Intermittency and number expansions for random interval maps

Zeegers, B.P.

Intermittency and Number Expansions for Random Interval Maps

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 14 februari 2023
klokke 15:00 uur

door

Benthen Pascal Zeegers
geboren te Alkmaar
in 1993

Promotores:

Prof. dr. W.Th.F. den Hollander	
Prof. dr. E.A. Verbitskiy	(Universiteit Leiden, Rijksuniversiteit Groningen)

Co-promotor:

Dr. C.C.C.J. Kalle

Promotiecommissie:

Prof. dr. H.J. Hupkes	
Prof. dr. F.M. Spieksma	
Prof. dr. W. Bahsoun	(Loughborough University)
Dr. C. González-Tokman	(University of Queensland)
Prof. dr. H. Peters	(Universiteit van Amsterdam)

Contents

CHAPTER 1

Introduction

§1.1 Motivation and context

In science, studies on systems that evolve in time are ubiquitous. In most of these studies mathematical models are essential to analyse and predict the behaviour of these systems. A large part of these models falls in the category of *discrete-time dynamical systems*. A state of the system under consideration is then represented as a point in some abstract space and the evolution of the system is described as moving from one point in this space to another in discrete time steps. This evolution can be modelled to be either deterministic or random.

More formally, a *deterministic* discrete-time dynamical system models the evolution by a single transformation T acting on a state space X, so that if the system starts at state $x_0 \in X$, then $x_1 = T(x_0)$ is the state of the system at time 1, and more generally,

$$x_n = T(x_{n-1}), \qquad n \in \mathbb{N} \tag{1.1}$$

is the state of the system at time n. Examples of deterministic discrete-time dynamical systems are *interval maps*, in which case the state space X is a bounded interval in $\mathbb{R}$. See Figure 1.1(a) for an example.

On the other hand, for a *random* discrete-time dynamical system a set $\mathcal{T} = \{T_i : X \to X\}_{i \in I}$ of transformations on X is considered and the evolution of the system starting at $x_0 \in X$ is given by

$$x_n = T_{i_n}(x_{n-1}), \qquad n \in \mathbb{N}, \tag{1.2}$$

where the sequence $\{i_n\}_{n \in \mathbb{N}}$ is drawn from $I^{\mathbb{N}}$ according to some probability law $\mathbb{P}$. If X is an interval, we then refer to the pair $(\mathcal{T}, \mathbb{P})$ as a *random interval map*. See Figure 1.1(b) for an example of a random interval map.

The past decades have seen an increasing interest in the mathematical properties of random interval maps. These systems have applications in electrical engineering, pseudo-random number generators, etc., and provide realistic models in many fields such as physics and population dynamics for studying real-world phenomena that evolve in time and involve noise. This dissertation consists of two parts, each of which considers a different research area related to random interval maps. We go into this in more detail in the next two subsections.

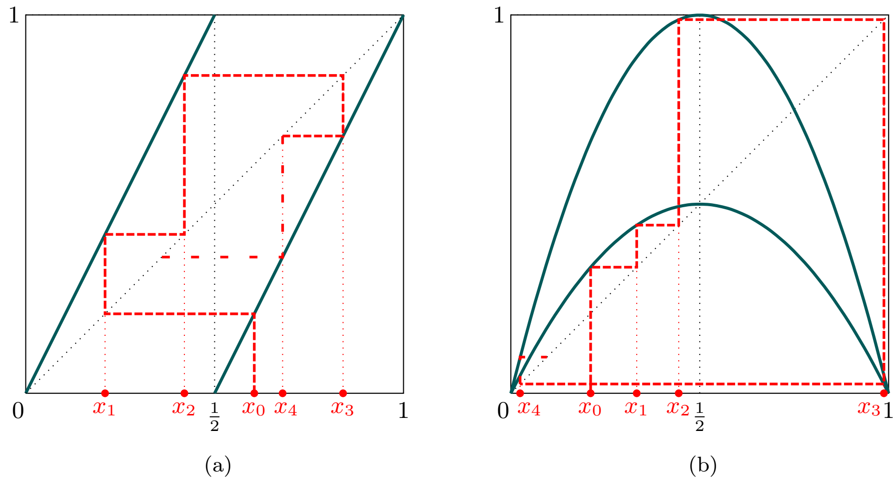


Figure 1.1: Example of (a) *an interval map, namely the doubling map* $x \mapsto 2x \bmod 1$, *and* (b) *a random interval map consisting of the two logistic maps* $x \mapsto 2x(1-x)$ *and* $x \mapsto 4x(1-x)$. *The dashed lines indicate parts of an orbit, which in* (a) *is of the form as in* (1.1) *and in* (b) *of the form as in* (1.2).

§1.1.1 Statistical properties of critically intermittent systems

Typically dynamical systems under consideration exhibit to a certain extent irregular or chaotic activity. See Figure 1.2. The long-term behaviour of these systems cannot be predicted by following single orbits, even if they are generated by a single deterministic transformation. On the other hand, systems with chaotic behaviour are usually statistically predictable and their long-term behaviour can be analyzed by using tools from Ergodic Theory. Of particular interest are absolutely continuous invariant measures, or *acim*'s for short, and mixing properties.

First of all, an *invariant measure* gives information about the long-term distribution of the orbits from (1.1) or (1.2) on part of the space, and this part is large if the invariant measure is *absolutely continuous* with respect to the Lebesgue measure. Furthermore, if this measure is finite, then on this part of the space mixing properties can be investigated. A dynamical system is *mixing* if any two observables on the system become uncorrelated from their initial values in the long run. In applications, observables are those quantities that can be measured from the states $\{x_n\}_{n\in\mathbb{N}}$ of the dynamical system given by (1.1) or (1.2), whereas the states themselves are usually not detectable. The speed of *decay of correlations* between any two observables gives an idea of the level of chaos within the system and can help deriving statistical limit laws for the system. Formal definitions of all these concepts are given in Section 1.2.

In the first part of this dissertation we are interested in random dynamical systems that are *intermittent*. This is a type of behaviour where the system alternates between

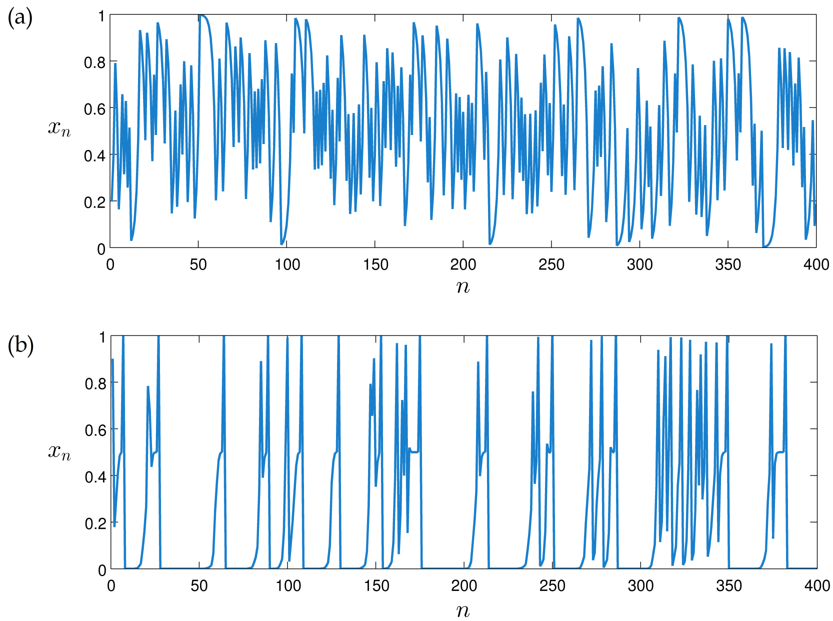


Figure 1.2: Time series of the systems from Figure 1.1: In (a) *part of an orbit under the doubling map*[1] *is shown, whereas* (b) *depicts a typical trajectory under the random map from Figure 1.1(b) if the two logistic maps are chosen with equal probability and independently at each time step.*

periods of either irregular activity or being in a seemingly steady state. An example is shown in Figure 1.2(b) where the system spends long periods near zero. This is in contrast to the system depicted in Figure 1.2(a) which only exhibits chaotic behaviour. In case an intermittent system admits an acim that is infinite, then this usually means that the steady state is more prevalent than the irregular activity and orbits tend to remain long near invariant sets of the system. If a finite acim exists and the intermittent system is mixing, then typically correlations decay subexponentially fast, which indicates a weak level of chaos.

The study on intermittent dynamical systems goes back to the seminal article [PM80] by Manneville and Pomeau, who used such systems to model intermittency in the context of transitions to turbulence in convective fluids, see also [MP80, BPV86], and distinguished several different types of intermittency. We are interested in so-called *critical intermittency* introduced recently in [AGH18, HPR21]. This is a kind of intermittent behaviour caused by an interplay of a superattracting fixed point and a repelling fixed point. To illustrate the concept, consider the random interval map consisting of the two logistic maps $L_2(x) = 2x(1-x)$ and $L_4(x) = 4x(1-x)$: for each

[1] Real numbers in $[0, 1)$ randomly generated by Matlab have 53 bits and thus after at most 53 applications of the doubling map end up in zero. To overcome this, we have used in Matlab a version of the doubling map rescaled to $[0, \pi]$ and then scaled the resulting plot to $[0, 1]$.

n, independently, let

$$x_n = \begin{cases} L_2(x_{n-1}), & \text{with prob. } p_2, \\ L_4(x_{n-1}), & \text{with prob. } p_4 = 1 - p_2. \end{cases}$$

The dynamics of these two maps individually are quite different: L_4 exhibits chaotic behaviour, admits a finite acim and correlations decay exponentially fast, while L_2 has $\frac{1}{2}$ as a superattracting fixed point with $(0,1)$ as its basin of attraction. Under random compositions of L_2 and L_4 the typical behaviour is the following: orbits are quickly attracted to $\frac{1}{2}$ by applications of L_2 and are then repelled first close to 1 and then close to 0 by one application of L_4 followed by an application of either L_2 or L_4. Since 0 is a repelling fixed point for both maps, orbits then leave a neighbourhood of 0 after a number of time steps, see Figure 1.1(b). This pattern occurs infinitely often in typical random orbits and is the result of the interplay between the exponential divergence from 0 under L_2 and L_4 and the superexponential convergence to $\frac{1}{2}$ under L_2. Figure 1.2(b) shows an orbit under random compositions of L_2 and L_4.

The dynamical behaviour of random compositions of the two logistic maps L_2 and L_4 was studied in [AD00, AS03, AGH18, C02, HPR21] among others. In [AGH18, HPR21] the authors investigated the existence and finiteness of an acim for this random system and for random systems consisting of rational maps on the Riemann sphere. One particular result from [AGH18] states that the random dynamical system generated by i.i.d. compositions of L_2 and L_4 chosen with probabilities p_2 and $p_4 = 1 - p_2$ admits a σ-finite acim that is infinite in case $p_2 > \frac{1}{2}$. An interesting question that was left open in [AGH18] is whether for $p_2 \leq \frac{1}{2}$ this measure is infinite or finite.

One of the first results in this dissertation answers this question. For a large family of random interval maps with critical intermittency that includes the random combination of L_2 and L_4 we show the existence of a unique acim. We give an explicit expression of a phase transition threshold in terms of the probabilities of choosing the maps as well as their strengths of convergence to the superstable fixed point, a threshold that separates the cases where the acim is either finite or infinite. In particular, the acim of the random interval map composed of L_2 and L_4 is finite if and only if the probability p_2 of choosing L_2 satisfies $p_2 < \frac{1}{2}$. Moreover, for a closely related class of critically intermittent random interval maps we show that in case the acim is finite correlations decay polynomially fast and we give bounds on the rate of this decay. Finally, we show that a similar phase transition holds for a similar class of intermittent random interval maps but where the superattracting fixed point is replaced by an attracting fixed point and the repelling fixed point is replaced by a weakly repelling fixed point. Among the techniques we will use are Perron-Frobenius operators, induced transformations and Young towers.

Most studies on random interval maps consider systems where the dynamics are dominated by only one type of behaviour, e.g. if the system is uniformly expanding as in [ANV15, M85b, P84], or if only one of the constituent maps governs the dynamics as is the case in [BBD14, BB16, NTV18]. The random interval maps we study go beyond this familiar setting and are composed of two types of maps that are very different from each other, one being chaotic and the other not. We prove statistical properties that depend on the features of both types of maps as well as on the probabilities of choosing the maps, a dependence that is new compared to other known results on random interval maps.

§1.1.2 Extensions of Lochs' Theorem to random systems

Besides modelling dynamical phenomena we can also use random interval maps to generate *number expansions*, which will be the main object of study in the second part of this dissertation. A number expansion of a real number is a representation of this number with a specific set of symbols or digits. For example, for each $x \in [0,1]$ there exists a sequence $(b_n)_{n\geq 1}$ in $\{0,1\}^{\mathbb{N}}$ such that

$$x = \sum_{n=1}^{\infty} \frac{b_n}{2^n},$$

a number expansion referred to as a *binary* (or *base 2*) expansion of x. Other classical examples of number expansions include decimal expansions, β-expansions and continued fraction expansions. These examples have in common that they can be generated by an interval map. For instance, let $T_2 : [0,1] \to [0,1]$ be the *doubling map* given by

$$T_2(x) = 2x \bmod 1 = \begin{cases} 2x, & \text{if } 0 \leq x < \frac{1}{2}, \\ 2x - 1, & \text{if } \frac{1}{2} \leq x \leq 1, \end{cases} \tag{1.3}$$

see Figure 1.1(a). Then the $n+1$-th symbol in the binary expansion of a point $x_0 \in [0,1]$ is zero if x_n as given in (1.1) is smaller than a half, and one otherwise. For example, the point x_0 in Figure 1.1(a) has binary expansion $(b_n)_{n\geq 1}$ with $b_1 = 1$, $b_2 = b_3 = 0$, etc. In a similar way we can use suitable random interval maps to generate number expansions by assigning symbols to subintervals and following orbits of the form in (1.2).

In [L64] Lochs compared the efficiency between representing real numbers in decimal expansions and regular continued fraction expansions. It is known that each irrational $x \in [0,1]$ has a unique *decimal expansion*

$$x = \sum_{n=1}^{\infty} \frac{d_n}{10^n},$$

where $(d_n)_{n\geq 1}$ is a sequence in $\{0,1,\ldots,9\}^{\mathbb{N}}$, and moreover that there exists a unique sequence $(a_n)_{n\geq 1}$ in $\mathbb{N}^{\mathbb{N}}$ such that

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \ddots}}},$$

which is referred to as the *regular continued fraction* expansion of x. In [L64] the following question is investigated: Suppose we know the first n decimal digits $d_1, \ldots, d_n$ of a further unknown irrational number $x \in [0,1]$. What is the largest number $m(n,x)$ of digits $a_1, \ldots, a_{m(n,x)}$ in the regular continued fraction expansion of x that can be determined from this information? In 1964, Lochs [L64] provided an answer for the limit $n \to \infty$ by proving that for Lebesgue almost every irrational $x \in [0,1]$

$$\lim_{n\to\infty} \frac{m(n,x)}{n} = \frac{6 \log 2 \log 10}{\pi^2} = 0.97027\cdots. \tag{1.4}$$

In other words, the first 100 decimal digits determine roughly 97 digits in the regular continued fraction expansion, which indicates that typically regular continued fraction expansions are slightly more efficient to represent real numbers than decimal expansions.

A natural question is whether something similar to Lochs' result can be said for other pairs of expansions as well. Interestingly, the right-hand side of (1.4) is the ratio of the *entropies* of the interval maps that generate these expansions. In [DF01] Lochs' result has been generalised to a large class of pairs of interval maps that generate number expansions. For such a pair the analogue of the right-hand side of (1.4) is then the ratio of the entropies of the interval maps with respect to their unique probability acim.

In this dissertation we further generalise this result and extend Lochs' result given in (1.4) to a wide class of pairs of random interval maps that produce number expansions. For this we generalise the method from [DF01] and apply existing theory on *fiber entropy* from [B93] as well as derive new theory on this topic. The random analog of the right-hand side of (1.4) is then a fraction of fiber entropies. Furthermore, under additional assumptions we also provide a corresponding Central Limit Theorem by using a method that is similar to the one from [H09] applied to the deterministic setting from [DF01].

We also study a question posed in [JMKA13] in the context of *β-encoders* that is closely related to Lochs' result. A β-encoder is an analog circuit that converts analog input signals into bitstreams and was first introduced in [DDGV02]. In [JMKA13] Jitsumatsu and Matsumura provide an algorithm that reads the output digits of a β-encoder with input $x \in [0,1)$ and converts them into digits that correspond with the binary expansion of x. The distribution of these base 2 digits are shown to be in some sense close to that of i.i.d. random variables, thus making it a suitable pseudo-random number generator. However, in order to produce m base 2 digits in this way, Jitsumatsu and Matsumura posed the question of what is the minimum required number $k(m)$ of output digits of the β-encoder. To approach this problem we provide multiple limit results as $m \to \infty$ for β-encoders. These results give an indication on the efficiency of the β-encoder being used in [JMKA13] as a potential source for pseudo-random number generation but also show that in the presence of amplification or scaling errors the proposed method in [JMKA13] is not optimal.

In the following sections we explain the aforementioned notions in more detail and provide some of the mathematical tools that are needed in the rest of this dissertation.

§1.2 Ergodic Theory

This section briefly covers some relevant concepts of Ergodic Theory. We refer the reader to [P89, W00, DK21] for a more extensive and complete introduction to Ergodic Theory. In a nutshell, Ergodic Theory is the study on the long-term average behaviour of systems over time. The state space of the system under consideration is assumed to be a measure space $(X, \mathcal{F}, m)$ with X a set, $\mathcal{F}$ a σ-algebra on X and m a measure on $(X, \mathcal{F})$. The evolution is given by a measurable transformation $T : X \to X$. We refer to the quadruple $(X, \mathcal{F}, m, T)$ as a *dynamical system.* Usually we assume T to satisfy the following property with respect to the reference measure m.

Definition 1.2.1 (Non-singularity). A measurable transformation T on a measure space $(X, \mathcal{F}, m)$ is said to be *non-singular* if for any $A \in \mathcal{F}$ we have $m(A) = 0$ if and only if $m(T^{-1}A) = 0$.

Suppose that there is some $A \in \mathcal{F}$ such that $T^{-1}A = A$. Then $T^{-1}(X\backslash A) = X\backslash A$, so in this case T can be decomposed into two transformations $T|_A : A \to A$ and $T|_{X\backslash A} : X\backslash A \to X\backslash A$. For this reason it is natural to study transformations that are indecomposable up to sets of measure zero.

Definition 1.2.2 (Ergodicity). A measurable transformation T on a measure space $(X, \mathcal{F}, m)$ is said to be *ergodic* if for any $A \in \mathcal{F}$ such that $T^{-1}A = A$ we have either have $m(A) = 0$ or $m(X\backslash A) = 0$.

We are interested in invariant measures for T, which is the topic of the next subsection.

§1.2.1 Invariant measures

Invariant measures can give an idea on the distribution of points in the orbits under T on part of the space. They are defined as follows:

Definition 1.2.3 (Invariant measures). Let $(X, \mathcal{F})$ be a measurable space and $T : X \to X$ be measurable. Then a measure μ on $(X, \mathcal{F})$ is called *invariant* with respect to T if

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

holds for all $A \in \mathcal{F}$, or equivalenty if

$$\int_X f \circ T d\mu = \int_X f d\mu$$

holds for all $f \in L^1(X, \mu)$. In this case we also say that T is *measure preserving* with respect to μ.

Among the invariant measures for T we focus usually on those that are absolutely continuous with respect to some reference measure m. We recall the definition:

Definition 1.2.4 (Absolute continuity and equivalence of measures). Let μ and m be two measures on a measurable space $(X, \mathcal{F})$. Then μ is *absolutely continuous* with respect to m if for each $A \in \mathcal{F}$ we have that $m(A) = 0$ implies $\mu(A) = 0$. In this case we use the notation $\mu \ll m$. Furthermore, we say that μ and m are *equivalent* if $\mu \ll m$ and $m \ll \mu$.

As is well known, it follows from the Radon-Nikodym Theorem that if $\mu \ll m$ and μ and m are σ-finite measures, then there exists an m-a.e. unique measurable function $\frac{d\mu}{dm} : X \to [0, \infty]$ called the *density* for which

$$\mu(A) = \int_A \frac{d\mu}{dm} dm$$

holds for all $A \in \mathcal{F}$. If μ is also invariant with respect to T, then we usually refer to $\frac{d\mu}{dm}$ as an *invariant density* for T and call μ an *absolutely continuous invariant measure*, or *acim* for short, for T. Furthermore, when we say that T has a unique acim μ, then we mean unique up to scalar multiplication, because in this case $c \cdot \mu$ is also an acim for each constant $c > 0$.

If μ is an invariant measure for T that is *finite*, i.e. $\mu(X) < \infty$, then Birkhoff's Ergodic Theorem gives the following characterization of the long-term average behaviour of orbits that are typical with respect to μ.

Theorem 1.2.5 (Birkhoff's Ergodic Theorem). *Let T be a measure preserving and ergodic transformation on a measure space $(X, \mathcal{F}, \mu)$ with finite measure μ. Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \frac{1}{\mu(X)} \int_X f d\mu \qquad \text{for } \mu\text{-a.e. } x \in X \tag{1.5}$$

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

The statement in (1.5) might hold only for a small or negligible part of the space, for instance when μ is a linear combination of Dirac measures. On the other hand, if μ is absolutely continuous with respect to a reference measure m for which sets of positive measures are considered large, e.g. the Lebesgue measure, then (1.5) holds for a non-trivial part of the space. This illustrates the significance of acim's.

In general a transformation can admit multiple acim's. The next result, which can be derived from Birkhoff's Ergodic Theorem as shown in [DK21, Theorem 3.1.2], gives conditions under which a transformation admits precisely one (up to scalar-multiplication) finite acim.

Theorem 1.2.6. *Let T be an ergodic transformation on a measure space $(X, \mathcal{F}, m)$. If μ is a finite acim of T that is equivalent to m, then μ is the only (up to scalar multiplication) finite acim of T.*

This result can be strengthened under additional conditions on the dynamical system $(X, \mathcal{F}, m, T)$, which is done below in Theorem 1.2.10. One of these conditions is that T is *conservative*, which means that almost every point in a set of positive measure will return to this set under iterations of T:

Definition 1.2.7 (Conservativity). A measurable transformation T on a measure space $(X, \mathcal{F}, m)$ is said to be *conservative* if $m(A \setminus \bigcup_{n\geq 1} T^{-n}A) = 0$ holds for each $A \in \mathcal{F}$ such that $m(A) > 0$.

The next result, Maharam's Recurrence Theorem, gives a sufficient condition for a transformation to be conservative with respect to an invariant measure. For this, we need the following definition.

Definition 1.2.8 (Sweep-out set). Let $(X, \mathcal{F}, m, T)$ be a dynamical system. A set $A \in \mathcal{F}$ is called a *sweep-out set* for T if $0 < m(A) < \infty$ and $m(X \setminus \bigcup_{n\geq 0} T^{-n}A) = 0$.

Theorem 1.2.9 (Maharam's Recurrence Theorem). *Let T be a measure preserving transformation on a measure space $(X, \mathcal{F}, \mu)$. If there exists a sweep-out set for T, then T is conservative.*

We remark that a transformation T that preserves a measure μ is *always* conservative with respect to μ if μ is finite. This is the content of the famous Poincaré Recurrence Theorem, which is one of the many results in Ergodic Theory where the dependence on the finiteness of the invariant measure is crucial. This dependence is also illustrated by Birkhoff's Ergodic Theorem, which does not apply if $\mu(X) = \infty$.

The following theorem can be found in e.g. [A97, Theorem 1.5.6].

Theorem 1.2.10. *Let T be a conservative, ergodic, non-singular transformation on a σ-finite measure space $(X, \mathcal{F}, m)$. Then T admits at most one (up to scalar multiplication) σ-finite acim.*

A commonly used technique to obtain an acim is by inducing the transformation on a suitable subset of the space. More precisely, let $(X, \mathcal{F}, m)$ be a measure space and $T : X \to X$ non-singular with respect to m. For a set $Y \in \mathcal{F}$ such that $0 < m(Y) < \infty$ and $m\big(Y \setminus \bigcup_{n\geq 1} T^{-n}Y\big) = 0$, the *first return time map* $\varphi_Y : Y \to \mathbb{N} \cup \{\infty\}$ given by

$$\varphi_Y(y) = \inf\{n \geq 1 \,:\, T^n(y) \in Y\} \tag{1.6}$$

is finite m-a.e. on Y, and moreover m-a.e. $y \in Y$ returns to Y infinitely often. If we remove from Y the m-null set of points that return to Y only finitely many times, and for convenience call this set Y again, then we can define the *induced transformation* $T_Y : Y \to Y$ by

$$T_Y(y) = T^{\varphi_Y(y)}(y).$$

The idea of inducing is to take a subset Y such that T_Y is easier to analyse than T and deduce properties on T via T_Y. This is illustrated by the following two results. The first one can be found in e.g. [A97, Proposition 1.5.2]. Note that this statement requires $m\big(X \setminus \bigcup_{n\geq 1} T^{-n}Y\big) = 0$, which immediately follows if T is non-singular and Y is a sweep-out set for T as is the case below.

Proposition 1.2.11. *Let T be a non-singular and conservative transformation on a measure space $(X, \mathcal{F}, m)$ and let $Y \in \mathcal{F}$ be such that $0 < m(Y) < \infty$. If Y is a sweep-out set for T and T_Y is ergodic with respect to $m|_Y$, then T is ergodic with respect to m.*

The next result can be found in e.g. [A97, Proposition 1.5.7]. Note that this statement asks for T to be conservative w.r.t. m. This is not used in the proof however and the condition $m\big(Y \setminus \bigcup_{n\geq 1} T^{-n}Y\big) = 0$ is enough to guarantee that the induced transformation is well defined.

Proposition 1.2.12. *Let T be a non-singular transformation on a measure space $(X, \mathcal{F}, m)$ and let $Y \in \mathcal{F}$ be such that $0 < m(Y) < \infty$ and $m\big(Y \setminus \bigcup_{n\geq 1} T^{-n}Y\big) = 0$. If $\nu \ll m|_Y$ is a finite invariant measure for the induced transformation T_Y, then the measure μ on $(X, \mathcal{F})$ defined by*

$$\mu(A) = \sum_{k\geq 0} \nu\Big(Y \cap T^{-k}A \setminus \bigcup_{j=1}^{k} T^{-j}Y\Big) \tag{1.7}$$

for $A \in \mathcal{F}$ is a σ-finite acim for T and $\mu|_Y = \nu$. Moreover, μ is finite if and only if $\int_Y \varphi_Y d\nu < \infty$.

The last statement of the previous result follows by taking $A = X$ in (1.7), which gives $\mu(X) = \int_Y \varphi_Y d\nu$. This formula holds true more generally:

Lemma 1.2.13 (Kac's Lemma). *Let T be a conservative, measure preserving and ergodic transformation on a measure space $(X, \mathcal{F}, \mu)$. Let $Y \in \mathcal{F}$ be such that $0 < \mu(Y) < \infty$. Then $\int_Y \varphi_Y d\mu = \mu(X)$.*

Kac's Lemma can serve as a powerful tool to show that a certain invariant measure μ is infinite, where the challenge is to find a suitable $Y \in \mathcal{F}$ such that $\int_Y \varphi_Y d\mu = \infty$. We will exploit this technique several times in this dissertation.

In Sections 1.3 and 1.4 we will discuss more techniques to find (finite) acim's in case the dynamical system under consideration is a (random) interval map.

§1.2.2 Mixing and statistical properties

Let T be a measure preserving transformation on a probability space $(X, \mathcal{F}, \mu)$. If T is ergodic, then it can be shown from Birkhoff's Ergodic Theorem that for each $f \in L^\infty(X, \mu)$ and $g \in L^1(X, \mu)$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} \int_X f \circ T^k \cdot g d\mu = \int_X f d\mu \cdot \int_X g d\mu.$$

In other words, *observables* (also called *test functions*) get uncorrelated on average over time. The following definition strengthens this property by getting rid of the averaging over time.

Definition 1.2.14 (Mixing). A measure preserving transformation T on a probability space $(X, \mathcal{F}, \mu)$ is said to be *mixing* if

$$\lim_{n\to\infty} \int_X f \circ T^n \cdot g d\mu = \int_X f d\mu \cdot \int_X g d\mu.$$

for each $f \in L^\infty(X, \mu)$ and $g \in L^1(X, \mu)$.

We define for each $n \in \mathbb{N}$, $f \in L^\infty(X,\mu)$ and $g \in L^1(X,\mu)$ the *correlation function*

$$\mathrm{Cor}_{n,T,\mu}(f,g) = \int_X f \circ T^n \cdot g d\mu - \int_X f d\mu \cdot \int_X g d\mu,$$

for which we will also sometimes just write $\mathrm{Cor}_n(f,g)$ if the context allows for it. Then T being mixing gives $\lim_{n\to\infty} \mathrm{Cor}_n(f,g) = 0$. The speed at which $\mathrm{Cor}_n(f,g)$ converges to zero, i.e. the speed at which $f \circ T^n$ and g get uncorrelated, depends on the degree of chaos within the dynamical system as well as on the regularity of the observables. It is in general always possible to find two observables $f \in L^\infty(X,\mu)$ and $g \in L^1(X,\mu)$ for which this loss of memory happens arbitrarily slowly. However, it is sometimes possible to obtain for a suitable class of observables that are sufficiently regular a rate on the decay of correlations that is uniform with respect to the observables in this class, thus being a good measure for the level of chaos within the system. We will discuss examples of this in Sections 1.3 and 1.4 for (random) interval maps.

A number of these examples obtain results on the rate of decay of correlations by constructing a suitable *Young tower*, a technique introduced in [Y98, Y99] by Young. We will exploit this technique as well in Chapter 3. The construction of a Young tower is rather technical and we explain this in more detail in Chapter 3, but let us briefly indicate here one of its powerful consequences. As we have seen in Proposition 1.2.12, sometimes an acim μ for a dynamical system $(X,\mathcal{F},m,T)$ can be obtained by finding a finite acim ν for an induced transformation T_Y with $Y \in \mathcal{F}$, and μ is then finite if and only if φ_Y is integrable with respect to ν. For the latter it is usually sufficient to verify that φ_Y is integrable with respect to $m|_Y$. If a Young tower can be constructed on Y, then T is mixing with respect to μ and more can be said depending on the tail of the distribution of φ_Y with respect to $m|_Y$. More precisely, if the tail is polynomial or exponential, then the correlation function for Hölder continuous functions with respect to μ decays polynomially or exponentially fast, respectively.

Mixing and correlation decay rates are statistical properties of the system that give an idea on the chaotic behaviour of the system under consideration. Another (but closely related) indicator for chaos in a dynamical system $(X,\mathcal{F},\mu,T)$ is the behaviour of processes of the form $\{f \circ T^n\}_{n\in\mathbb{N}}$ where $f: X \to \mathbb{R}$ is a measurable function. If the system is sufficiently chaotic such processes exhibit stochastic behaviour and there are many settings known in which these processes satisfy classical results from Probability Theory. For instance, if μ is a probability measure and T is measure preserving with respect to μ, then the processes $\{f \circ T^n\}_{n\in\mathbb{N}}$ are stationary on $(X,\mathcal{F},\mu)$, i.e. the joint probability distribution of $\{f \circ T^n\}_{n\in\mathbb{N}}$ does not change when shifted in time. Moreover, Birkhoff's Ergodic Theorem can be rephrased to saying that these processes satisfy the Law of Large Numbers if T is also ergodic with respect to μ. For such processes the classical Central Limit Theorem reads as follows:

Definition 1.2.15 (Central Limit Theorem). Let T be a transformation on a probability space $(X,\mathcal{F},\mu)$ that is measure preserving. We say that $f \in L^1(X,\mu)$ satisfies the *Central Limit Theorem (CLT)* if there exists a $\sigma > 0$ such that

$$\lim_{n\to\infty} \mu\Big(\Big\{x \in X : \frac{\sum_{k=0}^{n-1}(f\circ T^k(x) - \int f d\mu)}{\sigma\sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^u e^{-t^2/2} dt. \qquad (1.8)$$

The results in [L96] provide a large class of ergodic measure preserving transformations T on a probability space $(X, \mathcal{F}, \mu)$ and functions $f \in L^\infty(X, \mu)$ that satisfy the CLT if the correlation function $\mathrm{Cor}_n(f, f)$ decays fast enough as $n \to \infty$, namely if

$$\sum_{n\in\mathbb{N}} |\mathrm{Cor}_n(f, f)| < \infty. \tag{1.9}$$

See [L96] and references therein for more results on the CLT for dynamical systems.

Other statistical properties of dynamical systems include exactness and Bernoullicity and other statistical limit laws like the almost sure invariance principle and large deviation theorems. In this dissertation we restrict our attention to mixing and the Central Limit Theorem. Moreover, we only discuss these properties for dynamical systems $(X, \mathcal{F}, \mu, T)$ where μ is a probability (or finite) measure. For infinite measure systems statistical properties are usually harder to derive since we cannot immediately apply the tools from Probability Theory.

§1.2.3 Entropy

Shannon introduced the concept of entropy in information theory as a measure for randomness generated by an information source, see [S48]. This notion was then introduced to dynamical systems by Kolmogorov [K58] and Sinai [S59]. In this subsection we give a brief introduction on this topic.

Let $(X, \mathcal{F}, \mu)$ be a probability space. We call a collection $\mathcal{P}$ a *partition* of X if it is an at most countable collection of measurable sets, $\mathcal{P} \subseteq \mathcal{F}$, that are pairwise disjoint and satisfy $X = \bigcup_{P\in\mathcal{P}} P$, where both properties are considered modulo μ-null sets. The *entropy of a partition* $\mathcal{P}$ is defined as

$$H_\mu(\mathcal{P}) = -\sum_{P\in\mathcal{P}} \mu(P) \log \mu(P).$$

For two partitions $\mathcal{P}_1$ and $\mathcal{P}_2$ of X we use the notation $\mathcal{P}_1 \le \mathcal{P}_2$ to indicate that $\mathcal{P}_2$ is a *refinement* of $\mathcal{P}_1$, i.e. that for every $P \in \mathcal{P}_2$ there is a $Q \in \mathcal{P}_1$ such that $P \subseteq Q$. Moreover, we use $\mathcal{P}_1 \vee \mathcal{P}_2 := \{P \cap Q \,:\, P \in \mathcal{P}_1,\, Q \in \mathcal{P}_2\}$ to denote the *common refinement* of $\mathcal{P}_1$ and $\mathcal{P}_2$.

Definition 1.2.16 (Measure theoretic entropy). Let $(X, \mathcal{F}, \mu, T)$ be a dynamical system where μ is a T-invariant probability measure. For a partition $\mathcal{P}$ of X with finite entropy, i.e. $H_\mu(\mathcal{P}) < \infty$, the *entropy of* T *with respect to* $\mathcal{P}$ is given by

$$h_\mu(T, \mathcal{P}) := \lim_{n\to\infty} \frac{1}{n} H_\mu\Big(\bigvee_{k=0}^{n-1} T^{-k}\mathcal{P}\Big).$$

Furthermore, the *measure theoretic entropy of* T is given by

$$h_\mu(T) := \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}),$$

where the supremum is taken over all partitions $\mathcal{P}$ with finite entropy.

It is not very practical to calculate the measure theoretic entropy straight from its definition. We present some results that facilitate the computation of entropy. For a collection of measurable sets $\mathcal{E} \subseteq \mathcal{F}$ we use $\sigma(\mathcal{E})$ to denote the smallest sub σ-algebra of $\mathcal{F}$ containing $\mathcal{E}$. We say that a partition $\mathcal{P}$ of X is a *generator* for a transformation $T : X \to X$ if the sequence of partitions $\{P_n\}$ given by $\mathcal{P}_n = \bigvee_{k=0}^{n-1} T^{-k}\mathcal{P}$ satisfies $\sigma(\bigcup_{n\in\mathbb{N}} \mathcal{P}_n) = \mathcal{F}$ up to sets of μ-measure zero. If X is a Polish space and $\mathcal{F}$ the associated Borel σ-algebra, then according to [M57, Theorem 3.3] a sufficient condition for a partition $\mathcal{P}$ of X to be a generator for T is if $\{\mathcal{P}_n\}$ *separates* points, i.e. for each $x, y \in X$ with $x \neq y$ there exist $n \in \mathbb{N}$ and $A \in \mathcal{P}_n$ such that $x \in A$ and $y \notin A$.

Theorem 1.2.17 (Kolmogorov-Sinai Theorem). *Let $(X, \mathcal{F}, \mu, T)$ be a dynamical system where μ is a T-invariant probability measure and let $\mathcal{P}$ be a partition of X with finite entropy. If $\mathcal{P}$ is a generator with respect to T, then $h_\mu(T) = h_\mu(T, \mathcal{P})$.*

In applications of the Kolmogorov-Sinai Theorem, $\mathcal{P}$ usually consists of *invertibility domains of T*, which are measurable sets on which T is bijective to its image with measurable inverse. More generally, the following result, which can be found in e.g. [DK21, Proposition 9.3.1], holds.

Lemma 1.2.18. *Let $(X, \mathcal{F}, \mu, T)$ be a dynamical system where μ is a T-invariant probability measure. If $\mathcal{P}_1 \leq \mathcal{P}_2 \leq \dots$ is an increasing sequence of partitions of X with finite entropy and $\sigma(\bigcup_{n\in\mathbb{N}} \mathcal{P}_n) = \mathcal{F}$ up to sets of μ-measure zero, then $h_\mu(T) = \lim_{n\to\infty} h_\mu(T, \mathcal{P}_n)$.*

Another tool that can be useful for calculating the entropy is a result known as Rokhlin's Formula presented below, which relates entropy to the Jacobian of a transformation. The next result gives conditions under which the Jacobian function exists.

Proposition 1.2.19 (Jacobian). *Let $(X, \mathcal{F}, m, T)$ be a dynamical system where X is a Polish space, $\mathcal{F}$ the associated Borel σ-algebra, m a probability measure and T non-singular. Suppose that $\mathcal{P}$ is a partition of X consisting of invertibility domains of T. Then there exists a m-a.e. unique non-negative function $J_mT \in L^1(X, m)$, called the* Jacobian of T with respect to m, *such that*

$$m(T(B)) = \int_B J_mT dm$$

holds for each measurable $B \subseteq A$ and $A \in \mathcal{P}$.

In later chapters we will make use of the following change of variables formulae from [VO16, Lemma 9.7.4].

Lemma 1.2.20 (Change of variables formulae). *Under the assumptions of Proposition 1.2.19, for each measurable $B \subseteq A$ and $A \in \mathcal{P}$,*

(a) *$\int_{T(B)} \varphi dm = \int_B (\varphi \circ T) J_mT dm$ for any measurable function $\varphi : T(B) \to \mathbb{R}$ such that the integrals are defined (possibly $\pm\infty$),*

(b) $\int_B \psi dm = \int_{T(B)} (\psi/J_m T) \circ (T|_B)^{-1} dm$ *for any measurable function* $\psi : B \to \mathbb{R}$ *such that the integrals are defined (possibly* $\pm\infty$*).*

The following result goes back to Rokhlin [R52], see e.g. [VO16, Theorem 9.7.3].

Theorem 1.2.21 (Rokhlin Formula). *Let* $(X, \mathcal{F}, \mu, T)$ *be a dynamical system with* X *a Polish space,* $\mathcal{F}$ *the associated Borel* σ*-algebra,* μ *a probability measure and* T *measure preserving with respect to* μ*. Suppose that* $\mathcal{P}$ *is a partition of* X *with finite entropy consisting of invertibility domains of* T*, and that* $\mathcal{P}$ *is a generator for* T*. Then*

$$h_\mu(T) = \int_X \log J_\mu T d\mu.$$

Finally, we state the Shannon-McMillan-Breiman Theorem. For a partition $\mathcal{P}$ of X and $x \in X$ we denote by $\mathcal{P}(x)$ the partition element of $\mathcal{P}$ containing x.

Theorem 1.2.22 (Shannon-McMillan-Breiman Theorem). *Let* T *be a measure preserving and ergodic transformation on a probability space* $(X, \mathcal{F}, \mu)$ *and let* $\mathcal{P}$ *be a partition of* X *with finite entropy. Then*

$$\lim_{n\to\infty} -\frac{\log \mu\big(\bigvee_{k=0}^{n-1} T^{-k}\mathcal{P}(x)\big)}{n} = h_\mu(T, \mathcal{P}) \qquad \text{for } \mu\text{-a.e. } x \in X.$$

§1.3 Interval maps

An *interval map* is a dynamical system of the form $(X, \mathcal{B}, \lambda, T)$ where X is a bounded interval in $\mathbb{R}$, $\mathcal{B}$ the associated Borel σ-algebra and λ the Lebesgue measure on $(X, \mathcal{B})$. Without loss of generality we take in this section $X = [0, 1]$. There are numerous techniques showing the existence of acims for interval maps, but precise formulae for the densities are in general not known. In this section we review some of these techniques for the types of interval maps that will be relevant in this dissertation as well as some of their statistical properties. We refer to [L06] for a more complete overview.

In general, the question whether an interval map admits an acim is linked to the degree to which distances between points close to each other get expanded under iterations of the interval map. Usually an interval map $T : [0, 1] \to [0, 1]$ admits a finite acim with exponential decay of correlations if the map is sufficiently smooth and *uniformly expanding*, meaning that there exist constants $C > 0$ and $q > 1$ such that $|DT^n(x)| > Cq^n$ holds for all $n \in \mathbb{N}$ and all $x \in [0, 1]$. A sufficient condition to be uniformly expanding is if $\inf_{x\in[0,1]} |DT(x)| > 1$ holds and this is usually what is assumed when considering expanding interval maps.

The first important results on acims for uniformly expanding interval maps have been derived for so-called *Markov interval maps*, see e.g. [dMvS93, Section 5.2]. The main feature of a Markov interval map $T : [0, 1] \to [0, 1]$ is that T is piecewise monotonic and sufficiently smooth on a finite or countable partition $\{I_i\}$ such that for each I_j there exists a collection $\mathcal{A}_j \subseteq \{I_i\}$ such that $T(I_j) = \bigcup_{I \in \mathcal{A}_j} I$. The dynamics under T can then be modelled by a Markov chain where a state i can move to state j if and only if $I_j \subseteq T(I_i)$. A finite acim for T can then be obtained by e.g. applying the thermodynamic formalism to this Markov chain, see e.g. [S09]. Moreover, if the Markov chain is irreducible and aperiodic this is the only acim for T and it has very strong mixing properties under T, including that the correlation function for Hölder continuous observables decays exponentially fast. In particular, these observables satisfy the summability condition from (1.9) and therefore typically satisfy the Central Limit Theorem from (1.8). For more details see e.g. [dMvS93, Y98].

Example 1.3.1 (N-adic transformations). The simplest class of examples of a Markov interval map are the *N-adic transformations* $T_N : [0,1] \to [0,1]$ with integer $N \geq 2$ and

$$T_N(x) = Nx \bmod 1.$$

The doubling map from (1.3) is the N-adic transformation where $N = 2$, see Figure 1.1(a). It can be shown that, for each integer $N \geq 2$, T_N preserves the Lebesgue measure λ on $[0,1]$ and that λ is the only acim of T_N.

Example 1.3.2 (Gauss map). Another well-studied Markov interval map is the *Gauss map* $G : [0,1] \to [0,1]$ given by $G(0) = 0$ and for $x \neq 0$,

$$G(x) = \frac{1}{x} \bmod 1,$$

see Figure 1.3(a). It is well known that G preserves the *Gauss probability measure* μ_G on $[0,1]$ with density

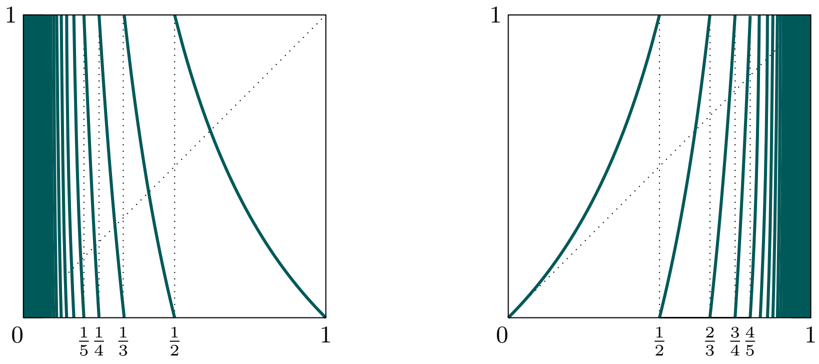
$$\frac{d\mu_G}{d\lambda}(x) = \frac{1}{\log 2} \frac{1}{x+1}, \qquad x \in [0,1],$$

and that μ_G is the only acim for G.

For interval maps that are uniformly expanding but not Markov, the lack of control on the images of intervals of monotonicity makes the study on acims more difficult. For such maps often a functional analytic approach is executed by considering the *Perron-Frobenius operator*, which for a piecewise strictly monotonic C^1 interval map T is an operator acting on non-negative measurable functions h on $[0,1]$ as

$$\mathcal{P}_T h(x) = \sum_{y \in T^{-1}\{x\}} \frac{h(y)}{|DT(y)|}. \tag{1.10}$$

A non-negative measurable function φ on $[0,1]$ is a fixed point of $\mathcal{P}_T$ if and only if it provides an acim μ for T by setting $\mu(B) = \int_B \varphi \, d\lambda$ for each $B \in \mathcal{B}$. The celebrated article by Lasota and Yorke [LY73] shows that the Perron-Frobenius operator of a uniformly expanding interval map that is piecewise C^2 and monotonic on a finite partition has a fixed point that is of bounded variation, thus yielding a finite acim for the interval map. Moreover, in the same article this result is extended to the case



(a) The map G

(b) The map R

Figure 1.3: In (a) *we see the graph of the Gauss map G and in* (b) *the graph of the Rényi map R is depicted.*

that the partition is countable under some additional conditions like long branches. Other extensions of this result can be found in e.g. [BG97] and references therein. Furthermore, in the setting of [LY73] the Perron-Frobenius operator restricted to the space of functions of bounded variations is quasi-compact and therefore has a spectral gap. This implies for instance for a large class of uniformly expanding (not necessarily Markov) interval maps that the correlation function for observables of bounded variation decays exponentially fast. See e.g. [B00, BG97] for more properties on the Perron-Frobenius operator.

One of the simplest ways to go beyond uniformly expanding interval maps is by considering interval maps that have derivative bigger than 1 everywhere except at some neutral fixed point. (A neutral fixed point for an interval map is a fixed point where the derivative is equal to 1.) Such maps are examples of *non-uniformly* expanding maps. The dynamics in the presence of a neutral fixed point are usually significantly different than that of uniformly expanding maps. The reason for this is that under uniformly expanding maps nearby points move away from each other exponentially fast, while points diverge from a neutral fixed point at a subexponential rate. Consequently, these maps typically exhibit intermittency and orbits spend long periods of time close to the neutral fixed point while behaving chaotically otherwise. Furthermore, if an acim exist, its density usually has a pole at the neutral fixed point and if the acim is infinite, then this means that typical orbits stay too long near the neutral fixed point for the acim to be normalisable. Also, if the acim is finite, the rate on the decay of correlations is typically subexponential, see [L06] and references therein. A common approach to study non-uniformly expanding interval maps is to induce the map on a suitable subset such that the induced transformation is a uniformly expanding (Markov) interval map.

Example 1.3.3 (Rényi map). The *Rényi map* $R : [0, 1] \to [0, 1]$ is an example of a non-uniformly expanding interval map with a neutral fixed point and is given by

$R(1) = 0$ and for $x \neq 1$,

$$R(x) = \frac{1}{1-x} \bmod 1,$$

which can be obtained by reflecting the Gauss map in the vertical line through $\frac{1}{2}$, see Figure 1.3(b). It is shown by Rényi in [R57a] that T admits no finite acim but does have a σ-finite acim μ_R on $[0,1]$ with density

$$\frac{d\mu_R}{d\lambda}(x) = \frac{1}{x}, \qquad x \in (0,1],$$

which is the only σ-finite acim for R.

Example 1.3.4 (LSV maps). For each $\alpha \in (0,\infty)$, let

$$S_\alpha : [0,1] \to [0,1], \quad S_\alpha(x) = \begin{cases} x(1+2^\alpha x^\alpha) & \text{if} \quad x \in [0,\frac{1}{2}], \\ 2x-1 & \text{if} \quad x \in (\frac{1}{2},1]. \end{cases} \tag{1.11}$$

The map S_α has a neutral fixed point at zero. The graph of S_α is shown in Figure 1.4(a). Members of the family $\{S_\alpha : \alpha \in (0,\infty)\}$ are called *Liverani-Saussol-Vaienti (LSV) maps* and were first introduced in [LSV99]. It follows from older results in e.g. [P80, T80] that S_α admits a unique acim μ_α with density that behaves like $x^{-\alpha}$. More precisely, there exist constants $C_2 > C_1 > 0$ such that

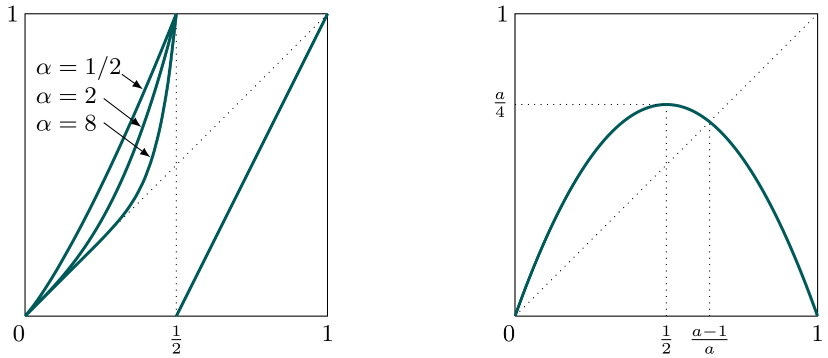
$$C_1 \cdot x^{-\alpha} \leq \frac{d\mu_\alpha}{d\lambda}(x) \leq C_2 \cdot x^{-\alpha}, \qquad x \in (0,1].$$

In particular, μ_α is finite if and only if $\alpha \in (0,1)$. Furthermore, it is shown in [H04, LSV99, Y99, G04] that correlations decay polynomially fast if $\alpha \in (0,1)$. More precisely, in e.g. [Y99] it is shown (in a more general fashion) by constructing a Young tower that for all $\alpha \in (0,1)$, $f \in L^\infty([0,1],\mu_\alpha)$ and $g : [0,1] \to \mathbb{R}$ Hölder continuous there are $C > 0$ and $N \in \mathbb{N}$ such that

$$\mathrm{Cor}_{n,S_\alpha,\mu_\alpha}(f,g) \leq C \cdot n^{1-1/\alpha}$$

for each integer $n \geq N$. In particular, for such an f the summability condition from (1.9) is satisfied if $\alpha \in (0,\frac{1}{2})$ and it is shown in [Y99] that these observables therefore satisfy the CLT. Moreover, it is shown in [G04] that for a class of observables that vanish near zero the rate $n^{1-1/\alpha}$ is in fact sharp and that such observables satisfy the CLT for all $\alpha \in (0,1)$.

The previous example falls in a more general class that is referred to as *Manneville-Pomeau maps*, which are transformations on $[0,1]$ consisting of two increasing branches onto $[0,1]$ with neutral fixed point at zero and everywhere else derivative bigger than 1. The intermittent behaviour of these maps was first studied by Manneville and Pomeau in [PM80, MP80, BPV86] to investigate intermittency of turbulent flows. The standard Manneville-Pomeau maps are given by $x \mapsto x + x^{1+\alpha} \bmod 1$ with $\alpha > 0$ and the LSV maps from Example 1.3.4 were introduced in [LSV99] as a simplification of the standard Manneville-Pomeau maps in the sense that the right branch of the LSV maps is linear. Typically one considers a one-parameter family of Manneville-Pomeau maps where the parameter determines the behaviour around zero, and typically there is a phase transition threshold for this parameter that separates the cases where the acim is either finite or infinite. For the LSV maps S_α parametrised by $\alpha \in (0, \infty)$ as in (1.11) this threshold is $\alpha = 1$.



(a) The map S_α (b) The map L_a

Figure 1.4: In (a) *we see the graph of the LSV map S_α for several values of α and in* (b) *the graph of the logistic map L_a is depicted.*

Another well-studied class of interval maps that are not uniformly expanding are smooth maps that have *critical points*, which are points where the derivative of the map is zero. We restrict our attention to critical points $c \in (0, 1)$ of a C^3 interval map T that are *non-flat*, i.e. there exist constants $0 < K \leq M$ and $\ell > 1$ such that for each $x \in [0, 1]$

$$K|x - c|^{\ell-1} \leq |DT(x)| \leq M|x - c|^{\ell-1}.$$

In this case ℓ is called the *(critical) order* of c. If T has only one critical point $c \in (0, 1)$ such that $DT > 0$ on $[0, c)$ and $DT < 0$ on $(c, 1]$, then T is said to be *unimodal*. Furthermore, $c_1 = T(c)$ is referred to as a *critical value* of T and the orbit of c_1 under T is called the *postcritical* orbit. Typically, orbits that get close to c remain close to this postcritical orbit for some time. It was recognised in [M81, CE83, NvS91, KN92, Y92, BLvS03, BSvS03] among others that the influence of a critical point c of T on the dynamics depends crucially on the expansion gained along the postcritical orbit, that is on the growth of $DT^n(c_1)$ as n grows. Indeed, orbits that get close to c will stay close to the postcritical orbit for some time due to the contracting behaviour near c. The faster $DT^n(c_1)$ grows as n grows, the faster the contraction gets compensated and the shorter orbits "shadow" the postcritical orbit. For the existence of an acim for unimodal maps it suffices in general to assume that $DT^n(c_1)$ stays large enough for all n sufficiently large and no growth of $DT^n(c_1)$ has to be assumed [BSvS03], but the rates on which correlations decay is strongly related with the growth of $DT^n(c_1)$ [BLvS03, Y92]. For instance, it was shown in [BLvS03] using a Young tower that, roughly speaking, if the derivative along all postcritical orbits of the map grows exponentially or polynomially, then correlations typically decay exponentially or polynomially, respectively.

Example 1.3.5 (Logistic maps). An important and well-studied class of unimodal maps is the family $\{L_a : a \in [0,4]\}$ of *logistic maps* given by

$$L_a : [0,1] \to [0,1], \quad L_a(x) = ax(1-x),$$

see Figure 1.4(b). It can in fact be shown that any unimodal map is semi-conjugate to a logistic map, see e.g. [dMvS93, Section 2.6] for more details. The dynamical behaviour of L_a depends crucially on the value of the parameter a. For $0 \leq a \leq 1$ all orbits under L_a converge to the attracting fixed point zero. If $a > 1$, then zero is a repelling fixed point for L_a and for $1 < a \leq 3$ orbits instead converge to the fixed point $\frac{a-1}{a}$ with a speed that is dependent on the value of a. For instance, if $a = 2$, then the critical value $\frac{a}{4}$ coincides with $\frac{a-1}{a}$ and there is superexponential convergence, whereas for $a = 3$ orbits converge sublinearly fast to $\frac{a-1}{a}$. For Lebesgue almost all $a \in (3,4)$ the map L_a either has an attracting periodic orbit or a finite acim [J81, L02]. Furthermore, it was shown by Young in [Y92] that there exists a subset in $(3,4)$ of positive Lebesgue measure such that for each a in this subset the map L_a has exponential decay of correlations for observables of bounded variation and also that such observables satisfy the CLT under L_a. Finally, the full branched logistic map L_4 admits an ergodic probability acim μ and is a special case in the sense that its invariant density is known explicitly, namely $\frac{d\mu}{d\lambda}(x) = \frac{2}{\pi}\frac{1}{\sqrt{1-x^2}}$. In this case it is clear what the postcritical orbit looks like, namely it consists of the points 1 and 0. Since L_4 is expanding in these two points this map has exponential decay of correlations, see e.g. [Y92]. For a more extensive discussion on logistic maps, see again [dMvS93].

Usually techniques for obtaining acims for an interval map T, including the ones discussed so far, require some control on the *distortion* of iterates of T. There are several definitions in the literature of bounded distortion of a C^1 map $T : I \to \mathbb{R}$ where I is an interval in $\mathbb{R}$. One way to define it is that there exists $K > 1$ such that for each $x, y \in I$,

$$\frac{1}{K} \leq \frac{DT(x)}{DT(y)} \leq K,$$

thus indicating that T cannot be too non-linear. This gives the following control on the sizes of images of intervals: Let $J \subseteq I$ be another interval. By the Mean Value Theorem there exists an $x \in J$ with $|DT(x)| = \frac{\lambda(T(J))}{\lambda(J)}$ and a $y \in I$ with $|DT(y)| = \frac{\lambda(T(I))}{\lambda(I)}$. Hence,

$$\frac{1}{K}\frac{\lambda(J)}{\lambda(I)} \leq \frac{DT(x)}{DT(y)}\frac{\lambda(J)}{\lambda(I)} = \frac{\lambda(T(J))}{\lambda(T(I))} \leq K\frac{\lambda(J)}{\lambda(I)}. \tag{1.12}$$

To guarantee for bounded distortion commonly interval maps where the branches have non-positive Schwarzian derivative are considered. For a C^3 map $T : I \to \mathbb{R}$ on an interval I the *Schwarzian derivative* of T at $x \in I$ with $DT(x) \neq 0$ is defined by

$$\boldsymbol{S}T(x) = \frac{D^3T(x)}{DT(x)} - \frac{3}{2}\Big(\frac{D^2T(x)}{DT(x)}\Big)^2.$$

We say that T has *non-positive* Schwarzian derivative on I if $DT(x) \neq 0$ and $\boldsymbol{S}T(x) \leq 0$ for all $x \in I$. A direct computation shows that the Schwarzian derivative of the composition of two C^3 maps $T_1 : I_1 \to \mathbb{R}$ and $T_2 : I_2 \to \mathbb{R}$ with $T_1(I_1) \subseteq I_2$ and I_1 and I_2 intervals satisfies

$$\boldsymbol{S}(T_2 \circ T_1)(x) = \boldsymbol{S}T_2\big(T_1(x)\big) \cdot |DT_1(x)|^2 + \boldsymbol{S}T_1(x). \tag{1.13}$$

Hence, $\boldsymbol{S}(T_2 \circ T_1) \leq 0$ provided $\boldsymbol{S}T_1 \leq 0$ and $\boldsymbol{S}T_2 \leq 0$. We will use the following two well-known properties of maps with non-positive Schwarzian derivative (see e.g. [dMvS93, Section 4.1]).

Theorem 1.3.6 (Koebe Principle). *For each $\rho > 0$ there exist $K^{(\rho)} > 1$ and $M^{(\rho)} > 0$ with the following property. Let $J \subseteq I$ be two intervals and suppose that $T : I \to \mathbb{R}$ has non-positive Schwarzian derivative. If both components of $T(I) \backslash T(J)$ have length at least $\rho \cdot \lambda(T(J))$, then*

$$\frac{1}{K^{(\rho)}} \leq \frac{DT(x)}{DT(y)} \leq K^{(\rho)}, \qquad \forall x, y \in J \tag{1.14}$$

and

$$\Big|\frac{DT(x)}{DT(y)} - 1\Big| \leq M^{(\rho)} \cdot \frac{|T(x) - T(y)|}{\lambda(T(J))}, \qquad \forall x, y \in J. \tag{1.15}$$

Note that the constants $K^{(\rho)}, M^{(\rho)}$ only depend on ρ and not on the map T.

In particular, Koebe's Principle implies bounded distortion on subintervals of the domain of a map with non-positive Schwarzian derivative.

Theorem 1.3.7 (Minimum Principle). *Let $I = [a, b]$ be a closed interval and suppose that $T : I \to \mathbb{R}$ has non-positive Schwarzian derivative. Then*

$$|DT(x)| \geq \min\{|DT(a)|, |DT(b)|\}, \quad \forall x \in [a, b].$$

A consequence of the Minimum Principle is that for any $T : I \to \mathbb{R}$ with non-positive Schwarzian derivative the absolute value of the derivative $|DT|$ has locally no strict minima in the interior of I.

§1.4 Random interval maps

Recall the definition of a random dynamical system $(\mathcal{T}, \mathbb{P})$ given in Section 1.1, where the evolution of the system is given by (1.2). Frequently random dynamical systems are studied by a deterministic map called the skew product. Given a family $\mathcal{T} = \{T_i : X \to X\}_{i \in I}$ the corresponding *skew product* is given by

$$F : I^{\mathbb{N}} \times X \to I^{\mathbb{N}} \times X,\ (\omega, x) \mapsto \big(\tau\omega, T_{\omega_1}(x)\big),$$

where $\omega = (\omega_1, \omega_2, \dots) \in I^{\mathbb{N}}$ and $\tau : I^{\mathbb{N}} \to I^{\mathbb{N}}$ is the *left shift* on $I^{\mathbb{N}}$, i.e. $\tau\omega = (\omega_2, \omega_3, \dots)$. We use the following notation for the compositions of the maps T_i. For each $\omega \in I^{\mathbb{N}}$ and each $n \in \mathbb{N}_0$ define

$$T_{\omega_1 \cdots \omega_n}(x) = T_\omega^n(x) = \begin{cases} x, & \text{if} \quad n = 0, \\ T_{\omega_n} \circ T_{\omega_{n-1}} \circ \cdots \circ T_{\omega_1}(x), & \text{for } n \geq 1. \end{cases} \tag{1.16}$$

With this notation, we can write the iterates of the skew product F as

$$F^n(\omega, x) = (\tau^n \omega, T_\omega^n(x)),$$

from which it becomes clear that F generates in its second coordinate the random orbits of the form in (1.2). For random dynamical systems one is usually interested in the *annealed* dynamics, where the behaviour is averaged over ω, as well as the *quenched* dynamics, where the behaviour is studied for fixed ω. In this dissertation we will mostly focus on obtaining annealed results and for this we will study the skew product.

For the random dynamical systems that we consider the index set I is assumed to be a Polish space and we write $\mathcal{B}_I$ for the corresponding Borel σ-algebra on I. Then the probability space $(I^{\mathbb{N}}, \mathcal{B}_I^{\mathbb{N}}, \mathbb{P})$ is referred to as the *base space* of the random dynamical system. Furthermore, for the case that I is finite or countable we introduce the following definitions. For any $n \in \mathbb{N}$ we use $\boldsymbol{u} \in I^n$ to denote a *word* $\boldsymbol{u} = u_1 \cdots u_n$. The set I^0 contains only the empty word, which we denote by ϵ. We write $I^* = \bigcup_{n \geq 0} I^n$ for the collection of all finite words with digits from I. We use the notation $|\boldsymbol{u}|$ for the length of $\boldsymbol{u} \in I^*$, so $|\boldsymbol{u}| = n$ for $\boldsymbol{u} \in I^n$, and for two words $\boldsymbol{u} \in I^n$ and $\boldsymbol{v} \in I^m$ the concatenation of $\boldsymbol{u}$ and $\boldsymbol{v}$ is denoted by $\boldsymbol{uv} \in I^{n+m}$. Furthermore, on $I^{\mathbb{N}}$ we use for each $\boldsymbol{u} \in I^n$ the notation

$$[\boldsymbol{u}] = [u_1 \cdots u_n] = \{\omega \in I^{\mathbb{N}} : \omega_1 = u_1, \dots, \omega_n = u_n\}$$

for the *cylinder set* corresponding to $\boldsymbol{u}$. We also use the notation from (1.16) for finite words $\boldsymbol{u} \in \Sigma^m$, $m \geq 1$, instead of sequences $\omega \in \Sigma^{\mathbb{N}}$, and with $n \leq m$.

In the first part of this dissertation we consider random interval maps $(\mathcal{T}, \mathbb{P})$ on $[0,1]$ where $\mathcal{T}$ consists of *finitely* many transformations, and for this setting we denote Σ instead of I for the index set of $\mathcal{T}$, i.e. $\mathcal{T} = \{T_j\}_{j \in \Sigma}$, to distinguish from the more general setup above. Moreover, we take $\mathbb{P}$ to be a Bernoulli measure. For a probability vector $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ the ***p**-Bernoulli measure* $m_{\boldsymbol{p}}$ on $(\Sigma^{\mathbb{N}}, \mathcal{B}_\Sigma^{\mathbb{N}})$ is defined on cylinder sets as

$$m_{\boldsymbol{p}}([\boldsymbol{u}]) = \prod_{i=1}^n p_{u_i}, \quad \boldsymbol{u} \in \Sigma^n, \quad n \in \mathbb{N}.$$

This defines $m_{\boldsymbol{p}}$ uniquely since the finiteness of Σ yields that the σ-algebra $\mathcal{B}_\Sigma^{\mathbb{N}}$ generated from the discrete topology on Σ is generated by the cylinder sets. We introduce the notation

$$p_{\boldsymbol{u}} = \prod_{i=1}^n p_{u_i}, \quad \boldsymbol{u} \in I^n, \quad n \in \mathbb{N}$$

to simply write $m_{\boldsymbol{p}}([\boldsymbol{u}]) = p_{\boldsymbol{u}}$. Taking the Bernoulli measure $\mathbb{P} = m_{\boldsymbol{p}}$ as probability law on the base space means that at each time step for each $j \in \Sigma$ the map T_j is applied with probability p_j independently from the maps that are applied at other time steps. In this case we refer to $(\mathcal{T}, \mathbb{P})$ as an *i.i.d.* random interval map.

Suppose $(\mathcal{T}, \mathbb{P})$ is a random dynamical system with index set I and where m is some reference measure on the state space X. When we speak of an acim ρ for F we mean that the absolute continuity of ρ is with respect to the reference measure $\mathbb{P} \times m$. Invariant measures of the skew product and in particular acims give information on the long-term average behaviour of orbits under a random dynamical system. For instance, suppose ρ is an ergodic probability acim for F. Then it follows from Birkhoff's Ergodic Theorem applied to (F, ρ) that for each bounded measurable function f on X there exists a measurable set $A \subseteq I^{\mathbb{N}} \times X$ with $\mathbb{P} \times m(A) > 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T_\omega^k(x)) = \int f(x) d\rho(\omega, x) \tag{1.17}$$

holds for all $(\omega, x) \in A$. In case $(\mathcal{T}, \mathbb{P})$ is an i.i.d. random interval map[2], then the finite acims of F are of a special form as the next lemma below will show.

Given a finite family $\mathcal{T} = \{T_j : [0,1] \to [0,1]\}_{j\in\Sigma}$ and probability vector $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ we say that a Borel measure μ on $[0,1]$ is *stationary* w.r.t. $(\mathcal{T}, \boldsymbol{p})$ if

$$\sum_{j\in\Sigma} p_j \mu(T_j^{-1} B) = \mu(B) \qquad \text{for all Borel sets } B \subseteq [0,1]. \tag{1.18}$$

It is easy to verify that $m_{\boldsymbol{p}} \times \mu$ is F-invariant if and only if μ is stationary. For brevity we call a stationary Borel measure μ on $[0,1]$ that is absolutely continuous w.r.t. λ an *acs* measure for $(\mathcal{T}, \boldsymbol{p})$.

Lemma 1.4.1 ([M85a], see also [F99, Lemma 3.2]). *Let $\{T_j : [0,1] \to [0,1]\}_{j\in\Sigma}$ be a finite family of non-singular transformations and let F denote the associated skew product. Furthermore, let $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ be a probability vector. Then the $m_{\boldsymbol{p}} \times \lambda$-absolutely continuous F-invariant finite measures are precisely the measures of the form $m_{\boldsymbol{p}} \times \mu$ where μ is a finite acs measure.*

Statistical properties of the skew product F yield statistical properties of the corresponding random dynamical system $(\mathcal{T}, \mathbb{P})$. For instance, if ρ is an invariant probability measure for F and f is a bounded measurable function on the state space X of the system that satisfies the Central Limit Theorem from Definition 1.2.15 w.r.t. (F, ρ) for some $\sigma > 0$, then this means

$$\lim_{n\to\infty} \rho\Big(\Big\{(\omega, x) : \frac{\sum_{k=0}^{n-1} \big(f(T_\omega^k(x)) - \int f(x) d\rho(\omega, x)\big)}{\sigma\sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt,$$

[2]Lemma 1.4.1 is formulated in [M85a, F99] for i.i.d. random interval maps but in fact holds for all random dynamical systems where the constituent maps are non-singular and where the probability law on the base space is a Bernoulli measure.

hence giving information on the fluctuations of the convergence as in (1.17). As another example, suppose the state space X is a Polish space, $\mathbb{P}$ is τ-invariant and ρ is again an invariant probability measure for F. Then there exists a family of probability measures $\{\rho_\omega\}_{\omega\in I^{\mathbb{N}}}$ on X such that for all bounded measurable functions f and g on X

$$\mathrm{Cor}_{n,F,\rho}(f,g) = \int_{I^{\mathbb{N}}} \mathrm{Cor}_{n,\omega}(f,g)d\mathbb{P}(\omega) + \mathrm{Cor}_{n,\tau,\mathbb{P}}(\bar{f},\bar{g}), \tag{1.19}$$

where

$$\mathrm{Cor}_{n,\omega}(f,g) = \int_X f\circ T_\omega^n \cdot g d\rho_\omega - \int_X f d\rho_{\tau^n\omega}\cdot \int_X g d\rho_\omega$$

are *(forward) fiberwise correlations* and $\bar{f}(\omega) = \int_X f d\rho_\omega$ and $\bar{g}(\omega) = \int_X g d\rho_\omega$ (see e.g. [B99, Subsection 0.2] or use (5.21) from Section 5.4 for a justification). In other words, the speed at which correlations decay with respect to (F,ρ) gives information on the decay of annealed correlations of the random system.[3]

Various results on acims and statistical properties for skew products that are associated with random interval maps have been found in the last decades. We briefly discuss some techniques and results for i.i.d. random interval maps and refer the reader to [ANV15] and references therein for a more extensive discussion.

Let $(\mathcal{T}, m_{\boldsymbol{p}})$ be an i.i.d. random interval map where $\mathcal{T} = \{T_j : [0,1]\to[0,1]\}_{j\in\Sigma}$ is a finite family of piecewise strictly monotonic C^1 interval maps. The *Perron-Frobenius operator* $\mathcal{P}_{\mathcal{T},\boldsymbol{p}}$ associated to $(\mathcal{T},\boldsymbol{p})$ is defined on the space of non-negative measurable functions h on $[0,1]$ by

$$\mathcal{P}_{\mathcal{T},\boldsymbol{p}}h(x) = \sum_{j\in\Sigma} p_j \mathcal{P}_{T_j}h(x), \tag{1.20}$$

where each $\mathcal{P}_{T_j}$ is as given in (1.10). Then a non-negative measurable function φ on $[0,1]$ is a fixed point of $\mathcal{P}_{\mathcal{T},\boldsymbol{p}}$ if and only if the Borel measure μ on $[0,1]$ given by $\mu(A) = \int \varphi d\lambda$ is an acs measure. Hence, it follows from Lemma 1.4.1 that the integrable fixed points of (1.20) provide precisely the finite acims of F. By applying techniques similar to the ones for the Perron-Frobenius operator from (1.10) discussed in the previous section it is possible to obtain analogous annealed results if the random interval map is uniformly expanding in some sense. We say $(\mathcal{T}, m_{\boldsymbol{p}})$ is *expanding on average* if

$$\sum_{j\in\Sigma} \frac{p_j}{\inf_{x\in[0,1]}|DT_j(x)|} < 1. \tag{1.21}$$

For a large class of i.i.d. random interval maps that satisfy this condition an acs probability measure can be obtained by applying to (1.20) an approach that is similar

[3] The second term on the right-hand side of (1.19) is a correlation function with respect to $(\tau,\mathbb{P})$, and its speed of decay is typically known. In Chapter 3, where we consider decay of correlations for the skew product, $\mathbb{P}$ is a Bernoulli measure, in which case correlations under τ decay exponentially fast.

to that of Lasota and Yorke for (1.10). Furthermore, exponential decay of annealed correlations and Central Limit Theorems can be obtained for functions of bounded variation using the quasi-compactness of (1.20). In [P84, I12] the weaker condition

$$\sup_{x\in[0,1]} \sum_{j\in\Sigma} \frac{p_j}{|DT_j(x)|} < 1, \tag{1.22}$$

is used to obtain for i.i.d. random interval maps the existence of acs probability measures among other results.

Example 1.4.2 (Random Gauss-Rényi map). We have seen that the Gauss map G from Example 1.3.2 admits a probability acim, whereas the Rényi map R from Example 1.3.3 admits no finite acim but does admit a σ-finite acim with density that has a pole at zero. Now let $(\{T_0, T_1\}, m_{\boldsymbol{p}})$ be the i.i.d. random interval map with $\boldsymbol{p} = \{p_0, p_1\}$ where the Gauss map $T_0 = G$ is chosen with probabity p_0 and the Rényi map $T_1 = R$ with probability $p_1 = 1-p_0$. This random interval map satisfies (1.22) if $p_0 \in (0,1)$, and it is shown in [KKV17] that for each $p_0 \in (0,1)$ it admits an invariant probability density of bounded variation, annealed correlations decay exponentially fast and an annealed CLT holds. In other words, if $p_0 \in (0,1)$ the annealed dynamical behaviour is not as much influenced by the presence of the neutral fixed point of T_1 at zero as compared to the $p_0 = 0$ case. On the other hand, it is shown in [KMTV22] that if $p_0 \in (0,1)$ the density of the acs measure is provably less smooth than the invariant density of the Gauss map.

We now give an example of an i.i.d. random interval map that does not possess any uniform expandingness.

Example 1.4.3 (Random LSV maps). In [BBD14, BB16, Z18, NTV18, BBR19, BQT21, NPPT21] i.i.d. random interval maps are considered that are composed of the LSV maps S_α from Example 1.3.4 where α is sampled from some fixed subset $A \subseteq (0,\infty)$. Because the maps S_α share the same neutral fixed point at zero such random interval maps are not uniformly expanding. Instead, it is shown that the annealed dynamics of such random interval maps are governed by the map with the fastest relaxation rate, i.e. the map $S_{\alpha_{\min}}$ where $\alpha_{\min}$ is the minimal value of A. In particular, it is shown in [BBD14] by means of a Young tower for the associated skew product that for the case that A is finite and a subset of $(0,1]$ and $\alpha_{\min} \in (0,1)$ an acs probability measure exists and annealed correlations decay as fast as $n^{1-1/\alpha_{\min}}$, a rate that in [BB16] is shown to be sharp for a class of observables that vanish near zero. In [Z18] it was shown that an acs probability measure also exists without the restriction $A \subseteq (0,1]$ as long as $A \subseteq (0,\infty)$ is finite and $\alpha_{\min}$ lies in $(0,1)$. This was later shown in [BQT21] as well without the finiteness condition on A provided there is a positive probability to choose a parameter < 1, and for this more general setting it is also shown in [BQT21] that annealed correlations decay polynomially fast with a rate that is close (and sometimes equal) to the rate found in [BBD14]. We refer the reader to [BBR19] for quenched results obtained for random LSV maps.

§1.5 Outline of dissertation

We conclude this chapter by giving a brief summary of the content of each chapter.

In **Chapter 2** we consider a wide class of critically intermittent random systems on $[0,1]$. By using the inducing technique from Proposition 1.2.12 we prove that these random interval maps admit a σ-finite acs measure that is either finite or infinite depending on the probabilities of choosing the constituent maps as well as their critical orders. The existence of an infinite acs measure is proven with the help of Kac's Lemma. On the other hand, a finite acs measure is derived by estimating the sizes of neighborhoods around points of the postcritical orbits, which is shown to be sufficient using an argument that involves the Koebe Principle and Minimum Principle.

In **Chapter 3** we derive several statistical properties of critically intermittent systems that are closely related to the ones from Chapter 2 but for which the corresponding skew product is easier to analyse using the Young tower technique. We show that these systems are mixing and that annealed correlations decay polynomially fast. We also provide sufficient conditions for an annealed CLT to hold for a class of Hölder continuous functions.

