RANDOM WALKS ON A LATTICE WITH DETERMINISTIC LOCAL DYNAMICS

A Dissertation

Presented to

the Faculty of the Department of Mathematics

University of Houston

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

By

Filiz Tümel

August 2012

RANDOM WALKS ON A LATTICE WITH DETERMINISTIC LOCAL DYNAMICS

Filiz Tümel

APPROVED:

Dr. Andrew Török, Committee Chair
Department of Mathematics, University of Houston

Dr. Michael J. Field
Department of Mathematics, University of Houston

Dr. William Ott
Department of Mathematics, University of Houston

Dr. Gemunu Gunaratne
Department of Physics, University of Houston

Dean, College of Natural Sciences and Mathematics

Acknowledgements

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the preparation and completion of this study.

First, my utmost gratitude to my advisor Dr. Andrew Török for the support and help he provided me throughout my graduate education;

My committee members, Dr. William Ott and Dr. Gemunu Gunaratne, for their participation, and Dr. Michael Field, for his thoughtful comments and wonderful support. I will never forget;

Dr. Manfred Denker, for illuminating me as to the relevance of this study to that of iterated function systems and for his encouragement and guidance toward a path in academia;

My friends, especially Dr. Aarti Jajoo who helped me to write my dissertation successfully, and Rebecca Chen, Nikita Agarwal, Kidist Terefe Zeleke, and Vandana Sharma, who have brightened many a dull lunchtime;

My deepest thanks go to Justin Schoneman for his love and understanding;

And my family for supporting me throughout all my studies, my dad who always believed in me, my mom for providing a nurturing environment, and my brother for his endless enthusiasm about my academic progress;

Finally, though he cannot possibly know how much of a help he has been, I would like to thank Cauchy, who comforted me after hard days of work.

Abstract

In this thesis we prove statistical properties of dynamical systems on a lattice with randomly occurring jumps. The original model of this type, called a *hybrid system*, was introduced by E. Kobre and L. S. Young in 2007. We use different methods to derive the drift rate and the averaged Central Limit Theorem. We generalize their results to piecewise uniformly expanding maps with *countable* partitions. We obtain an upper bound for the speed of convergence in the Central Limit Theorem and prove that the convergence is with tight maxima. We prove Large Deviation results. We also prove a quenched Central Limit Theorem, subject to a condition that can be verified following existing techniques for maps that are sufficiently expanding. Finally, we expand the drift rate results and averaged Central Limit Theorem to certain *non-uniformly* expanding systems.

Contents

List of Figures

CHAPTER 1

Introduction and Historical Remarks

1.1 Introduction

Statistical mechanics provides a framework for relating the microscopic properties of individual atoms and molecules to the macroscopic bulk properties of materials that can be observed in everyday life. The ability to make macroscopic predictions based on microscopic properties is the main goal of statistical mechanics.

Particle systems, as they appear in statistical mechanics, have been an important model motivating much development in the field of Dynamical Systems. While these are deterministic systems in the microscopic level, the evolution law is too complicated. Instead one uses a stochastic approach to such systems.

More generally ideas from statistical mechanics have been brought to the setting of dynamical systems by Y. Sinai [Sin68], D. Ruelle [Rue78] and R. Bowen [Bow70] in the 1970s. The objects they introduced are called SRB measures and they play an important role in the ergodic theory of dissipative dynamical systems.

On the other hand corresponding problems have been studied also in the context of the theory of random maps which was much developed by Y. Kifer, [Kif98] and L. Arnold, [Arn03]. The main idea of their approach is that evolution of many physical systems can be better described by compositions of different maps rather than by repeated application of exactly the same transformation. Y. Kifer proved the existence of equilibrium states for random uniformly expanding systems. This theory is applied to random networks, fractal dimensions of random sets and other models.

In this dissertation we combine these two approaches with a model given by E. Kobre and L. S. Young in [KY07]. Some other results in this direction can be found in [CD09], [Len06]. The common goal of their approach is closing the gap between deterministic and stochastic dynamics.

The model in [KY07] has an extended phase space given by a lattice structure and moving particles on that lattice. The dynamics of particles is defined by microscopic rules. In particular, the local dynamics is given by iterating the same piecewise uniformly expanding circle map. They introduce random jumps from one node of the lattice to another. The jumps give the macroscopic behavior of the particle and depend on the state of the local dynamics. The main question is "What can we say about the global behavior of the particle on the lattice by looking at the local behavior?". In that sense the model is an attempt to understand the particle systems.

One of the main goal of our work is generalizing the local dynamics of the model given in [KY07]. We are able to prove the result for processes with local dynamics that are given by piecewise uniformly expanding interval maps with countably many partition. We prove the drift rate and Central Limit Theorem for these maps. This extends the results to a more general class of maps. These maps may have non-expanding parts, but they induce uniformly expanding interval maps (See Chapter 2 for the definition).

We are also interested in obtaining more information about the statistical behavior of these models. We show the speed of the convergence in the Central Limit Theorem is $1/\sqrt{n}$. We prove that the convergence in the Central Limit Theorem has a property called "tight maxima" which is stronger than only converging to a normal distribution. We also give the Large Deviation estimate.

Our work uses a different method than [KY07]. We use Perturbation Theory which is in general used for deterministic dynamical systems. We show that with some modifications the traditional methods for deterministic dynamical systems can be used in the stochastic set up.

In this work we also attempt to give the Central Limit Theorem in the quenched sense. The averaged approach is the first possibility to understand a hard problem. The quenched result helps to understand more of the behaviour of the process. In Chapter 7, we are able to prove the quenched Central Limit Theorem for piecewise uniformly expanding random dynamical systems subject to a condition on the higher dimensional dynamics. We expect that the condition to be satisfied for our random maps by following the method of [KL05]. The higher dimensional maps that we consider are discussed at the end of Chapter 4.

CHAPTER 2

Setting and Results

In Chapter 2 we describe the model and state the results. This model was first introduced by L. S. Young and E. Kobre in [KY07]. Kobre and Young consider identical expanding circle maps as local dynamics of a lattice system. We describe first the general model of [KY07]. Then we introduce special cases of this model where our results apply. These special cases are more general than those considered in [KY07].

Remark 2.0.1. *The system that we consider lives on a lattice indexed by, say,* $\mathbb{Z}$*. After describing it, in order to keep the notation simpler, we continue with a lattice indexed by* $\mathbb{N}$*.* *See **Important Remark** on page 7.*

2.1 The General Hybrid System

The general hybrid model is given on a lattice where we place a local dynamics at each node. The local dynamics is given by a deterministic map and the global dynamics is a result of random jumps from one node to another. During the iteration of the deterministic map in one node, the point may jump to a neighbor node with some probability, depending on its present state. Below is the formal description.

Let $I := [0,1]$, $X := I^d$ and $\mathbb{Z}$ the set of integers. Consider a one dimensional lattice indexed by $\mathbb{Z}$ where each node is a copy of X. We denote the i^{th} copy of X in the lattice by X_i for $i \in \mathbb{Z}$. For each $i \in \mathbb{Z}$ the local (deterministic) dynamics is given by a map $\tau_i : X_i \to X_i$. For each $i \in \mathbb{Z}$ we define four subsets L_{ij} and R_{ij} of X_i, $j = 1, 2$, with R_{i1} and R_{i2} disjoint. The *jump maps* are $\varphi_{i1} : R_{i1} \subset X_i \to L_{i+1,1} \subset X_{i+1}$ and $\varphi_{i2} : R_{i2} \subset X_i \to L_{i-1,2} \subset X_{i-1}$. We call R_{ij} and L_{ij} the *outgoing,* respectively *incoming*, *jump sets.*

The *jump probabilities* are $0 < p_{ij} < 1$, $j = 1, 2$.

Here is how the system evolves. Let $x_0 \in X_i$, $i \in \mathbb{Z}$, be the initial state. One step of the dynamics consists of the following: we apply the map τ_i and check whether $\tau_i(x_0) \in X_i$ is in the jump set $R_{i1} \cup R_{i2} \subset X_i$. If not, then the evolution continues from $\tau_i(x_0)$. Otherwise, the point can jump to a neighboring site, according to the following rule. If $\tau_i(x_0) \in R_{i1}$ then there are two possibilities: with probability p_{i1} the point jumps to the right node to $\varphi_{i1}(\tau_i(x_0)) \in L_{i+1,1} \subset X_{i+1}$ and with probability $1 - p_{i1}$ it stays at $\tau_i(x_0) \in R_{i1} \subset X_i$. If $\tau_i(x_0) \in R_{i2}$ then with probability p_{i2} the point jumps to the left to $\varphi_{i2}(\tau_i(x_0)) \in L_{i-1,2} \subset X_{i-1}$ and with probability $1 - p_{i2}$ it stays at $\tau_i(x_0) \in R_{i2} \subset X_i$.

We keep iterating the system according to this rule. We study the statistical properties of the system on $\cup_{i\in\mathbb{Z}} X_i$.

More generally, one can define the process with more than two jump intervals, and the jump maps need not to be to the neighboring sites.

Important Remark (approach, notation from now on). *Our approach is to translate the hybrid dynamics into a random dynamical system, following [KY07]. For simplicity, we state the results and explain the method of proof below for dynamics on $\cup_{i\in\mathbb{N}} X_i$, with only one jump interval, to the right. Here $\mathbb{N} = \{0, 1, 2, \dots\}$ is the set of natural numbers. For this, one-sided case, the translation yields two maps, applied randomly. If there are T jump intervals then one needs 2^T maps.*

The statements of the results presented below adjust, mutatis mutandis, to the general case. The proofs are given already for arbitrary finite random systems, so need no modification.

Therefore, from now on the jump maps are $\varphi_i : R_i \subset X_i \to L_{i+1} \subset X_{i+1}$ with jump probabilities p_i, where $i \in \mathbb{N}$.

The process can be seen as a Markov process on the state space $\mathcal{X} = X_0 \cup X_1 \cup X_2 \cup \dots$. Since each point in $\mathcal{X}$ belongs to only one X_i which is just a copy of X, we identify $\mathcal{X}$ with $\mathbb{N} \times X$. Then each point on $\mathcal{X}$ can be given as (i, x) where $i \in \mathbb{N}$ indicates the location in the lattice and $x \in X$ indicates the particular point in X_i. When the process is at a point of X_i for some $i \in \mathbb{N}$, we say the system is at site i.

The transition probabilities $\mathbb{P}$ for the Markov process are given by

$$\mathbb{P}((i,x),(i,\tau_i(x))) = \begin{cases} 1, & \text{if } \tau_i(x) \notin R_i, \\ 1-p_i, & \text{if } \tau_i(x) \in R_i. \end{cases}$$

$$\mathbb{P}((i,x),(i+1,\varphi_i(\tau_i(x)))) = \begin{cases} 0, & \text{if } \tau_i(x) \notin R_i, \\ p_i, & \text{if } \tau_i(x) \in R_i. \end{cases}$$

and zero otherwise.

If a point (i,x) in site X_i moves to a site X_{i+1} under this process we say the point $x \in X_i$ "jumped to the right". Whenever a point $x \in X_i$ ends up in R_i we decide whether the point jumps to the right or not.

Definition 2.1.1. *Consider the Markov process on $\mathcal{X}$ defined above. For any point $(i,x) \in \mathcal{X}$ we introduce the random variable $\mathcal{J}_n(i,x)$ which is the number of times the point $x \in X_i$ has jumped to the right in the first n iterates of the Markov process.*

Definition 2.1.2 (drift rates)**.** *Let $(i,x) \in \mathcal{X}$ be the initial state of the Markov process. Define the pointwise drift rate of (i,x) to be $\alpha \in \mathbb{R}$ if*

$$\lim_{n\to\infty} \frac{\mathcal{J}_n(i,x)}{n} = \alpha \quad a.s.$$

Remark 2.1.3. *The* almost surely *in the above definition refers to the choices one makes when entering the jump intervals.*

To start the process we give an initial probability distribution μ_0 on the state space $\mathcal{X}$ of the Markov process. The measure of the i^{th} copy of the lattice under μ_0,

$\mu_0(X_i)$, is called *the weight of site* i with respect to μ_0. We are also interested in how the initial distribution μ_0 evolves under the Markov process. Denote by $\mu_1, \mu_2, \ldots$ the probability distributions on $\mathcal{X}$ after the $1^{st}, 2^{nd}, \ldots$ iterates of the Markov process. One can also describe the asymptotic behaviour of the center of mass:

Definition 2.1.4 (center of mass)**.** *Let μ be a probability distribution on $\mathcal{X}$. Define* the center of mass *of μ to be*

$$\mathfrak{C}(\mu) = \sum_i i\, \mu(X_i)$$

whenever the sum converges absolutely.

Definition 2.1.5 (absolutely continuous distribution)**.** *We say that μ is an* absolutely continuous *distribution on $\mathcal{X}$ if $\mu|_{X_i}$ is absolutely continuous with respect to the (normalized) Lebesgue measure on X_i for each i.*

Definition 2.1.6 (drift rate of the center of mass)**.** *Let μ_0 denote an initial probability distribution on $\mathcal{X}$, and $\mu_1, \mu_2, \ldots$ denote the distributions on $\mathcal{X}$ after the $1^{st}, 2^{nd}, \ldots$ iterates of the Markov process. Then* the drift rate of the center of mass *of μ_0 is the limit*

$$\lim_{n\to\infty} \frac{\mathfrak{C}(\mu_n)}{n}$$

whenever the limit exists.

2.2 Some Special Hybrid Systems

In this section we describe special cases of the general model given in the previous section. We give the results for these special cases in the next section.

We assume that the local dynamics is one-dimenensional and translation-invariant. That is, $X = I$ (so $d = 1$), and at each site i the maps τ_i are the same map τ, the jump intervals $R_i = R$ and $L_i = L$ are the same, $\varphi : R \to L$ are the same jump map, the jump probabilities are $p_i = p$ with $0 < p < 1$.

The properties of the local maps and jump maps are given in more detail for each model below.

2.2.1 Model I: Uniformly Expanding Maps, $\mathcal{T}_0$ and $\mathcal{T}_1$

First we define a set of functions that we call $\mathcal{T}_1$ and then we give the definition for Model I.

Definition 2.2.1 (the class $\mathcal{T}_1(Y)$)**.** *Let Y be a finite union of closed bounded disjoint intervals in $\mathbb{R}$, $\mathfrak{m}$ denote the normalized Lebesgue measure on Y and $\tau : U \to Y$ be a continuous map with $U \subset Y$ open and dense, and $\mathfrak{m}(U) = 1$. Let $S = Y \setminus U$. By taking the closure of each connected component of U we obtain a **countable** family β of closed intervals with disjoint interiors such that $\bigcup_{B\in\beta} B \supset U$ and $B \cap S$ consists exactly of the endpoints of B for each $B \in \beta$. Fix $\lambda > 1$.*

The functions $\mathcal{T}_1(Y)$ are described by the following properties:

(i) the restriction τ_B of τ to an interval $B \cap U$, $B \in \beta$, admits an extension to a homeomorphism of B, and τ_B is differentiable with $|\tau_B'| > \lambda > 1$;

(ii) the function $g(x)$ defined by

$$g(x) = \begin{cases} 1/|\tau'(x)|, & \text{if } x \in U, \\ 0, & \text{if } x \in S. \end{cases} \tag{2.1}$$

is of bounded variation on Y.

We give the above definition for a more general set Y since we define functions in the following chapters on sets other than I. But the local dynamics is given by a map defined on I as explained below.

Definition 2.2.2 (Model I)**.** *The Markov process given in the previous section is called* Model I *if the followings are satisfied:*

- *The local phase space is the 1-dimensional interval I.*
- *The local dynamics is given by $\tau : U \to U$ at each node $i \in \mathbb{N}$ where $\tau \in \mathcal{T}_1(I)$.*
- *The outgoing jump interval $R \subset U$ is arbitrary, but the same for each node $i \in \mathbb{N}$.*
- *The incoming jump interval L is U at each node $i \in \mathbb{N}$.*
- *The jump map $\varphi : R \to Y$ is differentiable with $|\tau'| \cdot |\varphi'| > \lambda > 1$ whenever τ' exists, where λ is given in Definition 2.2.1.*

The assumption of having a countable partition is the first generalization of the results in [KY07]. In their paper they use uniformly expanding maps with finite partitions.

Definition 2.2.3 (the class $\mathcal{T}_0$)**.** *We denote by $\mathcal{T}_0$ the set of maps in $\mathcal{T}_1$ with finite β.*

For some of the theorems, we need the corresponding random system to be weak-mixing. For the definition of weak-mixing, see Definition 6.0.3.

For the definition of the corresponding random dynamical system see Definition 9.1.1.

2.2.2 Model II: Non-uniformly Expanding Maps, $\mathcal{T}_2$

In this section first we define a more general set of maps that we call $\mathcal{T}_2$ and then give a more general Markov process than the Model I where the local dynamics is given by maps in $\mathcal{T}_2$.

Definition 2.2.4 (the class $\mathcal{T}_2(X)$)**.** *Let X be a finite union of closed bounded disjoint intervals in $\mathbb{R}$ and $\mathfrak{m}$ denote the normalized Lebesgue measure on X.*

The class $\mathcal{T}_2(X)$ consists of maps $\varsigma : X \to X$ for which the induced map τ on $Y \subset X$ is in $\mathcal{T}_1(Y)$, for $Y \subset X$ as in Definition 2.2.1.

We define next what it means for a map $\varsigma : X \to X$ to induce a mapping τ on $Y \subset X$. This class contains certain types of non-uniformly expanding maps like Pomeau-Manneville maps and maps that have a Young Tower structure as introduced in [You99].

Definition 2.2.5 (induced maps)**.** *Let $\bigcup_{i=1}^{\infty} Y_i \subset Y \subset X$ be a disjoint union of open intervals such that $\mathfrak{m}\left(\bigcup_{i=1}^{\infty} Y_i\right) = \mathfrak{m}(Y)$.*

A dynamical system $\varsigma : X \to X$ *is said to* induce *the map*

$$\tau : \bigcup_{i=1}^{\infty} Y_i \to Y \qquad \text{if} \qquad \tau(x) = \varsigma^{\mathfrak{R}(x)}(x)$$

where $\mathfrak{R} : \bigcup_{i=1}^{\infty} Y_i \to \mathbb{N}^+$ *is the* first return time *to* Y *and* $\mathfrak{R}$ *is constant on each partition element* Y_i*.*

We say that the induced system $\tau : \bigcup_{i=1}^{\infty} Y_i \to Y$ *has* summable return times *if*

$$\sum_{i=1}^{\infty} \mathfrak{R}_i \mathfrak{m}(Y_i) < \infty$$

where $\mathfrak{R}_i = \mathfrak{R}|_{Y_i}$*.*

Definition 2.2.6 (Model II)**.** *The Markov process given in the general hybrid model is called* Model II *if the following are satisfied:*

- *The local phase space is the 1-dimensional interval* I*.*
- *At each node the local map is the same map* $\varsigma : I \to I$ *with* $\varsigma \in \mathcal{T}_2(I)$*. Let* $\tau \in \mathcal{T}_1(Y)$ *be the induced map of* ς *on* $Y \subset X$*.*
- *At each node the outgoing jump interval is the same interval* $R \subset Y$ *and the incoming jump interval is* $L \subset Y$*.*
- *The jump map* $\varphi : R \to Y$ *is differentiable with* $|\tau'| \cdot |\varphi'| > \lambda > 1$ *whenever* τ' *exists, where the value* λ *corresponding to* τ *is from Definition 2.2.1.*
- *The return time of the induced system is summable.*

2.3 Results

For the special cases given in the previous sections note that the local maps τ_i and the jump probabilities p_i do not depend on the site number i. Then for Models I and II the random variable that counts the number of jumps in the first n iterates does not depend on the initial site. So we consider the random variable $\mathcal{J}_n(0, x)$ and denote it by $\mathcal{J}_n(x)$ for $x \in X_0$. The results are given for the random variable $\mathcal{J}_n(x)$. For simplicity we give the results for initial absolutely continuous probability distributions μ_0 with $\mu_0(X_0) = 1$ and we say μ_0 is concentrated on X_0. The results can be generalized to an absolutely continuous initial distribution μ_0 with $\mathfrak{C}(\mu_0) < \infty$.

Theorem 2.3.1 (Drift Rate)**.** *Let the Markov process on $\mathcal{X} = \mathbb{N} \times I$ be as in Model I or Model II, with initial site X_0.*

Then there is $\alpha \in \mathbb{R}$ such that for $\mathfrak{m}$-almost every initial state $x_0 \in X_0$ we have

$$\lim_{n\to\infty} \frac{\mathcal{J}_n(x_0)}{n} = \alpha \qquad a.s.$$

Let μ_0 be the absolutely continuous initial probability distribution on $\mathcal{X}$. We also give the drift rate result for the center of mass.

Theorem 2.3.2. *For the Markov process on $\mathcal{X} = \mathbb{N} \times I$ given in Model I or Model II, if μ_0 is an absolutely continuous initial probability distribution concentrated on X_0, then*

$$\lim_{n\to\infty} \frac{\mathfrak{C}(\mu_n)}{n} = \alpha$$

where $\alpha \in \mathbb{R}$ is the value given in Theorem 2.3.1. The same is true for absolutely continuous distributions μ_0 with finite center of mass.

Given an initial probability distribution concentrated on X_0, there exists a unique measure $\mathbb{P}_{\mu_0}$ on sequences of observations $\{x_0, x_1, x_2, \ldots\}$ where each $x_i \in X$, x_0 is chosen with respect to μ_0 and the sequence is given with respect to the transition probability $\mathbb{P}$ of the Markov process. In the following theorem we state that the random variable $\mathcal{J}_n$ satisfies the Central Limit Theorem with respect to the measure $\mathbb{P}_{\mu_0}$.

Theorem 2.3.3 (Central Limit Theorem)**.** *For a weak-mixing Markov process on $\mathcal{X} = \mathbb{N} \times I$ given in Model I or Model II, if μ_0 is an absolutely continuous initial probability distribution concentrated on X_0, then for every interval $J \subset \mathbb{R}$ we have*

$$\lim_{n\to\infty} \mathbb{P}_{\mu_0} \left\{ \frac{\mathcal{J}_n - n\alpha}{\sqrt{n}} \in J \right\} = \frac{1}{\sqrt{2\pi}\sigma} \int_J e^{-u^2/2\sigma} du$$

for some $\sigma > 0$, where $\mathbb{P}_{\mu_0}$ is the measure associated to the transition probability $\mathbb{P}$ of the Markov process.

The following results give more details about the convergence of the process to a normal distribution for local maps in $\mathcal{T}_1(I)$. The first theorem gives the rate of convergence to the normal distribution. The second theorem states that the convergence to a normal distribution is with tight maxima. And the last theorem is the Large Deviation estimate which measures the probability of outliers in the convergence of Theorem 2.3.1.

Theorem 2.3.4 (Speed for the Central Limit Theorem)**.** *For a weak-mixing Markov process on $\mathcal{X} = \mathbb{N} \times I$ given in Model I and μ_0 absolutely continuous, the convergence to a normal distribution given by Theorem 2.3.3 has speed $\mathcal{O}(n^{-1/2})$: there exists*

$C > 0$ such that for every interval $J \subset \mathbb{R}$ we have

$$\left| \mathbb{P}_{\mu_0} \left\{ \frac{\mathcal{J}_n - n\alpha}{\sqrt{n}} \in J \right\} - \frac{1}{\sqrt{2\pi}\sigma} \int_J e^{-u^2/2\sigma} du \right| \leq \frac{C}{\sqrt{n}}.$$

Theorem 2.3.5 (Tight Maxima)**.** *For the weak-mixing Markov process on $\mathcal{X} = \mathbb{N} \times I$ given in Model I and μ_0 absolutely continuous, the convergence to a normal distribution given in Theorem 2.3.3 is with tight maxima: for every $\epsilon > 0$ there exists $c > 0$ such that*

$$\mathbb{P}_{\mu_0} \left\{ \max_{1 \leq k \leq n} \frac{|\mathcal{J}_k - k\alpha|}{\sqrt{n}} > c \right\} \leq \epsilon, \text{ for every } n \geq 1.$$

Theorem 2.3.6 (Large Deviation)**.** *For the weak-mixing Markov process on $\mathcal{X} = \mathbb{N} \times I$ given in Model I and μ_0 absolutely continuous, there exists $A > 0$ such that for all $a \in (0, A)$*

$$\mathbb{P}_{\mu_0} \left\{ \frac{|\mathcal{J}_n - n\alpha|}{n} \geq a \right\} \leq Ce^{-Ca^2 n} \qquad \text{for some } C > 0.$$

For the weak-mixing Markov process on $\mathcal{X} = \mathbb{N} \times I$ we also obtain the quenched version of the Central Limit Theorem for Model I, if the local maps are in $\mathcal{T}_0(I)$. Recall that these are the maps used in [KY07]. The quenched Central Limit Theorem states that the Markov process satisfies the Central Limit Theorem for almost all jump choices that is made during the evolution of the process. The formal statement is given for the corresponding random dynamical system, see Theorem 7.4.6, and the correspondence to the Markov process is explained in Section 9.2.

CHAPTER 3

Preliminaries

In Chapter 3 we review the material needed for the following sections including functions of bounded variations in dimension one, the skew product realization of random dynamical systems and uniform ergodic theory.

3.1 Functions of Bounded Variation in 1D

The definitions and the notations in this section follow the book by Gora and Boyarski, see [KL05].

Definition 3.1.1. *For a function $f : I \to I$, the total variation of f is given by*

$$\bigvee_I (f) = \sup\{\sum_{i=1}^{r} |f(\zeta_i) - f(\zeta_{i-1})|\}$$

where supremum is taken over all finite partitions of I, $0 = \zeta_1, \zeta_2, \ldots, \zeta_r = 1$. We say that f is of bounded variation if $\bigvee_I (f) < \infty$ and denote the set of functions of bounded variation on I by $BV(I)$.

The set $BV(I)$ is clearly a vector subspace of $L^1(I)$. The set of all functions $f \in BV(I)$ with $\int_I f d\mathfrak{m} = 0$ is also a vector space in $BV(I)$. We denote the set of such functions by $BV_0(I)$ and restrict ourselves to $BV_0(I)$ to simplify the calculations in most of the proofs in following chapters.

Definition 3.1.2. *For each $f \in BV(I)$ define $var(f)$ by*

$$var(f) = \inf \left\{ \bigvee_I (\tilde{f}) : \tilde{f} = f \text{ for } \mathfrak{m}\text{-a.e.} \right\}$$

and the BV-norm by $\|f\|_{BV} := \|f\|_1 + var(f)$.

The space $BV(I)$ equipped with BV-norm is a Banach space.

Definition 3.1.3. *We say that a set of functions on I is strongly compact in L^1-norm if every sequence of functions has a convergent subsequence that converges in L^1-norm to an $L^1(I)$ function.*

Proposition 3.1.4. *If a set of functions of bounded variation is bounded with respect to the BV-norm, then the set of functions is strongly compact in L^1-norm.*

Proof. See [BG97], Chapter 2, Proposition 2.3.4. □

3.2 Uniform Ergodic Theory

The aim of the following section is to give the Uniform Ergodic Theorem which we use in Chapter 5 to give the spectral properties of the random Perron-Frobenius operator. First we fix the notation for the section and give the related definitions. We follow the book of Krengel and Brunel, see Section 2.2 of [KB75]. More details on spectral theory can be found in the Linear Operators, Part I by Dunford-Schwartz, see [DS09].

Definition 3.2.1. *The spectrum $\sigma(\mathcal{P})$ of a bounded linear operator $\mathcal{P}$ on a Banach space $\mathfrak{X}$ consists of all complex numbers λ such that $\lambda I - \mathcal{P}$ is not invertible. The complement of the spectrum is the* resolvent set, *$\varrho(\mathcal{P}) = \mathbb{C} \setminus \sigma(\mathcal{P})$ and $\rho(\mathcal{P}) = \sup\{|\lambda| : \lambda \in \sigma(\mathcal{P})\}$ is called the* spectral radius.

Note that if $\mathcal{P}$ is a bounded operator on a Banach space, then the inverse of $\lambda I - \mathcal{P}$ is bounded whenever it exists, by the Open Mapping Theorem.

Definition 3.2.2. *An isolated point λ_0 of $\sigma(\mathcal{P})$ is called a* pole of order n *if $\mathcal{S}(\lambda, \mathcal{P}) = (\lambda I - \mathcal{P})^{-1}$ has a Laurent expansion around λ_0 given by*

$$\mathcal{S}(\lambda, \mathcal{P}) = \sum_{k=-n}^{\infty} B_k (\lambda - \lambda_0)^k \quad \textit{with } B_{-n} \neq 0,$$

B_{-1} is called the residue *of $\mathcal{S}(\lambda, \mathcal{P})$ at λ_0.*

Definition 3.2.3. *A linear operator $\mathcal{P}$ is called* power bounded *if*

$$\sup_n \|\mathcal{P}^n\| < 1,$$

and Cesaro bounded *if*

$$\sup_n \|\frac{1}{n}\sum_{i=0}^{n-1}\mathcal{P}^i\| < 1.$$

where $\|\cdot\|$ *is the operator norm.*

Definition 3.2.4. *A linear operator* $\mathcal{P}$ *on a Banach space* $\mathfrak{X}$ *is called* compact *if the image under* $\mathcal{P}$ *of the unit sphere of* $\mathfrak{X}$ *is conditionally compact.* $\mathcal{P}$ *is called* quasi-compact *if there exists a compact operator* $\mathcal{K}$ *and* $m \in \mathbb{N}$ *such that*

$$\|\mathcal{P}^m - \mathcal{K}\| < 1.$$

Definition 3.2.5. *The operator* $\mathcal{P}$ *is called* uniformly ergodic *if there exists a finite dimensional projection* $\mathcal{K}$ *such that*

$$\lim_{n\to\infty}\left\|\frac{1}{n}\sum_{n=0}^{n-1}\mathcal{P}^n - \mathcal{K}\right\| = 0.$$

Set $A_n(\mathcal{P}) = \frac{1}{n}\sum_{i=0}^{n-1}\mathcal{P}^i$. Now we can give the main theorem of the section which is used in Chapter 5.

Theorem 3.2.6 (Uniform Ergodic Theorem)**.** *Let* $\mathcal{P}$ *be power bounded, quasi-compact linear operator in a Banach space* $\mathfrak{X}$*. Then each power has a representation*

$$\mathcal{P}^n = \sum_{i=1}^{k}\lambda_i^n\mathcal{P}_i + \mathcal{R}^n, n = 1, 2, \ldots$$

where $\lambda_1, \ldots, \lambda_k$ *are the finitely many poles of* $\mathcal{S}(., \mathcal{P})$ *with* $|\lambda_i| = 1$ *and finite multiplicity.* $\mathcal{P}_i$ *is the projection given as the limit of* $A_n(\lambda_i^{-1}\mathcal{P})$ *and* $\mathcal{R}$ *is the quasi-compact operator defined by*

$$\mathcal{R} = \mathcal{P} - \sum_{i=1}^{k}\lambda_i\mathcal{P}_i$$

satisfying

$$\mathcal{P}\mathcal{P}_i = \mathcal{P}_i\mathcal{P} = \lambda_i\mathcal{P}_i,$$
$$\mathcal{P}_i^2 = \mathcal{P}_i, \quad \mathcal{P}_i\mathcal{P}_j = 0,$$
$$\mathcal{P}_i\mathcal{R} = \mathcal{R}\mathcal{P}_i = 0, \quad \textit{for } i = 1, 2, \ldots, k,$$
$$\|\mathcal{R}\| < 1,$$

and

$$\|\mathcal{R}^n\| \leq \frac{M}{(1+\epsilon)^n}$$

where M and ϵ are positive constants independent from n.

Proof. See [KB75], Section 2.2, Theorem 2.8 for the proof. □

3.3 Random Maps

Let $(Y, \mathcal{B}, \mu)$ be a measure space with a probability measure μ on Y. Let $\{T_1, \ldots, T_N\}$ be a set of measurable functions on Y and let $\mathfrak{T}$ represent the random map chosen randomly from the set $\{T_1, \ldots, T_N\}$ with respect to the probability vector $(p_1, \ldots, p_N)$. We can define a deterministic system that gives the same dynamics of the random dynamical system on Y. Let $\{1, \ldots, N\}$ be the symbol set and Ω be the set of all sequences on that symbol set. We first give the definition of a Bernoulli shift on Ω. Then we define the corresponding deterministic map $\mathscr{F}$ on $\Omega \times Y$.

Definition 3.3.1. *Let $\Omega = \{\omega = (\omega_1, \omega_2, \ldots) : \omega_i \in \{1, 2, \ldots, N\}\}$ be the set of one sided sequences of symbols $\{1, 2, \ldots, N\}$. Let $\boldsymbol{\sigma} : \Omega \to \Omega$ be the left shift map defined by $(\boldsymbol{\sigma}(\omega))_j = \omega_{j+1}$. Let $(p_1, \ldots, p_N)$ be a probability vector on the symbol set*

$\{1, \ldots, N\}$. *Let* π *be the Bernoulli measure defined on* Ω *and* $\mathcal{A}$ *be the* σ*-algebra defined by the infinite product of the* σ*-algebras of the finite set* $\{1, \ldots, N\}$. *We call the measure preserving dynamical system* $(\Omega, \mathcal{A}, \pi, \boldsymbol{\sigma})$ *a* Bernoulli shift.

For the set of constituent functions $\{T_1, \ldots, T_N\}$, let $\Omega = \{1, \ldots, N\}^{\mathbb{N}}$ be the set of corresponding sequences of the symbol set. Let π be the Bernoulli measure on Ω where $(p_1, \ldots, p_N)$ is the probability vector on the random maps. The corresponding deterministic system of the random dynamical system $(\mathfrak{T}, \mu, Y)$ is given by $(\mathscr{F}, \pi \times \mu, \Omega \times Y)$ where

$$\mathscr{F}(\omega, x) = (\boldsymbol{\sigma}(\omega), T_{\omega_1}(x))$$

where $\boldsymbol{\sigma}$ is the left shift map on Ω and ω_1 is the first symbol of the sequence ω.

Definition 3.3.2. *Let* $\mathfrak{T}$ *be a random map on* Y *with the set of constituent functions* $\{T_1, \ldots, T_N\}$ *and probability vector* $(p_1, \ldots, p_N)$. *We say that a probability measure* μ *on* Y *is* stationary *for the random map* $\mathfrak{T}$ *if for every measurable set* $B \subset Y$ *we have*

$$\mu(B) = \sum_{i=1}^{N} p_i \mu(T_i^{-1}(B)).$$

Note that μ is stationary for the random map $\mathfrak{T}$ if and only if $\pi \times \mu$ is an invariant measure for the corresponding skew product realization $\mathscr{F}$, see [Kob05] page 17 for the proof or [Arn03], Example 1.4.7 to see when the product measure is invariant.

Definition 3.3.3. *Let* $\mathfrak{T}$ *be a random map on* Y *with constituent functions* $\{T_1, \ldots, T_N\}$, $\mathcal{B}$ *be the Borel* σ*-algebra on* Y. *We call a measurable set* $J \in \mathcal{B}$ *an* invariant set *under the random map* $\mathfrak{T}$ *if* $\mathfrak{m}(J \triangle \cup_{i=1}^{N} T_i(J)) = 0$.

Definition 3.3.4. *Let $\mathfrak{T}$ be a random map on Y, let $\mathcal{B}$ be the Borel σ-algebra on Y and let μ be a stationary measure for the random system. We call the random system $(\mathfrak{T}, \mathcal{B}, \mu)$* ergodic *if for any invariant measurable set $J \in \mathcal{B}$ we have $\mu(J) = 0$ or $\mu(J) = 1$.*

Again the stationary measure μ is ergodic for $\mathfrak{T}$ in the sense of Definition 3.3.4 if and only if $\mu \times \pi$ is ergodic for the deterministic map $\mathscr{F}$, see [Kob05] page 18 for the proof.

Definition 3.3.5. *Let $\mathfrak{T}$ be a random map on Y with constituent functions $\{T_1, \dots, T_N\}$ and probability vector $(p_1, \dots, p_N)$. The operator $\mathcal{P}_{\mathfrak{T}} : L^1(Y) \to L^1(Y)$ defined by*

$$\mathcal{P}_{\mathfrak{T}} = \sum_{i=1}^{N} p_i \mathcal{P}_{T_i}$$

is called the random Perron-Frobenius operator *of the random map $\mathfrak{T}$ where $\mathcal{P}_{T_i}$ is the Perron-Frobenious operator of the single map T_i given by*

$$\mathcal{P}_{T_i} f(x) = \sum_{y : T_i y = x} \frac{f(y)}{|T_i'(y)|}.$$

Note that for functions $f, g \in L^1(Y)$ and for arbitrary $h \in L^1(Y)$ we have

$$\begin{aligned}
\int \mathcal{P}_{\mathfrak{T}}(f)\, g h d\mathfrak{m} &= \int f \left(\sum_{i=1}^{N} p_i g h \circ T_i \right) d\mathfrak{m} \\
&= \int f \left(\sum_{i=1}^{N} p_i g \circ T_i \cdot h \circ T_i \right) d\mathfrak{m} \\
&= \sum_{i=1}^{N} p_i \int (f \cdot g \circ T_i)\, h \circ T_i d\mathfrak{m} \\
&= \sum_{i=1}^{N} p_i \int \mathcal{P}_{T_i}(f \cdot g \circ T_i)\, h d\mathfrak{m}.
\end{aligned}$$

Since h is arbitrary we get

$$g\mathcal{P}_{\mathfrak{T}}(f) = \sum_{i=1}^{N} p_i \mathcal{P}_{T_i}\left(f \cdot g \circ T_i\right). \tag{3.1}$$

where $\mathcal{P}_{T_i}$ is the Perron-Frobenius operator of the single map T_i.

3.4 Ornstein Theory

Ornstein Theory is about Bernoulli processes and their entropy. What we are interested in is a criteria for a dynamical system to be a Bernoulli process. First we give the definition of a Bernoulli process and then give the criteria we use in Chapter 6 to show that the dynamical system we are interested in is Bernoulli. For proofs of the results we refer the reader to Section 7 of the book by Donald S. Ornstein, see [Orn74].

Definition 3.4.1. *A dynamical system $(X, \mathcal{B}, \mu, T)$ is said to be* Bernoulli *if it is isomorphic to a Bernoulli shift $(\Omega, \mathcal{A}, \pi, \boldsymbol{\sigma})$, that is there exist measurable, measure preserving functions $\Phi : X \to \Omega$ and $\Psi : \Omega \to X$ such that*

- *for μ-almost all $x \in X$, $\Phi(Tx) = \boldsymbol{\sigma}(\Phi x)$ and $x = \Psi(\Phi x)$,*
- *for π-almost all $\omega \in \Omega$, $\omega = \Phi(\Psi \omega)$.*

Definition 3.4.2. *Let α and β be two measurable partitions. We define a binary operation $\mathfrak{d}$ on partitions with respect to a measure μ by*

$$\mathfrak{d}(\alpha, \beta) = \sum_{A \in \alpha, B \in \beta} |\mu(A \cap B) - \mu(A)\mu(B)|$$

$$= \sum_{A \in \alpha} \sum_{B \in \beta} |\mu(A|B) - \mu(A)|\mu(B).$$

Notation 3.4.3. *Let $T : X \to X$ with a measurable partition β of X. We denote the refined partition by*

$$\beta_n^m = \bigvee_n^m T^i(\beta) = T^n(\beta) \vee T^{n+1}(\beta) \vee \ldots \vee T^m(\beta)$$

for $n \leq m$.

