Intermittency and number expansions for random interval maps
Zeegers, B.P.

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.

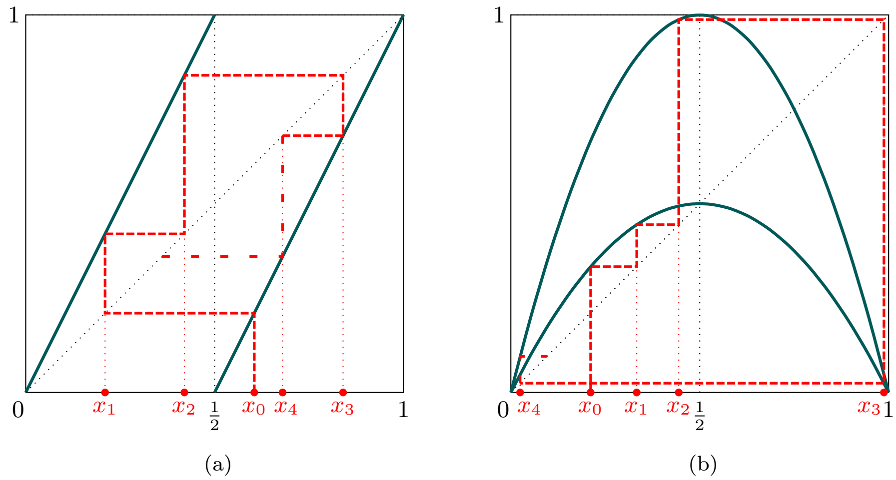


Figure 1.1: Example of (a) *an interval map, namely the doubling map* $x \mapsto 2x \bmod 1$*, and* (b) *a random interval map consisting of the two logistic maps* $x \mapsto 2x(1-x)$ *and* $x \mapsto 4x(1-x)$*. The dashed lines indicate parts of an orbit, which in* (a) *is of the form as in* (1.1) *and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

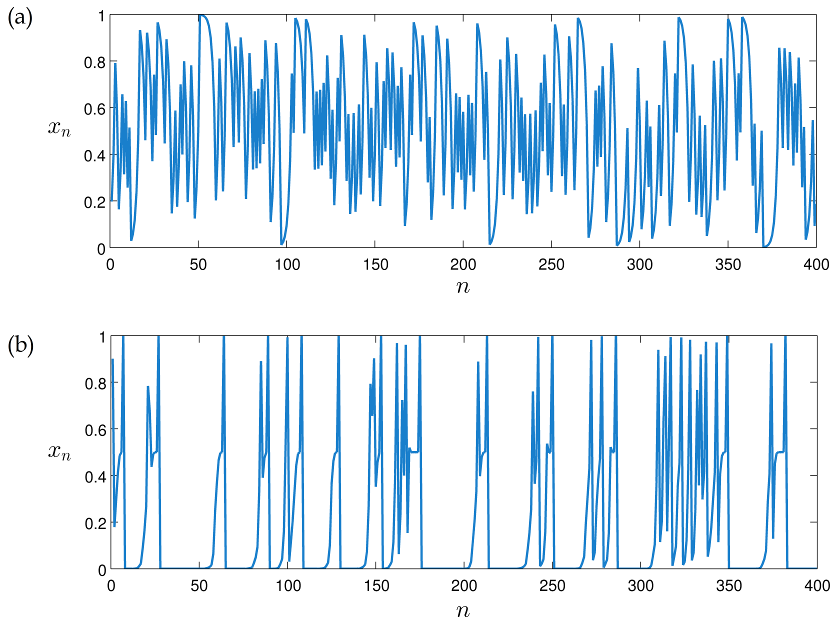


Figure 1.2: Time series of the systems from Figure 1.1: In (a) *part of an orbit under the doubling map*[1] *is shown, whereas* (b) *depicts a typical trajectory under the random map from Figure 1.1(b) if the two logistic maps are chosen with equal probability and independently at each time step.*

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 $\{0,1,\dots,9\}^{\mathbb{N}}$, 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 $d_1, \dots, d_n$ of a further unknown irrational number $x \in [0,1]$. What is the largest number $m(n,x)$ of digits $a_1, \dots, a_{m(n,x)}$ 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

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

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 $\mathcal{P}_1 \leq \mathcal{P}_2$ 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*

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

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) *$\int_{T(B)} \varphi dm = \int_B (\varphi \circ T) J_m T dm$ 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

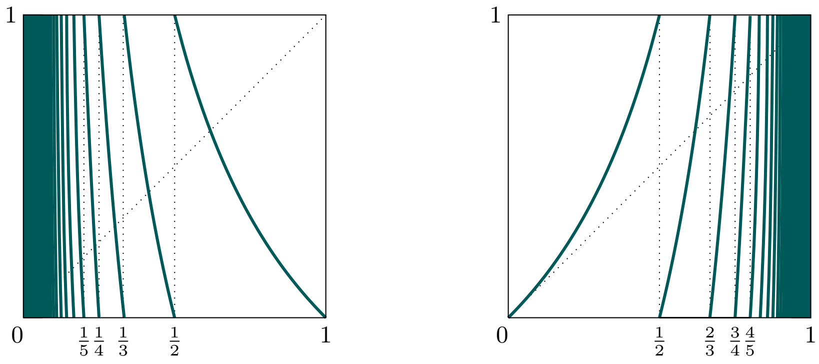
$$\frac{d\mu_G}{d\lambda}(x) = \frac{1}{\log 2} \frac{1}{x+1}, \qquad x \in [0,1],$$

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 $x \neq 1$,

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

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

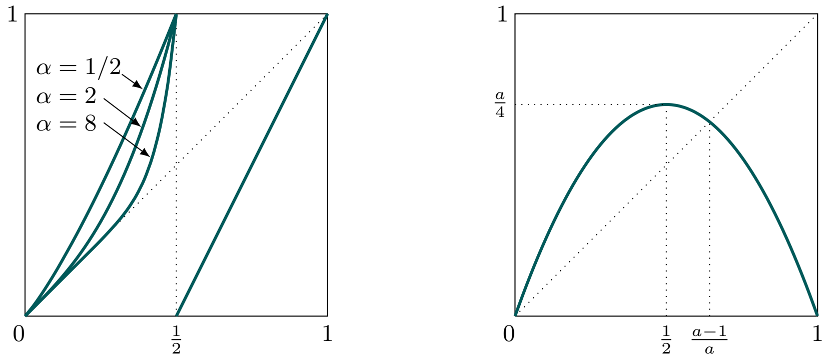
$$C_1 \cdot x^{-\alpha} \leq \frac{d\mu_\alpha}{d\lambda}(x) \leq C_2 \cdot x^{-\alpha}, \qquad x \in (0,1].$$

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 $x \mapsto x + x^{1+\alpha} \mod 1$ 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 $0 < K \le M$ and $\ell > 1$ such that for each $x \in [0,1]$

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

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 $\omega = (\omega_1, \omega_2, \ldots) \in I^{\mathbb{N}}$ and $\tau : I^{\mathbb{N}} \to I^{\mathbb{N}}$ is the *left shift* on $I^{\mathbb{N}}$, i.e. $\tau\omega = (\omega_2, \omega_3, \ldots)$. 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

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

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 *$\boldsymbol{p}$-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.