B.P. Zeegers

On invariant densities and Lochs' Theorem for random piecewise monotonic interval maps

Master's thesis

Supervisors:
prof. dr. H. Bruin (University of Vienna)
dr. C.C.C.J. Kalle (Leiden University)
prof. dr. E.A. Verbitskiy (Leiden University)

September 28, 2018

Mathematical Institute, Leiden University

Abstract

If m denotes the number of digits in the regular continued fraction expansion that can be determined from n digits in the decimal expansion, then Lochs' Theorem states that the fraction $\frac{m}{n}$ converges Lebesgue almost surely to a fraction of two entropies as $n \to \infty$. These are the entropies of the interval maps that generate these expansions. Lochs' Theorem has been generalized to pairs of interval maps that both belong to a class of piecewise monotonic transformations that generate expansions and that admit an invariant density with suitable ergodic properties. The first aim of this thesis is to review sufficient conditions on interval maps to belong to this class. For this, we first of all recover the famous existence result for invariant densities by Lasota and Yorke for expanding piecewise monotonic interval maps. As an example of a nonexpanding piecewise monotonic interval map, we also consider the Liverani-Saussol-Vaienti (LSV) map and provide a new proof of the already known result that such a map admits an invariant probability density if and only the corresponding parameter lies in $(0, 1)$. Motivated by the practical use of beta encoders, one of the main goals in this thesis is to extend Lochs' Theorem to expansions generated by a class of random piecewise monotonic interval maps. We review sufficient conditions on random interval maps to belong to this class. For two random interval maps T and S in this class, we show that, if m denotes the number of digits in the S-expansion that can be determined from n digits in the T-expansion, then, roughly speaking, the fraction $\frac{m}{n}$ converges Lebesgue almost surely to a fraction of two fiber entropies as $n \to \infty$. As a second important goal, we prove that the skew product of an LSV map with parameter in $(0, 1)$ and another LSV map with parameter in $[1, \infty)$ and with underlying Bernoulli shift admits an invariant probability density.

Contents

Chapter 1

Introduction

1.1 Motivation

1.1.1 Lochs' Theorem

It is a common known fact that each real number $x \in [0,1)$ has a decimal expansion

$$x = \sum_{k=1}^{\infty} \frac{d_k}{10^k}, \qquad d_k = d_k(x) \in \{0, 1, \ldots, 9\} \text{ for } k \in \mathbb{N}, \tag{1.1}$$

which is denoted as $x = 0.d_1 d_2 \ldots$ usually. Such a representation of x is unique, except for some real numbers for which the tail of the sequence can be expressed either with trailing 0's or 9's. We can generate the decimal expansions by iterating the decimal map $T : [0,1) \to [0,1)$ given by

$$Tx = 10x - d_1(x), \tag{1.2}$$

where

$$d_1(x) = \begin{cases} 0 & \text{if } x \in \left[0, \frac{1}{10}\right), \\ 1 & \text{if } x \in \left[\frac{1}{10}, \frac{2}{10}\right), \\ \vdots & \quad \vdots \\ 9 & \text{if } x \in \left[\frac{9}{10}, 1\right) \end{cases} \tag{1.3}$$

(see Figure 1.1). Indeed, rewriting (1.2) gives $x = \frac{d_1(x)}{10} + \frac{Tx}{10}$, and setting $d_n = d_n(x) = d_1(T^{n-1}x)$ for each $n \geq 1$ gives after n iterations

$$x = \frac{d_1}{10} + \frac{d_2}{10^2} + \cdots + \frac{d_n}{10^n} + \frac{T^n x}{10^n}. \tag{1.4}$$

Since $0 \leq T^n x < 1$, we obtain

$$\sum_{k=1}^{n} \frac{d_k}{10^k} \to x \qquad \text{as } n \to \infty, \tag{1.5}$$

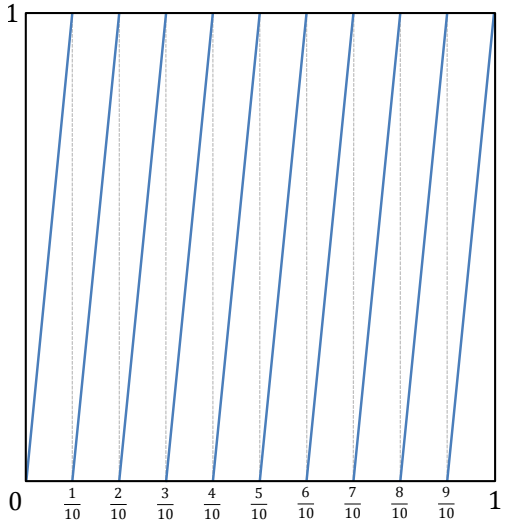

Figure 1.1: The decimal map T

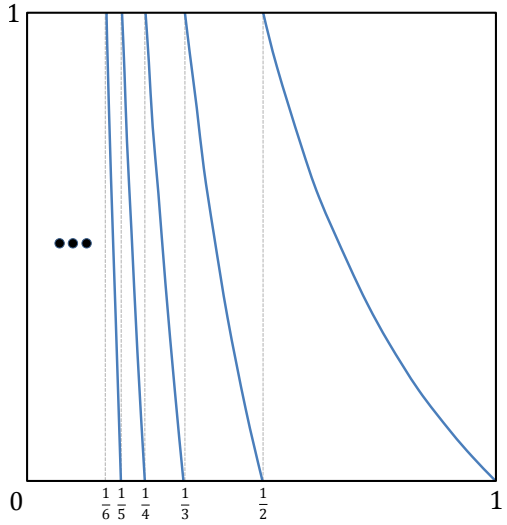

Figure 1.2: The Gauss map S

which is the decimal expansion of x in (1.1).[1]

Besides decimal expansions there are many more possible representations of real numbers in terms of a sequence of integers. As a second example, it is known (see e.g. [18]) that each irrational $x \in (0,1)$ can be represented in a unique way as

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}}, \qquad a_k = a_k(x) \in \mathbb{N} \text{ for each } k \in \mathbb{N}, \tag{1.6}$$

which is referred to as the *regular continued fraction* (RCF) expansion of x. These expansions can be generated from iterating the *Gauss map* $S : [0,1) \to [0,1)$ given by $S0 = 0$ and for $x \neq 0$

$$Sx = \frac{1}{x} \bmod 1 = \frac{1}{x} - a_1(x), \tag{1.7}$$

where

$$a_1(x) = \begin{cases} 1 & \text{if } x \in \left(\frac{1}{2}, 1\right), \\ n & \text{if } x \in \left(\frac{1}{n+1}, \frac{1}{n}\right] \text{ and } n \geq 2 \end{cases} \tag{1.8}$$

(see Figure 1.2). Namely, setting $a_n = a_n(x) = a_1(S^{n-1}x)$ for each $n \geq 1$, we obtain from (1.7) that

$$x = \frac{1}{a_1 + Sx} = \cdots = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n + S^n x}}}. \tag{1.9}$$

[1]The map T does not generate expansions with trailing 9's. Instead, defining d_1 in (1.3) as $d_1(x) = i$ if $x \in (\frac{i}{10}, \frac{i+1}{10}]$ with $i \in \{0, \ldots, 9\}$ yields expansions with no trailing 0's.

From this it can be shown (see e.g. Section 1.3 in [18]), writing

$$[0; a_1, \ldots, a_n] = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \ddots + \cfrac{1}{a_n}}}, \tag{1.10}$$

that

$$[0; a_1, \ldots, a_n] \to x \qquad \text{as } n \to \infty, \tag{1.11}$$

which is the RCF expansion of x in (1.6).[2]

It is natural to ask which of the two previous expansions is more efficient at representing real numbers. In other words, which of the two sequences in (1.5) and (1.11) converges faster to x as $n \to \infty$? The following question is related to this problem: Suppose we know only the first n decimal digits of an unknown irrational number $x \in (0, 1)$. How many digits in the RCF expansion of x does this determine? In 1964, Lochs [46] proved a surprising and elegant result answering this question for the limit $n \to \infty$.

More precisely, for each irrational $x \in (0, 1)$ and $n \in \mathbb{N}$, let $y_n = \sum_{k=1}^{n} \frac{d_k(x)}{10^k}$ and $z_n = y_n + 10^{-n}$. Then the interval $A_n(x) = [y_n, z_n)$ consists of x and all other real numbers of which the decimal expansion starts with the string $d_1(x), \ldots, d_n(x)$. Similarly, for each $m \in \mathbb{N}$, let $r_m = [0; a_1(x), \ldots, a_m(x)]$ and $s_m = [0; a_1(x), \ldots, a_{m-1}(x), a_m(x) + 1]$. Then, for m even (resp. m odd), one can derive that the interval $B_m(x) = [r_m, s_m)$ (resp. $B_m(x) = (s_m, r_m]$) consists of x and all other real numbers of which the RCF expansion starts with the string $a_1(x), \ldots, a_m(x)$. Putting

$$m(n, x) = \sup\{m \in \mathbb{N} : A_n(x) \subseteq B_m(x)\}, \tag{1.12}$$

Lochs proved [46] that, for Lebesgue almost every irrational $x \in (0, 1)$,

$$\lim_{n\to\infty} \frac{m(n, x)}{n} = \frac{6 \log 2 \log 10}{\pi^2} = 0.97027 \cdots . \tag{1.13}$$

Therefore, roughly 97 RCF digits are determined by 100 decimal digits. This indicates that the RCF expansion is slightly more efficient compared to the decimal expansion at representing irrational numbers.

1.1.2 Extension to expansions generated by other interval maps

Naturally, one can ask how the result by Lochs in (1.13) can be generalized to any two known expansions of numbers. For this, let us analyse Lochs' result in more detail. It appears that the right-hand side of (1.13) is the fraction of two *entropies*. As we review in Section 2.8, the entropy of a map is a nonnegative constant that measures

[2] The RCF expansion in (1.6) holds for irrational $x \in [0, 1)$ and consists of infinitely many digits a_k. In Section 4 of [31] it is shown that each rational $x \in (0, 1)$ has a *finite* RCF expansion of the form in (1.10). So a real number is rational if and only if it has an RCF expansion that is finite.

the average uncertainty about where the map moves the points in the system. The entropy $h(T)$ of the decimal map T satisfies

$$h(T) = \lim_{n\to\infty} -\frac{1}{n}\log\lambda(A_n(x)) = \log 10, \qquad \lambda\text{-a.e.}, \tag{1.14}$$

where λ denotes the Lebesgue measure on $[0,1)$. As we shall see, this is because (i) λ is *invariant* with respect to T in the sense that for each subinterval $[a,b) \subseteq [0,1)$ we have $\lambda\big(T^{-1}[a,b)\big) = \lambda([a,b))$,[3] which we see from

$$\lambda\big(T^{-1}[a,b)\big) = \lambda\left(\bigcup_{i=0}^{9}\Big[\frac{a+i}{10},\frac{b+i}{10}\Big)\right) = \sum_{i=0}^{9}\lambda\Big(\Big[\frac{a+i}{10},\frac{b+i}{10}\Big)\Big) = b-a, \tag{1.15}$$

and (ii) λ is *ergodic* with respect to T, meaning that $T^{-1}A = A$ implies $\lambda(A) \in \{0,1\}$ for each Borel set $A \subseteq [0,1)$. (For a proof of (ii), see e.g. [18]) On the other hand, it is easy to see that λ is not invariant with respect to the Gauss map S. However, it can be shown (see e.g. [18]) that the *Gauss measure* μ_G on $[0,1)$ given by

$$\mu_G(A) = \int_A \frac{1}{\log 2}\frac{1}{1+x}dx, \qquad A \subseteq [0,1) \text{ Borel} \tag{1.16}$$

is invariant with respect to S, and moreover that μ_G is ergodic with respect to S. As a consequence, one can derive (see e.g. [18]) that the entropy $h(S)$ of S satisfies

$$h(S) = \lim_{n\to\infty} -\frac{1}{n}\log\mu_G(B_n(x)) = \frac{\pi^2}{6\log 2}, \qquad \mu_G\text{-a.e.} \tag{1.17}$$

In [16], Dajani and Fieldsteel generalize Lochs' Theorem to expansions which are generated by surjective interval maps $R : [0,1) \to [0,1)$ that satisfy the following conditions:

1. There exists a finite or countable partition of $[0,1)$ into intervals such that R restricted to each interval is strictly monotonic and continuous,
2. There exists a Borel probability measure μ on $[0,1)$ that is invariant and ergodic w.r.t. R and is absolutely continuous w.r.t. λ such that

$$\exists M > 0 : \frac{1}{M} \le \frac{d\mu}{d\lambda} \le M. \tag{1.18}$$

Then, if R_1 and R_2 are any two such maps, it is shown in [16] that the number of digits $m_{R_1,R_2}(n,x)$ in the R_2-expansion of x that can be determined from knowing the first n digits in the R_1-expansion of x satisfies

$$\lim_{n\to\infty}\frac{m_{R_1,R_2}(n,x)}{n} = \frac{h(R_1)}{h(R_2)}, \qquad \lambda\text{-a.e.}, \tag{1.19}$$

where $h(R_1)$ (resp. $h(R_2)$) denotes the entropy of R_1 (resp. R_2). It is clear that the decimal expansions generated by T and the RCF expansions generated by S belong

[3] As we shall see in Section 2.1, this is equivalent to how T-invariance of λ is defined in Definition 2.1.

to this class, and it appears (see e.g. [16]) that almost all known expansions on $[0,1)$ generated by an interval map are members of this class.

1.1.3 Extension to expansions generated by random interval maps

So far we considered expansions generated by iterating points under a *single* interval map. Instead, let us now consider a *family* of interval maps $\{T_j : [0,1) \to [0,1)\}_{j\in E}$ where E is some index set. For given $\omega = (\omega_1, \omega_2, \ldots) \in E^{\mathbb{N}}$ and $x \in [0,1)$, we then consider the orbit

$$x \mapsto T_{\omega_1}x \mapsto T_{\omega_2}T_{\omega_1}x \mapsto T_{\omega_3}T_{\omega_2}T_{\omega_1}x \mapsto \ldots. \tag{1.20}$$

In other words, at time n we apply the transformation T_{ω_n} determined by the choice of $\omega \in E^{\mathbb{N}}$, and if we put a non-trivial probability measure $\mathbb{P}$ on $E^{\mathbb{N}}$ we can interpret (1.20) as iterating points under a *random* system of interval maps.

Let us consider an example for which orbits as in (1.20) generate expansions of points for each $\omega \in E^{\mathbb{N}}$. For that, let $E \subseteq (1,\infty)$ such that $\gamma = \inf E > 1$, and define for each $\beta \in E$ the map $T_\beta : [0,1) \to [0,1)$ as

$$T_\beta x = \beta x \bmod 1 = \beta x - b(\beta, x), \tag{1.21}$$

where

$$b(\beta,x) = \begin{cases} i & \text{if } x \in \left[\frac{i}{\beta}, \frac{i+1}{\beta}\right) \text{ and } i \in \{0,1\ldots,\lfloor\beta\rfloor - 1\}, \\ \lfloor\beta\rfloor & \text{if } x \in \left[\frac{\lfloor\beta\rfloor}{\beta}, 1\right). \end{cases} \tag{1.22}$$

Fix $(\beta_1, \beta_2, \ldots) \in E^{\mathbb{N}}$. For each $x \in [0,1)$, define $b_1(x) = b(\beta_1, x)$ and $b_k(x) = b(\beta_k, T_{\beta_{k-1}} \cdots T_{\beta_1}x)$. Then (1.21) gives $x = \frac{b_1}{\beta_1} + \frac{T_{\beta_1}x}{\beta_1}$. Similarly, $T_{\beta_1}x = \frac{b_2}{\beta_2} + \frac{T_{\beta_2}T_{\beta_1}x}{\beta_2}$ and after n iterations we see that

$$x = \frac{b_1}{\beta_1} + \frac{b_2}{\beta_1\beta_2} + \cdots + \frac{b_n}{\beta_1\cdots\beta_n} + \frac{T_{\beta_n}\cdots T_{\beta_1}x}{\beta_1\cdots\beta_n}. \tag{1.23}$$

We have $\frac{T_{\beta_n}\cdots T_{\beta_1}x}{\beta_1\cdots\beta_n} \le \frac{1}{\gamma^n} \to 0$ as $n \to \infty$, so for each $x \in [0,1)$ we obtain the expansion

$$x = \sum_{k=1}^{\infty} \frac{b_k}{\beta_1\cdots\beta_k}, \qquad b_k = b_k(x) \in \{0,1,\ldots,\lfloor\beta_k\rfloor\} \text{ for each } k \in \mathbb{N}. \tag{1.24}$$

As a motivation to generalize Lochs' Theorem to expansions such as (1.24) that are generated by a system of interval maps, let us consider a practical example. It is well known that each real number $x \in [0,1)$ has a binary expansion

$$x = \sum_{k=1}^{\infty} \frac{a_k}{2^k}, \qquad a_k = a_k(x) \in \{0,1\} \text{ for } k \in \mathbb{N}. \tag{1.25}$$

Just like the decimal expansion, such a representation is for each x essentially unique and can be generated by iterating the map $Tx = 2x \bmod 1$. On the other hand, for

$\beta > 1$ non-integer, it is known (see [14, 24, 62]) that Lebesgue almost every $x \in [0,1)$ has a *continuum* number of *β-expansions* of the form

$$x = \sum_{k=1}^{\infty} \frac{b_k}{\beta^k}, \qquad b_k = b_k(x) \in \{0, 1, \ldots, \lfloor \beta \rfloor\} \text{ for each } k \in \mathbb{N}. \tag{1.26}$$

Two well-known transformations that generate such representations are the so-called greedy β-transformation in (1.21) and the lazy β-transformation (see e.g. [17]), and a way to obtain other representations in a dynamical way is by superimposing these two transformations (see e.g. [17] for details). So-called beta encoders exploit the redundancy of the representations of the form (1.26) to encode information in analog-to-digital conversion more robustly compared to using binary expansions [67]. In practice, however, due to noise the value of β tends to vary while iterating. So if, for example, points are iterated under the greedy β-transformation from (1.21), we get in practice expansions of the form in (1.24) instead of (1.26). It is therefore relevant to ask how much information (e.g. in terms of the binary digits in (1.25)) can be determined once we know n digits of the expansion in (1.24). An extension of Lochs' Theorem to expansions generated by random interval maps would be helpful to address this problem.

1.2 Thesis Overview

In the next chapter we discuss the concepts and results in Ergodic Theory that will be relevant for the rest of this thesis. Prior knowledge of Ergodic Theory is not required, but we assume the reader has a basic understanding of Measure Theory.

Motivated by the extension of Lochs Theorem in [16] discussed in Subsection 1.1.2, we review in Chapter 3 results on the existence of *invariant densities* for piecewise monotonic transformations on the unit interval I: For a measurable transformation $T : I \to I$, we say $h \in L^1(\lambda)$ (with λ the Lebesgue measure on I) is an invariant density for T if

$$\int_A h d\lambda = \int_{T^{-1}A} h d\lambda, \qquad \text{for all } A \subseteq I \text{ Borel.} \tag{1.27}$$

We also say that in this case the measure μ on I given by $\mu(A) = \int_A h d\lambda$ is an *absolutely continuous invariant measure* (acim for short, or acipm if μ is a probability measure) for T. We review in Section 3.3 the famous result by Lasota and Yorke [43] that a transformation $T : I \to I$ which is piecewise C^2 and monotonic with respect to some finite partition and is *expanding* (i.e. $\inf_{x \in I} |T'(x)| > 1$) admits an invariant probability density. Moreover, we discuss in Section 3.4 among other results that an expanding piecewise monotonic interval map T admits nonzero but finitely many *ergodic* acipm's (originally proven in [44]). We shall see in Section 3.5, if we furthermore assume that T admits a suitable covering property as for example in the so-called Folklore Theorem,

that T in that case admits a *unique* acipm μ that moreover satisfies (1.18) and is ergodic.

As an example of a *nonexpanding* piecewise monotonic interval map, we consider in Section 3.6 the *Liverani-Saussol-Vaienti (LSV) map* $T_\alpha : I \to I$ with parameter $\alpha \in (0, \infty)$, defined as

$$T_\alpha(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & x \in [0, \frac{1}{2}], \\ 2x - 1 & x \in (\frac{1}{2}, 1] \end{cases} \tag{1.28}$$

(see Figure 3.4). It is well-known that T_α admits an acipm if $\alpha \in (0, 1)$ (see [45]) and an infinite *σ-finite acim* if $\alpha \geq 1$ (see e.g. [55]). We provide a new proof of these results by considering the expanding piecewise monotonic transformation obtained from T_α by inducing w.r.t. the *first passage time* in the interval $(\frac{1}{2}, 1]$. This method is based on Section 3 in [35].

As opposed to the deterministic setting in Chapter 3, we consider in Chapter 4 the setting in which points are iterated under a *random* piecewise monotonic interval map. Such a system is given by a family T of piecewise (sufficiently smooth) monotonic transformations on I and a probability law that describes which of these maps is chosen at each time step. In Sections 4.1-4.3, we take $T = \{T_j : I \to I\}_{j \in E}$ with E a Polish space (i.e. complete, separable metric space) which we assume to be countable most of the time, and we put a non-trivial probability measure $\mathbb{P}$ on the Borel sets in $E^{\mathbb{N}}$. We then consider the skew product

$$F_{\sigma,T} : E^{\mathbb{N}} \times I \to E^{\mathbb{N}} \times I, \quad (\omega, x) \mapsto (\sigma\omega, T_{\omega_1} x), \tag{1.29}$$

where $\omega = (\omega_1, \omega_2, \ldots)$ and $\sigma : E^{\mathbb{N}} \to E^{\mathbb{N}}$ is the left shift on $E^{\mathbb{N}}$, i.e. $\sigma\omega = (\omega_2, \omega_3, \ldots)$. Then iterating points (ω, x) under $F_{\sigma,T}$ yields (after projecting on I) random orbits of the form in (1.20). Similar as in Chapter 3, we review results on the existence of an *invariant density* $h \in L^1(\mathbb{P} \otimes \lambda)$ for $F_{\sigma,T}$, meaning in this case that

$$\int_A h d\mathbb{P} \otimes \lambda = \int_{F_{\sigma,T}^{-1} A} h d\mathbb{P} \otimes \lambda, \qquad \text{for all } A \subseteq E^{\mathbb{N}} \times I \text{ measurable.} \tag{1.30}$$

In Section 4.2 we discuss this for the setting that $\mathbb{P} = \pi^{\otimes \mathbb{N}}$ with π a probability measure on E. In this case, the map T_{ω_n} applied at time $n \in \mathbb{N}$ is randomly chosen from $\{T_j\}_{j \in E}$ independently from the maps that are applied at the other time points, and according to the same distribution π for all time points. This i.i.d. setting was first studied by Morita [47, 49] and Pelikan [53], who independently showed that there exists an invariant density for $F_{\sigma,T}$ if E is at most countable and the system is *expanding on average* in the sense that

$$\sum_{j \in E} \frac{\pi(j)}{\inf_{x \in I} |T_j'(x)|} < 1. \tag{1.31}$$

We review this in Section 4.2 as well as some ergodic properties of these invariant densities similar to the deterministic setting. Moreover, we consider in Section 4.4 an

extension of these results to the case that $\mathbb{P}$ is described by a Markov chain and review results from [40] and [28]. For both the i.i.d. case and Markov case, we shall see that if $\{T_j\}_{j\in E}$ satisfies a suitable random covering property, then an invariant density h for $F_{\sigma,T}$, if it exists, is (up to normalization) unique and satisfies

$$\exists M > 0 : \frac{1}{M} \leq h \leq M. \tag{1.32}$$

In Section 4.3 we consider the random i.i.d. compositions of two LSV maps T_α and T_β given by (1.28), where $\alpha \in (0,1)$ and $\beta \geq 1$. Letting $p \in (0,1)$ and setting $E = \{\alpha, \beta\}$ with $\pi(\alpha) = p$ and $\pi(\beta) = 1-p$, note that (1.31) is not satisfied because $T'_\alpha(0) = T'_\beta(0) = 1$. However, unlike the $p = 0$ case we can still show that there exists an invariant density for the skew product $F_{\sigma,T}$ in this case by generalizing the proof for the deterministic case (i.e. $p = 1$) discussed in [45]. Moreover, we propose a second way to prove this for $p \in (0,1]$ by extending the method of inducing w.r.t. the first passage time from Section 3.6.

In the last section of Chapter 4, we consider the setting where $(\Omega, \mathcal{F}, \mathbb{P})$ is some abstract probability space and $T = \{T_\omega : I \to I\}_{\omega\in\Omega}$ is a family of piecewise monotonic interval maps. Furthermore, we let $\varphi : \Omega \to \Omega$ be measurable and *invertible* (as opposed to the left shift σ on $E^{\mathbb{N}}$) and consider the skew product

$$F_{\varphi,T} : \Omega \times I \to \Omega \times I, \quad (\omega, x) \mapsto (\varphi\omega, T_\omega x). \tag{1.33}$$

We review the result by Buzzi [12] that if the system is in a certain way expanding on average (w.r.t. $\mathbb{P}$), then (under some additional assumptions on φ and T) there exists an invariant density $h \in L^1(\mathbb{P} \otimes \lambda)$ for $F_{\varphi,T}$ (in the sense of (1.30), replacing $F_{\sigma,T}$ and $E^{\mathbb{N}}$ with $F_{\varphi,T}$ and Ω, respectively). Furthermore, we discuss the result in [13] that this is (up to normalization) the only invariant density for $F_{\varphi,T}$ if T in addition satisfies a suitable covering property.

For skew products of the form in (1.33), Abramov and Rokhlin [1] introduced the notion of *fiber entropy*. For this, they assume there exists a Borel probability measure ρ on (in this case) I that for $\mathbb{P}$-a.a. $\omega \in \Omega$ is invariant w.r.t. T_ω, i.e. $\rho(T_\omega^{-1}A) = \rho(A)$ for all $A \subseteq I$ Borel. In Chapter 5 we generalize this to the weaker assumption that there exists a family of finite Borel measures $\{\rho_\omega\}_{\omega\in\Omega}$ on I that is *equivariant w.r.t.* (T, φ), meaning that

$$\rho_\omega(T_\omega^{-1}A) = \rho_{\varphi(\omega)}(A) \qquad \text{for all } A \subseteq I \text{ Borel} \tag{1.34}$$

for $\mathbb{P}$-a.a. $\omega \in \Omega$. In Sections 5.1 we give conditions under which such a family $\{\rho_\omega\}_{\omega\in\Omega}$ exists. In particular, we shall see that for each invariant density $h \in L^1(\mathbb{P} \otimes \lambda)$ for $F_{\varphi,T}$ the family $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A h(\omega, x) d\lambda(x)$ is equivariant w.r.t. (T, φ) if φ is invertible. We define the fiber entropy in Section 5.2 whose construction is similar to the construction of the "ordinary" (Kolmogorov-Sinai) entropy that we review in Section 2.8. Moreover, we give in Sections 5.3 and 5.4 the analogous theorems for

fiber entropy of the classical Kolmogorov-Sinai Theorem and the Shannon-McMillan-Breiman Theorem that we review in Section 2.8 as well.

In Sections 6.2 and 6.3 we provide the proof from [16] that shows the extension (1.19) of Lochs' Theorem to the piecewise monotonic interval maps considered in Subsection 1.1.2. We shall see that this proof is based on the Shannon-McMillan-Breiman Theorem, a result from which e.g. (1.14) and (1.17) (concerning the original Lochs Theorem) follow. Also, we discuss that any two piecewise monotonic interval maps studied in Section 3.5 satisfy (1.19), and we consider a central limit result from [32] associated with (1.19).

Finally, in Section 6.4 we formulate and prove a generalization of Lochs' Theorem to a class of random piecewise monotonic interval maps being of the form as in Chapter 4. We suppose that a member of this class has an invariant density h for the corresponding skew product (either of the form (1.29) or (1.33)) such that h satisfies (1.32). We shall see that this makes the class a natural generalization of the class of deterministic interval maps considered in Subsection 1.1.2. For two random interval maps T and S in this class, we show that, if m denotes the number of digits in the S-expansion that can be determined from n digits in the T-expansion, then, roughly speaking, the fraction $\frac{m}{n}$ converges with probability 1 to a fraction of fiber entropies as $n \to \infty$. Furthermore, we shall consider a corresponding central limit result.

Chapter 2

Preliminaries from Ergodic Theory

In this chapter we give a short introduction to Ergodic Theory and discuss the concepts that will be relevant for the rest of this thesis. Included are some proofs for convenience of the reader and some common examples. However, we refer to [10, 15] and Chapter 3 of [9] for a detailed and more complete introduction to Ergodic Theory.

In short, Ergodic Theory studies the long-term average behavior of systems over time. The states of the system under consideration form a space X, which we assume to be a probability space $(X, \mathcal{B}, \mu)$, and the evolution is given by a measurable tranformation $T : X \to X$. Furthermore, we usually suppose that the evolution T is *measure preserving*, which is the topic of the next section.

2.1 Measure Preserving Transformations

Definition 2.1. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be measurable. The map T is called* measure preserving *with respect to μ if $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{B}$. In this case we also say that μ is* invariant *with respect to T.*

The following proposition is very useful for verifying if a transformation is measure preserving.

Proposition 2.2. (see e.g. Theorem 2.1.2 in [9]) *Let T be a measurable transformation on a probability space $(X, \mathcal{B}, \mu)$. Let $\mathcal{A} \subseteq \mathcal{B}$ be a π-system that generates $\mathcal{B}$. If $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{A}$, then T is measure preserving with respect to μ.*

Example 2.3. Since the collection of all subintervals $[a, b) \subseteq [0, 1)$ forms a π-system that generates the Borel σ-algebra on $[0, 1)$, we see from (1.15) that the decimal map $Tx = 10x \bmod 1$ is measure preserving with respect to the Lebesgue measure λ on $[0, 1)$. In the same way one can show that for every $N \geq 2$ integer the N-adic transformation $T : [0, 1) \to [0, 1)$ given by $Tx = Nx \bmod 1$ is measure preserving with respect to λ.

Example 2.4. The Gauss map S from (1.7) is measure preserving with respect to the Gauss measure μ_G from (1.16) (see e.g. [18]).

Example 2.5. (Bernoulli shifts) Let $E \subseteq \mathbb{N}$, and let $\Omega_E = E^{\mathbb{N}}$ (or $\Omega_E = E^{\mathbb{Z}}$) be the space of one-sided (or two-sided) sequences in E. Furthermore, let $\mathcal{F}$ be the σ-algebra on Ω_E generated by all *cylinder sets* $\{\omega \in \Omega_E : \omega_i = z_i, \omega_{i+1} = z_{i+1}, \ldots, \omega_{i+n} = z_{i+n}\}$ where $i \in \mathbb{N}$ (or $\mathbb{Z}$) and $z_i, \ldots, z_{i+n} \in E$. We consider a probability vector $p = (p_j)_{j \in E}$, i.e. $p_j \geq 0$ for all $j \in E$ and $\sum_{j \in E} p_j = 1$. By Carathéodory's Extension Theorem, we obtain a measure $\mathbb{P}$ on $\mathcal{F}$ by specifying $\mathbb{P}$ on the cylinders as

$$\mathbb{P}(\{\omega \in \Omega_E : \omega_i = z_i, \ldots, \omega_{i+n} = z_{i+n}\}) = p_{z_i} \cdots p_{z_{i+n}}. \tag{2.1}$$

Let $\sigma : \Omega_E \to \Omega_E$ be the left-shift on Ω_E, i.e. $\sigma\omega = \tilde{\omega}$ where $\tilde{\omega}_n = \omega_{n+1}$. It is easy to verify that σ is measure preserving w.r.t. $\mathbb{P}$ by applying Proposition 2.2 to the collection of all cylinder sets, which is a π-system on Ω_E.

Example 2.6. (Markov shifts) Let $(\Omega_E, \mathcal{F}, \sigma)$ be as in the previous example. We assume E is finite, say $E = \{1, \ldots, r\}$. Let $W = (W_{ij})$ be a stochastic $r \times r$ matrix, and $q = (q_1, \ldots, q_r)$ a probability vector such that $qW = q$. Again by Carathéodory's Extension Theorem, we obtain a measure $\mathbb{P}$ on $\mathcal{F}$ by specifying $\mathbb{P}$ on the cylinders as

$$\mathbb{P}(\{\omega \in \Omega_E : \omega_i = z_i, \ldots, \omega_{i+n} = z_{i+n}\}) = q_{z_i} W_{z_i z_{i+1}} \cdots W_{z_{i+n-1} z_{i+n}}. \tag{2.2}$$

Again, one can derive that the left-shift σ on Ω_E is measure preserving w.r.t. $\mathbb{P}$ by applying Proposition 2.2 to the π-system consisting of all cylinder sets.

The following theorem gives an equivalent formulation of Definition 2.1.

Theorem 2.7. (see e.g. Theorem 3.1.2 in [9]) *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ measurable. Then T is measure preserving with respect to μ if and only if*

$$\int_X f d\mu = \int_X f \circ T d\mu \tag{2.3}$$

for any $f \in L^1(\mu)$.

2.2 Ergodicity

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be measurable. Suppose $T^{-1}B = B$ for some $B \in \mathcal{B}$. Then $T^{-1}(X \backslash B) = X \backslash B$, so the behavior of T splits into $T|_B$ and $T|_{X \backslash B}$. In the following definition, T is indecomposable μ-a.e.

Definition 2.8. *Let T be a measurable transformation on a probability space $(X, \mathcal{B}, \mu)$. Then T is said to be* ergodic *w.r.t. μ if for every $A \in \mathcal{B}$ such that $T^{-1}A = A$ we have $\mu(A) \in \{0, 1\}$. In this case, we also say that the pair (T, μ) is ergodic.*

It can be shown (see e.g. [15]) that the N-adic transformations from Example 2.3 are ergodic w.r.t. λ and that the Gauss map from Example 2.4 is ergodic w.r.t. the Gauss

measure. Moreover, the left shift σ from Example 2.5 is ergodic w.r.t. the probability measure given by (2.1) and σ is ergodic w.r.t. the probability measure given by (2.2) if and only if the Markov chain defined by the stochastic matrix W is irreducible (see e.g. Theorem 7.2.8 in [65]).

In general, it is difficult to determine ergodicity from Definition 2.8. In some cases, the following theorem may be useful.

Theorem 2.9. (see e.g. Theorem 3.2.3 in [9]) *Let $(X, \mathcal{F}, \mu)$ be a probability space, and $T : X \to X$ be measure preserving w.r.t. μ. The following statements are equivalent:*

(i) (T, μ) is ergodic,

(ii) If $f : X \to \mathbb{C}$ is measurable and $(f \circ T)(x) = f(x)$ for μ-a.e. x, then f is constant μ-a.e.,

(iii) If $f \in L^2(\mu)$ with $(f \circ T)(x) = f(x)$ for μ-a.e. x, then f is constant μ-a.e.

Theorem 2.10. *Let μ_1 and μ_2 be probability measures on a measurable space $(X, \mathcal{B})$, and let $T : X \to X$ be measure preserving with respect to μ_1 and μ_2.*

1. *If (T, μ_1) is ergodic and μ_2 is absolutely continuous w.r.t. μ_1, then $\mu_1 = \mu_2$.*
2. *If (T, μ_1) and (T, μ_2) are ergodic, then either $\mu_1 = \mu_2$ or μ_1 and μ_2 are mutually singular.*

Proof: For the proof of the first part we refer to Lemma 3.2.5 in [9] and for the proof of the second part we refer to Theorem 3.2.5 in [9]. □

2.3 Birkhoff's Ergodic Theorem

Let T be a measurable transformation on a probability space $(X, \mathcal{B}, \mu)$. For a nontrivial $A \in \mathcal{B}$, we can ask with what frequency the points of an orbit $\{x, Tx, T^2x, \ldots\}$ occur in A. Birkhoff's (Pointwise) Ergodic Theorem indicates the asymptotic behavior of the relative frequency $\frac{1}{n}\sum_{i=0}^{n-1} 1_A(T^i x)$ of points of $\{x, Tx, T^2x, \ldots\}$ in A.

Theorem 2.11. (Birkhoff's Ergodic Theorem) *Let T be a measure preserving transformation on a probability space $(X, \mathcal{B}, \mu)$. Then for any $f \in L^1(\mu)$,*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = f^*(x) \tag{2.4}$$

exists μ-a.e. and satisfies $f^* \circ T = f^*$ μ-a.e., and $\int_X f^* d\mu = \int_X f d\mu$. If furthermore (T, μ) is ergodic, then $f^* = \int_X f d\mu$ is constant μ-a.e.

Birkhoff's Ergodic Theorem is widely used and there are different proofs of this very important theorem (see Section 3.3 in [9] and references therein). The last statement of Theorem 2.11 follows from Theorem 2.9.

From Birkhoff's Ergodic Theorem, one can derive (see e.g. [15]) another characterization of ergodicity:

Corollary 2.12. *Let T be a measure preserving transformation on a probability space $(X, \mathcal{B}, \mu)$. Then (T, μ) is ergodic if and only if for all $A, B \in \mathcal{B}$ one has*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \mu(T^{-i}A \cap B) = \mu(A)\mu(B). \tag{2.5}$$

Remark 2.13. To prove ergodicity it suffices to show (2.5) for sets A and B that belong to a semi-algebra $\mathcal{A} \subseteq \mathcal{B}$ that generates $\mathcal{B}$. We refer to [15] for a proof.

Let X be a compact space and $\mathcal{B}$ the Borel sigma-algebra on X. Then the Banach space $C(X)$ of all continuous functions on X (under the supremum norm) is separable, i.e. $C(X)$ has a countable dense subset. In this case, we get the following strengthening of Birkhoff's Ergodic Theorem.

Theorem 2.14. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be measure preserving and ergodic w.r.t. μ. Furthermore, suppose that X is compact and that $\mathcal{B}$ is the Borel σ-algebra on X. Then there exists $Y \in \mathcal{B}$ such that $\mu(Y) = 1$ and*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f(T^i x) = \int_X f(x) d\mu(x) \tag{2.6}$$

for all $x \in Y$ and $f \in C(X)$.

Proof: Let $\{f_k\}_{k\in\mathbb{N}}$ be a countable dense subset $C(X)$. We obtain for each $k \in \mathbb{N}$ from Birkhoff's Ergodic Theorem a set $X_k \in \mathcal{B}$ such that $\mu(X_k) = 1$ and

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} f_k(T^i x) = \int_X f_k(x) d\mu(x) \tag{2.7}$$

for all $x \in X_k$. Taking $Y = \bigcap_{k=1}^{\infty} X_k$, we have $\mu(Y) = 1$ and (2.7) holds for all $x \in Y$ and $k \in \mathbb{N}$. Now, let $f \in C(X)$ and $\varepsilon > 0$. Then there exists $k \in \mathbb{N}$ such that $\|f - f_k\|_\infty < \frac{\varepsilon}{3}$. Let $x \in Y$, and take $N \in \mathbb{N}$ such that $|\frac{1}{n}\sum_{i=0}^{n-1} f_{k_n}(T^i x) - \int_X f_{k_n} d\mu| < \frac{\varepsilon}{3}$ for each $n \geq N$. Then for all $n \geq N$ we have

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{i=0}^{n-1} f(T^i x) - \int_X f d\mu\Big| &\leq \Big|\frac{1}{n}\sum_{i=0}^{n-1} \big[f(T^i x) - f_k(T^i x)\big]\Big| \\
&\quad + \Big|\frac{1}{n}\sum_{i=0}^{n-1} f_k(T^i x) - \int_X f_k d\mu\Big| + \Big|\int_X f_k - f d\mu\Big| \\
&< \varepsilon.
\end{aligned}$$

□

2.4 Mixing and Exactness

We see from Corollary 2.12 that ergodicity means average independence in the long-term. This motivates the following stronger notions of asymptotic independence.

Definition 2.15. *Let T be a measure preserving transformation on a probability space $(X, \mathcal{B}, \mu)$. Then*

1. *(T, μ) is called* weakly mixing *if for all $A, B \in \mathcal{B}$ one has*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \left|\mu(T^{-i}A \cap B) - \mu(A)\mu(B)\right| = 0. \tag{2.8}$$

2. *(T, μ) is called* strongly mixing *if for all $A, B \in \mathcal{B}$ one has*

$$\lim_{n\to\infty} \mu(T^{-n}A \cap B) = \mu(A)\mu(B). \tag{2.9}$$

Remark 2.16. Note that strongly mixing implies weakly mixing and that weakly mixing implies ergodicity. The converses are not true in general.

Remark 2.17. Again, to prove weak mixing (resp. mixing) it suffices (see e.g. [54]) to show (2.8) (resp. (2.9)) for sets A and B that belong to a semi-algebra $\mathcal{A} \subseteq \mathcal{B}$ that generates $\mathcal{B}$.

Example 2.18. Consider $(\Omega_E, \mathcal{F}, \sigma)$ from Example 2.5. For $\mathbb{P}$ given by (2.1), $A = \{\omega \in \Omega_E : \omega_i = z_i, \ldots, \omega_{i+n} = z_{i+n}\}$ and $B = \{\omega \in \Omega_E : \omega_j = w_j, \ldots, \omega_{j+m} = w_{j+m}\}$ it is clear that $\mathbb{P}(\sigma^{-n}A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ for all $n \geq |i| + j + m$. Since the cylinder sets form a semi-algebra that generates $\mathcal{F}$, we conclude that the Bernoulli shift is strongly mixing. Furthermore, if E is finite, then one can show (see e.g. Theorem 5.6 in [54]) that the Markov shift $(\mathbb{P}, \sigma)$ with $\mathbb{P}$ given by (2.2) is mixing if and only the Markov chain defined by the stochastic matrix W is irreducible and aperiodic.