Here is the criteria we use later in Chapter 6 to show that the system we work on is Bernoulli. See [Orn74], Section 7 for the proof.

Theorem 3.4.4. *Let $T : X \to X$ with partition β of X and measure μ on X. If the partition β satisfies*

$$\sup_{l,k \geq 1} \mathfrak{d}(\beta_0^n, \beta_{l+n}^{l+k+n}) \longrightarrow 0, \quad \text{as } n \to \infty$$

then the dynamical system (T, μ) has the Bernoulli property in the sense of Definition 3.4.1.

CHAPTER 4

Stationary Measures For Random Dynamical Systems

In Chapter 4 we assume that Y is a finite union of closed intervals in $\mathbb{R}$. Let $\mathfrak{m}$ be the normalized Lebesgue measure on Y. We prove the existence of absolutely continuous stationary measure for the random map $\mathfrak{T}$ when each of the constituent functions $T_1, T_2, \ldots, T_N$ belongs to the class $\mathcal{T}_1(Y)$ given in Model I, Section 2.2.

4.1 Existence of Stationary Measures

The aim of this section is to prove the existence of an absolutely continuous stationary measure of $\mathfrak{T}$. We use the method that is used by Rychlik to show the existence of an absolutely continuous invariant measure for a single map in $\mathcal{T}_1(Y)$ on the unit interval. See [Ryc83]. First we derive an inequality of the form of Equation (4.1) below for the random Perron-Frobenius operator associated to the random map $\mathfrak{T}$. The inequality is called the Lasota-York inequality for a single expanding maps. Then we use the standard compactness arguments of the fundamental paper of Lasota and Yorke [LY73] to show the existence of a stationary density. Here is the main result of Chapter 4.

Proposition 4.1.1. *Let $\mathfrak{T}$ be a random map defined on Y with constituent functions $T_1, T_2, \ldots, T_N$ and with probability vector $(p_1, p_2, \ldots, p_N)$. If each T_j belongs to the class $\mathcal{T}_1(Y)$ then $\mathfrak{T}$ has an absolutely continuous stationary measure μ.*

First we recall the Lasota-York inequality for a single map in $\mathcal{T}_1(Y)$ which arises from the paper of Rychlik [Ryc83].

Proposition 4.1.2. *Let $T : Y \to Y$ be a map in $\mathcal{T}_1(Y)$. Let $\mathcal{P}_T$ be the Perron-Frobenius operator of T. Then there exist constants $C, R > 0$ and $r \in (0, 1)$ such that for all $f \in BV(Y)$ and for every $n \geq 1$ we have*

$$\|\mathcal{P}_T^n(f)\|_{BV} \leq Cr^n \|f\|_{BV} + R\|f\|_1 \tag{4.1}$$

where $BV(Y)$ is the space of bounded variation functions on Y and where $\|.\|_{BV}$ denotes the BV-norm as in Definition 3.1.2.

Before we prove Proposition 4.1.1 we give some notations for the random dynamical system. The constituent functions belong to the class $\mathcal{T}_1(Y)$. For each T_i

with $i \in \{1, 2, \dots, N\}$, let $\mathcal{A}_i = \{a_{i1}, a_{i2}, \dots\}$ with $a_{ij} < a_{i,j+1}$ denote the countably many discontinuities of the map T_i. Let β_i denote the partition of Y with respect to the map T_i consisting of the closed intervals $I_{ij} = [a_{ij}, a_{i,j+1}]$. For our purposes, a partition will mean a countable family of closed intervals such that each two of them can have only endpoint in common and exhausting Y up to a set of measure zero. Let $U_i = \bigcup_{j=0}^{\infty} \mathrm{int}(I_{ij})$ where "int" is for interior and $S_i = Y \setminus U_i$ with $\mathfrak{m}(S_i) = 0$. The restriction of T_i to I_{ij}, namely T_{ij} satisfies $|T'_{ij}| > \lambda_i > 1$ for some λ_i.

First we show that any composition of the $\mathcal{T}_1(Y)$-maps is still a $\mathcal{T}_1(Y)$-map. For $\ell \in \mathbb{N}$ the ℓ composition of the maps $T_1, T_2, \dots, T_N$ has

$$|(T_{\omega_\ell} \circ \dots T_{\omega_2} \circ T_{\omega_1})'| > \lambda_\ell > 1$$

where $(\omega_1, \omega_2, \dots, \omega_\ell)$ is an arbitrary ℓ sequence with $\omega_k \in \{1, 2, \dots, N\}$ for every $k \in \{1, 2, \dots, \ell\}$. The derivative of the ℓ composition is defined in each partition element of the composition which is obtained by refining the partition after each iterate according to which function is applied at that iterate. Note that we have N^ℓ many different compositions. Let $\{\varphi_j\}$ with $1 \leq j \leq N^\ell$ represent the enumeration of all possible ℓ compositions of constituent maps for $\mathfrak{T}^\ell$, and let $\{q_\omega\}$ be the corresponding probabilities of occurrence. The partition for $\varphi_\omega = T_{\omega_\ell} \circ \dots \circ T_{\omega_2} \circ T_{\omega_1}$ is precisely given as follows:

Let $\mathcal{A}_{\omega_1} = \{a_1^{\omega_1}, a_2^{\omega_1}, \dots\}$ be the countably many discontinuity points of T_{ω_1}, $\mathcal{A}_{\omega_2} = \{a_1^{\omega_2}, a_2^{\omega_2}, \dots\}$ be the discontinuity points of T_{ω_2}, and so on. Let $S_{\varphi_\omega} = \mathcal{A}_{\omega_1} \cup T_{\omega_1}^{-1}(\mathcal{A}_{\omega_2}) \cup (T_{\omega_1}^{-1} \circ T_{\omega_2}^{-1})(\mathcal{A}_{\omega_3}) \cup \dots \cup (T_{\omega_1}^{-1} \circ \dots \circ T_{\omega_{\ell-1}}^{-1})(\mathcal{A}_{\omega_\ell})$ and let $U_{\varphi_\omega} = Y \setminus S_{\varphi_\omega}$. The map φ_ω is defined on U_{φ_ω}. Now we define the partition for φ_ω to be the refinement

β_ω which consists of all the sets of the form $B_1 \cap T_{\omega_1}^{-1}(B_2) \cap (T_{\omega_1}^{-1} \circ T_{\omega_2}^{-1})(B_3) \cap \ldots \cap (T_{\omega_1}^{-1} \circ \ldots \circ T_{\omega_{\ell-1}}^{-1})(B_\ell)$ where $B_1 \in \beta_{\omega_1}$, $B_2 \in \beta_{\omega_2} \ldots B_\ell \in \beta_{\omega_\ell}$.

We choose $\lambda > 1$ to be $\lambda_{\omega_1} \cdot \lambda_{\omega_2} \cdot \ldots \cdot \lambda_{\omega_\ell} > 1$ so the ℓ composition is uniformly expanding on partition β_ω. To conclude that $T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1} \in \mathcal{T}_1(Y)$ the last information we need is given by the following lemma.

Lemma 4.1.3. *For any ℓ composition of constituent $\mathcal{T}_1(Y)$-maps, namely for φ, the map $g_\varphi(x)$ defined by*

$$g_\varphi(x) = \begin{cases} 1/|\varphi'(x)|, & \text{if } x \in U_\varphi \\ 0, & \text{if } x \in S_\varphi \end{cases}$$

is of bounded variation where U_φ, S_φ is given as above so U_φ is of the form $\bigcup_{k=0}^{\infty}(x_k, x_{k+1})$ where $\{x_0, x_1, \ldots\}$ are the discontinuity points of φ and $S_\varphi = Y \setminus U_\varphi$.

Proof. Let $\varphi = T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1}$ be fixed with $T_{\omega_i} \in \{T_1, T_2, \ldots, T_N\}$ and let $\beta = \{I_k\}_{k=0}^{\infty}$ be the partition of Y where the endpoints are in S_φ.

$$\begin{aligned} \bigvee_Y g_\varphi &= \sum_{k=0}^{\infty} \bigvee_{I_k} g_\varphi \text{ since } g_\varphi = 0 \text{ at the endpoins of } I_k, \\ &= \sum_{k=0}^{\infty} \bigvee_{I_k} \frac{1}{|(T_{j_\ell} \circ \ldots \circ T_{j_2} \circ T_{j_1})'|} \\ &= \sum_{k=0}^{\infty} \bigvee_{I_k} \frac{1}{|(T'_{j_\ell}(T_{j_{\ell-1}} \circ \ldots \circ T_{j_1}))(T'_{j_{\ell-1}}(T_{j_{\ell-2}} \circ \ldots \circ T_{j_1}) \ldots (T'_{j_1})|} \end{aligned}$$

The first equality follows from an elementary result on functions of bounded variation given below, see [BG97], Chapter 2.

(*) Given a countable partition $\bigcup_{i=1}^{\infty} I_i$ of I and given $f \in BV(I)$ with $f(x) =$ constant for all $x \in I \setminus \bigcup_{i=1}^{\infty} \mathrm{int}(I_i)$ we have

$$\bigvee_I f = \sum_{I_i \in \mathcal{I}} \bigvee_{I_i} f.$$

By using (*) we see that it is enough to work on each partition element I_k separately to show the inequality. At this point we assume $\ell = 2$ to make the calculations easier and give the rest of the proof for the case $\ell = 2$. The general case can be obtained by induction on ℓ. Note that each map T_{ω_l}, $l = 1, 2$ satisfies $|T'_{\omega_l}| \geq \lambda_{\omega_l} > 1$, so $0 \leq \dfrac{1}{|T'_{\omega_l}(x)|} \leq \dfrac{1}{\lambda_{\omega_l}} < 1$ for any $x \in Y$. Therefore we have

$$\left\| \frac{1}{|T'_{j_l}|} \right\|_{\infty} = \sup_{x \in Y} \frac{1}{|T'_{j_l}(x)|} \leq \frac{1}{\lambda_{j_l}} < 1. \tag{4.2}$$

On each I_k the properties of functions of bounded variation yield the following inequality, see [BG97]:

$$\begin{aligned}
\bigvee_{I_k} \frac{1}{|T'_{\omega_2}(T_{\omega_1})||T'_{\omega_1}|} &\leq \left(\bigvee_{I_k} \frac{1}{|T'_{\omega_2}(T_{\omega_1})|} \right) \left\| \frac{\chi_{I_k}}{|T'_{\omega_1}|} \right\|_{\infty} + \left\| \frac{1}{|T'_{\omega_2}(T_{\omega_1})|} \right\|_{\infty} \left(\bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} \right) \\
&\leq \left(\bigvee_{T_{\omega_1}(I_k)} \frac{1}{|T'_{\omega_2}|} \right) \left(\bigvee_{I_k} \frac{2}{|T'_{\omega_1}|} \right) + \left\| \frac{1}{|T'_{\omega_2}|} \right\|_{\infty} \left(\bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} \right) \\
&\leq \left(\bigvee_{Y} \frac{1}{|T'_{\omega_2}|} \right) \left(\bigvee_{I_k} \frac{2}{|T'_{\omega_1}|} \right) + \left(\bigvee_{Y} \frac{2}{|T'_{\omega_2}|} \right) \left(\bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} \right) \\
&\leq 4 \left(\bigvee_{Y} \frac{1}{|T'_{\omega_2}|} \right) \left(\bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} \right)
\end{aligned}$$

Note that the restriction of T_{ω_1} to I_k is uniformly expanding, therefore it is continuous, monotonic with a continuous inverse $T_{\omega_1}|_{I_k}^{-1}$. Thus it is a homeomorphism and has a finite variation on I_k. Then to get the first term in the second inequality above we use the following result on functions of bounded variation, see [BG97]:

(**) If $f_1, f_2 \in BV(I_1)$ and if $f_1 : I_1 \to I_2$ is a homeomorphism for $I_2 \subset I_1$, then we have

$$\bigvee_{I_2} f_2 = \bigvee_{I_1} f_2 \circ f_1$$

We can continue by summing the variation over all partition elements $\{I_k\}$.

$$\begin{aligned}\sum_{I_k} \bigvee_{I_k} \frac{1}{|T'_{\omega_2}(T_{\omega_1})||T'_{\omega_1}|} &\leq 4 \left(\bigvee_{Y} \frac{1}{|T'_{\omega_2}|} \right) \left(\sum_{I_k} \bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} \right) \\ &= 4 \left(\bigvee_{Y} \frac{1}{|T'_{\omega_2}|} \right) \left(\bigvee_{Y} \frac{1}{|T'_{\omega_1}|} \right) < \infty\end{aligned} \tag{4.3}$$

Note that if $\{I_{j_1 k}\}$ is the partition corresponding to T_{ω_1} where the end points of $I_{\omega_1 k} = [a_{\omega_1 k}, a_{\omega_1 k+1}]$ consists of the discontinuity points of T_{ω_1}, then $I_{\omega_1 k} = [a_{\omega_1 k}, x_{k_1}] \cup [x_{k_1}, x_{k_2}] \cup \ldots \cup [x_{k_{f-1}}, x_{k_f}] \cup \ldots$ for some partition elements of $\{I_k\}$ where $x_{k_1}, x_{k_2}, \ldots$ are discontinuities only for T_{ω_2}. Let $[x_{k_{f-1}}, x_{k_f}]$ be denoted by I_{k_f}. Therefore, we have

$$\sum_{I_k} \bigvee_{I_k} \frac{1}{|T'_{\omega_1}|} = \sum_{I_k} \bigvee_{I_{k_1}} \frac{1}{|T'_{\omega_1}|} + \bigvee_{I_{k_2}} \frac{1}{|T'_{\omega_1}|} + \cdots = \sum_{I_{\omega_1 k}} \bigvee_{I_{\omega_1 k}} \frac{1}{|T'_{\omega_1}|} = \bigvee_{Y} \frac{1}{|T'_{\omega_1}|}$$

which gives the very last equality in Equation (4.3). By induction we can write

$$\begin{aligned}\sum_{k=0}^{\infty} \bigvee_{I_k} \frac{1}{|(T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1})'|} &\leq 4^{\ell-1} \left(\bigvee_{Y} \frac{1}{|T'_{\omega_1}|} \right) \left(\bigvee_{Y} \frac{1}{|T'_{\omega_2}|} \right) \cdots \left(\bigvee_{Y} \frac{1}{|T'_{\omega_\ell}|} \right) \\ &< \infty\end{aligned}$$

□

We give the last inequality we obtain in the proof of Lemma 4.1.3 as a corollary since it is needed later.

Corollary 4.1.4.

$$\bigvee_Y \frac{1}{|(T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1})'|} \leq 4^{\ell-1} \left(\bigvee_Y \frac{1}{|T'_{\omega_1}|} \right) \left(\bigvee_Y \frac{1}{|T'_{\omega_2}|} \right) \cdots \left(\bigvee_Y \frac{1}{|T'_{\omega_\ell}|} \right)$$

Lemma 4.1.5. *If φ is a fixed ℓ combination of the constituent maps and if β is the partition for φ then for $B \in \beta$ we have*

$$m(B) \leq \|g_\varphi\|_\infty \leq (\|g_{T_{\omega_1}}\|_\infty \cdot \ldots \cdot \|g_{T_{\omega_\ell}}\|_\infty)^\ell$$

Proof. We know that the Perron-Frobenious operator for $\varphi \in \mathcal{T}_1(Y)$ is defined by

$$(\mathcal{P}_\varphi f)(x) = \sum_{y:\varphi(y)=x} g_\varphi(y) f(y) \tag{4.4}$$

where g_φ is as in Definition 2.2.1. $\mathcal{P}_\varphi$ preserves the Lebesgue measure $\mathfrak{m}$, see [Bal00], page 73. So

$$\begin{aligned} \mathfrak{m}(B) &= \mathfrak{m}(\chi_B) = \mathfrak{m}(\mathcal{P}_\varphi \chi_B) \\ &= \mathfrak{m}(\sum_{\varphi(y)=x} g_\varphi(y)\chi_B(y)) \\ &\leq \mathfrak{m}(\varphi(B))\|g_\varphi\|_\infty \\ &\leq \|g_\varphi\|_\infty \end{aligned}$$

The rest follows from chain rule. □

Lemma 4.1.3 concludes that any ℓ iterates of the random map $\mathfrak{T} \in \mathcal{T}_1(Y)$, say $\varphi^{(\ell)}$ is still a $\mathcal{T}_1(Y)$-map. Therefore for every $\delta > 0$ there exists an ℓ such that for every x and for every ℓ composition of the constituent maps,

$$g_{\varphi^{(\ell)}} = \frac{1}{|(T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1})'|} < \delta$$

since $|(T_{\omega_\ell} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1})'| > \lambda_{\omega_\ell} \cdot \ldots \cdot \lambda_{\omega_2} \cdot \lambda_{\omega_1} \to \infty$ as $\ell \to \infty$. Hence for all $\varepsilon > 0$ there exists an ℓ such that for every ℓ composition of the constituent maps

$$2\|g_{\varphi^{(\ell)}}\|_\infty + \varepsilon < 1. \tag{4.5}$$

We give four lemmas for a single map $\tau \in \mathcal{T}_1(Y)$. We omit most of their proofs since they can be found in any paper or book where they prove the existence of invariant measures for uniformly expanding maps. We refer the reader to [Lit08], Rychlik Lemma 2-5.

Lemma 4.1.6. *Let $\tau \in \mathcal{T}_1(Y)$, $\tau : Y \to Y$ with countably many partition $\{I_k\}$. Then for the Perron-Frobenius operator $\mathcal{P}_\tau$ we have*

$$\sum_{I_k} \bigvee_Y \mathcal{P}_\tau(f \cdot \chi_{I_k}) = \bigvee_Y (f \cdot g_\tau)$$

where g_τ is defined as in Model I, Equation 2.1 for τ.

Lemma 4.1.7. *Given a finite partition $\mathfrak{Q}$ of Y and any $f \in BV(Y)$ we have for $\tau \in \mathcal{T}_1(Y)$ that*

$$\bigvee_Y (f \cdot g_\tau) \le A_\tau \bigvee_Y f + B_\tau \sum_{K \in \mathfrak{Q}} \int_K |f| d\mathfrak{m}$$

where

$$A_\tau = \|g_\tau\|_\infty + \max_{K \in \mathfrak{Q}} \{\bigvee_K g_\tau\},$$

$$B_\tau = \max_{K \in \mathfrak{Q}} \frac{\bigvee_K g_\tau}{\mathfrak{m}(K)}.$$

The following lemma is the key point to generalize the result of expanding maps on a finitely many partition to a countably many partition, so we also include the proof.

Lemma 4.1.8. *For any $\varepsilon > 0$ there exists a finite partition $\mathfrak{Q}$ for $\tau \in \mathcal{T}_1(Y)$ such that*

$$\max_{K \in \mathfrak{Q}} \bigvee_K g_\tau \leq \|g_\tau\|_\infty + \varepsilon.$$

Proof. The points of discontinuity of $V(x) = \bigvee_{[a,x]} g$ is same as those of $g : [a,b] \to \mathbb{R}$, see [BG97]. Since $0 \leq g \leq 1/\lambda < 1$ the magnitude of the discontinuity of g never exceeds $\|g\|_\infty$ and decays to zero sufficiently fast, so we have

$$\bigvee_Y g \leq M < \infty, \text{ for some } M \in \mathbb{R}.$$

Then for every $x \in Y$ there exists an open interval I_x containing x such that

$$\bigvee_{I_x} g \leq \|g\|_\infty + \varepsilon.$$

Since $\bigcup_{x \in Y} I_x$ covers the compact set Y there exists a finite subcover $\bigcup_{i=1}^{m} I_{x_i}$. Therefore, if we choose a finite partition $\mathfrak{Q}$ which is finer than the subcover $\cup_{i=1}^{m} I_{x_i}$ we get the result. $\square$

Lemma 4.1.9. *Let $g_\ell(x) = \frac{1}{|(\tau^\ell)'(x)|}$ where τ^ℓ is continuous and zero on the points of discontinuities of τ^ℓ. For $\varepsilon > 0$ and $L \in \mathbb{N}$ where L is the smallest integer among $\ell \in \mathbb{N}$ that satisfies*

$$2\,\|g_\ell\|_\infty + \varepsilon < 1$$

there exists a finite partition $\mathfrak{Q}_L$ such that

$$\max_{K\in\mathfrak{Q}_L} \bigvee_Y g_L \leq \|g_L\|_\infty + \varepsilon.$$

If we define A_L and B_L to be

$$A_L := \|g_L\|_\infty + \max_{K\in\mathfrak{Q}_L} \bigvee_Y g_L,$$

$$B_L := \max_{K\in\mathfrak{Q}_L} \frac{\bigvee_Y g_L}{\mathfrak{m}(K)},$$

then $A_L < 1$.

Now we return back to the random dynamical system. We fix $\varepsilon > 0$ and denote by L the smallest integer ℓ such that Equation 4.5 holds for every L composition of the constituent maps. Note that this is different from iterating the same map for L times, but still possible since each T_i for $i = 1, \ldots, N$ is expanding. Then we continue applying the previous lemmas to maps $\varphi^{(L)}$ which are L compositions of the constituent maps.

Lemma 4.1.10. *For L satisfying Equation 4.5 there exists a finite partition $\mathfrak{Q}_{\varphi^{(L)}}$ for each possible L composition of the constituent maps $\varphi^{(L)}$ so that*

$$\max_{J\in\mathfrak{Q}_{\varphi^{(L)}}} \bigvee_J g_{\varphi^{(L)}} \leq \|g_{\varphi^{(L)}}\|_\infty + \varepsilon.$$

We define

$$A_{\varphi^{(L)}} = \|g_{\varphi^{(L)}}\|_\infty + \max_{K\in\mathfrak{Q}_{\varphi^{(L)}}} \bigvee_K g_{\varphi^{(L)}},$$

$$B_{\varphi^{(L)}} = \max_{K\in\mathfrak{Q}_{\varphi^{(L)}}} \frac{\bigvee_K g_{\varphi^{(L)}}}{\lambda(K)},$$

then

$$A_{\varphi^{(L)}} < 1.$$

Proof. Proof is a result of Lemma 4.1.9 where the single map $\tau = \varphi^{(L)}$ and $\ell = 1$. $\square$

We define

$$A_L = \max_{\varphi^{(L)}}\{A_{\varphi^{(L)}}\} \text{ and } B_L = \max_{\varphi^{(L)}}\{B_{\varphi^{(L)}}\}$$

where maximum is taken over all possible L compositions of the constituent maps, so $A_L < 1$.

Remark 4.1.11. *For the fixed $\varepsilon > 0$, for each $i = 1, 2, \dots, L-1$ we know that we can find a finite partition $\mathfrak{Q}_{\varphi^{(i)}}$ for each i composition of the constituent maps by Lemma 4.1.8 such that*

$$\max_{J \in \mathfrak{Q}_{\varphi^{(i)}}} \bigvee_J g_{\varphi^{(i)}} \leq \|g_{\varphi^{(i)}}\|_\infty + \varepsilon.$$

Thus we define

$$A_{\varphi^{(i)}} = \|g_{\varphi^{(i)}}\|_\infty + \max_{K \in \mathfrak{Q}_{\varphi^{(i)}}} \bigvee_K g_{\varphi^{(i)}},$$

$$B_{\varphi^{(i)}} = \max_{K \in \mathfrak{Q}_{\varphi^{(i)}}} \frac{\bigvee_K g_{\varphi^{(i)}}}{\lambda(K)},$$

and

$$A_i = \max_{\varphi^{(i)}}\{A_{\varphi^{(i)}}\} \text{ and } B_i = \max_{\varphi^{(i)}}\{B_{\varphi^{(i)}}\}$$

for $i = 1, \dots, L-1$ where maximum is taken over all possible i compositions of the constituent maps, and then we define

$$C_1 = \max\{A_1, A_2, \dots, A_{L-1}\} \text{ and } C_2 = \max\{B_1, B_2, \dots, B_{L-1}\}.$$

Since each $\varphi^{(i)} \in \mathcal{T}_1(Y)$, we have $C_1, C_2 < \infty$.

Remark 4.1.12.

$$
\begin{aligned}
(\mathcal{P}_{T_{\omega_2}}\mathcal{P}_{T_{\omega_1}}f)(x) &= \mathcal{P}_{T_{\omega_2}}\left(\mathcal{P}_{T_{\omega_1}}f\right)(x)\\
&= \sum_{y:T_{\omega_2}y=x}\frac{(\mathcal{P}_{T_{\omega_1}}f)y}{|T'_{\omega_2}y|}\\
&= \sum_{y:T_{\omega_2}y=x}\frac{\displaystyle\sum_{z:T_{\omega_1}z=y}\frac{f(z)}{|T'_{\omega_1}z|}}{|T'_{\omega_2}y|}\\
&= \sum_{y:T_{\omega_2}y=x}\sum_{z:T_{\omega_1}z=y}\frac{f(z)}{|T'_{\omega_1}z||T'_{\omega_2}y|}\\
&= \sum_{z:T_{\omega_2}\circ T_{\omega_1}z=x}\frac{f(z)}{|(T_{\omega_2}\circ T_{\omega_1})'z|}\\
&= (\mathcal{P}_{T_{\omega_2}\circ T_{\omega_1}}f)(x).
\end{aligned}
$$

Let $U_{\varphi^{(i)}}$ be the partition of the i composition of the constituent maps $\varphi^{(i)}$. The map $\varphi^{(i)}$ is piecewise uniformly expanding on $U_{\varphi^{(i)}}$. Therefore for $i = 1, 2, \dots, L-1$ if ω run over all possible i compositions of the constituent maps which are N^i many and if $q_\omega^{(i)}$ is the probability of having such i composition $\varphi_\omega^{(i)}$ then we have

$$
\mathcal{P}_{T_{\omega_i}}\mathcal{P}_{T_{\omega_{i-1}}}\dots\mathcal{P}_{T_{\omega_1}} = \mathcal{P}_{T_{\omega_i}\circ T_{\omega_{i-1}}\circ\dots\circ T_{\omega_1}} = \mathcal{P}_{\varphi_\omega^{(i)}}
$$

by using the above argument. Then we get

$$
\begin{aligned}
\bigvee_Y \mathcal{P}^i_{\mathfrak{T}}(f) &= \bigvee_Y\sum_{\omega}^{N^i} q_\omega \mathcal{P}_{\varphi_\omega^{(i)}}(f)\\
&\le \sum_{\omega}^{N^i} q_\omega^{(i)}\bigvee_Y \mathcal{P}_{\varphi_\omega^{(i)}}(f)\\
&\le \sum_{\omega}^{N^i} q_\omega^{(i)}\sum_{J\in U_{\varphi^{(i)}}}\bigvee_Y \mathcal{P}_{\varphi_\omega^{(i)}}(f\chi_J)
\end{aligned}
$$

$$\begin{aligned}
&= \sum_{\omega}^{N^i} q_\omega^{(i)} \bigvee_Y (f g_{\varphi_\omega^{(i)}}) \text{ by Lemma 4.1.6,} \\
&\leq \sum_{\omega}^{N^i} q_\omega^{(i)} (A_{\varphi_\omega^{(i)}} \bigvee_Y f + B_{\varphi_j^{(i)}} \|f\|_1) \text{ by Lemma 4.1.7,} \\
&\leq \sum_{\omega}^{N^i} q_\omega^{(i)} (A_i \bigvee_Y f + B_i \|f\|_1) \text{ by Remark 4.1.11,} \\
&\leq (A_i \bigvee_Y f + B_i \|f\|_1)
\end{aligned}$$

since $\sum_{\omega}^{N^i} q_\omega^{(i)} = 1$.

Note that for any fixed i composition we have the same inequality as of the averaged i compositions by following the same steps. For fixed $\varphi^{(i)} = T_{\omega_i} \dots T_{\omega_1}$,

$$\begin{aligned}
\bigvee_Y \mathcal{P}_{\varphi^{(i)}}(f) &\leq \sum_{J \in U_{\varphi^{(i)}}} \bigvee_Y \mathcal{P}_{\varphi^{(i)}}(f \chi_J) \\
&= \bigvee_Y (f g_{\varphi^{(i)}}) \\
&\leq A_{\varphi^{(i)}} \bigvee_Y f + B_{\varphi^{(i)}} \|f\|_1 \\
&\leq A_i \bigvee_Y f + B_i \|f\|_1
\end{aligned}$$

where A_i, B_i as in Remark 4.1.11.

Lemma 4.1.13. *For all $f \in BV(Y)$ and for all $n \in \mathbb{N}$*

$$\bigvee_Y \mathcal{P}_{\mathfrak{T}}^n(f) \leq C_1 A_L^j \bigvee_Y f + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1$$

where $n = Lj + i$ with $i = 1, 2, \dots, L-1$.

Proof.

$$\bigvee_Y \mathcal{P}_{\mathfrak{T}}^n(f) = \bigvee_Y \mathcal{P}_{\mathfrak{T}^L}^j (\mathcal{P}_{\mathfrak{T}^i} f)$$

$$\leq A_L[\bigvee_Y \mathcal{P}_{\mathfrak{T}}^{j-1}(\mathcal{P}_{\mathfrak{T}^i} f)] + B_L\|f\|_1$$

$$\leq A_L[A_L[\bigvee_Y \mathcal{P}_{\mathfrak{T}}^{j-2}(\mathcal{P}_{\mathfrak{T}^i} f)] + B_L\|f\|_1] + B_L\|f\|_1$$

$$= A_L^2 \bigvee_Y \mathcal{P}_{\mathfrak{T}}^{j-2}(\mathcal{P}_{\mathfrak{T}^i} f) + (A_L+1)B_L\|f\|_1$$

$$\ldots$$

$$\leq A_L^j \bigvee_Y (\mathcal{P}_{\mathfrak{T}^i} f) + (\sum_{k=0}^{j-1} A_L^k) B_L\|f\|_1$$

$$\leq A_L^j (C_1 \bigvee_Y f + C_2\|f\|_1) + (\sum_{k=0}^{j-1} A_L^k) B_L\|f\|_1$$

$$\leq A_L^j (C_1 \bigvee_Y f + \frac{C_2}{A_L^j}\|f\|_1) + (\sum_{k=0}^{j-1} A_L^k) B_L\|f\|_1 \text{ since } A_L^j < 1,$$

$$= C_1 A_L^j \bigvee_Y f + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L\|f\|_1.$$

□

Remark 4.1.14. *Again note that we have the above equality for any fixed n sequence of constituent maps, say for $\omega = T_{\omega_n} \circ \ldots \circ T_{\omega_1}$ with $n = Lj + i$ we have*

$$\bigvee_Y \mathcal{P}_{T_{\omega_n}} \ldots \mathcal{P}_{T_{\omega_1}}(f) = \bigvee_Y \mathcal{P}_{T^{(j)}_{\omega_L}} \ldots \mathcal{P}_{T^{(j)}_{\omega_1}} \mathcal{P}_{T^{(j-1)}_{\omega_L}} \ldots \mathcal{P}_{T^{(j-1)}_{\omega_1}} \ldots \mathcal{P}_{T^{(1)}_{\omega_1}}(\mathcal{P}_{T_{\omega_i}} \ldots \mathcal{P}_{T_{\omega_1}} f)$$

$$\leq A_L[\bigvee_Y \mathcal{P}_{T^{(j-1)}_{\omega_L}} \ldots \mathcal{P}_{T^{(j-1)}_{\omega_1}} \ldots \mathcal{P}_{T^{(1)}_{\omega_1}}(\mathcal{P}_{T_{\omega_i}} \ldots \mathcal{P}_{T_{\omega_1}} f)]$$

$$+ B_L\|f\|_1$$

$$\leq A_L[A_L[\bigvee_Y \mathcal{P}_{T^{(j-2)}_{\omega_L}} \ldots \mathcal{P}_{T^{(j-2)}_{\omega_1}} \ldots \mathcal{P}_{T^{(1)}_{\omega_1}}(\mathcal{P}_{T_{\omega_i}} \ldots \mathcal{P}_{T_{\omega_1}} f)]$$

$$+ B_L\|f\|_1] + B_L\|f\|_1$$

$$= A_L^2 \bigvee_Y \mathcal{P}_{T^{(j-2)}_{\omega_L}} \ldots \mathcal{P}_{T^{(j-2)}_{\omega_1}} \ldots \mathcal{P}_{T^{(1)}_{\omega_1}}(\mathcal{P}_{T_{\omega_i}} \ldots \mathcal{P}_{T_{\omega_1}} f)$$

$$
\begin{aligned}
&+ (A_L + 1)B_L\|f\|_1 \\
&\dots \\
&\leq A_L^j \bigvee_Y (\mathcal{P}_{T_{\omega_i}} \dots \mathcal{P}_{T_{\omega_1}} f) + (\sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1 \\
&\leq A_L^j (C_1 \bigvee_Y f + C_2\|f\|_1) + (\sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1 \\
&\leq A_L^j (C_1 \bigvee_Y f + \frac{C_2}{A_L^j}\|f\|_1) + (\sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1 \ \ by\ A_L^j < 1, \\
&= C_1 A_N^j \bigvee_Y f + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_N \|f\|_1.
\end{aligned}
$$

The Lemma 4.1.15 below is the version for random dynamical system of the Lasota-York inequality for a single map given in Proposition 4.1.2. After proving Lemma 4.1.15 we proceed precisely as in the proof of Lasota-York inequality for a single map, see [LY73].

Lemma 4.1.15. *There exists constants $C, R > 0$ and $r \in (0, 1)$ such that for all $f \in BV(Y)$*

$$\|\mathcal{P}_{\mathfrak{T}}^n(f)\|_{BV} \leq Cr^n \|f\|_{BV} + R\|f\|_1. \tag{4.6}$$

Proof.

$$
\begin{aligned}
\|\mathcal{P}_{\mathfrak{T}}^n(f)\|_{BV} &= \|\mathcal{P}_{\mathfrak{T}}^n(f)\|_1 + \mathrm{var}_Y(\mathcal{P}_{\mathfrak{T}}^n(f)) \\
&= \|\mathcal{P}_{\mathfrak{T}}^n(f)\|_1 + \inf_{\hat{f}} \bigvee_Y (\mathcal{P}_{\mathfrak{T}}^n(\hat{f}))
\end{aligned}
$$

Since $\mathcal{P}_{T_i}(f)$ is a contraction for each map T_i, we have

$$\|\mathcal{P}_{\mathfrak{T}}(f)\|_1 \leq \sum_{i=1}^{N} p_i \|\mathcal{P}_{T_i}(f)\|_1 \leq \sum_{i=1}^{N} p_i \|f\|_1 = \|f\|_1,$$

which implies by induction that

$$\|\mathcal{P}^n_{\mathfrak{T}}(f)\|_1 \leq \|f\|_1.$$

Together with Lemma 4.1.13, we have for $n = Lj + i$ that

$$\begin{aligned} \|\mathcal{P}^n_{\mathfrak{T}}(f)\|_{BV} &\leq \|f\|_1 + \inf_{\hat{f}}(C_1 A_L^j \bigvee_Y \hat{f} + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L \|\hat{f}\|_1) \\ &= \|f\|_1 + C_1 A_L^j \inf_{\hat{f}} \bigvee_Y \hat{f} + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1 \\ &\leq \|f\|_1 + C_1 A_L^j \|f\|_{BV} + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L \|f\|_1 \end{aligned}$$

We choose $R = 1 + (C_2 + \sum_{k=0}^{j-1} A_L^k) B_L$ and to have $C_1 A_L^j = C r^n$ we choose $r = A_L^{1/L}$ so we can choose $C = C_1 r^{-L+1}$ which implies

$$\|\mathcal{P}^n_{\mathfrak{T}}(f)\|_{BV} \leq C r^n \|f\|_{BV} + R\|f\|_1.$$

□

Remark 4.1.16. *Note that we have the inequality in Equation (4.6) also for any fixed n-sequence of maps again by Remark 4.1.14, so say for $\omega = T_n \ldots T_1$ we have*

$$\|\mathcal{P}_{T_n} \ldots \mathcal{P}_{T_1}(f)\|_{BV} \leq C r^n \|f\|_{BV} + R\|f\|_1 \tag{4.7}$$

since the coefficients are chosen to be the maximum of all coefficients that works for all L possible compositions of constituent maps.

The above inequality in Equation 4.7 is not needed for this chapter but is used in Chapter 6 to prove a result for a fixed sequence of maps.

Proof of Proposition 4.1.1. For the given random dynamical system, let $\mathbf{1} = \chi_Y$ be the characteristic function on Y. We have $\mathbf{1} \in BV(Y)$. Then by Lemma 4.1.15 we have

$$\|\mathcal{P}_{\mathfrak{T}}^n(\mathbf{1})\|_{BV} \leq Cr^n\|\mathbf{1}\|_{BV} + R\|\mathbf{1}\|_1 = Cr^n + R \leq C + R.$$

Therefore the sequence $\{\mathcal{P}_{\mathfrak{T}}^n(\mathbf{1})\}_{n=0}^{\infty}$ is bounded by a constant in $BV(Y)$ so the time averages, namely

$$\left\{\frac{1}{n}\sum_{j=0}^{n-1}\mathcal{P}_{\mathfrak{T}}^j(\mathbf{1})\right\}_{n=1}^{\infty}.$$

Then by Proposition 3.1.4 we conclude that the set of time averages is strongly compact in L^1-norm as in Definition 3.1.3. Therefore there exists a subsequence of the set of time averages, say $\{f_{n_k}\}_{k=0}^{\infty}$ for

$$f_n = \frac{1}{n}\sum_{j=0}^{n-1}\mathcal{P}_{\mathfrak{T}}^j(\mathbf{1}),$$

and an L^1 function h so that

$$\|f_{n_k} - h\|_1 \to 0. \tag{4.8}$$

as $k \to \infty$.

