Approximation of Absolutely Continuous Invariant measures for Markov Compositions of Maps of an Interval

Chandra Nath Podder

A Thesis

in

The Department

of

Mathematics and Statistics

Presented in Partial Fulfilment of the Requirements

for the Degree of Master of Science at

Concordia University

Montreal, Quebec, Canada

September, 2004

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-94672-X
Our file Notre référence
ISBN: 0-612-94672-X

Abstract

Approximation of Absolutely Continuous Invariant measures for Markov Compositions of Maps of an Interval

Chandra Nath Podder

We study the approximation of absolutely continuous invariant measures of systems defined by random compositions of piecewise monotonic transformation (Lasota-Yorke maps). We discuss a generalization of Ulam's finite approximation conjecture to the situation where a family of piecewise monotonic transformations are composed according to a Markov law, and study an analogous convergence result. Also, we present bounds for the L^1 error of the Ulam's approximation.

Acknowledgments

I would like to express my gratitude to my supervisor, **Dr. Pawel Góra**, Department of Mathematics and Statistics of Concordia University, for accepting me as a masters student and for his valuable guidance and comments in the preparation of this thesis. I wish to thank Mr.Shafiqul Islam and Mr. Wael Bahsun, graduate students, for their valuable discussions with me. I am also grateful to the Department of Mathematics and Statistics of Concordia University and Concordia University for the financial support during my M.Sc. program.

Contents

Chapter 1

Introduction

General Introduction

A random map is a discrete time dynamical system under considering of a collection of transformations which are selected randomly by means of probabilities at each iteration. Let $\{T_k\}_{k=1}^{r}$ be a collection of nonsingular mappings from the unit interval I into itself. Given an intial point $x \in I$, and a random sequence $(k_o, k_1, \cdots,)$ with $k_N \in \{1, 2, \cdots, r\}$ for $N \geqq 0$, a random orbit by defining the N^{th} point in the orbit to be $x_N = x_N(k_{N-1}, \cdots, k_0, x) := T_{k_{N-1}} \circ \cdots \circ T_{k_1} \circ T_{k_0} x$. The map $T_{k_{N-1}}$ that is applied at time N is chosen so as to depend only on the map applied at the previous time step, and according to the same probability law for all time. In this situation, the indices $k_0, k_1, \cdots$ arise as random variables of a stationary first order Markov chain, and we call such a composition of maps a **Markov random composition**.

We study the asymptotic behaviour of such systems in situation where the orbits $\{x_N\}_{N=0}^{\infty} \subset I$ have the same asymptotic distribution on I for almost all sequences $k_0, k_1, \cdots$ and almost all starting points $x \in I$, and we discuss the method of Ulam [15] which produces a rigorous approximation method for absolutely continuous probability measures that are invariant on average under the action of the random systems. In the case of a single mapping (with $\inf_{x\in I \ \{b_0,\cdots,b_q\}} |T'(x)| > 1$) Li [9] first proved convergence of Ulam's approximation to the unique absolutely continuous invariant measure, following the Lasota and Yorke [8] proof of the existence of an absolutely continuous invariant measure. The existence of an absolutely continuous invariant measure for independent identical distributed (iid) random compositions of such mappings has been considered in this setting by Pelikan [13].
In the case of random compositions Ulam's conjecture has been studied by Froyland [3]. In our thesis, we follow Froyland [3], where we restrict ourselves by using Lasota-Yorke type.

Outline of the thesis

In Chapter 2, we introduce the Frobenius-Perron operator, which is a powerful tool to study the existence of absolutely continuous invariant measures for a large class of transformations τ, which are piecewise $\mathcal{C}^2$ and satisfy the condition $|\tau'| > 1$, where the derivative exists. In Section 2.4, we present some well-known results; the Kakutani-Yoshida theorem, Helly's Selection Principle and the Lasota-Yorke Theorem, which are needed for the existence theorem. Section 2.5 deals with approximating the fixed point of the Frobenius-Perron operator with fixed points of matrices.

In Chapter 3, we give the proof of the existence of absolutely continuous invariant measures and proof of convergence of Ulam's method. We begin by proving suitable inequalities regarding the variation of test functions and their images under an appropriate Frobenius-Perron operator. Then, we study a finite-dimensional approximation of this Frobenius-Perron operator and use the variational inequalities to prove convergence of Ulam's method. Under slightly stricter conditions, we discuss the rate of convergence, and in the case where each T_k is a $\mathcal{C}^2$ circle map, we study the bound for the error in our approximation, in terms of fundamental constants of the mappings T_k, $k = 1, 2 \cdots, r$.

Chapter 2

Invariant measures and The Frobenius-Perron Operator

2.1 Review of Functional Analysis and Statistics

In this section we will briefly state some well-known definitions from measure theory, dynamical systems and Markov processes. In this chapter we follow: [1],[2],[5],[6],[8],[9], [12],[13],[15].

Definition 2.1 *Let X be a set. A family $\mathcal{B}$ of subsets of X is called a* ***σ-algebra*** *iff it satisfies:*

(i) $X \in \mathcal{B}$

(ii) for any $B \in \mathcal{B}, X \backslash B \in \mathcal{B}$

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

Elements of $\mathcal{B}$ are known as ***measurable sets***

Definition 2.2 *A real valued function $\mu : \mathcal{B} \longrightarrow \mathbb{R}^+$ is called* ***measure*** *on a σ-algebra $\mathcal{B}$ if :*

(i) $\mu(\emptyset) = 0$

(ii) $\mu(B) \geqq 0$ *for all* $B \in \mathcal{B}$ *and*

(iii) for any sequence $\{B_n\}$ of disjoint measurable sets, $B_n \in \mathcal{B}$, $n = 1, 2, \cdots$,

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

Definition 2.3 *If $\mathcal{B}$ is a σ-algebra of subsets of X and μ is a measure on $\mathcal{B}$ then triple $(X, \mathcal{B}, \mu)$ is called a* ***measure space.***

Definition 2.4 *A measure space $(X, \mathcal{B}, \mu)$ is called finite if $\mu(X) < \infty$. In particular, if $\mu(X) = 1$, then the measure space is said to be a* ***normalised measure space*** *or* ***a probability space.***

Definition 2.5 *Let $(X, \mathcal{B}, \mu)$ be a measure space. The function $f : X \rightarrow \mathbb{R}$ is said to be* ***measurable*** *if for all $c \in \mathbb{R}$, $f^{-1}(c, \infty) \in \mathcal{B}$, or, equivalently, if $f^{-1}(B) \in \mathcal{B}$ for any Borel set.*

Definition 2.6 *Let $(X, \mathcal{B}, \mu)$ be a measure space. A transformation $T : X \rightarrow X$ is* ***measurable*** *if*

$$T^{-1}(A) = \{x \in X : \quad T(x) \in A\} \in \mathcal{B} \qquad \forall A \in \mathcal{B}.$$

Definition 2.7 *Let $(X, \mathcal{B}, \mu)$ be a normalised measure space and $T : X \to X$ be a transformation. Then T is **non-singular** if and only if $\mu(T^{-1}(A)) = 0$ wherever $\mu(A) = 0$, for all measurable subsets A of X.*

Definition 2.8 *We say the measurable transformation $T : X \to X$ **preserves measure** μ or that μ is T**-invariant** if $\mu(T^{-1}(A)) = \mu(A)$, for all $A \in \mathcal{B}$.*

Definition 2.9 *Let ν and μ be two measures on the same measure space $(X, \mathcal{B})$. We say that ν **is absolutely continuous with respect to** μ if for any $A \in \mathcal{B}$, such that $\mu(A) = 0$, it follows that $\nu(A) = 0$ and we write $\nu << \mu$.*

If $\nu << \mu$, then it is possible to represent ν in terms of μ.

Definition 2.10 *Let $(X, \mathcal{B}, \mu)$ be a measure space. By $(L^1, \|\cdot\|_1)$ we mean the family of all integrable functions f on X, i.e.,*

$$(L^1, \|\cdot\|_1) = \{f : X \to \mathbb{R} \ \textit{such that} \ \| f \|_1 = \int |f(x)| d\mu(x) < \infty\}.$$

By $(L^\infty, \|\cdot\|_\infty)$, we mean $(L^\infty, \|\cdot\|_\infty) =$ space of almost everywhere bounded measurable functions on $(X, \mathcal{B}, \mu)$ i.e.,

$$(L^\infty, \|\cdot\|_\infty) = \{f : X \to \mathbb{R} \ \textit{such that} \ \| f \|_\infty = \mathrm{essup}|\mathrm{f(x)}| < \infty\}$$

where $\mathrm{essup}|\mathrm{f(x)}| = \inf\{\mathrm{M} : \mu\{\mathrm{x} : |\mathrm{f(x)}| > \mathrm{M}\} = 0\}$.

Theorem 2.1 *Let $(X, \mathcal{B})$ be a space and ν and μ be two normalized measures on $(X, \mathcal{B})$. If $\nu << \mu$, then there exists a unique $f \in L^1(X, \mathcal{B}, \mu)$ such that for every $A \in \mathcal{B}$,*

$$\nu(A) = \int_A f d\mu.$$

f is called the ***Randon-Nikodym derivative*** *and is denoted by $\frac{d\nu}{d\mu}$.*

Definition 2.11 *Let $r \geqq 1$. $\mathcal{C}^r(X)$ denotes the space of all r-**times continuously differentiable** real functions $f : X \to \mathbb{R}$ with the norm*

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

where $f^{(k)}(x)$ is the k-th derivative of $f(x)$ and $f^{(0)}(x) = f(x)$.

Definition 2.12 *A transformation $T : X \to \mathbb{R}$ is called **piecewise** $\mathcal{C}^2$, if there exists a partition $a = a_0 < a_1 < \cdots < a_n = b$ of the closed interval $I = [0, 1]$ such that for each integer $i = 1, 2, \cdots, n$, the restriction T_i of T to the open interval (a_{i-1}, a_i) is a $\mathcal{C}^2$ function which can be extended to the closed interval $[a_{i-1}, a_i]$ as a $\mathcal{C}^2$ function. T need not be continuous at the points a_i.*

The Birkhoff Ergodic Theorem [2]

Let $\tau : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ be measure preserving and $E \in \mathcal{B}$. For $x \in X$, a question of physical interest is: with what frequency do the points of the orbit $\{x, \tau(x), \tau^2(x), \cdots\}$ occur in the set E?

Clearly, $\tau^i(x) \in E$ if and only if $\chi_E(\tau^i(x)) = 1$. Thus, the number of points of

$\{x, \tau(x), \tau^2(x), \cdots, \tau^{n-1}(x)\}$ in E is equal to $\sum_{k=0}^{n-1} \chi_E(\tau^k(x))$, and the relative frequency of elements of $\{x, \tau(x), \tau^2(x), \cdots, \tau^{n-1}(x)\}$ in E equals to $\frac{1}{n} \sum_{k=0}^{n-1} \chi_E(\tau^k(x))$. The following theorem is the first major result in ergodic theory and was proved in 1931 by G.D. Birkhoff.

Theorem 2.2 *Suppose $\tau : (X, \mathcal{B}, \mu) \to (X, \mathcal{B}, \mu)$ is measure preserving, where $(X, \mathcal{B}, \mu)$ is σ-finite, and $f \in L^1(\mu)$. Then there exists a function $f^* \in L^1(\mu)$ such that*

$$\frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k(x)) \to f^*, \quad \mu - a.e.$$

Furthermore, $f^* \circ \tau = f^*$ μ-a.e. and if $\mu(x) < \infty$, then $\int_X f^* d\mu = \int_X f d\mu$.

Theorem 2.3 (Brouwer Fixed-Point Theorem) *[5] Let S be the closed unit sphere in an n dimensional real Euclidian space; that is, $S = [x | x \text{ in } E_n \text{ and } \| x \| \leqq 1]$. Let K be a continuous mapping of S into itself so that if $\| x \| \leqq 1$, $\| K(x) \| \leqq 1$. Then K has at least one fixed point; that is, there is at least one x in S such that $K(x) = x$.*

Briefly from Statistics:

We consider a stochastic process $\{X_n, n = 0, 1, 2, \cdots\}$, that is, a family of random variables, defined on the space X of all possible values that the random variables can assume. The space X is called the **state space** of the process, and the elements $x \in X$, the different values that X_n can assume, are called the **states**.

We seek the conditional probability $\mathcal{P}\{X_{n+1} = x_{n+1} | X_n = x_n, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0 = 1\}$. If the structure of the stochastic process $\{X_n, n = 0, 1, \cdots\}$ is such that the conditional probability distribution of X_{n+1} depends only on the value of X_n and is independent of all previous values, we say that the process has a **Markov property** and call it a **Markov chain**. More precisely, $\mathcal{P}\{X_{n+1} = x_{n+1}|X_n = x_n, X_{n-1} = x_{n-1}, \cdots, X_0 = x_0 = 1\} = \mathcal{P}\{X_{n+1} = x_{n+1}|X_n = x_n\}$.

Let us now write

$$p_{ij} = \mathcal{P}\{X_{n+1} = j|X_n = i\}, i, j = 0, 1, 2, \cdots .$$

Definition 2.13 *Let p_{ij} be the probability of a transition from the state i to the state j, and call $P = (p_{ij})$ the matrix of* ***transition probabilities****:*

$$P = \begin{bmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{12} & p_{13} & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

P is a square matrix (of infinite order since the chain has a denumerable number of states) with nonnegative elements, since $p_{ij} \geqq 0$ for all i and j, and with row sums equal to unity, since $\sum_{j=0}^{\infty} p_{ij} = 1$ for all i. Such a matrix is called a ***stochastic*** *, or* ***Markov matrix****.*

Definition 2.14 *A* ***Markov chain*** *is completely defined by a matrix of transition probabilities P and a column vector, say $Q = (q(0), q(1), \cdots)$, which gives the probability distribution for the state $x = 0, 1, 2, \cdots$ at time zero.*

In addition to the so-called one-step transition probabilities p_{ij}, it is of interest to consider the higher, or n-step, transition probabilities denoted by $p_{ij}^{(n)}$. These express the probability of a transition from the state i to the state j in n $(n > 1)$ steps.

Definition 2.15 *A set of states $S \in X$ (state space) is called* ***closed*** *if no one-step transition is possible from any state in S to any state in $X - S$, the complement of the set S. Hence $p_{ij} = 0$ for $i \in S$ and $j \in X - S$. If the set S contains only one state, this state is called an* ***absorbing state****. It is clear that a necessary and sufficient condition for a state i to be an absorbing state is that $p_{ii} = 1$. If the state space X contains two or more closed sets, the chain is called* ***decomposable*** *or* ***reducible****. The Markov matrix associated with a decomposable chain can be written in the form of a partitioned matrix; for example,*

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}.$$

In the above, P_1 and P_2 represent Markov matrices which describe the transitions within the two closed sets of states. A chain, or matrix, which is not decomposable is called ***indecomposable*** *or* ***irreducible****, and a chain is indecomposable if and only if every state can be reached from every other state.*

Definition 2.16 *If $i \to i$, the greatest common divisor of the set of positive integers n such that $p_{ii}^{(n)} > 0$ is called the* ***period*** *of the state i. A state that is not periodic is called* ***aperiodic****.*

2.2 Absolutely Continuous Invariant Measures

Let $X = [0,1]$ and $\tau : X \to X$ (not necessarily one-to-one). For $A \subset X$, $\tau^{-1}(A) = \{x \in X : \tau(x) \in A\}$. We consider the average amount of time the orbit $\{\tau^n(x)\}_{n=0}^{\infty}$ spends in a set $B \subset X$. The number of times $\{\tau^n(x)\}_{n=0}^{\infty}$ is in B for n between 0 and N is

$$\sum_{n=0}^{N} \chi_B(\tau^n(x)).$$

The average time spent in B may be defined to be

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \chi_B(\tau^n(x)), \tag{2.1}$$

when limit exists.

A measure μ is an **absolutely continuous measure** iff there is a function $f : X \to [0,\infty)$, $f \in L^1(X)$, such that

$$\mu(B) = \int_B f(x)dx, \tag{2.2}$$

for every Lebesgue measurable set $B \subset X$. The density in (2.2) or the corresponding measure μ is called **invariant** (under τ) if $\mu(\tau^{-1}(A)) = \mu(A)$ for every measurable set A. The Birkhoff Ergodic Theorem (Theorem 2.2) says that if there exists an invariant density and the density is unique, then the limit in (2.1) exists for almost all x and furthermore

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} g(\tau^n(x)) = \int_0^1 g(x)f(x)dx \text{ a.e.,}$$

where g is integrable. In other words, except for x in a set B, $\mu(B) = 0$, the time average $\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} g(\tau^n(x))$ is equal to the space average $\int_0^1 g(x)f(x)dx$. Therefore, if one can find the absolutely continuous invariant measure (acim) μ for τ, then the problem of finding the limit in (2.1) is transformed into computing $\int_B g d\mu$. To find the acim μ for τ, let $g = \chi_B$, so

$$\mu(B) = \int_{[0,1]} \chi_B f(x)dx = \lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \chi_B(\tau^n(x)),$$

for almost every x in [0,1]. Hence, one might choose almost any $x \in [0,1]$ and calculate the average time for iterations $\tau^n(x)$ to recur in B.

2.3 The Frobenius-Perron Operator

2.3.1 Motivation

The Frobenius-Perron operator is a powerful tool to study absolutely continuous invariant measures.

Let X be a random variable on the space $I = [0,1]$ having probability density function f. Then, for any measurable set $A \subset I$,

$$\text{Prob}\{\text{x} \in \text{A}\} = \int_{\text{A}} \text{fdm},$$

where m is Lebesgue measure on I. Let $\tau : I \to I$ be a transformation. We would like to know the probability that x is in A after being transformed by τ. Thus, we write

$$\text{Prob}\{\tau(\text{x}) \in \text{A}\} = \text{Prob}\{\text{x} \in \tau^{-1}(\text{A})\} = \int_{\tau^{-1}(\text{A})} \text{fdm}.$$

Further, we would like to know if there exists a function ϕ such that

$$\text{Prob}\{\tau(\text{x}) \in \text{A}\} = \int_{\text{A}} \phi \text{dm}.$$

Obviously, if such a ϕ exists, it will depend both on f and on τ. Let us assume that τ is non-singular and define

$$\mu(A) = \int_{\tau^{-1}(A)} f dm,$$

where $f \in L_1$ and A is an arbitrary measurable set. Since τ is nonsingular, $m(A) = 0$ implies $m(\tau^{-1}(A)) = 0$, which in turn implies that $\mu(A) = 0$. Hence $\mu << m$. Then, by the Radon-Nikodym Theorem, there exists a $\phi \in L^1$ such that for all measurable sets A,

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

ϕ is unique a.e., and depends on τ and f. Set $\mathcal{P}_\tau f = \phi$. Thus, the probability density function f has been transformed to a new probability density function $\mathcal{P}_\tau f$. $\mathcal{P}_\tau$ obviously depends on the transformation τ and is an operator from the space of probability density functions on I into itself.

Thus, $\mathcal{P}_\tau$ maps L^1 into L^1. If we let $A = [a, x] \subset I$, then

$$\int_a^x \mathcal{P}_\tau f dm = \int_{\tau^{-1}[a,x]} f dm.$$

$\mathcal{P}_\tau$ is reffered to as the Frobenius-Perron operator associated wth τ. On differentiating both sides with respect to x, we obtain

$$\mathcal{P}_\tau f(x) = \frac{d}{dx} \int_{\tau^{-1}[a,x]} f dm.$$

Clearly f is invariant under τ if and only if $\mathcal{P}_\tau f = f$, i.e., f is a fixed point of the Frobenius-Perron operator. We study $\mathcal{P}_\tau$ because if there exists $f \in L^1$ with $\mathcal{P}_\tau f = f$ then the measure $\mu = \int f dm$ is invariant under τ. Thus, to find invariant measures for τ, we may instead find the fixed points of the Frobenius-Perron operator.