In **Chapter 4** we investigate what happens to the acs measure of the critically intermittent random systems when the system is modified in such a way that the superexponential convergence is replaced by exponential convergence and the exponential divergence is replaced by polynomial divergence. We show that a similar phase transition for the acs measure holds as the one found in Chapter 2, but for the proof we use techniques different from those executed in Chapter 2 and construct a suitable invariant set for the Perron-Frobenius operator.

In **Chapter 5** we give an extension of Lochs' result from (1.4) to a large class of pairs of random interval maps that produce number expansions. For this we generalise the method from [DF01] by applying fiberwise analogs of the Kolmogorov-Sinai Theorem and Shannon-McMillan-Breiman Theorem. To calculate the fiber entropy we also deduce a random analog of Rokhlin's Formula for entropy. Furthermore, we also provide a corresponding Central Limit Theorem. The chapter is concluded with applying the obtained general theory to common number expansions.

In **Chapter 6** we study the question posed in [JMKA13] that asks what is the minimum number $k(m)$ of output digits of the β-encoder needed to produce with this output m base 2 digits of the same input value. We provide several limit results as $m \to \infty$ in the case that the quantiser threshold fluctuates but the amplification factor and scaling factor are fixed. We end this chapter by observing that the method of [JMKA13] is not optimal for producing large pseudo-random numbers when the amplification factor or scaling factor of the β-encoder fluctuates as well.

PART I

STATISTICAL PROPERTIES OF CRITICALLY INTERMITTENT SYSTEMS

CHAPTER 2

Absolutely continuous invariant measures for critically intermittent systems

This chapter is based on: [HKR+22].

Abstract

Critical intermittency stands for a type of intermittent dynamics caused by an interplay of a superstable fixed point and a repelling fixed point. We consider a large class of random interval maps that exhibit critical intermittency and demonstrate the existence of a phase transition when varying probabilities, where the absolutely continuous invariant measure changes between finite and infinite. We discuss further properties of this measure and show that its density is not in L^q for any $q > 1$. This provides a theory of critical intermittency alongside the theory for the well-studied Manneville-Pomeau maps where the intermittency is caused by a neutral fixed point.

§2.1 Introduction

The concept of critical intermittency has been illustrated in Subsection 1.1.1 by random i.i.d. applications of the two logistic maps $L_2(x) = 2x(1-x)$ and $L_4(x) = 4x(1-x)$ on the unit interval: Orbits converge superexponentially fast to $\frac{1}{2}$ under applications of L_2, and as soon as L_4 is applied then diverge exponentially fast from the repelling fixed point, behaving chaotically again for some time once escaped. See Figures 1.1(b) and 1.2(b). The alternation between chaotic periods and being in a seemingly steady state is related to the probability of choosing the maps as well as the critical order $\ell_2 = 2$ of L_2, which determines the speed of convergence to $\frac{1}{2}$. In [AGH18] it is shown that this random system admits a σ-finite acs measure which is infinite if the probability p_2 of choosing L_2 satisfies $p_2 > \frac{1}{2}$. We will see in this chapter that this acs measure in fact is finite if and only if $p_2 \cdot \ell_2 < 1$. More generally, we will prove such a phase transition for the acs measure for a large family of random interval maps with critical intermittency.

In this chapter we consider critically intermittent systems on $[0,1]$ that are defined by random i.i.d. applications of so-called *bad* maps that share a globally superattracting fixed point $c \in (0,1)$ and *good* maps that map c into the common invariant and repelling set $\{0,1\}$. To be precise, the families of maps we consider are defined as follows.

Throughout the text we fix a point $c \in (0,1)$ that will represent the single critical point of our maps, both good and bad.

A map $T_g : [0,1] \to [0,1]$ is in the class of *good maps*, denoted by $\mathfrak{G}$, if

(G1) $T_g|_{[0,c)}$ and $T_g|_{(c,1]}$ are C^3 diffeomorphisms onto $[0,1)$ or $(0,1]$ and $T_g(c) \in \{T_g(c-), T_g(c+)\}$;

(G2) T_g has non-positive Schwarzian derivative on $[0,c)$ and $(c,1]$;

(G3) to T_g we can associate three constants $r_g \geq 1$, $0 < K_g < 1$ and $M_g > r_g$ such that, for each $x \in [0,1]$,

$$K_g|x-c|^{r_g-1} \leq |DT_g(x)| \leq M_g|x-c|^{r_g-1}; \tag{2.1}$$

(G4) we have $|DT_g(0)|, |DT_g(1)| > 1$.

These conditions imply in particular that $T_g(\{0,c,1\}) \subseteq \{0,1\}$, that at least one of the maps $T_g|_{[0,c]}$ or $T_g|_{[c,1]}$ is continuous, and that both branches of T_g are strictly monotone. Note also that the conditions $K_g < 1$ and $M_g > r_g$ are superfluous, since we can always choose a smaller constant K_g and larger constant M_g to satisfy (2.1), but we need these specific bounds in our estimates later. The critical point c is mapped to either 0 or 1 under each of the good maps and both 0 and 1 are (eventually) fixed points or periodic points (with period 2) by (G1) that are repelling by (G4). Furthermore, a consequence of the Minimum Principle is that $|DT_g|$ has locally no strict minima in the intervals $(0,c)$ and $(c,1)$. In particular, there cannot be any attracting fixed points for T_g in $(0,c)$ and $(c,1)$. Examples of good maps include the doubling map and surjective unimodal maps, see Figures 2.1(a)-(d).

The choice of conditions (G1)-(G4) is based on two factors: firstly, these conditions incorporate the most important properties of the 'good' logistic map $L_4(x) = 4x(1 - x)$, which is the primary motivating example for this chapter, and secondly, some of the techniques used in this chapter are motivated by the work of Nowicki and Van Strien [NvS91] where the following result has been proven. Let λ denote the Lebesgue measure on $[0, 1]$.

Theorem 2.1.1 (Main Theorem in [NvS91]). *Suppose that $T : [0, 1] \to [0, 1]$ is unimodal, C^3, has negative Schwarzian derivative and that the critical point of T is of order $r \geq 1$. Moreover assume that the growth rate of $|DT^n(c_1)|$, $c_1 = T(c)$, is so fast that*

$$\sum_{n=0}^{\infty} |DT^n(c_1)|^{-1/r} < \infty. \tag{2.2}$$

Then T has a unique probability acim μ which is ergodic and of positive entropy. Furthermore, there exists a positive constant K such that

$$\mu(B) \leq K \cdot \lambda(B)^{1/r}, \tag{2.3}$$

for any Borel set $B \subseteq [0, 1]$. Finally, the density $\frac{d\mu}{d\lambda}$ of the measure μ with respect to λ is an $L^{\tau-}$-function where $\tau = r/(r-1)$ and $L^{\tau-} = \bigcap_{1 \leq t < \tau} L^t$ and $L^t = \{f \in L^1([0, 1], \lambda) : \int_0^1 |f|^t d\lambda < \infty\}$.

Formally this result is not immediately applicable to the good maps we introduced. The difference, however, is not principal and the conclusion remains exactly the same, the main reason being that the conditions (G1) and (G4) imply the growth rate (2.2), and hence any good map admits a unique probability acim.

A map $T_b : [0, 1] \to [0, 1]$ is in the class of *bad maps*, denoted by $\mathfrak{B}$, if

(B1) $T_b|_{[0,c)}$ and $T_b|_{(c,1]}$ are C^3 diffeomorphisms onto $[0, c)$ or $(c, 1]$ and $T_b(c) = c$;

(B2) T_b has non-positive Schwarzian derivative on $[0, c)$ and $(c, 1]$;

(B3) to T_b we can associate three constants $\ell_b > 1$, $0 < K_b < 1$ and $M_b > \ell_b$ such that, for each $x \in [0, 1]$,

$$K_b|x - c|^{\ell_b - 1} \leq |DT_b(x)| \leq M_b|x - c|^{\ell_b - 1}; \tag{2.4}$$

(B4) we have $|DT_b(0)|, |DT_b(1)| > 1$.

In particular (B1) implies that $T_b(\{0, 1\}) \subseteq \{0, 1\}$, that T_b is continuous, and that T_b is strictly monotone on the intervals $[0, c]$ and $[c, 1]$. In contrast to (G3), note that in (B3) we have assumed that ℓ_b is not equal to one. This means that $DT_b(c) = 0$, so c is a superstable fixed point for each bad map. Furthermore, c has $(0, 1)$ as its basin of attraction, since it again follows from the Minimum Principle that $|DT_b|$ has locally no strict minima in the intervals $(0, c)$ and $(c, 1)$. Combining this with the Poincaré Recurrence Theorem, an immediate consequence is that the only finite T_b-invariant

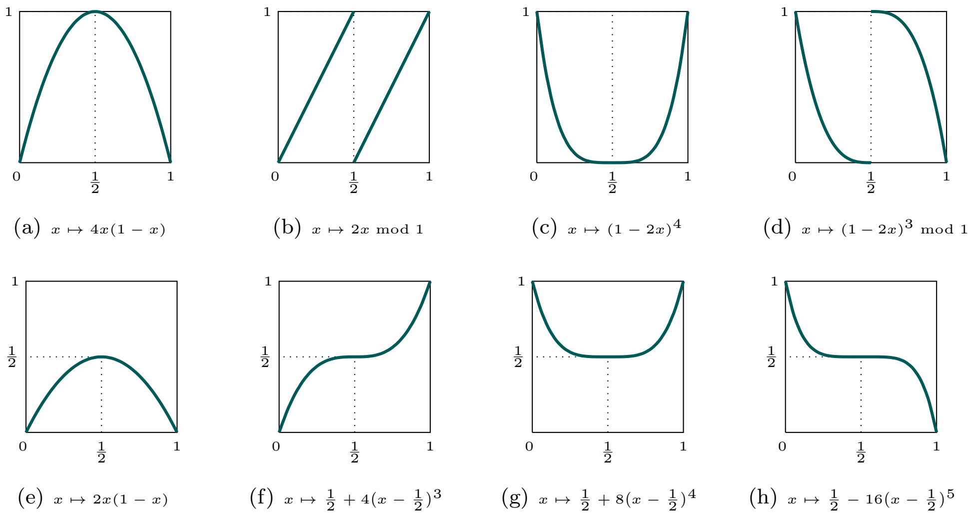


Figure 2.1: Eight maps with critical point $c = \frac{1}{2}$. The upper four graphs (a)-(d) *show good maps, while in* (e)-(h) *we see the graphs of four bad maps.*

measures are linear combinations of Dirac measures at $0, c$, and 1. For examples, see Figures 2.1(e)-(h).

The random systems we consider in this chapter are the following. Let $T_1, \dots, T_N \in \mathfrak{G} \cup \mathfrak{B}$ be a finite collection of good and bad maps. Write $\Sigma_G = \{1 \le j \le N : T_j \in \mathfrak{G}\}$ and $\Sigma_B = \{1 \le j \le N : T_j \in \mathfrak{B}\}$ for the index sets of the good and bad maps respectively and assume that $\Sigma_G, \Sigma_B \neq \emptyset$. Write $\Sigma = \{1, \dots, N\} = \Sigma_G \cup \Sigma_B$. Let F be the skew product associated to $\{T_j\}_{j \in \Sigma}$, i.e.

$$F : \Sigma^{\mathbb{N}} \times [0,1] \to \Sigma^{\mathbb{N}} \times [0,1], \ (\omega, x) \mapsto (\tau\omega, T_{\omega_1}(x)), \tag{2.5}$$

where τ denotes the left shift on sequences in $\Sigma^{\mathbb{N}}$. Let $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ be a probability vector representing the probabilities with which we choose the maps T_j, $j \in \Sigma$, and let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $\Sigma^{\mathbb{N}}$. Our main results are the following.

Theorem 2.1.2. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ a strictly positive probability vector.*

(a) *There exists a unique (up to scalar multiplication) σ-finite acs measure μ for $(\mathcal{T}, \boldsymbol{p})$. Moreover, F is ergodic w.r.t. $m_{\boldsymbol{p}} \times \mu$.*

(b) *The density $\frac{d\mu}{d\lambda}$ is bounded away from zero, is locally Lipschitz on $(0, c)$ and $(c, 1)$ and is not in L^q for any $q > 1$.*

Theorem 2.1.3. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ a strictly positive probability vector. Let μ be the unique acs measure from Theorem 2.1.2. Set $\theta = \sum_{b \in \Sigma_B} p_b \ell_b$. Then μ is finite if and only if $\theta < 1$. In this case, there exists a* *constant $C > 0$ such that*

$$\mu(B) \leq C \cdot \sum_{k=0}^{\infty} \theta^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}} \tag{2.6}$$

for any Borel set $B \subseteq [0,1]$, where $r_{\max} = \max\{r_g : g \in \Sigma_G\}$ and $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$.

As we shall see in (2.28) the bound in (2.6) can be improved by not bounding mixtures $\ell_{\boldsymbol{b}} r_g = \prod_{i=1}^k \ell_{b_i} r_g$ by their maximal value $\ell_{\max}^k r_{\max}$, but this improvement does not change the qualitative behaviour of the bound.

Theorem 2.1.3 shows that the system undergoes a phase transition where the acs measure changes between finite and infinite with threshold $\theta = 1$. Interestingly, this situation is significantly different than for the random LSV maps discussed in Example 1.4.3 where the existence of an acs probability measure only depends on whether there is a positive probability to choose an LSV map with parameter < 1.

It is also worth mentioning that it follows from Theorem 2.1.3 that not only does no finite acs measure exist if $\theta \geq 1$, but also that no physical measure[1] for F exists in this case. Indeed, this follows by applying Aaronson's Ergodic Theorem [A97, Theorem 2.4.2] to the infinite measure $m_{\boldsymbol{p}} \times \mu$.

Since μ is an acs measure, the density $\frac{d\mu}{d\lambda}$ is a fixed point of the associated Perron-Frobenius operator being of the form as in (1.20). Moreover, Theorem 2.1.2 tells that this density $\frac{d\mu}{d\lambda}$ is bounded away from zero. Using these two statements it is easy to see that $\frac{d\mu}{d\lambda}$ blows up to infinity at the points zero and one and also at least on one side of c. See Figure 2.2 for an example. Furthermore, Theorem 2.1.3 says that $\frac{d\mu}{d\lambda}$ is integrable if and only if θ is small enough, namely $\theta < 1$. This intuitively makes sense since for a smaller value of θ the attraction of orbits to c is weaker on average and consequently orbits typically spend less time near zero and one once a good map is applied.

The inequality (2.6) is the counterpart of the Nowicki-Van Strien inequality (2.3), and naturally gives a substantially worse bound due to the presence of bad maps. It is not immediately clear how much worse (2.6) is in comparison to (2.3). However, the following holds.

Corollary 2.1.4. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ a strictly positive probability vector. Suppose $\theta = \sum_{b\in\Sigma_B} p_b \ell_b < 1$. Then there exist $K > 0$ and $\varkappa > 0$ such that for any Borel set $B \subseteq [0,1]$ with $\lambda(B) \in (0,1)$ one has*

$$\mu(B) \leq K \frac{1}{\log^{\varkappa}(1/\lambda(B))}.$$

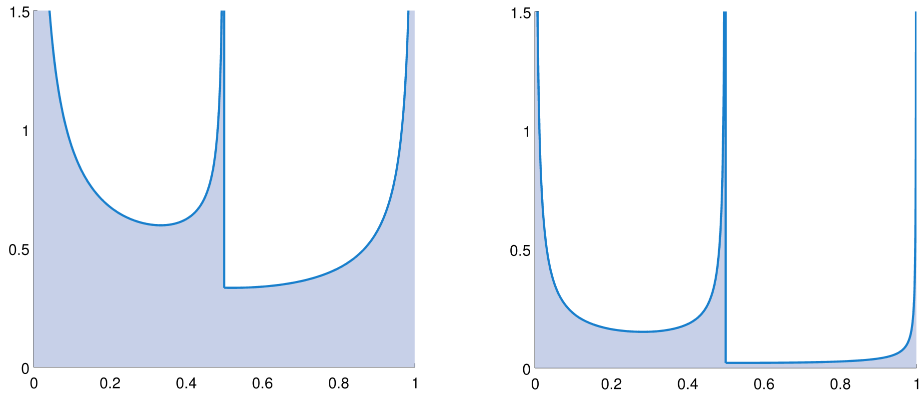
[1] A probability measure μ is a *physical measure* for a transformation $T : X \to X$ if it is T-invariant and there exists a set $U \subseteq X$ of positive Lebesgue measure (for our setting, this means $m_{\boldsymbol{p}} \times \lambda(U) > 0$) such that for each continuous function $f : X \to \mathbb{R}$ we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int_X f \, d\mu, \qquad \text{for each } x \in U.$$

Moreover, the acs measure from Theorem 2.1.2 depends continuously on the probability vector $\boldsymbol{p} \in \mathbb{R}^N$ as the next result shows. Here we write $\mu_{\boldsymbol{p}}$ for the acs probability measure that corresponds to the probability vector $\boldsymbol{p}$.

Corollary 2.1.5. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above. For each $n \geq 0$, let $\boldsymbol{p}_n = (p_{n,j})_{j\in\Sigma}$ be a strictly positive probability vector such that $\sup_n \sum_{b\in\Sigma_B} p_{n,b}\ell_b < 1$ and assume that $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ in $\mathbb{R}^N_+$. Then the sequence $\mu_{\boldsymbol{p}_n}$ converges weakly to $\mu_{\boldsymbol{p}}$.*

As discussed in Section 1.4 the problem of finding acs probability measures for random interval maps that are expanding on average is well studied and these results often rely on bounded variation techniques from Lasota and Yorke [LY73]. In [P84, Section 4] these techniques are extended to a class of i.i.d. random interval maps on the unit interval that are not expanding on average and are composed of a uniformly expanding map T and a contracting map S such that S contracts no faster than that T expands. An example is $T(x) = 2x \mod 1$ and $S(x) = \frac{x}{2}$. Under additional conditions on S and T and assuming that the probability of choosing T is bigger than S, Pelikan shows that such a random map admits an acs probability measure. The proof of this result relies heavily on the explicit expression that Pelikan has for a conjugacy map between this random map and another random map that is expanding on average and consists of two Lasota-Yorke type maps. This allows Pelikan to find the order of the density of the acs measure near the pole at 0. As we see in Theorem 2.1.2 and prove in Subsection 2.2.3, for our systems the density of the acs measure is not in L^q for any $q > 1$. This suggests that a similar conjugacy for our systems, if it exists, might not have such a nice explicit expression. Therefore, we resort to techniques similar to the ones used by Nowicki and Van Strien in their proof of Theorem 2.1.1 rather than the techniques introduced by Lasota and Yorke.



(a) (b)

Figure 2.2: Approximation of $\frac{d\mu}{d\lambda}$ in case $\Sigma_G = \{1\}$, $\Sigma_B = \{2\}$, $T_1(x) = L_4(x) = 4x(1-x)$ and $T_2(x) = L_2(x) = 2x(1-x)$ for two different values of p_1. Both pictures depict $P^{20}_{\mathcal{T},\boldsymbol{p}}(1)$ with Perron-Frobenius operator $P_{\mathcal{T},\boldsymbol{p}}$, where in (a) *we have taken $p_1 = \frac{3}{4}$ and in* (b) *$p_1 = \frac{1}{4}$.*

To be more precise, for the existence result from Theorem 2.1.2 we use an inducing scheme. This approach is inspired by [AGH18], but the choice of the inducing domain needed some care. With the help of Kac's Lemma we then obtain that the acs measure is infinite in case $\theta \geq 1$. To prove that this measure is finite for $\theta < 1$ we use an approach similar to the one employed in [NvS91] by estimating the sizes of preimages of neighborhoods around points in the postcritical orbits. For this we apply the Minimum Principle and Koebe Principle to iterates of the maps T_j, which is possible due to (G2) and (B2). The main difficulty to obtain these estimates is that it may take an arbitrarily long time before the superattracting fixed point is mapped onto the repelling orbit by one of the good maps, which decreases the regularity of the density of the acs measure. Furthermore, we will use the following key lemma.

Recall the constants ℓ_b, K_b and M_b from (2.4) in condition (B3) and set $\ell_{\min} = \min\{\ell_b : b \in \Sigma_B\}$ and $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$. We prove the next lemma in Subsection 2.2.3.

Lemma 2.1.6. *For all $n \in \mathbb{N}$, $\omega \in \Sigma_B^{\mathbb{N}}$ and $x \in [0,1]$,*

$$\left(\tilde{K}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}} \leq |T_\omega^n(x) - c| \leq \left(\tilde{M}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}},$$

with $\tilde{K} = \left(\frac{\min\{K_b : b \in \Sigma_B\}}{\ell_{\max}}\right)^{\frac{1}{\ell_{\min}-1}} \in (0,1)$ and $\tilde{M} = \left(\frac{\max\{M_b : b \in \Sigma_B\}}{\ell_{\min}}\right)^{\frac{1}{\ell_{\min}-1}} > 1$.

It follows that under iterations of bad maps the distance $|T_\omega^n(x) - c|$ is eventually decreasing superexponentially fast in n. In Section 2.3 we will use the upper bound on $|T_\omega^n(x) - c|$ that we obtained in Lemma 2.1.6 to prove that μ in Theorem 2.1.3 is infinite if $\theta \geq 1$. The lower bound from Lemma 2.1.6 will be used to show that μ is finite if $\theta < 1$.

In (B3) we have assumed that for any bad map T_b the corresponding value ℓ_b is not equal to one. Note that a bad map T_b for which we allow $\ell_b = 1$ satisfies $|DT_b(c)| > 0$, so in this case c is an attracting fixed point for T_b but not superattracting. It should not come as a surprise that results similar to Theorem 2.1.2 and Theorem 2.1.3 also hold in case some or all of the bad maps T_b have $\ell_b = 1$. The proofs presented for these theorems, however, do not immediately carry over. This has mainly to do with the constants $\tilde{K}$ and $\tilde{M}$ from Lemma 2.1.6, which are not well defined in case $\ell_{\min} = 1$. In the last section we explain how the results are affected in case some or all maps T_b satisfy $\ell_b = 1$ and what the necessary changes in the proofs are.

The remainder of this chapter is organised as follows. Section 2.2 is devoted to the proof of Theorem 2.1.2 and in Section 2.3 we prove Theorem 2.1.3. In Section 2.4 we prove Corollaries 2.1.4 and 2.1.5 and explain what the analogues of Theorem 2.1.2 and 2.1.3 are in case $\ell_b = 1$ for one or more $b \in \Sigma_B$ and how the proofs of Theorem 2.1.2 and 2.1.3 need to be modified to get these results. We end this chapter with some final remarks.

§2.2 Existence of a σ-finite acs measure

From now on we fix an integer $N \geq 2$ and consider a finite collection $T_1, \dots, T_N \in \mathfrak{G} \cup \mathfrak{B}$ of good and bad maps in the classes $\mathfrak{G}$ and $\mathfrak{B}$. As in Section 2.1 we write $\Sigma_G = \{1 \leq j \leq N : T_j \in \mathfrak{G}\}$ and $\Sigma_B = \{1 \leq j \leq N : T_j \in \mathfrak{B}\}$ for the corresponding index sets and assume that $\Sigma_G, \Sigma_B \neq \emptyset$. We write $\Sigma = \{1, 2, \dots, N\}$. In this section we prove Theorem 2.1.2, i.e. we establish the existence of an ergodic acs measure and several of its properties using an inducing scheme for the random system F. We fix the index $g \in \Sigma_G$ of one good map T_g and start by constructing an inducing domain that depends on this g. Throughout this section and the next ones we use the notations for words and compositions of the maps T_j introduced in Section 1.4.

§2.2.1 The induced system and first return time partition

The first lemma is needed to specify the set on which we induce. For each $k \in \mathbb{N}$ let x_k and x'_k in $(0, c)$ denote the critical points of T_g^k closest to 0 and c, respectively. Furthermore, let y_k and y'_k in $(c, 1)$ denote the critical points of T_g^k closest to 1 and c, respectively.

Lemma 2.2.1. *We have $x_k \downarrow 0$, $x'_k \uparrow c$, $y'_k \downarrow c$, $y_k \uparrow 1$ as $k \to \infty$.*

Proof. Let a and b denote the critical points of T_g^2 in $(0, c)$ and $(c, 1)$, respectively. Then at least one of the branches $T_g^2|_{(0,a)}$ and $T_g^2|_{(b,1)}$ is increasing. Suppose that $T_g^2|_{(0,a)}$ is increasing. It then follows from the Minimum Principle that $T_g^2(x) \geq \min\{\frac{x}{a}, DT_g^2(0) \cdot x\}$ for each $x \in [0, a]$. To see this, suppose there is an $x \in (0, a)$ with $T_g^2(x) < \min\{\frac{x}{a}, DT_g^2(0) \cdot x\}$. Then there must be a $y \in (0, x)$ with $DT_g^2(y) < \min\{DT_g^2(0), \frac{1}{a}\}$ and a $z \in [x, a)$ with $DT_g^2(z) > \frac{1}{a}$. On the other hand, by the Minimum Principle, $DT_g^2(y) \geq \min\{DT_g^2(0), DT_g^2(z)\}$, a contradiction. Combining this with $DT_g^2(0) > 1$ and defining $L : (0, 1) \to (0, a)$ by $L = (T_g^2|_{(0,a)})^{-1}$, we see that $L^k(a) \downarrow 0$ as $k \to \infty$. Furthermore, define $R : (0, 1) \to (b, 1)$ by $R = (T_g^2|_{(b,1)})^{-1}$. If $T_g^2|_{(b,1)}$ is increasing, we see that similarly $R^k(b) \uparrow 1$ as $k \to \infty$. On the other hand, if $T_g^2|_{(b,1)}$ is decreasing, we have $RL^k(a) \uparrow 1$ as $k \to \infty$. Finally, if $T_g^2|_{(0,a)}$ is decreasing, then $T_g^2|_{(b,1)}$ must be increasing, which yields $LR^k(b) \downarrow 0$ as $k \to \infty$. We conclude that $x_k \downarrow 0$ and $y_k \uparrow 1$ as $k \to \infty$. It follows from (G1) that c is a limit point of both of the sets $\bigcup_{k \in \mathbb{N}} (T_g|_{(0,c)})^{-1}(\{x_k, y_k\})$ and $\bigcup_{k \in \mathbb{N}} (T_g|_{(c,1)})^{-1}(\{x_k, y_k\})$. So $x'_k \uparrow c$, $y'_k \downarrow c$ as $k \to \infty$. $\square$

By the previous lemma and (G1), for $k \in \mathbb{N}$ large enough it holds that

$$\begin{aligned} &T_g(x'_k) \leq x'_k \text{ or } T_g(x'_k) \geq y'_k, \text{ and} \\ &T_g(y'_k) \leq x'_k \text{ or } T_g(y'_k) \geq y'_k, \end{aligned} \tag{2.7}$$

and, using also (G4), (B1) and (B4), for every $j \in \Sigma$,

$$\begin{aligned} &T_j\big([0, x_k] \cup [y_k, 1]\big) \subseteq [0, x'_k) \cup (y'_k, 1] \text{ and} \\ &|DT_j(x)| > d > 1 \quad \text{for all } x \in [0, x_k) \cup (y_k, 1] \text{ and some constant } d. \end{aligned} \tag{2.8}$$

Fix a $\kappa \in \mathbb{N}$ for which (2.7) and (2.8) hold. We introduce some notation. Let $t \in \Sigma$ be such that $t \neq g$, and define

$$\begin{aligned} C &= [\underbrace{g \cdots g}_{\kappa \text{ times}} t] = [g^\kappa t], \\ J_0 &= (x_\kappa, x'_\kappa), \quad J_1 = (y'_\kappa, y_\kappa), \quad J = J_0 \cup J_1, \\ Y &= C \times J. \end{aligned}$$

Lemma 2.2.2. *The set Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \lambda$.*

Proof. For $m_{\boldsymbol{p}}$-almost all $\omega \in \Sigma^{\mathbb{N}}$ we have $\tau^n \omega \in [g]$ for infinitely many $n \in \mathbb{N}$. For any such n and each $x \in (0,c) \cup (c,1)$ either $T_\omega^n(x) \in J$ or $T_\omega^n(x) \notin J$. If $T_\omega^n(x) \in (0, x_\kappa] \cup [y_\kappa, 1)$, then it follows from (2.8) that there is an $m \geq 1$ such that $T_\omega^{n+m}(x) \in J$. If $T_\omega^n(x) \in [x'_\kappa, c) \cup (c, y'_\kappa]$ it follows from (2.7) that $T_\omega^{n+1}(x) = T_g \circ T_\omega^n(x) \in (0, x'_\kappa] \cup [y'_\kappa, 1)$, which means that we are in the first case if $T_\omega^{n+1}(x) \notin J$. Hence, there exists a measurable set $A \subseteq \Sigma^{\mathbb{N}} \times [0,1]$ with $m_{\boldsymbol{p}} \times \lambda(A) = 1$ such that for each $(\omega, x) \in A$ we have $T_\omega^n(x) \in J$ for infinitely many $n \in \mathbb{N}$.

We define

$$\mathcal{E} = A \setminus \bigcup_{n=0}^{\infty} F^{-n} Y$$

and for each $x \in [0,1]$ we define

$$\mathcal{E}_x = \Big\{ \omega \in \Sigma^{\mathbb{N}} : (\omega, x) \in A \setminus \bigcup_{n=0}^{\infty} F^{-n} Y \Big\},$$

which is the x-section of $\mathcal{E}$. It follows from Fubini's Theorem that $\mathcal{E}_x$ is measurable for λ-almost all $x \in [0,1]$ and that

$$m_{\boldsymbol{p}} \times \lambda(\mathcal{E}) = \int_{[0,1]} m_{\boldsymbol{p}}(\mathcal{E}_x) d\lambda(x).$$

Combining this with $m_{\boldsymbol{p}} \times \lambda(A) = 1$, it remains to show that $m_{\boldsymbol{p}}(\mathcal{E}_x) = 0$ holds for λ-almost all $x \in [0,1]$ for which $\mathcal{E}_x$ is measurable.

Let $x \in [0,1]$ for which $\mathcal{E}_x$ is measurable. According to the Lebesgue Differentiation Theorem (see e.g. [T04]) we have that $m_{\boldsymbol{p}}$-almost all $\omega \in \Sigma^{\mathbb{N}}$ is a Lebesgue point of the function $1_{\mathcal{E}_x}$. Consider such an ω and suppose that $\omega \in \mathcal{E}_x$. Then $(\omega, x) \in A$, so there exists an increasing sequence $(n_j)_{j \in \mathbb{N}}$ in $\mathbb{N}$ that satisfies $T_\omega^{n_j}(x) \in J$ for each $j \in \mathbb{N}$. Recall that τ denotes the left shift on sequences. If $\omega' \in \tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]$, then $T_{\omega'}^{n_j}(x) = T_\omega^{n_j}(x) \in J$ and so $F^{n_j}(\omega', x) \in Y$, which gives $\omega' \notin \mathcal{E}_x$. So $\mathcal{E}_x$ and $\tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]$ are disjoint for each $j \in \mathbb{N}$, which together with ω being a Lebesgue point of $1_{\mathcal{E}_x}$ yields that

$$\begin{aligned} 1 \ &\geq \ \frac{m_{\boldsymbol{p}}\big((\mathcal{E}_x \cup \tau^{-n_j} C) \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} \\ &= \ \frac{m_{\boldsymbol{p}}\big(\mathcal{E}_x \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} + \frac{m_{\boldsymbol{p}}\big(\tau^{-n_j} C \cap [\omega_1 \cdots \omega_{n_j}]\big)}{m_{\boldsymbol{p}}\big([\omega_1 \cdots \omega_{n_j}]\big)} \xrightarrow{j \to \infty} 1_{\mathcal{E}_x}(\omega) + m_{\boldsymbol{p}}(C). \end{aligned}$$

Since $m_{\boldsymbol{p}}(C) > 0$, we find that $\omega \in \mathcal{E}_x$ gives a contradiction. We conclude that $m_{\boldsymbol{p}}(\mathcal{E}_x) = 0$. □

Since F is non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$, it follows from Lemma 2.2.2 that in particular

$$m_{\boldsymbol{p}} \times \lambda\Big(Y \setminus \bigcup_{n=1}^{\infty} F^{-n}Y\Big) \leq m_{\boldsymbol{p}} \times \lambda\Big(F^{-1}\Big(X \setminus \bigcup_{n=0}^{\infty} F^{-n}Y\Big)\Big) = 0.$$

Hence, the first return time map φ_Y of the form as in (1.6) and the induced transformation F_Y are well defined on the full measure subset of points in Y that return to Y infinitely often under iterations of F, which we call Y again. The set of points in Y that return to Y after n iterations of F can be described as

$$Y \cap F^{-n}(Y) = \bigcup_{\omega \in C \cap \tau^{-n}C} [\omega_1 \cdots \omega_n] \times (T_\omega^n|_J)^{-1}(J) \mod m_{\boldsymbol{p}} \times \lambda, \tag{2.9}$$

which is empty for $n \leq \kappa$. Note that in (2.9) in fact $[\omega_1 \cdots \omega_n] = [g^\kappa t \omega_{\kappa+2} \cdots \omega_n g^\kappa t]$ and that by construction each map $T_\omega^n|_J$ in (2.9) consists of branches that all have range $(0, c)$ or $(c, 1)$ or $(0, 1)$, since any branch of $T_\omega^\kappa|_J$ maps onto $(0, 1)$. Therefore, $Y \cap F^{-n}(Y)$ can be written as a finite union of products $A = [\boldsymbol{u} g^\kappa t] \times I$ of cylinders $[\boldsymbol{u} g^\kappa t] \subseteq C$ with $|\boldsymbol{u}| = n$ and open intervals $I \subseteq J$, each of which is mapped under F^n onto $C \times J_0$ or $C \times J_1$. Call the collection of these sets P_n and let $\alpha = \bigcup_{n > \kappa} P_n$. Let $m_{\boldsymbol{p},C}$ and λ_J denote the normalised restrictions of $m_{\boldsymbol{p}}$ to C and λ to J respectively.

Lemma 2.2.3.

(1) The collection α forms a countable first return time partition *of Y, i.e. $m_{\boldsymbol{p},C} \times \lambda_J(\bigcup_{A \in \alpha} A) = 1$, any two different sets $A, A' \in \alpha$ are disjoint and on any $A \in \alpha$ the first return time map φ_Y is constant.*

(2) Let π denote the canonical projection of $\Sigma^{\mathbb{N}} \times [0, 1]$ onto the second coordinate. Any $x \in J$ is contained in a set $\pi(A)$ for some set $A \in \alpha$.

Proof. The fact that $m_{\boldsymbol{p},C} \times \lambda_J(\bigcup_{A \in \alpha} A) = 1$ follows from Lemma 2.2.2. Furthermore, it is clear from the construction that the first return time map φ_Y is constant on any element of α once we know that any two distinct elements of α are disjoint. To show the latter, note that for $A, A' \in P_n$ this is clear. Suppose there are $1 \leq m < n$, $A = [\boldsymbol{u} g^\kappa t] \times I \in P_n$ and $A' = [\boldsymbol{v} g^\kappa t] \times I' \in P_m$ such that $A \cap A' \neq \emptyset$. Since $t \neq g$ we get $n \geq m + \kappa + 1$ and $[\boldsymbol{u} g^\kappa t] = [g^\kappa t v_{\kappa+2} \cdots v_m g^\kappa t u_{m+\kappa+2} \cdots u_n g^\kappa t]$. Moreover, $I \cap \partial I' \neq \emptyset$ or $I = I'$. In both cases, note that $F^{m+\kappa+1}([\boldsymbol{v} g^\kappa t] \times \partial I') \subseteq \Sigma^{\mathbb{N}} \times \{0, 1\}$, so by (G1) and (B1) also $F^n([\boldsymbol{v} g^\kappa t] \times \partial I') \subseteq \Sigma^{\mathbb{N}} \times \{0, 1\}$, contradicting that $F^n(A) \subseteq Y$. This proves (1).

For (2) note that, since α is a partition of Y, for each $x \in J$ it holds that there is an $A = [\boldsymbol{u} g^\kappa t] \times I \in \alpha$ with $x \in I$ or $x \in \partial I$. In the first case there is nothing to prove, so assume that $x \in \partial I$. Then $T_{\boldsymbol{u}}(x) \in \partial J_i$ for some $i \in \{0, 1\}$. From the first part of the proof of Lemma 2.2.2 it then follows that there is an $n > |\boldsymbol{u}|$ and an $\omega \in C$ such that $T_\omega^n(x) \in J$. If we write I' for the interval in $(T_\omega^n)^{-1}(J)$ containing x, then this means that there exists a set $A' = [\boldsymbol{v} g^\kappa t] \times I' \in \alpha$ with $x \in \pi(A')$. □

The second part of Lemma 2.2.3 shows that even though the partition elements of α are disjoint, their projections on the second coordinate are not. The same is true for the first coordinate as the same string $\boldsymbol{u}$ can lead points in J to J_0 and J_1.

§2.2.2 Properties of the induced transformation

It follows from (2.9) and Lemma 2.2.3 that for each $A \in \alpha$ we have either $F_Y(A) = C \times J_0$ or $F_Y(A) = C \times J_1$. For any $[\boldsymbol{u}g^\kappa t] \times I \in \alpha$, the transformation $T_{\boldsymbol{u}}|_I$ is invertible from I to one of the sets J_0 or J_1. Define the operator $\mathcal{P}_{\boldsymbol{u},I} : L^1(J, \lambda_J) \to L^1(J, \lambda_J)$ by

$$\mathcal{P}_{\boldsymbol{u},I}h(x) = \begin{cases} \dfrac{h(T_{\boldsymbol{u}}|_I^{-1}(x))}{\left|DT_{\boldsymbol{u}}|_I(T_{\boldsymbol{u}}|_I^{-1}(x))\right|}, & \text{if } T_{\boldsymbol{u}}|_I^{-1}\{x\} \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases}$$

The *random Perron-Frobenius-type operator* $\mathcal{P}_Y : L^1(J, \lambda_J) \to L^1(J, \lambda_J)$ is given by

$$\mathcal{P}_Y = \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}])\mathcal{P}_{\boldsymbol{u},I}. \tag{2.10}$$

Note that $\mathcal{P}_Y$ is not exactly of the same form as the usual Perron-Frobenius operator in (1.20). Nonetheless, we have the following result.

Lemma 2.2.4. *If $\varphi \in L^1(J, \lambda_J)$ is a fixed point of $\mathcal{P}_Y$, then the measure $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi d\lambda_J$ is invariant for F_Y.*

Proof. Suppose $\varphi \in L^1(J, \lambda_J)$ is a fixed point of $\mathcal{P}_Y$. For each cylinder $K \subseteq C$ and each Borel set $E \subseteq J$ we have

$$\begin{aligned} m_{\boldsymbol{p},C} \times \nu\big(F_Y^{-1}(K \times E)\big) &= \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}g^\kappa t] \cap \tau^{-|\boldsymbol{u}|}K)\nu(I \cap T_{\boldsymbol{u}}^{-1}E) \\ &= m_{\boldsymbol{p},C}(K) \sum_{[\boldsymbol{u}g^\kappa t] \times I \in \alpha} m_{\boldsymbol{p},C}([\boldsymbol{u}]) \int_E \mathcal{P}_{\boldsymbol{u},I}\varphi d\lambda_J \\ &= m_{\boldsymbol{p},C}(K) \int_E \mathcal{P}_Y \varphi d\lambda_J \\ &= m_{\boldsymbol{p},C} \times \nu(K \times E). \end{aligned}$$

This gives the result. □

In Lemma 2.2.5 below we show that a fixed point of $\mathcal{P}_Y$ exists. For $m \in \mathbb{N}$, set $\alpha_m = \bigvee_{j=0}^{m-1} F_Y^{-j}\alpha$. Atoms of this partition are the *m-cylinders of F_Y*. Introducing for each $Z = \bigcap_{j=0}^{m-1} F_Y^{-j}([\boldsymbol{u}_j g^\kappa t] \times I_j)$ in α_m the notation

$$C_Z = \bigcap_{j=0}^{m-1} \tau^{-\sum_{i=0}^{j-1}|\boldsymbol{u}_i|}[\boldsymbol{u}_j g^\kappa t] \quad \text{and} \quad J_Z = \bigcap_{j=0}^{m-1} T^{-1}_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{j-1}}(I_j), \tag{2.11}$$

we obtain $Z = C_Z \times J_Z$. Writing $\tau_Z = \tau^{\sum_{i=0}^{m-1}|\boldsymbol{u}_i|}|_{C_Z}$ and $T_Z = T_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{m-1}}|_{J_Z}$ we have $F_Y^m|_Z = \tau_Z \times T_Z$. Each T_Z has non-positive Schwarzian derivative, so we can apply the Koebe Principle. The image $T_Z(J_Z)$ either equals J_0 or J_1. Choose a $\bar{\rho} > 0$ such that $J_0' := [x_\kappa - \bar{\rho}, x_\kappa' + \bar{\rho}] \subseteq (0, c)$ and $J_1' := [y_\kappa' - \bar{\rho}, y_\kappa + \bar{\rho}] \subseteq (c, 1)$. There is a canonical way to extend the domain of each T_Z to an interval J_Z' containing J_Z, such that $T_Z(J_Z')$ equals either J_0' or J_1' and $\boldsymbol{S}(T_Z) \leq 0$ on J_Z'. Then by the Koebe Principle, i.e. (1.14) and (1.15), there exist constants $K^{(\bar{\rho})} > 1$ and $M^{(\bar{\rho})} > 0$ such that for all $m \in \mathbb{N}$, $Z \in \alpha_m$ and $x, y \in J_Z$,

$$\frac{1}{K^{(\bar{\rho})}} \leq \frac{DT_Z(x)}{DT_Z(y)} \leq K^{(\bar{\rho})}, \tag{2.12}$$

$$\Big|\frac{DT_Z(x)}{DT_Z(y)} - 1\Big| \leq \frac{M^{(\bar{\rho})}}{\min\{\lambda(J_0), \lambda(J_1)\}} \cdot |T_Z(x) - T_Z(y)|. \tag{2.13}$$

Note that for the random Perron-Frobenius-type operator from (2.10) we have for each $m \geq 1$ that

$$\mathcal{P}_Y^m = \frac{1}{m_{\boldsymbol{p}}(C)} \sum_{Z \in \alpha_m} m_{\boldsymbol{p},C}(C_Z)\mathcal{P}_{T_Z}, \tag{2.14}$$

where $\mathcal{P}_{T_Z}$ is of the form as in (1.10).

Lemma 2.2.5 (cf. Lemmas V.2.1 and V.2.2 of [dMvS93]). *$\mathcal{P}_Y$ admits a fixed point $\varphi \in L^1(J, \lambda_J)$ that is bounded, Lipschitz and bounded away from zero.*

Proof. For each $m \in \mathbb{N}$ and $x \in J$,

$$\mathcal{P}_Y^m 1(x) = \frac{1}{m_{\boldsymbol{p}}(C)} \sum_{\substack{Z \in \alpha_m: \\ x \in T_Z(J_Z)}} \frac{m_{\boldsymbol{p},C}(C_Z)}{|DT_Z(T_Z^{-1}x)|}.$$

Using the Mean Value Theorem, for all $m \in \mathbb{N}$ and $Z \in \alpha_m$ there exists a $\xi \in J_Z$ such that

$$\frac{\lambda\big(T_Z(J_Z)\big)}{\lambda(J_Z)} = |DT_Z(\xi)|. \tag{2.15}$$

Set $K_1 = \frac{\max\{K^{(\bar{\rho})}, M^{(\bar{\rho})}\}}{m_{\boldsymbol{p}}(C)\cdot\min\{\lambda(J_0),\lambda(J_1)\}}$, where $\bar{\rho}$ is as in (2.12) and (2.13). Since $DT_Z(\xi)$ and $DT_Z(y)$ have the same sign for any $y \in J_Z$, (2.15) together with (2.12) implies

$$\mathcal{P}_Y^m 1(x) \leq \sum_{Z \in \alpha_m} \frac{m_{\boldsymbol{p},C}(C_Z)}{m_{\boldsymbol{p}}(C)} \cdot K^{(\bar{\rho})} \frac{\lambda(J_Z)}{\lambda(T_Z(J_Z))} \leq K_1 \sum_{Z \in \alpha_m} m_{\boldsymbol{p},C} \times \lambda_J(C_Z \times J_Z) = K_1. \tag{2.16}$$

Moreover, if for $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ we take $x, y \in I$, then for any $Z \in \alpha_m$ it holds that $x \in T_Z(J_Z)$ if and only if $y \in T_Z(J_Z)$. For such Z, let $x_Z, y_Z \in J_Z$ be such that

$T_Z(x_Z) = x$ and $T_Z(y_Z) = y$. Then by (2.13)

$$\begin{aligned} |\mathcal{P}_Y^m 1(x) - \mathcal{P}_Y^m 1(y)| &\le \sum_{\substack{Z \in \alpha_m: \\ x \in T_Z(J_Z)}} \frac{m_{\boldsymbol{p},C}(C_Z)}{m_{\boldsymbol{p}}(C)} \left| \frac{1}{|DT_Z(x_Z)|} - \frac{1}{|DT_Z(y_Z)|} \right| \\ &\le \sum_{\substack{Z \in \alpha_m: \\ x \in T_Z(J_Z)}} m_{\boldsymbol{p},C}(C_Z) \frac{1}{|DT_Z(x_Z)|} K_1 |T_Z(x_Z) - T_Z(y_Z)| \\ &= K_1 \mathcal{P}_Y^m 1(x)|x-y|. \end{aligned} \tag{2.17}$$

Together (2.16) and (2.17) imply that the sequence $\big(\frac{1}{m}\sum_{j=0}^{m-1}\mathcal{P}_Y^j 1\big)_m$ is uniformly bounded and equicontinuous on I for each $A = [\boldsymbol{u}g^\kappa t] \times I$. By Lemma 2.2.3(2) it follows that the same holds on J. Hence, by the Arzela-Ascoli Theorem there exists a subsequence

$$\Big(\frac{1}{m_k}\sum_{j=0}^{m_k-1}\mathcal{P}_Y^j 1\Big)_{m_k}$$

converging uniformly to a function $\varphi : J \to [0,\infty)$ satisfying $\varphi \le K_1$ and for each $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ and $x, y \in I$,

$$|\varphi(x) - \varphi(y)| \le K_1 \varphi(x)|x-y|. \tag{2.18}$$

Hence, φ is bounded and by Lemma 2.2.3(2) it is clear that φ is Lipschitz (with Lipschitz constant bounded by K_1^2). It is readily checked that φ is a fixed point of $\mathcal{P}_Y$, so that $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi\, d\lambda$ is an invariant probability measure for F_Y.

Because of Lemma 2.2.3(2) and because φ is continuous, in order to obtain that φ is bounded away from zero on J it suffices to verify for each $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ that φ on I is bounded away from zero. Suppose that there is $A = [\boldsymbol{u}g^\kappa t] \times I \in \alpha$ for which $\inf_{x\in I}\varphi(x) = 0$. Then from (2.18) it follows that $\varphi(y) = 0$ for all $y \in I$, hence $\nu(I) = 0$. Either $I \subseteq J_0$ or $I \subseteq J_1$. If $I \subseteq J_0$, then for any set $A' = [\boldsymbol{v}g^\kappa t] \times I' \in \alpha$ with $T_{\boldsymbol{v}}(I') = J_0$ it holds that

$$m_{\boldsymbol{p},C} \times \lambda_J(A' \cap F_Y^{-1}A) > 0$$

and, by the F_Y-invariance of $m_{\boldsymbol{p},C} \times \nu$,

$$m_{\boldsymbol{p},C} \times \nu(A' \cap F_Y^{-1}A) \le m_{\boldsymbol{p},C} \times \nu(F_Y^{-1}A) = m_{\boldsymbol{p},C} \times \nu(A) = 0,$$

which together give $\inf_{x\in I'}\varphi(x) = 0$ and therefore, like before, $\nu(I') = 0$. There are sets $A' = [\boldsymbol{v}g^\kappa t] \times I'$ with $I' \subseteq J_1$ and $T_{\boldsymbol{v}}(I') = J_0$, so we can repeat the argument to show that also for any set $A'' = [\boldsymbol{v}g^\kappa t] \times I'' \in \alpha$ with $T_{\boldsymbol{v}}(I'') = J_1$ we have $\nu(I'') = 0$. So $m_{\boldsymbol{p},C} \times \nu(A) = 0$ for all $A \in \alpha$. If $I \subseteq J_1$ we come to the same conclusion. This gives a contradiction, so φ is bounded from below on each interval I. □

It follows from Lemma 2.2.4 that $m_{\boldsymbol{p},C} \times \nu$ with $\nu = \varphi d\lambda_J$ is a finite F_Y-invariant measure. To show that $m_{\boldsymbol{p},C} \times \lambda_J$ is F_Y-ergodic we need the following result, which states that the sets $\pi(A)$ for $A \in \alpha_m$ shrink uniformly to λ-null sets as $m \to \infty$.

Lemma 2.2.6. $\lim_{m\to\infty} \sup\{\lambda_J(J_Z) : Z \in \alpha_m\} = 0.$

Proof. Set $\delta = \sup\{\lambda_J(J_Z) : Z \in \alpha\} < 1$. Let $m \geq 2$ and $Z = \bigcap_{j=0}^{m-1} F_Y^{-j}([\boldsymbol{u}_j g^\kappa t] \times I_j) = C_Z \times J_Z \in \alpha_m$ as in (2.11). Set

$$\tilde{J}_Z = \bigcap_{j=0}^{m-2} T^{-1}_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{j-1}}(I_j),$$

so that $J_Z = \tilde{J}_Z \cap T^{-1}_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(I_{m-1})$. Let J_i, $i \in \{0,1\}$, be such that $T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z) = J_i$. It holds that $T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(J_Z) = I_{m-1}$, so $\lambda(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(J_Z)) \leq \delta$ and thus

$$\lambda(T_{\boldsymbol{u}_0\boldsymbol{u}_1\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z \setminus J_Z) \geq \lambda(J_i) - \delta.$$

Since $\tilde{J}_Z \setminus J_Z$ consists of at most two intervals, with (2.12) and (1.12) applied to this setting this gives

$$1 - \frac{\lambda_J(J_Z)}{\lambda_J(\tilde{J}_Z)} = \frac{\lambda_J(\tilde{J}_Z \setminus J_Z)}{\lambda_J(\tilde{J}_Z)} \geq \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z \setminus J_Z)}{\lambda_J(T_{\boldsymbol{u}_0\cdots\boldsymbol{u}_{m-2}}(\tilde{J}_Z))} \geq \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(J_i) - \delta}{\lambda_J(J_i)}.$$

Set $K_1 := \max\left\{1 - \frac{1}{K^{(\bar{\rho})}} \frac{\lambda_J(J_i)-\delta}{\lambda_J(J_i)} : i = 0,1\right\} \in (0,1)$. Then by repeating the same steps, we obtain

$$\lambda_J(J_Z) \leq K_1 \lambda_J(\tilde{J}_Z) \leq \cdots \leq K_1^{m-1} \lambda_J(I_0) < K_1^{m-1},$$

which proves the lemma. □

Lemma 2.2.7. *The measure $m_{\boldsymbol{p},C} \times \lambda_J$ is F_Y-ergodic.*

Proof. Suppose $E \subseteq Y$ with $m_{\boldsymbol{p},C} \times \lambda_J(E) > 0$ satisfies $F_Y^{-1}E = E \mod m_{\boldsymbol{p},C} \times \lambda_J$. We show that $m_{\boldsymbol{p},C} \times \lambda_J(E) = 1$. The Borel measure ρ on Y given by

$$\rho(V) = \int_V 1_E(\omega, x)\varphi(x) dm_{\boldsymbol{p},C}(\omega) d\lambda_J(x)$$

for Borel sets V is F_Y-invariant. According to Proposition 1.2.12 and Lemma 1.4.1 this yields a stationary measure $\tilde{\mu}$ on $[0,1]$ that is absolutely continuous w.r.t. λ and satisfies $(m_{\boldsymbol{p}} \times \tilde{\mu})|_Y = \rho$. Let $L := \{x \in J : \frac{d\tilde{\mu}}{d\lambda}(x) > 0\}$ denote the support of $\frac{d\tilde{\mu}}{d\lambda}|_J$. Since

$$m_{\boldsymbol{p}} \times \lambda(Y) \cdot \frac{d\tilde{\mu}}{d\lambda}(x) = \frac{d\rho}{dm_{\boldsymbol{p},C} \times \lambda_J}(\omega, x) = 1_E(\omega,x)\varphi(x), \qquad m_{\boldsymbol{p},C} \times \lambda_J\text{-a.e.}$$

we obtain, using that φ is bounded away from zero,

$$E = C \times L \mod m_{\boldsymbol{p},C} \times \lambda_J.$$

To obtain the result, it remains to show that $\lambda_J(J \setminus L) = 0$.

We have $C \times L = \bigcup_{Z \in \alpha_m} C_Z \times (J_Z \cap L)$ and $F_Y^{-m}(C \times L) = \bigcup_{Z \in \alpha_m} C_Z \times T_Z^{-1} L$. From the non-singularity of F_Y w.r.t. $m_{\boldsymbol{p},C} \times \lambda_J$ it follows that for each $m \in \mathbb{N}$,

$$C \times L = E = F_Y^{-m} E = F_Y^{-m}(C \times L) \mod m_{\boldsymbol{p},C} \times \lambda_J,$$

which yields

$$J_Z \cap L = T_Z^{-1} L \mod \lambda_J, \quad \text{for each } Z \in \alpha_m. \tag{2.19}$$

Let $\varepsilon > 0$. Since $\lambda_J(L) > 0$, it follows from Lemma 2.2.6 and the Lebesgue Density Theorem that there are $i \in \{0, 1\}$, $m' \in \mathbb{N}$ and $Z' \in \alpha_{m'}$ such that

$$T_{Z'}(J_{Z'}) = J_i \quad \text{and} \quad \lambda_J(J_{Z'} \cap L) \geq (1 - \varepsilon)\lambda_J(J_{Z'}).$$

By (2.19), $T_{Z'}^{-1}(J_i \backslash L) = J_{Z'} \backslash L \mod \lambda_J$. The Mean Value Theorem gives the existence of a $\xi \in J_{Z'}$ such that

$$\frac{\lambda_J(T_{Z'}(J_{Z'}))}{\lambda_J(J_{Z'})} = |DT_{Z'}(\xi)|,$$

and from (2.12) it follows that

$$\lambda_J(T_{Z'}(J_{Z'} \backslash L)) = \int_{J_{Z'} \backslash L} |DT_{Z'}| d\lambda \leq K^{(\bar{\rho})} |DT_{Z'}(\xi)| \lambda_J(J_{Z'} \backslash L).$$

Hence,

$$\frac{\lambda_J(J_i \backslash L)}{\lambda_J(J_i)} = \frac{\lambda_J(T_{Z'}(J_{Z'} \backslash L))}{\lambda_J(T_{Z'}(J_{Z'}))} \leq K^{(\bar{\rho})} \frac{\lambda_J(J_{Z'} \backslash L)}{\lambda_J(J_{Z'})} \leq K^{(\bar{\rho})} \varepsilon. \tag{2.20}$$

So, for each $\varepsilon > 0$ we can find $i = i(\varepsilon)$, $Z' = Z'(\varepsilon)$ and $m' = m'(\varepsilon)$ for which (2.20) holds. If for each $\varepsilon_0 > 0$ and each $i_0 \in \{0, 1\}$ there exists an $\varepsilon \in (0, \varepsilon_0)$ such that $i(\varepsilon) = i_0$, we obtain from (2.20) that $\lambda_J(J \backslash L) = 0$. Otherwise, there exists $\varepsilon_0 > 0$ and $i_0 \in \{0, 1\}$ such that $i(\varepsilon) = i_0$ for all $\varepsilon \in (0, \varepsilon_0)$. Without loss of generality, suppose that $i_0 = 0$. Then (2.20) gives $\lambda_J(J_0 \backslash L) = 0$. By the equivalence of ν and λ_J and the fact that every good map has full branches it follows that

$$m_{\boldsymbol{p},C} \times \nu\big((C \times J_0) \cap F_Y^{-1}(C \times J_1)\big) > 0.$$

Together with the Poincaré Recurrence Theorem this gives that

$$A = \{(\omega, x) \in C \times J_0 : F_Y^m(\omega, x) \in C \times J_1 \text{ for infinitely many } m \in \mathbb{N}\}$$

satisfies $m_{\boldsymbol{p},C} \times \nu(A) > 0$, and therefore $m_{\boldsymbol{p},C} \times \lambda_J(A) > 0$. Together with $\lambda_J(J_0 \backslash L) = 0$ it follows from the Lebesgue Density Theorem that there exists a Lebesgue point $x \in \pi(A) \cap L$ of $1_{\pi(A) \cap L}$. Since $x \in \pi(A)$, for infinitely many $m \in \mathbb{N}$ there exists $Z_m \in \alpha_m$ such that $x \in J_{Z_m}$ and $T_{Z_m}(J_{Z_m}) = J_1$. This again together with Lemma 2.2.6 yields that for each $\varepsilon > 0$ there exist $m \in \mathbb{N}$ and $Z \in \alpha_m$ such that

$$T_Z(J_Z) = J_1 \quad \text{and} \quad \lambda_J(J_Z \cap L) \geq (1 - \varepsilon)\lambda_J(J_Z).$$

Similar as before, this gives $\lambda_J(J_1 \backslash L) = 0$, so $\lambda_J(J \backslash L) = 0$. □

§2.2.3 Proof of Theorem 2.1.2

We will first give the proof of Lemma 2.1.6.

Proof of Lemma 2.1.6. It follows from (B3) that for any $j \in \Sigma_B$ and $x \in [0,1]$,

$$|T_j(x) - c| = |T_j(x) - T_j(c)| = \left| \int_c^x DT_j(y)dy \right| \geq \frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} |x-c|^{\ell_j}.$$

By induction we get that for each $n \in \mathbb{N}$ and $\omega \in \Sigma_B^{\mathbb{N}}$,

$$|T_\omega^n(x) - c| \geq \left(\frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} \right)^{1+\sum_{i=0}^{n-2} \ell_{\omega_n} \cdots \ell_{\omega_{n-i}}} \cdot |x-c|^{\ell_{\omega_1} \cdots \ell_{\omega_n}}.$$

From (B3) we see that $\frac{\min\{K_b \,:\, b \in \Sigma_B\}}{\ell_{\max}} < 1$. The lower bound now follows by observing that

$$\Big(1 + \sum_{i=0}^{n-2} \ell_{\omega_n} \cdots \ell_{\omega_{n-i}}\Big) / (\ell_{\omega_1} \cdots \ell_{\omega_n}) \leq \sum_{i=1}^{n} \frac{1}{\ell_{\min}^i} < \frac{1}{\ell_{\min} - 1}.$$

The result for the upper bound follows similarly, by noticing that in this case from (B3) it follows that $\frac{\max\{M_b \,:\, b \in \Sigma_B\}}{\ell_{\min}} > 1$. □

Together with the results from the previous two subsections we now collected all the ingredients necessary to prove Theorem 2.1.2.

Proof of Theorem 2.1.2. (1) We have constructed a finite F_Y-invariant measure $m_{\boldsymbol{p},C} \times \nu$ which is absolutely continuous with respect to $m_{\boldsymbol{p},C} \times \lambda_J$. Since F is non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$ we can therefore by Proposition 1.2.12 extend $m_{\boldsymbol{p},C} \times \nu$ to a σ-finite F-invariant measure $m_{\boldsymbol{p}} \times \mu$ that is absolutely continuous with respect to $m_{\boldsymbol{p}} \times \lambda$. According to Theorem 1.2.10 what is left to show is that F is conservative and ergodic w.r.t. $m_{\boldsymbol{p}} \times \lambda$.

Since $\mu \ll \lambda$ it follows from Lemma 2.2.2 that Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \mu$. Maharam's Recurrence Theorem then gives that F is conservative with respect to $m_{\boldsymbol{p}} \times \mu$. Furthermore, in the proof of part (2) below we will see that the density of $\frac{d\mu}{d\lambda}$ is bounded away from zero. Hence, $\lambda \ll \mu$ and therefore F is also conservative with respect to $m_{\boldsymbol{p}} \times \lambda$. Furthermore, combining the ergodicity of F_Y with respect to $m_{\boldsymbol{p},C} \times \lambda_J$ and the fact that Y is a sweep-out set for F with respect to $m_{\boldsymbol{p}} \times \lambda$ gives using Proposition 1.2.11 that F is ergodic with respect to $m_{\boldsymbol{p}} \times \lambda$.

(2) For the density $\psi := \frac{d\mu}{d\lambda}$ it holds that $\psi|_J = \varphi$. Since we can take κ in the definition of J as large as we want, ψ is locally Lipschitz on $(0,c)$ and $(c,1)$. Moreover, it is a fixed point of the Perron-Frobenius operator $\mathcal{P}_{\mathcal{T},\boldsymbol{p}}$ being of the form as in (1.20) and thus for all $x \in (0,1)$,

$$\psi(x) = \mathcal{P}_{\mathcal{T},\boldsymbol{p}}^\kappa \psi(x) \geq p_g^\kappa \sum_{y \in J \cap T_g^{-\kappa}\{x\}} \frac{\varphi(y)}{|DT_g^\kappa(y)|}.$$

From Lemma 2.2.5 we conclude that ψ is bounded from below by some constant $C > 0$. It remains to show that ψ is not in L^q for any $q > 1$. To see this, fix a $b \in \Sigma_B$. Since ψ is bounded from below by $C > 0$, we have for all $k \in \mathbb{N}_0$ and $x \in [0,1]$ that

$$\psi(x) = \mathcal{P}^{k+1}_{\mathcal{T},\boldsymbol{p}}\psi(x) \geq C \cdot p_g p_b^k \sum_{y \in (T_g T_b^k)^{-1}\{x\}} \frac{1}{|D(T_g T_b^k)(y)|}. \tag{2.21}$$

Let $\ell_b, M_b, r_g, M_g, K_g$ be as in (B3) and (G3). From (B3), (G3) and Lemma 2.1.6 we get

$$\begin{aligned}
|D(T_g T_b^k)(y)| &= |DT_g(T_b^k(y))| \prod_{i=1}^{k} |DT_b(T_b^{k-i}(y))| \\
&\leq M_g |T_b^k(y) - c|^{r_g - 1} \prod_{i=0}^{k-1} (M_b |T_b^i(y) - c|^{\ell_b - 1}) \\
&\leq M_g M_b^k (\tilde{M}|y-c|)^{\ell_b^k (r_g - 1)} \prod_{i=0}^{k-1} (\tilde{M}|y-c|)^{\ell_b^i(\ell_b - 1)} \\
&= K_1 |y - c|^{\ell_b^k r_g - 1},
\end{aligned} \tag{2.22}$$

for the positive constant $K_1 = M_g M_b^k \tilde{M}^{\ell_b^k r_g - 1}$. On the other hand, from (G3) we obtain for any $y \in (T_g T_b^k)^{-1}\{x\}$ as in the proof of Lemma 2.1.6 that

$$|x - T_g(c)| = |T_g T_b^k(y) - T_g(c)| \geq \frac{K_g}{r_g} |T_b^k(y) - c|^{r_g}$$

and then Lemma 2.1.6 yields

$$|x - T_g(c)| \geq K_2 |y - c|^{\ell_b^k r_g} \tag{2.23}$$

for the positive constant $K_2 = \frac{K_g}{r_g}\tilde{K}^{\ell_b r_g}$. Now for any $q > 1$ we can choose $k \in \mathbb{N}_0$ large enough so that $\tau := (1 - \ell_b^{-k} r_g^{-1})q \geq 1$. Combining (2.21), (2.22) and (2.23) we obtain

$$\begin{aligned}
\psi^q(x) &\geq \Big(\frac{C p_g p_b^k}{K_1}\Big)^q \Big(\sum_{y \in (T_g T_b^k)^{-1}\{x\}} |y - c|^{1 - \ell_b^k r_g} \Big)^q \\
&\geq K_3 |x - T_g(c)|^{-\tau}
\end{aligned}$$

for a positive constant K_3. This gives the result. □

Remark 2.2.8. The result from Theorem 2.1.2 still holds if we allow the critical order ℓ_b from (B3) to be equal to 1 for some b, as long as $\ell_{\max} > 1$. To see this, note that in the proof of Theorem 2.1.2 condition (B3) only plays a role in proving that $\frac{d\mu}{d\lambda} \notin L^q$ for any $q > 1$. Here we refer to Lemma 2.1.6 and the constants $\tilde{K}$ and $\tilde{M}$, which are not well defined if $\ell_{\min} = 1$. In (2.22) however, we use the estimates from

Lemma 2.1.6 only for one arbitrary fixed $b \in \Sigma_B$. By the same reasoning as in the proof of Lemma 2.1.6 it follows that

$$\left(\left(\frac{K_b}{\ell_b}\right)^{\frac{1}{\ell_b-1}}|x-c|\right)^{\ell_b^n} \leq |T_b^n(x)-c| \leq \left(\left(\frac{M_b}{\ell_b}\right)^{\frac{1}{\ell_b-1}}|x-c|\right)^{\ell_b^n}.$$

for any $b \in \Sigma_B$ with $\ell_b > 1$. Hence, if there exists at least one $b \in \Sigma_B$ with $\ell_b > 1$, then we can replace the bounds obtained from Lemma 2.1.6 in (2.22) and (2.23) by constants $K_1 = M_g M_b^k (\frac{K_b}{\ell_b})^{(\ell_b^k r_g - 1)/(\ell_b - 1)}$ and $K_2 = \frac{K_g}{r_g}(\frac{M_b}{\ell_b})^{\ell_b r_g/(\ell_b-1)}$ and obtain the same result. In case $\ell_{\max} = 1$, then most parts from Theorem 2.1.2 still remain valid with the exception that in that case we can only say that $\frac{d\mu}{d\lambda} \notin L^q$ if $q \geq \frac{r_{\max}}{r_{\max}-1}$. This follows from the above reasoning by taking $k = 0$ in the definition of τ in the proof of Theorem 2.1.2 and by noting that $\tau = (1 - r_{\max}^{-1})q \geq 1$ if $q \geq \frac{r_{\max}}{r_{\max}-1}$.

§2.3 Estimates on the acs measure

In this section we prove Theorem 2.1.3. Recall the definition of θ from Theorem 2.1.3:

$$\theta = \sum_{b \in \Sigma_B} p_b \ell_b.$$

§2.3.1 The case $\theta \geq 1$

To prove one direction of Theorem 2.1.3, namely that the unique acs measure μ from Theorem 2.1.2 is infinite if $\theta \geq 1$, we introduce another induced transformation. Furthermore, we use that there exists a $\delta > 0$ such that $|DT_b(x)| < 1$ for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B$. This implies

$$|T_b(x) - c| < |x - c| \tag{2.24}$$

for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B$.

Proposition 2.3.1. *Suppose $\theta \geq 1$. Then the unique acs measure μ from Theorem 2.1.2 is infinite.*

Proof. Fix a $b \in \Sigma_B$. Recall the definitions of $\tilde{M}$ from Lemma 2.1.6 and set $\gamma = \min\{\delta, \frac{1}{2}\tilde{M}^{-1}\}$ with δ as given above. Let $a \in [c-\gamma, c)$. Then there exists a $\xi \in (a, c)$ such that $T_b(a) > \xi$ and $T_b^2(a) > \xi$. Take $[bb] \times (a, \xi)$ as the inducing domain and let

$$\kappa(\omega, x) = \inf\{k \in \mathbb{N} : F^k(\omega, x) \in [bb] \times (a, \xi)\}$$

be the first return time to $[bb] \times (a, \xi)$ under F. If $m_{\boldsymbol{p}} \times \mu([bb] \times (a, \xi)) = \infty$, then there is nothing left to prove. If not, then we compute $\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu$ and use Kac's Lemma, i.e. Lemma 1.2.13, to prove the result.

So, assume that $m_{\boldsymbol{p}} \times \mu([bb] \times (a, \xi)) < \infty$. The conditions that $T_b(a) > \xi$ and $T_b^2(a) > \xi$ together with the fact that any bad map has c as a fixed point and is strictly monotone on the intervals $(0,c)$ and $(c,1)$, guarantee that for each $n \in \mathbb{N}$ and $\omega \in \Sigma_B^{\mathbb{N}} \cap [bb]$ we get

$$T_\omega^n((a,\xi)) \cap (a,\xi) = \emptyset. \tag{2.25}$$

For any $\omega \in [bb]$ and $x \in (a,\xi)$ it follows by (2.25) and (2.24) that $T_\omega^n(x)$ can only return to (a,ξ) after at least one application of a good map. Assume that $\omega \in [bb]$ is of the form

$$\omega = (b, b, \omega_3, \omega_4, \dots, \omega_n, g, \omega_{n+2}, \dots),$$

with $n \geq 2$, $\omega_i \in \Sigma_B$ for $3 \leq i \leq n$, $g \in \Sigma_g$, and $x \in (a,\xi)$. Then $\kappa(\omega, x) \geq n+1$. Lemma 2.1.6 yields that

$$|T_\omega^n(x) - c| \leq (\tilde{M}\gamma)^{\ell_{\omega_1}\cdots\ell_{\omega_n}} < 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n}}. \tag{2.26}$$

From (G3) and (2.26) we obtain that

$$|T_g T_\omega^n(x) - T_g(c)| = \left| \int_c^{T_\omega^n(x)} DT_g(y)\, dy \right| \leq \frac{M_g}{r_g} |T_\omega^n(x) - c|^{r_g} < \frac{M_g}{r_g} \cdot 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n} r_g}. \tag{2.27}$$

Set

$$\zeta = \sup\{|DT_j(x)| \,:\, j \in \Sigma,\, x \in [0,1]\}.$$

Then $\zeta > 1$ by (G4), (B4). Assume $\kappa(\omega, x) = m + n$ for some $m \geq 1$. Then $T_\omega^{m+n}(x) \in (a,\xi)$ so that by (G1),

$$|T_\omega^{m+n}(x) - T_g(c)| \geq \min\{a, 1-\xi\}.$$

Because of (2.27) this implies

$$\zeta^{m-1} \frac{M_g}{r_g} \cdot 2^{-\ell_{\omega_1}\cdots\ell_{\omega_n} r_g} \geq \min\{a, 1-\xi\}.$$

Solving for m yields

$$m \geq K_1 + K_2 \ell_{\omega_1} \cdots \ell_{\omega_n}$$

for constants $K_1 = 1 + \log\left(\frac{\min\{a,1-\xi\} r_g}{M_g}\right) / \log\zeta \in \mathbb{R}$ and $K_2 = \log(2^{r_g})/\log\zeta > 0$. Note that K_1, K_2 are independent of ω, x, m and n.

We obtain that for any $g \in \Sigma_G$,

$$\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu \geq \sum_{n \in \mathbb{N}_{\geq 2}} \sum_{\omega_3,\dots,\omega_n \in \Sigma_B} m_{\boldsymbol{p}}([bb\omega_3\cdots\omega_n g])\mu((a,\xi))\Big(n + K_1 + K_2 \ell_b^2 \prod_{i=3}^n \ell_{\omega_i}\Big).$$

Since

$$\sum_{n \in \mathbb{N}_{\geq 2}} \sum_{\omega_3,\dots,\omega_n \in \Sigma_B} m_{\boldsymbol{p}}([\omega_3\cdots\omega_n]) \prod_{i=3}^n \ell_{\omega_i} = 1 + \sum_{n\in\mathbb{N}} \theta^n = \infty,$$

we get $\int_{[bb]\times(a,\xi)} \kappa \, dm_{\boldsymbol{p}} \times \mu = \infty$ and from Lemma 1.2.13 we now conclude that μ is infinite. □

§2.3.2 The case $\theta < 1$

For the other direction of Theorem 2.1.3, assume $\theta < 1$. We first obtain a stationary probability measure $\tilde{\mu}$ for F as in (2.5) using a standard Krylov-Bogolyubov type argument. For this, let $\mathcal{M}$ denote the set of all finite Borel measures on $[0,1]$, and define the operator $\mathcal{P} : \mathcal{M} \to \mathcal{M}$ by

$$\mathcal{P}\nu = \sum_{j\in\Sigma} p_j \nu \circ T_j^{-1}, \qquad \nu \in \mathcal{M},$$

where $\nu \circ T_j^{-1}$ denotes the pushforward measure of ν under T_j. Then $\mathcal{P}$ is a *Markov-Feller* operator (see e.g. [LMS04]) with dual operator U on the space $BM([0,1])$ of all bounded Borel measurable functions given by[2] $Uf = \sum_{j\in\Sigma} p_j f \circ T_j$ for $f \in BM([0,1])$. As before, let λ denote the Lebesgue measure on $[0,1]$, and set $\lambda_n = \mathcal{P}^n \lambda$ for each $n \geq 0$. Furthermore, for each $n \in \mathbb{N}$ define the Cesáro mean $\mu_n = \frac{1}{n}\sum_{k=0}^{n-1} \lambda_k$. Since the space of probability measures on $[0,1]$ equipped with the weak topology is sequentially compact, there exists a subsequence $(\mu_{n_k})_{k\in\mathbb{N}}$ of $(\mu_n)_{n\in\mathbb{N}}$ that converges weakly to a probability measure $\tilde{\mu}$ on $[0,1]$. Using that a Markov-Feller operator is weakly continuous, it then follows from a standard argument that $\mathcal{P}\tilde{\mu} = \tilde{\mu}$, that is, $\tilde{\mu}$ is a stationary probability measure for F. The next theorem will lead to the estimate (2.6) from Theorem 2.1.3. For any $\boldsymbol{u} = u_1 \cdots u_k \in \Sigma^k$, $k \geq 0$, recall that we abbreviate $p_{\boldsymbol{u}} = \prod_{i=1}^k p_{u_i}$ and also let $\ell_{\boldsymbol{u}} = \prod_{i=1}^k \ell_{u_i}$ if $\boldsymbol{u} \in \Sigma_B^k$, where we use $p_{\boldsymbol{u}} = \ell_{\boldsymbol{u}} = 1$ in case $k = 0$.

Theorem 2.3.2. *There exists a constant $C > 0$ such that for all $n \in \mathbb{N}$ and all Borel sets $B \subseteq [0,1]$ we have*

$$\lambda_n(B) \leq C \cdot \sum_{g\in\Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot \lambda(B)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.$$

Before we prove this theorem, we first show how it gives Theorem 2.1.3.

Proof of Theorem 2.1.3. The first part of the statement follows from Proposition 2.3.1. For the second part, assume that $\theta < 1$ and that Theorem 2.3.2 holds. Let $B \subseteq [0,1]$ be a Borel set. Using the regularity of λ, for any $\delta > 0$ there exists an open set $G \subseteq [0,1]$ such that $B \subseteq G$ and $\lambda(G) \leq \lambda(B) + \delta$. Using that $(\mu_{n_k})_{k\in\mathbb{N}}$ converges weakly to $\tilde{\mu}$, we obtain from the Portmanteau Theorem together with Theorem 2.3.2 that

$$\begin{aligned}\tilde{\mu}(B) \leq \tilde{\mu}(G) &\leq \liminf_k \mu_{n_k}(G) \\ &\leq C \cdot \sum_{g\in\Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot (\lambda(B) + \delta)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.\end{aligned}$$

[2] By definition of a Markov-Feller operator, the space of bounded *continuous* functions is required to be invariant under the dual operator U. If there is a $g \in \Sigma_G$ for which T_g is discontinuous (namely at c), we then first identify $[0,1]$ with the unit circle S^1 so that T_g can be viewed as a continuous map on S^1. With the same identification any acs measure on S^1 then gives an acs measure on $[0,1]$.

Since $\theta < 1$, the sum is bounded and with the Dominated Convergence Theorem we can take the limit as $\delta \to 0$ to obtain

$$\tilde{\mu}(B) \le C \cdot \sum_{g \in \Sigma_G} p_g \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \cdot \lambda(B)^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}. \tag{2.28}$$

This proves that $\tilde{\mu}$ is absolutely continuous with respect to the Lebesgue measure on $[0,1]$. It follows that the probability measure $\tilde{\mu}$ is equal to the unique acs measure μ from Theorem 2.1.2. The estimate (2.6) follows directly from (2.28). □

It remains to give the proof of Theorem 2.3.2. We shall do this in a number of steps.