Now we show that the measure μ defined by $\mu(A) = \displaystyle\int_A h d\mathfrak{m}$ is a stationary measure for the random dynamical system so satisfies $\mathcal{P}_{\mathfrak{T}}(h) = h$.

$$\|\mathcal{P}_{\mathfrak{T}}(h) - h\|_1 \;\leq\; \|\mathcal{P}_{\mathfrak{T}}(h) - \mathcal{P}_{\mathfrak{T}}(f_{n_k})\|_1 + \|\mathcal{P}_{\mathfrak{T}}(f_{n_k}) - f_{n_k}\|_1 + \|f_{n_k} - h\|_1$$

where the first and the third term on the right hand side are converging to zero by Equation 4.8 since $\mathcal{P}_{\mathfrak{T}}$ is a contraction. For the second term we have

$$\|\mathcal{P}_{\mathfrak{T}}(f_{n_k}) - f_{n_k}\|_1 = \left\|\mathcal{P}_{\mathfrak{T}}\left(\frac{1}{n_k}\sum_{j=0}^{n_k-1}\mathcal{P}_{\mathfrak{T}}^j(\mathbf{1})\right) - \frac{1}{n_k}\sum_{j=0}^{n_k-1}\mathcal{P}_{\mathfrak{T}}^j(\mathbf{1})\right\|_1$$

$$
\begin{aligned}
&= \left\| \frac{1}{n_k} \sum_{j=1}^{n_k} \mathcal{P}_{\mathfrak{T}}^{j}(\mathbf{1}) - \frac{1}{n_k} \sum_{j=0}^{n_k-1} \mathcal{P}_{\mathfrak{T}}^{j}(\mathbf{1}) \right\|_1 \\
&\leq \frac{1}{n_k} \left\| \mathcal{P}_{\mathfrak{T}}^{n_k}(\mathbf{1}) - \mathbf{1} \right\|_1 \\
&\leq \frac{1}{n_k} \left\| \mathcal{P}_{\mathfrak{T}}^{n_k}(\mathbf{1}) \right\|_1 + \left\| \mathbf{1} \right\|_1 \\
&\leq \frac{2}{n_k}
\end{aligned}
$$

since $\mathcal{P}_{\mathfrak{T}}$ is a contraction. Therefore second term also converges to zero as $k \to \infty$ which concludes the result. □

4.2 Expanding the Stationary Measure to a 2D Random Dynamical System

In this section we introduce a random dynamical system where the constituent maps are defined on a torus. This section is not aiming to generalize the 1D results to the maps on a torus but needed to prove some limit theorems for 1D random dynamical system and the results apply only to some special type of maps on a torus. First we introduce these special maps.

Definition 4.2.1. *Let T be a piecewise expanding map on $I = [0, 1]$. We define the corresponding* torus map *of T to be $A_T : \mathbb{T} \to \mathbb{T}$ given by*

$$
A_T(x, y) = \begin{pmatrix} T & 0 \\ 0 & T \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (T(x), T(y))
$$

The random dynamical system on a torus is defined in a similar way. Let $\{A_{T_1}, \ldots, A_{T_N}\}$ be the constituent maps of the system on a torus where each A_{T_i} is the corresponding torus map of a piecewise expanding map T_i on $[0, 1]$, and let $(p_1, \ldots, p_N)$ be the probability distribution on maps A_{T_i}. Let A_i denote the map A_{T_i} to simplify the notation.

The Perron-Frobenius operator of each map A_i is defined in the usual way.

$$\mathcal{P}_{A_i} f(x) = \sum_{(y):A_i(y)=(x)} \frac{f(y)}{|JA_i|(y)}$$

for $f : \mathbb{T} \to \mathbb{R}$. Here $|JA_i|$ is the Jacobian of A_i and given by

$$|JA_i| = \left| \begin{pmatrix} \dfrac{\partial T_i}{\partial x_1} & 0 \\ 0 & \dfrac{\partial T_i}{\partial x_2} \end{pmatrix} \right|$$

so $|JA_i|(y) = |T_i'(y_1)||T_i'(y_2)|$. Note that if $A_i(y_1, y_2) = (x_1, x_2)$ then we simply have $T_i(y_1) = x_1$ and $T_i(y_2) = (x_2)$. Thus, the Perron-Frobenius operator of A_i is given by

$$\mathcal{P}_{A_i} f(x_1, x_2) = \sum_{y_1, y_2:T_i(y_j)=(x_j)} \frac{f(y_1, y_2)}{|T_i'(y_1)||T_i'(y_2)|}$$

Similarly, the random Perron-Frobenius operator is given by

$$\mathcal{P}_{\mathfrak{A}} = \sum_{i=1}^{N} p_i \mathcal{P}_{A_i}.$$

For any fixed sequence of constituent maps on a torus, say $A_{\omega_1} A_{\omega_2} A_{\omega_3} \ldots$ there is a unique sequence of constituent maps on I, namely $T_{\omega_1} T_{\omega_2} T_{\omega_3} \ldots$ so we keep using the same sequences with a symbol space $\{1, \ldots, N\}$ so $\omega_{\mathbb{T}} = \omega$ and denote the corresponding skew product system as follows

$$\mathscr{F}_{\mathbb{T}}^n(\omega_{\mathbb{T}}, (x, y)) = (\boldsymbol{\sigma}^n(\omega), (T_{\omega_n} \circ \ldots T_{\omega_1}(x), T_{\omega_n} \circ \ldots T_{\omega_1}(y))), \text{ for } (x, y) \in \mathbb{T},$$

where $\boldsymbol{\sigma}$ is the shift map defined on sequences $\omega \in \Omega$.

To prove that such systems also have absolutely continuous stationary measures and satisfies the limit theorems one can follow the methods of the paper by P. Gora and A. Boyarski, see [GB89]. All the arguments can be generalized to random dynamical systems on a torus given by the special maps by applying the same methods we use for maps on I like taking maximum or minimum of the coefficients over all possible combinations of the maps. The idea for higher dimensional maps is refining the partitions according to the finitely many constituent maps so the proofs work as in the deterministic system.

In our case we define the partition α_0 for a map A_i on a torus to be $\{B \times B : B \in \beta\}$ where β is the partition for the corresponding map T_i on I. The restriction of each A_i to a partition element is C^2, one to one and expanding that is

$$|JA_i|(y) = |T_i'(y_1)||T_i'(y_2)| > \lambda_A > 1 \text{ since } |T_i'| > \lambda > 1.$$

Now by defining the partition α notice that we have squares and rectangles inside the torus where each corner is a singular point. The condition given by Gora and Boyarski in [GB89] for the existence of invariant measure is a lower bound on the expansion rate that depends on the nature of the partition, specifically on the minimal angle on the boundaries of the regions in the partition. Particularly if $x \in \mathbb{T}$ is a singular point of one of the partition element say $a \in \alpha$ then let $\theta(x)$ be the angle of the corner of the possible biggest cone that can be drawn in the partition with the corner x and define $\gamma(a) = |\cos(\theta(x_0) + \pi/2)|$ where x_0 is a singular point in a that gives the minimum angle $\theta(x)$. In our case if we use the partition α then for every x we have

$\theta(x) = \pi/4$, see example in page 282 in [GB89]. Furthermore for a fixed partition $a \in \alpha$, $\theta(x)$ is the same angle for every singular point $x \in a$. Then we define γ to be

$$\gamma = \inf_{a\in\alpha} \{ |\cos(\theta(x) + \pi/2)| : x \text{ is a singular point in } a\} = \frac{1}{\sqrt{2}}$$

We use the following result to prove the quenched Central Limit Theorem for random dynamical systems given by maps in $\mathcal{T}_0$ in Chapter 7.

Claim 4.2.2. *Let each $A_i : \mathbb{T} \to \mathbb{T}$ be a piecewise C^2 expanding maps with a maximum expanding rate λ. If $\lambda^{-1}(1+\sqrt{2}) < 1$ then the random dynamical system admits an absolutely continuous stationary measure.*

One can also choose to follow the methods of the paper by Keller, G. and Liverani, C., see [KL05] to prove that such systems have absolutely continuous stationary measures and satisfy the limit theorems. In their paper they give the Lasota-York inequality for a single map that is exactly in the form of our maps on a torus defined above with only constraint that they have finite partition. They use bounded variation arguments as we do for a 1D case, so one can generalize their arguments to random dynamical systems. Again we give the quenched Central Limit Theorem for random dynamical systems given by maps in $\mathcal{T}_0$, so the following assumption is enough.

Claim 4.2.3. *Let each $A_i : \mathbb{T} \to \mathbb{T}$ be the corresponding torus map of a map T_i in $\mathcal{T}_1(I)$ with finite partition. Then the random dynamical system on torus admits an absolutely continuous stationary measure.*

In the following chapters we study the spectral properties of the random Perron-Frobenius operator and give the statistical results only for one dimensional random dynamical systems. However the random dynamical system on a torus given by the special maps A_i corresponding to $T_i \in \mathcal{T}_0$ satisfies the same spectral properties and therefore the limit theorems. In the derivation of quenched Central Limit Theorem we use these results without proof. The proof should follow from the approach of [KL05] or [GB89], for spectral properties of certain maps in higher dimension.

CHAPTER 5

Spectral Properties of the Random Perron-Frobenius Operator

In Chapter 5 we give the spectral properties of $\mathcal{P}_{\mathfrak{T}}$. For maps in $\mathcal{T}_0$, E. Kobre and L. S. Young gives the spectral properties in [Kob05]. Their method relies on the theorem of Ionescu-Tulcea-Marinescu, see [ITM50]. We apply the uniform ergodic theory and follow the steps of Rychlik in [Ryc83] to reproduce the same results.

5.1 Spectral Properties of $\mathcal{P}_{\mathfrak{T}}$

We first show that the random Perron-Frobenius operator is quasi-compact as in Definition 3.2.4.

Proposition 5.1.1. *There exists $L \geq 1$ and a finite dimensional operator $\mathcal{K}$ on $BV(Y)$ such that $\|\mathcal{P}_{\mathfrak{T}}^L - \mathcal{K}\|_{BV} < 1$.*

Proof. Choose L and a partition $\mathfrak{Q}_{\varphi^{(L)}}$ for each possible L composition of constituent maps $\varphi^{(L)}$ such that Lemma 4.1.7 holds with each $A_{\varphi^{(L)}} < 1/4$. Let $\mathfrak{Q}$ be the refined partition of partitions $\mathfrak{Q}_{\varphi^{(L)}}$ for each map $\varphi^{(L)}$ and $A = \max\{A_{\varphi^{(L)}}\}$. Let $\mathcal{E}(f) = \mathbb{E}[f|\mathfrak{Q}]$ where $\mathbb{E}[f|\mathfrak{Q}]$ is the conditional expectation of f with respect to $\mathfrak{Q}$ and $\mathcal{K} = \mathcal{P}_{\mathfrak{T}}^L\mathcal{E}$. We prove that the choice for $\mathcal{K}$ is good.

Let $f \in BV(Y)$ be fixed and take $f_1 = f - \mathbb{E}[f|\mathcal{Q}] = (I - \mathcal{E})(f)$. To prove that $\|\mathcal{P}_{\mathfrak{T}}^L - \mathcal{K}\| < 1$ it is enough to show that $\|\mathcal{P}_{\mathfrak{T}}^L(f_1)\|_{BV} < 2A\|f\|_{BV}$ since

$$\begin{aligned}
\|\mathcal{P}_{\mathfrak{T}}^L(f_1)\|_{BV} &= \|\mathcal{P}_{\mathfrak{T}}^L(I - \mathcal{E})(f)\|_{BV} \\
&= \|\mathcal{P}_{\mathfrak{T}}^L(f) - (\mathcal{P}_{\mathfrak{T}}^L\mathcal{E})(f)\|_{BV} \\
&= \|\mathcal{P}_{\mathfrak{T}}^L(f) - \mathcal{K}(f)\|_{BV} \\
&\leq 2A\|f\|_{BV} \text{ for an arbitrary } f,
\end{aligned}$$

implies $\|\mathcal{P}_{\mathfrak{T}}^L - \mathcal{K}\| \leq 2A$ with $A < 1/4$ where the last norm is the operator norm.

For every $K \in \mathfrak{Q}$, $\int_K f_1 d\mathfrak{m} = 0$ by definition of conditional expectation, so

$$\begin{aligned}
\bigvee_Y \mathcal{P}_{\mathfrak{T}}^L(f_1) &\leq A\bigvee_Y f_1 + B_L\|f_1\|_1 \text{ by Remark 4.1.12} \\
&\leq A\bigvee_Y f_1 \\
&= A\bigvee_Y (f - \mathbb{E}[f|\mathfrak{Q}]) \\
&\leq A\bigvee_Y f + A\bigvee_Y \mathbb{E}[f|\mathfrak{Q}] \\
&\leq 2A\bigvee_Y f \leq 2A\|f\|_{BV}.
\end{aligned}$$

since for any finite partition $\mathfrak{Q}$, we have $\bigvee_Y \mathbb{E}[f|\mathfrak{Q}] \leq \bigvee_Y f$.

We also have $\|\mathcal{P}_{\mathfrak{T}}^L f_1\|_1 \leq \|f_1\|_1$ by the contractive property of $\mathcal{P}_{\mathfrak{T}}$. Furthermore, by Lemma 4.1.5, we get

$$\mathfrak{m}(B) \leq \|g_{\varphi^{(L)}}\|_\infty \leq A_{\varphi^{(L)}} \leq A \tag{5.1}$$

since $A = \max\{A_{\varphi^{(L)}}\}$ and since for each $\varphi^{(L)}$ the constant $A_{\varphi^{(L)}}$ is chosen to be

$$A_{\varphi^{(L)}} = \|g_{\varphi^{(L)}}\|_\infty + \max_{K \in \mathfrak{Q}_{\varphi^{(L)}}} \bigvee_K g_{\varphi^{(L)}}.$$

Also

$$\begin{aligned} \|f_1\chi_B\|_\infty &= \|(f - \mathbb{E}[f|\mathfrak{Q}])\chi_B\|_\infty \\ &\leq \|f\chi_B\|_\infty + \|\mathbb{E}[f|\mathfrak{Q}]\chi_B\|_\infty \\ &\leq 2\|f\chi_B\|_\infty \leq 2\|f\chi_B\|_{BV} \end{aligned} \tag{5.2}$$

Then Equation 5.1 and Equation 5.2 together imply that

$$\begin{aligned} \|f_1\|_1 &= \sum_B \int_B |f_1| dm \text{ where } B \in \beta_{\varphi^{(L)}} \\ &\leq \sum_B \int_B \|f_1\chi_B\|_\infty d\mathfrak{m} \\ &= \sum_B \|f_1\chi_B\|_\infty \mathfrak{m}(B) \\ &\leq \sum_B 2\|f\chi_B\|_{BV} A = 2A\|f\|_{BV}. \end{aligned}$$

Thus, $\|\mathcal{P}_{\mathfrak{T}}^L f_1\|_{BV} = \|\mathcal{P}_{\mathfrak{T}}^L f_1\|_1 + \bigvee_Y \mathcal{P}_{\mathfrak{T}}^L f_1 \leq 4A\|f\|_{BV} < \|f\|_{BV}$ since $A < 1/4$. $\square$

Theorem 5.1.2. *The operator $\mathcal{P}_{\mathfrak{T}}$ on bounded variations has the following properties:*

(i) $\sigma(\mathcal{P}_{\mathfrak{T}}) \cap S^1$ consists of a finite number of simple poles of the resolvent of $\mathcal{P}_{\mathfrak{T}}$. Moreover, $\sigma(\mathcal{P}_{\mathfrak{T}}) \cap S^1$ is a union of full cyclic groups.

(ii) Other points of $\sigma(\mathcal{P}_{\mathfrak{T}})$ *are contained within a circle of radius* $\rho \in (0,1)$.

(iii) If $\sigma(\mathcal{P}_{\mathfrak{T}}) \cap S^1 = \{\zeta_1, \zeta_2, \ldots, \zeta_k\}$, *we denote the projection to the corresponding eigenspace for* $j = 1, 2, \ldots, k$ *by* $\mathcal{Q}^j_{\mathfrak{T}}$*; then* $\mathcal{P}_{\mathfrak{T}}$ *admits the representation*

$$\mathcal{P}_{\mathfrak{T}} = \sum_{j=1}^{k} \zeta_j \mathcal{Q}^j_{\mathfrak{T}} + R_{\mathfrak{T}},$$

where $R_{\mathfrak{T}} : BV \to BV$ *and spectral radius of* $R_{\mathfrak{T}}$ *is* $\rho(R_{\mathfrak{T}}) = \inf_n \|R_{\mathfrak{T}}\|^{1/n} < \rho$. *Operators* $\mathcal{Q}^j_{\mathfrak{T}}$ *and* $\mathcal{R}_{\mathfrak{T}}$ *commute, and* $\mathcal{Q}^i_{\mathfrak{T}}\mathcal{Q}^i_{\mathfrak{T}} = \mathcal{Q}^i_{\mathfrak{T}}$, $\mathcal{Q}^j_{\mathfrak{T}}\mathcal{Q}^i_{\mathfrak{T}} = 0$ *and* $\mathcal{Q}^j_{\mathfrak{T}}\mathcal{R}_{\mathfrak{T}} = 0$ *for* $i \neq j$, $i, j = 1, 2, \ldots, k$.

The facts given above are all implied by Proposition 5.1.1 and a lemma we give below which shows that $\mathcal{P}_{\mathfrak{T}}$ is power bounded. Then the rest is a consequence of theory of operators. However we still show the steps to prove the fact since they are used to prove one of the main results of the models.

Lemma 5.1.3. *There exists* $F > 0$ *such that the operator* $\mathcal{P}_{\mathfrak{T}}$ *satisfies*

$$\sup_n \|\mathcal{P}^n_{\mathfrak{T}}\| \leq 2F + 1,$$

therefore $\mathcal{P}_{\mathfrak{T}}$ *is power bounded.*

Proof. We show that for every $n \geq 1$ and $f \in BV$ we have

$$\sum_{B \in \beta^n} \|\mathcal{P}^n(f \cdot \chi_B)\|_{BV} \leq (2F+1)\|f\|_{BV}$$

which implies the result. So we can consider each term $\mathcal{P}^n(f \cdot \chi_B)$ separately. Now for such $B_1 \in \beta^n$ that satisfies

$$\|\mathcal{P}^n(f \cdot \chi_{B_1})\|_1 < \bigvee \mathcal{P}^n(f \cdot \chi_B)$$

we have

$$\begin{aligned}\textstyle\sum_{B_1\in\beta^n} \|\mathcal{P}^n(f\cdot\chi_{B_1})\|_{BV} &\leq \textstyle\sum_{B_1\in\beta^n} 2\bigvee\mathcal{P}^n(f\cdot\chi_B)\\ &\leq 2F\left(\bigvee f+\|f\|_1\right)=2F\|f\|_{BV}\end{aligned}$$

where F is greater than the coefficients of variation and L^1-norm in Lemma 4.1.13, the Lasota-York type inequality for the variation that we show before giving the inequality for BV-norm. The other type of sets are $B_2\in\beta^n$ so that

$$\bigvee\mathcal{P}^n(f\cdot\chi_{B_2})<\|\mathcal{P}^n(f\cdot\chi_{B_2})\|_1\leq\|f\cdot\chi_{B_2}\|_1$$

so simply all terms can give at most

$$\|f\|_1\leq\|f\|_{BV}.$$

Therefore we get

$$\begin{aligned}\sum_{B\in\beta^n}\|\mathcal{P}^n(f\cdot\chi_B)\|_{BV} &= \sum_{B_1\in\beta^n}\|\mathcal{P}^n(f\cdot\chi_{B_1})\|_{BV}+\sum_{B_2\in\beta^n}\|\mathcal{P}^n(f\cdot\chi_{B_2})\|_{BV}\\ &\leq 2F\|f\|_{BV}+\|f\|_{BV}\\ &= (2F+1)\|f\|_{BV}.\end{aligned}$$

□

The proof of Theorem 5.1.2 is mainly the result of the Uniform Ergodic Theorem proved by Yosida-Kakutani, see [YK41] for power bounded quasi-compact operators $\mathcal{P}$. The proof depends on spectral theory. We give the Uniform Ergodic Theorem in Chapter 3 and here we use the theorem to give the proof of the Theorem 5.1.2.

Proof of Theorem 5.1.2 . By using the Uniform Ergodic Theorem, we get all the facts of Theorem 5.1.2 except that $\sigma(\mathcal{P}_{\mathfrak{T}})\cap S^1$ is a union of full cyclic groups.

To show that an eigenvalue $\zeta \in \sigma(\mathcal{P}_{\mathfrak{T}}) \cap S^1$ generates a cyclic group we have to show that there exists $n_0 \in \mathbb{N}$ such that $\zeta^{n_0} = \zeta$. Assume the contrary and asuume $\zeta \neq 1$. Note that if $\zeta \in \sigma(\mathcal{P}_{\mathfrak{T}})$ then there exists $h \in BV(Y)$ such that $\mathcal{P}_{\mathfrak{T}}(h) = \zeta h$ where h is the eigenfunction of $\mathcal{Q}$, the projection operator to the eigenspace of *zeta* so $\mathcal{Q}^j_{\mathfrak{T}}(h) = \mathcal{R}_{\mathfrak{T}}(h) = 0$ where $\mathcal{Q}^j_{\mathfrak{T}}$ and $\mathcal{R}_{\mathfrak{T}}$ as in Theorem 5.1.2 for any $\zeta_j \neq \zeta$. We also have $\mathcal{P}^n_{\mathfrak{T}}(h) = \zeta^n h$ which implies $\lim\limits_{n\to\infty} \mathcal{P}^n_{\mathfrak{T}}(h) = \lim\limits_{n\to\infty} \zeta^n h$, so $\lim_{n\to\infty} \mathcal{Q}^n(h) = \mathcal{Q}(h) = \zeta h$ since $\mathcal{Q}$ is idempotent. It is not possible on S^1 for a point ζ to have the property $\lim_{n\to\infty} \zeta^n = \zeta$ which gives the contradiction. Then $\zeta^{n_0} = \zeta$ implies that $\zeta^{n_0-1} = 1$. □

Theorem 5.1.4. *Operators $\mathcal{Q}^j_{\mathfrak{T}}$ and $\mathcal{R}_{\mathfrak{T}}$ have unique expansions as operators to L^1 and $\mathcal{Q}^j_{\mathfrak{T}}$ is bounded as an operator from L^1 to BV, $\|\mathcal{Q}^j_{\mathfrak{T}}\|_1 \leq 1$ and $\sup_n \|\mathcal{R}_{\mathfrak{T}}\|_1 < \infty$. For every $f \in L^1$, $\lim_{n\to\infty} \mathcal{R}_{\mathfrak{T}}(f) = 0$.*

Proof. The proof depends on noticing that for $\zeta \in S^1$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} (\mathcal{P}_{\mathfrak{T}}/\zeta)^i = \begin{cases} 0, \text{ if } \zeta \notin \{\zeta_1, \dots, \zeta_k\} \\ \mathcal{Q}^j_{\mathfrak{T}}, \text{ if } \zeta = \zeta_j. \end{cases}$$

Also note that BV is dense in L^1 and $\mathcal{Q}^j_{\mathfrak{T}}$ can be defined in L^1 since $\|\mathcal{P}_{\mathfrak{T}}/\zeta\|_1 = 1$. □

5.2 $\mathcal{P}_{\mathfrak{T}}$ with only Eigenvalue 1

In this section we analyze the situation when $\sigma(\mathcal{P}_{\mathfrak{T}}) \cap S^1$ consists of only 1. This situation is enough to consider for such random dynamical systems because the random Perron-Frobenius operator has only finitely many eigenvalues on S^1 and each is a root of unity. Then if $M \in \mathbb{N}$ is a common multiple of orders of eigenvalues then the operator $\mathcal{P}_{\mathfrak{T}}^M$ has the only eigenvalue 1. From Theorem 5.1.2, we have $\mathcal{P}_{\mathfrak{T}} = \mathcal{Q}_{\mathfrak{T}} + \mathcal{R}_{\mathfrak{T}}$.

Notation 5.2.1.

$$A_n(\mathcal{P}) := \frac{1}{n}\sum_{i=1}^{n-1}\mathcal{P}^n.$$

Theorem 5.2.2. *There exists nonnegative functions $\varphi_1, \ldots, \varphi_s \in BV$ and $\psi_1, \ldots, \psi_s \in L^\infty$ such that*

(i) For every $f \in L^1$,

$$\mathcal{Q}_{\mathfrak{T}}(f) = \sum_{i=1}^{s}\left(\int(\psi_i.f)d\mathfrak{m}\right)\varphi_i.$$

(ii) $\mathcal{P}\varphi_i = \varphi_i$, $\displaystyle\sum_{j=1}^{N} p_j(\psi_i \circ T_j) = \psi_i$ for $i = 1, \ldots, s$.

(iii) $\displaystyle\int \varphi_i d\mathfrak{m} = 1$, $\displaystyle\int \varphi_i\psi_j dm = \delta_{ij}$, $\min\{\varphi_i, \varphi_j\} = \min\{\psi_i, \psi_j\} = 0$ for $i = 1, \ldots, s$.

(iv) There exists measurable sets $C_1, \ldots, C_s \subset Y$ such that $\psi_i = \chi_i$ a.e. for $i = 1, \ldots, s$ and $Y = \displaystyle\bigcup_{i=1}^{s} C_i$ a.e.

(v) Let $\mathcal{U}_{\mathfrak{T}}$ denote $\mathcal{P}_{\mathfrak{T}}^*$, so $\displaystyle\bigcap_{n=1}^{\infty}\mathcal{U}_{\mathfrak{T}}^n(L^1) = \bigcap_{n=1}^{\infty}\mathcal{U}_{\mathfrak{T}}^n(L^\infty) = span\{\psi_1, \ldots, \psi_s\}$.

(vi) For every $f \in L^1$ (or L^∞), $\mathcal{U}_{\mathfrak{T}}^n f \to \mathcal{Q}_{\mathfrak{T}}^*(f)$ in $\sigma(L^1, BV)$ topology (or $\sigma(L^\infty, BV)$ topology) as $n \to \infty$ and

$$\mathcal{Q}_{\mathfrak{T}}^*(f) = \sum_{i=1}^{s}\left(\int f\varphi_i d\mathfrak{m}\right)\psi_i.$$

Proof. Operator $\mathcal{Q}_{\mathfrak{T}}$ is a positive operator since it is given as the limit of $A_n(\zeta^{-1}\mathcal{P})$ for $\zeta = 1$, see Theorem 5.1.4. Let $Z = \ker(I - \mathcal{P}_{\mathfrak{T}})$ be the projection of $\mathcal{P}_{\mathfrak{T}}$ onto the eigenspace corresponding to the eigenvalue 1. Then for $f, g \in Z$ we have $\mathcal{Q}_{\mathfrak{T}}(\min\{f,g\}) \leq \min\{\mathcal{Q}_{\mathfrak{T}}f, \mathcal{Q}_{\mathfrak{T}}g\} = \min\{f,g\}$. On the other hand, $\mathcal{P}$ preserves the Lebesgue measure $\mathfrak{m}$ as given in Equation 4.4, so the limit of the average $A_n(\mathcal{P})$ which is $\mathcal{Q}_{\mathfrak{T}}$. Then we have

$$\int \mathcal{Q}_{\mathfrak{T}}(\min\{f,g\})d\mathfrak{m} = \int \min\{f,g\}d\mathfrak{m}$$

Therefore

$$0 \leq \min\{f,g\} - \mathcal{Q}_{\mathfrak{T}}(\min\{f,g\})$$

and

$$\int \min\{f,g\} - \mathcal{Q}_{\mathfrak{T}}(\min\{f,g\})d\mathfrak{m} = 0$$

implies that

$$\min\{f,g\} - \mathcal{Q}_{\mathfrak{T}}(\min\{f,g\}) = 0$$

$\mathfrak{m}$-a.e., so $\min\{f,g\} = \mathcal{Q}_{\mathfrak{T}}(\min\{f,g\})$ $\mathfrak{m}$-a.e.

Let

$$\Delta = \{\varphi \in Z : \int \varphi d\mathfrak{m} = 1 \text{ and } \varphi \geq 0\}$$

which is a convex and a compact set. Such a set has extreme points so let φ_1, φ_2 be two different extreme points of Δ. We have $\min\{\varphi_1, \varphi_2\} = 0$ so they are linearly independent and there are finitely many extreme points of Δ, say $\{\varphi_1, \ldots, \varphi_s\}$ with $s \leq \dim Z$. By Krein-Milman theorem we know that Δ is the closed convex hull of its extreme points so $\dim(\Delta) = s$. Furthermore, Δ spans Z so we have $\dim Z = s$.

Now, we have the basis for Z, namely $\{\varphi_1, \ldots, \varphi_s\}$. Therefore, for every $f \in Z$ we have

$$f = c_1\varphi_1 + \ldots + c_s\varphi_s$$

for some combinations c_i. If $\{\varphi_1^*, \ldots, \varphi_s^*\}$ is the dual basis defined by $\varphi_i^*(\varphi_j) = 1$ for $i = j$, and $\varphi_i^*(\varphi_j) = 0$ otherwise, then we have

$$\begin{aligned}
f &= c_1\varphi_1^*(\varphi_1)\varphi_1 + \ldots + c_s\varphi_s^*(\varphi_s)\varphi_s \\
&= \varphi_1^*(c_1\varphi_1)\varphi_1 + \ldots + \varphi_s^*(c_s\varphi_s)\varphi_s \\
&= \varphi_1^*(c_1\varphi_1 + \ldots + c_s\varphi_s)\varphi_1 + \ldots + \varphi_s^*(c_1\varphi_1 + \ldots + c_s\varphi_s)\varphi_s \\
&\qquad\qquad\qquad\qquad \text{since } \varphi_i^*(\varphi_j) = 0 \text{ for } i \neq j \\
&= \varphi_1^*(f)\varphi_1 + \ldots + \varphi_s^*(f)\varphi_s.
\end{aligned}$$

If we write $\mathcal{Q}_{\mathfrak{T}}(f)$ as a linear combination of the basis by using the dual basis as above we get

$$\begin{aligned}
\mathcal{Q}_{\mathfrak{T}}(f) &= \varphi_1^*(\mathcal{Q}_{\mathfrak{T}}(f))\varphi_1 + \ldots + \varphi_s^*(\mathcal{Q}_{\mathfrak{T}}(f))\varphi_s \\
&= (\mathcal{Q}_{\mathfrak{T}}^*\varphi_1^*)(f)\varphi_1 + \ldots + (\mathcal{Q}_{\mathfrak{T}}^*\varphi_s^*)(f)\varphi_s \text{ by the definition of } \mathcal{Q}_{\mathfrak{T}}^* \\
&= \mu_1(f)\varphi_1 + \ldots + \mu_s(f)\varphi_s \text{ by defining } \mathcal{Q}_{\mathfrak{T}}^*\varphi_i^* = \mu_i
\end{aligned}$$

where μ_i is a functional in L^1. So we can find $\psi_i \in L^\infty$ for each i so that

$$\mu_i(f) = \int f d\mu_i = \int f\psi_i d\mathfrak{m}.$$

Therefore, we get

$$\mathcal{Q}_{\mathfrak{T}}(f) = \left(\int f\psi_1 d\mathfrak{m}\right)\varphi_1 + \ldots + \left(\int f\psi_s d\mathfrak{m}\right)\varphi_s$$

which proves the first part of the theorem.

Now, for the second part it is clear that for $\varphi_i \in Z$ we have $\mathcal{P}_{\mathfrak{T}}\varphi_i = \varphi_i$ since Z is the eigenspace corresponding to the eigenvalue 1. Also we observe that $\mathcal{P}_{\mathfrak{T}}^* = \mathcal{U}_{\mathfrak{T}}$ is given by

$$\mathcal{U}_{\mathfrak{T}}(f) = \sum_{j=1}^{N} p_j f \circ T_j$$

where $T_1, \ldots, T_N$ are the constituent functions for the random dynamical system since

$$\begin{aligned} \mathcal{P}_{\mathfrak{T}} f(x) &= \textstyle\sum_{j=1}^{N} p_j \mathcal{P}_{T_j} f(x) \\ &= \textstyle\sum_{j=1}^{N} p_j \sum_{y \in T_j^{-1}(x)} g_{T_j}(y) f(y) \end{aligned}$$

implies for any $f_1, f_2 \in L^1$ that

$$\begin{aligned} < f_1, \mathcal{P}_{\mathfrak{T}} f_2 > &= \textstyle\sum_{j=1}^{N} p_j < f_1, \mathcal{P}_{T_j} f_2 > \\ &= \textstyle\sum_{j=1}^{N} p_j < \mathcal{P}_{T_j}^* f_1, f_2 > \\ &= \textstyle\sum_{j=1}^{N} p_j < f_1 \circ T_j, f_2 > \\ &= < \mathcal{U}_{\mathfrak{T}}(f_1), f_2 > . \end{aligned}$$

We have $\mathcal{Q}_{\mathfrak{T}}(f) = \mathcal{Q}_{\mathfrak{T}}\mathcal{P}_{\mathfrak{T}} f$ since $f \in Z$ where

$$\mathcal{Q}_{\mathfrak{T}}(\mathcal{P}_{\mathfrak{T}} f) = \sum_{i=1}^{s} \left(\int \mathcal{P}_{\mathfrak{T}}(f) \psi_i d\mathfrak{m} \right) \varphi_i$$

by the first part of the proof. If we set it equal to $\mathcal{Q}_{\mathfrak{T}}(f) = \sum_{i=1}^{s} \left(\int f \psi_i d\mathfrak{m} \right) \varphi_i$ together with $\mathcal{P}_{\mathfrak{T}}^* = \mathcal{U}_{\mathfrak{T}}$, we get

$$\int f(\mathcal{U}_{\mathfrak{T}}\psi_i) d\mathfrak{m} = \int (\mathcal{P}_{\mathfrak{T}} f) \psi_i d\mathfrak{m} = \int f \psi_i d\mathfrak{m}$$

for every $f \in L^1$, so $\mathcal{U}_{\mathfrak{T}}\psi_i = \psi_i$ for $i = 1, \ldots, s$ which proves second part of the theorem.

For the third part we already know that $\min\{\varphi_i, \varphi_j\} = 0$ because of the way they are defined. We leave the proof of $\min\{\psi_i, \psi_j\} = 0$ after proving part (vi). To show,

$\int \varphi_i \psi_j dm = \delta_{ij}$ we use the fact that the projection operator $\mathcal{Q}_{\mathfrak{T}}$ is idempotent, in other words $\mathcal{Q}_{\mathfrak{T}}^2 = \mathcal{Q}_{\mathfrak{T}}$. Then $\mathcal{Q}_{\mathfrak{T}}(f) = \sum_{i=1}^{s} \left(\int f \psi_i d\mathfrak{m} \right) \varphi_i$ implies

$$\begin{aligned}
\mathcal{Q}_{\mathfrak{T}}^2(f) &= \mathcal{Q}_{\mathfrak{T}} \left(\sum_{i=1}^{s} \left(\int f \psi_i d\mathfrak{m} \right) \varphi_i \right) \\
&= \sum_{i=1}^{s} \left(\int \sum_{j=1}^{s} \left(\int f \psi_j d\mathfrak{m} \right) \varphi_i \psi_j \right) \varphi_i \\
&= \sum_{i,j=1}^{s} \left(\int f \psi_j d\mathfrak{m} \right) \left(\int \varphi_i \psi_j d\mathfrak{m} \right) \varphi_i.
\end{aligned}$$

And by setting equal to $\mathcal{Q}_{\mathfrak{T}}(f) = \sum_{i=1}^{s} (\int f \psi_i d\mathfrak{m}) \varphi_i$ we get $\int \varphi_i \psi_j d\mathfrak{m} = \delta_{ij}$.

We can also prove (vi) easily by using $\mathcal{Q}_{\mathfrak{T}}(f) = \sum_{i=1}^{s} (\int f \psi_i dm) \varphi_i$. For every $g \in L^1$, we have

$$\begin{aligned}
\int (\mathcal{Q}_{\mathfrak{T}}^* f) g d\mathfrak{m} &= \int f \left(\mathcal{Q}_{\mathfrak{T}} g \right) d\mathfrak{m} \\
&= \int f \sum_{i=1}^{s} \left(\int g \psi_i d\mathfrak{m} \right) \varphi_i d\mathfrak{m} \\
&= \sum_{i=1}^{s} \left(\int f \varphi_i d\mathfrak{m} \right) \left(\int g \psi_i d\mathfrak{m} \right) \\
&= \int \left(\sum_{i=1}^{s} \left(\int f \varphi_i d\mathfrak{m} \right) \psi_i \right) g d\mathfrak{m}
\end{aligned}$$

which implies $\mathcal{Q}_{\mathfrak{T}}^* f = \sum_{i=1}^{s} \left(\int f \varphi_i d\mathfrak{m} \right) \psi_i$. Furthermore, for every $g \in L^1$

$$\int \left(\mathcal{U}_{\mathfrak{T}}^n f \right) g d\mathfrak{m} = \int f \left(\mathcal{P}_{\mathfrak{T}}^n g \right) d\mathfrak{m} \longrightarrow \int f \left(\mathcal{Q}_{\mathfrak{T}} g \right) d\mathfrak{m} = \int (\mathcal{Q}_{\mathfrak{T}}^* f) g d\mathfrak{m}$$

which implies $\mathcal{U}_{\mathfrak{T}}^n f \longrightarrow \mathcal{Q}_{\mathfrak{T}}^* f$.

Now we can prove the part we left in (iii) which is $\min\{\psi_i, \psi_j\} = 0$: We have

$\mathcal{U}_{\mathfrak{T}}^n f \to \mathcal{Q}_{\mathfrak{T}}^* f = \sum_{i=1}^{s} \left(\int f \varphi_i d\mathfrak{m} \right) \psi_i$ by part (iv). Choose f to be φ_j, so we get

$$\mathcal{U}_{\mathfrak{T}}^n \varphi_j \longrightarrow \mathcal{Q}_{\mathfrak{T}}^* \varphi_j = \sum_{i=1}^{s} \left(\int \varphi_j \varphi_i d\mathfrak{m} \right) \psi_i = \left(\int \varphi_j^2 d\mathfrak{m} \right) \psi_j$$

since $\varphi_j \varphi_i = 0$ for $i \neq j$. Then

$$\begin{aligned} \min\{\psi_i, \mathcal{U}_{\mathfrak{T}}^n \varphi_j\} &= \min\{\psi, \sum_{\omega}^{N^n} q_k \varphi_j \circ T_{\omega_n} \circ \ldots \circ T_{\omega_1}\} \text{ where } T_{\omega_n} \circ \ldots \circ T_{\omega_1} \\ &\qquad \text{is one of the } N^n \text{ many n-combinations of the} \\ &\qquad \text{constituent maps with probability } q_k, \\ &= \sum_{\omega}^{N} p_k \min\{\psi_i, \varphi_j\} \circ T_{\omega_n} \circ \ldots \circ T_{\omega_1} \\ &= 0 \text{ since } \int \psi_i \varphi_j d\mathfrak{m} = 0 \text{ for } i \neq j. \end{aligned}$$

And $\int (\mathcal{U}_{\mathfrak{T}}^n \varphi_j) \psi_i dm \to \int \left(\int \varphi_j^2 d\mathfrak{m} \psi_j \right) \psi_i d\mathfrak{m} = \left(\int \varphi_j^2 d\mathfrak{m} \right) \left(\int \psi_i \psi_j d\mathfrak{m} \right)$ implies $\int \psi_i \psi_j d\mathfrak{m} = 0$ since $\int \varphi_j^2 d\mathfrak{m} > 0$. Thus, $\min\{\psi_i, \psi_j\} = 0$ for $i \neq j$ so part (iii) is completed.

To prove part (iv) we use part (vi), that is $\mathcal{U}_{\mathfrak{T}}^n f \to \mathcal{Q}_{\mathfrak{T}}^* f$ by choosing f to be 1, so $\mathcal{U}_{\mathfrak{T}}^n 1 = 1 \to \mathcal{Q}_{\mathfrak{T}}^* 1$, so $\mathcal{Q}_{\mathfrak{T}}^* 1 = 1$. And $\mathcal{Q}_{\mathfrak{T}}^* f = \sum_{i=1}^{s} (\int f \varphi_i dm) \psi_i$ with $f = 1$ implies that $1 = \mathcal{Q}_{\mathfrak{T}}^* 1 = \sum_{i=1}^{s} (\int \varphi_i dm) \psi_i = \sum_{i=1}^{s} \psi_i$. Thus each ψ_i is of the form χ_{C_i} with $\bigcup_{i=1}^{s} C_i = Y$.