Definition 2.17 *Let $I = [a, b]$, $\mathcal{B}$ be the Borel σ-algebra of subsets of I and let m denote the normalized Lebesgue measure on I. Let $\tau : I \to I$ be a nonsingular transformation. We define the* ***Frobenius-Perron operator*** *$\mathcal{P}_\tau : L^1 \to L^1$ as follows:*

for any $f \in L^1$, $\mathcal{P}_\tau f$ is the unique (up to a.e. equivalence) function in L^1 such that

$$\int_A \mathcal{P}_\tau f dm = \int_{\tau^{-1}(A)} f dm$$

for any $A \in \mathcal{B}$. The validity of this definition, i.e., the existence and uniqueness of $\mathcal{P}_\tau f$, follows by the Theorem 2.1 (Radon-Nikodym).

2.3.2 Properties of the Frobenius-Perron Operator

Lemma 2.1 *(Linearity) $\mathcal{P}_\tau : L^1 \to L^1$ is a linear operator.*

Lemma 2.2 *(Positivity) Let $f \in L^1$ and assume that $f \geqq 0$. Then $\mathcal{P}_\tau f \geqq 0$.*

Lemma 2.3 *(Preservation of Integrals)*

$$\int_I \mathcal{P}_\tau f dm = \int_I f dm.$$

Lemma 2.4 *(Contraction Property)* $\mathcal{P}_\tau : L^1 \longrightarrow L^1$ *is a contraction, i.e.,*

$$\| \mathcal{P}_\tau f \|_1 \leqq \| f \|_1 \qquad \text{for any } f \in L^1.$$

Moreover, $\mathcal{P}_\tau : L^1 \rightarrow L^1$ *is continuous with respect to the norm topology since* $\| \mathcal{P}_\tau f - \mathcal{P}_\tau g \|_1 \leqq \| f - g \|_1$.

Lemma 2.5 *(Composition Property) Let* $\tau : I \longrightarrow I$ *and* $\sigma : I \longrightarrow I$ *be nonsingular. Then* $\mathcal{P}_{\tau \circ \sigma} f = \mathcal{P}_\tau \circ \mathcal{P}_\sigma f$. *In particular,* $\mathcal{P}_{\tau^n} f = \mathcal{P}_\tau^n f$.

Lemma 2.6 *Let* $\tau : I \longrightarrow I$ *be nonsingular. Then* $\mathcal{P}_\tau f^* = f^*$ *a.e., if and only if the measure* $\mu = f^* m$, *defined by* $\mu(A) = \int_A f^* dm$, *is* τ *invariant, i.e., if and only if* $\mu(\tau^{-1}(A)) = \mu(A)$ *for all measurable sets* A, *where* $f^* \geqq 0$, $f^* \in L^1$ *and* $\| f^* \|_1 = 1$.

2.3.3 Representation of the Frobenius-Perron Operator

Here we present the representation for the Frobenius-Perron operator for piecewise monotonic transformations.

By the definition of $\mathcal{P}_\tau$, we have

$$\int_A \mathcal{P}_\tau f dm = \int_{\tau^{-1}(A)} f dm$$

for any Borel set A in I. Since τ is monotonic on each $(a_{i-1}, a_i), i = 1, 2, \cdots, n$, we can define an inverse function for each $\tau|_{(a_{i-1}, a_i)}$. Let $\phi_i = \tau^{-1}|_{B_i}$, where $B_i = \tau([a_{i-1}, a_i])$. Then $\phi_i : B_i \longrightarrow [a_{i-1}, a_i]$ and

$$\tau^{-1}(A) = \bigcup_{i=1}^{n} \phi_i(B_i \cap A),$$

where the sets $\{\phi_i(B_i \cap A)\}_{i=1}^{n}$ are mutually disjoint. Note that, depending on A, $\phi_i(B_i \cap A)$ may be empty. We obtain

$$\begin{aligned} \int_A \mathcal{P}_\tau f dm &= \sum_{i=1}^{n} \int_{\phi_i(B_i \cap A)} f dm \\ &= \sum_{i=1}^{n} \int_{(B_i \cap A)} f(\phi_i(x))|\phi_i'(x)| dm, \end{aligned}$$

where we have used the change of variable formula for each i. Now

$$\begin{aligned} \int_A \mathcal{P}_\tau f dm &= \sum_{i=1}^{n} \int_A f(\phi_i(x))|\phi_i'(x)|\chi_{B_i}(x) dm \\ &= \int_A \sum_{i=1}^{n} \frac{f(\tau_i^{-1}(x))}{|\tau'(\tau_i^{-1}(x))|}\chi_\tau(a_{i-1}, a_i)(x) dm \end{aligned}$$

Since A is arbitrary,

$$\mathcal{P}_\tau f(x) = \sum_{i=1}^{n} \frac{f(\tau_i^{-1}(x))}{|\tau'(\tau_i^{-1}(x))|}\chi_\tau(a_{i-1}, a_i)(x)$$

for any $f \in L^1$.

2.3.4 Markov Transformation and Matrix Representation of the Frobenius-Perron Operator

Markov trasformation, which theory started with [Renyi, 1956], map each interval of the partition onto a union of intervals of the partition. Of particular importance are the piecewise linear Markov transformations whose invariant densities can be computed easily since the Frobenius-Perron operator can be represented by a finite-dimensional matrix. Furthermore, the piecewise linear Markov transformations can be used to approximate the long-term behaviour of more complicated transformations. Therefore, the fixed points of Frobenius-Perron operator associated with general transformations can be approximated by the fixed points of appropriate matrices.

Definition 2.18 *Let* $I = [a, b]$ *and let* $\tau : I \to I$. *Let* $\mathcal{P}$ *be a partition of* I *given by the points* $a = a_0 < a_1 < \cdots < a_n = b$. *For* $i = 1, 2, \cdots, n$, *let* $I_i = (a_{i-1}, a_i)$ *and denote the restriction of* τ *to* I_i *by* τ_i. *If* τ_i *is a homeomorphism from* I_i *onto some connected intervals* $(a_{j(i)}, a_{k(i)})$, *then* τ *is to said to be* ***Markov***.

Definition 2.19 *Let* $\tau : I \to I$ *be a piecewise monotonic transformation and let* $\mathcal{P} = \{I_i\}_{i=1}^n$ *be a partition of* I. *We define the* ***incidence matrix*** A_τ ***induced by*** τ *and* $\mathcal{P}$ *as follows : Let* $A_\tau = (a_{ij})_{1 \leqq i,j \leqq n}$, *where*

$$a_{ij} = \begin{cases} 1, & \text{if } I_j \subset \tau(I_i), \\ 0, & \text{otherwise.} \end{cases}$$

Now, let us fix a partition $\mathcal{P}$ of I and let S denote the class of all functions that are piecewise constant on the partition $\mathcal{P}$, i.e., the step functions on $\mathcal{P}$. Thus,

$$f \in S \qquad \text{if and only if} \qquad f = \sum_{i=1}^{n} \pi_i \chi_{I_i},$$

for some constants $\pi_1, \cdots, \pi_n$. Such an f will also be represented by the column vector $\pi^f = (\pi_1, \cdots, \pi_n)^T$, where T denotes transpose.

Theorem 2.4 *Let* $\tau : I \to I$ *be a piecewise linear Markov transformation on the partition* $\mathcal{P} = \{I_i\}_{i=1}^n$. *Then there exists an* $n \times n$ *matrix* M_τ *such that* $\mathcal{P}_\tau f = M_\tau^T \pi^f$

for every $f \in S$ *and* π^f *is the column vector obtained from* f.

The matrix M_τ *is of the form* $M_\tau = (m_{ij})_{1 \leqq i,j \leqq n}$, *where*

$$m_{ij} = \frac{a_{ij}}{|\tau_i'|} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)}, \qquad 1 \leqq i, j \leqq n,$$

where $A_\tau = (a_{ij})_{1 \leqq i,j \leqq n}$ *is the incidence matrix induced by* τ *and* $\mathcal{P}$.

2.4 Absolutely Continuous Invariant Measures for Piecewise Monotonic Transformation

Let $I = [a, b] \subset \mathbb{R}$ be a bounded interval and let m denote Lebesgue measure on I. For any sequence of points $a = x_0 < x_1 < \cdots < x_n = b, \quad n \geqq 1$, we define a partition $\mathcal{P} = \{I_i = (x_{i-1}, x_i) : i = 1, 2, \cdots, n\}$ of I. The points $\{x_0, \cdots, x_n\}$ are called endpoints of $\mathcal{P}$. Sometimes we will write $\mathcal{P} = \mathcal{P}\{x_0, x_1, \cdots, x_n\}$.

Definition 2.20 *Let* $f : I \to \mathbb{R}$ *and* $\mathcal{P} = \mathcal{P}\{x_0, x_1, \cdots, x_n\}$ *be a partition of* I. *If there exists a positive number* M *such that* $\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})| \leqq M$ *for all partitions* $\mathcal{P}$, *then* f *is said to be of* ***bounded variation*** *on* $[a, b]$. *In this case* $\sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$ *is called the* ***variation*** *of* f *with respect to* $\mathcal{P}$ *and we write* $\bigvee_a^b(f, \mathcal{P}) = \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|$. *The number* $\bigvee_{[a,b]} f = \sup \bigvee_a^b(f, \mathcal{P})$ *is called the* ***total variation*** *or simply the* ***variation*** *of* f *on* I.

Lemma 2.7 *If* $f \in \mathcal{C}^1[a, b]$ *with* $|f'| > 0$, *then* f *is monotonic on* $[a, b]$.

Lemma 2.8 *Let f and g be of bounded variation on $[a,b]$. Then*

$$\bigvee_{[a,b]}(f+g) \leqq \bigvee_{[a,b]} f + \bigvee_{[a,b]} g,$$

and

$$\bigvee_{[a,b]} \sum_{k=1}^{n} f_k \leqq \sum_{k=1}^{n} \bigvee_{[a,b]} f_k.$$

Lemma 2.9

$$x \in [a,b] \Rightarrow |f(a)| + |f(b)| \leqq \bigvee_{[a,b]} f + 2|f(x)|.$$

Lemma 2.10 *Let f_i be defined on $[\alpha_i, \beta_i] \subset [a,b]$ and*

$$\chi_i(x) = \begin{cases} 1, & x \in [\alpha_i, \beta_i] \\ 0, & \textit{otherwise.} \end{cases}$$

Then for $f = \sum_{i=1}^{n} f_i \chi_i$,

$$\bigvee_{[a,b]} f \leqq \sum_{i=1}^{n} \bigvee_{[\alpha_i,\beta_i]} f_i + \sum_{i=1}^{n} (|f_i(\alpha_i)| + |f_i(\beta_i)|).$$

Lemma 2.11 *If $\bigvee_{[0,1]} f \leqq a$ and $\| f \|_1 \leqq b$, where $\| f \|_1 = \int_0^1 |f| dm$, then*

$$|f(x)| \leqq a + b, \ \ \forall x \in [a,b].$$

Lemma 2.12 *Let $f : [a,b] \to \mathbb{R}$ have a continuous derivative f' on $[a,b]$. Then*

$$\bigvee_{[a,b]} f = \int_a^b |f'(x)| dm(x).$$

Lemma 2.13 *Let $f : [a,b] \to \mathbb{R}$ be of bounded variation. Let $x, y \in [a,b]$ and $x < y$. Then*

$$|f(x)| + |f(y)| \leqq \bigvee_{[x,y]} f + \frac{2}{|y-x|} \int_x^y |f(t)|dt.$$

$\square$

Let $S \subset L^1[0,1]$ be the space of functions of bounded variation on $[0,1]$; every $f \in S$ is differentiable almost everywhere with $f' \in L^1[0,1]$, and the total variation Var(f) of f is equal to the L^1 norm of f'. Let S_0 be the set of functions in S with integral zero.

Lemma 2.14 *For all $g \in S_0$*

$$(1) \qquad \| g \|_1 \leqq \frac{1}{2}(\sup g - \inf g) \leqq \frac{1}{2}\text{Var(g)}.$$

$$(2) \qquad \| g \|_\infty \leqq \text{Var(g)}.$$

Proof. When $g = 0$ a.e., then (1) and (2) are trivial. Since g has integral zero, it must take on both positive and negative values, and thus (2) follows immediately. Now to prove (1), let A and B be the sets on which g is positive and negative respectively; then

$$\begin{aligned} &\int_A g(x)dx = -\int_B g(x)dx \\ \Rightarrow &\int_A g(x)dx + \int_B g(x)dx = 0 \end{aligned} \tag{2.3}$$

and

$$\int_I |g(x)|dx = \int_A g(x)dx - \int_B g(x)dx$$

$$\Rightarrow \quad \| g \|_1 = \int_A g(x)dx - \int_B g(x)dx. \tag{2.4}$$

From (2.3) and (2.4), we obtain

$$\int_A g(x)dx = \frac{1}{2} \| g \|_1 \qquad \text{and} \qquad -\int_B g(x)dx = \frac{1}{2} \| g \|_1 .$$

Now, for $|E| = m(E) = \int_E dx$, we have

$$\begin{aligned} \mathrm{Var(g)} \geqq \max_{\mathrm{A}}(\mathrm{g}) - \max_{\mathrm{B}}(\mathrm{g}) &= \max_A(g) + \max_B(-g) \\ &\geqq \frac{\int_A g}{|A|} + \frac{\int_B (-g)}{|B|} = \frac{\| g \|_1}{2|A|} + \frac{\| g \|_1}{2|B|} \\ &= \frac{\| g \|_1}{2}\left(\frac{1}{|A|} + \frac{1}{|B|}\right) \geqq \frac{\| g \|_1}{2}(4) = 2 \| g \|_1, \end{aligned}$$

where, since $|A| + |B| \leqq 1$ then $\frac{1}{|A|} + \frac{1}{|B|} \geqq 4$ follows from: since $\sqrt{ab} \leqq \frac{a+b}{2}$ we have $ab \leqq \frac{(a+b)^2}{4}$, and then $\frac{1}{a} + \frac{1}{b} = \frac{a+b}{ab} \geqq \frac{(a+b).4}{(a+b)^2} \geqq \frac{4}{1} = 4$. Thus, (1) holds. □

Theorem 2.5 (Helly's Selection Principle) *Let $\mathcal{B}$ be a family of functions such that $f \in \mathcal{B} \Rightarrow \bigvee_{[a,b]} f \leqq \alpha$ and $|f(x)| \leqq \beta$, for any $x \in [a,b]$. Then there exists a sequence $\{f_n\} \subset \mathcal{B}$ such that $f_n \longrightarrow f^* \ \forall x \in [a,b]$ and $f^* \in BV[a,b]$.*

Theorem 2.6 (Kakutani-Yoshida) *Let $T : X \longrightarrow X$ be a bounded linear operator from a Banach sapce X into itself. Assume that there exists $M > 0$ such that $\| T^n \| \leqq M$, $n = 1, 2, \cdots$. Furthermore, if for any $f \in A \subset X$, the sequence $\{f_n\}$, where $f_n = \frac{1}{n}\sum_{k=1}^{n} T^k f$, contains a sub-sequence $\{f_{n_k}\}$ which converges weakly in X, then for any $f \in \bar{A}$, $\frac{1}{n}\sum_{k=1}^{n} T^k f \longrightarrow f^* \in X$ (norm convergence) and $T(f^*) = f^*$.*

Recall that a set $A \subset X$ of a Banach space X is called relatively compact if every infinite subset of A contains a sequence that converges to a point of X.

Theorem 2.7 (Lasota-Yorke) *Let $0 = b_1 < b_2 < \cdots < b_n = 1$ be the partition of $[0,1]$ for which the restriction τ_i of τ to the interval (b_{i-1}, b_i) is a C^2-function $(1 \leqq i \leqq n)$ such that $\inf |\tau'| > 1$. Then for any $f \in L^1[0,1]$ the sequence $\frac{1}{n}\sum_{k=1}^{n} \mathcal{P}_\tau^k f$ is convergent in norm to $f^* \in L^1[0,1]$. The limit function has the following properties:*

$(i) f \geqq 0 \Rightarrow f^* \geqq 0$;

$(ii) \int_0^1 f^* dm = \int_0^1 f dm$;

$(iii) \mathcal{P}_\tau f^* = f^*$ and consequently $d\mu^* = f^* dm$ is invariant under τ;

$(iv) f^* \in \mathrm{BV}[0,1]$. Moreover there exists c independent choice of initial f such that $\bigvee_{[0,1]} f^* \leqq c \parallel f \parallel_1$;

$(v) \bigvee_{[0,1]} \mathcal{P}_\tau f \leqq \alpha \parallel f \parallel_1 + \beta \bigvee_{[0,1]} f$, where $\alpha = K + h^{-1}$, $\beta = 2(\inf|\tau'|) < 1$, $K = \frac{\max_{i,x} |\sigma_i'(x)|}{\min_{i,x}(\sigma_i(x))}$, $\sigma = |(\tau_i^{-1})'|$, and $h = \min_i(b_{i-1}, b_i)$.

2.5 Finite Approximation of Invariant Measures

Let $[0,1]$ be divided into n equal subintervals $I_1, \cdots, I_n$ with $I_i = [a_{i-1}, a_i]$ and $m(I_i) = \frac{1}{n} = l$, $\forall i$. We define P_{ij} as the fraction of I_i which is mapped into interval I_j by τ. Let $A_{ij} = \{x \in I_i | \tau(x) \in I_j\}$. Then, $A_{ij} = I_i \cap \tau^{-1}(I_j)$. We see that $\tau(A_{ij}) = \tau(I_i \cap \tau^{-1}(I_j)) \subset \tau(I_i) \cap I_j \subset I_j$. Therefore,

$$P_{ij} = \frac{m(A_{ij})}{m(I_i)} = \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)}.$$

Let Δ_n be the n-dimensional linear subspace of L^1 which is the finite element space g generated by $\{\chi_i\}_{i=1}^n$, where χ_i is the characteristic function of I_i, and define $P_n(\tau) : \Delta_n \longrightarrow \Delta_n$ as a linear operator such that

$$P_n(\tau)\chi_i = \sum_{j=1}^{n} P_{ij}\chi_j.$$

We shall often write P_n for $P_n(\tau)$ when no clarification is needed.

Ulam conjectured that the sequence of fixed points f_n of P_n should converge to a fixed point of $\mathcal{P}_\tau$ as $n \longrightarrow \infty$ when $\mathcal{P}_\tau$ has a unique fixed point.

We will present some important Lemmas before the Theorem which gives a positive answer to this conjecture.