Proposition 2.3.3. *There exists a constant $K_1 > 0$ such that for all $n \in \mathbb{N}$, all $\boldsymbol{u} \in \Sigma^n$ and all Borel sets $A \subseteq [0,1]$ with $0 < 3\lambda(B) < \frac{1}{2}\min\{c, 1-c\}$ we have*

$$\lambda(T_{\boldsymbol{u}}^{-1}B) \le K_1\big(\lambda(T_{\boldsymbol{u}}^{-1}[0,3\eta)) + \lambda(T_{\boldsymbol{u}}^{-1}(c-3\eta, c+3\eta)) + \lambda(T_{\boldsymbol{u}}^{-1}(1-3\eta, 1])\big),$$

where $\eta = \lambda(B)$.

Proof. Let $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma^n$ and a Borel set $B \subseteq [0,1]$ with $0 < 3\lambda(B) < \frac{1}{2}\min\{c, 1-c\} < 1$ be given and write $\eta = \lambda(B)$. The map $T_{\boldsymbol{u}}$ has non-positive Schwarzian derivative on any of its intervals of monotonicity (see (1.13)) and the interior of the image of any such interval is $(0,c), (c,1)$ or $(0,1)$. Set $B_1 = (\eta, c-\eta)$ and $B_2 = (2\eta, c-2\eta)$. Let I be a connected component of $T_{\boldsymbol{u}}^{-1}B_1$, and set $f = T_{\boldsymbol{u}}|_I$ and $I^* = f^{-1}B_2$. The Minimum Principle yields

$$|Df(x)| \ge \min_{z \in \partial I^*} |Df(z)|, \qquad \text{for all } x \in I^*. \tag{2.29}$$

Suppose the minimal value is attained at $f^{-1}(2\eta)$ and set $B_3 = (2\eta, 3\eta)$ and $J = f^{-1}B_3$. By the condition on the size of B it follows from the Koebe Principle that

$$K^{(\eta)}|Df(f^{-1}(2\eta))| \ge |Df(x)|, \qquad \text{for all } x \in J. \tag{2.30}$$

Combining (2.29) and (2.30) gives

$$\begin{aligned}\lambda(f^{-1}(B \cap B_2)) &= \int_{B \cap B_2} \frac{1}{|Df(f^{-1}y)|}\, d\lambda(y) \le \lambda(B) \cdot \frac{1}{|Df(f^{-1}(2\eta))|} \\ &\le K^{(\eta)} \int_{B_3} \frac{1}{|Df(f^{-1}y)|}\, d\lambda(y) = K^{(\eta)}\lambda(f^{-1}(B_3)).\end{aligned}$$

We conclude that

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (2\eta, c-2\eta)\big)\big) \le K^{(\eta)}\lambda\big(T_{\boldsymbol{u}}^{-1}(2\eta, 3\eta)\big).$$

In case $\min_{z \in \partial I^*} |Df(z)| = f^{-1}(c-2\eta)$, a similar reasoning yields

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (2\eta, c-2\eta)\big)\big) \le K^{(\eta)}\lambda\big(T_{\boldsymbol{u}}^{-1}(c-3\eta, c-2\eta)\big).$$

Furthermore, a similar reasoning can be done for the interval $[c,1]$ to conclude that

$$\lambda\big(T_{\boldsymbol{u}}^{-1}\big(B \cap (c+2\eta, 1-2\eta)\big)\big) \le K^{(\eta)}\Big(\lambda\big(T_{\boldsymbol{u}}^{-1}(c+2\eta, c+3\eta)\big) + \lambda\big(T_{\boldsymbol{u}}^{-1}(1-3\eta, 1-2\eta)\big)\Big).$$

Hence, setting $K_1 = \max\{K^{(\eta)}, 1\}$ gives the desired result. □

Proposition 2.3.3 shows that to get the desired estimate from Theorem 2.3.2 it suffices to consider small intervals on the left and right of $[0,1]$ and around c, i.e. sets of the form

$$I_c(\varepsilon) := (c-\varepsilon, c+\varepsilon) \quad \text{and} \quad I_0(\varepsilon) := [0,\varepsilon) \cup (1-\varepsilon, 1]$$

for $\varepsilon > 0$. We first focus on estimating the measure of the intervals $I_c(\varepsilon)$.

Lemma 2.3.4. *There exists a constant $K_2 \geq 1$ such that for all $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma^{n-1} \times \Sigma_G$ and all $\varepsilon > 0$ we have*

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) \leq K_2 \varepsilon.$$

Proof. Let $n \in \mathbb{N}$ and $\boldsymbol{u} \in \Sigma^{n-1} \times \Sigma_G$. Let $\varepsilon > 0$. Suppose that $\varepsilon \geq \frac{1}{4}\min\{c, 1-c\}$. Then

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) \leq 1 \leq \frac{4\varepsilon}{\min\{c, 1-c\}}. \tag{2.31}$$

Now suppose $\varepsilon < \frac{1}{4}\min\{c, 1-c\}$. Again the map $T_{\boldsymbol{u}}$ has non-positive Schwarzian derivative on the interior of any of its intervals of monotonicity and since $u_n \in \Sigma_G$ the interior of the image of any such interval is $(0,1)$. Use $\mathcal{I}$ to denote the collection of connected components of $T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)$. Let $A \in \mathcal{I}$ and write $J = J_A$ and $I = I_A$ for the intervals that satisfy $A \subseteq J$, $A \subseteq I$ and

$$\begin{aligned} T_{\boldsymbol{u}}(J) &= \left[c - \frac{1}{2}\min\{c, 1-c\}, c + \frac{1}{2}\min\{c, 1-c\}\right], \\ T_{\boldsymbol{u}}(I) &= \left[c - \frac{3}{4}\min\{c, 1-c\}, c + \frac{3}{4}\min\{c, 1-c\}\right]. \end{aligned}$$

Also, write $f = T_{\boldsymbol{u}}|_I$. Since f has non-positive Schwarzian derivative, it follows from (1.12) applied to this setting

$$\frac{\lambda(A)}{\lambda(J)} \leq K^{(\frac{1}{4})} \frac{\lambda(f(A))}{\lambda(f(J))} = K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c, 1-c\}}.$$

We conclude that

$$\lambda(T_{\boldsymbol{u}}^{-1} I_c(\varepsilon)) = \sum_{A \in \mathcal{I}} \lambda(A) \leq K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c, 1-c\}} \sum_{A \in \mathcal{I}} \lambda(J_A) \leq K^{(\frac{1}{4})} \frac{2\varepsilon}{\min\{c, 1-c\}}. \tag{2.32}$$

Defining $K_2 = \frac{2\max\{2, K^{(\frac{1}{4})}\}}{\min\{c, 1-c\}}$, the desired result now follows from (2.31) and (2.32). □

To find $\lambda_n\big(I_c(\varepsilon)\big)$, first note that from Lemma 2.1.6 it follows that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{u} \in \Sigma_B^n$,

$$T_{\boldsymbol{u}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(\tilde{K}^{-1} \varepsilon^{\ell_{\boldsymbol{u}}^{-1}}\big). \tag{2.33}$$

By splitting Σ^n according to the final block of bad indices, we can then write using (2.33) and Lemma 2.3.4 that

$$\begin{aligned}\lambda_n\big(I_c(\varepsilon)\big) &= \sum_{k=0}^{n-1} \sum_{\boldsymbol{v}\in\Sigma^{n-k-1}} \sum_{g\in\Sigma_G} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{vgb}} \lambda\big(T_{\boldsymbol{vgb}}^{-1} I_c(\varepsilon)\big) + \sum_{\boldsymbol{b}\in\Sigma_B^n} p_{\boldsymbol{b}} \lambda\big(T_{\boldsymbol{b}}^{-1} I_c(\varepsilon)\big) \\ &\leq \sum_{k=0}^{n-1} \sum_{\boldsymbol{v}\in\Sigma^{n-k-1}} \sum_{g\in\Sigma_G} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{vgb}} \lambda\big(T_{\boldsymbol{vg}}^{-1} I_c(\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}})\big) + \sum_{\boldsymbol{b}\in\Sigma_B^n} p_{\boldsymbol{b}} \lambda\big(I_c(\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}})\big) \\ &\leq \sum_{k=0}^{n-1} \sum_{g\in\Sigma_G} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_g p_{\boldsymbol{b}} K_2 \tilde{K}^{-1} \varepsilon^{\ell_{\boldsymbol{b}}^{-1}} + \sum_{\boldsymbol{b}\in\Sigma_B^n} p_{\boldsymbol{b}} 2\tilde{K}^{-1}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}}.\end{aligned}$$

Taking $K_3 = \max\big\{K_2, 2\big(\sum_{g\in\Sigma_G} p_g\big)^{-1}\big\}\cdot\tilde{K}^{-1} \geq 1$ then gives

$$\lambda_n\big(I_c(\varepsilon)\big) \leq K_3 \sum_{g\in\Sigma_G} \sum_{k=0}^{n} \sum_{\boldsymbol{b}\in\Sigma_B^k} p_g p_{\boldsymbol{b}} \varepsilon^{\ell_{\boldsymbol{b}}^{-1}}. \tag{2.34}$$

We now focus on $I_0(\varepsilon) = [0,\varepsilon) \cup (1-\varepsilon, 1]$. Fix an $0 < \varepsilon_0 < \frac{1}{2}\min\{c, 1-c\}$ and a $t > 1$ that satisfy

$$|DT_j(x)| > t, \qquad \text{for all } x \in I_0(\varepsilon_0) \text{ and each } j \in \Sigma.$$

Such ε_0 and t exist because of (G4) and (B4). From (G3) it follows that for each $0 < \varepsilon < \varepsilon_0$ and $g \in \Sigma_G$,

$$|T_g(x) - T_g(c)| = \left|\int_c^x DT_g(y)dy\right| \geq \frac{K_g}{r_g}\cdot|x-c|^{r_g}.$$

Set $K_4 = \max\{(K_g^{-1}r_g)^{r_g^{-1}} : g \in \Sigma_G\} \geq 1$. Then (G1) implies that

$$T_g^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-1}) \cup I_c(K_4 \varepsilon^{r_g^{-1}}). \tag{2.35}$$

Furthermore, from (B1) it follows that for each $\varepsilon \in (0, \varepsilon_0)$ and $b \in \Sigma_B$,

$$T_b^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-1}). \tag{2.36}$$

Write each $\boldsymbol{u} \in \Sigma^n$ as

$$\boldsymbol{u} = \boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}} \tag{2.37}$$

for some $\tilde{s} \in \{1, \dots, n\}$, where for each i we have $\boldsymbol{b}_i = b_{i,1}\cdots b_{i,k_i} \in \Sigma_B^{k_i}$ and $\boldsymbol{g}_i = g_{i,1}\cdots g_{i,m_i} \in \Sigma_G^{m_i}$ for some $k_1, m_{\tilde{s}} \in \mathbb{N}_0$ and $k_2, \dots, k_{\tilde{s}}, m_1, \dots, m_{\tilde{s}-1} \in \mathbb{N}$. Define

$$s = \begin{cases} \tilde{s}, & \text{if } m_{\tilde{s}} \geq 1, \\ \tilde{s} - 1, & \text{if } m_{\tilde{s}} = 0. \end{cases}$$

Moreover, we introduce notation to indicate the length of the tails of the block $\boldsymbol{u}$:

$$\begin{aligned} d_i &= |\boldsymbol{b}_i \boldsymbol{g}_i \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}|, && i \in \{1, \dots, \tilde{s}\}, \\ q_{i,j} &= |g_{i,j+1} \cdots g_{i,m_i} \boldsymbol{b}_{i+1} \boldsymbol{g}_{i+1} \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}|, && i \in \{1, \dots, \tilde{s}\},\ j \in \{0, \dots, m_i\}. \end{aligned}$$

If necessary to avoid confusion, we write $s(\boldsymbol{u})$, $k_i(\boldsymbol{u})$, etc., to emphasise the dependence on $\boldsymbol{u}$.

Lemma 2.3.5. *There exists a constant $K_5 > 0$ such that for each $0 < \varepsilon < \varepsilon_0$, $n \in \mathbb{N}$ and $\boldsymbol{u} = \boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}} \in \Sigma^n$,*

$$\begin{aligned} T_{\boldsymbol{u}}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-d_1}) &\cup \bigcup_{i=1}^{s} T^{-1}_{\boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{i-1} \boldsymbol{g}_{i-1}} I_c\Big(K_5 (\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}}\Big) \\ &\cup \bigcup_{i=1}^{s} \bigcup_{j=2}^{m_i} T^{-1}_{\boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{i-1} \boldsymbol{g}_{i-1} \boldsymbol{b}_i g_{i,1} \cdots g_{i,j-1}} I_c(K_5 (\varepsilon t^{-q_{i,j}})^{r_{g_{i,j}}^{-1}}). \end{aligned}$$

Proof. We prove the statement by an induction argument for $\tilde{s}$. Let $\boldsymbol{u}$ be a word with symbols in Σ, and write $\boldsymbol{u} = \boldsymbol{b}_1 \boldsymbol{g}_1 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}$ for its decomposition as in (2.37). First suppose that $\tilde{s} = 1$. If $m_1 = 0$, then the statement immediately follows from repeated application of (2.36). If $m_1 \geq 1$, then repeated application of (2.35) gives

$$\begin{aligned} T_{\boldsymbol{g}_1}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-q_{1,0}}) &\cup I_c\Big(K_4 (\varepsilon t^{-q_{1,1}})^{r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_4 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

By setting $K_5 = \tilde{K}^{-1} K_4$, applying (2.33) and (2.36) then yields

$$\begin{aligned} T_{\boldsymbol{b}_1 \boldsymbol{g}_1}^{-1} I_0(\varepsilon) \subseteq I_0(\varepsilon t^{-d_1}) &\cup I_c\Big(K_5 (\varepsilon t^{-q_{1,1}})^{\ell_{\boldsymbol{b}_1}^{-1} r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{\boldsymbol{b}_1 g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_5 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

Note that this is true for the case that $k_1 = 0$ as well. This proves the statement if $\tilde{s} = 1$. Now suppose $\tilde{s}(\boldsymbol{u}) > 1$ and suppose that the statement holds for all words $\boldsymbol{v}$ with $\tilde{s}(\boldsymbol{v}) = \tilde{s}(\boldsymbol{u}) - 1$. In particular, the statement then holds for the word $\boldsymbol{b}_2 \boldsymbol{g}_2 \cdots \boldsymbol{b}_{\tilde{s}} \boldsymbol{g}_{\tilde{s}}$. Note that $m_1 \geq 1$. Again, by repeated application of (2.35) it follows that

$$\begin{aligned} T_{\boldsymbol{g}_1}^{-1} I_0(\varepsilon t^{-d_2}) \subseteq I_0(\varepsilon t^{-q_{1,0}}) &\cup I_c\Big(K_4 (\varepsilon t^{-q_{1,1}})^{r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_4 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

Furthermore, applying (2.33) and (2.36) then yields

$$\begin{aligned} T_{\boldsymbol{b}_1 \boldsymbol{g}_1}^{-1} I_0(\varepsilon t^{-d_2}) \subseteq I_0(\varepsilon t^{-d_1}) &\cup I_c\Big(K_5 (\varepsilon t^{-q_{1,1}})^{\ell_{\boldsymbol{b}_1}^{-1} r_{g_{1,1}}^{-1}}\Big) \\ &\cup \bigcup_{j=2}^{m_1} T^{-1}_{\boldsymbol{b}_1 g_{1,1} \cdots g_{1,j-1}} I_c\Big(K_5 (\varepsilon t^{-q_{1,j}})^{r_{g_{1,j}}^{-1}}\Big). \end{aligned}$$

This together with the statement being true for the word $\boldsymbol{b}_2\boldsymbol{g}_2\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}$ yields the statement for $\boldsymbol{u}$. □

Combining Lemma 2.3.4 and Lemma 2.3.5 gives

$$\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon)) \leq 2\varepsilon t^{-d_1} + \sum_{i=1}^{s} K_2K_5(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}} + \sum_{i=1}^{s}\sum_{j=2}^{m_i} K_2K_5(\varepsilon t^{-q_{i,j}})^{r_{g_{i,j}}^{-1}}.$$

Let $r_{\max} = \max\{r_g : g \in \Sigma_G\}$ and set $\alpha := t^{1/r_{\max}} > 1$. Then

$$\sum_{i=1}^{s}\sum_{j=2}^{m_i} \alpha^{-q_{i,j}} \leq \sum_{\ell=0}^{\infty} \alpha^{-\ell} = \frac{1}{1-1/\alpha},$$

so that

$$\begin{aligned}\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon)) &\leq 2\varepsilon^{1/r_{\max}} + K_2K_5\sum_{i=1}^{s}\sum_{j=2}^{m_i} \varepsilon^{1/r_{\max}}\alpha^{-q_{i,j}} + \sum_{i=1}^{s} K_2K_5(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}} \\ &\leq \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{i=1}^{s}(\varepsilon t^{-q_{i,1}})^{\ell_{\boldsymbol{b}_i}^{-1} r_{g_{i,1}}^{-1}}.\end{aligned} \tag{2.38}$$

Proposition 2.3.6. *There exists a constant $K_6 > 0$ such that for each $\varepsilon \in (0, \varepsilon_0)$ and $n \in \mathbb{N}$,*

$$\lambda_n(I_0(\varepsilon)) \leq K_6 \sum_{g\in\Sigma_G} p_g \sum_{k=0}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{b}}\ell_{\boldsymbol{b}} \cdot \varepsilon^{\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}.$$

Proof. Let $n \in \mathbb{N}$. Then with (2.38) we obtain

$$\begin{aligned}\lambda_n(I_0(\varepsilon)) &= \sum_{\boldsymbol{u}\in\Sigma^n} p_{\boldsymbol{u}}\lambda\big(T_{\boldsymbol{u}}^{-1}(I_0(\varepsilon))\big) \\ &\leq \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{\boldsymbol{u}\in\Sigma^n} p_{\boldsymbol{u}}\sum_{i=1}^{s(\boldsymbol{u})}\left(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\right)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}} \\ &= \left(2 + \frac{K_2K_5}{1-1/\alpha}\right)\varepsilon^{1/r_{\max}} \\ &\quad + K_2K_5\sum_{i=1}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\left(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\right)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}},\end{aligned} \tag{2.39}$$

where we defined $\tau = \lfloor \frac{n+1}{2} \rfloor$ which is the largest value $s(\boldsymbol{u})$ can take. Let us consider the term in (2.39). First of all, note that a word $\boldsymbol{u} \in \Sigma^n$ satisfies $s(\boldsymbol{u}) \geq 1$ if and only if $m_1(\boldsymbol{u}) \geq 1$. Therefore,

$$\{\boldsymbol{u} \in \Sigma^n : s(\boldsymbol{u}) \geq 1\} = \bigcup_{k=0}^{n-1} \Sigma_B^k \times \Sigma_G \times \Sigma^{n-k-1}.$$

Hence, defining the function χ on $\{0,\dots,n-1\}^2$ by

$$\chi(k,q) = \sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G} p_{\boldsymbol{b}}p_g\left(\varepsilon t^{-q}\right)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}, \qquad (k,q)\in\{0,\dots,n-1\}^2. \tag{2.40}$$

we can rewrite and bound the term with $i=1$ in (2.39) as follows:

$$\begin{aligned}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(1)p_{\boldsymbol{u}}\left(\varepsilon t^{-q_{1,1}(\boldsymbol{u})}\right)^{\ell_{\boldsymbol{b}_1(\boldsymbol{u})}^{-1}r_{g_{1,1}(\boldsymbol{u})}^{-1}} &= \sum_{k=0}^{n-1}\sum_{\boldsymbol{v}\in\Sigma^{n-k-1}} p_{\boldsymbol{v}}\chi(k,n-k-1)\\ &\le \varepsilon^{1/r_{\max}} + \sum_{k=1}^{n-1}\chi(k,n-k-1).\end{aligned} \tag{2.41}$$

Secondly, note that for each $i\in\{2,\dots,\tau\}$ a word $\boldsymbol{u}\in\Sigma^n$ satisfies $s(\boldsymbol{u})\ge i$ if and only if $m_{i-1}(\boldsymbol{u}),k_i(\boldsymbol{u}),m_i(\boldsymbol{u})\ge 1$. For each $k\in\{1,\dots,n-2\}$ and $q\in\{0,\dots,n-k-2\}$ and $i\in\{2,\dots,\tau\}$ we define

$$A_{i,k,q} = \{\boldsymbol{v}\in\Sigma^{n-k-q-1} : \tilde{s}(\boldsymbol{v}) = i-1, v_{n-k-q-1}\in\Sigma_G\}.$$

The set $A_{i,k,q}$ contains all words of length $n-k-q-1$ that can precede the word $\boldsymbol{b}_i\boldsymbol{g}_i\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}$ with $|\boldsymbol{b}_i| = k$ and $|g_{i,2}\cdots g_{i,m_i}\boldsymbol{b}_{i+1}\boldsymbol{g}_{i+1}\cdots\boldsymbol{b}_{\tilde{s}}\boldsymbol{g}_{\tilde{s}}| = q$. So

$$\{\boldsymbol{u}\in\Sigma^n : s(\boldsymbol{u})\ge i\} = \bigcup_{k=1}^{n-2}\bigcup_{q=0}^{n-k-2} A_{i,k,q}\times\Sigma_B^k\times\Sigma_G\times\Sigma^q, \qquad i\in\{2,\dots,\tau\}.$$

Hence, using (2.40) we can rewrite and bound the sum in (2.39) that runs from $i=2$ to τ as follows:

$$\begin{aligned}&\sum_{i=2}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\dots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\left(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\right)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1}r_{g_{i,1}(\boldsymbol{u})}^{-1}}\\ &\qquad = \sum_{i=2}^{\tau}\sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\sum_{\boldsymbol{v}_1\in A_{i,k,q}}\sum_{\boldsymbol{v}_2\in\Sigma^q} p_{\boldsymbol{v}_1}p_{\boldsymbol{v}_2}\chi(k,q)\\ &\qquad = \sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\chi(k,q)\sum_{i=2}^{\tau}\sum_{\boldsymbol{v}_1\in A_{i,k,q}}\sum_{\boldsymbol{v}_2\in\Sigma^q} p_{\boldsymbol{v}_1}p_{\boldsymbol{v}_2}\\ &\qquad \le \sum_{k=1}^{n-2}\sum_{q=0}^{n-k-2}\chi(k,q).\end{aligned} \tag{2.42}$$

Here the last step follows from the fact that

$$\sum_{i=2}^{\tau}\sum_{\boldsymbol{v}_1\in A_{i,k,q}} p_{\boldsymbol{v}_1} \le \sum_{\boldsymbol{v}\in\Sigma^{n-k-q-2}}\sum_{g\in\Sigma_G} p_{\boldsymbol{v}}p_g \le 1.$$

Combining (2.41) and (2.42) gives

$$\sum_{i=1}^{\tau}\sum_{\boldsymbol{u}\in\Sigma^n} 1_{\{1,\ldots,s(\boldsymbol{u})\}}(i)p_{\boldsymbol{u}}\left(\varepsilon t^{-q_{i,1}(\boldsymbol{u})}\right)^{\ell_{\boldsymbol{b}_i(\boldsymbol{u})}^{-1} r_{g_{i,1}(\boldsymbol{u})}^{-1}} \leq \varepsilon^{1/r_{\max}} + \sum_{k=1}^{n-1}\sum_{q=0}^{n-k-1}\chi(k,q). \tag{2.43}$$

Furthermore, for each $\boldsymbol{b}\in\Sigma_B^k$ and $g\in\Sigma_G$ we have again by setting $r_{\max}=\max\{r_j : j\in\Sigma_G\}$ and $\alpha = t^{1/r_{\max}}$ that

$$\sum_{q=0}^{n-k-1}(t^{-q})^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}} \leq \sum_{q=0}^{n-k-1}\left(\alpha^{-\ell_{\boldsymbol{b}}^{-1}}\right)^q \leq \frac{1}{1-\alpha^{-\ell_{\boldsymbol{b}}^{-1}}} \leq \frac{\alpha\ell_{\boldsymbol{b}}^{-1}}{\alpha^{\ell_{\boldsymbol{b}}^{-1}}-1}\ell_{\boldsymbol{b}} \leq \frac{\alpha}{\log(\alpha)}\ell_{\boldsymbol{b}}, \tag{2.44}$$

where the last step follows from the fact that $f(x)=\frac{x}{\alpha^x-1}$ is a decreasing function and $\lim_{x\downarrow 0} f(x) = \frac{1}{\log\alpha}$. Hence, combining (2.39), (2.43) and (2.44) gives

$$\begin{aligned}
&\lambda_n(I_0(\varepsilon))\\
&\leq \left(2+\frac{K_2K_5}{1-1/\alpha}+K_2K_5\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{k=1}^{n-1}\sum_{q=0}^{n-k-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G} p_{\boldsymbol{b}}p_g\left(\varepsilon t^{-q}\right)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}\\
&\leq \left(2+K_2K_5\frac{2\alpha-1}{\alpha-1}\right)\varepsilon^{1/r_{\max}} + K_2K_5\sum_{k=1}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}\sum_{g\in\Sigma_G} p_{\boldsymbol{b}}p_g\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}\frac{\alpha\ell_{\boldsymbol{b}}}{\log(\alpha)}\\
&\leq K_6\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{n-1}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}},
\end{aligned}$$

where $K_6 = \frac{1}{\min\{p_g : g\in\Sigma_G\}}\left(2+K_2K_5\frac{2\alpha-1}{\alpha-1}\right)+\frac{K_2K_5\alpha}{\log\alpha}$. □

We are now ready to prove Theorem 2.3.2.

Proof of Theorem 2.3.2. Let $B\subseteq[0,1]$ be a Borel set. First suppose that $\lambda(B)\geq\frac{\varepsilon_0}{3}$. Then there exists a constant $C=C(\varepsilon_0)>0$ such that

$$\lambda_n(B)\leq 1\leq C\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\cdot\lambda(B)^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}.$$

Now suppose that $\lambda(B)<\frac{\varepsilon_0}{3}$ and set $\varepsilon=3\lambda(B)$. It follows from Proposition 2.3.3 that for all $n\in\mathbb{N}$ and all $\boldsymbol{u}\in\Sigma^n$ we have

$$\lambda\left(T_{\boldsymbol{u}}^{-1}B\right)\leq K_1\left(\lambda(T_{\boldsymbol{u}}^{-1}I_0(\varepsilon))+\lambda(T_{\boldsymbol{u}}^{-1}I_c(\varepsilon))\right).$$

Together with (2.34) and Proposition 2.3.6 this yields for all $n\in\mathbb{N}$ that

$$\lambda_n(B)\leq K_1\cdot(K_3+K_6)\sum_{g\in\Sigma_G}p_g\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\ell_{\boldsymbol{b}}\cdot\varepsilon^{\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}.$$

This gives the result. □

§2.4 Further results and final remarks

§2.4.1 Proof of Corollaries 2.1.4 and 2.1.5

In this section we prove Corollaries 2.1.4 and 2.1.5.

Proof of Corollary 2.1.4. We use the bound (2.6) obtained in Theorem 2.1.3. For convenience, we set $\ell = \ell_{\max}$ and $x = \lambda(B)^{1/r_{\max}}$. The asymptotics are determined by the interplay between $\theta^k \searrow 0$ and $x^{1/\ell^k} \nearrow 1$. First suppose $\theta < x^{1/\ell}$. Then $\lambda(B) > \theta^{\ell r_{\max}}$, so for each $\varkappa > 0$ there exists $C > 0$ sufficiently large such that

$$\mu(B) \leq 1 \leq C \cdot \frac{1}{\log^{\varkappa}(1/\lambda(B))}.$$

Now suppose $\theta \geq x^{1/\ell}$. Note that $\theta^N \geq x^{1/\ell^N}$ if and only if

$$\log N + N \log \ell \leq \log\left(\frac{\log x}{\log \theta}\right).$$

Since $\log N \leq N$, this last inequality is satisfied if we take for example

$$N = \left\lfloor \frac{1}{1+\log \ell} \log\left(\frac{\log x}{\log \theta}\right) \right\rfloor = \left\lfloor \frac{1}{1+\log \ell} \log\left(\frac{\log(1/x)}{\log(1/\theta)}\right) \right\rfloor, \tag{2.45}$$

where $\lfloor y \rfloor$ denotes the largest integer not exceeding y. Taking N as in (2.45), note that it follows from $\theta \geq x^{1/\ell}$ that $N \geq 0$. Then $\theta^k \geq x^{1/\ell^k}$ for all $k \leq N$ as well, and hence

$$\begin{aligned}
\sum_{k=0}^{\infty} \theta^k x^{1/\ell^k} &= \sum_{k=0}^{N} \theta^k x^{1/\ell^k} + \sum_{k=N+1}^{\infty} \theta^k x^{1/\ell^k} \leq \sum_{k=0}^{N} \theta^k \cdot x^{1/\ell^N} + \sum_{k=N+1}^{\infty} \theta^k \cdot 1 \\
&\leq \frac{1}{1-\theta} x^{1/\ell^N} + \frac{\theta^{N+1}}{1-\theta} \leq \frac{1}{1-\theta}(1+\theta)\theta^N.
\end{aligned}$$

From (2.45) we see that $N \geq \frac{1}{1+\log \ell} \log\left(\frac{\log x}{\log \theta}\right) - 1$, thus

$$\begin{aligned}
\theta^N &= \exp(N \log \theta) \leq \exp\left(\left(\frac{1}{1+\log \ell} \log\left(\frac{\log(1/x)}{\log(1/\theta)}\right) - 1\right) \log \theta\right) \\
&= \exp\left(\frac{\log \theta}{1+\log \ell} \log\log(1/x) + C(\ell,\theta)\right) \\
&= \overline{C}(\ell,\theta)\big(\log(1/x)\big)^{\frac{\log \theta}{1+\log \ell}} = \overline{C}(\ell,\theta)\left(\frac{r_{\max}}{\log(1/\lambda(B))}\right)^{\varkappa},
\end{aligned}$$

where we set $\varkappa = \frac{\log(1/\theta)}{1+\log \ell} > 0$, and where $C(\ell,\theta) \in \mathbb{R}$ and $\overline{C}(\ell,\theta) > 0$ are constants that only depend on ℓ and θ. We conclude from the bound (2.6) that

$$\mu(B) \leq K \cdot \frac{1}{\log^{\varkappa}(1/\lambda(B))}$$

for some positive constant K. □

The proof of Corollary 2.1.5 consists of two steps. Firstly we show that any weak limit point of $\mu_{\boldsymbol{p}_n}$ is a stationary measure, i.e. satisfies (1.18), and secondly that any weak limit point of $\mu_{\boldsymbol{p}_n}$ is absolutely continuous with respect to the Lebesgue measure. The corollary then follows from the uniqueness of absolutely continuous stationary measures given by Theorem 2.1.2.

Proof of Corollary 2.1.5. For each $n \geq 0$, let $\boldsymbol{p}_n = (p_{n,j})_{j\in\Sigma}$ be a strictly positive probability vector such that $\sup_n \sum_{b\in\Sigma_B} p_{n,b}\ell_b < 1$ and assume that $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ in $\mathbb{R}^N_+$ for some strictly positive probability vector $\boldsymbol{p} = (p_j)_{j\in\Sigma}$. Let $\tilde{\mu}$ be a weak limit point of an arbitrary subsequence of $\mu_{\boldsymbol{p}_n}$. Again, note that such a $\tilde{\mu}$ exists because the space of probability measures on $[0,1]$ equipped with the weak topology is sequentially compact. After passing to a further subsequence we have for any continuous function $\varphi : [0,1] \to \mathbb{R}$ that

$$\lim_{n\to\infty} \int_{[0,1]} \varphi \, d\mu_{\boldsymbol{p}_n} = \int_{[0,1]} \varphi \, d\tilde{\mu}.$$

Moreover, by the stationarity of the measures $\mu_{\boldsymbol{p}_n}$ it follows that for each $n \geq 1$,

$$\int_{[0,1]} \varphi \, d\mu_{\boldsymbol{p}_n} = \sum_{j\in\Sigma} p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n}.$$

To prove that $\tilde{\mu}$ is stationary for $\boldsymbol{p}$, it is sufficient to show that for each $j \in \Sigma$,

$$\lim_{n\to\infty} p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n} = p_j \int_{[0,1]} \varphi \circ T_j \, d\tilde{\mu}. \tag{2.46}$$

If $j \in \Sigma_B$ this is obvious, since then $\varphi \circ T_j$ is continuous. For $j \in \Sigma_G$ the map $\varphi \circ T_j$ might have a discontinuity at c. In this case, we let φ_δ be the continuous function given by $\varphi_\delta(x) = \varphi \circ T_j(x)$ for $x \in I \setminus (c-\delta, c+\delta)$ and φ_δ is linear otherwise. Then we have

$$\lim_{n\to\infty} \left| p_{n,j} \int_{[0,1]} \varphi_\delta \, d\mu_{\boldsymbol{p}_n} - p_j \int_{[0,1]} \varphi_\delta \, d\tilde{\mu} \right| = 0,$$

by the weak convergence and since $p_{n,j} \to p_j$ as $n \to \infty$. Also, we have

$$\left| p_{n,j} \int_{[0,1]} \varphi \circ T_j \, d\mu_{\boldsymbol{p}_n} - p_{n,j} \int_{[0,1]} \varphi_\delta \, d\mu_{\boldsymbol{p}_n} \right| \leq C\mu_{\boldsymbol{p}_n}([c-\delta, c+\delta]) \to 0 \text{ as } \delta \to 0,$$

where the convergence is uniform in n because of (2.6). Similarly,

$$\left| p_j \int_{[0,1]} \varphi \circ T_j d\tilde{\mu} - p_j \int_{[0,1]} \varphi_\delta d\tilde{\mu} \right| \leq C\tilde{\mu}([c-\delta, c+\delta]) \to 0 \text{ as } \delta \to 0,$$

The last three relations imply (2.46).

To show that $\tilde{\mu}$ is absolutely continuous with respect to the Lebesgue measure λ we proceed as in the proof of Theorem 2.1.3. We set $\tilde{\theta} = \sup_n \sum_{b\in\Sigma_B} p_{n,b}\ell_b < 1$. Let $B \subseteq [0,1]$ be a Borel set. Every $\mu_{\boldsymbol{p}_n}$ satisfies the conclusion of Theorem 2.1.3, so

$$\mu_{\boldsymbol{p}_n}(B) \leq C_n \sum_{k=0}^{\infty} \tilde{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}},$$

where the constant C_n depends on $\left(\sum_{g\in\Sigma_G} p_{n,g}\right)^{-1}$ and $(\min\{p_{n,g} : g \in \Sigma_G\})^{-1}$ (and properties of the good and bad maps themselves that are not linked to the probabilities). Since $\boldsymbol{p}$ and each $\boldsymbol{p}_n$, $n \geq 0$, are strictly positive probability vector and $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$, both these quantities can be bounded from above and $\tilde{C} := \sup_n C_n < \infty$. From the weak convergence of $\mu_{\boldsymbol{p}_n}$ to $\tilde{\mu}$ we obtain as in (2.28) using the Portmanteau Theorem that

$$\tilde{\mu}(B) \leq \tilde{C} \sum_{k=0}^{\infty} \tilde{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}}.$$

Hence, $\tilde{\mu} \ll \lambda$. By Theorem 2.1.2 we know that $\mu_{\boldsymbol{p}}$ is the unique acs probability measure for $(\mathcal{T}, \boldsymbol{p})$. So, $\tilde{\mu} = \mu_{\boldsymbol{p}}$. □

§2.4.2 The non-superattracting case

With some modifications the results from Theorem 2.1.2 and Theorem 2.1.3 can be extended to the class $\mathfrak{B}^1 \supseteq \mathfrak{B}$ of bad maps whose critical order ℓ_b in (B3) is allowed to be equal to 1. We will list the modified statements and the necessary modifications to the proofs here. Note that for each $T \in \mathfrak{B}^1 \setminus \mathfrak{B}$, we have $DT(c) \neq 0$, and due to the minimal principle, $|DT(c)| < 1$. We consider $T_1, \ldots, T_N \in \mathfrak{G} \cup \mathfrak{B}^1$ with $\Sigma_B^1 = \{1 \leq j \leq N : T_j \in \mathfrak{B}^1\}$ and Σ_G, Σ_B as before and such that $\Sigma_G, \Sigma_B^1 \setminus \Sigma_B \neq \emptyset$. Furthermore, we write again $\Sigma = \{1, \ldots, N\} = \Sigma_G \cup \Sigma_B^1$. We again set $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$.

Theorem 2.4.1. *Let $\mathcal{T} = \{T_j : j \in \Sigma\}$ be as above and $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ a strictly positive probability vector.*

(a) *There exists a unique (up to scalar multiplication) σ-finite acs measure μ for $(\mathcal{T}, \boldsymbol{p})$. Moreover, F is ergodic w.r.t. $m_{\boldsymbol{p}} \times \mu$ and the density $\frac{d\mu}{d\lambda}$ is bounded away from zero and is locally Lipschitz on $(0, c)$ and $(c, 1)$.*

(b) *Suppose $\ell_{\max} > 1$.*

 (i) *The measure μ is finite if and only if $\theta = \sum_{b\in\Sigma_B^1} p_b \ell_b < 1$. In this case, for each $\hat{\theta} \in (\theta, 1)$ there exists a constant $C(\hat{\theta}) > 0$ such that*

$$\mu(B) \leq C(\hat{\theta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}^k \lambda(B)^{\ell_{\max}^{-k} r_{\max}^{-1}}$$

 for any Borel set $B \subseteq [0, 1]$, where $r_{\max} = \max\{r_g : g \in \Sigma_G\}$.

 (ii) *The density $\frac{d\mu}{d\lambda}$ is not in L^q for any $q > 1$.*

(c) *Suppose $\ell_{\max} = 1$.*

 (i) *The measure μ is finite, and for each $\boldsymbol{\eta} = (\eta_b)_{b\in\Sigma_B^1}$ such that $\eta_b > 1$ for each $b \in \Sigma_B^1$ and $\hat{\theta}(\boldsymbol{\eta}) = \sum_{b\in\Sigma_B^1} p_b \eta_b < 1$ there exists a constant $C(\boldsymbol{\eta}) > 0$ such that*

$$\mu(B) \leq C(\boldsymbol{\eta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}(\boldsymbol{\eta})^k \lambda(B)^{\eta_{\max}^{-k} r_{\max}^{-1}}$$

for any Borel set $B \subseteq [0,1]$, where $\eta_{\max} = \max\{\eta_b : b \in \Sigma_B^1\}$. If $\sum_{b\in\Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, so if the bad maps are expanding on average at the point c, then we can get the estimate

$$\mu(B) \leq C \cdot \lambda(B)^{r_{\max}^{-1}} \tag{2.47}$$

for some constant $C > 0$ and any Borel set $B \subseteq [0,1]$.

(ii) If $r_{\max} > 1$, then $\frac{d\mu}{d\lambda} \notin L^q$ for any $q \geq \frac{r_{\max}}{r_{\max}-1}$. If also $\sum_{b\in\Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, then $\frac{d\mu}{d\lambda} \in L^q$ for all $1 \leq q < \frac{r_{\max}}{r_{\max}-1}$.

(iii) If $r_{\max} = 1$ and $\sum_{b\in\Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, then $\frac{d\mu}{d\lambda} \in L^\infty$.

The main issue we need to deal with in order to get Theorem 2.4.1 is adapting Lemma 2.1.6, i.e. finding suitable bounds for $|T_\omega^n(x) - c|$, since the constants $\tilde{K}$ and $\tilde{M}$ from Lemma 2.1.6 are not well defined in case $\ell_{\min} = 1$. This is done in the next two lemmas. For the upper bound of $|T_\omega^n(x) - c|$ we assume $\ell_{\max} > 1$ since we only need it for the proof of part (b)(i).

Lemma 2.4.2. *Let $\{T_j : j \in \Sigma\}$ be as above. Suppose $\ell_{\max} > 1$. There are constants $\hat{M} > 1$ and $\delta > 0$ such that for all $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $x \in [c-\delta, c+\delta]$ we have*

$$|T_\omega^n(x) - c| \leq \left(\hat{M}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}}.$$

Proof. Similar as in the proof of Lemma 2.1.6 it follows that there exists an $M > 1$ such that for any $b \in \Sigma_B$ and $x \in [0,1]$ we have

$$|T_b(x) - c| \leq M|x-c|^{\ell_b}. \tag{2.48}$$

Furthermore, there exists a $\delta > 0$ such that $|DT_b(x)| < 1$ for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B^1$. This implies

$$|T_b(x) - c| < |x - c| \tag{2.49}$$

for all $x \in [c-\delta, c+\delta]$ and $b \in \Sigma_B^1$. Note that $\Sigma_B \neq \emptyset$ because $\ell_{\max} > 1$. We set $\upsilon = \min\{\ell_b : b \in \Sigma_B\} > 1$ and $\hat{M} = M^{\frac{1}{\upsilon-1}}$. For each $n \in \mathbb{N}$ and $\omega \in (\Sigma_B^1)^{\mathbb{N}}$, write

$$m(n,\omega) = \#\{1 \leq \omega_i \leq n : \ell_{\omega_i} > 1\}.$$

The statement follows by showing that for all $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $x \in [c-\delta, c+\delta]$ we have

$$|T_\omega^n(x) - c| \leq \left(M^{(1-\upsilon^{-m(n,\omega)})/(\upsilon-1)}|x-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_n}}. \tag{2.50}$$

We prove (2.50) by induction. From (2.48) and (2.49) it follows that (2.50) holds for $n = 1$. Now suppose (2.50) holds for some $n \in \mathbb{N}$. Let $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $y \in [c-\delta, c+\delta]$. If $\ell_{\omega_{n+1}} = 1$, then the desired result follows by applying (2.49) with $b = \omega_{n+1}$ and $x = T_\omega^n(y)$. Suppose $\ell_{\omega_{n+1}} > 1$. Then, using (2.48),

$$\begin{aligned}|T_\omega^{n+1}(y) - c| &\leq M|T_\omega^n(y) - c|^{\ell_{\omega_{n+1}}} \\ &\leq \left(M^{(1-\upsilon^{-m(n,\omega)})/(\upsilon-1)+\upsilon^{-m(n+1,\omega)}}|y-c|\right)^{\ell_{\omega_1}\cdots\ell_{\omega_{n+1}}}.\end{aligned}$$

Using that

$$\upsilon^{-m(n+1,\omega)} = \frac{\upsilon^{-m(n,\omega)} - \upsilon^{-m(n+1,\omega)}}{\upsilon - 1},$$

the desired result follows. □

Lemma 2.4.3. *Let $\{T_j : j \in \Sigma\}$ be as above. Let $\boldsymbol{\eta} = (\eta_b)_{b\in\Sigma_B^1}$ be a vector such that $\eta_b > 1$ for each $b \in \Sigma_B^1$. Set $\hat{\eta}_b = \max\{\eta_b, \ell_b\}$ for each $b \in \Sigma_B^1$. Then there exists a constant $\hat{K}(\boldsymbol{\eta}) \in (0,1)$ such that for all $n \in \mathbb{N}$, $\omega \in (\Sigma_B^1)^{\mathbb{N}}$ and $x \in [0,1]$ we have*

$$\left(\hat{K}(\boldsymbol{\eta})|x-c|\right)^{\hat{\eta}_{\omega_1}\cdots\hat{\eta}_{\omega_n}} \le |T_\omega^n(x) - c|.$$

Proof. Note from (B3) that for each $b \in \Sigma_B^1$ we have

$$K_b|x-c|^{\hat{\eta}_b - 1} \le K_b|x-c|^{\ell_b - 1} \le |DT_b(x)|.$$

By setting $\hat{\eta}_{\min} = \min\{\hat{\eta}_b : b \in \Sigma_B^1\}$, $\hat{\eta}_{\max} = \max\{\hat{\eta}_b : b \in \Sigma_B^1\}$ and $\hat{K}(\boldsymbol{\eta}) = \left(\frac{\min\{K_b : b\in\Sigma_B^1\}}{\hat{\eta}_{\max}}\right)^{\frac{1}{\hat{\eta}_{\min}-1}}$ the result now follows in the same way as in the proof of Lemma 2.1.6. □

Proof of Theorem 2.4.1. Firstly, note that (a), (b)(ii) and the first part of (c)(ii) immediately follow from Remark 2.2.8. Moreover, as in [dMvS93, Section 5.4] it can be shown that (2.47) implies that $\frac{d\mu}{d\lambda}$ is in L^q if $r_{\max} > 1$ and $1 \le q < \frac{r_{\max}}{r_{\max}-1}$, giving the remainder of 3(ii). It is immediate that (2.47) implies that $\frac{d\mu}{d\lambda}$ is in L^∞ if $r_{\max} = 1$, so (3)(iii) holds. Hence, it remains to prove (b)(i) and (c)(i).

Suppose $\theta = \sum_{b\in\Sigma_B^1} p_b\ell_b \ge 1$, which means that $\ell_{\max} > 1$. The proof that in this case $\mu_{\boldsymbol{p}}$ is infinite follows by the same reasoning as in Subsection 2.3.1 by now taking $\gamma = \min\{\delta, \frac{1}{2}\hat{M}^{-1}\}$ with δ and $\hat{M}$ as in the proof of Lemma 2.4.2. Now suppose $\theta < 1$. Let $\boldsymbol{\eta} = (\eta_b)_{b\in\Sigma_B^1}$ be a vector such that $\eta_b > 1$ for each $b \in \Sigma_B^1$ and $\hat{\theta}(\boldsymbol{\eta}) = \sum_{b\in\Sigma_B^1} p_b\hat{\eta}_b < 1$ with again $\hat{\eta}_b = \max\{\eta_b, \ell_b\}$. Applying Lemma 2.4.3 yields that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{b} \in (\Sigma_B^1)^n$,

$$T_{\boldsymbol{b}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(\hat{K}(\boldsymbol{\eta})^{-1}\varepsilon^{\hat{\eta}_{\boldsymbol{b}}^{-1}}\big), \tag{2.51}$$

where we used the notation $\hat{\eta}_{\boldsymbol{b}} = \hat{\eta}_{b_1}\cdots\hat{\eta}_{b_n}$ for a word $\boldsymbol{b} = b_1\cdots b_n$. Following the line of reasoning in Subsection 2.3.2 with (2.51) instead of (2.33), we obtain that there exists a constant $C(\boldsymbol{\eta}) > 0$ such that

$$\mu(B) \le C(\boldsymbol{\eta}) \cdot \sum_{k=0}^{\infty} \hat{\theta}(\boldsymbol{\eta})^k \lambda(B)^{\hat{\eta}_{\max}^{-k} r_{\max}^{-1}} \tag{2.52}$$

for any Borel set $B \subseteq [0,1]$. In case $\ell_{\max} > 1$ we can choose $\boldsymbol{\eta}$ to satisfy $\hat{\eta}_{\max} = \ell_{\max}$ and such that $\hat{\theta}(\boldsymbol{\eta}) - \theta > 0$ is arbitrarily small, which yields (b)(i). In case $\ell_{\max} = 1$, then $\hat{\eta}_{\max} = \eta_{\max}$, so this together with (2.52) yields the first part of (c)(i).

Finally, for the second part of (c)(i), suppose $\ell_{\max} = 1$ and $\Lambda = \sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$. Setting $K_{\boldsymbol{b}} = |DT_{\boldsymbol{b}}(c)|$ for each $\boldsymbol{b} \in (\Sigma_B^1)^n$ and $n \in \mathbb{N}$, note that for all $\varepsilon > 0$, $n \in \mathbb{N}$, $\boldsymbol{b} \in (\Sigma_B^1)^n$,

$$T_{\boldsymbol{b}}^{-1}\big(I_c(\varepsilon)\big) \subseteq I_c\big(K_{\boldsymbol{b}}^{-1}\varepsilon\big). \tag{2.53}$$

By using (2.53) instead of (2.33), letting $\tilde{p}_{\boldsymbol{b}} = K_{\boldsymbol{b}}^{-1} p_{\boldsymbol{b}}$ play the role of $p_{\boldsymbol{b}}$ in the reasoning of Subsection 2.3.2 and noting that $\Lambda^k = \sum_{\boldsymbol{b} \in (\Sigma_B^1)^k} \tilde{p}_{\boldsymbol{b}}$, we arrive similarly as for Theorem 2.3.2 to the conclusion that there exists a constant $\tilde{C} > 0$ such that for all $n \in \mathbb{N}$ and all Borel sets $B \subseteq [0,1]$,

$$\lambda_n(B) \leq \tilde{C} \cdot \sum_{g \in \Sigma_G} p_g \Big(\sum_{k=0}^{\infty} \Lambda^k \Big) \lambda(B)^{r_g^{-1}}.$$

This proves the remaining part of (c)(i). □

§2.4.3 Final remarks

The results from Theorem 2.4.1 contain one possible extension of our main results to another set of conditions (G1)–(G4), (B1)–(B4). In this section we discuss some of the questions that our main results brought up in this respect, i.e. about whether or not some of the conditions (G1)–(G4), (B1)–(B4) can be relaxed, and questions about other possible future extensions.

A condition that plays a fundamental role in the proofs of Theorem 2.1.2 and Theorem 2.1.3 is the fact that the critical point is mapped to a point that is a common repelling fixed point for all maps T_j. We considered whether this condition can be relaxed, for instance by assuming that the branches of one of the good maps are not full. However, in this case the critical values of the random system are not just $0, c, 1$ but contain all the values of all possible postcritical orbits of c. This has several consequences:

- An invariant density (if it exists) clearly cannot be locally Lipschitz on $(0,c)$ and $(c,1)$.

- Proposition 2.3.3 and all subsequent arguments fail, since it is not sufficient to restrict to neighbourhoods around only 0, c and 1. One might try to solve this issue by requiring that (on average) the postcritical orbits 'gain enough expansion' as is done in for instance [NvS91] for deterministic maps (see Theorem 2.1.1). An analogous condition for random systems, however, would be much more difficult to verify since it would involve all possible random orbits of c.

- The argument using Kac's Lemma might fail, because in that case there exist words $\boldsymbol{u}$ with symbols in Σ and neighbourhoods U of c such that $T_{\boldsymbol{u}}(x)$ is bounded away from 0 and 1 uniformly in $x \in U$.

There are also some additional questions that our main results raise. It would be interesting for example to study further statistical properties of the random systems such as mixing properties and if possible mixing rates in case the acs measure is finite.

As is well known from e.g. [KN92, Y92, Y98, BLvS03] the good maps individually have exponential decay of correlations. But since trajectories in the random systems spend long periods of time near the points 0 or 1 or both, polynomial mixing rates are expected rather than exponential. This will be the topic of Chapter 3.

The dynamical behaviour of the system is governed by the interplay between the superexponential convergence to c and the exponential divergence from 0 and 1. In this chapter we fixed the exponential divergence away from 0 and 1 and the two regimes $\theta < 1$ and $\theta \geq 1$ in Theorem 2.1.3 only refer to the convergence to c: For smaller (bigger) θ orbits are less (more) attracted to c. It would be interesting to see under what other conditions on the rates of convergence to c and divergence from 0 and 1 the system admits an acs measure. Could one for example

- take exponential convergence to c and polynomial divergence from 0 and 1? We will investigate this question in Chapter 4.

- replace the conditions (G4) and (B4) stating that all good and bad maps are expanding at 0 and 1 by the condition that the random system is expanding on average at a sufficiently large neighbourhood of 0 and 1?

Finally, in Theorems 2.1.2 and 2.4.1 we have seen that the regularity of the density $\frac{d\mu}{d\lambda}$ depends on whether or not there is a bad map for which c is superstable: If $\ell_{\max} > 1$, then $\frac{d\mu}{d\lambda}$ is not in L^q for any $q > 1$. On the other hand, if $\ell_{\max} = 1$ and the bad maps are expanding on average at c, i.e. $\sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} < 1$, then the density has the same regularity as in the setting of Theorem 2.1.1 by Nowicki and Van Strien. Indeed, in this case, if $r_{\max} > 1$, we have $\frac{d\mu}{d\lambda} \in L^q$ if and only if $1 \leq q < \frac{r_{\max}}{r_{\max}-1}$ and in the case that $r_{\max} = 1$ we have $\frac{d\mu}{d\lambda} \in L^q$ for all $q \in [1, \infty]$. In view of this, one could wonder for which $q > 1$ we have $\frac{d\mu}{d\lambda} \in L^q$ in the intermediate case that $\ell_{\max} = 1$ and $\sum_{b \in \Sigma_B^1} \frac{p_b}{|DT_b(c)|} \geq 1$, i.e. if c is not superattracting for any bad map and the bad maps are not expanding on average at c.

CHAPTER 3

Decay of correlations for critically intermittent systems

This chapter is based on: [KZ22].

Abstract

For a family of random intermittent dynamical systems with a superattracting fixed point we prove that a phase transition occurs for the existence of an absolutely continuous invariant probability measure depending on the randomness parameters and the orders of the maps at the superattracting fixed point. In case the systems have an absolutely continuous invariant probability measure, we show that the systems are mixing and that correlations decay polynomially even though some of the deterministic maps present in the system have exponential decay of correlations. This contrasts other known results, where random systems adopt the best decay rate of the deterministic maps in the systems.

§3.1 Introduction

For the random systems from Chapter 2 that have an acs probability measure, one can naturally wonder about the mixing properties and decay of correlations. In this chapter we explore this further, but instead of the random maps from the previous chapter we work with adapted versions. We take a subclass of the maps from Chapter 2 and replace the right branches with the right branch of the doubling map. Under random compositions of these maps orbits then converge superexponentially fast to $\frac{1}{2}$ and diverge exponentially fast from 1, see Figure 3.1(a). The way in which we have adapted the systems from Chapter 2 very much resembles the way in which the LSV maps from (1.11) are adaptations of the standard Manneville-Pomeau maps. We work with these adaptations because they allow us to build a suitable Young tower and use the corresponding results, while preserving the main dynamical properties of the maps from Chapter 2.

We describe the systems we consider in more detail. Just as in the previous chapter we distinguish between so-called good and bad maps. We call a map $T_g : [0,1] \to [0,1]$ *good* if it is given by

$$T_g(x) = \begin{cases} 1 - 2^{r_g}(\frac{1}{2} - x)^{r_g} & \text{if} \quad x \in [0, \frac{1}{2}), \\ 2x - 1 & \text{if} \quad x \in [\frac{1}{2}, 1], \end{cases} \tag{3.1}$$

for some $r_g \geq 1$ and denote the class of good maps by $\mathfrak{G}$. A map $T_b : [0,1] \to [0,1]$ is called *bad* if it is given by

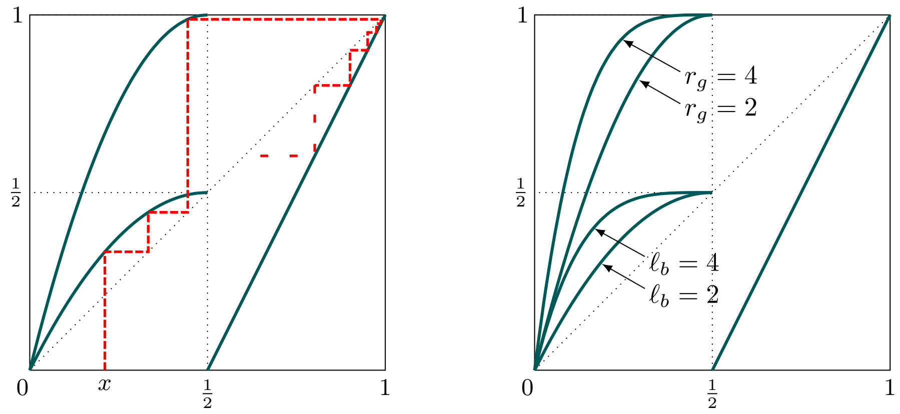
$$T_b(x) = \begin{cases} \frac{1}{2} - 2^{\ell_b - 1}(\frac{1}{2} - x)^{\ell_b} & \text{if} \quad x \in [0, \frac{1}{2}), \\ 2x - 1 & \text{if} \quad x \in [\frac{1}{2}, 1], \end{cases} \tag{3.2}$$

for some $\ell_b > 1$ and we denote the class of bad maps by $\mathfrak{B}$. The graphs of T_g and T_b are shown in Figure 3.1(b). Note that if $r_g = 1$, then T_g is equal to the doubling map. Furthermore, T_g with $r_g = 2$ and T_b with $\ell_b = 2$ are on $[0, \frac{1}{2})$ equal to the logistic maps $x \mapsto 4x(1-x)$ and $x \mapsto 2x(1-x)$, respectively. One easily computes that each good map T_g and each bad map T_b have non-positive Schwarzian derivative when restricted to $[0, \frac{1}{2})$ or $[\frac{1}{2}, 1]$. Just like in the previous chapter, let $\{T_1, \dots, T_N\} \subseteq \mathfrak{G} \cup \mathfrak{B}$ be a finite collection of good and bad maps, write $\Sigma_G = \{1 \leq j \leq N : T_j \in \mathfrak{G}\}$ and $\Sigma_B = \{1 \leq j \leq N : T_j \in \mathfrak{B}\}$ for the index sets of the good and bad maps, respectively. We assume that $\Sigma_G, \Sigma_B \neq \emptyset$. We write $\Sigma = \{1, \dots, N\} = \Sigma_G \cup \Sigma_B$ and let F be the skew product associated to $\{T_j\}_{j \in \Sigma}$ given by

$$F : \Sigma^{\mathbb{N}} \times [0,1] \to \Sigma^{\mathbb{N}} \times [0,1],\ (\omega, x) \mapsto (\tau\omega, T_{\omega_1}(x)),$$

where again τ denotes the left shift on sequences in $\Sigma^{\mathbb{N}}$.

Let $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ be a probability vector with strictly positive entries representing the probabilities with which we choose the maps from $\mathcal{T} = \{T_j\}_{j \in \Sigma}$. Furthermore, let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $\Sigma^{\mathbb{N}}$. Our first two main results establish that there exists a phase transition for the existence of an acs probability measure for $(\mathcal{T}, \boldsymbol{p})$ that is similar to the phase transition found in Chapter 2. Set $\theta = \sum_{b \in \Sigma_B} p_b \ell_b$.



(a) (b)

Figure 3.1: In (a) *we see the critically intermittent system consisting of the maps given by* (3.1) *with* $r_g = 2$ *and* (3.2) *with* $\ell_b = 2$*. The dashed lines indicate part of a random orbit of* x*. In* (b) *the graphs of* (3.1) *and* (3.2) *are depicted for several values of* r_g *and* ℓ_b*.*

Theorem 3.1.1. *If* $\theta \geq 1$*, then no acs probability measure exists for* $(\mathcal{T}, \boldsymbol{p})$*.*

Theorem 3.1.2. *If* $\theta < 1$*, then there exists a unique acs probability measure* μ *for* $(\mathcal{T}, \boldsymbol{p})$*. Moreover,* F *is mixing with respect to* $m_{\boldsymbol{p}} \times \mu$ *and the density* $\frac{d\mu}{d\lambda}$ *is bounded away from zero.*

Since μ is an acs measure, the density $\frac{d\mu}{d\lambda}$ is a fixed point of the associated Perron-Frobenius operator $P_{\mathcal{T},\boldsymbol{p}}$ being of the form as in (1.20). Moreover, Theorem 3.1.2 tells that this density $\frac{d\mu}{d\lambda}$ is bounded away from zero. Using these two statements it is easy to see that $\frac{d\mu}{d\lambda}$ blows up to infinity when approaching the points $\frac{1}{2}$ and 1 from below. See Figure 3.2 for an example.

Our second set of main results involves the decay of correlations in case $\theta < 1$. Equip $\Sigma^{\mathbb{N}} \times [0,1]$ with the metric

$$d\big((\omega, x), (\omega', y)\big) = 2^{-\min\{i \in \mathbb{N}\,:\,\omega_i \neq \omega_i'\}} + |x - y|. \tag{3.3}$$

For $\alpha \in (0,1)$, let $\mathcal{H}_\alpha$ be the class of α-Hölder continuous functions on $\Sigma^{\mathbb{N}} \times [0,1]$, i.e.

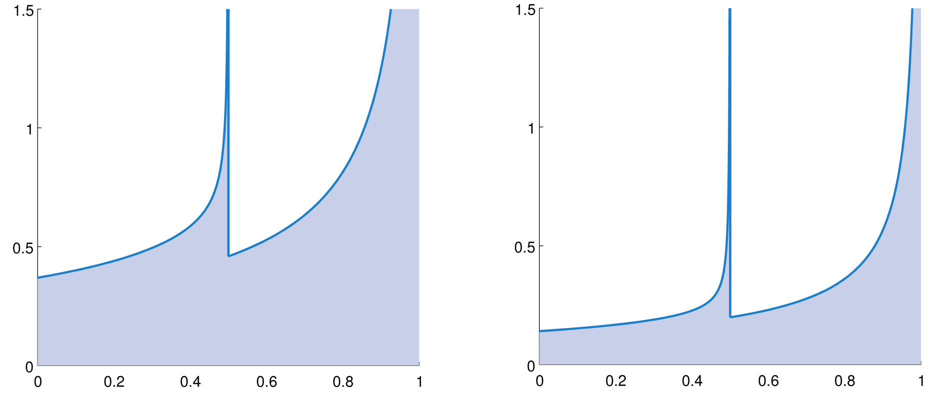
$$\mathcal{H}_\alpha = \Big\{ h : \Sigma^{\mathbb{N}} \times [0,1] \to \mathbb{R} \,\Big|\, \sup\Big\{ \frac{|h(z_1) - h(z_2)|}{d(z_1, z_2)^\alpha} : z_1, z_2 \in \Sigma^{\mathbb{N}} \times [0,1], z_1 \neq z_2 \Big\} < \infty \Big\},$$

and set

$$\mathcal{H} = \bigcup_{\alpha \in (0,1)} \mathcal{H}_\alpha.$$

For $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$ the *correlations* are defined by

$$\mathrm{Cor}_n(f, h) = \int_{\Sigma^{\mathbb{N}} \times [0,1]} f \circ F^n \cdot h \, dm_{\boldsymbol{p}} \times \mu - \int_{\Sigma^{\mathbb{N}} \times [0,1]} f \, dm_{\boldsymbol{p}} \times \mu \int_{\Sigma^{\mathbb{N}} \times [0,1]} h \, dm_{\boldsymbol{p}} \times \mu.$$



(a) (b)

Figure 3.2: Approximation of $\frac{d\mu}{d\lambda}$ in case $\Sigma_G = \{1\}$, $\Sigma_B = \{2\}$, $p_1 = \frac{7}{10}$ and $r_1 = 2$ for two different values of ℓ_2. Both pictures depict $P^{50}_{\mathcal{T},\boldsymbol{p}}(1)$ with Perron-Frobenius operator $P_{\mathcal{T},\boldsymbol{p}}$, where in (a) *we have taken $\ell_2 = \frac{3}{2}$ and in* (b) *$\ell_2 = 3$.*

Set $\ell_{\max} = \max\{\ell_b : b \in \Sigma_B\}$ and

$$\gamma_1 = \frac{\log \theta}{\log \ell_{\max}}. \tag{3.4}$$

The following result says that correlations decay at least polynomially fast with degree arbitrarily close to γ_1. Here and in the rest of this chapter the notation $f(n) = O(g(n))$ means that there exists a constant $C > 0$ and integer $N \in \mathbb{N}$ such that for each integer $n \geq N$ we have $f(n) \leq C \cdot g(n)$.

Theorem 3.1.3. *Assume that $\theta < 1$. If $\gamma \in (\gamma_1, 0)$, $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$, then $|\mathrm{Cor}_n(f,h)| = O(n^\gamma)$.*

For each $b \in \Sigma_B$ set

$$\pi_b = \sum_{j \in \Sigma_B : \ell_j \geq \ell_b} p_j \tag{3.5}$$

and let

$$\gamma_2 = 1 + \max\left\{\frac{\log \pi_b}{\log \ell_b} : b \in \Sigma_B\right\}. \tag{3.6}$$

Note that

$$\gamma_2 = \max\left\{\frac{\log(\pi_b \cdot \ell_b)}{\log \ell_b} : b \in \Sigma_B\right\} \leq \max\left\{\frac{\log \theta}{\log \ell_b} : b \in \Sigma_B\right\} = \gamma_1.$$

In our final result we show that, under additional assumptions on the parameters of the random systems, the class of observables $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$

contains functions for which the correlation decay is not faster than polynomially with degree γ_2. The notation $f(n) = \Omega(g(n))$ means that there exists a constant $C > 0$ and integer $N \in \mathbb{N}$ such that for each integer $n \geq N$ we have $f(n) \geq C \cdot g(n)$.

Theorem 3.1.4. *Assume that $\theta < 1$. Furthermore, assume that $\gamma_2 > \gamma_1 - 1$ if $\gamma_1 < -1$ and $\gamma_2 > 2\gamma_1$ if $-1 \leq \gamma_1 < 0$. Let $f \in L^\infty(\Sigma^\mathbb{N} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$ be such that both f and h are identically zero on $\Sigma^\mathbb{N} \times \big([0, \frac{1}{2}] \cup [\frac{3}{4}, 1]\big)$ and such that*

$$\int_{\Sigma^\mathbb{N} \times [0,1]} f \, dm_{\boldsymbol{p}} \times \mu \cdot \int_{\Sigma^\mathbb{N} \times [0,1]} h \, dm_{\boldsymbol{p}} \times \mu > 0.$$

Then

$$|\mathrm{Cor}_n(f,h)| = \Omega(n^{\gamma_2}).$$

In Section 3.4 we provide examples of values of ℓ_b and probability vectors $\boldsymbol{p}$ that satisfy the conditions of Theorem 3.1.4.

The proof of Theorem 3.1.2 we present below also carries over to the case that $\Sigma_B = \emptyset$. Applying Theorem 3.1.2 to the case that $\Sigma_B = \emptyset$ and $\Sigma_G = \{g\}$ contains one element yields together with [LM13, Theorem 1.5] the following result on the good maps $T_g \in \mathfrak{G}$.

Corollary 3.1.5. *For any $T_g : [0,1] \to [0,1] \in \mathfrak{G}$ the following hold.*

(a) *T_g admits an invariant probability measure μ_g that is mixing and absolutely continuous with respect to Lebesgue measure λ.*

(b) *There exists a constant $a > 0$ such that for each $f \in L^\infty([0,1], \mu_g)$ and each function $h : [0,1] \to \mathbb{R}$ of bounded variation we have*

$$|\mathrm{Cor}_{n,T_g,\mu_g}(f,h)| = O(e^{-an}),$$

where

$$\mathrm{Cor}_{n,T_g,\mu_g}(f,h) = \int f \circ T_g^n \cdot h \, d\mu_g - \int f d\mu_g \cdot \int h \, d\mu_g.$$

This result is expected in view of the exponential decay of correlations found for the unimodal maps from [KN92, Y92, Y98]. We see that test functions that fall within the scope of both this corollary and Theorems 3.1.3 and 3.1.4 have exponential decay of correlations under a single good map while under the random system with $\Sigma_B \neq \emptyset$ they have polynomial decay of correlations.[1] This indicates that a system of good maps loses its exponential decay of correlations when mixed with bad maps and instead adopts polynomial mixing rates. This is different from what has been observed for other random systems in e.g. [BBD14, BB16, BQT21], where the random systems of LSV maps under consideration adopt the highest decay rate from the rates of the individual LSV maps present in the system. See Example 1.4.3.

[1] Examples of such test functions are Lipschitz continuous functions that vanish outside of $\Sigma^\mathbb{N} \times (\frac{1}{2}, \frac{3}{4})$.

The remainder of this chapter is organised as follows. In Section 3.2 we list some preliminaries. In Section 3.3 we prove Theorem 3.1.1. The method of proof for this part is reminiscent of that in Chapter 2 in the sense that we introduce an induced system and apply Kac's Lemma on the first return times. In Section 3.3 we also give estimates on the first return times and the induced map that we use later in Section 3.4. For Theorem 3.1.2 the approach from Subsection 2.3.2 no longer works, because we have introduced a discontinuity for the bad maps (at $\frac{1}{2}$). Instead we prove Theorem 3.1.2, as well as Theorem 3.1.3 and Theorem 3.1.4, by constructing a Young tower on the inducing domain from Section 3.3 and by applying the general theory from [Y99] and [G04]. This is done in Section 3.4 and is inspired by the methods from [BBD14, BB16]. Section 3.5 contains some further results. More specifically, we show that for a specific class of test functions the upper bound from Theorem 3.1.3 can be improved and we obtain a Central Limit Theorem. We end with some final remarks.

§3.2 Preliminaries

We use this section first of all to give some properties of good and bad maps. Secondly, we give the construction of a Young tower, list some of the results from [Y99, G04] and present this in an adapted form, rephrased to our setting and only referring to the parts that are relevant for our purposes.

§3.2.1 Properties of good and bad maps

As in Section 3.1, let $T_1, \dots, T_N \in \mathfrak{G} \cup \mathfrak{B}$ be a finite collection of good and bad maps, and write $\Sigma_G = \{1 \le j \le N : T_j \in \mathfrak{G}\} \neq \emptyset$, $\Sigma_B = \{1 \le j \le N : T_j \in \mathfrak{B}\} \neq \emptyset$ and $\Sigma = \{1, \dots, N\} = \Sigma_G \cup \Sigma_B$. For ease of notation, we refer to the left branch of a map T_j by L_j, i.e. for $x \in [0, \frac{1}{2})$ we write $L_j(x) = T_j(x)$, $j \in \Sigma$. We will use $R : [\frac{1}{2}, 1] \to [0, 1]$, $x \mapsto 2x - 1$ to denote the right branch of the maps T_j.

Throughout this section and the next ones we use the notations for words and compositions of the maps T_j introduced in Section 1.4. We use the same notations for compositions of the left branches L_j, so for $\omega \in \Sigma^{\mathbb{N}}$, $n \in \mathbb{N}_0$ and $x \in [0, \frac{1}{2})$ such that $L_{\omega_j} \circ \cdots \circ L_{\omega_1}(x) \in [0, \frac{1}{2})$ for each $j = 1, \dots, n-1$ we write

$$L_{\omega_1 \cdots \omega_n}(x) = L_\omega^n(x) = \begin{cases} x, & \text{if } \quad n = 0, \\ L_{\omega_n} \circ L_{\omega_{n-1}} \circ \cdots \circ L_{\omega_1}(x), & \text{for } n \ge 1. \end{cases} \tag{3.7}$$

Also in this situation we use the same notation for finite words $\boldsymbol{u} \in \Sigma^m$, $m \ge 1$, and $n \le m$. Furthermore, for any $\boldsymbol{u} = u_1 \cdots u_k \in \Sigma^k$, $k \ge 0$, recall that we abbreviate $p_{\boldsymbol{u}} = \prod_{i=1}^k p_{u_i}$ and also, as in Chapter 2, let $\ell_{\boldsymbol{u}} = \prod_{i=1}^k \ell_{u_i}$ if $\boldsymbol{u} \in \Sigma_B^k$, where we use $p_{\boldsymbol{u}} = \ell_{\boldsymbol{u}} = 1$ in case $k = 0$. Finally, recall the notation $\Sigma^* = \bigcup_{n \ge 0} \Sigma^n$ for the collection of all finite words (including the empty word) with digits from Σ. Similarly we define Σ_G^* and Σ_B^*.

For future reference we give two lemmas on the properties of good and bad maps. The first one involves compositions of bad maps and implies in particular that orbits starting in $[0, \frac{1}{2})$ converge superexponentially fast to $\frac{1}{2}$ under iterations of bad maps.

Lemma 3.2.1. *For each* $x \in [0, \frac{1}{2})$ *and* $\boldsymbol{b} \in \Sigma_B^*$ *we have*

(i) $L_{\boldsymbol{b}}(x) = \frac{1}{2}\big(1 - (1-2x)^{\ell_{\boldsymbol{b}}}\big)$,

(ii) $L_{\boldsymbol{b}}^{-1}(x) = \frac{1}{2}\big(1 - (1-2x)^{\ell_{\boldsymbol{b}}^{-1}}\big)$.

Proof. Part (i) holds trivially if $|\boldsymbol{b}| = 0$. Now suppose for some $k \geq 0$ that (i) holds for all $\boldsymbol{b} \in \Sigma_B^*$ with $|\boldsymbol{b}| \leq k$. Let $\boldsymbol{b}b_{k+1} \in \Sigma_B^{k+1}$. Then $L_{\boldsymbol{b}}(x) \in [0, \frac{1}{2})$ and

$$\begin{aligned} L_{\boldsymbol{b}b_{k+1}}(x) &= \frac{1}{2}\Big(1 - (1 - 2L_{\boldsymbol{b}}(x))^{\ell_{b_{k+1}}}\Big) \\ &= \frac{1}{2}\Big(1 - \big((1-2x)^{\ell_{\boldsymbol{b}}}\big)^{\ell_{b_{k+1}}}\Big) = \frac{1}{2}\Big(1 - (1-2x)^{\ell_{\boldsymbol{b}b_{k+1}}}\Big). \end{aligned}$$

This proves (i), and (ii) follows easily from (i). □

For $g \in \Sigma_G$ the map $L_g : [0, \frac{1}{2}) \to [0, 1)$ is invertible and for $b \in \Sigma_B$ the map $L_b : [0, \frac{1}{2}) \to [0, \frac{1}{2})$ is invertible. For all $j \in \Sigma$ the map L_j is strictly increasing with continuous and decreasing derivative and, if T_j is not the doubling map (i.e. $j \in \Sigma_B$ or $j \in \Sigma_G$ with $r_j > 1$),

$$\max_{x \in [0, \frac{1}{2})} DL_j(x) = DT_j(0) > 1 \quad \text{and} \quad \lim_{x \uparrow \frac{1}{2}} DL_j(x) = 0.$$

This allows us to define for each $g \in \Sigma_G$ with $r_g > 1$ the point x_g as the point in $(0, \frac{1}{2})$ for which $DL_g(x_g) = 1$ and for each $b \in \Sigma_B$ the point x_b as the point in $(0, \frac{1}{2})$ for which $DL_b(x_b) = 1$.

Lemma 3.2.2. *The following hold.*

(i) *For each* $g \in \Sigma_G$ *it holds that* $L_g^{-1}\big(\frac{1}{2}\big) \leq \frac{1}{4}$.

(ii) *For each* $g \in \Sigma_G$ *with* $r_g > 1$ *it holds that* $L_g^{-1}\big(\frac{1}{2}\big) < x_g$.

(iii) *For each* $b \in \Sigma_B$ *it holds that* $L_b^{-1}\big(\frac{1}{4}\big) < x_b$.

Proof. One can compute that

$$L_g^{-1}\Big(\frac{1}{2}\Big) = \frac{1}{2}\big(1 - 2^{-1/r_g}\big) \tag{3.8}$$

for all $g \in \Sigma_G$. If $r_g > 1$, then

$$x_g = \frac{1}{2} - \big(r_g \cdot 2^{r_g}\big)^{1/(1-r_g)}.$$

For each $b \in \Sigma_B$ we have

$$L_b^{-1}\Big(\frac{1}{4}\Big) = \frac{1}{2}\big(1 - 2^{-1/\ell_b}\big) \quad \text{and} \quad x_b = \frac{1}{2}\big(1 - \ell_b^{1/(1-\ell_b)}\big). \tag{3.9}$$

Since $r_g \geq 1$ and thus $2^{-1/r_g} \geq \frac{1}{2}$ for each $g \in \Sigma_G$, (i) follows. Furthermore, one can show that $2^{-1/x-1} > (x \cdot 2^x)^{1/(1-x)}$ and $2^{-1/x} > x^{1/(1-x)}$ hold for all $x > 1$, which give (ii) and (iii), respectively. □

§3.2.2 Young towers

Let $\mathcal{F}$ be the product σ-algebra on $\Sigma^{\mathbb{N}} \times [0,1]$ given by the Borel σ-algebra's on both coordinates. We call the set on which we will induce $Y \in \mathcal{F}$. It will be defined later and will be such that the first return time map φ on Y associated to the skew product F given by

$$\varphi(\omega, x) = \inf\{n \geq 1 : F^n(\omega, x) \in Y\}$$

satisfies $\varphi(\omega, x) < \infty$ for all $(\omega, x) \in Y$. The induced transformation $F_Y : Y \to Y$ as given in Subsection 1.2.1 by $F_Y(\omega, x) = F^{\varphi(\omega,x)}(\omega, x)$ is then well defined.

