$

Then we define the variation of f to be the variation of $\tilde{f}$ by any of the three formulae, that is,

$$\begin{aligned}
\mathrm{var}(f) &:= \mathrm{var}(\tilde{f}), \\
\mathrm{var}_s(f) &:= \mathrm{var}_s(\tilde{f}), \\
\mathrm{var}_{ac}(f) &:= \mathrm{var}_{ac}(\tilde{f}).
\end{aligned}$$

In the case of var_{ac} we can streamline the definition slightly so that there is no need to make reference to the extension $\tilde{f}$ or concern ourselves with test functions supported outside I. In particular we will use test functions taken from the following set.

Definition A.1.25. Let *Lip* denote the set of all absolutely continuous functions $\psi\colon I \to \mathbb{R}$ such that $D\psi \in L^\infty(I, \lambda)$.

The key observation is the following lemma.

Lemma A.1.26. *If $\psi \in Lip$ and $|\psi| \leq 1$ then there exists a function $\Psi \in Lip_c$ that extends ψ and $|\Psi| \leq 1$.*

Proof. Given $\psi \in Lip$ chose points $a < 0$ and $b > 1$ and let Ψ be defined by

$$\Psi(x) = \begin{cases} 0, & x < a \\ -\frac{\psi(0)}{a}\, x, & x \in [a, 0] \\ \psi(x), & x \in I \\ -\frac{\psi(1)}{b-1}\, x, & x \in [1, b] \\ 0, & x > b \end{cases}$$

It is elementary to check that Ψ satisfies all of the requirements. □

With this lemma in hand we see that for $\psi \in Lip$

$$\int_I f\, D\psi\, d\lambda = \int_I f\, D\Psi\, d\lambda = \int_{\mathbb{R}} \tilde{f}\, D\Psi\, d\lambda$$

and for all $\zeta \in Lip_c$

$$\int_{\mathbb{R}} \tilde{f}\, D\zeta\, d\lambda = \int_I f\, D\zeta\, d\lambda = \int_I f\, D\zeta|_I\, d\lambda.$$

Therefore,

$$\begin{aligned} \operatorname{var}_{ac}(f) &= \sup\left\{ \int_{\mathbb{R}} \tilde{f}\, D\zeta\, d\lambda : \zeta \in Lip_c, |\zeta| \leq 1 \right\} \\ &= \sup\left\{ \int_I f\, D\psi\, d\lambda : \psi \in Lip, |\zeta| \leq 1 \right\}. \end{aligned} \tag{A.1.1}$$

A similar restriction to I can be made for var_s. Consider the set

$$C = \left\{ \phi \in C_0 : \exists \varphi \in C_c^1, |\varphi| \leq 1 : \psi = \varphi|_I \right\}.$$

Note that $C \subset C_0 \subset Lip$ and hence for $f\colon I \to \mathbb{R}$

$$\begin{aligned}
\operatorname{var}_s(f) = \operatorname{var}_s(\tilde{f}) &= \operatorname{var}_{ac}(\tilde{f}) = \operatorname{var}_{ac}(f) \\
&\geq \sup\left\{\int_I f\, D\phi\, d\lambda : \phi \in C_0, |\phi| \leq 1\right\} \\
&\geq \sup\left\{\int_I f\, D\phi\, d\lambda : \phi \in C\right\} \\
&= \sup\left\{\int_{\mathbb{R}} \tilde{f}\, D\varphi\, d\lambda : \varphi \in C_c^1, |\varphi| \leq 1\right\} \\
&= \operatorname{var}_s(\tilde{f}) = \operatorname{var}_s(f).
\end{aligned}$$

Therefore,

$$\operatorname{var}_s(f) = \sup\left\{\int_I f\, D\phi\, d\lambda : \phi \in C_0, |\phi| \leq 1\right\}. \tag{A.1.2}$$

A.2 Measure Theory

A.2.1 σ-algebra

Let X denote a set and define the *power set* of X to be the set consisting of all subsets of X,

$$\wp(X) = \{A : A \subseteq X\}.$$

It is taken as an axiom of Zermelo-Fraenkel set theory that for any given set X there is a set which contains every subset of X.