Lemma 2.15 *Let* $\Delta_n^1 = \{\sum_{i=1}^n a_i\chi_i \mid a_i \geqq 0$ *and* $\sum_{i=1}^n a_i = 1\}$. *Then*

$$P_n : \Delta_n^1 \longrightarrow \Delta_n^1.$$

Proof. Let $f = \sum_{i=1}^n a_i\chi_i$ and $\sum_{i=1}^n a_i = 1$. Then $f \in \Delta_n^1$, and

$$P_n f = P_n(\sum_{i=1}^{n} a_i\chi_i) = \sum_{i=1}^{n} a_i(P_n\chi_i) = \sum_{j=1}^{n}(\sum_{i=1}^{n} a_i P_{ij})\chi_j.$$

But,

$$\sum_{i=1}^{n} P_{ij} = \sum_{i=1}^{n} \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)} = 1 \qquad \text{for all } i = 1, \cdots, n.$$

Hence,

$$\sum_{j=1}^{n}(\sum_{i=1}^{n} a_i P_{ij}) = \sum_{i=1}^{n} a_i(\sum_{j=1}^{n} P_{ij}) = \sum_{i=1}^{n} a_i = 1.$$

Therefore, $P_n f \in \Delta_n^1$. $\square$

Since $P_n(\Delta_n^1) \subset \Delta_n^1$, by the Brouwer Fixed Point Theorem 2.3, there exists a point $g_n \in \Delta_n^1$ for which $P_n g_n = g_n$. Let $f_n = n g_n$. Then $f_n \in \Delta_n$ and

$$\begin{aligned} \| f_n \| &= \| n g_n \| = n \| \sum_{i=1}^{n} a_i \chi_i \| = n \int_0^1 | \sum_{i=1}^{n} a_i \chi_i | \\ &= n \int_0^1 \sum_{i=1}^{n} a_i \chi_i = n \sum_{i=1}^{n} a_i \int_0^1 \chi_i = n \frac{1}{n} \sum_{i=1}^{n} a_i = 1. \end{aligned}$$

Definition 2.21 *For $f \in L^1$ and, for every positive integer n, we define $Q_n : L^1 \to \Delta_n$ by*

$$Q_n f = \sum_{i=1}^{n} c_i \chi_i \qquad \text{where } c_i = \frac{1}{m(I_i)} \int_{I_i} f(s) ds.$$

We see that $f \geqq 0 \Rightarrow Q_n f \geqq 0$ and that $Q_n(af + bg) = aQ_n f + bQ_n g$. Hence $Q_n f = Q_n(f^+ - f^-)$ and $|Q_n f| \leqq Q_n f^+ + Q_n f^-$.

Lemma 2.16 *For $f \in L^1$, the sequence $Q_n f$ converges in L^1 to f as $n \to \infty$.*

Proof. Since $f \in L^1$, for any $\epsilon > 0$ there exists a continuous function g such that $\| f - g \| < \frac{\epsilon}{3}$. Since g is continuous in $[0, 1]$, g is uniformly continuous. We can choose N large enough such that for $n > N$we have $|g(x_1) - g(x_2)| < \frac{\epsilon}{3}$ for all $x_1, x_2 \in I_i$, $\forall i \in \{1, \cdots, n\}$. It follows that,

$$\begin{aligned} \int_{I_i} |(Q_n g)(s) - g(s)| ds &= \int_{I_i} | \sum_{j=1}^{n} \left(\frac{1}{m(I_j)} \int_{I_j} g(t) dt \right) \chi_j(s) - g(s)| ds \\ &= \int_{I_i} | \frac{1}{m(I_i)} \int_{I_i} g(t) dt \chi_i(s) - g(s)| ds \end{aligned}$$

since $s \in I_i$. Therefore,

$$\begin{aligned}\int_{I_i} |(Q_n g)(s) - g(s)| ds &\leqq \int_{I_i} \frac{1}{m(I_i)} \int_{I_i} |g(t) - g(s)| dt ds \\ &< \frac{1}{m(I_i)} \int_{I_i} \int_{I_i} \frac{\epsilon}{3} dt ds = m(I_i)\frac{\epsilon}{3}.\end{aligned}$$

Hence,

$$\| Q_n g - g \| = \int_0^1 |Q_n g - g| = \sum_{i=1}^{n} \int_{I_i} |Q_n g - g| < \sum_{i=1}^{n} m(I_i)\frac{\epsilon}{3} = \frac{\epsilon}{3}.$$

And for $\phi \in L^1$,

$$\begin{aligned}\int_1^1 Q_n \phi &= \int_0^1 \sum_{i=1}^{n} \left(\frac{1}{m(I_i)} \int_{I_i} \phi(t) dt \right) \chi_i(s) ds \\ &= \sum_{i=1}^{n} \frac{1}{m(I_i)} \int_{I_i} \phi(t) \int_0^1 \chi_i(s) ds dt = \sum_{i=1}^{n} \frac{1}{m(I_i)} \int_{I_i} m(I_i)\phi(t) dt \\ &= \sum_{i=1}^{n} \int_{I_i} \phi(t) dt = \int_0^1 \phi.\end{aligned}$$

Then,

$$\| Q_n \phi \| \leqq \int_0^1 Q_n \phi^+ + \int_0^1 Q_n \phi^- = \int_0^1 \phi^+ + \int_0^1 \phi^- = \| \phi \| .$$

Hence,

$$\| Q_n (f - g) \| \leqq \| f - g \| .$$

Thus,

$$\begin{aligned}\| Q_n f - f \| &\leqq \| Q_n f - Q_n g \| + \| Q_n g - g \| + \| g - f \| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon.\end{aligned}$$

□

Lemma 2.17 *For $f \in \Delta_n$ we have $P_n f = Q_n \mathcal{P}_\tau f$*

Proof. By Definition 2.17, we have

$$\int_{I_j} \mathcal{P}_\tau \chi_i = \int_{\tau^{-1}(I_j))} \chi_i(s)ds.$$

Therefore,

$$\begin{aligned} Q_n(\mathcal{P}_\tau \chi_i) &= \sum_{j=1}^{n} \left[\frac{1}{m(I_j)} \int_{I_j} (\mathcal{P}_\tau \chi_i)(x)dx \right] \chi_j \\ &= \sum_{j=1}^{n} \left[\frac{1}{m(I_j)} \int_{\tau^{-1}(I_j))} \chi_i(s)ds \right] \chi_j. \end{aligned}$$

Since $m(I_i) = m(I_j) = \frac{1}{n} \quad \forall i, j$, we have,

$$Q_n(\mathcal{P}_\tau \chi_i) = \sum_{j=1}^{n} \frac{m(I_i \cap \tau^{-1}(I_j))}{m(I_i)} \chi_j = \sum_{j=1}^{n} P_{ij} \chi_j = P_n \chi_i.$$

And so for $f = \sum_{k=1}^{n} c_k \chi_k$, we have

$$\begin{aligned} Q_n \mathcal{P}_\tau f &= Q_n \mathcal{P}_\tau (\sum_{k=1}^{n} c_k \chi_k) = \sum_{k=1}^{n} c_k Q_n \mathcal{P}_\tau \chi_k \\ &= \sum_{k=1}^{n} c_k P_n \chi_k = P_n (\sum_{k=1}^{n} c_k \chi_k) = P_n f. \end{aligned}$$

□

Lemma 2.18 *For $f \in \Delta_n$, the sequence $P_n f$ converges to $\mathcal{P}_\tau f$ in L_1 as $n \to \infty$.*

Proof. By Lemma 2.17, $f \in \Delta_n, \Rightarrow P_n f = Q_n \mathcal{P}_\tau f$.

By Lemma 2.16, $Q_n \mathcal{P}_\tau f \to \mathcal{P}_\tau f$. Thus the statement. □

Lemma 2.19 *If $f \in L_1$, then $V_0^1 Q_n f \leqq V_0^1 f$.*

Proof. Let $c_i = (\frac{1}{l}) \int_{I_i} f$. Then

$$V_0^1 Q_n f = V_0^1 \left(\sum_{i=1}^{n} c_i \chi_i \right) = \sum_{i=1}^{n} (\frac{1}{l}) | \int_{I_i} f - \int_{I_{i+1}} f |.$$

For every $1 \leqq i \leqq n$, there exists m_i and M_i in $[a_{i-1}, a_i]$ such that

$$f(m_i) \leqq (\frac{1}{l}) \int_{I_i} f \leqq f(M_i).$$

For simplicity we assume $m_i \leqq M_i$ for all i, the other case being almost identical.

There are two cases to consider, first

$$\frac{1}{l} \int_{I_i} f < \frac{1}{l} \int_{I_{i+1}} f.$$

and second, the same equation with the inequality reversed. For case 1

$$\begin{aligned} |\frac{1}{l} \int_{I_{i+1}} f - \frac{1}{l} \int_{I_i} f| &\leqq |f(m_i) - f(M_{i+1})| \\ &\leqq |f(m_i) - f(M_i)| + |f(M_i) - f(m_{i+1})| + |f(m_{i+1}) - f(M_{i+1})|, \end{aligned}$$

while for case 2,

$$|\frac{1}{l} \int_{I_i} f - \frac{1}{l} \int_{I_{i+1}} f| \leqq |f(M_i) - f(m_{i+1})|.$$

Hence, in either case, we have

$$\begin{aligned} V_0^1 Q_n f &\leqq \sum_{i=1}^{n} (|f(m_i) - f(M_i)| + |f(M_i) - f(m_{i+1})| + |f(m_{i+1}) - f(M_{i+1})|) \\ &\leqq V_0^1 f. \end{aligned}$$

□

Lemma 2.20 *If τ is piecewise C^2 for partition $\{b_0, \cdots, b_n\}$ and $s = \inf|\tau'| > 2$, then $\{V_0^1 f_n\}_{n=1}^{\infty}$ is bounded, where $\mathcal{P}_n f_n = f_n$.*

Proof. By Lemma 2.17,

$$f_n = P_n f_n = Q_n \mathcal{P}_\tau f_n, \quad \forall n.$$

By Lemma 2.8,

$$V_0^1 Q_n \mathcal{P}_\tau f_n \leqq V_0^1 \mathcal{P}_\tau f_n.$$

By Theorem 2.8 (Lasota-Yorke),

$$V_0^1 \mathcal{P}_\tau f_n \leqq (K + h^{-1}) \parallel f_n \parallel + \beta V_0^1 f_n,$$

with $K = \frac{\max_{i,x}|\sigma_i'(x)|}{\min_{i,x}(\sigma_i(x))}$, $\sigma_i = |(\tau_i^{-1})'|$, $h = \min_i(b_i - b_{i-1})$ and $\beta = 2s^{-1} < 1$. Since $\parallel f_n \parallel = 1$, we have

$$V_0^1 f_n \leqq (K + h^{-1}) + \beta V_0^1 f_n.$$

Since $f_n \in \Delta_n$, $V_0^1 f_n < \infty$. Hence,

$$(1 - \beta) V_0^1 f_n \leqq K + h^{-1}$$

and

$$V_0^1 f_n \leqq \frac{K + h^{-1}}{1 - \beta}.$$

□

Theorem 2.8 *Let $\tau : [0,1] \to [0,1]$ be a piecewise C^2 function with $s = \inf |\tau'| > 2$. Suppose $\mathcal{P}_\tau$ has a unique fixed point. Then, for any positive integer n, P_n has a fixed point f_n in $\triangle_n$ with $\| f_n \| = 1$ and the sequence $\{f_n\}$ converges to the fixed point of $\mathcal{P}_\tau$.*

Proof. By Lemma 2.20 and Lemma 2.11, and by Theorem 2.5 (Helly's Selection Principle), the set $\{f_n\}$ is relatively compact. Let $\{f_{n_k}\} \subset \{f_n\}$ be a convergent subsequence and let $f = \lim_{k\to\infty} f_{n_k}$. Then,

$$\begin{aligned} \| f - \mathcal{P}_\tau f \| &\leqq \| f - f_{n_k} \| + \| f_{n_k} - Q_{n_k}\mathcal{P}_\tau f_{n_k} \| \\ &+ \| Q_{n_k}\mathcal{P}_\tau f_{n_k} - Q_{n_k}\mathcal{P}_\tau f \| + \| Q_{n_k}\mathcal{P}_\tau f - \mathcal{P}_\tau f \| . \end{aligned}$$

By Lemma 2.10,

$$\| f_{n_k} - Q_{n_k}\mathcal{P}_\tau f_{n_k} \| = \| P_{n_k} f_{n_k} - Q_{n_k}\mathcal{P}_\tau f_{n_k} \| = 0.$$

Also,

$$\| Q_{n_k}\mathcal{P}_\tau (f_{n_k} - f) \| \leqq \| Q_{n_k} \| \, \| \mathcal{P}_\tau \| \, \| f_{n_k} - f \| \to 0, \qquad \text{as } f_{n_k} \to f,$$

and by Lemma 2.9, $Q_{n_k}\mathcal{P}_\tau f \to \mathcal{P}_\tau f$. Hence $\mathcal{P}_\tau f = f$.

Any convergent subsequence of $\{f_n\}$ converges to a fixed point of $\mathcal{P}_\tau$. By assumption, $\mathcal{P}_\tau$ has a unique fixed point and so we must have $f_n \to f$. □

Chapter 3

Invariant Measures for Markov Compositions of Maps of an Interval

3.1 Definitions and Notations

Let $S_i = \{1, \cdots, r\}, i \geqq 0$, and $\Omega = \prod_{i=0}^{\infty} S_i$. We select a probability measure $\mathbb{P}$ on Ω that is invariant under the **left shift** $\sigma : \Omega \to \Omega$ (i.e., $(\sigma(\omega_i))_j = \omega_{j+1}$). The space Ω contains infinite sequences of indices for the maps $T(T_1, \cdots, T_r)$, and the shift invariant probability measure $\mathbb{P}$ governs the stationary stochastic process that generates a random index at each time step. We select a stochastic $r \times r$ matrix $\mathcal{W}$ with invariant (normalised) left eigenvector $(w_1, \cdots, w_r)$ and define a probability measure ρ on $S_i, i \geqq 0$, by $\rho(\{k\}) = w_k$. Denote $[a_0, \cdots, a_s] = \{\omega \in \Omega : \omega_t = a_0, \omega_{t+1} = a_1, \cdots, \omega_{t+s} = a_s\}$, and define $\mathbb{P}([a_0, \cdots, a_s]) = w_{a_0} \mathcal{W}_{a_0,a_1} \cdots \mathcal{W}_{a_{s-1},a_s}$, consistently extending $\mathbb{P}$ to all of Ω.

Let $I = [0, 1]$. Define the **skew product** $\tau : \Omega \times I \to \Omega \times I$ by $\tau(\omega, x) = (\sigma\omega, T_{\omega_0} x)$. We form a random dynamical system by considering the orbit $\{\mathrm{Proj_I}(\tau^{\mathrm{N}}(\omega, \mathrm{x}))\}_{\mathrm{N}=0}^{\infty}$ on I where $\omega \in \Omega$, $x \in I$, and $\tau^N(\omega, x) = (\sigma^N(\omega), T_{\omega_{N-1}} \circ \cdots \circ T_{\omega_1} \circ T_{\omega_0} x)$. By putting $x_N = \mathrm{Proj_I}(\tau^{\mathrm{N}}(\omega, \mathrm{x}))$, we see that $x_N = T_{\omega_{N-1}} \circ \cdots \circ T_{\omega_0} x$ for $N \geqq 1$, with $x_0 = x$. Thus the **orbit** x_N is defined by a random composition of mappings $T_1, \cdots, T_r$; the orbit is random in the sense that the sequence of maps $T_{\omega_{N-1}} \circ \cdots \circ T_{\omega_0}$ has probabilty $\mathbb{P}([\omega_0, \cdots, \omega_{N-1}])$ of occuring. We want to discuss the asymptotic behaviour of the orbit x_N. Here we follow: [3],[4],[7],[8],[9],[10],[11],[13],[14],[15],[16].

Definition 3.1 *We say that an interval map $T : I \to I$ is a* ***Lasota-Yorke map*** *if*

(i) there is a finite partition $0 = b_0 < b_1 < \cdots < b_q = 1$ of I such that $T|_{(b_{l-1},b_l)}$, is a $\mathcal{C}^2$ function and may be extended to a $\mathcal{C}^2$ function on $[b_{l-1}, b_l]$ for $l = 1, \cdots, q$, and

(ii) $\inf_{x \in I \setminus \{b_0, \cdots, b_q\}} |T'(x)| > 0$.

We denote the partition for the map T_k by $0 = b_0^k < b_1^k < \cdots < b_{q_k}^k = 1$.

Denote by $T_k(b_l^{k,-})$ and $T_k(b_l^{k,+})$, the values that T_k takes on either side of the break point b_l^k, $l = 1, \cdots, q_k - 1$. We define the numbers $\theta_{k,l}$, $l = 1, \cdots, q_k - 1$, as follows:

$$\theta_{k,l} = \begin{cases} 0, & \text{if } T_k(b_l^{k,-}) = 0 \text{ or } 1, \text{ and } T_k(b_l^{k,+}) = 0 \text{ or } 1, \\ 2, & \text{if } T_k(b_l^{k,-}) \neq 0 \text{ or } 1, \text{ and } T_k(b_l^{k,+}) \neq 0 \text{ or } 1, \\ 1, & \text{otherwise} \end{cases}$$

For $l = 0$ and $l = q_k$, we put $\theta_{k,l} = 0$ if $T_k(b_l^k) = 0$ or 1, and $\theta_{k,l} = 1$ otherwise.

There exists a minimal partition $0 = b_0^* < b_1^* < \cdots < b_{q^*}^* = 1$ such that for each $k = 1, \cdots, r$ and all $l = 1, \cdots, q^*$, $T_k|_{(b_{l-1}^*, b_l^*)}$, is a $\mathcal{C}^2$ function and may be extended to a $\mathcal{C}^2$ function on $[b_{l-1}^*, b_l^*]$. This number q^* will be used in the main theorem.

Definition 3.2 *We call a piecewise onto Lasota-Yorke map a* ***circle map****.*

Definition 3.3 *We call* $\mathcal{C}^{1+\text{Lip}}$ ***map*** *is such a map whose first time derivative satisfies the Lipschitz condition.*

3.2 Invariant measures of Markov compositions

We assume that I is a metric space and assume that each of the T_k is a Borel measurable mapping on I. Let $\mathcal{M}(\Omega \times I)$ be the space of Borel probability measures on $\Omega \times I$. We will define invarint measure for our random maps after the following Lemmas.

Definition 3.4 *We shall say that a probability measure* $\tilde{\mu} \in \mathcal{M}(\Omega \times I)$ *is* τ*-* ***invariant*** *if*

(i) $\tilde{\mu} \circ \tau^{-1} = \tilde{\mu}$ *and*

(ii) $\tilde{\mu}(E \times I) = \mathbb{P}(E)$ *for all measurable* $E \subset \Omega$*.*

We say that $\mu \in \mathcal{M}(I)$ is **invariant** on **average**, or simple **invariant**, if there exists a τ-invariant probability measure $\tilde{\mu}$ such that $\mu(A) = \tilde{\mu}(\Omega \times A)$ for all measurable $A \subset I$.

We seek to approximate invariant measures μ that are absolutely continuous with respect to Lebesgue measure m on I.

Definition 3.5 *Define an operator* $\widehat{\mathcal{D}^*} : C(S_0 \times I, \mathbb{R}) \to C(S_0 \times I, \mathbb{R})$ *by*

$$(\widehat{\mathcal{D}^*}g)(\omega_0, x) = \sum_{\omega_1=1}^{r} g(\omega_1, T_{\omega_0}x)\mathcal{W}_{\omega_0\omega_1}.$$

We call probability measure $\xi \in \mathcal{M}(S_0 \times I)$ $\widehat{\mathcal{D}}$-invariant if

$$\int_{S_0 \times I} g(\omega_0, x) d\xi(\omega_0, x) = \int_{S_0 \times I} (\widehat{\mathcal{D}^*}g)(\omega_0, x) d\xi(\omega_0, x) \text{ for all } g \in C(S_0 \times I, \mathbb{R}). \quad (3.1)$$

The following lemma characterises τ-invariant measures on $\Omega \times I$ in terms of $\widehat{\mathcal{D}}$-invariant measures on the simpler space $S_0 \times I$.