On the inducing domain Y one can construct a Young tower for F. For the convenience of the reader we briefly give this construction, which is outlined in a more general fashion and in more detail in [Y99]. Let $\mathcal{P}$ be a countable partition of Y into measurable sets on which the first return time function φ is constant, called a *first return time partition*. Let φ_P be the constant value that φ assumes on P. The tower is defined by

$$\Delta = \{(z, n) \in Y \times \{0, 1, 2 \ldots\} \,:\, n < \varphi(z)\}.$$

The l-th level of the tower is $\Delta_l := \Delta \cap \{n = l\}$ and for each $P \in \mathcal{P}$ let $\Delta_{l,P} = \Delta_l \cap \{z \in P\}$. We equip Δ with the Borel σ-algebra $\mathcal{B}$ and let m be the unique measure on Δ that corresponds to $m_{\boldsymbol{p}} \times \lambda$ on each level of the tower. On Δ define the function

$$G(z, n) = \begin{cases} (z, n+1), & \text{if } n + 1 < \varphi_Y(z), \\ (F_Y(z), 0), & \text{otherwise.} \end{cases}$$

Let the induced map $G^{\varphi} : \Delta_0 \to \Delta_0$ be defined by $G^{\varphi}(v) = G^{\varphi(z)}(v)$, where $z \in Y$ is such that $v = (z, 0)$. We identify G^{φ} with F^{φ} by identifying Δ_0 with Y and using the correspondence $G^{\varphi}(z, 0) = (F^{\varphi}(z), 0)$. The *separation time* is defined for each $(z_1, l_1), (z_2, l_2) \in \Delta$ by $s((z_1, l_1), (z_2, l_2)) = 0$ if $l_1 \neq l_2$ and otherwise letting $s((z_1, l_1), (z_2, l_2)) = s(z_1, z_2)$ be given by

$$s(z_1, z_2) = \inf\{n \geq 0 \,:\, (G^{\varphi})^n(z_1, 0),\, (G^{\varphi})^n(z_2, 0) \text{ lie in distinct } \Delta_{0,P},\ P \in \mathcal{P}\}. \tag{3.10}$$

The setup from [Y99] assumes the following conditions on Δ and G:

(t1) $gcd\{\varphi_P : P \in \mathcal{P}\} = 1$;

(t2) All the sets in the construction above are measurable and $m(Y) < \infty$.

(t3) For each $P \in \mathcal{P}$ the top level $\Delta_{\varphi_P - 1, P}$ above P is mapped bijectively onto $\Delta_0 = Y \times \{0\}$ under the map G;

(t4) The partition $\eta = \{\Delta_{l,P} \,:\, P \in \mathcal{P},\, 0 \leq l \leq \varphi_P - 1\}$ generates $\mathcal{B}$.

(t5) The restrictions $G^\varphi|_{\Delta_{0,P}} : \Delta_{0,P} \to \Delta_0$ and their inverses are non-singular with respect to m, so that the Jacobian $J_m G^\varphi$ with respect to m exists and is > 0 m-a.e.

(t6) There are constants $C > 0$ and $\beta \in (0,1)$ such that for each $P \in \mathcal{P}$ and all $(z_1, 0), (z_2, 0) \in \Delta_{0,P}$,

$$\left| \frac{J_m G^\varphi(z_1, 0)}{J_m G^\varphi(z_2, 0)} - 1 \right| \leq C \beta^{s(G^\varphi(z_1,0), G^\varphi(z_2,0))}.$$

Under these conditions [Y99] yields the following results on the existence of invariant measures for the map G on the tower Δ and decay of correlations. Define for all $\delta \in (0,1)$ the following function space on Δ:

$$\mathcal{C}_\delta = \left\{ \hat{f} : \Delta \to \mathbb{R} \,\middle|\, \sup \left\{ \frac{|\hat{f}(\upsilon_1) - \hat{f}(\upsilon_2)|}{\delta^{s(\upsilon_1, \upsilon_2)}} : \upsilon_1, \upsilon_2 \in \Delta, \upsilon_1 \neq \upsilon_2 \right\} < \infty \right\}. \tag{3.11}$$

For $\upsilon \in \Delta$ let $\hat{\varphi}(\upsilon) := \inf\{n \geq 0 \,:\, G^n(\upsilon) \in \Delta_0\}$.

Theorem 3.2.3 (Theorem 1 and 3 from [Y99]). *If (t1)–(t6) hold and we have $\int_Y \varphi \, dm < \infty$, then the following statements hold.*[2]

(i) $G : \Delta \to \Delta$ admits an invariant probability measure ν that is absolutely continuous w.r.t. m;

(ii) The density $\frac{d\nu}{dm}$ is bounded away from zero and is in $\mathcal{C}_\beta$ with β as in (t6). Moreover, there is a constant $C^+ > 0$ such that for each $\Delta_{l,P}$ and each $\upsilon_1, \upsilon_2 \in \Delta_{l,P}$

$$\left| \frac{\frac{d\nu}{dm}(\upsilon_1)}{\frac{d\nu}{dm}(\upsilon_2)} - 1 \right| \leq C^+ \beta^{s(\upsilon_1, \upsilon_2)}. \tag{3.12}$$

(iii) G is exact, hence ergodic and mixing.

(iv) (Polynomial decay of correlations) If, moreover, $m(\{\upsilon \in \Delta \,:\, \hat{\varphi}(\upsilon) > n\}) = O(n^{-\alpha})$ for some $\alpha > 0$, then for all $\hat{f} \in L^\infty(\Delta, m)$, all $\delta \in (0,1)$ and all $\hat{h} \in \mathcal{C}_\delta$,

$$\left| \int_\Delta \hat{f} \circ G^n \cdot \hat{h} \, d\nu - \int_\Delta \hat{f} \, d\nu \int_\Delta \hat{h} \, d\nu \right| = O(n^{-\alpha}).$$

We will also use [G04, Theorem 6.3] by Gouëzel which, when adapted to our setting, says the following.

Theorem 3.2.4 (Theorem 6.3 from [G04]). *Let ρ be an invariant and mixing probability measure for F on $\Sigma^{\mathbb{N}} \times [0,1]$. Let $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], \rho)$ and $h \in L^1(\Sigma^{\mathbb{N}} \times [0,1], \rho)$ be such that both f and h are identically zero on $(\Sigma^{\mathbb{N}} \times [0,1]) \setminus Y$. Assume that there is a $\delta \in (0,1)$ such that the following three conditions hold.*

[2] The version in [Y99, Theorem 3] of statement (iv) above only states the result for $\delta = \beta$. Note however that statement (iv) for the case $0 < \delta < \beta$ follows from $\mathcal{C}_\delta \subseteq \mathcal{C}_\beta$ and that if $\beta < \delta < 1$, then the bound in (t6) where β is replaced by δ also holds and therefore [Y99, Theorem 3] can be applied with δ taking the role of β.

(g1) There is a constant $C^* > 0$ *such that for each* $n \geq 0$ *and* $z_1, z_2 \in \bigvee_{k=0}^{n-1} F_Y^{-n}\mathcal{P}$,

$$\Big|\log \frac{J_\rho F^\varphi(z_1)}{J_\rho F^\varphi(z_2)}\Big| \leq C^* \cdot \delta^n.$$

(g2) There is a $\zeta > 1$ *such that* $\rho(\varphi > n) = O(n^{-\zeta})$.

(g3) $\sup\left\{\frac{|h(z_1)-h(z_2)|}{\delta^{s(z_1,z_2)}} \,:\, z_1, z_2 \in Y,\, z_1 \neq z_2\right\} < \infty$.

Then there is a constant $\tilde{C} > 0$ *such that*

$$\Big|\mathrm{Cor}_{n,F,\rho}(f,h) - \Big(\sum_{k>n} \rho(\varphi > k)\Big) \int f\, d\rho \int h\, d\rho\Big| \leq \tilde{C} \cdot K_\zeta(n),$$

where

$$K_\zeta(n) = \begin{cases} n^{-\zeta}, & \text{if } \zeta > 2, \\ \frac{\log n}{n^2}, & \text{if } \zeta = 2, \\ n^{2-2\zeta}, & \text{if } \zeta \in (1,2). \end{cases}$$

Chapter 3

§3.3 Inducing the random map on $(\frac{1}{2}, \frac{3}{4})$

§3.3.1 The induced system

Define

$$\tilde{\Omega} = \{\omega \in \Sigma^{\mathbb{N}} : \omega_i \in \Sigma_G \text{ for infinitely many } i \in \mathbb{N}\},$$
$$Y = \Big\{(\omega, x) \in \tilde{\Omega} \times \Big(\frac{1}{2}, \frac{3}{4}\Big) : T_\omega^n(x) \neq \frac{1}{2} \text{ for all } n \in \mathbb{N}\Big\}.$$

The set Y, which equals $\Sigma^{\mathbb{N}} \times (\frac{1}{2}, \frac{3}{4})$ up to a set of measure zero, will be our inducing domain. Recall the definition of the first return time function

$$\varphi : Y \to \mathbb{N},\ (\omega, x) \mapsto \inf\{n \geq 1 \,:\, F^n(\omega, x) \in Y\}.$$

For each $(\omega, x) \in Y$, the following happens under iterations of F. Firstly, $T_\omega(x) = R(x) \in (0, \frac{1}{2})$. By Lemma 3.2.2 there is an $a > 1$ such that $DT_g(y) \geq a$ for all $y \in (0, L_g^{-1}(\frac{1}{2})]$ and $g \in \Sigma_G$. Combining this with Lemma 3.2.1 yields by definition of Y that

$$\kappa(\omega, x) := \inf\Big\{n \geq 1 \,:\, T_\omega^n(x) > \frac{1}{2}\Big\} < \infty. \tag{3.13}$$

Note that $\omega_{\kappa(\omega,x)} \in \Sigma_G$. Then, again since $(\omega, x) \in Y$, so $T_\omega^n(x) \neq \frac{1}{2}$ for all $n \in \mathbb{N}$,

$$l(\omega, x) := \inf\Big\{n \geq 0 \,:\, T_\omega^{\kappa(\omega,x)+n} \in \Big(\frac{1}{2}, \frac{3}{4}\Big)\Big\} < \infty \tag{3.14}$$

and $T_\omega^{\kappa(\omega,x)+l(\omega,x)}(x) = R^{l(\omega,x)} \circ T_\omega^{\kappa(\omega,x)}(x)$. Thus

$$\varphi(\omega, x) = \kappa(\omega, x) + l(\omega, x) < \infty \tag{3.15}$$

for all $(\omega, x) \in Y$. We will first derive an estimate for $l(\omega, x)$.

For each $g \in \Sigma_G$ let $J_g = (L_g^{-1}(\frac{1}{2}), \frac{1}{2})$. The intervals J_g are such that if $T_\omega^n(x) \in J_g$ and $\omega_{n+1} = g$, then $T_\omega^{n+1}(x) > \frac{1}{2}$. For each $b \in \Sigma_B$ let $J_b = [x_b, \frac{1}{2})$ with x_b as defined above Lemma 3.2.2. Define

$$m(\omega, x) := \inf\{n \geq 1 \,:\, T_\omega^n(x) \in J_{\omega_{n+1}}\} < \kappa(\omega, x). \tag{3.16}$$

If $\omega_{m(\omega,x)+1} \in \Sigma_G$, then $T_\omega^{m(\omega,x)+1}(x) > \frac{1}{2}$ by the definition of the intervals J_g. If $\omega_{m(\omega,x)+1} \in \Sigma_B$, then $T_\omega^{m(\omega,x)+1}(x) \in (\frac{1}{4}, \frac{1}{2})$ by Lemma 3.2.2(iii). We then see by Lemma 3.2.2(i) that the number $m(\omega, x)$ is such that after this time any application of a good map will bring the orbit of x in the interval $\left(\frac{1}{2}, 1\right)$. We have the following estimate on l.

Lemma 3.3.1. *Let* $(\omega, x) \in Y$. *Write* $\boldsymbol{d} = \omega_{m(\omega,x)+1} \cdots \omega_{\kappa(\omega,x)-1} \in \Sigma_B^*$ *and* $g = \omega_{\kappa(\omega,x)} \in \Sigma_G$. *Then*

$$l(\omega, x) \geq \ell_{\boldsymbol{d}} r_g \frac{\log\left((1 - 2 \cdot T_\omega^{m(\omega,x)}(x))^{-1}\right)}{\log 2} - 2.$$

Proof. Set $y = T_\omega^{m(\omega,x)}(x)$. It follows from Lemma 3.2.1 that

$$T_\omega^{\kappa(\omega,x)}(x) = L_g \circ L_{\boldsymbol{d}}(y) = 1 - (1 - 2y)^{\ell_{\boldsymbol{d}} r_g}.$$

By definition of R, it follows that $l(\omega, x)$ is equal to the minimal $l \in \mathbb{N}_0$ such that

$$2^l \cdot (1 - 2y)^{\ell_{\boldsymbol{d}} r_g} > \frac{1}{4}. \tag{3.17}$$

Solving for l gives the lemma. □

With this estimate we can now apply Kac's Lemma to prove Theorem 3.1.1.

Proof of Theorem 3.1.1. Suppose μ is an acs probability measure for $(\mathcal{T}, \boldsymbol{p})$. We first show that $\mu((\frac{1}{2}, \frac{3}{4})) > 0$. Since $F^2(Y)$ equals $\Sigma^{\mathbb{N}} \times [0, 1]$ up to some set of $m_{\boldsymbol{p}} \times \lambda$-measure zero and by (3.15) all $(\omega, x) \in Y$ have a finite first return time to Y under F, it follows that $\Sigma^{\mathbb{N}} \times [0, 1]$ equals $\bigcup_{n=0}^{\infty} F^{-n} Y$ up to some set of $m_{\boldsymbol{p}} \times \lambda$-measure zero. Since μ is absolutely continuous with respect to λ, we obtain that

$$1 = m_{\boldsymbol{p}} \times \mu\big(\Sigma^{\mathbb{N}} \times [0, 1]\big) \leq \sum_{n=0}^{\infty} m_{\boldsymbol{p}} \times \mu(F^{-n} Y) = \sum_{n=0}^{\infty} \mu\Big(\Big(\frac{1}{2}, \frac{3}{4}\Big)\Big),$$

from which we see that we indeed must have $\mu((\frac{1}{2}, \frac{3}{4})) > 0$. Using the continuity of the measure, we know there exists an $a > \frac{1}{2}$ such that $\mu((a, \frac{3}{4})) > 0$. Note that for all $(\omega, x) \in \Sigma^{\mathbb{N}} \times (a, \frac{3}{4}) \cap Y$ we have

$$T_\omega^{m(\omega,x)}(x) \geq R(a). \tag{3.18}$$

Fix a $g \in \Sigma_G$ and consider the subsets $A_{\boldsymbol{b}} = [\boldsymbol{b}g] \times (a, \frac{3}{4}) \cap Y$, $\boldsymbol{b} \in \Sigma_B^*$, of $\Sigma^{\mathbb{N}} \times (a, \frac{3}{4}) \cap Y$. Set $p_B = \sum_{b \in \Sigma_B} p_b$. Then by (3.15), (3.18) and Lemma 3.3.1 we get

$$\begin{aligned}
\int_Y \varphi \, dm_{\boldsymbol{p}} \times \mu &\geq \sum_{\boldsymbol{b} \in \Sigma_B^*} \int_{A_{\boldsymbol{b}}} l(\omega, x) \, dm_{\boldsymbol{p}} \times \mu \\
&\geq \mu\Big(\Big(a, \frac{3}{4}\Big)\Big) \cdot \sum_{\boldsymbol{b} \in \Sigma_B^*} p_{\boldsymbol{b}g} \cdot \Big(\ell_{\boldsymbol{b}} r_g \cdot \frac{\log\big((1 - 2 \cdot R(a))^{-1}\big)}{\log 2} - 2\Big) \\
&= \mu\Big(\Big(a, \frac{3}{4}\Big)\Big) \cdot p_g \cdot \Big(r_g \cdot \frac{\log\big((1 - 2 \cdot R(a))^{-1}\big)}{\log 2} \sum_{\boldsymbol{b} \in \Sigma_B^*} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} - 2 \sum_{\boldsymbol{b} \in \Sigma_B^*} p_{\boldsymbol{b}}\Big) \\
&= M_1 \cdot \sum_{k \geq 0} \theta^k - M_2 \cdot \frac{1}{1 - p_B},
\end{aligned}$$

with $M_1 = \mu((a, \frac{3}{4})) \cdot p_g r_g \cdot \frac{\log((1-2 \cdot R(a))^{-1})}{\log 2} > 0$ and $M_2 = 2\mu((a, \frac{3}{4})) \cdot p_g$. It now follows from $\theta \geq 1$ that

$$\int_Y \varphi \, dm_{\boldsymbol{p}} \times \mu = \infty. \tag{3.19}$$

On the other hand, since μ is a probability measure by assumption, we obtain from the Ergodic Decomposition Theorem, see e.g. [EW11, Theorem 6.2], that there exists a probability space $(E, \mathcal{E}, \nu)$ and a measurable map $e \mapsto \mu_e$ with μ_e an F-invariant ergodic probability measure for ν-a.e. $e \in E$, such that

$$\int_Y \varphi dm_{\boldsymbol{p}} \times \mu = \int_E \Big(\int_Y \varphi d\mu_e\Big) d\nu(e).$$

For each $e \in E$ for which μ_e is an F-invariant ergodic probability measure we have $\int_Y \varphi d\mu_e = \mu_e(X) = 1$ if $\mu_e(Y) > 0$ by Kac's Lemma, i.e. Lemma 1.2.13, and we have $\int_Y \varphi d\mu_e = 0$ if $\mu_e(Y) = 0$. This gives

$$\int_Y \varphi dm_{\boldsymbol{p}} \times \mu \leq \nu(E) = 1,$$

which is in contradiction with (3.19). □

§3.3.2 Estimates on the first return time

From now on we only consider the case $\theta < 1$. We first define a first return time partition for F to Y. For any $u \in \Sigma$, $g \in \Sigma_G$ and $\boldsymbol{s}, \boldsymbol{w} \in \Sigma^*$ write

$$P_{u\boldsymbol{s}g\boldsymbol{w}} := ([u\boldsymbol{s}g\boldsymbol{w}] \cap \tilde{\Omega}) \times (R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{s}} \circ R)^{-1}\Big(\frac{1}{2}, \frac{3}{4}\Big)$$

and define the collection of sets

$$\mathcal{P} = \Big\{P_{u\boldsymbol{s}g\boldsymbol{w}} \, : \, u \in \Sigma, \, g \in \Sigma_G, \, \boldsymbol{s}, \boldsymbol{w} \in \Sigma^*\Big\}. \tag{3.20}$$

See Figure 3.3 for an illustration.

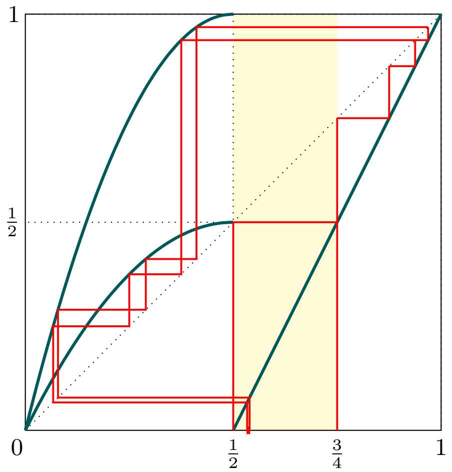


Figure 3.3: Example of a first return time partition element $P_{\boldsymbol{usgw}}$ with $|\boldsymbol{s}| = 2$ and $|\boldsymbol{w}| = 3$ projected onto $[0, 1]$. The yellow area indicates the inducing domain Y projected onto $[0, 1]$.

Proposition 3.3.2. *The collection $\mathcal{P}$ is a* first return time partition *of F to Y and for all (ω, x) in a set $P_{\boldsymbol{usgw}}$ it holds that $\kappa(\omega, x) = 2 + |\boldsymbol{s}|$ and $l(\omega, x) = |\boldsymbol{w}|$, so*

$$\varphi(\omega, x) = 2 + |\boldsymbol{s}| + |\boldsymbol{w}|. \tag{3.21}$$

Proof. Let $(\omega, x) \in Y$. Since $\kappa(\omega, x), l(\omega, x) < \infty$ it is clear that we can find suitable $u \in \Sigma$, $g \in \Sigma_G$ and $\boldsymbol{s}, \boldsymbol{w} \in \Sigma^*$ so that $(\omega, x) \in P_{\boldsymbol{usgw}}$, so $\mathcal{P}$ covers Y. Now fix a set $P_{\boldsymbol{usgw}} \in \mathcal{P}$. By the definition of the set $(R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{s}} \circ R)^{-1}(\frac{1}{2}, \frac{3}{4})$ one has for any $(\omega, x) \in P_{\boldsymbol{usgw}}$ that

$$\begin{aligned} T_\omega^n(x) &< \tfrac{1}{2}, \quad 1 \le n \le 1 + |\boldsymbol{s}|, \\ T_\omega^n(x) &> \tfrac{3}{4}, \quad 2 + |\boldsymbol{s}| \le n \le 1 + |\boldsymbol{s}| + |\boldsymbol{w}|, \\ T_\omega^{2+|\boldsymbol{s}|+|\boldsymbol{w}|}(x) &\in (\tfrac{1}{2}, \tfrac{3}{4}). \end{aligned}$$

Hence, $\kappa(\omega, x) = 2 + |\boldsymbol{s}|$ and $l(\omega, x) = |\boldsymbol{w}|$ for each $(\omega, x) \in P_{\boldsymbol{usgw}}$, and (3.21) follows from (3.15). From this we immediately obtain that the sets in $\mathcal{P}$ are disjoint. To see this, suppose there are two different sets $P_{\boldsymbol{usgw}}, P_{\tilde{u}\tilde{\boldsymbol{s}}\tilde{g}\tilde{\boldsymbol{w}}} \in \mathcal{P}$ with $P_{\boldsymbol{usgw}} \cap P_{\tilde{u}\tilde{\boldsymbol{s}}\tilde{g}\tilde{\boldsymbol{w}}} \neq \emptyset$ and let $(\omega, x) \in P_{\boldsymbol{usgw}} \cap P_{\tilde{u}\tilde{\boldsymbol{s}}\tilde{g}\tilde{\boldsymbol{w}}}$. Then without loss of generality we can assume that $[\tilde{u}\tilde{\boldsymbol{s}}\tilde{g}\tilde{\boldsymbol{w}}] \subseteq [u\boldsymbol{s}g\boldsymbol{w}]$, where the inclusion is strict. But this would give that

$$\varphi(\omega, x) = 2 + |\boldsymbol{s}| + |\boldsymbol{w}| < 2 + |\tilde{\boldsymbol{s}}| + |\tilde{\boldsymbol{w}}| = \varphi(\omega, x),$$

a contradiction. □

For the estimates we give below, we split the word $\boldsymbol{s}$ into two parts $\boldsymbol{s} = \boldsymbol{vb}$, where $\boldsymbol{b}$ specifies the string of bad digits that immediately precedes $\omega_{\kappa(\omega,x)}$. In other words, if we write

$$A_G = \{\boldsymbol{v} \in \Sigma^* \, : \, v_{|\boldsymbol{v}|} \in \Sigma_G\}$$

for the set of words that end with a good digit, then for any $\boldsymbol{s} \in \Sigma^*$ there are unique $\boldsymbol{v} \in A_G$ and $\boldsymbol{b} \in \Sigma_B^*$ such that $\boldsymbol{s} = \boldsymbol{vb}$. Recall γ_1 and γ_2 from (3.4) and (3.6), respectively. In the remainder of this subsection we prove the following result.

Proposition 3.3.3. *Suppose $\theta < 1$. Then the following statements hold.*

(i) $\int_Y \varphi \, dm_{\boldsymbol{p}} \times \lambda < \infty$.

(ii) $m_{\boldsymbol{p}} \times \lambda(\varphi > n) = O(n^{\gamma-1})$ *for any* $\gamma \in (\gamma_1, 0)$.

(iii) $m_{\boldsymbol{p}} \times \lambda(\varphi > n) = \Omega(n^{\gamma_2 - 1})$.

For the proof of Proposition 3.3.3 we will first prove three lemmas. Write

$$s = \Big(\min\Big\{\Big\{DL_g\Big(L_g^{-1}\Big(\frac{1}{2}\Big)\Big) : g \in \Sigma_G\Big\} \cup \Big\{DL_b\Big(L_b^{-1}\Big(\frac{1}{4}\Big)\Big) : b \in \Sigma_B\Big\}\Big)^{-1}. \quad (3.22)$$

The number $\frac{1}{s}$ will serve below as a lower bound on the derivative of the maps T_j in some situation. Using Lemma 3.2.2, we see that $s \in (0,1)$.

Lemma 3.3.4. *For each $n \in \mathbb{N}$ we have*

$$m_{\boldsymbol{p}} \times \lambda(\varphi > n) \le \frac{1}{4} \cdot \sum_{j=0}^{\infty} s^j \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}},$$

$$m_{\boldsymbol{p}} \times \lambda(\varphi > n) \ge \frac{1}{4} \cdot \min\{p_g : g \in \Sigma_G\} \cdot \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\min}^{-1}}}.$$

Proof. For $P = P_{\boldsymbol{uvbgw}} \in \mathcal{P}$ we know from Proposition 3.3.2 that the first return time is constant on P and equal to $\varphi_P = 2 + |\boldsymbol{v}| + |\boldsymbol{b}| + |\boldsymbol{w}|$. Let $n \in \mathbb{N}$. Then

$$m_{\boldsymbol{p}} \times \lambda(\varphi > n) = \sum_{P : \varphi_P > n} m_{\boldsymbol{p}} \times \lambda(P).$$

To obtain the desired lower bound on $m_{\boldsymbol{p}} \times \lambda(\varphi > n)$ we only consider those $P = P_{\boldsymbol{uvbgw}} \in \mathcal{P}$ where $\boldsymbol{v} = \epsilon$ is the empty word. From Lemma 3.2.1 we get

$$(R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{b}})^{-1}\Big(\frac{1}{2}, \frac{3}{4}\Big) = \Big(\frac{1}{2}\Big(1 - \frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}}\Big), \frac{1}{2}\Big(1 - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}}\Big)\Big). \quad (3.23)$$

Since R has derivative 2, we then have

$$\lambda\Big((R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{b}} \circ R)^{-1}\Big(\frac{1}{2}, \frac{3}{4}\Big)\Big) = \frac{1}{4}\Big(\frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}} - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1} r_g^{-1}}}\Big)$$

and thus

$$\begin{aligned}
&m_{\boldsymbol{p}} \times \lambda(\varphi > n)\\
&\geq \sum_{u\in\Sigma}\sum_{g\in\Sigma_G}\sum_{\boldsymbol{b}\in\Sigma_B^*}\sum_{\substack{\boldsymbol{w}\in\Sigma^*:\\ |\boldsymbol{w}|\geq\max\{n-2-|\boldsymbol{b}|,0\}}} \frac{m_{\boldsymbol{p}}([u\boldsymbol{b}g\boldsymbol{w}])}{4}\Big(\frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big)\\
&= \sum_{g\in\Sigma_G}\frac{p_g}{4}\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k} p_{\boldsymbol{b}}\sum_{l=\max\{n-2-k,0\}}^{\infty}\Big(\frac{1}{2^{(l+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(l+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big) \qquad (3.24)\\
&\geq \frac{1}{4}\cdot\min\{p_g \,:\, g\in\Sigma_G\}\cdot\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}\frac{p_{\boldsymbol{b}}}{2^{\max\{n-1-k,1\}\ell_{\boldsymbol{b}}^{-1}r_{\min}^{-1}}}.
\end{aligned}$$

For the upper bound, we look for the smallest derivative to bound the length of $(R^{|\boldsymbol{w}|}\circ L_g\circ L_{\boldsymbol{b}}\circ L_{\boldsymbol{v}}\circ R)^{-1}(\frac{1}{2},\frac{3}{4})$. If $\boldsymbol{v}=\epsilon$, then from above we see

$$\lambda\Big((R^{|\boldsymbol{w}|}\circ L_g\circ L_{\boldsymbol{b}}\circ R)^{-1}\Big(\frac{1}{2},\frac{3}{4}\Big)\Big) = \frac{s^{|\boldsymbol{v}|}}{4}\Big(\frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big).$$

On the other hand, if $\boldsymbol{v}=v_1\cdots v_j$ with $j\geq 1$, then $v_j\in\Sigma_G$. We have

$$(R^{|\boldsymbol{w}|}\circ L_g\circ L_{\boldsymbol{b}})^{-1}\Big(\frac{1}{2},\frac{3}{4}\Big)\subseteq\Big(0,\frac{1}{2}\Big).$$

As follows from before s from (3.22) represents the smallest possible shrinkage factor when applying $L_{v_j}^{-1}$. If $j\geq 2$, then by Lemma 3.2.2(i) we have $L_{v_{j-1}}^{-1}(L_{v_j}^{-1}(\frac{1}{2}))\leq L_{v_{j-1}}^{-1}(\frac{1}{4})$. Hence, $s^{-|\boldsymbol{v}|}$ is a lower bound for the derivative of $L_{\boldsymbol{v}}$ for any $\boldsymbol{v}\in A_G$ on $(R^{|\boldsymbol{w}|}\circ L_g\circ L_{\boldsymbol{b}})^{-1}(\frac{1}{2},\frac{3}{4})$. It then follows from (3.23) that

$$\lambda\Big((R^{|\boldsymbol{w}|}\circ L_g\circ L_{\boldsymbol{b}}\circ L_{\boldsymbol{v}}\circ R)^{-1}\Big(\frac{1}{2},\frac{3}{4}\Big)\Big)\leq\frac{s^{|\boldsymbol{v}|}}{4}\Big(\frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big).$$

Writing $f(n,j,k)=\max(n-2-j-k,0)$, we thus obtain that

$$\begin{aligned}
&m_{\boldsymbol{p}}\times\lambda(\varphi>n)\\
&\leq \sum_{g\in\Sigma_G}\sum_{\boldsymbol{v}\in A_G}\sum_{\boldsymbol{b}\in\Sigma_B^*}\sum_{\substack{\boldsymbol{w}\in\Sigma^*:\\ |\boldsymbol{w}|\geq f(n,|\boldsymbol{v}|,|\boldsymbol{b}|)}}\frac{p_{\boldsymbol{v}\boldsymbol{b}g\boldsymbol{w}}s^{|\boldsymbol{v}|}}{4}\Big(\frac{1}{2^{(|\boldsymbol{w}|+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(|\boldsymbol{w}|+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big)\\
&\leq \frac{1}{4}\sum_{j=0}^{\infty}s^j\sum_{k=0}^{\infty}\sum_{g\in\Sigma_G}\sum_{\boldsymbol{b}\in\Sigma_B^k}p_{\boldsymbol{b}}\sum_{l=f(n,j,k)}^{\infty}\Big(\frac{1}{2^{(l+1)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}} - \frac{1}{2^{(l+2)\ell_{\boldsymbol{b}}^{-1}r_g^{-1}}}\Big) \qquad (3.25)\\
&\leq \frac{1}{4}\sum_{j=0}^{\infty}s^j\sum_{k=0}^{\infty}\sum_{\boldsymbol{b}\in\Sigma_B^k}\frac{p_{\boldsymbol{b}}}{2^{(f(n,j,k)+1)\ell_{\boldsymbol{b}}^{-1}r_{\max}^{-1}}}.
\end{aligned}$$

This gives the result. □

The next lemma gives estimates for the last part of the expression on the right-hand side of the first inequality from Lemma 3.3.4 for an initial range of values of j. The number of values j for which we obtain an upper bound for the double sum grows logarithmically with n.

Lemma 3.3.5. *Let $\gamma \in (\gamma_1, 0)$, and for each $n \in \mathbb{N}$ define $j(n) = \lfloor \frac{(\gamma-1)\log n}{\log s} \rfloor$. Then there exist $C_1 > 0$ and $n_1 \in \mathbb{N}$ such that for all integers $n \geq n_1$ and $j = 0, 1, \ldots, j(n)$ we have*

$$\sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}} \leq C_1 \cdot n^{\gamma-1}. \tag{3.26}$$

Proof. We will split the sum over k in (3.26) into three pieces: $0 \leq k \leq k_1(n)$, $k_1(n)+1 \leq k \leq k_2(n)$ and $k > k_2(n)$. To define $k_1(n)$, let $a = \frac{\gamma \log \ell_{\max}}{\log \theta} \in (0,1)$. Then for each $n \in \mathbb{N}$ set $k_1(n) = \lfloor \frac{\log(n^a)}{\log \ell_{\max}} \rfloor = \lfloor \frac{\gamma \log n}{\log \theta} \rfloor$. The values a and $k_1(n)$ are such that $\theta^{k_1(n)+1} \leq n^{\gamma}$ and $\ell_{\max}^{k_1(n)} \leq n^a$. Since $j(n) + k_1(n) = O(\log n)$, we can find an $N_0 \in \mathbb{N}$ and a constant $K_1 > 0$ such that for all integers $n \geq N_0$ we have $n - 1 - j(n) - k_1(n) > 1$ and $(n-1-j(n)-k_1(n)) \cdot r_{\max}^{-1} \geq K_1 \cdot n$. Then, for all $n \geq N_0$, $0 \leq j \leq j(n)$ and $0 \leq k \leq k_1(n)$,

$$\max(n-1-j-k, 1) \geq \max(n-1-j(n)-k_1(n), 1) = n-1-j(n)-k_1(n)$$

and $\ell_{\boldsymbol{b}} \leq \ell_{\max}^{k_1(n)} \leq n^a$ for all $\boldsymbol{b} \in \Sigma_B^k$. Setting $p_B = \sum_{b \in \Sigma_B} p_b$ as before, this together gives

$$\begin{aligned} \sum_{k=0}^{k_1(n)} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}} &\leq \sum_{k=0}^{\infty} \frac{p_B^k}{2^{(n-1-j(n)-k_1(n))n^{-a} r_{\max}^{-1}}} \\ &\leq \frac{2^{-K_1 \cdot n^{1-a}}}{1-p_B}. \end{aligned} \tag{3.27}$$

Secondly, for each $n \in \mathbb{N}$ set $k_2(n) = \lceil \frac{1}{2}(n-1-j(n)) \rceil$ and take an integer $N_1 \geq N_0$ and constant $K_2 > 0$ such that for all integers $n \geq N_1$ we have $k_2(n) \geq k_1(n) + 1$ and $\frac{1}{2}(n-1-j(n)) - 1 \geq K_2 \cdot n$. Noting that for each $d > 1$ the function f on $\mathbb{R}$ given by $f(x) = \frac{x}{d^x}$ has maximal value $\frac{1}{e \log d} < \frac{1}{\log d}$, we obtain for all integers $n \geq N_1$,

$$\begin{aligned} \sum_{k=k_1(n)+1}^{k_2(n)} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}} &= \sum_{k=k_1(n)+1}^{k_2(n)} \sum_{\boldsymbol{b} \in \Sigma_B^k} p_{\boldsymbol{b}} \ell_{\boldsymbol{b}} \frac{\ell_{\boldsymbol{b}}^{-1}}{(2^{(n-1-j-k) r_{\max}^{-1}})^{\ell_{\boldsymbol{b}}^{-1}}} \\ &\leq \sum_{k=k_1(n)+1}^{k_2(n)} \theta^k \frac{r_{\max}}{(n-1-j-k)\log 2} \\ &\leq \frac{\theta^{k_1(n)+1}}{1-\theta} \cdot \frac{r_{\max}}{(n-1-j(n)-k_2(n))\log 2} \\ &\leq \frac{r_{\max}}{(1-\theta)K_2 \log 2} \cdot n^{\gamma-1}. \end{aligned} \tag{3.28}$$

Finally, for each $n \geq N_1$ we have

$$\sum_{k=k_2(n)+1}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}} \leq \sum_{k=k_2(n)+1}^{\infty} p_B^k = \frac{p_B^{k_2(n)+1}}{1-p_B} \leq \frac{p_B^{K_2 \cdot n}}{1-p_B}. \tag{3.29}$$

Combining (3.27), (3.28) and (3.29) yields

$$\sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-j-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\max}^{-1}}} \leq \frac{2^{-K_1 \cdot n^{1-a}} + p_B^{K_2 \cdot n}}{1 - p_B} + \frac{r_{\max}}{(1-\theta)K_2 \log 2} n^{\gamma - 1}.$$

Since the first term on the right-hand side decreases superpolynomially fast in n, this yields the existence of a constant $C_1 > 0$ and integer $n_1 \geq N_1$ for which the statement of the lemma holds. □

Lemma 3.3.6. *There exist $C_2 > 0$ and $n_2 \in \mathbb{N}$ such that for each integer $n \geq N_2$ we have*

$$\sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\min}^{-1}}} \geq C_2 \cdot n^{\gamma_2 - 1}.$$

Proof. Let $b \in \Sigma_B$ be such that $\gamma_2 = 1 + \frac{\log \pi_b}{\log \ell_b}$ with π_b as in (3.5). For each $k \in \mathbb{N}$ let

$$A_k = \{\boldsymbol{b} = b_1 \cdots b_k \in \Sigma_B^k : \ell_{b_j} \geq \ell_b \text{ for each } j = 1, \ldots, k\}.$$

Then $\sum_{\boldsymbol{b} \in A_k} p_{\boldsymbol{b}} = \pi_b^k$ and for each $\boldsymbol{b} \in A_k$ we have $\ell_{\boldsymbol{b}} \geq \ell_b^k$. This gives

$$\begin{aligned} \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in \Sigma_B^k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\min}^{-1}}} &\geq \sum_{k=0}^{\infty} \sum_{\boldsymbol{b} \in A_k} \frac{p_{\boldsymbol{b}}}{2^{\max(n-1-k,1)\ell_{\boldsymbol{b}}^{-1} r_{\min}^{-1}}} \\ &\geq \sum_{k=0}^{\infty} \frac{\pi_b^k}{2^{\max(n-1-k,1)\ell_b^{-k} r_{\min}^{-1}}}. \end{aligned} \tag{3.30}$$

For each $n \in \mathbb{N}$ we define $k_3(n) = \lceil \frac{(\gamma_2 - 1)\log n}{\log \pi_b} \rceil = \lceil \frac{\log n}{\log \ell_b} \rceil$. Then $\pi_b^{k_3(n)-1} \geq n^{\gamma_2 - 1}$ and $\ell_b^{-k_3(n)} \leq n^{-1}$. We take $N_2 \in \mathbb{N}$ and $K_3 > 0$ such that for each integer $n \geq N_2$ we have $n - 1 - k_3(n) \geq 1$ and $(n - 1 - k_3(n)) \cdot n^{-1} \cdot r_{\min}^{-1} \leq K_3$. Then we get

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{\pi_b^k}{2^{\max(n-1-k,1)\ell_b^{-k} r_{\min}^{-1}}} &\geq \pi_b^{k_3(n)} \frac{1}{2^{(n-1-k_3(n))\ell_b^{-k_3(n)} r_{\min}^{-1}}} \\ &\geq \pi_b \cdot n^{\gamma_2 - 1} \frac{1}{2^{(n-1-k_3(n)) \cdot n^{-1} \cdot r_{\min}^{-1}}} \\ &\geq \pi_b \cdot 2^{-K_3} \cdot n^{\gamma_2 - 1} \end{aligned} \tag{3.31}$$

for each $n \geq N_2$. Combining (3.30) and (3.31) now yields the result with $C_2 = \pi_b \cdot 2^{-K_3}$. □

Proof of Proposition 3.3.3. First of all, note that

$$\int_Y \varphi \, dm_{\boldsymbol{p}} \times \lambda = \sum_{n=2}^{\infty} n \cdot m_{\boldsymbol{p}} \times \lambda(\varphi = n) \leq 2 \cdot \sum_{n=1}^{\infty} m_{\boldsymbol{p}} \times \lambda(\varphi > n).$$

CHAPTER 3

Since for each $\gamma < 0$ we have $\sum_{n=1}^{\infty} n^{\gamma-1} < \infty$, (i) follows from (ii). For (ii), let $\gamma \in (\gamma_1, 0)$. It follows from Lemma 3.3.4 and Lemma 3.3.5 that for each integer $n \geq n_1$ we have

$$\begin{aligned} m_{\boldsymbol{p}} \times \lambda(\varphi > n) &\leq \frac{C_1}{4} \cdot n^{\gamma-1} \cdot \sum_{j=0}^{j(n)} s^j + \frac{1}{4} \sum_{j=j(n)+1}^{\infty} s^j \sum_{k=0}^{\infty} p_B^k \\ &\leq \frac{C_1}{4(1-s)} \cdot n^{\gamma-1} + \frac{1}{4(1-s)(1-p_B)} \cdot s^{j(n)+1}. \end{aligned}$$

By the definition of $j(n)$, we have $s^{j(n)+1} \leq n^{\gamma-1}$, which gives (ii).

Finally, it follows from Lemma 3.3.4 and Lemma 3.3.6 that for each integer $n \geq n_2$ we have

$$m_{\boldsymbol{p}} \times \lambda(\varphi > n) \geq \frac{\min\{p_g : g \in \Sigma_G\} \cdot C_2}{4} \cdot n^{\gamma_2 - 1}.$$ □

§3.3.3 Estimates on the induced map

Recall that $F^\varphi(\omega, x) = (\tau^{\varphi(\omega,x)}\omega, T_\omega^{\varphi(\omega,x)}(x))$. The second part of the next lemma shows in particular that F^φ projected on the second coordinate is expanding.

Lemma 3.3.7. *Let $(\omega, x) \in Y$.*

(i) For each $j = 1, \ldots, \varphi(\omega, x) - 1$ we have $DT_{\tau^j\omega}^{\varphi(\omega,x)-j}(T_\omega^j(x)) \geq \frac{1}{2}$.

(ii) $DT_\omega^{\varphi(\omega,x)}(x) \geq 2$.

Proof. Let $(\omega, x) \in Y$. Recall the definitions of $\kappa = \kappa(\omega, x)$ from (3.13), $l = l(\omega, x)$ from (3.14) and $m = m(\omega, x)$ from (3.16). Write[3]

$$\begin{aligned} \boldsymbol{u} &= \omega_2 \cdots \omega_m \in \Sigma^*, \\ \boldsymbol{d} &= \omega_{m+1} \cdots \omega_{\kappa-1} \in \Sigma_B^*, \\ g &= \omega_\kappa \in \Sigma_G. \end{aligned}$$

Then

$$T_\omega^{\varphi(\omega,x)}(x) = R^l \circ L_g \circ L_{\boldsymbol{d}} \circ L_{\boldsymbol{u}} \circ R(x).$$

We have $DL_b(y) \geq 1$ for all $y \in [0, x_b)$ and all $b \in \Sigma_B$. For $v \in \Sigma_G$ with $r_v > 1$ we obtained in Lemma 3.2.2(ii) that $x_v > L_v^{-1}(\frac{1}{2})$ and hence $DL_v(y) \geq 1$ for all $v \in \Sigma_G$ and $y \in [0, L_v^{-1}(\frac{1}{2}))$. It follows from the definition of m that, for each $j \in \{1, \ldots, m-1\}$,

$$DL_{u_j \cdots u_{m-1}}(L_{u_1 \cdots u_{j-1}} \circ R(x)) = \prod_{i=j}^{m-1} DL_{u_i}(L_{u_1 \cdots u_{i-1}} \circ R(x)) \geq 1. \tag{3.32}$$

[3] We use different letters here than for the partition elements $P_{\boldsymbol{uvbgw}}$ from $\mathcal{P}$, since the subdivision here is different (and (ω, x)-dependent).

Let $\boldsymbol{q}, \boldsymbol{t} \in \Sigma_B^*$ be any two words such that $\boldsymbol{d} = \boldsymbol{qt}$. Using Lemma 3.2.1 we find that for each $y \in [0, \frac{1}{2})$,

$$D(L_g \circ L_{\boldsymbol{t}})(y) = 2\ell_{\boldsymbol{t}} r_g (1-2y)^{\ell_{\boldsymbol{t}} r_g - 1}. \tag{3.33}$$

Furthermore, from (3.17) we see that

$$2^l \geq \frac{1}{4}(1 - 2L_{\boldsymbol{u}} \circ R(x))^{-\ell_{\boldsymbol{d}} r_g} \tag{3.34}$$

and applying Lemma 3.2.1 to $L_{\boldsymbol{q}}$ gives

$$(1 - 2L_{\boldsymbol{q}} \circ L_{\boldsymbol{u}} \circ R(x))^{\ell_{\boldsymbol{t}} r_g - 1} = (1 - 2L_{\boldsymbol{u}} \circ R(x))^{\ell_{\boldsymbol{d}} r_g - \ell_{\boldsymbol{q}}}. \tag{3.35}$$

Combining (3.33), (3.34) and (3.35) yields

$$\begin{aligned} &D(R^l \circ L_g \circ L_{\boldsymbol{t}})(L_{\boldsymbol{q}} \circ L_{\boldsymbol{u}} \circ R(x)) \\ &\quad= 2^l \cdot 2\ell_{\boldsymbol{t}} r_g (1 - 2L_{\boldsymbol{q}} \circ L_{\boldsymbol{u}} \circ R(x))^{\ell_{\boldsymbol{t}} r_g - 1} \\ &\quad\geq \frac{1}{4}(1 - 2L_{\boldsymbol{u}} \circ R(x))^{-\ell_{\boldsymbol{d}} r_g} \cdot 2\ell_{\boldsymbol{t}} r_g (1 - 2L_{\boldsymbol{u}} \circ R(x))^{\ell_{\boldsymbol{d}} r_g - \ell_{\boldsymbol{q}}} \\ &\quad= \frac{1}{2}\ell_{\boldsymbol{t}} r_g (1 - 2L_{\boldsymbol{u}} \circ R(x))^{-\ell_{\boldsymbol{q}}}, \end{aligned} \tag{3.36}$$

which we can lower bound by $\frac{1}{2}$. To prove (i), for any $j \in \{1, \dots, m-1\}$ taking $\boldsymbol{q} = \epsilon$ (which means $\ell_{\boldsymbol{q}} = 1$) and $\boldsymbol{t} = \boldsymbol{d}$ we obtain using (3.32) and (3.36) that

$$DT_{\tau^j \omega}^{\varphi(\omega,x)-j}(T_\omega^j x) = D(R^l \circ L_g \circ L_{\boldsymbol{d}})(L_{\boldsymbol{u}} \circ R(x)) \cdot DL_{u_j \cdots u_{m-1}}(L_{u_1 \cdots u_{j-1}} \circ R(x)) \geq \frac{1}{2}.$$

For $j \in \{m, \dots, \kappa - 1\}$ we take $\boldsymbol{q} = \omega_{m+1} \cdots \omega_j$ and $\boldsymbol{t} = \omega_{j+1} \cdots \omega_{\kappa-1}$ (so $\boldsymbol{q} = \epsilon$ in case $j = m$ and $\boldsymbol{t} = \epsilon$ in case $j = \kappa - 1$) and get

$$DT_{\tau^j \omega}^{\varphi(\omega,x)-j}(T_\omega^j(x)) = D(R^l \circ L_g \circ L_{\boldsymbol{t}})(L_{\boldsymbol{q}} \circ L_{\boldsymbol{u}} \circ R(x)) \geq \frac{1}{2}$$

by (3.36). Finally, if $j \in \{\kappa, \dots, \varphi(\omega, x) - 1\}$, then

$$DT_{\tau^j \omega}^{\varphi(\omega,x)-j}(T_\omega^j(x)) = 2^{\kappa + l - j} \geq \frac{1}{2}.$$

This proves (i). For (ii), we write

$$DT_\omega^{\varphi(\omega,x)}(x) = D(R^l \circ L_g \circ L_{\boldsymbol{d}})(L_{\boldsymbol{u}} \circ R(x)) \cdot DL_{\boldsymbol{u}}(R(x)) \cdot DR(x).$$

We have $DR(x) = 2$ and by (3.32) with $j = 1$ we get $DL_{\boldsymbol{u}}(R(x)) \geq 1$. What is left is to estimate the first factor. From (3.36) with $\boldsymbol{q} = \epsilon$ and $\boldsymbol{t} = \boldsymbol{d}$ we see that

$$D(R^l \circ L_g \circ L_{\boldsymbol{d}})(L_{\boldsymbol{u}} \circ R(x)) \geq \frac{1}{2}\ell_{\boldsymbol{d}} r_g (1 - 2L_{\boldsymbol{u}} \circ R(x))^{-1}.$$

Note that by the definition of m we have $L_{\boldsymbol{u}} \circ R(x) \in (L_g^{-1}(\frac{1}{2}), \frac{1}{2})$ if $m = \kappa - 1$, so if $\boldsymbol{d} = \epsilon$, and $L_{\boldsymbol{u}} \circ R(x) \in [x_{d_1}, \frac{1}{2})$ if $m < \kappa - 1$. In case $m = \kappa - 1$ we obtain that

$$\ell_{\boldsymbol{d}} r_g (1 - 2L_{\boldsymbol{u}} \circ R(x))^{-1} \geq r_g \Big(1 - 2L_g^{-1}\Big(\frac{1}{2}\Big)\Big)^{-1} = r_g \cdot 2^{1/r_g} \geq 2,$$

where we used the expression for $L_g^{-1}(\frac{1}{2})$ from (3.8) and the fact that $x \cdot 2^{1/x} \geq 2$ for all $x \geq 1$. In case $m < \kappa - 1$, we have

$$\ell_{\boldsymbol{d}} r_g (1 - 2L_{\boldsymbol{u}} \circ R(x))^{-1} \geq \ell_{d_1}(1 - 2x_{d_1})^{-1} = \ell_{d_1}^{1+(\ell_{d_1}-1)^{-1}} \geq 2,$$

where we used (3.9) and the fact that $x^{1+(x-1)^{-1}} > 2$ for all $x > 1$. Hence, in all cases

$$DT_\omega^{\varphi(\omega,x)}(x) \geq \frac{1}{2} \cdot 2 \cdot 1 \cdot 2 = 2. \qquad \square$$

Recall the first return time partition $\mathcal{P}$ from (3.20). For $P = P_{\boldsymbol{usgw}} \in \mathcal{P}$ set

$$\pi_2(P) := (R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{s}} \circ R)^{-1}\Big(\frac{1}{2}, \frac{3}{4}\Big)$$

and write S_P for the restriction of the map $T_\omega^{\varphi(\omega,x)}$ to $\pi_2(P)$, so

$$S_P := T_\omega^{\varphi(\omega,x)}|_{\pi_2(P)} = R^{|\boldsymbol{w}|} \circ L_g \circ L_{\boldsymbol{s}} \circ R|_{\pi_2(P)}.$$

We give two lemmas on the maps S_P, that will be useful when verifying (t6) for the Young tower in the next section.

Lemma 3.3.8. *There exists a constant $C_3 > 0$ such that for each $P \in \mathcal{P}$ and all $(\omega, x), (\omega', y) \in P$ we have*

$$\Big| \frac{J_{m_{\boldsymbol{p}} \times \lambda} F^\varphi(\omega, x)}{J_{m_{\boldsymbol{p}} \times \lambda} F^\varphi(\omega', y)} - 1 \Big| \leq C_3 \cdot \big| S_P(x) - S_P(y) \big|.$$

Proof. For each $P \in \mathcal{P}$ and all $(\omega, x), (\omega', y) \in P$ we have $\varphi(\omega, x) = \varphi(\omega', y) = \varphi_P$ and $\omega_j = \omega_j'$ for all $1 \leq j \leq \varphi_P$. Hence, for each measurable set $A \subseteq P$ we have

$$m_{\boldsymbol{p}} \times \lambda(F^\varphi(A)) = \int_A \Big(\prod_{j=1}^{\varphi(\omega,x)} p_{\omega_j}^{-1} \Big) DT_\omega^{\varphi(\omega,x)}(x) \, dm_{\boldsymbol{p}} \times \lambda(\omega, x).$$

By Proposition 1.2.19 we obtain

$$J_{m_{\boldsymbol{p}} \times \lambda} F^\varphi(\omega, x) = \Big(\prod_{j=1}^{\varphi(\omega,x)} p_{\omega_j}^{-1} \Big) DT_\omega^{\varphi(\omega,x)}(x),$$

which, for each $P \in \mathcal{P}$ and all $(\omega, x), (\omega', y) \in P$, gives

$$\Big| \frac{J_{m_{\boldsymbol{p}} \times \lambda} F^\varphi(\omega, x)}{J_{m_{\boldsymbol{p}} \times \lambda} F^\varphi(\omega', y)} - 1 \Big| = \Big| \frac{DT_\omega^{\varphi(\omega,x)}(x)}{DT_{\omega'}^{\varphi(\omega',y)}(y)} - 1 \Big| = \Big| \frac{DS_P(x)}{DS_P(y)} - 1 \Big|.$$

Let $c > 0$. As compositions of good and bad maps each S_P has non-positive Schwarzian derivative and by Lemma 3.3.7(ii) each S_P satisfies $DS_P \geq 2$. For this reason for each $P \in \mathcal{P}$ we can extend the domain $\pi_2(P)$ of S_P on both sides to an interval $I_P \supseteq \pi_2(P)$ such that there exists an extension $\tilde{S}_P : I_P \to \mathbb{R}$ of S_P, i.e. $\tilde{S}_P|_{\pi_2(P)} =$

$S_P|_{\pi_2(P)}$, that has non-positive Schwarzian derivative and for which both components of $\tilde{S}_P(I_P)\backslash(\frac{1}{2},\frac{3}{4})$ have length at least $\frac{c}{4}$. Applying for each $P \in \mathcal{P}$ the Koebe Principle (1.15) with $I = I_P$ and $J = \pi_2(P)$ then gives a constant $C_3 > 0$ that depends only on c such that for each $P \in \mathcal{P}$ and each $x, y \in \pi_2(P)$ we have

$$\left|\frac{DS_P(x)}{DS_P(y)} - 1\right| \leq C_3 \cdot |S_P(x) - S_P(y)|.$$

This gives the lemma. □

Recall the definition of the separation time from (3.10). We have the following lemma.

Lemma 3.3.9. *Let* $(\omega, x), (\omega', y) \in Y$. *Then*

$$|x - y| \leq 2^{-s((\omega,x),(\omega',y))}. \tag{3.37}$$

Furthermore, if $(\omega, x), (\omega', y) \in P$ *for some* $P \in \mathcal{P}$*, then*

$$\big|S_P(x) - S_P(y)\big| \leq 2^{-s(F^\varphi(\omega,x),F^\varphi(\omega',y))}. \tag{3.38}$$

Proof. Write $n = s((\omega, x), (\omega', y))$ and, for each $k \in \{0, 1, \ldots, n-1\}$, let $P^{(k)} \in \mathcal{P}$ be such that $(F^\varphi)^k(\omega, x), (F^\varphi)^k(\omega', y) \in P^{(k)}$. Then for each $k \in \{0, 1, \ldots, n-1\}$ the points $(S_{P^{(k-1)}} \circ \cdots \circ S_{P^{(0)}})(x)$ and $(S_{P^{(k-1)}} \circ \cdots \circ S_{P^{(0)}})(y)$ lie in the domain $\pi_2(P^{(k)})$ of $S_{P^{(k)}}$, so it follows from Lemma 3.3.7(ii) together with the Mean Value Theorem that

$$\frac{|S_{P^{(k)}} \circ \cdots \circ S_{P^{(0)}}(x) - S_{P^{(k)}} \circ \cdots \circ S_{P^{(0)}}(y)|}{|S_{P^{(k-1)}} \circ \cdots \circ S_{P^{(0)}}(x) - S_{P^{(k-1)}} \circ \cdots \circ S_{P^{(0)}}(y)|} \geq \inf DS_{P^{(k)}} \geq 2.$$

We conclude that

$$\begin{aligned}|x - y| &\leq 2^{-1}|S_{P^{(0)}}(x) - S_{P^{(0)}}(y)| \\ &\leq \cdots \leq 2^{-n}|S_{P^{(n-1)}} \circ \cdots \circ S_{P^{(0)}}(x) - S_{P^{(n-1)}} \circ \cdots \circ S_{P^{(0)}}(y)| \leq 2^{-n},\end{aligned}$$

which gives the first part of the lemma. For the second part, note that if $(\omega, x), (\omega', y) \in P$ for some $P \in \mathcal{P}$, then (3.38) follows by applying (3.37) to the points $F^\varphi(\omega, x) = (\tau^{\varphi_P}\omega, S_P(x))$ and $F^\varphi(\omega', y) = (\tau^{\varphi_P}\omega', S_P(y))$. □

§3.4 A Young tower for the random map

§3.4.1 The acs probability measure

We are now in the position to construct a Young tower for the skew product F according to the setup from [Y99, Section 1.1] that we outlined in Subsection 3.2.2.

As the base for the Young tower we take the set Y. The Young tower Δ, the l^{th} levels of the tower Δ_l and the tower map $G : \Delta \to \Delta$ are defined in Subsection 3.2.2, as well as the reference measure m and the partition η on Δ. Following the general setup in [Y99], inducing the map G on $\Delta_0 = Y \times \{0\}$ yields a transformation G^φ on Δ_0 given by $G^\varphi(z,0) = G^{\varphi(z)}(z,0)$. Recall that we identify G^φ with F^φ by identifying Δ_0 with Y and using the correspondence $G^\varphi(z,0) = (F^\varphi(z),0)$. We check that the conditions (t1)–(t6) from Section 3.2.2 hold for this construction.

Proposition 3.4.1. *The conditions (t1)–(t6) hold for the map G on the Young tower Δ defined above.*

Proof. From the first return time partition $\mathcal{P}$ it is clear that (t1), (t2), (t3) and (t5) hold.

For (t4) it is enough to show that the collection

$$\bigvee_{n\geq 0} G^{-n}\eta = \{E_0 \cap G^{-1}E_1 \cap \cdots \cap G^{-n}E_n \,:\, E_i \in \eta,\, 1 \leq i \leq n,\, n \geq 0\}$$

separates points. To show this, let $(z_1,l_1),(z_2,l_2) \in \Delta$ be two points. If $z_1 = z_2$ and $l_1 \neq l_2$ and $P \in \mathcal{P}$ is such that $z_1 \in P$, then $\Delta_{l_1,P}, \Delta_{l_2,P} \in \eta$ are sets that separate (z_1,l_1) and (z_2,l_2). Assume that $z_1 \neq z_2$. Lemma 3.3.7(ii) implies that the map F^φ is expanding on Y, so there exist an $N \geq 0$ and two disjoint sets $A, E \in \bigvee_{n=0}^{N}(F^\varphi)^{-n}\mathcal{P}$ such that $z_1 \in A$ and $z_2 \in E$. On these sets the first N first return times to Y are constant, meaning that if $K > 0$ is such that $G^K(z_1,0) = ((F^\varphi)^N(z_1),0)$, then $A \times \{l_1\} \in \bigvee_{n=0}^{K+l_1} G^{-n}\eta$ and if $L > 0$ such that $G^L(z_2,0) = ((F^\varphi)^N(z_2),0)$, then $E \times \{l_2\} \in \bigvee_{n=0}^{L+l_2} G^{-n}\eta$. Note that $(z_1,l_1) \in A \times \{l_1\}$ and $(z_2,l_2) \in E \times \{l_2\}$ and $(A \times \{l_1\}) \cap (E \times \{l_2\}) = \emptyset$. Hence, (t4) holds.

Finally, from Lemma 3.3.8 and (3.38) we obtain that

$$\left|\frac{J_{m_p\times\lambda}F^\varphi(z_1)}{J_{m_p\times\lambda}F^\varphi(z_2)} - 1\right| \leq C_3 \cdot 2^{-s(F^\varphi(z_1),F^\varphi(z_2))} \tag{3.39}$$

for each $P \in \mathcal{P}$ and all $z_1, z_2 \in P$. This gives (t6) with $\beta = \frac{1}{2}$ and the proposition follows. ☐

Now Proposition 3.3.3(i) and Theorem 3.2.3 imply the existence of a probability measure ν on $(\Delta, \mathcal{B})$ that is G-invariant, exact and absolutely continuous with respect to m with a density that is bounded (because it is in $\mathcal{C}_\beta$) and bounded away from zero and that satisfies (3.12). We use this to construct the invariant measure for F that is promised in Theorem 3.1.2. Define

$$\pi : \Delta \to \Sigma^{\mathbb{N}} \times [0,1],\ (z,l) \mapsto F^l(z).$$

Then

$$\begin{aligned} \pi(G(z,l)) &= \pi(z,l+1) = F^{l+1}(z) = F(\pi(z,l)), && l < \varphi(z) - 1, \\ \pi(G(z,l)) &= \pi(F^\varphi(z),0) = F^\varphi(z) = F(\pi(z,l)), && l = \varphi(z) - 1. \end{aligned}$$

So $\pi \circ G = F \circ \pi$. Let $\rho = \nu \circ \pi^{-1}$ be the pushforward measure of ν under π.

Lemma 3.4.2. *The probability measure ρ satisfies the following properties.*

(i) *F is measure preserving and mixing with respect to* ρ.

(ii) ρ *is absolutely continuous with respect to* $m_{\boldsymbol{p}} \times \lambda$.

(iii) *We have*

$$\rho(A \cap Y) = \nu((A \cap Y) \times \{0\}), \qquad A \in \mathcal{F}.$$

Proof. Part (i) immediately follows from the properties of the measure ν and the fact that $\pi \circ G = F \circ \pi$. For (ii), let $A \in \mathcal{F}$ be such that $m_{\boldsymbol{p}} \times \lambda(A) = 0$. Using that F is non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$, we obtain that

$$m(\pi^{-1}(A)) = m\Big(\Delta \cap \Big(\bigcup_{l \geq 0} F^{-l}(A) \times \{l\}\Big)\Big) \leq \sum_{l \geq 0} m_{\boldsymbol{p}} \times \lambda(F^{-l}(A)) = 0.$$

Since ν is absolutely continuous with respect to m, it follows that $\rho(A) = \nu(\pi^{-1}A) = 0$. For (iii) let $A \in \mathcal{F}$. We have

$$\pi^{-1}(A \cap Y) = \bigcup_{P \in \mathcal{P}} \bigcup_{l=0}^{\varphi_P - 1} (F^{-l}(A \cap Y) \cap P) \times \{l\}.$$

By definition of φ_P, we have $F^l(z) \notin Y$ for each $z \in P$ and each $l \in \{1, \ldots, \varphi_P - 1\}$. Therefore

$$\pi^{-1}(A \cap Y) = \bigcup_{P \in \mathcal{P}} (A \cap P) \times \{0\} = (A \cap Y) \times \{0\}.$$

□

Combining Lemma 3.4.2 with Lemma 1.4.1 yields that there exists a probability measure μ that is absolutely continuous with respect to λ and such that $\rho = m_{\boldsymbol{p}} \times \mu$. In other words, μ is an acs measure for $(\mathcal{T}, \boldsymbol{p})$. We will now prove Theorem 3.1.2, which shows that μ is in fact the only acs measure for $(\mathcal{T}, \boldsymbol{p})$.

Proof of Theorem 3.1.2. It follows from Lemma 3.4.2(i) that F is mixing with respect to $m_{\boldsymbol{p}} \times \mu$. Hence, to obtain that μ is the only acs probability measure for $(\mathcal{T}, \boldsymbol{p})$, it suffices according to Theorem 1.2.6 to show that $\frac{d\mu}{d\lambda} > 0$ holds λ-a.e. Theorem 3.2.3 asserts that there is a constant $C_4 \geq 1$ such that

$$\frac{1}{C_4} \leq \frac{d\nu}{dm} \leq C_4. \tag{3.40}$$

Let $B \subseteq (\frac{1}{2}, \frac{3}{4})$ be a Borel set. Lemma 3.4.2(iii) and (3.40) imply that

$$\mu(B) = \nu(((\tilde{\Omega} \times B) \cap Y) \times \{0\}) \geq C_4^{-1} \cdot m(((\tilde{\Omega} \times B) \cap Y) \times \{0\}) = C_4^{-1} \cdot \lambda(B).$$

Since B was arbitrary, we have $\frac{d\mu}{d\lambda}(x) \geq C_4^{-1}$ for λ-a.e. $x \in (\frac{1}{2}, \frac{3}{4})$. Recall that the density $\frac{d\mu}{d\lambda}$ is a fixed point of the Perron-Frobenius operator $\mathcal{P}_{\mathcal{T},\boldsymbol{p}}$ being of the form as in (1.20). Fix some $g \in \Sigma_G$. The map $T_g^2|_{(\frac{1}{2},\frac{3}{4})} : (\frac{1}{2}, \frac{3}{4}) \to (0,1)$ is a measurable bijection with measurable inverse. For each $x \in (0,1)$ let y_x be the unique element in $(\frac{1}{2}, \frac{3}{4})$ that satisfies $x = T_g^2(y_x)$. Furthermore, note that $\sup_{y \in (\frac{1}{2},\frac{3}{4})} DT_g^2(y) \leq 2DT_g(0)$. We conclude that for λ-a.e. $x \in (0,1)$

$$\frac{d\mu}{d\lambda}(x) = \mathcal{P}_{\mathcal{T},\boldsymbol{p}}^2 \frac{d\mu}{d\lambda}(x) \geq p_g^2 \frac{\frac{d\mu}{d\lambda}(y_x)}{DT_g^2(y_x)} \geq \frac{p_g^2 C_4^{-1}}{2DT_g(0)} > 0.$$

This gives that $\frac{d\mu}{d\lambda}$ is bounded away from zero and therefore that μ is the unique acs measure. □

Remark 3.4.3. Besides Theorems 3.1.1 and 3.1.2 it can also be shown that all the results from Theorem 2.1.2 carry over. Namely, by following the same steps as in Section 2.2 it can be shown that $(\mathcal{T}, \boldsymbol{p})$ admits, independent of the value of θ, a unique (up to scalar multiplication) acs measure that is σ-finite and ergodic and for which the density is bounded away from zero, is locally Lipschitz on $(0, \frac{1}{2})$ and $[\frac{1}{2}, 1)$ and is not in L^q for any $q > 1$. This measure is infinite if $\theta \geq 1$ and coincides with μ if $\theta < 1$. Similar as in Chapter 2, if $\theta \geq 1$, it follows from applying Aaronson's Ergodic Theorem [A97, Theorem 2.4.2] to this infinite measure that no physical measure (see footnote 1 on page 39) for F exists.

§3.4.2 Decay of correlations

Recall that for $\alpha \in (0,1)$ we have set $\mathcal{H}_\alpha$ for the set of α-Hölder continuous functions on $\Sigma^{\mathbb{N}} \times [0,1]$ with metric d as in (3.3). Also recall the definition of the function spaces $\mathcal{C}_\delta$ on Δ from (3.11).

Lemma 3.4.4. *Let $\alpha \in (0,1)$ and $h \in \mathcal{H}_\alpha$. Then $h \circ \pi \in \mathcal{C}_{1/2^\alpha}$.*

Proof. Since $h \in \mathcal{H}_\alpha$, there exists a constant $C_5 > 0$ such that

$$|h(z_1) - h(z_2)| \leq C_5 \cdot d(z_1, z_2)^\alpha \quad \text{for all } z_1, z_2 \in \Sigma^{\mathbb{N}} \times [0,1]. \tag{3.41}$$

From this it is easy to see that $\|h\|_\infty < \infty$. Let $\upsilon_1 = (z_1, l_1), \upsilon_2 = (z_2, l_2) \in \Delta$. If $l_1 \neq l_2$ or if z_1 and z_2 lie in different elements of $\mathcal{P}$, then $s(\upsilon_1, \upsilon_2) = 0$ and

$$|h \circ \pi(\upsilon_1) - h \circ \pi(\upsilon_2)| = |h(F^{l_1}(z_1)) - h(F^{l_2}(z_2))| \leq 2\|h\|_\infty = 2\|h\|_\infty \cdot 2^{-\alpha s(\upsilon_1, \upsilon_2)}. \tag{3.42}$$

Hence, to prove that $h \circ \pi \in \mathcal{C}_{1/2^\alpha}$, it remains to consider the case that $z_1, z_2 \in P$ for some $P \in \mathcal{P}$ and $l_1 = l_2 = l \in \{0, \ldots, \varphi_P - 1\}$. Write $z_1 = (\omega, x)$ and $z_2 = (\omega', y)$. Note that $\omega_j = \omega'_j$ for each $j \in \{1, 2, \ldots, \varphi_P\}$. Hence,

$$2^{-\min\{i \in \mathbb{N} : \omega_{l+i} \neq \omega'_{l+i}\}} \leq 2^{-\min\{i \in \mathbb{N} : \omega_{\varphi_P - 1 + i} \neq \omega'_{\varphi_P - 1 + i}\}} \leq 2^{-s(z_1, z_2)}.$$

Furthermore, it follows from the Mean Value Theorem together with Lemma 3.3.7(i) that

$$\frac{|T_\omega^{\varphi_P}(x) - T_\omega^{\varphi_P}(y)|}{|T_\omega^l(x) - T_\omega^l(y)|} = \frac{|T_{\tau^l \omega}^{\varphi_P - l}(T_\omega^l(x)) - T_{\tau^l \omega}^{\varphi_P - l}(T_\omega^l(y))|}{|T_\omega^l(x) - T_\omega^l(y)|} \geq \frac{1}{2}.$$

Combining this with (3.38) yields that

$$|T_\omega^l(x) - T_\omega^l(y)| \leq 2 \cdot 2^{-s(F^\varphi(z_1), F^\varphi(z_2))} = 4 \cdot 2^{-s(z_1, z_2)}$$

and hence by (3.41),

$$|h(F^l(z_1)) - h(F^l(z_2))| \leq C_5 (2^{-s(z_1, z_2)} + 4 \cdot 2^{-s(z_1, z_2)})^\alpha = 5^\alpha C_5 \cdot 2^{-\alpha s(z_1, z_2)}.$$

Together with (3.42) this gives the result. □

We now have all the ingredients to prove Theorem 3.1.3.

Proof of Theorem 3.1.3. To prove the theorem, we would like to use Theorem 3.2.3(iv), which requires us to bound $m(\hat{\varphi} > n)$, where

$$\hat{\varphi} : \Delta \to \mathbb{N}_0,\ v \mapsto \inf\{n \geq 0 \,:\, G^n(v) \in \Delta_0\}.$$

Since

$$\{\hat{\varphi} = 0\} = \Delta_0 = \bigcup_{P \in \mathcal{P} : \varphi_P > 0} \Delta_{0,P},$$

$$\{\hat{\varphi} = n\} = \bigcup_{P \in \mathcal{P} : \varphi_P > n} \Delta_{\varphi_P - n, P}, \qquad n \geq 1,$$

we have for each $n \geq 0$ that

$$m(\hat{\varphi} = n) = \sum_{P \in \mathcal{P} : \varphi_P > n} m_{\boldsymbol{p}} \times \lambda(P) = m_{\boldsymbol{p}} \times \lambda(\varphi > n).$$

It follows from Proposition 3.3.3(ii) that for each $\gamma \in (\gamma_1, 0)$ there is an $M > 0$ and an $N \geq 1$ such that for each $n \geq N$,

$$m_{\boldsymbol{p}} \times \lambda(\varphi > n) \leq M \cdot n^{\gamma - 1}.$$

Thus, for all $n \geq N$,

$$m(\hat{\varphi} > n) = \sum_{k > n} m(\hat{\varphi} = k) \leq M \sum_{k \geq n} k^{\gamma - 1} \leq M \cdot n^{\gamma - 1} + M \int_n^\infty x^{\gamma - 1} dx. \tag{3.43}$$

So, $m(\hat{\varphi} > n) = O(n^\gamma)$. Combining Proposition 3.3.3(i), Proposition 3.4.1 and Theorem 3.2.3(iv) now gives that for each $\gamma \in (\gamma_1, 0)$, $\hat{f} \in L^\infty(\Delta, \nu)$, $\delta \in (0,1)$ and $\hat{h} \in \mathcal{C}_\delta$,

$$\Big| \int_\Delta \hat{f} \circ G^n \cdot \hat{h} \, d\nu - \int_\Delta \hat{f} \, d\nu \int_\Delta \hat{h} \, d\nu \Big| = O(n^\gamma). \tag{3.44}$$

Now, let $\gamma \in (\gamma_1, 0)$, $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$. Using that $m_{\boldsymbol{p}} \times \mu = \nu \circ \pi^{-1}$ and $\pi \circ G = F \circ \pi$, it then follows that

$$|\mathrm{Cor}_n(f,h)| = \Big| \int_\Delta (f \circ \pi) \circ G^n \cdot (h \circ \pi) \, d\nu - \int_\Delta f \circ \pi \, d\nu \int_\Delta h \circ \pi \, d\nu \Big|.$$

Since $h \in \mathcal{H}$, it holds that $h \in \mathcal{H}_\alpha$ for some $\alpha \in (0,1)$, so $h \circ \pi \in \mathcal{C}_{1/2^\alpha}$ by Lemma 3.4.4. Since also $f \circ \pi \in L^\infty(\Delta, \nu)$, we obtain the result from (3.44) with $\hat{f} = f \circ \pi$ and $\hat{h} = h \circ \pi$. □

In order to prove Theorem 3.1.4, we need the following lemma.

Lemma 3.4.5. *There exists $C_6 > 0$ such that for each $P \in \mathcal{P}$ and $z_1, z_2 \in P$,*

$$\Big| \log \frac{J_{m_{\boldsymbol{p}} \times \mu} F^\varphi(z_1)}{J_{m_{\boldsymbol{p}} \times \mu} F^\varphi(z_2)} \Big| \leq C_6 \cdot 2^{-s(z_1, z_2)}.$$

Proof. From the definition of the Jacobian we see that $J_m G^\varphi|_{\Delta_0} = J_{m_{\boldsymbol{p}}\times\lambda} F^\varphi$ with the identification of Δ_0 and Y. Lemma 3.4.2(iii) and Lemma 1.2.20(a) give us that for each $P \in \mathcal{P}$ and each measurable set $A \subseteq P$,

$$\begin{aligned} m_{\boldsymbol{p}} \times \mu(F^\varphi(A)) &= \nu(G^\varphi(A\times\{0\})) \\ &= \int_{A\times\{0\}} \Big(\frac{d\nu}{dm}\circ G^\varphi\Big) J_m G^\varphi \, dm \\ &= \int_A \frac{d\nu}{dm}(F^\varphi(z),0)\cdot J_{m_{\boldsymbol{p}}\times\lambda}F^\varphi(z)\cdot \frac{dm}{d\nu}(z,0)\, dm_{\boldsymbol{p}}\times\mu(z). \end{aligned}$$

This gives

$$J_{m_{\boldsymbol{p}}\times\mu}F^\varphi(z) = \frac{d\nu}{dm}(F^\varphi(z),0)\cdot J_{m_{\boldsymbol{p}}\times\lambda}F^\varphi(z)\cdot\frac{dm}{d\nu}(z,0), \qquad z\in Y,$$

and thus, for each $z_1, z_2 \in Y$,

$$\Big|\log\frac{J_{m_{\boldsymbol{p}}\times\mu}F^\varphi(z_1)}{J_{m_{\boldsymbol{p}}\times\mu}F^\varphi(z_2)}\Big| \le \Big|\log\frac{\frac{d\nu}{dm}(F^\varphi(z_1),0)}{\frac{d\nu}{dm}(F^\varphi(z_2),0)}\Big| + \Big|\log\frac{J_{m_{\boldsymbol{p}}\times\lambda}F^\varphi(z_1)}{J_{m_{\boldsymbol{p}}\times\lambda}F^\varphi(z_2)}\Big| + \Big|\log\frac{\frac{d\nu}{dm}(z_2,0)}{\frac{d\nu}{dm}(z_1,0)}\Big|.$$

Combining Proposition 3.3.3(i), Proposition 3.4.1 and Theorem 3.2.3(ii) gives the existence of a constant $C^+ > 0$ such that, for each $\Delta_{l,P} \in \eta$ and $\upsilon_1, \upsilon_2 \in \Delta_{l,P}$,

$$\Big|\frac{\frac{d\nu}{dm}(\upsilon_1)}{\frac{d\nu}{dm}(\upsilon_2)} - 1\Big| \le C^+ \cdot 2^{-s(\upsilon_1,\upsilon_2)}. \tag{3.45}$$

Using that $|\log x| \le \max\{|x-1|, |x^{-1}-1|\}$ for all $x > 0$, we obtain from (3.39) and (3.45) that

$$\Big|\log\frac{J_{m_{\boldsymbol{p}}\times\mu}F^\varphi(z_1)}{J_{m_{\boldsymbol{p}}\times\mu}F^\varphi(z_2)}\Big| \le C^+\cdot 2^{-s(F^\varphi(z_1),F^\varphi(z_2))} + C_3\cdot 2^{-s(F^\varphi(z_1),F^\varphi(z_2))} + C^+\cdot 2^{-s(z_1,z_2)}$$

for all $z_1, z_2 \in P$, $P\in\mathcal{P}$. The lemma thus holds with $C_6 = 3C^+ + 2C_3$. □

Proof of Theorem 3.1.4. Let $f \in L^\infty(\Sigma^{\mathbb{N}}\times[0,1], m_{\boldsymbol{p}}\times\mu)$ and $h \in \mathcal{H}$ be such that both f and h are identically zero on $\Sigma^{\mathbb{N}} \times \big([0,\frac12]\cup[\frac34,1]\big)$ and such that $\int f\, dm_{\boldsymbol{p}} \times \mu \cdot \int h\, dm_{\boldsymbol{p}}\times\mu > 0$. Let $\gamma \in (\gamma_1, \min\{\gamma_2+1,-1\})$ if $\gamma_1 < -1$ and $\gamma\in(\gamma_1,\frac{\gamma_2}{2})$ if $-1\le\gamma_1<0$. This is possible by assumption. Our strategy is to apply Theorem 3.2.4 with Y as before. For this, we verify (g1), (g2) and (g3).[4]

For (g3), $h\in\mathcal{H}$ implies that $h\in\mathcal{H}_\alpha$ for some $\alpha\in(0,1)$ and thus $h\circ\pi\in\mathcal{C}_{1/2^\alpha}$ by Lemma 3.4.4. In particular this yields (g3) with $\delta = 2^{-\alpha}$. For (g2), Lemma 3.4.2(iii) and (3.40) give

$$m_{\boldsymbol{p}}\times\mu(\varphi>n) = \int_{\{\varphi>n\}\times\{0\}}\frac{d\nu}{dm}\, dm \le C_4\cdot m_{\boldsymbol{p}}\times\lambda(\varphi>n).$$

[4]More precisely, we apply Theorem 3.2.4 to versions of f and h that are also zero on $\big(\Sigma^{\mathbb{N}}\times(\frac12,\frac34)\big)\backslash Y$.