The set $\wp(X)$ is naturally equipped with the binary operations[1] of intersection ($\cap$) and union ($\cup$) as well as the involution[2] of complementation ($\cdot^c$). Set inclusion $\subseteq$ is a partial order[3] on $\wp(X)$. A set together with a partial order is a *partially ordered set.*

Given a subset of Y of a partially ordered set $(Z, \leq)$ an element $u \in Z$ is an *upper bound* for Y if for all $y \in Y$, $y \leq u$. Lower bounds are defined similarly. A *least upper bound* l for Y is an upper bound for Y such that for any upper bound u of Y, $l \leq u$.

[1] A binary operation on a set Z is a function mapping $Z \times Z$ into Z.

[2] An involution on a set Z is a function mapping Z onto Z that is its own inverse

[3] A partial order $\leq$ on a set Z is a relation with the following properties for all $x, y, z \in Z$: reflexivity ($x \leq x$), antisymmetry ($x \leq y$ and $y \leq x$ implies $x = y$), transitivity ($x \leq y$ and $y \leq z$ implies $x \leq z$).

Greatest lower bounds are defined similarly.

Given any collection[4] $\mathcal{C} \subseteq \wp(X)$ of subsets of X the least upper bound of $\mathcal{C}$ is the union $\cup\mathcal{C}$ of all sets in $\mathcal{C}$. The greatest lower bound of $\mathcal{C}$ is the intersection $\cap\mathcal{C}$. A partially ordered set in which every subset has a least upper bound and a greatest lower bound is called a *complete lattice*.

Definition A.2.1. A σ-algebra on a set X is a collection $\mathcal{C} \subseteq \wp(X)$ satisfying the following properties:

1. $X \in \mathcal{C}$.

2. If $A \in \mathcal{C}$, then $A^c \in \mathcal{C}$.

3. If for each j in $\mathbb{N}$, $A_j \in \mathcal{C}$, then $\cup_{j\in\mathbb{N}} A_j \in \mathcal{C}$.

Definition A.2.2. Given a set X let $\Sigma(X) \subseteq \wp^2(X)$ denote the set of all σ-algebra on X.

Lemma A.2.3. *Given a set X, $\Sigma(X)$ is a complete sub-lattice of $(\wp^2(X), \subseteq)$. The greatest lower bound of a collection $\mathfrak{C} \subseteq \Sigma(X)$ is given by the intersection $\cap\mathfrak{C}$. The least upper bound of $\mathfrak{C}$ contains $\cup\mathfrak{C}$ and is denoted by $\bigvee_\sigma \mathfrak{C}$.*

Definition A.2.4. Given a set X let $\sigma\colon \wp^2(X) \to \Sigma(X)$ by

$$\sigma(\mathcal{C}) = \bigcap \{\mathcal{A} \in \Sigma(X) : \mathcal{C} \subseteq \mathcal{A}\}.$$

Lemma A.2.5. *If $\mathcal{C} \in \wp^2(X)$ and $\mathcal{D} \in \Sigma(X)$, then*

$$\sigma(\mathcal{C} \cap \mathcal{D}) = \sigma(\mathcal{C}) \cap \mathcal{D}.$$

Proof. First note that by Definition A.2.4 we have $\mathcal{C} \subseteq \sigma(\mathcal{C})$, so $\mathcal{C} \cap \mathcal{D} \subseteq \sigma(\mathcal{C}) \cap \mathcal{D}$. Note that by Lemma A.2.3 $\sigma(\mathcal{C}) \cap \mathcal{D} \in \Sigma(X)$. By applying Definition A.2.4 again we see that

$$\sigma(\mathcal{C} \cap \mathcal{D}) \subseteq \sigma(\mathcal{C}) \cap \mathcal{D}.$$

[4]Note that $\mathcal{C} \in \wp^2(X)$, the power set of the power set or *second power set*. We will have occasion to refer to elements of X, $\wp(X)$, $\wp^2(X)$ and $\wp^3(X)$. We will use the following the conventions to clarify the hierarchy: elements of a set will be denoted by lower case letters ($x \in X$), elements of the power set will be denoted by uppercase letters ($A \in \wp(X)$), elements of the second power set will be denoted by uppercase calligraphic letters ($\mathcal{C} \in \wp^2(X)$), and elements of the third power set will be denoted by uppercase fraktur letters ($\mathfrak{C} \in \wp^3(X)$) or uppercase Greek letters.