The following notion is even stronger than mixing and is introduced by Rokhlin [58].

Definition 2.19. *Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be measure preserving w.r.t. μ. We say (T, μ) is* exact *if $\bigcap_{n=0}^{\infty} T^{-n}\mathcal{B}$ consists of sets $B \in \mathcal{B}$ such that $\mu(B) \in \{0, 1\}$.*

Proposition 2.20. *Let T be a measure preserving transformation on a probability space $(X, \mathcal{B}, \mu)$. Then (T, μ) is exact if and only if for any $A \in \mathcal{B}$ such that $\mu(A) > 0$ and $T^n A \in \mathcal{B}$ for any $n \geq 0$ we have*

$$\lim_{n\to\infty} \mu(T^n A) = 1. \tag{2.10}$$

Proof: We refer to Theorem 3.4.3 in [9] for the proof that this condition is necessary. Let us show that it is sufficient. Suppose that $A \in \bigcap_{n=0}^{\infty} T^{-n}\mathcal{B}$, i.e. for each $n \geq 0$ there exists $A_n \in \mathcal{B}$ such that $A = T^{-n}A_n$. Then $T^n A \subseteq A_n$, which gives $T^{-n}(T^n A) \subseteq A$. Also, it holds for all $B \subseteq X$ that $B \subseteq T^{-n}(T^n B)$, so $A = T^{-n}(T^n A)$ and thus $\mu(A) = \mu(T^{-n}(T^n A)) = \mu(T^n A)$. So $\mu(A) > 0$ implies $\mu(A) = \lim_{n\to\infty} \mu(T^n A) = 1$. $\square$

Remark 2.21. One can show that exactness implies strongly mixing. However, the converse is not true in general.

Remark 2.22. A measurable transformation T on a probability measure $(X, \mathcal{B}, \mu)$ that is invertible (i.e. T is one-to-one and T^{-1} is measurable) cannot be exact. Indeed, in this case for all $A \in \mathcal{B}$ with $\mu(A) < 1$ we have $\mu(T^n A) = \mu(T^{-n}T^n A) = \mu(A) < 1$ for all $n \in \mathbb{N}$.

Example 2.23. The one-sided Bernoulli shift from Example 2.5 is exact. Also, the one-sided Markov shift from Example 2.6 is exact if and only if the Markov chain defined by the stochastic matrix W is irreducible and aperiodic. (For a proof, see the solution of Exercise 9.5.5 in [65].)

2.5 The Koopman Operator

Let $(X, \mathcal{B}, \mu)$ be a probability space. Recall that the space $L^2(\mu)$ of complex-valued square-integrable functions is a Hilbert space w.r.t. the inner product

$$\langle f, g \rangle = \int_X f \bar{g} d\mu. \tag{2.11}$$

A measurable transformation $T : X \to X$ induces an operator $U_{T,\mu} : L^2(\mu) \to L^2(\mu)$ defined by

$$U_{T,\mu} f = f \circ T, \tag{2.12}$$

which is called the *Koopman operator* for T. We now give some results regarding the relation between the spectrum of $U_{T,\mu}$ and the ergodic properties of T w.r.t. μ.

Recall that $\lambda \in \mathbb{C}$ is an eigenvalue for $U_{T,\mu}$ if there exists a nonzero $f \in U_{T,\mu}$ such that $U_{T,\mu} f = \lambda f$. Note that $\lambda = 1$ is always an eigenvalue for any constant function.

Proposition 2.24. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then an eigenvalue λ of $U_{T,\mu}$ satisfies $|\lambda| = 1$. Furthermore, if $\lambda \neq 1$ is an eigenvalue for $U_{T,\mu}$ corresponding to an eigenfunction $f \in U_{T,\mu}$, then $\int_X f d\mu = 0$.*

Proof: Let $\lambda \in \mathbb{C}$ and $f \in U_{T,\mu}$ nonzero such that $U_{T,\mu} f = \lambda f$. Then the first claim follows from

$$\langle f, f \rangle = \langle U_{T,\mu} f, U_{T,\mu} f \rangle = \langle \lambda f, \lambda f \rangle = |\lambda|^2 \langle f, f \rangle. \tag{2.13}$$

The second statement follows from

$$\lambda \int_X f d\mu = \int_X U_{T,\mu} f d\mu = \int_X f d\mu. \tag{2.14}$$

□

We recall that an eigenvalue is called simple if the corresponding eigenspace is 1-dimensional. Hence, we can reformulate the equivalence (i) $\Leftrightarrow$ (iii) in Theorem 2.9 as follows:

Theorem 2.25. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then T is ergodic w.r.t. μ if and only if 1 is a simple eigenvalue for $U_{T,\mu}$.*

We can characterize weak mixing as follows.

Theorem 2.26. (see e.g. Theorem 3.5.2 in [9]) *Let $T : X \to X$ be measure preserving w.r.t. μ. Then the following statements are equivalent:*

(i) T is weakly mixing w.r.t. μ,

(ii) T is ergodic w.r.t. μ and 1 is the only eigenvalue of $U_{T,\mu}$,

(iii) Every eigenfunction of $U_{T,\mu}$ is constant.

2.6 The Transfer Operator

Let $(X, \mathcal{B}, \mu)$ be a probability space, and $T : X \to X$ be measurable. The transfer operator (or Ruelle-Perron-Frobenius operator) was first introduced in [41, 42] and describes how functions in $L^1(\mu)$ transform under T. We shall see in Chapter 3 that this operator serves as a powerful tool for determining the invariant densities for piecewise monotonic transformations on the unit interval. In this section we define the transfer operator and state its basic properties, and in Section 3.2 we give an explicit formula for the transfer operator for piecewise monotonic interval maps.

Definition 2.27. *Let $T : X \to X$ measurable. We say that T is* nonsingular *w.r.t. μ if and only if for any $A \in \mathcal{B}$ with $\mu(A) = 0$ we have $\mu(T^{-1}A) = 0$.*

Remark 2.28. Note that if $T : X \to X$ is measure preserving w.r.t. μ, then T is nonsingular w.r.t. μ.

Definition 2.29. *Let $T : X \to X$ be nonsingular w.r.t. μ. For any $f \in L^1(\mu)$, write $P_{T,\mu}f$ for the unique element in $L^1(\mu)$ such that, for each $A \in \mathcal{B}$,*

$$\int_A P_{T,\mu} f d\mu = \int_{T^{-1}A} f d\mu. \tag{2.15}$$

We call $P_{T,\mu} : L^1(\mu) \to L^1(\mu)$ the transfer operator *for T.*

Remark 2.30. The existence and uniqueness of $P_{T,\mu}f$ as in the above definition is justified as follows: Consider the measure ν given by

$$\nu(A) = \int_{T^{-1}A} f d\mu, \qquad A \in \mathcal{B}. \tag{2.16}$$

Using the nonsingularity of T, $\mu(A) = 0$ implies $\mu(T^{-1}A) = 0$, which in turn implies $\nu(A) = 0$. This gives $\nu \ll \mu$, so by the Radon-Nikodym Theorem, there exists a unique element in $L^1(\mu)$ denoted as $P_{T,\mu}f$ such that

$$\nu(A) = \int_A P_{T,\mu} f d\mu, \qquad A \in \mathcal{B}. \tag{2.17}$$

Because the Radon-Nikodym Theorem applies as well if ν is a positive σ-finite measure, note that we can extend the definition of $P_{T,\mu}$ to measurable functions $f : X \to [0, \infty]$ for which ρ given by $\rho(A) = \int_A f d\mu$ is σ-finite (in that case, ν given by (2.16) is also σ-finite). This will be relevant in Section 3.6.

The following basic properties of the transfer operator in the next two propositions are easy to show. We refer to Section 4.2 in [9] for the proofs.

Proposition 2.31. *Let $T : X \to X$ be nonsingular w.r.t. μ. Then*

(a) $P_{T,\mu}$ is linear,

(b) The integral is preserved by $P_{T,\mu}$, i.e.

$$\int_X P_{T,\mu} f d\mu = \int_X f d\mu, \qquad f \in L^1(\mu), \tag{2.18}$$

(c) $P_{T,\mu}$ is a positive operator: if $f \in L^1(\mu)$ is such that $f \geq 0$, then $P_{T,\mu} f \geq 0$,

(d) $P_{T,\mu}$ is a contraction on $L^1(\mu)$, i.e.

$$\|P_{T,\mu} f\|_{1,\mu} \leq \|f\|_{1,\mu}, \qquad f \in L^1(\mu). \tag{2.19}$$

Proposition 2.32. *Let $T : X \to X$ and $S : X \to X$ be nonsingular w.r.t. μ. Then $P_{T \circ S,\mu} = P_{T,\mu} \circ P_{S,\mu}$. In particular, $P_{T^n,\mu} = P^n_{T,\mu}$ for each $n \in \mathbb{N}$.*

For a nonsingular transformation $T : X \to X$ w.r.t. μ, the following proposition gives a one-to-one correspondence between the fixed points of $P_{T,\mu}$ and the measures that are T-invariant and absolutely continuous w.r.t. μ.

Proposition 2.33. *Let $T : X \to X$ be nonsingular w.r.t. μ and $h \in L^1(\mu)$. Then $P_{T,\mu} h = h$ if and only if T is measure preserving with respect to the measure ν given by*

$$\nu(A) = \int_A h d\mu, \qquad A \in \mathcal{B}. \tag{2.20}$$

Proof: This follows immediately from

$$\nu(T^{-1}A) = \int_{T^{-1}A} h d\mu = \int_A P_{T,\mu} h d\mu. \tag{2.21}$$

□

Proposition 2.34. *Let $T : X \to X$ be nonsingular w.r.t. μ and ν be the measure given by (2.20) for some $h \in L^1(\mu)$. Then T is nonsingular w.r.t. ν and*

$$P_{T,\nu} f = \frac{P_{T,\mu}(f \cdot h)}{h}, \qquad f \in L^1(\nu) \tag{2.22}$$

where $P_{T,\nu} f$ can be taken as a version in $L^1(\nu)$ such that $P_{T,\nu} f(x) = 0$ whenever $h(x) = 0$.

Proof: Let $A \in \mathcal{B}$ and $f \in \mathcal{L}^1(\mu)$. We have

$$\int_A \frac{P_{T,\mu}(f \cdot h)}{h} d\nu = \int_A P_{T,\mu}(f \cdot h) d\mu = \int_{T^{-1}A} f d\nu. \tag{2.23}$$

Taking $f \equiv 1$, we see that T is nonsingular w.r.t. ν, and (2.22) follows from (2.23). □

Recall that a sequence $\{f_n\}_{n\in\mathbb{N}}$ in $L^1(\mu)$ converges weakly to a function $f \in L^1(\mu)$ if $\int_X f_n g d\mu \to \int_X f g d\mu$ as $n \to \infty$ for each $g \in L^\infty(\mu)$, which we denote as $f_n \overset{w}{\to} f$. On the other hand, we write $f_n \overset{L^1(\mu)}{\to} f$ for convergence w.r.t. the $L^1(\mu)$-norm. Using this we can formulate the concepts of ergodicity, (weak) mixing and exactness in terms of the transfer operator as in the following theorem.

Theorem 2.35. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then T is nonsingular w.r.t. μ, and*

1. *T is ergodic w.r.t. μ if and only if for all $f \in L^1(\mu)$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} P_{T,\mu}^k f \overset{w}{\to} \int_X f d\mu, \quad n \to \infty, \tag{2.24}$$

2. *T is weakly mixing w.r.t. μ if and only if for all $f \in L^1(\mu)$,*

$$\frac{1}{n}\sum_{k=0}^{n-1} \left| P_{T,\mu}^k f - \int_X f d\mu \right| \overset{w}{\to} 0, \quad n \to \infty, \tag{2.25}$$

3. *T is strongly mixing w.r.t. μ if and only if for all $f \in L^1(\mu)$,*

$$P_{T,\mu}^n f \overset{w}{\to} \int_X f d\mu, \quad n \to \infty, \tag{2.26}$$

4. *T is exact w.r.t. μ if and only if for all $f \in L^1(\mu)$,*

$$P_{T,\mu}^n f \overset{L^1(\mu)}{\to} \int_X f d\mu, \quad n \to \infty. \tag{2.27}$$

We refer to Propositions 4.2.10 and 4.2.11 in [9] for a proof of Theorem 2.35.

In case $T : X \to X$ is measure preserving w.r.t. μ, we have the following strengthening of part (d) in Proposition 2.31 (see Corollary 4.2.1 in [9]).

Proposition 2.36. *Let $T : X \to X$ be measure preserving w.r.t. μ. For each $p \in [1, \infty]$, $P_{T,\mu}$ is a contraction on $L^p(\mu)$, i.e.*

$$\|P_{T,\mu} f\|_{p,\mu} \le \|f\|_{p,\mu}, \qquad f \in L^p(\mu). \tag{2.28}$$

In the rest of this section we suppose that μ is T-invariant and consider the restriction of $P_{T,\mu}$ to $L^2(\mu)$, which is possible by the previous proposition.

Proposition 2.37. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then the adjoint of $P_{T,\mu} : L^2(\mu) \to L^2(\mu)$ is the Koopman operator $U_{T,\mu}$.*

Proof: Let $f \in L^2(\mu)$ and set $g = 1_A, A \in \mathcal{B}$. Then

$$\langle P_T f, g\rangle = \int_{T^{-1}A} f d\mu = \int_X f(1_A \circ T) d\mu = \langle f, U_{T,\mu} g\rangle. \tag{2.29}$$

Since the linear combinations of indicator functions are dense in $L^2(\mu)$, one can derive the statement $\langle P_T f, g\rangle = \langle f, U_{T,\mu} g\rangle$ for all $f, g \in L^2(\mu)$. □

Lemma 2.38. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then $P_{T,\mu} U_{T,\mu} f = f$ for all $f \in L^2(\mu)$.*

Proof: For all $f, g \in L^2(\mu)$ we have

$$\langle P_{T,\mu} U_{T,\mu} f, g\rangle = \langle U_{T,\mu} f, U_{T,\mu} g\rangle = \langle f, g\rangle. \tag{2.30}$$

□

Proposition 2.39. *Let $T : X \to X$ be measure preserving w.r.t. μ. Then*

$$U_{T,\mu} f = \lambda f \Leftrightarrow P_{T,\mu} \overline{f} = \lambda \overline{f} \text{ and } |\lambda| = 1 \tag{2.31}$$

for $f \in L^2(\mu)$ and $\lambda \in \mathbb{C}$. In particular, the set of eigenvalues of $U_{T,\mu}$ equals the set of eigenvalues of $P_{T,\mu}$ with modulus 1.

Proof: Suppose $U_{T,\mu} f = \lambda f$. Then $|\lambda| = 1$ and $U_{T,\mu} \overline{f} = \overline{U_{T,\mu} f} = \overline{\lambda f}$, so $\lambda^{-1} = \overline{\lambda}$ and $\lambda U_{T,\mu} \overline{f} = \overline{f}$. This gives $P_{T,\mu} \overline{f} = P_{T,\mu}(\lambda U_{T,\mu} \overline{f}) = \lambda \overline{f}$. Conversely, suppose $P_{T,\mu} \overline{f} = \lambda \overline{f}$ with $|\lambda| = 1$. Then

$$\begin{aligned}
\langle \lambda U_{T,\mu} \overline{f} - \overline{f}, \lambda U_{T,\mu} \overline{f} - \overline{f}\rangle &= |\lambda|^2 \langle U_{T,\mu} \overline{f}, U_{T,\mu} \overline{f}\rangle - \lambda \langle U_{T,\mu} \overline{f}, \overline{f}\rangle - \overline{\lambda} \langle \overline{f}, U_{T,\mu} \overline{f}\rangle + \langle \overline{f}, \overline{f}\rangle \\
&= 2\langle \overline{f}, \overline{f}\rangle - \lambda \overline{\langle P_{T,\mu} \overline{f}, \overline{f}\rangle} - \overline{\lambda} \langle P_{T,\mu} \overline{f}, \overline{f}\rangle \\
&= 2\langle \overline{f}, \overline{f}\rangle - 2|\lambda|^2 \langle \overline{f}, \overline{f}\rangle = 0.
\end{aligned}$$

Hence, $\lambda U_{T,\mu} \overline{f} = \overline{f}$, or equivalently $U_{T,\mu} f = \lambda f$. □

2.7 Measure Preserving Isomorphisms and Lebesgue Spaces

Let $(X, \mathcal{B}, \mu)$ be a probability space and let $T : X \to X$ be measure preserving. We call the quadruple $(X, \mathcal{B}, \mu, T)$ a *dynamical system*. Such a system is characterized by its measure structure given by $(X, \mathcal{B}, \mu)$ modulo sets of measure zero, and by its dynamical structure given by T. For this reason, we have the following definition that classifies two dynamical systems as identical.

Definition 2.40. *Two dynamical systems $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ are* isomorphic *if there exist $N \in \mathcal{B}$ with $\mu(N) = 0$ and $T(X \backslash N) \subseteq X \backslash N$, $M \in \mathcal{C}$ with $\nu(M) = 0$ and $S(Y \backslash M) \subseteq Y \backslash M$, and a measurable and invertible transformation $\psi : X \backslash N \to Y \backslash M$ such that $\psi \circ T = S \circ \psi$ on $X \backslash N$ and $\mu(\psi^{-1} A) = \nu(A)$ for all measurable $A \subseteq Y \backslash M$. The map ψ is called an* isomorphism.

Example 2.41. Let $([0,1), \mathcal{B}, \lambda)$ be the unit interval with associated Borel σ-algebra $\mathcal{B}$ and Lebesgue measure λ. Let $N \in \mathbb{N}$, and let $T : [0,1) \to [0,1)$ be given by $Tx = Nx \bmod 1$. Then, similar as the decimal map considered in Subsection 1.1.1, T generates N-adic expansions so that each $x \in [0,1)$ can be written as $x = \sum_{k=1}^{\infty} \frac{a_k}{N^k}$ with $a_k \in \{0, 1, \ldots, N-1\}$. Furthermore, let $(\Omega_E, \mathcal{F}, \mathbb{P}, \sigma)$ be the one-sided Bernoulli shift from Example 2.5 with $E = \{0, 1, \ldots, N-1\}$ and a probability vector $p = (p_j)_{j \in E}$ given by $p_j = \frac{1}{N}$ for each $j \in E$. Let

$$N = \{\omega \in \Omega_E : \exists k \geq 1 : \omega_i = N - 1 \text{ for all } i \geq k\}. \tag{2.32}$$

Then one can show that $([0,1), \mathcal{B}, \lambda, T)$ and $(\Omega_E, \mathcal{F}, \mathbb{P}, \sigma)$ are isomorphic with an isomorphism $\psi : \Omega_E \backslash N \to [0,1)$ given by

$$\psi(\omega) = \sum_{k=1}^{\infty} \frac{\omega_k}{N^k}. \tag{2.33}$$

The following proposition is obvious.

Proposition 2.42. *Suppose $(X, \mathcal{B}, \mu, T)$ and $(Y, \mathcal{C}, \nu, S)$ are two isomorphic dynamical systems. Then (T, μ) is ergodic (resp. weakly mixing, mixing, exact) if and only if (S, ν) is ergodic (resp. weakly mixing, mixing, exact).*

In this thesis, we mostly work on probability spaces that are Lebesgue spaces. These non-pathological probability spaces are introduced by [59] and can be thought of as the union of an interval and an at most countable number of atoms. This is made more precise in the next definition. First, recall that a set $A \in \mathcal{B}$ in a probability space $(X, \mathcal{B}, \mu)$ is called an *atom* if $\mu(A) > 0$ and if for each $B \in \mathcal{B}$ with $B \subseteq A$ and $\mu(B) < \mu(A)$ we have $\mu(B) = 0$.

Definition 2.43. *(Definition 4.5 in [54]) We call a probability space $(X, \mathcal{B}, \mu)$ a* Lebesgue space *if there exists an at most countable union $X_0 = \cup_i A_i$ of atoms $A_i \in \mathcal{B}$ such that, writing $\tilde{X} = X \backslash X_0$, $\tilde{\mathcal{B}} = \mathcal{B} \cap \tilde{X}$ and $\tilde{\mu}(\cdot) = \frac{\mu(\cdot)}{\mu(\tilde{X})}$, the dynamical system $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{\mu}, id_{\tilde{X}})$ is isomorphic to $([0,1), \mathcal{B}([0,1)), \lambda, id_{[0,1)})$. Here, $\mathcal{B}([0,1))$ is the Lebesgue σ-algebra on $[0,1)$ and λ the Lebesgue measure.*

Theorem 2.44. *(Theorem 4.6 in [54]) Let X be a Polish space and $\mathcal{B}$ the corresponding Borel σ-algebra on X. Let μ be a probability measure on $(X, \mathcal{B})$. Suppose $(X, \mathcal{B}, \mu)$ is complete, i.e. if $A \in \mathcal{B}$ such that $\mu(A) = 0$, then $B \in \mathcal{B}$ for all $B \subseteq A$. Then $(X, \mathcal{B}, \mu)$ is a Lebesgue space.*

Example 2.45. Let E be a Polish space. Then one can show that $\Omega_E = E^{\mathbb{N}}$ (or $\Omega_E = E^{\mathbb{Z}}$) is a Polish space as well. Let $\mathcal{F}$ be the Borel σ-algebra on Ω_E, which in case E is countable corresponds to the σ-algebra generated by the cylinder sets. Let $\mathbb{P}$ be a Borel probability measure on $(\Omega_E, \mathcal{F})$, and write $\mathcal{F}_{\mathbb{P}}$ for the completion of $\mathcal{F}$ w.r.t. $\mathbb{P}$. Then $(\Omega_E, \mathcal{F}_{\mathbb{P}}, \mathbb{P})$ is a Lebesgue space.

2.8 Entropy

The notion of entropy in information theory was introduced by Shannon [61] to quantify the amount of randomness produced by an information source. In [38], Kolmogorov introduced entropy in dynamical systems, which was made rigorous by Sinai [63]. In this section we briefly review this very important concept of (Kolmogorov-Sinai) entropy in Ergodic Theory. Let us fix a dynamical system $(X, \mathcal{B}, \mu, T)$.

We say $\alpha = \{A_i : i \in I\}$ is a *partition* of X if X is the disjoint union (up to sets of μ-measure zero) of the sets A_i, where $A_i \in \mathcal{B}$ for each $i \in I$ and where I is a finite or countable index set. For a partition α of X, we define the *entropy of the partition* α as

$$H_\mu(\alpha) = \sum_{A \in \alpha} \mu(A) \log \mu(A). \tag{2.34}$$

Also, we define

$$T^{-1}\alpha := \{T^{-1}A : A \in \alpha\}, \tag{2.35}$$

which is a partition of X as well.

Furthermore, for two partitions α and β of X, we define the *conditional entropy of* α *given* β as

$$H_\mu(\alpha|\beta) = -\sum_{A \in \alpha} \sum_{B \in \beta} \log \Big(\frac{\mu(A \cap B)}{\mu(B)} \Big) \mu(A \cap B). \tag{2.36}$$

Also, we define

$$\alpha \vee \beta := \{A \cap B : A \in \alpha, B \in \beta\}, \tag{2.37}$$

which is a partition of X as well and is called the *common refinement* of α and β.

The following properties are easy to show and will be needed in Chapter 5.

Proposition 2.46. *Let α, β and γ be partitions of X. Then*

(a) $H_\mu(T^{-1}\alpha) = H_\mu(\alpha)$,

(b) $H_\mu(\alpha \vee \beta) = H_\mu(\alpha) + H_\mu(\beta|\alpha)$,

(c) $H_\mu(\beta|\alpha) \leq H_\mu(\beta)$,

(d) $H_\mu(\alpha \vee \beta) \leq H_\mu(\alpha) + H_\mu(\beta)$.

Let α be a partition of X, and define for each $n \in \mathbb{N}$ the partition

$$\alpha_n = \bigvee_{k=0}^{n-1} T^{-k}\alpha = \Big\{ \bigcap_{k=0}^{n-1} T^{-k}A_k : A_k \in \alpha, k = 0, 1, \ldots, n-1 \Big\}. \tag{2.38}$$

In order to define the entropy of T with respect to the partition α, we need the following analytic lemma.

Lemma 2.47. (Fekete's Subadditive Lemma) *Suppose a sequence* $\{a_n\}_{n\in\mathbb{N}}$ *of real numbers is subadditive, i.e.* $a_{n+m} \leq a_n + a_m$ *for all* $n, m \in \mathbb{N}$. *Then* $\lim_{n\to\infty} \frac{a_n}{n}$ *exists and is equal to* $\inf_{n\in\mathbb{N}} \frac{a_n}{n}$.

Proposition 2.48. *Let* α *be a partition of* X *such that* $H_\mu(\alpha) < \infty$. *Then the sequence* $\{H_\mu(\alpha_n)\}_{n\in\mathbb{N}}$ *is subadditive.*

Proof: For all $n, m \in \mathbb{N}$ we have

$$H_\mu(\alpha_{n+m}) \leq H_\mu(\alpha_n) + H_\mu\Big(\bigvee_{k=n}^{n+m-1} T^{-k}\alpha\Big) \leq H_\mu(\alpha_n) + H_\mu(\alpha_m), \tag{2.39}$$

where we applied parts (d) and (a) of Proposition 2.46, respectively. □

Definition 2.49. *Let* α *be a partition of* X *such that* $H_\mu(\alpha) < \infty$. *Then the* entropy of T w.r.t. α *given by*

$$h_\mu(\alpha, T) = \lim_{n\to\infty} \frac{1}{n} H_\mu(\alpha_n) \tag{2.40}$$

exists and is finite by the previous two results. Finally, the entropy *of* T *is defined as*

$$h_\mu(T) = \sup\{h_\mu(\alpha, T) |\ \alpha \text{ partition of } X \text{ such that } H_\mu(\alpha) < \infty\}. \tag{2.41}$$

The following theorem shows that entropy is an isomorphism invariant. We refer to Theorem 5.2.2 in [15] for a proof.

Theorem 2.50. *Suppose* $(X, \mathcal{B}, \mu, T)$ *and* $(Y, \mathcal{C}, \nu, S)$ *are two isomorphic dynamical systems. Then* $h_\mu(T) = h_\nu(S)$.

In general, it does not seem possible to calculate the entropy straight from its definition. The next theorem is an important tool for calculating the entropy. First, let us define that a partition α of X is a *generator* with respect to a non-invertible transformation T if

$$\sigma\Big(\bigvee_{i=0}^{\infty} T^{-i}\alpha\Big) = \mathcal{B} \quad \text{up to sets of } \mu\text{-measure zero.} \tag{2.42}$$

If T is invertible, then α is called a generator w.r.t. T if $\sigma(\bigvee_{i=-\infty}^{\infty} T^{-i}\alpha) = \mathcal{B}$ up to sets of μ-measure zero.

Theorem 2.51. (Kolmogorov-Sinai) *Let* α *be a partition of* X *such that* $H_\mu(\alpha) < \infty$. *If* α *is a generator w.r.t.* T, *then* $h_\mu(T) = h_\mu(T, \alpha)$.

Example 2.52. Let $(\Omega_E, \mathcal{F}, \mathbb{P}, \sigma)$ be the Bernoulli shift from Example 2.5. By definition of $\mathcal{F}$, note that the partition $\alpha = \{A_j : j \in E\}$ given by $A_j = \{\omega \in \Omega_E : \omega_1 = j\}$ is a generator w.r.t. σ. Furthermore, since $\mathbb{P}$ is a product measure on Ω_E, we can derive

$$H_\mathbb{P}(\alpha_n) = \sum_{i=0}^{n-1} H_\mathbb{P}(\sigma^{-i}\alpha) = nH_\mathbb{P}(\alpha) = -n\sum_{j\in E} p_j \log p_j. \tag{2.43}$$

It follows from Theorem 2.51 that

$$h_{\mathbb{P}}(\sigma) = \lim_{n\to\infty} \frac{1}{n} H_{\mathbb{P}}(\alpha_n) = -\sum_{j\in E} p_j \log p_j. \tag{2.44}$$

Moreover, combining this with Example 2.41 and Theorem 2.50 yields that $T : [0,1) \to [0,1)$ given by $Tx = Nx \bmod 1$ ($N \in \mathbb{N}$) has entropy $h_\lambda(T) = \log N$ w.r.t. the Lebesgue measure λ on $[0,1)$.

Finally, we state the classical Shannon-McMillan-Breiman Theorem. For this, we define the *information function* associated to a partition α of X as

$$I_\alpha(x) = -\sum_{A\in\alpha} 1_A(x) \log \mu(A). \tag{2.45}$$

Note that $H_\mu(\alpha) = \int_X I_\alpha(x) d\mu(x)$. Denoting $\alpha(x)$ for the atom of α containing x, we can also write

$$I_\alpha(x) = -\log \mu(\alpha(x)). \tag{2.46}$$

Theorem 2.53. (Shannon-McMillan-Breiman) *Let α be a partition of X s.t. $H_\mu(\alpha) < \infty$. Suppose that T is ergodic w.r.t. μ. Then*

$$\lim_{n\to\infty} -\frac{I_{\alpha_n}(x)}{n} = h_\mu(T,\alpha), \qquad \mu\text{-a.e.} \tag{2.47}$$

Chapter 3

Invariant Densities for Piecewise Monotonic Interval Maps

3.1 Introduction

In this and the next chapter, we work on the probability space $(I, \mathcal{B}, \lambda)$, where $I = [0, 1]$ is the unit interval, $\mathcal{B}$ the Borel σ-algebra on I and λ the Lebesgue measure restricted to I.

Let $T : I \to I$ be measure preserving and ergodic with respect to some probability measure μ on I. We know from Theorem 2.14 that there exists $B \in \mathcal{B}$ with $\mu(B) = 1$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int_I f d\mu \qquad \text{for all } x \in B \text{ and } f \in C(I). \tag{3.1}$$

In case μ is absolutely continuous with respect to λ we have $\lambda(B) > 0$. Hence, the existence of an ergodic *absolutely continuous invariant probability measure* (acipm) for T implies in weak sense a characterization of the long-term average behavior of points in a set of *at least* positive Lebesgue measure. This raises the question under what conditions T admits an ergodic acipm.

We address this question in this chapter for transformations $T : I \to I$ that are finitely or countably *piecewise C^k-monotonic* ($k \geq 1$), i.e. there exists a finite or countable partition $\{I_i\}$ of I such that each I_i is an interval and the restriction of T to I_i is C^k, monotone and injective (see Figure 3.1). Note that such a transformation T is nonsingular w.r.t. λ, so T admits a corresponding transfer operator $P_{T,\lambda} : L^1(\lambda) \to L^1(\lambda)$ that we simply denote as P_T.

We know from Proposition 2.33 that $h \in L^1(\lambda)$ is a fixed point of P_T (i.e. h is an *invariant density* for T in the sense of (1.27)) if and only if the (complex) measure μ

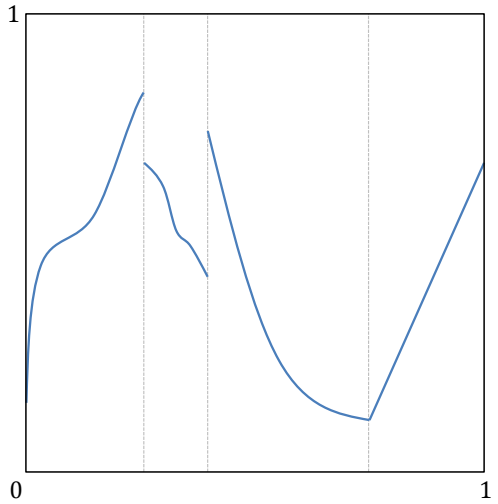

Figure 3.1: Example of a piecewise monotonic transformation on I

given by $\mu(A) = \int_A h d\lambda$ ($A \in \mathcal{B}$) is T-invariant (i.e. μ is an acim for T). Hence, we are interested in the fixed points of P_T.

We derive in Section 3.2 for a piecewise monotonic transformation $T : I \to I$ that

$$P_T f(x) = \sum_{y \in T^{-1}x} \frac{f(y)}{|T'(y)|} \qquad \lambda\text{-a.e.} \tag{3.2}$$

for all $f \in L^1(\lambda)$. Furthermore, we discuss in Section 3.3 the celebrated theorem by Lasota and Yorke that any finitely piecewise C^2-monotonic map $T : I \to I$ that is *expanding* (i.e. $\inf_{x \in I} |T'(x)| > 1$) admits an acipm. The main technical step in the proof is to obtain the *Lasota-Yorke inequality*: There exists $k \in \mathbb{N}$, $\rho \in (0, 1)$ and $L \in (0, \infty)$ such that

$$\mathrm{Var}_I(P_T^k f) \leq \rho \mathrm{Var}_I(f) + L\|f\|_1 \qquad \text{for all } f \in BV(I). \tag{3.3}$$

In (3.3), $\mathrm{Var}_I(\cdot)$ denotes the *variation* of a function on I and $BV(I)$ is the space of functions of *bounded variation* on I. We recall these definitions in Appendix A. We shall see in Section 3.4 that (3.3) implies that P_T is *quasi-compact* (see Appendix B.3) on $BV(I)$ as a consequence of the famous Ionescu-Tulcea and Marinescu Theorem, from which several ergodic properties of T can be derived. If we in addition assume that T admits a suitable covering property such as in the so-called Folklore Theorem, we shall see in Section 3.5 that T admits a unique acipm μ, that T is exact with respect to this μ and that μ satisfies

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\mu}{d\lambda} \leq M. \tag{3.4}$$

Finally, we recover in Section 3.6 by a method based on Section 3 in [35] that the LSV map from (1.28) admits a unique acipm if $\alpha \in (0, 1)$ and an infinite *σ-finite acim* if $\alpha \geq 1$.

3.2 Representation of the Transfer Operator

Let $T : I \to I$ be piecewise monotonic. We derive (3.2) by following Section 4.3 in [9]. For simplicity, let us assume T is *finitely* piecewise monotonic, i.e. there exists a finite partition of I, $0 = a_0 < a_1 < \cdots < a_n = 1$ such that

(1) $T_i := T|_{(a_{i-1},a_i)}$ is C^k $(k \geq 1)$ and has a C^k extension to $[a_{i-1}, a_i]$, $i = 1, \ldots, n$,

(2) $\theta(T) := \inf_{x \in I} |T'(x)| > 0$ (where we take the one-sided derivates at the points where T is not differentiable).

Let $A \in \mathcal{B}$ and $f \in L^1(\lambda)$. By definition, we have

$$\int_A P_T f d\lambda = \int_{T^{-1}A} f d\lambda = \sum_{i=1}^{n} \int_{T_i^{-1}A} f d\lambda. \tag{3.5}$$

Recall the change of variable formula for (Riemann) integration: If g is differentiable over $[a, b]$ such that g' is integrable over $[a, b]$ and if h is integrable over $g([a, b])$, then

$$\int_{g(a)}^{g(b)} h(x)dx = \int_a^b h(g(y))g'(y)dy. \tag{3.6}$$

For each $i = 1, \ldots, n$, this formula with $h(x) = \frac{1_A(x)f(T_i^{-1}(x))}{T_i'(T_i^{-1}x)}$ and $g(y) = T_i(y)$ gives

$$\begin{aligned}
\int_{T_i^{-1}A} f d\lambda &= \int_{a_{i-1}}^{a_i} 1_A(T_i y) f(y) \frac{T_i'(y)}{T_i'(y)} dy \\
&= \int_{T_i(a_{i-1})}^{T_i(a_i)} \frac{1_A(x) f(T_i^{-1}(x))}{T_i'(T_i^{-1}x)} dx \\
&= \int_{A \cap T(a_{i-1},a_i)} \frac{f(T_i^{-1}x)}{|T'(T_i^{-1}x)|} d\lambda(x),
\end{aligned}$$

where in the last step we account for the fact that T_i is either increasing or decreasing. Combining this with (3.5) yields

$$\int_A P_T f d\lambda = \int_A \sum_{i=1}^{n} \frac{f(T_i^{-1}x)}{|T'(T_i^{-1}x)|} 1_{T(a_{i-1},a_i)}(x) d\lambda(x). \tag{3.7}$$

Since (3.7) holds for each $A \in \mathcal{B}$, we obtain

$$P_T f(x) = \sum_{i=1}^{n} \frac{f(T_i^{-1}x)}{|T'(T_i^{-1}x)|} 1_{T(a_{i-1},a_i)}(x) \qquad \lambda\text{-a.e.} \tag{3.8}$$

for all $f \in L^1(\lambda)$. Note that we can rewrite this to

$$P_T f(x) = \sum_{y \in T^{-1}x} \frac{f(y)}{|T'(y)|} \qquad \lambda\text{-a.e.} \tag{3.9}$$

for all $f \in L^1(\lambda)$. Observe that if T is piecewise monotonic with respect to a *countable* partition, we can derive (3.9) using the same arguments (in addition, we need to apply

the Monotone Convergence Theorem to interchange integral and series).

3.3 The Lasota-Yorke Inequality

We say that a piecewise monotonic transformation $T : I \to I$ is *expanding* if $\theta(T) = \inf_{x\in I} |T'(x)| > 1$. We have the following famous result due to Lasota and Yorke [43].

Theorem 3.1. (Lasota-Yorke) *Let $T : I \to I$ be finitely piecewise C^2-monotonic and expanding. Then T admits an acipm whose density is of bounded variation.*

In order to prove Theorem 3.1, we follow Section 10 of [10]. The key to this proof is the Lasota-Yorke Inequality (3.3) that we shall obtain with the following technical lemma. We refer the reader to Appendix A for a review on functions of bounded variation.

Lemma 3.2. *Let $T : I \to I$ be finitely piecewise C^2-monotonic. Then*

$$\mathrm{Var}_I(P_T f) \le \frac{2}{\theta(T)}\mathrm{Var}_I(f) + L(T)\|f\|_1 \qquad \text{for all } f \in BV(I), \tag{3.10}$$

where $\theta(T) = \inf_{x\in I} |T'(x)|$ and $L(T)$ is a finite positive constant depending only on T.

Proof: Write $0 = a_0 < a_1 < \cdots < a_n = 1$ for the (minimal) partition on which T is finitely piecewise C^2-monotonic. We have $|T''| \le K$ for some $K > 0$, so

$$\Big|\frac{d}{dx}\frac{1}{T'(x)}\Big| = \frac{|T''(x)|}{(T'(x))^2} \le \frac{K}{\theta^2} \qquad \text{for all } x \in I, \tag{3.11}$$

where $\theta = \theta(T) > 0$. Now, let $f \in BV(I)$. Then from (3.8) it follows that

$$P_T f(x) = \sum_{i=1}^{n} \frac{f(T_i^{-1}x)}{|T'(T_i^{-1}x)|} 1_{T([a_{i-1},a_i])}(x) \qquad \lambda\text{-a.e.}, \tag{3.12}$$

where we changed the right-hand side on a finite number of points x and now write T_i for its C^2-extension to $[a_{i-1}, a_i]$. For each $i = 1, \ldots, n$, Yorke's Inequality (Theorem A.17) gives

$$\begin{aligned}
\mathrm{Var}_I\Big(\frac{f\circ T_i^{-1}}{|T'\circ T_i^{-1}|}1_{T([a_{i-1},a_i])}\Big) &\le 2\mathrm{Var}_{T([a_{i-1},a_i])}\Big(\frac{f}{|T'|}\circ T_i^{-1}\Big) \\
&\quad + \frac{2}{T(a_i)-T(a_{i-1})}\int_{T(a_{i-1})}^{T(a_i)} \frac{|f(T_i^{-1}(x))|}{|T'(T_i^{-1}(x))|}dx \\
&\le 2\mathrm{Var}_{[a_{i-1},a_i]}\Big(\frac{f}{|T'|}\Big) + \frac{2}{|T(a_i)-T(a_{i-1})|}\int_{a_{i-1}}^{a_i} |f(y)|dy,
\end{aligned}$$

where the last step follows from (A.7) and the change of variable formula (3.6). Moreover, for each $i = 1, \ldots, n$ it follows from Proposition A.14 that

$$\begin{aligned}
\mathrm{Var}_{[a_{i-1},a_i]}\Big(\frac{f}{|T'|}\Big) &\le \frac{1}{\theta}\mathrm{Var}_{[a_{i-1},a_i]}(f) + \int_{a_{i-1}}^{a_i} |f(y)|\Big|\frac{d}{dy}\frac{1}{T'(y)}\Big|dy \\
&\le \frac{1}{\theta}\mathrm{Var}_{[a_{i-1},a_i]}(f) + \frac{K}{\theta^2}\int_{a_{i-1}}^{a_i} |f(y)|dy,
\end{aligned}$$

where the last step follows from (3.11). We conclude

$$\mathrm{Var}_I(P_T f) \leq \frac{2}{\theta}\sum_{i=1}^{n} \mathrm{Var}_{[a_{i-1},a_i]}(f) + \sum_{i=1}^{n} \frac{2}{\theta}\Big(\frac{K}{\theta} + \frac{1}{a_i - a_{i-1}}\Big)\int_{a_{i-1}}^{a_i} |f(s)|ds. \tag{3.13}$$

Using that $\sum_{i=1}^{n} \mathrm{Var}_{[a_{i-1},a_i]}(f) \leq \mathrm{Var}_I(f)$ (see Lemma A.5), we obtain (3.10) with $L(T) = \frac{2}{\theta}(\frac{K}{\theta} + \max_i \frac{1}{a_i - a_{i-1}})$. □

Proof (Theorem 3.1): Note that, for each $k \in \mathbb{N}$, T^k is also finitely piecewise C^2-monotonic. Moreover, we have $\theta(T^k) \geq \theta^k$ with $\theta = \theta(T)$, because $(T^k)'(x) = \prod_{i=0}^{k-1} T'(T^i x)$ for each $x \in I$ by the chain rule. Let us fix a $k \in \mathbb{N}$ such that $\theta^k > 2$. Then from Lemma 3.2 we obtain the Lasota-Yorke inequality

$$\mathrm{Var}_I(P_T^k f) \leq \rho \mathrm{Var}_I(f) + L\|f\|_1 \qquad \text{for all } f \in BV(I), \tag{3.14}$$

where $\rho := \frac{2}{\theta(T^k)} \in (0,1)$ and $L := L(T^k) \in (0,\infty)$.