Together with Proposition 3.3.3(ii) this implies that

$$m_{\boldsymbol{p}} \times \mu(\varphi > n) = O(n^{\gamma-1}).$$

Finally, (g1) with $\delta = 2^{-\alpha}$ follows from Lemma 3.4.5 by setting $C^* = C_6$ and noting that $\frac{1}{2^\alpha} > \frac{1}{2}$. Hence, we satisfy all the conditions of Theorem 3.2.4 with $\delta = 2^{-\alpha}$ and $\zeta = 1-\gamma$. Note that

$$K_{1-\gamma}(n) = \begin{cases} n^{\gamma-1}, & \text{if } 1-\gamma > 2, \\ \frac{\log n}{n^2}, & \text{if } 1-\gamma = 2, \\ n^{2\gamma}, & \text{if } 1-\gamma \in (1,2). \end{cases}$$

If $\gamma_1 < -1$, then $1-\gamma \in (\max\{-\gamma_2, 2\}, 1-\gamma_1) \subseteq (2,\infty)$ and if $-1 \le \gamma_1 < 0$, then $1-\gamma \in (1-\frac{\gamma_2}{2}, 1-\gamma_1) \subseteq (1,2)$. We can thus conclude from Theorem 3.2.4 that

$$\left|\mathrm{Cor}_n(f,h) - \Big(\sum_{k>n}^{\infty} m_{\boldsymbol{p}} \times \mu(\varphi > k)\Big) \int f \, dm_{\boldsymbol{p}} \times \mu \int h \, dm_{\boldsymbol{p}} \times \mu\right| = O(n^{\xi}),$$

where $\xi = \gamma - 1$ if $\gamma_1 < -1$ and $\xi = 2\gamma$ if $-1 \le \gamma_1 < 0$. As above it follows from Proposition 3.3.3(iii) combined with Lemma 3.4.2(iii) and (3.40) that $m_{\boldsymbol{p}} \times \mu(\varphi > n) = \Omega(n^{\gamma_2 - 1})$ and thus

$$\sum_{k>n}^{\infty} m_{\boldsymbol{p}} \times \mu(\varphi > k) = \Omega(n^{\gamma_2}).$$

The result now follows from observing that $\gamma_2 > \xi$. □

We provide some examples of combinations of parameters for which the conditions of Theorem 3.1.4 hold. As before set $\ell_{\min} = \min\{\ell_b : b \in \Sigma_B\}$ and $p_B = \sum_{j\in\Sigma_B} p_j$ and set $\pi_B = \sum_{j\in\Sigma_B : \ell_j = \ell_{\max}} p_j$. Examples that satisfy the conditions of Theorem 3.1.4 include the following.

- If Σ_B consists of one element, then $\gamma_1 = \gamma_2$.

- If $p_B^{-1/3} < \ell_{\min} \le \ell_{\max} < p_B^{-1/2}$, or equivalently $\ell_{\min} > \ell_{\max}^{2/3}$ and $p_B \in (\ell_{\min}^{-3}, \ell_{\max}^{-2})$, then $\theta \le p_B \cdot \ell_{\max} < \ell_{\max}^{-1}$, so $\theta < 1$ and $\gamma_1 < -1$, and

$$\gamma_2 \ge 1 + \frac{\log p_B}{\log \ell_{\min}} > -2 \ge \gamma_1 - 1.$$

- If $\pi_B > p_B^{4/3}$ (or equivalently $\pi_B^{-1/2} < p_B^{-2}\pi_B$) and $\ell_{\max} \in [\pi_B^{-1/2}, p_B^{-2}\pi_B)$, then $\theta \le p_B \cdot \ell_{\max} \le p_B^{-1}\pi_B < 1$ and $\theta \ge \pi_B \cdot \ell_{\max} \ge \ell_{\max}^{-1}$, i.e. $\gamma_1 \ge -1$, and

$$\gamma_2 \ge 1 + \frac{\log \pi_B}{\log \ell_{\max}} > \frac{2\log(p_B \cdot \ell_{\max})}{\log \ell_{\max}} \ge 2\gamma_1.$$

§3.5 Further results and final remarks

We can obtain more information from the results from [G04]. First of all, using the last part of [G04, Theorem 6.3] the upper bound in Theorem 3.1.3 can be improved for a specific class of test functions.

Theorem 3.5.1. *Assume that $\theta < 1$. Let $f \in L^\infty(\Sigma^{\mathbb{N}} \times [0,1], m_{\boldsymbol{p}} \times \mu)$ and $h \in \mathcal{H}$ be such that both f and h are identically zero on $\Sigma^{\mathbb{N}} \times \big([0, \frac{1}{2}] \cup [\frac{3}{4}, 1]\big)$ and $\int h \, dm_{\boldsymbol{p}} \times \mu = 0$. Let $\gamma \in (\gamma_1, 0)$. Then*

$$|\mathrm{Cor}_n(f,h)| = O(n^{\gamma-1}).$$

Proof. The statement follows by applying the last part of [G04, Theorem 6.3]. For this, (g1), (g2) and (g3) need to be verified. This is done before in the proof of Theorem 3.1.4. □

In [G04, Theorem 6.13] a Central Limit Theorem is derived for a specific class of functions in $\mathcal{H}$ with zero integral. This result immediately carries over to our setting and is given in the next theorem.

Theorem 3.5.2 (cf. Theorem 6.13 in [G04]). *Assume that $\theta < 1$. Let $h \in \mathcal{H}$ be identically zero on $\Sigma^{\mathbb{N}} \times \big([0, \frac{1}{2}] \cup [\frac{3}{4}, 1]\big)$ and with $\int h \, dm_{\boldsymbol{p}} \times \mu = 0$. Then the sequence $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} h \circ F^k$ converges in distribution with respect to $m_{\boldsymbol{p}} \times \mu$ to a normally distributed random variable with zero mean and finite variance σ^2 given by*

$$\sigma^2 = -\int h^2 \, dm_{\boldsymbol{p}} \times \mu + 2 \sum_{n=0}^{\infty} \int h \cdot h \circ F^n \, dm_{\boldsymbol{p}} \times \mu.$$

Furthermore, we have $\sigma = 0$ if and only if there exists a measurable function ψ on $\Sigma^{\mathbb{N}} \times [0,1]$ such that $h \circ F = \psi \circ F - \psi$. Such a function ψ then satisfies $\sup_{z_1, z_2 \in Y} \frac{|\psi(z_1) - \psi(z_2)|}{(2^{-\alpha})^{s(z_1,z_2)}} < \infty$ and $\psi(F^j(z)) = \psi(z)$ for each $z \in Y$ and each $j = 0, 1, \ldots, \varphi(z) - 1$.

Using [Y99, Theorem 4] we can also derive a Central Limit Theorem, this time for a more general class of functions in $\mathcal{H}$ with zero integral but under the more restrictive assumption that $\theta < \ell_{\max}^{-1}$.

Theorem 3.5.3. *Assume that $\theta < \ell_{\max}^{-1}$. Let $h \in \mathcal{H}$ be such that $\int h \, dm_{\boldsymbol{p}} \times \mu = 0$. Then the sequence $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} h \circ F^k$ converges in distribution with respect to $m_{\boldsymbol{p}} \times \mu$ to a normally distributed random variable with zero mean and finite variance σ^2. Furthermore, we have $\sigma = 0$ if and only if there exists a measurable function ψ on Δ such that $h \circ \pi \circ G = \psi \circ G - \psi$.*

Proof. The result from [Y99, Theorem 4] gives a statement for G on Δ. We have already seen that $h \in \mathcal{H}$ implies $h \circ \pi \in \mathcal{C}_{1/2^\alpha}$ for some $\alpha \in (0,1)$. The assumption that $\theta < \ell_{\max}^{-1}$ implies $\gamma_1 < -1$. Take $\gamma \in (\gamma_1, -1)$. We saw in (3.43) that $m(\hat{\varphi} > n) = O(n^\gamma)$. It then follows from [Y99, Theorem 4] that $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} h \circ \pi \circ G^k$ converges in distribution with respect to ν to a normally distributed random variable with zero mean and finite variance σ^2, with $\sigma > 0$ if and only if $h \circ \pi \circ G \neq \psi \circ G - \psi$ for any measurable function ψ on Δ. Since $m_{\boldsymbol{p}} \times \mu = \nu \circ \pi^{-1}$ and $F \circ \pi = \pi \circ G$, we get for any $u \in \mathbb{R}$ that

$$m_{\boldsymbol{p}} \times \mu\Big(\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} h \circ F^k \leq u\Big) = \nu\Big(\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} (h \circ \pi) \circ G^k \leq u\Big).$$

The result now follows. □

Under an additional assumption on $r_{\min} = \min\{r_g : g \in \Sigma_G\}$ and $\ell_{\min} = \{\ell_b : b \in \Sigma_B\}$ we can weaken the assumption that the test functions in Theorem 3.1.4, Theorem 3.5.1 and Theorem 3.5.2 should be identically zero on $\Sigma^{\mathbb{N}} \times \big([0, \frac{1}{2}] \cup [\frac{3}{4}, 1]\big)$. Namely, if for an integer $l \geq 2$ we have

$$2^{-l} \cdot \min\left\{r_{\min} \cdot 2^{1/r_{\min}}, \ell_{\min}^{1+1/(\ell_{\min}-1)}\right\} \geq 1,$$

then it suffices to assume that these test functions are identically zero on $\Sigma^{\mathbb{N}} \times \big([0, \frac{1}{2}] \cup [1 - \frac{1}{2^{l+1}}, 1]\big)$. Indeed, in this case Lemma 3.3.7 and also Proposition 3.3.3 still carry over if we induce the random map on $(\frac{1}{2}, 1 - \frac{1}{2^{l+1}})$ instead, and the result then follows by applying [G04, Theorem 6.3 and Theorem 6.13] to this induced system in the same way as has been done in the proofs of Theorem 3.1.4, Theorem 3.5.1 and Theorem 3.5.2. The step in Lemma 3.4.5 where (3.45) is applied will then be replaced by applying

$$\left|\frac{\frac{d\mu}{d\lambda}(x)}{\frac{d\mu}{d\lambda}(y)} - 1\right| \leq C \cdot |x - y|, \qquad \forall x, y \in \left[\frac{1}{2}, 1 - \frac{1}{2^{l+1}}\right], \qquad \text{for some } C > 0,$$

which can be shown using that $\frac{d\mu}{d\lambda}$ is bounded away from zero and is locally Lipschitz on $[\frac{1}{2}, 1)$. As remarked in Remark 3.4.3, the latter can be shown by following the same steps as in Section 2.2. It in particular shows that $\frac{d\mu}{d\lambda}$ is bounded on $[\frac{1}{2}, 1 - \frac{1}{2^{l+1}}]$, which replaces the step in the proof of Theorem 3.1.4 where Lemma 3.4.2(iii) and (3.40) are applied.

We can extend the results in this chapter to the following more general classes of good and bad maps. Fix a $c \in (0, 1)$, and let the class of good maps $\mathfrak{G}$ consist of maps $T_g : [0, 1] \to [0, 1]$ given by

$$T_g(x) = \begin{cases} 1 - c^{-r_g}(c - x)^{r_g}, & \text{if} \quad x \in [0, c), \\ \frac{x-c}{1-c}, & \text{if} \quad x \in [c, 1], \end{cases}$$

where $r_g \geq 1$, and the class of bad maps $\mathfrak{B}$ consist of maps $T_b : [0, 1] \to [0, 1]$ given by

$$T_b(x) = \begin{cases} c - c^{-\ell_b+1}(c - x)^{\ell_b}, & \text{if} \quad x \in [0, c), \\ \frac{x-c}{1-c}, & \text{if} \quad x \in [c, 1], \end{cases}$$

where $\ell_b > 1$. Again, one easily computes that each map from $\mathfrak{G} \cup \mathfrak{B}$ has non-positive Schwarzian derivative when restricted to $[0, \frac{1}{2})$ or $[\frac{1}{2}, 1]$. For these collections of maps Lemma 3.2.2 carries over, replacing $\frac{1}{2}$ with c and $\frac{1}{4}$ with c^2, under the additional assumptions that $c < r_g \cdot (1-c)^{1-r_g^{-1}}$ holds for all $g \in \Sigma_G$, and that $1 - c > \ell_b^{\ell_b/(1-\ell_b)}$ holds for all $b \in \Sigma_B$. Furthermore, Lemma 3.3.7 carries over with lower bounds $\frac{(1-c)^2}{c}$ and $\frac{1}{1-c}$ in (i) and (ii) instead of $\frac{1}{2}$ and 2, respectively, under the additional assumption that

$$\frac{(1-c)^2}{c} \cdot \min\left\{ \min_{b \in \Sigma_B} \{r_b \cdot (1-c)^{-1/r_b}\}, \ell_{\min}^{1+1/(\ell_{\min}-1)} \right\} \geq 1.$$

By equipping $\Sigma^{\mathbb{N}}$ with the metric $d_{\Sigma^{\mathbb{N}}}(\omega, \omega') = (1-c)^{\min\{i \in \mathbb{N} : \omega_i \neq \omega_i'\}}$, it can be shown that under these additional conditions all the results formulated in Sections 3.1 and 3.5 carry over and are proven in the same way.

Finally, polynomial decay of correlations is expected to hold for a more general class of good and bad maps for which random compositions show critical intermittency, but the proofs may become more cumbersome. Our assumption that all maps are identical on the interval $[\frac{1}{2}, 1]$ made it easier to find a suitable inducing domain, but does not seem necessary. Furthermore, the linearity of this right branch and the explicit forms of the left branches of the good and bad maps made the series in (3.24) and (3.25) telescopic. A first step to generalise our results to a more general class might be to require this explicit form of the left branch only close to c, though any generalisations will inevitably make the calculations more complicated.

CHAPTER 4

Intermittency generated by attracting and weakly repelling fixed points

This chapter is based on: [Z].

Abstract

In Chapter 2 for a class of critically intermittent random systems a phase transition was found for the finiteness of the absolutely continuous invariant measure. The systems for which this result holds are characterised by the interplay between a superexponentially attracting fixed point and an exponentially repelling fixed point. In this chapter we consider a closely related family of random systems with instead exponentially fast attraction to and polynomially fast repulsion from two fixed points, and show that such a phase transition still exists. The method of the proof however is different and relies on the construction of a suitable invariant set for the Perron-Frobenius operator.

§4.1 Introduction

For the critically intermittent random systems studied in Chapter 2 we asked in Subsection 2.4.3 the question what happens to the absolutely continuous invariant measure, if it exists, when the superexponential convergence to c is replaced by exponential convergence to c and the exponential divergence from 0 and 1 is replaced by polynomial divergence from 0 and 1. In this chapter we investigate this by considering a random system that generates i.i.d. random compositions of a finite fixed number of maps of two types: Type 1 consists of the LSV maps from (1.11) and type 2 consists of LSV maps where the right branch is replaced by increasing branches that map $(\frac{1}{2}, 1]$ to itself and for which the derivative close to $\frac{1}{2}$ is smaller than 1. The random orbits then converge exponentially fast to $\frac{1}{2}$ under applications of maps of type 2, and as soon as a map of type 1 is applied then diverge polynomially fast from 0, see Figure 4.1(a). We will show that such random systems exhibit a phase transition similar to the ones found in Chapters 2 and 3 in the sense that it depends on the features of the maps as well as on the probabilities of choosing the maps whether the system admits a finite absolutely continuous invariant measure or not.

We define the class $\mathfrak{S} = \{S_\alpha : \alpha \in (0, \infty)\}$ where S_α is the LSV map from (1.11), and the class $\mathfrak{R} = \{R_{\alpha,K} : \alpha \in (0, \infty), K \in (0, 1)\}$ where

$$R_{\alpha,K}(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \text{if} \quad x \in [0, \frac{1}{2}], \\ \frac{1}{2} + K(x - \frac{1}{2}) + 2(1 - K)(x - \frac{1}{2})^2 & \text{if} \quad x \in (\frac{1}{2}, 1]. \end{cases} \tag{4.1}$$

The graph of $R_{\alpha,K}$ is shown in Figure 4.1(b). The right branch of $R_{\alpha,K}$ is defined in such a way that $\frac{1}{2}$ and 1 are fixed points for $R_{\alpha,K}$ and that under $R_{\alpha,K}$ orbits eventually approach $\frac{1}{2}$ from above. The rate of this convergence to $\frac{1}{2}$ is determined by K. Let $T_1, \ldots, T_N \in \mathfrak{S} \cup \mathfrak{R}$ be a finite collection. Similar as in the previous chapters we write

$$\begin{aligned} \Sigma_S &= \{1 \le j \le N : T_j \in \mathfrak{S}\}, \\ \Sigma_R &= \{1 \le j \le N : T_j \in \mathfrak{R}\}, \\ \Sigma &= \{1, \ldots, N\} = \Sigma_S \cup \Sigma_R. \end{aligned}$$

We assume that $\Sigma_S, \Sigma_R \neq \emptyset$. For each $j \in \Sigma$ we write $\alpha_j \in (0, \infty)$ if $T_j(x) = x(1 + 2^{\alpha_j} x^{\alpha_j})$ for $x \in [0, \frac{1}{2}]$. For $j \in \Sigma_R$ we moreover write $K_j \in (0, 1)$ if $T_j(x) = \frac{1}{2} + K_j(x - \frac{1}{2}) + 2(1 - K_j)(x - \frac{1}{2})^2$ for $x \in (\frac{1}{2}, 1]$.

Let F be the skew product associated to $\{T_j\}_{j \in \Sigma}$, i.e.

$$F : \Sigma^{\mathbb{N}} \times [0, 1] \to \Sigma^{\mathbb{N}} \times [0, 1],\ (\omega, x) \mapsto (\tau\omega, T_{\omega_1}(x)), \tag{4.2}$$

where τ denotes the left shift on sequences in $\Sigma^{\mathbb{N}}$. Let $\boldsymbol{p} = (p_j)_{j \in \Sigma}$ be a probability vector with strictly positive entries representing the probabilities with which we choose the maps from $\mathcal{T} = \{T_j\}_{j \in \Sigma}$. Let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $\Sigma^{\mathbb{N}}$. Since each of the maps T_j $(j \in \Sigma)$ has zero as a neutral fixed point, orbits under $(\mathcal{T}, \boldsymbol{p})$ exhibit intermittent behaviour in the sense that periods of chaotic behaviour are followed by

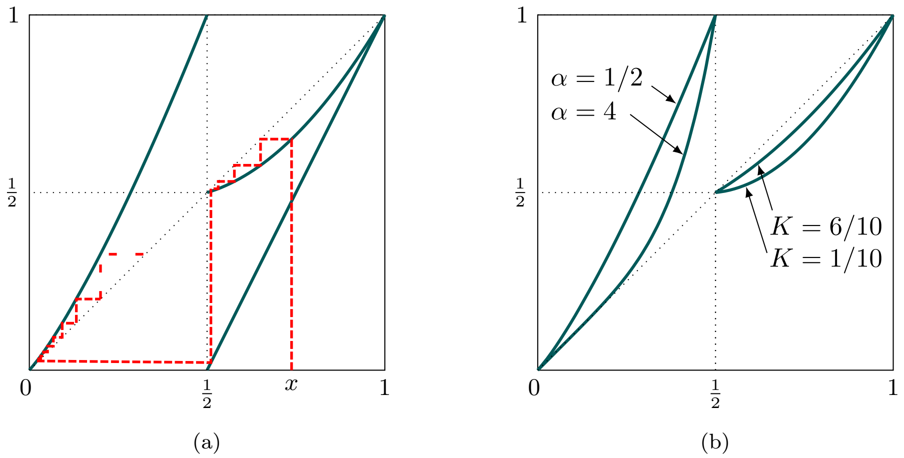


Figure 4.1: In (a) *we see the critically intermittent system consisting of the maps* S_α *and* $R_{\alpha,K}$ *given by* (1.11) *and* (4.1)*, respectively. The dashed lines indicate part of a random orbit of* x*. In* (b) *the graph of* $R_{\alpha,K}$ *is depicted for several values of* α *and* K*.*

periods of spending time near zero. The amount of time spent near zero generally increases for larger values of p_j $(j \in \Sigma_R)$, smaller values of K_j $(j \in \Sigma_R)$ and larger values of α_j $(j \in \Sigma)$.

We set $\alpha_{\min} = \min\{\alpha_j : j \in \Sigma\}$. Throughout this chapter we assume the following:

Assumption: $\alpha_{\min} < 1$.

Furthermore, we set

$$\eta = \sum_{r \in \Sigma_R} p_r K_r^{-\alpha_{\min}},$$
$$\gamma = \sup\{\delta \geq 0 : \sum_{r \in \Sigma_R} p_r K_r^{-\delta} < 1\}.$$

Note that if $\eta < 1$, then $\gamma > \alpha_{\min}$. We have the following main results.

Theorem 4.1.1. *Suppose* $\eta > 1$*. Then no acs probability measure exists for* $(\mathcal{T}, \boldsymbol{p})$*.*

Theorem 4.1.2. *Suppose* $\eta < 1$*.*

(1) There exists a unique acs probability measure μ *for* $(\mathcal{T}, \boldsymbol{p})$*. Moreover,* F *is ergodic with respect to* $m_{\boldsymbol{p}} \times \mu$*.*

(2) The density $\frac{d\mu}{d\lambda}$ *with respect to the Lebesgue measure* λ *is bounded away from zero and on the intervals* $(0, \frac{1}{2}]$ *and* $(\frac{1}{2}, 1]$ *is decreasing and locally Lipschitz. Furthermore, for each* $\beta \in (\alpha_{\min}, \gamma) \cap (0, 1]$ *there exist* $a_1, a_2 > 0$ *such that*

$$\frac{d\mu}{d\lambda}(x) \leq a_1 \cdot x^{-\alpha_{\min}-1+\beta}, \qquad x \in \Big(0, \frac{1}{2}\Big], \tag{4.3}$$
$$\frac{d\mu}{d\lambda}(x) \leq a_2 \cdot \Big(x - \frac{1}{2}\Big)^{-1+\beta}, \qquad x \in \Big(\frac{1}{2}, 1\Big]. \tag{4.4}$$

See Figure 4.2 for a plot of $\frac{d\mu}{d\lambda}$. In addition to Theorem 4.1.1 we argue in Section 4.3 that if $\eta > 1$ then an infinite acs measure exists and no physical measure (see footnote 1 on page 39) for F exists. Together with Theorem 4.1.2 this shows that the random system undergoes a phase transition with threshold $\eta = 1$. It is not clear if an acs probability measure exists when $\eta = 1$. We discuss this in Section 4.3 as well. Note that if $\sum_{r\in\Sigma_R} p_r K_r^{-1} < 1$, then $\gamma > 1$. So in this case we can take $\beta = 1$, and then Theorem 4.1.2(2) says that there exists $a > 0$ such that

$$\frac{d\mu}{d\lambda}(x) \le a \cdot x^{-\alpha_{\min}}, \qquad x \in (0,1]. \tag{4.5}$$

This bound is also found in [LSV99] where only one LSV map $T_1 \in \mathfrak{G}$ with $\alpha_1 \in (0,1)$ is considered and no maps in $\mathfrak{R}$. This suggest that in case $\sum_{r\in\Sigma_R} p_r K_r^{-1} < 1$ the attraction by the maps $\{T_j\}_{j\in\Sigma_R}$ to $\frac{1}{2}$ does not change the order of the pole of the invariant density at zero. Note however that the density in the setting of [LSV99] is shown to be continuous on $(0,1)$, which in general is not the case for the density in the setting of Theorem 4.1.2. See Figure 4.2(b).

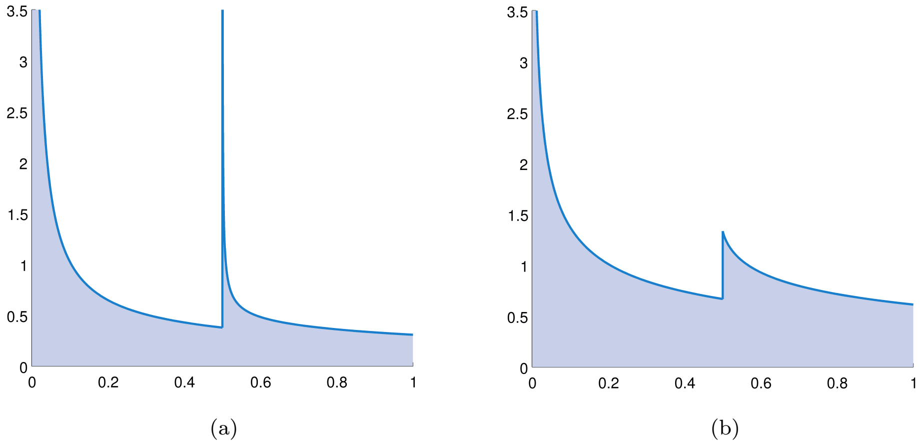


Figure 4.2: Approximation of $\frac{d\mu}{d\lambda}$ *in case* $\Sigma_S = \{1\}$, $\Sigma_R = \{2\}$, $p_1 = \frac{7}{10}$ *and* $\alpha_1 = \alpha_2 = \frac{1}{2}$ *for two different values of* K_2. *Both pictures depict* $P^{100}(1)$ *with* P *as in* (4.17), *where in* (a) *we have taken* $K_2 = \frac{1}{10}$ *(so* $\eta < 1 < p_2 K_2^{-1}$*) and in* (b) $K_2 = \frac{6}{10}$ *(so* $\eta < p_2 K_2^{-1} < 1$*).*

With Theorem 4.1.2 we can derive the following result, which says that the density $\frac{d\mu_{\boldsymbol{p}}}{d\lambda}$ in $L^1(\lambda) = L^1([0,1],\lambda)$ depends continuously on the probability vector $\boldsymbol{p} \in \mathbb{R}^N$ w.r.t. the $L^1(\lambda)$-norm. Here we write $\mu_{\boldsymbol{p}}$ for the acs probability measure that corresponds to the probability vector $\boldsymbol{p}$.

Corollary 4.1.3. *For each* $n \in \mathbb{N}$, *let* $\boldsymbol{p}_n = (p_{n,j})_{j\in\Sigma}$ *be a strictly positive probability vector such that* $\sup_n \sum_{r\in\Sigma_R} p_{n,r} K_r^{-\alpha_{\min}} < 1$ *and assume that* $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ *in* $\mathbb{R}^N_+$. *Then*

$$\lim_{n\to\infty} \left\| \frac{d\mu_{\boldsymbol{p}_n}}{d\lambda} - \frac{d\mu_{\boldsymbol{p}}}{d\lambda} \right\|_{L^1(\lambda)} = 0.$$

Note that the convergence in Corollary 4.1.3 is stronger than in Corollary 2.1.5 where only weak convergence for the acs measure is derived.

Let us briefly give a heuristic explanation of why the value of η determines whether $(\mathcal{T}, \boldsymbol{p})$ admits an acs probability measure or not. We do this by referring to techniques involving inducing. First of all, as results like Kac's Lemma, Proposition 1.2.12 and the Young tower technique discussed in Chapter 3 indicate, given a subset Y in which orbits stay relatively short, the expected time[1] to first return to Y after leaving Y is finite typically if and only if an acs probability measure exists. Let us now argue for our random systems that for $\eta < 1$ ($\eta > 1$) the expected time to first return to $Y = \Sigma^{\mathbb{N}} \times (\frac{1}{2}, 1)$ after leaving Y is finite (infinite), thus suggesting the result of Theorem 4.1.1 and the first part of Theorem 4.1.2. If $\Sigma_R = \emptyset$, then our system is an i.i.d. random LSV map and as explained in Example 1.4.3 we know that in this case the existence of an acs probability measure depends on how long orbits stick close to zero. As follows from the results in [BBD14, BB16, Z18, BQT21], this stickiness at zero is governed by the LSV map with the fastest relaxation rate, i.e. having parameter $\alpha_{\min}$. In particular, a point $(\omega, x) \in Y$ with x close to $\frac{1}{2}$ typically needs of the order $(x - \frac{1}{2})^{-\alpha_{\min}}$ iterations under F to first return to Y as shown in [BB16, Theorem 1.1]. In this case the expected return time to Y behaves roughly as $\kappa := \int_{\frac{1}{2}}^{1} (x - \frac{1}{2})^{-\alpha_{\min}} dx$, which is finite if and only if $\alpha_{\min} < 1$. If $\Sigma_R \neq \emptyset$, then the influence of the stickiness at zero is enhanced because points in $(\frac{1}{2}, 1)$ close to $\frac{1}{2}$ are sent closer to zero when first a number of times maps from $\mathfrak{R}$ are applied before a map from $\mathfrak{G}$ is applied. In this case the expected return time to Y after leaving Y behaves roughly like

$$\sum_{m=0}^{\infty} \sum_{r_1 \in \Sigma_R} \cdots \sum_{r_m \in \Sigma_R} \Big(\prod_{j=1}^{m} p_{r_j} \Big) \int_{\frac{1}{2}}^{1} \Big(T_{r_m} \circ \cdots \circ T_{r_1}(x) - \frac{1}{2} \Big)^{-\alpha_{\min}} dx. \tag{4.6}$$

First of all, for all $r \in \Sigma_R$ and $x \in (\frac{1}{2}, 1)$ we have $T_r(x) \geq \frac{1}{2} + K_r(x - \frac{1}{2})$ and so the quantity in (4.6) can be bounded from above by $\kappa \sum_{m=0}^{\infty} \eta^m$. Hence, if $\eta < 1$, then it is reasonable to expect that the expected return time to Y after leaving Y is finite. On the other hand, if $\eta > 1$, then there exists $\varepsilon > 0$ small enough such that $\eta_\varepsilon := \sum_{r \in \Sigma_R} p_r (K_r + \varepsilon)^{-\alpha_{\min}} > 1$ as well. Since for $x \in (\frac{1}{2}, 1)$ sufficiently close to $\frac{1}{2}$ we have $T_r(x) \leq \frac{1}{2} + (K_r + \varepsilon)(x - \frac{1}{2})$, the quantity in (4.6) can be bounded from below by $\tilde{\kappa} \sum_{m=0}^{\infty} \eta_\varepsilon^m = \infty$ with $\tilde{\kappa} \in (0, \kappa]$. Hence, if $\eta > 1$, then this suggests that the expected return time to Y after leaving Y is infinite.

For the proof of Theorem 4.1.1 we will work out the above sketch in more detail and obtain the result using Kac's Lemma. On the other hand, for the proof of Theorem 4.1.2 we will not make use of an inducing technique. The first reason is that working out in precise detail the above sketch of bounding the expected return time still requires additional work that is not straightforward. Secondly, as we have seen in Subsection 2.2 and Chapter 3, inducing techniques often require the induced transformation to satisfy certain bounded distortion conditions, which are hard to obtain for the random systems in this chapter since the first branch of the maps T_j

[1] Here we mean with 'expected' that we take the expectation with respect to a reference measure, which in our random system is $m_{\boldsymbol{p}} \times \lambda$.

can have positive Schwarzian derivative (namely if $\alpha_j > 1$). For this reason but also because the maps T_r ($r \in \Sigma_R$) have a discontinuity, we cannot use the method from Subsection 2.3.2 either. Furthermore, we remark that not only in Chapters 2 and 3 but also in this chapter we cannot use the technique from Pelikan in [P84, Section 4] discussed in Section 2.1. The main reason is that the constituent maps S and T from Pelikan have competing behaviour at the same fixed point, whereas our systems are characterised by the interplay between the behaviour at two different fixed points.

Instead, for the case $\eta < 1$ we show the existence of an acs probability measure by considering a suitable set of functions that is invariant with respect to the Perron-Frobenius operator of the random system. We will then apply the Arzelà-Ascoli Theorem to prove that this set has a fixed point. This approach is similar to the one in Section 2 of [LSV99] where only one LSV map is considered.

The remainder of this chapter is organised as follows. Section 4.2 concentrates on proving Theorems 4.1.1 and 4.1.2 and Corollary 4.1.3. This chapter will be concluded in Section 4.3 with some final remarks.

§4.2 Phase transition for the acs measure

As in Section 4.1, let $T_1, \ldots, T_N \in \mathfrak{S} \cup \mathfrak{R}$ be a finite collection, write $\Sigma_S = \{1 \le j \le N : T_j \in \mathfrak{S}\}$, $\Sigma_R = \{1 \le j \le N : T_j \in \mathfrak{R}\}$ and $\Sigma = \{1, \ldots, N\} = \Sigma_S \cup \Sigma_R$ and assume that $\Sigma_S, \Sigma_R \neq \emptyset$ and $\alpha_{\min} < 1$. Furthermore, we again denote by F the skew product associated to $\mathcal{T} = \{T_j\}_{j\in\Sigma}$ given by (4.2), let $\boldsymbol{p} = (p_j)_{j\in\Sigma}$ be a probability vector with strictly positive entries and let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $\Sigma^{\mathbb{N}}$. Also, recall that

$$\eta = \sum_{r\in\Sigma_R} p_r K_r^{-\alpha_{\min}}.$$

§4.2.1 The case $\eta > 1$

In this subsection we prove Theorem 4.1.1, namely that any acs measure for $(\mathcal{T}, \boldsymbol{p})$ must be infinite if $\eta > 1$. Throughout this subsection we use the notations for words and compositions of the maps T_j introduced in Section 1.4. Furthermore, we will use the following well-known results.

Let $j \in \Sigma$ and define the sequence $\{x_n(j)\}$ in $(0, \frac{1}{2}]$ by

$$x_1(j) = \frac{1}{2} \quad \text{and} \quad x_n(j) = T_j|_{[0,\frac{1}{2}]}^{-1}\big(x_{n-1}(j)\big) \text{ for each integer } n \ge 2.$$

As explained in e.g. the beginning of Section 6.2 of [Y99] there exists a constant $C_j > 1$ such that for each $n \in \mathbb{N}$

$$C_j^{-1} n^{-\frac{1}{\alpha_j}} \le x_n(j) \le C_j n^{-\frac{1}{\alpha_j}}. \tag{4.7}$$

Furthermore, we define for each $\omega \in \Sigma^{\mathbb{N}}$ the random sequence $\{x_n(\omega)\}$ in $(0, \frac{1}{2}]$ by

$$x_1(\omega) = \frac{1}{2} \quad \text{and} \quad x_n(\omega) = T_{\omega_1}|_{[0,\frac{1}{2}]}^{-1}(x_{n-1}(\tau\omega)) \text{ for each integer } n \geq 2.$$

Then, for each $\omega \in \Sigma^{\mathbb{N}}$ and $n \in \mathbb{N}$,

$$T_\omega^{n-1}((x_{n+1}(\omega), x_n(\omega)]) = \Big(x_2(\tau^{n-1}\omega), \frac{1}{2}\Big]. \tag{4.8}$$

Letting $i \in \Sigma$ be such that $\alpha_i = \alpha_{\min}$, it has been shown in [BBD14, Lemma 4.4] that for each $\omega \in \Sigma^{\mathbb{N}}$ and $n \in \mathbb{N}$ we have

$$x_n(i) \leq x_n(\omega). \tag{4.9}$$

Proof of Theorem 4.1.1. Suppose that $\eta > 1$ and that μ is an acs probability measure for $(\mathcal{T}, \boldsymbol{p})$. We will use Kac's Lemma to arrive at a contradiction. Define

$$\begin{aligned} A_j &= \Big(x_2(j), T_j|_{[0,\frac{1}{2}]}^{-1}\Big(\frac{3}{4}\Big)\Big), \qquad j \in \Sigma, \\ B_j &= \Big(\frac{3}{4}, T_j|_{(\frac{1}{2},1]}^{-1}\Big(\frac{3}{4}\Big)\Big), \qquad j \in \Sigma, \\ Y &= \bigcup_{j \in \Sigma} [j] \times (A_j \cup B_j). \end{aligned}$$

We consider the first return time map φ_Y to Y under F as defined in (1.6). Since $\eta > 1$, there exists $\delta > 0$ small enough such that

$$\zeta := \sum_{r \in \Sigma_R} p_r M_r^{-\alpha_{\min}} \geq 1, \qquad \text{where } M_r := K_r + 2(1-K_r) \cdot \delta \;\; \forall r \in \Sigma_R. \tag{4.10}$$

For each $x \in (\frac{1}{2}, \frac{1}{2} + \delta)$ we have

$$T_r(x) = \frac{1}{2} + \Big(K_r + 2(1-K_r)\Big(x - \frac{1}{2}\Big)\Big)\Big(x - \frac{1}{2}\Big) \leq \frac{1}{2} + M_r\Big(x - \frac{1}{2}\Big). \tag{4.11}$$

For $\boldsymbol{r} = (r_1, \ldots, r_n) \in \Sigma_R^n$ we write $M_{\boldsymbol{r}} = \prod_{l=1}^n M_{r_l}$ with $M_{\boldsymbol{r}} = 1$ if $n = 0$. Furthermore, fix $t \in \Sigma_R$. It is easy to see that $\lim_{n\to\infty} T_t^n(\frac{3}{4}) = \frac{1}{2}$, so there exists an integer $k \geq 0$ such that $T_t^k(\frac{3}{4}) \in (\frac{1}{2}, \frac{1}{2} + \delta)$ holds.

Let $(\omega, x) \in Y$ and t and k be as above. Furthermore, fix $s \in \Sigma_S$. Suppose that

$$\omega \in [u \underbrace{t \cdots t}_{k \text{ times}} \boldsymbol{r} s] = [u t^k \boldsymbol{r} s], \quad \text{for some } u \in \Sigma,\ \boldsymbol{r} \in \Sigma_R^n, n \geq 0.$$

We then have $T_\omega^l(x) \in (\frac{1}{2}, \frac{3}{4})$ for all $1 \leq l \leq 1 + k + n$. It follows from $T_{\omega_1}(x) \leq \frac{3}{4}$, $T_t^k(\frac{3}{4}) \in (\frac{1}{2}, \frac{1}{2} + \delta)$ and (4.11) that

$$T_\omega^{1+k+n}(x) \leq T_{\tau\omega}^{k+n}\Big(\frac{3}{4}\Big) \leq \frac{1}{2} + M_{\boldsymbol{r}}\Big(T_t^k\Big(\frac{3}{4}\Big) - \frac{1}{2}\Big),$$

which gives

$$T_\omega^{2+k+n}(x) \leq M_{\boldsymbol{r}}\Big(2T_t^k\Big(\frac{3}{4}\Big)-1\Big) \tag{4.12}$$

Fix $i \in \Sigma$ such that $\alpha_i = \alpha_{\min}$. There exists an $m \in \mathbb{N}$ such that $T_\omega^{2+k+n}(x) \in (x_{m+1}(i), x_m(i)]$. It follows from (4.8) and (4.9) that

$$\varphi_Y(\omega, x) \geq 2+k+n+m. \tag{4.13}$$

We give a lower bound for m in terms of $\boldsymbol{r}$. It follows from (4.7) and (4.12) that

$$C_i^{-1}(m+1)^{-\frac{1}{\alpha_i}} \leq M_{\boldsymbol{r}}\Big(2T_t^k\Big(\frac{3}{4}\Big)-1\Big).$$

Solving for m yields

$$m \geq D_1 \cdot M_{\boldsymbol{r}}^{-\alpha_i} - 1, \tag{4.14}$$

where we defined $D_1 = C_i^{-\alpha_i} \cdot (2T_t^k(\frac{3}{4})-1)^{-\alpha_i}$. Combining (4.13) and (4.14) yields

$$\begin{aligned}\int_Y \varphi_Y \, dm_{\boldsymbol{p}} \times \mu &\geq \sum_{u\in\Sigma}\sum_{n=0}^{\infty}\sum_{\boldsymbol{r}\in\Sigma_R^n}\int_{[ut^k\boldsymbol{r}s]\times(A_u\cup B_u)} \varphi_Y \, dm_{\boldsymbol{p}} \times \mu \\ &\geq \sum_{u\in\Sigma}\sum_{n=0}^{\infty}\sum_{\boldsymbol{r}\in\Sigma_R^n} m_{\boldsymbol{p}}\big([ut^k\boldsymbol{r}s]\big)\int_{A_u\cup B_u} D_1 \cdot M_{\boldsymbol{r}}^{-\alpha_i}\, d\mu(x) \\ &= D_2 \cdot \sum_{n=0}^{\infty}\zeta^n,\end{aligned} \tag{4.15}$$

where

$$D_2 = D_1 \cdot p_t^k p_s \cdot \sum_{u\in\Sigma} p_u \mu(A_u \cup B_u) = D_1 \cdot p_t^k p_s \cdot m_{\boldsymbol{p}} \times \mu(Y).$$

Almost every orbit that starts in $\Sigma^{\mathbb{N}} \times [0,1]$ will eventually enter $\Sigma^{\mathbb{N}} \times (\frac{1}{2}, \frac{3}{4})$, either via $\bigcup_{j\in\Sigma}[j] \times A_j$ or via $\bigcup_{j\in\Sigma}[j]\times B_j$. Hence, we have $\bigcup_{n=0}^{\infty} F^{-n}Y = \Sigma^{\mathbb{N}} \times [0,1]$ up to some set of measure zero, i.e. Y is a sweep-out set. This together with the F-invariance of $m_{\boldsymbol{p}} \times \mu$ yields

$$1 = m_{\boldsymbol{p}} \times \mu(\Sigma^{\mathbb{N}} \times [0,1]) \leq \sum_{n=0}^{\infty} m_{\boldsymbol{p}} \times \mu(F^{-n}Y) = \sum_{n=0}^{\infty} m_{\boldsymbol{p}} \times \mu(Y).$$

This gives $m_{\boldsymbol{p}} \times \mu(Y) > 0$ and so $D_2 > 0$. Hence, from (4.15) and $\zeta \geq 1$ it now follows that

$$\int_Y \varphi_Y \, dm_{\boldsymbol{p}} \times \mu = \infty. \tag{4.16}$$

On the other hand, since μ is a probability measure by assumption, we obtain from the Ergodic Decomposition Theorem and Kac's Lemma in a similar way as in Subsection 3.3.1 that

$$\int_Y \varphi_Y \, dm_{\boldsymbol{p}} \times \mu \leq 1,$$

which is in contradiction with (4.16). □

§4.2.2 The case $\eta < 1$

In this subsection we will prove Theorem 4.1.2 and Corollary 4.1.3. For this we will identify a suitable set of functions which is preserved by the Perron-Frobenius operator $P = P_{\mathcal{T},\boldsymbol{p}}$ associated to $(\mathcal{T},\boldsymbol{p})$ being of the form as in (1.20). We will do this in a number of steps in a way that is similar to the approach of Section 2 in [LSV99].

Suppose $\eta < 1$. On $[0,1]$ we define for each $j \in \Sigma$ the functions $x \mapsto y_j(x)$ and $x \mapsto \xi_j(x)$ by $y_j(x) = (T_j|_{[0,\frac{1}{2}]})^{-1}(x)$ and $\xi_j(x) = (2y_j(x))^{\alpha_j}$. Furthermore, we define on $[0,1]$ the function $z(x) = \frac{x+1}{2}$ and on $(\frac{1}{2},1]$ we define for each $r \in \Sigma_R$ the function $z_r(x) = (T_r|_{(\frac{1}{2},1]})^{-1}(x)$. Whenever convenient, we will just write y_j for $y_j(x)$ and similarly for ξ_j, z and z_r. Writing $p_S = \sum_{s\in\Sigma_S} p_s$, we then have

$$Pf(x) = \begin{cases} \sum_{j\in\Sigma} p_j \frac{f(y_j)}{1+(\alpha_j+1)\xi_j} + p_S \frac{f(z)}{2}, & x \in [0,\frac{1}{2}] \\ \sum_{j\in\Sigma} p_j \frac{f(y_j)}{1+(\alpha_j+1)\xi_j} + p_S \frac{f(z)}{2} + \sum_{r\in\Sigma_R} p_r \frac{f(z_r)}{DR_{\alpha,K}(z_r)}, & x \in (\frac{1}{2},1]. \end{cases} \tag{4.17}$$

Note that $x \mapsto y_j(x)$, $x \mapsto \xi_j(x)$, $x \mapsto z(x)$ and $x \mapsto z_r(x)$ are increasing and continuous on $(0,\frac{1}{2}]$ and $(\frac{1}{2},1]$. This in combination with the fact that $R_{\alpha,K}$ is C^1 on $(\frac{1}{2},1]$ with increasing derivative gives that the set

$$\mathcal{C}_0 = \left\{ f \in L^1(\lambda) : f \geq 0, f \text{ decreasing and continuous on } \left(0,\frac{1}{2}\right] \text{ and } \left(\frac{1}{2},1\right]\right\}$$

is preserved by P, i.e. $P\mathcal{C}_0 \subseteq \mathcal{C}_0$.

Since $\eta < 1$, we have $\gamma = \sup\{\delta \geq 0 : \sum_{r\in\Sigma_R} p_r K_r^{-\delta} < 1\} > \alpha_{\min}$, so $(\alpha_{\min},\gamma)$ is non-empty. In the remainder of this subsection we fix a $\beta \in (\alpha_{\min},\gamma) \cap (0,1]$. We set $\alpha_{\max} = \max\{\alpha_j : j \in \Sigma\}$ and $d = \alpha_{\max} + 2$. We need the following two lemmas.

Lemma 4.2.1. *For each $\alpha > 0$ the function $x \mapsto \frac{(1+x)^d}{1+(\alpha+1)x}$ is increasing on $[0,1]$.*

Proof. Set

$$f_\alpha(x) = \frac{(1+x)^d}{1+(\alpha+1)x}, \qquad x \in [0,1].$$

Furthermore, set $g(x) = (1+x)^d$ and $h_\alpha(x) = 1 + (\alpha+1)x$ where $x \in [0,1]$. Then

$$f'_\alpha(x) = \frac{h_\alpha(x)g'(x) - g(x)h'_\alpha(x)}{h_\alpha(x)^2}.$$

We have

$$\begin{aligned} h_\alpha(x)g'(x) &= (1+(\alpha+1)x) \cdot d(1+x)^{d-1} \\ &\geq (1+x)^d \cdot d \geq (1+x)^d \cdot (\alpha+1) \\ &= g(x)h'_\alpha(x), \end{aligned}$$

so $f'_\alpha(x) \geq 0$ holds for all $x \in [0,1]$. □

CHAPTER 4

Define for each $K > 0$ and $b \geq 0$ the function $H_{K,b} : (\frac{1}{2}, 1] \to \mathbb{R}$ by

$$H_{K,b}(x) = \frac{(K + 2(1-K)(x - \frac{1}{2}))^b}{K + 4(1-K)(x - \frac{1}{2})}, \qquad x \in \Big(\frac{1}{2}, 1\Big].$$

Lemma 4.2.2. *Let $K > 0$ and $b \geq 0$.*

(i) If $b \geq 2$, then $H_{K,b}$ is increasing.

(ii) If $b \leq 1$, then $H_{K,b}$ is decreasing.

Proof. Set $f_K(x) = K + 2(1-K)(x - \frac{1}{2})$ and $g_K(x) = K + 4(1-K)(x - \frac{1}{2})$ where $x \in (\frac{1}{2}, 1]$. Note that $g_K'(x) = 2f_K'(x)$. Then for $x \in (\frac{1}{2}, 1)$

$$\begin{aligned} H_{K,b}'(x) &= \frac{g_K(x) \cdot b \cdot f_K(x)^{b-1} f_K'(x) - f_K(x)^b \cdot g_K'(x)}{g_K(x)^2} \\ &= \frac{f_K(x)^b \cdot f_K'(x)\big(b \cdot \frac{g_K(x)}{f_K(x)} - 2\big)}{g_K(x)^2}. \end{aligned}$$

If $b \geq 2$, then

$$b \cdot \frac{g_K(x)}{f_K(x)} - 2 \geq 2 \cdot \frac{g_K(x)}{f_K(x)} - 2 \geq 2 \cdot \frac{f_K(x)}{f_K(x)} - 2 = 0$$

and thus $H_{K,b}'(x) \geq 0$. This proves (i). If $b \leq 1$, then

$$b \cdot \frac{g_K(x)}{f_K(x)} - 2 \leq \frac{g_K(x)}{f_K(x)} - 2 \leq \frac{2f_K(x)}{f_K(x)} - 2 = 0$$

and thus $H_{K,b}'(x) \leq 0$. This proves (ii). □

We can now prove the following lemma.

Lemma 4.2.3. *The set*

$$\mathcal{C}_1 = \Big\{ f \in \mathcal{C}_0 : x \mapsto x^d f(x) \text{ incr. on } \Big(0, \frac{1}{2}\Big],\ x \mapsto \Big(x - \frac{1}{2}\Big)^d f(x) \text{ incr. on } \Big(\frac{1}{2}, 1\Big] \Big\}$$

is preserved by P.

Proof. Let $f \in \mathcal{C}_1$. Let $x \in (0, \frac{1}{2}]$. Using that for each $j \in \Sigma$ we have $x = y_j(1 + \xi_j)$ and that $z(x) - \frac{1}{2} = \frac{x}{2}$, we obtain

$$\begin{aligned} x^d Pf(x) &= \sum_{j \in \Sigma} p_j \Big(\frac{x}{y_j}\Big)^d \frac{y_j^d f(y_j)}{1 + (\alpha_j + 1)\xi_j} + \frac{p_S}{2} \Big(\frac{x}{z - \frac{1}{2}}\Big)^d \Big(z - \frac{1}{2}\Big)^d f(z) \\ &= \sum_{j \in \Sigma} p_j \frac{(1 + \xi_j)^d}{1 + (\alpha_j + 1)\xi_j} \cdot y_j^d f(y_j) + p_S \cdot 2^{d-1} \cdot \Big(z - \frac{1}{2}\Big)^d f(z). \end{aligned}$$

Because $x \mapsto \xi_j(x)$ is increasing for each $j \in \Sigma$ it follows from Lemma 4.2.1 that $x \mapsto \frac{(1+\xi_j(x))^d}{1+(\alpha_j+1)\xi_j(x)}$ is increasing for each $j \in \Sigma$. Combining this with the fact that $f \in \mathcal{C}_1$, that $y_j \in (0, \frac{1}{2}]$ for each $j \in \Sigma$ and that $z \in (\frac{1}{2}, 1]$ we conclude that $x \mapsto x^d Pf(x)$ is increasing on $(0, \frac{1}{2}]$.

Now let $x \in (\frac{1}{2}, 1]$. Then

$$\begin{aligned}\Big(x-\frac{1}{2}\Big)^d Pf(x) &= \Big(\frac{x-\frac{1}{2}}{x}\Big)^d \sum_{j\in\Sigma} p_j \Big(\frac{x}{y_j}\Big)^d \frac{y_j^d f(y_j)}{1+(\alpha_j+1)\xi_j} \\ &+ \frac{p_S}{2}\Big(\frac{x-\frac{1}{2}}{z-\frac{1}{2}}\Big)^d \Big(z-\frac{1}{2}\Big)^d f(z) \\ &+ \sum_{r\in\Sigma_R} \frac{p_r}{DR_{\alpha_r,K_r}(z_r)} \Big(\frac{x-\frac{1}{2}}{z_r-\frac{1}{2}}\Big)^d \Big(z_r-\frac{1}{2}\Big)^d f(z_r).\end{aligned}$$

Using again that for each $j \in \Sigma$ we have $x = y_j(1+\xi_j)$, that $z - \frac{1}{2} = \frac{x}{2}$ and also that $x - \frac{1}{2} = K_r(z_r - \frac{1}{2}) + 2(1-K_r)(z_r - \frac{1}{2})^2$ for each $r \in \Sigma_R$, we obtain

$$\begin{aligned}\Big(x-\frac{1}{2}\Big)^d Pf(x) &= \Big(1-\frac{1}{2x}\Big)^d \sum_{j\in\Sigma} p_j \frac{(1+\xi_j)^d}{1+(\alpha_j+1)\xi_j} \cdot y_j^d f(y_j) \\ &+ \frac{p_S}{2}\Big(2-\frac{1}{x}\Big)^d \Big(z-\frac{1}{2}\Big)^d f(z) \\ &+ \sum_{r\in\Sigma_R} p_r \frac{(K_r+2(1-K_r)(z_r-\frac{1}{2}))^d}{K_r+4(1-K_r)(z_r-\frac{1}{2})} \Big(z_r-\frac{1}{2}\Big)^d f(z_r).\end{aligned}$$

Note that $x \mapsto (1-\frac{1}{2x})^d$ and $x \mapsto (2-\frac{1}{x})^d$ are positive and increasing on $(\frac{1}{2}, 1]$. Combining this with Lemma 4.2.1 and Lemma 4.2.2(i) and with the fact that $f \in \mathcal{C}_1$ we conclude that $x \mapsto (x-\frac{1}{2})^d Pf(x)$ is increasing on $(\frac{1}{2}, 1]$. □

We set $t_1 = \alpha_{\min} + 1 - \beta$ and $t_2 = 1 - \beta$. It follows from $\beta \in (\alpha_{\min}, 1]$ that $t_1 \in [\alpha_{\min}, 1)$ and $t_2 \in [0, 1-\alpha_{\min})$.

Lemma 4.2.4. *For sufficiently large $a_1, a_2 > 0$, the set*

$$\mathcal{C}_2 = \Big\{f \in \mathcal{C}_1 : f(x) \le a_1 x^{-t_1} \text{ on } \Big(0,\frac{1}{2}\Big], f(x) \le a_2\Big(x-\frac{1}{2}\Big)^{-t_2} \text{ on } \Big(\frac{1}{2},1\Big], \int_0^1 f d\lambda = 1\Big\}$$

is preserved by P.

Proof. Let $f \in \mathcal{C}_2$. First, let $x \in (\frac{1}{2}, 1]$. For each $j \in \Sigma$ we have $y_j \le \frac{1}{2}$ and thus, using that $f \in \mathcal{C}_1$,

$$y_j^d f(y_j) \le 2^{-d} f\Big(\frac{1}{2}\Big) \le 2^{-d} \cdot 2 \cdot \int_0^{\frac{1}{2}} f(u)du \le 2^{-d+1}.$$

Furthermore, for each $j \in \Sigma$ we have

$$T_j\Big(\frac{1}{4}\Big) = \frac{1}{4}(1+2^{-\alpha_j}) \le \frac{1}{4}(1+1) = \frac{1}{2},$$

which gives $y_j \in (\frac{1}{4}, \frac{1}{2}]$. Setting $M := 2^{d+1}$ we obtain for each $j \in \Sigma$ that

$$\frac{f(y_j)}{1+(\alpha_j+1)\xi_j} = y_j^d f(y_j) \cdot \frac{y_j^{-d}}{1+(\alpha_j+1)\cdot(2y_j)^{\alpha_j}} \leq 2^{-d+1} \cdot 4^d = M. \tag{4.18}$$

It also follows from $f \in \mathcal{C}_1$ that

$$\left(z-\frac{1}{2}\right)^d f(z) \leq \left(1-\frac{1}{2}\right)^d f(1) \leq 2^{-d} \cdot 2 \cdot \int_{\frac{1}{2}}^1 f(u)du \leq 2^{-d+1}.$$

Using that $z \in (\frac{3}{4}, 1]$, this gives

$$f(z) \leq 2^{-d+1} \cdot \left(z-\frac{1}{2}\right)^{-d} \leq 2^{-d+1} \cdot \left(\frac{3}{4}-\frac{1}{2}\right)^{-d} = M. \tag{4.19}$$

Combining (4.17), (4.18) and (4.19) and using that $f \in \mathcal{C}_2$ gives

$$Pf(x) \leq M + \frac{p_S}{2} \cdot M + \sum_{r \in \Sigma_R} \frac{p_r}{DR_{\alpha_r,K_r}(z_r)} \cdot a_2 \left(z_r - \frac{1}{2}\right)^{-t_2}.$$

For each $r \in \Sigma_R$ we have $x - \frac{1}{2} = K_r(z_r - \frac{1}{2}) + 2(1-K_r)(z_r - \frac{1}{2})^2$ and therefore

$$\frac{1}{DR_{\alpha_r,K_r}(z_r)} \left(\frac{x-\frac{1}{2}}{z_r-\frac{1}{2}}\right)^{t_2} = \frac{(K_r + 2(1-K_r)(z_r-\frac{1}{2}))^{t_2}}{K_r + 4(1-K_r)(z_r - \frac{1}{2})},$$

which by Lemma 4.2.2(ii) can be bounded from above by $H_{K_r,t_2}(\frac{1}{2}) = K_r^{t_2-1}$. Furthermore, since $t_2 \geq 0$ we have $(x-\frac{1}{2})^{t_2} \leq 2^{-t_2}$. We obtain

$$Pf(x) \leq \left\{ \frac{M(1+\frac{p_S}{2}) \cdot 2^{-t_2}}{a_2} + \sum_{r \in \Sigma_R} p_r \cdot K_r^{t_2-1} \right\} \cdot a_2 \cdot \left(x - \frac{1}{2}\right)^{-t_2}. \tag{4.20}$$

We have $t_2 - 1 = -\beta$ and $\beta < \gamma$, so

$$\sum_{r \in \Sigma_R} p_r \cdot K_r^{t_2-1} = \sum_{r \in \Sigma_R} p_r \cdot K_r^{-\beta} < 1.$$

Hence, there exists an $a_2 > 0$ sufficiently large such that the term in braces in (4.20) is bounded by 1.

Now let $x \in (0, \frac{1}{2}]$. Using that $f \in \mathcal{C}_2$, it follows from (4.17) that

$$Pf(x) \leq \sum_{j \in \Sigma} p_j \frac{a_1 \cdot y_j^{-t_1}}{1+(\alpha_j+1)\xi_j} + \frac{p_S \cdot a_2}{2} \cdot \left(z - \frac{1}{2}\right)^{-t_2}. \tag{4.21}$$

For each $j \in \Sigma$ we have, using that $x = y_j(1+\xi_j)$ and that $t_1 \in (0,1)$,

$$\frac{y_j^{-t_1}}{1+(\alpha_j+1)\xi_j} = \frac{x^{-t_1}(1+\xi_j)^{t_1}}{1+(\alpha_j+1)\xi_j} \leq \frac{x^{-t_1}(1+t_1\xi_j)}{1+(\alpha_j+1)\xi_j} \leq x^{-t_1}. \tag{4.22}$$

Fix an $i \in \Sigma$ with $\alpha_i = \alpha_{\min}$. Applying for each $j \in \Sigma \backslash \{i\}$ the bound (4.22) to (4.21) and using that $z - \frac{1}{2} = \frac{x}{2}$ yields

$$Pf(x) \le \Big\{ p_i \Big(\frac{x}{y_i}\Big)^{t_1} \cdot \frac{1}{1 + (\alpha_i + 1)\xi_i} + (1 - p_i) + \frac{p_S \cdot a_2 \cdot 2^{t_2 - 1}}{a_1} \cdot x^{t_1 - t_2} \Big\} \cdot a_1 \cdot x^{-t_1}. \tag{4.23}$$

It remains to find a_1 sufficiently large such that the term in braces in (4.23) is bounded by 1. First of all, using again that $x = y_i(1 + \xi_i)$ and that $t_1 \in (0, 1)$ we get

$$\Big(\frac{x}{y_i}\Big)^{t_1} \cdot \frac{1}{1 + (\alpha_i + 1)\xi_i} = \frac{(1 + \xi_i)^{t_1}}{1 + (\alpha_i + 1)\xi_i} \le \frac{1 + t_1 \xi_i}{1 + (\alpha_i + 1)\xi_i}. \tag{4.24}$$

Furthermore, we have

$$x^{t_1 - t_2} = x^{\alpha_{\min}} = y_i^{\alpha_i}(1 + \xi_i)^{\alpha_i} \le y_i^{\alpha_i} \cdot 2^{\alpha_i} = \xi_i. \tag{4.25}$$

It follows from (4.24) and (4.25) that the term in braces in (4.23) is bounded by

$$p_i \frac{1 + t_1 \xi_i + \frac{p_i^{-1} \cdot p_S \cdot a_2 \cdot 2^{t_2 - 1}}{a_1} \cdot \xi_i \cdot (1 + (\alpha_i + 1)\xi_i)}{1 + (\alpha_i + 1)\xi_i} + (1 - p_i). \tag{4.26}$$

Using that $1 + (\alpha_i + 1)\xi_i \le \alpha_i + 2$ we get that the numerator in (4.26) is bounded by

$$1 + \Big(t_1 + \frac{p_i^{-1} \cdot p_S \cdot a_2 \cdot 2^{t_2 - 1}(\alpha_i + 2)}{a_1}\Big)\xi_i.$$

Taking $a_1 > 0$ sufficiently large such that $t_1 + \frac{p_i^{-1} \cdot p_S \cdot a_2 \cdot 2^{t_2 - 1}(\alpha_i + 2)}{a_1} \le 1 \le \alpha_i + 1$ now yields the result. □

Lemma 4.2.5. *The set $\mathcal{C}_2$ is compact with respect to the $L^1(\lambda)$-norm.*

Proof. For each $f \in \mathcal{C}_2$ let ϕ_f denote the continuous extension of $(0, \frac{1}{2}] \ni x \mapsto x^d f(x)$ to $[0, \frac{1}{2}]$ and let ψ_f denote the continuous extension of $(\frac{1}{2}, 1] \ni x \mapsto (x - \frac{1}{2})^d f(x)$ to $[\frac{1}{2}, 1]$. Furthermore, we define $\mathcal{A}_1 = \{\phi_f : f \in \mathcal{C}_2\}$ and $\mathcal{A}_2 = \{\psi_f : f \in \mathcal{C}_2\}$. For each $f \in \mathcal{C}_2$ we have, for $x, y \in [0, \frac{1}{2}]$ with $x \ge y$, that

$$\begin{aligned} 0 \le \phi_f(x) - \phi_f(y) \le f(x)(x^d - y^d) &\le a_1 x^{-t_1} \cdot d \int_y^x t^{d-1} dt \\ &\le a_1 x^{d-1-t_1} \cdot d|x - y| \le a_1 \cdot 2^{-d+1+t_1} \cdot d|x - y|. \end{aligned} \tag{4.27}$$

and for $x, y \in [\frac{1}{2}, 1]$ with $x \ge y$, that

$$\begin{aligned} 0 \le \psi_f(x) - \psi_f(y) &\le f(x)\Big(\Big(x - \frac{1}{2}\Big)^d - \Big(y - \frac{1}{2}\Big)^d\Big) \\ &\le a_2 \Big(x - \frac{1}{2}\Big)^{-t_2} \cdot d \int_y^x \Big(t - \frac{1}{2}\Big)^{d-1} dt \\ &\le a_2 \Big(x - \frac{1}{2}\Big)^{d-1-t_2} \cdot d|x - y| \le a_2 \cdot 2^{-d+1+t_2} \cdot d|x - y|. \end{aligned} \tag{4.28}$$

Also, from the definition of $\mathcal{C}_2$ in Lemma 4.2.4 and the fact that $d > \max\{t_1, t_2\}$ we see that $\phi_f(0) = \psi_f(\frac{1}{2}) = 0$ holds for each $f \in \mathcal{C}_2$. It follows that $\mathcal{A}_1$ and $\mathcal{A}_2$ are uniformly bounded and equicontinuous, so from the Arzelà-Ascoli Theorem we obtain that $\mathcal{A}_1$ and $\mathcal{A}_2$ are compact in $C([0, \frac{1}{2}])$ and $C([\frac{1}{2}, 1])$, respectively, w.r.t. the supremum norm.

Now let $\{f_n\}$ be a sequence in $\mathcal{C}_2$. It follows from the above that $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ such that $\{\phi_{f_{n_k}}\}$ converges uniformly to some $\phi^* \in C([0, \frac{1}{2}])$ and $\{\psi_{f_{n_k}}\}$ converges uniformly to some $\psi^* \in C([\frac{1}{2}, 1])$ (for this we take a suitable subsequence of a subsequence of $\{f_n\}$). Now define the measurable function f^* on $(0, 1]$ by

$$f^*(x) = \begin{cases} x^{-d}\phi^*(x) & \text{if} \quad x \in (0, \frac{1}{2}], \\ (x - \frac{1}{2})^{-d}\psi^*(x) & \text{if} \quad x \in (\frac{1}{2}, 1]. \end{cases}$$

Then f^* is continuous on $(0, \frac{1}{2}]$ and $(\frac{1}{2}, 1]$. Moreover, $\{f_{n_k}\}$ converges pointwise to f^*. First of all, this gives $f^* \in \mathcal{C}_1$ once we know $f^* \in L^1(\lambda)$. Secondly, this gives combined with

$$\sup_{k \in \mathbb{N}} f_{n_k}(x) \le a_1 x^{-t_1} \text{ for } x \in \Big(0, \frac{1}{2}\Big], \qquad \sup_{k \in \mathbb{N}} f_{n_k}(x) \le a_2 \Big(x - \frac{1}{2}\Big)^{-t_2} \text{ for } x \in \Big(\frac{1}{2}, 1\Big]$$

and

$$\int_0^{\frac{1}{2}} x^{-t_1} dx < \infty, \qquad \int_{\frac{1}{2}}^1 \Big(x - \frac{1}{2}\Big)^{-t_2} dx < \infty,$$

that $f^*(x) \le a_1 x^{-t_1}$ for $x \in (0, \frac{1}{2}]$ and $f^*(x) \le a_2(x - \frac{1}{2})^{-t_2}$ for $x \in (\frac{1}{2}, 1]$, and that

$$\lim_{k \to \infty} \|f^* - f_{n_k}\|_1 = 0 \quad \text{and so} \quad \int_0^1 f^* d\lambda = 1$$

using the Dominated Convergence Theorem. We conclude that $f^* \in \mathcal{C}_2$ and that f^* is a limit point of $\{f_n\}$ with respect to the $L^1(\lambda)$-norm. □

Using the previous lemmas we are now ready to prove Theorem 4.1.2.

Proof of Theorem 4.1.2. (1) Take $f \in \mathcal{C}_2$ and define the sequence of functions $\{f_n\}$ by $f_n = \frac{1}{n}\sum_{i=0}^{n-1} P^i f$. Using that P preserves $\mathcal{C}_2$ and that the average of a finite collection of elements of $\mathcal{C}_2$ is also an element of $\mathcal{C}_2$, we obtain that $\{f_n\}$ is a sequence in $\mathcal{C}_2$. It follows from Lemma 4.2.5 that $\{f_n\}$ has a subsequence $\{f_{n_k}\}$ that converges w.r.t. the $L^1(\lambda)$-norm to some $f^* \in \mathcal{C}_2$. As is standard, we then obtain that $Pf^*(x) = f^*(x)$ holds for λ-a.e. $x \in [0, 1]$ by noting that

$$\begin{aligned} \|Pf^* - f^*\|_1 &\le \|Pf^* - Pf_{n_k}\|_1 + \|Pf_{n_k} - f_{n_k}\|_1 + \|f_{n_k} - f^*\|_1 \\ &\le 2\|f_{n_k} - f^*\|_1 + \Big\|\frac{1}{n_k}\sum_{i=0}^{n_k-1} P^{i+1}f - \frac{1}{n_k}\sum_{i=0}^{n_k-1} P^i f\Big\|_1 \\ &\le 2\|f_{n_k} - f^*\|_1 + \frac{1}{n_k}\|P^{n_k}f - f\|_1 \\ &\le 2\|f_{n_k} - f^*\|_1 + \frac{2}{n_k}\|f\|_1 \to 0, \qquad k \to \infty. \end{aligned}$$

Hence, $(\mathcal{T}, \boldsymbol{p})$ admits an acs probability measure μ with $\frac{d\mu}{d\lambda} \in \mathcal{C}_2$. It follows from the properties of $\mathcal{C}_2$ that $\frac{d\mu}{d\lambda}$ has full support on $[0,1]$, i.e. there is a version such that $\frac{d\mu}{d\lambda}(x) > 0$ for all $x \in [0,1]$, so we obtain from Theorem 1.2.6 that μ is the only acs probability measure once we know that F is ergodic with respect to $m_{\boldsymbol{p}} \times \mu$. So let $A \subseteq \Sigma^{\mathbb{N}} \times [0,1]$ be Borel measurable such that $F^{-1}A = A$. Suppose $m_{\boldsymbol{p}} \times \mu(A) > 0$. The probability measure ρ on $\Sigma^{\mathbb{N}} \times [0,1]$ given by

$$\rho(B) = \frac{m_{\boldsymbol{p}} \times \mu(A \cap B)}{m_{\boldsymbol{p}} \times \mu(A)}$$

for Borel measurable sets $B \subseteq \Sigma^{\mathbb{N}} \times [0,1]$ is F-invariant and absolutely continuous with respect to $m_{\boldsymbol{p}} \times \lambda$ with density

$$\frac{d\rho}{dm_{\boldsymbol{p}} \times \lambda}(\omega, x) = \frac{1}{m_{\boldsymbol{p}} \times \mu(A)} 1_A(\omega, x) \frac{d\mu}{d\lambda}(x), \qquad m_{\boldsymbol{p}} \times \lambda\text{-a.e.} \tag{4.29}$$

According to Lemma 1.4.1 this yields an acs measure $\tilde{\mu}$ for $(\mathcal{T}, \boldsymbol{p})$ such that $\rho = m_{\boldsymbol{p}} \times \tilde{\mu}$. From this we see that also

$$\frac{d\rho}{dm_{\boldsymbol{p}} \times \lambda}(\omega, x) = \frac{d\tilde{\mu}}{d\lambda}(x), \qquad m_{\boldsymbol{p}} \times \lambda\text{-a.e.} \tag{4.30}$$

Write L for the support of $\frac{d\tilde{\mu}}{d\lambda}$, i.e. $L := \{x \in [0,1] : \frac{d\tilde{\mu}}{d\lambda}(x) > 0\}$. Combining (4.29) and (4.30) and using that $\frac{d\mu}{d\lambda}$ has full support on $[0,1]$, we obtain

$$A = \Sigma^{\mathbb{N}} \times L \mod m_{\boldsymbol{p}} \times \lambda. \tag{4.31}$$

Using the non-singularity of F with respect to $m_{\boldsymbol{p}} \times \lambda$, we also obtain from this that

$$F^{-1}A = F^{-1}(\Sigma^{\mathbb{N}} \times L) \mod m_{\boldsymbol{p}} \times \lambda. \tag{4.32}$$

Combining (4.31) and (4.32) with

$$\Sigma^{\mathbb{N}} \times L = \bigcup_{j \in \Sigma} [j] \times L \quad \text{and} \quad F^{-1}(\Sigma^{\mathbb{N}} \times L) = \bigcup_{j \in \Sigma} [j] \times T_j^{-1} L$$

yields

$$L = T_j^{-1} L \mod \lambda$$

for each $j \in \Sigma$. For all $i \in \Sigma$ with $\alpha_i < 1$, in particular for $i \in \Sigma$ with $\alpha_i = \alpha_{\min}$, we have that T_i is ergodic with respect to λ, see e.g. [Y99, Theorem 5]. In particular we have $\lambda(L) \in \{0,1\}$. Together with (4.31) this shows that $m_{\boldsymbol{p}} \times \lambda(A) \in \{0,1\}$. Since $\mu \ll \lambda$, it follows from the assumption $m_{\boldsymbol{p}} \times \mu(A) > 0$ that $m_{\boldsymbol{p}} \times \mu(A) = 1$. We conclude that F is ergodic with respect to $m_{\boldsymbol{p}} \times \mu$.

(2) Since $\frac{d\mu}{d\lambda} \in \mathcal{C}_2$, it follows that $\frac{d\mu}{d\lambda}$ is bounded away from zero, is decreasing on the intervals $(0, \frac{1}{2}]$ and $(\frac{1}{2}, 1]$, and satisfies (4.3) and (4.4) with $a_1, a_2 > 0$ as in

Lemma 4.2.4. Furthermore, applying the last three inequalities in (4.27) with $f = \frac{d\mu}{d\lambda}$ yields, for $x, y \in (0, \frac{1}{2}]$ with $x \geq y$,

$$\begin{aligned} 0 \leq \frac{d\mu}{d\lambda}(y) - \frac{d\mu}{d\lambda}(x) &= y^{-d}\Big(y^d \frac{d\mu}{d\lambda}(y) - y^d \frac{d\mu}{d\lambda}(x)\Big) \\ &\leq y^{-d} \cdot \frac{d\mu}{d\lambda}(x)(x^d - y^d) \\ &\leq y^{-d} \cdot a_1 \cdot 2^{-d+1+t_1} \cdot d|x-y| \end{aligned}$$

and likewise applying the last three inequalities in (4.28) with $f = \frac{d\mu}{d\lambda}$ yields for $x, y \in (\frac{1}{2}, 1]$ with $x \geq y$,

$$\begin{aligned} 0 \leq \frac{d\mu}{d\lambda}(y) - \frac{d\mu}{d\lambda}(x) &= \Big(y - \frac{1}{2}\Big)^{-d}\Big(\Big(y - \frac{1}{2}\Big)^d \frac{d\mu}{d\lambda}(y) - \Big(y - \frac{1}{2}\Big)^d \frac{d\mu}{d\lambda}(x)\Big) \\ &\leq \Big(y - \frac{1}{2}\Big)^{-d} \cdot \frac{d\mu}{d\lambda}(x)\Big(\Big(x - \frac{1}{2}\Big)^d - \Big(y - \frac{1}{2}\Big)^d\Big) \\ &\leq \Big(y - \frac{1}{2}\Big)^{-d} \cdot a_2 \cdot 2^{-d+1+t_2} \cdot d|x-y| \end{aligned}$$

Hence, $\frac{d\mu}{d\lambda}$ is locally Lipschitz on the intervals $(0, \frac{1}{2}]$ and $(\frac{1}{2}, 1]$. □

CHAPTER 4

We conclude this subsection with the proof of Corollary 4.1.3.