Now note that for all $\mathcal{B} \in \Sigma(X)$ such that $\mathcal{C} \subseteq \mathcal{B}$ we have that $\mathcal{C} \cap \mathcal{D} \subseteq \mathcal{B} \cap \mathcal{D}$ and $\mathcal{B} \cap \mathcal{D} \in \Sigma(X)$. Therefore,

$$\{\mathcal{B} \cap \mathcal{D} : \mathcal{B} \in \Sigma(X), \mathcal{C} \subseteq \mathcal{B}\} \subseteq \{\mathcal{C} \in \Sigma(X) : \mathcal{C} \cap \mathcal{D} \subseteq \mathcal{C}\} .$$

Note that

$$\bigcap \{\mathcal{B} \cap \mathcal{D} : \mathcal{B} \in \Sigma(X), \mathcal{C} \subseteq \mathcal{B}\} = \left(\bigcap \{\mathcal{B} : \mathcal{B} \in \Sigma(X), \mathcal{C} \subseteq \mathcal{B}\}\right) \cap \mathcal{D} = \sigma(\mathcal{C}) \cap \mathcal{D}$$

Combining the two displayed equations above shows that

$$\sigma(\mathcal{C}) \cap \mathcal{D} \subseteq \sigma(\mathcal{C} \cap \mathcal{D})$$

We conclude that

$$\sigma(\mathcal{C}) \cap \mathcal{D} = \sigma(\mathcal{C} \cap \mathcal{D})$$

as desired. □

Definition A.2.6. A partition on a set X is a collection $\mathcal{Q} \subseteq \wp(X)$ satisfying the following properties:

1. $\cup \mathcal{Q} = X$,

2. For any $P \neq Q \in \mathcal{Q}$, $P \cap Q = \emptyset$.

Definition A.2.7. Given a set X let $\Pi(X) \subseteq \wp^2(x)$ denote the set of all partitions of X.

Definition A.2.8. Given a partition $\mathcal{Q} \in \Pi(X)$ define the associated *saturation map* $\mathcal{Q}\colon \wp(X) \circlearrowleft$ by

$$\mathcal{Q}(E) = \{x \in X : \exists Q \in \mathcal{Q} \text{ s.t. } Q \cap E \neq \emptyset \text{ and } x \in Q\} .$$

Lemma A.2.9. *For all $\mathcal{Q} \in \Pi(X)$ and $E \in \wp(X)$, $E \subseteq \mathcal{Q}(E)$.*

Proof. Fix $x \in E$. Since $\cup \mathcal{Q} = X$ there exists $Q \in \mathcal{Q}$ such that $x \in Q$. Since $x \in E \cap Q$ we see that $E \cap Q \neq \emptyset$ and therefore by Definition A.2.8 $x \in \mathcal{Q}(E)$. Since $x \in E$ was arbitrary we conclude that $E \subseteq \mathcal{Q}(E)$ as desired. □

Lemma A.2.10. *For any partition $\mathcal{Q} \in \Pi(X)$ the associated saturation map is idempotent, that is for every $E \in \wp(X)$, $\mathcal{Q}^2(E) = \mathcal{Q}(E)$.*

Proof. By Lemma A.2.9 we have $\mathcal{Q}(E) \subseteq \mathcal{Q}^2(E)$ so it suffices to show that $\mathcal{Q}^2(E) \subseteq \mathcal{Q}(E)$. Fix an element $x \in \mathcal{Q}^2(E)$. By Definition A.2.8 there exists $Q \in \mathcal{Q}$ such that

$$x \in Q$$

and $Q \cap \mathcal{Q}(E) \neq \emptyset$. Select $y \in Q \cap \mathcal{Q}(E)$. Since $y \in \mathcal{Q}(E)$ there exists $P \in \mathcal{Q}$ such that $y \in P$ and $P \cap E \neq \emptyset$. Since $y \in P \cap Q$ and $P, Q \in \mathcal{Q}$ we conclude that $P = Q$ and thus

$$Q \cap E \neq \emptyset.$$

The displayed equations above show that $x \in \mathcal{Q}(E)$. Since $x \in \mathcal{Q}^2(E)$ was arbitrary we conclude that $\mathcal{Q}^2(E) \subseteq \mathcal{Q}(E)$ as desired. □