We now construct a fixed point of P_T with (3.14). Let $f \in BV(I)$. Iterating (3.14), it follows that, for each $n \in \mathbb{N}$,

$$\mathrm{Var}_I(P_T^{kn} f) \leq \rho^n \mathrm{Var}_I f + L\|f\|_1 \sum_{i=0}^{n-1} \rho^i. \tag{3.15}$$

Because $\mathrm{Var}_I(\cdot)$ is a seminorm, we thus obtain that the sequence $\{f_n\}$ given by

$$f_n = \frac{1}{n}\sum_{i=1}^{n} P_T^{ki} f \tag{3.16}$$

satisfies $\mathrm{Var}_I(f_n) \leq M$ for each $n \in \mathbb{N}$, where $M := \mathrm{Var}_I(f) + \frac{L\|f\|_1}{1-\rho}$. We also have $\|f_n\|_1 \leq \|f\|_1$ for each $n \in \mathbb{N}$, so Lemma A.4 gives $\sup_{n\in\mathbb{N}} \|f_n\|_\infty \leq M + \|f\|_1$. From Helly's First Theorem (Theorem A.18) we now obtain a subsequence $\{f_{n_j}\}$ that converges pointwise to some $f^* \in BV(I)$ that satisfies (using Lemma A.2 and (3.15))

$$\mathrm{Var}_I(f^*) \leq \liminf_j \frac{1}{n_j}\sum_{i=0}^{n_j-1} \mathrm{Var}_I(P_T^{ki} f) \leq \limsup_{n\to\infty} \mathrm{Var}_I(P_T^{kn} f) \leq \frac{L\|f\|_1}{1-\rho}. \tag{3.17}$$

Combining this with $\sup_{j\in\mathbb{N}} \|f_{n_j}\|_\infty < \infty$ yields with the Dominated Convergence Theorem that

$$\begin{aligned}
\|P_T f^* - f^*\|_1 &\leq \|P_T f^* - P_T f_{n_j}\|_1 + \|P_T f_{n_j} - f_{n_j}\|_1 + \|f_{n_j} - f^*\|_1 \\
&\leq 2\|f_{n_j} - f^*\|_1 + \Big\|\frac{1}{n_j}\sum_{i=0}^{n_j-1} P_T^{ki+1} f - \frac{1}{n_j}\sum_{i=0}^{n_j-1} P_T^{ki} f\Big\|_1 \\
&\leq 2\|f_{n_j} - f^*\|_1 + \frac{1}{n_j}\|P_T^{k(n_j-1)+1} f - f\|_1 \\
&\leq 2\|f_{n_j} - f^*\|_1 + \frac{2}{n_j}\|f\|_1 \to 0, \quad j \to \infty.
\end{aligned} \tag{3.18}$$

The result now follows from Proposition 2.33. (Note that f^* is a probability density if we take for instance $f \equiv 1$.) □

Remark 3.3. The proof of Lemma 3.2 does not directly carry over to the case that T is piecewise monotonic with respect to a *countable* partition $\{I_i\}$, because in that case $L(T)$ would be infinite. This can be solved by assuming as in [43] that $\sup|T''| < \infty$ and that $T_i(\overline{I_i}) = I$ for all but finitely many i. Indeed, $\eta := \max(\frac{1}{\lambda(I_i)} : i \text{ s.t. } T_i(\overline{I_i}) \neq I)$ is then finite, so that we can estimate the problematic term

$$\sum_i \frac{1}{\lambda(T_i(I_i))} \int_{I_i} |f(y)|dy \leq \max\left(1, \frac{\eta}{\theta}\right) \cdot \|f\|_1. \tag{3.19}$$

Another sufficient assumption besides $\sup|T''| < \infty$ is to require as in [10] that there exists $\gamma > 0$ such that $\lambda(T_i(I_i)) \geq \gamma$ for all i. Indeed, in this case the left-hand side of (3.19) is bounded by $\gamma^{-1}\|f\|_1$.

Remark 3.4. In [60], Rychlik found that the Lasota-Yorke inequality and therefore the conclusion of Theorem 3.1 also holds for piecewise monotonic and expanding maps of the following form: Let $T : I \to I$ be piecewise monotonic and expanding w.r.t. a finite or countable interval partition $\{I_i\}$. Write $U = \bigcup_i \mathrm{Int}(I_i)$ where $\mathrm{Int}(I_i)$ denotes the interior of I_i. It is shown in [60] (see also [9]) that if $g : I \to \mathbb{R}$ given by

$$g(x) = \begin{cases} \frac{1}{|T'(x)|} & \text{if } x \in U, \\ 0 & \text{if } x \in I \backslash U \end{cases} \tag{3.20}$$

satisfies $\mathrm{Var}_I(g) < \infty$, then the Lasota-Yorke inequality (3.14) holds for some $k \in \mathbb{N}$, $\rho \in (0, 1)$ and $L \in (0, \infty)$. Other generalizations of Theorem 3.1 can be found in e.g. Section 6 of [9].

Remark 3.5. Note that the proof of Theorem 3.1 only shows there exists an acipm and does not show how to construct it. Explicit formula's of the acipm's are derived in e.g. [29, 39] for the case that T is piecewise linear and expanding.

Remark 3.6. The assumption in Theorem 3.1 that T is expanding can be weakened to some extent (see e.g. Theorem 3 in [43]), but it is certainly not possible to omit it completely. An example is considered in Section 3.6. As an example of a countably piecewise monotonic transformation, consider the Rényi map $R : I \to I$ given by

$$R(x) = \begin{cases} 0 & \text{if } x = 1, \\ \frac{1}{1-x} \mod 1 & \text{otherwise,} \end{cases} \tag{3.21}$$

which can be obtained by reflecting the Gauss map in Figure 1.2 over the vertical line through $\frac{1}{2}$ (see Figure 3.2). Clearly, R is not expanding because $x = 0$ is a *neutral* fixed point of R, i.e. $|R'(0)| = 1$. It was proved by Rényi in [56] that R has no acipm, but does have a σ-finite acim with density $h(x) = \frac{1}{x}$. It appears that a piecewise monotonic transformation T with a neutral fixed point and $|T'| > 1$ elsewhere typically has an invariant density of the type $\frac{1}{x}$ (see Remark 5.3.2 in [9]) and we shall consider another example of such a T in Section 3.6.

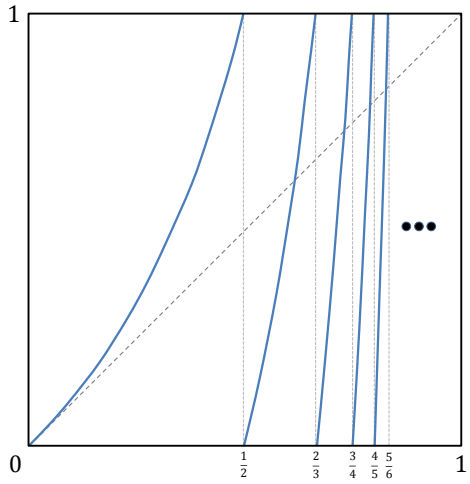

Figure 3.2: The Rényi map R

3.4 Quasi-compactness of the Transfer Operator

In this section we let $T : I \to I$ be a finitely piecewise C^2-monotonic and expanding map.[1] Moreover, we fix $k \in \mathbb{N}$ such that P_T satisfies the Lasota-Yorke inequality

$$\mathrm{Var}_I(P_T^k f) \leq \rho \mathrm{Var}_I(f) + L\|f\|_1 \qquad \text{for all } f \in BV(I). \tag{3.22}$$

for some $\rho \in (0,1)$ and $L \in (0,\infty)$. Let us prove that each fixed point of P_T is of bounded variation. For this, we need the following lemma. The proof below is essentially the one from [43].

Lemma 3.7. *For each $f \in L^1(\lambda)$, the sequence $\{f_n\}$ given by*

$$f_n = \frac{1}{n}\sum_{i=1}^{n} P_T^{ki} f \tag{3.23}$$

converges in $L^1(\lambda)$-norm to some $\tilde{f} \in L^1(\lambda)$.

Proof: In the proof of Theorem 3.1 we showed that for each $f \in BV(I)$ the sequence $\{f_n\}$ contains a subsequence $\{f_{n_j}\}$ that converges pointwise to some $\tilde{f} \in BV(I)$ and satisfies $\sup_{j\in\mathbb{N}} \|f_{n_j}\|_\infty < \infty$, in which case we obtain from the Dominated Convergence Theorem that

$$\lim_{j\to\infty} \int_I f_{n_j} g d\lambda = \int_I \tilde{f} g d\lambda \qquad \text{for each } g \in L^\infty(\lambda). \tag{3.24}$$

Since $BV(I)$ is dense in $(L^1(\lambda), \|\cdot\|_1)$ (see Corollary A.16), the result now follows from the Kakutani-Yosida Theorem (Theorem B.1). □

[1] The results in this section also hold for the piecewise monotonic expanding maps in Remarks 3.3 and 3.4 with countably many branches.

As in [43], let us define the bounded linear operator

$$Q : L^1(\lambda) \to L^1(\lambda), \qquad Qf = \tilde{f} \tag{3.25}$$

with $\tilde{f}$ as in Lemma 3.7. Since P_T is a contraction on $(L^1(\lambda), \|\cdot\|_1)$ (see part (d) of Proposition 2.31), it follows that also Q is a contraction on $(L^1(\lambda), \|\cdot\|_1)$.

Proposition 3.8. *Each fixed point of P_T is an element of $BV(I)$.*

Proof: Let $f \in L^1(\lambda)$ be s.t. $P_T f = f$. Since $BV(I)$ is dense in $(L^1(\lambda), \|\cdot\|_1)$, there exists a sequence $\{g_m\} \subseteq BV(I)$ such that $g_m \xrightarrow{L^1} f$. So $M := \sup_{m\in\mathbb{N}} \|g_m\|_1 < \infty$ and

$$Qg_m \xrightarrow{L^1} Qf = f, \qquad m \to \infty. \tag{3.26}$$

For each $m \in \mathbb{N}$ we know from the proof of Theorem 3.1 that $\{\frac{1}{n}\sum_{i=1}^n P_T^{ki} g_m\}_{n\in\mathbb{N}}$ contains a subsequence that converges pointwise and in L^1 to some $g_m^* \in BV(I)$ that satisfies $\mathrm{Var}_I(g_m^*) \leq \frac{L\|g_m\|_1}{1-\rho}$. Obviously, we have $g_m^* = Qg_m$ for each $m \in \mathbb{N}$, so

$$\sup_{m\in\mathbb{N}} \mathrm{Var}_I(Qg_m) \leq \frac{LM}{1-\rho}. \tag{3.27}$$

From Lemma A.4 it now follows that $\sup_{m\in\mathbb{N}} \|Qg_m\|_\infty \leq M + \frac{LM}{1-\rho}$, which together with (3.27) and Helly's First Theorem (Theorem A.18) yields

$$\exists g^* \in BV(I)\ \forall x \in I : Qg_m(x) \to g^*(x), \qquad m \to \infty. \tag{3.28}$$

Combining (3.26) and (3.28) yields $f = g^*$ λ-a.e. □

Corollary 3.9. (Theorem 1(vi) in [33]) *The probability measure $\tilde{\mu}$ on $(I, \mathcal{B})$ given by*

$$\tilde{\mu}(A) = \int_A Q1 d\lambda, \qquad A \in \mathcal{B} \tag{3.29}$$

is the biggest acipm of T in the following sense: If μ is an acipm of T, then μ is absolutely continuous w.r.t. $\tilde{\mu}$.

Proof: Let $f \in L^1(\lambda)$ be such that $P_T f = f$. Then $f \in BV(I)$ by the previous proposition, so $|f|$ is bounded by some constant $M > 0$. This gives

$$|f| = \frac{1}{n}\sum_{i=1}^n |P_T^{ki} f| \leq M \cdot \frac{1}{n}\sum_{i=1}^n P_T^{ki} 1 \tag{3.30}$$

and therefore $|f| \leq MQ1$, from which the result follows. □

Motivated by the result of Proposition 3.8, let us consider the restriction of P_T to $BV(I)$, which we denote as $P_{T,BV}$. We see from Lemma 3.2 that $BV(I)$ is preserved by P_T, so $P_{T,BV} : BV(I) \to BV(I)$. The next well-known theorem shows that $P_{T,BV}$ is a quasi-compact operator. In Appendix B.3 we briefly recall the definition of a quasi-compact operator on a general complex Banach space and state the Ionescu-Tulcea and Marinescu Theorem on which the proof below is based. Recall from Proposition A.13 that $BV(I)$ is a complex Banach space with respect to the norm

$$\|f\|_{BV} = \mathrm{Var}_I f + \|f\|_1, \qquad f \in BV(I). \tag{3.31}$$

Theorem 3.10. *The operator $P_{T,BV} : BV(I) \to BV(I)$ is quasi-compact and the set of eigenvalues of $P_{T,BV}$ with modulus 1 has only a finite number of elements, say $\lambda_1, \dots, \lambda_m$. That is, there are bounded linear operators $Q_1, \dots, Q_m$ and S on $BV(I)$ such that*

$$\begin{aligned}
&P^n_{T,BV} = \sum_{i=1}^{m} \lambda_i^n Q_i + S^n, && \text{for all } n \in \mathbb{N},\\
&Q_i Q_j = 0 && \text{if } i \neq j,\\
&Q_i^2 = Q_i && \text{for all } i = 1, \dots, m,\\
&Q_i S = S Q_i = 0 && \text{for all } i = 1, \dots, m,\\
&Q_i BV(I) = E(\lambda_i) && \text{for all } i = 1, \dots, m,\\
&\rho(S) < 1,
\end{aligned}$$

where $E(\lambda_i) = \{f \in BV(I) : P_{T,BV} f = \lambda_i f\}$ is the eigenspace *of $P_{T,BV}$ associated to λ_i, and $\rho(S) = \lim_{n\to\infty} \|S^n\|_{BV}^{1/n}$ is the* spectral radius *of S. Moreover, for each $i = 1, \dots, m$ the eigenspace $E(\lambda_i)$ associated to λ_i is finite-dimensional.*

Proof: It suffices to check the conditions in the Ionescu-Tulcea and Marinescu Theorem (Theorem B.5). Note that the second condition in Theorem B.5 follows from the fact that $P_{T,BV}$ is a contraction on $(BV(I), \|\cdot\|_1)$ and that the third condition follows directly from the Lasota-Yorke inequality (3.22). The first and fourth condition do not depend on T and for their proof we refer the reader to Proposition 7.2.1 in [9]. □

We can now prove the following result.

Theorem 3.11. *The set $M_{ac}(I, T)$ of acim's of T is a non-empty finite-dimensional vector space generated by the ergodic acipm's of T.*

Proof: First of all, non-emptiness of $M_{ac}(I, T)$ follows from Theorem 3.1. Moreover, it is clear that $M_{ac}(I, T)$ is a finite-dimensional vector space, because $E(1)$ in Theorem 3.10 is finite-dimensional. So there exists $n \in \mathbb{N}$ such that $\dim M_{ac}(I, T) = \dim E(1) = n$. Let $\tilde{\mu}$ be the measure from Corollary 3.9 given by (3.29). We now show the following claim for each $k = 1, \dots, n$: There exists a partition of I into sets $A_1, \dots, A_k \in \mathcal{B}$ such that $T^{-1}A_i = A_i$ and $\tilde{\mu}(A_i) > 0$ for each $i = 1, \dots, k$.

- For $k = 1$, just take $A_1 = I$.
- Suppose the claim holds for some $k \in \{1, \dots, n-1\}$ with corresponding partition $\{A_1, \dots, A_k\}$. Let us show the claim for $k + 1$. Assume that for each $i = 1, \dots, k$ and each $A \in \mathcal{B}$ such that $T^{-1}A = A$ we have $\frac{\tilde{\mu}(A \cap A_i)}{\tilde{\mu}(A_i)} \in \{0, 1\}$. Then the measures

$$\tilde{\mu}_{A_i}(B) = \frac{\tilde{\mu}(B \cap A_i)}{\tilde{\mu}(A_i)} \qquad B \in \mathcal{B}, \qquad i = 1, \dots, k \tag{3.32}$$

  are ergodic acipm's of T. Note that for each acim μ of T and each $i = 1, \dots, k$ the measure $\frac{\mu(\cdot \cap A_i)}{\mu(A_i)}$ is also an acipm of T and absolutely continuous w.r.t. $\tilde{\mu}_{A_i}$ (because of Corollary 3.9), which combined with the first part of Theorem 2.10 yields that $\frac{\mu(\cdot \cap A_i)}{\mu(A_i)} = \tilde{\mu}_{A_i}(\cdot)$. This gives the contradiction $\dim M_{ac}(I, T) = k < n$.

So there exist $i \in \{1, \dots, k\}$ and $A \in \mathcal{B}$ such that $T^{-1}A = A$ and $\frac{\tilde{\mu}(A \cap A_i)}{\tilde{\mu}(A_i)} \in (0,1)$. We obtain the claim for $k+1$ with the partition $\{A_1, \dots, A \cap A_i, A^{\complement} \cap A_i, \dots, A_k\}$.

The measures in (3.32) are mutually singular and therefore linearly independent, so the claim does not hold for $k > n$. It follows that for $k = n$ the measures in (3.32) are ergodic acipm's of T. □

Remark 3.12. The result in Theorem 3.11 for finitely piecewise C^2-monotonic maps $T : I \to I$ has first been derived in [44] (and without any arguments involving the quasi-compactness of $P_{T,BV}$). In particular, it is shown in [44] that the dimension of $M_{ac}(I,T)$ is bounded by the number of discontinuities of T.

Using the quasi-compactness of $P_{T,BV}$ it is possible to obtain a number of ergodic properties of T (see e.g. Section 7.2 in [9]). As an example, we show that if $(T, \tilde{\mu})$ is weak mixing (with $\tilde{\mu}$ given by (3.29)), then $(T, \tilde{\mu})$ is exact. Let us first prove the following lemma.

Lemma 3.13. *Suppose $f \in BV(I)$ and $\lambda \in \mathbb{C}$ satisfy $P_{T,BV} f = \lambda f$ and $|\lambda| = 1$. Then $Rf := \frac{f}{Q1}$ can be taken as a version in $\mathcal{L}^1(\tilde{\mu})$ s.t. $Rf(x) = 0$ whenever $Q1(x) = 0$, in which case $U_{T,\tilde{\mu}}(\overline{Rf}) = \lambda \overline{Rf}$. In particular, the set of eigenvalues of $P_{T,BV}$ with modulus 1 is contained in the set of eigenvalues of $U_{T,\tilde{\mu}}$.*

Proof: Suppose $P_{T,BV} f = \lambda f$ and $|\lambda| = 1$. We have $f \in BV(I)$, so $|f|$ is bounded by some constant $M > 0$. This gives $|f| = \frac{1}{n} \sum_{i=1}^n |P_T^{ki} f| \leq M \cdot \frac{1}{n} \sum_{i=1}^n P_T^{ki} 1$ and therefore $|f| \leq MQ1$. This indeed gives $Rf \in \mathcal{L}^1(\tilde{\mu})$. From Proposition 2.34 it now follows that $P_{T,\tilde{\mu}}(Rf) = \lambda Rf$, which together with Proposition 2.39 gives $U_{T,\tilde{\mu}} \overline{Rf} = \lambda \overline{Rf}$. □

Proposition 3.14. *Suppose $(T, \tilde{\mu})$ is weakly mixing. Then for each $n \in \mathbb{N}$ we have*

$$P_{T,BV}^n g = \Big(\int_I g d\lambda \Big) Q1 + S^n g, \qquad g \in BV(I) \tag{3.33}$$

where for some $q \in (0,1)$ and $M > 0$ we have for each $n \in \mathbb{N}$ that $\|S^n\|_{BV} \leq Mq^n$.

Proof: First of all, since $(T, \tilde{\mu})$ is weakly mixing, we know from Theorem 2.26 that 1 is the only eigenvalue of $U_{T,\tilde{\mu}}$. Combining this with Theorem 3.10 and Lemma 3.13 gives

$$P_{T,BV}^n g = Q_1 g + S^n g, \qquad g \in BV(I),\ n \in \mathbb{N}, \tag{3.34}$$

where Q_1 and S are bounded linear operators on $BV(I)$ such that $Q_1(BV(I)) = E(1)$ and $\rho(S) \in (0,1)$. The latter implies that, for some $q \in (0,1)$ and $M > 0$, $\|S^n\|_{BV} \leq Mq^n$ for each $n \in \mathbb{N}$. Furthermore, since $(T, \tilde{\mu})$ is ergodic, it follows from Theorem 2.25 and Lemma 3.13 that $E(1)$ is 1-dimensional. So $Q_1 g = \varphi(g) Q1$ for each $g \in BV(I)$, where $\varphi : BV(I) \to \mathbb{C}$ is a bounded linear map. By the Hahn-Banach Theorem, we can extend φ to a bounded linear map $\psi : L^1(\lambda) \to \mathbb{C}$. Since $(L^1(\lambda))^*$ is isomorphic to $L^\infty(\lambda)$ via the correspondence $\theta(g) = \int_I g h d\lambda$, $h \in L^\infty(\lambda)$, it follows that there exists $h \in L^\infty(\lambda)$ such that $\psi(g) = \int_I g h d\lambda$. We conclude

$$Q_1 g = \Big(\int_I g h d\lambda \Big) Q1, \qquad g \in BV(I). \tag{3.35}$$

It remains to show that h is λ-a.e. equal to 1. Indeed, for each $A \in \mathcal{B}$ we have

$$\int_A h d\lambda = \int_I \Big(\int_I h 1_A d\lambda \Big) Q1 d\lambda = \int_I (Q1_A) d\lambda = \lim_{n\to\infty} \int_I P^n_{T,BV} 1_A d\lambda = \int_A 1 d\lambda.$$

□

Corollary 3.15. *Suppose $(T, \tilde{\mu})$ is weakly mixing. Then $(T, \tilde{\mu})$ is exact.*

Proof: From Proposition 3.14 it follows that, for each $n \in \mathbb{N}$,

$$\Big\| P^n_T g - \Big(\int_I g d\lambda \Big) Q1 \Big\|_1 \leq M q^n \|g\|_{BV}, \qquad g \in BV(I), \tag{3.36}$$

where $q \in (0,1)$ and $M > 0$. First, let $g \in BV(I)$. Then $Q1 \cdot g \in BV(I)$. Applying (3.36) to $Q1 \cdot g$ gives

$$\Big\| P^n_{T,\tilde{\mu}} g - \int_I g d\tilde{\mu} \Big\|_{\tilde{\mu},1} = \Big\| P^n_T (Q1 \cdot g) - \Big(\int_I Q1 \cdot g d\lambda \Big) Q1 \Big\|_1 \to 0, \quad n \to \infty. \tag{3.37}$$

Now let $g \in L^1(\tilde{\mu})$ and $\{g_m\} \subseteq BV(I)$ s.t. $g_m \overset{L^1(\tilde{\mu})}{\to} g$. Then

$$\Big\| P^n_{T,\tilde{\mu}} g - \int_I g d\tilde{\mu} \Big\|_{\tilde{\mu},1} \leq \|g - g_m\|_{\tilde{\mu},1} + \Big\| P^n_{T,\tilde{\mu}} g_m - \int_I g_m d\tilde{\mu} \Big\|_{\tilde{\mu},1} + \Big| \int_I g_m d\tilde{\mu} - \int_I g d\tilde{\mu} \Big|,$$

where we used part (d) of Proposition 2.31. Since the right-hand side converges to zero by first taking $n \to \infty$ and then $m \to \infty$, the result now follows from Theorem 2.35. □

3.5 Covering Property and Folklore Theorem

In this section we give some conditions for which a piecewise monotonic expanding transformation $T : I \to I$ admits a unique acipm μ, (T, μ) is exact and $\frac{d\mu}{d\lambda}$ is bounded and bounded away from zero. The proof of the next proposition uses a standard technique that can be found in e.g. [4, 36].

Proposition 3.16. *Let $T : I \to I$ be finitely piecewise C^2-monotonic and expanding. Furthermore, suppose that T satisfies the following* covering property*: For each nontrivial subinterval $J \subseteq I$ there exist $n \in \mathbb{N}$ and a finite set $I_0 \subseteq I$ such that $T^n J = I \backslash I_0$. Then $\tilde{\mu}$ from Corollary 3.9 is the only acipm of T and satisfies*

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\tilde{\mu}}{d\lambda} \leq M. \tag{3.38}$$

Moreover, $(T, \tilde{\mu})$ is ergodic.

Proof: Let $f \in BV(I)$ nonzero and real valued such that $f \geq 0$ and $P_T f = f$. First of all, since f is of bounded variation it follows that f is bounded. Moreover, by Corollary A.11, we may assume that f is lower semicontinuous. Then there exist $\alpha > 0$ and a nontrivial interval $J \subseteq I$ such that $f \geq \alpha 1_J$. Now take $n \in \mathbb{N}$ and $I_0 \subseteq I$

finite such that $T^n J = I \backslash I_0$. Since T^n is finitely piecewise C^2-monotonic, we have $K := \sup_{x \in I} |(T^n)'(x)| < \infty$. Hence, for all $x \in I \backslash I_0$ we obtain

$$f(x) = P_T^n f(x) \geq \alpha P_{T^n} 1_J(x) = \alpha \sum_{y \in T^{-n}x} \frac{1_J(y)}{|(T^n)'(y)|} \geq \frac{\alpha}{K}, \tag{3.39}$$

because for each $x \in I \backslash I_0$ there exists $y \in J$ such that $T^n y = x$. It follows from (A.10) that $f(x) \geq \frac{\alpha}{K}$ for all $x \in I_0$ as well. We conclude that f is bounded away from zero. Therefore, every acipm of T has full support, which together with the second part of Theorem 2.10 yields that the space $M_{ac}(I, T)$ from Theorem 3.11 has dimension 1. □

Remark 3.17. In fact, one can show with the second part of Theorem 7.2.1 in [9] that $(T, \tilde{\mu})$ is exact under the assumptions of Proposition 3.16.

Example 3.18. Let $\beta > 1$ and consider $T_\beta : I \to I$ given by $T_\beta x = \beta x \bmod 1$. Defining $A_i = [\frac{i}{\beta}, \frac{i+1}{\beta})$ for $i \in \{0, 1, \ldots, \lfloor \beta \rfloor - 1\}$ and $A_{\lfloor \beta \rfloor} = [\frac{\lfloor \beta \rfloor}{\beta}, 1]$, we can as well write

$$T_\beta x = \begin{cases} \beta x & \text{if } x \in A_0, \\ \beta x - 1 & \text{if } x \in A_1, \\ \vdots & \qquad \vdots \\ \beta x - \lfloor \beta \rfloor & \text{if } x \in A_{\lfloor \beta \rfloor}. \end{cases} \tag{3.40}$$

Now let $J \subseteq I$ be a nontrivial subinterval. For each $m \in \mathbb{N}$, we have that $T_\beta^{m-1} J \subseteq A_i$ for some $i \in \{0, 1 \ldots, \lfloor \beta \rfloor\}$ implies $\lambda(T_\beta^m J) = \beta \cdot \lambda(T_\beta^{m-1} J)$. From this it follows that there exists $k \in \mathbb{N}$ such that $T_\beta^k J$ contains an endpoint in $(0, 1)$ of one of the intervals $A_1, \ldots, A_{\lfloor \beta \rfloor}$. Then there exists $a \in (0, 1)$ such that $[0, a) \subseteq T_\beta^{k+1} J$, from which we conclude $T_\beta^n J = [0, 1)$ for $n \geq k + 1$ sufficiently large. We conclude that T_β meets the assumptions of Proposition 3.16.

In 1957, Rényi [57] proved the first important result on the existence of an acipm for piecewise onto transformations. This result is now considered to be a folklore theorem. Below we state a version of this theorem that is based on Theorem 2.2 in Chapter 5 of [19]. We need the following definition.

Definition 3.19. *Let $T : I \to I$ be piecewise C^k-monotonic and expanding w.r.t. a finite or countable interval partition $\{I_i\}$. We say that T is* Markov *if $k \geq 2$ and the following conditions are met:*

(i) There exist $C > 0$ and $l \in \{1, \ldots, k-1\}$ such that for each i and $x, y \in I_i$ we have

$$\left| \frac{T'(x)}{T'(y)} - 1 \right| \leq C \cdot |T(x) - T(y)|^l, \tag{3.41}$$

(ii) There exists $\gamma > 0$ such that $\lambda(T(I_i)) \geq \gamma$ for each i,

(iii) If $I_i \cap T(I_j) \neq \emptyset$, then $I_i \subseteq T(I_j)$.

Remark 3.20. Condition (iii) implies that for each j there exists a collection $\mathcal{A} \subseteq \{I_i\}$ such that $T(I_j) = \bigcup_{I \in \mathcal{A}} I$. In other words, an interval cannot be mapped only partly

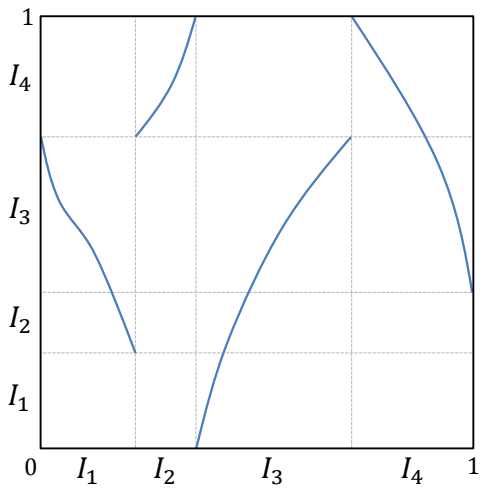

Figure 3.3: Example of a Markov transformation on I

to some other interval, which is why we call T Markov. See Figure 3.3 for a typical example of a Markov transformation.

Theorem 3.21. (Folklore Theorem) *Let $T : I \to I$ be a Markov transformation w.r.t. a finite or countable interval partition $\{I_i\}$. Furthermore, suppose that for every i and j there exists $n \in \mathbb{N}$ such that $I_i \subseteq T^n(I_j)$. Then T admits a unique acipm μ. Furthermore, μ satisfies*

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\mu}{d\lambda} \leq M \tag{3.42}$$

and (T, μ) is exact.

Proof: See e.g. Theorem 2.2 in Chapter 5 of [19]. □

Remark 3.22. The conclusion of Theorem 3.21 remains true under the conditions (i'), (ii') and (iii), where

(i') $\sup_{I_i} \sup_{x,y \in I_i} \left| \frac{T''(x)}{T'(y)^2} \right| < \infty$,

(ii') The set $\{T(I_i)\}$ is finite,

and where (iii) is as in Definition 3.19. This result can be found in e.g. [2, 3, 7, 8]. Condition (i') is known as *Adler's condition*. Moreover, conditions (i) and (i') are each sometimes referred to as T having *bounded distortion*. In fact, there are many variations in the literature on the definition of bounded distortion of a transformation (see e.g. Section 2.2 in [19] for an overview), each with a corresponding version of the Folklore Theorem. Finally, note that (i'), (ii) and (ii') are always satisfied if $\{I_i\}$ consists of finitely many intervals.

3.6 The invariant density for the LSV map

For each $\alpha \in (0, \infty)$, let $T_\alpha : I \to I$ be given by

$$T_\alpha(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & x \in [0, \frac{1}{2}], \\ 2x - 1 & x \in (\frac{1}{2}, 1] \end{cases} \tag{3.43}$$

(see Figure 3.4). Note that T_α is nonexpanding and has a neutral fixed point at zero. Members of this family are called Liverani-Saussol-Vaienti (LSV) maps because they were first studied in [45]. In this paper, Liverani, Saussol and Vaienti showed (among other results) that T_α admits an acipm if $\alpha \in (0, 1)$ with corresponding density $h_\alpha = O(x^{-\frac{1}{\alpha}})$ for x near zero. (In particular, h_α is not of bounded variation. Compare with Proposition 3.8 for the expanding case.) Moreover, this is the only acipm for T_α (see Theorem 1 in [50]). On the other hand, for the case that $\alpha \in [1, \infty)$ it follows from e.g. [55] that T_α admits an infinite σ-finite acim with again corresponding density $h_\alpha = O(x^{-\frac{1}{\alpha}})$ for x near zero.

In this section we recover the above results (except the asymptotic behavior of h_α near zero) using a method based on Section 3 in [35]. More precisely, let $\mathcal{H}$ denote the set of measurable functions $f : I \to [0, \infty]$ (modulo being Lebesgue almost equal everywhere) for which the measure μ given by $\mu(A) = \int_A f d\lambda$ is σ-finite. Recall from Remark 2.30 that $P_{T_\alpha} : \mathcal{H} \to \mathcal{H}$ is well defined. We shall prove the following theorem.

Theorem 3.23. *For each $\alpha \in (0, \infty)$, there exists $h_\alpha \in \mathcal{H}$ such that*

$$\{f \in \mathcal{H} : P_{T_\alpha} f = f\} = \{a h_\alpha : a \geq 0\}. \tag{3.44}$$

Moreover, $h_\alpha \in L^1(\lambda)$ if and only if $\alpha \in (0, 1)$.

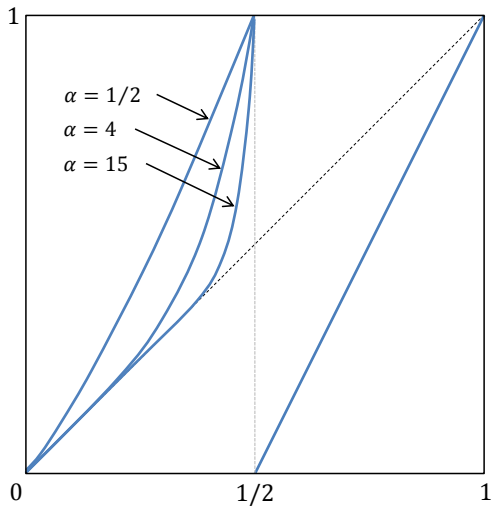

Figure 3.4: The LSV map T_α for several values of α (adapted from [11])

Write L_α and R for the left and right branch of T_α, respectively. That is,

$$\begin{cases} L_\alpha(x) = x(1+2^\alpha x^\alpha) & x \in [0, \frac{1}{2}], \\ R(x) = 2x - 1 & x \in (\frac{1}{2}, 1]. \end{cases} \tag{3.45}$$

We view R as the nice expanding branch and L_α as the complex nonexpanding branch. In order to obtain an invariant density for T_α, the idea is now to construct an *expanding* transformation S_α by properly composing iterations of L_α with R. Applying the theory of the previous sections to S_α, it then remains to find a one-to-one correspondence between the invariant densities of T_α and those of S_α.

More precisely, let $\{I_n^\alpha\}_{n\geq 1}$ be the countable interval partition of I given by $I_1^\alpha = I_1 = (\frac{1}{2}, 1]$ and $I_n^\alpha = (L_\alpha^{-n+1}\frac{1}{2}, L_\alpha^{-n+2}\frac{1}{2}]$ for $n \geq 2$ (see Figure 3.5). Putting $I_0^\alpha = (0, 1]$, then obviously $T_\alpha(I_n^\alpha) = I_{n-1}^\alpha$ for all $n \geq 1$. The *first passage time* $\tau : I \to \mathbb{N}$ *in* I_1 is given by

$$\tau(x) = 1 + \min\{n \geq 0 : T_\alpha^n(x) \in I_1^\alpha\}. \tag{3.46}$$

Note that $I_n^\alpha = \{x \in I : \tau(x) = n\}$ for all $n \geq 1$. Now define $S_\alpha : I \to I$ as $S_\alpha(0) = 0$, $S_\alpha(1) = 1$ and

$$S_\alpha(x) = T_\alpha^{\tau(x)}(x) = R \circ L_\alpha^{n-1}(x) \quad \text{for } x \in I_n^\alpha \text{ and } n \geq 1 \tag{3.47}$$

(see Figure 3.6). We have the following lemma.

Lemma 3.24. *For each $\alpha \in (0, \infty)$, there exists $f_\alpha \in L^1(\lambda)$ such that*

$$\{f \in \mathcal{H} : P_{S_\alpha} f = f\} = \{a f_\alpha : a \geq 0\}. \tag{3.48}$$

Moreover, f_α is bounded and bounded away from zero.

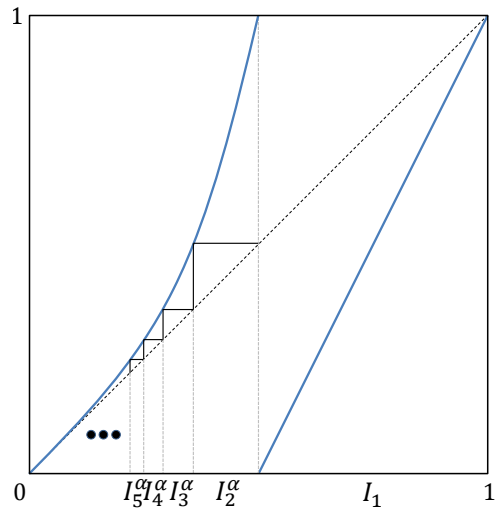

Figure 3.5: Sketch of the partition $\{I_n^\alpha\}_{n\geq 1}$

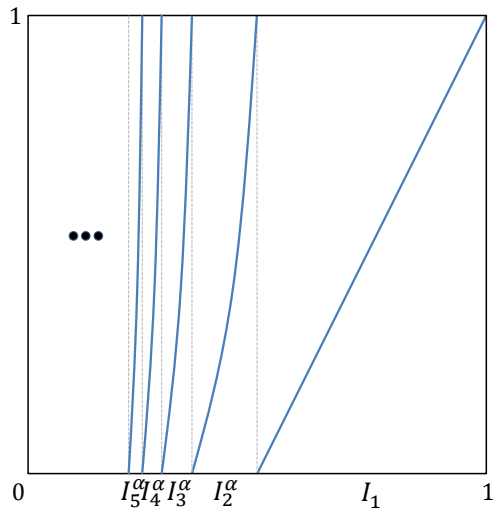

Figure 3.6: Sketch of the map S_α given in (3.47)

Proof: First of all, we can extend the proof of Proposition 3.8 to conclude that every fixed point of P_{S_α} in $\mathcal{H}$ is of bounded variation, so in particular is an element of $L^1(\lambda)$. Hence, it would be sufficient if we could apply the Folklore Theorem to S_α. Note that the only difficult task is to show that S_α has bounded distortion in the sense of (3.41). This has been shown in the proof of Proposition 3.3 in [45] (see regime (2)) for the case that $\alpha \in (0,1)$. Here it is used that the points $y_n = L_\alpha^{-n+1}(\frac{1}{2})$ with $n \geq 1$ satisfy $y_n \leq Cn^{-\frac{1}{\alpha}}$ for some $C > 0$ (see Lemma 3.2 in [45]). In fact, this bound for y_n can be derived for all $\alpha \in (0,\infty)$ using that $y_n = \frac{1}{2(\alpha n)^{1/\alpha}} + O(\frac{\log n}{n^{1+1/\alpha}})$ (see the proof of Theorem 31 in [10]), so (3.41) can be derived for $\alpha \in [1,\infty)$ as well using the method in the proof of Proposition 3.3 in [45]. The result now follows from the Folklore Theorem. □

We define the operators $B_\alpha, A : \mathcal{H} \to \mathcal{H}$ by

$$B_\alpha h(x) = \frac{h(L_\alpha^{-1}x)}{|L_\alpha'(L_\alpha^{-1}x)|}, \qquad Ah(x) = \frac{h(R^{-1}x)}{|R'(R^{-1}x)|}. \tag{3.49}$$

Then we have

$$P_{T_\alpha} = A + B_\alpha. \tag{3.50}$$

We have the following two lemmata.

Lemma 3.25. *For each $\alpha \in (0,\infty)$, the transfer operator P_{S_α} of S_α satisfies*

$$P_{S_\alpha} f = \sum_{k=0}^{\infty} AB_\alpha^k f, \qquad \lambda\text{-a.e.} \tag{3.51}$$

for each $f \in \mathcal{H}$. Moreover, each $f \in \mathcal{H}$ satisfies $\sum_{k=0}^{\infty} B_\alpha^k f(x) < \infty$ for λ-a.e. $x \in I$.