Proof of Corollary 4.1.3. For each $n \in \mathbb{N}$, let $\boldsymbol{p}_n = (p_{n,j})_{j \in \Sigma}$ be a strictly positive probability vector such that $\sup_n \sum_{r \in \Sigma_R} p_{n,r} K_r^{-\alpha_{\min}} < 1$ and assume that $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ in $\mathbb{R}_+^N$. In order to conclude that $\frac{d\mu_{\boldsymbol{p}_n}}{d\lambda}$ converges in $L^1(\lambda)$ to $\frac{d\mu_{\boldsymbol{p}}}{d\lambda}$ we will show that each subsequence of $\{\frac{d\mu_{\boldsymbol{p}_n}}{d\lambda}\}$ has a further subsequence that converges in $L^1(\lambda)$ to $\frac{d\mu_{\boldsymbol{p}}}{d\lambda}$.

Let $\{\boldsymbol{q}_k\}$ be a subsequence of $\{\boldsymbol{p}_n\}$, and for convenience write $f_k = \frac{d\mu_{\boldsymbol{q}_k}}{d\lambda}$ for each $k \in \mathbb{N}$. First of all, observe that from $\sup_n \sum_{r \in \Sigma_R} p_{n,r} K_r^{-\alpha_{\min}} < 1$ and $\lim_{n\to\infty} \boldsymbol{p}_n = \boldsymbol{p}$ it follows from the proof of Lemma 4.2.4 that there exist sufficiently large $a_1, a_2 > 0$ and $\beta \in (\alpha_{\min}, \gamma)$ sufficiently close to $\alpha_{\min}$ such that $\mathcal{C}_2 = \mathcal{C}_2(a_1, a_2, \beta)$ from Lemma 4.2.4 contains the sequence $\{f_k\}$. Hence, it follows from Lemma 4.2.5 that $\{f_k\}$ has a subsequence $\{f_{k_m}\}$ that converges with respect to the $L^1(\lambda)$-norm to some $\tilde{f} \in \mathcal{C}_2$. We have

$$\begin{aligned} \|P_{F,\boldsymbol{p}}\tilde{f} - \tilde{f}\|_1 &\leq \|P_{F,\boldsymbol{p}}\tilde{f} - P_{F,\boldsymbol{q}_{k_m}}\tilde{f}\|_1 + \|P_{F,\boldsymbol{q}_{k_m}}\tilde{f} - f_{k_m}\|_1 + \|f_{k_m} - \tilde{f}\|_1 \\ &\leq \sum_{j\in\Sigma} |p_j - q_{k_m,j}| \cdot \|P_{T_j}\tilde{f}\|_1 + \|P_{F,\boldsymbol{q}_{k_m}}\tilde{f} - P_{F,\boldsymbol{q}_{k_m}} f_{k_m}\|_1 + \|f_{k_m} - \tilde{f}\|_1 \\ &\leq \sum_{j\in\Sigma} |p_j - q_{k_m,j}| \cdot \|\tilde{f}\|_1 + 2\|f_{k_m} - \tilde{f}\|_1. \end{aligned}$$

Since we have $\lim_{m\to\infty} \boldsymbol{q}_{k_m} = \boldsymbol{p}$ in $\mathbb{R}_+^N$ and $\lim_{m\to\infty} \|f_{k_m} - \tilde{f}\|_1 = 0$ we obtain that $P_{F,\boldsymbol{p}}\tilde{f}(x) = \tilde{f}(x)$ holds for λ-a.e. $x \in [0,1]$. It follows from Theorem 4.1.2 that $(\mathcal{T}, \boldsymbol{p})$ admits only one acs probability measure, so we conclude that $\tilde{f} = \frac{d\mu_{\boldsymbol{p}}}{d\lambda}$ holds λ-a.e. Hence, $\{f_{k_m}\}$ converges in $L^1(\lambda)$ to $\frac{d\mu_{\boldsymbol{p}}}{d\lambda}$. □

§4.3 Final remarks

Suppose $\eta > 1$. Then Theorem 4.1.1 says that no acs probability measure exists. We claim that in fact no physical measure for F can exist in this case but that an infinite acs measure exists with a density that has full support on $[0,1]$. Indeed, for the case that $\alpha_j \le 1$ holds for each $j \in \Sigma$, the existence of a σ-finite acs measure can, regardless of the value of η, be proven by applying the same steps of the inducing technique as in Section 2.2 for the inducing domain $Y = \bigcup_{j\in\Sigma}[j] \times (x_2(j), \frac{1}{2})$ with $x_2(j)$ as in Subsection 4.2.1 and using that in this case the maps T_j have non-positive Schwarzian derivative. If $\eta > 1$, this acs measure then must be infinite by Theorem 4.1.1. Furthermore, it is clear that the corresponding density must have full support because Y is a sweep-out set. Since for an i.i.d. random LSV map the dynamics are dominated by the LSV map with the fastest relaxation rate, the existence of such an infinite acs measure can therefore also be expected under the conditions $\eta > 1$ and $\alpha_{\min} < 1$ without assuming $\alpha_j \le 1$ for each $j \in \Sigma$. Similar as in Chapters 2 and 3, Aaronson's Ergodic Theorem [A97, Theorem 2.4.2] applied to this infinite acs measure then yields that no physical measure for F exists if $\eta > 1$.

It does not become clear from the results of Theorems 4.1.1 and 4.1.2 if an acs probability measure exists if $\eta = 1$. The proof of Theorem 4.1.1 does not work for $\eta = 1$ because in this case there exists no $\delta > 0$ such that ζ from (4.10) is at least 1, and the proof of Theorem 4.1.2 fails for $\eta = 1$ because in this case we have $\gamma = \alpha_{\min}$ and therefore the set $(\alpha_{\min}, \gamma) \cap (0,1]$ from which we pick β is empty. For each $\delta > 0$ the result and proof of Theorem 4.1.1 do however carry over if $\eta = 1$ and the maps T_r ($r \in \Sigma_R$) are slightly adapted such that on $(\frac{1}{2}, \frac{1}{2} + \delta)$ they would be linear with derivative K_r. Indeed, in that case the bound in (4.11) can be replaced with $T_r(x) = \frac{1}{2} + K_r(x - \frac{1}{2})$ where $x \in (\frac{1}{2}, \frac{1}{2} + \delta)$. We therefore conjecture that if $\eta = 1$ then no acs probability measure exists and a possible approach is to work with a sharper bound on the term $K_r + 2(1 - K_r)(x - \frac{1}{2})$ in (4.11) that is not uniform in $x \in (\frac{1}{2}, \frac{1}{2} + \delta)$ as opposed to the upper bound M_r in (4.11).

The proof of Theorem 4.1.2 immediately carries over to the case that $\Sigma_R = \emptyset$ by taking $\beta = 1$, thus recovering the result from [Z18] that a random system generated by i.i.d. random compositions of finitely many LSV maps admits a unique absolutely continuous invariant probability measure if $\alpha_{\min} < 1$ with density as in (4.5) for some $a > 0$. To show that in case $\Sigma_R = \emptyset$ this density is decreasing and continuous on the whole interval $(0,1]$ similar arguments as in Subsection 4.2.2 can be used with the sets $\mathcal{C}_0$, $\mathcal{C}_1$ and $\mathcal{C}_2$ replaced by

$$\mathcal{K}_0 = \Big\{ f \in L^1(\lambda) : f \ge 0, f \text{ decreasing and continuous on } (0,1] \Big\},$$
$$\mathcal{K}_1 = \Big\{ f \in \mathcal{K}_0 : x \mapsto x^{\alpha_{\max}+1} f(x) \text{ increasing on } (0,1] \Big\},$$
$$\mathcal{K}_2 = \Big\{ f \in \mathcal{K}_1 : f(x) \le a x^{-\alpha_{\min}} \text{ on } (0,1], \int_0^1 f d\lambda = 1 \Big\} \quad \text{with } a > 0 \text{ large enough.}$$

This has been done in [Z18].

CHAPTER 4

It would be interesting to study further statistical properties of the random systems from Theorem 4.1.2. As discussed in Example 1.4.3 the annealed dynamics of random systems of LSV maps are dominated by the LSV map with the fastest relaxation rate, namely $S_{\alpha_{\min}}$, and annealed correlations decay as fast as $n^{1-1/\alpha_{\min}}$. This behaviour is significantly different from the behaviour of the random systems from Theorem 4.1.2 where the annealed dynamics are determined by the interplay between the exponentially fast attraction to $\frac{1}{2}$ and polynomially fast repulsion from zero. We conjecture that the random systems from Theorem 4.1.2 are mixing and that in case $\Sigma_R = \{1\}$ the decay of annealed correlations is at least polynomially fast with degree $1 - \frac{1}{\alpha_{\min}} \cdot \min\{\frac{\log p_1}{\log K_1}, 1\}$.

Finally, a natural question is whether the results of Theorems 4.1.1 and 4.1.2 can be extended to a more general class of one-dimensional random dynamical systems that exhibit this interplay between two fixed points, one to which orbits converge exponentially fast and one from which orbits diverge polynomially fast. First of all, if being C^1 and having $\frac{1}{2}$ as attracting fixed point are the only conditions we put on the right branches of the maps in $\mathfrak{R}$, then it can be shown in a similar way as in the proof of Theorem 4.1.1 that $(\mathcal{T}, \boldsymbol{p})$ admits no acs probability measure if

$$\sum_{r \in \Sigma_R} p_r \cdot \Big(\lim_{x \downarrow \frac{1}{2}} |DT_r(x)| \Big)^{-\alpha_{\min}} > 1$$

by applying Kac's Lemma. Secondly, we used in the proofs of Lemma 4.2.3 and Lemma 4.2.4 that $\frac{1}{DR_{\alpha_r,K_r}(z_r)} \big(\frac{x-\frac{1}{2}}{z_r - \frac{1}{2}} \big)^d$ is increasing and $\frac{1}{DR_{\alpha_r,K_r}(z_r)} \big(\frac{x-\frac{1}{2}}{z_r - \frac{1}{2}} \big)^{t_2}$ is decreasing, respectively, by means of the results on $H_{K,b}$ in Lemma 4.2.2. However, for other maps that have the property that $\frac{1}{2}$ and 1 are fixed points and that orbits are attracted to $\frac{1}{2}$ exponentially fast this is not true in general. Still a phase transition is to be expected, but different techniques are needed to prove this. This is also the case when we drop the condition that 1 is a fixed point of the maps in $\mathfrak{R}$, for instance by taking $R_{\alpha,K}(x) = \frac{1}{2} + K(x - \frac{1}{2})$ if $x \in (\frac{1}{2}, 1]$, in which case Lemma 4.2.3 would not hold. Thirdly, the results of Theorems 4.1.1 and 4.1.2 might carry over if we allow the left branches to only satisfy the conditions on the left branch of the maps $\{T_\alpha : [0,1] \to [0,1]\}_{\alpha \in (0,1)}$ considered in [M05] or Section 5 of [LSV99]. Each map T_α then satisfies $T_\alpha(0) = 0$ and $DT_\alpha(x) = 1 + Cx^\alpha + o(x^\alpha)$ for x close to zero and where $C > 0$ is some constant.

PART II

EXTENSIONS OF LOCHS' THEOREM TO RANDOM SYSTEMS

CHAPTER 5

A Lochs Theorem for random interval maps

This chapter is based on: [KVZ22].

Abstract

In 1964 Lochs proved a theorem on the number of continued fraction digits of a real number x that can be determined from just knowing its first n decimal digits. In 2001 this result was generalised to a dynamical systems setting by Dajani and Fieldsteel, where it compares sizes of cylinder sets for different transformations. In this chapter we prove a version of Lochs' Theorem for a broad class of random dynamical systems and under additional assumptions we prove a corresponding Central Limit Theorem as well. The main ingredient for the proof is an estimate on the asymptotic size of the cylinder sets of the random system in terms of the fiber entropy. To compute this entropy we provide a random version of Rokhlin's formula for entropy.

§5.1 Introduction

§5.1.1 Extension of Lochs' Theorem to number theoretic fibered maps

Real numbers can be represented in many different ways, e.g. by binary, decimal or continued fraction expansions, and one can wonder about the amount of information that each one of these expansions carries. In 1964 Lochs considered a specific question of this form: Given the first n decimal digits of a further unknown irrational number $x \in (0,1)$, what is the largest number $m = m(n,x)$ of regular continued fraction digits of x is that can be determined from this information. Lochs answered this question in [L64] for the limit $n \to \infty$ by showing that for Lebesgue almost every $x \in (0,1)$,

$$\lim_{n\to\infty} \frac{m(n,x)}{n} = \frac{6\log 2 \log 10}{\pi^2}. \tag{5.1}$$

Over the years Lochs' result has been refined and generalised in many directions. Let λ denote the Lebesgue measure on $[0,1)$. In [F98] Faivre established a Central Limit Theorem associated to Lochs' Theorem:

$$\lim_{n\to\infty} \lambda\Bigg(\Bigg\{x \in (0,1) : \frac{m(n,x) - n\frac{6\log 2\log 10}{\pi^2}}{\sigma\sqrt{n}} \le u\Bigg\}\Bigg) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt \tag{5.2}$$

holds for some constant $\sigma > 0$. See [F97, F01, W06, W08] for other results related to the limit in (5.1) and [LW08, BI08, FWL16, FWL19] for results where the decimal expansions in (5.1) are replaced by β-expansions.

In [BDK99] Bosma, Dajani and Kraaikamp highlighted that Lochs' Theorem can be seen as a dynamical statement. This viewpoint was further developed in [DF01], where Dajani and Fieldsteel gave the dynamical equivalent of the local limit statement from (5.1) for what they called *number theoretic fibered maps (NTFM)*. An NTFM is a triple (T,μ,α) where $T : [0,1) \to [0,1)$ is a surjective map, μ is a Borel measure on $[0,1)$ and $\alpha = \{A_j : j \in D\}$ is an at most countable interval partition of $[0,1)$ indexed by some set D, such that

(n1) $T|_{A_j}$ is continuous and strictly monotone for each $j \in D$;

(n2) μ is an ergodic invariant probability measure for T that is equivalent to λ with a density that is bounded and bounded away from zero;

(n3) the partition α generates the Borel σ-algebra $\mathcal{B}$ on $[0,1)$ in the sense that if for each n we use

$$\alpha_n = \bigvee_{k=0}^{n-1} T^{-k}\alpha = \{A_{j_1} \cap T^{-1}A_{j_2} \cap \cdots \cap T^{-(n-1)}A_{j_n} \,:\, A_{j_k} \in \alpha,\, 1 \le k \le n\} \tag{5.3}$$

to denote the *level n cylinders* of T, then the smallest σ-algebra containing all these sets for all $n \ge 1$, denoted by $\sigma\big(\bigcup_{n\in\mathbb{N}} \alpha_n\big)$, equals $\mathcal{B}$ up to sets of Lebesgue measure zero;

Chapter 5

(n4) the entropy $-\sum_{A\in\alpha}\mu(A)\log\mu(A)$ of α with respect to μ is finite.

The name NTFM refers to the fact that an NTFM generates for each $x \in [0,1)$ a digit sequence $(d_n^T(x))_{n\geq 1}$ with digits in D by setting

$$d_n^T(x) = j \quad \text{if } T^{n-1}x \in A_j,\ j \in D, \tag{5.4}$$

and for certain maps T these sequences correspond to well-known number expansions. The procedure from (5.4) can in words be described as follows: After assigning digit j to subinterval A_j for each $j \in D$, for each $x \in [0,1)$ the digit sequence from (5.4) is obtained by writing down in order the digits corresponding to the elements of α that the orbit of x under T visits.

Example 5.1.1 (N-adic transformations). Let $T_N : [0,1) \to [0,1)$ with integer $N \geq 2$ be the N-adic transformation from Example 1.3.1, λ the Lebesgue measure on $[0,1)$ and $\alpha = \{A_j : j \in \{0,1,\dots,N-1\}\}$ the partition given by

$$A_j = \Big[\frac{j}{N}, \frac{j+1}{N}\Big), \quad j \in \{0,1,\dots,N-1\}. \tag{5.5}$$

Then (T_N, λ, α) is an NTFM and digit sequences $(d_n^{T_N}(x))_{n\geq 1}$ in $\{0,1,\dots,N-1\}^{\mathbb{N}}$ can then be obtained by following the procedure from (5.4). This is illustrated in Figure 5.1 for the case that $N=2$. For each $x \in [0,1)$ this sequence $(d_n^{T_N}(x))_{n\geq 1}$ yields the *expansion in integer base N of x* given by

$$x = \sum_{n=1}^{\infty} \frac{d_n^{T_N}(x)}{N^n}. \tag{5.6}$$

Indeed, note that

$$T_N^n(x) = N \cdot T_N^{n-1}(x) - d_n^{T_N}(x)$$

holds for each $n \in \mathbb{N}$, so that recursively, this gives

$$x = \frac{d_1^{T_N}(x)}{N} + \frac{d_2^{T_N}(x)}{N^2} + \cdots + \frac{d_n^{T_N}(x)}{N^n} + \frac{T_N^n(x)}{N^n}.$$

Since $0 \leq T_N^n x \leq 1$, this yields (5.6) by taking $n \to \infty$.

Example 5.1.2 (Gauss map). Let $G : [0,1) \to [0,1)$ and μ_G be the Gauss map and Gauss probability measure from Example 1.3.2, respectively, and let $\alpha = \{A_j : j \in \mathbb{N}\}$ be the partition of $(0,1]$ given by

$$A_j = \Big(\frac{1}{j+1}, \frac{1}{j}\Big], \quad j \in \mathbb{N}.$$

The triple (G, μ_G, α) is an NTFM and for each irrational $x \in [0,1)$ the digit sequence $(d_n^G(x))_{n\geq 1}$ in $\mathbb{N}^{\mathbb{N}}$ generated as in (5.4) gives the *regular continued fraction expansion*

$$x = \cfrac{1}{d_1^G(x) + \cfrac{1}{d_2^G(x) + \cfrac{1}{d_3^G(x) + \ddots}}}.$$

See e.g. [DK21, Section 8.1] for a justification.

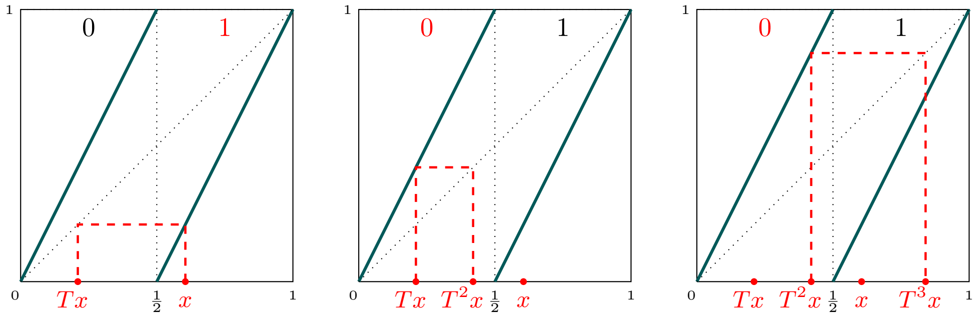


Figure 5.1: Illustration of the generation of digit sequences $(d_n^T(x))_{n\geq 1}$ given by (5.4) for the doubling map. The point x depicted satisfies $d_1^T(x) = 1$, $d_2^T(x) = 0$, $d_3^T(x) = 0$, etc.

Other examples of number expansions that can be obtained by an NTFM include β-expansions, various Lüroth-type expansions and various types of continued fraction expansions.

If for two NTFM's (T, μ, α) and $(S, \tilde{\mu}, \gamma)$ we define the number

$$m_{T,S}(n, x) = \sup\{m \in \mathbb{N} : \alpha_n(x) \subseteq \gamma_m(x)\},$$

where $\alpha_n(x)$ and $\gamma_m(x)$ denote the elements of the partitions α_n and γ_m as in (5.3) that contain x, respectively, then one can interpret $m_{T,S}(n, x)$ as the largest m so that the level m cylinder for S containing x can be determined from knowing only the level n cylinder for T that contains x. Equivalently, if we use $(d_k^T(x))_{k\geq 1}$ and $(d_k^S(x))_{k\geq 1}$ to denote the digit sequences produced by T and S, $m_{T,S}(n, x)$ is the largest m such that the digits $d_1^S(x), \ldots, d_m^S(x)$ can be determined from knowing $d_1^T(x), \ldots, d_n^T(x)$ of a further unknown $x \in [0, 1)$. The authors of [DF01] proved that for any two NTFM's (T, μ, α) and $(S, \tilde{\mu}, \gamma)$ with measure theoretic entropies $h_\mu(T), h_{\tilde{\mu}}(S) > 0$ it holds that

$$\lim_{n\to\infty} \frac{m_{T,S}(n, x)}{n} = \frac{h_\mu(T)}{h_{\tilde{\mu}}(S)} \qquad \lambda\text{-a.e.} \tag{5.7}$$

Lochs' original result given in (5.1) can be recovered by taking for T the map T_N from Example 5.1.1 with $N = 10$ and for S the Gauss map S from Example 5.1.2. Similar to the result in (5.2) by Faivre, Herczegh [H09] proved a Central Limit Theorem for the statement in (5.7) for a specific class of pairs of NTFM's. For two NTFM's (T, μ, α) and $(S, \tilde{\mu}, \gamma)$ that satisfy $h_\mu(T), h_{\tilde{\mu}}(S) > 0$ and several additional conditions, he proved in [H09, Corollary 2.1] that for each $u \in \mathbb{R}$,

$$\lim_{n\to\infty} \tilde{\mu}\left(\left\{x \in (0,1) : \frac{m_{T,S}(n, x) - n\frac{h_\mu(T)}{h_{\tilde{\mu}}(S)}}{\sigma\sqrt{n}} \leq u\right\}\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt \tag{5.8}$$

for some appropriate constant $\sigma > 0$.

§5.1.2 Number expansions generated by random interval maps

Random interval maps that generate number expansions have received increasing attention in the past few decades. This is partly due to the fact that often such systems generate for a typical real number not just one, but uncountably many different number expansions of a given type. We describe the general procedure to obtain such expansions.

Let $\mathcal{T} = \{T_i : [0,1) \to [0,1)\}_{i\in I}$ be a family of interval maps on $[0,1)$ and suppose for each $i \in I$ there is a partition α_i of $[0,1)$ into finitely or countably many subintervals. For each $i \in I$ we assign to each subinterval of α_i a digit and write D_i for the collection of these digits. We then write $A_{i,j}$ for the subinterval corresponding to the digit $j \in D_i$, so that for each $i \in I$ we can write $\alpha_i = \{A_{i,j} : j \in D_i\}$. We now define digit sequences by following random orbits and at each time step recording, before a map T_i is applied, the digit that corresponds to the partition element of α_i in which the random orbit at that moment is located. More precisely, for each $(\omega, x) \in I^{\mathbb{N}} \times [0,1)$ we define the digit sequence $(d_n^{\mathcal{T}}(\omega,x))_{n\geq 1}$ where $d_n^{\mathcal{T}}(\omega,x) \in D_{\omega_n}$ is the label of the partition element of α_{ω_n} in which $T_\omega^{n-1}(x)$ lies, i.e.

$$d_n^{\mathcal{T}}(\omega,x) = j \qquad \text{if } T_\omega^{n-1}(x) \in A_{\omega_n,j}, \quad j \in D_{\omega_n}. \tag{5.9}$$

We clarify this method with an example.

Example 5.1.3 (Random integer base transformations). Let $I = \{2,3\}$ and $\mathcal{T} = \{T_i\}_{i\in I}$ where $T_2, T_3 : [0,1] \to [0,1]$ are as in Example 5.1.1, i.e. $T_2(x) = 2x \bmod 1$ and $T_3(x) = 3x \bmod 1$. For the method described above, we associate to T_2 and T_3 the partitions $\alpha_2 = \{A_{2,0}, A_{2,1}\}$ and $\alpha_3 = \{A_{3,0}, A_{3,1}, A_{3,2}\}$ given by (5.5), meaning that before applying T_2 we assign digits 0 and 1 to $[0,\frac{1}{2})$ and $[\frac{1}{2},1)$, respectively, and before applying T_3 we assign digits 0,1 and 2 to $[0,\frac{1}{3})$, $[\frac{1}{3},\frac{2}{3})$ and $[\frac{2}{3},1)$, respectively. For each $(\omega,x) \in I^{\mathbb{N}} \times [0,1]$ we then obtain a digit sequence $(d_n^{\mathcal{T}}(\omega,x))_{n\geq 1}$ as given by (5.9). This is illustrated in Figure 5.2. For each $(\omega,x) \in I^{\mathbb{N}} \times [0,1]$ and $n \in \mathbb{N}$ we have

$$T_\omega^n(x) = \omega_n \cdot T_\omega^{n-1}(x) - d_n^{\mathcal{T}}(\omega,x),$$

so that similar as in Example 5.1.1 this iteratively yields

$$x = \frac{d_1^{\mathcal{T}}(\omega,x)}{\omega_1} + \frac{d_2^{\mathcal{T}}(\omega,x)}{\omega_1\omega_2} + \cdots + \frac{d_n^{\mathcal{T}}(\omega,x)}{\omega_1\cdots\omega_n} + \frac{T_\omega^n(x)}{\omega_1\cdots\omega_n}.$$

From $\lim_{n\to\infty} \left|\frac{T_\omega^n(x)}{\omega_1\cdots\omega_n}\right| \leq \lim_{n\to\infty} \frac{1}{2^n} = 0$, it follows that

$$x = \sum_{n=1}^{\infty} \frac{d_n^{\mathcal{T}}(\omega,x)}{2^{c_n(\omega)}3^{n-c_n(\omega)}},$$

where $c_n(\omega) = \#\{1 \leq k \leq n : \omega_k = 2\}$. Thus, the procedure from (5.9) with $\mathcal{T} = \{T_2, T_3\}$ generates number expansions in mixed integer base 2 and 3.

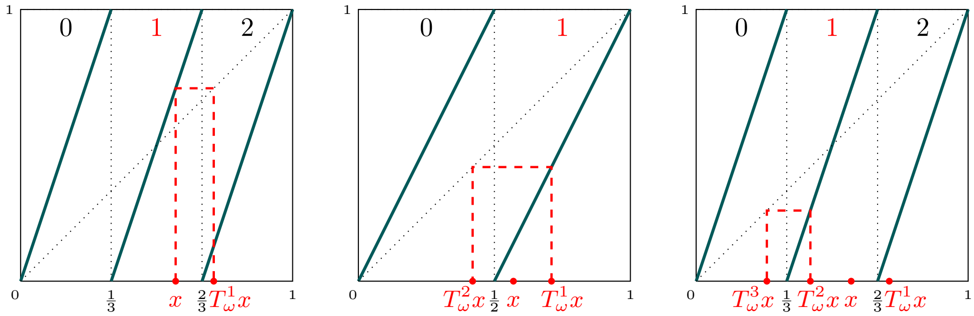


Figure 5.2: Illustration of the generation of digit sequences $(d_n^{\mathcal{T}}(\omega, x))_{n\geq 1}$ *given by* (5.9) *for the random map from Example 5.1.3 given by random compositions of* T_2 *and* T_3*. The point* x *depicted satisfies, for* $\omega = (3, 2, 3, \ldots)$, $d_1^{\mathcal{T}}(\omega, x) = 1$, $d_2^{\mathcal{T}}(\omega, x) = 1$, $d_3^{\mathcal{T}}(\omega, x) = 1$, *etc.*

There are various other random dynamical systems like the above example related to number expansions. The random β-transformation was first introduced in [DK03] and then further investigated in [DdV05, DdV07, DK07, DK13, K14, BD17, DJ17, S19, DKM21]. Interesting features of this system are its relation to Bernoulli convolutions, see e.g. [JSS11, DK13, K14], and to β-encoders, see Chapter 6 and e.g. [DDGV02, DGWY10, G12, KHTA12, JMKA13, MIS$^+$15, SJO15, JM16] and the references therein. A random system producing binary expansions was studied in [DK20], random dynamical systems related to continued fraction expansions appear in [KKV17, DO18, BRS20, AFGTV21, DKM21, KMTV22] and random Lüroth maps are considered in [KM22a, KM22b].

In this chapter we extend the version of Lochs' Theorem in (5.7) to the setting of random dynamical systems and give a corresponding Central Limit Theorem in the spirit of (5.8). The class of random dynamical systems we consider, which we call *random number systems*, contains the class of deterministic NTFM's and all of the random dynamical systems related to number expansions mentioned above. A random number system consists of a family of maps $\mathcal{T} = \{T_i : [0,1) \to [0,1)\}_{i\in I}$, where the index set I is a possibly uncountable Polish space, each map $T_i : [0,1) \to [0,1)$, $i \in I$, admits an appropriate partition $\alpha_i = \{A_{i,0}, A_{i,1}, \ldots\}$ of $[0,1)$ and there exists an appropriate probability measure μ on $I^{\mathbb{N}} \times [0,1)$. (The precise definition will be given in Section 5.2.) Thus a random number system is a quintuple $(I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$, where $\mathbb{P}$ is the probability law on $I^{\mathbb{N}}$ determining the random choices. For ease of notation we also write $\mathcal{T}$ for this quintuple.

Generating a digit sequence for x by means of (5.9) gives information about the location of x in the following way:

$$(d_1^{\mathcal{T}}(\omega, x), \ldots, d_n^{\mathcal{T}}(\omega, x)) = (j_1, \ldots, j_n) \quad \Rightarrow \quad x \in \bigcap_{k=1}^{n} T^{-1}_{\omega_1\cdots\omega_{k-1}} A_{\omega_k, j_k}. \tag{5.10}$$

As the right side shows, this information about the location of x provided by the digit sequence is under the assumption that $\omega = (\omega_1, \omega_2, \ldots) \in I^{\mathbb{N}}$ is known. This setting is interesting from a number theoretic point of view and will be investigated in this chapter, but for more practical applications where $\omega \in I^{\mathbb{N}}$ models e.g. some noise in the system, it is not very feasible to assume that ω is known or can be observed with full accuracy. In such cases the following model might be more realistic:

$$(j_1, \dots, j_n) \text{ observed for unknown } x \quad \Rightarrow \quad x \in \bigcup_{(\omega_1,\dots,\omega_n)\in I^n} \bigcap_{k=1}^{n} T^{-1}_{\omega_1\cdots\omega_{k-1}} A_{\omega_k, j_k}, \tag{5.11}$$

where $A_{i,j} = \emptyset$ if $j \notin D_i$. We consider the setting of (5.11) in Chapter 6 for a specific random interval map that models the A/D conversion in a β-encoder and where fluctuations in the system are due to noise and thus unknown, see footnote 1 on page 164. Finally, note that the models (5.10) and (5.11) coincide if the corresponding digit sets $\{D_i\}_{i\in I}$ are pairwise disjoint.

§5.1.3 Main results

For a random number system $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ we define analogous to (5.3) the *random level n cylinders* as follows: For each $\omega \in I^{\mathbb{N}}$ and $n \in \mathbb{N}$ we define the partition

$$\begin{aligned}\alpha_{\omega,n} &= \bigvee_{k=0}^{n-1} (T^k_\omega)^{-1} \alpha_{\omega_{k+1}} \\ &= \{A_{\omega_1,j_1} \cap T^{-1}_{\omega_1} A_{\omega_2,j_2} \cap \cdots \cap T^{-1}_{\omega_1\cdots\omega_{n-1}} A_{\omega_n,j_n} \,:\, A_{\omega_k,j_k} \in \alpha_{\omega_k},\ 1 \le k \le n\}.\end{aligned} \tag{5.12}$$

Furthermore, we write $\alpha_{\omega,n}(x)$ for the random (n,ω)-cylinder that contains x. Given two random number systems $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ and $\mathcal{S} = (J, \mathbb{Q}, \{S_j\}_{j\in J}, \rho, \{\gamma_j\}_{j\in J})$, for each $n \in \mathbb{N}$, $\omega \in I^{\mathbb{N}}$, $\tilde{\omega} \in J^{\mathbb{N}}$ and $x \in [0,1)$, let

$$m_{\mathcal{T},\mathcal{S}}(n, \omega, \tilde{\omega}, x) = \sup\{m \in \mathbb{N} : \alpha_{\omega,n}(x) \subseteq \gamma_{\tilde{\omega},m}(x)\}. \tag{5.13}$$

This quantity can be interpreted as follows: For given $\omega \in I^{\mathbb{N}}$ and $\tilde{\omega} \in J^{\mathbb{N}}$, $m_{\mathcal{T},\mathcal{S}}(n, \omega, \tilde{\omega}, x)$ is the largest level m for which we can determine the random $(m,\tilde{\omega})$-cylinder for $\mathcal{S}$ containing x from knowing only the random (n,ω)-cylinder for $\mathcal{T}$ that contains x. Alternatively, it is the largest m such that $d^{\mathcal{S}}_1(\tilde{\omega}, x), \dots, d^{\mathcal{S}}_m(\tilde{\omega}, x)$ can be determined from knowing the digits $d^{\mathcal{T}}_1(\omega, x), \dots, d^{\mathcal{T}}_n(\omega, x)$ of a further unknown $x \in [0,1)$. In this chapter we obtain the following Random Lochs' Theorem, where the measure theoretic entropy from (5.7) is replaced by *fiber entropy*, which for a random number system $\mathcal{T}$ is a quantity $h^{\text{fib}}(\mathcal{T}) \in [0,\infty)$ and will be defined in Section 5.5.

Theorem 5.1.4. *Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$, $\mathcal{S} = (J, \mathbb{Q}, \{S_j\}_{j\in J}, \rho, \{\gamma_j\}_{j\in J})$ be two random number systems. If $h^{\text{fib}}(\mathcal{T}), h^{\text{fib}}(\mathcal{S}) > 0$, then*

$$\lim_{n\to\infty} \frac{m_{\mathcal{T},\mathcal{S}}(n, \omega, \tilde{\omega}, x)}{n} = \frac{h^{\text{fib}}(\mathcal{T})}{h^{\text{fib}}(\mathcal{S})} \qquad \lambda\text{-a.e.}$$

for $\mathbb{P} \times \mathbb{Q}$-a.a. $(\omega, \tilde{\omega}) \in I^{\mathbb{N}} \times J^{\mathbb{N}}$.

We like to make two remarks about this result. Firstly, the quotient of measure theoretic entropies that appears as the value of the limit in the deterministic setting has been replaced by a quotient of fiber entropies in the random setting. Secondly, the setup allows for the index set I of the family $\{T_i\}_{i\in I}$ to be uncountable, so that the results apply to e.g. random β-transformations where the value of β can range over a whole interval, see Example 5.7.5 below. This makes the proofs more involved.

In [DF01] an essential ingredient to prove (5.7) is the following general result on interval partitions. If $\mathcal{P} = \{P_n\}_{n=1}^{\infty}$ is a sequence of interval partitions and $c \geq 0$, we say that $\mathcal{P}$ *has entropy* c λ*-a.e.* if

$$\lim_{n\to\infty} -\frac{\log \lambda\big(P_n(x)\big)}{n} = c \qquad \lambda\text{-a.e.},$$

where $P_n(x)$ denotes the element of the partition P_n containing x.

Theorem 5.1.5 (Theorem 4 of [DF01]). *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*

$$m_{\mathcal{P},\mathcal{Q}}(n,x) = \sup\{m \in \mathbb{N} : P_n(x) \subseteq Q_m(x)\}.$$

Suppose that $\mathcal{P}$ has entropy $c \in (0,\infty)$ λ-a.e. and $\mathcal{Q}$ has entropy $d \in (0,\infty)$ λ-a.e. Then

$$\lim_{n\to\infty} \frac{m_{\mathcal{P},\mathcal{Q}}(n,x)}{n} = \frac{c}{d} \qquad \lambda\text{-a.e.}$$

The proof of (5.7) goes roughly along the following lines. An application of the Kolmogorov-Sinai Theorem and of the Shannon-McMillan-Breiman Theorem to the NTFM's T and S provides the appropriate asymptotics for the size of the cylinder sets from (5.3) for both maps T and S to establish the positive entropy conditions and then Theorem 5.1.5 completes the proof. To achieve Theorem 5.1.4 we also employ Theorem 5.1.5 and therefore the main achievement here is obtaining the right asymptotics for the size of the random cylinder sets from (5.12). More precisely Theorem 5.1.4 will appear as a corollary of the following theorem.

Theorem 5.1.6. *Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system. The following hold,*

(i) For $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$ we have

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\alpha_{\omega,n}(x)\big)}{n} = h^{\text{fib}}(\mathcal{T}), \qquad \lambda\text{-a.e.}$$

(ii) Let ν denote the marginal of μ on $I^{\mathbb{N}}$. Furthermore, let F be the skew product on $I^{\mathbb{N}} \times [0,1)$ given by $F(\omega,x) = (\tau\omega, T_{\omega_1}(x))$, where τ denotes the left shift on $I^{\mathbb{N}}$. If $h_\nu(\tau) < \infty$, then

$$h^{\text{fib}}(\mathcal{T}) = h_\mu(F) - h_\nu(\tau).$$

(iii) If for each $i \in I$ and $A \in \alpha_i$ the restriction $T_i|_A$ is differentiable, then

$$h^{\text{fib}}(\mathcal{T}) = \int_{I^{\mathbb{N}} \times [0,1)} \log |DT_{\omega_1}(x)| \, d\mu(\omega, x).$$

The first part of this theorem gives the required estimates for the asymptotic sizes of the cylinder sets from (5.12) and, when combined with Theorem 5.1.5, leads to Theorem 5.1.4. The limit from Theorem 5.1.4 is expressed in terms of the fiber entropies of the two random number systems. Parts (ii) and (iii) of Theorem 5.1.6 give different ways to determine this limit. The second part works in case the entropy of the marginal of μ on $I^{\mathbb{N}}$ is finite. The third part gives a random version of Rokhlin's Formula for entropy.

We also prove a Central Limit Theorem for Theorem 5.1.4 in case we compare the digits obtained from a random number system $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i \in I}, \mu, \{\alpha_i\}_{i \in I})$ to those from an NTFM $(S, \tilde{\mu}, \gamma)$ under additional assumptions on both systems. To be more specific, for such systems we obtain that for all $u \in \mathbb{R}$,

$$\lim_{n \to \infty} \mu\Big(\Big\{(\omega, x) \in I^{\mathbb{N}} \times [0,1) : \frac{m_{\mathcal{T},S}(n, \omega, x) - n\frac{h^{\text{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)}}{\kappa\sqrt{n}} \leq u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2} dt, \tag{5.14}$$

for an appropriate constant $\kappa > 0$.

The remainder of this chapter is organised as follows. In Section 5.2 we give a precise definition of random number systems. We then consider in Section 5.3 a special but wide class of random number systems for which the proof of Theorem 5.1.6 is relatively easy. In Sections 5.4 and 5.5 we provide some preliminaries for the proof of the general case and give a precise definition of fiber entropy. We prove Theorem 5.1.6 in Section 5.6 and obtain Theorem 5.1.4 as a corollary and we prove the Central Limit Theorem from (5.14). In Section 5.7 we provide some examples.

§5.2 Random number systems

In this section we define the dynamical systems that we are interested in. Let $(I^{\mathbb{N}}, \mathcal{B}_I^{\mathbb{N}}, \mathbb{P})$ be a base space where I is a Polish space with associated Borel σ-algebra $\mathcal{B}_I$ and where $\mathbb{P}$ is a Borel probability measure on the product σ-algebra $\mathcal{B}_I^{\mathbb{N}}$ such that the left shift τ on $I^{\mathbb{N}}$ is non-singular with respect to $\mathbb{P}$, i.e. $\mathbb{P}(\tau^{-1}A) = 0$ if and only if $\mathbb{P}(A) = 0$ for all $A \in \mathcal{B}_I^{\mathbb{N}}$. For each $i \in I$, let $T_i : [0,1) \to [0,1)$ be a Borel measurable transformation. Let $\mathcal{B}$ denote the Borel σ-algebra on $[0,1)$ and λ the Lebesgue measure on $[0,1)$. Associated to the family $\{T_i : [0,1) \to [0,1)\}_{i \in I}$ let F be the skew product transformation

$$F : I^{\mathbb{N}} \times [0,1) \to I^{\mathbb{N}} \times [0,1), \; (\omega, x) \mapsto \big(\tau\omega, T_{\omega_1}(x)\big).$$

Let μ be an invariant probability measure for F on $I^{\mathbb{N}} \times [0,1)$. For each $i \in I$ let $\alpha_i = \{A_{i,0}, A_{i,1}, \ldots\}$ be a partition of $[0,1)$ by countably many subintervals of

$[0,1)$, possibly containing empty sets.[1] A *random number system* is a collection $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ on $[0,1)$ that satisfies the following conditions.

(r1) The map $I \times [0,1) \ni (i,x) \mapsto T_i(x) \in [0,1)$ is measurable.

(r2) For each $i \in I$ and $A \in \alpha_i$, $T_i|_A$ is strictly monotone and continuous.

(r3) The partition $\Delta = \{\Delta(j) \,:\, j \geq 0\}$ of $I^{\mathbb{N}} \times [0,1)$ given by

$$\Delta(j) = \{(\omega, x) \,:\, x \in A_{\omega_1,j}\} = \bigcup_{i\in I} [i] \times A_{i,j} \quad \text{for each } j \geq 0 \tag{5.15}$$

is measurable, i.e. $\Delta(j) \in \mathcal{B}_I^{\mathbb{N}} \times \mathcal{B}$ for all $j \geq 0$.

(r4) For $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$ we have that, for all $B \in \mathcal{B}$, $\lambda\big(T_{\omega_1}^{-1} B\big) = 0$ if $\lambda(B) = 0$.

(r5) For $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$ we have $\sigma\big(\bigcup_{n\in\mathbb{N}} \alpha_{\omega,n}\big) = \mathcal{B}$ up to sets of λ-measure zero.

(r6) The F-invariant measure μ is ergodic and equivalent to $\mathbb{P} \times \lambda$.

(r7) The entropy of Δ w.r.t. μ, i.e. $H_\mu(\Delta) = -\sum_{j\geq 0} \mu(\Delta(j)) \log \mu(\Delta(j))$, is finite.

Most of the conditions (r1)–(r7) are easily verified in specific applications and not very restrictive. We give some comments on them.

- Conditions (r1) and (r3) are typical measurability conditions and are immediate in case I is at most countable (and equipped with the discrete topology). It easily follows from (r1) that the skew product F is measurable.

- Condition (r2) is needed to get digit sequences $(d_n^{\mathcal{T}}(\omega, x))_{n\geq 1}$ as in (5.9). It follows from (r5) that, for $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$, knowing $(d_n^{\mathcal{T}}(\omega, x))_{n\geq 1}$ determines $x \in [0,1)$ uniquely λ-a.e.

- Condition (r4) is a form of fiberwise non-singularity and from (r6) it follows that μ is the only probability measure that is both F-invariant and absolutely continuous w.r.t. $\mathbb{P} \times \lambda$ as can be seen from Theorem 1.2.6. In case I is countable, then it is easy to verify that (r4) already follows from only assuming (r6).

- If we let $\pi_I : I^{\mathbb{N}} \times [0,1) \to I^{\mathbb{N}}$ be the canonical projection onto the first coordinate and write $\nu = \mu \circ \pi_I^{-1}$ for the marginal of the invariant measure μ on $I^{\mathbb{N}}$, then from (r6) it follows that ν is τ-invariant, ergodic and equivalent to $\mathbb{P}$. In particular, again by Theorem 1.2.6, if $\mathbb{P}$ is τ-invariant, then $\nu = \mathbb{P}$.

- Condition (r7) guarantees that the fiber entropy defined later on is well defined. Note that if Δ is a finite set (that is, if Δ contains a finite number of non-empty elements), then (r7) is automatically satisfied.

We now present some classes of systems for which the assumptions from the definition of random number system are satisfied.

[1] For notational convenience we take $\{0,1,2,\ldots\}$ as digit set for each T_i. If for a map T_i the digit set would naturally be a finite set, then we take for α_i a collection that contains countably many empty sets.

- A class of maps that satisfy (r2) and (r4) and that are well studied is the class of *Lasota-Yorke type maps.* A Lasota-Yorke type map is a map $T : [0,1) \to [0,1)$ that is piecewise monotone C^2 and non-singular with $|DT(x)| > 0$ for all x where the derivative is defined. In that case an obvious candidate for the partition from (r2) is the partition of $[0,1)$ given by the maximal intervals on which T is monotone.

- Given a family $\{T_i\}_{i\in I}$ of Lasota-Yorke type maps for some appropriate index set I, a sufficient condition for Δ to be a generator in the sense of (r5) is that $\inf_{(i,x)} |DT_i(x)| > 1$. In case I is finite, this is equivalent to the condition that each T_i is expanding. We can allow for neutral fixed points as well and still get (r5) if we assume that the branches of the maps are full, i.e. map onto the whole interval $(0,1)$, and expanding outside each neighborhood of the neutral fixed point. Examples include the random Gauss-Rényi map from Example 1.4.2 that we again will encounter in Example 5.7.4 below and random Manneville-Pomeau maps.

- There exist various sets of conditions under which the existence of an invariant measure μ for the skew product F that satisfies (r6) is guaranteed. See Section 1.4 and the references mentioned there for some results in this direction for the case that $\mathbb{P}$ is a Bernoulli measure.

- The results from [KM22a] give an algorithm for determining explicit formulae for invariant probability measures of the form $\mathbb{P} \times \rho$ with $\rho \ll \lambda$ in case all maps T_i are piecewise linear Lasota-Yorke type maps satisfying some further conditions. Having an explicit formula facilitates the computation of the entropy of Δ and the verification of (r7).

Remark 5.2.1. If I consists of only one element, then the random number system reduces to an interval map. In this case, conditions (r2), (r5) and (r7) are equivalent with assuming that (n1), (n3) and (n4) hold for this interval map, respectively. Moreover, it follows from (r6) that the interval map is onto $[0,1)$ up to some λ-measure zero set and it follows from (r6) that this interval map satisfies (n2) except that the density does not necessarily have to be bounded and bounded away from zero. Thus in particular, each NTFM is an example of a random number system where the index set consists of only one element. Furthermore, note that in case $I = \{1\}$ contains only one element, then $h_\nu(\tau) = 0$ and Theorem 5.1.6(ii) gives that $h^{\text{fib}}(\mathcal{T}) = h_\rho(T_1)$, where ρ is the ergodic invariant probability measure for T_1 equivalent to λ. Hence, Theorem 5.1.4 is an extension of the result in (5.7) from [DF01] and shows that (5.7) remains true for two NTFM's for which the condition in (n2) on the bounds on the density is dropped.

§5.3 A special case

Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system. In this section we put the following additional four conditions on $\mathcal{T}$:

(s1) The index set I is countable.

(s2) Writing $\iota = \{[i] : i \in I\}$ for the countable partition of $I^{\mathbb{N}}$ given by the 1-cylinders, we assume that the entropy of ι w.r.t. the marginal ν of μ on $I^{\mathbb{N}}$ is finite, i.e. $H_\nu(\iota) < \infty$.

(s3) For all $\omega \in I^{\mathbb{N}}$ we have $\sigma\big(\bigcup_{n\in\mathbb{N}} \alpha_{\omega,n}\big) = \mathcal{B}$.

(s4) The density $\frac{d\mu}{d\nu\times\lambda}$ is bounded and bounded away from zero.

Note that (s3) is a slight strengthening of (r5). A class of random interval maps can be derived from [P84, I12, M85a, M85b, P84] that satisfy the conditions of a random number system as well as the additional conditions (s1)-(s4). These random interval maps are expanding on average and are composed of Lasota-Yorke type maps such that the corresponding invariant density is bounded, and admit a suitable covering property such that the density is bounded away from zero, thus satisfying (s4). Moreover, the base space then consists of a finite or countable index set I equipped with a Bernoulli measure or Markov measure.

Let $\mathcal{T}$ be as above satisfying (s1)-(s4). We will prove Theorem 5.1.6 for $\mathcal{T}$. It is clear that ι is a generator for the left shift τ on $I^{\mathbb{N}}$ w.r.t. ν. Furthermore, we define $\bar{\iota} = \{[i] \times [0,1) : i \in I\}$ being the partition ι embedded into $I^{\mathbb{N}} \times [0,1)$, and we define the countable partition ξ of $I^{\mathbb{N}} \times [0,1)$ given by

$$\xi = \big\{[i] \times A_{i,j} : i \in I, j \in \mathbb{N}_0\big\},$$

where as before $A_{i,j}$ are the partition elements of α_i. Note that ξ is the common refinement of $\bar{\iota}$ and the partition Δ given by (5.15), i.e. $\xi = \bar{\iota} \vee \Delta$.

Lemma 5.3.1. *The partition ξ is a generator for F.*

Proof. We write $\xi_n = \bigvee_{k=0}^{n-1} F^{-k}\xi$ and $\iota_n = \bigvee_{k=0}^{n-1} \tau^{-k}\iota$ for each $n \in \mathbb{N}$. Then for each $(\omega, x) \in I^{\mathbb{N}} \times [0,1)$ and $n \in \mathbb{N}$ we have

$$\xi_n(\omega, x) = [\omega_1 \cdots \omega_n] \times \alpha_{\omega,n}(x) = \iota_n(\omega) \times \alpha_{\omega,n}(x), \tag{5.16}$$

where $\xi_n(\omega, x)$ denotes the partition element of ξ_n containing x, and a similar meaning for $\iota_n(\omega)$ and $\alpha_{\omega,n}(x)$. Let $(\omega, x), (\tilde{\omega}, y) \in I^{\mathbb{N}} \times [0,1)$. If $\omega \neq \tilde{\omega}$, then there exists $n \in \mathbb{N}$ such that $\iota_n(\omega) \neq \iota_n(\tilde{\omega})$ and thus $\xi_n(\omega, x) \neq \xi_n(\tilde{\omega}, y)$. If $\omega = \tilde{\omega}$ and $x \neq y$, then according to (s3) there exists $n \in \mathbb{N}$ such that $\alpha_{\omega,n}(x) \neq \alpha_{\omega,n}(y)$ and thus $\xi_n(\omega, x) \neq \xi_n(\tilde{\omega}, y)$. Hence $\{\xi_n\}$ separates points, so ξ is a generator for F. □

It follows from (s2) that $h_\nu(\tau) < \infty$. Furthermore, we obtain from (s2) and (r7) that

$$H_\mu(\xi) \leq H_\mu(\bar{\iota}) + H_\mu(\Delta) = H_\nu(\iota) + H_\mu(\Delta) < \infty,$$

so that by Lemma 5.3.1 we have $h_\mu(F) < \infty$. We have the following two results:

Proposition 5.3.2. *For $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$ we have*

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\alpha_{\omega,n}(x)\big)}{n} = h_\mu(F) - h_\nu(\tau), \qquad \lambda\text{-a.e.}$$

Proof. Combining Lemma 5.3.1 and $H_\mu(\xi) < \infty$ with the Kolmogorov-Sinai Theorem (Theorem 1.2.17) and the Shannon-McMillan-Breiman Theorem (Theorem 1.2.22) yields that

$$\lim_{n\to\infty} -\frac{\log \mu\big(\xi_n(\omega,x)\big)}{n} = h_\mu(F), \qquad \mu\text{-a.e.} \tag{5.17}$$

By (r6) we can replace in (5.17) μ-a.e. with $\mathbb{P}\times\lambda$-a.e. and by (s4) we can replace $\mu\big(\xi_n(\omega,x)\big)$ with $\nu\times\lambda\big(\xi_n(\omega,x)\big)$, i.e.

$$\lim_{n\to\infty} -\frac{\log \nu\times\lambda\big(\xi_n(\omega,x)\big)}{n} = h_\mu(F), \qquad \mathbb{P}\times\lambda\text{-a.e.} \tag{5.18}$$

Also, we obtain from the Kolmogorov-Sinai Theorem and the Shannon-McMillan-Breiman Theorem together with the equivalence between ν and $\mathbb{P}$ that

$$\lim_{n\to\infty} -\frac{\log \nu\big(\iota_n(\omega)\big)}{n} = h_\nu(\tau), \qquad \mathbb{P}\text{-a.e.} \tag{5.19}$$

Combining (5.16), (5.18) and (5.19) now yields the result. □

Proposition 5.3.3. *If for each $i \in I$ and $A \in \alpha_i$ the restriction $T_i|_A$ is differentiable, then*

$$h_\mu(F) = \int_{I^{\mathbb{N}}\times[0,1)} \log|DT_{\omega_1}(x)|d\mu(\omega,x) + h_\nu(\tau).$$

Proof. Note that the partition ι of $I^{\mathbb{N}}$ consists of invertibility domains of τ. It follows from the Rokhlin Formula (Theorem 1.2.21) that

$$h_\nu(\tau) = \int_{I^{\mathbb{N}}} \log J_\nu\tau d\nu,$$

where $J_\nu\tau$ is the Jacobian of τ w.r.t. ν. Furthermore, note that ξ is a partition of $I^{\mathbb{N}}\times[0,1)$ by invertibility domains of F. For each $i \in I$, $j \in \mathbb{N}_0$, $C \in \mathcal{B}_I^{\mathbb{N}} \cap [i]$ and $D \in \mathcal{B} \cap A_{i,j}^T$ we have

$$\begin{aligned}\nu\times\lambda(F(C\times D)) &= \nu\times\lambda(\tau(C)\times T_i(D)) \\ &= \int_{C\times D} J_\nu\tau(\omega)|DT_{\omega_1}(x)|d\nu\times\lambda(\omega,x).\end{aligned}$$

Using standard arguments we can show from this that for each $A \in \xi$ and each measurable $B \subseteq A$ we have

$$\nu\times\lambda(F(B)) = \int_B J_\nu\tau(\omega)|DT_{\omega_1}(x)|d\nu\times\lambda(\omega,x),$$

so the Jacobian $J_{\nu\times\lambda}F$ of F w.r.t. $\nu\times\lambda$ exists and is given by

$$J_{\nu\times\lambda}F = J_\nu\tau(\omega)|DT_{\omega_1}(x)|, \qquad \text{for } \nu\times\lambda\text{-a.e. } (\omega,x) \in I^{\mathbb{N}}\times[0,1).$$

Furthermore, by the change of variables formula from Lemma 1.2.20(a) this gives for each $A \in \xi$ and each measurable $B \subseteq A$ that

$$\begin{aligned}\mu(F(B)) &= \int_{F(B)} \frac{d\mu}{d\nu \times \lambda} d\nu \times \lambda \\ &= \int_B \Big(\frac{d\mu}{d\nu \times \lambda} \circ F\Big) J_{\nu\times\lambda} F \frac{d\nu \times \lambda}{d\mu} d\mu,\end{aligned}$$

so the Jacobian $J_\mu F$ of F w.r.t. μ exists and is given by

$$J_\mu F = \big(\frac{d\mu}{d\nu \times \lambda} \circ F\big) J_{\nu\times\lambda} F \frac{d\nu \times \lambda}{d\mu}, \qquad \mu\text{-a.e.}$$

Using Lemma 5.3.1 and $H_\mu(\xi) < \infty$, it follows from the Rokhlin Formula that

$$h_\mu(F) = \int_{I^{\mathbb{N}} \times [0,1)} \log J_\mu F(\omega, x) d\mu(\omega, x).$$

We conclude that

$$\begin{aligned} h_\mu(F) &= \int \log\Big(\frac{d\mu}{d\nu \times \lambda} \circ F\Big) d\mu + \int \log\big(J_{\nu\times\lambda} F\big) d\mu + \int \log\Big(\frac{d\nu \times \lambda}{d\mu}\Big) d\mu \\ &= \int \log\Big(\frac{d\mu}{d\nu \times \lambda}\frac{d\nu \times \lambda}{d\mu}\Big) d\mu + \int \log |DT_{\omega_1}(x)| d\mu(\omega, x) \\ &\quad + \int \log\big(J_\nu \tau(\omega)\big) d\mu(\omega, x) \\ &= \int \log |DT_{\omega_1}(x)| d\mu(\omega, x) + h_\nu(\tau). \end{aligned}$$

So the above two propositions prove Theorem 5.1.6 in the special case that $\mathcal{T}$ satisfies (s1)-(s4).

Remark 5.3.4. We used condition (s4) for replacing $\mu\big(\xi_n(\omega, x)\big)$ in (5.17) with $\nu \times \lambda\big(\xi_n(\omega, x)\big)$. For this purpose, instead of assuming (s4), it is also sufficient to assume that $\nu \times \lambda$-a.a. $(\omega, x) \in I^{\mathbb{N}} \times [0, 1)$ is a *density point* for $\frac{d\mu}{d\nu\times\lambda}$ w.r.t. $\nu \times \lambda$, i.e.

$$\lim_{r \downarrow 0} \frac{1}{\nu \times \lambda(B((\omega, x), r))} \int_{B((\omega,x),r)} \frac{d\mu}{d\nu \times \lambda} d\nu \times \lambda = \frac{d\mu}{d\nu \times \lambda}(\omega, x), \tag{5.20}$$

where $B((\omega, x), r)$ denotes the open ball centered at (ω, x) with radius $r > 0$ with respect to a compatible metric, e.g. the one of the form as in (3.3). Then combined with Lemma 5.3.1 it follows that for $\mathbb{P} \times \lambda$-a.a. (ω, x) and every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for each integer $n \geq N$ we have

$$\frac{d\mu}{d\nu \times \lambda}(\omega, x) - \varepsilon \leq \frac{\mu\big(\xi_n(\omega, x)\big)}{\nu \times \lambda\big(\xi_n(\omega, x)\big)} \leq \frac{d\mu}{d\nu \times \lambda}(\omega, x) + \varepsilon.$$

Using this and the fact that $\frac{d\mu}{d\nu\times\lambda}(\omega, x) \in (0, \infty)$ holds for $\mathbb{P}\times\lambda$-a.e. $(\omega, x) \in I^{\mathbb{N}} \times [0, 1)$ by (r6), we can indeed then replace $\mu\big(\xi_n(\omega, x)\big)$ in (5.17) with $\nu \times \lambda\big(\xi_n(\omega, x)\big)$. An example when (5.20) is satisfied $\nu \times \lambda$-a.e. is when

$$\limsup_{r\downarrow 0} \frac{\nu \times \lambda\big(B((\omega,x),2r)\big)}{\nu \times \lambda\big(B(x,r)\big)} < \infty, \qquad \nu \times \lambda\text{-a.e. } (\omega,x) \in I^{\mathbb{N}} \times [0,1),$$

see e.g. [HKST15, Section 3.4]. This is for instance the case if $\mathbb{P}$ is a Bernoulli measure or Markov measure (recall that $\nu = \mathbb{P}$ if $\mathbb{P}$ is τ-invariant).

Note that the reasoning in this section does not work for proving Theorem 5.1.6 in full generality, because the classical Kolmogorov-Sinai Theorem, Shannon-McMillan-Breiman Theorem and Rokhlin Formula as applied in this section require a partition that is finite or countable with finite entropy. To overcome this problem, we will apply instead a fiberwise version of those theorems. However, the fiberwise Shannon-McMillan-Breiman Theorem and Rokhlin Formula we will use require that the underlying base map of the skew product is invertible. So we first need to extend the dynamics on the base space so that it becomes invertible, which we will do in the next section.

§5.4 Invertible base maps and invariant measures

Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system. One of the consequences of (r1)–(r7) is that $\mathcal{T}$ admits a *system of conditional measures*, i.e. a family of measures $\{\mu_\omega\}_{\omega\in I^{\mathbb{N}}}$ such that

- μ_ω is a probability measure on $([0,1),\mathcal{B})$ for ν-a.a. $\omega \in I^{\mathbb{N}}$,
- for any $f \in L^1(I^{\mathbb{N}} \times [0,1), \mu)$ the map $\omega \mapsto \int_{[0,1)} f(\omega,x)\,d\mu_\omega(x)$ on $I^{\mathbb{N}}$ is measurable and

$$\int_{I^{\mathbb{N}}\times[0,1)} f\,d\mu = \int_{I^{\mathbb{N}}} \Big(\int_{[0,1)} f(\omega,x)\,d\mu_\omega(x)\Big) d\nu(\omega). \tag{5.21}$$

Moreover, if $\{\tilde{\mu}_\omega\}_{\omega\in I^{\mathbb{N}}}$ is another system of conditional measures for μ, then $\mu_\omega = \tilde{\mu}_\omega$ for ν-a.a. $\omega \in I^{\mathbb{N}}$. (See [A97, Theorem 1.0.8] together with [VO16, Proposition 5.1.7] for a justification.)

The dynamics on the base space I of a random number system is given by the left shift τ on the set $I^{\mathbb{N}}$, which is not invertible. This setup corresponds to the setup for random systems associated to number expansions that is adopted in most of the references mentioned in the introduction. To prove Theorem 5.1.4, however, we employ known theory on random systems and fiber entropy that is available for skew products with invertible dynamics on the base space. One can easily extend the one-sided shift in the first coordinate of F to a two-sided (thus invertible) shift and as we shall see next, this has no profound effect on the invariant measures.

Let $\hat{\tau}$ denote the left shift on $I^{\mathbb{Z}}$ and extend the skew product F to a map $\hat{F}$ that is invertible in the first coordinate by setting

$$\hat{F} : I^{\mathbb{Z}} \times [0,1) \to I^{\mathbb{Z}} \times [0,1),\ (\hat{\omega},x) \mapsto (\hat{\tau}(\hat{\omega}), T_{\hat{\omega}_1}(x)).$$

Let $\mathcal{B}_I^{\mathbb{Z}}$ denote the Borel σ-algebra on $I^{\mathbb{Z}}$. Use $\pi : I^{\mathbb{Z}} \to I^{\mathbb{N}}$ to denote the canonical projection. To keep notation simple for two-sided sequences $\hat{\omega} \in I^{\mathbb{Z}}$ and $n \geq 0$ we use the same notation for $T_{\hat{\omega}}^n$, $\alpha_{\hat{\omega},n}$ and $\mu_{\hat{\omega}}$ as for one-sided sequences, i.e.

$$T_{\hat{\omega}}^n = T_{\pi(\hat{\omega})}^n, \quad \alpha_{\hat{\omega},n} = \alpha_{\pi(\hat{\omega}),n}, \quad \mu_{\hat{\omega}} = \mu_{\pi(\hat{\omega})}.$$

The skew product $\hat{F}$ is measurable due to (r1). The next proposition gives a relation between the (ergodic) invariant measures of F and those of $\hat{F}$. It can be found in a slightly more restrictive setting in Appendix A of [GH17] (see also the references therein), but the proof carries over unchanged to our setting. We reproduce the statement here for our setting for convenience. Use $\hat{\pi}_I : I^{\mathbb{Z}} \times [0,1) \to I^{\mathbb{Z}}$ and $\Pi : I^{\mathbb{Z}} \times [0,1) \to I^{\mathbb{N}} \times [0,1)$ to denote the respective canonical projections.

Proposition 5.4.1 ([GH17, Proposition A.1 and Remark A.2]). *Let μ be an F-invariant probability measure with marginal $\nu = \mu \circ \pi_I^{-1}$ and system of conditional measures $\{\mu_\omega\}_{\omega \in I^{\mathbb{N}}}$. Then the following statements hold.*

(i) There exists an $\hat{F}$-invariant probability measure $\hat{\mu}$ with marginal $\hat{\nu} = \hat{\mu} \circ \hat{\pi}_I^{-1}$ and a system of conditional measures $\{\hat{\mu}_{\hat{\omega}}\}_{\hat{\omega} \in I^{\mathbb{Z}}}$ such that, for $\hat{\nu}$-a.a. $\hat{\omega} \in I^{\mathbb{Z}}$,

$$\hat{\mu}_{\hat{\omega}}(B) = \lim_{n \to \infty} \mu_{\hat{\tau}^{-n}\hat{\omega}}\Big((T_{\hat{\tau}^{-n}\hat{\omega}}^n)^{-1}(B)\Big), \qquad B \in \mathcal{B}.$$

(ii) Conversely, let $\hat{\mu}$ be an $\hat{F}$-invariant probability measure with marginal $\hat{\nu} = \hat{\mu} \circ \hat{\pi}_I^{-1}$. Then the probability measure

$$\tilde{\mu} = \hat{\mu} \circ \Pi^{-1}$$

is F-invariant and has marginal $\tilde{\nu} = \hat{\nu} \circ \pi^{-1}$.

(iii) The correspondence $\mu \leftrightarrow \hat{\mu}$ given by (i) and (ii) is one-to-one and has the property that μ is ergodic for F if and only if $\hat{\mu}$ is ergodic for $\hat{F}$.

From Proposition 5.4.1 and (r6) we obtain an $\hat{F}$-invariant and ergodic probability measure $\hat{\mu}$ with a system of conditional measures $\{\hat{\mu}_{\hat{\omega}}\}_{\hat{\omega} \in I^{\mathbb{Z}}}$ for the marginal $\hat{\nu} = \hat{\mu} \circ \hat{\pi}_I^{-1}$. The following lemma will be used later.

Lemma 5.4.2. *For $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$ it holds that $\hat{\mu}_{\hat{\omega}} \ll \lambda$. Moreover, $\hat{\mu} \ll \hat{\nu} \times \lambda$ and for $\hat{\nu} \times \lambda$-a.e. $(\hat{\omega}, x) \in I^{\mathbb{Z}} \times [0,1)$ we have*

$$\frac{d\hat{\mu}}{d\hat{\nu} \times \lambda}(\hat{\omega}, x) = \frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x).$$

Proof. Combining that τ is non-singular with respect to $\mathbb{P}$ with (r4) gives that for each $n \geq 1$,

$$\mathbb{P}(\{\omega \in I^{\mathbb{N}} : \exists B \in \mathcal{B} \quad \lambda(B) = 0 \text{ and } \lambda((T_\omega^n)^{-1}B) > 0\}) = 0.$$

Recall that ν and $\mathbb{P}$ are equivalent. Furthermore, the measure $\hat{\nu}$ is invariant with respect to $\hat{\tau}$ and $\hat{\nu} \circ \pi^{-1} = \nu$. This gives for each $m \in \mathbb{Z}$ that

$$\hat{\nu}(\hat{\tau}^{-m}\pi^{-1}\{\omega \in I^{\mathbb{N}} : \exists B \in \mathcal{B} \quad \lambda(B) = 0 \text{ and } \lambda((T_\omega^n)^{-1}B) > 0\}) = 0.$$

Taking $m = -n$ it follows for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$ and $n \geq 1$ that

$$\lambda(B) = 0 \quad \Rightarrow \quad \lambda((T^n_{\hat{\tau}^{-n}(\hat{\omega})})^{-1}B) = 0, \qquad \forall B \in \mathcal{B}. \tag{5.22}$$

The measures μ and $\nu \times \lambda$ are equivalent by condition (r6). Let $A \in \mathcal{B}_I^{\mathbb{N}}$ and $B \in \mathcal{B}$. Then by (5.21)

$$\int_A \mu_\omega(B)\, d\nu(\omega) = \mu(A \times B) = \int_A \int_B \frac{d\mu}{d\nu \times \lambda}(\omega, x)\, d\lambda(x)\, d\nu(\omega),$$

which means that for any $B \in \mathcal{B}$ we can find a ν-null set N_B, such that for all $\omega \in I^{\mathbb{N}} \setminus N_B$ we have

$$\mu_\omega(B) = \int_B \frac{d\mu}{d\nu \times \lambda}(\omega, x)\, d\lambda(x).$$

Since $\mathcal{B}$ is countably generated, we can find a ν-null set N, such that for each $\omega \in I^{\mathbb{N}} \setminus N$ and all $B \in \mathcal{B}$,

$$\mu_\omega(B) = \int_B \frac{d\mu}{d\nu \times \lambda}(\omega, x)\, d\lambda(x).$$

Hence, $\mu_\omega \ll \lambda$ for ν-a.e. $\omega \in I^{\mathbb{N}}$ and for those ω,

$$\frac{d\mu_\omega}{d\lambda}(x) = \frac{d\mu}{d\nu \times \lambda}(\omega, x) \qquad \lambda\text{-a.e.}$$

It immediately follows that $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$ satisfies $\mu_{\hat{\omega}} \ll \lambda$ and since $\hat{\nu}$ is $\hat{\tau}$-invariant, we get that for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$, $\mu_{\hat{\tau}^{-n}\hat{\omega}} \ll \lambda$ for each n. Combining this with (5.22) and Proposition 5.4.1(i) gives that $\hat{\mu}_{\hat{\omega}} \ll \lambda$ for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$. Since $\{\hat{\mu}_{\hat{\omega}}\}$ is a system of conditional invariant measures, we get that for each $A \in \mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B}$ (cf. (5.21)),

$$\hat{\mu}(A) = \int_{I^{\mathbb{Z}}} \hat{\mu}_{\hat{\omega}}(A_{\hat{\omega}})\, d\hat{\nu}(\hat{\omega}) = \int_A \frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x)\, d\hat{\nu} \times \lambda(\hat{\omega}, x),$$

where $A_{\hat{\omega}} = \{x \in [0,1) : (\hat{\omega}, x) \in A\}$. This means that $\hat{\mu} \ll \hat{\nu} \times \lambda$ and that for $\hat{\nu} \times \lambda$-a.e. $(\hat{\omega}, x)$ it holds that $\frac{d\hat{\mu}}{d\hat{\nu} \times \lambda}(\hat{\omega}, x) = \frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x)$. □

§5.5 Fiber entropy

The concept of fiber entropy was introduced in [AR66]. Here, as well as in the later works [B82] and [M86], the entropy of a skew product is studied for the case that the associated transformations are all measure preserving with respect to the same measure. In [K86] and [LY88], the notion fiber entropy is considered for skew products of transformations with a Bernoulli measure on the base space. These two settings are extended in the works [B92] and [BC92], where the invariant measure of the skew product admits a system of conditional measures.

Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i \in I}, \mu, \{\alpha_i\}_{i \in I})$ be a random number system. Here we introduce fiber entropy for $\mathcal{T}$ following the approach of Bogenschütz [B93]. Two standing assumptions in [B93] (and one of the reasons why we extended F to $\hat{F}$) are that the dynamics on the base space are invertible and that the σ-algebra considered on the first coordinate is countably generated. By our definition of $\hat{F}$ and the assumption that I is a Polish space, we satisfy both these assumptions.

Consider the sub-σ-algebra $\mathcal{A} := \mathcal{B}_I^{\mathbb{Z}} \times [0,1)$ of $\mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B}$. From $\hat{\tau} \circ \hat{\pi}_I = \hat{\pi}_I \circ \hat{F}$ we see that $\hat{F}^{-1}\mathcal{A} \subseteq \mathcal{A}$ and in this situation we can define the *conditional entropy* of a partition $\mathcal{P}$ of $I^{\mathbb{Z}} \times [0,1)$ given $\mathcal{A}$ as in [B93] using the conditional expectation by

$$H_{\hat{\mu}}(\mathcal{P}|\mathcal{A}) = -\int_{I^{\mathbb{Z}}\times[0,1)} \sum_{P\in\mathcal{P}} \mathbb{E}_{\hat{\mu}}(1_P|\mathcal{A}) \log \mathbb{E}_{\hat{\mu}}(1_P|\mathcal{A})\, d\hat{\mu}.$$

Then $H_{\hat{\mu}}(\mathcal{P}|\mathcal{A}) \le H_{\hat{\mu}}(\mathcal{P})$ (see e.g. [K86, Lemma II.1.2(vi)]). The *fiber entropy* of $\mathcal{T}$ is defined in [B93, Definition 2.2.1] as

$$h^{\text{fib}}(\mathcal{T}) := \sup_{\mathcal{P}} \lim_{n\to\infty} \frac{1}{n} H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\mathcal{P}|\mathcal{A}\Big),$$

where the supremum is taken over all partitions $\mathcal{P}$ satisfying $H_{\hat{\mu}}(\mathcal{P}|\mathcal{A}) < \infty$.

As usual for entropy it is often not very practical to compute the fiber entropy of a system directly from the definition. One way to determine $h^{\text{fib}}(\mathcal{T})$ follows from the main theorem of [BC92], which gives the *Abramov-Rokhlin Formula*

$$h_{\hat{\mu}}(\hat{F}) = h_{\hat{\nu}}(\hat{\tau}) + h^{\text{fib}}(\mathcal{T}). \tag{5.23}$$

This leads to an expression for $h^{\text{fib}}(\mathcal{T})$ in case $h_{\hat{\nu}}(\hat{\tau}) < \infty$. Furthermore, the literature provides two versions of the Kolmorogov-Sinai Theorem for random systems, namely [K86, Lemma II.1.5] and [B93, Theorem 2.3.3]. The first of these results requires a generating partition $\mathcal{P}$ of the product space $I^{\mathbb{Z}} \times [0,1)$ in the sense that

$$\sigma\Big(\bigvee_{k\ge 0} \hat{F}^{-k}\mathcal{P}\Big) \vee \mathcal{A} = \mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B} \;\; \text{(up to sets of } \hat{\mu}\text{-measure zero)}. \tag{5.24}$$

The latter requires a partition γ of $[0,1)$ that satisfies for $\hat{\nu}$-a.a. $\hat{\omega} \in I^{\mathbb{Z}}$

$$\sigma\Big(\bigvee_{k\ge 0} (T_{\hat{\omega}}^k)^{-1}\gamma\Big) = \mathcal{B} \;\; \text{(up to sets of } \hat{\mu}_{\hat{\omega}}\text{-measure zero)}.$$

For random number systems a natural candidate for a generating partition is provided by the family of partitions $\{\alpha_i\}_{i\in I}$ and the corresponding partition Δ on $I^{\mathbb{N}} \times [0,1)$ from (5.15). Δ is easily extended to a partition $\hat{\Delta}$ of $I^{\mathbb{Z}} \times [0,1)$ by setting

$$\hat{\Delta} = \{\Pi^{-1}\Delta(j) : j \ge 0\}.$$

The property (r5) is not enough to guarantee that the conditions of [K86, Lemma II.1.5] are satisfied by $\hat{\Delta}$ and the conditions of [B93, Theorem 2.3.3] are not satisfied either, since we consider a family of partitions $\{\alpha_i\}_{i\in I}$ on $[0,1)$ rather than a single partition γ. Nevertheless, since from (r7) we have

$$H_{\hat{\mu}}(\hat{\Delta}|\mathcal{A}) \le H_{\hat{\mu}}(\hat{\Delta}) = H_{\mu}(\Delta) < \infty, \tag{5.25}$$

and from Lemma 5.4.2 we know that $\hat{\mu}_{\hat{\omega}} \ll \lambda$ holds for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$, by a reasoning completely analogous to the proof of [B93, Theorem 2.3.3] we still obtain that

$$h^{\text{fib}}(\mathcal{T}) = \lim_{n\to\infty} \frac{1}{n} H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}|\mathcal{A}\Big). \tag{5.26}$$

Hence, $\hat{\Delta}$ serves enough as a generating partition to have a Kolmogorov-Sinai type result and therefore an expression of $h^{\text{fib}}(\mathcal{T})$ in terms of $\hat{\Delta}$. The reason we have put (r5) as a property of random number systems and not a condition like (5.24) is that compared to condition (5.24) from [K86, Lemma II.1.5], condition (r5) is easier to verify.

The sequence $(a_n)_{n\in\mathbb{N}}$ given by $a_n = H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}|\mathcal{A}\Big)$ is subadditive (see e.g. [K86, Theorem II.1.1]), thus it follows from Fekete's Subadditive Lemma together with (5.26) that

$$h^{\text{fib}}(\mathcal{T}) = \inf_{n\in\mathbb{N}} \frac{1}{n} H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}|\mathcal{A}\Big).$$

In particular, (r7) implies via (5.25) that $h^{\text{fib}}(\mathcal{T}) < \infty$.