Definition A.2.11. Given a set X and a partition $\mathcal{Q} \in \Pi(X)$ define $\wp(X, \mathcal{Q}) \subseteq \wp(X)$ by

$$\wp(X, \mathcal{Q}) = \{E \subseteq X : \mathcal{Q}(E) = E\}.$$

A set $E \in \wp(X, \mathcal{Q})$ is referred to as a *$\mathcal{Q}$-saturated set.*

Lemma A.2.12. *For every partition $\mathcal{Q} \in \Pi(X)$ the collection $\wp(X, \mathcal{Q})$ is a complete algebra, that is*

1. $X \in \wp(X, \mathcal{Q})$;

2. *for all* $E \in \wp(X, \mathcal{Q})$, $E^c \in \wp(X, \mathcal{Q})$;

3. *for all* $\mathcal{C} \subseteq \wp(X, \mathcal{Q})$, $\cap\mathcal{C} \in \wp(X, \mathcal{Q})$.

Proof. We begin by verifying condition 2. Suppose that $x \in \mathcal{Q}(E^c) \cap E$. By Definition A.2.8 there exists $Q \in \mathcal{Q}$ such that $x \in Q$ and $Q \cap E^c \neq \emptyset$. Note that $x \in Q \cap E$ Select $y \in Q \cap E^c$. Since $y \in Q$ and $Q \cap E \neq \emptyset$ we have $y \in \mathcal{Q}(E)$. Since $E \in \wp(X, \mathcal{Q})$ we have $\mathcal{Q}(E) = E$ and thus $y \in E$. There fore $y \in E \cap E^c$, and we arrive at a contradiction. We conclude that $\mathcal{Q}(E^c) \cap E = \emptyset$ and thus that $\mathcal{Q}(E^c) \subseteq E^c$. By Lemma A.2.9 $E^c \subseteq \mathcal{Q}(E^c)$ so $\mathcal{Q}(E^c) = E^c$ and therefore $E^c \in \wp(X, \mathcal{Q})$.

Next we verify condition 3. Fix $x \in \mathcal{Q}(\cap\mathcal{C})$. By Definition A.2.8 there exists Q such that $x \in Q$ and $Q \cap (\cap\mathcal{C}) \neq \emptyset$. For all $E \in \mathcal{C}$, $Q \cap E \neq \emptyset$, so $x \in \mathcal{Q}(E)$. Since $(E) = E$ for all $E \in \mathcal{C}$ we see that $x \in \cap\mathcal{C}$. Since $x \in \mathcal{Q}(\cap\mathcal{C})$ was arbitrary we have $\mathcal{Q}(\cap\mathcal{C}) \subseteq \cap\mathcal{C}$. An application of Lemma A.2.9 provides the other inclusion and so condition 3 is verified.

Finally we verify condition 1. By de Morgans laws and conditions 2 and 3 we see that $\wp(X, \mathcal{Q})$ is also closed under arbitrary unions. It is not hard to check that $\mathcal{Q} \subseteq \wp(X, \mathcal{Q})$. We conclude that $\mathcal{Q}(X) = \mathcal{Q}(\cup\mathcal{Q}) = \cup\mathcal{Q} = X$. □

Corollary A.2.12.1. *For every partition* $\mathcal{Q} \in \Pi(X)$ *the collection* $\wp(X, \mathcal{Q})$ *is a* σ*-algebra, that is* $\wp(X, \mathcal{Q}) \in \Sigma(X)$.

Definition A.2.13. Given a partition $\mathcal{Q} \in \Pi(X)$ and a collection $\mathcal{C} \subset \wp(X)$, define

$$\mathcal{C}/\mathcal{Q} = \mathcal{C} \cap \wp(X, \mathcal{Q}).$$

Lemma A.2.14. *If* $\mathcal{F} \in \Sigma(X)$*, then* $\mathcal{F}/\mathcal{Q} \in \Sigma(X)$

Proof. By Corollary A.2.12.1 and Lemma A.2.3 we see that $\mathcal{F}/\mathcal{Q}$ is indeed a σ-algebra. □

Lemma A.2.15. *If* $\mathcal{C} \subseteq \wp(X)$ *and* $\mathcal{Q} \in \Pi(X)$*, then* $\sigma(\mathcal{C}/\mathcal{Q}) = \sigma(\mathcal{C})/\mathcal{Q}$.