Proof: Note that $S_\alpha^{-1}x = \{L_\alpha^{-k}R^{-1}x : k \geq 0\}$ for each $x \in I$. Let $f \in \mathcal{H}$. For λ-a.e. $x \in I$ we have

$$\begin{aligned} P_{S_\alpha} f(x) &= \sum_{y \in S^{-1}x} \frac{f(y)}{|S'(y)|} = \sum_{k=0}^{\infty} \frac{f(L_\alpha^{-k}R^{-1}x)}{|(RL_\alpha^k)'(L_\alpha^{-k}R^{-1}x)|} \\ &= \sum_{k=0}^{\infty} \frac{1}{|R'(R^{-1}x)|} \cdot \frac{f(L_\alpha^{-k}R^{-1}x)}{|(L_\alpha^k)'(L_\alpha^{-k}R^{-1}x)|} \\ &= \sum_{k=0}^{\infty} \frac{(B_\alpha^k f)(R^{-1}x)}{|R'(R^{-1}x)|} = \sum_{k=0}^{\infty} A(B_\alpha^k f)(x). \end{aligned}$$

Because $P_{S_\alpha} f \in \mathcal{H}$, we see from this that $\sum_{k=0}^{\infty} B_\alpha^k f(R^{-1}x) < \infty$ for λ-a.e. $x \in I$, i.e. $\sum_{k=0}^{\infty} B_\alpha^k f(x) < \infty$ for λ-a.e. $x \in I_1$. Suppose now that for some $n \in \mathbb{N}$ we have $\sum_{k=0}^{\infty} B_\alpha^k f(x) < \infty$ for λ-a.e. $x \in I_n$. Then

$$\sum_{k=0}^{\infty} \frac{B_\alpha^k f(L_\alpha^{-1}x)}{|L_\alpha'(L_\alpha^{-1}x)|} = \sum_{k=1}^{\infty} B_\alpha^k f(x) < \infty \tag{3.52}$$

for λ-a.e. $x \in I_n$, which gives $\sum_{k=0}^{\infty} B_\alpha^k f(x) < \infty$ for λ-a.e. $x \in I_{n+1}$. □

Lemma 3.26. *Let $\alpha \in (0,\infty)$, and let f_α be as in Lemma 3.24. Then $h_\alpha := \sum_{k=0}^{\infty} B_\alpha^k f_\alpha$ is an element of $\mathcal{H}$. Moreover, $h_\alpha \in L^1(\lambda)$ if and only if $\alpha \in (0,1)$.*

Proof: Define $y_0 = 1$ and $y_n = L_\alpha^{-n+1}(\frac{1}{2})$ for $n \geq 1$ (for convenience, we omit in the notation that y_n depends on α for $n \geq 2$). Then $I_n^\alpha = [y_n, y_{n-1})$ for each $n \geq 1$. From the proof of Theorem 31 in [10] it follows that

$$y_n = \frac{1}{2(\alpha n)^{1/\alpha}} + O\Big(\frac{\log n}{n^{1+1/\alpha}}\Big), \tag{3.53}$$

from which we can derive

$$T'_\alpha y_n = 1 + \xi\frac{1}{n} + O\Big(\frac{\log n}{n}\Big), \tag{3.54}$$

where $\xi = \frac{\alpha+1}{\alpha}$. Furthermore, since L_α is convex, we have

$$L'_\alpha(y_n) \leq L'_\alpha(x) \leq L'_\alpha(y_{n-1}), \qquad x \in I_n^\alpha = [y_n, y_{n-1}) \tag{3.55}$$

for each $n \geq 1$. Now let $M_\alpha > 0$ such that

$$\frac{1}{M_\alpha} \leq f_\alpha \leq M_\alpha. \tag{3.56}$$

Combining (3.55) and (3.56) yields

$$B_\alpha^k f_\alpha(x) \leq \frac{M_\alpha}{|L'_\alpha(y_{n+1})| \cdots |L'_\alpha(y_{n+k})|} \qquad \text{for } x \in I_n^\alpha,\ n \geq 1 \text{ and } k \geq 1. \tag{3.57}$$

Moreover, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$ and $k \geq 1$ it follows that

$$\begin{aligned}
\prod_{i=1}^k \frac{1}{|L'_\alpha(y_{n+i})|} &= e^{-\sum_{i=1}^k \log\left(1+\xi\frac{1}{n+i}+O\left(\frac{\log(n+i)}{n+i}\right)\right)} && \big(\text{using (3.54)}\big)\\
&\leq Ce^{-\xi\sum_{i=1}^k \frac{1}{n+i}} && \big(\text{using } \log(1+x) = x + O(x^2)\big)\\
&\leq C'e^{-\xi(\log(n+(k+1))-\log(n+1))} && \Big(\text{using } \sum_{i=1}^k \frac{1}{n+i} \geq \int_1^{k+1} \frac{1}{n+x}dx\Big)\\
&= C'\Big(\frac{n+1}{n+(k+1)}\Big)^\xi
\end{aligned}$$

for suitable constants $C, C' \in (0, \infty)$. This gives together with (3.57) and the fact that $\xi > 1$ for each $x \in I_n^\alpha$ and $n \geq n_0$ that

$$\begin{aligned}
h_\alpha(x) = \sum_{k=0}^\infty B_\alpha^k f(x) &\leq M_\alpha + M_\alpha C'(n+1)^\xi \sum_{k=n}^\infty \frac{1}{(k+1)^\xi}\\
&\leq M_\alpha + M_\alpha C'(n+1)^\xi\Big((n+1)^{-\xi} + \frac{1}{\xi-1}(n+1)^{-\xi+1}\Big) \leq C''n
\end{aligned} \tag{3.58}$$

and for each $x \in I_n^\alpha$ and $n = 1, \ldots, n_0 - 1$ that

$$h_\alpha(x) \leq \Big(\sum_{k=1}^{n_0-n-1} \prod_{i=1}^k \frac{M_\alpha}{|L'_\alpha(y_{n+i})|}\Big) + \frac{C''n_0}{|L'_\alpha(y_{n+1})| \cdots |L'_\alpha(y_{n_0-1})|} \tag{3.59}$$

for a suitable constant $C'' \in (0, \infty)$. In particular, it follows that $\int_{I_n^\alpha} h_\alpha d\lambda < \infty$ for each $n \geq 1$, so $h_\alpha \in \mathcal{H}$.

Furthermore, because

$$\frac{1}{n^{1/\alpha}} - \frac{1}{(n+1)^{1/\alpha}} = \frac{1}{\alpha}\int_n^{n+1} x^{-(1+1/\alpha)}dx \leq \frac{1}{\alpha}n^{-(1+1/\alpha)}, \tag{3.60}$$

we get $\lambda(I_n) = y_n - y_{n+1} = O(\frac{\log n}{n^{1+1/\alpha}})$. Together with (3.58) and (3.59) we obtain that there exist $M > 0$ and $n_1 \in \mathbb{N}$ such that for all $n \geq n_1$ we have $\int_{I_n^\alpha} h_\alpha d\lambda \leq M\frac{\log n}{n^{1/\alpha}}$, so

$$\int_I h_\alpha d\lambda \leq \sum_{n=1}^{\infty}\int_{I_n^\alpha} h_\alpha d\lambda < \infty \tag{3.61}$$

for $\alpha \in (0,1)$. On the other hand, using the lower bounds in (3.55) and (3.56) we can in the same way derive that there exists a constant $M' > 0$ such that for sufficiently large n we have $\int_{I_n^\alpha} h_\alpha d\lambda \geq M'\frac{\log n}{n^{1/\alpha}}$, and thus, for $\alpha \geq 1$,

$$\int_I h_\alpha d\lambda \geq \sum_{n=1}^{\infty}\int_{I_n^\alpha} h_\alpha d\lambda = \infty. \tag{3.62}$$

□

We can now complete the proof of Theorem 3.23:

Proof (Theorem 3.23): Let $\alpha \in (0,\infty)$. First of all, it follows from the previous lemmata that h_α is a fixed point of P_{T_α}:

$$\begin{aligned} P_{T_\alpha}h_\alpha = (A+B_\alpha)h_\alpha &= A\Big(\sum_{k=0}^{\infty} B_\alpha^k f_\alpha\Big) + B_\alpha\Big(\sum_{k=0}^{\infty} B_\alpha^k f_\alpha\Big) \\ &= \Big(\sum_{k=0}^{\infty} AB_\alpha^k\Big)f_\alpha + \sum_{k=1}^{\infty} B_\alpha^k f_\alpha = f_\alpha + \sum_{k=1}^{\infty} B_\alpha^k f_\alpha = h_\alpha. \end{aligned}$$

Now let $g \in \mathcal{H}$ be any fixed point of P_{T_α}. Then $Ag = g - B_\alpha g \in \mathcal{H}$ is a fixed point of P_{S_α}, because

$$P_{S_\alpha}(Ag) = P_{S_\alpha}(g) - P_{S_\alpha}(B_\alpha g) = \sum_{k=0}^{\infty} AB_\alpha^k g - \sum_{k=1}^{\infty} AB_\alpha^k g = Ag, \qquad \lambda\text{-a.e.} \tag{3.63}$$

Lemma 3.24 yields $Ag = af_\alpha$ for a certain $a \geq 0$, so with Lemma 3.25 we get

$$\begin{aligned} g &= \sum_{k=0}^{\infty} B_\alpha^k g - \sum_{k=1}^{\infty} B_\alpha^k g = \sum_{k=0}^{\infty} B_\alpha^k (g - B_\alpha g) \\ &= \sum_{k=0}^{\infty} B_\alpha^k Ag = a\sum_{k=0}^{\infty} B_\alpha^k f_\alpha = ah_\alpha, \qquad \lambda\text{-a.e.} \end{aligned}$$

Combining this with Lemma 3.26 yields the result. □

Remark 3.27. It is shown in Section 2 of [50] that for $\alpha \in (0,1)$ the acipm of T_α with (normalized) density h_α is ergodic. Moreover, the result in Theorem 3.23 is proven in [50] for a more general class $\mathcal{T}_\alpha$ $(0 < \alpha < 1)$ of transformations on I by generalizing the method of Liverani, Saussol and Vaienti in [45]. Each transformation $T_\alpha : I \to I$ in this class consists of two branches, both increasing, convex, C^1 and onto I, with $T_\alpha(0) = 0$ and $T'_\alpha(x) = 1 + Cx^\alpha + o(x^\alpha)$ for x close to zero.

Chapter 4

Invariant Densities for Random Piecewise Monotonic Interval Maps

4.1 Introduction

Let us now consider a *random* dynamical system on $(I, \mathcal{B}, \lambda)$. For that, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space (the *base*) and let $\varphi : \Omega \to \Omega$ (the *base map*) be measure preserving and ergodic w.r.t. $\mathbb{P}$. For each $\omega \in \Omega$ we consider a piecewise monotonic interval map $T_\omega : I \to I$ and we suppose throughout this chapter that the map $T : \Omega \times I \to I$ given by $T(\omega, x) = T_\omega x$ is measurable. We are then considering orbits of the form

$$x \mapsto T_\omega x \mapsto T_{\varphi\omega} T_\omega x \mapsto T_{\varphi^2\omega} T_{\varphi\omega} T_\omega x \mapsto \dots . \tag{4.1}$$

As in the deterministic situation, we are interested in the long-term average behavior of these random orbits. However, in general there is no measure on I that is simultaneously invariant under all the maps $\{T_\omega : \omega \in \Omega\}$. As an analogue of (3.1), we instead consider *S.R.B. measures* (Sinai-Ruelle-Bowen measures) as in [12]:

A probability measure ν on I is S.R.B. for the random dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$ if for $\mathbb{P}$-a.a. $\omega \in \Omega$ the set $B_\omega(\nu)$ of points $x \in I$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T_{\varphi^{k-1}\omega} \circ \dots \circ T_\omega x) = \int_I f d\nu \quad \text{for all } f \in C(I) \tag{4.2}$$

satisfies $\lambda\big(B_\omega(\nu)\big) > 0$.

To obtain such measures, we consider the skew product

$$F = F_{\varphi,T} : \Omega \times I \to \Omega \times I, \quad (\omega, x) \mapsto (\varphi\omega, T_\omega x). \tag{4.3}$$

Note that F is measurable because T is measurable. Iterating (ω, x) under F, observe that we obtain the random orbit in (4.1) by projecting on I. The following theorem is proven by Buzzi [12].

Theorem 4.1. (Proposition 4.1 in [12]) *Suppose Ω is a compact space and $\mathcal{F}$ is the Borel σ-algebra on Ω.*[1] *Furthermore, suppose μ is a probability measure on $\Omega \times I$ that is invariant and ergodic w.r.t. F and absolutely continuous w.r.t. $\mathbb{P} \otimes \lambda$. Then the projection of μ on I given by $\nu(A) = \mu(\Omega \times A)$ for $A \in \mathcal{B}$ is an S.R.B. measure for $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$.*

This motivates the question under what conditions there exists a probability measure μ on $\Omega \times I$ that is invariant w.r.t. F and absolutely continuous w.r.t. $\mathbb{P} \otimes \lambda$. Similar as in Chapter 3, we say in this case that μ is an acipm for F (or acim if μ is a complex measure). We need the following proposition:

Proposition 4.2. *The skew product F is nonsingular w.r.t. $\mathbb{P} \otimes \lambda$.*

Proof: Let $A \subseteq \Omega \times I$ Borel. Put

$$A_\omega = \{x \in I : (\omega, x) \in A\}, \qquad \omega \in \Omega. \tag{4.4}$$

Note that $A_\omega \in \mathcal{B}$ for each $\omega \in \Omega$. Then $A = \bigcup_{\omega \in \Omega} \{\omega\} \times A_\omega$, and

$$\mathbb{P} \otimes \lambda(A) = \int_\Omega \int_I 1_A(\omega, x) d\lambda(x) d\mathbb{P}(\omega) = \int_\Omega \lambda(A_\omega) d\mathbb{P}(\omega). \tag{4.5}$$

Furthermore, we have

$$F^{-1}A = \bigcup_{\omega \in \Omega} F^{-1}\big(\{\omega\} \times A_\omega\big) = \bigcup_{\omega \in \Omega} \Big(\bigcup_{\tilde{\omega} \in \varphi^{-1}\omega} \{\tilde{\omega}\} \times T_{\tilde{\omega}}^{-1} A_\omega \Big) = \bigcup_{\omega \in \Omega} \{\omega\} \times T_\omega^{-1} A_{\varphi\omega}, \tag{4.6}$$

which gives

$$\mathbb{P} \otimes \lambda(F^{-1}A) = \int_\Omega \lambda(T_\omega^{-1} A_{\varphi\omega}) d\mathbb{P}(\omega). \tag{4.7}$$

Now suppose $\mathbb{P} \otimes \lambda(A) = 0$. Then combining (4.5) with Theorem 2.7 yields $\lambda(A_{\varphi\omega}) = 0$ for $\mathbb{P}$-a.a. $\omega \in \Omega$. Since each T_ω is nonsingular w.r.t. λ, this gives $\lambda(T_\omega^{-1} A_{\varphi\omega}) = 0$ for $\mathbb{P}$-a.a. $\omega \in \Omega$. Together with (4.7) it follows that $\mathbb{P} \otimes \lambda(F^{-1}A) = 0$. □

Hence, F admits a corresponding transfer operator $P_{F,\mathbb{P}\otimes\lambda} : L^1(\mathbb{P} \otimes \lambda) \to L^1(\mathbb{P} \otimes \lambda)$ that we simply denote as P_F. In view of Proposition 2.33 we are thus interested in the fixed points of P_F, which are the *invariant densities* for F in the sense of (1.30).

In Sections 4.2-4.4 we consider the setting that the base Ω is equal to the product space $\Omega_E = E^{\mathbb{N}}$ with E a Polish space. Moreover, writing $\omega = (\omega_1, \omega_2, \ldots)$ for $\omega \in \Omega_E$, we assume that $T_\omega = T_{\omega_1}$ for each $\omega \in \Omega_E$, and that the base map φ is equal to the left shift σ on Ω_E, i.e. $\sigma\omega = \tilde{\omega}$ where $\tilde{\omega}_n = \omega_{n+1}$.

We consider in Section 4.2 the case that $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is a Bernoulli shift and discuss results from Morita [47, 49] and Pelikan [53] that generalize the results in Sections 3.3 and 3.4 to random i.i.d. compositions of piecewise monotonic maps that are *expanding in mean*. Moreover, we give the random covering property from [4] that implies that an

[1] For the proof of Theorem 4.1 we apply the result of Theorem 2.14 to the pair (μ, F). It is in this step that we need the compactness of Ω.

invariant density h for $F_{\sigma,T}$, if it exists, is (up to normalization) unique and is bounded and bounded away from zero.

As an example of a random system that is not expanding in mean, we consider in Section 4.3 the random i.i.d. compositions of two LSV maps T_α and T_β given by (3.43), where $\alpha \in (0,1)$ and $\beta \geq 1$. Here, at each time point T_α is applied with probability $p \in (0,1)$ and T_β is applied with probability $1-p$. We prove that there exists an invariant density for the corresponding skew product by generalizing the proof for the deterministic case (i.e. $p = 1$) discussed in [45]. Moreover, we propose a second way to prove this with the method of inducing w.r.t. the first passage time from Section 3.6.

In Section 4.4 we generalize the results from Section 4.2 to the case that $(\Omega, \mathcal{F}, \mathbb{P}, \sigma)$ is a Markov shift. Part of these results are from Kowalski [40] and Froyland [28].

Finally, in Section 4.5 we consider the setting where the base $(\Omega, \mathcal{F}, \mathbb{P})$ is an abstract probability space and the base map φ is invertible. We review conditions from Buzzi [12] under which $F_{\varphi,T}$ admits an invariant density and an additional covering property from Buzzi [13] under which this is the only invariant density for $F_{\varphi,T}$.

4.2 One-sided Bernoulli Shift as Base

Let E be a Polish space and $\mathcal{E}$ the Borel σ-algebra on E. For each $z \in E$, let $T_z : I \to I$ be a piecewise monotonic interval map. We define the skew product

$$F : \Omega_E \times I \to \Omega_E \times I, \quad (\omega, x) \mapsto (\sigma\omega, T_{\omega_1} x). \tag{4.8}$$

where σ is the left shift on $\Omega_E = E^{\mathbb{N}}$. We assume that the map $E \times I \ni (z, x) \mapsto T_z x$ is measurable (which always holds if E is countable), so that $T : \Omega \times I \to I$ given by $T(\omega, x) = T_{\omega_1} x$ is measurable. Also, let π be a probability measure on $(E, \mathcal{E})$, and take $\mathbb{P} = \pi^{\otimes \mathbb{N}}$ as a probability measure on $(\Omega_E, \mathcal{F})$, where $\mathcal{F}$ is the Borel σ-algebra on $E^{\mathbb{N}}$.

We first recover the result from Morita [49] (see also Lemma 3.2 in [28]) that each acim of F has the form $\mathbb{P} \otimes \nu$, where ν is absolutely continuous w.r.t. λ with density $\frac{d\nu}{d\lambda}$ that is a fixed point of the operator $P_T : L^1(I) \to L^1(I)$ given by

$$P_T f(x) = \int_E P_{T_z} f(x) d\pi(z), \qquad \lambda\text{-a.e.} \tag{4.9}$$

for each $f \in L^1(I)$. In (4.9), P_{T_z} is the transfer operator that is associated to T_z. In this case, we see from (4.9) with $f = \frac{d\nu}{d\lambda}$ together with Fubini's Theorem that such a ν satisfies

$$\nu(A) = \int_E \nu(T_z^{-1} A) d\pi(z), \qquad A \in \mathcal{B}, \tag{4.10}$$

which is a natural generalization of the definition of invariance of a measure w.r.t. a single transformation. We first need two lemmata.

For $k \geq 1$, let $\mathcal{I}_k$ be the linear span of characteristic functions of sets $A \in \mathcal{E}^k$, i.e.

$$\mathcal{I}_k = \Big\{ \sum_{i=1}^{n} a_i 1_{A_i} : a_i \in \mathbb{C}, A_i \in \mathcal{E}^k, i = 1, \dots, n, n \geq 1 \Big\}. \tag{4.11}$$

Furthermore, we define $\mathcal{A}_0 \subseteq L^1(\mathbb{P} \otimes \lambda)$ as

$$\mathcal{A}_0 = \bigcup_{k \geq 1} \{\Omega_E \times I \ni (\omega, x) \mapsto \psi(\omega_1, \dots, \omega_k)\phi(x) : \psi \in \mathcal{I}_k, \phi \in L^1(\lambda)\}. \tag{4.12}$$

Lemma 4.3. *$\mathcal{A}_0$ is dense in $L^1(\mathbb{P} \otimes \lambda)$.*

Proof: We know that the linear span of characteristic functions of $\mathcal{I} \subseteq \mathcal{F} \otimes \mathcal{B}$ with

$$\mathcal{I} = \Big\{ \big(A_1 \times \cdots \times A_k \times E \times E \times \cdots\big) \times B : B \in \mathcal{B}, A_i \in \mathcal{E}, i = 1, \dots, k, k \geq 1 \Big\}$$

is dense in $L^1(\mathbb{P} \otimes \lambda)$ because of e.g. Theorem 4.12 in [20]. (Note that $\mathcal{I}$ is a semiring that generates $\mathcal{F} \otimes \mathcal{B}$.) Now observe that $\mathcal{A}_0$ contains the linear span of characteristic functions of $\mathcal{I}$. □

Lemma 4.4. (Lemma 4.1 in [49]) *Let $\Phi \in \mathcal{A}_0$ be given by $\Phi(\omega, x) = \psi(\omega_1, \dots, \omega_k)\phi(x)$ for some $k \geq 1$, $\psi \in \mathcal{I}_k$ and $\phi \in L^1(\lambda)$. Then for all $n \geq k$ we have*

$$P_F^n \Phi(\omega, x) = \int_{E^n} \psi(z_1, \dots, z_k) \cdot (P_{T_{z_n}} \cdots P_{T_{z_1}} \phi)(x) d\pi^n(z_1, \dots, z_n), \qquad \mathbb{P} \otimes \lambda\text{-a.e.}$$

So if $n \geq k$, then the value of $P_F^n \Phi(\omega, x)$ does not depend on ω for $\mathbb{P}$-a.a. ω.

Proof: For all $n \geq k$, $A \in \mathcal{F}$ and $B \in \mathcal{B}$ we have

$$\begin{aligned}
\int_{A \times B} P_F^n \Phi d\mathbb{P} \otimes \lambda &= \int_{F^{-n}(A \times B)} \Phi d\mathbb{P} \otimes \lambda \\
&= \int_{\Omega_E} \int_I \phi(x)\psi(z_1, \dots, z_k) 1_A(\sigma^n z) 1_B(T_{z_n} \cdots T_{z_1} x) d\lambda(x) d\mathbb{P}(z) \\
&= \mathbb{P}(A) \int_{E^n} \int_I \phi(x)\psi(z_1, \dots, z_k) 1_B(T_{z_n} \cdots T_{z_1} x) d\lambda(x) d\pi^n(z_1, \dots, z_n) \\
&= \mathbb{P}(A) \int_{E^n} \psi(z_1, \dots, z_k) \Big(\int_B (P_{T_{z_n}} \cdots P_{T_{z_1}} \phi)(x) d\lambda(x) \Big) d\pi^n(z_1, \dots, z_n) \\
&= \int_{A \times B} \Big(\int_{E^n} \psi(z_1, \dots, z_k) \cdot (P_{T_{z_n}} \cdots P_{T_{z_1}} \phi)(x) d\pi^n(z_1, \dots, z_n) \Big) d\mathbb{P} \otimes \lambda.
\end{aligned}$$

□

Theorem 4.5. *Let $h \in L^1(\mathbb{P} \otimes \lambda)$. Then $P_F h = h$ if and only if there exists $\tilde{h} \in L^1(I)$ such that $h(\omega, x) = \tilde{h}(x)$ for $\mathbb{P} \otimes \lambda$-a.e. (ω, x) and $\tilde{h}$ is fixed under the operator $P_T : L^1(I) \to L^1(I)$ given by (4.9).*

Proof: Suppose $h(\omega, x) = \tilde{h}(x)$ for $\mathbb{P} \otimes \lambda$-a.e. (ω, x) and $\tilde{h}$ is fixed under P_T. Then from Lemma 4.4 it follows that

$$P_F h(\omega, x) = \int_E P_{T_z} \tilde{h}(x) d\pi(z) = P_T \tilde{h}(x) = \tilde{h}(x) = h(\omega, x), \qquad \mathbb{P} \otimes \lambda\text{-a.e.}$$

Conversely, suppose $P_F h = h$. From Lemma 4.3 it follows that there exist a sequence $\{\Phi_m\}$ in $\mathcal{A}_0$ that converges in $L^1(\mathbb{P} \otimes \lambda)$ to h. For each $m \in \mathbb{N}$, take k_m such that $h_m(\omega, x) := P_F^{k_m} \Phi_m(\omega, x)$ does not depend on ω for $\mathbb{P}$-a.a. ω. Then

$$\begin{aligned}\int_{\Omega_E} \lim_{m\to\infty} \Big(\int_I |h - h_m| d\lambda \Big) d\mathbb{P} &= \lim_{m\to\infty} \int_{\Omega_E \times I} |P_F^{k_m} h - P_F^{k_m} \Phi_m| d\mathbb{P} \otimes \lambda \\ &\leq \lim_{m\to\infty} \int_{\Omega_E \times I} |h - \Phi_m| d\mathbb{P} \otimes \lambda = 0,\end{aligned}$$

so for $\mathbb{P}$-a.a. $\omega \in \Omega_E$ we obtain $h_m(\omega, \cdot)$ converges in $L^1(\lambda)$ to $h(\omega, \cdot)$ as $m \to \infty$. Combining that this limit is λ-a.e. unique with the fact that, for each $m \in \mathbb{N}$, $h_m(\omega, x)$ does not depend on ω for $\mathbb{P}$-a.a. ω yields that $h(\omega, x) = \tilde{h}(x)$ $\mathbb{P} \otimes \lambda$-a.e. for some $\tilde{h} \in L^1(I)$. In particular, $h \in \mathcal{A}_0$ and we obtain from Lemma 4.4 that

$$\tilde{h}(x) = h(\omega, x) = P_F h(\omega, x) = \int_E P_{T_z} \tilde{h}(x) d\pi(z) = P_T \tilde{h}(x), \quad \mathbb{P} \otimes \lambda\text{-a.e.}$$

□

So according to Theorem 4.5 there is a one-to-one relation between the acim's for F and the fixed points of P_T in (4.9). By deriving a Lasota-Yorke inequality for P_T, we shall obtain similar results as in Sections 3.3 and 3.4.

For simplicity, let us from now on assume that E is countable and write $p_j := \pi(j)$ for each $j \in E$. Furthermore, we assume that each T_j ($j \in E$) is finitely piecewise C^2-monotonic. We say that T is *expanding on average* w.r.t. $(p_j)_{j\in E}$ if

$$\Lambda_T := \sum_{j\in E} \frac{p_j}{\theta(T_j)} < 1, \tag{4.13}$$

where $\theta(T_j) = \inf_{x\in I} |T_j'(x)| > 0$.

Theorem 4.6. (Proposition 2.3 in [4]) *Suppose T is expanding on average w.r.t. $(p_j)_{j\in E}$. Then there exist $k \in \mathbb{N}$, $\rho \in (0, 1)$ and $L \in (0, \infty)$ such that*

$$\mathrm{Var}_I(P_T^k f) \leq \rho \mathrm{Var}_I(f) + L\|f\|_1 \qquad \textit{for all } f \in BV(I). \tag{4.14}$$

Consequently, F admits an acipm whose density is of bounded variation.

Proof: From (4.9) it follows that, for each $k \in \mathbb{N}$,

$$P_T^k = \sum_{(\omega_1,\ldots,\omega_k)\in E^k} p_{\omega_1} \cdots p_{\omega_k} P_{T_{\omega_k}} \circ \cdots \circ P_{T_{\omega_1}} = \sum_{(\omega_1,\ldots,\omega_k)\in E^k} p_{\omega_1} \cdots p_{\omega_k} P_{T_{\omega_k} \circ \cdots \circ T_{\omega_1}}. \tag{4.15}$$

Using the subadditivity of $\mathrm{Var}_I(\cdot)$ we thus find for each $k \in \mathbb{N}$ and $f \in BV(I)$, applying Lemma 3.2 to each $T_{\omega_k} \circ \cdots \circ T_{\omega_1}$,

$$\mathrm{Var}_I(P_T^k f) \leq \rho_k \mathrm{Var}_I(f) + L_k \|f\|_{L^1} \tag{4.16}$$

with $\rho_k = \sum_{(\omega_1,\ldots,\omega_k)\in E^k} \frac{2p_{\omega_1} \cdots p_{\omega_k}}{\theta(T_{\omega_k} \circ \cdots \circ T_{\omega_1})}$ and $L_k = \sum_{(\omega_1,\ldots,\omega_k)\in E^k} p_{\omega_1} \cdots p_{\omega_k} L(T_{\omega_k} \circ \cdots \circ T_{\omega_1})$. It follows by the chain rule that $\theta(T_{\omega_k} \circ \cdots \circ T_{\omega_1}) \geq \theta(T_{\omega_1}) \cdots \theta(T_{\omega_k})$, so $\rho_k \leq 2\Lambda_T^k$. Therefore, we can take k large enough so that $\rho := \rho_k < 1$. In exactly the same way as

in the proof of Theorem 3.1 we can construct from (4.14) a fixed point of P_T in $BV(I)$, which yields the result. □

Remark 4.7. Note that (4.13) is satisfied if each T_j is expanding (i.e. $\theta(T_j) > 1$). Moreover, it is possible that T is expanding on average if there exists $j \in E$ such that $\theta(T_j) < 1$ by choosing a suitable probability vector $(p_j)_{j\in E}$.

Remark 4.8. Pelikan [53] showed for the case that E is finite that the Lasota-Yorke inequality (4.14) still holds if

$$\sup_{x\in I} \sum_{j\in E} \frac{p_j}{|T_j'(x)|} < 1. \tag{4.17}$$

Note that this is a weaker condition than requiring that T is expanding on average. An extension of this result is given in [30] to the setting where E is finite and each T_j is allowed to be piecewise monotonic on a *countable* partition $\{I_{i,j}\}$ such that

$$g_j(x) = \begin{cases} \frac{1}{|T_j'(x)|} & \text{if } x \in \bigcup_i \operatorname{Int}(I_{i,j}), \\ 0 & \text{if } x \in I \setminus \bigcup_i \operatorname{Int}(I_{i,j}) \end{cases} \tag{4.18}$$

is of bounded variation (compare with Remark 3.4).[2]

Remark 4.9. The case that E is a general Polish space was first considered by Morita [49], who showed that for a probability measure π on $(E, \mathcal{E})$ the result in Theorem 4.6 is still valid if $E \ni z \mapsto \theta(T_z)^{-1}$ is an element of $L^1(\pi)$ that satisfies

$$\int_E \frac{1}{\theta(T_z)} d\pi(z) < 1. \tag{4.19}$$

This result is further generalized in [48] to the case that each T_z is allowed to have countably many branches under some additional technical assumptions. (For instance, the distortion of random i.i.d. compositions of $\{T_z\}$ should be integrable w.r.t. π.)

Remark 4.10. Finally, the result in Remark 4.8 is further extended by Inoue [34] to the setting where $(E, \mathcal{E}, \pi)$ is a general probability space and where each T_z $(z \in E)$ is allowed to be piecewise monotonic on a countable partition $\{I_{i,z}\}$.[3] For this, it is assumed that there exists a constant $M > 0$ such that

$$g_z(x) = \begin{cases} \frac{1}{|T_z'(x)|} & \text{if } x \in \bigcup_i \operatorname{Int}(I_{i,z}), \\ 0 & \text{if } x \in I \setminus \bigcup_i \operatorname{Int}(I_{i,z}) \end{cases} \tag{4.20}$$

satisfies $\operatorname{Var}_I(g_z) < M$ for π-a.a. $z \in E$, and such that

$$\sup_{x\in I} \int_E g_z(x) d\pi(z) < 1. \tag{4.21}$$

[2]The result by Bahsoun and Góra in [30] is even more general, because $p_j(x)$ is allowed to change as a function of x. That is, it is assumed that $\{p_j(x)\}_{j\in E}$ is a set of position dependent measurable probabilities.

[3]It is furthermore allowed in [34] that π is position dependent being a measure on $E \times I$ such that $d\pi(z, x) = p(z, x) d\nu(z)$ for some probability density function $p : E \times I \to [0, \infty)$ and some measure ν on $(E, \mathcal{E})$. We only consider the simplified version that $p(z, x) = 1$.

Example 4.11. Let $T_0 : I \to I$ denote the Gauss map from (1.7) (see Figure 1.2). We learned in Subsection 1.1.2 that T_0 admits an invariant probability density $h_0 = \frac{1}{\log 2}\frac{1}{1+x}$. Furthermore, let $T_1 : I \to I$ denote the Rényi map from (3.21) (see Figure 3.2). We know from Remark 3.6 that T_1 does not admit an invariant probability density but does admit a σ-finite acim with density $h_1(x) = \frac{1}{x}$. Now let $E = \{0, 1\}$, $p_0 = p \in (0, 1)$ and $p_1 = 1 - p$. It is shown in Proposition 3.1 of [36] that $\{T_0, T_1, p_0, p_1\}$ satisfies the conditions in Remark 4.8. (Note that T_0 and T_1 are not expanding and that T is not expanding on average, but that (4.17) is satisfied.) Hence, if $p \in [0, 1)$, then the corresponding skew product F as given in (4.8) admits an invariant probability density h_p, which is not the case if $p = 1$.

Most of the results in Section 3.4 carry over to the setting of Theorem 4.6, as the next theorem states.

Theorem 4.12. *Suppose that E is countable, and let $(p_j)_{j\in E}$ be a probability vector. Furthermore, assume that $\{T_j\}_{j\in E}$ is a collection of finitely piecewise C^2-monotonic interval maps and that T is expanding on average w.r.t. $(p_j)_{j\in E}$. Then*

1. *the fixed points of P_T are elements of $BV(I)$,*
2. *there exists a biggest acipm $\tilde{\mu} = \mathbb{P} \otimes \tilde{\nu}$ of F in the sense that if μ is an acipm of F, then μ is absolutely continuous w.r.t. $\tilde{\mu}$.*
3. *the restriction $P_{T,BV}$ of P_T to $BV(I)$ satisfies $P_{T,BV} : BV(I) \to BV(I)$ and is quasi-compact,*
4. *the set $M_{ac}(\Omega_E \times I, F)$ of acim's of F is a non-empty finite-dimensional vector space generated by the ergodic acipm's of F.*

Proof: Using the Lasota-Yorke inequality (4.14), the above statements follow in exactly the same way as the proofs of Proposition 3.8, Corollary 3.9, Theorem 3.10 and Theorem 3.11, respectively. □

Remark 4.13. Under the assumptions of Theorem 4.12, it follows both from Lemma 5.4 in [53] and Corollary 7 in [49] that if there exists a T_j that is expanding, then the dimension of $M_{ac}(\Omega_E \times I, F)$ is bounded by the number of discontinuities of T_j.

As in the deterministic setting, we can from the quasi-compactness of $P_{T,BV}$ deduce a number of ergodic properties of F. As an example, we show that if $(F, \tilde{\mu})$ is weakly mixing, then $(F, \tilde{\mu})$ is mixing.

Proposition 4.14. *In addition to the assumptions in Theorem 4.12, assume that $(F, \tilde{\mu})$ is weakly mixing. Then for each $n \in \mathbb{N}$ we have*

$$P^n_{T,BV} g = \Big(\int_I g d\lambda \Big) \frac{d\tilde{\nu}}{d\lambda} + S^n g, \qquad g \in BV(I) \tag{4.22}$$

where for some $q \in (0, 1)$ and $M > 0$ we have for each $n \in \mathbb{N}$ that $\|S^n\|_{BV} \leq Mq^n$.

Proof: This can be shown similarly as done in the proof of Proposition 3.14. □

Corollary 4.15. *Under the assumptions of Proposition 4.14, $(F, \tilde{\mu})$ is mixing.*

Proof: We use the following notation for cylinders of the form

$$[j_1 \cdots j_n] = \{\omega \in \Omega_E : \omega_1 = j_1, \ldots, \omega_n = j_n\}. \tag{4.23}$$

Furthermore, let us write $\tilde{h} = \frac{d\tilde{\nu}}{d\lambda}$. Then for all cylinders $[j_1 \cdots j_n], [l_1, \ldots, l_m] \in \mathcal{F}$ and $A, B \in \mathcal{B}$ we have for all $N > m$ that (with $k = N - m$)

$$\begin{aligned}
&\tilde{\mu}\Big(F^{-N}\big([j_1 \cdots j_n] \times A\big) \cap [l_1 \cdots l_m] \times B\Big) \\
&\quad = \mathbb{P} \times \tilde{\nu}\Big(\bigcup_{i_1 \cdots i_k} [l_1 \cdots l_m i_1 \cdots i_k j_1 \cdots j_n] \times \big((T_{l_1}^{-1} \cdots T_{l_m}^{-1} T_{i_1}^{-1} \cdots T_{i_k}^{-1} A) \cap B\big)\Big) \\
&\quad = \mathbb{P}([j_1 \cdots j_n])\mathbb{P}([l_1 \cdots l_m]) \sum_{i_1 \cdots i_k} p_{i_1} \cdots p_{i_k} \int_I 1_B \cdot 1_{T_{l_1}^{-1} \cdots T_{l_m}^{-1} T_{i_1}^{-1} \cdots T_{i_k}^{-1} A} \tilde{h} d\lambda.
\end{aligned}$$

Moreover, from Proposition 4.14 it follows that

$$\begin{aligned}
&\lim_{k \to \infty} \int_A \sum_{i_1 \cdots i_k} p_{i_1} \cdots p_{i_k} P_{i_k} \cdots P_{i_1}\big(P_{l_m} \cdots P_{l_1}(1_B \tilde{h})\big) d\lambda \\
&\quad = \lim_{k \to \infty} \int_A P_T^k\big(P_{l_m} \cdots P_{l_1}(1_B \tilde{h})\big) d\lambda \\
&\quad = \int_A \Big(\int_I P_{l_m} \cdots P_{l_1}(1_B \tilde{h}) d\lambda\Big) \tilde{h} d\lambda \\
&\quad = \tilde{\nu}(A)\tilde{\nu}(B),
\end{aligned}$$

so we obtain

$$\lim_{N \to \infty} \tilde{\mu}\Big(F^{-N}\big([j_1 \cdots j_n] \times A\big) \cap [l_1 \cdots l_m] \times B\Big) = \tilde{\mu}([j_1 \cdots j_n] \times A) \cdot \tilde{\mu}([l_1 \cdots l_m] \times B).$$

□

Finally, the next proposition is a generalization of Proposition 3.16. The proof remains similar in spirit.

Proposition 4.16. *In addition to the assumptions in Theorem 4.12, assume that $p_j > 0$ for each $j \in E$. Furthermore, suppose that the following* random covering property *holds: For each non-trivial subinterval $J \subseteq I$ there exist $n \in \mathbb{N}$ and a finite set $I_0 \subseteq I$ and $(\omega_1, \ldots, \omega_n) \in E^n$ such that $T_{\omega_n} \circ \cdots \circ T_{\omega_1}(J) = I \backslash I_0$. Then $\tilde{\mu}$ is the only acipm of F and satisfies*

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\tilde{\mu}}{d\lambda} \leq M. \tag{4.24}$$

Moreover, $(F, \tilde{\mu})$ is ergodic.

Remark 4.17. In fact, because the Bernoulli shift $(\sigma, \mathbb{P})$ is exact (see Example 2.23), one can show with the fourth part of Theorem 2.1 in [47] that $(F, \tilde{\mu})$ is exact under the assumptions of Proposition 4.16.

Example 4.18. Any countable family $\{T_j : I \to I\}_{j \in E}$ given by $T_j x = \beta_j x \bmod 1$ and with $\inf_{j \in E} \beta_j > 1$, together with any probability vector $\{p_j\}_{j \in E}$ such that $p_j > 0$ for

each $j \in E$ satisfies the assumptions of Proposition 4.16. This can be shown in a similar way as the deterministic case in Example 3.18.

4.3 Random i.i.d. Compositions of Two LSV Maps

Let us return to the LSV maps discussed in Section 3.6. We now consider *two* LSV maps $\{T_\alpha, T_\beta\}$ with $\alpha \in (0,1)$ and $\beta \geq 1$. At each time step we apply T_α with probability p and T_β with probability $1-p$, independently from the maps that are applied at the other time steps. That is, we consider iterations of the skew product $F : \{\alpha,\beta\}^{\mathbb{N}} \times I \to \{\alpha,\beta\}^{\mathbb{N}} \times I$ given by $F(\omega, x) = (\sigma\omega, T_{\omega_1}x)$ and we put on $\{\alpha,\beta\}^{\mathbb{N}}$ the Bernoulli measure $\mathbb{P}$ with corresponding probability vector $(p, 1-p)$. Recall from Section 3.6 that T_α admits an invariant probability density (the $p = 1$ case) and that T_β does not (the $p = 0$ case). We want to know if in the intermediate region $p \in (0,1)$ the skew product F admits an invariant probability density, or equivalently if the operator $P_T := pP_{T_\alpha} + (1-p)P_{T_\beta}$ has a fixed point in $L^1(I)$. As opposed to the case in Example 4.11, note that we now cannot apply the result in Remark 4.8 because $|T'_\alpha(0)| = |T'_\beta(0)| = 1$. Still, we have the following result:

Theorem 4.19. *Let $\alpha \in (0,1)$, $\beta \geq 1$ and $p \in (0,1)$. Then there exists a locally Lipschitz function $f^* \in L^1(I)$ s.t. $P_T f^* = f^*$ and $f^*(x) \leq ax^{-\alpha}$ with $a \geq 2^\beta p^{-1}(\alpha+2)$.*

Remark 4.20. As explained in Remark 4.2 in [6], a fixed point of P_T can also be obtained using the techniques in [5, 6] (that is, using Young towers) when a *linearized* version of the LSV maps T_α and T_β is considered.