Using standard arguments (see e.g. [VO16, Lemma 9.1.12]) one can deduce that

$$\lim_{n\to\infty} \frac{1}{n} H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}|\mathcal{A}\Big) = \lim_{n\to\infty} H_{\hat{\mu}}\Big(\hat{\Delta}|\sigma\Big(\bigvee_{k=1}^{n-1} \hat{F}^{-k}\hat{\Delta}\Big) \vee \mathcal{A}\Big).$$

This leads to the following expression for the fiber entropy of $\mathcal{T}$:

$$h^{\text{fib}}(\mathcal{T}) = \lim_{n\to\infty} H_{\hat{\mu}}\Big(\hat{\Delta}|\sigma\Big(\bigvee_{k=1}^{n-1} \hat{F}^{-k}\hat{\Delta}\Big) \vee \mathcal{A}\Big).$$

For a partition $\mathcal{P}$ of $I^{\mathbb{Z}} \times [0,1)$ we can for each $\hat{\omega} \in I^{\mathbb{Z}}$ obtain a partition $\mathcal{P}_{\hat{\omega}}$ of $[0,1)$ by intersecting it with the fiber $\{\hat{\omega}\} \times [0,1)$, i.e. $\mathcal{P}_{\hat{\omega}} = \{Z_{\hat{\omega}} : Z \in \mathcal{P}\}$ where $Z_{\hat{\omega}} = \{x \in [0,1) : (\hat{\omega}, x) \in Z\}$. With this notation, note that

$$\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}\Big)_{\hat{\omega}} = \alpha_{\hat{\omega},n} \quad \text{and} \quad \Big(\bigvee_{k=1}^{n-1} \hat{F}^{-k}\hat{\Delta}\Big)_{\hat{\omega}} = T_{\hat{\omega}_1}^{-1}\alpha_{\hat{\tau}\hat{\omega},n}.$$

It now follows from [B93, Lemma 2.2.3] that

$$H_{\hat{\mu}}\Big(\bigvee_{k=0}^{n-1} \hat{F}^{-k}\hat{\Delta}|\mathcal{A}\Big) = \int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}(\alpha_{\hat{\omega},n})\, d\hat{\nu}(\hat{\omega}) \tag{5.27}$$

and

$$H_{\hat{\mu}}\Big(\hat{\Delta}|\sigma\Big(\bigvee_{k=1}^{n-1} \hat{F}^{-k}\hat{\Delta}\Big) \vee \mathcal{A}\Big) = \int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}\big(\alpha_{\hat{\omega}_1}\big|\sigma(T_{\hat{\omega}_1}^{-1}\alpha_{\hat{\tau}\hat{\omega},n})\big)d\hat{\nu}(\hat{\omega}).$$

Hence, $h^{\text{fib}}(\mathcal{T})$ can be rewritten as

$$h^{\text{fib}}(\mathcal{T}) = \lim_{n\to\infty} \frac{1}{n} \int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}(\alpha_{\hat{\omega},n})\, d\hat{\nu}(\hat{\omega}) \tag{5.28}$$

$$= \lim_{n\to\infty} \int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}\big(\alpha_{\hat{\omega}_1} \big| \sigma(T_{\hat{\omega}_1}^{-1}\alpha_{\hat{\tau}\hat{\omega},n})\big) d\hat{\nu}(\hat{\omega}). \tag{5.29}$$

Remark 5.5.1. Condition (r7) ensures that $H_{\hat{\mu}}(\hat{\Delta}|\mathcal{A}) < \infty$ as follows from (5.25), which enables us to rewrite $h^{\text{fib}}(\mathcal{T})$ as in (5.26). It follows from (5.27) that $H_{\hat{\mu}}(\hat{\Delta}|\mathcal{A}) < \infty$ is also ensured by requiring

$$\int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}(\alpha_{\hat{\omega}_1})\, d\hat{\nu}(\hat{\omega}) < \infty, \tag{5.30}$$

which is in general weaker than condition (r7). In view of this, we will call a collection $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ on $[0,1)$ satisfying (r1)-(r6) together with (5.30) a random number system as well. All our results also hold in this case.

§5.6 Proofs of the main results

§5.6.1 Asymptotic size of cylinder sets

In this and the next subsection we prove the three statements of Theorem 5.1.6 and we obtain Theorem 5.1.4 as a corollary. We start with Theorem 5.1.6(i) on the asymptotic size of cylinder sets. Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system.

Theorem 5.6.1. *For $\mathbb{P}$-a.a. $\omega \in I^{\mathbb{N}}$ we have*

$$\lim_{n\to\infty} -\frac{\log \lambda\big(\alpha_{\omega,n}(x)\big)}{n} = h^{\text{fib}}(\mathcal{T}) \qquad \lambda\text{-a.e.}$$

Proof. Since $I^{\mathbb{Z}}$ is a Polish space on which $\hat{\tau}$ is invertible and the measure $\hat{\mu}$ is $\hat{F}$-invariant and ergodic, we are in the position to apply [Z08, Proposition 2.2(3)], which gives the Shannon-McMillan-Breiman Theorem for random dynamical systems. We apply it to the partition $\hat{\Delta}$ of $I^{\mathbb{Z}} \times [0,1)$, which is an at most countable partition satisfying $H_{\hat{\mu}}(\hat{\Delta}|\mathcal{A}) < \infty$. Note that [Z08, Proposition 2.2] considers random compositions of continuous transformations and requires the partition under consideration to be finite. However, the continuity of the transformations is not used in the proof, and the necessary condition on the partition is that it is at most countable and satisfies

$$\int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}(\hat{\Delta}_{\hat{\omega}})\, d\hat{\nu}(\hat{\omega}) < \infty. \tag{5.31}$$

As we already observed in Remark 5.5.1, we obtain $\int_{I^{\mathbb{Z}}} H_{\hat{\mu}_{\hat{\omega}}}(\hat{\Delta}_{\hat{\omega}})\, d\hat{\nu}(\hat{\omega}) = H_{\hat{\mu}}(\hat{\Delta}|\mathcal{A})$ from (5.27), so the condition from (5.31) is satisfied using (5.25). Thus, using the formula for $h^{\text{fib}}(\mathcal{T})$ from (5.28), [Z08, Proposition 2.2(3)] gives that

$$\lim_{n\to\infty} -\frac{\log \hat{\mu}_{\hat{\omega}}\big(\alpha_{\hat{\omega},n}(x)\big)}{n} = h^{\text{fib}}(\mathcal{T}), \qquad \hat{\mu}\text{-a.e. } (\hat{\omega}, x) \in I^{\mathbb{Z}} \times [0,1). \tag{5.32}$$

Let $C \in \mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B}$ be the set of points $(\hat{\omega}, x)$ with the following three properties:

Chapter 5

(i) the limit statement from (5.32) holds;

(ii) $\frac{d\hat{\mu}}{d\hat{\nu}\times\lambda}(\hat{\omega},x) = \frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x)$;

(iii) x is a Lebesgue density point for $\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}$.

It follows from the Lebesgue Differentiation Theorem that for each $\hat{\omega} \in I^{\mathbb{Z}}$ such that $\hat{\mu}_{\hat{\omega}} \ll \lambda$ and thus $\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda} \in L^1([0,1),\lambda)$ we have that λ-a.e. $x \in [0,1)$ is a Lebesgue density point of $\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}$. Together with (5.32) and Lemma 5.4.2 we deduce that $\hat{\mu}(C) = 1$. Define

$$A = \Big\{(\omega,x) \in I^{\mathbb{N}} \times [0,1) : \frac{d\hat{\mu}}{d\hat{\nu}\times\lambda}\Big|_{\Pi^{-1}\{(\omega,x)\}\cap C} = 0\Big\}.$$

Then by Proposition 5.4.1,

$$\mu(A) = \hat{\mu}\big(\Pi^{-1}A \cap C\big) = \int_{\Pi^{-1}A\cap C} \frac{d\hat{\mu}}{d\hat{\nu}\times\lambda}\, d\hat{\nu}\times\lambda = 0,$$

thus $\mathbb{P}\times\lambda(A) = 0$ by (r6). Hence, for $\mathbb{P}\times\lambda$-a.e. (ω,x) there exists an $\hat{\omega} \in I^{\mathbb{Z}}$ such that $(\hat{\omega},x) \in C$ and $\pi(\hat{\omega}) = \omega$ and $\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x) > 0$, where this last part follows from property (ii) of C. For each such (ω,x) and $\hat{\omega}$ it follows from property (iii) of elements in C that for every $\varepsilon > 0$ there exists an $N \in \mathbb{N}$ such that for each integer $n \geq N$,

$$\Big(\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x) - \varepsilon\Big)\lambda\big(\alpha_{\hat{\omega},n}(x)\big) \leq \hat{\mu}_{\hat{\omega}}\big(\alpha_{\hat{\omega},n}(x)\big) \leq \Big(\frac{d\hat{\mu}_{\hat{\omega}}}{d\lambda}(x) + \varepsilon\Big)\lambda\big(\alpha_{\hat{\omega},n}(x)\big).$$

Here we used (r5). Combining this with (5.32) yields the result. □

From this result the Random Lochs' Theorem from Theorem 5.1.4 immediately follows.

Proof of Theorem 5.1.4. We have $h^{\text{fib}}(\mathcal{T}), h^{\text{fib}}(\mathcal{S}) \in (0,\infty)$, so Theorem 5.1.5 and Theorem 5.1.6(i) together yield the desired result. □

§5.6.2 The Random Rokhlin Formula

The last item of Theorem 5.1.6 is the Random Rokhlin Formula relating fiber entropy to the Jacobian of the transformations. We first prove an auxiliary lemma, for which we introduce some notation.

Assume as in Theorem 5.1.6(iii) that for each $i \in I$ and $A \in \alpha_i$ the restriction $T_i|_A$ is differentiable. Then for each $i \in I$ the Jacobian JT_i of T_i with respect to λ exists and is equal to $JT_i = |DT_i(x)|$ for λ-a.e. $x \in [0,1)$. From Lemma 5.4.2 together with the $\hat{\tau}$-invariance of $\hat{\nu}$ it follows that for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$, $\hat{\mu}_{\hat{\tau}\hat{\omega}} \ll \lambda$. By (r5) we obtain that for $\hat{\nu}$-a.e. $\hat{\omega} \in I^{\mathbb{Z}}$,

$$\sigma\big(\lim_{n\to\infty} \alpha_{\hat{\tau}\hat{\omega},n}\big) = \mathcal{B} \text{ up to sets of } \hat{\mu}_{\hat{\tau}\hat{\omega}}\text{-measure zero.} \tag{5.33}$$

Since the standing assumptions of [B93] are satisfied, [B93, Lemma 1.1.2] gives that

$$\hat{\mu}_{\hat{\omega}} \circ T_{\hat{\omega}_1}^{-1} = \hat{\mu}_{\hat{\tau}\hat{\omega}}, \qquad \hat{\nu}\text{-a.a. } \hat{\omega} \in I^{\mathbb{Z}}. \tag{5.34}$$

Let

$$E = \{\hat\omega \in I^{\mathbb{Z}} : \hat\mu_{\hat\tau\hat\omega} \ll \lambda \text{ and } (5.33) \text{ and } (5.34) \text{ hold}\}.$$

Then $\hat\nu(E) = 1$. For $\hat\omega \in E$ let $C_{\hat\omega} = \{x \in [0,1) : \frac{d\hat\mu_{\hat\tau\hat\omega}}{d\lambda}(x) > 0\}$, define $h_{\hat\omega} : [0,1) \to [0,\infty)$ by

$$h_{\hat\omega}(y) = \begin{cases} \left(\frac{d\hat\mu_{\hat\tau\hat\omega}}{d\lambda}(y)\right)^{-1}, & \text{if } y \in C_{\hat\omega}, \\ 1, & \text{if } y \in [0,1)\backslash C_{\hat\omega}, \end{cases}$$

and define

$$\mathcal{L}_{\hat\omega}\psi(y) = h_{\hat\omega}(y) \sum_{z \in T_{\hat\omega_1}^{-1}\{y\}} \frac{\psi(z)}{JT_{\hat\omega_1}(z)} \cdot \frac{d\hat\mu_{\hat\omega}}{d\lambda}(z),$$

which, as we will see in the proof of Lemma 5.6.2 below, is well defined as an operator from $L^1(\hat\mu_{\hat\omega}) = L^1([0,1), \hat\mu_{\hat\omega})$ to $L^1(\hat\mu_{\hat\tau\hat\omega})$.

Lemma 5.6.2. *For $\hat\omega \in E$, $\psi \in L^1(\hat\mu_{\hat\omega})$ and $n \geq 1$ the following hold.*

(i) $\mathbb{E}_{\hat\mu_{\hat\omega}}(\psi|\sigma(T_{\hat\omega_1}^{-1}\alpha_{\hat\tau\hat\omega,n}))(x) = \mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega}\psi|\sigma(\alpha_{\hat\tau\hat\omega,n}))(T_{\hat\omega_1}x)$ *for $\hat\mu_{\hat\omega}$-a.e. x.*

(ii) $\lim_{n\to\infty} \mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega}\psi|\sigma(\alpha_{\hat\tau\hat\omega,n})) = \mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega}\psi|\mathcal{B}) = \mathcal{L}_{\hat\omega}\psi$ *for $\hat\mu_{\hat\tau\hat\omega}$-a.e. x.*

Proof. By (5.34) and the definition of $C_{\hat\omega}$, $\hat\mu_{\hat\omega}(T_{\hat\omega_1}^{-1}C_{\hat\omega}) = \hat\mu_{\hat\tau\hat\omega}(C_{\hat\omega}) = 1$. Using the change of variables formula from Lemma 1.2.20(b) we obtain for all $B \in \alpha_{\hat\tau\hat\omega,n}$ that

$$\begin{aligned}
\int_B \mathcal{L}_{\hat\omega}\psi\, d\hat\mu_{\hat\tau\hat\omega} &= \int_{B\cap C_{\hat\omega}} \mathcal{L}_{\hat\omega}\psi \frac{d\hat\mu_{\hat\tau\hat\omega}}{d\lambda}\, d\lambda = \int_{B\cap C_{\hat\omega}} \sum_{z\in T_{\hat\omega_1}^{-1}\{y\}} \frac{\psi(z)}{JT_{\hat\omega_1}(z)} \cdot \frac{d\hat\mu_{\hat\omega}}{d\lambda}(z)\, d\lambda(y) \\
&= \sum_{A\in\alpha_{\hat\omega_1}} \int_{T_{\hat\omega_1}(A)\cap B\cap C_{\hat\omega}} \left(\frac{\psi\cdot\frac{d\hat\mu_{\hat\omega}}{d\lambda}}{JT_{\hat\omega_1}}\right) \circ (T_{\hat\omega_1}|_A)^{-1}(y)\, d\lambda(y) \\
&= \sum_{A\in\alpha_{\hat\omega_1}} \int_{A\cap T_{\hat\omega_1}^{-1}B\cap T_{\hat\omega_1}^{-1}C_{\hat\omega}} \psi\cdot\frac{d\hat\mu_{\hat\omega}}{d\lambda}\, d\lambda \\
&= \int_{T_{\hat\omega_1}^{-1}B} \psi\, d\hat\mu_{\hat\omega}.
\end{aligned}$$

In particular, $\mathcal{L}_{\hat\omega} : L^1(\hat\mu_{\hat\omega}) \to L^1(\hat\mu_{\hat\tau\hat\omega})$ is well defined. We obtain that

$$\begin{aligned}
\mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega}\psi|\sigma(\alpha_{\hat\tau\hat\omega,n}))(T_{\hat\omega_1}x) &= \sum_{B\in\alpha_{\hat\tau\hat\omega,n}} 1_B(T_{\hat\omega_1}x)\frac{\int_B \mathcal{L}_{\hat\omega}\psi\, d\hat\mu_{\hat\tau\hat\omega}}{\hat\mu_{\hat\tau\hat\omega}(B)} \\
&= \sum_{B\in\alpha_{\hat\tau\hat\omega,n}} 1_{T_{\hat\omega_1}^{-1}B}(x)\frac{\int_{T_{\hat\omega_1}^{-1}B} \psi\, d\hat\mu_{\hat\omega}}{\hat\mu_{\hat\omega}(T_{\hat\omega_1}^{-1}B)} \\
&= \sum_{A\in T_{\hat\omega_1}^{-1}\alpha_{\hat\tau\hat\omega,n}} 1_A(x)\frac{\int_A \psi\, d\hat\mu_{\hat\omega}}{\hat\mu_{\hat\omega}(A)} \\
&= \mathbb{E}_{\hat\mu_{\hat\omega}}(\psi|\sigma(T_{\hat\omega_1}^{-1}\alpha_{\hat\tau\hat\omega,n}))(x)
\end{aligned}$$

for $\hat\mu_{\hat\tau\hat\omega}$-a.e. x. This gives (i). Part (ii) follows from combining (5.33) and Lévy's Upward Theorem. □

This lemma is enough to prove the following Random Rokhlin's Formula.

Theorem 5.6.3. *If for each $i \in I$ and $A \in \alpha_i$ the restriction $T_i|_A$ is differentiable, then*

$$h^{\mathrm{fib}}(\mathcal{T}) = \int_{I^{\mathbb{N}} \times [0,1)} \log |DT_{\omega_1}(x)| \, d\mu(\omega, x).$$

Proof. We follow the reasoning of the proof of [VO16, Theorem 9.7.3]. Write $\phi(x) = -x \log x$. As before, set $\mathcal{A} = \mathcal{B}_I^{\mathbb{Z}} \times [0,1)$. Then using (5.31) and the Dominated Convergence Theorem, we see from (5.29) that

$$\begin{aligned} h^{\mathrm{fib}}(\mathcal{T}) &= \int_{I^{\mathbb{Z}}} \lim_{n\to\infty} H_{\hat\mu_{\hat\omega}}\big(\alpha_{\hat\omega_1} \big| \sigma(T_{\hat\omega_1}^{-1} \alpha_{\hat\tau\hat\omega,n})\big) d\hat\nu(\hat\omega) \\ &= \int_{I^{\mathbb{Z}}} \lim_{n\to\infty} \int_{[0,1)} \sum_{A \in \alpha_{\hat\omega_1}} \phi\big(\mathbb{E}_{\hat\mu_{\hat\omega}}(1_A | \sigma(T_{\hat\omega_1}^{-1} \alpha_{\hat\tau\hat\omega,n}))\big)(x)\, d\hat\mu_{\hat\omega}(x) d\hat\nu(\hat\omega) \\ &= \int_{I^{\mathbb{Z}}} \sum_{A \in \alpha_{\hat\omega_1}} \int_{[0,1)} \phi\big(\lim_{n\to\infty} \mathbb{E}_{\hat\mu_{\hat\omega}}(1_A | \sigma(T_{\hat\omega_1}^{-1} \alpha_{\hat\tau\hat\omega,n}))\big)(x)\, d\hat\mu_{\hat\omega}(x) d\hat\nu(\hat\omega), \end{aligned} \tag{5.35}$$

where for the last equality we used the Dominated Convergence Theorem again as well as the continuity of ϕ. From Lemma 5.6.2 with $\psi = 1_A$ we get for each $\hat\omega \in E$ from (i) that

$$\mathbb{E}_{\hat\mu_{\hat\omega}}(1_A | \sigma(T_{\hat\omega_1}^{-1} \alpha_{\hat\tau\hat\omega,n}))(x) = \mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega} 1_A | \sigma(\alpha_{\hat\tau\hat\omega,n}))(T_{\hat\omega_1} x) \qquad \hat\mu_{\hat\omega}\text{-a.e.} \tag{5.36}$$

and from (ii) that

$$\lim_{n\to\infty} \mathbb{E}_{\hat\mu_{\hat\tau\hat\omega}}(\mathcal{L}_{\hat\omega} 1_A | \sigma(\alpha_{\hat\tau\hat\omega,n})) = \mathcal{L}_{\hat\omega} 1_A \qquad \hat\mu_{\hat\tau\hat\omega}\text{-a.e.} \tag{5.37}$$

Recall that $C_{\hat\omega} = \{x \in [0,1) : \frac{d\hat\mu_{\hat\tau\hat\omega}}{d\lambda}(x) > 0\}$ satisfies $\hat\mu_{\hat\tau\hat\omega}(C_{\hat\omega}) = 1$. Combining (5.36) and (5.37) with (5.35) and using $\hat\mu_{\hat\tau\hat\omega} = \hat\mu_{\hat\omega} \circ T_{\hat\omega_1}^{-1}$ and the change of variables formula from Lemma 1.2.20(b) we conclude that

$$\begin{aligned} h^{\mathrm{fib}}(\mathcal{T}) &= \int_{I^{\mathbb{Z}}} \sum_{A \in \alpha_{\hat\omega_1}} \int_{[0,1)} \phi\big(\mathcal{L}_{\hat\omega} 1_A(x)\big)\, d\hat\mu_{\hat\tau\hat\omega}(x)\, d\hat\nu(\hat\omega) \\ &= -\int_{I^{\mathbb{Z}}} \sum_{A \in \alpha_{\hat\omega_1}} \left[\int_{T_{\hat\omega_1}(A) \cap C_{\hat\omega}} \Big(\frac{1}{JT_{\hat\omega_1}} \frac{d\hat\mu_{\hat\omega}}{d\lambda} \log\Big(\frac{h_{\hat\omega} \circ T_{\hat\omega_1}}{JT_{\hat\omega_1}} \frac{d\hat\mu_{\hat\omega}}{d\lambda}\Big)\Big) \circ (T_{\hat\omega_1}|A)^{-1}\, d\lambda \right] d\hat\nu(\hat\omega) \\ &= -\int_{I^{\mathbb{Z}}} \sum_{A \in \alpha_{\hat\omega_1}} \int_{A \cap T_{\hat\omega_1}^{-1} C_{\hat\omega}} \log\Big(\frac{h_{\hat\omega} \circ T_{\hat\omega_1}}{JT_{\hat\omega_1}} \frac{d\hat\mu_{\omega}}{d\lambda}\Big)(x)\, d\hat\mu_{\hat\omega}(x) d\hat\nu(\hat\omega) \\ &= \int_{I^{\mathbb{Z}} \times [0,1)} \log JT_{\hat\omega_1}(x)\, d\hat\mu(\hat\omega, x) - \int_{I^{\mathbb{Z}} \times [0,1)} \log h_{\hat\omega} \circ T_{\hat\omega_1}(x)\, d\hat\mu(\hat\omega, x) \\ &\quad - \int_{I^{\mathbb{Z}} \times [0,1)} \log \frac{d\hat\mu_{\hat\omega}}{d\lambda}(x)\, d\hat\mu(\omega, x). \end{aligned}$$

For each $(\hat\omega, x) \in I^{\mathbb{Z}} \times [0,1)$, set

$$\eta(\hat\omega, x) = \begin{cases} \big(\frac{d\hat\mu_{\hat\omega}}{d\lambda}(x)\big)^{-1}, & \text{if } \hat\mu_{\hat\omega} \ll \lambda \text{ and } \frac{d\hat\mu_{\hat\omega}}{d\lambda}(x) > 0, \\ 1, & \text{otherwise.} \end{cases}$$

Then for each $\hat{\omega} \in E$ and $x \in T_{\hat{\omega}_1}^{-1}(C_{\hat{\omega}})$,

$$h_{\hat{\omega}} \circ T_{\hat{\omega}_1}(x) = \eta \circ \hat{F}(\hat{\omega}, x).$$

With the $\hat{F}$-invariance of $\hat{\mu}$ this yields

$$\begin{aligned}\int_{I^{\mathbb{Z}}\times[0,1)} \log h_{\hat{\omega}} \circ T_{\hat{\omega}_1}(x)\, d\hat{\mu}(\hat{\omega}, x) &= \int_E \int_{T_{\hat{\omega}_1}^{-1} C_{\hat{\omega}}} \log \eta \circ \hat{F}(\hat{\omega}, x)\, d\hat{\mu}_{\hat{\omega}}(x) d\hat{\nu}(\hat{\omega}) \\ &= \int_{I^{\mathbb{Z}}\times[0,1)} \log \eta \, d\hat{\mu}.\end{aligned}$$

The result now follows. □

Proof of Theorem 5.1.6. Parts (i) and (iii) are given by Theorem 5.6.1 and Theorem 5.6.3, respectively. By (5.23) for part (ii) it is enough to show that $h_{\hat{\mu}}(\hat{F}) = h_{\mu}(F)$ and $h_{\hat{\nu}}(\hat{\tau}) = h_{\nu}(\tau)$. The latter follows immediately, because $(I^{\mathbb{Z}}, \mathcal{B}_I^{\mathbb{Z}}, \hat{\nu}, \hat{\tau})$ is the *natural extension* of the system $(I^{\mathbb{N}}, \mathcal{B}_I^{\mathbb{N}}, \nu, \tau)$, i.e. $(I^{\mathbb{Z}}, \mathcal{B}_I^{\mathbb{Z}}, \hat{\nu}, \hat{\tau})$ is the smallest (in the sense of the σ-algebra) invertible system that has $(I^{\mathbb{N}}, \mathcal{B}_I^{\mathbb{N}}, \nu, \tau)$ as a factor, and entropy is preserved under this construction (see e.g. [DK21, Chapter 5]). For the first part, note that $(I^{\mathbb{N}} \times [0,1), \mathcal{B}_I^{\mathbb{N}} \times \mathcal{B}, \mu, F)$ is a factor of $(I^{\mathbb{Z}} \times [0,1), \mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B}, \hat{\mu}, \hat{F})$. Hence, $h_{\hat{\mu}}(\hat{F}) \geq h_{\mu}(F)$. Conversely, let $\mathcal{C} = \{C_1, C_2, \ldots\}$ be a countable set that generates $\mathcal{B}_I$, so $\sigma(\mathcal{C}) = \mathcal{B}_I$. Define for each $n \in \mathbb{N}$ the partition γ_n of I given by $\gamma_n = \{C_i \backslash \bigcup_{j=1}^{i-1} C_j \,:\, i = 1, \ldots, n\} \cup \{X \backslash \bigcup_{i=1}^{n} C_i\}$. Then $\sigma(\gamma_n) \subseteq \sigma(\gamma_{n+1})$ for each $n \geq 1$ and $\sigma(\lim_{n\to\infty} \gamma_n) = \mathcal{B}_I$. Similarly, since $\mathcal{B}$ is countably generated there exists a sequence of finite partitions $(\beta_n)_{n\geq 1}$ of $[0,1)$ such that $\sigma(\lim_{n\to\infty} \beta_n) = \mathcal{B}$. Then

$$\xi_n = \{(\cdots \times I \times A_{-n} \times A_{-n+1} \times \cdots \times A_n \times I \times \cdots) \times B : A_i \in \gamma_n, i = -n, \ldots, n, B \in \beta_n\}$$

(where A_0 is on the 0-th position) is a finite partition of $I^{\mathbb{Z}} \times [0,1)$ for each $n \in \mathbb{N}$ such that $\sigma(\lim_{n\to\infty} \xi_n) = \mathcal{B}_I^{\mathbb{Z}} \times \mathcal{B}$. Note that $\hat{F}^{-n-1}\xi_n$ specifies sets in positions 1 to $2n+1$ in the first coordinate. Recall that $\Pi : I^{\mathbb{Z}} \times [0,1) \to I^{\mathbb{N}} \times [0,1)$ denotes the canonical projection. We conclude using Lemma 1.2.18 that

$$\begin{aligned} h_{\hat{\mu}}(\hat{F}) &= \lim_{n\to\infty} h_{\hat{\mu}}(\hat{F}, \xi_n) = \lim_{n\to\infty} h_{\hat{\mu}}(\hat{F}, \hat{F}^{-n-1}\xi_n) \\ &= \lim_{n\to\infty} h_{\mu}(F, \Pi(\hat{F}^{-n-1}\xi_n)) \leq h_{\mu}(F). \end{aligned}$$

This finishes the proof. □

§5.6.3 The Central Limit Theorem

In case we compare a random number system to an NTFM, then under additional assumptions on both systems we can obtain a Central Limit Theorem for Theorem 5.1.4 in a way comparable to what has been done in [H09] for two NTFM's. Given a random number system $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ and an NTFM $(S, \tilde{\mu}, \gamma)$, for each $n \in \mathbb{N}$, $\omega \in I^{\mathbb{N}}$ and $x \in [0,1)$, let

$$m_{\mathcal{T},S}(n, \omega, x) = \sup\{m \in \mathbb{N} : \alpha_{\omega,n}(x) \subseteq \gamma_m(x)\}.$$

This is the analog of (5.13) for comparing two random number systems. We first introduce additional assumptions that we put on the systems. The following property can be found in [H09, Property 2.1].

Definition 5.6.4. Let $(S, \tilde{\mu}, \gamma)$ be an NTFM. We say that S satisfies the *zero-property* if

$$\lim_{n\to\infty} \frac{-\log\lambda(\gamma_n(x)) - nh_{\tilde{\mu}}(S)}{\sqrt{n}} = 0 \qquad \lambda\text{-a.e.}$$

The zero-property is rather strong, but [H09, Section 3.2] contains several examples of NTFM's that satisfy it, including the N-adic transformations from Example 5.1.1 and β-transformations $S(x) = \beta x \bmod 1$ with $\beta > 1$ a so-called *Parry number*, i.e. a number β for which the set $\{S^n(\beta - 1) : n \geq 0\}$ is finite. The following lemma, which compares to [H09, Lemma 2.1], yields a consequence of the zero-property.

Lemma 5.6.5 (cf. Lemma 2.1 in [H09]). *Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system and let $(S, \tilde{\mu}, \gamma)$ be an NTFM that satisfies the zero-property. Assume that $h^{\text{fib}}(\mathcal{T}), h_{\tilde{\mu}}(S) \in (0, \infty)$. Then*

$$\log\left(\frac{\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x))}{\lambda(\alpha_{\omega,n}(x))}\right) = o(\sqrt{n}) \quad \textit{in } \mu\textit{-probability.}$$

That is, for each $\varepsilon > 0$,

$$\lim_{n\to\infty} \mu\big(\big\{(\omega, x) : \big|\log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) - \log\lambda(\alpha_{\omega,n}(x))\big| > \varepsilon\sqrt{n}\big\}\big) = 0.$$

Proof. By definition of $m_{\mathcal{T},S}(n, \omega, x)$ we see that $\alpha_{\omega,n}(x) \subseteq \gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)$. Since $\mu \ll \mathbb{P}\times\lambda$, it suffices to show that for all $\varepsilon, \tilde{\varepsilon} > 0$ there exists an $N \in \mathbb{N}$ such that for all $n \geq N$ we have

$$\mathbb{P}\times\lambda\big(\big\{(\omega, x)\ :\ \log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) - \log\lambda(\alpha_{\omega,n}(x)) > \varepsilon\sqrt{n}\big\}\big) < \tilde{\varepsilon}.$$

Fix $\varepsilon, \tilde{\varepsilon} > 0$ and set $\eta = \varepsilon\sqrt{\frac{h_{\tilde{\mu}}(S)}{3h^{\text{fib}}(\mathcal{T})}}$. For each $n \in \mathbb{N}$, we put

$$A_n = I^{\mathbb{N}} \times \big\{x \in [0, 1]\ :\ \exists\, k \geq n \text{ s.t. } |\log\lambda(\gamma_k(x)) + kh_{\tilde{\mu}}(S)| > \frac{1}{2}\eta\sqrt{k}\,\big\}.$$

Because S satisfies the zero-property, we know that there exists an $n_0 \in \mathbb{N}$ such that $\mathbb{P}\times\lambda(A_{n_0}) \leq \frac{\tilde{\varepsilon}}{3}$. Note that if $(\omega, x) \notin A_{n_0}$, then for each $n > n_0$ we have

$$\log\lambda(\gamma_{n-1}(x)) \leq \frac{1}{2}\eta\sqrt{n} - (n-1)h_{\tilde{\mu}}(S),$$
$$\log\lambda(\gamma_n(x)) \geq -\frac{1}{2}\eta\sqrt{n} - nh_{\tilde{\mu}}(S),$$

which when combined leads to

$$\log\lambda(\gamma_{n-1}(x)) - \log\lambda(\gamma_n(x)) \leq h_{\tilde{\mu}}(S) + \eta\sqrt{n}. \tag{5.38}$$

For each $n \geq 1$ put

$$B_n = \Big\{ (\omega, x) \in I^{\mathbb{N}} \times [0,1) \,:\, \frac{1}{2} < \frac{m_{\mathcal{T},S}(k,\omega,x)}{k} \frac{h_{\tilde{\mu}}(S)}{h^{\mathrm{fib}}(\mathcal{T})} < 2 \quad \forall\, k \geq n \Big\}. \tag{5.39}$$

From Theorem 5.1.4 it follows that there exists an integer $n_1 > 2n_0 \frac{h_{\tilde{\mu}}(S)}{h^{\mathrm{fib}}(\mathcal{T})}$ such that $\mathbb{P} \times \lambda(B_{n_1}) > 1 - \frac{\tilde{\varepsilon}}{3}$. Note from (5.39) that for each $(\omega, x) \in B_{n_1}$ we have $n_0 < \frac{1}{2} n_1 \frac{h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)} < m_{\mathcal{T},S}(n_1,\omega,x)$. Therefore, it follows from (5.38) that, for each $(\omega,x) \in B_{n_1} \setminus A_{n_0}$ and for each $n > n_1$,

$$\log \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) \leq h_{\tilde{\mu}}(S) + \eta \sqrt{m_{\mathcal{T},S}(n,\omega,x)+1} + \log \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)). \tag{5.40}$$

For each $n \geq 1$ and interval $E \in \gamma_n$ use ∂E to denote the boundary of E, i.e. the collection of its two endpoints, and use $\mathrm{dist}(x, \partial E) = \inf\{|x-a| \,:\, a \in \partial E\}$ to denote the distance from x to the nearest boundary point of E. For each $n \in \mathbb{N}$ and $(\omega, x) \in I^{\mathbb{N}} \times [0,1)$ we have that $\alpha_{\omega,n}(x) \nsubseteq \gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)$, so $\mathrm{dist}(x, \partial\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)) \leq \lambda(\alpha_{\omega,n}(x))$ and with (5.40) we obtain for each $(\omega,x) \in B_{n_1} \setminus A_{n_0}$ and each $n > n_1$ that

$$\begin{aligned} &\frac{\log \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) - \log \lambda(\alpha_{\omega,n}(x))}{\sqrt{n}} \\ &\quad \leq \frac{\log \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)) - \log \mathrm{dist}(x, \partial\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x))}{\sqrt{n}} + \frac{h_{\tilde{\mu}}(S)}{\sqrt{n}} \\ &\qquad + \eta \sqrt{\frac{m_{\mathcal{T},S}(n,\omega,x)+1}{n}}. \end{aligned} \tag{5.41}$$

For each $n \in \mathbb{N}$ and interval $E \in \gamma_n$, we define a new interval E' by removing from both ends of E an interval of length $\frac{\tilde{\varepsilon}}{\pi^2 n^2}\lambda(E)$. Furthermore, we define

$$C = [0,1) \setminus \Big(\bigcup_{n \in \mathbb{N}} \bigcup_{E \in \gamma_n} E \setminus E' \Big).$$

Then

$$\lambda(C) \geq 1 - \sum_{n \in \mathbb{N}} \sum_{E \in \gamma_n} \frac{2\tilde{\varepsilon}}{\pi^2 n^2} \lambda(E) = 1 - \frac{\tilde{\varepsilon}}{3},$$

so $\mathbb{P} \times \lambda\big((B_{n_1} \cap (I^{\mathbb{N}} \times C)) \setminus A_{n_0} \big) \geq 1 - \tilde{\varepsilon}$. For each $(\omega, x) \in I^{\mathbb{N}} \times C$ and $n \in \mathbb{N}$ we have the bound

$$\mathrm{dist}(x, \partial\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)) \geq \frac{\tilde{\varepsilon}}{\pi^2 n^2} \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)+1}(x)).$$

Combining this with (5.39) and (5.41) gives for each integer $n > n_1$ and each $(\omega,x) \in \big(B_{n_1} \cap (I^{\mathbb{N}} \times C)\big) \setminus A_{n_0}$ that

$$\frac{\log \lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) - \log \lambda(\alpha_{\omega,n}(x))}{\sqrt{n}} \leq \frac{|\log \frac{\pi^2 n^2}{\tilde{\varepsilon}}| + h_{\tilde{\mu}}(S)}{\sqrt{n}} + \eta \sqrt{\frac{2h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)} + \frac{1}{n}}. \tag{5.42}$$

We now take an integer $N > n_1$ large enough such that for each integer $n \geq N$ the right-hand side of (5.42) is bounded by $\eta\sqrt{\frac{3h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)}} = \varepsilon$. This yields the lemma. □

Furthermore, we consider random number systems with the following property.

Definition 5.6.6. Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system. We say that $\mathcal{T}$ satisfies the *CLT-property* if there is some $\sigma > 0$ such that

$$\lim_{n\to\infty} \mu\Big(\Big\{(\omega, x) \in I^{\mathbb{N}}\times[0,1) \colon \frac{-\log\lambda(\alpha_{\omega,n}(x)) - nh^{\mathrm{fib}}(\mathcal{T})}{\sigma\sqrt{n}} \leq u\Big\}\Big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}dt$$

holds for each $u \in \mathbb{R}$.

With the properties from Definitions 5.6.4 and 5.6.6 we obtain the following result.

Theorem 5.6.7. *Let $\mathcal{T} = (I, \mathbb{P}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ be a random number system satisfying the CLT-property with constant $\sigma > 0$ and let $(S, \tilde{\mu}, \gamma)$ be an NTFM that satisfies the zero-property. Assume that $h^{\mathrm{fib}}(\mathcal{T}), h_{\tilde{\mu}}(S) \in (0, \infty)$. Then for all $u \in \mathbb{R}$,*

$$\lim_{n\to\infty} \mu\Big(\Big\{(\omega, x) \in I^{\mathbb{N}} \times [0,1) : \frac{m_{\mathcal{T},S}(n,\omega,x) - n\frac{h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)}}{\kappa\sqrt{n}} \leq u\Big\}\Big) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}\, dt,$$

where $\kappa = \frac{\sigma}{h_{\tilde{\mu}}(S)}$.

Proof. Fix some $u \in \mathbb{R}$. We rewrite

$$\begin{aligned} \frac{m_{\mathcal{T},S}(n,\omega,x) - n\frac{h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)}}{\kappa\sqrt{n}} &= \frac{-\log\lambda(\alpha_{\omega,n}(x)) - nh^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}} \\ &\quad + \frac{\log\lambda(\alpha_{\omega,n}(x)) - \log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x))}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}} \\ &\quad + \frac{\log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) + h_{\tilde{\mu}}(S)m_{\mathcal{T},S}(n,\omega,x)}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}}. \end{aligned} \tag{5.43}$$

The last term can be written as

$$\begin{aligned} &\frac{\log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) + h_{\tilde{\mu}}(S)m_{\mathcal{T},S}(n,\omega,x)}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}} \\ &\quad = \frac{1}{h_{\tilde{\mu}}(S)\kappa}\sqrt{\frac{m_{\mathcal{T},S}(n,\omega,x)}{n}}\frac{\log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) + h_{\tilde{\mu}}(S)m_{\mathcal{T},S}(n,\omega,x)}{\sqrt{m_{\mathcal{T},S}(n,\omega,x)}}. \end{aligned} \tag{5.44}$$

From Theorem 5.1.4 we know that $\lim_{n\to\infty}\frac{m_{\mathcal{T},S}(n,\omega,x)}{n} = \frac{h^{\mathrm{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)} < \infty$ for $\mathbb{P}\times\lambda$-a.e. $(\omega, x) \in I^{\mathbb{N}} \times [0,1)$. Since $\lim_{n\to\infty} m_{\mathcal{T},S}(n,\omega,x) = \infty$ for $\mathbb{P}\times\lambda$-a.e. $(\omega,x) \in I^{\mathbb{N}} \times [0,1)$, it then follows from the zero-property of S that (5.44) converges to 0 as $n \to \infty$ $\mathbb{P}\times\lambda$-a.e. and thus also μ-a.e. Hence it converges in μ-probability as well. Furthermore, we know from Lemma 5.6.5 that the second term on the right-hand side of (5.43) converges to 0 as $n\to\infty$ in μ-probability.

CHAPTER 5

Define three sequences of random variables $(X_n)_{n\geq 1}, (Y_n)_{n\geq 1}$ and $(Z_n)_{n\geq 1}$ on $I^{\mathbb{N}} \times [0,1)$ by setting

$$\begin{aligned} X_n &= \frac{m_{\mathcal{T},S}(n,\omega,x) - n\frac{h^{\text{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)}}{\kappa\sqrt{n}}, \\ Y_n &= \frac{-\log\lambda(\alpha_{\omega,n}(x)) - nh^{\text{fib}}(\mathcal{T})}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}}, \\ Z_n &= \frac{\log\lambda(\alpha_{\omega,n}(x)) - \log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x))}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}} \\ &\quad + \frac{\log\lambda(\gamma_{m_{\mathcal{T},S}(n,\omega,x)}(x)) + h_{\tilde{\mu}}(S)m_{\mathcal{T},S}(n,\omega,x)}{h_{\tilde{\mu}}(S)\kappa\sqrt{n}}. \end{aligned}$$

Then by the above for each $\varepsilon > 0$, $\lim_{n\to\infty}\mu(|Z_n| > \varepsilon) = 0$ and since $\mathcal{T}$ satisfies the CLT-property for each $u \in \mathbb{R}$ it holds that

$$\lim_{n\to\infty}\mu(Y_n \leq u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-t^2/2}dt.$$

Fix some $u \in \mathbb{R}$ and some $\varepsilon > 0$. We are interested in $\lim_{n\to\infty}\mu(X_n \leq u)$. From (5.43) we see that

$$\begin{aligned} \mu(X_n \leq u) &= \mu(Y_n \leq u - Z_n) \\ &= \mu(Y_n \leq u - Z_n,\, |Z_n| \leq \varepsilon) + \mu(Y_n \leq u - Z_n,\, |Z_n| > \varepsilon). \end{aligned}$$

Since $\lim_{n\to\infty}\mu(Y_n \leq u - Z_n,\, |Z_n| > \varepsilon) = 0$ we get

$$\begin{aligned} \limsup_{n\to\infty}\mu(X_n \leq u) &= \limsup_{n\to\infty}\mu(Y_n \leq u - Z_n,\, |Z_n| \leq \varepsilon) \\ &\leq \limsup_{n\to\infty}\mu(Y_n \leq u + \varepsilon) \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u+\varepsilon} e^{-t^2/2}dt. \end{aligned}$$

On the other hand,

$$\begin{aligned} \liminf_{n\to\infty}\mu(X_n \leq u) &= \liminf_{n\to\infty}\mu(Y_n \leq u - Z_n,\, |Z_n| \leq \varepsilon) \\ &\geq \liminf_{n\to\infty}\mu(Y_n \leq u - \varepsilon,\, |Z_n| \leq \varepsilon) \\ &\geq \liminf_{n\to\infty}\big(\mu(Y_n \leq u - \varepsilon) - \mu(|Z_n| > \varepsilon)\big) \\ &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u-\varepsilon} e^{-t^2/2}dt. \end{aligned}$$

Since this holds for all $\varepsilon > 0$ we get the result. □

We will now identify a class of random number systems that satisfies the CLT property. Let $\{T_i : [0,1) \to [0,1)\}_{i\in I}$ be a countable collection of transformations that satisfy the following two properties:

(p1) For each $i \in I$ there exists an interval partition α_i of $[0,1)$, such that for each $A \in \alpha_i$, $T_i|_A$ has non-positive Schwarzian derivative and $T_i(A) = [0,1)$.

(p2) There exist $1 < K \leq M < \infty$ such that, for all $x \in [0,1)$ and all $i \in I$,

$$K \leq |DT_i(x)| \leq M. \tag{5.45}$$

In particular, each T_i is expanding and has finitely many onto branches (so each partition α_i has at most finitely many non-empty intervals) and $T_i|_A$ is C^3 for each $A \in \alpha_i$. We let $\mathcal{D}$ denote the set of all collections $\{T_i\}_{i\in I}$ satisfying conditions (p1) and (p2) for some countable index set I and show that each element of $\mathcal{D}$ gives a random number system, that under some additional assumptions satisfies the CLT-property.

Proposition 5.6.8. *Let $\{T_i\}_{i\in I} \in \mathcal{D}$, $\{\alpha_i\}_{i\in I}$ the set of partitions given by Property (p1), let $\boldsymbol{p} = (p_i)_{i\in I}$ be a strictly positive probability vector and let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $I^{\mathbb{N}}$. Then there exists a unique measure μ such that $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ is a random number system. Moreover, $\mathcal{T}$ satisfies the CLT-property if and only if for each measurable function $\psi : I^{\mathbb{N}} \times [0,1) \to \mathbb{R}$ we have $\varphi \circ F \neq \psi \circ F - \psi$, where F is the skew product associated to $\{T_i\}_{i\in I}$ and $\varphi : I^{\mathbb{N}} \times [0,1) \to \mathbb{R}$ is given by*

$$\varphi(\omega, x) = \log |DT_{\omega_1}(x)| - h^{\text{fib}}(\mathcal{T}). \tag{5.46}$$

To prove Proposition 5.6.8 we use two theorems by Young [Y99], both of which we have already used in Chapter 3. The results from [Y99] are formulated for Young towers, i.e. extensions of induced systems for suitable return time maps. We will, however, apply them to the system itself. That is, we will take the whole space $I^{\mathbb{N}} \times [0,1)$ as the inducing domain and as a consequence the return time function $R : I^{\mathbb{N}} \times [0,1) \to \mathbb{N}$ will have $R(\omega, x) = 1$ for each (ω, x). In particular $\int_{I^{\mathbb{N}}\times[0,1)} R \, dm_{\boldsymbol{p}} \times \lambda = 1$. For the convenience of the reader, we will reformulate the results from [Y99] that are relevant for the proof of Proposition 5.6.8 here for our setting, together with the necessary conditions.

The skew product F maps each element $[i] \times A$, $i \in I$ and $A \in \alpha_i$, bijectively onto $I^{\mathbb{N}} \times [0,1)$. Moreover, both $F|_{[i]\times A}$ and its inverse are non-singular with respect to $m_{\boldsymbol{p}} \times \lambda$ (giving (r4)). Hence, the Jacobian $J_{m_{\boldsymbol{p}}\times\lambda} F$ exists and is positive $m_{\boldsymbol{p}} \times \lambda$-a.e. By condition (p2) the collection $\{[i] \times A : i \in I, A \in \alpha_i\}$ generates the σ-algebra $\mathcal{B}_I^{\mathbb{N}} \times \mathcal{B}$ (giving (r5)). For each $(\omega, x), (\tilde{\omega}, y) \in I^{\mathbb{N}} \times [0,1)$ write $s\big((\omega, x), (\tilde{\omega}, y)\big)$ for the smallest $n \geq 0$ such that $F^n(\omega, x)$ and $F^n(\tilde{\omega}, y)$ lie in distinct sets $[i] \times A$. The results from [Y99, Theorem 1] then imply, among other things, the following: if there are $C_1 > 0$, $\eta \in (0,1)$ such that for all $[i] \times A$ and all $(\omega, x), (\tilde{\omega}, y) \in [i] \times A$ it holds that

$$\left| \frac{J_{m_{\boldsymbol{p}}\times\lambda} F(\omega, x)}{J_{m_{\boldsymbol{p}}\times\lambda} F(\tilde{\omega}, y)} - 1 \right| \leq C_1 \eta^{s(F(\omega,x), F(\tilde{\omega},y))}, \tag{5.47}$$

then F admits an invariant and ergodic invariant probability measure μ that is absolutely continuous with respect to $m_{\boldsymbol{p}} \times \lambda$ with a density that is bounded away from 0. We will use this to show that each $\{T_i\}_{i\in I} \in \mathcal{D}$ yields a random number system.

For the statement about the CLT-property in Proposition 5.6.8 we apply [Y99, Theorem 4] to φ from (5.46). For this we need to verify that $\int_{I^{\mathbb{N}}\times[0,1)} \varphi \, d\mu = 0$ and that there is a constant $C_2 > 0$ such that

$$|\varphi(\omega, x) - \varphi(\tilde{\omega}, y)| \le C_2 \eta^{s((\omega,x),(\tilde{\omega},y))} \tag{5.48}$$

for all $(\omega, x), (\tilde{\omega}, y) \in I^{\mathbb{N}} \times [0,1)$, where η is the constant from (5.47). Under these conditions [Y99, Theorem 4] states that the sequence $\left(\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ F^i\right)_n$ converges in distribution with respect to μ to a normal distribution with mean 0 and variance σ^2 for some $\sigma > 0$ if and only if $\varphi \circ F \ne \psi \circ F - \psi$ for any measurable function $\psi : I^{\mathbb{N}} \times [0,1) \to \mathbb{R}$.

Proof of Proposition 5.6.8. Let $\{T_i\}_{i\in I} \in \mathcal{D}$, $\{\alpha_i\}_{i\in I}$ the set of partitions from Property (p1), $\boldsymbol{p} = (p_i)_{i\in I}$ be a strictly positive probability vector and $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $I^{\mathbb{N}}$. A suitable invariant measure μ for the skew product F is obtained from [Y99, Theorem 1] once we show that (5.47) holds. First note that (r1), (r2), (r3) follow straightforwardly and (r4), (r5) were already addressed above. By Property (p2) the partition Δ is finite, yielding (r7) once we have μ. Hence, we focus on (r6).

Since each branch of each T_i has non-positive Schwarzian derivative and we have $\inf_{(i,x)} |DT_i(x)| > 1$, it follows from the Koebe Principle, i.e. (1.14) and (1.15), that there exists $K_1, K_2 > 0$ such that for each $\omega \in I^{\mathbb{N}}$, $n \in \mathbb{N}$, $A \in \alpha_{\omega,n}$ and $x, y \in A$ we have

$$\frac{1}{K_1} \le \frac{DT_\omega^n(x)}{DT_\omega^n(y)} \le K_1 \tag{5.49}$$

and

$$\left|\frac{DT_\omega^n(x)}{DT_\omega^n(y)} - 1\right| \le K_2 \cdot \frac{|T_\omega^n(x) - T_\omega^n(y)|}{\lambda(T_\omega^n(A))} = K_2 \cdot |T_\omega^n(x) - T_\omega^n(y)|. \tag{5.50}$$

For each $i \in I$ and $A \in \alpha_i$ we have for all measurable sets $E \subseteq [i]$ and $B \subseteq A$ that

$$m_{\boldsymbol{p}} \times \lambda\big(F(E \times B)\big) = \int_{E\times B} \frac{1}{p_{\omega_1}} DT_{\omega_1}(x) \, dm_{\boldsymbol{p}} \times \lambda(\omega, x),$$

from which it follows that

$$J_{m_{\boldsymbol{p}}\times\lambda} F(\omega, x) = \frac{1}{p_{\omega_1}} DT_{\omega_1}(x). \tag{5.51}$$

Combining this with (5.50) yields for i and A and $(\omega, x), (\tilde{\omega}, y) \in [i] \times A$ that

$$\left|\frac{J_{m_{\boldsymbol{p}}\times\lambda} F(\omega, x)}{J_{m_{\boldsymbol{p}}\times\lambda} F(\tilde{\omega}, y)} - 1\right| = \left|\frac{DT_i(x)}{DT_i(y)} - 1\right| \le K_2 \cdot |T_i(x) - T_i(y)|. \tag{5.52}$$

Assume $s\big(F(\omega, x), F(\tilde{\omega}, y)\big) = n$. Then for each $2 \le k \le n+1$ we have $\omega_k = \tilde{\omega}_k$ and that $T_\omega^{k-1}(x)$ and $T_{\tilde{\omega}}^{k-1}(y) = T_\omega^{k-1}(y)$ are in the same interval of the partition α_{ω_k}, and thus from the Mean Value Theorem and property (p2) that

$$K \le \min |DT_{\omega_k}| \le \left|\frac{T_{\omega_k}(T_\omega^{k-1}(x)) - T_{\omega_k}(T_\omega^{k-1}(y))}{T_\omega^{k-1}(x) - T_\omega^{k-1}(y)}\right| = \left|\frac{T_\omega^k(x) - T_\omega^k(y)}{T_\omega^{k-1}(x) - T_\omega^{k-1}(y)}\right|.$$

We conclude that

$$|T_i(x) - T_i(y)| \le K^{-1}|T_\omega^2(x) - T_\omega^2(y)| \le \cdots \le K^{-n}|T_\omega^{n+1}(x) - T_\omega^{n+1}(y)| \le K^{-n}.$$

Together with (5.52) this shows that (5.47) holds with $\eta = K^{-1}$. Hence, we obtain an F-invariant and F-ergodic measure μ that is equivalent to $m_{\boldsymbol{p}} \times \lambda$. This implies that $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ is a random number system.

What is left is to prove the statement on the CLT-property. Note that

$$\frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} \varphi \circ F^i(\omega, x) = \frac{\log |DT_\omega^n(x)| - nh^{\text{fib}}(\mathcal{T})}{\sqrt{n}}.$$

Since $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i\in I}, \mu, \{\alpha_i\}_{i\in I})$ is a random number system, Theorem 5.1.6(iii) implies that $\int_{I^{\mathbb{N}}\times[0,1)} \varphi\, d\mu = 0$. Assume for a moment that also condition (5.48) holds, i.e. that we satisfy the conditions of [Y99, Theorem 4]. It then follows that there is some $\sigma > 0$ such that

$$\lim_{n\to\infty} \mu\Big(\Big\{(\omega, x) \in I^{\mathbb{N}} \times [0,1) \,:\, \frac{\log |DT_\omega^n(x)| - nh^{\text{fib}}(\mathcal{T})}{\sigma\sqrt{n}} \le u\Big\}\Big) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt$$

if and only if $\varphi \circ F \neq \psi \circ F - \psi$ for any measurable function $\psi : I^{\mathbb{N}} \times [0,1) \to \mathbb{R}$. From (5.49) it follows that for each $\omega \in I^{\mathbb{N}}$, $n \in \mathbb{N}$ and $x \in [0,1)$,

$$\begin{aligned} \lambda(\alpha_{\omega,n}(x)) &\le \frac{1}{\inf_{y\in\alpha_{\omega,n}(x)} |DT_\omega^n(y)|} = \frac{|DT_\omega^n(x)|}{\inf_{y\in\alpha_{\omega,n}(x)} |DT_\omega^n(y)|} \cdot |DT_\omega^n(x)|^{-1} \\ &\le K_1 \cdot |DT_\omega^n(x)|^{-1} \end{aligned}$$

and similarly

$$\lambda(\alpha_{\omega,n}(x)) \ge \frac{1}{\sup_{y\in\alpha_{\omega,n}(x)} |DT_\omega^n(y)|} \ge \frac{1}{K_1} \cdot |DT_\omega^n(x)|^{-1}.$$

Hence, if for each $n \ge 1$ we write

$$X_n(\omega, x) = \frac{-\log \lambda(\alpha_{\omega,n}(x)) - nh^{\text{fib}}(\mathcal{T})}{\sigma\sqrt{n}},$$

then

$$\frac{-\log K_1}{\sigma\sqrt{n}} + \frac{\log |DT_\omega^n(x)| - nh^{\text{fib}}(\mathcal{T})}{\sigma\sqrt{n}} \le X_n(\omega, x) \le \frac{\log K_1}{\sigma\sqrt{n}} + \frac{\log |DT_\omega^n(x)| - nh^{\text{fib}}(\mathcal{T})}{\sigma\sqrt{n}},$$

and we see that to prove the last part of the proposition, it is enough to show that (5.48) with $\eta = K^{-1}$ holds.

Let $(\omega, x), (\tilde\omega, y) \in I^{\mathbb{N}} \times [0,1)$. We first consider the case that $s\big((\omega, x), (\tilde\omega, y)\big) > 0$. Let $i \in I$ and $A \in \alpha_i$ be such that $(\omega, x), (\tilde\omega, y) \in [i] \times A$. It then follows from (5.51) that

$$|\varphi(\omega, x) - \varphi(\tilde\omega, y)| = \Big| \log \Big| \frac{DT_{\omega_1}(x)}{DT_{\tilde\omega_1}(y)} \Big| \Big| = \Big| \log \Big| \frac{J_{m_{\boldsymbol{p}}\times\lambda} F(\omega, x)}{J_{m_{\boldsymbol{p}}\times\lambda} F(\tilde\omega, y)} \Big| \Big|.$$

CHAPTER 5

Recall from the first part of the proof that

$$\Big|\frac{J_{m_{\boldsymbol{p}}\times\lambda}F(\omega,x)}{J_{m_{\boldsymbol{p}}\times\lambda}F(\tilde{\omega},y)}-1\Big|\leq K_2\cdot K^{-s(F(\omega,x),F(\tilde{\omega},y))}.$$

Combining this with the fact that for all $x>0$,

$$|\log x|\leq \max\{x-1,x^{-1}-1\}\leq \max\{|x-1|,|x^{-1}-1|\},$$

yields that

$$\Big|\log\Big|\frac{J_{m_{\boldsymbol{p}}\times\lambda}F(\omega,x)}{J_{m_{\boldsymbol{p}}\times\lambda}F(\tilde{\omega},y)}\Big|\Big|\leq K_2\cdot K^{-s(F(\omega,x),F(\tilde{\omega},y))}\leq \tilde{K}_2\cdot K^{-s((\omega,x),(\tilde{\omega},y))},$$

where $\tilde{K}_2=K_2\cdot K$. In case $s\big((\omega,x),(\tilde{\omega},y)\big)=0$ we just notice from (5.45) that

$$\Big|\log\Big|\frac{DT_{\omega_1}(x)}{DT_{\tilde{\omega}_1}(y)}\Big|\Big|\leq \log M-\log K.$$

By taking $C_2=\max\{\tilde{K}_2,\log M-\log K\}$ we obtain the result. □

Remark 5.6.9. (i) A natural question is whether a Central Limit Theorem would hold when comparing two random number systems. For two random number systems $\mathcal{T}=(I,\mathbb{P},\{T_i\}_{i\in I},\mu,\{\alpha_i\}_{i\in I})$ and $\mathcal{S}=(J,\mathbb{Q},\{S_j\}_{j\in J},\rho,\{\gamma_j\}_{j\in J})$ the limit statement for an annealed result with respect to $\mathcal{S}$ would describe a subset of $I^{\mathbb{N}}\times J^{\mathbb{N}}\times[0,1)$. One might expect that a Central Limit Theorem with measure $\mathbb{P}\times\mathbb{Q}\times\lambda$ holds for random number systems $\mathcal{T}$ and $\mathcal{S}$ with invariant measures $\mu=\mathbb{P}\times\lambda$ and $\rho=\mathbb{Q}\times\lambda$, respectively. For a quenched result with respect to $\mathcal{S}$ the arguments for Theorem 5.6.7 might work if $\mathcal{S}$ satisfies a random zero-property, where in Definition 5.6.4 we replace $\gamma_n(x)$ and $h_{\tilde{\mu}}(S)$ by $\gamma_{\hat{\omega},n}(x)$ and $h^{\text{fib}}(\mathcal{S})$, respectively, and ask for the limit to hold for $\mathbb{Q}$-a.e. $\tilde{\omega}$.

(ii) The Central Limit Theorem from [H09, Corollary 2.1] derived for the quantity $m_{T,S}(n,x)$ for two NTFM's T and S asks for T to satisfy the zero-property and for S to satisfy a property called the weak invariance principle, which seems to be quite strong. By asking that S satisfies the zero-property, we have obtained the Central Limit Theorem under the somewhat less restrictive CLT-property on the random number system $\mathcal{T}$. This implies that the Central Limit Theorem from [H09, Corollary 2.1] also holds under the assumptions that the NTFM T has the CLT-property and S has the zero-property.

§5.7 Examples involving well-known number expansions

Below we consider some specific examples of random number systems with relations to number expansions.

Example 5.7.1 (Random integer base expansions). In this example we generalize the setting of Example 5.1.3. Consider a sequence of integers $2 \leq N_1 < N_2 < N_3 < \dots$. Set $I = \{N_1, N_2, \dots\}$, let $\boldsymbol{p} = (p_i)_{i\in I}$ be a strictly positive probability vector and let $m_{\boldsymbol{p}}$ be the $\boldsymbol{p}$-Bernoulli measure on $I^{\mathbb{N}}$. Assume that

$$\sum_{i\in I} p_i \log^2 i < \infty. \tag{5.53}$$

For each $i \in I$, let $T_i(x) = ix \bmod 1$. The maps T_i are non-singular and piecewise strictly monotonic and C^1 with respect to the partitions $\alpha_i = \{A_{i,j}\}_{j\geq 0}$ given by

$$A_{i,j} = \begin{cases} \left[\frac{j}{i}, \frac{j+1}{i}\right), & \text{if } 0 \leq j \leq i-1, \\ \emptyset, & \text{otherwise.} \end{cases}$$

Thus, conditions (r2) and (r4) are satisfied. The countability of I accounts for (r1) and (r3). The invariance of $m_{\boldsymbol{p}} \times \lambda$ for the skew product F follows directly (using that all maps T_i preserve λ) and its ergodicity follows from standard results, such as e.g. [M85b, Theorem 5.1]. Hence, we get (r6). All maps are expanding, since $DT_i(x) \geq 2$ for all $x \in [0,1)$ and $i \in I$, which implies (r5). Finally, the $\hat{F}$-invariant measure that is obtained by applying Proposition 5.4.1 to $m_{\boldsymbol{p}} \times \lambda$ is $\hat{m}_{\boldsymbol{p}} \times \lambda$, where $\hat{m}_{\boldsymbol{p}}$ is the $\boldsymbol{p}$-Bernoulli measure on $I^{\mathbb{Z}}$. Hence, it follows from (5.53) that condition (5.30) is satisfied (because $(\hat{m}_{\boldsymbol{p}} \times \lambda)_{\hat{\omega}} = \lambda$ for $\hat{m}_{\boldsymbol{p}}$-a.e. $\hat{\omega}$ by Fubini's Theorem), so $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i\in I}, m_{\boldsymbol{p}} \times \lambda, \{\alpha_i\}_{i\in I})$ is a random number system and thus Theorem 5.1.6 applies. Combining Theorem 5.1.6(i), (iii) and (5.53) then gives for $m_{\boldsymbol{p}}$-a.e. ω and λ-a.e. x that

$$\begin{aligned} \lim_{n\to\infty} -\frac{\log \lambda\big(\alpha_{\omega,n}(x)\big)}{n} &= h^{\mathrm{fib}}(\mathcal{T}) = \int_{I^{\mathbb{N}}} \int_{[0,1)} \log |DT_{\omega_1}(x)|\, d\lambda(x) dm_{\boldsymbol{p}}(\omega) \\ &= \int_{I^{\mathbb{N}}} \log \omega_1 dm_{\boldsymbol{p}}(\omega) = \sum_{i\in I} p_i \log i < \infty. \end{aligned}$$

Note that the right-hand side is the weighted sum of the entropies $h_\lambda(T_i) = \log i$.

Note also that the collection $\{T_i\}_{i\in I}$ does not necessarily fall into the set $\mathcal{D}$ from the previous section, since there does not need to be a uniform upper bound on the derivatives of the maps T_i. We show that $\mathcal{T}$ satisfies the CLT-property nonetheless. For each $j \in \mathbb{N}$, define the random variable X_j on $I^{\mathbb{N}} \times [0,1)$ as

$$X_j(\omega, x) = -\sum_{A\in\alpha_{\omega_j}} 1_{(T_\omega^{j-1})^{-1}A}(x) \log \lambda(A) = -\log \lambda\big(\alpha_{\omega_j}(T_\omega^{j-1}(x))\big).$$

Then $\{X_j\}_{j\geq 1}$ is an i.i.d. sequence on $(I^{\mathbb{N}} \times [0,1), \mathcal{B}_I^{\mathbb{N}} \times \mathcal{B}, m_{\boldsymbol{p}} \times \lambda)$. Since each map T_i preserves λ, we obtain

$$\begin{aligned} \mathbb{E}_{m_{\boldsymbol{p}}\times\lambda}(X_j) &= -\int_{I^{\mathbb{N}}} \int_{[0,1)} \log \lambda\big(\alpha_{\omega_j}(T_\omega^{j-1}(x))\big)\, d\lambda(x) dm_{\boldsymbol{p}}(\omega) \\ &= -\int_{I^{\mathbb{N}}} \int_{[0,1)} \log \lambda\big(\alpha_{\omega_j}(x)\big)\, d\lambda(x) dm_{\boldsymbol{p}}(\omega) \\ &= \int_{I^{\mathbb{N}}} \log \omega_j \, dm_{\boldsymbol{p}}(\omega) = h^{\mathrm{fib}}(\mathcal{T}). \end{aligned}$$

Similarly,

$$\sigma^2 = \mathrm{Var}(X_j) = \int_{I^{\mathbb{N}}} (\log \omega_j - h^{\mathrm{fib}}(\mathcal{T}))^2 dm_{\boldsymbol{p}}(\omega) = \sum_{i \in I} p_i \log^2 i - \Big(\sum_{i \in I} p_i \log i\Big)^2. \tag{5.54}$$

It follows from (5.53) that $\sigma^2 \in (0, \infty)$. Also $-\log \lambda(\alpha_{\omega,n}(x)) = \sum_{j=1}^n X_j(\omega, x)$, hence from the Central Limit Theorem we get

$$\frac{-\log \lambda\big(\alpha_{\omega,n}(x)\big) - n h^{\mathrm{fib}}(\mathcal{T})}{\sigma \sqrt{n}} = \frac{\sum_{j=1}^n X_j - n \mathbb{E}_{m_{\boldsymbol{p}} \times \lambda}(X_j)}{\sigma \sqrt{n}} \to \mathcal{N}(0,1),$$

where the convergence is in distribution with respect to $m_{\boldsymbol{p}} \times \lambda$. Hence, $\mathcal{T}$ satisfies the CLT-property with variance σ^2 and with respect to $m_{\boldsymbol{p}} \times \lambda$.

Recall that the digit sequence $(d_n^{\mathcal{T}}(\omega, x))_{n \geq 1}$ was defined in (5.9) by setting $d_n^{\mathcal{T}}(\omega, x) = j_n$ if $T_\omega^{n-1}(x) \in A_{\omega_n, j_n}$. In a similar way as in Example 5.1.3 it can be shown that this yields a number expansion of x given by

$$x = \sum_{n=1}^\infty \frac{d_n^{\mathcal{T}}(\omega, x)}{\prod_{i \in I} i^{c_{n,i}(\omega)}},$$

where $c_{n,i}(\omega) = \#\{1 \leq j \leq n : \omega_j = i\}$. Hence, the random number system $\mathcal{T}$ produces number expansions of numbers $x \in [0,1)$ in mixed integer bases $N_1, N_2, \ldots$.

Consider the random number system $\mathcal{T}$ from above and another random number system of this form for $2 \leq M_1 < M_2 < \ldots$ with $J = \{M_1, M_2, \ldots\}$ and a probability vector $\boldsymbol{q} = (q_j)_{j \in J}$ specifying the Bernoulli measure $m_{\boldsymbol{q}}$ on $J^{\mathbb{N}}$. Assume that $\boldsymbol{q}$ satisfies the condition (5.53) as well. Write $\{S_j\}_{j \in J}$ for the maps $S_j(x) = jx \bmod 1$ and let $\{\gamma_j\}_{j \in J}$ be the corresponding partitions into maximal intervals on which the maps S_j are monotone. For the random number systems $\mathcal{T} = (I^{\mathbb{N}}, m_{\boldsymbol{p}}, \{T_i\}_{i \in I}, m_{\boldsymbol{p}} \times \lambda, \{\alpha_i\}_{i \in I})$ and $\mathcal{S} = (J^{\mathbb{N}}, m_{\boldsymbol{q}}, \{S_j\}_{j \in J}, m_{\boldsymbol{q}} \times \lambda, \{\gamma_j\}_{j \in J})$ we obtain

$$h^{\mathrm{fib}}(\mathcal{T}) = \sum_{i \in I} p_i \log i, \quad h^{\mathrm{fib}}(\mathcal{S}) = \sum_{j \in J} q_j \log j.$$

The Random Lochs' Theorem from Theorem 5.1.4 then states that for $m_{\boldsymbol{p}} \times m_{\boldsymbol{q}}$-a.e. $(\omega, \tilde{\omega}) \in I^{\mathbb{N}} \times J^{\mathbb{N}}$ it holds that

$$\lim_{n \to \infty} \frac{m_{\mathcal{T},\mathcal{S}}(n, \omega, \tilde{\omega}, x)}{n} = \frac{\sum_{i \in I} p_i \log i}{\sum_{j \in J} q_j \log j} \quad \lambda\text{-a.e.}$$

In other words, given $\omega, \tilde{\omega}$ and the first n digits $d_1^{\mathcal{T}}(\omega, x), \ldots, d_n^{\mathcal{T}}(\omega, x)$ of an unknown x, then typically we can determine approximately the first $n \frac{\sum_{i \in I} p_i \log i}{\sum_{j \in J} q_j \log j}$ digits of x in mixed integer bases $M_1, M_2, \ldots$ generated by the random system $\mathcal{S}$.

Moreover, if we take the NTFM $S(x) = Mx \bmod 1$ for some integer $M \geq 2$, then from Theorem 5.1.4 we obtain for $m_{\boldsymbol{p}}$-a.e. $\omega \in I^{\mathbb{N}}$,

$$\lim_{n \to \infty} \frac{m_{\mathcal{T},S}(n, \omega, x)}{n} = \frac{\sum_{i \in I} p_i \log i}{\log M} \quad \lambda\text{-a.e.}$$

From [H09, Section 3.2] we know that S has the zero-property, so that Theorem 5.6.7 gives for each $u \in \mathbb{R}$ that

$$\lim_{n\to\infty} m_{\boldsymbol{p}} \times \lambda\Big(\Big\{(\omega, x) \in I^{\mathbb{N}} \times [0,1) \,:\, \frac{m_{\mathcal{T},S}(n,\omega,x) - n\frac{\sum_{i\in I} p_i \log i}{\log M}}{\frac{\sigma}{\log M}\sqrt{n}} \le u\Big\}\Big)$$
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u} e^{-t^2/2}\, dt,$$

where σ is as in (5.54).

Remark 5.7.2. Arguments almost identical to the ones in Example 5.7.1 hold for any random system consisting of GLS-transformations. A GLS-transformation, see [BBDK96], is a piecewise linear map $T : [0,1) \to [0,1)$ specified by an at most countable interval partition of $[0,1)$ and for each of these intervals an orientation, such that T maps each interval linearly onto $[0,1)$ with the specified orientation. For example, we can fix some $N \in \{2,3,\ldots\} \cup \{\infty\}$ (the number of branches of the map), $\eta \in (0,1)$ (providing a lower bound on the slope of the branches) and $q = (q_j)_{j=0}^{N-1} \in (0,\eta)^N$ with $\sum_{j=0}^{N-1} q_j = 1$ (the sizes of the intervals). The transformation $T_q : [0,1) \to [0,1)$ given by

$$T_q(x) = \sum_{j=0}^{N-1} \frac{1}{q_j}\Big(x - \sum_{k=0}^{j-1} q_k\Big) 1_{\left[\sum_{k=0}^{j-1} q_k, \sum_{k=0}^{j} q_k\right)}(x)$$

is a GLS-transformation mapping each interval $\big[\sum_{k=0}^{j-1} q_k, \sum_{k=0}^{j} q_k\big)$ linearly and orientation preservingly onto $[0,1)$. We can set $I = \{q = (q_j)_{j=0}^{N-1} \in (0,\eta)^N \,:\, \sum_{j=0}^{N-1} q_j = 1\}$, let $\mathbb{P}$ be a τ-invariant probability measure on $I^{\mathbb{N}}$ and consider the family $\{T_q\}_{q\in I}$. (Note that contrary to in Example 5.7.1 this I is not countable.) We assume that

$$-\int_{I^{\mathbb{N}}} \int_{[0,1)} \log \lambda\big(\alpha_{\omega_1}(x)\big) d\lambda(x) d\mathbb{P}(\omega) \in (0,\infty)$$

and that the skew product F associated to $\{T_q\}_{q\in I}$ is ergodic. Then it can in a similar way as in Example 5.7.1 be shown that $\mathcal{T} = (I^{\mathbb{N}}, \mathbb{P}, \{T_q\}_{q\in I}, \mathbb{P}\times\lambda, \{\alpha_q\}_{q\in I})$ is a random number system and that

$$h^{\text{fib}}(\mathcal{T}) = \int_{I^{\mathbb{N}}} h_\lambda(T_{\omega_1})\, d\mathbb{P}(\omega).$$

If we furthermore assume that $\mathbb{P} = \pi^{\mathbb{N}}$ with π a non-trivial probability measure on I and that

$$\int_I \int_{[0,1)} \log^2 \lambda\big(\alpha_i(x)\big)\, d\lambda(x) d\pi(i) < \infty,$$

then $\mathcal{T}$ satisfies the CLT-property with variance

$$\sigma^2 = \int_I \int_{[0,1)} \Big(\log \lambda\big(\alpha_i(x)\big) + h^{\text{fib}}(\mathcal{T})\Big)^2 d\lambda(x) d\pi(i) \in (0,\infty).$$

Number expansions obtained from this system are random versions of what are called *generalised Lüroth series expansions*. A particular instance of this class was studied in [KM22b].

Chapter 5

Remark 5.7.3. The family of GLS-transformations from Remark 5.7.2 provides examples of random number systems that do not satisfy (s1) from Section 5.3 by having an uncountable index set. We can also construct from these transformations random number systems that satisfy (s1) but not (s2). As an easy example, set $I = \{2, 3, \ldots\}$, define the probability vector $\boldsymbol{p} = (p_i)_{i \in I}$ by

$$p_i = C \cdot \frac{1}{i \cdot (\log i)^2}$$

with normalisation $C^{-1} = \sum_{n=2}^{\infty} \frac{1}{n \cdot (\log n)^2}$, and again write $m_{\boldsymbol{p}}$ for the $\boldsymbol{p}$-Bernoulli measure on $I^{\mathbb{N}}$. Furthermore, for each $i \in I$, let

$$T_i : [0,1) \to [0,1), \qquad T_i(x) = \begin{cases} \frac{1}{\frac{1}{2} - \frac{1}{i+1}} x, & \text{if } x \in [0, \frac{1}{2} - \frac{1}{i+1}), \\ \frac{1}{\frac{1}{2} + \frac{1}{i+1}} \left(x - \frac{1}{2} + \frac{1}{i+1}\right), & \text{if } x \in [\frac{1}{2} - \frac{1}{i+1}, 1) \end{cases}$$

and $\alpha_i = \{A_{i,0}, A_{i,1}\}$ with $A_{i,0} = [0, \frac{1}{2} - \frac{1}{i+1})$ and $A_{i,1} = [\frac{1}{2} - \frac{1}{i+1}, 1)$. Then $\mathcal{T} = (I^{\mathbb{N}}, m_{\boldsymbol{p}}, \{T_i\}_{i \in I}, m_{\boldsymbol{p}} \times \lambda, \{\alpha_i\}_{i \in I})$ is a random number system that meets (r1)-(r7) and (s1). However, (s2) is not satisfied, since the entropy $H_{m_{\boldsymbol{p}}}(\iota)$ of the partition $\iota = \{[i] : i \in I\}$ satisfies

$$H_{m_{\boldsymbol{p}}}(\iota) = -\sum_{i \in I} p_i \log p_i = C \cdot \sum_{i=2}^{\infty} \frac{1}{i \cdot (\log i)^2} \cdot \log(i \cdot (\log i)^2) - \tilde{C}$$

with $\tilde{C} = C \cdot \log(C) \cdot \sum_{i=2}^{\infty} \frac{1}{i \cdot (\log i)^2} = \log(C) \in \mathbb{R}$, and thus

$$H_{m_{\boldsymbol{p}}}(\iota) \geq C \cdot \sum_{i=2}^{\infty} \frac{1}{i \cdot (\log i)^2} \cdot \log i - \tilde{C} = C \cdot \sum_{i=2}^{\infty} \frac{1}{i \cdot \log i} - \tilde{C} = \infty.$$

This indicates that the class of random systems is considerably bigger than the class of random number systems that meets the additional conditions (s1)-(s4).