Proof. This follows directly from Lemma A.2.5, Corollary A.2.12.1, and Definition A.2.13. □

Bibliography

[1] Jon Aaronson and Manfred Denker. Local limit theorems for partial sums of stationary sequences generated by gibbs–markov maps. *Stochastics and Dynamics*, 1(02):193–237, 2001.

[2] Viviane Baladi. *Positive transfer operators and decay of correlations*, volume 16. World scientific, 2000.

[3] Viviane Baladi and Masato Tsujii. Anisotropic hölder and sobolev spaces for hyperbolic diffeomorphisms. In *Annales de l'institut Fourier*, volume 57, pages 127–154, 2007.

[4] Michael Blank, Gerhard Keller, and Carlangelo Liverani. Ruelle–perron–frobenius spectrum for anosov maps. *Nonlinearity*, 15(6):1905, 2002.

[5] Christopher Bose and Rua Murray. Polynomial decay of correlations in the generalized baker's transformation. *International Journal of Bifurcation and Chaos*, 23(08):1350130, 2013.

[6] Christopher J. Bose. Generalized baker's transformations. *Ergodic Theory and Dynamical Systems*, 9:1–17, 3 1989.

[7] Mark Demers and Carlangelo Liverani. Stability of statistical properties in two-dimensional piecewise hyperbolic maps. *Transactions of the American Mathematical Society*, 360(9):4777–4814, 2008.

[8] Mark F. Demers. Functional norms for young towers. *Ergodic Theory and Dynamical Systems*, 30:1371–1398, 10 2010.

[9] Jiu Ding, Q Du, and TY Li. The spectral analysis of frobenius-perron operators. *Journal of Mathematical Analysis and Applications*, 184(2):285–301, 1994.

[10] Gerald B. Folland. *Real analysis*. Pure and Applied Mathematics (New York). John Wiley & Sons, Inc., New York, 1984. Modern techniques and their applications, A Wiley-Interscience Publication.

[11] A. O. Gelfond. An estimate for the remainder term in a limit theorem for recurrent events. *Theory of Probability & Its Applications*, 9(2):299–303, 1964.

[12] David Gilbarg and Neil S Trudinger. *Elliptic partial differential equations of second order*. springer, 2015.

[13] Sébastien Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel Journal of Mathematics*, 139(1):29–65, 2004.

[14] Sébastien Gouëzel and Carlangelo Liverani. Banach spaces adapted to anosov systems. *Ergodic Theory and Dynamical Systems*, 26:189–217, 2 2006.

[15] Hubert Hennion and Loïc Hervé. *Limit theorems for Markov chains and stochastic properties of dynamical systems by quasi-compactness*, volume 1766 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2001.

[16] A. Lasota and James A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Transactions of the American Mathematical Society*, 186:pp. 481–488, 1973.

[17] Carlangelo Liverani. Invariant measures and their properties. a functional analytic point of view. *Dynamical systems. Part II*, pages 185–237, 2004.

[18] Carlangelo Liverani and Dalia Terhesiu. Mixing for some non-uniformly hyperbolic systems. *Annales Henri Poincar*, pages 1–48, 2015.

[19] Ian Melbourne and Dalia Terhesiu. Operator renewal theory and mixing rates for dynamical systems with infinite measure. *Inventiones mathematicae*, 189(1):61–110, 2012.

[20] Roger D Nussbaum et al. The radius of the essential spectrum. *Duke math. J*, 37:473–478, 1970.

[21] David Ruelle. A measure associated with axiom-a attractors. *American Journal of Mathematics*, 98(3):pp. 619–654, 1976.

[22] Marek Rychlik. Bounded variation and invariant measures. *Studia Mathematica*, 76(1):69–80, 0 1983.

[23] Omri Sarig. Subexponential decay of correlations. *Inventiones mathematicae*, 150(3):629–653, 2002.

[24] Lai-Sang Young. Statistical properties of dynamical systems with some hyperbolicity. *Annals of Mathematics*, 147(3):pp. 585–650, 1998.

[25] Lai-Sang Young. Recurrence times and rates of mixing. *Israel Journal of Mathematics*, 110(1):153–188, 1999.