We prove Theorem 4.19 by closely following Section 2 in [45] where the result is shown for $p = 1$. In the following, we set $T_\alpha^{-1}x = \{y_\alpha, y_0\}$ with $y_\alpha \leq y_0$ and $T_\beta^{-1}x = \{y_\beta, y_0\}$, and $\xi_\alpha = (2y_\alpha)^\alpha$ and $\xi_\beta = (2y_\beta)^\beta$. Writing L_α, L_β and R as in (3.45), we then have

$$\begin{aligned} P_T f(x) &= p\frac{f(y_\alpha)}{L'_\alpha(y_\alpha)} + (1-p)\frac{f(y_\beta)}{L'_\beta(y_\beta)} + \frac{f(y_0)}{R'(y_0)} \\ &= p\frac{f(y_\alpha)}{1+(\alpha+1)\xi_\alpha} + (1-p)\frac{f(y_\beta)}{1+(\beta+1)\xi_\beta} + \frac{f(y_0)}{2}. \end{aligned}$$

Let us define the set $\mathcal{C}_0 = \{f \in C^0((0,1]) : f \geq 0, f \text{ decreasing}\}$. Since $x \mapsto y_0(x)$, $x \mapsto y_s(x)$ and $x \mapsto \xi_s(x)$ are increasing for each $s \in \{\alpha,\beta\}$, it follows that $\mathcal{C}_0$ is preserved by P_T, i.e. $P_T\mathcal{C}_0 \subseteq \mathcal{C}_0$. As in [45], we need the following two lemmata.

Lemma 4.21. *The set $\mathcal{C}_1 = \{f \in \mathcal{C}_0 : x \mapsto x^{\beta+1}f(x)$ increasing$\}$ is preserved by P_T.*

Proof: We have

$$\begin{aligned} x^{\beta+1}P_Tf(x) &= p\Big(\frac{L_\alpha y_\alpha}{y_\alpha}\Big)^{\beta+1}\frac{y_\alpha^{\beta+1}f(y_\alpha)}{1+(\alpha+1)\xi_\alpha}+(1-p)\Big(\frac{L_\beta y_\beta}{y_\beta}\Big)^{\beta+1}\frac{y_\beta^{\beta+1}f(y_\beta)}{1+(\beta+1)\xi_\beta} \\ &\quad +\Big(\frac{Ry_0}{y_0}\Big)^{\beta+1}\frac{y_0^{\beta+1}f(y_0)}{2} \\ &= p\frac{(1+\xi_\alpha)^{\beta+1}}{1+(\alpha+1)\xi_\alpha}y_\alpha^{\beta+1}f(y_\alpha)+(1-p)\frac{(1+\xi_\beta)^{\beta+1}}{1+(\beta+1)\xi_\beta}y_\beta^{\beta+1}f(y_\beta) \\ &\quad +\frac{1}{2}\Big(2-\frac{1}{y_0}\Big)^{\beta+1}y_0^{\beta+1}f(y_0). \end{aligned}$$

The result follows by noting that also $\xi\mapsto\frac{(1+\xi)^{\beta+1}}{1+(\alpha+1)\xi}$, $\xi\mapsto\frac{(1+\xi)^{\beta+1}}{1+(\beta+1)\xi}$ and $y\mapsto(2-\frac{1}{y})^{\beta+1}$ are increasing functions. □

Lemma 4.22. *The set* $\mathcal{C}_2=\{f\in\mathcal{C}_1\cap L^1(I): f(x)\le ax^{-\alpha},\int_I fd\lambda=1\}$ *is preserved by* P_T*, provided* a *is chosen large enough.*

Proof: Let $f\in\mathcal{C}_2$. First of all, we have $\int_I P_Tfd\lambda=\int_I fd\lambda=1$ by part (b) of Proposition 2.31. Since $x\mapsto x^{\beta+1}f(x)$ is increasing and f is decreasing, we have

$$x^{\beta+1}f(x)\le f(1)\le\int fd\lambda=1. \tag{4.25}$$

Combining this with $f(x)\le ax^{-\alpha}$ yields

$$\begin{aligned} P_Tf(x) &= p\frac{f(y_\alpha)}{L'_\alpha(y_\alpha)}+(1-p)\frac{f(y_\beta)}{L'_\beta(y_\beta)}+\frac{f(y_0)}{R'(y_0)} \\ &\le p\frac{ay_\alpha^{-\alpha}}{L'_\alpha(y_\alpha)}+(1-p)\frac{ay_\beta^{-\alpha}}{L'_\beta(y_\beta)}+\frac{y_0^{-\beta-1}}{R'(y_0)} \\ &\le\Big\{p\Big(\frac{x}{y_\alpha}\Big)^\alpha\frac{1}{L'_\alpha(y_\alpha)}+(1-p)\Big(\frac{x}{y_\beta}\Big)^\alpha\frac{1}{L'_\beta(y_\beta)}+\frac{1}{a}\frac{x^\alpha}{y_0^{\beta+1}R'(y_0)}\Big\}ax^{-\alpha}. \end{aligned} \tag{4.26}$$

We need to find a such that the term in curly brackets is bounded by 1. First of all, we know for each $\xi\ge-1$ that $(1+\xi)^\alpha\le1+\alpha\xi$ (using that $\alpha\in(0,1)$), so

$$\Big(\frac{x}{y_\beta}\Big)^\alpha\frac{1}{L'_\beta(y_\beta)}=\frac{(1+\xi_\beta)^\alpha}{1+(\beta+1)\xi_\beta}\le\frac{1+\alpha\xi_\beta}{1+(\beta+1)\xi_\beta}\le1. \tag{4.27}$$

Moreover, we have $y_0\ge\frac{1}{2}$ and $\xi_\alpha\le1$, so

$$\frac{x^\alpha}{y_0^{\beta+1}R'(y_0)}\le\frac{(y_\alpha)^\alpha(1+\xi_\alpha)^\alpha}{2^{-\beta-1}\cdot2}\le2^\beta(y_\alpha)^\alpha2^\alpha=2^\beta\xi_\alpha. \tag{4.28}$$

It follows from (4.28) that

$$\begin{aligned} p\Big(\frac{x}{y_\alpha}\Big)^\alpha\frac{1}{L'_\alpha(y_\alpha)}+\frac{1}{a}\frac{x^\alpha}{y_0^{\beta+1}R'(y_0)} &\le p\frac{(1+\xi_\alpha)^\alpha}{1+(\alpha+1)\xi_\alpha}+\frac{2^\beta}{a}\xi_\alpha \\ &\le p\frac{1+\alpha\xi_\alpha+\frac{2^\beta}{ap}\xi_\alpha(1+(\alpha+1)\xi_\alpha)}{1+(\alpha+1)\xi_\alpha}\le p\frac{1+\big(\alpha+\frac{2^\beta(\alpha+2)}{ap}\big)\xi_\alpha}{1+(\alpha+1)\xi_\alpha}, \end{aligned} \tag{4.29}$$

where in the last step we use $1 + (\alpha + 1)\xi_\alpha \leq \alpha + 2$. Combining (4.26), (4.27) and (4.29), we obtain for $a \geq \frac{2^\beta(\alpha+2)}{p}$ that $P_T f(x) \leq ax^{-\alpha}$, which yields the result. □

Remark 4.23. For (4.27) and the second bound in (4.29) we use that $(1+\xi)^\alpha \leq 1+\alpha\xi$, which for positive α is the case if and only if $\alpha \in (0, 1]$. For this reason, we cannot extend this result to the case that $\alpha > 1$ for any $p \in [0, 1]$.

Proof of Theorem 4.19: Let us define $S \subseteq C^0([0, 1])$ as follows:

$$S = \{[0, 1] \ni x \mapsto x^{1+\beta} f(x) : f \in \mathcal{C}_2\}. \tag{4.30}$$

Let $\phi \in S$ be given by $\phi(x) = x^{1+\beta} f(x)$ with $f \in \mathcal{C}_2$. Then for $x \geq y$ we get

$$\begin{aligned} 0 \leq \phi(x) - \phi(y) &\leq (x^{1+\beta} - y^{1+\beta}) f(x) \leq ax^{-\alpha}(1 + \beta) \int_y^x t^\beta dt \\ &\leq a(1 + \beta)|x - y|. \end{aligned} \tag{4.31}$$

From this we see that S is bounded and equicontinuous, so from the Arzelà-Ascoli Theorem (Theorem B.6) it follows that S is compact in $C^0([0, 1])$ w.r.t. the supremum norm. Using that P_T preserves $\mathcal{C}_2$ and that a weighted average of elements in $\mathcal{C}_2$ is also an element of $\mathcal{C}_2$, we therefore obtain that the sequence $\{\phi_n\} \subseteq S$ given by $\phi_n(x) = x^{1+\beta} f_n(x)$ with $f_n = \frac{1}{n} \sum_{i=0}^{n-1} P_T^i f$ has a subsequence $\{\phi_{n_k}\}$ that converges uniformly to some $\phi^* \in C^0([0, 1])$. Now define $f^* \in C^0((0, 1])$ as $f^*(x) = x^{-1-\beta}\phi^*(x)$. Then $\{f_{n_k}\}$ converges pointwise to f^*, and since

$$\sup_{k\in\mathbb{N}} |f_{n_k}(x)| \leq ax^{-\alpha} \text{ and } \int_0^1 x^{-\alpha} dx < \infty, \tag{4.32}$$

it follows that $f^*(x) \leq ax^{-\alpha}$ and that

$$\int_0^1 f^*(x)dx = \lim_{k\to\infty} \int f_{n_k}(x)dx = 1 \tag{4.33}$$

using the Dominated Convergence Theorem. We conclude that $f^* \in \mathcal{C}_2$. In exactly the same way as in (3.18) it can now be shown that $P_T(f^*) = f^*$, and that f^* is locally Lipschitz follows from the fact that for $x \geq y$ we have

$$0 \leq f^*(y) - f^*(x) \leq x^{-1-\beta}(x^{1+\beta} - y^{1+\beta}) f^*(y) \leq x^{-1-\beta} a(1 + \beta)|x - y|, \tag{4.34}$$

where we used that f^* is decreasing, that $x \mapsto x^{1+\beta} f^*(x)$ is increasing and (4.31), respectively. □

Remark 4.24. We already observed in Remark 4.23 that the above proof does not work for $\alpha > 1$. It follows from (4.32) that the same is true for $\alpha = 1$, because $\int_0^1 x^{-1} dx = \infty$.

Remark 4.25. The proof of Theorem 4.19 is almost the same as the one in Section 2 of [45]. Compared to this $p = 1$ case, the new idea in the proof of Theorem 4.19 is how the expression in curly brackets in (4.26) is bounded. Namely, consider the $p = 0$ case: If $p = 0$, then the first term in (4.26) is zero, and bounding the second term by

1 as in (4.27) requires that the third term is zero, or equivalently, that $a = \infty$. On the other hand, if $p \in (0,1)$, we can bound the second term by $1-p$ and we don't need to require that the third term is zero. This allows us to bound the third term together with the first term by p, from which Lemma 4.22 follows. It is worthwhile to investigate if this idea can be used for proving the same result with two maps on I from the class considered in [50] (see Remark 3.27, where we take one map from $\mathcal{T}_\alpha$ and one map from $\mathcal{T}_\beta$, with $0 < \alpha < 1 \leq \beta < \infty$). The main motivation for this is that the $p = 1$ case in [45] has been extended in [50] to maps in $\mathcal{T}_\alpha$ using essentially the same ideas as in [45].

Proposition 4.26. *Let $\alpha \in (0,1)$, $\beta \geq 1$ and $p \in (0,1)$. Then $f^* \in L^1(I)$ from Theorem 4.19 is the unique fixed point of P_T in $L^1(I)$ and the corresponding acipm μ on $\{0,1\}^{\mathbb{N}} \times I$ is ergodic w.r.t. the skew product F.*

Proof: Since f^* has full support on I (because $f^* \in \mathcal{C}_2$), note from the first part of Theorem 2.10 that the result follows if we show that μ is F-ergodic. So suppose that $F^{-1}A = A$ for some $A \in \mathcal{F} \otimes \mathcal{B}$, where $\mathcal{F}$ is the σ-algebra generated by the cylinders in $\Omega_E = \{\alpha, \beta\}^{\mathbb{N}}$. Then it is easy to see that $P_F(1_A f^*) = 1_A f^*$, so it follows from Theorem 4.5 that there exists $C \in \mathcal{B}$ such that

$$1_A = 1_{\Omega_E \times C}, \qquad \mathbb{P} \otimes \lambda\text{-a.e.} \tag{4.35}$$

Then also

$$1_{F^{-1}A} = 1_{F^{-1}(\Omega_E \times C)}, \qquad \mathbb{P} \otimes \lambda\text{-a.e.} \tag{4.36}$$

Using that $F^{-1}A = A$ and that $F^{-1}(\Omega_E \times C) = [\alpha] \times T_\alpha^{-1}C \cup [\beta] \times T_\beta^{-1}C$ (using the notation from (4.23)), (4.35) and (4.36) together yield

$$1_C = 1_{T_s^{-1}C}, \qquad \lambda\text{-a.e.} \tag{4.37}$$

for both $s \in \{\alpha, \beta\}$. From Remark 3.27 we know that T_α admits an ergodic acipm with full support, so it follows from e.g. Theorem 1.6.1 in [15] that $\lambda(C) \in \{0,1\}$. Together with (4.35) we conclude that $\mu(A) \in \{0,1\}$. □

We now state some ideas how to generalize the method in Section 3.6 to obtain in a second way the existence of the fixed point $f^* \in L^1(I)$ for P_T as in Theorem 4.19. Set $E = \{\alpha, \beta\}$ and $\Omega_E = E^{\mathbb{N}}$. For each $\omega \in \Omega_E$, we define the sequence $\{y_{n,\omega}\}_{n\geq 1}$ given by $y_{1,\omega} = \frac{1}{2}$ and $y_{n+1,\omega} = L_{\omega_n}^{-1}(y_{n,\omega})$ for $n \geq 1$. Also, for each $\omega \in \Omega_E$ we let $\{I_{n,\omega}\}_{n\geq 1}$ be the countable interval partition of I given by $I_{1,\omega} = I_1 = (\frac{1}{2}, 1]$ and $I_{n,\omega} = (y_{n,\omega}, y_{n-1,\omega}]$ for $n \geq 2$.

Let $S : \Omega_E \times I \to I$ be given by $S(\omega, x) = S_\omega x$, where each $S_\omega : I \to I$ is piecewise monotonic and given by

$$S_\omega(x) = R \circ L_{\omega_1} \circ \cdots \circ L_{\omega_{n-1}}(x) \qquad \text{for } x \in I_{n,\omega} \text{ and } n \geq 1. \tag{4.38}$$

Note that Ω_E is a Polish space. Now consider the skew product

$$\tilde{F} : (\Omega_E)^{\mathbb{N}} \times I \to (\Omega_E)^{\mathbb{N}} \times I, \quad (\tilde{\omega}, x) \mapsto (\tilde{\sigma}\tilde{\omega}, S_{\tilde{\omega}_1} x), \tag{4.39}$$

where $\tilde{\sigma}$ denotes the left shift on $(\Omega_E)^{\mathbb{N}}$. We take $\tilde{\mathbb{P}} = \mathbb{P}^{\otimes\mathbb{N}}$ as a probability measure on $((\Omega_E)^{\mathbb{N}}, \tilde{\mathcal{F}})$, where $\tilde{\mathcal{F}}$ is the Borel σ-algebra on $(\Omega_E)^{\mathbb{N}}$. From Theorem 4.5 we know that each acim of $\tilde{F}$ has the form $\tilde{\mathbb{P}} \otimes \nu$, where ν is absolutely continuous w.r.t. λ with density $\frac{d\nu}{d\lambda}$ that is a fixed point of the operator $P_S : L^1(I) \to L^1(I)$ given by

$$P_S f(x) = \int_{\Omega_E} P_{S_\omega} f(x) d\mathbb{P}(\omega), \qquad \lambda\text{-a.e.} \tag{4.40}$$

Now, let us write A and B_α for the operators as given in (3.49). We define $B := pB_\alpha + (1-p)B_\beta$. Then we have

$$P_T = A + B. \tag{4.41}$$

Let us set $p_\alpha := p$ and $p_\beta := 1 - p$. We have the following lemma.

Lemma 4.27. *The operator P_S sastisfies*

$$P_S f = \sum_{k=0}^{\infty} AB^k f, \qquad \lambda\text{-a.e.} \tag{4.42}$$

for each $f \in L^1(I)$. Also, each $f \in L^1(I)$ satisfies $\sum_{k=0}^{\infty} B^k f(x) < \infty$ for λ-a.e. $x \in I$.

Proof: Let $f \in L^1(I)$. For λ-a.e. $x \in I$ we have

$$\begin{aligned}
P_S f(x) &= \int_{\Omega_E} P_{S_\omega} f(x) d\mathbb{P}(\omega) = \int_{\Omega_E} \sum_{y \in S_\omega^{-1} x} \frac{f(y)}{|S'_\omega(y)|} d\mathbb{P}(\omega) \\
&= \int_{\Omega_E} \sum_{k=0}^{\infty} \frac{f(L_{\omega_k}^{-1} \cdots L_{\omega_1}^{-1} R^{-1} x)}{|(RL_{\omega_1} \cdots L_{\omega_k})'(L_{\omega_k}^{-1} \cdots L_{\omega_1}^{-1} R^{-1} x)|} d\mathbb{P}(\omega) \\
&= \sum_{k=0}^{\infty} \int_{\Omega_E} A(B_{\omega_1} \cdots B_{\omega_k} f)(x) d\mathbb{P}(\omega) \\
&= \sum_{k=0}^{\infty} \sum_{\omega_1 \cdots \omega_k} p_{\omega_1} \cdots p_{\omega_k} A(B_{\omega_1} \cdots B_{\omega_k} f)(x) = \sum_{k=0}^{\infty} AB^k f(x),
\end{aligned}$$

where the interchange of integral and series is justified by applying the Monotone Convergence Theorem to the positive and negative part of the real and imaginary part of the integrand. The second statement can be shown similarly as done in the proof of Lemma 3.25. □

Using Theorem 4.19 we have the following two results about the operator P_S.

Proposition 4.28. *There exists a real and nonnegative $f \in L^1(I)$ such that $P_S f = f$.*

Proof: Let f^* be as in Theorem 4.19. Then $f := Af^* = f^* - Bf^* \in L^1(I)$ is a fixed point of P_S, because

$$P_S(f) = P_S(f^*) - P_S(Bf^*) = \sum_{k=0}^{\infty} AB^k f^* - \sum_{k=1}^{\infty} AB^k f^* = Af^*, \qquad \lambda\text{-a.e.} \tag{4.43}$$

□

Proposition 4.29. *There exists $f \in L^1(I)$ for which $P_S f = f$ such that $h := \sum_{k=0}^{\infty} B^k f$ is an element of $L^1(I)$ and such that h is a fixed point of P_T.*

Proof: Take as in (4.43) the fixed point $f = Af^*$ of P_S. Then with Lemma 4.27 we get

$$f^* = \sum_{k=0}^{\infty} B^k f^* - \sum_{k=1}^{\infty} B^k f^* = \sum_{k=0}^{\infty} B^k A f^* = \sum_{k=0}^{\infty} B^k f. \tag{4.44}$$

This gives $h = f^* \in L^1(I)$. □

Let us now state some ideas how to prove the above two propositions in a similar way as the proofs of Lemma 3.24 and Lemma 3.26 (so without making use of the result of Theorem 4.19), thus indicating an alternative proof of Theorem 4.19.

First of all, it is clear that S is expanding on average in the sense of (4.19) (taking Ω_E for E and $\mathbb{P}$ for π). As we know from Remark 4.9, the result of Proposition 4.28 now follows if S has suitable distortion bounds such as in [48]. If true, this can be viewed as a natural generalization of the proof of Lemma 3.24. Alternatively, in view of Remark 4.10, one can check if the function $g_\omega(x)$ from (4.20) associated to S has bounded variation uniformly on $\mathbb{P}$-a.a. $\omega \in \Omega_E$. It is not clear either if this is true.

Secondly, suppose that f is as in Proposition 4.28. Then if $h := \sum_{k=0}^{\infty} B^k f$ satisfies $h \in L^1(I)$, then it follows from Lemma 4.27 that h is a fixed point of P_T:

$$\begin{aligned} P_T h = (A+B)h &= A\Big(\sum_{k=0}^{\infty} B^k f\Big) + B\Big(\sum_{k=0}^{\infty} B^k f\Big) \\ &= \Big(\sum_{k=0}^{\infty} AB^k\Big) f + \sum_{k=1}^{\infty} B^k f = f + \sum_{k=1}^{\infty} B^k f = h. \end{aligned} \tag{4.45}$$

Hence, Proposition 4.29 follows if we show that $h \in L^1(I)$. Suppose that there exists $M > 0$ such that $f \leq M$. This is for instance the case if S satisfies the conditions in Remark 4.9 or Remark 4.10 (taking $(\Omega_E, \mathcal{F}, \mathbb{P})$ for $(E, \mathcal{E}, \pi)$), because similar as in the first part of Theorem 4.12 it then follows that f has bounded variation. Then

$$\begin{aligned} \int_I h d\lambda &= \int_I \sum_{k=0}^{\infty} B^k f d\lambda = \int_I \int_{\Omega_E} \sum_{k=0}^{\infty} B_{\omega_1} \cdots B_{\omega_k} f d\mathbb{P}(\omega) d\lambda(x) \\ &= \sum_{n=1}^{\infty} \int_{\Omega_E} \int_{I_{n,\omega}} \sum_{k=0}^{\infty} B_{\omega_1} \cdots B_{\omega_k} f d\lambda d\mathbb{P}(\omega) \\ &\leq M \sum_{n=1}^{\infty} c_n, \end{aligned} \tag{4.46}$$

where we used Fubini's Theorem, and where

$$\begin{aligned} c_n &= \int_{\Omega_E} \int_{I_{n,\omega}} \sum_{k=0}^{\infty} \frac{1}{|(L_{\omega_1} \cdots L_{\omega_k})'(L_{\omega_k}^{-1} \cdots L_{\omega_1}^{-1} x)|} d\lambda d\mathbb{P}(\omega) \\ &\leq \int_{\Omega_E} \lambda(I_{n,\omega}) \Big\{ \sum_{k=0}^{\infty} \frac{1}{L'_{\omega_1}(L_{\omega_1}^{-1} y_{n,\omega}) \cdots L'_{\omega_k}(L_{\omega_k}^{-1} \cdots L_{\omega_1}^{-1} y_{n,\omega})} \Big\} d\mathbb{P}(\omega). \end{aligned} \tag{4.47}$$

We know from Theorem 1.1 in [6] that $y_{n,\omega} \sim \frac{1}{2}(\alpha pn)^{-1/\alpha}$ for $\mathbb{P}$-a.a $\omega \in \Omega$, i.e.

$$\lim_{n\to\infty} \frac{\frac{1}{2}(\alpha pn)^{-1/\alpha}}{y_{n,\omega}} = 1, \qquad \mathbb{P}\text{-a.a. } \omega \in \Omega.$$

Therefore, as for the $p = 1$ case in Section 3.6, it seems reasonable to expect that $\lambda(I_{n,\omega}) = O(\frac{\log n}{n^{1+1/\alpha}})$. Moreover, in a similar way as in the proof of Lemma 3.26 we can argue that the expression in curly brackets in (4.47) is $O(n)$ for each $\omega \in \Omega$. This together with (4.46) and the assumption that $\lambda(I_{n,\omega}) = O(\frac{\log n}{n^{1+1/\alpha}})$ suggests that indeed $\int_I h d\lambda < \infty$. However, we don't even have a bound on $\lambda(I_{n,\omega})$ uniform in $\mathbb{P}$-a.a. $\omega \in \Omega_E$, so the proof seems to be more delicate.

4.4 One-sided Markov Shift as Base

Let us now assume that E is finite, say $E = \{1, \dots, r\}$, and that $(\Omega_E, \mathcal{F}, \sigma, \mathbb{P})$ with $\Omega_E = E^{\mathbb{N}}$ is a one-sided Markov shift (see Example 2.6) given by an irreducible, aperiodic stochastic matrix W and a probability vector $q = (q_1, \dots, q_r)$ such that $qW = q$. Furthermore, as in [28] we write $W^*_{lk} = \frac{W_{kl}q_k}{q_l}$ (which are the entries of the transition matrix of the time-reversed Markov chain). For each $z \in E$, let $T_z : I \to I$ be a piecewise monotonic interval map. For simplicity, we assume that each T_z is finitely piecewise C^2-monotonic. Again, the skew product $F : \Omega_E \times I \to \Omega_E \times I$ is given by $F(\omega, x) = (\sigma\omega, T_{\omega_1}x)$.

We first recover the result from Kowalski [40] (see also Lemma 4.2 in [28]) that each acim μ of F has the form

$$\mu(A) = \int_{\Omega_E} \mu_{\omega_1}(A_\omega) d\mathbb{P}(\omega), \qquad A \in \mathcal{F} \otimes \mathcal{B}, \tag{4.48}$$

where $A_\omega = \{x \in I : (\omega, x) \in A\}$ and where each μ_j $(j \in E)$ is absolutely continuous w.r.t. λ with density h_j such that $(h_1, \dots, h_r)$ is a fixed point of the operator $P_T : \prod_{j=1}^r L^1(\lambda) \to \prod_{j=1}^r L^1(\lambda)$ given by

$$P_T(f_1, \dots, f_r) = \Big(\sum_{k=1}^r W^*_{1k} P_{T_k} f_k, \sum_{k=1}^r W^*_{2k} P_{T_k} f_k, \dots, \sum_{k=1}^r W^*_{rk} P_{T_k} f_k\Big). \tag{4.49}$$

We first need two lemmata.

Again, for $k \geq 1$, let $\mathcal{I}_k$ denote the linear span of characteristic functions of sets $A \in \mathcal{E}^k$, where $\mathcal{E} = 2^E$ is the power set of E. We define $\mathcal{A}_1 \subseteq L^1(\mathbb{P} \otimes \lambda)$ as

$$\mathcal{A}_1 = \bigcup_{k\geq 1} \{\Omega_E \times I \ni (\omega, x) \mapsto \psi(\omega_1, \dots, \omega_k)\phi_{\omega_1}(x) : \psi \in \mathcal{I}_k, \phi_z \in L^1(\lambda), z \in E\}. \tag{4.50}$$

Lemma 4.30. *$\mathcal{A}_1$ is dense in $L^1(\mathbb{P} \otimes \lambda)$.*

Proof: This follows from Lemma 4.3 combined with the fact that $\mathcal{A}_0 \subseteq \mathcal{A}_1$. □

Lemma 4.31. *Let $\Phi \in \mathcal{A}_1$ be given by $\Phi(\omega, x) = \psi(\omega_1, \dots, \omega_k)\phi_{\omega_1}(x)$ for some $k \geq 1$, $\psi \in \mathcal{I}_k$ and $\phi_z \in L^1(\lambda)$ for each $z \in E$. Then for all $n \geq k$ we have*

$$P_F^n \Phi(\omega, x) = \sum_{j_1 \cdots j_n} \psi(j_2, \dots, j_k) W^*_{\omega_1 j_n} W^*_{j_n j_{n-1}} \cdots W^*_{j_2 j_1} P_{T_{j_n}} \cdots P_{T_{j_1}} \phi_{j_1}(x), \quad \mathbb{P} \otimes \lambda\text{-a.e.}$$

So if $n \geq k$, then the value of $P_F^n \Phi(\omega, x)$ does not depend on $(\omega_2, \omega_3, \cdots)$ for $\mathbb{P}$-a.a. ω.

Proof: First of all, we have (using the notation from (4.23))

$$\begin{aligned}
\mathbb{P}([j_1 \cdots j_n l_1 \cdots l_m]) &= q_{j_1} W_{j_1 j_2} \cdots W_{j_{n-1} j_n} W_{j_n l_1} W_{l_1 l_2} \cdots W_{l_{m-1} l_m} \\
&= \frac{q_{j_1} W_{j_1 j_2}}{q_{j_2}} \frac{q_{j_2} W_{j_2 j_3}}{q_{j_3}} \cdots \frac{q_{j_{n-1}} W_{j_{n-1} j_n}}{q_{j_n}} \frac{q_{j_n} W_{j_n l_1}}{q_{l_1}} \cdot q_{l_1} W_{l_1 l_2} \cdots W_{l_{m-1} l_m} \\
&= W^*_{j_2 j_1} W^*_{j_3 j_2} \cdots W^*_{j_n j_{n-1}} W^*_{l_1 j_n} \cdot \mathbb{P}([l_1 \cdots l_m]).
\end{aligned}$$

From this it follows that for all $n \geq k$, $A = [l_1 \cdots l_m] \in \mathcal{F}$ and $B \in \mathcal{B}$ we have

$$\begin{aligned}
&\int_{A \times B} P_F^n \Phi d\mathbb{P} \otimes \lambda \\
&\quad = \int_{F^{-n}(A \times B)} \Phi d\mathbb{P} \otimes \lambda \\
&\quad = \int_{\Omega_E} \int_I \psi(\omega_2, \dots, \omega_k) \phi_{\omega_1}(x) 1_A(\sigma^n \omega) 1_B(T_{\omega_n} \cdots T_{\omega_1} x) d\lambda(x) d\mathbb{P}(\omega) \\
&\quad = \sum_{j_1 \cdots j_n} \int_{[j_1 \cdots j_n]} \int_I \psi(j_2, \dots, j_k) \phi_{j_1}(x) 1_A(\sigma^n \omega) 1_B(T_{j_n} \cdots T_{j_1} x) d\lambda(x) d\mathbb{P}(\omega) \\
&\quad = \sum_{j_1 \cdots j_n} \mathbb{P}([j_1 \cdots j_n l_1 \cdots l_m]) \int_I \psi(j_2, \dots, j_k) \phi_{j_1}(x) 1_B(T_{j_n} \cdots T_{j_1} x) d\lambda(x) \\
&\quad = \sum_{j_1 \cdots j_n} W^*_{l_1 j_n} \cdots W^*_{j_2 j_1} \mathbb{P}([l_1 \cdots l_n]) \int_B \psi(j_2, \dots, j_k) P_{T_{j_n}} \cdots P_{T_{j_1}} \phi_{j_1}(x) d\lambda(x) \\
&\quad = \int_{A \times B} \sum_{j_1 \cdots j_n} \psi(j_2, \dots, j_k) W^*_{\omega_1 j_n} \cdots W^*_{j_2 j_1} P_{T_{j_n}} \cdots P_{T_{j_1}} \phi_{j_1}(x) d\lambda(x) d\mathbb{P}(\omega).
\end{aligned}$$

□

Theorem 4.32. *Let $h \in L^1(\mathbb{P} \otimes \lambda)$. Then $P_F h = h$ if and only if there exists $(\tilde{h}_1, \dots, \tilde{h}_r) \in \prod_{j=1}^r L^1(\lambda)$ such that $h(\omega, x) = \tilde{h}_{\omega_1}(x)$ for $\mathbb{P} \otimes \lambda$-a.e. (ω, x) and $(\tilde{h}_1, \dots, \tilde{h}_r)$ is fixed under the operator P_T given by* (4.49).

Proof: Suppose $h(\omega, x) = \tilde{h}_{\omega_1}(x)$ for $\mathbb{P} \otimes \lambda$-a.e. (ω, x) and $(\tilde{h}_1, \dots, \tilde{h}_r)$ is fixed under P_T. Then from Lemma 4.31 it follows that

$$P_F h(\omega, x) = \sum_{j_1 \in \{1, \dots, r\}} W^*_{\omega_1 j_1} P_{T_{j_1}} \tilde{h}_{j_1}(x) = \tilde{h}_{\omega_1}(x) = h(\omega, x) \qquad \mathbb{P} \otimes \lambda\text{-a.e.}$$

The converse can be proven in the same way as in Theorem 4.5, where we now use Lemma 4.30 and Lemma 4.31. □

So according to Theorem 4.32 there is a one-to-one relation between the acim's for F and the fixed points of P_T in (4.49). We now define $\hat{BV} = \prod_{j=1}^r BV(I)$ and endow it with the norm $\|(f_1,\dots,f_r)\|_{BV} = \max_{j=1,\dots,r} \|f_j\|_{BV}$. Similarly, we endow $\hat{L} = \prod_{j=1}^r L^1(I)$ with the norm $\|(f_1,\dots,f_r)\|_1 = \max_{j=1,\dots,r} \|f_j\|_1$. Furthermore, we define

$$\alpha = \max_{l=1,\dots,r} \alpha_l \quad \text{with} \quad \alpha_l = \sum_{k=1}^r \frac{W^*_{lk}}{\theta(T_k)},$$

where again $\theta(T_j) = \inf_{x\in I} |T_j'(x)| > 0$.

Theorem 4.33. (see Remark 4.4(i) in [28]) *Suppose that $\alpha < 1$. Then there exist $k \in \mathbb{N}$, $\rho \in (0,1)$ and $L \in (0,\infty)$ such that*

$$\|P_T^k(f_1,\dots,f_r)\|_{BV} \le \rho\|(f_1,\dots,f_r)\|_{BV} + L\|(f_1,\dots,f_r)\|_1 \tag{4.51}$$

for all $(f_1,\dots,f_r) \in \hat{BV}$. As a consequence, P_T admits a fixed point in $\hat{BV}$.

Proof: For each $n \in \mathbb{N}$ and $(f_1,\dots,f_r) \in \hat{BV}$, the i-th coordinate of $P_T^n(f_1,\dots,f_r)$ denoted by $\{P_T^n(f_1,\dots,f_r)\}_i$ equals

$$\{P_T^n(f_1,\dots,f_r)\}_i = \sum_{j_1\cdots j_n} W^*_{ij_n} W^*_{j_n j_{n-1}} \cdots W^*_{j_2 j_1} P_{T_{j_n}} \cdots P_{T_{j_1}} f_{j_1}. \tag{4.52}$$

Applying Lemma 3.2 to each $P_{T_{j_n}\circ\cdots\circ T_{j_1}}$ and using that $\theta(T_{j_n}\circ\cdots\circ T_{j_1}) \ge \theta(T_{j_n})\cdots\theta(T_{j_1})$ gives

$$\begin{aligned}
&\mathrm{Var}_I(\{P_T^n(f_1,\dots,f_r)\}_i)\\
&\quad\le 2\cdot \sum_{j_1\cdots j_n} \frac{W^*_{ij_n}}{\theta(T_{j_n})}\cdots\frac{W^*_{j_2j_1}}{\theta(T_{j_1})}\mathrm{Var}_I(f_{j_1}) + \sum_{j_1\cdots j_n} W^*_{ij_n}\cdots W^*_{j_2j_1} L(T_{j_n}\circ\cdots\circ T_{j_1})\|f_{j_1}\|_1\\
&\quad\le 2\cdot\alpha^n\|(f_1,\dots,f_r)\|_{BV} + \sum_{j_1\cdots j_n} W^*_{ij_n}\cdots W^*_{j_2j_1} L(T_{j_n}\circ\cdots\circ T_{j_1})\|(f_1,\dots,f_r)\|_1.
\end{aligned}$$

Therefore, the first statement follows with $\rho := 2\cdot\alpha^n$ for n sufficiently large such that $\rho < 1$, and with $L = 1 + \sum_{j_1\cdots j_n} W^*_{ij_n}\cdots W^*_{j_2j_1} L(T_{j_n}\circ\cdots\circ T_{j_1})$.

We can now construct a fixed point of P_T with (4.51). Let $(f_1,\dots,f_r) \in \hat{BV}$. Iterating (4.51), it follows that, for each $n \in \mathbb{N}$,

$$\|P_T^{kn}(f_1,\dots,f_r)\|_{BV} \le \rho^n\|(f_1,\dots,f_r)\|_{BV} + L\|(f_1,\dots,f_r)\|_1 \sum_{i=0}^{n-1}\rho^i \le M, \tag{4.53}$$

where $M = \|(f_1,\dots,f_r)\|_{BV} + \frac{L\|(f_1,\dots,f_r)\|_1}{1-\rho}$. Now for each $n \in \mathbb{N}$, define

$$(g_1^{(n)},\dots,g_r^{(n)}) = \frac{1}{n}\sum_{i=1}^n P_T^{ki}(f_1,\dots,f_r). \tag{4.54}$$

Combining (4.53) with Lemma A.4 and Helly's First Theorem (Theorem A.18) yields that $\{g_1^{(n)}\}$ contains a subsequence $\{g_1^{(n_l)}\}$ that converges pointwise to some $g_1 \in BV(I)$. In the same way we obtain that $\{g_2^{n_l}\}$ contains a further subsequence that converges pointwise to some $g_2 \in BV(I)$. Continuing in this manner, we conclude

that $\{(g_1^{(n)}, \dots, g_r^{(n)})\}$ contains a subsequence that converges pointwise to some $g = (g_1, \dots, g_r) \in \hat{BV}$. The rest of the proof is analogous to that of Theorem 3.1. □

We have the following analogue of Theorem 4.12.

Theorem 4.34. *Suppose that $\alpha < 1$. Then*

1. *the fixed points of P_T are elements of $\hat{BV}$,*
2. *there exists a biggest acipm $\tilde{\mu}$ of F in the sense that if μ is an acipm of F, then μ is absolutely continuous w.r.t. $\tilde{\mu}$.*
3. *the restriction $P_{T,BV}$ of P_T to $\hat{BV}$ satisfies $P_{T,BV} : \hat{BV} \to \hat{BV}$ and is quasi-compact,*
4. *the set $M_{ac}(\Omega_E \times I, F)$ of acim's of F is a non-empty finite-dimensional vector space generated by the ergodic acipm's of F.*

Proof: The first two statements can be shown by mimicking the proof of Lemma 3.7, Proposition 3.8 and Corollary 3.9 (Note that we need to apply the Kakutani-Yosida Theorem (Theorem B.1) with $X = \hat{L}$, $P = P_T$ and $A = \hat{BV}$.) The quasi-compactness of $P_{T,BV}$ can be shown with the Ionescu-Tulcea and Marinescu Theorem (Theorem B.5) with $V = \hat{L}$ and $W = \hat{BV}$. (The first and fourth condition can be shown in a similar way as has been done in the proof of Proposition 7.2.1 in [9] by r applications of Helly's First Theorem as in the proof of Theorem 4.32.) Finally, the last statement follows in exactly the same way as the proof of Theorem 3.11. □

As in Section 4.2, we now show that if $(F, \tilde{\mu})$ is weakly mixing, then $(F, \tilde{\mu})$ is mixing. We write $\tilde{\mu}(A) = \int_{\Omega_E} \tilde{\mu}_{\omega_1}(A_\omega) d\mathbb{P}(\omega)$ and $\tilde{h} = (\tilde{h}_1, \dots, \tilde{h}_r) \in \hat{BV}$ where each $\tilde{\mu}_j$ ($j \in \{1, \dots, r\}$) is absolutely continuous w.r.t. λ with density $\tilde{h}_j$.

Proposition 4.35. *Suppose that $\alpha < 1$ and that $(F, \tilde{\mu})$ is weakly mixing. Then for each $n \in \mathbb{N}$ we have*

$$P_{T,BV}^n g = \Big(\sum_{i=1}^r q_i \int_I g_i d\lambda \Big) \tilde{h} + S^n g, \qquad g \in \hat{BV}, \tag{4.55}$$

where for some $q \in (0,1)$ and $M > 0$ we have for each $n \in \mathbb{N}$ that $\|S^n\|_{\hat{BV}} \leq M q^n$.