Lemma 3.1 *Let* $A \in \mathcal{B}(S_0 \times I)$ *(the* σ*-algebra of Borel measurable sets on* $S_0 \times I$*) and B*$\in \mathcal{B}(\Omega \times I)$*. Define the sections* $A_{\omega_0} = \{x \in I : (\omega_0, x) \in A\}$ *and* $B_\omega = \{x \in I : (\omega, x) \in B\}$*. Let* $\{\mu_{\omega_0}\}_{\omega_0=1}^{r}$ *be a collection of Borel probability measures on* I*. Define a probability measure* $\xi \in \mathcal{M}(S_0 \times I)$ *by*

$$\xi(A) = \int_{S_0} \mu_{\omega_0}(A_{\omega_0}) d\rho(\omega_0), \quad (3.2)$$

and a probability measure $\tilde{\mu} \in \mathcal{M}(\Omega \times I)$ *by*

$$\tilde{\mu}(B) = \int_{\Omega} \mu_{\omega_0}(B_\omega) d\mathbb{P}(\omega). \quad (3.3)$$

Then ξ *is* $\widehat{\mathcal{D}}$*-invariant if and only if* $\tilde{\mu}$ *is* τ*-invariant.*

Proof. Let $g : \Omega \times I \to \mathbb{R}$ be any continuous function and define

$$\widehat{g}(\omega_0, x) = \left(\int_{[\omega_0]} g(\omega, x) d\mathbb{P}(\omega) \right) / \mathbb{P}([\omega_0]),$$

where

$$[a_0] = \{\omega \in \Omega : \omega_0 = a_0\}.$$

Now,

$$\begin{aligned}
& \int_{\Omega \times I} g(\tau((\omega_0 \omega_1 \omega_2 ...), x)) d\tilde{\mu}(\omega, x) \\
= & \int_\Omega \int_I g((\omega_1 \omega_2 ...), T_{\omega_0} x) d\mu_{\omega_0}(x) d\mathbb{P}(\omega) = \int_I [\int_\Omega g((\omega_1 \omega_2 ...), T_{\omega_0} x) d\mathbb{P}(\omega)] d\mu_{\omega_0}(x) \\
= & \int_I \left[\sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \int_{[\omega_0 \omega_1]} g((\omega_1 \omega_2 ...), T_{\omega_0} x) d\mathbb{P}(\omega) \right] d\mu_{\omega_0}(x) \\
= & \int_I \left[\sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \frac{1}{\mathbb{P}([\omega_0 \omega_1])} \int_{[\omega_0 \omega_1]} g((\omega_1 \omega_2 ...), T_{\omega_0} x) d\mathbb{P}(\omega) \right] d\mu_{\omega_0}(x) \mathbb{P}([\omega_0 \omega_1]) \\
= & \int_I \left[\sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \widehat{g}(\omega_1, T_{\omega_0} x) \right] d\mu_{\omega_0}(x) \mathbb{P}([\omega_0 \omega_1]) \\
= & \int_I \sum_{\omega_0=1}^{r} \sum_{\omega_1=1}^{r} \widehat{g}(\omega_1, T_{\omega_0} x) d\mu_{\omega_0}(x) w_{\omega_0} \mathcal{W}_{\omega_0 \omega_1} \\
= & \int_I \sum_{\omega_0=1}^{r} \left(\sum_{\omega_1=1}^{r} \widehat{g}(\omega_1, T_{\omega_0} x) \mathcal{W}_{\omega_0 \omega_1} \right) d\mu_{\omega_0}(x) w_{\omega_0} = \sum_{\omega_0=1}^{r} \int_I (\widehat{\mathcal{D}}^* \widehat{g})(\omega_0, x) d\mu_{\omega_0}(x) w_{\omega_0} \\
= & \int_{S_0} \int_I (\widehat{\mathcal{D}}^* \widehat{g})(\omega_0, x) d\mu_{\omega_0}(x) d\rho(\omega_0) = \int_{S_0 \times I} (\widehat{\mathcal{D}}^* \widehat{g})(\omega_0, x) d\xi(\omega_0, x)
\end{aligned}$$

Since $\int_{\Omega \times I} g(\tau((\omega_0 \omega_1 \omega_2 ...), x)) d\tilde{\mu}(\omega, x)$ equals to $\int_{\Omega \times I} g(\omega, x) d\tilde{\mu}(\omega, x)$ iff $\tilde{\mu}$ is τ-invariant and $\int_{S_0 \times I} (\widehat{\mathcal{D}}^* \widehat{g})(\omega_0, x) d\xi(\omega_0, x)$ is equal to $\int_{S_0 \times I} \widehat{g}(\omega_0, x) d\xi(\omega_0, x)$ iff ξ is $\widehat{\mathcal{D}}$-invariant. Thus, ξ is $\widehat{\mathcal{D}}$-invariant iff $\tilde{\mu}$ is τ-invariant. □

Lemma 3.2 *Let $\{\mu_k\}_{k=1}^r$ be a familly of Borel probability measures on I. Define the section $B_\omega = \{x \in I : (\omega, x) \in B\}$ where $B \in \mathcal{B}(\Omega \times I)$ is a Borel measurable subset of $\Omega \times I$. A measure $\tilde{\mu}$ defined by*

$$\tilde{\mu}(B) = \int_\Omega \mu_{\omega_0}(B_\omega) d\mathbb{P}(\omega), \qquad \textit{for all} \qquad B \in \mathcal{B}(\Omega \times I) \tag{3.4}$$

is τ-invariant iff the family of measures $\{\mu_k\}_{k=1}^r$ is fixed under the transformation

$$(\nu_1, ..., \nu_r) \mapsto \left(\sum_{k=1}^r \nu_k \circ T_k^{-1} \mathcal{W}_{1k}^*, ..., \sum_{k=1}^r \nu_k \circ T_k^{-1} \mathcal{W}_{rk}^* \right), \qquad \nu_k \in \mathcal{M}(I) \tag{3.5}$$

where $\mathcal{W}_{lk}^* = \mathcal{W}_{kl} w_k / w_l$ is the transition matrix for the reversed Markov chain.

Proof. We show that the measure ξ in (3.2) is $\widehat{\mathcal{D}}$-invariant iff the family $\{\mu_{\omega_0}\}_{\omega 0=1}^r$ is fixed under the transformation (3.5). The result will then follow from Lemma 3.1. Suppose that ξ is $\widehat{\mathcal{D}}$-invariant, and choose $g(\omega_0, x) = \chi_{\{j\}\times A}(\omega_0, x)$ for some $j \in S_1$ and $A \in \mathcal{B}(I)$. On one hand, we have:

$$\begin{aligned} & \int_{S_1 \times I} g(\omega_1, x) d\xi(\omega_1, x) \\ = & \int_{S_1 \times I} \chi_{\{j\}\times A}(\omega_1, x) d\xi(\omega_1, x) = \int_{S_1} \int_I \chi_{\{j\}\times A}(\omega_1, x) d\mu_{\omega_1}(x) d\rho(\omega_1) \\ = & \left(\int_I \chi_{\{j\}\times A}(\omega_1, x) d\mu_{\omega_1}(x) \right) \left(\int_{\{j\}} d\rho(\omega_1) \right) = \mu_j(A) w_j. \end{aligned} \tag{3.6}$$

On the other hand, we have

$$
\begin{aligned}
&\int_{S_0\times I}(\widehat{\mathcal{D}}^*\widehat{g})(\omega_0,x)d\xi(\omega_0,x)\\
=\;& \int_{S_0\times I}\widehat{\mathcal{D}}^*\chi_{\{j\}\times A}(\omega_0,x)d\xi(\omega_0,x)=\int_{S_0}\int_I\sum_{\omega_1=1}^{r}\chi_{\{j\}\times A}(\omega_1,T_{\omega_0}x)\mathcal{W}_{\omega_0\omega_1}d\xi(\omega_0,x)\\
=\;& \int_{S_0}\int_I\sum_{\omega_1=1}^{r}\chi_{\{j\}\times A}(\omega_1,T_{\omega_0}x)\mathcal{W}_{\omega_0\omega_1}d\mu_{\omega_0}(x)d\rho(\omega_0)\\
=\;& \sum_{\omega_0=1}^{r}\int_I\sum_{\omega_1=1}^{r}\chi_{\{j\}\times A}(\omega_1,T_{\omega_0}x)\mathcal{W}_{\omega_0\omega_1}d\mu_{\omega_0}(x)w_{\omega_0}\\
=\;& \sum_{\omega_0=1}^{r}\left(\int_I\chi_{\{j\}\times A}(\omega_1,T_{\omega_0}x)d\mu_{\omega_0}(x)\right)\mathcal{W}_{\omega_0 j}w_{\omega_0}=\sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}_{\omega_0 j}w_{\omega_0}\\
=\;& \sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}^*_{j\omega_0}w_j
\end{aligned}
\tag{3.7}
$$

Now from (3.6) and (3.7), we have

$$
\begin{aligned}
\mu_j(A)w_j &= \sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}^*_{j\omega_0}w_j\\
\Rightarrow \mu_j(A) &= \sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}^*_{j\omega_0}.
\end{aligned}
$$

Since any continuous function can be approximated by the limit of simple functions, property (3.1) also holds for continuous function and thus the above result is true for all $g(\omega_0,x)$.

Thus $\{\mu_{\omega_0}\}_{\omega 0=1}^{r}$ is fixed under the transformation (3.5).

For the converse, suppose that the family $\{\mu_{\omega_0}\}_{\omega 0=1}^{r}$ is fixed under the transformation (3.5) and consider $g \in \mathcal{C}(S_0 \times I, \mathbb{R})$. Now choose $g(\omega_0, x) = \chi_{\{j\}\times A}(\omega_0, x)$. Then,

$$\begin{aligned}\int_{S_0\times I}(\widehat{\mathcal{D}}^*\widehat{g})(\omega_0,x)d\xi(\omega_0,x) &= \int_{S_0\times I}\widehat{\mathcal{D}}^*\chi_{\{j\}\times A}(\omega_0,x)d\xi(\omega_0,x)\\ &= \sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}_{\omega_0 j}w_{\omega_0} = \sum_{\omega_0=1}^{r}\mu_{\omega_0}(T_{\omega_0}^{-1}A)\mathcal{W}_{j\omega_0}^{*}w_j\\ &= \mu_j(A)w_j.\end{aligned}$$

Again,

$$\begin{aligned}\int_{S_0\times I} g(\omega_0,x)d\xi(\omega_0,x) &= \int_{S_0\times I}\chi_{\{j\}\times A}(\omega_0,x)d\xi(\omega_0,x)\\ &= \int_{S_0}\int_{I}\chi_{\{j\}\times A}(\omega_0,x)d\mu_{\omega_0}(x)d\rho(\omega_0)\\ &= \left(\int_{I}\chi_{\{j\}\times A}(\omega_0,x)d\mu_{\omega_0}(x)\right)\left(\int_{\{j\}}d\rho(\omega_0)\right) = \mu_j(A)w_j.\end{aligned}$$

Thus,

$$\int_{S_0\times I}\widehat{\mathcal{D}}^*\chi_{\{j\}\times A}(\omega_0,x)d\xi(\omega_0,x) = \int_{S_0\times I}\chi_{\{j\}\times A}(\omega_0,x)d\xi(\omega_0,x),$$

i.e., the above equation is true for any simple function and since any continuous function can be approximated by the limit of simple functions, so for any continuous $g \in \mathcal{C}(S_0 \times I, \mathbb{R})$

$$\int_{S_0\times I}(\widehat{\mathcal{D}}^*\widehat{g})(\omega_0,x)d\xi(\omega_0,x) = \int_{S_0\times I} g(\omega_0,x)d\xi(\omega_0,x).$$

Thus $\{\mu_{\omega_0}\}_{\omega 0=1}^{r}$ being fixed by (3.5) implies $\widehat{\mathcal{D}}$-invariance of ξ. The fact that $\mu = \sum_{k=1}^{r} w_k\mu_k$ is invariant follows immediately from Definition 3.3 □

Definition 3.6 *Any probability measure on I of the form:*

$$\mu = \sum_{k=1}^{r} w_k \mu_k \tag{3.8}$$

that arises as a projection onto $\mathcal{M}(I)$ of a measure of the form (3.4), is ***invariant on average*** *iff the μ_k, $k = 1, \cdots, r$ are fixed under the transformation (3.5).*

The following Lemma says that all absolutely continuous (with respect to $\mathbb{P} \times m$) τ invariant measures may be written in the simple form (3.4):

Lemma 3.3 *(Kowalski[7]): Assume that each T_k, $k = 1, \cdots, r$ is non-singular with respect to m. Then $\mathbb{P} \times m$ absolutely continuous τ-invariant measure may be written in the form (3.4).*

Thus, finding an absolutely continuous probability measure μ of the form (3.8) with the $\{\mu_k\}_{k=1}^{r}$ being fixed under the action of (3.5) is the only way to construct an absolutely continuous τ-invariant measure on $\Omega \times I$.

3.3 Frobenius-Perron operator and fundamental results for Frobenius-Perron operator

We denote the density of ν_j with respect to Lebesgue measure (as in (3.5)) by $f^{(j)}$. Let $\widehat{BV} = \prod_{i=1}^{r} BV$ denote the r-fold product of the space of functions of bounded variation. We endow the space $\widehat{BV}$ with the norm:

$$\| (f^{(1)}, \cdots, f^{(r)}) \| = \max_{1 \leqq k \leqq r} \| f^{(k)} \| = \max_{1 \leqq k \leqq r} \left\{ \max\{\mathrm{Var}(\mathrm{f}^{(k)}), \| \mathrm{f}^{(k)} \|_1\} \right\}.$$

Denote by $\mathcal{P}_k : BV \to BV$, the standard **Perron-Frobenious operator** for the map T_k. Following (3.5), we define an operator $\widehat{\mathcal{P}} : \widehat{BV} \to \widehat{BV}$ by

$$\widehat{\mathcal{P}}(f^{(1)}, \cdots, f^{(r)}) = (\sum_{k=1}^{r} \mathcal{W}^*_{1k}\mathcal{P}_k f^{(k)}, \sum_{k=1}^{r} \mathcal{W}^*_{2k}\mathcal{P}_k f^{(k)}, \cdots, \sum_{k=1}^{r} \mathcal{W}^*_{rk}\mathcal{P}_k f^{(k)}) \tag{3.9}$$

By Lemma 3.2, we may construct an absolutely continuous invariant probability measure μ from a collection $(h^{(1)}, \cdots, h^{(r)})$ of densities that is fixed by $\widehat{\mathcal{P}}$. We will call the density of μ, $h = \sum_{k=1}^{r} w_k h^{(k)}$ an invariant probability density for our Markov random compositions.

The following are the fundamental inequalities for Frobenius-Perron operator:

Lemma 3.4 *Let $\widehat{f} = (f^{(1)}, f^{(2)}, ..., f^{(r)}) \in \widehat{BV}$. Suppose that each T_k, $k = 1, 2, ..., r$, is a Lasota-Yorke map, and set q^* as in Definition 3.1.*
Set $\widehat{BV_0} = \{\widehat{f} \in \widehat{BV} : \int f^{(k)} dm = 0 \quad for\ all\ k = 1, 2, ..., r\}$
Define

$$\begin{aligned}
\alpha'_l &:= \sum_{k=1}^{r} W^*_{lk} \frac{1}{\inf_{x\in I} |T'_k(x)|}; \\
\beta'_l &:= \sum_{k=1}^{r} W^*_{lk} \frac{\sup_{x\in I} |T''_k(x)|}{\inf_{x\in I} |T'_k(x)|^2}; \\
\eta'_l &:= \sum_{k=1}^{r} W^*_{lk} \frac{\sum_{l=0}^{q_k} \theta_{k,l}}{\inf_{x\in I} |T'_k(x)|},
\end{aligned}$$

with $\alpha' = \max_{1\leqq l\leqq r} \alpha'_l$ and $\beta' = \max_{1\leqq l\leqq r} \beta'_l$. Then

$$\begin{aligned}
(i) \max_{1\leqq k\leqq r} \mathrm{Var}(\widehat{\mathcal{P}}\widehat{\mathrm{f}})^{(\mathrm{k})} &\leqq 2\alpha' \max_{1\leqq k\leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) \\
&+ \max_{1\leqq l\leqq r} (2q^*\alpha'_l + \beta'_l) \max_{1\leqq k\leqq r} \|f^{(k)}\|_1, \widehat{f} \in \widehat{BV};
\end{aligned}$$

$$(ii)\ \|\widehat{\mathcal{P}}\widehat{f}\| \leqq (\max_{1\leqq l\leqq r}(\alpha_l' + \eta_l') + \beta'/2)\|\widehat{f}\| \qquad for\ \widehat{f} \in \widehat{BV_0}.$$

If in addition, if each T_K is a $\mathcal{C}^2$ circle map,

$$(iii)\ \|\widehat{\mathcal{P}}\widehat{f}\| \leqq (\alpha' + \beta'/2)\|\widehat{f}\| \qquad for\ \widehat{f} \in \widehat{BV_0}.$$

Proof. Let $B_k := \{[b_0^k, b_1^k], ..., [b_{q_k-1}^k, b_{q_k}^k]\}$ and $B^* := \{[b_0^*, b_1^*], ..., [b_{q_k-1}^*, b_{q_k}^*]\}$ be as in Definition 3.1. We have

$$\widehat{\mathcal{P}}(f^{(1)}, f^{(2)}, ..., f^{(r)}) = (\sum_{k=1}^{r} W_{1k}^*\mathcal{P}_k f^{(k)}, \sum_{k=1}^{r} W_{2k}^*\mathcal{P}_k f^{(k)}, ..., \sum_{k=1}^{r} W_{rk}^*\mathcal{P}_k f^{(k)}),$$

so

$$(\widehat{\mathcal{P}}\widehat{f})^{(l)} = \sum_{k=1}^{r} W_{lk}^*\mathcal{P}_k f^{(k)},$$

where $\mathcal{P}_k : L^1(I, m) \to L^1(I, m)$ denotes the standard Perron-Frobenious operator for the map T_k, namely

$$\mathcal{P}_k f^{(k)}(x) = \sum_{l=1}^{r} f^{(k)}(\psi_l(x))\sigma_l(x)\chi_{H_l}(x),$$

where $H_l = T_k(B_l),\ B_l \in B^*,\ \psi_l = (T_k|_{B_l})^{-1},\ \sigma_l = |\psi_l'|$ and χ_{H_l} is the characteristic function of the set $H_l = T_k([b_{l^*-1}^*, b_{l_*}^*]) = T_k(B_l)$. Thus,

$$\mathrm{Var}(\widehat{\mathcal{P}}\widehat{\mathrm{f}})^{(\mathrm{l})} = \mathrm{Var}(\sum_{\mathrm{k}=1}^{\mathrm{r}} \mathrm{W}_{\mathrm{lk}}^*\mathcal{P}_{\mathrm{k}}\mathrm{f}^{(\mathrm{k})}) \leqq \sum_{\mathrm{k}=1}^{\mathrm{r}} \mathrm{W}_{\mathrm{lk}}^*\mathrm{Var}(\mathcal{P}_{\mathrm{k}}\mathrm{f}^{(\mathrm{k})}). \tag{3.10}$$

and we proceed to bound $\mathrm{Var}(\mathcal{P}_{\mathrm{k}}\mathrm{f}^{(\mathrm{k})}),\ \mathrm{k} = 1, 2, ..., \mathrm{r}$, individually.