Now only part (v) is left. It is clear that $\text{span}\{\psi_1, \ldots, \psi_s\} \subset \bigcap_{i=1}^{\infty} \mathcal{U}_{\mathfrak{T}}^n (L^1)$ since each $\psi_j \in L^1$ and $\mathcal{U}_{\mathfrak{T}}(\psi_j) = \psi_j$. To show $\text{span}\{\psi_1, \ldots, \psi_s\} = \bigcap_{i=1}^{\infty} \mathcal{U}_{\mathfrak{T}}^n (L^1)$ let $f \in \bigcap_{i=1}^{\infty} \mathcal{U}_{\mathfrak{T}}^n (L^1)$. Since $f \in \mathcal{U}_{\mathfrak{T}}^n (L^1)$ for every $n \in \mathbb{N}$, choose $f_n \in L^1$ for every $n \in \mathbb{N}$ such that $f = \mathcal{U}_{\mathfrak{T}}^n (f_n)$. Note that $\|f\|_\infty = \|\mathcal{U}_{\mathfrak{T}}^n (f_n)\|_\infty \geq \|f_n\|_\infty$. The sequence $\{\mathcal{Q}_{\mathfrak{T}}^* f_n\}_{n=1}^{\infty}$

is bounded in span$\{\psi_1, \ldots, \psi_s\}$, so there exists a subsequence $\{\mathcal{Q}_{\mathfrak{T}}^* f_{n_k}\}_{k=1}^{\infty}$ such that $\mathcal{Q}_{\mathfrak{T}}^* f_n \to f_0$ where $f_0 \in$ span$\{\psi_1, \ldots, \psi_s\}$ as $k \to \infty$. If we show that $f_0 = f$ we are done.

Let $g \in L^1$ be arbitrary. For any $k \in \mathbb{N}$ we have

$$\begin{aligned} \int (gf) d\mathfrak{m} &= \int g(\mathcal{U}_{\mathfrak{T}}^{n_k} f_{n_k}) d\mathfrak{m} \\ &= \int (\mathcal{P}_{\mathfrak{T}}^{n_k} g) f_{n_k} d\mathfrak{m} \end{aligned}$$

$$\begin{aligned} &= \int ((\mathcal{P}_{\mathfrak{T}}^{n_k} - \mathcal{Q}_{\mathfrak{T}}) g) f_{n_k} d\mathfrak{m} + \int (\mathcal{Q}_{\mathfrak{T}} g) f_{n_k} d\mathfrak{m} \\ &= \int ((\mathcal{P}_{\mathfrak{T}}^{n_k} - \mathcal{Q}_{\mathfrak{T}}) g) f_{n_k} d\mathfrak{m} + \int g(\mathcal{Q}_{\mathfrak{T}}^* f_{n_k}) d\mathfrak{m} \\ &\to 0 + \int g f_0 d\mathfrak{m} \text{ as } k \to \infty. \end{aligned}$$

Since $g \in L^1$ is arbitrary we have $f = f_0$ $\mathfrak{m}$-a.e. $\square$

CHAPTER 6

Mixing Properties of the Random Dynamical System

In Chapter 6, we show that random dynamical system $(\mathfrak{T}, \mu)$ has a Bernoulli scheme. We give the definition of what it means for a random dynamical system to be Bernoulli. Before that we give the theorem below which makes it clear that why analyzing an operator with only eigenvalue 1 is enough.

Definition 6.0.3 (weak-mixing)**.** *We say that the random dynamical system* $(\mathfrak{T}, \mu)$ *is* weak-mixing *if the random Perron-Frobenius operator* $\mathcal{P}_{\mathfrak{T}}$ *has only the eigenvalue 1 on the unit circle, and it has multiplicity one (that is,* $\dim\ker(\mathcal{P}_{\mathfrak{T}} - I) = 1$*).*

In the rest of the thesis we assume that the dimension of the eigenspace corresponding to 1 is one; that is, the corresponding skew product realization of the random dynamical system is weak-mixing.

Theorem 6.0.4. *Let $\mathcal{P}_{\mathfrak{T}}$ be the random Perron-Frobenius operator of the random dynamical system with functions from $\mathcal{T}_1(Y)$ as given before. Fix an integer M such that $\sigma(\mathcal{P}_{\mathfrak{T}}^M) \cap S^1 = \{1\}$. Let $\varphi_1, \ldots, \varphi_s$ and $\psi_1 = \chi_{C_1}, \ldots, \psi_s = \chi_{C_s}$ be as in Theorem 5.2.2 applied to $\mathcal{P}_{\mathfrak{T}}^M$. Then there exists a permutation $\boldsymbol{\pi}$ of the set $\{1, 2, \ldots, s\}$ such that*

$$\mathcal{P}_{\mathfrak{T}}(\varphi_i) = \varphi_{\boldsymbol{\pi}(i)},$$
$$\mathcal{U}_{\mathfrak{T}}(\psi_{\boldsymbol{\pi}(i)}) = \psi_i \qquad \text{for } i = 1, 2, \ldots, s.$$

6.1 Bernoulli Property

First we define the Bernoulli property for the random dynamical system $(\{\mathfrak{T}\}, \mu)$ where 1 is the only eigenvalue of $\mathcal{P}_{\mathfrak{T}}$ on the unit circle and there exists only one $h \in L^1$ such that $\mathcal{P}_{\mathfrak{T}}(h) = h$ with $\int \varphi dm = 1$. We prove that the skew product realization of the random dynamical system $(\mathfrak{T}, \mu)$ where $\mu = hm$ has the Bernoulli property, so it is mixing, implying that the random dynamical system is ergodic. We use results of Ornstein Theory for deterministic maps.

Definition 6.1.1. *Let $\mathfrak{T}$ be a random map defined on Y with constituent maps $T_i : Y \to Y$, $i = 1, \ldots, N$ and probability distribution $(p_1, \ldots, p_N)$ on the maps, let μ be a stationary measure for the random dynamical system. Let $\mathscr{F}$ be the associated skew product map. If $(\mathscr{F}, \pi \times \mu)$ has the Bernoulli property in the sense of Definition 3.4.1 then we say that the random dynamical system $(\mathfrak{T}, \mu)$ has the Bernoulli property.*

Let us review the notation for the $\mathcal{T}_1(Y)$ types of maps first. The map $T_i \in \mathcal{T}_1(Y)$ is piecewise with countably many partition say β_i. For finitely many functions $\{T_1, \ldots, T_N\}$ we refine the partitions $\beta_1, \ldots, \beta_N$ to get a common partition β on which all the maps are piecewise expanding. Furthermore, $\bigvee_{i=0}^{n} \mathfrak{T}^{-i}(\beta)$ represents the refinement $\beta \vee T_{\omega_1}^{-1}(\beta) \vee (T_{\omega_2}^{-1} \circ T_{\omega_1}^{-1})(\beta) \vee \ldots \vee (T_{\omega_i}^{-1} \circ \ldots \circ T_{\omega_2}^{-1} \circ T_{\omega_1}^{-1})(\beta)$ for every i combination $T_{\omega_i} \circ \ldots \circ T_{\omega_1}$ of constituent maps for $i = 1, \ldots, n$.

Let $\mathfrak{T}$ be a random map on Y with constituent functions $\{T_1, \ldots, T_N\} \subset \mathcal{T}_1(Y)$ and β be the common countably many partition so that each T_i is piecewise expanding. Let $\mathcal{F}_m^n$ denote the σ-algebra generated by the partition $\beta_m^n = \bigvee_{i=m}^{n} \mathfrak{T}^{-i}$.

We give the criteria below for the random dynamical system $(\mathfrak{T}, \mu)$ to has the Bernoulli property by using the criteria for a single map to have Bernoulli property, see Definition 3.4.2 for the distance operation $\mathfrak{d}$ used in the following Proposition.

Proposition 6.1.2. *If the partition β for the random dynamical system $(\mathfrak{T}, \mu)$ satisfies*

$$\sup_{l,k \geq 1} \mathfrak{d}(\beta_0^l, \beta_{l+n}^{l+k+n}) \to 0 \text{ as } n \to \infty$$

then the random dynamical system $(\mathfrak{T}, \mu)$ has the Bernoulli property in the sense of Definition 6.1.1.

Proof. The proof is a result of Ornstein Theory, see 3.4, applied to the corresponding skew product system.

Let $\mathscr{F} : \Omega \times Y \to \Omega \times Y$ be the corresponding skew product realization of the random dynamical system $(\mathfrak{T}, \mu)$ where Ω is the set of sequences with the the symbol space $\{,\ldots,N\}$ on each entry and $\mathscr{F}$ is defined by

$$\mathscr{F}(\omega,x)=(\boldsymbol{\sigma}(\omega),T_{\omega_1}(x))$$

where $\boldsymbol{\sigma}$ is the left shift function. Let π be the product measure on Ω obtained by the distribution $(p_1,\ldots,p_N)$ on constituent maps. Define the 1-cylinders C_i for $i=1,\ldots,N$ to be

$$C_i=\{\omega=(w_1w_2\ldots)\in\Omega:T_{\omega_1}=T_i\}.$$

Define a partition γ on $\Omega\times Y$ such that for any $G\in\gamma$, $G=C_i\times B$ for some $i\in\{1,\ldots,N\}$ and $B\in\beta$. Note that $\mathscr{F}^{-1}(G)=\mathscr{F}^{-1}(C_i\times B)=\{(\omega_1\omega_2\ldots,x)\in\Omega\times Y:\omega_2=T_i,x\in\omega_1^{-1}(B)\}$, so for any other $H\in\gamma$, say $H=C_j\times D$ we have $H\cap\mathscr{F}^{-1}(G)=\{(\omega_1\omega_2\ldots,x)\in\Sigma\times Y:\omega_1=T_j,\omega_2=T_i,\text{ and }x\in D\cap T_j^{-1}(B)\}=C_{ji}\times E$ where C_{ji} is a 2-cylinder and $E\in\beta_0^1$. If we apply the same argument to all sets of the partition γ and all n^{th} inverse images we see that the refined partition γ_0^n consists of elements in the form of $C_{(i_1\ldots i_{n+1})}\times B$ where $C_{(i_1\ldots i_{n+1})}$ is an (n+1)-cylinder and $B\in\beta_0^n$.

Now it is enough to show that the deterministic system $(\mathscr{F},\pi\times\mu)$ has the Bernoulli property which is implied by

$$\sup_{l,k\geq 1}\mathfrak{d}(\gamma_0^l,\gamma_{l+n}^{l+k+n})\to 0\text{ as }n\to\infty \tag{6.1}$$

by Ornstein Theory , Corollary 3.4.4. Then we only need to show that the assumption of the Proposition 6.1.2 implies the Equation 6.1. Let $G\in\gamma_0^l$ and $H\in\gamma_{l+n}^{l+k+n}$ for some fixed n,m and $k\in\mathbb{N}$. From the above argument we know that $G=C_{(i_1\ldots i_n)}\times B$

where $C_{(i_1 \dots i_n)}$ is an n-cylinder for some fixed n sequence of maps $(T_{i_1} \dots T_{i_n})$ and $B \in \beta_0^n$, and $H = C_{(i_{l+n+1} \dots i_{l+k+n+1})} \times D$ where

$$C_{(i_{l+n+1} \dots i_{l+k+n+1})} = \{\omega \in \Omega : T_{\omega_{l+n+1}} = T_{i_{l+n+1}}, \dots, T_{\omega_{l+k+n+1}} = T_{i_{l+k+n+1}}\}$$

for some fixed k-sequence of maps $(T_{i_{l+n+1}}, \dots, T_{i_{l+k+n+1}})$ and $D \in \beta_{l+n}^{l+k+n}$. Then we get the following

$$\begin{aligned}
\mathfrak{d}(\gamma_0^l, \gamma_{l+n}^{l+k+n}) &= \sum_{G \in \gamma_0^l, H \in \gamma_{l+n}^{l+k+n}} |\pi \times \mu(G \cap H) - \pi \times \mu(G)\pi \times \mu(H)| \\
&= \sum |\pi \times \mu(C_{(i_1 \dots i_n)} \times B \cap C_{(i_{l+n+1} \dots i_{l+k+n+1})} \times D) \\
&\quad - [\pi \times \mu(C_{(i_1 \dots i_n)} \times B)][\pi \times \mu(C_{(i_{l+n+1} \dots i_{l+k+n+1})} \times D)]
\end{aligned}$$

where the sum is over $B \in \beta_0^n, D \in \beta_{l+n}^{l+k+n}, C_{(i_1 \dots i_n)}$ and $C_{(i_{l+n+1} \dots i_{l+k+n+1})}$. Therefore

$$\begin{aligned}
\mathfrak{d}(\gamma_0^l, \gamma_{l+n}^{l+k+n}) &= \sum |\pi \times \mu(C_{(i_1 \dots i_n; i_{l+n+1} \dots i_{l+k+n+1})} \times B \cap D) \\
&\quad -[\pi \times \mu(C_{(i_1 \dots i_n)} \times B)][\pi \times \mu(C_{(i_{l+n+1} \dots i_{l+k+n+1})} \times D)] \\
&= \sum |p_{i_1} \dots p_{i_n} p_{i_{l+n+1}} \dots p_{i_{l+k+n+1}} \mu(B \cap D) \\
&\quad -[p_{i_1} \dots p_{i_n} \mu(B)][p_{i_{l+n+1}} \dots p_{i_{l+k+n+1}} \mu(D)]
\end{aligned}$$

where p_{i_j} is the probability of choosing the map T_{i_j}, and since we are summing over all possible n and k sequences of maps the sum of probabilities is 1 and we get

$$\begin{aligned}
\mathfrak{d}(\gamma_0^l, \gamma_{l+n}^{l+k+n}) &= \sum_{B \in \beta_0^n, D \in \beta_{l+n}^{l+k+n}} |\mu(B \cap D) - \mu(B)\mu(D)| \\
&= \mathfrak{d}(\beta_0^l, \beta_{l+n}^{l+k+n})
\end{aligned}$$

which concludes the proof. $\square$

To show that the random dynamical system $(\mathfrak{T}, \mu)$ has the Bernoulli property we see that by Proposition 6.1.2 it is enough to show that

$$\sup_{l,k\geq 1} \mathfrak{d}(\beta_0^l, \beta_{l+n}^{l+k+n}) \to 0 \text{ as } n \to 0 \tag{6.2}$$

which we do by first defining an equal quantity to Equation (6.2) given in the following section and then we show that the new quantity converges to zero.

6.2 Decay of Correlation

Definition 6.2.1. *Let $(\mathfrak{T}, \mu)$ be a random dynamical system with the common infinitely many partition β. We define the n^{th} correlation of the system by*

$$\mathrm{Corr}(n) = \sup_{l\geq 1} \mathbb{E}_\mu \left[\sup_{A\in\mathcal{F}_{l+n}^\infty} \left|\mu(A|\mathcal{F}_0^l) - \mu(A)\right| \right].$$

where $\mathcal{F}_{l+n}^\infty$ is the σ-algebra generated by $\beta_{l+n}^\infty = \bigvee_{i=l+n}^\infty \mathfrak{T}^{-i}$.

The following Lemma gives the equivalent quantity to Equation (6.2).

Lemma 6.2.2. *Let $\mathfrak{d}$ and $\mathrm{Corr}(n)$ be as above, and β be the common partition for the random dynamical system $(\mathfrak{T}, \mu)$. Then we have*

$$\mathrm{Corr}(n) = \frac{1}{2} \sup_{l,k\geq 1} \mathfrak{d}(\beta_0^l, \beta_{l+n}^{l+k+n}).$$

Proof.

$$\begin{aligned}
\sup_{l,k\geq 1}\mathfrak{d}(\beta_0^l,\beta_{l+n}^{l+k+n}) &= \sup_{l,k\geq 1}\sum_{A\in\beta_0^l, B\in\beta_{l+n}^{l+k+n}} |\mu(A\cap B)-\mu(A)\mu(B)| \\
&= \sup_{l\geq 1}\sum_{A\in\beta_0^l}\sup_{k\geq 1}\sum_{B\in\beta_{l+n}^{l+k+n}} |\mu(B|A)-\mu(B)|\mu(A) \\
&= \sup_{l\geq 1}\sum_{A\in\beta_0^l}\sup_{B\in\mathcal{F}_{l+n}^{\infty}} |\mu(B|A)-\mu(B)|\mu(A) \\
&= \sup_{l\geq 1}\mathbb{E}_\mu\left[\sup_{B\in\mathcal{F}_{l+n}^{\infty}} |\mu(B|\mathcal{F}_0^l)-\mu(B)|\right] \\
&= \mathrm{Corr}(n).
\end{aligned}$$

□

The following Theorem concludes that the equivalent quantity of Equation (6.2) converges to zero so the Bernoulli property follows.

Theorem 6.2.3. *There exists $K\geq 0$ and $\rho\in(0,1)$ such that*

$$\mathrm{Corr}(n)\leq K\rho^n \textit{ for } n=1,2,\ldots$$

Proof. Let $A\in\mathcal{F}_{l+n}$ for $l,n\geq 1$. There exists $B_\omega\in\mathcal{F}_0^\infty$ such that $A=\mathfrak{T}^{-(l+n)}(B_\omega)$ where $\mathfrak{T}$ represents the choice is random. Note that such B_ω exists since we refine the original partition by using the inverse images of every possible choice of the constituent map but the set depends on the sequence we choose so we denote the dependence with the subindex ω. Now we have

$$\begin{aligned}
\mu(A|\mathcal{F}_0^l) &= \frac{1}{\mu(C)}\int_A \chi_C d\mu \quad \text{for } C\in\beta^l \\
&= \frac{1}{\mu(C)}\int_{B_\omega}\mathcal{P}_{\omega_{l+n}}\ldots\mathcal{P}_{\omega_1}(\chi_C h)d\mathfrak{m}
\end{aligned}$$

If we take the integral of both sides with respect to the measure π we get

$$\mu(A|\mathcal{F}_0^l) = \frac{1}{\mu(C)} \int_\Omega \int_{B_\omega} \mathcal{P}_{\omega_{l+n}} \dots \mathcal{P}_{\omega_1}(\chi_C h) d\mathfrak{m} d\pi$$

where $h \in BV$ is the invariant density so that $d\mu = h\mathfrak{m}$.

Since $\pi \times \mu$ is the invariant measure for the skew product realization of the random dynamical system for $\mathscr{F} : \Omega \times Y \to \Omega \times Y$ we have

$$\begin{aligned} \pi \times \mu(\Omega \times A) &= \mu(A) = \pi \times \mu(\Omega, B_\omega) \\ &= \sum_\omega \pi([\omega]_{l+n})\mu(B_\omega) \end{aligned}$$

where $[\omega]_{l+n}$ is an (l+n)-cylinder and B_ω is the inverse image of A obtained by the cylinder $[\omega]_{l+n}$ so $\mathscr{F}([\omega]_{l+n}, B_\omega) = \Omega \times A$ and the sum is over all possible (l+n)-cylinders. Then we get

$$\begin{aligned} \left|\mu(A|\mathcal{F}_0^l) - \mu(A)\right| &= \left|\frac{1}{\mu(C)} \int_\Omega \int_{B_\omega} \mathcal{P}_{\omega_{l+n}} \dots \mathcal{P}_{\omega_1}(\chi_C h) d\mathfrak{m}\, d\pi - \int_\Omega \int_{B_\omega} h\, d\mathfrak{m}\, d\pi\right| \\ &\leq \int_\Omega \int_{B_\omega} \left|\frac{1}{\mu(C)} \mathcal{P}_{\omega_{l+n}} \dots \mathcal{P}_{\omega_1}(\chi_C h) - h\right| d\mathfrak{m}\, d\pi \\ &\leq \int_\Omega \int_Y \left|\frac{1}{\mu(C)} \mathcal{P}_{\omega_{l+n}} \dots \mathcal{P}_{\omega_1}(\chi_C h) - h\right| d\mathfrak{m}\, d\pi \\ &= \int_Y \left|\frac{1}{\mu(C)} \mathcal{P}_{\mathfrak{T}}^{l+n}(\chi_C h) - h\right| d\mathfrak{m} \\ &= \left\|\frac{\mathcal{P}_{\mathfrak{T}}^{l+n}(\chi_C h)}{\mu(C)} - h\right\|_1 \end{aligned}$$

which implies

$$\begin{aligned} \mathrm{Corr}(n) &= \sup_{l \geq 1} \mathbb{E}_\mu \left[\sup_{A \in \mathcal{F}_{l+n}^\infty} |\mu(A|\mathcal{F}_0^l) - \mu(A)|\right] \\ &\leq \sup_{l \geq 1} \sum_{C \in \beta^l} \mu(C) \cdot \left[\frac{\left\|\mathcal{P}_{\mathfrak{T}}^{l+n}(\chi_C h) - h\mu(C)\right\|_1}{\mu(C)}\right] \\ &= \sup_{l \geq 1} \sum_{C \in \beta^l} \left\|\mathcal{P}_{\mathfrak{T}}^{l+n}(\chi_C h) - h\mu(C)\right\|_1. \end{aligned}$$

We have $\mathcal{P}_{\mathfrak{T}} = \mathcal{Q} + \mathcal{R}$ where $\inf \|\mathcal{R}^n\|^{1/n} < \rho < 1$ for some $\rho \in \mathbb{R}$. Also note that the projection of $\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)$ to the eigenspace corresponding to the eigenvalue 1 is

$$\left(\int_Y \mathcal{P}^l_{\mathfrak{T}}(\chi_C h) d\mathfrak{m}\right) h = \left(\int_Y \chi_C h\, \mathcal{U}^l_{\mathfrak{T}}(1)\, d\mathfrak{m}\right) h = \left(\int_Y \chi_C h\, 1\, d\mathfrak{m}\right) h = \mu(C)h.$$

So if we consider the above quantity as the following way

$$\left\|\mathcal{P}^{l+n}_{\mathfrak{T}}(\chi_C h) - h\mu(C)\right\|_1 \quad = \quad \left\|\mathcal{P}^n_{\mathfrak{T}}(\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)) - h\mu(C)\right\|_1$$

then by applying $\mathcal{P}^n_{\mathfrak{T}} = \mathcal{Q} + \mathcal{R}^n$ to the term $\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)$ and then by subtracting the projection $\mathcal{Q}(\mathcal{P}^l_{\mathfrak{T}}(\chi_C h))$ we end up with

$$\left\|\mathcal{P}^{l+n}_{\mathfrak{T}}(\chi_C h) - h\mu(C)\right\|_1 = \left\|\mathcal{R}^n(\mathcal{P}^l_{\mathfrak{T}}(\chi_C h))\right\|_1.$$

Therefore,

$$\begin{aligned}\left\|\mathcal{P}^{l+n}_{\mathfrak{T}}(\chi_C h) - h\mu(C)\right\|_1 &\leq \|\mathcal{R}^n\|\|\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)\|_1 \\ &\leq K_1\rho^n\|\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)\|_{BV}\end{aligned}$$

implying

$$\begin{aligned}\mathrm{Corr}(n) &\leq \sup_{l\geq 1}\sum_{C\in\beta^l}\left\|\mathcal{P}^{l+n}_{\mathfrak{T}}(\chi_C h) - h\mu(C)\right\|_1 \\ &\leq K_1\rho^n\sup_{l\geq 1}\sum_{C\in\beta^l}\|\mathcal{P}^l_{\mathfrak{T}}(\chi_C h)\|_{BV} \\ &\leq K_1\rho^n(2F+1)\|h\|_{BV}.\end{aligned}$$

by using the proof of Lemma 5.1.3. We conclude the proof by setting $K = K_1(2F + 1)\|h\|_{BV}$. □

Corollary 6.2.4. *The random dynamical system $(\mathfrak{T}, \mu)$ is isomorphic to some Bernoulli shift, so $(\mathfrak{T}, \mu)$ is mixing, so ergodic.*

CHAPTER 7

Limit Theorems for Random Dynamical Systems

By having the Bernoulli property in Chapter 6 we could conclude that the limit theorems hold in the form proposed by Hofbauer and Keller [HK82] in the averaged sense since the corresponding skew realization has the Bernoulli property. The averaged Central Limit Theorem is also given by L. S. Young and E. Kobre for uniformly expanding maps with finite partition in [KY07] where they use martingale arguments. However we continue with perturbation methods to prove the averaged Central Limit Theorem since we also need the speed of convergence to get the quenched Central Limit Theorem. Our main reference for the Perturbation Theory results is [RE83] where the Central Limit Theorem is given for a single uniformly expanding map with the speed of convergence.

7.1 Characteristic Function Operators

Let us introduce some notation before we start giving the limit theorems. In Chapter 7 we define $\mathcal{P}$ to be the conjugate random Perron-Frobenius operator of $\mathcal{P}_{\mathfrak{T}}$.

Let $h \in BV$ be the eigenfunction corresponding to the unique eigenvalue $1 \in S^1$ so $\mathcal{P}_{\mathfrak{T}}(h) = h$. We define $\mathcal{M}_h$ to be the multiplication operator so $\mathcal{M}_h(f) = fh$. Then the conjugate operator $\mathcal{P}$ is defined by

$$\mathcal{P}(f) = \mathcal{M}_h^{-1}\mathcal{P}_{\mathfrak{T}}\mathcal{M}_h(f) = \mathcal{P}_{\mathfrak{T}}(fh)/h.$$

Note that $1 \in BV$ is the eigenfunction of $\mathcal{P}$ corresponding to the eigenvalue 1 since

$$\mathcal{P}(1) = \mathcal{P}_{\mathfrak{T}}(1h)/h = \mathcal{P}_{\mathfrak{T}}(h)/h = h/h = 1$$

which implies that the stationary measure for the system is $\mathfrak{m}$. The choice $\mathcal{P}$ is good for Chapter 7 since such terms $\displaystyle\sum_{i=1}^{N} p_i \sum_{y:T_i(y)=x} \frac{1}{|T_i'(y)|} = (\mathcal{P}1)(x)$ is simply equal to 1 which makes the calculations easier.

Furthermore it is also a right choice to work with since if we consider the corresponding skew product realization of the random process $\mathscr{F} : \Omega \times Y \to \Omega \times Y$ then the adjoint operator of the composition operator with $\mathscr{F}$ is denoted by $\mathcal{P}_{\mathscr{F}}$ and given by

$$\mathcal{P}_{\mathscr{F}}(f) = \frac{1}{h}\mathcal{P}_{\mathfrak{T}}(fh) = \mathcal{P}(f).$$

for $f \in BV(Y)$. See Lemma 16, page 75 in [Kob05] for the proof.

For $f \in C^\infty(Y, \mathbb{R})$, we denote $S_n^\omega f$ to be the sum of the evaluations of $T_{\omega_k} \circ \ldots \circ T_{\omega_2} \circ T_{\omega_1}(x)$ under f, namely

$$S_n^\omega f(x) = f(x) + f(T_{\omega_1}(x)) + f(T_{\omega_2} \circ T_{\omega_1}(x)) + \ldots + f(T_{\omega_{n-1}} \ldots T_{\omega_2} \circ T_{\omega_1}(x)).$$

The same process can also be given by using the corresponding skew product realization that is

$$S_n^\omega f(x) = f(x) + f \circ \mathscr{F}(\omega, x) + f \circ \mathscr{F}^2(\omega, x) + \ldots + f \circ \mathscr{F}^n(\omega, x)$$

Therefore to prove the Central Limit Theorem for the process $S_n(\omega, x)$ we use the operator $\mathcal{P}$ since f does not depend on the first coordinate $\omega \in \Omega$. Note that we use the notation $S_n^\omega f(x)$ to denote that the sum depends on ω not the function f itself.

Definition 7.1.1. *Define for $f \in C^\infty(Y, \mathbb{R})$ the operators $\mathcal{P}_{\mathfrak{T},t,f}$ and $\mathcal{P}_{t,f}$ to be*

$$\mathcal{P}_{\mathfrak{T},t,f}(g) = \mathcal{P}_{\mathfrak{T}}(e^{itf} g)$$

and

$$\mathcal{P}_{t,f}(g) = \mathcal{P}(e^{itf} g)$$

respectively where $\mathcal{P}_{\mathfrak{T}}$ is the random Perron-Frobenius operator and $\mathcal{P}$ is the conjugate random Perron-Frobenious operator of the random dynamical system $\mathfrak{T}$, so

$$\mathcal{P}_{\mathfrak{T},t,f}(g) = \sum_{j=1}^{N} p_j \mathcal{P}_{T_j}(e^{itf} g) = \sum_{j=1}^{N} p_j \mathcal{P}_{T_j,t,f}(g)$$

where $\mathcal{P}_{T_j,t,f}$ is the characteristic operator of the single map T_j given by $\mathcal{P}_{T_j,t,f}(g) = \mathcal{P}_{T_j}(e^{itf} g)$ as in [RE83], Section 1.5. We call $\mathcal{P}_{\mathfrak{T},t,f}$ the random characteristic operator *with respect to the observable f and $\mathcal{P}_{t,f}$* the characteristic operator *of the skew product system $\mathscr{F}$ with respect to the observable.*

The idea to prove the averaged Central Limit Theorem is to show that the characteristic function of the process

$$\frac{S_n^\omega f(x) - n\displaystyle\int f d\mu}{\sqrt{n}}$$

converges to the characteristic function of the normal distribution. We assume that $\displaystyle\int f d\mu = 0$ to make the calculations easier. Then the characteristic function of the process $\dfrac{S_n^\omega f(x)}{\sqrt{n}}$ is given by

$$\mu\left(\mathbb{E}_\Omega\left[e^{i\theta\frac{S_n^\omega f(x)}{\sqrt{n}}}\right]\right) \text{ for } \theta \in \mathbb{R}. \tag{7.1}$$

where $\mathbb{E}_\Omega$ is the expected value with respect to the measure π. Let $t = \frac{\theta}{\sqrt{n}} \in \mathbb{R}$ so our main interest is the function $\Psi(t) = \mu(\mathbb{E}_\Omega\left[e^{itS_n^\omega f(x)}\right])$. The following lemma gives how the characteristic function of the random process is related to the characteristic operator $\mathcal{P}_{t,f}^n$. Later we take $g = 1$ to prove the result.

Lemma 7.1.2. *For every $n \geq 1$ and $t \in \mathbb{R}$, we have*

$$\mu(\mathcal{P}_{t,f}^n(g)) = \mu(\mathbb{E}_\Omega\left[e^{itS_n f(\omega,x)}g(x)\right]).$$

Proof. First we find how the random characteristic operator and the characteristic operator are related. We have $\mathcal{P}(f) = \dfrac{\mathcal{P}_{\mathfrak{T}}(fh)}{h}$ which implies

$$\begin{aligned}
\mathcal{P}_{t,f}(g) &= \mathcal{P}(e^{itf}g) \\
&= \frac{\mathcal{P}_{\mathfrak{T}}(e^{itf}gh)}{h} \\
&= \frac{\mathcal{P}_{\mathfrak{T},t,f}(gh)}{h}
\end{aligned}$$

and similarly

$$\begin{aligned}\mathcal{P}^2_{t,f}(g) &= \mathcal{P}_{t,f}(\mathcal{P}_{t,f}(g))\\ &= \mathcal{P}_{t,f}\left(\frac{\mathcal{P}_{\mathfrak{T},t,f}(gh)}{h}\right)\\ &= \frac{\mathcal{P}^2_{\mathfrak{T},t,f}(gh)}{h}\end{aligned}$$

which generalizes to $\mathcal{P}^n_{t,f}(g) = \dfrac{\mathcal{P}^n_{\mathfrak{T},t,f}(gh)}{h}$. If we first find how the characteristic function of the process is related to the random characteristic operator $\mathcal{P}^n_{\mathfrak{T},t,f}$ then we can use that

$$\mu(\mathcal{P}^n_{t,f}(g)) = \mu(\frac{\mathcal{P}^n_{\mathfrak{T},t,f}(gh)}{h}) = \mathfrak{m}(\mathcal{P}^n_{\mathfrak{T},t,f}(gh)).$$

Therefore we need to find the expression for $\mathcal{P}^n_{\mathfrak{T},t,f}(g)$:

$$\begin{aligned}\mathfrak{m}(\mathcal{P}^n_{\mathfrak{T},t,f}(g)) &= \mathfrak{m}(\mathcal{P}_{\mathfrak{T}}(e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots)))) \text{ with n many } \mathcal{P}_{\mathfrak{T}}\text{s},\\ &= \mathfrak{m}(e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots))) \text{ with (n-1) many } \mathcal{P}_{\mathfrak{T}}\text{s}\\ &\quad \text{since } \mathfrak{m} \text{ is invariant,}\\ &= \mathfrak{m}(\mathcal{U}_{\mathfrak{T}}(e^{itf})e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots)) \text{ with (n-2) many } \mathcal{P}_{\mathfrak{T}}\text{s}\\ &\quad \text{since } \mathcal{U}_{\mathfrak{T}} \text{ is the adjoint operator,}\\ &= \mathfrak{m}(\sum_{j=1}^{N} p_j(e^{itf\circ T_j})e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots)) \text{ by Definition of } \mathcal{U}_{\mathfrak{T}},\\ &= \mathfrak{m}(\sum_{j=1}^{N} p_j e^{it(f+f\circ T_j)}\mathcal{P}_{\mathfrak{T}}(e^{itf}\ldots\mathcal{P}(e^{itf}g)\ldots))\\ &= \mathfrak{m}(\mathcal{U}(\sum_{j=1}^{N} p_j e^{it(f+f\circ T_j)})e^{itf}\ldots\mathcal{P}(e^{itf}g)\ldots) \text{ with (n-3) many } \mathcal{P}\text{s},\\ &= \mathfrak{m}(\sum_{j,k=1}^{N} p_k p_j e^{it(f\circ T_k+f\circ T_j\circ T_k)}e^{itf}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots)\\ &= \mathfrak{m}(\sum_{j,k=1}^{N} p_k p_j e^{it(f+f\circ T_k+f\circ T_j\circ T_k)}\ldots\mathcal{P}_{\mathfrak{T}}(e^{itf}g)\ldots)\end{aligned}$$

$$= \mathfrak{m}(\sum_{j=1}^{M_2} q_j^2 e^{itS_3^{\omega_j^2} f} \ldots \mathcal{P}_{\mathfrak{T}}(e^{itf} g) \ldots) \text{ where the sum is over all 2-cylinders}$$

$\ldots$

$$= \mathfrak{m}(\sum_{j=1}^{M_{n-1}} q_j^{n-1} e^{itS_n^{\omega_j^{n-1}} f} g) \text{ where the sum is over all (n-1)-cylinders, so}$$

$$= \mathfrak{m}(\mathbb{E}_\Omega \left[e^{itS_n^\omega f} g \right])$$

Here M_i denotes the number of possible combinations of i-cylinders, ω_j^i is one of the i-cylinders and q_j^i is the probability of the corresponding i-cylinder. The sum $S_{i+1}^{\omega_j^i} f$ is equal to $f + f \circ T_{\omega_1} + \ldots + f \circ T_{\omega_i} \circ \ldots T_{\omega_1}$ where $[T_{\omega_1}, \ldots, T_{\omega_i}]$ is the fixed i-cylinder denoted by ω_j^i.

Therefore, we get

$$\begin{aligned} \mu(\mathcal{P}_{t,f}^n(g)) &= \mathfrak{m}(\mathcal{P}_{\mathfrak{T},t,f}^n(gh)) \\ &= \mathfrak{m}(\mathbb{E}_\Omega \left[e^{itS_n^\omega f} gh \right]) \\ &= \mu(\mathbb{E}_\Omega \left[e^{itS_n^\omega f} g \right]). \end{aligned}$$

□

Proposition 7.1.3. *For every $t \in \mathbb{R}$, the characteristic operator $\mathcal{P}_{t,f}$ is continuous on $(BV, \|\cdot\|_{BV})$ and on $(L^1, \|\cdot\|_1)$. Furthermore, the function $t \to \mathcal{P}_{t,f}$ is analytic.*

Proof.

$$\|\mathcal{P}_{t,f}(g)\|_{BV} = \|\mathcal{P}(e^{itf} g)\|_{BV} \leq 2\|\mathcal{P}\|_{BV} \|e^{itf}\|_{BV} \|g\|_{BV}$$

since we have

$$\|fg\|_{BV} = \bigvee fg + \|fg\|_1 \leq |f|_\infty \bigvee g + |g|_\infty \bigvee f + |f|_\infty |g|_1 + |g|_\infty |f|_1$$

$$= |f|_\infty \|g\|_{BV} + |g|_\infty \|f\|_{BV}$$
$$\leq 2\|f\|_{BV}\,\|g\|_{BV}$$

since $|f|_\infty \leq \inf |f| + \bigvee f \leq \|f\|_1 + \bigvee f$. Furthermore

$$\begin{aligned}\|e^{itf}\|_{BV} &= \bigvee e^{itf} + \|e^{itf}\|_1 \\ &= \bigvee \cos(tf) + \bigvee \sin(tf) + 1 \\ &= 2|t| \bigvee f + 1.\end{aligned}$$

So $\|\mathcal{P}_{t,f}(g)\|_{BV} \leq C(t)\|g\|_{BV}$ which implies the continuity on $(BV, \|\cdot\|_{BV})$.

Similarly

$$\begin{aligned}\|\mathcal{P}_{t,f}(g)\|_1 &= \|\mathcal{P}(e^{itf}g)\|_1 \\ &\leq \|e^{itf}g\|_1 \leq \|g\|_1\end{aligned}$$

which implies the continuity on $(L^1\|\cdot\|_1)$.

To check the analyticity consider

$$\begin{aligned}\mathcal{P}_{t,f}(g) &= \mathcal{P}(e^{itf}g) \\ &= \mathcal{P}\left(\sum_{n=0}^{\infty} \frac{(itf)^n}{n!} g\right) \\ &= \sum_{n=0}^{\infty} \frac{(it)^n}{n!} \mathcal{P}(f^n g).\end{aligned}$$

For each term in the sum we have

$$\frac{|t|^n}{n!}\|\mathcal{P}(f^n g)\|_{BV} \leq \frac{2|t|^n}{n!}\|\mathcal{P}\|_{BV}\|f\|_{BV}^n\|g\|_{BV}.$$

which implies that $t \to \mathcal{P}_{t,f}$ is infinitely many differentiable with respect to t which implies the analyticity. $\square$

The following proposition is the main result that is used to prove the Central Limit Theorem. The proof depends on the spectral properties of the operator $\mathcal{P}$ and the proof for a Perron-Frobenius operator of a single map can be applied exactly the same way to the random Perron-Frobenius operator. So we give only a sketch for the proof and refer the reader to [RE83], Proposition 4 or to the book of Dunford and Schwartz, Part I, see [DS09].

The notation is as in Chapter 5 but for the operator $\mathcal{P}$ that has only 1 as its eigenvalue on S^1 so $\mathcal{P} = \mathcal{Q} + \mathcal{R}$.