Proof: One can check that the dual of $\hat{L}$ consists of all bounded linear functionals $\psi : \hat{L} \to \mathbb{C}$ such that

$$\psi(g_1, \dots, g_r) = \sum_{i=1}^r \int_I f_i g_i d\lambda, \tag{4.56}$$

where $f_i \in L^\infty(I)$ for each $i \in \{1, \dots, r\}$. Hence, using the same reasoning as in the proof of Proposition 3.14 one can show that

$$P_{T,BV}^n g = \Big(\sum_{i=1}^r \int_I f_i g_i d\lambda \Big) \tilde{h} + S^n g, \qquad g \in \hat{BV} \tag{4.57}$$

where $f_i \in L^\infty(I)$ for each $i \in \{1, \dots, r\}$ and where S is a bounded linear operator on $BV(I)$ such that, for each $n \in \mathbb{N}$, $\|S^n\|_{\hat{BV}} \leq M q^n$ for some $q \in (0,1)$ and $M > 0$. We

now prove that each f_i must be equal to q_i λ-a.e. Fix $i \in \{1, \dots, r\}$, let $A \in \mathcal{B}$ and take $g_i = 1_A$ and $g_l = 0$ for all $l \neq i$. From (4.57) it follows that

$$\Big\{ \lim_{n\to\infty} P^n_{T,BV}(0, \dots, 0, 1_A, 0, \dots, 0) \Big\}_l = \Big(\int_I f_i 1_A d\lambda \Big) \tilde{h}_l, \tag{4.58}$$

where we have in the left-hand side of (4.58) convergence in $\hat{L}$ and 1_A is on the i-th coordinate. Hence, combining (4.52) and (4.58), we have for each $l \in \{1, \dots, r\}$ that

$$\begin{aligned}
\int_A f_i d\lambda &= \int_I \Big(\int_I f_i 1_A d\lambda \Big) \tilde{h}_l d\lambda \\
&= \lim_{n\to\infty} \int_I \sum_{j_2 \cdots j_n} W^*_{l j_n} W^*_{j_n j_{n-1}} \cdots W^*_{j_2 i} P_{T_{j_n}} \cdots P_{T_{j_2}} P_{T_i} 1_A d\lambda \\
&= \lim_{n\to\infty} \sum_{j_2 \cdots j_n} \frac{q_i}{q_l} W_{i j_2} \cdots W_{j_{n-1} j_n} W_{j_n l} \int_A 1 d\lambda = \frac{q_i}{q_l} \lim_{n\to\infty} W^{n-1}_{il} \cdot \int_A 1 d\lambda = \int_A q_i d\lambda.
\end{aligned}$$

We have defined $W^{n-1}_{ij} := \sum_{j_2 \cdots j_n} W_{i j_2} \cdots W_{j_{n-1} j_n} W_{j_n l}$, which is the probability to go from state i to l in $n-1$ steps, and we know that this converges to q_l as $n \to \infty$. □

Corollary 4.36. *Under the assumptions of Proposition 4.35, $(F, \tilde{\mu})$ is mixing.*

Proof: For all cylinders $[j_1 \cdots j_n], [l_1 \cdots l_m] \in \mathcal{F}$ and $A, B \in \mathcal{B}$ we have for all $N > m$ that (with $k = N - m$)

$$\begin{aligned}
&\tilde{\mu}\Big(F^{-N}\big([j_1 \cdots j_n] \times A\big) \cap [l_1 \cdots l_m] \times B \Big) \\
&\quad = \tilde{\mu}\Big(\bigcup_{i_1 \cdots i_k} [l_1 \cdots l_m i_1 \cdots i_k j_1 \cdots j_n] \times \big((T^{-1}_{l_1} \cdots T^{-1}_{l_m} T^{-1}_{i_1} \cdots T^{-1}_{i_k} A) \cap B \big) \Big) \\
&\quad = \sum_{i_1 \cdots i_k} \mathbb{P}([l_1 \cdots l_m i_1 \cdots i_k j_1 \cdots j_n]) \tilde{\mu}_{l_1}\big((T^{-1}_{l_1} \cdots T^{-1}_{l_m} T^{-1}_{i_1} \cdots T^{-1}_{i_k} A) \cap B \big),
\end{aligned}$$

where in the last step we use $\tilde{\mu}(C) = \int_{\Omega_E} \tilde{\mu}_{\omega_1}(C_\omega) d\mathbb{P}(\omega)$. We have

$$\mathbb{P}([l_1 \cdots l_m i_1 \cdots i_k j_1 \cdots j_n]) = \mathbb{P}([l_1 \cdots l_m]) \mathbb{P}([j_1 \cdots j_n]) W^*_{j_1 i_k} W^*_{i_k i_{k-1}} \cdots W^*_{i_2 i_1} \frac{W_{l_m i_1}}{q_{i_1}}.$$

Moreover, from Proposition 4.35 it follows that

$$\begin{aligned}
&\lim_{k\to\infty} \int_A \sum_{i_1 \cdots i_k} W^*_{j_1 i_k} W^*_{i_k i_{k-1}} \cdots W^*_{i_2 i_1} P_{i_k} \cdots P_{i_1} \Big(\frac{W_{l_m i_1}}{q_{i_1}} P_{l_m} \cdots P_{l_1}(1_B \tilde{h}_{l_1}) \Big) d\lambda \\
&\quad = \lim_{k\to\infty} \int_A \Big\{ P^k_T \Big(\frac{W_{l_m 1}}{q_1} P_{l_m} \cdots P_{l_1}(1_B \tilde{h}_{l_1}), \cdots, \frac{W_{l_m r}}{q_r} P_{l_m} \cdots P_{l_1}(1_B \tilde{h}_{l_1}) \Big) \Big\}_{j_1} d\lambda \\
&\quad = \int_A \Big(\sum_{i=1}^r q_i \int_I \frac{W_{l_m i}}{q_i} P_{l_m} \cdots P_{l_1}(1_B \tilde{h}_{l_1}) d\lambda \Big) \cdot \tilde{h}_{j_1} d\lambda \\
&\quad = \tilde{\mu}_{j_1}(A) \cdot \tilde{\mu}_{l_1}(B),
\end{aligned}$$

so we obtain

$$\lim_{N\to\infty} \tilde{\mu}\Big(F^{-N}\big([j_1 \cdots j_n] \times A\big) \cap [l_1 \cdots l_m] \times B \Big) = \tilde{\mu}([j_1 \cdots j_n] \times A) \cdot \tilde{\mu}([l_1 \cdots l_m] \times B).$$

□

Finally, the next proposition is the analogue of Proposition 4.16.

Proposition 4.37. *Suppose that $\alpha < 1$ and that the entries of W are strictly positive. Furthermore, suppose that the following* random covering property *holds: For each non-trivial subinterval $J \subseteq I$ and $\omega_1 \in \{1, \dots, r\}$ there exist $n \in \mathbb{N}$, $I_0 \subseteq I$ finite and $(\omega_2, \dots, \omega_n) \in \{1, \dots, r\}^{n-1}$ such that $T_{\omega_n} \circ \cdots \circ T_{\omega_1}(J) = I \backslash I_0$. Then $\tilde{\mu}$ is the only acipm for F and satisfies*

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\tilde{\mu}}{d\lambda} \leq M. \tag{4.59}$$

Moreover, $(F, \tilde{\mu})$ is ergodic.

Proof: This follows using similar arguments as in the proof of Proposition 3.16 for each of the r coordinates (that is why we require the above random covering property to hold for each $\omega_1 \in \{1, \dots, r\}$). □

Remark 4.38. In fact, because the Markov shift $(\sigma, \mathbb{P})$ is exact (see Example 2.23), one can show with the fourth part of Theorem 2.1 in [47] that $(F, \tilde{\mu})$ is exact under the assumptions of Proposition 4.37.

4.5 Automorphism as Base

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\varphi : \Omega \to \Omega$ be an automorphism, which means that φ is measure preserving and invertible (i.e. φ is one-to-one and φ^{-1} is measurable). Furthermore, suppose that $(\varphi, \mathbb{P})$ is ergodic. For each $\omega \in \Omega$, let $T_\omega : I \to I$ be a finitely piecewise monotonic interval map w.r.t. a partition $\{I_{i,\omega}\}$, and suppose that $T : \Omega \times I \to I$ given by $T(\omega, x) = T_\omega x$ is measurable. For each $\omega \in \Omega$, we write $N(\omega)$ for the minimal possible number of elements in $\{I_{i,\omega}\}$ and we set $\theta(\omega) = \inf_{x \in I} |T'_\omega(x)|$. We consider the skew product

$$F : \Omega \times I \to \Omega \times I, \quad (\omega, x) \mapsto (\varphi\omega, T_\omega x). \tag{4.60}$$

The following result is proven in [12] by Buzzi.

Theorem 4.39. (Theorem 0.3 in [12]) *Suppose that the following conditions are satisfied:*

1. $\Omega \ni \omega \mapsto \Big(\theta(\omega), N(\omega), \mathrm{Var}_I\Big(\dfrac{1}{|T'_\omega|}\Big)\Big)$ *is measurable,*
2. $\displaystyle\lim_{K \to \infty} \int_\Omega \log\big(\min\big(\theta(\omega), K\big)\big) d\mathbb{P}(\omega) > 0,$
3. $\Omega \ni \omega \mapsto \log^+ \dfrac{N(\omega)}{\theta(\omega)}$ *is in* $L^1(\mathbb{P})$,
4. $\Omega \ni \omega \mapsto \log^+ \mathrm{Var}_I\Big(\dfrac{1}{|T'_\omega|}\Big)$ *is in* $L^1(\mathbb{P})$.

Then the set $M_{ac}(\Omega \times I, F)$ of acim's of F is a non-empty finite-dimensional vector space generated by the ergodic acipm's of F. Moreover, each invariant density h of F satisfies $\mathrm{Var}_I(h_\omega) < \infty$ for $\mathbb{P}$-a.a. $\omega \in \Omega$, where each $h_\omega : I \to \mathbb{C}$ is given by $h_\omega(x) = h(\omega, x)$.

Note that condition 2 in Theorem 4.39 implies that T expands distances on average w.r.t. $\mathbb{P}$. Also, under the above conditions, P_F is not even bounded in $L^\infty(\mathbb{P}\otimes\lambda)$ (see Remark 0.6 in [12]). For this reason, the Ionescu-Tulcea and Marinescu Theorem (Theorem B.5) cannot be applied to obtain the result in Theorem 4.39. Instead, the result is obtained in [12] using *fiberwise* Lasota-Yorke type inequalities (see Proposition 1.4 in [12]).

Remark 4.40. As is remarked in [13], in this setting $h \in L^1(\mathbb{P}\otimes\lambda)$ is an invariant density for F if and only if $P_{T_\omega}h_\omega = h_{\varphi(\omega)}$ for $\mathbb{P}$-a.a. $\omega\in\Omega$, where again $h_\omega(x) = h(\omega, x)$. We can deduce this also from Proposition 5.6.

Like for the deterministic, Bernoulli and Markov case, we have the following strengthening if we assume a suitable covering property for T:

Theorem 4.41. (Part 1 of Main Theorem in [13]) *In addition to the assumptions of Theorem 4.39, suppose that* $\sup_{x\in I}|T'_\omega(x)| < \infty$ *for* $\mathbb{P}$*-a.a.* $\omega\in\Omega$ *and that* T *satisfies the following* covering property*: For each non-trivial subinterval* $J\subseteq I$ *and for* $\mathbb{P}$*-a.a.* $\omega\in\Omega$ *there exist* $n\in\mathbb{N}$ *and a finite set* $I_0\subseteq I$ *such that* $T_{\varphi^{n-1}(\omega)}\circ\cdots\circ T_\omega(J) = I\backslash I_0$*. Then there exists a unique acipm for* F*.*

Furthermore, the following result can be viewed as an analogue of the results in Remarks 4.17 and 4.38.

Proposition 4.42. (Comment 0.2.4 in [13]) *Under the assumptions of Theorem 4.41, if* $(\varphi,\mathbb{P})$ *is mixing, then the unique acipm is mixing w.r.t.* F*.*

Example 4.43. Let $\Omega\subseteq(1,\infty)$ with corresponding Borel σ-algebra $\mathcal{F}$, and let $\mathbb{P}$ be a probability measure on $(\Omega,\mathcal{F})$. Also, let $\varphi:\Omega\to\Omega$ be an automorphism such that $(\varphi,\mathbb{P})$ is ergodic. Furthermore, let $T:\Omega\times I\to I$ be given by $T(\omega,x) = T_\omega x = \omega x \bmod 1$. Then according to Example 0.4 in [12] the conditions in Theorem 4.39 are satisfied. Now suppose that $\inf\Omega > 1$. In a similar way as in Example 3.18 it can be shown that the assumptions in Theorem 4.41 are also satisfied. So F admits a unique invariant probability density $h\in L^1(\mathbb{P}\otimes\lambda)$. Let us show that for $\mathbb{P}$-a.a. $\omega\in\Omega$ we have

$$\exists M_\omega > 0 : \frac{1}{M_\omega} \le h_\omega \le M_\omega, \tag{4.61}$$

where again $h_\omega(x) = h(\omega,x)$. First of all, it is clear that $A = \{\omega\in\Omega : h_\omega \neq 0\}$ satisfies $\mathbb{P}(A) > 0$. Take $\omega\in A$. By Theorem 4.39, we may assume $\mathrm{Var}_I(h_\omega) < \infty$. Hence, by Corollary A.11, we may assume that h_ω is lower semicontinuous. Then there exist $\alpha > 0$ and a nontrivial interval $J\subseteq I$ such that $h_\omega \ge \alpha 1_J$. Now take $n\in\mathbb{N}$ and $I_0\subseteq I$ finite such that $T_{\varphi^{n-1}(\omega)}\circ\cdots\circ T_\omega(J) = I\backslash I_0$. For convenience, write $T^n_\omega = T_{\varphi^{n-1}(\omega)}\circ\cdots\circ T_\omega$. It is clear that $K_\omega := \sup_{x\in I}|(T^n_\omega)'(x)| < \infty$. Hence, for all $x\in I\backslash I_0$ we obtain (using Remark 4.40)

$$h_{\varphi^n\omega}(x) = P_{T^n_\omega}h_\omega(x) \ge \alpha P_{T^n_\omega}1_J(x) = \alpha\sum_{y\in(T^n_\omega)^{-1}x}\frac{1_J(y)}{|(T^n_\omega)'(y)|} \ge \frac{\alpha}{K_\omega}, \tag{4.62}$$

because for each $x \in I \backslash I_0$ there exists $y \in J$ such that $T_\omega^n y = x$. Since for each $m > n$ we have that $L_m := \sup_{x \in I} |(T_{\varphi^n \omega}^{m-n})'(x)| < \infty$ and that the map $T_{\varphi^{m-1}(\omega)}$ is surjective modulo a finite set, we obtain with

$$h_{\varphi^m \omega}(x) = P_{T_{\varphi^{m-1}\omega} \circ \cdots \circ T_{\varphi^n \omega}} h_{\varphi^n \omega} = \sum_{y \in (T_{\varphi^n \omega}^{m-n})^{-1} x} \frac{h_{\varphi^n \omega}(y)}{|(T_{\varphi^n \omega}^{m-n})'(y)|} \tag{4.63}$$

that for each $m > n$ there exists a finite set $I_m \subseteq I$ such that $h_{\varphi^m(\omega)}(x) \geq \frac{\alpha}{K_\omega L_m}$ for each $x \in I \backslash I_m$. Again, by Theorem 4.39 (and using that $\mathbb{P}$ is φ-invariant), we may assume that $\mathrm{Var}_I(h_{\varphi^m(\omega)}) < \infty$ holds for all $m \geq n$. Hence, it follows from (A.10) that $h_{\varphi^m(\omega)}$ is bounded away from zero for each $m \geq n$. We conclude that the set

$$B = \bigcup_{n \in \mathbb{N}} \bigcap_{m \geq n} \left\{ \omega \in \Omega : \inf_{x \in I} h_{\varphi^m(\omega)}(x) > 0 \right\} \tag{4.64}$$

satisfies $A \subseteq B$ and therefore $\mathbb{P}(B) > 0$. Now suppose that the set

$$C = \left\{ \omega \in \Omega : \inf_{x \in I} h_\omega(x) = 0 \right\} \tag{4.65}$$

satisfies $\mathbb{P}(C) > 0$. Since $(\varphi, \mathbb{P})$ is ergodic, we then know that $\mathbb{P}$-almost all points in Ω visit the set C infinitely often under iterations of φ (see e.g. Remark 1.6.1.3 in [15]). This is in contradiction with $\mathbb{P}(B) > 0$. We conclude that $\mathbb{P}(C) = 0$, and together with Theorem 4.39 we conclude that (4.61) holds for $\mathbb{P}$-a.a. $\omega \in \Omega$.

Chapter 5

Fiber Entropy

Let $(\Omega, \mathcal{F})$ and $(X, \mathcal{B})$ be measurable spaces. In [1], Abramov and Rokhlin introduced the notion of *fiber entropy* for skew products of the form

$$F : \Omega \times X \to \Omega \times X, (\omega, x) \mapsto (\varphi\omega, T_\omega x). \tag{5.1}$$

For this, they assume there exist probability measures $\mathbb{P}$ and ρ on Ω and X, respectively, such that $\varphi : \Omega \to \Omega$ is an automorphism on $(\Omega, \mathcal{F}, \mathbb{P})$ and such that $T_\omega : X \to X$ is measure preserving w.r.t. ρ for $\mathbb{P}$-a.a. $\omega \in \Omega$. In that case, F is measure preserving w.r.t. $\mathbb{P} \otimes \rho$.

Over the years the definition of fiber entropy has been extended to more general settings such as in Section 2.6 of [22]. A special case of the setting in [22] is considered in Section 2.4 of [64], where instead of the product measure $\mathbb{P} \otimes \rho$ a general F-invariant probability measure μ on $\Omega \times X$ is considered. One can then associate a fiber entropy to F if μ *disintegrates* into an *equivariant system of conditional measures*, which is the topic of the next section.

5.1 Equivariant System of Conditional Measures

Let $(\Omega, \mathcal{F})$ and $(X, \mathcal{B})$ be measurable spaces. Let μ be a probability measure on $(\Omega \times X, \mathcal{F} \otimes \mathcal{B})$ and let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$.

Definition 5.1. *A* system of conditional measures for μ over $\mathbb{P}$ *is a family of measures* $\{\rho_\omega\}_{\omega \in \Omega}$ *such that*

1. ρ_ω *is a positive finite measure on* $(X, \mathcal{B})$ *for* $\mathbb{P}$*-a.a.* $\omega \in \Omega$*,*
2. *For any* $f \in L^1(\mu)$*, the map* $\Omega \ni \omega \mapsto \int_X f(\omega, x) d\rho_\omega(x)$ *is measurable and* $\int_{\Omega \times X} f d\mu = \int_\Omega (\int_X f d\rho_\omega) d\mathbb{P}$*.*

In this case, we say that μ disintegrates *over* $\mathbb{P}$ *on the fibers* $\{\omega\} \times X$ *as* $\mu = \int \rho_\omega d\mathbb{P}$*.*

Remark 5.2. In particular, for any $A \in \mathcal{F} \otimes \mathcal{B}$, the map $\Omega \ni \omega \mapsto \rho_\omega(A_\omega)$ is measurable and $\mu(A) = \int_\Omega \rho_\omega(A_\omega) d\mathbb{P}(\omega)$, where $A_\omega = \{x \in I : (\omega, x) \in A\}$.

The next theorem is a simplified version of an important result due to Rokhlin [59].

Theorem 5.3. (Rokhlin) *Suppose that Ω and X are compact metric spaces, and that $\mathcal{F}$ and $\mathcal{B}$ are the corresponding Borel σ-algebra's, respectively. Let $\pi : \Omega \times X \to \Omega$ be the projection on Ω, i.e. $\pi(\omega, x) = \omega$, and suppose that $\mathbb{P}$ is the pushforward measure of μ under π, i.e. $\mathbb{P} = \mu \circ \pi^{-1}$. Then there exists a system of conditional measures $\{\rho_\omega\}_{\omega \in \Omega}$ for μ over $\mathbb{P}$. Also, ρ_ω is a probability measure for $\mathbb{P}$-a.a. $\omega \in \Omega$. Moreover, if $\{\tilde{\rho}_\omega\}_{\omega \in \Omega}$ is another system of conditional measures for μ over $\mathbb{P}$, then $\rho_\omega = \tilde{\rho}_\omega$ for $\mathbb{P}$-a.a. $\omega \in \Omega$.*

The following result is an easy consequence of Fubini's Theorem.

Proposition 5.4. *Let ρ be a probability measure on $(X, \mathcal{B})$, and suppose that μ is absolutely continuous w.r.t. $\mathbb{P} \otimes \rho$ with corresponding density denoted by $h \in L^1(\mathbb{P} \otimes \rho)$. Then the family $\{\rho_\omega\}_{\omega \in \Omega}$ given by $\rho_\omega(A) = \int_A h(\omega, x) d\rho(x)$ is a system of conditional measures for μ over $\mathbb{P}$.*

Let $T : \Omega \times X \to X$, $T(\omega, x) = T_\omega x$ be measurable, and let $\varphi : \Omega \to \Omega$ be measure preserving w.r.t. $\mathbb{P}$. Then the skew product F given by

$$F : \Omega \times X \to \Omega \times X, \quad (\omega, x) \mapsto (\varphi\omega, T_\omega x) \tag{5.2}$$

is measurable.

Definition 5.5. *We say that a system of conditional measures $\{\rho_\omega\}_{\omega \in \Omega}$ for μ over $\mathbb{P}$ is* equivariant *w.r.t. (T, φ) if for each $B \in \mathcal{B}$ there exists $C \in \mathcal{F}$ with $\mathbb{P}(C) = 1$ such that $\rho_\omega(T_\omega^{-1} B) = \rho_{\varphi(\omega)}(B)$ holds for all $\omega \in C$.*

Proposition 5.6. *Let $\{\rho_\omega\}_{\omega \in \Omega}$ be a system of conditional measures for μ over $\mathbb{P}$.*

1. *If $\{\rho_\omega\}_{\omega \in \Omega}$ is equivariant w.r.t. (T, φ), then μ is F-invariant.*
2. *If μ is F-invariant and φ is invertible, then $\{\rho_\omega\}_{\omega \in \Omega}$ is equivariant w.r.t. (T, φ).*

Proof: Let $A \in \mathcal{F}$ and $B \in \mathcal{B}$. First of all, because $\mathbb{P}$ is φ-invariant, we have

$$\mu(A \times B) = \int_A \rho_\omega(B) d\mathbb{P}(\omega) = \int_{\varphi^{-1} A} \rho_{\varphi(\omega)}(B) d\mathbb{P}(\omega). \tag{5.3}$$

Moreover, we have

$$F^{-1}(A \times B) = \bigcup_{\omega \in \varphi^{-1} A} \{\omega\} \times T_\omega^{-1} B \tag{5.4}$$

and so

$$\mu(F^{-1}(A \times B)) = \int_{\varphi^{-1} A} \rho_\omega(T_\omega^{-1} B) d\mathbb{P}(\omega). \tag{5.5}$$

Hence, the first part follows from (5.3) and (5.5) combined with the fact that $\{\rho_\omega\}_{\omega\in\Omega}$ is equivariant w.r.t. (T,φ). For the second part, we obtain from the F-invariance of μ combined with (5.3) and (5.5) that

$$\int_{\varphi^{-1}A} \rho_\omega(T_\omega^{-1}B)d\mathbb{P}(\omega) = \int_{\varphi^{-1}A} \rho_{\varphi(\omega)}(B)d\mathbb{P}(\omega). \tag{5.6}$$

Since φ is invertible we have $\varphi^{-1}\mathcal{F} = \mathcal{F}$, so this indeed yields that $\{\rho_\omega\}_{\omega\in\Omega}$ is equivariant w.r.t. (T,φ). □

Example 5.7. Consider the case that $X = I = [0,1]$, $\mathcal{B}$ the Borel σ-algebra on I and λ the Lebesgue measure restricted to I. Furthermore, suppose that φ is an automorphism on $(\Omega,\mathcal{F},\mathbb{P})$. Then under the assumptions of Theorem 4.39, there exists an acipm μ of F with density, say, $h \in L^1(\mathbb{P}\otimes\lambda)$. From Proposition 5.4 together with the second part of Proposition 5.6 it follows that the family $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A h(\omega,x)d\lambda(x)$ is an equivariant system of conditional measures for μ over $\mathbb{P}$.

5.2 Definition of Fiber Entropy

Let $(\Omega,\mathcal{F},\mathbb{P})$ and $(X,\mathcal{B},\rho)$ be Lebesgue spaces. Furthermore, let $T : \Omega\times X \to X$, $T(\omega,x) = T_\omega x$ be measurable, and let $\varphi : \Omega \to \Omega$ be measure preserving w.r.t. $\mathbb{P}$. Suppose that μ is a probability measure on $(\Omega\times X,\mathcal{F}\otimes\mathcal{B})$ that is absolutely continuous w.r.t. $\mathbb{P}\otimes\rho$ with density, say, $h \in L^1(\mathbb{P}\otimes\rho)$. We know from Proposition 5.4 that the family $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A h(\omega,x)d\rho(x)$ is a system of conditional measures for μ over $\mathbb{P}$, and we furthermore suppose that $\{\rho_\omega\}_{\omega\in\Omega}$ is equivariant w.r.t. (T,φ). For this setting we shall define the fiber entropy of (T,φ).

Remark 5.8. The next construction resembles the construction of fiber entropy in Section 2.4 of [64] (and the references therein). However, in [64] one works with the system of conditional measures for μ over the pushforward measure $\mu\circ\pi^{-1}$ as in Theorem 5.3. Moreover, it is assumed in [64] that μ is F-invariant and φ is an automorphism, which is in general stronger than our assumption that $\{\rho_\omega\}_{\omega\in\Omega}$ is equivariant w.r.t. (T,φ) (see Proposition 5.6).

Let ξ be a (finite or countable) partition of $\Omega\times X$. For each $n\in\mathbb{N}$ we write ξ_n for the partition of $\Omega\times X$ given by

$$\xi_n = \bigvee_{k=0}^{n-1} F^{-k}\xi. \tag{5.7}$$

Also, for each $\omega\in\Omega$, we write $\xi_\omega = \{Z_\omega : Z\in\xi\}$ for the partition of X where $Z_\omega = \{x\in X : (\omega,x)\in Z\}$. Furthermore, for each $\omega\in\Omega$ and $n\in\mathbb{N}$ we define the partition $\xi_{\omega,n}$ of X given by

$$\xi_{\omega,n} = \xi_\omega \vee \bigvee_{k=1}^{n-1} T_\omega^{-1}T_{\varphi(\omega)}^{-1}\cdots T_{\varphi^{k-1}(\omega)}^{-1}\xi_{\varphi^k(\omega)}. \tag{5.8}$$

Note that $(\xi_n)_\omega = \xi_{\omega,n}$ for each $\omega \in \Omega$ and $n \in \mathbb{N}$.

We define the *fiber entropy of the partition* ξ by

$$H_\mu(\{\xi_\omega\}) = \int_\Omega H_{\rho_\omega}(\xi_\omega) d\mathbb{P}(\omega) = \int_\Omega \sum_{A \in \xi_\omega} \rho_\omega(A) \log \rho_\omega(A) d\mathbb{P}(\omega), \tag{5.9}$$

which is well defined by noting that $H_{\rho_\omega}(\xi_\omega) = \sum_{Z \in \xi} \rho_\omega(Z_\omega) \log \rho_\omega(Z_\omega)$ from which it together with Remark 5.2 follows that the map $\Omega \ni \omega \mapsto H_{\rho_\omega}(\xi_\omega)$ is measurable.

Proposition 5.9. *Let ξ be a partition of $\Omega \times X$ s.t. $H_\mu(\{\xi_\omega\}) < \infty$. Then the* fiber entropy of (T, φ) w.r.t. ξ *given by*

$$h^\varphi(\xi, T) = \lim_{n \to \infty} \frac{1}{n} \int_\Omega H_{\rho_\omega}(\xi_{\omega,n}) d\mathbb{P}(\omega), \tag{5.10}$$

exists and is finite.

Proof: We are done if we show that the sequence $\{a_n\}$ given by $a_n = \int_\Omega H_{\rho_\omega}(\xi_{\omega,n}) d\mathbb{P}(\omega)$ is subadditive, i.e. $a_{n+m} \leq a_n + a_m$ for all $n, m \in \mathbb{N}$, because then $\lim_{n\to\infty} \frac{a_n}{n} = \inf_{m \in \mathbb{N}} \frac{a_m}{m}$ (see Lemma 2.47). Indeed, we have

$$\begin{aligned}
a_{n+m} &= \int_\Omega H_{\rho_\omega}(\xi_{\omega,n+m}) d\mathbb{P}(\omega) \\
&\leq \int_\Omega H_{\rho_\omega}(\xi_{\omega,n}) + H_{\rho_\omega}\Big(T_\omega^{-1} \cdots T_{\varphi^{n-1}(\omega)}^{-1} \bigvee_{k=n}^{m+n-1} T_{\varphi^n(\omega)}^{-1} \cdots T_{\varphi^{k-1}(\omega)}^{-1} \xi_{\varphi^k(\omega)}\Big) d\mathbb{P}(\omega) \\
&= a_n + \int_\Omega H_{\rho_{\varphi^n(\omega)}}(\xi_{\varphi^n(\omega),m}) d\mathbb{P}(\omega) \\
&= a_n + a_m,
\end{aligned}$$

where the last two steps follow from the equivariance of $\{\rho_\omega\}_{\omega \in \Omega}$ w.r.t. (T, φ) and from the invariance of $\mathbb{P}$ w.r.t. φ, respectively. □

Definition 5.10. *The* fiber entropy of (T, φ) *is defined as*

$$h^\varphi(T) = \sup\{h^\varphi(\xi, T) |\ \xi \text{ partition of } \Omega \times X \text{ s.t. } H_\mu(\{\xi_\omega\}) < \infty\}. \tag{5.11}$$

Lemma 5.11. *Let ξ and ζ each be a partition of $\Omega \times X$. Then the sequence $\{a_n\}$ given by $a_n = \int_\Omega H_{\rho_\omega}(\xi_{\omega,n} | \zeta_{\omega,n}) d\mathbb{P}(\omega)$ is subadditive.*

Proof: This follows in a similar way as the proof of Proposition 5.9. □

Let α be a partition of X. We define $\bar{\alpha} = \{\Omega \times A : A \in \alpha\}$ as the extension of α to a partition of $\Omega \times X$.

Lemma 5.12. *If $\alpha_1 \leq \alpha_2 \leq \cdots$ is an increasing sequence of finite partitions on X such that $\sigma(\bigvee_n \alpha_n) = \mathcal{B}$ up to sets of ρ-measure zero, then $h^\varphi(T) = \lim_{k\to\infty} h^\varphi(\bar{\alpha}_k, T)$.*

Proof: It is sufficient to show that for any partition ζ such that $H_\mu(\{\zeta_\omega\}) < \infty$ we have $h^\varphi(\zeta, T) \leq \lim_{k\to\infty} h^\varphi(\bar{\alpha}_k, T)$. For each $\omega \in \Omega$ and $n, k \in \mathbb{N}$ we have

$$H_{\rho_\omega}(\zeta_{\omega,n}) \leq H_{\rho_\omega}((\bar{\alpha}_k)_{\omega,n}) + H_{\rho_\omega}(\zeta_{\omega,n}|(\bar{\alpha}_k)_{\omega,n}). \tag{5.12}$$

By Lemma 5.11 and Lemma 2.47, we know for each $k \in \mathbb{N}$ that

$$\lim_{n\to\infty} \frac{1}{n} \int_\Omega H_{\rho_\omega}(\zeta_{\omega,n}|(\bar{\alpha}_k)_{\omega,n}) d\mathbb{P}(\omega) \leq \int_\Omega H_{\rho_\omega}(\zeta_\omega|\alpha_k) d\mathbb{P}(\omega), \tag{5.13}$$

and combined with (5.12) this gives

$$h^\varphi(\zeta, T) \leq h^\varphi(\bar{\alpha}_k, T) + \int_\Omega H_{\rho_\omega}(\zeta_\omega|\alpha_k) d\mathbb{P}(\omega). \tag{5.14}$$

It follows from the Dominated Convergence Theorem and the Martingale Convergence Theorem that, for $\mathbb{P}$-a.a. $\omega \in \Omega$,

$$\begin{aligned} \lim_{k\to\infty} H_{\rho_\omega}(\zeta_\omega|\alpha_k) &= -\int_X \lim_{k\to\infty} \log E_{\rho_\omega}(1_{\xi_\omega(x)}|\sigma(\alpha_k))(x) d\rho_\omega(x) \\ &= -\int_X \log E_{\rho_\omega}(1_{\xi_\omega(x)}|\mathcal{B})(x) d\rho_\omega(x) = 0. \end{aligned} \tag{5.15}$$

Note that we use here that $\rho_\omega \ll \rho$ for $\mathbb{P}$-a.a. $\omega \in \Omega$. Moreover, combining the Dominated Convergence Theorem with $H_{\rho_\omega}(\zeta_\omega|\alpha_k) \leq H_{\rho_\omega}(\zeta_\omega)$ for each $k \in \mathbb{N}$ and $H_\mu(\{\zeta_\omega\}) < \infty$, we obtain

$$\lim_{k\to\infty} \int_\Omega H_{\rho_\omega}(\zeta_\omega|\alpha_k) d\mathbb{P}(\omega) = \int_\Omega \lim_{k\to\infty} H_{\rho_\omega}(\zeta_\omega|\alpha_k) d\mathbb{P}(\omega). \tag{5.16}$$

From (5.14), (5.15) and (5.16) the desired result follows. □

Corollary 5.13. *We have*

$$h^\varphi(T) = \sup\{h^\varphi(\bar{\alpha}, T) |\ \alpha \text{ partition of } X \text{ s.t. } H_\mu(\{\bar{\alpha}_\omega\}) < \infty\}. \tag{5.17}$$

Proof: Because $(X, \mathcal{B}, \rho)$ is a standard Lebesgue space, we know there exists a sequence of partitions $\{\alpha_n\}$ as in Lemma 5.12. □

Proposition 5.14. *We have*

$$h_\mu(F) = h_{\mathbb{P}}(\varphi) + h^\varphi(T). \tag{5.18}$$

Proof: Using Corollary 5.13, this follows in an analogous manner as the proof of Proposition 1.3 in Chapter 6 of [54]. Note that in the second part of the proof of this proposition an increasing sequence of partitions $\beta_1 \leq \beta_2 \leq \cdots$ on Ω is chosen such that $\sigma(\bigvee_n \beta_n) = \mathcal{F}$ up to sets of $\mathbb{P}$-measure zero. The existence of such a sequence of partitions $\beta_1, \beta_2, \ldots$ is guaranteed by the assumption that $(\Omega, \mathcal{F}, \mathbb{P})$ is a Lebesgue space. □

5.3 Analogue of the Kolmogorov-Sinai Theorem

Let us remain in the setting of the previous section. We need the following definition.

Definition 5.15. *Let ξ be a partition of $\Omega \times X$. For $\omega \in \Omega$, we say that ξ is* ω-generating *w.r.t. (φ, T) if*

$$\sigma\Big(\xi_\omega \vee \bigvee_{k=1}^{\infty} T_\omega^{-1} T_{\varphi(\omega)}^{-1} \cdots T_{\varphi^{k-1}(\omega)}^{-1} \xi_{\varphi^k(\omega)}\Big) = \mathcal{B} \tag{5.19}$$

up to sets of ρ-measure zero.

We have the following analogue of the Kolmogorov-Sinai Theorem (Theorem 2.51).

Theorem 5.16. *Let ξ be a partition of $\Omega \times X$ such that $H_\mu(\{\xi_\omega\}) < \infty$. Suppose that ξ is ω-generating w.r.t. (φ, T) for $\mathbb{P}$-a.a. $\omega \in \Omega$. Then $h^\varphi(T) = h^\varphi(\xi, T)$.*

Proof: This follows in a similar way as the proof of Lemma 5.12, replacing $\bar{\alpha}_k$ with $\xi_{\omega,k}$ and noting that

$$\lim_{n\to\infty} \frac{1}{n} \int_\Omega H_{\rho_\omega}(\xi_{\omega,n+k}) d\mathbb{P}(\omega) = \lim_{n\to\infty} \frac{n+k}{n} \lim_{n\to\infty} \frac{1}{n+k} \int_\Omega H_{\rho_\omega}(\xi_{\omega,n+k}) d\mathbb{P}(\omega) = h^\varphi(\xi, T).$$

□

5.4 Analogue of the Shannon-McMillan-Breiman Theorem

Again, we remain in the setting of Section 5.3. The next theorem is an analogue of the Shannon-McMillan-Breiman Theorem (Theorem 2.53). For this, we define the *information function*

$$I_\alpha(\omega, x) = -\sum_{A\in\alpha} 1_A(x) \log \rho_\omega(A), \tag{5.20}$$

where α is a partition of X. Note that $H_{\rho_\omega}(\alpha) = \int_X I_\alpha(\omega, x) d\rho_\omega(x)$. Furthermore, for two partitions α and β of X, we define the *conditional information function*

$$I_{\alpha|\beta}(\omega, x) = -\sum_{B\in\beta} \sum_{A\in\alpha} 1_{A\cap B}(x) \log\Big(\frac{\rho_\omega(A\cap B)}{\rho_\omega(B)}\Big). \tag{5.21}$$

Then $H_{\rho_\omega}(\alpha|\beta) = \int_X I_{\alpha|\beta}(\omega, x) d\rho_\omega(x)$. Also, note that

$$I_{\alpha\vee\beta}(\omega, x) = I_\beta(\omega, x) + I_{\alpha|\beta}(\omega, x). \tag{5.22}$$

Denoting $\alpha(x)$ for the atom of α containing x, we can also write

$$\begin{aligned} I_\alpha(\omega, x) &= -\log \rho_\omega(\alpha(x)), \\ I_{\alpha|\beta}(\omega, x) &= -\log E_{\rho_\omega}(1_{\alpha(x)}|\sigma(\beta)). \end{aligned}$$

Theorem 5.17. *Let ξ be a partition of $\Omega \times X$ such that $H_\mu(\{\xi_\omega\}) < \infty$. Suppose F is ergodic w.r.t. μ. Then*

$$\lim_{n\to\infty} -\frac{I_{\xi_{\omega,n}}(\omega,x)}{n} = h^\varphi(\xi,T), \qquad \mu\text{-a.e. } (\omega,x) \in \Omega \times X. \tag{5.23}$$

Proof: We follow the idea of the proof of the Shannon-McMillan-Breiman Theorem in Section 6.2 of [54]. Write $\tilde{\xi}_{\omega,n} = T_\omega^{-1} \bigvee_{k=1}^{n-1} T_{\varphi(\omega)}^{-1} \cdots T_{\varphi^{k-1}(\omega)}^{-1} \xi_{\varphi^k(\omega)}$ for $n \geq 2$. Then $\xi_{\omega,n} = \xi_\omega \vee \tilde{\xi}_{\omega,n}$. Furthermore, define $f_n(\omega,x) = I_{\xi_\omega|\tilde{\xi}_{\omega,n}}(\omega,x)$ for $n \geq 2$ and $f_1(\omega,x) = I_{\xi_\omega}(\omega,x)$. Then, using (5.22),

$$\begin{aligned}
I_{\xi_{\omega,n}}(\omega,x) &= I_{\tilde{\xi}_{\omega,n}}(\omega,x) + I_{\xi_\omega|\tilde{\xi}_{\omega,n}}(\omega,x) \\
&= I_{\xi_{\varphi(\omega),n-1}}(F(\omega,x)) + f_n(\omega,x) \\
&= I_{\tilde{\xi}_{\varphi(\omega),n-1}}(F(\omega,x)) + I_{\xi_{\varphi(\omega)}|\tilde{\xi}_{\varphi(\omega),n-1}}(F(\omega,x)) + f_n(\omega,x) \\
&= I_{\xi_{\varphi^2(\omega),n-2}}(F^2(\omega,x)) + f_{n-1}(F(\omega,x)) + f_n(\omega,x) \\
&\vdots \\
&= f_1(F^{n-1}(\omega,x)) + \cdots + f_{n-1}(F(\omega,x)) + f_n(\omega,x).
\end{aligned} \tag{5.24}$$

First of all, this gives

$$\begin{aligned}
\sum_{k=0}^{n-1} \int f_{n-k} \circ F^k d\mu &= \int I_{\xi_{\omega,n}}(\omega,x) d\mu(\omega,x) \\
&= \int_\Omega \Big(\int_X I_{\xi_{\omega,n}}(\omega,x) d\rho_\omega(x) \Big) d\mathbb{P}(\omega) \\
&= \int_\Omega H_{\rho_\omega}(\xi_{\omega,n}) d\mathbb{P}(\omega).
\end{aligned}$$

Using that μ is F-invariant (see Proposition 5.6), we obtain

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \int f_k d\mu = h^\varphi(\xi,T). \tag{5.25}$$

Furthermore, note that the sequence $\big\{H_{\rho_\omega}(\xi_\omega|\tilde{\xi}_{\omega,n})\big\}_{n\geq 1}$ is bounded from below and is non-increasing for each $\omega \in \Omega$, so $\lim_{n\to\infty} H_{\rho_\omega}(\xi_\omega|\tilde{\xi}_{\omega,n})$ exists for each $\omega \in \Omega$. Hence, from the Dominated Convergence Theorem follows that $\lim_{n\to\infty} \int f_n d\mu$ exists. Because the Cesàro means $\{\frac{1}{n}\sum_{k=1}^n x_k\}_{n\in\mathbb{N}}$ of a convergent sequence $\{x_n\}_{n\in\mathbb{N}}$ converge to the same limit as the sequence itself, we obtain from (5.25) that

$$\lim_{n\to\infty} \int f_n d\mu = h^\varphi(\xi,T). \tag{5.26}$$

By the Martingale Convergence Theorem we know for each $\omega \in \Omega$ that $\lim_{n\to\infty} f_n(\omega,\cdot)$ exists ρ_ω-a.e., so $f = \lim_{n\to\infty} f_n$ exists μ-a.e. Furthermore, this is an element of $L^1(\mu)$, because (5.26) together with the Dominated Convergence Theorem yield $\int f d\mu =$

$h^{\varphi}(\xi, T)$. Also, from (5.24) we see that

$$\frac{I_{\xi_{\omega,n}}(\omega, x)}{n} = \frac{1}{n}\sum_{k=0}^{n-1} f(F^k(\omega, x)) + \frac{1}{n}\sum_{k=0}^{n-1}(f_{n-k} - f)(F^k(\omega, x)). \tag{5.27}$$

By Birkhoff's Ergodic Theorem (Theorem 2.11),

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} f(F^k(\omega, x)) = \int f d\mu = h^{\varphi}(\xi, T), \qquad \mu\text{-a.e.} \tag{5.28}$$

Hence, it remains to prove that

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1}(f_{n-k} - f)(F^k(\omega, x)) = 0, \qquad \mu\text{-a.e.} \tag{5.29}$$

We refer the reader to Section 6.2 in [54] for a proof of (5.29), where the same limit as in (5.29) is shown for the proof of the Shannon-McMillan-Breiman Theorem. □

Chapter 6

Lochs' Theorem and Extensions

6.1 Introduction

Let us reformulate Lochs' Theorem from Subsection 1.1.1.