$$\begin{aligned}
\mathrm{Var}(\mathcal{P}_{\mathrm{k}}\mathrm{f}^{(\mathrm{k})}) &= \mathrm{Var}\sum_{\mathrm{l}=1}^{\mathrm{r}} \mathrm{f}^{(\mathrm{k})}(\psi_{\mathrm{l}}(\mathrm{x}))\sigma_{\mathrm{l}}(\mathrm{x})\chi_{\mathrm{H_l}}(\mathrm{x}) \leqq \sum_{\mathrm{B_l}\in\mathrm{B}^*} \mathrm{Var}_{\mathrm{H_l}}(\mathrm{f}^{(\mathrm{k})})(\psi_{\mathrm{l}}(\mathrm{x}))\sigma_{\mathrm{l}}(\mathrm{x}) \\
&+ \sum_{l=1}^{q^*} \left(|f^{(k)}(b_{l-1}^*)|\sigma_k(b_{l-1}^*) + |f^{(k)}(b_l^*)|\sigma_k(b_l^*)\right) \quad \text{(by Lemma 2.10)} \\
&= \sum_{B_l\in B^*} \mathrm{Var}_{\mathrm{H_l}}(\mathrm{f}^{(\mathrm{k})})(\psi_{\mathrm{l}}(\mathrm{x}))\sigma_{\mathrm{l}}(\mathrm{x}) + \sum_{\mathrm{l}=1}^{\mathrm{q}^*} \left(|\frac{\mathrm{f}^{(\mathrm{k})}(\mathrm{b}_{\mathrm{l}-1}^*)}{\mathrm{T}_{\mathrm{k}}'(\mathrm{b}_{\mathrm{l}-1}^*)}| + |\frac{\mathrm{f}^{(\mathrm{k})}(\mathrm{b}_{\mathrm{l}}^*)}{\mathrm{T}_{\mathrm{k}}'(\mathrm{b}_{\mathrm{l}}^*)}|\right)
\end{aligned} \tag{3.11}$$

1st term of (3.11):

$$\mathrm{Var_{H_l}(f^{(k)})(\psi_l(x))\sigma_l(x)} = \int_{\mathrm{H_l}} \mathrm{|d(f^{(k)} \circ \psi_l(x))\sigma_l(x)|} \text{ (by Lemma 2.12)}$$

$$\leqq \int_{H_l} |f^{(k)} \circ \psi_l(x)||\sigma_l'(x)|dm + \int_{H_l} |\sigma_l(x)||d(f^{(k)} \circ \psi_l(x))|. \qquad (3.12)$$

Here

$$\psi_l = (T_k|_{B_l})^{-1} \Rightarrow T_k(\psi_l(x)) = x$$

$$\Rightarrow \quad T_k'(\psi_l(x))\psi_l'(x) = 1 \Rightarrow \psi_l'(x) = \sigma_l = \frac{1}{T_k'(\psi_l(x))}$$

$$\Rightarrow \quad \sigma_l' = -\frac{T_k''(\psi_l(x))\psi_l'(x)}{(T_k'(\psi_l(x)))^2}.$$

Now we consider the 1st term of (3.12):

$$\int_{H_l} \frac{|f^{(k)} \circ \psi_l(x)||T_k''(\psi_l(x))||\psi_l'(x)|}{|T_k'(\psi_l(x))|^2} dx.$$

Changing the variables we obtain:

$$\int_{B_l} \frac{|f^{(k)}(x)||T_k''(x)|\frac{1}{|T_k'(x)|}T_k'(x)}{|T_k'(x)|^2} dx = \int_{B_l} |f^{(k)}(x)|\frac{|T_k''(x)|}{|T_k'(x)|^2} dm.$$

Also changing the variable in the 2nd term of (3.12) we obtain:

$$\int_{H_l} |\sigma_l(x)||d(f^{(k)} \circ \psi_l(x))| = \int_{B_l} \frac{1}{|T_k'(x)|}|df^{(k)}(x)|.$$

Thus from (3.12) we obtain:

$$\begin{aligned}
\mathrm{Var_{H_l}(f^{(k)})(\psi_l(x))\sigma_l(x)} &\leqq \int_{B_l} |f^{(k)}(x)|\frac{|T_k''(x)|}{|T_k'(x)|^2} dm + \int_{B_l} \frac{1}{|T_k'(x)|}|df^{(k)}(x)| \\
&\leqq \frac{\sup_{B_l} |T''(x)|}{\inf_{B_l} |T_k'(x)|^2} \int_{B_l} |f^{(k)}(x)|dm + \frac{1}{\inf_{B_l} |T_k'(x)|} \int_{B_l} |df^{(k)}(x)| \\
&= \frac{\sup_{B_l} |T''(x)|}{\inf_{B_l} |T_k'(x)|^2} \int_{B_l} |f^{(k)}(x)|dm \\
&+ \frac{1}{\inf_{B_l} |T_k'(x)|} \mathrm{Var_{B_l}(f^{(k)})}. \text{ (by Lemma 2.12)}
\end{aligned}$$

Thus,

$$\begin{aligned}
&\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\sum_{B_l\in B^*}\mathrm{Var_{H_l}(f^{(k)})(\psi_l(x))\sigma_l(x)}\\
\leqq\;&\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\sum_{B_l\in B^*}\frac{\sup_{B_l}|T''(x)|}{\inf_{B_l}|T_k'(x)|^2}\int_{B_l}|f^{(k)}(x)|dm+\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\sum_{B_l\in B^*}\frac{1}{\inf_{B_l}|T_k'(x)|}\mathrm{Var_{B_l}(f^{(k)})}\\
\leqq\;&\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\sum_{B_l\in B^*}\frac{\sup_{B_l}|T''(x)|}{\inf_{B_l}|T_k'(x)|^2}\|f^{(k)}(x)\|_1+\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\frac{1}{\inf_{B_l}|T_k'(x)|}\mathrm{Var(f^{(k)})}. \qquad (3.13)
\end{aligned}$$

Now for the 2nd term of (3.11):

$$\begin{aligned}
\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\;\sum_{l=1}^{q^*}\;\left(\left|\frac{f^{(k)}(b_{l-1}^*)}{T_k'(b_{l-1}^*)}\right|+\left|\frac{f^{(k)}(b_l^*)}{T_k'(b_l^*)}\right|\right)&\leqq\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\frac{1}{\inf_{x\in I}|T_k'(x)|}\sum_{l=1}^{q^*}|f^k(b_{l-1}^*)|+|f^{(k)}(b_l^*)|\\
&\leqq\;\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\frac{1}{\inf_{x\in I}|T_k'(x)|}(\mathrm{Var(f^{(k)})}+2\mathrm{q}^*\|\mathrm{f^{(k)}}\|_1)\text{ (by Lemma 2.13).} \qquad (3.14)
\end{aligned}$$

Thus, from (3.10) by considering (3.13) and (3.14) above, we obtain

$$\begin{aligned}
\mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(l)}\;&\leqq\;2\left(\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\frac{1}{\inf_{x\in I}|T_k'(x)|}\right)\max_{1\leqq k\leqq r}\mathrm{Var(f^{(k)})}\\
&+\;\left(\sum_{k=1}^{r}\mathcal{W}_{lk}^{*}\left(\sum_{B_l\in B^*}\frac{\sup_{B_l}|T''(x)|}{\inf_{B_l}|T_k'(x)|^2}\|f^{(k)}(x)\|_1+\frac{2q^*}{\inf_{x\in I}|T_k'(x)|}\right)\right)\max_{1\leqq k\leqq r}\|f^{(k)}\|_1\\
&=\;2\alpha_l'\max_{1\leqq k\leqq r}\mathrm{Var(f^{(k)})}+(2\mathrm{q}^*\alpha_l'+\beta_l')\max_{1\leqq \mathrm{k}\leqq \mathrm{r}}\|\mathrm{f^{(k)}}\|_1,\qquad \text{for } \widehat{\mathrm{f}}\in\widehat{\mathrm{BV}}.
\end{aligned}$$

Thus,

$$\max_{1\leqq l\leqq r}\mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(l)}\leqq 2\alpha'\max_{1\leqq \mathrm{k}\leqq \mathrm{r}}\mathrm{Var(f^{(k)})}+\max_{1\leqq \mathrm{l}\leqq \mathrm{r}}(2\mathrm{q}^*\alpha_l'+\beta_l')\max_{1\leqq \mathrm{k}\leqq \mathrm{r}}\|\mathrm{f^{(k)}}\|_1,$$

for $\widehat{f}\in\widehat{BV}$ as required.

To prove Theorem 3.1(ii) we need to reduce the combination of both the coefficients of $\mathrm{Var}(\mathrm{f}^{(\mathrm{k})})$ and $\|f\|_1$. We use a modification of the inequality (3.11):

$$\begin{aligned}
\mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} &= \mathrm{Var}(\sum_{\mathrm{k}=1}^{\mathrm{r}} \mathcal{W}^*_{\mathrm{lk}} \mathcal{P}_{\mathrm{k}} \mathrm{f}^{(\mathrm{k})}) \\
&\leqq \sum_{k=1}^{r} \mathcal{W}^*_{lk} \sum_{B_l \in B^k} \mathrm{Var}_{\mathrm{H}_\mathrm{l}}(\mathrm{f}^{(\mathrm{k})})(\psi_\mathrm{l}(\mathrm{x}))\sigma_\mathrm{l}(\mathrm{x}) \\
&+ \sum_{l=0}^{q_k} \theta_{k,l} \max\left\{ \left|\frac{f^{(k)}(b_l^{k,-})}{T'_k(b_l^{k,-})}\right|, \left|\frac{f^{(k)}(b_l^{k,+})}{T'_k(b_l^{k,+})}\right| \right\}. \qquad (3.15)
\end{aligned}$$

The first term is bounded as before. From the second term we obtain:

$$\begin{aligned}
&\sum_{k=1}^{r} \mathcal{W}^*_{lk} \sum_{l=0}^{q_k} \theta_{k,l} \max\left\{ \left|\frac{f^{(k)}(b_l^{k,-})}{T'_k(b_l^{k,-})}\right|, \left|\frac{f^{(k)}(b_l^{k,+})}{T'_k(b_l^{k,+})}\right| \right\} \\
\leqq\ & \left(\sum_{k=1}^{r} \mathcal{W}^*_{lk} \sum_{l=0}^{q_k} \frac{\theta_{k,l}}{min\{|T'_k(b_l^{k,-})|, |T'_k(b_l^{k,+})|\}} \right) \max_{1\leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}), \\
&\text{where we use } \|f^{(k)}\|_\infty \leqq \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) \text{ as } \widehat{\mathrm{f}} \in \widehat{\mathrm{BV}_0} \text{ (by Lemma 2.14)} \\
\leqq\ & \left(\sum_{k=1}^{r} \mathcal{W}^*_{lk} \frac{\sum_{l=0}^{q_k} \theta_{k,l}}{\inf_{x\in I} |T'_k(x)|} \right) \max_{1\leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}). \qquad (3.16)
\end{aligned}$$

Thus, from (3.15):

$$\begin{aligned}
\mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} &\leqq \left(\sum_{k=1}^{r} \mathcal{W}^*_{lk} \frac{1+\sum_{l=0}^{q_k} \theta_{k,l}}{\inf_{x\in I} |T'_k(x)|} \right) \max_{1\leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) \\
&+ \left(\sum_{k=1}^{r} \mathcal{W}^*_{lk} \sum_{B_l \in B^*} \frac{\sup_{B_l} |T''(x)|}{\inf_{B_l} |T'_k(x)|^2} \right) \max_{1\leqq k \leqq r} \|f^{(k)}(x)\|_1 \\
&= (\alpha'_l + \eta'_l) \max_{1\leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) + \beta'_\mathrm{l} \max_{1\leqq \mathrm{k} \leqq \mathrm{r}} \|\mathrm{f}^{(\mathrm{k})}\|_1.
\end{aligned}$$

So,

$$\max_{1\leqq k \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} \leqq \max_{1\leqq \mathrm{k} \leqq \mathrm{r}} (\alpha'_\mathrm{l} + \eta'_\mathrm{l}) \max_{1\leqq \mathrm{k} \leqq \mathrm{r}} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) + \beta' \max_{1\leqq \mathrm{k} \leqq \mathrm{r}} \|\mathrm{f}^{(\mathrm{k})}\|_1.$$

For $\widehat{f} \in \widehat{BV_0}$, we have $\|f^{(k)}\|_1 \leqq \frac{1}{2}\mathrm{Var}(\mathrm{f}^{(\mathrm{k})})$ by Lemma 2.14. So

$$\|\widehat{\mathcal{P}}\widehat{f}\| = \max_{1 \leqq l \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} \text{ and } \|\widehat{\mathrm{f}}\| = \max_{1 \leqq \mathrm{k} \leqq \mathrm{r}} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}).$$

Now,

$$\begin{aligned}
\|\widehat{\mathcal{P}}\widehat{f}\| &= \max\left\{\max_{1 \leqq l \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})}, \|\widehat{\mathcal{P}\mathrm{f}}\|_1\right\} = \max_{1 \leqq l \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} \\
&\leqq \max_{1 \leqq l \leqq r}(\alpha_l' + \eta_l') \max_{1 \leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) + \frac{\beta'}{2} \max_{1 \leqq \mathrm{k} \leqq \mathrm{r}} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) \\
&= \left(\max_{1 \leqq l \leqq r}(\alpha_l' + \eta_l') + \frac{\beta'}{2}\right) \max_{1 \leqq k \leqq r} f^{(k)} \\
&= \left(\max_{1 \leqq l \leqq r}(\alpha_l' + \eta_l') + \frac{\beta'}{2}\right) \|\widehat{f}\|, \qquad \text{for } \widehat{f} \in \widehat{BV_0}.
\end{aligned}$$

as required.

To prove part (iii), since each T_k is $\mathcal{C}^2$ circle map, we use the bound of part (ii), and delete the contributions from the branches of monotonicity not being onto (the second term in the preceding argument). This leaves us with

$$\max_{1 \leqq l \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} \leqq \max_{1 \leqq \mathrm{l} \leqq \mathrm{r}} \alpha_{\mathrm{l}}' \max_{1 \leqq \mathrm{k} \leqq \mathrm{r}} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) + \max_{1 \leqq \mathrm{l} \leqq \mathrm{r}} \beta_{\mathrm{l}}' \max_{1 \leqq \mathrm{k} \leqq \mathrm{r}} \|\mathrm{f}^{(\mathrm{k})}\|_1,$$

and so

$$\begin{aligned}
\|\widehat{\mathcal{P}}\widehat{f}\| &= \max_{1 \leqq l \leqq r} \mathrm{Var}(\widehat{\mathcal{P}\mathrm{f}})^{(\mathrm{l})} \\
&\leqq \max_{1 \leqq l \leqq r} \alpha_l' \max_{1 \leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) + \max_{1 \leqq \mathrm{l} \leqq \mathrm{r}} \beta_{\mathrm{l}}' \frac{1}{2} \max_{1 \leqq \mathrm{k} \leqq \mathrm{r}} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) \\
&= (\alpha' + \frac{\beta'}{2}) \max_{1 \leqq k \leqq r} \mathrm{Var}(\mathrm{f}^{(\mathrm{k})}) = (\alpha' + \frac{\beta'}{2})\|\widehat{\mathrm{f}}\|, \quad \text{for } \widehat{\mathrm{f}} \in \widehat{\mathrm{BV}_0}
\end{aligned}$$

as required. □

3.4 Approximation of ACIM for Markov compositions

Here we discuss the approximation of absolutely continuous invariant measures for Markov random compositions by means of convergence of Ulam's finite approximation scheme. First part of the Theorem assures the existence and then it shows the approximation of absolutely continuous invariant measures. Second and third part give us the bounds.

Theorem 3.1 *Let $\{T_1, \cdots, T_r\}$ be a collection of Lasota-Yorke maps, and assume that the Markov composition has a unique invariant density h. Equipartition the unit interval into n subintervals $I_i = [(i-1)/n, i/n]$, $i = 1, \cdots, n$ and define r stochastic matrices $P_n(k)$, $k = 1, \cdots, r$, by*

$$P_{n,ij}(k) = \frac{m(I_i \cap T_k^{-1} I_j)}{m(I_i)}.$$

Further, define the $rn \times rn$ matrix

$$S_n = \begin{pmatrix} \mathcal{W}_{11}^* P_n(1) & \mathcal{W}_{21}^* P_n(1) & \cdots & \mathcal{W}_{r1}^* P_n(1) \\ \mathcal{W}_{12}^* P_n(2) & \mathcal{W}_{22}^* P_n(2) & \cdots & \mathcal{W}_{r2}^* P_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{W}_{1r}^* P_n(r) & \mathcal{W}_{2r}^* P_n(r) & \cdots & \mathcal{W}_{rr}^* P_n(r) \end{pmatrix},$$

and let $s_n = [s_n^{(1)} | s_n^{(2)} | \cdots | s_n^{(r)}]$ be a fixed left eigenvector of S_n, where each $s_n^{(k)}$, $k = 1, \cdots, r$, is a vector of length n satisfying $\sum_{i=1}^{n} s_{n,i}^{(k)} = 1$. Define the approximate *invariant density*

$$h_n = \sum_{i=1}^{n} \left(\frac{\sum_{k=1}^{r} w_k s_{n,i}^{(k)}}{m(I_i)} \right) \chi_{I_i}.$$

Then,

(i) If $\alpha' < 1/2$, $\|h_n - h\|_1 \to 0$ *as* $n \to \infty$;

(ii) If $\max_{1 \leqq l \leqq r}(\alpha'_l + \eta'_l) + \beta'/2 < 1$, *and the endpoints of the partition* $\{I_1, \cdots, I_n\}$ *contain all points where there is a break in the* C^1 *behaviour of any* $h^{(k)}$ *(densities),* $k = 1, \cdots, r$, *then there exists a constant* $C < \infty$ *such that* $\|h_n - h\|_1 \leqq C \log\ n/n$.

(iii) If $\alpha' + \beta'/2 < 1$ *and each* T_k *is a* $\mathcal{C}^2$ *circle map, then the constant* C *above may be written in terms of fundamental constants of the maps* T_k. *Set* $C = \max_{1 \leqq k \leqq r} \mathrm{Lip}(\log\ |\mathrm{T}'_\mathrm{k}|)$ *where* $\mathrm{Lip}(\log\ |\mathrm{T}'_\mathrm{k}|)$ *is the Lipschitz constant of* $\log\ |T'_k|$ *and* $\lambda = \min_{1 \leqq k \leqq r} \inf_x |T'_k(x)|$ *(assuming* $\lambda > 1$ *). Then,*

$$\begin{aligned} \|h_n - h\|_1 \ &\leqq\ (e^{C/(\lambda-1)n} - 1) \left(\max_{1 \leqq k \leqq r} \sum_{l=1}^{r} \mathcal{W}^*_{lk} \right) \\ &\times\ \left(\inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left([\frac{\log(4rn/\delta)}{-\log(\alpha' + \beta'/2)}] + 1 \right) - 1 \right\} + 1/2 \right), \end{aligned}$$

where $[\cdot]$ denotes the integer part and α'_l, β'_l, η'_l and α', β' are defined as before.