Proposition 7.1.4. *There exists a real number $a > 0$ such that whenever $|t| < a$ we have*

(i) for every $g \in BV$ and $n \geq 1$,

$$\mathcal{P}^n_{t,f}(g) = \lambda^n(it)\mathcal{N}_t(g) + \mathcal{M}^n_t(g)$$

where $\lambda(it)$ is the unique greatest eigenvalue of $\mathcal{P}_{t,f}$ and $|\lambda(it)| > (2 + \rho(\mathcal{R})/3$ where $\rho(\mathcal{R})$ is the spectral radius of the operator $\mathcal{R}$, and $\mathcal{N}_t$ is the projection onto the eigenspace $\mathscr{E}_t$ corresponding to the eigenvalue $\lambda(it)$. $\mathcal{M}_t$ is an operator on BV with spectral radius $\rho(\mathcal{M}_t) \leq (1 + 2\rho(\mathcal{R}))/3$, and $\mathcal{M}_t(\mathscr{E}_t) = 0$.

(ii) the functions $t \to \lambda(it)$, $t \to \mathcal{N}_t$ and $t \to \mathcal{M}_t$ are analytic,

(iii) $\|\mathcal{M}^n_t(\mathbf{1})\|_{BV} \leq C|t|((1 + 2\rho(\mathcal{R}))/3)^n$ where C is a positive constant.

Proof. We start the proof by giving some definitions. Given the random Perron-Frobenius operator $\mathcal{P} = \mathcal{Q} + \mathcal{R}$ we define the resolvent of $\mathcal{P}$ to be the operator $\mathcal{S}(z)$

on BV defined by

$$\mathcal{S}(z) = 1/(zI - \mathcal{P}) = \mathcal{Q}/(z-1) + \sum_{n=0}^{\infty} \mathcal{R}^n / z^{n-1}$$

whenever $|z| > \rho(\mathcal{R})$ and $z \neq 1$. Then we define the resolvent of $\mathcal{P}_{t,f}$ to be

$$\mathcal{S}_t(z) = \mathcal{S}(z) \sum_{n=0}^{\infty} ((\mathcal{P}_{t,f} - \mathcal{P})\mathcal{S}(z))^n.$$

whenever $\|\mathcal{P}_{t,f} - \mathcal{P}\|_{BV} < 1/\|\mathcal{S}(z)\|_{BV}$ so that the series above converges.

Let $\mathscr{S}_1$ and $\mathscr{S}_2$ be the circles with center 1 and 0, and with radii $\rho_1 = (1-\rho(\mathcal{R}))/3$ and $\rho_2 = (1+2\rho(\mathcal{R}))/3$ respectively. Let $0 < \delta < \rho_1$, so

$$\rho(\mathcal{R}) + \delta < \rho_2$$

since $\rho_2 - \rho_1 = \rho(\mathcal{R})$. Let $M_\delta = \sup \|\mathcal{S}(z)\|_{BV}$ where the supremum is taken over $|z| > \rho(\mathcal{R}) + \delta$ and $|z-1| < \delta$. If $\|\mathcal{P}_{t,f} - \mathcal{P}\|_{BV} < 1/M_\delta$ then the circles $\mathscr{S}_1$ and $\mathscr{S}_2$ are in the resolvent set of $\mathcal{P}_{t,f}$. Then the projection operators are

$$\mathcal{N}_t = \frac{1}{2\pi i} \int_{\mathscr{S}_1} \mathcal{S}_t(z) dz$$

$$\mathcal{M}'_t = \frac{1}{2\pi i} \int_{\mathscr{S}_2} \mathcal{S}_t(z) dz$$

For $\|\mathcal{N}_t - \mathcal{Q}\|_{BV} < 1$ the image of $\mathcal{N}_t$, say $\mathscr{E}_t$ is one dimensional. So for any g_t that generates $\mathscr{E}_t$ we have

$$\mathcal{P}_{t,f}\mathcal{N}_t(g_t) = \mathcal{N}_t\mathcal{P}_{t,f}(g_t) = \lambda(it)g_t.$$

Then for any $n \geq 1$,

$$\mathcal{P}_{t,f}^n = \mathcal{P}_{t,f}^n\mathcal{N}_t + \mathcal{P}_{t,f}^n\mathcal{M}'_t = \lambda^n(it)\mathcal{N}_t + \mathcal{M}_t^n$$

where

$$\mathcal{M}_t^n = \frac{1}{2\pi i} \int_{\mathscr{S}_2} z^n \mathcal{S}_t(z) dz.$$

There exists $a \in \mathbb{R}$ such that for $|t| < a$ we have

$$\mathcal{S}_t(z) = \mathcal{S}(z) + it\mathcal{S}_t^{(1)}(z),$$

implying

$$\begin{aligned} \mathcal{M}_t^n(1) &= \frac{1}{2\pi i} \int_{\mathscr{S}_2} z^n \mathcal{S}(z) dz + \frac{it}{2\pi i} \int_{\mathscr{S}_2} z^n \mathcal{S}_t^{(1)}(z) dz \\ &= \frac{t}{2\pi} \int_{\mathscr{S}_2} z^n \mathcal{S}_t^{(1)}(z) dz, \end{aligned}$$

implying $\|\mathcal{M}_t^n(1)\|_{BV} \leq C|t|\rho_2^n$ where

$$C = \frac{1}{2\pi} \sup_{|z|=\rho_2, |t|<a} \|\mathcal{S}_t^{(1)}(z)\|_{BV}.$$

□

Then again the spectral properties of the operator $\mathcal{P}_{t,f}$ for every real number t is provided by Theorem 5.1.2:

Proposition 7.1.5. *For $t \in \mathbb{R}$ with $|t|$ small enough, the operator $\mathcal{P}_{t,f}$ has only finitely many eigenvalues of modulo 1. For each such eigenvalue, say $\zeta \in \sigma(\mathcal{P}_{t,f}) \cap S^1$ the corresponding eigenspace $\mathscr{E}_\zeta$ is finite dimensional and contained in BV. We have*

$$\mathcal{P}_{t,f}^n = \sum_{j=1}^{k} \zeta_j^n \mathcal{Q}_t^{(j)} + \mathcal{R}_t^n, \text{ for } n \geq 1$$

where $\mathcal{Q}_t^{(j)}$ is the projection to the eigenspace $\mathscr{E}_{\zeta_j}$. Furthermore

$$\mathcal{Q}_t^{(j)} \mathcal{Q}_t^{(i)} = 0 \text{ for } i \neq j, \quad (\mathcal{Q}_t^{(j)})^2 = \mathcal{Q}_t^{(j)}, \quad \mathcal{Q}_t^{(j)} \mathcal{R}_t = \mathcal{R}_t \mathcal{Q}_t^{(j)} = 0.$$

And $\mathcal{R}_t(BV) \subset BV$ with $\rho(\mathcal{R}_t) < 1$.

Proof. We only need to check if the operator $\mathcal{P}_{t,f}$ satisfies the Lasota-York inequality, the rest is the result of Theorem 5.1.2. First note that we have

$$\mathcal{P}^n_{\mathfrak{T},t,f}(g) = \sum_{i=1}^{M} q_i \mathcal{P}_{T^{(i)}_{\omega_n}} \dots \mathcal{P}_{T^{(i)}_{\omega_1}}(e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}g)$$

where q_i is the probability of the ith n-cylinder given by $T^{(i)}_{\omega_n} \dots T^{(i)}_{\omega_1}$ and the sum is over all possible n-cylinders which are $M = N^n$ many. The equality is obtained by applying the Equation 3.1 to maps e^{itf} and $e^{itf}g$ so

$$e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}g) = \sum_{i=1}^{N} p_i \mathcal{P}_{T_i}\left(e^{itf}g \cdot e^{itf\circ T_i}\right) = \sum_{i=1}^{N} p_i \mathcal{P}_{T_i}\left(g \cdot e^{it(f+f\circ T_i)}\right)$$

implying for $n = 2$ that

$$\begin{aligned}
\mathcal{P}^2_{\mathfrak{T},t,f}(g) &= \mathcal{P}_{\mathfrak{T}}(e^{itf}\mathcal{P}_{\mathfrak{T}}(e^{itf}g)) \\
&= \mathcal{P}_{\mathfrak{T}}\left(\textstyle\sum_{i=1}^{N} p_i \mathcal{P}_{T_i}\left(g \cdot e^{it(f+f\circ T_i)}\right)\right) \\
&= \textstyle\sum_{i,j=1}^{N} p_j p_i \mathcal{P}_{T_j}\mathcal{P}_{T_i}\left(g \cdot e^{it(f+f\circ T_i)}\right) \\
&= \textstyle\sum_{i,j=1}^{N} p_j p_i \mathcal{P}_{T_j\circ T_i}\left(g \cdot e^{it(f+f\circ T_i)}\right)
\end{aligned}$$

We use the same rule above for any $n \geq 1$. For $\|\mathcal{P}^n_{\mathfrak{T},t,f}(g)\|_{BV}$ we use the inequality 4.7 in Remark 4.1.16 to get

$$\begin{aligned}
\|\mathcal{P}^n_{\mathfrak{T},t,f}(g)\|_{BV} &= \left\|\sum_{i=1}^{M} q_i \mathcal{P}_{T^{(i)}_{\omega_n}} \dots \mathcal{P}_{T^{(i)}_{\omega_1}}(e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}g)\right\|_{BV} \\
&\leq \sum_{i=1}^{M} q_i \left\|\mathcal{P}_{T^{(i)}_{\omega_n}} \dots \mathcal{P}_{T^{(i)}_{\omega_1}}(e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}g)\right\|_{BV} \\
&\leq \sum_{i=1}^{M} q_i \left(Cr^n \left\|e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}g\right\|_{BV}\right. \\
&+ \left. R\|e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}g\|_1\right) \\
&\leq \sum_{i=1}^{M} q_i \left(Cr^n 2\left\|e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})}\right\|_{BV}\|g\|_{BV} + R\|g\|_1\right)
\end{aligned}$$

$$\leq \sum_{i=1}^{M} q_i(Cr^n 2(\bigvee_Y e^{it(f+f\circ T^{(i)}_{\omega_1}+\ldots+f\circ T^{(i)}_{\omega_{i-1}}\circ\ldots\circ T^{(i)}_{\omega_1})} + 1)\|g\|_{BV}$$

$$+ R\|g\|_1)$$

$$\leq \sum_{i=1}^{M} q_i(Cr^n 2(2|t|\bigvee_Y (f + f\circ T^{(i)}_{\omega_1} + \ldots + f\circ T^{(i)}_{\omega_{i-1}} \circ \ldots \circ T^{(i)}_{\omega_1})$$

$$+ 1)\|g\|_{BV} + R\|g\|_1)$$

$$\leq \sum_{i=1}^{M} q_i \left(Cr^n 2(2|t|n \bigvee_Y f + 1)\|g\|_{BV} + R\|g\|_1 \right)$$

$$\leq Cr^n 2 \left(2|t|n \bigvee_Y f + 1 \right) \|g\|_{BV} + R\|g\|_1 \text{ since } \sum_{i=1}^{M} q_i = 1,$$

Now, for every $t \in \mathbb{R}$ there exists n_0 such that $Cr^{n_0} 2(2|t|n_0 \bigvee_Y f + 1) < 1$ which gives the Lasota-York inequality for $\mathcal{P}_{\mathfrak{T},t,f}$ but then $\mathcal{P}^n_{t,f}(g) = \dfrac{\mathcal{P}^n_{\mathfrak{T},t,f}(gh)}{h}$ implies

$$\begin{aligned}
\|\mathcal{P}^n_{t,f}(g)\|_{BV} &\leq 2\left\|\frac{1}{h}\right\|_{BV} \left\|\mathcal{P}^n_{\mathfrak{T},t,f}(gh)\right\|_{BV}, \\
&\leq 2\left\|\frac{1}{h}\right\|_{BV} \left(Cr^n 2 \left(2|t|n \bigvee_Y f + 1 \right) \|gh\|_{BV} + R\|gh\|_1 \right) \\
&\leq 2\left\|\frac{1}{h}\right\|_{BV} 2 \left(Cr^n 2 \left(2|t|n \bigvee_Y f + 1 \right) \|g\|_{BV}\|h\|_{BV} + R\|g\|_1\|h\|_\infty \right)
\end{aligned}$$

where $\left\|\frac{1}{h}\right\|_{BV}$, $\|h\|_{BV}$ and $\|h\|_\infty$ are all finite so the characteristic operator $\mathcal{P}_{t,f}$ satisfies the Lasota-York inequality, too. □

7.2 Central Limit Theorem

The condition on the function $f \in BV$ is that the equation

$$f = k + \varphi \circ T_i - \varphi \text{ for every } i \in \{1, \ldots, L\} \tag{7.2}$$

admits no solution for $\varphi \in BV$ and $k \in \mathbb{R}$.

Theorem 7.2.1. *(Averaged Central Limit Theorem) The random dynamical system $(\mathfrak{T}, \mathfrak{m})$ satisfies the averaged Central Limit Theorem for every $f \in C^\infty(Y, \mathbb{R})$ with the above condition (7.2), meaning*

$$\lim_{n\to\infty} \pi \times \mu \left\{ (\omega, x) : \frac{S_n^\omega f(x) - n\displaystyle\int f d\mu}{\sqrt{n}} < c \right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{c} e^{-t^2/2\sigma^2} dt.$$

where the variance σ^2 is given by

$$\sigma^2 = \mu(f^2) + 2\sum_{n=1}^{\infty} \mu(f\mathcal{P}^n f). \tag{7.3}$$

We assume that $\displaystyle\int f d\mu = 0$ to simplify the calculations. First, note that for $|t| < a$ as in Proposition 7.1.4 we have

$$\begin{aligned}
\int_Y \int_\Omega e^{iS_n^\omega f} d\pi d\mu &= \int_Y \mathbb{E}_\Omega \left[e^{itS_n^\omega f} \right] d\mu \\
&= \int_Y \mathcal{P}_{t,f}^n(1) d\mu \text{ by Lemma 7.1.2,} \\
&= \lambda^n(it) \int_Y \mathcal{N}_t(1) d\mu + \int_Y \mathcal{M}_t^n(1) d\mu.
\end{aligned}$$

Lemma 7.2.2. *For λ as in Proposition 7.1.4 we have*

$$\lambda'(0) = \mu(f).$$

Proof. We know that

$$\int_Y \mathbb{E}_\Omega \left[e^{i\frac{t}{n}S_n^\omega f} \right] d\mu = \int_Y \mathcal{P}_{\frac{t}{n},f}(1) d\mu,$$

and if we take the limit over n of the left-hand side we get

$$\lim_{n\to\infty} \int_Y \mathbb{E}_\Omega[e^{i\frac{t}{n}S_n^\omega f}] d\mu = e^{it\mu(f)}$$

since $\lim_{n\to\infty} \mathbb{E}_\Omega[S_n^\omega f] = \mu(f)$. Then the aim is to find the limit of right-hand side and set it equal to $e^{it\mu(f)}$. By Proposition 7.1.4 we know that the right-hand side is equal to

$$\int_Y \mathbb{E}_\Omega[e^{i\frac{t}{n}S_n^\omega f}]d\mu = \lambda^n(i\frac{t}{n})\int_Y \mathcal{N}_{\frac{t}{n}}(1)d\mu + \int_Y \mathcal{M}^n_{\frac{t}{n}}(1)d\mu.$$

Then we need to find the limit of each term as $n \to \infty$. First,

$$\left|\int_Y \mathcal{M}^n_{\frac{t}{n}}(1)d\mu\right| \leq \|\mathcal{M}^n_{\frac{t}{n}}(1)\|_{BV} \leq C\frac{|t|}{n}\rho_2^n$$

with $\rho_2 < 1$ so $\lim_{n\to\infty} \left|\int_Y \mathcal{M}^n_{\frac{t}{n}}(1)d\mu\right| = 0$. On the other hand,

$$\mathcal{N}_{\frac{t}{n}}(1) = \mathcal{Q} + \frac{it}{n}\mathcal{N}^{(1)} - \frac{t^2}{n^2}\mathcal{N}^{(2)} + \frac{t^2}{n^2}\overline{\mathcal{N}}_{\frac{t}{n}} \tag{7.4}$$

where $\mathcal{N}^{(1)}, \mathcal{N}^{(2)}$ and $\overline{\mathcal{N}}_{\frac{t}{n}}$ are operators on BV such that

$$\lim_{n\to\infty} \|\overline{\mathcal{N}}_{\frac{t}{n}}\|_{BV} = 0.$$

Therefore

$$\lim_{n\to\infty} \int_Y \mathcal{N}_{\frac{t}{n}}(1)d\mu = \int_Y \mathcal{Q}(1)d\mu = 1.$$

Similarly,

$$\lambda(i\frac{t}{n}) = \lambda(0) + \frac{it}{n}\lambda'(0) + \frac{(it)^2}{n^2}\lambda''(0) + \frac{t^2}{n^2}\overline{\lambda}(i\frac{t}{n}), \tag{7.5}$$

with $\lambda(0) = 1$, $\lim_{n\to\infty} \overline{\lambda}(i\frac{t}{n}) = 0$, and $\lim_{n\to\infty} \lambda^n(i\frac{t}{n}) = e^{it\lambda'(0)}$. Thus the limit of right-hand side as $n \to \infty$ is $e^{it\lambda'(0)}$ which is equal to the left-hand side $e^{it\mu(f)}$ implying $\lambda'(0) = \mu(f)$. □

We assume $\mu(f) = 0$ so $\lambda'(0) = 0$.

Lemma 7.2.3. *For λ as in Proposition 7.1.4 we have*

$$\lambda''(0) = \lim_{n\to\infty} \int_Y \mathbb{E}_\Omega \left[\left(\frac{S_n^\omega f}{\sqrt{n}} \right)^2 \right] d\mu.$$

Proof. The idea is similar to the previous proof: we give how $\lim_{n\to\infty} \int_Y \mathbb{E}_\Omega[(\frac{S_n^\omega f}{\sqrt{n}})^2] d\mu$ is related to the random characteristic operator and then use Proposition 7.1.4 and calculate each term.

Note that $\int_Y \mathbb{E}_\Omega[(\frac{S_n^\omega f}{\sqrt{n}})^2] d\mu$ can be obtained by using the characteristic function $\int_Y \mathbb{E}_\Omega[e^{(it/\sqrt{n})S_n^\omega f}] d\mu$ by taking the derivative twice with respect to t and then evaluating at $t = 0$:

$$\frac{\partial^2}{\partial t^2} \left\{ \int_Y \mathbb{E}_\Omega[e^{(it/\sqrt{n})S_n^\omega f}] d\mu \right\}_{t=0} = - \int_Y \mathbb{E}_\Omega[(\frac{S_n^\omega f}{\sqrt{n}})^2] d\mu.$$

Again we have that

$$\int_Y \mathbb{E}_\Omega[e^{i\frac{t}{\sqrt{n}}S_n^\omega f}] d\mu = \int_Y \mathcal{P}^n_{\frac{t}{\sqrt{n}},f}(1) d\mu,$$

implies together with Proposition 7.1.4 that

$$\int_Y \mathbb{E}_\Omega[e^{i\frac{t}{\sqrt{n}}S_n^\omega f}] d\mu = \lambda^n(i\frac{t}{\sqrt{n}}) \int_Y \mathcal{N}_{\frac{t}{\sqrt{n}}}(1) d\mu + \int_Y \mathcal{M}^n_{\frac{t}{\sqrt{n}}}(1) d\mu$$

as in previous proof except that we have $\sqrt{n}$ instead of n in the denominator of the perturbation value. Therefore we need to find the second derivative of each term on the right-hand side with respect to t and then evaluate at $t = 0$ and take the limit as $n \to \infty$.

Now, we have

$$\mathcal{M}^n_{\frac{t}{\sqrt{n}}}(1) = \frac{1}{2\pi i} \int_{S_2} z^n \mathcal{S}_{it/\sqrt{n}}(z)(1) dz.$$

For n sufficiently large and $|z| = \rho_2$ for $\mathcal{S}_{it/\sqrt{n}}(z)$ we have

$$\mathcal{S}_{it/\sqrt{n}}(z) = \mathcal{S}(z) + \frac{it}{\sqrt{n}}\mathcal{S}^{(1)}(z) - \frac{t^2}{2n}\mathcal{S}^{(2)}(z) + \frac{t^2}{n}\overline{\mathcal{S}}_{it/\sqrt{n}}(z)$$

where $\mathcal{S}^{(1)}$,$\mathcal{S}^{(2)}$ and $\overline{\mathcal{S}}_{it/\sqrt{n}}$ are operators on BV and $\lim_{n\to\infty} \|\overline{\mathcal{S}}_{it/\sqrt{n}}|_{BV} = 0$. Therefore,

$$\begin{aligned} \mathcal{M}^n_{\frac{t}{\sqrt{n}}}(1) &= \frac{1}{2\pi i}\int_{S_2} z^n \left(\mathcal{S}(z) + \frac{it}{\sqrt{n}}\mathcal{S}^{(1)}(z) - \frac{t^2}{2n}\mathcal{S}^{(2)}(z) + \frac{t^2}{n}\overline{\mathcal{S}}_{it/\sqrt{n}}(z)\right)(1)dz \\ &= \frac{t}{2\pi\sqrt{n}}\int_{S_2} z^n \mathcal{S}^{(1)}_{it/\sqrt{n}}(z)(1)dz - \frac{t^2}{4i\pi n}\int_{S_2} z^n\mathcal{S}^{(2)}(z)(1)dz \\ &\quad + \frac{t^2}{2i\pi n}\int_{S_2} z^n\overline{\mathcal{S}}_{it/\sqrt{n}}(z)(1)dz, \end{aligned}$$

implying

$$\lim_{n\to\infty}\frac{\partial^2}{\partial t^2}\left\{\int_Y \mathcal{M}^n_{\frac{t}{\sqrt{n}}}(1)d\mu\right\}_{t=0} = \lim_{n\to\infty}\frac{-1}{2i\pi n}\int_{S_2} z^n\mathcal{S}^{(2)}(z)(1)dz = 0.$$

Now for the other term $\lambda^n(i\frac{t}{\sqrt{n}})\int_Y \mathcal{N}_{\frac{t}{\sqrt{n}}}(1)d\mu$ we use the results from the previous proof, namely the Equations (7.4) and (7.5), and replace $\sqrt{n}$ with n. Then we take the second derivative and evaluate at $t = 0$ and get

$$\frac{\partial^2}{\partial t^2}\left\{\lambda^n(i\frac{t}{\sqrt{n}})\int_Y \mathcal{N}_{\frac{t}{\sqrt{n}}}(1)d\mu\right\}_{t=0} = -\lambda''(0) - \frac{1}{n}\mathcal{N}^{(2)}(1),$$

and if we take the limit as $n \to \infty$ the sequence $-\int_Y \mathbb{E}_\Omega[(\frac{S^\omega_n f}{\sqrt{n}})^2]d\mu$ converges to $-\lambda''(0)$ so the result follows. □

The following lemma gives the representation of the variance σ^2 in terms of the conjugate random Perron-Frobenius operator:

Lemma 7.2.4. *If* $\sigma^2 = \lim_{n\to\infty}\int_Y \mathbb{E}_\Omega\left[\left(\frac{S^\omega_n f}{\sqrt{n}}\right)^2\right]d\mu$, *then*

$$\sigma^2 = \int_Y \mathcal{P}(g^2) - (\mathcal{P}g)^2 d\mu$$

where $\mathcal{P}$ *is the random Perron-Frobenius operator and* $g = (I - \mathcal{P})^{-1}f$.

Proof. Let us first give in detail how $\int_Y \mathbb{E}_\Omega \left[\left(\frac{S_n^\omega f}{\sqrt{n}} \right)^2 \right] d\mu$ behaves for some fixed n values:

For $n = 2$, $S_2^\omega f = f + f \circ T_1$ with probability p_1 and $S_2^\omega f = f + f \circ T_2$ with probability p_2 and so on. So

$$\begin{aligned}
\int_Y \mathbb{E}_\Omega \left[\left(\frac{S_2^\omega f}{\sqrt{2}} \right)^2 \right] d\mu &= \int_Y \left(\frac{\sum_{i=1}^N p_i (f + f \circ T_i)^2}{2} \right) h \, d\mathfrak{m} \\
&= \int_Y \frac{\sum_{i=1}^N p_i (f^2 + 2f \cdot f \circ T_i + (f \circ T_i)^2)}{2} h \, d\mathfrak{m} \\
&= \frac{1}{2} \left[\int_Y f^2 d\mu + 2 \int_Y \sum_{i=1}^N p_i f \cdot (f \circ T_i) h \, d\mathfrak{m} \right. \\
&\quad \left. + \sum_{i=1}^N p_i \int_Y f \circ T_i \cdot (f \circ T_i) h \, d\mathfrak{m} \right] \\
&= \frac{1}{2} \left[\int_Y f^2 d\mu + 2 \int_Y \mathcal{P}_{\mathfrak{T}}(fh) \cdot f d\mathfrak{m} \right. \\
&\quad \left. + \sum_{i=1}^N p_i \int_Y \mathcal{P}_{T_i}(h \cdot f \circ T_i) \cdot f \, d\mathfrak{m} \right] \\
&= \frac{1}{2} \left[\int_Y f^2 d\mathfrak{m} + 2 \int_Y \mathcal{P}_{\mathfrak{T}}(fh) \cdot f d\mathfrak{m} + \int_Y f \mathcal{P}_{\mathfrak{T}}(h) \cdot f d\mathfrak{m} \right] \\
&= \frac{1}{2} \left[\int_Y f^2 d\mathfrak{m} + 2 \int_Y \mathcal{P}_{\mathfrak{T}}(fh)/h \cdot f d\mu + \int_Y f^2 h d\mathfrak{m} \right]
\end{aligned}$$

since we have

$$\sum_{i=1}^N p_i \mathcal{P}_{T_i}(f_1 \cdot f \circ T_i) = f \mathcal{P}_{\mathfrak{T}}(f_1)$$

so we take $f_1 = h$ for the above equality, see Equation 3.1. Thus,

$$\int_Y \mathbb{E}_\Omega \left[\left(\frac{S_2^\omega f}{\sqrt{2}} \right)^2 \right] d\mu = \int_Y f^2 d\mu + \int_Y \mathcal{P}(f) \cdot f d\mu.$$

since $\mathcal{P}_{\mathfrak{T}}(h) = h$ and $\mathcal{P}_{\mathfrak{T}}(fh)/h = \mathcal{P}(f)$.

For $n = 3$, $S_3^\omega f = f + f \circ T_1 + f \circ T_1 \circ T_1$ with probability p_1^2, $S_3^\omega f = f + f \circ T_1 + f \circ T_2 \circ T_1$ with probability $p_1 p_2$ and $S_3^\omega f = f + f \circ T_2 + f \circ T_1 \circ T_2$ with probability $p_2 p_1$ and so on. Then we get

$$\begin{aligned}
&\int_Y \mathbb{E}_\Omega\left[\left(\frac{S_3^\omega f}{\sqrt{3}}\right)^2\right] d\mu = \int_Y \left(\frac{\sum_{i,j=1}^N p_i p_j (f + f\circ T_i + f\circ T_j \circ T_i)^2}{3}\right) h d\mathfrak{m}\\
=&\int_Y \frac{\sum_{i,j=1}^N p_i p_j (f^2 + (f\circ T_i)^2 + (f\circ T_j\circ T_i)^2}{3} h\, d\mathfrak{m}\\
&+ \int_Y \frac{\sum_{i,j=1}^N p_i p_j 2((f\cdot f\circ T_i + (f\cdot f\circ T_j \circ T_i + (f\circ T_i \cdot f\circ T_j\circ T_i))}{3} h\, d\mathfrak{m}\\
=&\frac{1}{3}\left[\int_Y f^2 h\, d\mathfrak{m} + \int_Y \sum_{i=1}^N p_i (f\circ T_i)^2 h d\mathfrak{m} + \int_Y \sum_{i,j=1}^N p_i p_j (f\circ T_j \circ T_i)^2 h d\mathfrak{m}\right]\\
&+\frac{2}{3}\left[\int_Y f \sum_{i=1}^N p_i (f\circ T_i) h d\mathfrak{m} + \int_Y f \sum_{i,j=1}^N p_i p_j (f\circ T_j\circ T_i) h d\mathfrak{m}\right.\\
&\left.+\int_Y \sum_{i,j=1}^N p_i p_j (f\circ T_i)(f\circ T_j \circ T_i) h d\mathfrak{m}\right]\\
=&\frac{1}{3}\left[\int_Y f^2 d\mu + \int_Y \sum_{i=1}^N p_i \mathcal{P}_{T_i}(h\cdot f\circ T_i)\cdot f\, d\mathfrak{m}\right.\\
&\left.+\int_Y \sum_{i,j=1}^N p_i p_j \mathcal{P}_{T_i}(h\cdot f\circ T_j \circ T_i)(f\circ T_j) d\mathfrak{m}\right] + \frac{2}{3}\left[\int_Y \mathcal{P}_{\mathfrak{T}}(fh)\cdot f d\mathfrak{m}\right.\\
&\left.+\int_Y \mathcal{P}_{\mathfrak{T}}(fh)\sum_{j=1}^N p_j (f\circ T_j) d\mathfrak{m} + \int_Y \sum_{i,j=1}^N p_j p_i \mathcal{P}_{T_i}(h\cdot f\circ T_i)\cdot(f\circ T_j) d\mathfrak{m}\right]\\
=&\frac{1}{3}\left[\int_Y f^2 d\mu + \int_Y fh\cdot f d\mathfrak{m} + \int_Y \sum_{j=1}^N p_j (h\cdot f\circ T_j)(f\circ T_j) d\mathfrak{m}\right]\\
&+\frac{2}{3}\left[\int_Y \mathcal{P}_{\mathfrak{T}}(fh)\cdot f d\mathfrak{m} + \int_Y \mathcal{P}_{\mathfrak{T}}^2(fh)\cdot f d\mathfrak{m} + \int_Y \sum_{i,j=1}^N p_j fh\cdot f\circ T_j d\mathfrak{m}\right]\\
=&\frac{1}{3}\left[\int_Y f^2 d\mu + \int_Y f^2 d\mu + \int_Y \sum_{i=1}^N p_i \mathcal{P}_{T_i}(h\cdot f\circ T_i)\cdot f d\mathfrak{m}\right]\\
&+\frac{2}{3}\left[\int_Y \mathcal{P}_{\mathfrak{T}}(fh)\cdot f d\mathfrak{m} + \int_Y \mathcal{P}_{\mathfrak{T}}^2(fh)\cdot f d\mathfrak{m} + \int_Y \mathcal{P}_{\mathfrak{T}}(fh)\cdot f d\mathfrak{m}\right]\\
=&\frac{1}{3}\left[3\int_Y f^2 d\mu\right] + \frac{2}{3}\left[2\int_Y \mathcal{P}(f)\cdot f d\mu + \int_Y \mathcal{P}^2(f)\cdot f d\mu\right]
\end{aligned}$$

since $\mathcal{P}_{\mathfrak{T}}^2(fh)/h = \mathcal{P}^2(fh)/h$. Thus,

$$\int_Y \mathbb{E}_\Omega\left[\left(\frac{S_3^\omega f}{\sqrt{3}}\right)^2\right] d\mu = \int_Y f^2 d\mu + \frac{4}{3}\int_Y \mathcal{P}(f)\cdot f d\mu + \frac{2}{3}\int_Y \mathcal{P}^2(f)\cdot f d\mu.$$

Here the main step is when we multiply the two forms $f\circ T_j\circ T_i$ and $f\circ T_i$ we get

$$\begin{aligned}\int_Y \sum_{i,j=1}^N p_i p_j (f\circ T_i)(f\circ T_j\circ T_i) h \, d\mathfrak{m} &= \int_Y \sum_{i,j=1}^N p_j p_i \mathcal{P}_{T_i}(h\cdot f\circ T_i)(f\circ T_j) d\mathfrak{m} \\ &= \int_Y \sum_{j=1}^N p_j h f (f\circ T_j) d\mathfrak{m} \\ &= \int_Y \mathcal{P}_{\mathfrak{T}}(hf) f d\mathfrak{m} \\ &= \int_Y \mathcal{P}(f) f d\mu.\end{aligned}$$

And by using the same methods in more general combinations like the ones below, we get the integral of a product of second iterates and fourth iterates of random maps which is given by

$$\begin{aligned}&\int_Y \sum_{i,j,k,l=1}^N p_i p_j p_k p_l (f\circ T_j\circ T_i)(f\circ T_l\circ T_k\circ T_j\circ T_i) h \, d\mathfrak{m} \\ &= \int_Y \sum_{i,j,k,l=1}^N p_i p_j p_k p_l \mathcal{P}_{T_i}(h\cdot f\circ T_j\circ T_i)(f\circ T_l\circ T_k\circ T_j) d\mathfrak{m} \\ &= \int_Y \sum_{j,k,l=1}^N p_j p_k p_l (h\cdot f\circ T_j)(f\circ T_l\circ T_k\circ T_j) d\mathfrak{m} \\ &= \int_Y \sum_{j,k,l=1}^N p_j p_k p_l \mathcal{P}_{T_j}(h\cdot f\circ T_j)(f\circ T_l\circ T_k) d\mathfrak{m} \\ &= \int_Y \sum_{k,l=1}^N p_k p_l (h\cdot f)(f\circ T_l\circ T_k) d\mathfrak{m} \\ &= \int_Y \sum_{l=1}^N p_l \mathcal{P}_{\mathfrak{T}}(fh)(f\circ T_l) d\mathfrak{m} \\ &= \int_Y \mathcal{P}_{\mathfrak{T}}^2(fh) f d\mathfrak{m} \\ &= \int_Y \mathcal{P}^2(f) f d\mu.\end{aligned}$$

For $n = 4$ we have

$$
\begin{aligned}
&\int_Y \mathbb{E}_\Omega\left[\left(\frac{S_4^\omega f}{\sqrt{4}}\right)^2\right] h\, d\mathfrak{m} \\
&= \int_Y \left(\frac{\sum_{i,j,k=1}^N p_i p_j p_k (f + f\circ T_i + f\circ T_j \circ T_i + f\circ T_k \circ T_j \circ T_i)^2}{4}\right) h d\mathfrak{m} \\
&= \frac{1}{4}\int_Y \sum_{i,j=1}^N p_i p_j p_k (f^2 + (f\circ T_i)^2 + (f\circ T_j\circ T_i)^2 + (f\circ T_k \circ T_j \circ T_i)^2 h d\mathfrak{m} \\
&+ \frac{2}{4}\int_Y \sum_{i,j,k=1}^N p_i p_j p_k (f\cdot f\circ T_i) + (f\cdot f\circ T_j \circ T_i) + (f\cdot f\circ T_k\circ T_j\circ T_i) h d\mathfrak{m} \\
&+ \frac{2}{4}\int_Y \sum_{i,j,k=1}^N p_i p_j p_k (f\circ T_i \cdot f\circ T_j\circ T_i) + (f\circ T_i\cdot f\circ T_k\circ T_j\circ T_i) h d\mathfrak{m} \\
&+ \frac{2}{4}\int_Y \sum_{i,j,k=1}^N p_i p_j p_k (f\circ T_j\circ T_i \cdot f\circ T_k\circ T_j\circ T_i) h d\mathfrak{m} \\
&= \frac{1}{4}\int_Y 4f^2 d\mu + \frac{2}{4}\int_Y \mathcal{P}(f)f + \mathcal{P}^2(f)f + \mathcal{P}^3(f)f d\mu + \frac{2}{4}\int_Y \mathcal{P}(f)f \\
&+ \mathcal{P}^2(f)f d\mu + \tfrac{2}{4}\int_Y \mathcal{P}(f)f d\mu \\
&= \int_Y f^2 d\mu + \frac{6}{4}\int_Y \mathcal{P}(f)f d\mu + \frac{4}{4}\int_Y \mathcal{P}^2(f)f d\mu + \frac{2}{4}\int_Y \mathcal{P}^3(f)f d\mu.
\end{aligned}
$$

By using the same idea above we can write for any n that

$$
\begin{aligned}
\int_Y \mathbb{E}_\Omega\left[\left(\frac{S_n^\omega f}{\sqrt{n}}\right)^2\right] d\mu &= \int_Y f^2 d\mu + \frac{2(n-1)}{n}\int_Y \mathcal{P}(f)f d\mu \\
&+ \frac{2(n-2)}{n}\int_Y \mathcal{P}^2(f)f d\mu \\
&+ \frac{2(n-3)}{n}\int_Y \mathcal{P}^3(f)f d\mu + \dots \\
&+ \frac{2}{n}\int_Y \mathcal{P}^{n-1}(f)f d\mu \\
&= \int_Y f^2 d\mu + 2\sum_{k=1}^{n-1}\frac{n-k}{n}\int_Y \mathcal{P}^k(f)f d\mu
\end{aligned}
$$

And if we take the limit as $n \to \infty$ we get

$$
\lim_{n\to\infty}\int_Y \mathbb{E}_\Omega\left[\left(\frac{S_n^\omega f}{\sqrt{n}}\right)^2\right] d\mu = \lim_{n\to\infty}\int_Y f^2 d\mu + 2\sum_{k=1}^{n-1}\frac{n-k}{n}\int_Y \mathcal{P}^k(f)f d\mu
$$

$$
\begin{aligned}
&= \int_Y f^2 d\mu + 2\int_Y \lim_{n\to\infty} \sum_{k=1}^{n-1} \frac{n-k}{n} \mathcal{P}^k(f) f d\mu \\
&= \int_Y f^2 d\mu + 2\int_Y \lim_{n\to\infty} \left(\sum_{k=1}^{n-1} \mathcal{P}^k(f) - \frac{1}{n} \sum_{k=1}^{n-1} k\mathcal{P}^k(f) \right) f d\mu \\
&= \int_Y f^2 d\mu + 2\int_Y \sum_{k=1}^{\infty} \mathcal{P}^k(f) f d\mu
\end{aligned}
$$

since $\sum_{k=1}^{\infty} \mathcal{P}^k(f) = \sum_{k=1}^{\infty} \mathcal{R}^k(f)$ and $\sum_{k=1}^{\infty} k\mathcal{P}^k(f) = \sum_{k=1}^{\infty} k\mathcal{R}^k(f)$ with $\rho(\mathcal{R}) < 1$ are finite. This gives the Equation 7.3.