We put $A_i = [\frac{i}{10}, \frac{i+1}{10})$ for each $i \in \{1, \ldots, 9\}$. The decimal map $T : [0,1) \to [0,1)$ is given by $Tx = 10x - d_1(x)$, where $d_1(x) = i$ if $x \in A_i$. Recall that, for each $x \in [0,1)$,

$$x = \sum_{k=1}^{\infty} \frac{d_k(x)}{10^k}, \tag{6.1}$$

where $d_k(x) = d_1(T^{k-1}x)$ for each $k \geq 1$. For each $n \in \mathbb{N}$, the *cylinders of order n w.r.t. T* are

$$A_{i_0 \cdots i_{n-1}} = \bigcap_{k=0}^{n-1} T^{-k} A_{i_k}, \quad (i_0, \ldots, i_{n-1}) \in \{0, 1, \ldots, 9\}^n, \tag{6.2}$$

which are 10^n equally sized disjoint intervals and cover $[0,1)$. Note that

$$x \in A_{i_0 \cdots i_{n-1}} \Leftrightarrow d_k(x) = i_{k-1} \text{ for all } k = 1, \ldots, n. \tag{6.3}$$

Similarly, we put $B_1 = (\frac{1}{2}, 1)$ and $B_i = (\frac{1}{i+1}, \frac{1}{i}]$ for $i \geq 2$. The Gauss map $S : [0,1) \to [0,1)$ is given by $S0 = 0$ and $Sx = \frac{1}{x} - a_1(x)$, where $a_1(x) = i$ if $x \in B_i$. Recall that, for each $x \in (0,1)$ irrational,

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}}, \tag{6.4}$$

where $a_k(x) = a_1(S^{k-1}x)$ for each $k \geq 1$. Again, for each $n \in \mathbb{N}$, the cylinders of order n w.r.t. S are

$$B_{i_0 \cdots i_{n-1}} = \bigcap_{k=0}^{n-1} S^{-k} B_{i_k}, \quad (i_0, \ldots, i_{n-1}) \in \mathbb{N}^n, \tag{6.5}$$

which are countably many disjoint intervals and cover $(0, 1)$. Again, note that

$$x \in B_{i_0 \cdots i_{n-1}} \Leftrightarrow a_k(x) = i_{k-1} \text{ for all } k = 1, \ldots, n. \tag{6.6}$$

Define the partitions $\alpha = \{A_i : i \in \{0, \ldots, 9\}\}$ and $\beta = \{B_i : i \in \mathbb{N}\}$, and write $\alpha_n(x)$ (resp. $\beta_n(x)$) for the decimal cylinder of order n (resp. RCF cylinder of order n) containing x. Now, put

$$m(n, x) = \sup\{m \in \mathbb{N} : \alpha_n(x) \subseteq \beta_m(x)\}. \tag{6.7}$$

As we see from (6.3) and (6.6), $m(n, x)$ is the number of digits in the RCF expansion (6.4) of x that are determined by knowing n digits of the decimal expansion (6.1) of x. In [46], Lochs proved the following law of large numbers result for $m(n, x)$.

Theorem 6.1. (Lochs) *For λ-a.e. $x \in [0, 1)$*

$$\lim_{n \to \infty} \frac{m(n, x)}{n} = \frac{6 \log 2 \log 10}{\pi^2} = 0.97027 \cdots . \tag{6.8}$$

Furthermore, setting $z_0 = \frac{6 \log 2 \log 10}{\pi^2}$, Faivre [26] obtained a corresponding central limit theorem:

Theorem 6.2. (Faivre) *There exists $\sigma > 0$ such that for all $u \in \mathbb{R}$ we have*

$$\lim_{n \to \infty} \lambda\Big(\Big\{x \in I : \frac{m(n, x) - nz_0}{\sigma\sqrt{n}} \leq u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt. \tag{6.9}$$

Also, Faivre [25] obtained a large deviation result associated with (6.8), which is extended in [27] to the case that $T : [0, 1) \to [0, 1)$ is given by $Tx = \beta x \bmod 1$ for any $\beta > 1$. Moreover, as explained in Subsection 1.1.2, the result by Lochs in Theorem 6.1 has been generalized by Dajani and Fieldsteel in [16] to a wide class of interval maps that generate expansions. Members of this class are so-called *number-theoretic fibered maps* (NTFM) and we give a precise definition of such interval maps in Section 6.3. In [32], Herczegh proved a central limit theorem associated with the extension of Lochs' Theorem to any pair of NTFM's.

In the next two sections we provide the proof from [16] that shows the extension of Lochs' Theorem to any pair of NTFM's. Specifically, we review in Section 6.2 that Lochs' Theorem holds for any two sequences of interval partitions on $[0, 1)$ that both satisfy the conclusion of the Shannon-McMillan-Breiman Theorem. Using this, we shall see for any two NTFM's T and S that the number of digits $m_{T,S}(n, x)$ in the

S-expansion of x that can be determined from knowing the first n digits in the T-expansion of x satisfies

$$\lim_{n\to\infty} \frac{m_{T,S}(n,x)}{n} = \frac{h_{\mu_T}(T)}{h_{\mu_S}(S)}, \qquad \lambda\text{-a.e.} \tag{6.10}$$

In (6.10), $h_{\mu_T}(T)$ (resp. $h_{\mu_S}(T)$) denotes the entropy of T (resp. S) with respect to the measure μ_T (resp. μ_S) that (as we shall see in Section 6.3) is the unique acipm of T (resp. S). Moreover, we formulate in Section 6.3 the central limit theorem obtained by Herczegh in [32] that is associated with the law of large numbers result in (6.10).

Finally, in Section 6.4 we generalize the concept of an NTFM to so-called *random number-theoretic fibered systems* (RNTFS), which form a class of random piecewise monotonic interval maps being of the form as in Chapter 4. For each RNTFS given by $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$, we shall see that iterations of $F_{\varphi,T}(\omega,\cdot)$ generate (after projecting on $[0,1)$) expansions of points in $[0,1)$ for $\mathbb{P}$-a.a. $\omega \in \Omega$, where $F_{\varphi,T}$ denotes the skew product as given in (4.3). We shall prove for any two RNTFS's given by $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \psi, S)$ where each of the two underlying bases is a one-sided Bernoulli shift, a one-sided Markov shift or an automorphism, that for $\mathbb{P} \otimes \tilde{\mathbb{P}}$-a.a. $(\omega, \tilde{\omega}) \in \Omega \times \tilde{\Omega}$

$$\lim_{n\to\infty} \frac{m_{T,S}(n,\omega,\tilde{\omega},x)}{n} = \frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi)}, \qquad \lambda\text{-a.e.} \tag{6.11}$$

In (6.11), $m_{T,S}(n,\omega,\tilde{\omega},x)$ is the number of digits in the $F_{\psi,S}(\tilde{\omega},\cdot)$-expansion of x that can be determined from knowing the first n digits in the $F_{\varphi,T}(\omega,\cdot)$-expansion of x. Moreover, $h_{\mu_T}(F_{\varphi,T})$ (resp. $h_{\mu_S}(F_{\psi,S})$) denotes the entropy of $F_{\varphi,T}$ (resp. $F_{\psi,S}$) with respect to the measure μ_T (resp. μ_S) that (as we shall see in Section 6.4) is the unique acipm of $F_{\varphi,T}$ (resp. $F_{\psi,S}$). Furthermore, we derive from the results in [32] a central limit theorem associated with (6.11).

6.2 Equipartition of Interval Partitions

Like in [16], we introduce the following definitions.

Definition 6.3. *We say that P is an* interval partition *if it consists of finitely or countably many subintervals of $[0,1)$ that together form a partition of $[0,1)$. For an interval partition P and $x \in [0,1)$, we write $P(x)$ for the interval in P that contains x.*

Definition 6.4. *Let $\mathcal{P} = \{P_n\}_{n=1}^{\infty}$ be a sequence of interval partitions. Let $c \geq 0$. We say that $\mathcal{P}$* has entropy c λ-a.e. *if*

$$\lim_{n\to\infty} -\frac{\log \lambda(P_n(x))}{n} = c, \qquad \lambda\text{-}a.e. \tag{6.12}$$

Remark 6.5. Note that in Definition 6.4 we do not assume that each P_n is refined by P_{n+1}. In other words, we do not assume that for every interval $A \in P_{n+1}$ there exists an interval $B \in P_n$ such that $A \subseteq B$ (up to sets of Lebesgue measure zero).

The next theorem will be essential for the rest of this chapter.

Theorem 6.6. (Theorem 4 in [16]) *Let $\mathcal{P} = \{P_n\}_{n=1}^{\infty}$ and $\mathcal{Q} = \{Q_n\}_{n=1}^{\infty}$ be two sequences of interval partitions. For each $n \in \mathbb{N}$ and $x \in [0,1)$, put*

$$m_{\mathcal{P},\mathcal{Q}}(n,x) = \sup\{m \in \mathbb{N} : P_n(x) \subseteq Q_m(x)\}. \tag{6.13}$$

Suppose that for some constants $c > 0$ and $d > 0$, $\mathcal{P}$ has entropy c λ-a.e. and $\mathcal{Q}$ has entropy d λ-a.e. Then

$$\lim_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} = \frac{c}{d}, \quad \lambda - a.e. \tag{6.14}$$

For completeness, we include the proof of Theorem 6.6. We follow the proof in [16], which is based on general measure-theoretic covering arguments and not on the dynamics of specific maps.

Proof: Let us first show that

$$\limsup_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} \leq \frac{c}{d}, \qquad \lambda\text{-a.e.} \tag{6.15}$$

By assumption, for λ-a.e. $x \in I$ we have

$$\lim_{n\to\infty} -\frac{\log \lambda(P_n(x))}{n} = c, \qquad \lim_{n\to\infty} -\frac{\log \lambda(Q_n(x))}{n} = d. \tag{6.16}$$

We take such an $x \in I$. Let $\varepsilon > 0$, and take $\eta > 0$ such that $\frac{c+\eta}{c-\frac{c}{d}\eta} < 1+\varepsilon$. It follows from (6.16) that there exists $N \in \mathbb{N}$ such that for all $n \geq N$ we have[1]

$$\lambda(P_n(x)) > 2^{-n(c+\eta)}, \quad \lambda(Q_n(x)) < 2^{-n(d-\eta)}. \tag{6.17}$$

Choose $n \geq N$ such that $\min\{n, \frac{c}{d}n\} \geq N$, and let m' be any integer greater than $(1+\varepsilon)\frac{c}{d}n$. Then

$$\lambda(P_n(x)) > 2^{-n(1+\varepsilon)(c-\frac{c}{d}\eta)} > 2^{-m'(d-\eta)} > \lambda(Q_{m'}(x)), \tag{6.18}$$

from which it follows that $P_n(x)$ is not contained in $Q_{m'}(x)$. For this reason, we obtain

$$m_{\mathcal{P},\mathcal{Q}}(n,x) \leq (1+\varepsilon)\frac{c}{d}n \tag{6.19}$$

and therefore

$$\limsup_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} \leq (1+\varepsilon)\frac{c}{d}. \tag{6.20}$$

Since (6.20) holds for each $\varepsilon > 0$, (6.15) follows.

We now show that

$$\liminf_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} \geq \frac{c}{d}, \qquad \lambda\text{-a.e.} \tag{6.21}$$

[1]Equation (6.17) holds since the logarithm function in Definition 6.4 (as well as in Section 2.8 and Chapter 5) is by convention with respect to base 2. This is because information is usually measured in bits.

Let $\varepsilon > 0$, and take $\eta > 0$ such that $\zeta := \varepsilon c - \eta(1 + (1-\varepsilon)\frac{c}{d}) > 0$. For each $n \in \mathbb{N}$, we define $\bar{m}(n) = \lfloor (1-\varepsilon)\frac{c}{d} n \rfloor$ and

$$D_n(\eta) = \big\{x : \lambda(P_n(x)) < 2^{-n(c-\eta)}, \lambda(Q_{\bar{m}(n)}(x)) > 2^{-\bar{m}(n)(d+\eta)}, P_n(x) \not\subseteq Q_{\bar{m}(n)}(x)\big\}.$$

The number of intervals $A \in Q_{\bar{m}(n)}$ for which $\lambda(A) > 2^{-\bar{m}(n)(d+\eta)}$ is bounded by $2^{\bar{m}(n)(d+\eta)}$. Moreover, for each such $A \in Q_{\bar{m}(n)}$, there exists $x \in D_n(\eta) \cap A \cap B$ only for those $B \in P_n$ that contain a boundary point of A (of which there are 2) and satisfy $\lambda(B) < 2^{-n(c-\eta)}$. We conclude that

$$\lambda(D_n(\eta)) \leq 2 \cdot 2^{-n(c-\eta)} \cdot 2^{\bar{m}(n)(d+\eta)} \leq 2 \cdot 2^{-n\zeta}, \tag{6.22}$$

which gives $\sum_{n=1}^{\infty} \lambda(D_n(\eta)) < \infty$. From the Borel-Cantelli Lemma it follows that

$$\lambda\big(\{x \in [0,1) : x \in D_n(\eta) \text{ for infinitely many } n \in \mathbb{N}\}\big) = 0, \tag{6.23}$$

and therefore

$$\lambda\big(\{x \in [0,1) |\ \exists N \in \mathbb{N} \text{ s.t. } \forall n \geq N : x \notin D_n(\eta)\}\big) = 1. \tag{6.24}$$

Combining this with (6.16) (and using that $\bar{m}(n) \to \infty$ as $n \to \infty$) yields that for λ-a.e. $x \in [0,1)$ there exists an $N \in \mathbb{N}$ such that $m_{\mathcal{P},\mathcal{Q}}(n,x) \geq \bar{m}(n)$ for all $n \geq N$. This gives

$$\liminf_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} \geq (1-\varepsilon)\frac{c}{d}, \qquad \lambda\text{-a.e.} \tag{6.25}$$

Since $\varepsilon > 0$ was arbitrary, this concludes the proof. □

The next theorem is a central limit result associated with the law of large numbers result in (6.14). It is proven in [32] for the case that $\mathcal{P}$ and $\mathcal{Q}$ are both sequences of interval partitions consisting of cylinders of all orders w.r.t. some NTFM. We state this result in the next section, but the proof of this result (namely Corollary 2.1 in [32]) immediately carries over to all pairs of sequences of interval partitions $\mathcal{P}$ and $\mathcal{Q}$ that have the following properties:

Definition 6.7. *Let $\mathcal{P} = \{P_n\}_{n=1}^{\infty}$ be a sequence of interval partitions and suppose that for some constant $c > 0$, $\mathcal{P}$ has entropy c λ-a.e. We say that $\mathcal{P}$ satisfies the* 0-property *if*

$$\lim_{n\to\infty} \frac{-\log \lambda\big(P_n(x)\big) - nc}{\sqrt{n}} = 0, \qquad \lambda\text{-a.e.} \tag{6.26}$$

Definition 6.8. *Let $\mathcal{Q} = \{Q_n\}_{n=1}^{\infty}$ be a sequence of interval partitions and suppose that for some constant $d > 0$, $\mathcal{Q}$ has entropy d λ-a.e. For each $x \in [0,1)$ and $m \in \mathbb{N}$ we put $W_{m,x}(0) = 0$ and*

$$W_{m,x}\Big(\frac{l}{m}\Big) = \frac{-\log \lambda\big(Q_l(x)\big) - ld}{\sigma\sqrt{m}}, \qquad l \in \{1,\ldots,m\}, \tag{6.27}$$

for some $\sigma > 0$, and we extend this linearly on the subintervals $\{[\frac{l-1}{m}, \frac{l}{m}] : 1 \leq l \leq m\}$ so that $W_{m,x} \in C[0,1)$. We say that $\mathcal{Q}$ satisfies the weak invariance principle with variance

σ^2 and w.r.t. some (Borel) probability measure ν on $[0,1)$ *if the process* $t \mapsto W_m(t)$ *converges in law w.r.t.* ν *to the Brownian motion on* $[0,1)$ *as* $m \to \infty$.

Theorem 6.9. (cf. Corollary 2.1 in [32]) *Let* $\mathcal{P} = \{P_n\}_{n=1}^{\infty}$ *and* $\mathcal{Q} = \{Q_n\}_{n=1}^{\infty}$ *be sequences of interval partitions. Suppose that for some constants* $c > 0$ *and* $d > 0$, $\mathcal{P}$ *has entropy* c λ*-a.e. and* $\mathcal{Q}$ *has entropy* d λ*-a.e. Furthermore, suppose that* $\mathcal{P}$ *satisfies the* 0*-property and that* $\mathcal{Q}$ *satisfies the weak invariance principle with variance* σ^2 *and w.r.t. some probability measure* ν *on* $[0,1)$. *Then for all* $u \in \mathbb{R}$

$$\lim_{n\to\infty} \nu\Big(\Big\{x \in [0,1) : \frac{m_{\mathcal{P},\mathcal{Q}}(n,x) - n\frac{c}{d}}{\sigma_1\sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt, \tag{6.28}$$

where $\sigma_1 = \sqrt{\frac{c}{d^3}}\sigma$.

6.3 Extension to Number-Theoretic Fibered Maps

In [16], Dajani and Fieldsteel introduce the following class of interval maps:

Definition 6.10. *A surjective map* $T : [0,1) \to [0,1)$ *is said to be a* number-theoretic fibered map (NTFM) *if it satisfies the following conditions:*

1. *There exists a finite or countable partition* α *of* $[0,1)$ *into intervals such that* T *restricted to each interval is strictly monotonic and continuous. Furthermore,* α *is a generator w.r.t.* T *in the sense of* (2.42).

2. *There exists a Borel probability measure* μ_T *on* $[0,1)$ *that is invariant and ergodic w.r.t.* T *and is absolutely continuous w.r.t.* λ *such that*

$$\exists M > 0 : \frac{1}{M} \le \frac{d\mu_T}{d\lambda} \le M. \tag{6.29}$$

Remark 6.11. Note that in Definition 6.10 a sufficient condition for α to be a generator w.r.t. T is when T is expanding.

Remark 6.12. From the first part of Theorem 2.10 it follows that an NTFM admits a unique invariant probability measure μ_T that satisfies (6.29).

Remark 6.13. Let us write $\alpha = \{A_i : i \in D\}$ for the partition in Definition 6.10, where D is a finite or countable index set. The requirement that α is a generator w.r.t. T implies that for λ-a.e. pair of different points $x, y \in [0,1)$ there exists $n \in \mathbb{N}$ such that x and y are contained in different cylinders of order n w.r.t. T. Hence, knowing all the digits $\{i_k \in D : k = 0, 1, \ldots\}$ for which $T^k x \in A_{i_k}$ determines $x \in [0,1)$ uniquely λ-a.e. For this reason, if T is not too complex, iterations of T generate representations of points in terms of a sequence of digits in D. It appears (see [16]) that almost all known expansions on $[0,1)$ are generated by an NTFM.

Example 6.14. Let $T : [0,1) \to [0,1)$ be such that it (has an extension to $[0,1]$ that) satisfies the conditions in Proposition 3.16. (For example, $Tx = \beta x \bmod 1$ with $\beta > 1$ as in Example 3.18.) Then T is an NTFM with corresponding partition α that is equal to the partition on which T is piecewise monotonic. Note that α is a generator w.r.t. T because T is expanding. It follows in a similar way from the Folklore Theorem (Theorem 3.21) that each interval map $T : [0,1) \to [0,1)$ that can be extended to a Markov transformation on $[0,1]$ is an NTFM.

Let T be an NTFM with corresponding partition α and measure μ_T. For each $n \in \mathbb{N}$ we define the interval partition

$$\alpha_n = \bigvee_{i=0}^{n-1} T^{-i}\alpha, \tag{6.30}$$

which consists of the cylinders of order n w.r.t. T. Suppose that the entropy of the partition α w.r.t. μ_T is finite, i.e. $H_{\mu_T}(\alpha) < \infty$. It follows from the Kolmogorov-Sinai Theorem (Theorem 2.51) and the Shannon-McMillan-Breiman Theorem (Theorem 2.53) that

$$\lim_{n\to\infty} -\frac{\log \mu_T\big(\alpha_n(x)\big)}{n} = h_{\mu_T}(T), \qquad \mu_T\text{-a.e.} \tag{6.31}$$

Because of (6.29), we can replace μ_T in (6.31) by λ so that we get

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\alpha_n(x)\big)}{n} = h_{\mu_T}(T), \qquad \lambda\text{-a.e.} \tag{6.32}$$

In other words, the sequence of interval partitions $\{\alpha_n\}_{n=1}^{\infty}$ has entropy $h_{\mu_T}(T)$ λ-a.e. Hence, Theorem 6.6 applies, which proves the following theorem:

Theorem 6.15. (Theorem 5 in [16]) *Let T and S be two NTFM's with corresponding partitions α and β, respectively, and measures μ_T and μ_S, respectively. For each $n \in \mathbb{N}$ and $x \in [0,1)$, put*

$$m_{T,S}(n,x) = \sup\{m \in \mathbb{N} : \alpha_n(x) \subseteq \beta_m(x)\}. \tag{6.33}$$

Suppose that $h_{\mu_T}(T), h_{\mu_S}(S) \in (0,\infty)$. Then

$$\lim_{n\to\infty} \frac{m_{T,S}(n,x)}{n} = \frac{h_{\mu_T}(T)}{h_{\mu_S}(S)}, \qquad \lambda\text{-a.e.} \tag{6.34}$$

Remark 6.16. Recall from Subsection 1.1.2 that λ is invariant and ergodic w.r.t. the decimal map T, and that $h_\lambda(T) = \log 10$. Morover, we discussed that the Gauss measure μ_G on $[0,1)$ given by

$$\mu_G(A) = \int_A \frac{1}{\log 2}\frac{1}{1+x}dx, \qquad A \subseteq [0,1) \text{ Borel} \tag{6.35}$$

is invariant and ergodic w.r.t. the Gauss map S, and that $h_{\mu_G}(S) = \frac{\pi^2}{6\log 2}$. From Theorem 6.15 we now obtain Lochs' result in Theorem 6.1.

Let us now state the central limit result in Theorem 6.9 for NTFM's.

Definition 6.17. *Let T be an NTFM with corresponding partition α and measure μ_T.*

- *We say that T satisfies the* 0-property *if $\{\alpha_n\}_{n=1}^{\infty}$ satisfies the 0-property in the sense of Definition 6.7, with $c = h_{\mu_T}(T)$.*
- *We say that T satisfies the* weak invariance principle with variance σ^2 and w.r.t. some probability measure ν on $[0,1)$ *if $\{\alpha_n\}_{n=1}^{\infty}$ satisfies the weak invariance principle with variance σ^2 and w.r.t. ν in the sense of Definition 6.8, with $d = h_{\mu_T}(T)$.*

Remark 6.18. Because of (6.29), we can replace λ by μ_T in both (6.26) and (6.27) applied to $\{\alpha_n\}_{n=1}^{\infty}$.

The following theorem is now an easy consequence of Theorem 6.9:

Theorem 6.19. (Corollary 2.1 in [32]) *Let T be an NTFM that satisfies the* 0*-property and S be an NTFM that satisfies the weak invariance principle with variance σ^2 and w.r.t. to some probability measure ν on $[0,1)$. Then for all $u \in \mathbb{R}$*

$$\lim_{n\to\infty} \nu\left(\left\{x \in [0,1) : \frac{m_{T,S}(n,x) - n\frac{h_{\mu_T}(T)}{h_{\mu_S}(S)}}{\sigma_1\sqrt{n}} \leq u\right\}\right) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}dt, \tag{6.36}$$

where $\sigma_1 = \sqrt{\frac{h_{\mu_T}(T)}{h_{\mu_S}(S)^3}}\sigma$.

Example 6.20. (from Section 3.2 in [32]) As an example of an NTFM that satisfies the 0-property, consider $Tx = mx \mod 1$ with $m \geq 2$ integer. It is clear that T together with the partition $\alpha = \left\{\left[\frac{k}{m}, \frac{k+1}{m}\right), k = 0, 1, \ldots, m-1\right\}$ and λ defines an NTFM. Furthermore, for each $n \in \mathbb{N}$ and $x \in I$ we have $\lambda\big(\alpha_n(x)\big) = m^{-n}$. Hence, from the Kolmogorov-Sinai Theorem and the Shannon-McMillan-Breiman Theorem we get $h_{\mu_T}(T) = \log m$. It is now easy to see that T satisfies the 0-property. Another example of an NTFM that satisfies the 0-property is $Tx = \beta x \mod 1$ with β the golden mean, i.e. $\beta = \frac{1+\sqrt{5}}{2}$, which is shown in [32].

6.4 Extension to Random Number-Theoretic Fibered Systems

Analogous to the definition of an NTFM, we define the following class of random interval maps:

Definition 6.21. *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Lebesgue space, and let $\varphi : \Omega \to \Omega$ be measure preserving w.r.t. $\mathbb{P}$. A measurable map $T : \Omega \times [0,1) \to [0,1)$ given by $T(\omega, x) = T_\omega x$ is a* random number-theoretical fibered system (RNTFS) *w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ if*

1. *There exists a finite or countable partition ξ of $\Omega \times [0,1)$ such that, for each $\omega \in \Omega$ and $Z \in \xi$, $Z_\omega := \{x \in [0,1) : (\omega, x) \in Z\}$ is an interval and T_ω restricted to Z_ω is strictly monotonic and continuous. Furthermore, for $\mathbb{P}$-a.a. $\omega \in \Omega$, ξ is ω-generating w.r.t. (φ, T) in the sense of* (5.19).

2. *Letting $F_{\varphi,T}(\omega,x) = (\varphi\omega, T_\omega x)$, there exists an $F_{\varphi,T}$-invariant and $F_{\varphi,T}$-ergodic probability measure μ_T on $\Omega \times [0,1)$ that is absolutely continuous w.r.t. $\mathbb{P} \otimes \lambda$ such that*

$$\exists M > 0 : \frac{1}{M} \leq \frac{d\mu_T}{d\mathbb{P} \otimes \lambda} \leq M. \tag{6.37}$$

Remark 6.22. Note that in Definition 6.21 a sufficient condition for ξ to be ω-generating w.r.t. (φ, T) for each $\omega \in \Omega$ is when each T_ω is surjective and T is *expanding*, i.e. each T_ω is piecewise C^1 and $\inf_{(\omega,x)} |T'_\omega(x)| > 1$.

Remark 6.23. Again, from the first part of Theorem 2.10 it follows that $F_{\varphi,T}$ in Definition 6.21 admits a unique invariant probability measure μ_T that satisfies (6.37).

Remark 6.24. Note that φ in Definition 6.21 is ergodic w.r.t. $\mathbb{P}$. Indeed, suppose that $A \in \mathcal{F}$ satisfies $\varphi^{-1}A = A$. Then $F^{-1}_{\varphi,T}(A \times [0,1)) = \varphi^{-1}A \times [0,1) = A \times [0,1)$, which implies $\mu(A \times [0,1)) \in \{0,1\}$. From (6.37) it follows that $\mathbb{P}(A) \in \{0,1\}$.

For an RNTFS given by $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$ with corresponding partition ξ, we define for each $\omega \in \Omega$ and $n \in \mathbb{N}$ the interval partition as in (5.8) by

$$\xi_{\omega,n} = \xi_\omega \vee \bigvee_{k=1}^{n-1} T_\omega^{-1} T_{\varphi(\omega)}^{-1} \cdots T_{\varphi^{k-1}(\omega)} \xi_{\varphi^k(\omega)}, \tag{6.38}$$

where $\xi_\omega = \{Z_\omega : Z \in \xi\}$ and $Z_\omega = \{x \in I : (\omega, x) \in Z\}$. Similar as for NTFM's, the elements of $\xi_{\omega,n}$ are called the *cylinders of order n with respect to $F_{\varphi,T}(\omega,\cdot)$*.

Example 6.25. Let E be countable, and for each $j \in E$, let $T_j : [0,1) \to [0,1)$ be surjective and finitely piecewise C^2-monotonic on some interval partition α_j. Assume that $\inf_{(j,x)} |T'_j(x)| > 1$ and that the random covering property from Proposition 4.16 holds for $\{T_j\}_{j\in E}$. Furthermore, let $(p_j)_{j\in E}$ be a probability vector such that $p_j > 0$ holds for each $j \in E$. Then Proposition 4.16 yields a probability measure μ_T on $\Omega_E \times I$ with $\Omega_E = E^{\mathbb{N}}$ that meets the conditions in the second property of Definition 6.21, where $\mathbb{P}$ is in this case the Bernoulli measure on Ω_E associated with $(p_j)_{j\in E}$. Therefore, $T : \Omega_E \times [0,1) \to [0,1)$ given by $T(\omega, x) = T_{\omega_1} x$ is an RNTFS w.r.t. the Bernoulli shift $(\Omega_E, \mathcal{F}_{\mathbb{P}}, \mathbb{P}, \sigma)$ with corresponding partition ξ given by

$$\xi = \bigcup_{j\in E} \{[j] \times A : A \in \alpha_j\}, \tag{6.39}$$

where we use the notation from (4.23). Note, for each $\omega \in \Omega_E$, that $\xi_\omega = \alpha_{\omega_1}$ and that

$$\xi_{\omega,n} = \alpha_{\omega_1} \vee \bigvee_{k=1}^{n-1} T_{\omega_1}^{-1} T_{\omega_2}^{-1} \cdots T_{\omega_k}^{-1} \alpha_{\omega_{k+1}}. \tag{6.40}$$

In a similar way, we can with Proposition 4.37 obtain an RNTFS with corresponding partition as in (6.39) if the underlying base is a Markov shift.

Example 6.26. An explicit example for a family $\{T_j\}_{j\in E}$ as in Example 6.25 is given in Example 4.18, where $T_j = \beta_j x \bmod 1$ and $\inf_{j\in E} \beta_j > 1$. As we have seen in Subsection

1.1.3, such a family $\{T_j\}_{j\in E}$ generates for each $\omega \in \Omega_E$ expansions of the form

$$x = \sum_{k=1}^{\infty} \frac{b(\omega_k, x)}{\beta_{\omega_1} \cdots \beta_{\omega_k}}, \qquad b(\omega_k, x) \in \{0, 1, \ldots, \lfloor \beta_{\omega_k} \rfloor\} \text{ for each } k \in \mathbb{N}. \tag{6.41}$$

Remark 6.27. Let us write $\xi_\omega = \{A_{\omega,i} : i \in D_\omega\}$ for each $\omega \in \Omega$, where ξ is the partition from Definition 6.21 and each D_ω is a finite or countable index set. The requirement that ξ is ω-generating w.r.t. (φ, T) implies that for λ-a.e. pair of different points $x, y \in [0, 1)$ there exists $n \in \mathbb{N}$ such that x and y are contained in different cylinders of order n w.r.t. $F_{\varphi,T}(\omega, \cdot)$. Hence, knowing all the digits $\{i_k \in D_{\varphi^k(\omega)} : k = 0, 1, \ldots\}$ for which $x \in A_{\omega,i_0}$ and $T_{\varphi^{k-1}(\omega)} \cdots T_\omega x \in A_{\varphi^k(\omega),i_k}$ $(k \geq 1)$ determines $x \in [0, 1)$ uniquely λ-a.e. For this reason, if (like in the previous example) T is not too complex, iterations of $F_{\varphi,T}(\omega, \cdot)$ generate (after projecting on $[0, 1)$) expansions of points in $[0, 1)$ where the nth digit is in $D_{\varphi^n(\omega)}$.

In order to apply Theorem 6.6 to any pair of RNTFS's that are of the form such as in Example 6.25, we need the following proposition:

Proposition 6.28. *Let E be countable, $\Omega_E = E^{\mathbb{N}}$, $\mathcal{F}$ the Borel σ-algebra on Ω_E and $\mathbb{P}$ a probability measure on $(\Omega_E, \mathcal{F})$ that is invariant w.r.t. the left shift σ on Ω_E. Let $T : \Omega_E \times I \to I$ given by $T(\omega, x) = T_{\omega_1}(x)$ be an RNTFS w.r.t. $(\Omega_E, \mathcal{F}, \mathbb{P}, \sigma)$ such that the corresponding partition ξ is of the form*

$$\xi = \bigcup_{j\in E} \{[j] \times A : A \in \alpha_j\} \tag{6.42}$$

with α_j an interval partition for each $j \in E$. Then

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\xi_{\omega,n}(x)\big)}{n} = h_{\mu_T}(F_{\sigma,T}) - h_{\mathbb{P}}(\sigma), \qquad \lambda\text{-a.e.} \tag{6.43}$$

for $\mathbb{P}$-a.a. $\omega \in \Omega_E$. In other words, for $\mathbb{P}$-a.a. $\omega \in \Omega_E$, the sequence of interval partitions $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ has entropy $h_{\mu_T}(F_{\sigma,T}) - h_{\mathbb{P}}(\sigma)$ λ-a.e.

Proof: It is clear that $\{[j] : j \in E\}$ is a generator w.r.t. σ. Combining this with the fact that ξ is ω-generating w.r.t. (σ, T) for $\mathbb{P}$-a.a. $\omega \in \Omega_E$, note that for $\mathbb{P} \otimes \lambda$-a.e. pair of different points $(\omega, x), (\tilde{\omega}, y) \in \Omega_E \times I$ there exists $n \in \mathbb{N}$ such that (ω, x) and $(\tilde{\omega}, y)$ are in different elements of the partition $\bigvee_{i=0}^{n-1} F_{\sigma,T}^{-i}\xi$. It follows (see e.g. the remark in Section 7.5 of [66]) that ξ is a generating partition for $F_{\sigma,T}$. Hence, from the Shannon-McMillan-Breiman Theorem (Theorem 2.53) we obtain that

$$\lim_{n\to\infty} -\frac{\log \mu_T\big(\bigvee_{i=0}^{n-1} F_{\sigma,T}^{-i}\xi(\omega, x)\big)}{n} = h_{\mu_T}(F_{\sigma,T}), \qquad \mu_T\text{-a.e.} \tag{6.44}$$

Because of (6.37), we can interchange μ_T for $\mathbb{P} \otimes \lambda$, i.e.

$$\lim_{n\to\infty} -\frac{\log \mathbb{P} \otimes \lambda\big(\bigvee_{i=0}^{n-1} F_{\sigma,T}^{-i}\xi(\omega, x)\big)}{n} = h_{\mu_T}(F_{\sigma,T}), \qquad \mathbb{P} \otimes \lambda\text{-a.e.} \tag{6.45}$$

Note that

$$\bigvee_{i=0}^{n-1} F_{\sigma,T}^{-i}\xi(\omega,x) = [\omega_1^n] \times \xi_{\omega,n}(x) = [\omega_1^n] \times \Big(\alpha_{\omega_1} \vee \bigvee_{k=1}^{n-1} T_{\omega_1}^{-1} T_{\omega_2}^{-1} \cdots T_{\omega_k}^{-1} \alpha_{\omega_{k+1}}\Big)(x), \quad (6.46)$$

so that inserting this in (6.45) yields (6.43) (using Remark 6.24 and Theorem 2.53). □

Remark 6.29. For a general RNTFS given by $(\Omega, \mathcal{F}, \mathbb{P}, \varphi, T)$, note from (4.6) that $\bigvee_{i=0}^{n-1} F_{\varphi,T}^{-k}\xi$ consists of sets of the form

$$\bigcup_{\omega\in\Omega} \{\omega\} \times \Big((Z_0)_\omega \cap \Big(\bigcap_{k=1}^{n-1} T_\omega^{-1} T_{\varphi(\omega)}^{-1} \cdots T_{\varphi^{k-1}(\omega)}^{-1}((Z_k)_{\varphi^k(\omega)})\Big)\Big) \quad (6.47)$$

where $Z_0, \ldots, Z_{n-1} \in \xi$. In particular, the set in (6.47) is not of the form $A \times B$ with $A \in \mathcal{F}$ and $B \in \mathcal{B}$ Borel like in (6.46). Hence, with the line of reasoning in the proof of Proposition 6.28 we cannot extend the result in (6.43) to a general RNTFS.

Example 6.30. Let us again consider the setting of Example 6.26. We take $E = \{0, 1\}$ and let $T_0, T_1 : [0, 1) \to [0, 1)$ be given by $T_0 x = Nx \bmod 1$ and $T_1 x = Mx \bmod 1$ with $M, N \geq 2$ integers. Furthermore, let $p \in (0, 1)$. Writing $(\Omega_E, \mathcal{F}, \mathbb{P}, \sigma)$ for the one-sided Bernoulli shift with corresponding probability vector $(p_0, p_1) = (p, 1 - p)$, it is easy to see that the skew product $F_{\sigma,T}(\omega, x) = (\sigma\omega, T_{\omega_1} x)$ is measure preserving w.r.t $\mathbb{P} \otimes \lambda$. Moreover, from Example 4.18 it follows that $(F_{\sigma,T}, \mathbb{P} \otimes \lambda)$ is ergodic. We know from (6.41) that $F_{\sigma,T}(\omega, \cdot)$ generates for each $\omega \in \Omega_E$ expansions of the form

$$x = \sum_{k=1}^{\infty} \frac{b(\omega_k, x)}{N^{k-c_k(\omega)} M^{c_k(\omega)}}, \quad (6.48)$$

where $b(0, x) \in \{0, 1, \ldots, N - 1\}$, $b(1, x) \in \{0, 1, \ldots, M - 1\}$ and $c_k(\omega) = \sum_{i=1}^k \omega_i$. Now let $(\Omega_{E'}, \mathcal{F}', \mathbb{P}', \sigma')$ denote the Bernoulli shift with $E' = \{0, 1, \ldots, N + M - 1\}$ and corresponding probability vector $(p'_0, \ldots, p'_{N+M-1})$ with $p'_i = \frac{p_0}{N}$ for $i = 0, 1, \ldots, N - 1$ and $p'_i = \frac{p_1}{M}$ for $i = N, \ldots, N + M - 1$. In a similar way as in Example 3.1.2 of [15] one can derive that the dynamical systems $(\Omega_E \times [0, 1), \mathcal{F} \otimes \mathcal{B}, \mathbb{P} \otimes \lambda, F_{\sigma,T})$ and $(\Omega_{E'}, \mathcal{F}', \mathbb{P}', \sigma')$ are isomorphic with an isomorphism $\psi(\omega, x) = \{\omega_k \cdot N + b(\omega_k, x)\}_{k=1}^{\infty}$. Hence, combining this with Example 2.52 and Theorem 2.50 yields that

$$h_{\mathbb{P}\otimes\lambda}(F_{\sigma,T}) = -\sum_{i=0}^{N+M-1} p'_i \log p'_i = -p_0 \log\Big(\frac{p_0}{N}\Big) - p_1 \log\Big(\frac{p_1}{M}\Big). \quad (6.49)$$

Furthermore, Example 2.52 gives that $h_{\mathbb{P}}(\sigma) = -p_0 \log(p_0) - p_1 \log(p_1)$. Defining $\xi = \{[0] \times [\frac{i}{N}, \frac{i+1}{N}) : i = 0, \ldots, N-1\} \cup \{[1] \times [\frac{i}{M}, \frac{i+1}{M}) : i = 0, \ldots, M-1\}$, we conclude from Proposition 6.28 that

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\xi_{\omega,n}(x)\big)}{n} = p_0 \log(N) + p_1 \log(M), \qquad \lambda\text{-a.e.} \quad (6.50)$$

for $\mathbb{P}$-a.a. $\omega \in \Omega_E$. Note that the right-hand side of (6.50) is the weighted sum of the entropies of T_0 and T_1 w.r.t. λ.

Proposition 6.31. *Let T be an RNTFS w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ with corresponding partition ξ and measure μ_T. Suppose that the family $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A \frac{d\mu_T}{d\mathbb{P}\otimes\lambda}(\omega, x)d\lambda(x)$ is equivariant w.r.t. (φ, T). Then*

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\xi_{\omega,n}(x)\big)}{n} = h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi), \qquad \lambda\text{-a.e.} \tag{6.51}$$

for $\mathbb{P}$-a.a. $\omega \in \Omega$. That is, $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ has entropy $h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$ λ-a.e.

Proof: It follows from Proposition 5.14, Theorem 5.16 and Theorem 5.17 that

$$\lim_{n\to\infty} -\frac{\log \rho_\omega\big(\xi_{\omega,n}(x)\big)}{n} = h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi), \qquad \mu_T\text{-a.e.} \tag{6.52}$$

Because of (6.37), we know that $\mathbb{P}\otimes\lambda$ is absolutely continuous w.r.t. μ_T and that there exists $M > 0$ such that $\frac{1}{M} \leq \frac{d\rho_\omega}{d\lambda} \leq M$ for $\mathbb{P}$-a.a. $\omega \in \Omega$. Hence, we can in (6.52) replace ρ_ω with λ and μ_T-a.e. with $\mathbb{P}\otimes\lambda$-a.e. This yields (6.51). □

Remark 6.32. Note that for the proof of Proposition 6.31 we can weaken the assumption in (6.37) by merely requiring that

$$\exists M_\omega > 0 : \frac{1}{M_\omega} \leq \frac{d\rho_\omega}{d\lambda} \leq M_\omega \tag{6.53}$$

for $\mathbb{P}$-a.a. $\omega \in \Omega$.

Example 6.33. Let T be an RNTFS w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ with corresponding partition ξ and measure μ_T. Suppose there exists a probability measure ρ on $[0, 1)$ that is absolutely continuous w.r.t. λ and such that T_ω is invariant w.r.t. ρ for $\mathbb{P}$-a.a. $\omega \in \Omega$. Then for each $A \in \mathcal{F}$ and $B \in \mathcal{B}$ we have

$$\begin{aligned}\mathbb{P}\otimes\rho(F^{-1}(A\times B)) &= \int_{\varphi^{-1}A} \rho(T_\omega^{-1}B)d\mathbb{P} = \int_\Omega 1_A(\varphi\omega)\rho(B)d\mathbb{P}\\ &= \mathbb{P}\otimes\rho(A\times B),\end{aligned}$$

so it follows from the first part of Theorem 2.10 that $\mu_T = \mathbb{P}\otimes\rho$. Then obviously $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A \frac{d\mu_T}{d\mathbb{P}\otimes\lambda}(\omega, x)d\lambda(x) = \rho(A)$ is equivariant w.r.t. (φ, T). We conclude from Proposition 6.31 that, for $\mathbb{P}$-a.a. $\omega \in \Omega$, $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ has entropy $h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$ λ-a.e.