Proof. Consider $F_n = \{f_n \in BV : \ f_n = n \sum_{i=1}^{n} f_{n,i} \chi_{I_i}, \ \text{ for some } f_{n,i} \in \mathbb{R}\}$. Denote $\widehat{F}_n = \prod_{k=1}^{r} F_n$ and define the projection $\widehat{\Pi}_n : \widehat{BV} \to \widehat{F}_n$ by $\widehat{\Pi}_n\left((f^{(1)}, \cdots, f^{(r)})\right) = \left(\Pi_n(f^{(1)}), \cdots, \Pi_n(f^{(r)})\right)$, where $f^{(k)} \in F_n$, and $\widehat{\Pi}_n(f^{(k)}) = n \sum_{i=1}^{n} (\int_{I_i} f^{(k)} dm) \chi_{I_i}$. Note that the matrix representation of $[\widehat{\Pi}_n \widehat{\mathcal{P}}]$ with respect to the basis $\Pi_{k=1}^{r}\{\chi_{I_i}, \cdots, \chi_n\}$ is simply S_n. By Lemma 2.17 we have $[\widehat{\Pi}_n \widehat{\mathcal{P}}]_{ij} = P_{n,ij}$ and so

$$\left(h_n^{(1)}, \cdots, h_n^{(r)}\right) = P_n \left(h_n^{(1)}, \cdots, h_n^{(r)}\right) = \widehat{\Pi}_n \widehat{\mathcal{P}} \left((h_n^{(1)}, \cdots, h_n^{(r)})\right).$$

Now,

$$\max_{1\leqq k\leqq r} \mathrm{Var}(\mathrm{h_n^{(k)}}) = \max_{1\leqq k\leqq r} \mathrm{Var}\left(\widehat{\Pi}_{\mathrm{n}}\widehat{\mathcal{P}}(\mathrm{h_n^{(1)}},\cdots,\mathrm{h_n^{(r)}})\right)^{(\mathrm{k})}$$

$$\leqq \max_{1\leqq k\leqq r} \mathrm{Var}\widehat{\mathcal{P}}(\mathrm{h_n^{(1)}},\cdots,\mathrm{h_n^{(r)}})^{(\mathrm{k})}, \qquad \text{(by Lemma 2.19)}$$

$$\leqq 2\alpha' \max_{1\leqq k\leqq r} \mathrm{Var}(\mathrm{h_n^{(k)}}) + \max_{1\leqq \mathrm{l}\leqq \mathrm{r}}(2\mathrm{q}^*\alpha_{\mathrm{l}}' + \beta_{\mathrm{l}}') \max_{1\leqq \mathrm{k}\leqq \mathrm{r}} \|\mathrm{h_n^{(k)}}\|_1, \qquad \text{(by Lemma 3.4).}$$

Thus,

$$\begin{aligned} (1-2\alpha') \max_{1\leqq k\leqq r} \mathrm{Var}(\mathrm{h_n^{(k)}}) &\leqq \max_{1\leqq l\leqq r} (2q^*\alpha_l' + \beta_l') \max_{1\leqq k\leqq r} \|h_n^{(k)}\|_1 \\ \Rightarrow \max_{1\leqq k\leqq r} \mathrm{Var}(\mathrm{h_n^{(k)}}) &\leqq \left(\max_{1\leqq l\leqq r} (2q^*\alpha_l' + \beta_l') \,/\, (1-2\alpha')\right) \max_{1\leqq k\leqq r} \|h_n^{(k)}\|_1. \end{aligned}$$

Thus, the sequence $\{\mathrm{Var}(\mathrm{h_n^{(k)}})\}$ is bounded. So by Helly's Selection Principle (Theorem 2.5), the set $C = \{(h_n^{(1)},\cdots,h_n^{(r)}); n = 1,2,\cdots,\}$ is sequencetially compact (in $\prod_{k=1}^r L^1$).

Let $\{(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\}$ be any convergent subsequence of C and let $\{(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\}$ converge to $(h^{(1)},\cdots,h^{(r)})$ as $k\to\infty$. Then,

$$\begin{aligned} & \|(h^{(1)},\cdots,h^{(r)}) - \widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)})\| \\ \leqq\ & \|(h^{(1)},\cdots,h^{(r)}) - (h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\| + \|(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - \widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\| \\ +\ & \|\widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - \widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)})\| \\ +\ & \|\widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)}) - \widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)})\|. \end{aligned} \tag{3.17}$$

Taking into account that $(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})$ is a fixed point of P_{n_k} and by Lemma 2.17, we obtain

$$\|(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - \widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\| = \|P_{n_k}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - \widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)})\| = 0$$

Also

$$\|\widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - \widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)})\|$$

$$\leqq \|\widehat{\Pi}_{n_k}\|\|\widehat{\mathcal{P}}\|\|(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) - (h^{(1)},\cdots,h^{(r)})\| \to 0,$$

as $(h_{n_k}^{(1)},\cdots,h_{n_k}^{(r)}) \to (h^{(1)},\cdots,h^{(r)})$, and also by Lemma 2.16,

$$\widehat{\Pi}_{n_k}\widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)}) \to \widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)}).$$

Thus, by (3.17), $(h^{(1)},\cdots,h^{(r)}) = \widehat{\mathcal{P}}(h^{(1)},\cdots,h^{(r)})$.

Therefore any convergent subsequence of C converges to a fixed point of $\widehat{\mathcal{P}}$. By assumption, $\widehat{\mathcal{P}}$ has a unique fixed point h, that is $\|h - h_n\|_1 \to 0$ as $n \to \infty$. □

To prove the parts (ii) and (iii) of the above theorem we need to prove some inequalities. The following subsections are devoted to those.

3.4.1 Sensitivity of finite Markov chains

In this section we give error estimates for eigenvectors of stochastic matrices. The sensitivity of a finite Markov chain is a measure of how much the invariant density changes in response to a perturbation in the elements of the transition matrix. Whenever talking about norms on vectors, we shall denote the standard L^1 vector norm as $\|\cdot\|_m$ to avoid confusion with the L^1 norm on functions, which will be denoted by $\|\cdot\|_1$.

Our invariant measure μ may be decomposed as $\sum_{k=1}^{r} w_k\mu_k$ where the μ_k are fixed under (3.5). We construct matrices

$$\tilde{P}_n(k) = \frac{\mu_k(I_i \cap T_k^{-1} I_j)}{\mu_k(I_i)},$$

and form

$$\tilde{S}_n = \begin{pmatrix} \mathcal{W}_{11}^* \tilde{P}_n(1) & \mathcal{W}_{21}^* \tilde{P}_n(1) & \cdots & \mathcal{W}_{r1}^* \tilde{P}_n(1) \\ \mathcal{W}_{12}^* \tilde{P}_n(2) & \mathcal{W}_{22}^* \tilde{P}_n(2) & \cdots & \mathcal{W}_{r2}^* \tilde{P}_n(2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{W}_{1r}^* \tilde{P}_n(r) & \mathcal{W}_{2r}^* \tilde{P}_n(r) & \cdots & \mathcal{W}_{rr}^* \tilde{P}_n(r) \end{pmatrix}.$$

Let $\tilde{S}_{n,ij}^{(kl)} = \mathcal{W}_{lk}^* \tilde{P}(k)_{ij}$, $1 \leqq i, j \leqq n$, $1 \leqq k, l \leqq r$ be the $(i,j)^{\text{th}}$ entry of the $(k,l)^{\text{th}}$ block.

Let

$$\tilde{s}_n := [\mu_1(I_1), \cdots, \mu_1(I_n), \mu_2(I_1), \cdots, \mu_2(I_n), \cdots, \mu_r(I_1), \cdots, \mu_r(I_n)].$$

By $\tilde{s}_{n,i}^{(k)} := \mu_k(I_i)$, we denote the i^{th} entry of the k^{th} block of $\tilde{s}_n$. Here we have

$$\begin{aligned} & \sum_{k=1}^{r} \sum_{i=1}^{n} \tilde{S}_{n,ij}^{(kl)} \tilde{s}_{n,i}^{(k)} = \sum_{i=1}^{n} \sum_{k=1}^{r} \mu_k(I_i) \mathcal{W}_{lk}^* \tilde{P}(k)_{ij} \\ = & \sum_{i=1}^{n} \sum_{k=1}^{r} \mu_k(I_i) \mathcal{W}_{lk}^* \frac{\mu_k(I_i \cap T_k^{-1} I_j)}{\mu_k(I_i)} = \sum_{i=1}^{n} \sum_{k=1}^{r} \mathcal{W}_{lk}^* \mu_k(I_i \cap T_k^{-1} I_j) \\ = & \sum_{k=1}^{r} \mathcal{W}_{lk}^* \mu_k(T_k^{-1} I_j) = \sum_{k=1}^{r} \frac{\mathcal{W}_{kl} w_k}{w_l} \mu_k(I_j), \qquad (\text{by } T_k - \text{invariance of } \mu_k) \\ = & \mu_l(I_j) = \tilde{s}_{n,j}^{(l)}. \end{aligned}$$

Thus, the vector $\tilde{s}_n$ is a fixed left eigenvector of $\tilde{S}_n$. Let $\tilde{S}_n$ and S_n be two matrices and $\tilde{s}_n$ and s_n their eigenvectors correspondingly.

An important inequality by Paul J. Schweitzer [13] is:

$$\| \tilde{s}_n - s_n \|_m \leqq \| \tilde{S}_n - S_n \|_m \| (I_{rn} - S_n + S_n^{\infty})^{-1} \|_m$$

where $(S_n^{\infty})_{ij}^{(l)} = s_j^{(l)}$, time average transition probability matrix $S_n^{\infty} = \lim_{m\to+\infty}[S_n + \cdots + (S_n)^m]/m$ exists, $S_n\widehat{Z}_n = \widehat{Z}_n S_n = S_n^{\infty} + \widehat{Z}_n - I_{rn}$, and the fundamental matrix $\widehat{Z}_n \equiv (I_{rn} - S_n + S_n^{\infty})^{-1} = \sum_{k=0}^{\infty}(S_n - S_n^{\infty})^k = I_{rn} + \sum_{k=1}^{\infty}(S_n - S_n^{\infty})^k$ with $S_n^{\infty} = S_n S_n^{\infty} = \widehat{Z}_n S_n^{\infty} = S_n^{\infty}\widehat{Z}_n$.

In the following we will bound $\| \Pi_n(h) - h_n \|_1$ and in the later sections we will bound $\| \tilde{S}_n - S_n \|_m$ and $\| \widehat{Z}_n \|_m$.

For $\| \Pi_n(h) - h_n \|_1$, we have:

$$\begin{aligned}
&\|\Pi_n(h) - h_n\|_1 = \sum_{i=1}^{n} \int_{I_i} |\Pi_n(h) - h_n| dm \\
= &\sum_{i=1}^{n} \int_{I_i} \left| n \left(\sum_{i=1}^{n} \int_{I_i} hdm \right) \chi_{I_i} - \sum_{i=1}^{n} \left(\frac{\sum_{k=1}^{r} w_k s_{n,i}^{(k)}}{m(I_i)} \right) \chi_{I_i} \right| dm \\
= &\sum_{i=1}^{n} \int_{I_i} \left| n \int_{I_i} hdm\chi_{I_i} - \frac{\sum_{k=1}^{r} w_k s_{n,i}^{(k)}}{m(I_i)} \chi_{I_i} \right| dm \\
= &\sum_{i=1}^{n} \int_{I_i} \left| n \int_{I_i} \left(\sum_{k=1}^{r} w_k h^{(k)} \right) dm\chi_{I_i} - \frac{\sum_{k=1}^{r} w_k s_{n,i}^{(k)}}{m(I_i)} \chi_{I_i} \right| dm \\
= &\sum_{i=1}^{n} \int_{I_i} \frac{\left| \int_{I_i} \left(\sum_{k=1}^{r} w_k h^{(k)} \right) dm - \sum_{k=1}^{r} w_k s_{n,i}^{(k)} \right|}{m(I_i)} \chi_{I_i} dm \\
= &\sum_{i=1}^{n} \left| \int_{I_i} \left(\sum_{k=1}^{r} w_k h^{(k)} \right) dm - \sum_{k=1}^{r} w_k s_{n,i}^{(k)} \right| = \sum_{i=1}^{n} \left| \left(\sum_{k=1}^{r} w_k (\int_{I_i} h^{(k)}) dm - s_{n,i}^{(k)} \right) \right| \\
= &\sum_{i=1}^{n} \left| \sum_{k=1}^{r} w_k (\mu_k(I_i) - s_{n,i}^{(k)}) \right| = \sum_{i=1}^{n} \left| \sum_{k=1}^{r} w_k (\tilde{s}_{n,i}^{(k)} - s_{n,i}^{(k)}) \right| \\
\leqq &\sum_{i=1}^{n} \sum_{k=1}^{r} \left| \tilde{s}_{n,i}^{(k)} - s_{n,i}^{(k)} \right| = \| \tilde{s}_n - s_n \|_m \leqq \| \tilde{S}_n - S_n \|_m \ \| \widehat{Z}_n \|_m .
\end{aligned}$$

3.4.2 Renyi estimates for the invariant density

In this section we derive the necessary bounds for the regularity of the invariant density h in terms of fundamental constants of the maps T_k, when each T_k is a $\mathcal{C}^{1+\mathrm{Lip}}$ expanding circle map.

Lemma 3.5 *Suppose that each T_k is an expanding $\mathcal{C}^{1+\mathrm{Lip}}$ circle map. Define $\lambda = \min_{1\leqq k\leqq r}\inf_{x\in I}|T_k'(x)|$, and $C = \max_{1\leqq k\leqq r}\mathrm{Lip}(\log|\mathrm{T_k'}|)$. Then*

$$\frac{h^{(k)}(x)}{h^{(k)}(y)} \leqq \exp^{C|x-y|/(\lambda-1)}$$

for all $x \in I$ and each k=1,2,...,r.

Proof. Since each T_k is expanding, there exists $\epsilon > 0$ such that

$|x-y| < \epsilon \Rightarrow |T_kx - T_ky| \geqq \lambda|x-y|$ for all $x, y \in I$ and k=1,2,...,r.

We have,

$$\begin{aligned}
& \log\frac{|(T_{k_{N-1}}\circ\cdots\circ T_{k_0})'(x)|}{|(T_{k_{N-1}}\circ\cdots\circ T_{k_0})'(y)|} \\
= & \log\frac{|T'_{k_{N-1}}(T_{k_{N-2}}\circ\cdots\circ T_{k_0})(x)\cdot T'_{k_{N-2}}(T_{k_{N-3}}\circ\cdots\circ T_{k_0})(x)\cdots T'_{k_1}(T_{k_0})(x)\cdot T'_{k_0}(x)|}{|T'_{k_{N-1}}(T_{k_{N-2}}\circ\cdots\circ T_{k_0})(y)\cdot T'_{k_{N-2}}(T_{k_{N-3}}\circ\cdots\circ T_{k_0})(y)\cdots T'_{k_1}(T_{k_0})(y)\cdot T'_{k_0}(y)|} \\
= & \log|T'_{k_{N-1}}(T_{k_{N-2}}\circ\cdots\circ T_{k_0})(x)| - \log|T'_{k_{N-1}}(T_{k_{N-2}}\circ\cdots\circ T_{k_0})(y)| + \cdots \\
+ & \log|T'_{k_1}(T_{k_0}(x))| - \log|T'_{k_1}(T_{k_0}(y))| + \log|T'_{k_0}(x)| - \log|T'_{k_0}(y)| \\
= & \sum_{i=0}^{N-1}\log|T'_{k_i}(T_{k_{i-1}}\circ\cdots\circ T_{k_0})(x)| - \log|T'_{k_i}(T_{k_{i-1}}\circ\cdots\circ T_{k_0})(y)|
\end{aligned}$$

$$\begin{aligned}
&\Rightarrow \log \frac{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(x)|}{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(y)|} \\
&\leqq \sum_{i=0}^{N-1} \left|\log |T'_{k_i}(T_{k_{i-1}} \circ \cdots \circ T_{k_0})(x)| - \log |T'_{k_i}(T_{k_{i-1}} \circ \cdots \circ T_{k_0})(y)|\right| \\
&= \sum_{i=0}^{N-1} \left|\log |T'_k(T_{k_{i-1}} \circ \cdots \circ T_{k_0})(x)| - \log |T'_k(T_{k_{i-1}} \circ \cdots \circ T_{k_0})(y)|\right| \\
&= \sum_{i=0}^{N-1} \mathrm{Lip}\,(\log |T'_k|) \left|T_{k_{i-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{i-1}} \circ \cdots \circ T_{k_0}(y)\right| \\
&\leqq \sum_{i=0}^{N-1} C \left|T_{k_{i-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{i-1}} \circ \cdots \circ T_{k_0}(y)\right|. \qquad (3.18)
\end{aligned}$$

Now, since

$$\begin{aligned}
& |T_k x - T_k y| \geqq \lambda |x - y| \\
\Rightarrow\ & |(T_{k_{N-1}} \circ \cdots \circ T_{k_0})(x) - (T_{k_{N-1}} \circ \cdots \circ T_{k_0})(y)| \\
\geqq\ & \lambda |(T_{k_{N-2}} \circ \cdots \circ T_{k_0})(x) - (T_{k_{N-2}} \circ \cdots \circ T_{k_0})(y)| \\
\geqq\ & \lambda^2 |(T_{k_{N-3}} \circ \cdots \circ T_{k_0})(x) - (T_{k_{N-3}} \circ \cdots \circ T_{k_0})(y)| \\
=\ & \lambda^{N-i} |(T_{k_{i-1}} \circ \cdots \circ T_{k_0})(x) - (T_{k_{i-1}} \circ \cdots \circ T_{k_0})(y)|.
\end{aligned}$$

Thus, by (3.18),

$$\begin{aligned}
& \log \frac{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(x)|}{|(T_{k_{N-1}} \circ \cdots \circ T_{k_0})'(y)|} \\
\leqq\ & \sum_{i=0}^{N-1} C\lambda^{-(N-i)} |T_{k_{N-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{N-1}} \circ \cdots \circ T_{k_0}(y)|, \\
& \qquad (\text{provided } |T_{k_{i-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{i-1}} \circ \cdots \circ T_{k_0}(y)| < \epsilon) \\
=\ & \frac{C}{\lambda - 1} |T_{k_{N-1}} \circ \cdots \circ T_{k_0}(x) - T_{k_{N-1}} \circ \cdots \circ T_{k_0}(y)|. \qquad (3.19)
\end{aligned}$$

Let $\phi_0 \equiv 1$ be an initial density that is to be pushed forward and denote by $\phi^i_{k_{N-1},\cdots,k_0}$ the push forward of ϕ_0 under $T_{k_{N-1}} \circ \cdots \circ T_{k_0}$ along one of the inverse branches of $T_{k_{N-1}} \circ \cdots \circ T_{k_0}$. By (3.19), we have

$$\log \frac{\phi^i_{k_{N-1},\cdots,k_0}(x)}{\phi^i_{k_{N-1},\cdots,k_0}(y)} \leqq \frac{C}{\lambda - 1}|x-y|$$
$$\Rightarrow \quad \frac{\phi^i_{k_{N-1},\cdots,k_0}(x)}{\phi^i_{k_{N-1},\cdots,k_0}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|}.$$

For a fixed sequence $k_{N-1}, \cdots, k_0$, we may sum over i to obtain

$$\frac{\sum_{i=0}^{N-1} \phi^i_{k_{N-1},\cdots,k_0}(x)}{\sum_{i=0}^{N-1} \phi^i_{k_{N-1},\cdots,k_0}(y)} = \frac{\phi_{k_{N-1},\cdots,k_0}(x)}{\phi_{k_{N-1},\cdots,k_0}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|}.$$

We may combine the contribution from each of the sequences $k_{N-1}, \cdots, k_0$ to obtain

$$\frac{\sum_{k_0,\cdots,k_{N-1}=1}^{r} \mathcal{W}^*_{k_{N-1}} \cdots \mathcal{W}^*_{k_1 k_0} \phi_{k_{N-1},\cdots,k_0}(x)}{\sum_{k_0,\cdots,k_{N-1}=1}^{r} \mathcal{W}^*_{k_{N-1}} \cdots \mathcal{W}^*_{k_1 k_0} \phi_{k_{N-1},\cdots,k_0}(y)} := \frac{\phi_N^{(k)}(x)}{\phi_N^{(k)}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|}.$$

Here, $\phi_N^{(k)}(x) = (\widehat{\mathcal{P}}^N \phi_0)^{(k)}(x)$, so that we have a bound on the distortion of the uniform density after being pushed forward N times under the Perron-Frobenious operator.