Note that $\mathcal{S}(z) = 1/(zI - \mathcal{P}) = \mathcal{Q}/(z-1) + \sum_{n=0}^{\infty} \mathcal{R}^n/z^n$ and we have $\mathcal{Q}(f) = 0$ so $\mathcal{S}(z)(f) = (zI - \mathcal{P})^{-1}(f) = \sum_{n=0}^{\infty} \mathcal{R}^n/z^n(f)$ and if we evaluate at $z = 1$ we get $(I - \mathcal{P})^{-1}(f) = \sum_{n=0}^{\infty} \mathcal{R}^n(f) < \infty$ since $\rho(\mathcal{R}) < 1$. Therefore, $g = (I - \mathcal{P})^{-1} f$ is well defined and

$$
\begin{aligned}
\sigma^2 &= \lim_{n\to\infty} \int_Y \mathbb{E}_\Omega \left[\left(\frac{S_n^\omega f}{\sqrt{n}} \right)^2 \right] d\mu = \sum_{k=-\infty}^{\infty} \int_Y \mathcal{P}^{|k|}(f) f d\mu \\
&= 2\int_Y \left(\sum_{k=0}^{\infty} \mathcal{P}^k(f) f d\mu \right) - \int_Y f^2 d\mu \\
&= 2\int_Y (I - \mathcal{P})^{-1}(f) \cdot f d\mu - \int_Y f^2 d\mu \\
&= \int_Y 2gf - f^2 d\mu \\
&= \int_Y (2g - f) f d\mu \\
&= \int_Y (g + \mathcal{P}g)(g - \mathcal{P}g) d\mu \\
&= \int_Y g^2 - (\mathcal{P}g)^2 d\mu \\
&= \int_Y \mathcal{P}(g^2) - (\mathcal{P}g)^2 d\mu.
\end{aligned}
$$

since $g = (I - \mathcal{P})^{-1}(f) = \sum_{k=0}^{\infty} \mathcal{P}^k(f)$, so we have $g + \mathcal{P}g = \sum_{k=0}^{\infty} \mathcal{P}^k(f) + \sum_{k=1}^{\infty} \mathcal{P}^k(f) = 2\sum_{k=0}^{\infty} \mathcal{P}^k(f) - f = 2g - f$ and $g - \mathcal{P}g = \sum_{k=0}^{\infty} \mathcal{P}^k(f) - \sum_{k=1}^{\infty} \mathcal{P}^k(f) =$

f with the fact that

$$\begin{aligned}
\int_Y g^2 d\mu &= \int_Y g^2 h d\mathfrak{m} \\
&= \int_Y g^2 h \sum_{i=1}^{N} p_i 1 \circ T_i d\mathfrak{m} \\
&= \int_Y \mathcal{P}_{\mathfrak{T}}(g^2 h) d\mathfrak{m} \\
&= \int_Y \mathcal{P}(g^2) d\mu.
\end{aligned}$$

Note that $\mathcal{P}(g^2) - (\mathcal{P}g)^2 \geq 0$ since we have for every $i = 1, \ldots, N$

$$\begin{aligned}
(\mathcal{P}_{T_i}(gh))^2 &= \left(\sum_{y:T_i y = x} \frac{gh(y)}{|T_i'(y)|} \right)^2 \\
&= \left(\sum_{y:T_i y = x} \frac{(h(y))^{1/2}}{|T_i'(y)|^{1/2}} \cdot \frac{(h(y))^{1/2} g(y)}{|T_i'(y)|^{1/2}} \right)^2 \\
&\leq \left(\sum_{y:T_i y = x} \frac{h(y)}{|T_i'(y)|} \right) \left(\sum_{y:T_i y = x} \frac{hg^2(y)}{|T_i'(y)|} \right) \text{ by Cauchy's inequality,} \\
&= \mathcal{P}_{T_i}(h) \mathcal{P}_{T_i}(hg^2) x,
\end{aligned}$$

then for $i \neq j$ we get

$$(\mathcal{P}_{T_i}(gh))^2 (\mathcal{P}_{T_j}(gh))^2 \leq \mathcal{P}_{T_i}(h) \mathcal{P}_{T_1}(hg^2) \cdot \mathcal{P}_{T_j}(h) \mathcal{P}_{T_j}(hg^2)$$

implying

$$\begin{aligned}
(\mathcal{P}_{T_i}(gh))(\mathcal{P}_{T_j}(gh)) &\leq \sqrt{\mathcal{P}_{T_i}(h) \mathcal{P}_{T_j}(hg^2) \cdot \mathcal{P}_{T_j}(h) \mathcal{P}_{T_i}(hg^2)} \\
&\leq \frac{\mathcal{P}_{T_i}(h) \mathcal{P}_{T_j}(hg^2) + \mathcal{P}_{T_j}(h) \mathcal{P}_{T_i}(hg^2)}{2}.
\end{aligned} \tag{7.6}$$

Therefore we have

$$(\mathcal{P}g)^2 = \frac{1}{h^2} \sum_{i,j=1}^{N} p_i p_j (\mathcal{P}_{T_i}(gh))(\mathcal{P}_{T_j}(gh))$$

$$\leq \frac{1}{h^2}\sum_{i,j=1}^{N} p_i p_j \frac{\mathcal{P}_{T_i}(h)\mathcal{P}_{T_j}(hg^2)+\mathcal{P}_{T_j}(h)\mathcal{P}_{T_i}(g^2)}{2} \text{ by Equation (7.6),}$$
$$= \frac{1}{2h^2}\sum_{j=1}^{N} p_j \left(\sum_{i=1} p_i \mathcal{P}_{T_i}(h)\right)\mathcal{P}_{T_j}(hg^2) + \frac{1}{2h^2}\sum_{i=1}^{N} p_i \left(\sum_{j=1}^{N} p_j \mathcal{P}_{T_j}(h)\right)\mathcal{P}_{T_i}(hg^2)$$
$$= \frac{1}{2h}\sum_{j=1}^{N} p_j \mathcal{P}_{T_j}(hg^2) + \frac{1}{2h}\sum_{i=1}^{N} p_i \mathcal{P}_{T_i}(hg^2)$$
$$= \frac{1}{2h}\mathcal{P}_{\mathfrak{T}}(hg^2) + \frac{1}{2h}\mathcal{P}_{\mathfrak{T}}(hg^2) = \mathcal{P}(g^2)$$

since $\sum_{i=1} p_i \mathcal{P}_{T_i}(h) = \mathcal{P}_{\mathfrak{T}}(h) = h$ □

Lemma 7.2.5.

$$\lim_{n\to\infty}\int_Y \mathcal{P}^n_{\frac{t}{\sqrt{n}},f}(1)d\mu = e^{-t^2\sigma^2/2}$$

Proof. From the proof of Lemma 7.2.2 we have that

$$\int_Y \mathcal{P}_{\frac{t}{n},f}(1)d\mu = \lambda^n(i\frac{t}{n})\int_Y \mathcal{N}_{\frac{t}{n}}(1)d\mu + \int_Y \mathcal{M}^n_{\frac{t}{n}}(1)d\mu$$

where $\lim_{n\to\infty}\left|\int_Y \mathcal{M}^n_{\frac{t}{n}}(1)d\mu\right| = 0$ and

$$\mathcal{N}_{\frac{t}{n}}(1) = \mathcal{Q} + \frac{it}{n}\mathcal{N}^{(1)} - \frac{t^2}{n^2}\mathcal{N}^{(2)} + \frac{t^2}{n^2}\overline{\mathcal{N}}_{\frac{t}{n}}$$

and

$$\lambda(i\frac{t}{n}) = \lambda(0) + \frac{it}{n}\lambda'(0) + \frac{(it)^2}{n^2}\lambda''(0) + \frac{t^2}{n^2}\overline{\lambda}(i\frac{t}{n}).$$

We replace $\frac{t}{n}$ with $\frac{t}{\sqrt{n}}$ in the equations and proceed in a same way. If we take $n\to\infty$ together with the assumption that $\lambda'(0)=0$ we get

$$\lim_{n\to\infty}\int_Y \mathcal{N}_{\frac{t}{\sqrt{n}}}(1)d\mu = \lim_{n\to\infty}\mathcal{Q}d\mu = 1$$

and

$$\lim_{n\to\infty} \lambda^n(i\frac{t}{\sqrt{n}}) = \lambda(0) + \frac{it}{\sqrt{n}}\lambda'(0) + \frac{(it)^2}{n^2}\lambda''(0) + \frac{t^2}{n}\overline{\lambda}(i\frac{t}{\sqrt{n}}) = e^{-\lambda''(0)t^2/2}$$

which is obtained by taking the logarithm of the equation above so we get

$$\begin{aligned}
\lim_{n\to\infty} n\ln(\lambda(i\frac{t}{\sqrt{n}})) &= \lim_{n\to\infty} n\ln(1-\lambda''(0)\frac{t^2}{2n}) \\
&= \lim_{n\to\infty} \frac{\ln(1-\lambda''(0)\frac{t^2}{2n})}{1/n} \\
&= \lim_{n\to\infty} \frac{(\lambda''(0)t^2/2n^2)\,/\,(1-\lambda''(0)t^2/2n)}{-1/n^2} \quad \text{by L'hopitals rule,} \\
&= -\lambda''(0)t^2/2
\end{aligned}$$

since $\lim_{n\to\infty} \overline{\lambda}(i\frac{t}{\sqrt{n}}) = 0$. Then the result follows from

$$\int_Y \mathcal{P}_{\frac{t}{\sqrt{n}},f}(1)d\mu = \lambda^n(i\frac{t}{\sqrt{n}})\int_Y \mathcal{N}_{\frac{t}{\sqrt{n}}}(1)d\mu + \int_Y \mathcal{M}^n_{\frac{t}{\sqrt{n}}}(1)d\mu$$

$\square$

Lemma 7.2.6. *$\sigma^2 > 0$ if and only if f is not of the form $f = k + \varphi\circ\mathfrak{T} - \varphi$ for some function φ and number k for the random map $\mathfrak{T}$.*

Proof. We have $\mu(f) = k$ and we also assume that $\mu(f) = 0$ so we have $k = 0$. If $\sigma^2 = 0$ then in the proof of Lemma 7.2.4 the inequality $(\mathcal{P}g)^2 \le \mathcal{P}(g^2)$ is an equality. So we have

$$\begin{aligned}
\frac{1}{h^2}\sum_{i,j=1}^{N} p_ip_j(\mathcal{P}_{T_i}(gh))(\mathcal{P}_{T_j}(gh)) &= \frac{1}{2h^2}\sum_{i,j=1}^{N} p_ip_j\left(\mathcal{P}_{T_i}(h)\mathcal{P}_{T_j}(hg^2) + \mathcal{P}_{T_j}(h)\mathcal{P}_{T_i}(hg^2)\right) \\
&= \frac{1}{h^2}\sum_{i,j=1}^{N} p_ip_j\mathcal{P}_{T_i}(h)\mathcal{P}_{T_j}(hg^2),
\end{aligned}$$

in other words

$$\left(\sum_{i=1}^{N} p_i \sum_{y:T_iy=x} \frac{(h(y))^{1/2}}{|T_i'(y)|^{1/2}}\frac{(h(y))^{1/2}g(y)}{|T_i'(y)|^{1/2}}\right)^2 = \left(\sum_{i=1}^{N} p_i \sum_{y:T_iy=x} \frac{h(y)}{|T_i'(y)|}\right)\left(\sum_{i=1}^{N} p_i \sum_{y:T_iy=x} \frac{hg^2(y)}{|T_i'(y)|}\right)$$

implying that for every $i = 1, \dots, N$ and for every $y \in Y$ with $T_i(y) = x$

$$\frac{c(h(y))^{1/2}}{|T_i'(y)|^{1/2}} = \frac{(h(y))^{1/2} g(y)}{|T_i'(y)|^{1/2}}$$

for some constant $c \in \mathbb{R}$ by Cauchy's inequality given in Proposition A.0.1 in an averaged version, see Example A.0.2. That means for some fixed $x \in Y$ for every $i = 1, \dots, N$ and for every y with $T_i(y) = x$ the value $g(y)$ is constant and does not depend on i and y. Therefore,

$$\begin{aligned}
\mathcal{P}g(x) &= \frac{\mathcal{P}_{\mathfrak{T}}(gh)}{h}(x) \\
&= \frac{1}{h(x)} \sum_{i=1}^{N} p_i \sum_{y:T_i y = x} \frac{gh(y)}{|T_i'(y)|} \\
&= \frac{g(y_i)}{h(x)} \sum_{i=1}^{N} p_i \sum_{y:T_i y = x} \frac{h(y)}{|T_i'(y)|} \\
&= \frac{g(y_i)}{h(x)} \mathcal{P}_{\mathfrak{T}} h(x) = g(y_i)
\end{aligned}$$

for some $i \in \{1, \dots, N\}$ and for some $y_i \in Y$ with $T_i(y_i) = x$ since $\mathcal{P}_{\mathfrak{T}} h = h$. And since $f = g - \mathcal{P}g$ we get

$$\begin{aligned}
f(x) &= g(x) - (\mathcal{P}g)x = g(x) - g(y_i) \\
f(T_i(y_i)) &= g(T_i(y_i)) - g(y_i) \\
f(y_i) &= g(T_i(y_i)) - f(T_i(y_i)) - g(y_i) + f(y_i), \quad \text{by adding } f(y_i) \\
f(y_i) &= (g - f) \circ T_i(y_i) - (g - f)(y_i) \\
&= \varphi \circ T_i(y_i) - \varphi(y_i).
\end{aligned}$$

Since the choice of the map is not important we write

$$f(y) = \varphi(\mathfrak{T}(y)) - \varphi(y)$$

where $\mathfrak{T}$ is the random map. $\square$

Proposition 7.2.7. *If we have* $f = \chi_A$ *for some borel subset* A *of* Y *with* $0 < \mu(A) < 1$ *then* $\sigma^2 > 0$ *.*

Proof. Assume for a contradiction that $\sigma^2 = 0$. Then $\chi_A \circ \mathfrak{T}(x) = \mu(A) + \varphi \circ \mathfrak{T}(x) - \varphi(x)$ for some function φ, implying

$$e^{2\pi i \chi_A \circ \mathfrak{T}(x)} = e^{2\pi i (\mu(A) + \varphi \circ \mathfrak{T}(x) - \varphi(x))}$$

The left-hand side can have 0 or 2π as its exponent that is independent from the choice of $\mathfrak{T}$ so it is 1 implying

$$e^{2\pi i \varphi \circ \mathfrak{T}(x)} = e^{-2\pi i \mu(A)} e^{2\pi i \varphi(x)}.$$

We can consider $\mathfrak{T}(x)$ as $\mathscr{F}(\omega, x) = (\boldsymbol{\sigma}(\omega), T_{\omega_1}(x))$ since φ only depends on the second coordinate. Here $\mathscr{F}$ is the corresponding skew product realization of the random dynamical system which is mixing by Corollary 6.2.4. So for $G(\omega, x) = e^{2\pi i \varphi(\pi_2(\omega, x))}$ where π_2 is the projection function to the second coordinate we have $G \circ \mathscr{F}^n = \lambda G$ with $\lambda = e^{-2\pi i \mu(A)}$ implying that G is constant since $\mathscr{F}$ is mixing, so $\lambda = e^{-2\pi i \mu(A)} = 1$ implying $\mu(A)$ is 0 or 1 which contradicts the assumption. □

7.3 Speed of Convergence

One of the advantages of using Perturbation Theory is that it allows us to calculate the speed of convergence of the limit in the Central Limit Theorem. In probability theory the result for independent identically distributed random variables is known as the Berry-Essen's Theorem. The proof depends on the Essen's inequality given in Theorem A.0.3.

Theorem 7.3.1. *For the random dynamical system $(\mathfrak{T}, \mu)$ defined before there exists a constant $C > 0$ such that for every $v \in \mathbb{R}$ we have*

$$\left|\mu \times \pi \left\{(x, \omega) : \frac{S_n^\omega(f)x - n\mu(f)}{\sigma\sqrt{n}}\right\} - \frac{1}{2\pi}\int_{-\infty}^{v} e^{-u^2/2} du\right| \leq \frac{C}{\sqrt{n}}$$

By the proof of Essen's inequality A.0.3 for every $U > 0$ and for every $n \geq 1$ we have

$$\sup_{v \in \mathbb{R}} \left|\mu \times \pi \left\{\frac{S_n^\omega(f) - n\mu(f)}{\sigma\sqrt{n}}\right\} - \frac{1}{2\pi}\int_{-\infty}^{v} e^{-u^2/2} du\right|$$

$$\leq \frac{K}{U} + \frac{1}{\pi}\int_{-U}^{U} \frac{1}{|u|}\left|\int_\Omega \int_Y e^{iu\frac{S_n^\omega(f)}{\sigma\sqrt{n}}} - e^{-u^2/2} d\mu d\pi\right| du.$$

To prove Theorem 7.3.1 first we need to estimate the following term

$$\left|\int_\Omega \int_Y e^{iu\frac{S_n^\omega(f)}{\sigma\sqrt{n}}} - e^{-u^2/2} d\mu d\pi\right|$$

which is given by the lemma below.

The lemma below is the corresponding version of Lemma 1 in [Pet85], page 109 for the process $\dfrac{S_n^\omega(f)}{\sigma\sqrt{n}}$.

Lemma 7.3.2. *There exists a real number $a > 0$ such that for every $|u| < a\sqrt{n}$ we have*

$$\left|\int_\Omega \int_Y e^{iu\frac{S_n^\omega(f)}{\sigma\sqrt{n}}} - e^{-u^2/2} d\mu d\pi\right| \leq e^{-u^2/4}(2A\frac{|u|^3}{\sigma^3\sqrt{n}} + B\frac{|u|}{\sigma\sqrt{n}}) + (C\frac{|u|}{\sigma\sqrt{n}})\rho_2^n$$

where $\rho_2 = (1 + 2\rho(\mathcal{R}))/3$, for some positive constants A, B and C.

Proof. Clearly by Lemma 7.1.2 we have

$$\begin{aligned}\left|\int_\Omega \int_Y e^{iu\frac{S_n^\omega(f)}{\sigma\sqrt{n}}} - e^{-u^2/2} d\mu d\pi\right| &= \left|\int_\Omega \int_Y \mathcal{P}^n_{\frac{u\sigma}{\sqrt{n}}, f}(1) - e^{-u^2/2} d\mu d\pi\right| \\ &\leq \int_\Omega \int_Y \left|\mathcal{P}^n_{\frac{u\sigma}{\sqrt{n}}, f}(1) - e^{-u^2/2}\right| d\mu d\pi.\end{aligned}$$

Again by using Proposition 7.1.4 with $\theta = \dfrac{u\sigma}{\sqrt{n}}$ we get

$$\begin{aligned}
\mathcal{P}^n_{f,\theta} &= \lambda^n(i\theta)\mathcal{N}_\theta + \mathcal{M}^n_\theta \\
&= \Big(1 + i\theta\lambda'(0) - \tfrac{\theta^2}{2}\lambda''(0) - \tfrac{i\theta^3}{6}\lambda^{(3)} + \theta^3\overline{\lambda}(i\theta)\Big)^n \\
&\cdot \Big(\mu + i\theta\mathcal{N}^{(1)} - \tfrac{\theta^2}{2}\mathcal{N}^{(2)} + \theta^2\overline{\mathcal{N}}_{f,\theta}\Big) + \mathcal{M}^n_\theta \\
&= e^{n(-\dfrac{\theta^2}{2}\sigma^2 + iA_1\theta^3 + \theta^3\varepsilon(\theta))}\left(\mu + i\theta\mathcal{N}^{(1)} - \frac{\theta^2}{2}\mathcal{N}^{(2)} + \theta^2\overline{\mathcal{N}}_{f,\theta}\right) + \mathcal{M}^n_\theta
\end{aligned}$$

for some constant A_1 and $\lim_{\theta\to 0}\varepsilon(\theta) = 0$. Then if we go back to the integration we want to approximate and replace $\theta = u/\sigma\sqrt{n}$ we get

$$\int_\Omega\int_Y \left|\mathcal{P}^n_{\frac{u\sigma}{\sqrt{n}},f}(1) - e^{-u^2/2}\right| d\mu d\pi \leq A_n(u) + B_n(u) + (C|u|\sigma/\sqrt{n})\rho_2^n$$

where

$$\begin{aligned}
A_n(u) = &\ e^{-u^2/2}\left|e^{iA_1u^3/\sigma^3\sqrt{n} + (u^3/\sigma^3\sqrt{n})\varepsilon(u/\sigma\sqrt{n})} - 1\right|, \\
B_n(u) = &\ e^{-u^2/2}e^{iA_1u^3/\sigma^3\sqrt{n} + (u^3/\sigma^3\sqrt{n})\varepsilon(u/\sigma\sqrt{n})} \\
&\cdot \int_\Omega\int_Y \Big(\frac{|u|}{\sigma\sqrt{n}}\left|i\mathcal{N}^{(1)}(1) - \frac{u}{2\sigma\sqrt{n}}\mathcal{N}^{(2)}(1) + \frac{u}{2\sigma\sqrt{n}}\overline{\mathcal{N}}_{\frac{u}{\sigma\sqrt{n}}}(1)\right| d\mu d\pi.
\end{aligned}$$

For the first term $A_n(u)$ we use the fact that $|e^z - 1| \leq |z|e^{|z|}$ and approximate the corresponding z term that is

$$\left|iA_1u^3/\sigma^3\sqrt{n} + (u^3/\sigma^3\sqrt{n})\varepsilon(u/\sigma\sqrt{n})\right| \leq |u|2Au^2/\sigma^3\sqrt{n}$$

where $A = |A_1|$. We choose the real number $a > 0$ that satisfies $2Aa/\sigma^3 < 1/4$ so for every $|u| < a\sqrt{n}$ we have

$$\left|iA_1u^3/\sigma^3\sqrt{n} + (u^3/\sigma^3\sqrt{n})\varepsilon(u/\sigma\sqrt{n})\right| \leq u^2/4.$$

Also let B be a BV bound for the terms in the intergral of the term $B_n(u)$ that is

$$\left\|i\mathcal{N}^{(1)}(1) - \frac{u}{2\sigma\sqrt{n}}\mathcal{N}^{(2)}(1) + \frac{u}{2\sigma\sqrt{n}}\overline{\mathcal{N}}_{\frac{u}{\sigma\sqrt{n}}}(1)\right\|_{BV} \leq B.$$

With the bounds above the result follows. □

Proof of Theorem 7.3.1 . By using the result of Lemma 7.3.2 with $U = a\sqrt{n}$ we get

$$\sup_{v\in\mathbb{R}} \left| \mu \times \pi \left\{ \frac{S_n^\omega(f) - n\mu(f)}{\sigma\sqrt{n}} \right\} - \frac{1}{2\pi} \int_{-\infty}^{v} e^{-u^2/2} du \right|$$

$$\leq \frac{K}{a\sqrt{n}} + \frac{1}{\sqrt{n}} \int_{-a\sqrt{n}}^{a\sqrt{n}} e^{-u^2/2}(2Au^2 + \frac{B}{\sigma}) + C\frac{\rho_2^n}{\sigma\sqrt{n}} du.$$

□

7.4 More Limit Theorems

The main goal of this section is the quenched Central Limit Theorem. To prove the quenched Central Limit Theorem first we need to prove two more limit theorems. The first one is the Averaged Large Deviation Estimate which is a very standard step if one is working on statistical properties of some dynamical systems. And the next limit theorem gives that the process as in Theorem 7.2.1 not only converges to the normal distribution but it also converges with tight maxima. The necessary definitions are given below.

Theorem 7.4.1. *(Averaged Large Deviation Estimate) Let $S_n^\omega f/n$ be the random process as before. There exists a real number $A > 0$ such that for all $a \in (0, A)$ we have the following estimate*

$$\pi \times \mu \left\{ (\omega, x) : \left| \frac{1}{n} S_n^\omega f(x) \right| \geq a \right\} \leq Ce^{-Ca^2 n} \tag{7.7}$$

Proof. To prove the averaged large deviation result again we use the perturbation theory, specifically the same result we used to prove the averaged Central Limit Theorem which is mainly the Lemma 7.1.2 and Proposition 7.1.4.

For any random variable X with a probability distribution $\mathbb{P}$ for any $\theta > 0$ we have

$$\begin{aligned}\mathbb{P}\{X \geq a\} &\leq \mathbb{E}_{\mathbb{P}}\left[\chi_{\{X\geq a\}}\, e^{\theta(X-a)}\right] \\ &\leq e^{-\theta a}\mathbb{E}_{\mathbb{P}}\left[e^{\theta X}\right]\end{aligned}$$

Since

$$\mathbb{P}\{|X| \geq a\} = \mathbb{P}\{X \geq a\} + \mathbb{P}\{X \leq -a\}$$

it is enough to give the result only for $\mathbb{P}\{X \geq a\}$.

If we consider the random variable X being the random process $S_n^\omega f(x)/n$ for each n, then the inequality becomes

$$\pi \times \mu\left\{\frac{1}{n}S_n^\omega f(x) \geq a\right\} \leq e^{-at/n}\,\mu\left(\mathbb{E}_\Omega\left[e^{tS_n^\omega f(x)/n}\right]\right) \text{ for } \theta = t/n, \text{ for every } t > 0.$$

This is by Lemma 7.1.2 equivalent to the quantity below

$$\pi \times \mu\left\{\frac{1}{n}S_n^\omega f(x) \geq a\right\} \leq e^{-at/n}\,\mu\left\{\mathcal{P}^n_{\frac{t}{n},f}(1)\right\}.$$

Again we can use Proposition 7.1.4 so that

$$\mathcal{P}^n_{\frac{t}{n},f}(1) = \lambda^n(i\frac{t}{n})\mathcal{N}_{\frac{t}{n}}(1) + \mathcal{M}^n_{\frac{t}{n}}(1)$$

and by using the approximations of each term in the proof of Lemma 7.2.2 we can deduce that

$$\mathcal{P}^n_{\frac{t}{n},f}(1) = \lambda^n\left(i\frac{t}{n}\right)\left(1 + \mathscr{O}\left(\frac{|t|}{n}\right)\right) + \mathscr{O}\left(\rho_2^n\right).$$

We define $\mathscr{A}(a) = \sup_{|\theta|\le C} a\theta - \ln(\lambda(i\theta))$ for some $C > 0$. Let θ_0 be the value at which $\mathscr{A}$ attains its maximum so we have

$$\ln(\lambda(i\theta_0)) = a\theta_0 - \sup\{\mathscr{A}(a)\} \le a\theta_0 - \mathscr{A}(a).$$

By considering $\lambda^n\left(i\frac{t}{n}\right) = e^{n\ln\left(\lambda\left(i\frac{t}{n}\right)\right)}$ if we choose t to be $n\theta_0$ then we have

$$\begin{aligned}\pi \times \mu\left\{\tfrac{1}{n}S_n^\omega f(x) \ge a\right\} &\le e^{-a(\theta_0 n)/n} e^{n(a\theta_0 - \mathscr{A}(a))}(1 + C|\theta_0|) + \mathscr{O}\left(\rho_2^n\right) \\ &= e^{-n\mathscr{A}(a)}(1 + C|\theta_0|) + \mathscr{O}\left(\rho_2^n\right).\end{aligned}$$

Now we need to analyze $\mathscr{A}(a)$. So by using the Equation 7.5 we have

$$\lambda(i\theta) = 1 + i\theta\lambda'(0) + \frac{(i\theta)^2}{2}\lambda''(0) + \theta^2\overline{\lambda}(i\theta),$$

with $\lambda'(0) = \mu(f) = 0$ because of the assumption on f. This implies

$$a\theta - \ln\left(\lambda\left(i\theta\right)\right) = a\theta - \frac{\theta^2}{2}\lambda''(0) + \mathscr{O}(\theta^3). \tag{7.8}$$

To find the point θ_0 where the maximum is taken we take the derivative and set it equal to zero. So we get $a - \theta\lambda''(0) = 0$ implying $\theta_0 = a/\lambda''(0)$. Note that $\lambda''(0) = \sigma^2$ by Lemma 7.2.4. Thus we have

$$\pi \times \mu\left\{\frac{1}{n}S_n^\omega f(x) \ge a\right\} \le 2e^{-a^2 n(1-\epsilon)/2\sigma^2}$$

for $a \le C\epsilon$ where $C\epsilon$ is small enough. $\square$

First we give the general definition of convergence to a distribution with tight maxima. Then we show that the random dynamical system $S_n^\omega f$ given before converges to normal distribution with tight maxima by showing that the random dynamical system is a reverse martingale difference.

Definition 7.4.2. *Let S_n be a sequence of random variables on a probability space and B_n be a renormalizing sequence, meaning $B_n = B(n)$ as a function $B(x) : \mathbb{R}^+ \to \mathbb{R}^+$ is of the form $B(x) = x^d L(x)$ where $d > 0$ and the function $L : \mathbb{R}^+ \to \mathbb{R}^+$ is L^1, slowly varying that is, for every $\alpha > 0$ $\lim_{x\to\infty} \frac{L(\alpha x)}{L(x)} = 1$ and normalized that is, $L'(x) = (L(x)/x)$. Then we say that $(S_n/B_n, B_n)$ converges with tight maxima to a random variable $\mathscr{X}$ if S_n/B_n converges in law to $\mathscr{X}$ and if the sequence*

$$M_n = \left(\max_{1\le k\le n} |S_k| \right) / B_n$$

is tight, that is for every $\epsilon > 0$ there exists $c > 0$ such that for every $n \ge 1$ we have

$$\mathbb{P}\left\{ \max_{1\le k\le n} \frac{|S_k|}{B_n} > c \right\} \le \epsilon.$$

Now the process S_n in our case is $S_n^\omega f(x)$ on probability space $\Omega \times Y$ with distribution $\pi \times \mu$. The corresponding renormalizing sequence B_n is $\sqrt{n}$. We already know that the process $S_n^\omega f(x)/\sqrt{n}$ converges in distribution to the normal distribution $\mathscr{N}(0, \sigma)$ for $\int f d\mu = 0$. We show below that the convergence is in fact with tight maxima.

Theorem 7.4.3. *(Convergence with Tight Maxima) Let $S_n^\omega f(x)/\sqrt{n}$ be the random process as defined before with constituent maps in $\mathcal{T}_1(Y)$. Then for every $\epsilon > 0$ there exists $c > 0$ such that for every $n \ge 1$ we have*

$$\pi \times \mu \left\{ \max_{1\le k\le n} \frac{|S_k^\omega f(x)|}{\sqrt{n}} > c \right\} \le \epsilon.$$

Note that this property is not only about the process $S_n^\omega f(x)/\sqrt{n}$, the choice of the sequence $\sqrt{n}$ is also very important. Before giving the proof of Theorem

7.4.3 first we give a tight maxima result for reverse martingale differences. Then to prove Theorem 7.4.3 we only need to show that the process is a reverse martingale difference.

Proposition 7.4.4. *Let $Z_0, Z_1, \ldots$ be a sequence of reverse martingale differences. Let $S_n = \sum_{k=0}^{n-1} Z_n$. If $S_n/\sqrt{n}$ converges to normal distribution $\mathscr{N}(0,\sigma)$ and if $S_n/\sqrt{n}$ is L^1-bounded then $S_n/\sqrt{n}$ converges to $\mathscr{N}(0,\sigma)$ with tight maxima.*

Proof. By using the Martingale Maximal Inequality A.0.5 we have for every $\alpha > 0$ and $n \in \mathbb{N}$ that

$$\mathbb{P}\left\{\max_{1\leq k\leq n} |S_k| \geq \alpha\right\} \leq \frac{C}{\alpha}\mathbb{E}[|S_n|]$$

for some constant $C > 0$. By choosing $\alpha = c\sqrt{n}$ we get

$$\mathbb{P}\left\{\max_{1\leq k\leq n} \frac{|S_k|}{\sqrt{n}} \geq \alpha\right\} \leq \frac{C}{c\sqrt{n}}\mathbb{E}[|S_n|] \leq \frac{C_1}{c}$$

since $\dfrac{|S_k|}{\sqrt{n}}$ is bounded in L^1. Thus for every $\epsilon > 0$ choose $c > 0$ to be $\dfrac{C_1}{c} < \epsilon$. $\square$

Proof of Theorem 7.4.3. Let $f \in BV$ with $\int f d\mu = 0$. We consider all maps on BV as maps on the product space $\Omega \times Y$ that only depends on the second coordinate so $f(\omega, x) = f(x)$. Let $\mathcal{U}_{\mathscr{F}}$ be the composition operator with $\mathscr{F}$ and let $\mathcal{P}_{\mathscr{F}}$ be the adjoint operator of $\mathcal{U}_{\mathscr{F}}$. Therefore the process can be given by

$$S_n(f) = \sum_{i=0}^{n-1} f \circ \mathscr{F}^i = \sum_{i=0}^{n-1} \mathcal{U}_{\mathscr{F}}^n(f).$$

To show that the process $S_n/\sqrt{n}$ has a tight maxima we first show that the process defined by

$$Z_i = \mathcal{U}_{\mathscr{F}}^i(f) - \mathcal{U}_{\mathscr{F}}^i(g) + \mathcal{U}_{\mathscr{F}}^{i-1}(g)$$

is a reverse martingale difference, so $\left(\sum_{i=0}^{n-1} Z_i\right)/\sqrt{n}$ has a tight maxima. And then we show that the sequence $(g-\mathcal{U}^n_{\mathscr{F}}(g))/\sqrt{n}$ also has a tight maxima which concludes the result.

Here the function $g=\sum_{n=0}^{\infty}\mathcal{P}_{\mathfrak{T}}(f)$ is as in Lemma 7.2.4 which is convergent by the spectral gap. Also note that we have

$$\mathcal{P}_{\mathscr{F}}(f)=\frac{1}{h}\mathcal{P}_{\mathfrak{T}}(fh)$$

where h is the stationary density so $\sum_{n=0}^{\infty}\mathcal{P}^n_{\mathscr{F}}(f)$ is also convergent. For $i=1$ in the above equation we have

$$Z_1=\mathcal{U}_{\mathscr{F}}(f)-\mathcal{U}_{\mathscr{F}}(g)+g$$

implying

$$\mathcal{P}_{\mathscr{F}}(Z_1)=f-g+\mathcal{P}_{\mathscr{F}}(g)=f-g+\mathcal{P}(g)=0.$$

Thus for the filtration $\mathscr{F}^{-1}(\mathcal{B})$ we have

$$(\pi\times\mu)[Z_1|\mathscr{F}^{-1}(\mathcal{B})] \quad = \quad 0$$

implying that Z_i is a reverse martingale difference. Then by summing the random variable Z_i over all i values, we conclude that the process we get below has a tight maxima.

$$\frac{\sum_{i=0}^{n-1} Z_i}{\sqrt{n}}=\frac{S_n(f)}{\sqrt{n}}+\frac{g-\mathcal{U}^n_{\mathscr{F}}(g)}{\sqrt{n}}.$$

Finally we need to show that $(g-\mathcal{U}^n_{\mathscr{F}}(g))/\sqrt{n}$ has a tight maxima to finish the proof. But this is easy since the function g is bounded so we have

$$\lim_{n\to\infty}\frac{g-\mathcal{U}^n_{\mathscr{F}}(g)}{\sqrt{n}}=0.$$

That concludes the proof. $\square$

We use the fact that the averaged Central Limit Theorem on the induced system also satisfies the averaged Large Deviation result by Theorem 7.4.1 to show the quenched Central Limit Theorem. The proof is based on the paper of Ayyer, Liverani and Stenlund [ALS09] where the constituent functions are toral automorphisms. Here we use the Borel-Cantelli argument they have in their proof for quenched Central Limit Theorem.

Again we have the averaged Central Limit Theorem by Theorem 7.2.1. Let σ^2 be the variance of the process that satisfies the Central Limit Theorem. We show that for π almost every sequence ω we have the CLT with variance σ^2. For that first we show that the previous limit theorems can also be given by using the Lebesgue measure $\mathfrak{m}$.

Theorem 7.4.5. *(Averaged CLT for non-stationary measures) The random dynamical system $(\mathfrak{T}, \mathfrak{m})$ satisfies the averaged CLT for every $f \in C^\infty(Y, \mathbb{R})$ with respect to the normalized Lebesgue measure $\mathfrak{m}$, that is*

$$\lim_{n\to\infty} (\pi \times \mathfrak{m}) \left\{ (\omega, x) : \frac{f_n^\omega(x) - n \int f d\mathfrak{m}}{\sqrt{n}} < c \right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{c} e^{-t^2/2\sigma^2} dt.$$

Proof. We give a lower and an upper bound for the process

$$\pi \times \mathfrak{m} \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_n^\omega f(x) \in J \right\}$$

for some $J \in \mathbb{R}$. Let $\epsilon > 0$ be small. We know that the process satisfies the Central Limit Theorem with respect to μ so there exists J_0, J_1 with $J_0 \subset J \subset J_1$ such that for n large enough say $n \geq N_1$ we have

$$\pi \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_n^\omega f(x) \in J_0 \right\} > \frac{1}{\sqrt{2\pi\sigma}} \int_{J_0} e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}$$

and similarly

$$\pi \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_n^\omega f(x) \in J_1 \right\} < \frac{1}{\sqrt{2\pi\sigma}} \int_{J_1} e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}$$

And if we take both integrals over the set J as given above we get

$$\pi \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_n^\omega f(x) \in J_0 \right\} > \frac{1}{\sqrt{2\pi\sigma}} \int_{J} e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}$$

and similarly

$$\pi \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_n^\omega f(x) \in J_1 \right\} < \frac{1}{\sqrt{2\pi\sigma}} \int_{J} e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}.$$

We have an expanding system so we want to find a suitable number of iterates say M that expands the process from J_0 to J or from J to J_1 and the probability difference caused by that many number of iterates is less than $\frac{\epsilon}{2}$. For that first choose $M_1 \in \mathbb{N}$ so that for every $n > M_1$ we have

$$\int |\mathcal{P}_{\mathfrak{X}}^n(1) - h| d\mathfrak{m}| < \frac{\epsilon}{2} \tag{7.9}$$

where h is the stationary density for μ.

Then if we choose $M > N_1, M_1$ large enough so that

$$\frac{1}{\sqrt{M}} \sum_{i=M_1}^{M_1+M-1} S_i^\omega f(x) \in J_0 \text{ implies } \frac{1}{\sqrt{M+M_1}} \sum_{i=0}^{M_1+M-1} S_i^\omega f(x) \in J, \qquad \text{and}$$

$$\frac{1}{\sqrt{M+M_1}} \sum_{i=0}^{M_1+M-1} S_i^\omega f(x) \in J \text{ implies } \frac{1}{\sqrt{M}} \sum_{i=M_2}^{M_1+M-1} S_i^\omega f(x) \in J_1$$

then we have for every $n > M$ that

$$\pi_{[\omega]_{n+M_1}} \times \mathfrak{m} \left\{ \frac{1}{\sqrt{n+M_1}} \sum_{i=0}^{M_1+n-1} S_i^\omega f(x) \in J \right\}$$

$$
\begin{aligned}
&\geq \pi_{[\omega]_n} \times \mathfrak{m}^{(M_1)} \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_i^{\omega} f(x) \in J_0 \right\} \\
&\geq \pi_{[\omega]_n} \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_i^{\omega} f(x) \in J_0 \right\} + \frac{\epsilon}{2} \text{ by Equation (7.9),} \\
&> \frac{1}{\sqrt{2\pi}\sigma} \int_J e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}, \qquad \text{and}
\end{aligned}
$$

$$
\begin{aligned}
&\pi_{[\omega]_{n+M_1}} \times \mathfrak{m} \left\{ \frac{1}{\sqrt{n+M_1}} \sum_{i=0}^{M_1+n-1} S_i^{\omega} f(x) \in J \right\} \\
&\leq \pi_{[\omega]_n} \times \mathfrak{m}^{(n)} \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_i^{\omega} f(x) \in J_1 \right\} \\
&\leq \pi_{[\omega]_n} \times \mu \left\{ \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} S_i^{\omega} f(x) \in J_1 \right\} + \frac{\epsilon}{2} \text{ by Equation (7.9),} \\
&< \frac{1}{\sqrt{2\pi}\sigma} \int_J e^{-t^2/2\sigma} dt - \frac{\epsilon}{2}
\end{aligned}
$$

where $\mathfrak{M}^{(M_1)}$ is the measure with density $\mathcal{P}_{\mathfrak{T}}^{M_1}(1)$ and $\pi_{[\omega]_n}$ denotes that π only changes with n-cylinders. And that concludes the proof. $\square$

Large Deviations and speed of the convergence results can also be given in $\mathfrak{m}$ with the same argument.