Example 6.34. Let T be an RNTFS w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ with corresponding partition ξ and measure μ_T. Suppose that $\varphi : \Omega \to \Omega$ is invertible. Then it follows from the second part of Proposition 5.6 that $\{\rho_\omega\}_{\omega\in\Omega}$ given by $\rho_\omega(A) = \int_A \frac{d\mu_T}{d\mathbb{P}\otimes\lambda}(\omega, x)d\lambda(x)$ is equivariant w.r.t. (φ, T). We obtain from Proposition 6.31 that, for $\mathbb{P}$-a.a. $\omega \in \Omega$, $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ has entropy $h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$ λ-a.e.

Example 6.35. Let us give an explicit case of the previous example. Let $\gamma = 1+\varepsilon$ with $\varepsilon > 0$ small, and take $\Omega = [\gamma, \infty)$ with corresponding Lebesgue σ-algebra $\mathcal{F}$. Let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$ and $\varphi : \Omega \to \Omega$ be an automorphism such that $(\varphi, \mathbb{P})$ is ergodic. Furthermore, let $T : \Omega\times[0, 1) \to [0, 1)$ be given by $T(\omega, x) = T_\omega x = \omega x \bmod 1$.

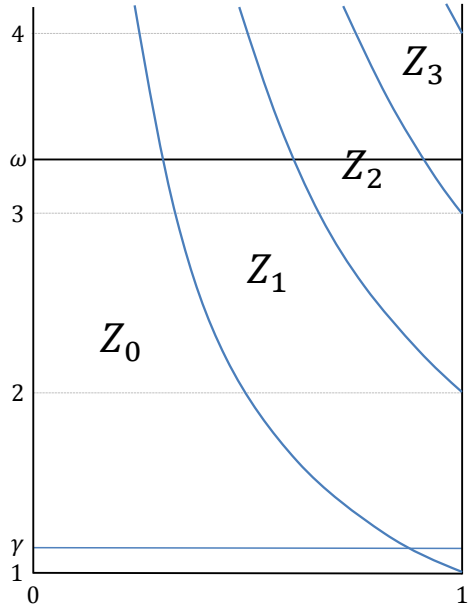

Figure 6.1: Representation of the partition ξ from Example 6.35. The fiber ξ_ω that is visualized is clearly an interval partition.

Define the partition $\xi = \{Z_i : i = 0, 1, \ldots\}$ of $\Omega \times [0, 1)$ as

$$\begin{aligned} Z_0 &= \left\{(x, y) : 0 < x < 1, \gamma \leq y < \frac{1}{x}\right\}, \\ Z_1 &= \left\{(x, y) : 0 < x < 1, \max\left(\gamma, \frac{1}{x}\right) \leq y < \frac{2}{x}\right\}, \\ Z_i &= \left\{(x, y) : 0 < x < 1, \frac{i}{x} \leq y < \frac{i+1}{x}\right\}, \qquad i \geq 2 \end{aligned}$$

(see Figure 6.1). Then for each $\omega \in \Omega$ we have $\xi_\omega = \{(Z_i)_\omega : i = 0, 1, \ldots\}$ with $(Z_i)_\omega = [\frac{i}{\omega}, \frac{i+1}{\omega})$ if $i \in \{0, 1, \ldots, \lfloor\omega\rfloor - 1\}$, $(Z_{\lfloor\omega\rfloor})_\omega = [\frac{\lfloor\omega\rfloor}{\omega}, 1)$ and $(Z_i)_\omega = \emptyset$ for $i > \omega$. It is therefore clear that T_ω is piecewise monotonic on ξ_ω for each $\omega \in \Omega$. Furthermore, ξ is ω-generating w.r.t. (φ, T) for each $\omega \in \Omega$ because $\inf_{(\omega,x)} |T'_\omega(x)| = \gamma > 1$. Moreover, according to Example 4.43 there exists an $F_{\varphi,T}$-invariant and $F_{\varphi,T}$-ergodic probability measure μ_T on $\Omega \times I$ that is absolutely continuous w.r.t. $\mathbb{P} \otimes \lambda$ such that

$$\exists M_\omega > 0 : \frac{1}{M_\omega} \leq \frac{d\mu_T}{d\mathbb{P} \otimes \lambda}(\omega, \cdot) \leq M_\omega \tag{6.54}$$

for $\mathbb{P}$-a.a. $\omega \in \Omega$. Together with Remark 6.32 we conclude that, for $\mathbb{P}$-a.a. $\omega \in \Omega$, $\{\xi_{\omega,n}\}_{n=1}^\infty$ has entropy $h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$ λ-a.e.

The following theorem is an easy consequence of Theorem 6.6.

Theorem 6.36. *Let T and S be RNTFS's w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \psi)$, respectively, each satisfying the conditions of Proposition 6.28 or Proposition 6.31 (so we distinguish four cases), and with corresponding partitions ξ and ζ, respectively, and measures μ_T and*

μ_S, respectively. For each $n \in \mathbb{N}$, $\omega \in \Omega$, $\tilde{\omega} \in \tilde{\Omega}$ and $x \in [0,1)$, put

$$m_{T,S}(n,\omega,\tilde{\omega},x) = \sup\{m \in \mathbb{N} : \xi_{\omega,n}(x) \subseteq \zeta_{\tilde{\omega},m}(x)\}. \tag{6.55}$$

Suppose that $h_{\mathbb{P}}(\varphi) < h_{\mu_T}(F_{\varphi,T}) < \infty$ and $h_{\tilde{\mathbb{P}}}(\psi) < h_{\mu_S}(F_{\psi,S}) < \infty$. Then

$$\lim_{n\to\infty} \frac{m_{T,S}(n,\omega,\tilde{\omega},x)}{n} = \frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi)}, \qquad \lambda\text{-a.e.} \tag{6.56}$$

for $\mathbb{P} \otimes \tilde{\mathbb{P}}$-a.a. $(\omega,\tilde{\omega}) \in \Omega \times \tilde{\Omega}$.

Remark 6.37. The right-hand side of (6.56) is in general hard to calculate, but in practice it can be approximated using the convergence in (6.43) and (6.51).

For fixed $(\omega,\tilde{\omega}) \in \Omega \times \tilde{\Omega}$ we can formulate a *quenched* central limit result associated with Theorem 6.36, which directly follows from Theorem 6.9.

Theorem 6.38. *Let T and S be RNTFS's w.r.t. $(\Omega,\mathcal{F},\mathbb{P},\varphi)$ and $(\tilde{\Omega},\tilde{\mathcal{F}},\tilde{\mathbb{P}},\psi)$, respectively, with corresponding partitions ξ and ζ, respectively, and measures μ_T and μ_S, respectively. Let $(\omega,\tilde{\omega}) \in \Omega \times \tilde{\Omega}$ and suppose that $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ satisfies the 0-property in the sense of Definition 6.7 with $c = h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$, and that $\{\zeta_{\omega,n}\}_{n=1}^{\infty}$ satisfies the weak invariance principle with variance σ^2 and w.r.t. some probability measure ν on $[0,1)$ in the sense of Definition 6.8 with $d = h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi)$. Then for all $u \in \mathbb{R}$*

$$\lim_{n\to\infty} \nu\Big(\Big\{x \in [0,1) : \frac{m_{T,S}(n,\omega,\tilde{\omega},x) - nz_0}{\sigma_1\sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt, \tag{6.57}$$

where $z_0 = \frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi)}$ and $\sigma_1 = \sqrt{\frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{(h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi))^3}}\sigma$.

We can as well formulate an *averaged* central limit result corresponding to Theorem 6.36. For this we need the following definition:

Definition 6.39. *Let T be an RNTFS w.r.t. $(\Omega,\mathcal{F},\mathbb{P},\varphi)$ with corresponding partition ξ and measure μ_T. For each $\omega \in \Omega$, $x \in [0,1)$ and $m \in \mathbb{N}$ we put $W_{m,\omega,x}(0) = 0$ and*

$$W_{m,\omega,x}\Big(\frac{l}{m}\Big) = \frac{-\log\lambda\big(\xi_{\omega,l}(x)\big) - l\big(h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)\big)}{\sigma\sqrt{m}}, \quad l \in \{1,\dots,m\} \tag{6.58}$$

for some $\sigma > 0$, and we extend this linearly on the subintervals $\{[\frac{l-1}{m},\frac{l}{m}] : 1 \le l \le m\}$ so that $W_{m,\omega,x} \in C[0,1)$. We say that T satisfies the averaged weak invariance principle with variance σ^2 and w.r.t. some probability measure ν on $(\Omega \times [0,1), \mathcal{F} \otimes \mathcal{B})$ *if the process $t \mapsto W_m(t)$ converges in law w.r.t. ν to the Brownian motion on $[0,1)$ as $m \to \infty$.*

Example 6.40. Let us again consider the RNTFS given by $(\Omega_E,\mathcal{F},\mathbb{P},\sigma,T)$ from Example 6.30, where $E = \{0,1\}$, $T_0x = Nx \bmod 1$, $T_1x = Mx \bmod 1$ ($M,N \ge 2$ integers) and $\mathbb{P}$ the Bernoulli measure with corresponding probability vector (p_0,p_1). Since the corresponding partition ξ is equal to $\{[0] \times [\frac{i}{N},\frac{i+1}{N}) : i = 0,\dots,N-1\} \cup \{[1] \times [\frac{i}{M},\frac{i+1}{M}) : i = 0,\dots,M-1\}$, it is clear that $\lambda(\xi_{\omega,l}(x)) = N^{c_l(\omega)-l}M^{-c_l(\omega)}$ with $c_l(\omega) = \sum_{i=1}^{l}\omega_i$ for

each $\omega \in \Omega_E$ and $x \in [0,1)$. For each $i \in \mathbb{N}$, let us define the random variable X_i on $\Omega_E \times [0,1)$ as

$$X_i(\omega, x) = (1-\omega_i)\log(N) + \omega_i \log(M). \tag{6.59}$$

Clearly, $\{X_i\}_{i=1}^{\infty}$ is an i.i.d. sequence on $(\Omega_E \times [0,1], \mathcal{F} \otimes \mathcal{B}, \mathbb{P} \otimes \lambda)$. Also, we have

$$\mathbb{E}_{\mathbb{P}\otimes\lambda}(X_i) = p_0 \log(N) + p_1 \log(M) = h_{\mathbb{P}\otimes\lambda}(F_{\sigma,T}) - h_{\mathbb{P}}(\sigma). \tag{6.60}$$

Setting $\sigma^2 = \mathrm{Var}(X_i)$, $X_i' = \frac{X_i - \mathbb{E}_{\mathbb{P}\otimes\lambda}(X_i)}{\sigma}$ and $S_l = \sum_{i=1}^{l} X_i'$, we obtain that W_m as defined in Definition 6.39 in this case takes the form

$$W_m(t) = \frac{1}{\sqrt{m}} \sum_{l=1}^{m} \Big(S_{l-1} + m\Big(t - \frac{l-1}{m}\Big) X_l' \Big) 1_{[\frac{l-1}{m}, \frac{l}{m})}(t). \tag{6.61}$$

We conclude from Donsker's Theorem (see e.g. Theorem 1.4 in [37]) that $t \mapsto W_m(t)$ converges in law w.r.t. $\mathbb{P} \otimes \lambda$ to the Brownian motion on $[0,1)$ as $m \to \infty$, so T satisfies the averaged weak invariance principle with variance σ^2 and w.r.t. $\mathbb{P} \otimes \lambda$.

Theorem 6.41. *Let T and S be RNTFS's w.r.t. $(\Omega, \mathcal{F}, \mathbb{P}, \varphi)$ and $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}}, \psi)$, respectively, with corresponding partitions ξ and ζ, respectively, and measures μ_T and μ_S, respectively. Let $\omega \in \Omega$ and suppose that $\{\xi_{\omega,n}\}_{n=1}^{\infty}$ satisfies the 0-property in the sense of Definition 6.7 with $c = h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)$. Furthermore, suppose that S satisfies the averaged weak invariance principle with variance σ^2 and w.r.t. some probability measure ν on $(\tilde{\Omega} \times [0,1), \tilde{\mathcal{F}} \otimes \mathcal{B})$. Then for all $u \in \mathbb{R}$*

$$\lim_{n\to\infty} \nu\Big(\Big\{(\tilde{\omega}, x) \in \tilde{\Omega} \times [0,1) : \frac{m_{T,S}(n, \omega, \tilde{\omega}, x) - nz_0}{\sigma_1 \sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt, \tag{6.62}$$

where $z_0 = \frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi)}$ *and* $\sigma_1 = \sqrt{\frac{h_{\mu_T}(F_{\varphi,T}) - h_{\mathbb{P}}(\varphi)}{(h_{\mu_S}(F_{\psi,S}) - h_{\tilde{\mathbb{P}}}(\psi))^3}}\sigma$.

Proof: This follows in exactly the same way as the proof of Corollary 2.1 in [32]. This is because in Section 2.2 of [32] it is nowhere used that the defined processes W, K, K', M' and M have underlying probability space $([0,1), \mathcal{B}, \lambda)$ and not just an arbitrary Lebesgue space. For this reason, the proof also holds if we instead work with $(\tilde{\Omega} \times [0,1), \tilde{\mathcal{F}} \otimes \mathcal{B}, \nu)$ as underlying probability space for the process W. □

Appendix A

Functions of Bounded Variation

In this appendix we briefly review the theory of functions of bounded variation. The following results are well-known and we refer to e.g. [52] and [9] for a more detailed discussion on this topic.

Definition A.1. *Let* $[a,b] \subseteq \mathbb{R}$ *and* $f : [a,b] \to \mathbb{C}$. *The* variation of f *is defined as*

$$\mathrm{Var}_{[a,b]}(f) = \sup \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|, \tag{A.1}$$

where the supremum runs over all finite partitions generated by the points $a = x_0 < x_1 < \cdots < x_n = b$. *The space*

$$BV([a,b]) = \{f : [a,b] \to \mathbb{C} : \mathrm{Var}_{[a,b]}(f) < \infty\} \tag{A.2}$$

is called the space of functions of bounded variation *on* $[a,b]$.

The variation measures the oscillation of a function. Note that $\mathrm{Var}_{[a,b]}(\cdot)$ is a seminorm, because it is positive, and homogeneous and subadditive in the sense that

$$\mathrm{Var}_{[a,b]}(t \cdot f) = |t|\mathrm{Var}_{[a,b]}(f) \quad \text{for every } t \in \mathbb{C}, \tag{A.3}$$

$$\mathrm{Var}_{[a,b]}(f_1 + f_2) \leq \mathrm{Var}_{[a,b]}(f_1) + \mathrm{Var}_{[a,b]}(f_2). \tag{A.4}$$

It is not a norm because $\mathrm{Var}_{[a,b]}(f + C) = \mathrm{Var}_{[a,b]}(f)$ for every constant $C \in \mathbb{C}$. Also, $BV([a,b])$ is closed under taking products, because

$$\mathrm{Var}_{[a,b]}(f_1 \cdot f_2) \leq \|f_2\|_\infty \mathrm{Var}_{[a,b]}(f_1) + \|f_1\|_\infty \mathrm{Var}_{[a,b]}(f_2), \tag{A.5}$$

where $\|f_i\|_\infty = \sup |f_i|$ is the supremum norm.

Furthermore, $f : [a,b] \to \mathbb{C}$ is constant if and only if $\mathrm{Var}_{[a,b]}(f) = 0$. For $f : [a,b] \to \mathbb{R}$ it is easy to see that

$$\sup f - \inf f \leq \mathrm{Var}_{[a,b]}(f), \tag{A.6}$$

with equality if and only if f is monotone. Therefore, $|f|$ is bounded if $\mathrm{Var}_{[a,b]}(f) < \infty$.

We state some more well-known properties of variation.

Lemma A.2. *Let $\{f_n\}$ be a sequence of complex functions on $[a,b]$ converging pointwise to $f : [a,b] \to \mathbb{C}$. Then $\mathrm{Var}_{[a,b]}(f) \leq \liminf_n \mathrm{Var}_{[a,b]}(f_n)$.*

Lemma A.3. *If $f_1 : [a,b] \to [c,d]$ is monotone and $f_2 : [c,d] \to \mathbb{C}$, then*

$$\mathrm{Var}_{[a,b]}(f_2 \circ f_1) \leq \mathrm{Var}_{[c,d]}(f_2). \tag{A.7}$$

Lemma A.4. *Let $f : [a,b] \to \mathbb{C}$ be integrable. Then*

$$\|f\|_\infty \leq \mathrm{Var}_{[a,b]}(f) + \frac{1}{b-a}\int_a^b |f(x)| d\lambda(x). \tag{A.8}$$

The following lemma is Problem 5.4.1 in [9].

Lemma A.5. *Let $\{I_i\}$ be a finite or countable partition of $[a,b]$ into intervals. Let $f \in BV([a,b])$. Then*

$$\sum_i \mathrm{Var}_{I_i}(f) \leq \mathrm{Var}_{[a,b]}(f). \tag{A.9}$$

Theorem A.6. (Jordan decomposition of BV functions) *Let $f : [a,b] \to \mathbb{R}$. Then f is of bounded variation if and only if it can be represented as $f = u - v$ where u and v are two real-valued increasing functions on $[a,b]$.*

Remark A.7. Note that this decomposition is not unique, because we can instead take $u + g$ and $v + g$ for any increasing function $g : [a,b] \to \mathbb{R}$. In particular, u and v can be taken positive by adding a sufficiently large constant.

Remark A.8. It follows from Theorem A.6 that $f : [a,b] \to \mathbb{C}$ is of bounded variation if and only if it can be representated as $f = (u_r - v_r) + i(u_i - v_i)$ where u_r, v_r, u_i, v_i are real-valued increasing functions on $[a,b]$.

Since a monotonic function $f : [a,b] \to \mathbb{R}$ can only have jump discontinuities (i.e. points $x \in [a,b]$ for which $L = \lim_{y\uparrow x} f(y)$ and $M = \lim_{y\downarrow x} f(y)$ exist, but $L \neq M$) and since each such discontinuity can be associated with a rational number, we obtain the following corollary:

Corollary A.9. *Let $f \in BV([a,b])$. Then the set of discontinuities of f is at most countable and consists of jump discontinuities.*

Corollary A.10. *Let $f \in BV([a,b])$. Then f is Lebesgue integrable and Riemann integrable. In fact, $f \in \mathcal{L}^p([a,b],\lambda)$ holds for all $p \geq 1$.*

Proof: Both statements follow from Corollary A.9 and the fact that f is bounded. □

We call a function $f : [a,b] \to \mathbb{R}$ *lower semicontinuous* if for any $x \in [a,b]$ we have $f(x) \leq \liminf_{y\to x} f(y)$, i.e. if for all $\varepsilon > 0$ there exists $\delta > 0$ such that $f(y) \geq f(x) - \varepsilon$ for all $y \in (x-\delta, x+\delta) \cap [a,b]$.

Corollary A.11. *Let $f : [a, b] \to \mathbb{R}$ be of bounded variation. Then it can be redefined on a countable set to become lower semicontinuous. In that case, f takes its minimum on $[a, b]$.*

Proof: From Corollary A.9 we know that the set S of points in $[a, b]$ where f is discontinuous is at most countable, and that f has one-sided limits at every point in $[a, b]$. For each $x \in S$ we redefine f as

$$f(x) = \min\Big(\lim_{y\uparrow x} f(y), \lim_{y\downarrow x} f(y)\Big). \tag{A.10}$$

Note that in this way f becomes lower semicontinuous. The second statement is a well-known property of lower semi-continuous functions, see e.g. Theorem 8.1.1 in [9]. □

Let $f \in BV([a, b])$. Then from Corollary A.10 we know that $f \in \mathcal{L}^1([a, b], \lambda)$, so we can define

$$\|f\|_{BV} = \mathrm{Var}_{[a,b]}(f) + \|f\|_1. \tag{A.11}$$

Recall that $\|\cdot\|_1$ defined on $\mathcal{L}^1([a, b], \lambda)$ as $\|f\|_1 = \int_a^b |f|dx$ is a seminorm but not a norm, because $\|f\|_1 = 0$ only implies $f = 0$ λ-a.e. For that reason, $\|\cdot\|_{BV}$ defined on $BV([a, b])$ as in (A.11) is also a seminorm but not a norm. This problem can be circumvented by defining the equivalence relation $f \sim g$ if and only if $f = g$ λ-a.e. and considering the quotient space $L^1([a, b], \lambda) = \mathcal{L}^1([a, b], \lambda)/\sim$.

Definition A.12. *Let $f \in L^1([a, b], \lambda)$. The* variation of f *is defined to be*

$$\mathrm{Var}_{[a,b]}(f) = \inf\{\mathrm{Var}_{[a,b]}(g) : g \in \mathcal{L}^1([a, b]), g = f \ \lambda\text{-a.e.}\} \tag{A.12}$$

If $Var_{[a,b]}(f) < \infty$, then we say that f is of bounded variation *on $[a, b]$ and we let*

$$\widetilde{BV}([a, b]) = \{f \in L^1([a, b], \lambda) : \mathrm{Var}_{[a,b]}(f) < \infty\}. \tag{A.13}$$

We usually just write $BV([a, b])$ for the space $\widetilde{BV}([a, b])$.

Note that $\mathrm{Var}_{[a,b]}(\cdot)$ in (A.12) is a seminorm on $\widetilde{BV}([a, b])$. On the other hand, $\|\cdot\|_1$ induces a norm on $L^1([a, b], \lambda)$, so $\|\cdot\|_{BV} : \widetilde{BV}([a, b]) \to [0, \infty)$ defined as

$$\|f\|_{BV} = \mathrm{Var}_{[a,b]}(f) + \|f\|_1, \quad f \in \widetilde{BV}([a, b]) \tag{A.14}$$

is a norm on $\widetilde{BV}([a, b])$.

Proposition A.13. *The space $\widetilde{BV}([a, b])$ equipped with the norm $\|\cdot\|_{BV}$ is a complex Banach space.*

Proof: See Lemma 5(ii) in [33]. □

Let us now prove that $BV([a, b])$ contains all C^1 functions on $[a, b]$.

Proposition A.14. (see (15) in [10]) *Let $f_1 \in BV([a,b])$ and $f_2 \in C^1([a,b])$. Then*

$$\mathrm{Var}_{[a,b]}(f_1 f_2) \leq \|f_2\|_\infty \mathrm{Var}_{[a,b]}(f_1) + \int_a^b |f_1(s) f_2'(s)| ds. \tag{A.15}$$

Proof: From the Mean Value Theorem it follows that

$$\begin{aligned}
\mathrm{Var}_{[a,b]}(f_1 f_2) &= \sup \sum_{i=1}^n |f_1(x_i) f_2(x_i) - f_1(x_{i-1}) f_2(x_{i-1})| \\
&\leq \sup \sum_{i=1}^n |f_2(x_i)||f_1(x_i) - f_1(x_{i-1})| + |f_1(x_{i-1})||f_2(x_i) - f_2(x_{i-1})| \\
&\leq \|f_2\|_\infty \mathrm{Var}_{[a,b]}(f_1) + \sup \sum_{i=1}^n |f_1(x_{i-1}) f_2'(\xi_i)||x_i - x_{i-1}| \\
&= \|f_2\|_\infty \mathrm{Var}_{[a,b]}(f_1) + \int_a^b |f_1(s) f_2'(s)| ds,
\end{aligned}$$

where the last step follows by definition of the Riemann integral. □

Corollary A.15. *Let $f \in C^1([a,b])$. Then $f \in BV([a,b])$ and*

$$\mathrm{Var}_{[a,b]}(f) \leq \int_a^b |f'(s)| ds. \tag{A.16}$$

Proof: Apply Proposition A.14 with $f_1 \equiv 1$ and $f_2 = f$. □

Corollary A.16. *For any finite Borel measure μ on $[a,b]$, the space $BV([a,b])$ is dense in $(L^1(\mu), \|\cdot\|_{1,\mu})$.*

Proof: The result follows from the previous corollary combined with the fact that $C^1([a,b])$ is dense in $\mathcal{L}^1(\mu)$. □

As a preparation for proving the existency result by Lasota and Yorke in Section 3.3, we also need the following two important theorems:

Theorem A.17. (Yorke's Inequality) *Let $f \in BV([a,b])$ and $[c,d] \subseteq [a,b]$. Then*

$$\mathrm{Var}_{[a,b]}(f 1_{[c,d]}) \leq 2\mathrm{Var}_{[c,d]}(f) + \frac{2}{d-c} \int_c^d |f(s)| ds. \tag{A.17}$$

Proof: For any $\xi \in [c,d]$ we have

$$\begin{aligned}
\mathrm{Var}_{[a,b]}(f 1_{[c,d]}) &\leq \mathrm{Var}_{[c,d]}(f) + |f(c)| + |f(d)| \\
&\leq \mathrm{Var}_{[c,d]}(f) + |f(c) - f(\xi)| + |f(d) - f(\xi)| + 2|f(\xi)| \\
&\leq 2\mathrm{Var}_{[c,d]}(f) + 2|f(\xi)|.
\end{aligned}$$

We can choose ξ such that $|f(\xi)| \leq \frac{1}{d-c} \int_c^d |f(s)| ds$ by the Mean Value Theorem for integrals, which gives the result. □

Theorem A.18. (Helly's First Theorem) *Let $\mathcal{C}$ be a collection of infinitely many functions $f \in BV([a,b])$ for which there exists $M > 0$ such that*

$$\|f\|_\infty \leq M, \quad \mathrm{Var}_{[a,b]}(f) \leq M, \qquad \textit{for all } f \in \mathcal{C}.$$

Then there exists a sequence $\{f_n\} \subseteq \mathcal{C}$ that converges pointwise to some $f^* \in BV([a,b])$ that satisfies $\mathrm{Var}_{[a,b]}(f^*) \leq M$.

Appendix B

Some Results from Functional Analysis

In this appendix we state some results from Functional Analysis that will be needed in Chapters 3 and 4.

B.1 The Kakutani-Yosida Theorem

The next theorem can be found in e.g. Section 2.2 of [9].

Theorem B.1. (Kakutani-Yosida) *Let X be a Banach space and $P : X \to X$ be a bounded linear operator. Assume there exists $c > 0$ such that $\|P^n\| \leq c$ for each $n \in \mathbb{N}$ Moreover, if for any $f \in A \subseteq X$, the sequence $\{f_n\}$ given by*

$$f_n = \frac{1}{n} \sum_{k=1}^{n} P^k f \tag{B.1}$$

contains a subsequence $\{f_{n_k}\}$ which converges weakly in X, then for any $f \in \overline{A}$,

$$\frac{1}{n} \sum_{k=1}^{n} P^k f \to \tilde{f} \in X \tag{B.2}$$

(convergence in norm) and $P(\tilde{f}) = \tilde{f}$.

B.2 Quasi-Compact Operators

Let $(V, \|\cdot\|_V)$ be a complex Banach space and $P : V \to V$ a bounded linear operator.

Definition B.2. *Let $B_1(0) = \{f \in V : \|f\| < 1\}$ denote the open unit ball in V. We say that P is* compact *if the closure of $P(B_1(0))$ is compact in V.*

Compact operators are an important class of bounded linear operators in Functional Analysis. We have the following related notion of *quasi-compactness* of a bounded linear operator.

Definition B.3. *We say that P is* quasi-compact *if there exists a compact operator $R : V \to V$ and $k \in \mathbb{N}$ such that*

$$\|P^k - R\|_V < 1. \tag{B.3}$$

There are several equivalent definitions for the quasi-compactness of a bounded linear operator. One of them is formulated as follows (see [9, 21, 23]).

Theorem B.4. *The operator P is quasi-compact if and only if there are bounded linear operators $\{Q_\lambda : \lambda \in \Lambda\}$ and S on V such that*

$$\begin{aligned}
&P^n = \sum_{\lambda\in\Lambda} \lambda^n Q_\lambda + S^n, && \text{for all } n \in \mathbb{N},\\
&Q_\lambda Q_{\lambda'} = 0 && \text{if } \lambda \neq \lambda',\\
&Q_\lambda^2 = Q_\lambda && \text{for all } \lambda \in \Lambda,\\
&Q_\lambda S = S Q_\lambda = 0 && \text{for all } \lambda \in \Lambda,\\
&Q_\lambda V = E(\lambda) && \text{for all } \lambda \in \Lambda,\\
&\rho(S) < 1,
\end{aligned}$$

where Λ is the set of eigenvalues of P with modulus 1, $E(\lambda) = \{f \in V : Pf = \lambda f\}$ is the eigenspace *of P corresponding to $\lambda \in \Lambda$, and $\rho(S) = \lim_{n\to\infty} \|S^n\|_V^{1/n}$ is the* spectral radius *of S.*

The next theorem gives a useful sufficient condition for a bounded linear operator to be quasi-compact.

Theorem B.5. (Ionescu-Tulcea and Marinescu Theorem) *Let $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ be two complex Banach spaces such that $V \subseteq W$. Let $P : V \to V$ be a linear operator that is bounded with respect to both $\|\cdot\|_V$ and the restriction of $\|\cdot\|_W$ to V. Assume that*

1. *If $f_n \in V$ for $n \in \mathbb{N}$, $f \in W$, $\lim_{n\to\infty} \|f_n - f\|_W = 0$ and $\|f_n\|_V \leq K$ for $n \in \mathbb{N}$, then $f \in V$ and $\|f\|_V \leq M$, where M is a constant,*
2. $\sup_{n\geq 0}\{\|P^n f\|_W / \|f\|_W : f \in V, f \neq 0\} < \infty$,
3. *There exist $k \in \mathbb{N}$, $\rho \in (0, 1)$ and $L > 0$ such that*

$$\|P^k f\|_V \leq \rho \|f\|_V + L\|f\|_W \tag{B.4}$$

   for all $f \in V$,

4. *If $U \subseteq V$ is bounded w.r.t. $\|\cdot\|_V$, then the closure of $P^k U$ w.r.t. $\|\cdot\|_W$ is compact in $(W, \|\cdot\|_W)$.*

Then $P : (V, \|\cdot\|_V) \to (V, \|\cdot\|_V)$ is quasi-compact, the set Λ of eigenvalues of P with modulus 1 is finite and for each $\lambda \in \Lambda$ the eigenspace $E(\lambda)$ associated to λ is finite-dimensional.

B.3 The Arzelà-Ascoli Theorem

As usual, write $C^0([0,1])$ for the space of all continuous functions on $[0,1]$ with values in $\mathbb{C}$ or $\mathbb{R}$. We equip $C^0([0,1])$ with the supremum norm $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$. We say that a set $S \subseteq C^0([0,1])$ is *bounded* if there exists $M \in (0,\infty)$ such that $\|f\|_\infty \leq M$ for all $f \in S$. Moreover, we call S equicontinuous if for each $\varepsilon > 0$ there exists $\delta > 0$ such that for $x, y \in [0,1]$:

$$|x - y| < \delta \Rightarrow \sup_{f\in S} |f(x) - f(y)| < \varepsilon. \tag{B.5}$$

The following famous theorem can be found in e.g. [51].

Theorem B.6. (Arzelà-Ascoli) *If $S \subseteq C^0([0,1])$ is bounded and equicontinuous, then for any sequence $\{f_n\} \subseteq S$ there exists a subsequence $\{f_{n_k}\}$ that converges w.r.t. $\|\cdot\|_\infty$ to some $f^* \in C^0([0,1])$.*

The Arzelà-Ascoli Theorem can also be extended to general compact metric spaces (see e.g. Section 4.6 in [23]).

Bibliography

[1] L. Abramov and V. Rokhlin. *Entropy of a skew product of mappings with invariant measure* (russian). Vestnik Leningrad. Univ., 17 (1962), 5-13.

[2] R.L. Adler. *Continued fractions and Bernoulli Trials.* In: Ergodic theory. Moser, J., Phillips, E., Varadhan, S. (eds.). Lecture notes. New York: Courant Inst. Math. Sci. 1975.

[3] R.L. Adler. *F-expansions revisited.* Springer lecture notes 318, 1-5 (1973).

[4] R. Aimino, M. Nicol, and S. Vaienti. *Annealed and quenched limit theorems for random expanding dynamical systems.* Probability Theory and Related Fields 162.1-2 (2015): 233-274.

[5] W. Bahsoun and C. Bose. *Decay of correlation for random intermittent maps.* Nonlinearity. 27 (2014) 1543–1554.

[6] W. Bahsoun and C. Bose. *Mixing rates and limit theorems for random intermittent maps.* Nonlinearity 29.4 (2016): 1417.

[7] R. Bowen. *Invariant measures for Markov maps of the interval.* Comm. Math. Phys. (1979).

[8] R. Bowen. *Markov maps associated with Fuchsian groups.* I.H.E.S. Publications 50 (1979).

[9] A. Boyarsky and P. Góra. *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension.* Springer Science & Business Media, 2012.

[10] H. Bruin. *Notes on Ergodic Theory.* University of Vienna, `http://www.mat.univie.ac.at/~bruin/ET3.pdf`, accessed on August 13, 2018.

[11] H. Bruin and P. Oprocha. *On observable Li–Yorke tuples for interval maps.* Nonlinearity 28.6 (2015): 1675.

[12] J. Buzzi. *Absolutely continuous S.R.B. measures for random Lasota-Yorke maps.* Trans. Amer. Math. Soc., 352(7):3289-3303,2000.

[13] J. Buzzi. *Exponential decay of correlations for random Lasota-Yorke maps.* Comm. Math. Phys., 208(1):25-54, 1999.

[14] K. Dajani and M. de Vries. *Invariant densities for random β-expansions.* J. Eur. Math. Soc. (JEMS), 2007.

[15] K. Dajani and S. Dirksin. *A simple introduction to ergodic theory.* University of Utrecht, `http://www.staff.science.uu.nl/~kraai101/lecturenotes2009.pdf`, accessed on August 13, 2018.

[16] K. Dajani and A. Fieldsteel. *Equipartition of interval partitions and an application to number theory.* Proceedings of the American Mathematical Society 129.12 (2001): 3453-3460.

[17] K. Dajani and C. Kraaikamp. *Random β-expansions.* Ergodic Theory Dynam. Systems, 23 (2003), no.2, 461-479.

[18] K. Dajani, C. Kraaikamp, and B. Solomyak. *Ergodic theory of numbers.* Carus Mathematical Monographs, 29. Mathematical Association of America, Washington, DC 2002.

[19] W. De Melo and S. van Strien. *One-Dimensional Dynamics.* Springer-Verlag, New York (1993).

[20] J. Dereziński. Measure theory. Warsaw University, `https://pdfs.semanticscholar.org/683d/e51a87337e46d55c7130555ace486205ff31.pdf`, accessed on August 13, 2018.

[21] J. Ding and A. Zhou. *Statistical properties of deterministic systems.* Springer, 2010.

[22] T. Downarowicz. *Entropy in dynamical systems.* Cambridge University Press, New Mathematical Monographs 18, Cambridge 2011.

[23] N. Dunford and L. Schwartz. *Linear Operators, Part I.* Amsterdam: Interscience, 1957.

[24] P. Erdös, I. Joó, and V. Komornik. *Characterization of the unique expansions $1 = \sum_{i=1}^{\infty} q^{-n_i}$ and related problems.* Bull. Soc. Math. France 118 (1990), no. 3, 377-390.

[25] C. Faivre. *A central limit theorem related to decimal and continued fraction expansion.* Arch. Math. 70 (1998), no. 6, 455–463.

[26] C. Faivre. *On decimal and continued fraction expansions of a real number.* Acta Arith. 82 (1997), no. 2, 119–128.

[27] L. Fang, M. Wu, and B. Li. *Beta-expansion and continued fraction expansion of real numbers.* arXiv:1603.01081.

[28] G. Froyland. *Ulam's method for random interval maps.* Nonlinearity, 12(4):1029-1052,1999.

[29] P. Góra. *Invariant densities for piecewise linear maps of the unit interval.* Ergodic Theory Dynam. Systems, 29(5):1549-1583, 2009.

[30] P. Góra and A. Boyarsky. *Absolutely continuous invariant measures for random maps with position dependent probabilities.* J. Math. Anal. Appl. 278 (2003), 225-242.

[31] W.F. Hammond. *Continued Fractions and the Euclidean Algorithm.* University of Albany, `https://www.math.u-bordeaux.fr/~pjaming/M1/exposes/MA2.pdf`.

[32] A. Herczegh. *Central Limit Theorems in Ergodic Theory.* MSc thesis, Eötvös Loránd University, 2009.

[33] F. Hofbauer and G. Keller. *Ergodic properties of invariant measures for piecewise monotonic transformations.* Math. Zeitschrift 180 (1982a), 119-140.

[34] T. Inoue. *Invariant measures for position dependent random maps with continuous random parameters.* Studia Math., 208(1):11-29,2012.

[35] S. Isola. *On the spectrum of Farey and Gauss maps.* Nonlinearity 15 (2002), 1521-1539.

[36] C. Kalle, T. Kempton, and E. Verbitskiy. *The random continued fraction transformation.* Nonlinearity 30.3 (2017): 1182.

[37] D. Khoshnevisan. *Lecture notes on Donsker's Theorem.* University of Utah, `http://www.math.utah.edu/~davar/ps-pdf-files/donsker.pdf`, accessed on August 13, 2018.

[38] A.N. Kolmogorov. *New Metric Invariant of Transitive Dynamical Systems and Endomorphisms of Lebesgue Spaces.* Doklady of Russian Academy of Sciences, (1958), 119, N5, 861-864.

[39] C. Kopf. *Invariant measures for piecewise linear transformations of the interval.* Applied Mathematics and Computation, 1990.

[40] Z.S. Kowalski. *Ergodic properties of skew products with Lasota–Yorke type maps in the base.* Stud. Math.106 45–57,1993.

[41] R.O. Kuzmin. *A problem of Gauss.* Dokl. Akad. Nauk, Ser. A (1928), 375-380.

[42] R.O. Kuzmin. *Atti de Congresso Internazionale de Mathematici.* vol. VI, Bologna (1928).

[43] A. Lasota and J.A. Yorke. *On the existence of invariant measures for piecewise monotonic transformations.* Trans. Amer. Math. Soc. 186 (1973), 481-488.

[44] T. Li and J. Yorke. *Ergodic transformations from an interval into itself.* Trans. Amer. Math. Soc. 235 (1978), 183-192.

[45] C. Liverani, B. Saussol, and S. Vaienti. *A probabilistic approach to intermittency.* Ergodic Theory Dynam. Systems 19 (1999), 671-685.

[46] G. Lochs. *Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch.* Abh. Math. Sem. Hamburg, 27 (1964), 142-144. MR 29:57.

[47] T. Morita. *Asymptotic behavior of one-dimensional random dynamical systems.* J. Math. Soc. Japan 37 651–63.

[48] T. Morita. *Asymptotic behavior of one-dimensional random dynamical systems - a rivisit.* Kyoto University Research Information Repository (2015), 1942: 31-43.

[49] T. Morita. *Random iteration of one-dimensional transformation.* Osaka J. Math. 22 489–518.

[50] R. Murray. *Approximation of invariant measures for a class of maps with indifferent fixed points.* University of Waikato, Mathematics Research Report Series II No. 106 (2005). Hamilton, New Zealand: University of Waikato.

[51] J. Nachbar. *The Arzelà-Ascoli Theorem.* Washington University, `https://pages.wustl.edu/files/pages/imce/nachbar/arzela-ascoli.pdf`, accessed on August 13, 2018.

[52] I.P. Natanson. *Theory of Functions of a Real Variable.* Frederick Ungar Publ. Co., 1955.

[53] S. Pelikan. *Invariant densities for random maps of the interval.* Trans. Am. Math. Soc. 281 813–25,1984.

[54] K. Petersen. *Ergodic Theory.* Cambridge Studies in Advanced Mathematics, 2. Cambridge University Press, Cambridge, 1989.

[55] G. Pianigiani. *First return map and invariant measures.* Israel J., Math. 35 (1980), 32-48.

[56] A. Rényi. *On algorithms for the generation of real numbers.* Magyar Tud. Akad. Mat. Fiz. Oszt.Közl., 7:265–293, 1957.

[57] A. Rényi. *Representation for real numbers and their ergodic properties.* Acta Math. Acad. Sci. Hungar. 8 (1957), 477-493.

[58] V.A. Rokhlin. *Exact endomorphisms of Lebesgue spaces. Amer. Math. Soc. Transl. 2 (1964), 1-36.*

[59] V.A. Rokhlin. *On the fundamental ideas of measure theory.* Transl. Amer. Math. Soc., Series 1, 10 (1962), 1–52.

[60] M.R. Rychlik. *Bounded variation and invariant measures.* Studia Mathematica 76 (1983), 69-80.

[61] C.E. Shannon. *Mathematical Theory of Communication.* Bell Systems Technical Journal. 27., July and October, pp. 379-423 and 623-656.

[62] N. Sidorov. *Almost every number has a continuum of β-expansion.* Amer. Math. Monthly, 110 (2003), no. 9, 838-842.

[63] Ya.G. Sinai. On the notion of entropy of a dynamical system. Doklady of Russian Academy of Sciences, (1959), 124, 768-771.

[64] P. Sun. *Entropy and invariant measures for skew product maps.* PhD thesis, The Pennsylvania State University, 2010.

[65] M. Viana and K. Oliveira. *Foundations of Ergodic Theory.* Volume 151 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2016.

[66] C. Walkden. Lecture 7: Entropy. The University of Manchester, `http://www.maths.manchester.ac.uk/~cwalkden/magic/lecture07.pdf`, accessed on August 13, 2018.

[67] R. Ward. *On robustness properties of beta encoders and golden ration encoders.* IEEE Transactions on Information Theory 54.9 (2008): 4323-4334.