Since $\int_I \phi_N^{(k)} dx = 1$, $\exists x \in I$, $\phi_N^{(k)}(x) \geqq 1$ and $\exists y \in I$ such that $\phi_N^{(k)}(y) \leqq 1$. We have

$$\frac{\phi_N^{(k)}(x)}{\phi_N^{(k)}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|} \leqq A,$$

since $|x-y| \leqq 1$, where $A = e^{\frac{C}{\lambda-1}}$. That is, $\phi_N^{(k)}(x) \leqq A\phi_N^{(k)}(y) \Rightarrow \phi_N^{(k)}(y) \geqq B$ where $B = e^{\frac{-C}{\lambda-1}}$. If $\phi_N^{(k)}(x) \geqq 1$ we have $\phi_N^{(k)}(y) \geqq \frac{1}{A}$ for all y. If $\phi_N^{(k)}(y) \leqq 1$ we obtain $\phi_N^{(k)}(x) \leqq A$ for all $x \in I$. Thus $\frac{1}{A} \leqq \phi_N^{(k)}(x) \leqq A$.

Let $\phi^{(k)}$ be the limit of the sequence $\frac{1}{N}\sum_{i=0}^{N-1} \phi_N^{(k)}(x)$ as $N \to \infty$, we see that ϕ is fixed by $\widehat{\mathcal{P}}$ and is bounded above and below by A and $\frac{1}{A}$ respectively.

Furthermore, $\frac{\phi_N^{(k)}(x)}{\phi_N^{(k)}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|}$. By uniqueness, $\phi^{(k)} = h^{(k)}$.

Thus, $\frac{h^{(k)}(x)}{h^{(k)}(y)} \leqq e^{\frac{C}{\lambda-1}|x-y|}$. □

3.4.3 Bounding $\| \tilde{S}_n - S_n \|_m$ and $\| \tilde{Z}_n \|_m$

Lemma 3.6 *Let S_n and $\tilde{S}_n$ be as defined before. Under the assumptions of Theorem 3.1, we have*

(i) $\|\tilde{S}_n - S_n\|_m \leqq \max_{1\leqq k\leqq r}\left((\sum_{l=1}^r \mathcal{W}^*_{lk})(\mathrm{Lip}(\mathrm{h}^{(k)})/\inf_{\mathrm{x}\in \mathrm{I}}\mathrm{h}^{(k)})\right)/n,$

if each T_k is a general Lasota-Yorke map, and the partition $\{I_1, \cdots, I_n\}$ contains all points of non-Lipschitzness of every T_k, $k = 1, \cdots, r$.

(ii) $\|\tilde{S}_n - S_n\|_m \leqq (\max_{1\leqq k\leqq r}\sum_{l=1}^r \mathcal{W}^*_{lk})(e^{C/(1-\lambda)n} - 1),$

if each T_k is a C^{1+Lip} circle map.

Proof. We treat case (ii) first

$$
\begin{aligned}
|P_{n,ij}(k) - \tilde{P}_{n,ij}(k)| &= \left|\frac{m(I_i\cap T_k^{-1}I_j)}{m(I_i)} - \frac{\mu_k(I_i\cap T_k^{-1}I_j)}{\mu_k(I_i)}\right| \\
&= \frac{m(I_i\cap T_k^{-1}I_j)}{m(I_i)}\left|1 - \frac{\mu_k(I_i\cap T_k^{-1}I_j)}{m(I_i\cap T_k^{-1}I_j)}\frac{m(I_i)}{\mu_k(I_i)}\right| \\
&= P_{n,ij}(k)\left|1 - \left(\frac{1}{m(I_i\cap T_k^{-1}I_j)}\int_{I_i\cap T_k^{-1}I_j} h^{(k)}dm\right)\left(\frac{1}{m(I_i)}\int_{I_i} h^{(k)}dm\right)^{-1}\right| \\
&\leqq P_{n,ij}(k)\left|1 - \left(\sup_{x\in I_i\cap T_k^{-1}I_j} h^{(k)}(x)\right)\left(\inf_{x\in I_i} h^{(k)}(x)\right)^{-1}\right| \\
&\leqq P_{n,ij}(k)\left|1 - \left(\sup_{x\in I_i} h^{(k)}(x)\right)\left(\inf_{x\in I_i} h^{(k)}(x)\right)^{-1}\right|
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&\|P_n(k) - \tilde{P}_n(k)\|_m \\
\leqq\ & \max_{1\leqq i\leqq n}\sum_{j=1}^n P_{n,ij}(k)\left|\left(\sup_{x\in I_i} h^{(k)}(x)\right)\left(\inf_{x\in I_i} h^{(k)}(x)\right)^{-1} - 1\right| \qquad (3.20) \\
\leqq\ & \max_{1\leqq i\leqq n}\sum_{j=1}^n P_{n,ij}(k)(e^{C/(\lambda-1)n} - 1) = e^{C/(\lambda-1)n} - 1,
\end{aligned}
$$

and

$$
\begin{aligned}
\|S_n - \tilde{S}_n\|_m &= \max_{1\leqq k\leqq r} \sum_{l=1}^{r} \mathcal{W}^*_{lk} \|P_n(k) - \tilde{P}_n(k)\|_m \\
&\leqq \left(\max_{1\leqq k\leqq r} \sum_{l=1}^{r} \mathcal{W}^*_{lk} \right) \left(e^{C/(\lambda-1)n} - 1 \right).
\end{aligned}
$$

For the proof of (i), we have from (3.20)

$$
\begin{aligned}
\|P_n(k) - \tilde{P}_n(k)\|_m &\leqq \max_{1\leqq i\leqq n} \sum_{j=1}^{n} P_{n,ij}(k) \left| \frac{\sup_{x\in I} h^{(k)}(x) - \inf_{x\in I} h^{(k)}(x)}{\inf_{x\in I} h^{(k)}(x)} \right| \\
&\leqq \max_{1\leqq i\leqq n} \sum_{j=1}^{n} P_{n,ij}(k) \frac{\mathrm{Lip}(h^{(k)})\frac{1}{n}}{\inf_{x\in I} h^{(k)}(x)} = \left(\frac{\mathrm{Lip}(h^{(k)})}{\inf_{x\in I} h^{(k)}(x)} \right) /n,
\end{aligned}
$$

where $\mathrm{Lip}(h^{(k)})$ is understood to be the maximum Lipschitz constant calculated over each of the Lipschitz pieces of $h^{(k)}$ seperately. Thus,

$$
\begin{aligned}
\|S_n - \tilde{S}_n\|_m &= \max_{1\leqq k\leqq r} \sum_{l=1}^{r} \mathcal{W}^*_{lk} \|P_n(k) - \tilde{P}_n(k)\|_m \\
&\leqq \max_{1\leqq k\leqq r} \left(\sum_{l=1}^{r} \mathcal{W}^*_{lk} \cdot \left(\frac{\mathrm{Lip}(h^{(k)})}{\inf_{x\in I} h^{(k)}(x)} \right) /n \right).
\end{aligned}
$$

□

Bounding $\| \widehat{Z}_n \|_m$:

Converting norms: We wish to study the rate of convergence of S_n^N to the limiting matrix S_n^∞ (defined before) as

$$
\| \tilde{s}_n - s_n \|_m \leqq \| \tilde{S}_n - S_n \|_m \| (I_{rn} - S_n + S_n^\infty)^{-1} \|_m \quad \text{as } N \to \infty,
$$

in terms of the $\| \cdot \|_m$ norm where norm on $F_n = (f_{n,1}, \cdots, f_{n,n})$ will be $\| f_n \|_m = \sum_{i=1}^{n} |f_{n,i}|$ and $f_{n,i}$ define f_n. At the moment, we have the information regarding the convergence of $\widehat{\mathcal{P}}^N \mid_{\widehat{BV}_0}$ to $\widehat{0}$ from Lemma 3.4 and in this section, we link these two types of convergence.

We define an intermediate vector norm $\| \cdot \|_{m'}$ as $\| S_n \|_{m'} = \max_{1 \leqq k \leqq r} \| S_n^{(k)} \|_m$. But we defined $\| \cdot \|_m = \sum | \cdot |$. That is, $\| \cdot \|_m = \sum_1^r | \cdot | \leqq r \| \cdot \|_{m'} \Rightarrow \| \cdot \|_m \leqq r \| \cdot \|_{m'}$ and $\| \cdot \|_m = \sum_1^r | \cdot | \geqq \max_{1 \leqq k \leqq r} \| \cdot \|_m = \| \cdot \|_{m'} \Rightarrow \| \cdot \|_m \geqq \| \cdot \|_{m'}$. That is,

$$\| \cdot \|_{m'} \leqq \| \cdot \|_m \leqq r \| \cdot \|_{m'} . \tag{3.21}$$

Lemma 3.7 *For $nr-$tuple*

$$\tilde{f}_n = \left(f_{n,1}^{(1)}, \cdots, f_{n,n}^{(1)}, f_{n,1}^{(2)}, \cdots, f_{n,n}^{(2)}, \cdots, f_{n,1}^{(r)}, \cdots, f_{n,n}^{(r)} \right)$$

representing a $1 \times nr$ vector and an element of $\widehat{BV}$, we have the relations

$$\| \widehat{f}_n \|_{m'} \leqq n \| \widehat{f}_n \| \quad \text{and} \quad \| \widehat{f}_n \| \leqq 2 \| \widehat{f}_n \|_{m'} .$$

Proof. We know that

$$\| \widehat{f}_n \|_{m'} = \max_{1 \leqq k \leqq r} \| f_n^{(k)} \|_m \quad \text{and} \quad \| \widehat{f}_n \| = \max_{1 \leqq k \leqq r} \| f_n^{(k)} \| .$$

Also, we note that

$$\| \widehat{f}_n \|_{m'} \leqq \| \widehat{f}_n \|_m = n \| \widehat{f}_n \|_1 \quad \text{as} \quad \| \widehat{f}_n \|_m = \sum_{i=1}^n |\widehat{f}_{n,i}|$$

$$\Rightarrow \| \widehat{f}_n \|_{m'} \leqq n \| \widehat{f}_n \|_1 .$$

So,

$$\| \widehat{f}_n \|_{m'} \leqq n \| \widehat{f}_n \|_1 \leqq \max \left\{ \mathrm{Var}(\mathrm{n}\widehat{\mathrm{f}}_\mathrm{n}), \mathrm{n} \| \widehat{\mathrm{f}}_\mathrm{n} \|_1 \right\} = n \| \widehat{f}_n \| .$$

Again,

$$\begin{aligned}
\mathrm{Var}(\widehat{\mathrm{f}}_{\mathrm{n}}) &= \max_{1\leqq k\leqq r}\left\{|f^{(k)}_{n,2}-f^{(k)}_{n,1}|+\cdots+|f^{(k)}_{n,n}-f^{(k)}_{n,n-1}|\right\}\\
&\leqq \max_{1\leqq k\leqq r}\left\{2|f^{(k)}_{n,1}|+2|f^{(k)}_{n,2}|+\cdots+2|f^{(k)}_{n,n}|\right\}\\
&= \max_{1\leqq k\leqq r} 2 \parallel f^{(k)}_n \parallel_m = 2 \parallel \widehat{f}_n \parallel_{m'} .
\end{aligned}$$

So,

$$\begin{aligned}
\parallel \widehat{f}_n \parallel &= \max\left\{\mathrm{Var}(\widehat{\mathrm{f}}_{\mathrm{n}}),\ \parallel \widehat{\mathrm{f}}_{\mathrm{n}} \parallel_1\right\}\\
&\leqq \max\left\{2 \parallel \widehat{f}_n \parallel_{m'},\ \frac{1}{n} \parallel \widehat{f}_n \parallel_{m'}\right\} = 2 \parallel \widehat{f}_n \parallel_{m'} .
\end{aligned}$$

□

Lemma 3.8

$$\parallel S^N_n - S^\infty_n \parallel_m \leqq 4rn \parallel \widehat{\mathcal{P}} |_{\widehat{BV}_0} \parallel^N .$$

Proof. Let $\widehat{F}$ and $\widehat{\Pi}_n$ be as in the proof of Theorem 3.1(i). Define

$$\widehat{F}_{n,0} = \left\{ (f^{(1)}_n, \cdots, f^{(r)}_n) \in \widehat{F}_n \ : \ \sum_{i=1}^n f^{(k)}_{n,i} = 0 \ \text{ for all } \ k = 1,2,\cdots,r \right\}.$$

We begin by relating $\parallel S^N_n - S^\infty_n \parallel_m$ and $\parallel S^N_n |_{\hat{F}_{n,0}} \parallel_m$. In what follows, we simultaneously consider $\widehat{f}_n = (f^{(1)}, \cdots, f^{(r)}_n)$ as a step function, and as the $n-$tuple $[f^{(1)}_{n,1}, \cdots, f^{(r)}_{n,n}]$; in the latter case the action of matrices is understood to be the left multiplication. Now,

$$\begin{aligned}
\| S_n^N - S_n^\infty \|_m &= \sup_{\hat{f}_n \in \hat{F}_n} \frac{\| (S_n^N - S_n^\infty)\widehat{f}_n \|_m}{\| \widehat{f}_n \|_m} \\
&= \sup_{\hat{f}_n \in \hat{F}_n} \frac{\| S_n^N(\widehat{f}_n - S_n^\infty \widehat{f}_n) \|_m}{\| \widehat{f}_n \|_m} \qquad (\text{as} \qquad S_n^N S_n^\infty = S_n^\infty) \\
&\leqq \sup_{\hat{f}_n \in \hat{F}_n} \frac{\| S_n^N \|_m \| \widehat{f}_n - S_n^\infty \widehat{f}_n \|_m}{\| \widehat{f}_n \|_m} = \| S_n^N |_{\hat{F}_{n,0}} \|_m \sup_{\hat{f}_n \in \hat{F}_n} \frac{\| \widehat{f}_n - S_n^\infty \widehat{f}_n \|_m}{\| \widehat{f}_n \|_m} \\
&\leqq \| S_n^N |_{\hat{F}_{n,0}} \|_m \sup_{\hat{f}_n \in \widehat{F}_n} \frac{\| \widehat{f}_n \|_m + \| S_n^\infty \widehat{f}_n \|_m}{\| \widehat{f}_n \|_m} \\
&\leqq \| S_n^N |_{\hat{F}_{n,0}} \|_m \sup_{\hat{f}_n \in \widehat{F}_n} \frac{\| \widehat{f}_n \|_m + \| \widehat{f}_n \|_m}{\| \widehat{f}_n \|_m} \\
&= 2 \| S_n^N |_{\hat{F}_{n,0}} \|_m .
\end{aligned}$$

Now we link this result with the bounds that we have for the Perron-Frobenious operator. Recall that the matrix form of $\widehat{\Pi}_n \widehat{\mathcal{P}}$ with respect to the basis $\Pi_{k=1}^r \{\chi_{I_1}, \cdots, \chi_{I_n}\}$ is simply S_n.

So

$$\begin{aligned}
2 \| S_n^N |_{\hat{F}_{n,0}} \|_m &= 2 \sup_{\hat{f}_{n,0} \in \hat{F}_{n,0}} \frac{\| [\widehat{\Pi}_n \widehat{\mathcal{P}}]^N \widehat{f}_{n,0} \|_m}{\| \widehat{f}_{n,0} \|_m} \\
&\leqq 2 \sup_{\hat{f}_{n,0} \in \hat{F}_{n,0}} \frac{r \| [\widehat{\Pi}_n \widehat{\mathcal{P}}]^N \widehat{f}_{n,0} \|_{m'}}{\| \widehat{f}_{n,0} \|_{m'}} \qquad (\text{by } 3.21\ ) \\
&\leqq 2r \sup_{\hat{f}_{n,0} \in \hat{F}_{n,0}} \frac{n \| [\widehat{\Pi}_n \widehat{\mathcal{P}}]^N \widehat{f}_{n,0} \|}{\| \widehat{f}_{n,0} \| /2} \qquad (\text{by Lemma } 3.7) \\
&\leqq 4rn \| [\widehat{\Pi}_n \widehat{\mathcal{P}}]^N |_{\hat{F}_{n,0}} \| \leqq 4rn \| \widehat{\Pi}_n \widehat{\mathcal{P}} |_{\hat{F}_{n,0}} \|^N \\
&\leqq 4rn \| \widehat{\mathcal{P}} |_{\hat{F}_{n,0}} \|^N \leqq 4rn \| \widehat{\mathcal{P}} |_{\hat{BV}_0} \|^N .
\end{aligned}$$

□

Corollary 3.1 *Under the hypotheses of Theorem 3.1(ii) or (iii),*

$\| S_n^N - S_n^\infty \|_m \leqq 4rn\gamma^N$, *where* $\gamma = \max_{1\leqq l \leqq r}(\alpha_l' + \eta_l') + \beta'/2$ *or* $\gamma = \alpha' + \beta'/2$ *respectively.*

We now state a result from [10] (Theorem16.2.4) to prove the next Lemma:

Lemma 3.9 *Suppose that* P_n *is an* $n \times n$ *irreducible, aperiodic stochastic matrix with fixed left eigenvector* p_n. *Define* $P_{n,ij}^\infty = p_{n,j}$. *Select a number* $0 < \delta < 1$ *and let* m_n *be such that*

$$P_{n,ij}^{m_n} \geqq (1-\delta)p_{n,j} \text{ for all } 1 \leqq i,j \leqq n. \tag{3.22}$$

Then

$$\| P_n^N - P_n^\infty \|_m \leqq \begin{cases} 2, & \text{if } N < m_n; \\ \delta^{[N/m_n]}, & \text{if, } N \geqq m_n. \end{cases}$$

Lemma 3.10 *Under the hypotheses of Theorem 3.1(ii) or (iii), setting*

$$\gamma = \max_{1\leqq l \leqq r}(\alpha_l' + \eta_l') + \beta'/2 \text{ or } \gamma = \alpha' + \beta'/2 \text{ respectively, we have}$$

$$\| \widehat{Z}_n \|_m \leqq \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1 \right) - 1 \right\}$$

where $[\cdot]$ is the integer part.