Theorem 7.4.6. *(Quenched CLT) The random dynamical system $(\mathfrak{T}, \mathfrak{m})$ with maps in $\mathcal{T}_0$ satisfies the quenched Central Limit Theorem for every $f \in C^{\infty}(Y, \mathbb{R})$, that is for π-almost every sequence ω*

$$
\lim_{n\to\infty} \mathfrak{m} \left\{ x \in Y : \frac{S_n^{\omega} f(x) - n \int f d\mu}{\sqrt{n}} < c \right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{c} e^{-t^2/2\sigma^2} dt.
$$

if the corresponding random dynamical system on a torus satisfies the averaged Central Limit Theorem.

Proof. To simplify the proof assume that $\int f d\mathfrak{m} = 0$ since we can generalize the result by plugging in $f = f_1 - \int f_1 d\mathfrak{m}$ for a general function f_1. We know that the random dynamical system $(\mathfrak{T}, \mathfrak{m})$ satisfies the averaged Central Limit Theorem for every $f \in C^\infty(Y, \mathbb{R})$ with $\int f d\mathfrak{m} = 0$, so the mean of the process $\mathfrak{m}\left(\frac{S_n^\omega f(x)}{\sqrt{n}}\right)$ with respect to the Bernoulli distribution π converges to the normal distribution $\mathscr{N}(0, \sigma^2)$, so the characteristic functions:

$$\lim_{n\to\infty} \int_{\omega\in\Omega^{\mathbb{N}}} \mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right) d\pi = e^{-\frac{1}{2}t^2\sigma^2}.$$

To simplify the notation let $Z_n = \mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right)$ and Z_σ be the characteristic function of the centered normal distribution with variance σ^2. If $\lim\limits_{n\to\infty} \mathbb{E}_\pi[Z_n] = Z_\sigma$ where $\mathbb{E}_\Omega$ is the expectation with respect to the Bernoulli measure π on the random sequences in Ω, then we can compute an L^2 estimate:

$$\begin{aligned}
\mathbb{E}_\Omega(|Z_n - Z_\sigma|^2) &= \mathbb{E}_\Omega(|Z_n|^2 + Z_\sigma^2 - 2Z_\sigma|Z_n|) \\
&= \mathbb{E}_\Omega(|Z_n|^2 - Z_\sigma^2 + 2Z_\sigma^2 - 2Z_\sigma|Z_n|) \\
&= \mathbb{E}_\Omega(|Z_n|^2 - Z_\sigma^2 + 2Z_\sigma(Z_\sigma - |Z_n|)) \\
&= \mathbb{E}_\Omega(|Z_n|^2) - Z_\sigma^2 + 2Z_\sigma(Z_\sigma - \mathbb{E}_\Omega(|Z_n|))
\end{aligned}$$

Now, we give a bound for the right-hand side, first for $Z_\sigma - \mathbb{E}_\pi(|Z_n|)$ and second for $\mathbb{E}_\Omega(|Z_n|^2) - Z_\sigma^2$.

First bound: For $Z_\sigma - \mathbb{E}_\Omega(|Z_n|)$ we already know that the process $|Z_n|$ satisfies the averaged Central Limit Theorem with a speed given in Theorem 7.3.1. The Lemma 7.3.2 used in the proof gives the order of the convergence of the characteristic functions. Thus we get

$$Z_\sigma - \mathbb{E}_\pi(|Z_n|) = \mathscr{O}\left(\frac{1+|t|^3}{\sqrt{n}}\right)$$

Second bound: Since Y is a finite union of intervals in $\mathbb{R}$ we can consider it as $[0,1]$ for now to easily calculate the numbers. Consider $|Z_n|^2$ as a random variable itself of a Cartesian product system since

$$\begin{aligned}
|Z_n|^2 &= \left[\mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right)\right]^2 \\
&= \left(\int_0^1 e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}dx\right)\left(\int_0^1 e^{it\frac{S_n^\omega f(y)}{\sqrt{n}}}dy\right) \\
&= \int_0^1\int_0^1 e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}e^{it\frac{S_n^\omega f(y)}{\sqrt{n}}}dxdy \\
&= \int_0^1\int_0^1 e^{it\frac{S_n^\omega f(x)+S_n^\omega f(y)}{\sqrt{n}}}dxdy
\end{aligned}$$

where

$$S_n^\omega f(x)+S_n^\omega f(y) = f(x)+f(y)+f\circ T_1(x)+f\circ T_1(y)\ldots f\circ T_{n-1}\circ T_1(x)+f\circ T_{n-1}\circ T_1(y)$$

so can be considered as

$$F_n^\omega(x,y) = F(x,y)+F(T_1(x),T_1(y))+\ldots+F(T_{n-1}\circ\ldots T_1(x),T_{n-1}\circ\ldots T_1(y))$$

where $F(x,y)=f(x)+f(y)$. Furthermore, we can define $(T_1(x),T_1(y))$ by $(T_1\times T_1)(x,y)$, and $(T_2\circ T_1(x),T_2\circ T_1(y))$ by $(T_2\times T_2)\circ(T_1\times T_1)(x,y)$ and so on.

Now we iterate the space $[0,1]\times[0,1]$ with maps $T_i\times T_i$ which can be considered as maps on a torus by $A_i=\begin{pmatrix} T_i & 0 \\ 0 & T_i\end{pmatrix}$. Therefore, $|Z_n|^2$ is the characteristic function of the process $S_n^\omega F(x,y)$ on a torus. Such maps are defined in Chapter 4 and we discuss in Theorem 4.2.3 that they satisfy the averaged Central Limit Theorem, too. Since we have not gone over each step to give the proof we prefer to give it as a claim in the theorem. So by the assumption given in the statement of the theorem we have

$$\lim_{n\to\infty}\mathbb{E}_\Omega(|Z_n|^2) = \lim_{n\to\infty}\mathbb{E}_\Omega\left[\mathfrak{m}_2\left(e^{it\frac{S_n^\omega F(x,y)}{\sqrt{n}}}\right)\right] = e^{-\frac{1}{2}t^2\sigma_{\mathbb{T}}^2}$$

where $\mathfrak{m}_2$ is the Lebesgue measure on a torus and $\sigma^2_{\mathbb{T}}$ is the variation of the normal distribution that the process on a torus converges. Now we need to calculate how it is related to the variance of the process Z_n. We know again from the straightforward computation in Theorem 7.2.1 that

$$\sigma^2 = \mathfrak{m}(f^2) + 2\sum_{n=1}^{\infty} \mathfrak{m}(f\mathcal{P}^n f).$$

Similarly on a torus we have

$$\begin{aligned}
\sigma^2_{\mathbb{T}} &= \mathfrak{m}_2(F^2) + 2\sum_{n=1}^{\infty} \mathfrak{m}_2(F\mathcal{P}^n_{\mathbb{T}} F) \\
&= \int_0^1\int_0^1 F(x,y)^2 dxdy + 2\sum_{n=1}^{\infty}\int_0^1\int_0^1 F(x,y)\mathcal{P}^n_{\mathbb{T}}(F(x,y))dxdy \\
&= \int_0^1\int_0^1 (f(y)+f(x))^2 dxdy \\
&+ 2\sum_{n=1}^{\infty}\int_0^1\int_0^1 \sum_{\omega}(F)\circ T_{\omega_n}\circ\ldots T_{\omega_1}(x,y)F(x,y)dxdy \\
&= \int_0^1\int_0^1 (f^2(y)+2f(y)f(x)+f^2(x))dxdy \\
&\quad +2\sum_{n=1}^{\infty}\int_0^1\int_0^1 (f\circ T_n\circ\ldots T_1(y)+f\circ T_n\circ\ldots T_1(x))(f(y)+f(x))dxdy \\
&= \int_0^1 f^2(y)dy + 0 + \int_0^1 f^2(x)dx \\
&\quad +2\sum_{n=1}^{\infty}\int_0^1 f\circ T_n\circ\ldots T_1(y)f(y)dy + \int_0^1 f\circ T_n\circ\ldots T_1(x)f(x)dx \\
&\quad \text{since } \int_0^1\int_0^1 f\circ T_n\circ\ldots T_1(x)f(y)dxdy \\
&\quad = \int_0^1 f\circ T_n\circ\ldots T_1(x)\int_0^1 f(y)dydx = 0 \text{ by assumption that } \mathfrak{m}(f)=0, \\
&= 2\mathfrak{m}(f^2) + 4\sum_{n=1}^{\infty}\mathfrak{m}(f\circ T_n\circ\ldots T_1\cdot f) \\
&= 2\left(\mathfrak{m}(f^2) + 2\sum_{n=1}^{\infty}\mathfrak{m}(f\mathcal{P}^n f)\right) \\
&= 2\sigma^2.
\end{aligned}$$

where σ^2 is the variance of the random process on $[0,1]$. Therefore we have the first term to be

$$\begin{aligned}
\mathbb{E}_\Omega(|Z_n|^2) - Z_\sigma^2 &= \mathbb{E}_\Omega\left[\mathfrak{m}_2\left(\exp(it\frac{F_n^\omega(x,y)}{\sqrt{n}})\right)\right] - \left[\exp(-\frac{1}{2}t^2\sigma^2)\right]^2 \\
&= \mathbb{E}_\Omega\left[\mathfrak{m}_2\left(\exp(it\frac{F_n^\omega(x,y)}{\sqrt{n}})\right)\right] - \exp(-\frac{1}{2}t^2(2\sigma^2)) \\
&= \mathbb{E}_\Omega\left[\mathfrak{m}_2\left(\exp(it\frac{F_n^\omega(x,y)}{\sqrt{n}})\right)\right] - \exp(-\frac{1}{2}t^2\sigma_{\mathbb{T}}^2)
\end{aligned}$$

which converges to zero for each t as $n \to \infty$ with again a speed of $\mathscr{O}\left(\frac{1+|t|^3}{\sqrt{n}}\right)$ by Theorem 7.3.1. Therefore we have

$$\mathbb{E}_\Omega\left[\left|\mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right) - e^{-\frac{1}{2}t^2\sigma^2}\right|^2\right] \leq C\frac{1+|t|^3}{\sqrt{n}}$$

which implies by Chebyshev's Inequality A.0.4 that

$$\pi\left\{\left|\mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right) - e^{-\frac{1}{2}t^2\sigma^2}\right| \geq \epsilon\right\} \leq C\epsilon^{-2}\frac{1+|t|^3}{\sqrt{n}}. \tag{7.10}$$

The infinite sum of the probabilities on the left-hand side of the Equation (7.10) is bounded by $\sum_{i=1}^{\infty} C\epsilon^{-2}\frac{1+|t|^3}{\sqrt{n}}$ which is not finite and also depends on $t \in \mathbb{R}$. If it was finite and existed for every $t \in \mathbb{R}$ then we could apply the Borel-Cantelli Lemma A.0.6 to deduce that

$$\pi\left\{\lim_{n\to\infty}\left|\mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right) - e^{-\frac{1}{2}t^2\sigma^2}\right| \geq \epsilon\right\} = 0$$

which means the sequences for which we do not have the Central Limit Theorem has zero π-measure which would conclude the quenched Central Limit Theorem.

Let $E_n(t)$ denote the sequence of events

$$E_n(t) = \left\{\lim_{n\to\infty}\left|\mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}}\right) - e^{-\frac{1}{2}t^2\sigma^2}\right| \geq \epsilon\right\}$$

Now we define subsequences of events $E_n(t)$ say $E_{n_k}(t_k)$ and first show that Borel-Cantelli argument can be applied to each subsequence and if we sum over these subsequences to get back the original sequence we show that the sum is finite which finishes the proof.

The subsequences are defined the following way: Define the set $J_k \subset \mathbb{N} \times \mathbb{R}$ to be

$$J_k = \{(n,t) : 2^k \leq n \leq 2^{k+1}, |t| \leq k\}$$

Then $E_{n_k}(t_k)$ is equal to the terms of $E_n(t)$ where (n,t) belongs to J_k. Then we have

$$\sum_{n=1}^{\infty} \pi(E_n(t)) = \sum_{k=1}^{\infty} \sum_{(n,t)\in J_k} \pi(E_n(t)) \leq \sum_{k=1}^{\infty} \sup_{(n,t)\in J_k} \pi(E_n(t)).$$

Thus we reduced the problem to showing that if the terms

$$\sup_{(n,t)\in J_k} \pi(E_n(t)) = \pi\left\{ \sup_{(n,t)\in J_k} \left\{ \left| \mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}} \right) - e^{-\frac{1}{2}t^2\sigma^2} \right| \geq \epsilon \right\} \right\} \tag{7.11}$$

are summable over k. The rest of the proof looks for a suitable upper bound for the term in Equation (7.11).

First we define new sets that cover the J_ks. We choose reference points first then define smaller sets around these reference points so that the union of these small sets cover J_k. Now fix the J_k so we have points (n,t) with $2^k \leq n \leq 2^{k+1}$ and $|t| \leq k$. Denote these points by $A_k = \{2^k, 2^k+1, \ldots, 2^{k+1}\}$ and $B_k = [-k,k]$. We divide A_k and B_k into smaller sets as follows. First we choose reference points to be

$$R(A_k) = \{2^k, 2^k + [2^{3k/4}], 2^k + [2 \cdot 2^{3k/4}], 2^k + [3 \cdot 2^{3k/4}], \ldots, 2^k + [2^{k/4}2^{3k/4}] = 2^{k+1}\}$$

Similarly we choose reference points for B_k that are finitely many which are

$$R(B_k) = \{-k, -k + \frac{1}{k}, -k + \frac{2}{k}, \ldots, -k + \frac{k^2}{k} = k\}$$

Now according to these reference points we can define the small sets that each lie in A_k and B_k accordingly. For each fixed $n \in R(A_k)$ and $t \in R(B_k)$ we define

$$A_k(n) = \{m \in \mathbb{N} : |n - m| \leq 2^{k/4} + 1\}$$

and

$$B_k(t) = \{u \in \mathbb{R} : |u - t| \leq 1/k\}.$$

We can consider the reference points as the centers of the small sets we define. If we can consider all the sets with the given centers we cover all possibilities for n and t. Here we use that for every n if $|t| < \log_2 n$ then $(n, t) \in J_k$ for $k = \lfloor \log_2 n \rfloor$ to define the reference points. As one can notice the union of these small sets contains more than the original sets so we have for $R(J_k) = R(A_k) \times R(B_k)$ and $J_k(n, t) = A_k(n) \times B_k(t)$ that

$$J_k = A_k \times B_k \subset \bigcup_{(n,t)\in R(J_k)} J_k(n, t).$$

Therefore the quantity in Equation 7.11 is less then

$$\sum_{(n,t)\in R(J_k)} \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left\{ \left| \mathfrak{m}\left(e^{it_0 \frac{S^{\omega}_{n_0} f(x)}{\sqrt{n_0}}} \right) - e^{-\frac{1}{2}t_0^2\sigma^2} \right| \geq \epsilon \right\} \right\}$$

Now we work on each term inside the sum and apply triangle inequality to the terms obtained by adding and subtracting the terms that have n instead of n_0 and t instead of t_0 and combine them suitably so we get

$$\begin{aligned} & \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left| \mathfrak{m}\left(e^{it_0 \frac{S^{\omega}_{n_0} f(x)}{\sqrt{n_0}}} \right) - e^{-\frac{1}{2}t_0^2\sigma^2} \right| \geq \epsilon \right\} \\ \leq\ & \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left| \mathfrak{m}\left(e^{it_0 \frac{S^{\omega}_{n_0} f(x)}{\sqrt{n_0}}} \right) - \mathfrak{m}\left(e^{it_0 \frac{S^{\omega}_{n} f(x)}{\sqrt{n}}} \right) \right| \geq \epsilon/4 \right\} \\ +\ & \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left| \mathfrak{m}\left(e^{it_0 \frac{S^{\omega}_{n} f(x)}{\sqrt{n}}} \right) - \mathfrak{m}\left(e^{it \frac{S^{\omega}_{n} f(x)}{\sqrt{n}}} \right) \right| \geq \epsilon/4 \right\} \end{aligned}$$

$$+ \quad \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left| \mathfrak{m}\left(e^{it\frac{S_n^\omega f(x)}{\sqrt{n}}} \right) - e^{-\frac{1}{2}t^2\sigma^2} \right| \geq \epsilon/4 \right\}$$
$$+ \quad \pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left| e^{-\frac{1}{2}t^2\sigma^2} - e^{-\frac{1}{2}t_0^2\sigma^2} \right| \geq \epsilon/4 \right\}$$

Now we work on each term. Note that the third term does not include (n_0, t_0) so the supremum is simply

$$\pi \left\{ \left| \mathfrak{m}\left(e^{it_0\frac{S_n^\omega f(x)}{\sqrt{n}}} \right) - e^{-\frac{1}{2}t^2\sigma^2} \right| \geq \epsilon/4 \right\}$$

and by using the bound on such terms we have from Equation 7.10 we get

$$\pi \left\{ \left| \mathfrak{m}\left(e^{it_0\frac{S_n^\omega f(x)}{\sqrt{n}}} \right) - e^{-\frac{1}{2}t^2\sigma^2} \right| \geq \epsilon/4 \right\} \leq C(\epsilon/4)^{-2}\frac{1+|t|^3}{\sqrt{n}}.$$

The last term is independent from n and n_0 so from the measure π. For k large enough we have $\sigma/k \leq \epsilon/4$ and the terms t_0 varies with $|t-t_0| \leq 1/k$ so we have $\left| e^{-\frac{1}{2}t^2\sigma^2} - e^{-\frac{1}{2}t_0^2\sigma^2} \right| \leq \epsilon/4$ which gives an empty set so the term vanishes.

The first term has t_0 fixed so we check the how the process itself differs with a change in n:

$$\begin{aligned}
& \pi \times \mathfrak{m} \left\{ \left| \frac{1}{\sqrt{n}} S_n^\omega f(x) - \frac{1}{\sqrt{n+m}} S_{n+m}^\omega f(x) \right| \geq \epsilon/4 \right\} \\
= \quad & \pi \times \mathfrak{m} \left\{ \left| \frac{\sqrt{1+m/n}-1}{\sqrt{n+m}} S_{n+m}^\omega f(x) - \frac{1}{\sqrt{n}} S_{n,n+m}^\omega f(x) \right| \geq \epsilon/4 \right\} \\
\leq \quad & \pi \times \mathfrak{m} \left\{ \left| \frac{1}{n+m} S_{n+m}^\omega f(x) \right| \geq \frac{\epsilon}{2\sqrt{n+m}(\sqrt{1+m/n}-1)} \right\} \\
+ \quad & \pi \times \mathfrak{m} \left\{ \left| \frac{1}{m} S_m^\omega f(x) \right| \geq \frac{\epsilon\sqrt{n}}{2m} \right\} \\
\leq \quad & Ce^{-C\epsilon^2 n/m}
\end{aligned}$$

by averaged Large Deviation result, Theorem 7.4.1 where $S_{n,n+m}^\omega$ represents the sum is from n to $n+m$. Furthermore we can estimate the term

$$\pi \left\{ \sup_{(n_0,t_0)\in J_k(n,t)} \left\{ \left| \mathfrak{m}\left(e^{it_0\frac{S_{n_0}^\omega f(x)}{\sqrt{n_0}}} \right) - \mathfrak{m}\left(e^{it_0\frac{S_n^\omega f(x)}{\sqrt{n}}} \right) \right| \geq \epsilon/4 \right\} \right\}$$

with

$$
\begin{aligned}
& \frac{1}{\epsilon}(\pi\times\mathfrak{m})\left\{\sup_{(n_0,t_0)\in J_k(n,t)}\left|\left(e^{it_0\frac{S^\omega_{n_0}f(x)}{\sqrt{n_0}}}\right)-\left(e^{it_0\frac{S^\omega_{n}f(x)}{\sqrt{n}}}\right)\right|^2\geq\epsilon/8\right\}\\
\leq\ & \frac{1}{\epsilon}(\pi\times\mathfrak{m})\left\{\sup_{(n_0,t_0)\in J_k(n,t)}|t_0|\left|\frac{S^\omega_{n_0}f(x)}{\sqrt{n_0}}-\frac{S^\omega_{n}f(x)}{\sqrt{n}}\right|^2\geq\epsilon/8\right\}\\
\leq\ & \frac{1}{\epsilon}\sum_{|n-n_0|\leq 2^{3k/4+1}}(\pi\times\mathfrak{m})\left\{\left|\frac{S^\omega_{n_0}f(x)}{\sqrt{n_0}}-\frac{S^\omega_{n}f(x)}{\sqrt{n}}\right|^2\geq\epsilon/16k\right\}\\
\leq\ & C\sum_{|n-n_0|\leq 2^{3k/4+1}}e^{-C\frac{n}{|n-n_0|}\epsilon^2/k^2}.
\end{aligned}
$$

The second term has n fixed while there is both t and t_0 terms in exponents so we can easily estimate that

$$
\left|e^{it_0\frac{S^\omega_{n}f(x)}{\sqrt{n}}}-e^{it\frac{S^\omega_{n}f(x)}{\sqrt{n}}}\right|\leq\frac{|t-t_0|}{\sqrt{n}}|S^\omega_n f(x)|.
$$

And by using the above estimation we get

$$
\begin{aligned}
& \pi\left\{\sup_{(n_0,t_0)\in J_k(n,t)}\left|\mathfrak{m}\left(e^{it_0\frac{S^\omega_{n}f(x)}{\sqrt{n}}}\right)-\mathfrak{m}\left(e^{it\frac{S^\omega_{n}f(x)}{\sqrt{n}}}\right)\right|\geq\epsilon/4\right\}\\
\leq\ & \frac{1}{\epsilon}(\pi\times\mathfrak{m})\left\{\sup_{(n_0,t_0)\in J_k(n,t)}\left|e^{it_0\frac{S^\omega_{n}f(x)}{\sqrt{n}}}-e^{it\frac{S^\omega_{n}f(x)}{\sqrt{n}}}\right|\geq\epsilon/8\right\}\\
\leq\ & \frac{1}{\epsilon}(\pi\times\mathfrak{m})\left\{\sup_{(n_0,t_0)\in J_k(n,t)}\frac{|t-t_0|}{\sqrt{n}}|S^\omega_n f(x)|\geq\epsilon/8\right\}\\
\leq\ & \frac{1}{\epsilon}(\pi\times\mathfrak{m})\left\{\frac{1}{\sqrt{n}}|S^\omega_n f(x)|\geq k\epsilon/8\right\}\\
\leq\ & Ce^{-Ck^2\epsilon^2}/\epsilon.
\end{aligned}
$$

Now if we add all the terms over all reference points we get

$$
\begin{aligned}
& \sum_{(n,t)\in R(J_k)}\pi\left\{\sup_{(n_0,t_0)\in J_k(n,t)}\left\{\left|\mathfrak{m}\left(e^{it_0\frac{S^\omega_{n_0}f(x)}{\sqrt{n_0}}}\right)-e^{-\frac{1}{2}t_0^2\sigma^2}\right|\geq\epsilon\right\}\right\}\\
\leq\ & C\sum_{(n,t)\in R(J_k)}\left((\epsilon/4)^{-2}\frac{1+|t|^3}{\sqrt{n}}+e^{-C\frac{n}{|n-n_0|}\epsilon^2/k^2}+\frac{e^{-Ck^2\epsilon^2}}{\epsilon}\right)\\
\leq\ & Ck^2 2^{k/4}/\epsilon\left(k^3 2^{-k/2}/\epsilon+2^{3k/4}e^{-C2^{k/4}\epsilon^2/k^2}+e^{-Ck^2\epsilon^2}\right)
\end{aligned}
$$

where $n \in R(A_k)$ and $t \in R(B_k)$. The sum over k is finite which makes it possible to apply the Borel-Cantelli argument. $\square$

CHAPTER 8

Random Induced Maps

The main goal of Chapter 8 is proving the Central Limit Theorem that induces $\mathcal{T}_1$ maps if the induced random dynamical system satisfies the Central Limit Theorem. We use the method of [CG07] which they prove the result for a single map. We generalize the method and the result to random dynamical systems.

8.1 Induced Maps of Deterministic Systems

Before introducing the system induced by random dynamical system first we review the induced map of a dynamical system given by a single map and we give the

Induced Map Theorem. The definition below is already given before for a single map in Definition 2.2.5 to introduce the class $\mathcal{T}_2(Y)$ but here we define it again with a more suitable notation so when we switch to the random dynamical system we can use the same notation, that is T for the induced maps and S for the maps that induces the map T.

Definition 8.1.1. *A one-dimensional dynamical system given by the map $S : X \to X$ is said to* induce a map

$$T : Y \to Y$$

such that $Y \subset X$ where $Y = \bigcup_{i=1}^{\infty} Y_i$ is a disjoint union of intervals, and there exists a function $R : Y \to \mathbb{N}^+$ such that $R|_{Y_i} = R(i)$ is constant and gives the number of iterations of S that must be applied to $x \in Y_i$ to get $T(x)$, so for all $i \in \mathbb{N}$ and for all $x \in Y_i$ we have

$$T(x) = S^{R(i)}(x)$$

The function R is the return time and the induced map T is the return map on Y. We say $S : Y \to Y$ has summable return times if

$$\int_Y R d\mathfrak{m} = \sum_{i=1}^{\infty} R(i)\mathfrak{m}(Y_i) < \infty$$

where R is the return time.

If a map $S : X \to X$ induces $T : Y \to Y$ then an abstract tower model can be constructed on $T : Y \to Y$ which gives the same dynamics of $S : X \to X$. Here Y is called the base of the tower. Now we give the notation for an abstract tower model as introduced in [You99].

Let F be a map defined on a space Δ where

$$\Delta := \{(x, \ell) \in Y \times \{0, 1, 2, \ldots\} : \ell < R(x)\}.$$

together with a reference measure $\mathfrak{m}$. Here Y is as above, the domain of the induced system partitioned into $\{Y_{0,i} = Y_i\}_{i=1,2,\ldots}$ and $R : Y \to \mathbb{N}^+$ is the return time function that is constant on each $Y_{0,i}$.

The function F takes (x, ℓ) simply to $(x, \ell + 1)$ if $\ell + 1 < R(x)$ and it maps each $Y_{R_i - 1, i}$ which is the top level of the tower directly above $Y_{0,i}$ bijectively onto Y, where $R_i = R(x)$ for $x \in Y_{0,i}$. We define the correspondence of two systems $F : \Delta \to \Delta$ and $S : X \to X$ by the map $\pi : \Delta \to X$ defined by

$$\pi(x, \ell) = S^\ell(x)$$

and it is easy to see that the picture below commutes:

$$\begin{array}{cccc} F: & \Delta & \longrightarrow & \Delta \\ & \downarrow \pi & & \downarrow \pi \\ S: & X & \longrightarrow & X \end{array}$$

8.2 Induced Maps of Random Dynamical Systems

In the previous section, we see that for some systems we have the corresponding abstract tower model that gives the same dynamics so it is enough to work on the tower model to explain the dynamics of the original map $S : X \to X$. Here we introduce a random dynamical system $\mathfrak{S} : X \to X$ with some conditions on the constituent maps $\{S_1, \ldots, S_N\}$ so that similarly we can introduce the induced random dynamical system of $\mathfrak{S} : X \to X$, namely $\mathfrak{T} : Y \to Y$ and then define the corresponding abstract tower model for the random dynamical system $\mathfrak{T} : Y \to Y$ with constituent maps $\{T_1, \ldots, T_N\}$.

Definition 8.2.1. *Let $\mathfrak{S} : X \to X$ be a random dynamical system with constituent maps $\{S_1, \ldots, S_N\}$ and with probability density $(p_1, \ldots, p_N)$. The random map $\mathfrak{S} : X \to X$ is said to* induce a random map

$$\mathfrak{T} : \bigcup_{i=1}^{\infty} Y_i \to Y$$

such that $\bigcup_{i=1}^{\infty} Y_i \subset Y \subset X$ where $\bigcup_{i=1}^{\infty} Y_i$ is a disjoint union of intervals such that $\mathfrak{m}(\bigcup_{i=1}^{\infty} Y_i) = \mathfrak{m}(Y)$, and there exists a function $R : \mathbb{N} \to \mathbb{N}^+$ such that $R(i)$ gives the number of iterations of $\mathfrak{S}$ that must be applied to $x \in Y_i$ to get $\mathfrak{T}(x)$ which is independent from the first $R(i)-1$ choices of $\mathfrak{S}$ meaning if $S_i(x) \notin Y$ then $S_j(x) \notin Y$ and $S_i(x) = S_j(x)$ for every $j = 1, \ldots, N$, so for all $i \in \mathbb{N}$ and for all $x \in Y_i$ we have

$$\mathfrak{T}(x) = \mathfrak{S}^{R(i)}(x)$$

which means in particular

$$T_j(x) = S_j^{R(i)}(x) \text{ for every } j \in \{1, \ldots, N\}$$

so the random dynamical system $\mathfrak{T}$ has constituent functions $\{T_1 = S_1^R, \ldots, T_N = S_N^R\}$ with probability density $(p_1, \ldots, p_N)$ where S^R is defined by $S^{R(x)}(x) = S^{R(i)}(x)$ for $x \in Y_i$.

From now on we call $\mathfrak{S} : X \to X$ a hybrid system since it is deterministic as long as it does not end up in Y and the number of iterations that is needed to come back to Y does not depend on which constituent map is applied, it is only given by the partition $\{Y_i\}$. Now assume that the hybrid system $\mathfrak{S} : X \to X$ induces the random dynamical system $\mathfrak{T} : Y \to Y$. We denote the sequences of symbols on X to be Ω and on Y to be Ω_0 to emphasize the maps they represent although they are the same sequences.

We define the corresponding hybrid maps $F_1, F_2, \ldots, F_N$ on an abstract tower Δ,

$$\Delta := \{(x, \ell) \in Y \times \{0, 1, 2, \ldots\} : \ell < R(x)\}.$$

which is defined to be exactly the same way of a tower of a single map since the return map $R : Y \to \mathbb{N}^+$ does not depend on which constituent map is applied. We define each F_i to be identical on (x, ℓ) if $\ell + 1 < R(x)$ which simply maps (x, ℓ) to $(x, \ell + 1)$. If $\ell + 1 = R(x)$, then we define F_i on $(x, R(x) - 1)$ to be $F_i(x, \ell) = (T_i(x), 0)$. Let $\{F_1, F_2, \ldots, F_n\}$ be the set of constituent maps and $\mathfrak{F}$ denote the hybrid map on Δ with the set of sequences of sumbols Ω_Δ. We assign probability density $(p_1, p_2, \ldots, p_N)$ to each symbol which gives the probability of choosing each map respectively when we are iterating the system.

We claim that the hybrid maps system $\mathfrak{F}$ on the tower Δ gives the same dynamics of $\mathfrak{S}$ on X. For that we define $\pi : \Omega_\Delta \times \Delta \to \Omega \times X$ by

$$\pi(\omega, x, \ell) = (\omega, S_{\omega'_\ell} \circ \ldots \circ S_{\omega'_1}(x)) \in \Omega \times X$$

for any finite ℓ sequence $[\omega']$ of functions $\mathfrak{S}$ since $\ell < R(x)$ so we have

$$\begin{array}{lccc}
\mathfrak{F}: & (\omega, x, \ell) & \longrightarrow & (\boldsymbol{\sigma}(\omega), F_{\omega_1}(x, \ell)) \\
& \downarrow \boldsymbol{\pi} & & \downarrow \boldsymbol{\pi} \\
\mathfrak{S}: & (\omega, S_{\omega'_\ell} \circ \ldots \circ S_{\omega'_1}(x)) & \longrightarrow & (\boldsymbol{\sigma}(\omega), S_{\omega_1} S_{\omega'_\ell} \circ \ldots \circ S_{\omega'_1}(x))
\end{array}$$

which commutes since if $\ell+1 < R(x)$ then the right upper corner is simply $(\boldsymbol{\sigma}(\omega), x, \ell+1)$ which maps to $(\boldsymbol{\sigma}(\omega), S_{\omega_1} S_{\omega'_\ell} \circ \ldots \circ S_{\omega'_1}(x))$ under $\boldsymbol{\pi}$ where $\omega_1 \omega'_\ell \ldots \omega'_1$ is an arbitraty $\ell+1$ sequence, and if $\ell+1 = R(x)$ then the right upper corner is $(\boldsymbol{\sigma}(\omega), T_{\omega_1} x, 0)$ which maps to $(\boldsymbol{\sigma}(\omega), S_{\omega_1} S_{\omega'_\ell} \circ \ldots \circ S_{\omega'_1}(x)) = (\boldsymbol{\sigma}(\omega), S_{\omega_1}^{\ell+1}(x)) = (\boldsymbol{\sigma}(\omega), T_{\omega_1}(x))$ since the S functions are identical if the process is not returning to Y so we can use S_{ω_1} for every step before the return which is decided by the sequence ω.

Theorem 8.2.2. *Assume $\int R d\mathfrak{m} < \infty$. If the random dynamical system $\mathfrak{T}$ on Δ_0 has a finite stationary measure μ_0 whose density is uniformly bounded then the random dynamical system $\mathfrak{F}$ on Δ has a finite stationary measure.*

Proof. We simply push forward the measure defined on Δ_0 to the tower. Let μ_0 be the stationary measure for the system $\{T_1, T_2, \ldots, T_N\}$ on Δ_0 with probability vector $(p_1, p_2, \ldots, p_n)$. Define $\mu'_i = \sum_{\ell=0}^{\infty} (F_i)_*^\ell (\mu_0 | \{R > \ell\})$ where $(F_i)_*^\ell \mu_0(E) = \mu_0(F_i^{-\ell} E)$. Since $\dfrac{d\mu_0}{d\mathfrak{m}}$ is uniformly bounded, and $\displaystyle\int R d\mathfrak{m} < \infty$ we have $\mu'_i(\Delta) < \infty$ for every $i = 1, 2, \ldots, n$. Then define $\mu' = \sum_{i=1}^{N} p_i \mu'_i$. We normalize μ' to give the desired stationary measure μ to the system with functions $\{F_1, F_2, \ldots, F_n\}$ on Δ. $\square$

We need the following theorem for the initial model we define in Chapter 2. The connection of the theorem to the initial model is given in Chapter 9 . It is mainly expanding the averaged Central Limit Theorem from the induced system to the original system and is the main theorem of the section.

Theorem 8.2.3. *Let $\mathfrak{T} : \Delta_0 \to \Delta_0$ be a random dynamical system with constituent maps $T_1, T_2, \ldots, T_N$, a probability vector $(p_1, p_2, \ldots, p_N)$ and a stationary measure*

μ_0. Let Δ be the tower defined as above with respect to the return time function R with $\int R d\mathfrak{m} < \infty$.

Define $\hat{f} : \Delta \to \{0, 1\}$ to be a function on Δ and set $S_n^\omega \hat{f} = \sum_{j=0}^{n-1} \hat{f} \circ \mathfrak{F}^j$ where $\mathfrak{F}$ is randomly chosen from $\{F_1, F_2, \ldots, F_N\}$ and $F_i^R = T_i$ for $i = 1, 2, \ldots, N$. We define $f : \Delta_0 \to \{0, 1\}$ to be the function on Δ_0 given by

$$f(y) = \sum_{k=0}^{R(y)-1} \hat{f} \circ \mathfrak{F}^k$$

where $\mathfrak{F}$ is randomly chosen but does not affect the value of $f(y)$ since every constituent map just takes the point to one level up for $0 \leq k \leq R(y) - 1$. Set $S_n^\omega f = \sum_{j=0}^{n-1} f \circ \mathfrak{T}^j$ where $\mathfrak{T}$ is randomly chosen according to the constituent maps sequence ω. For almost every realization of constituent maps $\mathfrak{T}$ on Δ_0 if we have

$$\lim_{n \to \infty} (\pi \times \mu_0) \left\{ \frac{f_n - n \int f d\mu_0}{\sqrt{n}} < c \right\} = \frac{1}{\sqrt{2\pi\sigma_0^2}} \int_{-\infty}^{c} e^{-x^2/2\sigma_0^2} dx,$$

then on Δ we have

$$\lim_{n \to \infty} (\pi \times \mu) \left\{ \frac{\hat{f}_n - n \int \hat{f} d\mu}{\sqrt{n}} < c \right\} = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{c} e^{-x^2/2\sigma^2} dx.$$

The proof has two main steps. First we give how the process on the tower is related to its induced process. Then by using the fact that the convergence to normal distribution is in fact tight as given in Theorem 7.4.3 we expand the result to the tower model.

First, let us give the drift of the original system in terms of the drift of the induced system.

Lemma 8.2.4. *If the functions $\hat{f}$ on Δ, f on Δ_0 and measures μ, μ_0 and return time R as above in Theorem 8.2.3 then*

$$\int \hat{f} d\mu = \frac{\int f d\mu_0}{\int R d\mu_0}.$$

Proof. Let us give the proof for the characteristic functions on $A \subset \Delta_0$ since it is easier to calculate but the general idea is similar for any map f on Δ_0.

$$\begin{aligned}
\int \hat{\chi}_A d\mu &= \mu(A) \\
&= \frac{\mu'(A)}{\mu'(\Delta)} \\
&= \left(\sum_{i=1}^{N} p_i \mu_0(A)\right) / \left(\sum_{i=1}^{N} p_i \sum_{\ell=0}^{\infty} (F_i)_*^\ell \mu_0|_{R>\ell}(\Delta)\right) \\
&= \mu_0(A) / \left(\sum_{i=1}^{N} p_i \sum_{\ell=1}^{\infty} \sum_{j=\ell}^{\infty} \mu_0(\Delta_{0,j})\right)
\end{aligned}$$

$$\begin{aligned}
&= \mu_0(A) / \left(\sum_{\ell=1}^{\infty} \sum_{j=\ell}^{\infty} \mu_0(\Delta_{0,j})\right) \\
&= \mu_0(A) / \left(\sum_{j=0}^{\infty} \sum_{\ell=1}^{j} \mu_0(\Delta_{0,j})\right) \\
&= \mu_0(A) / \left(\sum_{j=0}^{\infty} j \mu_0(\Delta_{0,j})\right) \\
&= \mu_0(A) / \left(\int R d\mu_0\right) \\
&= \int \chi_A d\mu_0 / \left(\int R d\mu_0\right).
\end{aligned}$$

Therefore we have

$$\int \hat{\chi}_A d\mu = \frac{\int \chi_A d\mu_0}{\int R d\mu_0} \tag{8.1}$$

□

Here is how two processes are related:

Theorem 8.2.5. *For each $\hat{f}$ on Δ and for every sequence of maps , say $\mathcal{F}$ on Δ and for μ-almost every point $(x,\ell)\in\Delta$, there exists a function f on Δ_0, a sequence of maps, say $\mathcal{T}$ on Δ_0 such that for every $n\in\mathbb{N}$ there exists $m=m(x,\mathcal{F})\in\mathbb{N}$ such that*

$$S_n^{\mathcal{F}}\hat{f}(x,\ell)=S_m^{\mathcal{T}}f(x)$$

where $S_n^{\mathcal{F}}\hat{f}$ is the sum of the first n iterations of $\mathcal{F}$ evaluated under $\hat{f}$ on Δ while $S_m^{\mathcal{T}}f$ is the sum of the first m iterations of $\mathcal{T}$ evaluated under f on Δ_0, that is

$$S_m^{\mathcal{T}}f=\sum_{j=0}^{m-1}f\circ T_{\omega'_j}\circ\ldots\circ T_{\omega'_1}$$

and

$$S_n^{\mathcal{F}}\hat{f}=\sum_{j=0}^{n-1}\hat{f}\circ F_{\omega_j}\circ\ldots\circ F_{\omega_1}$$

if $\mathcal{T}=(T_{\omega'_1},T_{\omega'_2},\ldots)$ and $\mathcal{F}=(F_{\omega_1},F_{\omega_2},\ldots)$.