Example 5.7.4 (Random continued fraction expansions). Let $(\{T_0, T_1\}, m_{\boldsymbol{p}})$ be the random Gauss-Rényi map from Example 1.4.2, where $T_0 = G$ is the Gauss map from Example 1.3.2, $T_1 = R$ is the Rényi map from Example 1.3.3 and $\boldsymbol{p} = \{p_0, p_1\}$ with $p_0 = p \in (0,1)$ and $p_1 = 1 - p$ gives the probabilities of choosing G and R, respectively. The partitions $\alpha_0 = \{A_{0,j}\}_{j \geq 0}$ and $\alpha_1 = \{A_{1,j}\}_{j \geq 0}$ are given by

$$A_{0,j} = \left(\frac{1}{j+2}, \frac{1}{j+1}\right] \quad \text{and} \quad A_{1,j} = \left[\frac{j}{j+1}, \frac{j+1}{j+2}\right), \quad j \geq 0.$$

In [KKV17] it was proven that there exists a measure ρ_p equivalent to λ such that $m_{\boldsymbol{p}} \times \rho_p$ is invariant and ergodic with respect to the skew product F with a density $\frac{d\rho_p}{d\lambda}$ that is bounded away from zero and is of bounded variation, so in particular bounded. For the system $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i=0,1}, m_{\boldsymbol{p}} \times \rho_p, \{\alpha_i\}_{i=0,1})$ conditions (r1)–(r6) follow straightforwardly. For (r7), let $M > 1$ be such that $M^{-1} \leq \frac{d\rho_p}{d\lambda} \leq M$ and note that

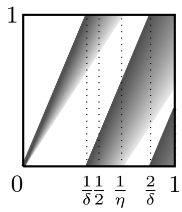
$$H_{m_{\boldsymbol{p}} \times \rho_p}(\Delta) \leq M \log M + M \sum_{j \geq 0} \frac{\log((j+1)(j+2))}{(j+1)(j+2)}.$$

It follows that $H_{m_{\boldsymbol{p}}\times\rho_p}(\Delta) < \infty$. Thus (r7) holds and $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i=0,1}, m_{\boldsymbol{p}} \times \rho_p, \{\alpha_i\}_{i=0,1})$ is a random number system. For the fiber entropy we obtain from Theorem 5.1.6(iii) that

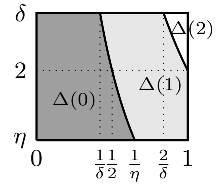
$$\begin{aligned} h^{\mathrm{fib}}(\mathcal{T}) &= \int_{I^{\mathbb{N}}} \int_{[0,1)} \log|DT_{\omega_1}(x)|\, d\rho_p(x) dm_{\boldsymbol{p}}(\omega) \\ &= -2\int_0^1 p\log x + (1-p)\log(1-x)\, d\rho_p(x). \end{aligned}$$

It follows from equation (8) in [KKV17] that the digit sequences $(d_n^{\mathcal{T}}(\omega,x))_{n\geq 1}$ from this random number system give the semi-regular continued fraction expansions of real numbers $x \in [0,1]$ by

$$x = \omega_1 + \cfrac{(-1)^{\omega_1}}{d_1^{\mathcal{T}}(\omega,x) + 1 + \omega_2 + \cfrac{(-1)^{\omega_2}}{d_2^{\mathcal{T}}(\omega,x) + 1 + \omega_3 + \ddots}}.$$



(a) S_β for all $\eta \leq \beta \leq \delta$.



(b) The partition Δ.

Figure 5.3: In (a) we see the graphs of all the maps S_β for $\beta \in [\eta, \delta]$ from Example 5.7.5 for some $1 < \eta < 2 < \delta < 3$. Each shade of grey corresponds to one graph. In (b) we see the elements of the partition Δ for values η and δ as in (a).

Example 5.7.5. (Random β-expansions in alternate base) Fix two constants $1 < \eta < \delta$ and let $J = [\eta, \delta]$. For each $\beta \in J$, let $S_\beta : [0,1) \to [0,1)$, $x \mapsto \beta x \bmod 1$ be the β-transformation, which is the piecewise linear map with slope β on the partition $\gamma_\beta = \{C_{\beta,j}\}_{j\geq 0}$ given by

$$C_{\beta,j} = \begin{cases} \left[\frac{j}{\beta}, \frac{j+1}{\beta}\right), & \text{if } 0 \leq j < \lceil\beta\rceil - 1 \\ \left[\frac{\lceil\beta\rceil-1}{\beta}, 1\right), & \text{if } j = \lceil\beta\rceil - 1, \\ \emptyset, & \text{otherwise,} \end{cases}$$

where $\lceil\beta\rceil$ indicates the smallest integer not smaller than β. See Figure 5.3(a) for some graphs. The study of β-transformations was initiated by Rényi in [R57b]. The β-transformations are related to β-expansions of real numbers, which are expressions of the form

$$x = \sum_{n\geq 1} \frac{b_n}{\beta^n}, \quad b_n \in \{0, 1, \dots, \lceil\beta\rceil - 1\}.$$

Fix some periodic sequence

$$u = (u_1, u_2, \dots, u_m, u_1, u_2, \dots, u_m, u_1, \dots) \in J^{\mathbb{N}}$$

of period length $m \geq 2$ and consider the shift invariant measure $\mathbb{Q}$ on $J^{\mathbb{N}}$ defined by

$$\mathbb{Q} = \frac{1}{m} \sum_{i=1}^{m} \delta_{\tau^{i-1} u}, \tag{5.55}$$

where δ_y denotes the Dirac measure at the point $y \in J^{\mathbb{N}}$. We define the map $\psi : \{1, 2, \dots, m\} \times [0, 1) \to J^{\mathbb{N}} \times [0, 1)$ by

$$\psi(i, x) = \big((u_i, u_{i+1}, \dots, u_m, u_1, u_2, \dots, u_m, u_1, \dots), x\big),$$

which is measurable if we put on $\{1, 2, \dots, m\} \times [0, 1)$ the σ-algebra

$$\mathcal{A} = \Big\{ \bigcup_{i=1}^{m} \{i\} \times B_i \, : \, B_1, \dots, B_m \in \mathcal{B} \Big\}.$$

Defining the transformation $T_u : \{1, 2, \dots, m\} \times [0, 1) \to \{1, 2, \dots, m\} \times [0, 1)$ by

$$T_u(i, x) = \big((i + 1) \bmod m, T_{u_i}(x)\big),$$

it follows from [CCD21] that there are probability measures $\mu_{u,1}, \dots, \mu_{u,m}$ on $([0, 1), \mathcal{B})$ such that the probability measure μ_u on $(\{1, \dots, m\} \times [0, 1), \mathcal{A})$ given by

$$\mu_u\Big(\bigcup_{i=1}^{m} \{i\} \times B_i\Big) = \frac{1}{m} \sum_{i=1}^{m} \mu_{u,i}(B_i), \qquad B_1, \dots, B_m \in \mathcal{B},$$

is an ergodic invariant measure for T_u that is equivalent to the probability measure λ_u on $(\{1, \dots, m\} \times [0, 1), \mathcal{A})$ given by

$$\lambda_u\Big(\bigcup_{i=1}^{m} \{i\} \times B_i\Big) = \frac{1}{m} \sum_{i=1}^{m} \lambda(B_i), \qquad B_1, \dots, B_m \in \mathcal{B}.$$

Define the probability measure ρ on $J^{\mathbb{N}} \times [0, 1)$ by

$$\rho = \frac{1}{m} \sum_{i=1}^{m} \delta_{\tau^{i-1} u} \times \mu_{u,i}.$$

Since $\mathbb{Q} \times \lambda = \lambda_u \circ \psi^{-1}$ and ψ is an isomorphism between the dynamical systems $(\{1, \dots, m\} \times [0, 1), \mathcal{A}, \mu_u, T_u)$ and $(J^{\mathbb{N}} \times [0, 1), \mathcal{B}_J^{\mathbb{N}} \times \mathcal{B}, \rho, F)$ with F the skew product associated to $\{S_\beta\}_{\beta \in J}$, it follows that ρ is an ergodic invariant measure for F and that ρ is equivalent to $\mathbb{Q} \times \lambda$. Since δ is an upper bound for J, the collection Δ is finite and thus $H_\rho(\Delta) < \infty$. Figure 5.3(b) illustrates the sets $\Delta(j)$ for $\eta \in (1, 2)$ and $\delta \in (2, 3)$. It follows that $\mathcal{S} = (J, \mathbb{Q}, \{S_\beta\}_{\beta \in J}, \rho, \{\gamma_\beta\}_{\beta \in J})$ is a random number system. From Theorem 5.1.6(iii) we get

$$h^{\text{fib}}(\mathcal{S}) = \int_{J^{\mathbb{N}}} \int_{[0,1)} \log \omega_1 \, d\rho(\omega, x) = \frac{1}{m} \sum_{i=1}^{m} \log u_i.$$

By the definition of the digit sequence $(d_n^{\mathcal{S}}(\omega, x))_{n\geq 1}$ we can write for each $\omega \in J^{\mathbb{N}}$, $x \in [0,1)$ and $n \geq 1$ that

$$S_\omega^n(x) = \omega_n S_\omega^{n-1}(x) - d_n^{\mathcal{S}}(\omega, x),$$

so that

$$x = \frac{d_1^{\mathcal{S}}(\omega, x)}{\omega_1} + \frac{d_2^{\mathcal{S}}(\omega, x)}{\omega_1\omega_2} + \cdots + \frac{d_n^{\mathcal{S}}(\omega, x)}{\omega_1 \cdots \omega_n} + \frac{S_\omega^n(x)}{\omega_1 \cdots \omega_n}.$$

Since $\lim_{n\to\infty} \frac{S_\omega^n(x)}{\omega_1\cdots\omega_n} \leq \lim_{n\to\infty} \frac{1}{\eta^n} = 0$, for each $x \in [0,1)$ and $\omega \in J^{\mathbb{N}}$ we obtain the *random mixed β-expansion*

$$x = \sum_{n\geq 1} \frac{d_n^{\mathcal{S}}(\omega, x)}{\omega_1 \cdots \omega_n}.$$

With our choice of $\mathbb{Q}$ from (5.55) it holds that for $\mathbb{Q}$-a.e. $\omega \in J^{\mathbb{N}}$ with $\omega_1 = u_1$ the random mixed β-expansions produced by the system $\mathcal{S}$ are the *greedy* $(u_1, \ldots, u_m)$-*expansions in alternate base* that are the object of study in [CCD21].

With Theorem 5.1.4 we can compare the semi-regular continued fraction digits from the random continued fraction map from Example 5.7.4 with the alternate base greedy β-expansions. If we let $\mathcal{T} = (I, m_{\boldsymbol{p}}, \{T_i\}_{i=0,1}, m_{\boldsymbol{p}} \times \rho_p, \{\alpha_i\}_{i=0,1})$ be the system from Example 5.7.4 and let $(J, \mathbb{Q}, \{S_\beta\}_{\beta\in J}, \rho, \{\gamma_\beta\}_{\beta\in J})$ be the system from Example 5.7.5, then Theorem 5.1.4 tells us that for $m_{\boldsymbol{p}} \times \mathbb{Q}$-a.e. $(\omega, \tilde{\omega}) \in I^{\mathbb{N}} \times J^{\mathbb{N}}$,

$$\lim_{n\to\infty} \frac{m_{\mathcal{T},\mathcal{S}}(n, \omega, \tilde{\omega}, x)}{n} = \frac{-2\int_{[0,1)} p \log x + (1-p)\log(1-x)\, d\rho_p(x)}{\frac{1}{m}\sum_{i=1}^m \log u_i} \quad \lambda\text{-a.e.}$$

CHAPTER 6

A Lochs Theorem for pseudo-random number generation with β-encoders

This chapter is joint work with Charlene Kalle and Evgeny Verbitskiy.

Abstract

The β-encoder is an analog circuit that converts an input signal $x \in [0, 1)$ into a finite bitstream $b_1, \ldots, b_m$ that corresponds to a representation of x in non-integer base $\beta \in (1, 2)$. In this chapter we study a question posed by Jitsumatsu and Matsumura on the number of output digits from the β-encoder that are necessary to correctly determine the first n base 2 digits of the original input x. We confirm the lower bound established by Jitsumatsu and Matsumura, we provide an upper bound and give two different limit results for this value, the last one of which is reminiscent of Lochs' Theorem.

§6.1 Introduction

Since the work [DDGV02] from 2002 by Daubechies et al. the advantages and disadvantages of the β-encoder as a replacement for the commonly used Pulse Code Modulation (PCM) in analog-to-digital (A/D) conversion have been considered. Given a real number $\beta \in (1, 2)$, the scale-adjusted β-encoder converts an analog input signal $x = x_0 \in [0, 1)$ into a bitstream $b_1, \ldots, b_k$ of specified length k by using a circuit consisting of an *amplifier* with amplification factor β, a *scale adjuster* with scaling factor $\beta - 1$, and a *quantiser*

$$Q_u(y) = \begin{cases} 0, & \text{if } y < u, \\ 1, & \text{if } y \geq u, \end{cases}$$

with threshold value $u \in [\beta - 1, 1]$. The bits b_n are produced iteratively by setting $x_n = \beta x_{n-1} - (\beta - 1)b_n$ and $b_n = Q_u(\beta x_{n-1})$. This algorithm is depicted in Figure 6.1 and is set up in such a way that it provides a β-*expansion* of the number $\frac{x}{\beta-1}$, i.e. x can be represented as

$$x = (\beta - 1) \sum_{n=1}^{\infty} \frac{b_n}{\beta^n},$$

and thus finite truncations $b_1, \ldots, b_k$ of the sequence $(b_n)_{n \geq 1}$ give bitstreams that approximate x well. The PCM works similarly but with multiplication factor 2 and threshold value 1 (and no scale adjuster).

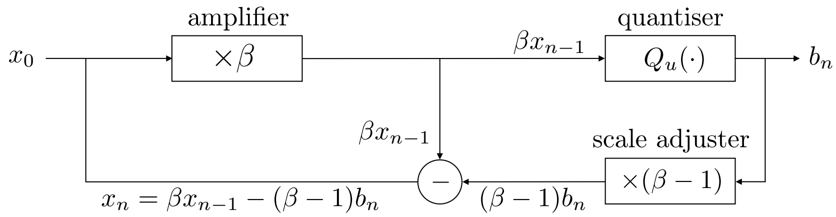


Figure 6.1: Iteration process of the scale-adjusted β-encoder.

Due to noise in the circuit the value of u fluctuates while running the iteration process. Since each number x has an essentially unique base 2 representation, PCM is not robust against these quantisation errors, see e.g. [DDGV06, G12]. It was proven in [EJK90, S03] that, contrary to binary expansions, for each $\beta \in (1, 2)$ Lebesgue almost every x has uncountably many different β-expansions. The advantage of using a β-encoder over PCM for A/D-conversion thus lies in the fact that if the quantisation threshold u is chosen well inside the interval $[\beta - 1, 1]$, then the β-encoder can recover from a quantisation error and still produce good approximations of the original input signal x by shifting from one β-expansion of x to another. The robustness of the β-encoder in the A/D-conversion process has been studied in e.g. [DDGV06, JW09, KHTA12, KHA12, SKM$^+$13, MIS$^+$15].

In recent years the β-encoder was considered as a source for random number generation, see [JMKA13, SJO15, JM16, KJ16]. An issue with this application was that the successive bits in the output of a β-encoder are strongly correlated. Jitsumatsu and Matsumura proposed in [JM16] to remove this dependence between bits by adding an algorithm to the process that converts the output bits from the β-encoder into the base 2 digits of the number it represents. A natural question asked in [JM16] is the following: If we use $\boldsymbol{u} = (u_n)_{n\geq 1}$ to denote the consecutive threshold values u_n used at each time step of the approximation algorithm, what is the number $k(m, \boldsymbol{u}, x)$ of bits from the β-encoder that are necessary to obtain m base 2 digits of the number x via this process? In [JM16] the lower bound $k(m, \boldsymbol{u}, x) \geq \frac{m \log 2}{\log \beta}$ was found. The authors of [JM16] remarked that a theoretical analysis of the expected value of $k(m, \boldsymbol{u}, x)$ is relevant as an indication of the efficiency of the proposed pseudo-random number generator.

It is the purpose of this chapter to address the question posed in [JM16], which is reminiscent of the considerations of Lochs in [L64] and related settings from [BDK99, DF01] discussed in Subsection 5.1.1. Unfortunately these results as well as the results obtained in Chapter 5 do not immediately apply to the question from [JM16] due to the uncertainty in the threshold value u, even though as we will see below the iteration process of the β-encoder can be described with a random dynamical system. In particular, the results of Chapter 5, which are in the context of random dynamical systems, do not apply because fluctuations in u are assumed to be unknown. (This is discussed at the end of Subsection 5.1.2.) Indeed, in reality we do not know, given some output $b_1, \dots, b_k$ of the β-encoder, which $u_1, \dots, u_k$ resulted in this output. Nevertheless, in our first main result of this chapter we recover the lower bound from [JM16] and we obtain a statement on an upper bound for $k(m, \boldsymbol{u}, x)$. More precisely, we obtain the following results. Again λ denotes the one-dimensional Lebesgue measure.

Theorem 6.1.1. *Let $\beta \in (1, 2)$ and $\boldsymbol{u} = (u_n)_{n\geq 1} \in [\beta - 1, 1]^{\mathbb{N}}$. For all $x \in [0, 1)$ and all $m \in \mathbb{N}$ it holds that*

$$k(m, \boldsymbol{u}, x) \geq \frac{m \log 2}{\log \beta}. \tag{6.1}$$

Moreover, for each $\varepsilon \in (0, 1)$ there exists a constant $C(\varepsilon) > 0$ such that for all $m \in \mathbb{N}$

$$\lambda\Big(\Big\{x \in [0, 1) \, : \, k(m, \boldsymbol{u}, x) - \frac{m \log 2}{\log \beta} > C(\varepsilon)\Big\}\Big) < \varepsilon. \tag{6.2}$$

From these bounds we obtain the following corollary on the asymptotic behaviour of the sequences $(k(m, \boldsymbol{u}, x))_{m\geq 1}$.

Corollary 6.1.2. *For any real positive sequence $(a_m)_{m\in\mathbb{N}}$ with $\lim_{m\to\infty} a_m = \infty$, each $\boldsymbol{u} \in [\beta - 1, 1]^{\mathbb{N}}$ and each $\varepsilon > 0$ it holds that*

$$\lim_{m\to\infty} \lambda\Big(\Big\{x \in [0, 1) \, : \, \frac{1}{a_m}\Big|k(m, \boldsymbol{u}, x) - \frac{m \log 2}{\log \beta}\Big| > \varepsilon\Big\}\Big) = 0,$$

i.e. the sequence $\big(\frac{1}{a_m}\big(k(m, \boldsymbol{u}, x) - \frac{m \log 2}{\log \beta}\big)\big)_{m\geq 1}$ converges to 0 in λ-probability.

In particular, the above corollary has the following implications:
- Taking $a_m = \sqrt{m}$ for each m gives a Central Limit Theorem result where the limiting distribution has zero variance;
- Taking $a_m = m$ for each m we retrieve a limit statement in the spirit of (5.1), but with convergence in probability instead of almost surely.

By adjusting the setup from [DF01] to suit our purposes, we obtain the stronger result of almost sure convergence for the specific sequence $(a_m)_{m\geq 1}$ with $a_m = m$ for each m that is stated in the next theorem.

Theorem 6.1.3. *For each $\boldsymbol{u} \in [\beta - 1, 1]^{\mathbb{N}}$ it holds that*

$$\lim_{m\to\infty} \frac{k(m, \boldsymbol{u}, x)}{m} = \frac{\log 2}{\log \beta} \qquad \lambda\text{-}a.e.$$

More specifically, for typical x and large m one needs approximately $\frac{m \log 2}{\log \beta}$ output bits of the β-encoder to obtain m correct base 2 digits.

The remainder of this chapter is organised as follows. In the next section we introduce the necessary notation and preliminaries on binary and β-expansions. In the third section we prove our main results. We conclude with some final remarks. Here we discuss in particular what happens if not only the threshold value u but also the values of the amplification factor β or the scaling factor $\beta - 1$ fluctuate over time, a fact that has been observed for β-encoders and discussed in e.g. [DY06, W08, DGWY10].

§6.2 Preliminaries

If A is an interval in the real line, then we write ∂A for the set containing the two boundary points of A and we use A^- and A^+ to denote the lower and upper endpoint of A, respectively.

For each m the collection of *dyadic intervals of order* m is given by

$$\mathcal{D}_m = \Big\{ \Big[\frac{k}{2^m}, \frac{k+1}{2^m}\Big) \, : \, 0 \leq k \leq m-1 \Big\}.$$

If we write the point $\frac{k}{2^m} = \sum_{i=1}^{m} \frac{d_i}{2^i}$, $d_i \in \{0,1\}$, in its binary expansion, then we see that the interval $\big[\frac{k}{2^m}, \frac{k+1}{2^m}\big)$ contains precisely those $x \in [0,1)$ that have $d_1, \ldots, d_m$ as their first m base 2 digits. For each $x \in [0,1)$ and each $m \geq 1$ there is a unique element of $\mathcal{D}_m$ that contains x. We denote this by $\mathcal{D}_m(x)$. Then

$$\lambda(\mathcal{D}_m(x)) = 2^{-m}. \tag{6.3}$$

Hence, each collection $\mathcal{D}_m$ is a *partition* of $[0,1)$ by intervals of length 2^{-m}.

Usually A/D-converters rely on binary expansions of numbers to produce good approximations of the input signal. The β-encoder is based on β-expansions instead. Fix a value of $\beta \in (1,2)$. An expression of the form

$$x = \sum_{i\geq 1} \frac{b_i}{\beta^i}, \quad b_i \in \{0,1\},$$

is called a β-expansion of x, see also Example 5.7.5. One easily sees that if x has such a β-expansion, then $x \in \left[0, \frac{1}{\beta-1}\right]$. The β-encoder as described in [JM16] considers as input signal a number $x \in [0,1)$ and thus has rescaled the setup by a factor $\beta - 1$. Below we briefly explain how one can get a β-expansion of a number $\frac{x}{\beta-1}$ for $x \in [0,1)$ from the β-encoder given in the introduction but with varying threshold values u_n.

For each $u \in [\beta - 1, 1]$ define the interval map $T_u : [0,1) \to [0,1)$ by

$$T_u(y) = \begin{cases} \beta y, & \text{if } y < \frac{u}{\beta}, \\ \beta y - (\beta - 1), & \text{if } y \geq \frac{u}{\beta}. \end{cases}$$

The graph of such a map is shown in Figure 6.2. If we let u_n denote the threshold value of the quantiser at time n, then the dynamics of the β-encoder can be represented by

$$x_n = T_{u_n}(x_{n-1}) = T_{u_n} \circ \cdots \circ T_{u_1}(x), \quad n \geq 1.$$

For each $n \geq 1$, set $b_n = b_n(x) = 0$ if $\beta x_{n-1} < u_n$ and 1 otherwise. Putting $x_0 = x$, then for each $n \geq 1$,

$$T_{u_n}(x_{n-1}) = \beta x_{n-1} - (\beta - 1) b_n,$$

so that

$$x = (\beta - 1) \sum_{i=1}^{n} \frac{b_i}{\beta^i} + \frac{T_{u_n} \circ \cdots \circ T_{u_1}(x)}{\beta^n}.$$

Since $T_{u_n} \circ \cdots \circ T_{u_1}(x) \in [0,1)$ for each n, we obtain that $x = (\beta - 1) \sum_{i=1}^{\infty} \frac{b_i}{\beta^i}$. From Figure 6.2 it becomes clear that each threshold value u_n must lie in the interval $[\beta - 1, 1]$ to obtain a recursive process and bits that correspond to β-expansions.

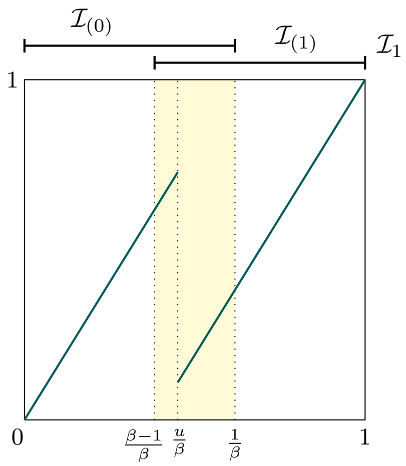


Figure 6.2: The graph of one of the maps T_u is shown for $\beta = \frac{1+\sqrt{5}}{2}$, the golden mean. The yellow area in the middle relates to the interval in which the threshold value u may be chosen. At the top we see the two intervals $\mathcal{I}_{(0)}$ and $\mathcal{I}_{(1)}$ that are the elements of the cover $\mathcal{I}_1$.

Chapter 6

Given the first k output bits $b_1, \dots, b_k$ of the β-encoder, we know that the input signal x has to satisfy

$$\frac{x}{\beta-1} \in \Big[\sum_{i=1}^{k} \frac{b_i}{\beta^i}, \sum_{i=1}^{k} \frac{b_i}{\beta^i} + \sum_{i\geq k+1} \frac{1}{\beta^i}\Big] = \Big[\sum_{i=1}^{k} \frac{b_i}{\beta^i}, \sum_{i=1}^{k} \frac{b_i}{\beta^i} + \frac{1}{\beta^k(\beta-1)}\Big].$$

For each $b_1, \dots, b_k \in \{0,1\}$ define

$$\mathcal{I}_{(b_1,\dots,b_k)} = \Big[(\beta-1)\sum_{i=1}^{k} \frac{b_i}{\beta^i}, (\beta-1)\sum_{i=1}^{k} \frac{b_i}{\beta^i} + \frac{1}{\beta^k}\Big].$$

Comparable to the partitions $\mathcal{D}_m$ for binary expansions, we consider for each $k \geq 1$ the *cover* $\mathcal{I}_k$ of $[0,1)$ associated to β-expansions given by

$$\mathcal{I}_k = \{\mathcal{I}_{(b_1,\dots,b_k)} \,:\, b_i \in \{0,1\},\, 1 \leq i \leq k\}.$$

See Figure 6.2 for an illustration of $\mathcal{I}_1 = \{\mathcal{I}_{(0)}, \mathcal{I}_{(1)}\}$.

If for $k \geq 1$ the first k output bits of the β-encoder for an input signal $x \in [0,1)$ and a threshold value sequence $\boldsymbol{u} \in [\beta-1,1]^{\mathbb{N}}$ are $b_1, \dots, b_k$, then we set

$$\mathcal{I}_k(\boldsymbol{u}, x) = \mathcal{I}_{(b_1,\dots,b_k)},$$

since the information that the bits $b_1, \dots, b_k$ give us is that x is contained in this interval.[1] Note that

$$\lambda(\mathcal{I}_k(\boldsymbol{u}, x)) = \beta^{-k}. \tag{6.4}$$

Furthermore,

$$k(m, \boldsymbol{u}, x) = \inf\{k \geq 1 \,:\, \mathcal{I}_k(\boldsymbol{u}, x) \subseteq \mathcal{D}_m(x)\}. \tag{6.5}$$

§6.3 Proofs of the main results

In this section we prove the main results. We start with the proof of Theorem 6.1.1, which provides bounds for the quantities $k(m, \boldsymbol{u}, x)$. This proof is inspired by the proof of [H09, Theorem 2.3].

[1] This is the setting of (5.11). Indeed, defining $A_{u,0} = [0, \frac{u}{\beta})$, $A_{u,1} = [\frac{u}{\beta}, 1)$ for each $u \in [\beta-1, 1]$ and setting $\chi(0) = 1$,$\chi(1) = \beta - 1$ and $I = [\beta-1, 1]$, note that

$$\mathcal{I}_{(b_1,\dots,b_k)} = \bigcap_{i=1}^{k} T^{-1}_{\chi(b_1)\cdots\chi(b_{i-1})} A_{\chi(b_i),b_i} = \bigcup_{(\omega_1,\dots,\omega_k)\in I^k} \bigcap_{i=1}^{k} T^{-1}_{\omega_1\cdots\omega_{i-1}} A_{\omega_i,b_i}.$$

Proof of Theorem 6.1.1. Fix $\boldsymbol{u} = (u_n)_{n\geq 1} \in [\beta-1,1]^{\mathbb{N}}$. For all $m \in \mathbb{N}$ and $x \in [0,1)$ we find using (6.3) and (6.4) that

$$\begin{aligned} k(m,\boldsymbol{u},x) - \frac{m\log 2}{\log\beta} &= \frac{k(m,\boldsymbol{u},x)\log\beta + \log\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)}(\boldsymbol{u},x))}{\log\beta} \\ &\quad + \frac{-\log\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)}(\boldsymbol{u},x)) + \log\lambda(\mathcal{D}_m(x))}{\log\beta} \\ &\quad + \frac{-\log\lambda(\mathcal{D}_m(x)) - m\log 2}{\log\beta} \\ &= \frac{1}{\log\beta}\cdot\log\left(\frac{\lambda(\mathcal{D}_m(x))}{\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)}(\boldsymbol{u},x))}\right). \end{aligned} \tag{6.6}$$

By the definition of $k(m,\boldsymbol{u},x)$ we have $\mathcal{I}_{k(m,\boldsymbol{u},x)}(\boldsymbol{u},x) \subseteq \mathcal{D}_m(x)$ and thus the above yields

$$k(m,\boldsymbol{u},x) \geq \frac{m\log 2}{\log\beta}.$$

This gives (6.1).

For (6.2) let $\varepsilon \in (0,1)$ and fix some integer $m \geq 1$. By the definition of $k(m,\boldsymbol{u},x)$ we have that $\mathcal{I}_{k(m,\boldsymbol{u},x)-1}(\boldsymbol{u},x) \not\subseteq \mathcal{D}_m(x)$. Hence, the distance between x and the nearest boundary point of $\mathcal{D}_m(x)$, denoted by $|x - \partial\mathcal{D}_m(x)|$, is at most equal to $\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)-1}(\boldsymbol{u},x))$. Furthermore, we have

$$\log\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)-1}(\boldsymbol{u},x)) - \log\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)}(\boldsymbol{u},x)) = \log\beta.$$

Together this gives that

$$\log\left(\frac{\lambda(\mathcal{D}_m(x))}{\lambda(\mathcal{I}_{k(m,\boldsymbol{u},x)}(x))}\right) \leq \log\lambda\big(\mathcal{D}_m(x)\big) + \log\beta - \log|x - \partial\mathcal{D}_m(x)|. \tag{6.7}$$

We slightly adjust the intervals in $\mathcal{D}_m$ by removing small intervals at the endpoints: For each $m \in \mathbb{N}$ and interval $J \in \mathcal{D}_m$, let J' be the interval obtained by removing on both ends of J an interval of length $\frac{\varepsilon}{6}\cdot 2^{-m}$ and let $C_m = \bigcup_{J\in\mathcal{D}_m} J'$. Then $\lambda(J') = \big(1-\frac{\varepsilon}{3}\big)\cdot 2^{-m}$ and $\lambda(C_m) = 1 - \frac{\varepsilon}{3}$. For $x \in C_m$ we have the bound $|x - \partial\mathcal{D}_m(x)| \geq \frac{\varepsilon}{6}\lambda(\mathcal{D}_m(x))$. Combining this with (6.6) and (6.7) gives for each integer $m \in \mathbb{N}$ and each $x \in C_m$ that

$$k(m,\boldsymbol{u},x) - \frac{m\log 2}{\log\beta} \leq \frac{\log\frac{6}{\varepsilon}}{\log\beta} + 1.$$

Hence, we obtain (6.2) with constant $C(\varepsilon) = \frac{\log\frac{6}{\varepsilon}}{\log\beta} + 1$. □

Theorem 6.1.1 gives bounds on the value of $k(m,\boldsymbol{u},x)$ and immediately leads to the statement on the asymptotics of the sequence $(k(m,\boldsymbol{u},x))_{m\geq 1}$ from Corollary 6.1.2 that we prove next.

Proof of Corollary 6.1.2. Let $(a_m)_{m\geq 1}$ be a sequence that satisfies the conditions of the corollary. From (6.1) we get that for each $x \in [0,1)$ and $m \in \mathbb{N}$,

$$\frac{1}{a_m}\left(k(m,\boldsymbol{u},x) - \frac{m\log 2}{\log\beta}\right) \geq 0.$$

Hence, it suffices to show that for all $\delta, \varepsilon > 0$ there exists an $M \in \mathbb{N}$ such that for all $m \geq M$ we have

$$\lambda\Big(\Big\{x \in [0,1) \,:\, \frac{1}{a_m}\Big(k(m,\boldsymbol{u},x) - \frac{m \log 2}{\log \beta}\Big) > \delta\Big\}\Big) < \varepsilon.$$

This immediately follows from (6.2) by taking $M \in \mathbb{N}$ big enough such that $\frac{C(\varepsilon)}{a_m} \leq \delta$ for all $m \geq M$, which is possible because $\lim_{m\to\infty} a_m = \infty$. □

As we saw in the introduction, by choosing $a_m = m$ for all $m \geq 1$ Corollary 6.1.2 gives a limit statement reminiscent of Lochs' Theorem, but with convergence in probability. Our final result, Theorem 6.1.3 which we prove next, shows that this limit statement also holds almost surely. This proof is inspired by the proof of [DF01, Theorem 4].

Proof of Theorem 6.1.3. Fix some $\boldsymbol{u} \in [\beta - 1, 1]^{\mathbb{N}}$. It follows from (6.1) that for all $x \in [0,1)$

$$\liminf_{m\to\infty} \frac{k(m,\boldsymbol{u},x)}{m} \geq \frac{\log 2}{\log \beta}.$$

Conversely, let $\varepsilon \in (0,1)$ and for each $m \geq 1$ define $\bar{k}(m) = \lceil (1+\varepsilon)\frac{m \log 2}{\log \beta} \rceil$. Let

$$\begin{aligned}
\mathcal{P}_m &= \{x \in [0,1) \,:\, \mathcal{I}_{\bar{k}(m)}(\boldsymbol{u},x) \not\subseteq \mathcal{D}_m(x)\} \\
&\subseteq \bigcup_{B \in \mathcal{D}_m} \bigcup_{A \in \mathcal{I}_{\bar{k}(m)} \,:\, A \not\subseteq B} A \cap B \\
&\subseteq \bigcup_{B \in \mathcal{D}_m} [B^-, B^- + \beta^{-(1+\varepsilon)\frac{m \log 2}{\log \beta}}] \cup [B^+ - \beta^{-(1+\varepsilon)\frac{m \log 2}{\log \beta}}, B^+],
\end{aligned}$$

where B^- and B^+ denote the lower and upper endpoint of B, respectively. Since $\mathcal{D}_m$ has $\beta^{\frac{m \log 2}{\log \beta}}$ elements, we have

$$\lambda(\mathcal{P}_m) \leq \beta^{\frac{m \log 2}{\log \beta}} \cdot 2 \cdot \beta^{-(1+\varepsilon)\frac{m \log 2}{\log \beta}} \leq 2 \cdot \beta^{-\varepsilon \frac{m \log 2}{\log \beta}},$$

which gives that $\sum_{m=1}^{\infty} \lambda(\mathcal{P}_m) < \infty$. From the Borel-Cantelli Lemma it follows that

$$\lambda(\{x \in [0,1) \,:\, x \in \mathcal{P}_m \text{ for infinitely many } m \in \mathbb{N}\}) = 0.$$

Hence,

$$\lambda(\{x \in [0,1) \,:\, \exists M \in \mathbb{N} \text{ s.t. } \forall m \geq M\ \mathcal{I}_{\bar{k}(m)}(\boldsymbol{u},x) \subseteq \mathcal{D}_m(x)\}) = 1,$$

or in other words, for Lebesgue almost all $x \in [0,1]$ there exists an $M \in \mathbb{N}$ such that for all $m \geq M$ it holds that $k(m,\boldsymbol{u},x) \leq \bar{k}(m)$. This gives

$$\limsup_{m\to\infty} \frac{k(m,\boldsymbol{u},x)}{m} \leq \limsup_{m\to\infty} \frac{\bar{k}(m)}{m} = (1+\varepsilon)\frac{\log 2}{\log \beta}, \qquad \lambda\text{-a.e.}$$

Since $\varepsilon > 0$ was arbitrary, this concludes the proof. □

Remark 6.3.1. Note that the first part of the previous proof holds for all $x \in [0,1]$. It is the second part that only holds Lebesgue almost everywhere.

§6.4 Final remarks

In practice it is not only the threshold value u that is subject to fluctuations due to noise on the circuit, but also the amplification factor β and the scaling factor $\beta - 1$. This issue and possible solutions to it were discussed in [DY06, W08, DGWY10]. Here we discuss the consequences for the value $k(m, \boldsymbol{u}, x)$.

Assume that the amplification factor and scaling factor fluctuate within intervals $[\beta_{\min}, \beta_{\max}], [\tilde{\beta}_{\min}, \tilde{\beta}_{\max}] \subseteq (1, 2)$, respectively. We use $\boldsymbol{\beta} = (\beta_n)_{n\geq 1} \in [\beta_{\min}, \beta_{\max}]^{\mathbb{N}}$ and $\tilde{\boldsymbol{\beta}} = (\tilde{\beta}_n)_{n\geq 1} \in [\tilde{\beta}_{\min}, \tilde{\beta}_{\max}]^{\mathbb{N}}$ to denote the sequence of consecutive amplification factors β_n and scaling factors $\tilde{\beta}_n - 1$, respectively. For an input value $x = x_0 \in [0, 1)$, the bits b_n are produced iteratively by setting $x_n = \beta_n x_{n-1} - (\tilde{\beta}_n - 1)b_n$ and

$$b_n = \begin{cases} 0, & \text{if } \beta_n x_{n-1} < u_n, \\ 1, & \text{if } \beta_n x_{n-1} \geq u_n. \end{cases}$$

This gives, for each n,

$$x = \sum_{i=1}^{n} \frac{(\tilde{\beta}_i - 1)b_i}{\prod_{j=1}^{i} \beta_j} + \frac{x_n}{\prod_{j=1}^{n} \beta_j}.$$

Note that

$$\sum_{i=1}^{\infty} \frac{(\tilde{\beta}_i - 1)b_i}{\prod_{j=1}^{i} \beta_j} \leq (\tilde{\beta}_{\max} - 1) \sum_{i=1}^{\infty} \frac{1}{\beta_{\min}^i} < \infty,$$

so

$$\varkappa := \lim_{n\to\infty} \frac{x_n}{\prod_{j=1}^{n} \beta_j} = x - \sum_{i=1}^{\infty} \frac{(\tilde{\beta}_i - 1)b_i}{\prod_{j=1}^{i} \beta_j}$$

is finite.

If for $k \geq 1$ the first k output bits of the β-encoder are $b_1, \ldots, b_k$ for an input signal $x \in [0, 1)$, an amplification sequence $\boldsymbol{\beta} \in [\beta_{\min}, \beta_{\max}]^{\mathbb{N}}$, a scaling sequence $\tilde{\boldsymbol{\beta}} \in [\tilde{\beta}_{\min}, \tilde{\beta}_{\max}]^{\mathbb{N}}$ and a threshold sequence $\boldsymbol{u} \in [0, 1]^{\mathbb{N}}$, then optimally (that means, knowing the value of $\varkappa$) we know that x lies in the interval

$$\tilde{\mathcal{I}}_k(\boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x) = \Big[(\tilde{\beta}_{\min} - 1) \sum_{i=1}^{k} \frac{b_i}{\beta_{\max}^i} + \varkappa, (\tilde{\beta}_{\max} - 1) \sum_{i=1}^{k} \frac{b_i}{\beta_{\min}^i} + \frac{\tilde{\beta}_{\max} - 1}{\beta_{\min} - 1} \frac{1}{\beta_{\min}^k} + \varkappa\Big].$$

The lower and upper endpoints of the sets $\{\tilde{\mathcal{I}}_k(\boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x)\}$ form an increasing and decreasing sequence, respectively, so

$$\bigcap_{k\geq 1} \tilde{\mathcal{I}}_k(\boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x) = \Big[(\tilde{\beta}_{\min} - 1) \sum_{i=1}^{\infty} \frac{b_i}{\beta_{\max}^i} + \varkappa, (\tilde{\beta}_{\max} - 1) \sum_{i=1}^{\infty} \frac{b_i}{\beta_{\min}^i} + \varkappa\Big], \tag{6.8}$$

which has length > 0 if $x \neq 0$ and either $\beta_{\min} < \beta_{\max}$ or $\tilde{\beta}_{\min} < \tilde{\beta}_{\max}$ or both are the case. (Indeed, note that in case $x \neq 0$ there exists $i \in \mathbb{N}$ such that $b_i = 1$.) Similar as to (6.5), we define $k(m, \boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x)$ to be the number of bits from the β-encoder

Chapter 6

subject to noise in the amplification, the scaling and the threshold that are necessary to obtain m base 2 digits of the number x, i.e.

$$k(m, \boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x) = \inf\{k \geq 1 \,:\, \tilde{\mathcal{I}}_k(\boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x) \subseteq \mathcal{D}_m(x)\}.$$

We see from (6.8) that if $\beta_{\min} < \beta_{\max}$ or $\tilde{\beta}_{\min} < \tilde{\beta}_{\max}$, then for all $x \in (0,1)$, all $\boldsymbol{\beta} \in [\beta_{\min}, \beta_{\max}]^{\mathbb{N}}$, all $\tilde{\boldsymbol{\beta}} \in [\tilde{\beta}_{\min}, \tilde{\beta}_{\max}]^{\mathbb{N}}$ and all $\boldsymbol{u} \in [0,1]^{\mathbb{N}}$ there exists $M > 0$ such that for all integers $m > M$ we have

$$k(m, \boldsymbol{u}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, x) = \infty.$$

In other words, the expected number of bits from the β-encoder that are necessary to obtain m base 2 digits of the number x is infinite for large m. Hence, this indicates that the proposed pseudo-random number generator of [JM16] is not efficient for generating large pseudo-random numbers if the β-encoder in this process is subject to noise in the amplification or scaling as well.

Bibliography

[A97] J. Aaronson. *An Introduction to Infinite Ergodic Theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 1997.

[AGH18] N. Abbasi, M. Gharaei, and A. J. Homburg. Iterated function systems of logistic maps: synchronization and intermittency. *Nonlinearity*, 31(8):3880–3913, 2018.

[AR66] L. M. Abramov and V. A. Rokhlin. The entropy of a skew product of measure-preserving transformations. *Amer. Math. Soc. Transl. Ser.*, 2(48):255–265, 1966.

[ANV15] R. Aimino, M. Nicol, and S. Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields*, 162(1-2):233–274, 2015.

[AD00] K. B. Athreya and J. Dai. Random logistic maps. I. *J. Theoret. Probab.*, 13(2):595–608, 2000.

[AS03] K. B. Athreya and H.-J. Schuh. Random logistic maps. II. The critical case. *J. Theoret. Probab.*, 16(4):813–830 (2004), 2003.

[AFGTV21] J. Atnip, G. Froyland, C. González-Tokman, and S. Vaienti. Thermodynamic formalism for random weighted covering systems. *Comm. Math. Phys.*, 386(2):819–902, 2021.

[BB16] W. Bahsoun and C. Bose. Mixing rates and limit theorems for random intermittent maps. *Nonlinearity*, 29(4):1417–1433, 2016.

[BBD14] W. Bahsoun, C. Bose, and Y. Duan. Decay of correlation for random intermittent maps. *Nonlinearity*, 27(7):1543–1554, 2014.

[BBR19] W. Bahsoun, C. Bose, and M. Ruziboev. Quenched decay of correlations for slowly mixing systems. *Trans. Amer. Math. Soc.*, 372(9):6547–6587, 2019.

[BRS20] W. Bahsoun, M. Ruziboev, and B. Saussol. Linear response for random dynamical systems. *Adv. Math.*, 364:107011, 44, 2020.

[BD17] S. Baker and K. Dajani. Induced random β-transformation. *Acta Arith.*, 178(1):1–14, 2017.

[B00] V. Baladi. *Positive Transfer Operators and Decay of Correlations*, volume 16. World scientific, 2000.

[BI08] L. Barreira and G. Iommi. Partial quotients of continued fractions and β-expansions. *Nonlinearity*, 21(10):2211–2219, 2008.

[BBDK96] J. Barrionuevo, R. M. Burton, K. Dajani, and C. Kraaikamp. Ergodic properties of generalized Lüroth series. *Acta Arith.*, 74(4):311–327, 1996.

[BPV86] P. Bergé, Y. Pomeau, and C. Vidal. *Order within Chaos.* J. Wiley, New York, 1986.

[B92] T. Bogenschütz. Entropy, pressure, and a variational principle for random dynamical systems. *Random Comput. Dynam.*, 1(1):99–116, 1992.

[B93] T. Bogenschütz. *Equilibrium States for Random Dynamical Systems.* PhD thesis, Bremen University, 1993.

[BC92] T. Bogenschütz and H. Crauel. The Abramov-Rokhlin formula. In *Ergodic Theory and Related Topics, III (Güstrow, 1990)*, volume 1514 of *Lecture Notes in Math.*, pages 32–35. Springer, Berlin, 1992.

[BQT21] C. Bose, A. Quas, and M. Tanzi. Random composition of L-S-V maps sampled over large parameter ranges. *Nonlinearity*, 34(6):3641–3675, 2021.

[BDK99] W. Bosma, K. Dajani, and C. Kraaikamp. Entropy and counting correct digits. Technical Report 9925, University of Nijmegen, 1999.

[BG97] A. Boyarsky and P. Góra. *Laws of Chaos: Invariant Measures and Dynamical Systems in One Dimension.* Springer Science & Business Media, 1997.

[B82] M. G. Branton. *The Dynamics of Random Maps.* PhD thesis, University of North Carolina at Chapel Hill, 1982.

[BLvS03] H. Bruin, S. Luzzatto, and S. van Strien. Decay of correlations in one-dimensional dynamics. *Ann. Sci. École Norm. Sup. (4)*, 36(4):621–646, 2003.

[BSvS03] H. Bruin, W. Shen, and S. van Strien. Invariant measures exist without a growth condition. *Comm. Math. Phys.*, 241(2-3):287–306, 2003.

[B99] J. Buzzi. Exponential decay of correlations for random Lasota-Yorke maps. *Comm. Math. Phys.*, 208(1):25–54, 1999.

[C02] N. Carlsson. A contractivity condition for iterated function systems. *J. Theoret. Probab.*, 15(3):613–630, 2002.

[CCD21] E. Charlier, C. Cisternino, and K. Dajani. Dynamical behavior of alternate base expansions. *Ergodic Theory Dynam. Systems*, pages 1–34, 2021.

[CE83] P. Collet and J.-P. Eckmann. Positive Liapunov exponents and absolute continuity for maps of the interval. *Ergodic Theory Dynam. Systems*, 3(1):13–46, 1983.

[DF01] K. Dajani and A. Fieldsteel. Equipartition of interval partitions and an application to number theory. *Proc. Amer. Math. Soc.*, 129(12):3453–3460, 2001.

[DJ17] K. Dajani and K. Jiang. Shrinking random β-transformation. *Indag. Math. (N.S.)*, 28(1):74–83, 2017.

[DK07] K. Dajani and C. Kalle. Random β-expansions with deleted digits. *Discrete Contin. Dyn. Syst.*, 18(1):199–217, 2007.

[DK13] K. Dajani and C. Kalle. Local dimensions for the random β-transformation. *New York J. Math.*, 19:285–303, 2013.

[DK20] K. Dajani and C. Kalle. Invariant measures, matching and the frequency of 0 for signed binary expansions. *Publ. Res. Inst. Math. Sci.*, 56(4):701–742, 2020.

[DK21] K. Dajani and C. Kalle. *A First Course in Ergodic Theory.* CRC Press, 2021.

[DKM21] K. Dajani, C. Kalle, and M. Maggioni. Matching for random systems with an application to minimal weight expansions. *Nonlinearity*, 34(6):3676–3708, 2021.

[DK03] K. Dajani and C. Kraaikamp. Random β-expansions. *Ergodic Theory Dynam. Systems*, 23(2):461–479, 2003.

[DO18] K. Dajani and M. Oomen. Random N-continued fraction expansions. *J. Approx. Theory*, 227:1–26, 2018.

[DdV05] K. Dajani and M. de Vries. Measures of maximal entropy for random β-expansions. *J. Eur. Math. Soc. (JEMS)*, 7(1):51–68, 2005.

[DdV07] K. Dajani and M. de Vries. Invariant densities for random β-expansions. *J. Eur. Math. Soc. (JEMS)*, 9(1):157–176, 2007.

[DDGV02] I. Daubechies, R. A. DeVore, C. S. Güntürk, and V. A. Vaishampayan. Beta expansions: a new approach to digitally corrected A/D conversion. *Proc. IEEE Int. Symp. Circ. Syst.*, 2:784–787, 2002.

[DDGV06] I. Daubechies, R. A. DeVore, C. S. Güntürk, and V. A. Vaishampayan. A/D conversion with imperfect quantizers. *IEEE Trans. Inform. Theory*, 52(3):874–885, 2006.

[DGWY10] I. Daubechies, S. Güntürk, Y. Wang, and Ö. Yılmaz. The golden ratio encoder. *IEEE Trans. Inform. Theory*, 56(10):5097–5110, 2010.

[DY06] I. Daubechies and Ö. Yilmaz. Robust and practical analog-to-digital conversion with exponential precision. *IEEE Trans. Inform. Theory*, 52(8):3533–3545, 2006.

[EW11] M. Einsiedler and T. Ward. Ergodic theory with a view towards number theory, volume 259 of Graduate Texts in Mathematics. Springer-Verlag London, Ltd., London, 2011

[EJK90] 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(3):377–390, 1990.

[F97] C. Faivre. On decimal and continued fraction expansions of a real number. *Acta Arith.*, 82(2):119–128, 1997.

[F98] C. Faivre. A central limit theorem related to decimal and continued fraction expansion. *Arch. Math. (Basel)*, 70(6):455–463, 1998.

[F01] C. Faivre. On calculating a continued fraction expansion from a decimal expansion. *Acta Sci. Math. (Szeged)*, 67(3-4):505–519, 2001.

[FWL16] L. Fang, M. Wu, and B. Li. Limit theorems related to beta-expansion and continued fraction expansion. *J. Number Theory*, 163:385–405, 2016.

[FWL19] L. Fang, M. Wu, and B. Li. Beta-expansion and continued fraction expansion of real numbers. *Acta Arith.*, 187(3):233–253, 2019.

[F99] G. Froyland. Ulam's method for random interval maps. *Nonlinearity*, 12(4):1029–1052, 1999.

[GH17] M. Gharaei and A. J. Homburg. Random interval diffeomorphisms. *Discrete Cont. Dynam. Syst. Ser. S*, 10(2):241–272, 2017.

[G04] S. Gouëzel. Sharp polynomial estimates for the decay of correlations. *Israel J. Math.*, 139:29–65, 2004.

[G12] C.S. Güntürk. Mathematics of analog-to-digital conversion. *Comm. Pure Appl. Math.*, 65(12):1671–1696, 2012.

[HKST15] J. Heinonen, P. Koskela, N. Shanmugalingam, and J. T. Tyson. *Sobolev Spaces on Metric Measure Spaces. An approach based on upper gradients*, volume 27 of New Mathematical Monographs. Cambridge University Press, Cambridge, 2015.

[H09] A. Herczegh. Central limit theorems in ergodic theory. Master's thesis, Eötvös Loránd University, 2009.

[HKR$^+$22] A. J. Homburg, C. Kalle, M. Ruziboev, E. Verbitskiy, and B. Zeegers. Critical intermittency in random interval maps. *Comm. Math. Phys.*, 394(1):1–37, 2022.

[HPR21] A. J. Homburg, H. Peters, and V. Rabodonandrianandraina. Critical intermittency in rational maps. Preprint, available at arXiv:1909.05559v2, 2021.

[H04] H. Hu. Decay of correlations for piecewise smooth maps with indifferent fixed points. *Ergodic Theory and Dynamical Systems*, 24(2):495–524, 2004.

[I12] T. Inoue. Invariant measures for position dependent random maps with continuous random parameters. *Studia Math.*, 208(1):11–29, 2012.

[J81] M. V. Jakobson. Absolutely continuous invariant measures for one-parameter families of one-dimensional maps. *Comm. Math. Phys.*, 81(1):39–88, 1981.

[JW09] D. Jiménez and Y. Wang. The $\beta\alpha$-encoders for robust A/D conversion. *Acta Appl. Math.*, 107(1-3):313–323, 2009.

[JM16] Y. Jitsumatsu and K. Matsumura. A β-ary to binary conversion for random number generation using a β encoder. *Nonlinear Theory and Its Applications, IEICE*, 38–55, 2016.

[JMKA13] Y. Jitsumatsu, K. Matsumura, T. Kohda, and K. Aihara. Pseudo-random number generator using beta-encoder cmos circuit. *The 3rd Int. Symp. Innovative Mathematical Modelling*, page 107, 2013.

[JSS11] T. Jordan, P. Shmerkin, and B. Solomyak. Multifractal structure of Bernoulli convolutions. *Math. Proc. Cambridge Philos. Soc.*, 151(3):521–539, 2011.

[KKV17] C. Kalle, T. Kempton, and E. Verbitskiy. The random continued fraction transformation. *Nonlinearity*, 30(3):1182–1203, 2017.

[KM22a] C. Kalle and M. Maggioni. Invariant densities for random systems of the interval. *Ergodic Theory Dynam. Systems*, 42(1):141–179, 2022.

[KM22b] C. Kalle and M. Maggioni. On approximation by random Lüroth expansions. *Int. J. Number Theory*, 18(5):1013–1046, 2022.

[KMTV22] C. Kalle, V. Matache, M. Tsujii, and E. Verbitskiy. Invariant densities for random continued fractions. *J. Math. Anal. Appl.*, 512(2):Paper No. 126163, 28, 2022.

[KVZ22] C. Kalle, E. Verbitskiy, and B. Zeegers. Random Lochs' Theorem. *Studia Math.*, 267(2):201–239, 2022.

[KZ22] C. Kalle and B. Zeegers. Decay of correlations for critically intermittent systems. Preprint, available at arXiv:2206.07601, 2022.

[KN92] G. Keller and T. Nowicki. Spectral theory, zeta functions and the distribution of periodic points for Collet-Eckmann maps. *Comm. Math. Phys.*, 149(1):31–69, 1992.

[K14] T. Kempton. On the invariant density of the random β-transformation. *Acta Math. Hungar.*, 142(2):403–419, 2014.

[K86] Y. Kifer. *Ergodic Theory of Random Transformations*, volume 10 of *Progress in Probability and Statistics*. Birkhäuser Boston, Inc., Boston, MA, 1986.

[KHA12] T. Kohda, Y. Horio, and K. Aihara. Beta-expansion attractors observed in A/D converters. *Chaos*, 22(4):1–18, 047512, 2012.

[KHTA12] T. Kohda, Y. Horio, Y. Takahashi, and K. Aihara. Beta encoders: symbolic dynamics and electronic implementation. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 22(9):1–55, 1230031, 2012.

[KJ16] I. Koji and Y. Jitsumatsu. Random number generation using outputs from multiple beta encoders. In *Proceedings of Nonlinear Theory and Its Applications 2016*, pages 249–252, 2016.

[K58] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. In *Doklady of Russian Academy of Sciences*, volume 119, pages 861–864, 1958.

[LMS04] A. Lasota, J. Myjak, and T. Szarek. Markov operators and semifractals. In *Fractal geometry and stochastics III*, volume 57 of *Progr. Probab.*, pages 3–22. Birkhäuser, Basel, 2004.

[LY73] A. Lasota and J. A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488 (1974), 1973.

[LY88] F. Ledrappier and L.-S. Young. Entropy formula for random transformations. *Probab. Theory Related Fields*, 80(2):217–240, 1988.

[L96] C. Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems (Montevideo, 1995)*, volume 362 of *Pitman Res. Notes Math. Ser.*, pages 56–75. Longman, Harlow, 1996.

[LSV99] C. Liverani, B. Saussol, and S. Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999.

[LW08] B. Li and J. Wu. Beta-expansion and continued fraction expansion. *J. Math. Anal. Appl.*, 339(2):1322–1331, 2008.

[L64] G. Lochs. Vergleich der Genauigkeit von Dezimalbruch und Kettenbruch. *Abh. Math. Sem. Univ. Hamburg*, 27:142–144, 1964.

[L06] S. Luzzatto. Stochastic-like behaviour in nonuniformly expanding maps. In *Handbook of Dynamical Systems. Vol. 1B*, pages 265–326. Elsevier B. V., Amsterdam, 2006.

[LM13] S. Luzzatto and I. Melbourne. Statistical properties and decay of correlations for interval maps with critical points and singularities. *Comm. Math. Phys.*, 320(1):21–35, 2013.

[L02] M. Lyubich. Almost every real quadratic map is either regular or stochastic. *Ann. of Math. (2)*, 156(1):1–78, 2002.

[M57] G. W. Mackey. Borel structure in groups and their duals. *Trans. Amer. Math. Soc.*, 85:134–165, 1957.

[MIS$^+$15] T. Makino, Y. Iwata, K. Shinohara, Y. Jitsumatsu, M. Hotta, H. San, and K. Aihara. Rigorous estimates of quantization error for A/D converters based on beta-map. *Nonlinear Theory and Its Applications*, 6(1):99–111, 2015.

[MP80] P. Manneville and Y. Pomeau. Different ways to turbulence in dissipative dynamical systems. *Phys. D*, 1(2):219–226, 1980.

[dMvS93] W. de Melo and S. van Strien. *One-Dimensional Dynamics*, Springer-Verlag, New York, 1993.

[M81] M. Misiurewicz. Absolutely continuous measures for certain maps of an interval. *Inst. Hautes Études Sci. Publ. Math.*, (53):17–51, 1981.

[M85a] T. Morita. Asymptotic behavior of one-dimensional random dynamical systems. *J. Math. Soc. Japan*, 37(4):651–663, 1985.

[M85b] T. Morita. Random iteration of one-dimensional transformations. *Osaka Journal of Mathematics*, 22(3):489–518, 1985.

[M86] T. Morita. Entropy of random dynamical systems. *Proc. Japan Acad. Ser. A Math. Sci.*, 62(4):121–124, 1986.

[M05] R. Murray. Approximation of invariant measures for a class of maps with indifferent fixed points. *Mathematics Research Report Series II No. 106.* University of Waikato, 2005.

[NPPT21] M. Nicol, F. Perez Pereira, and A. Török. Large deviations and central limit theorems for sequential and random systems of intermittent maps. *Ergodic Theory Dynam. Systems*, 41(9):2805–2832, 2021.

[NTV18] M. Nicol, A. Török, and S. Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dynam. Systems*, 38:1127–1153, 2018.

[NvS91] T. Nowicki and S. van Strien. Invariant measures exist under a summability condition for unimodal maps. *Invent. Math.*, 105(1):123–136, 1991.

[P84] S. Pelikan. Invariant densities for random maps of the interval. *Transactions of the American Mathematical Society*, 281(2):813–825, 1984.

[P89] K. Petersen. *Ergodic Theory*, volume 2 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1989.

[P80] G. Pianigiani. First return map and invariant measures. *Israel J. Math.*, 35(1-2):32–48, 1980.

[PM80] Y. Pomeau and P. Manneville. Intermittent transition to turbulence in dissipative dynamical systems. *Comm. Math. Phys.*, 74(2):189–197, 1980.

[R57a] A. Rényi. On algorithms for the generation of real numbers. *Magyar Tud. Akad. Mat. Fiz. Oszt. Közl.*, 7:265–293, 1957.

[R57b] A. Rényi. Representations for real numbers and their ergodic properties. *Acta Math. Acad. Sci. Hungar*, 8:477–493, 1957.

[R52] V. A. Rokhlin. On the fundamental ideas of measure theory. *Amer. Math. Soc. Translation*, 1952(71):55, 1952.

[SKM$^+$13] H. San, T. Kato, T. Maruyama, K. Aihara, and M. Hotta. Non-binary pipeline analog-to-digital converter based on beta-expansion. *IEICE Trans. Fundamentals*, 96(2):415–421, 2013.

[S09] O. M. Sarig. Lecture notes on thermodynamic formalism for topological markov shifts. *Penn State*, 2009.

[S48] C. E. Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3):379–423, 1948.

[SJO15] Y. Shu, Y. Jitsumatsu, and K. Oda. Performance evaluation of a random number generation using a beta encoder. *NOLTA2015*, pages 511–514, 2015.

[S03] N. Sidorov. Almost every number has a continuum of β-expansions. *Amer. Math. Monthly*, 110(9):838–842, 2003.

[S59] Ya. G. Sinai. On the notion of entropy of a dynamical system. *Doklady of Russian Academy of Sciences*, 124:768–771, 1959.

[S19] S. Suzuki. Invariant density functions of random β-transformations. *Ergodic Theory Dynam. Systems*, 39(4):1099–1120, 2019.

[T80] M. Thaler. Estimates of the invariant densities of endomorphisms with indifferent fixed points. *Israel J. Math.*, 37(4):303–314, 1980.

[T04] R. Toledano. A note on the Lebesgue differentiation theorem in spaces of homogeneous type. *Real Anal. Exchange*, 29(1):335–339, 2003/04.

[VO16] M. Viana and K. Oliveira. *Foundations of Ergodic Theory*, volume 151 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2016.

[W00] P. Walters. *An Introduction to Ergodic Theory*, volume 79. Springer Science & Business Media, 2000.

[W08] R. Ward. On robustness properties of beta encoders and golden ratio encoders. *IEEE Trans. Inform. Theory*, 54(9):4324–4334, 2008.

[W06] J. Wu. Continued fraction and decimal expansions of an irrational number. *Adv. Math.*, 206(2):684–694, 2006.

[W08] J. Wu. An iterated logarithm law related to decimal and continued fraction expansions. *Monatsh. Math.*, 153(1):83–87, 2008.

[Y92] L.-S. Young. Decay of correlations for certain quadratic maps. *Comm. Math. Phys.*, 146(1):123–138, 1992.

[Y98] L.-S. Young. Statistical properties of dynamical systems with some hyperbolicity. *Ann. of Math.*, 147:585–650, 1998.

[Y99] L.-S. Young. Recurrence times and rates of mixing. *Israel J. Math.*, 110:153–188, 1999.

[Z] B. Zeegers. Intermittency generated by attracting and weakly repelling fixed points. To appear in *Indag. Math.* DOI: 10.1016/j.indag.2022.12.002.

[Z18] B. Zeegers. On invariant densities and Lochs' Theorem for random piecewise monotonic interval maps. Master's thesis, Leiden University, 2018.

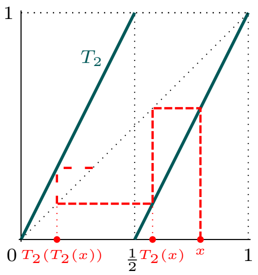
[Z08] Y. Zhu. On local entropy of random transformations. *Stoch. Dyn.*, 8(2):197–207, 2008.

Samenvatting

In dit proefschrift staat het gedrag van stochastische intervalafbeeldingen centraal. Een intervalafbeelding $T : [0,1] \to [0,1]$ is een functie die ieder getal x in $[0,1]$ stuurt naar een getal $T(x)$ in $[0,1]$ (dat afhangt van x). Door een intervalafbeelding herhaald toe te passen krijgen we zogeheten banen. Dit zijn rijtjes van de vorm

$$x, \quad T(x), \quad T(T(x)), \quad T(T(T(x))), \quad \cdots .$$

Een voorbeeld van een intervalafbeelding is de functie T_2 gegeven door $T_2(x) = 2x$ voor $0 \leq x < \frac{1}{2}$ en $T_2(x) = 2x-1$ voor $\frac{1}{2} \leq x \leq 1$. De baan onder T_2 van bijvoorbeeld $x = \frac{15}{19}$ kan dan als volgt worden gevisualiseerd:

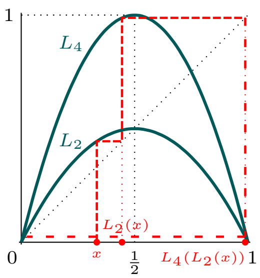


Intervalafbeeldingen zijn voorbeelden van *deterministische* discrete dynamische systemen. Dit zijn systemen waarvan de toestand verandert in tijdstappen en wordt bepaald door één functie T. Een toestand x wordt dan opgevolgd door toestand $T(x)$.

Een *stochastische* intervalafbeelding daarentegen is een dynamisch systeem waarbij er een *collectie* functies beschikbaar is en waaruit per tijdstap één functie wordt gekozen met een bepaalde kans. Een voorbeeld van zo'n systeem: kies per tijdstap met kans p de intervalafbeelding L_2 gegeven door $L_2(x) = 2x(1-x)$ voor $0 \leq x \leq 1$ en met kans $1-p$ de intervalafbeelding L_4 gegeven door $L_4(x) = 4x(1-x)$ voor $0 \leq x \leq 1$. Dit kan bijvoorbeeld de volgende baan opleveren:

$$x, \quad L_2(x), \quad L_4(L_2(x)), \quad L_2(L_4(L_2(x))), \quad \cdots ,$$

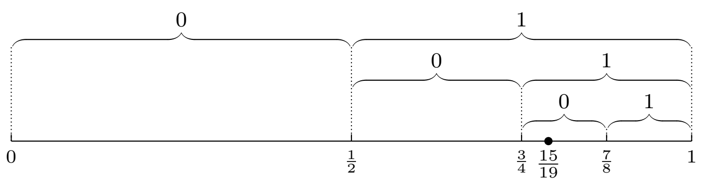
wat voor bijvoorbeeld $x = \frac{1}{3}$ als volgt kan worden gevisualiseerd:



Dit voorbeeld van een stochastische intervalafbeelding is speciaal in de zin dat het *kritisch intermitterend* gedrag vertoont, een verschijnsel waar we in het eerste deel van dit proefschrift in geïnteresseerd zijn. In dat geval wisselen banen tussen periodes van chaotisch en stabiel gedrag als gevolg van de wisselwerking tussen een *superaantrekkend vast punt* en een *afstotend vast punt* in het systeem. In bovenstaand voorbeeld is 0 een afstotend vast punt onder L_2 en L_4 omdat $L_2(0) = L_4(0) = 0, L_2'(0) > 1, L_4'(0) > 1$ en is $\frac{1}{2}$ een superaantrekkend vast punt onder L_2 omdat $L_2(\frac{1}{2}) = \frac{1}{2}$, $L_2'(\frac{1}{2}) = 0$. Doordat $L_4(\frac{1}{2}) = 1$ en $L_2(1) = L_4(1) = 0$ kent een baan onder dit systeem gewoonlijk periodes van chaotisch gedrag die worden afgewisseld met stabiele periodes dicht bij 0.

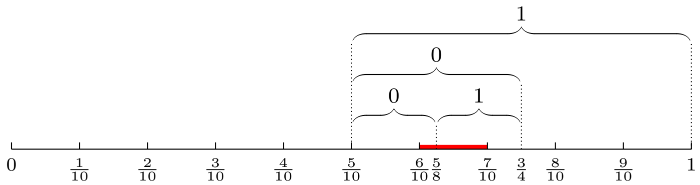
In Hoofdstuk 2 beschouwen we een brede klasse van stochastische intervalafbeeldingen die kritisch intermitterend zijn. We bewijzen onder meer onder welke voorwaarden er wel of geen *absoluut continue invariante kansmaat* bestaat. Het al dan niet bestaan hiervan geeft aan dat respectievelijk het chaotische dan wel stabiele gedrag in het systeem overheerst. In het bovenstaande voorbeeld met L_2 en L_4 bestaat zo'n kansmaat bijvoorbeeld dan en slechts dan als de kans p dat L_2 wordt gekozen kleiner is dan $\frac{1}{2}$. In Hoofdstuk 3 leiden we voor een klasse kritisch intermitterende systemen statistische eigenschappen af die karakteristiek zijn voor intermitterende systemen. Zo vinden we voor een klasse van *testfuncties* dat indien een absoluut continue invariante kansmaat bestaat *correlaties* polynomiaal snel naar nul gaan en geven we voorwaarden wanneer voor deze systemen een bepaalde vorm van de *Centrale Limietstelling* geldt. Tenslotte bewijzen we in Hoofdstuk 4 dat soortgelijke resultaten als in Hoofdstuk 2 over het bestaan van een absoluut continue invariante kansmaat ook gelden voor een intermitterend systeem waarbij de mate van zowel aantrekking als afstoting van de vaste punten met een orde van grootte is verlaagd.

In het tweede deel van dit proefschrift kijken we naar getalsontwikkelingen gegenereerd door stochastische intervalafbeeldingen. Getalsontwikkelingen zijn een manier om getallen uit te drukken door middel van een specifieke verzameling symbolen of cijfers. Het meeste bekende voorbeeld is de decimale ontwikkeling, waarbij getallen worden uitgedrukt met de cijfers $0, 1, \ldots, 9$. De decimale ontwikkeling van bijvoorbeeld $\frac{15}{19}$ is $0.78947\cdots$. Een ander bekend voorbeeld is de binaire ontwikkeling, waarbij alleen de cijfers 0 en 1 worden gebruikt. De cijfers in de binaire ontwikkeling van een getal x in $[0, 1)$ kunnen als volgt worden bepaald. Deel het interval $[0, 1)$ op in twee helften. Dan is het eerste cijfer in de binaire ontwikkeling van x gelijk aan 0 als x kleiner is dan het grenspunt $\frac{1}{2}$ en anders 1. Deel nu de helft waarin x zich bevindt verder op in twee helften. Dan is het tweede cijfer gelijk aan 0 als x kleiner is dan het grenspunt dat de twee helften scheidt en anders 1. Enzovoorts. Op deze manier vinden we bijvoorbeeld dat de eerste drie cijfers in de binaire ontwikkeling van $\frac{15}{19}$ gelijk zijn aan respectievelijk 1, 1 en 0:



In plaats van bovenstaande methode kunnen de cijfers in de binaire ontwikkeling van een getal x in $[0,1)$ ook worden bepaald door de baan van x onder de intervalafbeelding T_2 te volgen. Het eerste cijfer is dan 0 als x kleiner is dan $\frac{1}{2}$ en anders 1, het tweede cijfer is 0 als $T_2(x)$ kleiner is dan $\frac{1}{2}$ en anders 1, het derde cijfer is 0 als $T_2(T_2(x))$ kleiner is dan $\frac{1}{2}$ en anders 1, enzovoorts. Voor bijvoorbeeld $x = \frac{15}{19}$ zien we (door twee pagina's terug te bladeren) dat dit inderdaad opnieuw de ontwikkeling $1, 1, 0, \cdots$ oplevert. We zeggen dat T_2 de binaire ontwikkeling *genereert*. Behalve de binaire ontwikkeling kunnen veel soorten getalsontwikkelingen worden gegenereerd door intervalafbeeldingen, waaronder ook de decimale ontwikkeling. Ook stochastische intervalafbeeldingen kunnen getalsontwikkelingen genereren door op eenzelfde manier de banen te volgen en cijfers toe te kennen aan specifieke deelintervallen.

Als we een deel van de cijfers in een getalsontwikkeling van een getal x in $[0,1)$ kennen, dan geeft dit informatie over de locatie van x in $[0,1)$. Als we bijvoorbeeld weten dat het eerste cijfer in de decimale ontwikkeling van een verder onbekend getal x in $[0,1)$ gelijk is aan 6, dan weten we dat x zich bevindt tussen $\frac{6}{10}$ en $\frac{7}{10}$. Voor dit voorbeeld kunnen we dan de eerste twee cijfers in de binaire ontwikkeling van x bepalen, maar niet meer cijfers:



Bovenstaande laat zien dat x in het deelinterval ligt dat bestaat uit getallen waarvan de binaire ontwikkeling begint met respectievelijk de cijfers 1 en 0, maar dat het niet duidelijk is of het derde cijfer in de binaire ontwikkeling van x een 0 of een 1 is omdat we niet weten of x kleiner of groter is dan $\frac{5}{8}$. Meer in het algemeen kunnen we de volgende vraag stellen: Stel we weten de eerste n cijfers in een getalsontwikkeling van een verder onbekend getal x in $[0,1)$. Hoeveel cijfers kunnen we dan bepalen in een andere getalsontwikkeling van x? Dit vraagstuk is tot op zekere hoogte opgelost voor het geval dat n heel groot is en decimale ontwikkelingen worden vergeleken met zogeheten reguliere kettingbreuken. Dit staat bekend als de *Stelling van Lochs*.

In Hoofdstuk 5 breiden we het resultaat van de stelling van Lochs uit naar een brede klasse van paren getalsontwikkelingen die worden gegenereerd door stochastische intervalafbeeldingen. Bovendien bewijzen we een bijbehorende Centrale Limietstelling. Tenslotte bestuderen we in Hoofdstuk 6 een vraagstuk dat gerelateerd is aan de stelling van Lochs in de context van zogeheten β-encoders. De resultaten in dit hoofdstuk geven inzicht in de mate van geschiktheid van β-encoders als *pseudo-toevalsgeneratoren*.

Acknowledgements

This dissertation could not have been written without the help of numerous people around me.

First of all, I am very grateful to my supervisor Charlene for giving me the opportunity to do this PhD. Charlene, you supervised me during my master's project and I was happy that we could continue this collaboration for my PhD. Your expertise allowed me to grow as a mathematician and I very much enjoyed working together on our projects. Thank you for always being enthusiastic, for your patience and for your positive attitude. I also appreciated the freedom and independence you gave me, both in doing research as well as sharing it abroad.

I also thank my promotor Evgeny and my other co-authors Ale Jan and Marks. Marks, thank you for the interesting discussions but also for the laughs we shared.

I want to express my gratitude to my promotor Frank and to the members of the reading committee for reading and assessing my dissertation.

Wael Bahsoun, thank you for inviting me to Loughborough. I had a great stay and enjoyed our mathematical discussions. Thank you for spending time in helping me with my future career.

I am grateful to have been part of the POD group. I shared an office with only nice people and I enjoyed the time we spent together. Moreover, I want to thank my fellow organizers of the iPOD seminar. It was always a pleasure to work together. Furthermore, I would like to thank the members of the reading club, where we learned great mathematics presented in a comfortable environment.

I want to thank my colleagues at the Mathematical Institute and in particular the PhD students for all the fun and pleasant lunches, coffee breaks, borrels and more generally for the great atmosphere. I am in particular grateful to those who made the covid lockdown more lively by maintaining the necessary after work drinks, by running to the beach or doing calisthenics (regardless of snow or hail), and more of which I have nice memories. And speaking of sports, thanks to all my sporty colleagues with whom I enjoyed the half marathon, the triathlon and trainings at the USC, as well as our fun tennis, volleyball and football matches. The latter resulted in Mathletico recently finishing first in the highly esteemed BSP football competition, an achievement that will be talked about for a long time in the corridors of the MI.

I would like to thank my housemates and my friends outside the Mathematical Institute as well. In particular Wim, the help and distraction you provided me during my PhD has been essential. The conversations we have are always interesting and fun and I look back at an amazing bike trip to Paris. En Daniël, jij bewijst dat een verre vriend soms beter is dan een goede buur. Bedankt voor alle fantastische reizen.

Tussen de klippen doorzeilend hebben we veel lol getrapt, menig tafel de nodige eer bewezen (soms gekauwd als een aap op knikkers), en meer. Ik kijk reikhalzend uit naar het volgende avontuur.

Finally, I want to thank my parents Marianne and Siem, my sister Sascha and my brothers Guido, Esben and Mathé for their continuous love and support throughout my life. I am also grateful to Gilles for helping me out multiple times. Aan mijn ouders, zonder al jullie hulp en aanmoedigingen zou ik nooit zo ver zijn gekomen en zou dit proefschrift er nooit zijn gekomen. Bedankt dat ik bij jullie altijd een luisterend oor en wijze raad heb gekregen. En Mathé, bedankt voor alle keren dat je me hebt geholpen van basisschool tot en met PhD.

Curriculum Vitae

Benthen Pascal Zeegers was born in Alkmaar in 1993. He obtained his gymnasium diploma in 2012 at secondary school Han Fortmann in Heerhugowaard and continued with the bachelors Mathematics and Physics at Leiden University. He completed both bachelors cum laude in 2015 and alongside finished the Honours College programme Bèta and Life Science. His bachelor thesis contains new results on the Kuramoto model in the context of a hierarchical lattice structure. Benthen contributed as a co-author to an article that builds upon these results.

In 2018 he received his master's degree in Mathematics summa cum laude at Leiden University. Benthen carried out part of his master's project at the University of Vienna on an acquired Erasmus+ grant. This project led to new results in several topics on random dynamical systems and part of these results became the starting point for the work presented in the second part of his doctoral dissertation.

Benthen started his PhD studies in November 2018 at the Mathematical Institute of Leiden University under the supervision of Dr. C. Kalle. His doctoral dissertation *Intermittency and Number Expansions for Random Interval Maps* covers results obtained during five different research projects on random dynamical systems. Benthen presented his research projects in many local and international seminars, workshops and conferences. He has been a teaching assistant for various courses taught at the Mathematical Institute in Leiden and supervised students of the Leiden PRE-University programme. For over two years he co-organized the Leiden seminar on probability theory, operations research and dynamical systems.