Proof. We now find an appropriate m_n to satisfy (3.22) for S_n. A sufficient condition for (3.22) to be satisfied is that

$$|S_{n,ij}^{m_n} - s_{n,j}| \leqq \delta s_{n,j} \text{ for all } 1 \leqq i,j \leqq n.$$

Summing over j and maximizing over i gives

$$\max_{1 \leqq i \leqq n} \sum_{j=1}^{n} |S_{n,ij}^{m_n} - s_{n,j}| \leqq \sum_{j=1}^{n} \delta s_{n,j} \Rightarrow \max_{1 \leqq i \leqq n} \sum_{j=1}^{n} |S_{n,ij}^{m_n} - S_n^{\infty}| \leqq \delta$$
$$\Rightarrow \quad \| S_n^{m_n} - S_n^{\infty} \|_m \leqq \delta$$

which implies (3.22) holds. From corollary 3.1 and above, we see that

$$\| S_n^{m_n} - S_n^{\infty} \|_m \leqq 4rn\gamma^{m_n} \quad \text{and}$$

$$\| S_n^{m_n} - S_n^{\infty} \|_m \leqq \delta.$$

That is, provided $4rn\gamma^N \leqq \delta$, (3.22) will hold.

Now we have to find a condition on m_n. Suppose $4rn\gamma^{m_n} \leqq \delta$.

Then, we have

$$\begin{aligned}
& m_n \log\ \gamma + \log\ 4rn \leqq \log\ \delta \\
\Rightarrow \quad & -m_n \log\ \gamma \geqq \log\ 4rn - \log\ \delta \\
\Rightarrow \quad & -m_n \log\ \gamma \geqq \log\ \frac{4rn}{\delta} \\
\Rightarrow \quad & m_n \geqq \frac{\log\ \frac{4rn}{\delta}}{-\log\ \gamma}, \qquad (\text{since } \ 0 < \gamma < 1) \\
\Rightarrow \quad & m_n \geqq [\frac{\log\ \frac{4rn}{\delta}}{-\log\ \gamma}] + 1.
\end{aligned}$$

Thus, $4rn\gamma^{m_n} \leqq \delta$ if

$$m_n \geqq [\frac{\log\ \frac{4rn}{\delta}}{-\log\ \gamma}] + 1, \tag{3.23}$$

where $[\cdot]$ denotes the integer part.

Thus,

$$\begin{aligned}
\| \widehat{Z}_n \|_m = \| I_{rn} - S_n + S_n^{\infty} \| = \| I + \sum_{N=1}^{\infty}(S_n^N - S_n^{\infty}) \| \\
&= 1 + \sum_{N=1}^{m_n-1} \| S_n^N - S_n^{\infty} \| + \sum_{N=m_n}^{\infty} \| S_n^N - S_n^{\infty} \| \leqq 1 + \sum_{N=1}^{m_n-1} 2 + \sum_{N=m_n}^{\infty} \delta^{[N/m_n]} \\
&= 1 + 2(m_n - 1) + [\delta + \cdots + \delta(m_n \text{times}) + \cdots + \delta^r + \cdots + \delta^r(m_n \text{times}) + \cdots] \\
&= 2m_n - 1 + \delta m_n[1 + \delta + \delta^2 + \cdots] = 2m_n - 1 + m_n\delta\frac{1}{1-\delta} = (2 + \frac{\delta}{1-\delta})m_n - 1.
\end{aligned}$$

Thus,

$$\| \widehat{Z}_n \|_m \leqq \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta}\right) \left([\frac{\log(4rn/\delta)}{-\log\, \gamma}] + 1\right) - 1 \right\} \quad \text{by (3.23).}$$

□

3.4.4 The difference $\|h - \Pi_n(h)\|_1$:

Lemma 3.11 *Under the assumptions of Theorem 3.1*

(i) $\|h - \Pi_n(h)\|_1 \leqq \sum_{k=1}^{r} w_k \mathrm{Lip}(\mathrm{h}^{(\mathrm{k})})/2\mathrm{n}$, *if each* $\mathrm{T_k}$ *is a general Lasota-Yorke map,*

(ii) $\|h - \Pi_n(h)\|_1 \leqq \left(e^{C/(\lambda-1)n}\right)/2$, *if each* T_k *is a* $C^{1+\mathrm{Lip}}$ *circle map.*

Proof. First assume that each T_k is $C^{1+\mathrm{Lip}}$ circle map.

$$\begin{aligned}
\|h - \Pi_n(h)\|_1 &= \| \sum_{k=1}^{r} w_k h^{(k)} - \Pi_n \left(\sum_{k=1}^{r} w_k h^{(k)} \right) \|_1 \\
&\leqq \sum_{k=1}^{r} w_k \|h^{(k)} - \Pi_n h^{(k)}\|_1 = \sum_{k=1}^{r} w_k \sum_{i=i}^{n} \int_{I_i} \left| h^{(k)} - n \int_{I_i} h^{(k)} dm \right| dm. \quad (3.24)
\end{aligned}$$

Here,

$$\int_{I_i}(h^{(k)} - n\int_{I_i} h^{(k)}dm)dm = \int_{I_i} h^{(k)} - \int_{I_i}\left(n\int_{I_i} h^{(k)}dm\right)dm$$
$$= \int_{I_i} h^{(k)}dm - \left(n\int_{I_i} h^{(k)}dm\right)\left(\int_{I_i} dm\right) = \int_{I_i} h^{(k)}dm - n\int_{I_i} h^{(k)}dm \ \ m(I_i)$$
$$= \int_{I_i} h^{(k)}dm - \int_{I_i} h^{(k)}dm = 0.$$

Since $h^{(k)}$ has integral zero, then by Lemma 2.14, we have

$$n\int_{I_i} |h^{(k)}|dm \leqq \frac{1}{2}\left(\sup_{x\in I_i} h^{(k)}(x) - \inf_{x\in I_i} h^{(k)}(x)\right)$$
$$\Rightarrow \int_{I_i} |h^{(k)}|dm \leqq \frac{1}{n}\left(\sup_{x\in I_i} h^{(k)}(x) - \inf_{x\in I_i} h^{(k)}(x)\right)/2.$$

Thus, by (3.24)

$$\|h - \Pi_n(h)\|_1 \leqq \sum_{k=1}^{r} w_k \sum_{i=1}^{n} \frac{1}{n}\left(\sup_{x\in I_i} h^{(k)}(x) - \inf_{x\in I_i} h^{(k)}(x)\right)/2 \tag{3.25}$$
$$= \sum_{k=1}^{r} w_k \frac{1}{n}\sum_{i=1}^{n} \inf_{x\in I_i} h^{(k)}(x)\left(\frac{\sup_{x\in I_i} h^{(k)}(x)}{\inf_{x\in I_i} h^{(k)}(x)} - 1\right)/2$$
$$\leqq \sum_{k=1}^{r} w_k \left(e^{C/(\lambda-1)n}\right)/2 \qquad \text{(by Lemma 3.5)}$$
$$= \left(e^{C/(\lambda-1)n}\right)/2. \tag{3.26}$$

In case of general Lasota-Yorke maps, by (3.25), we have

$$\|h - \Pi_n(h)\|_1 \leqq \sum_{k=1}^{r} w_k \mathrm{Lip}(h^{(k)})/2n.$$

□

3.4.5 Proof of (ii) and (iii) of Theorem 3.1

Part (ii):

$$\begin{aligned}
\| h_n - h \|_1 &\leqq \| h_n - \Pi_n(h) \|_1 + \| \Pi_n(h) - h \|_1 \\
&\leqq \| \tilde{S}_n - S_n \|_m \| \tilde{Z}_n \|_m + \| \Pi_n(h) - h \|_1 \\
&\leqq \max_{1 \leqq k \leqq r} \left((\sum_{l=1}^{r} \mathcal{W}_{lk}^{*})(\mathrm{Lip}(h^{(k)}) / \inf_{x \in I} h^{(k)}) \right) / n \\
&\times \inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta} \right) \left([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1 \right) - 1 \right\} + \sum_{k=1}^{r} w_k \mathrm{Lip}(h^{(k)}/2n) \\
&= \frac{1}{n} \max_{1 \leqq k \leqq r} \left((\sum_{l=1}^{r} \mathcal{W}_{lk}^{*})(\mathrm{Lip}(h^{(k)}) / \inf_{x \in I} h^{(k)}) \right) \\
&\times \left(\inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta} \right) \left([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1 \right) - 1 \right\} + \sum_{k=1}^{r} w_k \mathrm{Lip}(h^{(k)}/2) \right) \\
&\leqq \frac{1}{n} C_1 \frac{\log n}{C_2} + C_3,
\end{aligned}$$

where

$$\inf_{0<\delta<1} \left\{ \left(2 + \frac{\delta}{1-\delta} \right) \left([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1 \right) - 1 \right\} \leqq \frac{\log n}{C_2},$$

and $C_1 = \max_{1 \leqq k \leqq r} \left((\sum_{l=1}^{r} \mathcal{W}_{lk}^{*})(\mathrm{Lip}(h^{(k)}) / \inf_{x \in I} h^{(k)}) \right)$, C_2 is a constant.

Thus,

$$\| h_n - h \|_1 \leqq C \frac{\log n}{n},$$

where $C = \frac{C_1}{C_2}$ and for large n we can neglect C_3.

Part(iii):

$$
\begin{aligned}
& \| h_n - h \|_1 \leqq \| h_n - \Pi_n(h) \|_1 + \| \Pi_n(h) - h \|_1 \\
\leqq \ & \| \tilde{S}_n - S_n \|_m \| \tilde{Z}_n \|_m + \| \Pi_n(h) - h \|_1 \leqq (\max_{1 \leqq k \leqq r} \sum_{l=1}^{r} \mathcal{W}_{lk}^{*})(e^{c/(1-\lambda)n)} - 1) \\
\times \ & \inf_{0<\delta<1} \{(2 + \frac{\delta}{1-\delta})([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1) - 1\} + (e^{c/(1-\lambda)n)} - 1)/2 \\
= \ & (e^{c/(1-\lambda)n)} - 1) \times \left((\max_{1 \leqq k \leqq r} \sum_{l=1}^{r} \mathcal{W}_{lk}^{*}) \inf_{0<\delta<1} \left\{ (2 + \frac{\delta}{1-\delta}) \left([\frac{\log(4rn/\delta)}{-\log \gamma}] + 1 \right) - 1 \right\} + 1/2 \right).
\end{aligned}
$$

3.4.6 Appendix

Example: Let τ_1 and τ_2 be defined by

$$\tau_1(x) = 6x^3 - 9x^2 + 8x(mod\ 1) \qquad \text{and}$$

$$\tau_2(x) = \begin{cases} 3x + x^2, & 0 \leqq x \leqq \frac{-3+\sqrt{13}}{2}, \\ (\frac{9}{4} - \frac{\sqrt{13}}{2})(x - \frac{3}{4}) + 1, & \frac{-3+\sqrt{13}}{2} \leqq x \leqq \frac{3}{4}, \\ 4x - 3, & \frac{3}{4} \leqq x \leqq 1. \end{cases}$$

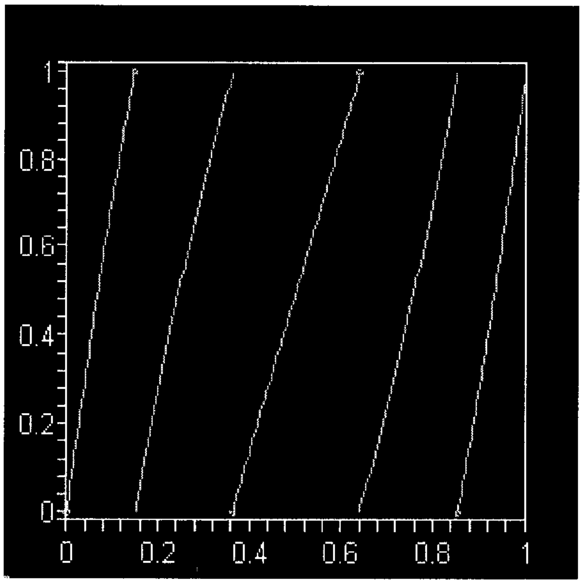

Figure 3.1: Graph of τ_1

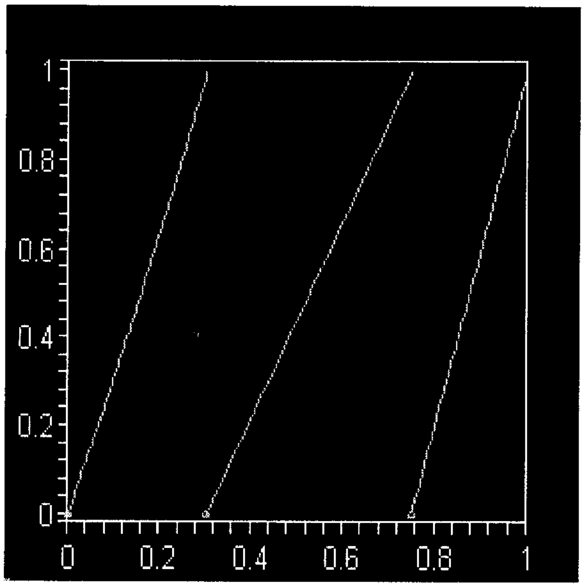

Figure 3.2: Graph of τ_2

Here, $\inf_{x\in I}|\tau_1'(x)| = 3.5$ and $\inf_{x\in I}|\tau_2'(x)| = \frac{1}{\frac{9}{4}-\frac{\sqrt{13}}{2}}$.

Consider the matrix

$$W = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 1 & 0 \end{pmatrix}$$

and its fixed left eigenvector $w = [\frac{2}{3}, \ \frac{1}{3}]$. So, our $\alpha_1' = 0.37$ and $\alpha_2' = 0.28$ and $\alpha' = 0.37 < \frac{1}{2}$.

Thus, according to the condition (i) of Theorem, we are guaranteed that there exists a unique invariant density for our maps.

Now, we will approximate the invariant density for $n = 8$.

The transition matrices for τ_1 and τ_2 are respectively

$$P_8(1) = \begin{bmatrix} 0.33 & 0.32 & 0.31 & 0.04 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .27 & 0.30 & 0.29 & 0.14 & 0 \\ 0.45 & 0.13 & 0 & 0 & 0 & 0 & 0.14 & 0.28 \\ 0 & 0.32 & 0.45 & 0.24 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.21 & 0.45 & 0.34 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.11 & 0.45 & 0.45 \\ 0.25 & 0.25 & 0.25 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0.25 & 0.25 & 0.25 \end{bmatrix} \text{ and}$$

$$P_8(2) = \begin{bmatrix} 0.13 & 0.13 & 0.14 & 0.14 & 0.15 & 0.16 & 0.16 & 0 \\ 0.18 & 0.19 & 0.20 & 0.21 & 0.05 & 0 & 0.01 & 0.17 \\ 0.11 & 0 & 0 & 0 & 0.17 & 0.23 & 0.24 & 0.25 \\ 0.16 & 0.27 & 0.28 & 0.29 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.29 & 0.28 & 0.27 & 0.16 \\ 0.25 & 0.24 & 0.23 & 0.17 & 0 & 0 & 0 & 0.11 \\ 0.17 & 0.01 & 0 & 0.05 & 0.21 & 0.20 & 0.19 & 0.18 \\ 0 & 0.16 & 0.16 & 0.15 & 0.14 & 0.14 & 0.13 & 0.13 \end{bmatrix}$$

and then

$$S_8 = \begin{bmatrix}
.06 & .06 & .06 & .07 & .07 & .07 & .07 & 0 & .06 & .06 & .06 & .07 & .07 & .07 & .07 & 0 \\
.08 & .09 & .09 & .10 & .02 & 0 & .002 & .08 & .08 & .09 & .09 & .10 & .02 & 0 & .002 & .08 \\
.05 & 0 & 0 & 0 & .08 & .11 & .12 & .12 & .05 & 0 & 0 & 0 & .08 & .11 & .12 & .12 \\
.07 & .13 & .14 & .14 & 0 & 0 & 0 & 0 & .07 & .13 & .14 & .14 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .14 & .14 & .13 & .07 & 0 & 0 & 0 & 0 & .14 & .14 & .13 & .07 \\
.12 & .12 & .11 & .08 & 0 & 0 & 0 & .05 & .12 & .12 & .11 & .08 & 0 & 0 & 0 & .05 \\
.08 & .002 & 0 & .02 & .10 & .09 & .09 & .08 & .08 & .002 & 0 & .02 & .10 & .09 & .09 & .08 \\
0 & .07 & .07 & .07 & .07 & .06 & .06 & .06 & 0 & .07 & .07 & .070 & .070 & .06 & .06 & .06 \\
.32 & .32 & .31 & .03 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .26 & .29 & .29 & .14 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
.44 & .13 & 0 & 0 & 0 & 0 & .14 & .28 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & .31 & .44 & .23 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & .21 & .44 & .34 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & .10 & .44 & .44 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
.25 & .25 & .25 & .25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & .25 & .25 & .25 & .25 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.$$

Its fixed left eigenvector

$$s_8 = [1.35\ 0.66|1.35\ 0.67|1.34\ 0.67|1.34\ 0.67|1.33\ 0.67|1.32\ 0.67|1.32\ 0.67|1.31\ 0.66].$$

Hence, the approximate invariant density $h_8 = [8.97, 8.95, 8.94, 8.93, 8.88, 8.85, 8.82, 8.78]$.

Bibliography

[1] A.T. Bharucha-Reid. *Elements of the Theory of Markov Process and Their Applications.*

[2] Abraham Boyarsky and Pawel Góra. *Laws of Chaos: Invariant measures and dynamical systems in one dimension*, Birkhaüser, New York, 1997.

[3] Gary Froyland. *Ulam's method for random interval maps.* Nonlinearity, 12: 1029-1052, 1999.

[4] Gary Froyland. *Approximating physical invariant measures of mixing dynamical systems in higher dimensions.* Nonlinear Analysis, Theory, Methods and Applications, 32(7):831-860,1998.

[5] H. Hochstadt. *Integral Equations,* John Wiley and Sons, New Yorke, London, Sydney and Toronto, 1973, 262-264.

[6] Brian R. Hunt. *Estimating invariant measures and Lyapunov exponents.* Ergodic Theory and Dynamical Systems, 16:735-749, 1996.

[7] Zbigniew S. Kowalski. *Ergodic properties of skew porducts with Lasota-Yorke type maps in the base.* Studia Mathematica, 106(1):45-57, 1993.

[8] Andrzej Lasota and James A. Yorke. *On the existence of invariant measures for piecewise monotonic transformations.* Transactions of the American Mathematical Society, 186:481-488, 1973.

[9] Tien-Yien Li. *Finite approximation for the Frobenius-Perron Operator.* A solution to Ulam's conjecture. Journal of Approximation Theory, 17: 177-186, 1976.

[10] Sean P. Meyn and Richard L. Tweedie. *Markov Chains and Stochastic Stability.* Communications and Control Engineering. Springer-Verlag, London, 1993.

[11] Takehito Morita. *Asymptotic behavior of one-dimensional random dynamical systems.* Journal of the Mathecal Society of Japan, 37(4):651-663, 1985.

[12] Steve Papadakis. *The exsistence of absolutely continuous invariant measures for a class of piecewise monotonic transformations.* M.Sc. thesis, Concordia University, 1979.

[13] S. Pelikan. *Invariant densities for random maps of the intervals.* Transactions of the American Mathematical Society, 281(2):813-825, 1984.

[14] Paul J. Schweitzer. *Perturbation theory and finite Markov chains.* Journal of Applied Probability, 5(3):401-404, 1968.

[15] S. M. Ulam. *Problems in Modern Mathematics.* Interscience, 1964.

[16] Lai-Sang Young. *Ergodic theory of differentiable dynamical systems.* In B. Branner and P. Hjorth, editors, Real and Complex Dynamical Systems, 293-336, Dordrecht, 1995.