Proof. Fix one realization of the constituent maps say$\mathcal{F}=(F_{\omega_1},F_{\omega_2},F_{\omega_3},\ldots)$. The sum of the first n iteration of $\hat{f}$ with respect to the fixed sequence $\mathcal{F}$ is denoted by $S_n^{\mathcal{F}}\hat{f}$ and given by

$$S_n^{\mathcal{F}}\hat{f}(x,0)=\hat{f}(x,0)+\hat{f}\circ F_{\omega_1}(x,0)+\hat{f}\circ F_{\omega_2}\circ F_{\omega_1}(x,0)+\cdots+\hat{f}\circ F_{\omega_{n-1}}\circ\ldots\circ F_{\omega_1}(x,0).$$

Here we choose the special point $(x,0)\in\Delta$ instead of (x,ℓ) to show the result. It is enough to do that since for (x,ℓ) with $\ell>0$ we can add ℓ-many of the first map of the sequence $\mathcal{F}$ to itself and define $\mathcal{F}'=(F_{\omega_1},F_{\omega_1},\ldots,F_{\omega_1},F_{\omega_2},\ldots)$ and start iterating at $(x,0)$. Since this modification adds finitely many iterations and since it does not change the point x, it is enough to show the result by starting at point $(x,0)$.

We can also choose to iterate with the same function we start iterating from the base when the image is on the levels of tower since the functions are identical as long as the image is not in Δ_0. Therefore we can write

$$F_{\omega_j} \circ \ldots \circ F_{\omega_2} \circ F_{\omega_1}(x,0) = F_{\omega_1} \circ \ldots \circ F_{\omega_1} \circ F_{\omega_1}(x,0)$$

where F_{ω_1} is iterated j times on the right-hand side of the above equation as long as $F_{\omega_k} \circ \ldots \circ F_{\omega_1}(x,0) \notin \Delta_0$ for $k = 1, \ldots, j$, that is $R(x) < j$.

To simplify the notation let $x_0 = (x,0)$, $x_1 = F^{R(x_0)}_{\omega_{R(x_0)}}(x_0)$, $x_2 = F^{R(x_1)}_{\omega_{R(x_1)+R(x_0)}}(x_1)$ and so on, where $R(x_i)$ is the return time of $x_i \in \Delta_0$ and does not depend on $\mathcal{F}$. When we write $F^{R(x_1)}_{\omega_{R(x_1)+R(x_0)}}(x_1)$ the subindex gives the position of the map F on the sequence $\mathcal{F}$ so it decides which constituent map is used to iterate the system while the upper index gives how many times the function is iterated. Then the sum becomes

$$\begin{aligned}
S_n^{\mathcal{F}} \hat{f}(x_0) &= \hat{f}(x_0) + \ldots + \hat{f} \circ F_{\omega_1}^{R(x_0)-1}(x_0) \\
&+ \hat{f} \circ F^{R(x_0)}_{\omega_{R(x_0)}}(x_0) + \ldots + \hat{f} \circ F^{R(x_1)-1}_{\omega_{R(x_0)}}(x_1) \\
&+ \hat{f} \circ F^{R(x_1)}_{\omega_{R(x_1)+R(x_0)}}(x_1) + \ldots + \hat{f} \circ F^{R(x_2)-1}_{\omega_{R(x_1)+R(x_0)}}(x_2) \\
&+ \ldots \\
&+ \hat{f} \circ F^{R(x_m)}_{\omega_{R(x_m)+\cdots+R(x_0)}}(x_m), \\
&= \hat{f}(x_0) + \ldots + \hat{f} \circ F_{\omega_1}^{R(x_0)-1}(x_0) + \hat{f}(x_1) + \ldots + \hat{f} \circ F^{R(x_1)-1}_{\omega_{R(x_0)}}(x_1) \\
&+ \hat{f}(x_2) + \ldots + \hat{f} \circ F^{R(x_2)-1}_{\omega_{R(x_1)+R(x_0)}}(x_2) + \cdots + \hat{f}(x_{m+1}).
\end{aligned}$$

where $m = m(n, x_0, \mathcal{F})$ is the number of returns to the base in first n iterations of the constituent maps sequence $\mathcal{F}$ starting at $x_0 \in \Delta_0$. For now let us assume that we choose n so that the process ends up in the base after the n iterations of $\mathcal{F}$. It is easy to show that the difference between a process with any n and a process with such n that end up in the base converges to zero in $\pi \times \mu$.

In the sum above note that the functions are of the form

$$\hat{f}(x_i) + \ldots + \hat{f} \circ F^{R(x_i)-1}_{\omega_{R(x_{i-1})}}(x_i) = f(x_i)$$

for some $1 \leq i \leq m$ which is a function on Δ_0 since the sum only depends on $x_i \in \Delta_0$. Also the functions of the form $\mathfrak{F}^R$ can be considered as follows:

$$F^R_{\omega_{R(x_0)}} = T_{\omega'_0}, \quad F^R_{\omega_{R(x_1)+R(x_0)}} = T_{\omega'_1}, \quad F^R_{\omega_{R(x_2)+R(x_1)+R(x_0)}} = T_{\omega'_2}, \ldots$$

so we have

$$x_1 = T_{\omega'_0}(x_0), \quad x_2 = T_{\omega'_1}(x_1), \quad x_3 = T_{\omega'_2}(x_2) \ldots$$

Then we define the sequence $\mathcal{T} = (T_{\omega'_0}, T_{\omega'_1}, T_{\omega'_2}, \ldots)$ that depends on $\mathcal{F}$ and x_0. Therefore for each $\mathcal{F}$ and for each $x_0 \in \Delta_0$ there exists a sequence ω' in Ω_0 so that there exists a corresponding sequence $\mathcal{T}$ of functions $\mathfrak{T}$ such that for every $n \in \mathbb{N}$ there exists $m = m(n, x_0, \mathcal{F})$ with $S^{\mathcal{F}}_n \hat{f}(x_0) = S^{\mathcal{T}}_m f(x_0)$ where $\mathcal{T} = (T_{\omega'_0}, T_{\omega'_1}, T_{\omega'_2}, \ldots)$. $\square$

Now after analyzing how the original and the induced systems are related let us go back to what we want to show. We want to prove the system on the tower satisfies the Central Limit Theorem so we fix a sequence $\omega \in \Omega_\Delta$ and denote the sequence of maps obtained by the fixed ω by $\mathcal{F}$ so $\mathcal{F} = (F_{\omega_1}, F_{\omega_2} \ldots)$. We rewrite the process on the tower obtained by the sequence $\mathcal{F}$ by using the corresponding process $\mathcal{T}$ on Δ_0.

Again assume that $\int \hat{f} d\mu = 0$ to simplify the calculations. Therefore the process

$$\frac{S^{\mathcal{F}}_n \hat{f}}{\sqrt{n}}(x, 0)$$

can be approximated by

$$\frac{S_m^{\mathcal{T}} f}{\sqrt{n}}(x)$$

where m is the largest integer such that $R_m^{\mathcal{T}}(x) \leq n$ where R_m is the sum of the return times of first n iterates of $\mathcal{T}$.

Since μ_0 is ergodic we also have by Birkhoff's Ergodic Theorem that

$$\frac{1}{m} R_m^{\mathcal{T}}(x) = \frac{1}{m} R \circ \mathscr{F}^m(\mathcal{T}, x) \to \int R(x) d(\pi \times \mu_0) = \int R d\mu_0$$

and since the system is ergodic this is equal to $\mu(\Delta_0)$ by Kač's Lemma A.0.7. Then we can approximate $n = R_m^{\mathcal{T}}(x)$ with $\mu(\Delta_0)m$. Therefore it is enough to show that the process

$$\frac{S_m^{\mathcal{T}} f}{\sqrt{n}}(x)$$

is converging to the normal distribution $\mathscr{N}(0, \sigma)$ with the measure μ' given by $d\mu' = \chi_{\Delta_0} R d\mu$. We prove it in the following theorem which concludes also the Theorem 8.2.3. Note that the sequence m_n given in the following theorem corresponds to $m(x, \mathcal{F}, n)$ defined above and is approximated by $\lfloor n/\mu(\Delta_0) \rfloor$ for our case so m_n/n converges to one.

Theorem 8.2.6. *Let $(Y, \mu_0, \mathfrak{T})$ be an ergodic probability preserving random dynamical system, and let $f : Y \to \mathbb{R}$. Assume the process $S_n^\omega f/\sqrt{n}$ converges to normal distribution $\mathscr{N}$ with tight maxima in average. Let $m_1, m_2, \ldots$ be a sequence of integer valued functions on $\Omega \times X$ such that m_n/n converges to 1 in probability. Then $S_{m_n}^\omega f/\sqrt{n}$ also converges to $\mathscr{N}$ for any absolutely continuous measure μ' with respect to μ_0.*

Proof. We show that for any $\epsilon > 0$, $\delta > 0$ we have

$$(\pi \times \mu) \left\{ (\omega, x) : \left| \frac{S^{\omega}_{m_n} f(x) - S^{\omega}_{n} f(x)}{\sqrt{n}} \right| \geq \epsilon \right\} \leq 2\delta$$

Then convergence to the normal distribution $\mathscr{N}$ with respect to $\pi \times \mu'$ follows from Eagleson's Theorem [Eag76].

We have that the convergence of $S^{\omega}_{n} f/\sqrt{n}$ to normal distribution is with tight maxima so there exists $c > 0$ such that for all $n \in \mathbb{N}$ we have

$$(\pi \times \mu_0) \left\{ \max_{1 \leq k \leq n} |S^{\omega}_{n} f| \geq c\sqrt{n} \right\} \leq \delta.$$

Let $r \in (0, 1)$ be small enough so that $\lceil \sqrt{2rn} \rceil \leq \dfrac{\epsilon\sqrt{n}}{2c}$ for all large enough n and define a sequence s_n to be $s_n = \lceil (1 - r)n \rceil$. Since m_n/n converges to 1 in probability we have for large n that $(\pi \times \mu_0)\{(\omega, x) : |m_n(\omega, x) - n| > rn\} \leq \delta$. Then we get

$$\begin{aligned} & (\pi \times \mu_0) \left\{ \left| \frac{S^{\omega}_{m_n} f(x) - S^{\omega}_{n} f(x)}{\sqrt{n}} \right| \geq \epsilon \right\} \\ \leq \quad & \delta + (\pi \times \mu_0) \left\{ \left| \frac{S^{\omega}_{m_n} f(x) - S^{\omega}_{n} f(x)}{\sqrt{n}} \right| \geq \epsilon \text{ with } m_n \in [(1 - r)n, (1 + r)n] \right\}. \end{aligned}$$

Let (ω, x) belong to that last set. Then there exists $m \in [(1 - r)n, (1 + r)n]$ so that $|S^{\omega}_{m} f(x) - S^{\omega}_{n} f(x)| \geq \epsilon\sqrt{n}$. Since $s_n = \lceil (1 - r)n \rceil \geq (1 - r)n$ and it is the smallest integer in that set so $s_n \leq m$ so we should have $\left|S^{\omega}_{m} f(x) - S^{\omega}_{s_n} f(x)\right| \geq \epsilon\sqrt{n}/2$ or $\left|S^{\omega}_{n} f(x) - S^{\omega}_{s_n} f(x)\right| \geq \epsilon\sqrt{n}/2$. Then we choose the maximum over all possibilities so we get

$$\begin{aligned} & (\pi \times \mu_0) \left\{ \left| \frac{S^{\omega}_{m_n} f(x) - S^{\omega}_{n} f(x)}{\sqrt{n}} \right| \geq \epsilon \right\} \\ \leq \quad & \delta + (\pi \times \mu_0) \left\{ \max_{0 \leq k \leq \lceil 2rn \rceil} \left| S^{\omega}_{s_n + k} f(x) - S^{\omega}_{s_n} f(x) \right| \geq \epsilon\sqrt{n}/2 \right\} \end{aligned}$$

and then by the choice of r we have $\epsilon\sqrt{n}/2 \geq c\sqrt{2rn}$ so

$$\begin{aligned}
&(\pi \times \mu_0)\left\{\left|\frac{S^\omega_{m_n} f(x) - S^\omega_n f(x)}{\sqrt{n}}\right| \geq \epsilon\right\} \\
\leq\ & \delta + (\pi \times \mu_0)\left\{\max_{0\leq k\leq \lceil 2rn\rceil} \left|S^{\omega'}_k f(y)\right| \geq \epsilon\sqrt{n}/2\right\}
\end{aligned}$$

which is less than 2δ since c is the constant that satisfies the tight maxima property.

$\square$

CHAPTER 9

Corresponding Models

9.1 The Corresponding Random Dynamical System for the Hybrid Model

In Chapter 9 we show that the Markov process defined in Chapter 2 can be realized as a random dynamical system. The idea is to append the jump interval to $[0,1]$ and on this new space define two maps, one which simulates the situation when there is a jump and the other map which simulates the no jump situation. We assign probability p to the jump map and probability $1-p$ to the no jump map. Then to study the drift of the system to the right on the lattice we simply study the number of visits to the jump interval.

Definition 9.1.1. *For $I = [0,1]$, let $\tau : I \to I$ be the local map that belongs to $\mathcal{T}_1(I)$, U be the jump interval and $\varphi : U \to [0,1]$ be the jump map and p is the jump probability for the Markov process as defined in Chapter 2. Let $Y = [0,1] \cup \tilde{U}$ where $\tilde{U}$ is simply the copy of the jump interval so $\mathfrak{m}(\tilde{U}) = \mathfrak{m}(U)$. Define $id : U \to \tilde{U}$ to be the identification and $\tilde{\varphi} : \tilde{U} \to [0,1]$ so that $\varphi = \tilde{\varphi} \circ id$, $\mathfrak{m}$ denotes the Lebesgue measure on Y. We define the constituent maps T_0, T_1 first on $[0,1]$ as follows*

$$T_0(x) = \tau(x) \text{ for } x \in [0,1],$$

$$T_1(x) = \begin{cases} \tau(x) & \text{if } \tau(x) \notin U, \\ id \circ \tau(x) & \text{if } \tau(x) \in U. \end{cases}$$

And for $x \in \tilde{U}$ we define both of the maps by first moving the point to $[0,1]$ with $\tilde{\varphi}$ and define the same way as they are defined in $[0,1]$, namely $T_i(x) = \tilde{\varphi} \circ T_i(x)$ for $i = 0,1$ if $x \in \tilde{U}$. Here T_1 is the jump map and T_0 is the no jump map with probabilities p and $1-p$ respectively. We call the random dynamical system given by constituent maps T_0 and T_1 with Bernoulli measure obtained by the probability vector $(1-p, p)$ the corresponding random dynamical system *of the Markov process defined in Chapter 2.*

Now we are ready to show that the new maps system consists of maps in $\mathcal{T}_1(Y)$ with respect to a common partition β. For that assume α is the partition for the original map τ with the properties given in Model I. We define a new partition on $[0,1]$ by using the partition element of α and refining it if a partition element intersects the jump interval U and on $\tilde{U}$ again by using the partition α since $\tilde{U}$ is only an identification of U and refining it with the map $\tilde{\varphi}$. Since $\tau \in \mathcal{T}_1(I)$ both of the maps T_0 and T_1 are piecewise uniformly expanding because we assume that $|\tau'| \cdot |\tilde{\varphi}'| > \lambda > 1$.

The random dynamical system on Y given by the constituent maps $\{T_0, T_1\}$ represents the Markov process $\mathcal{J}_n(x)$ on $\mathbb{N} \times [0,1]$ given in Model I. The transitions can simply be given by

$$(i,x) \to \begin{cases} (i, \mathfrak{T}(x)) & \text{if } x \in [0,1] \\ (i+1, \mathfrak{T}(x)) & \text{if } x \in \tilde{U} \end{cases}$$

where $\mathfrak{T}$ is chosen from $\{T_0, T_1\}$ with probability distribution $(1-p, p)$.

9.2 The Drift Rates and Limit Theorems of the Hybrid Models

Assume for simplicity that we start at site 0. Note that the Markov process can be given by iterating the corresponding random dynamical system and by counting how many times the system ends up in the jump interval, so we have

$$\mathcal{J}_n(x) = \sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ \mathfrak{T}^i(x).$$

Two random variables $\mathcal{J}_n$ and $\sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ \mathfrak{T}^i$ are in fact same as functions of x. Furthermore, the sample space of $\mathcal{J}_n$ consists of the jump sequences. Note that the jump sequences do not depend on x, we can first fix a jump sequence and then start iterating the process. We check the next entry of a jump sequence only if the process ends up in the jump interval R. Let $\omega \in \Sigma$ be a fixed jump sequence and $x \in I$ be fixed. Let n_k be the subsequence of n so that the n_k iterates of the process is in R, so k is the index of the jump sequence. Since x and ω are fixed the subsequence is well-defined. Then there exists a sequence of random maps ω' that depends on both ω and x such that $T_{n_k} = T_0$ if $k = 0$, and $T_{n_k} = T_1$ if $k = 1$. The other entries of the random maps sequence can be chosen to be T_0 or T_1. Note that the ρ measure of any fixed cylinder of jump sequences is same as the π measure of the set of corresponding random map sequences. Therefore, for any fixed jump sequence $\omega \in \Sigma$ if we consider $\mathcal{J}_n^{\omega}(x)$ also as a function of ω we have

$$\mathcal{J}_n^{\omega}(x) = \sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ T_{\omega'_k} \circ \ldots \circ T_{\omega'_1}(x)$$

where $\omega' \in \Omega$ is the corresponding random maps sequence of $\omega \in \Sigma$. Thus, whenever we prove a result for almost every jump choices of the process $\mathcal{J}_n(x)$ we can prove the same result for the process $\sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ \mathfrak{T}^i(x)$ for π almost every random maps sequence.

First we prove the drift rate. By using Theorem 4.1.1 together with Corollary 6.2.4 we know that we have an ergodic stationary measure μ for the random dynamical system. By Ergodic Theorem we conclude that for μ-almost every $x \in Y$ and for π-almost every sequence of maps $\mathfrak{T}$ we have

$$\frac{1}{n} \sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ \mathfrak{T}^i(x) \to \mu(\tilde{U})$$

so $\alpha = \mu(\tilde{U})$ is the Drift Rate. Since μ has the positive density $h(x)$ and $\mathfrak{T}$ is mixing the drift rate result also holds for $\mathfrak{m}$-almost every x. This proves Theorem 2.3.1 for

Model I. To give the drift rate result for the center of mass

$$\lim_{n\to\infty} \frac{\mathfrak{C}(\mu_n)}{n} = \lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n} i\, \mu_n(X_i)$$

again we consider the corresponding random dynamical system. Here $\sum_{i=0}^{n} i\, \mu_n(X_i)$ is the expected number of jumps in the first n iteration which can be easily given for the corresponding random dynamical system by

$$\int_\Omega \int_Y \sum_{i=0}^{n} \chi_{\tilde{U}} \circ \mathscr{F}^i(\omega, x) d\pi d\mu = n\mu(\tilde{U})$$

since $\pi \times \mu$ is $\mathscr{F}$-invariant. Then in fact $\dfrac{\mathfrak{C}(\mu_n)}{n} = \alpha$ for every n.

The Central Limit Theorem for Model I follows from Theorem 7.4.5 which is given for $\mathfrak{m}$-almost every $x \in Y$ together with the equality of the Markov process and the random dynamical system given above. Note that we need the Central Limit Theorem for non-stationary measures since the process $\mathscr{J}_n$ is not restricted to the stationary measure of the random dynamical system. However, the initial distribution is absolutely continuous so it is enough to give the results for the corresponding random dynamical system with respect to Lebesgue.

The maps given in Model II have induced maps that lie in $\mathcal{T}_1(I)$, so say $\varsigma \in \mathcal{T}_2(I)$ with $\varsigma : I \to I$ and let $\tau : Y \to Y$ be the map induced by ς with return time map $R : Y \to \mathbb{N}$, for some measurable set $Y \subset I$. The only extra condition for such Markov processes is that the jump interval is placed to the base Y and the jump map is $\varphi : U \to Y$ so it is mapped to the base Y. This condition is not that restrictive since we know that most of the maps that are modeled with such induced systems can have the whole set I as their base, see [You99] for an example. Then we can consider the Markov process only on the base so with local dynamics given by $\tau : Y \to Y$. We define the constituent maps T_0, T_1 in a similar way. If we define S_0, S_1 to be T_0^R, T_1^R respectively then the new set of constituent maps $\{S_0, S_1\}$ on I satisfies the properties of maps given in Chapter 8, Definition 8.2.1. Therefore we can continue with the results given for such maps.

First the stationary measure μ is given as the pushforward measure of the stationary measure of the base μ_0 and the jump set is in the base Y therefore the drift rate of the random dynamical system $\mathfrak{S}$ given by the constituent maps S_0, S_1 is

$$\frac{1}{n}\sum_{k=0}^{n-1} \chi_{\tilde{U}} \circ \mathfrak{S}^i(x) \to \mu(\tilde{U}) = \mu_0(\tilde{U}) / \int R d\mu_0 = \mu_0(\tilde{U})/\mu(Y) = \alpha/\mu(Y)$$

where α is the drift rate of the induced system. The drift rate for the center of mass can be given as before and the Central Limit Theorem for the random dynamical system $\mathfrak{S}$ is a result of Theorem 8.2.3.

The other limit theorems for the random variable $\mathcal{J}_n(x)$ can be given with the same idea, including the rate of convergence to the normal distribution, tight maxima and large deviation estimate. For the quenched Central Limit Theorem again recall the correspondence of jump sequences and the random maps sequences, then the result follows.

9.3 Example: Pomeau-Manneville Maps

Lastly we give a concrete example for a hybrid system and give the corresponding random dynamical system. We have a Pomeau-Manneville map as the local dynamics defined by

$$\varsigma(x) = \begin{cases} x + 2^\alpha x^{\alpha+1} \text{ if } x \in [0, \frac{1}{2}) \\ 2x - 1, \text{ if } x \in [\frac{1}{2}, 1] \end{cases}$$

where $x_0 = 0$ is the only neutral point with $f'(0) = 1$. Let $Z = [0, \frac{1}{2})$ be the neighborhood of $x_0 = 0$, and divide $Y = [\frac{1}{2}, 1]$ into partitions:

$Y_1 = [q_1, 1]$ where $\varsigma(q_1) = p_1$ with $p_1 = \dfrac{1}{2}$, so $q_1 = \dfrac{3}{4}$

$Y_2 = [q_2, q_1)$ where $\varsigma(q_2) = p_2$ with $p_1 = \varsigma(p_2)$, so $p_2 \approx 0.2850$ and $q_2 \approx 0.6425$

$Y_3 = [q_3, q_2)$ where $\varsigma(q_3) = p_3$ with $p_2 = \varsigma(p_3)$, so $p_3 \approx 0.1784$ and $q_3 \approx 0.5892$

$\vdots$

The return map corresponding to the above partition is $\tau = \varsigma^R$ where $R|_{Y_j} = j$ and the graph of τ is given below. Let $U = [\frac{13}{16}, \frac{14}{16}]$ and $V = [\frac{10}{16}, \frac{11}{16}]$ be the jump

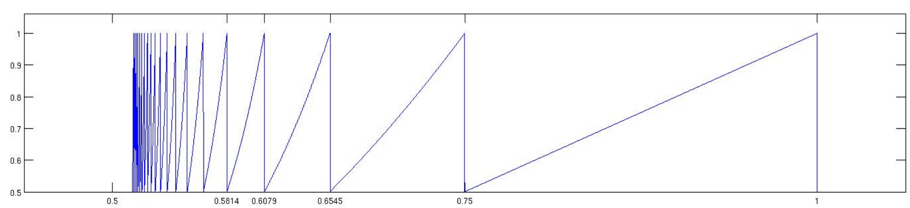

Figure 9.1: Return function τ with $\alpha = 0.5$ with respect to the partition $\{Y_j\}$

intervals to right and left respectively. In previous chapters we define the hybrid models only with a jump to right but it can be generalized to systems with jumps to a finite distance sites as long as the jump intervals do not intersect. Here we have $U \cap V = \emptyset$. Let $\phi : U \to [\frac{1}{2}, 1]$ defined by $\phi(x) = x^3$ and $\psi : V \to [\frac{1}{2}, 1]$ defined by $\psi(x) = x^3 + \frac{1}{2}$. The isometries are given by $i_U(x) = x + \frac{3}{16}$ and $i_V(x) = x - \frac{3}{16}$ so $\hat{U} = [1, \frac{17}{16}]$ and $\hat{V} = [\frac{7}{16}, \frac{1}{2}]$ and $\mathbf{Y} = [\frac{7}{16}, \frac{17}{16}]$.

We define the random maps. The map T_{UV} simulates the situation of jumps for both intervals U and V and is given on $[\frac{1}{2}, 1]$ by

$$T_{UV}(x) = \begin{cases} \tau(x) + \frac{3}{16}, & \text{if } x \in [\frac{1}{2}, 1] \text{ and } \tau(x) \in U\,, \\ \tau(x) - \frac{3}{16}, & \text{if } x \in [\frac{1}{2}, 1] \text{ and } \tau(x) \in V\,, \\ \tau(x), & \text{if } x \in [\frac{1}{2}, 1], \text{ but } \tau(x) \notin U \cup V\,, \end{cases}$$

and $T_{UV}(x) = T_{UV}(x^3)$ for $x \in \hat{U}$ and $T_{UV}(x) = T_{UV}(x^3 + \frac{1}{2})$ for $x \in \hat{V}$. The graph of T_{UV} on $[\frac{7}{16}, \frac{17}{16}]$ is given below.

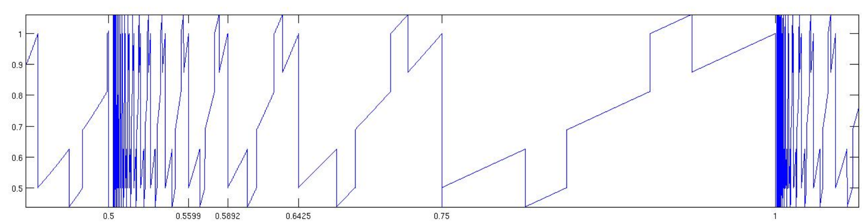

Figure 9.2: The corresponding $T_{UV} : [\frac{7}{16}, \frac{17}{16}] \circlearrowleft$ of the return map τ

If we look at the graphs of the map T_{UV} restricted to the right jump interval $\hat{U}$ and restricted to the left jump interval $\hat{V}$ respectively below we see the same graph that we have on the set $[\frac{1}{2}, 1]$ but with a different scaling.

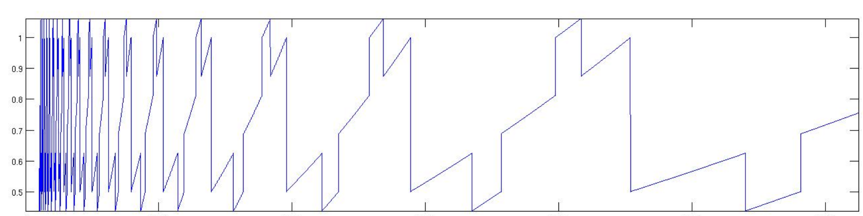

Figure 9.3: T_{UV} map restricted to $\hat{U} = [1, \frac{17}{16}]$

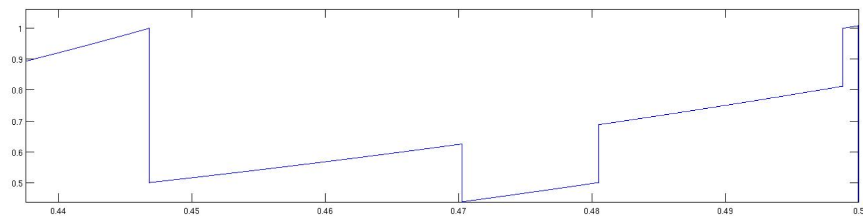

Figure 9.4: T_{UV} map restricted to $\hat{V} = [\frac{7}{16}, \frac{1}{2}]$

The map T_U simulates the situation when there is jump on U to right but no jump on V to left. Similarly T_V is a jump on V and no jump on U map. The map T_* is the no jump map for both intervals U and V. It is more clear with the pictures that the new maps we define for the random dynamical system $\{T_U, T_V, T_{UV}, T_*\}$ are countably piecewise expanding. Here are the explicit definitions of the other maps:

$$T_U(x) = \begin{cases} \tau(x) + \frac{3}{16}, & \text{if } x \in [\frac{1}{2}, 1] \text{ and } \tau^R(x) \in U\,, \\ \tau(x), & \text{if } x \in [\frac{1}{2}, 1], \text{ but } \tau(x) \notin U\,, \end{cases}$$

and $T_U(x) = T_U(x^3)$ for $x \in \hat{U}$, and $T_U(x) = T_U(x^3 + \frac{1}{2})$ for $x \in \hat{V}$.

$$T_V(x) = \begin{cases} \tau(x) - \frac{3}{16}, & \text{if } x \in [\frac{1}{2}, 1] \text{ and } \tau(x) \in V\,, \\ \tau(x), & \text{if } x \in [\frac{1}{2}, 1], \text{ but } \tau(x) \notin V\,, \end{cases}$$

and $T_V(x) = T_V(x^3)$ for $x \in \hat{U}$, and $T_V(x) = T_V(x^3 + \frac{1}{2})$ for $x \in \hat{V}$. As one can notice that on $\hat{U} \cup \hat{V}$ each constituent map first takes the point back to $[\frac{1}{2}, 1]$ by jump maps then acts on the point as defined on $[\frac{1}{2}, 1]$.

$$T_*(x) = \begin{cases} \tau(x) - \frac{3}{16}, & \text{if } x \in [\frac{1}{2}, 1]\,, \end{cases}$$

and $T_*(x) = T_*(x^3)$ for $x \in \hat{U}$, and $T_*(x) = T_*(x^3 + \frac{1}{2})$ for $x \in \hat{V}$.

If the probability of jumping right is given to be $p \in (0,1)$ and probability of jumping to left is $q \in (0,1)$ then the probability distribution on the constituent maps $\{T_U, T_V, T_{UV}, T_*\}$ is given by $p(1-q), q(1-p), pq, (1-p)(1-q)$ respectively.

APPENDIX A

Some Equations and Calculations

Proposition A.0.1 (Cauchy's Inequality)**.** *For vectors $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$ we have*

$$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} x_i^2\right)\left(\sum_{i=1}^{n} y_i^2\right)$$

where the equality holds if $x = cy$ for some constant $c \in \mathbb{R}$.

Example A.0.2. *We give an averaged version of Cauchy's Inequality. Our main interest is in fact the equality. So let x, y be a tuple of vector with probability p_1 and a, b be another tuple of vector with probability p_2 where $p_1 + p_2 = 1$. It is easy to* *show that*

$$\left(p_1\sum_{i=1}^n x_iy_i + p_2\sum_{i=1}^n a_ib_i\right)^2 \leq \left(p_1\sum_{i=1}^n x_i^2 + p_2\sum_{i=1}^n a_i^2\right)\left(p_1\sum_{i=1}^n y_i^2 + p_2\sum_{i=1}^n b_i^2\right)$$

by using the original Cauchy's Inequality. If we distribute the terms and combine in a suitable way to check the condition for an equality we see that there exists $c \in \mathbb{R}$ *such that* $\dfrac{x}{y} = \dfrac{a}{b} = c$.

The Berry-Essen's Theorem is stated in different ways, as it is proved independently by two mathematicians, Andrew C. Berry (in 1941) and Carl-Gustav Esseen (1942). We use the notation given in the book of V. V. Petrov, see [Pet85] and give only one inequality which is used to prove the Berry-Essen's Theorem.

Theorem A.0.3 (Essen's Inequality)**.** *Let* $X_1, \ldots, X_n$ *be independent random variables such that* $E[X_j] = 0$, $E|X_j|^3 < \infty$ *for* $j = 1, \ldots, n$. *If* $\sigma_j^2 = E[X_j^2]$, $B_n = \sum_{j=1}^n \sigma_j^2$, *and* $F_n(x) = \mathbb{P}\left\{B_n^{-1/2}\sum_{j=1}^n X_j < x\right\}$, $L_n = B_n^{-3/2}\sum_{j=1}^n E|X_j|^3$ *then*

$$\sup_x |F_n(x) - \Phi(x)| \leq AL_n \tag{A.1}$$

for some constant A, *where* $\Phi(x) = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{x} e^{-t^2/2}dt$.

Now we give some inequalities from probability theory.

Theorem A.0.4 (Chebychev's Inequality)**.** *Let* X *be a random variable in some probability space. Then for any real number* $\lambda \in \mathbb{R}$,

$$\mathbb{P}(|X| \geq \lambda) \leq \frac{\mathbb{E}(X^2)}{\lambda^2}.$$

Theorem A.0.5. *Let $\{X_n\}$ be a nonnegative submartingale and $\lambda > 0$. Then for any $n \geq 0$*

$$\mathbb{P}(\max_{0 \leq k \leq n} \geq \lambda) \leq \frac{\mathbb{E}(X_n)}{\lambda}$$

Lemma A.0.6 (Borel-Cantelli I)**.** *Let (E_n) be a sequence of events in some probability space. If the sum of the probabilities of the events E_n is finite*

$$\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty$$

then the probability that infinitely many of them occur is zero, that is

$$\mathbb{P}\left(\limsup_{n \to \infty} E_n\right) = 0.$$

Lemma A.0.7 (Kač's Lemma)**.** *Let $(\Omega, \mathcal{B}, \mu)$ be a probability space and $A \in \mathcal{B}$ be of positive measure. Let μ_A be the conditional probability measure defined by*

$$\mu_A(B) := \frac{\mu(A \cap B)}{\mu(A)}, \quad \forall B \in \mathcal{B}.$$

Let $T : \Omega \to \Omega$ be a measure preserving map and let $\tau_A(x)$ be the first return time of x to A Then we have

$$\mathbb{E}(\tau_A) := \int_A \tau_A \, d\mu = \mu(\{\tau_A < +\infty\}).$$

In particular, when the system is ergodic we have

$$\mathbb{E}_A(\tau_A) := \int_A \tau_A \, d\mu_A = \frac{1}{\mu(A)}.$$

Bibliography

[ALS09] A. Ayyer, C. Liverani, and M. Stenlund. Quenched CLT for random toral automorphism. *Discrete Contin. Dyn. Syst.*, 24(2):331–348, 2009.

[Arn03] L. Arnold. *Random Dynamical Systems*. Springer, 2003.

[Bal00] V. Baladi. *Positive Transfer Operators and Decay of Correlations*. World Scientific, 2000.

[BG97] A. Boyarsky and P. Gora. *Laws of Chaos : Invariant Measures and Dynamical Systems in One Dimension*. Birkhauser, Boston, Mass., 1997.

[Bow70] R. Bowen. Markov Partitions for Axiom A Diffeomorphisms. *Amer. J. Math.*, 92:725–747, 1970.

[CD09] N. Chernov and D. Dolgopyat. The Galton Board: Limit Theorems and Recurrence. *J. Amer. Soc*, 22:821–858, 2009.

[CG07] J.-R. Chazottes and S. Gouëzel. On almost-sure versions of classical limit theorems for dynamical systems. *Probab. Theory Related Fields*, 138(1-2):195–234, 2007.

[DS09] N. Dunford and J. T. Schwartz. *Linear Operators, Part 1, General Theory*. John Wiley and Sons, 2009.

[Eag76] G. K. Eagleson. Some Simple Conditions for Limit Theorems to be Mixing. *Teor. Veroyatnost. i Primenen.*, 21(3):653–660, 1976.

[GB89] P. Gora and A. Boyarski. Absolutely Continuous Invariant Measures for Piecewise Expanding C^2 Transformations in R^n. *Israel Journal of Mathematics*, 67(3):272–286, 1989.

[HK82] F. Hofbauer and G. Keller. Ergodic Properties of Invariant Measures for Piecewise Monotonic Transformations. *Mathematische Zeitschrift*, 180:119–140, 1982. 10.1007/BF01215004.

[ITM50] C. T. Ionescu Tulcea and G. Marinescu. Theorie Ergodique Pour Des Classes D'Operations Non Completement Continues. *The Annals of Mathematics*, 52(1):pp. 140–147, 1950.

[KB75] U. Krengel and A. Brunel. *Ergodic Theorems.* De Gruyter, Walter, Inc., 1975.

[Kif98] Y. Kifer. *Random Perturbations of Dynamical Systems.* Birkhauser, 1998.

[KL05] G. Keller and C. Liverani. A Spectral Gap for a One-dimensional Lattice of Coupled Piecewise Expanding Interval Maps. *Lecture Notes in Physics*, 671:115–151, 2005.

[Kob05] E. J. Kobre. *Rates of Diffusion in Dynamical Systems with Random Jumps.* PhD thesis, New York University, 2005.

[KY07] E. Kobre and L. S. Young. Extended Systems with Deterministic Local Dynamics and Random Jumps. *Comm. Math. Phys.*, 275(3):709–720, 2007.

[Len06] M. Lenci. Typicality of Recurrence for Lorentz Gases. *Ergodic Theory Dynam. Systems*, 26:799–820, 2006.

[Lit08] C. Little. Absolutely Continuous Invariant Measures for Interval Maps. Master's thesis, Imperial College London, 2008.

[LY73] A. Lasota and J. A. Yorke. On The Existence of Invariant Measures for Piecewise Monotonic Transformations. *Trans. Amer. Math. Soc.*, 186:481–488, 1973.

[Orn74] D. S. Ornstein. *Ergodic Theory, Randomness and Dynamical Systems.* Yale University Press, 1974.

[Pet85] V. V. Petrov. *Sums of Independent Random Variables.* Springer-Verlag, 1985.

[RE83] J. Rousseau-Egele. Un Théorème de la Limite Locale Pour une Classe de Transformations Dilatantes et Monotones par Morceaux. *Ann. Probab.*, 11(3):772–788, 1983.

[Rue78] D. Ruelle. Thermodynamic Formalism. *Addison Wesley, Reading, MA*, 1978.

[Ryc83] M. Rychlik. Bounded Variation and Invariant Measures. *Studia Math.*, 76(1):69–80, 1983.

[Sin68] Ya. G. Sinai. Markov Partitions and C-Diffeomorphisms. *Func. Anal. and its Appl.*, 2:64–89, 1968.

[YK41] K. Yosida and S. Kakutani. Operator-Theoretical Treatment of Markoff's Process and Mean Ergodic Theorem. *The Annals of Mathematics*, 42(1):pp. 188–228, 1941.

[You99] L. S. Young. Recurrence Times and Rates of Mixing. *Israel Journal of Mathematics*, 110(1):153–188, 1999.